# Chapter 2
# THE FORMALISM OF QUANTUM MECHANICS

The purpose of this section is twofold: (1) to enumerate the basic postulates of quantum mechanics in a form which is manifestly *representation independent,* and which is potentially applicable to general quantum mechanical systems of one or more particles; and (2) to provide the mathematical framework necessary to understand the contents of these postulates. To this end it is useful to adopt an approach which mixes these two tasks. Thus, we start out by simply stating the first postulate, followed by a general discussion of its mathematical implications. This procedure will be repeated for the remaining postulates. Following the general structure introduced earlier, we begin the postulates by describing the means by which an arbitrary dynamical state of a quantum mechanical system is specified.

## 2.1 Postulate I: Specification of the Dynamical State

> The dynamical state of a quantum mechanical system is at each instant of time associated with a **state vector** $|\psi\rangle$. Possible state vectors are elements of a complex **linear vector space** $S$, referred to as the state space or Hilbert space of the system.

Obviously, a prerequisite to our understanding of this postulate and its implications is a knowledge of linear vector spaces.

### *2.1.1 Properties of Linear Vector Spaces*

A set $S$ of objects $\{|\psi\rangle, |\zeta\rangle, |\xi\rangle, \ldots\}$ forms a *linear vector space* (LVS) if it is closed under two mutually distributive operations: (1) an associative and commutative law of **vector addition**, and (2) multiplication by elements of an associated **field** $F$ of scalars $\{\lambda, \mu, \nu, \ldots\}$. This operation of vector addition is assumed to satisfy the properties enumerated below.

1. For all vectors $|\psi\rangle$ and $|\xi\rangle$ in $S$ there exists a vector $|\chi\rangle$ in $S$ such that $|\chi\rangle = |\psi\rangle + |\xi\rangle$. (Closure)
2. $|\psi\rangle + |\xi\rangle = |\xi\rangle + |\psi\rangle$. (Commutivity)
3. $|\psi\rangle + (|\xi\rangle + |\chi\rangle) = (|\psi\rangle + |\xi\rangle) + |\chi\rangle$. (Associativity)
4. There exists a unique null vector 0 in $S$ such that $0 + |\xi\rangle = |\xi\rangle$.
5. For each $|\xi\rangle$ in $S$ there exists an element $-|\xi\rangle$, such that $|\xi\rangle + (-|\xi\rangle) = |\xi\rangle - |\xi\rangle = 0$. (Additive inverse)

The field with respect to which the space is defined is an associated set $F$ of numbers (usually the set $R$ of real numbers or the set $C$ of complex numbers) which we may use to multiply the elements of the space itself. This operation involving multiplication of elements of $S$ by elements of $F$ is assumed to have the following properties:

1. For all vectors $|\psi\rangle$ in $S$ and scalars $\lambda$ in $F$ there exists a vector $|\chi\rangle$ in $S$ such that $|\chi\rangle = \lambda|\xi\rangle$. (Closure).

2. $\lambda[|\psi\rangle + |\xi\rangle] = \lambda|\xi\rangle + \lambda|\psi\rangle$ (Distribution of scalar multiplication over vector addition)

3. $\lambda_1|\psi\rangle + \lambda_2|\psi\rangle = (\lambda_1 + \lambda_2)|\psi\rangle$ (Distribution of scalar addition over multiplication of vectors by scalars.)

4. $\lambda_1(\lambda_2|\psi\rangle) = (\lambda_1\lambda_2)|\psi\rangle$ (Associativity of scalar multiplication.)

5. There exists an identity element **1** in $F$ such that $\mathbf{1}|\xi\rangle = |\xi\rangle$ for each $|\xi\rangle$ in $S$.

Some of these properties are actually associated with definitions of the term *field*. The reader is no doubt familiar with many examples of linear vector spaces, a common example being the set of displacement vectors $\vec{r}$ in $R^3$, which is a linear vector space with respect to the field $R$ of real numbers (i.e., it is a **real vector space**). Obviously, the notation that we are using (which is due to Dirac) to distinguish vectors $|\xi\rangle$ from scalars $\lambda$ (elements of the field) has a direct parallel in the common use of arrows, boldface symbols, etc. commonly used to distinguish vectors in $R^3$ from their scalar counterparts. Indeed, once one grows accustomed to this notation, it is not hard to imagine an introductory physics course making use of the Dirac notion to express the ubiquitous second law of Newton in the form $|F\rangle = m|a\rangle$. Other examples of vector spaces more relevant to quantum mechanics are listed below. In all cases of quantum mechanical interest the relevant field is the set $C$ of complex numbers, i.e., we are interested in **complex vector spaces**. The interested reader is encouraged to verify the properties listed above with regard to each of these spaces:

1. The set $C^N$ of column vectors of length $N$ having complex elements. The field is that of the complex numbers and vector addition is defined through the addition of components as it is in $R^3$. It is isomorphic to the set of complex row vectors having the same number of elements.

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix} + \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \begin{pmatrix} b_1 + c_1 \\ b_2 + c_2 \\ \vdots \\ b_N + c_N \end{pmatrix} \tag{2.1}$$

$$\lambda \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = \begin{pmatrix} \lambda c_1 \\ \lambda c_2 \\ \vdots \\ \lambda c_N \end{pmatrix} \tag{2.2}$$

2. The set $L^2(R^n)$ of complex-valued square-integrable functions $\psi(\vec{r})$ on $R^n$. Thus, the function $\psi(x)$ is in $L^2(R)$ if the integral

$$\int_{-\infty}^{\infty} \psi^* \psi \, dx = M, \tag{2.3}$$

exists and is non-infinite. Again the field is that of the complex numbers and vector addition is associated with the pointwise addition of functions, i.e., for any two functions $\psi_1(x)$ and $\psi_2(x)$ in this set we can produce the function

$$\psi(x) = \psi_1(x) + \psi_2(x) \tag{2.4}$$

which is also an element of the set. Note that all we are doing here is "adding by components", as in the example above, except that in this circumstance the components are labeled by a continuous index $x$, instead of a discrete one.

3. It is often useful to work in a larger space than that defined in example 2, namely a space which includes the plane waves and delta functions. We will loosely define this space to be the Hilbert space $F(R^n)$ of Fourier-transformable functions on $R^n$. This space includes, as a subset, the square-integrable functions of $L^2(R^n)$. It also includes the planes waves (whose Fourier transforms are the delta functions) and the delta functions (whose Fourier transforms are the plane waves), neither of which are themselves square-integrable. Since $L^2(R^n)$ is a subset of $F(R^n)$ and is, by itself, a linear vector space under the same operations of vector addition and scalar multiplication as $F(R^n)$, it is referred to as a **subspace** of the larger vector space.

### *2.1.2 Additional Definitions*

**Span** - A set of vectors $\{|\phi_i\rangle\}$ is said to *span* a vector space $S$ if every vector $|\psi\rangle$ in $S$ can be written as a linear combination

$$|\psi\rangle = \sum_i c_i |\phi_i\rangle \tag{2.5}$$

of the elements of the set. In this expression the $c_i$ are complex constants (i.e., elements of the field.). The vectors $\{\hat{\imath}, \hat{\jmath}, 2\hat{\imath}+\hat{\jmath}, \hat{k}\}$ span the set of vectors in $R^3$. There are obviously more than we need, however. This leads us to the next definition.

**Linear Independence** - A set of vectors $\{|\phi_i\rangle\}$ is *linearly independent* if no element in the set can be written as a linear combination of the other elements in the set. Otherwise the members of the set are linearly dependent.

**Comment:** This definition turns out to be equivalent to the statement that *if* the set $\{|\phi_i\rangle\}$ is linearly independent, then the only solution to the equation

$$\sum_i c_i |\phi_i\rangle = 0 \tag{2.6}$$

is $c_i \equiv 0$ for all $i$.

**Proof:** Assume that the set $\{|\phi_i\rangle\}$ are linearly independent, and assume that there existed a solution to the equation above with at least one coefficient $c_j$ not equal to zero. Then we could solve for the associated vector $|\phi_j\rangle$ as a function of the others, i.e., it would follow that

$$c_j|\phi_j\rangle + \sum_{i\neq j} c_i |\phi_i\rangle = 0 \tag{2.7}$$

and so

$$|\phi_j\rangle = -\sum_{i\neq j}\left(\frac{c_i}{c_j}\right)|\phi_i\rangle = \sum_{i\neq j} b_i|\phi_i\rangle \tag{2.8}$$

which violates the assumption of linear independence. Hence we conclude that there can be no such nonzero coefficient.

**Dimension** - If a linear vector space $S$ contains a subset of $N$ linearly independent vectors, but not a subset of $N+1$ linearly independent vectors it has dimension $N$, or is an $N$-dimensional space, If there exist subsets of $S$ with $N$ linearly independent vectors for all $N$ it is said to be infinite dimensional.

**Basis** - A linearly independent set of vectors $\{|\phi_i\rangle\}$ which spans a space forms a *basis* for the space.

A basis, in a certain sense, forms the minimal set of vectors that is necessary to form all other vectors in the space by linear combination. A basis set is often referred to as being **complete** with respect to the vector space that it spans. Note that by definition, any spanning set of vectors contains a basis as a subset. (We just have to weed out the unnecessary vectors which can themselves be written as linear combinations of the remaining ones.) Perhaps the most useful distinction between a basis set and a spanning set, is that the expansion coefficients expressing any element of the space in terms of the members of a *basis* are *unique*. That is, if the set of vectors $\{|\phi_i\rangle\}$ forms a basis for $S$, and $|\psi\rangle$ is an arbitrary element of $S$ it can be written as a linear combination

$$|\psi\rangle = \sum_i b_i|\phi_i\rangle \tag{2.9}$$

of the elements of the basis in one and only one way. This is easy to prove. Assume that there existed another expansion for this vector of the form

$$|\psi\rangle = \sum_i c_i|\phi_i\rangle. \tag{2.10}$$

The difference of these two expansions gives the null vector, which we can write as

$$0 = \sum_i (b_i - c_i)|\phi_i\rangle \tag{2.11}$$

Due to the assumed linear independence of the vectors $\{|\psi_i\rangle\}$ this can only be satisfied if the coefficients $(b_i - c_i)$ vanish. This implies that $c_i = b_i$ for all $i$ and hence the expansion is unique.

### *2.1.3 Continuous Bases and Continuous Sets*

It often arises that a set of vectors $\{|\phi_\alpha\rangle\}$ is labeled by a continuous index $\alpha$. Examples include the plane waves

$$\phi_k(x) = \frac{e^{ikx}}{\sqrt{2\pi}}; \quad k \in R \tag{2.12}$$

or

$$\phi_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}; \quad \vec{k} \in R^3 \tag{2.13}$$

We therefore extend our previous definitions developed for discrete sets of vectors as follows:

**Span** - A continuously indexed set of vectors $\{|\phi_\alpha\rangle\}$ is said to *span* a vector space $S$ if every vector $|\psi\rangle$ in $S$ can be written as a continuous linear combination

$$|\psi\rangle = \int d\alpha\, \zeta(\alpha)|\phi_\alpha\rangle \tag{2.14}$$

of the elements of the set. In this expression the function $\zeta(\alpha)$ gives the complex value of the expansion coefficient of $|\psi\rangle$ associated with the state $|\phi_\alpha\rangle$ of the spanning set.

**Linear Independence** - A continuously indexed set of vectors $\{|\phi_\alpha\rangle\}$ is *linearly independent* if the only solution to the equation

$$\int d\alpha\, C(\alpha)|\phi_\alpha\rangle = 0 \tag{2.15}$$

is that the function $C(\alpha) \equiv 0$ for all $\alpha$.

**Basis** - A linearly independent set of continuously indexed vectors $\{|\phi_\alpha\rangle\}$ that spans a linear vector space is said to form a basis for the space. In this case, the space is necessarily infinite dimensional, since it clearly contains an infinite number of vectors in any region where the index $\alpha$ takes on values.

### *2.1.4 Inner Products*

Another important property associated with the linear vector spaces of quantum mechanics is that they are inner product spaces. A linear vector space $S$ is an **inner product space** if there exists an assignment to each pair of vectors $|\phi\rangle$ and $|\psi\rangle$ in $S$, a complex number (i.e., element of the field), referred to as the **inner product** of $|\phi\rangle$ and $|\psi\rangle$ and denoted by the symbol $\langle\phi|\psi\rangle$, obeying the following properties:

1. $\langle\phi|\phi\rangle$ is real and non-negative, i.e., $\langle\phi|\phi\rangle \geq 0$. Moreover, $\langle\phi|\phi\rangle = 0$, if and only if $|\phi\rangle$ is the null vector.
2. $\langle\phi|[|\psi_1\rangle + |\psi_2\rangle] = \langle\phi|\psi_1\rangle + \langle\phi|\psi_2\rangle$. Thus, the inner product distributes itself over vector addition.
3. $\langle\phi|[\lambda|\psi\rangle] = \lambda\langle\phi|\psi\rangle$
4. $\langle\phi|\psi\rangle = (\langle\psi|\phi\rangle)^*$. Thus the order of the inner product is important for complex vector spaces.

**Comment:** In complex vector spaces, the inner product $\langle\phi|\psi\rangle$ is linear in $|\psi\rangle$, but antilinear or *symplectic* in $|\phi\rangle$. The first half of this comment stems from the observation that

$$\langle\phi|[\lambda_1|\psi_1\rangle + \lambda_2|\psi_2\rangle] = \lambda_1\langle\phi|\psi_1\rangle + \lambda_2\langle\phi|\psi_2\rangle \tag{2.16}$$

which follows from (2) and (3), while the second stems from the fact that if $|\phi\rangle = \lambda_1|\phi_1\rangle + \lambda_2|\phi_2\rangle$ then the inner product of $|\phi\rangle$ with $|\psi\rangle$ can be written

$$\langle\phi|\psi\rangle = \lambda_1^*\langle\phi_1|\psi\rangle + \lambda_2^*\langle\phi_2|\psi\rangle. \tag{2.17}$$

This follows straightforwardly from (2), (3), and (4). The proof is left as an exercise. It is convenient, and not entirely wrong, to *think* of each vector $|\psi\rangle$ as a column vector containing elements $\psi_i$, and to think of $\langle\phi|$ as a row vector whose elements are the complex conjugates $\phi_i^*$ of the those associated with the column vector representing $|\phi\rangle$. As we will see, such a *representation* for the kets and bras is usually possible (but not essential). In this way the inner product can be viewed as the "dot product"

$$\begin{pmatrix} \phi_1^* & \phi_2^* & \cdots & \phi_N^* \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_N \end{pmatrix} = \sum_i \phi_i^*\psi_i \tag{2.18}$$

This is, of course, the inner product commonly associated with $C^N$. The complex conjugated row vectors associated with the symbols $\{\langle\phi|\}$ thus form a vector space $S^*$ of their own which is isomorphic (or **dual** or **adjoint**) to the original space $S$ having elements $\{|\phi\rangle\}$. The vectors of $S^*$ are in 1-1 correspondence with the vectors of $S$. In the Dirac notation, an element $|\phi\rangle$ of $S$ is referred to as a **ket**, while an elements $\langle\phi|$ of $S^*$ is referred to as a **bra**. The combination $\langle\phi|\psi\rangle$ forms a "bracket", which in the Dirac formalism is always a number.

**Examples:**

1. In the space of displacement vectors $\vec{r}$ in $R^3$ the inner product is just the familiar "dot product".
2. As discussed above the inner product in $C^N$ is obtained by "dotting" a complex conjugated row vector into an unconjugated column vector.

3. In functional spaces, the inner product usually involves the continuous analog of a summation over components, namely an integral. Thus, in the space of Fourier transformable function on $R^3$ we "associate" with each function $\psi(\vec{r})$ a vector $|\psi\rangle$. The inner product of two functions then takes the form

$$\langle\psi|\phi\rangle = \int d^3r\, \psi^*(\vec{r})\phi(\vec{r}) \tag{2.19}$$

where the integral is over all space.

The concept of an inner product allows us to make several new definitions:

**Norm** - The positive-definite real quantity $||\psi|| \equiv \sqrt{\langle\psi|\psi\rangle}$ is referred to as the *norm*, or the *length* of the vector $|\psi\rangle$. A vector $|\psi\rangle$ is said to be *square-normalized*, have *unit norm*, or be a *unit vector* if $||\psi|| = 1$. Any vector having finite norm can be normalized. That is, if $||\psi||$ is not infinite, then the vector

$$|\phi\rangle = \frac{|\psi\rangle}{||\psi||} = \frac{|\psi\rangle}{\sqrt{\langle\psi|\psi\rangle}} \tag{2.20}$$

is a unit vector along the same direction in the space as $|\psi\rangle$

**Orthogonality** - Two vectors $|\psi\rangle$ and $|\phi\rangle$ are *orthogonal* if

$$\langle\psi|\phi\rangle = \langle\phi|\psi\rangle = 0, \tag{2.21}$$

i.e., if their inner product vanishes. We loosely say that the vectors have no overlap, or that $|\psi\rangle$ has no component along $|\phi\rangle$ and vice versa.

**Orthonormal Set**

1. A *discrete* set of vectors $\{|\phi_i\rangle\}$ forms an orthonormal set if

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}, \tag{2.22}$$

that is, if they are a set of unit-normalized vectors which are mutually orthogonal.

2. A *continuously-indexed* set of vectors $\{|\phi_\alpha\rangle\}$ forms an orthonormal set if they obey the Dirac normalization condition

$$\langle\phi_\alpha|\phi_{\alpha'}\rangle = \delta(\alpha - \alpha'). \tag{2.23}$$

Note that the members of such a set have *infinite* norm, and are *not* square-normalizable.

**Orthonormal Basis** - An orthonormal set of linearly independent vectors which spans a space is referred to as an *orthonormal basis* (which we will abbreviate as ONB).

It is straightforward to show that any set of mutually orthogonal vectors not containing the null vector is linearly independent. Thus, any orthonormal set of vectors which span a space also forms an orthonormal basis for the space. It is also possible to show that the number of basis vectors in any orthonormal basis is equal to the dimension of the vector space. To do this we first show that from any set of $N$ linearly independent vectors it is possible to construct a set of $N$ mutually orthogonal vectors, and thus, generally speaking, an orthonormal set of such vectors. The following explicit algorithm for doing this is referred to as the **Gram-Schmidt orthogonalization** procedure.

Let $\{|\chi_1\rangle, |\chi_2\rangle, \cdots, |\chi_N\rangle\}$ be a set of linearly independent vectors of finite length. Set

$$|\psi_1\rangle = |\chi_1\rangle, \qquad |\phi_1\rangle = \frac{|\psi_1\rangle}{||\psi_1||}, \tag{2.24}$$

to produce a unit normalized vector $|\phi_1\rangle$ pointing along the same direction as $|\chi_1\rangle$. Now construct the second vector to be orthogonal, by subtracting off that part of it which lies along the direction of the first vector: set

$$|\psi_2\rangle = |\chi_2\rangle - |\phi_1\rangle\langle\phi_1|\chi_2\rangle, \qquad |\phi_2\rangle = \frac{|\psi_2\rangle}{||\psi_2||}. \tag{2.25}$$

Note, that by construction

$$\langle\phi_1|\psi_2\rangle = \langle\phi_1|\chi_2\rangle - \langle\phi_1|\phi_1\rangle\langle\phi_1|\chi_2\rangle = 0, \tag{2.26}$$

so that $|\psi_2\rangle$ and $|\phi_2\rangle$ are orthogonal to $|\phi_1\rangle$. We now proceed in this fashion, constructing each vector orthogonal to those previously constructed. Thus, we set

$$|\psi_3\rangle = |\chi_3\rangle - \sum_{i=1}^{2} |\phi_i\rangle\langle\phi_i|\chi_3\rangle, \qquad |\phi_3\rangle = \frac{|\psi_3\rangle}{||\psi_3||} \tag{2.27}$$

and, more generally, the $n$th such vector takes the form

$$|\phi_n\rangle = |\chi_n\rangle - \sum_{i=1}^{n-1} |\phi_i\rangle\langle\phi_i|\chi_n\rangle, \qquad |\phi_n\rangle = \frac{|\psi_n\rangle}{||\psi_n||}. \tag{2.28}$$

The only way this process could stop is if one of the resulting vectors $|\phi_n\rangle$ turned out to be the null vector. A close inspection of the process reveals that this can't happen if the original set is linearly independent, as we have assumed. Thus, in this way we construct an orthonormal set of vectors $\{|\phi_n\rangle\}$ equal in number to the original. It follows, that given any basis for the space we can construct an orthonormal basis with an equal number of vectors. Thus, for a finite dimensional space there exists at least one ONB with the same number of members as the dimension of the space. It turns out that there can't exist any bases with any fewer members, because then we could ultimately end up solving for one of the members of the larger set in terms of the remaining members, which would contradict their linear independence. The proof is left as an exercise. For now, we will simply observe that orthonormal bases are extremely useful due to the ease with which they allow arbitrary vectors to be expressed. We explore this below.

### *2.1.5 Expansion of a Vector on an Orthonormal Basis*

**Discrete Bases** - Let the set $\{|\phi_i\rangle\}$ form an orthonormal basis (or ONB) for the space $S$, so that $\langle\phi_i|\phi_j\rangle = \delta_{ij}$,and let $|\chi\rangle$ be an arbitrary element of the space. By assumption there exists an expansion of the form

$$|\chi\rangle = \sum_i \chi_i \, |\phi_i\rangle \tag{2.29}$$

for a unique set of expansion coefficients $\chi_i$. How do we determine what these expansion coefficients are? Consider the inner product

$$\langle\phi_j|\chi\rangle = \sum_i \chi_i \, \langle\phi_j|\phi_i\rangle = \sum_i \chi_i \, \delta_{ij} = \chi_j \tag{2.30}$$

of the vector $|\chi\rangle$ with an arbitrary element $|\phi_j\rangle$ of this basis. This shows that the expansion coefficient $\chi_j$ is just the inner product of the vector of interest with the unit vector along that direction in Hilbert space. Thus $\chi_i = \langle\phi_i|\chi\rangle$. Since the order in which we write the product of a number and a vector is unimportant, we will often write this in the form

$$|\chi\rangle = \sum_i |\phi_i\rangle \, \chi_i = \sum_i |\phi_i\rangle \, \langle\phi_i|\chi\rangle \tag{2.31}$$

for reasons which will become clearer later on.

**Extension to Continuous Bases** - Let the set $\{|\phi_\alpha\rangle\}$ form a continuous orthonormal basis for the space $S$, so that

$$\langle\phi_\alpha|\phi_{\alpha'}\rangle = \delta(\alpha - \alpha'), \tag{2.32}$$

and let $|\chi\rangle$ be an arbitrary element of the space. By assumption there exists an expansion

$$|\chi\rangle = \int d\alpha\ \chi(\alpha)|\phi_\alpha\rangle \tag{2.33}$$

for some unique expansion function $\chi(\alpha)$. How do we determine what this expansion function is? Consider the inner product

$$\langle\phi_{\alpha'}|\chi\rangle = \langle\phi_{\alpha'}|\left[\int d\alpha\ \chi(\alpha)|\phi_\alpha\rangle\right] = \int d\alpha\ \chi(\alpha)\langle\phi_{\alpha'}|\phi_\alpha\rangle = \int d\alpha\ \chi(\alpha)\delta(\alpha-\alpha') = \chi(\alpha') \tag{2.34}$$

of the vector $|\chi\rangle$ with an arbitrary element $|\phi_{\alpha'}\rangle$ of the basis. This shows that, as in the discrete case, the expansion coefficient $\chi(\alpha')$ is just the inner product of the vector of interest with the basis vector along that direction in Hilbert space. Thus $\chi(\alpha) = \langle\phi_\alpha|\chi\rangle$. We will refer to the function $\chi(\alpha)$ as the "wave function" representing $|\chi\rangle$ in the $\alpha$-basis or $\alpha$ representation. Again, since the order in which we write the product of a number and a vector is unimportant, we will write this often in the form

$$|\chi\rangle = \int d\alpha\ |\phi_\alpha\rangle\chi(\alpha) = \int d\alpha\ |\phi_\alpha\rangle\langle\phi_\alpha|\chi\rangle. \tag{2.35}$$

**Comment:** It is clear that when we talk about ONB's, such as $\{|\phi_i\rangle\}$ or $\{|\phi_\alpha\rangle\}$, the important information appearing inside the ket which distinguishes the different basis vectors from one another is the label or index: $i$ or $j$ in the discrete case, $\alpha$ or $\alpha'$ in the continuous case. The symbols $\phi$ just sort of come along for the ride. From this point on we will acknowledge this by using the abbreviated notations

$$|i\rangle = |\phi_i\rangle \tag{2.36}$$

and

$$|\alpha\rangle = |\phi_\alpha\rangle. \tag{2.37}$$

In this way the expansion of an arbitrary ket can be written

$$|\chi\rangle = \sum_i \chi_i\,|i\rangle = \sum_i |i\rangle\,\langle i|\chi\rangle \tag{2.38}$$

and

$$|\chi\rangle = \int d\alpha\ \chi(\alpha)\,|\alpha\rangle = \int d\alpha\ |\alpha\rangle\langle\alpha|\chi\rangle. \tag{2.39}$$

### *2.1.6 Calculation of Inner Products Using an Orthonormal Basis*

**Discrete Bases** - Let the set $\{|i\rangle\}$ form an orthonormal basis (or ONB) for the space $S$, so that $\langle i|j\rangle = \delta_{ij}$,and let $|\chi\rangle, |\psi\rangle$ be arbitrary elements of the space. By assumption there exists expansions of the form

$$|\chi\rangle = \sum_i \chi_i\,|i\rangle = \sum_i |i\rangle\langle i|\chi\rangle \tag{2.40}$$

$$|\psi\rangle = \sum_i \psi_i \, |i\rangle = \sum_i |i\rangle\langle i|\psi\rangle. \tag{2.41}$$

The inner product of these two vectors can be written

$$\langle\psi|\chi\rangle = \langle\psi| \left\{ \sum_i \chi_i \, |i\rangle \right\} = \sum_i \chi_i \, \langle\psi|i\rangle. \tag{2.42}$$

But $\langle\psi|i\rangle = (\langle i|\psi\rangle)^* = \psi_i^*$. Thus, we can write the inner product in the form

$$\langle\psi|\chi\rangle = \sum_i \psi_i^* \chi_i \, . \tag{2.43}$$

But this is the form obtained by taking the "dot product" of a complex-conjugated row vector and a column vector. This justifies our earlier loose association of kets and bras with column and row vectors, and, in fact, makes clear the conditions under which such a picture is justified, i.e.,

> Any discrete ONB for a space $S$ *induces* (or generates or defines) a **row-vector/column vector representation** for the space, i.e., it gives us a natural way of associating each abstract ket $|\psi\rangle$ in $S$ with a complex-valued column vector having components $\psi_i$, and each bra $\langle\psi|$ in $S^*$ with a complex-valued row vector having components $\psi_i^*$.

**Extension to Continuous Bases** - Let the set $\{|\alpha\rangle\}$ form a continuous orthonormal basis for the space $S$, so that

$$\langle\alpha|\alpha'\rangle = \delta(\alpha - \alpha'), \tag{2.44}$$

and let $|\chi\rangle$,$|\psi\rangle$ be arbitrary elements of the space. By assumption there exist expansions of the form

$$|\chi\rangle = \int d\alpha \; \chi(\alpha)|\alpha\rangle = \int d\alpha \; |\alpha\rangle\langle\alpha|\chi\rangle \tag{2.45}$$

$$|\psi\rangle = \int d\alpha \; \psi(\alpha)|\alpha\rangle = \int d\alpha \; |\alpha\rangle\langle\alpha|\psi\rangle. \tag{2.46}$$

The inner product of these two vectors can be written

$$\langle\psi|\chi\rangle = \langle\psi| \left\{ \int d\alpha \; \chi(\alpha)|\alpha\rangle \right\} = \int d\alpha \; \chi(\alpha)\langle\psi|\alpha\rangle. \tag{2.47}$$

But $\langle\psi|\alpha\rangle = (\langle\alpha|\psi\rangle)^* = \psi^*(\alpha)$. Thus, we can write the inner product in the form

$$\langle\psi|\chi\rangle = \int d\alpha \; \chi(\alpha)\psi^*(\alpha) = \int d\alpha \; \psi^*(\alpha)\chi(\alpha). \tag{2.48}$$

This is also interesting, because it looks just like the inner product which appears in "functional" linear vector spaces, such as the set of Fourier transformable functions on $R^3$. This suggests the following important point:

> Any continuous ONB for a space $S$ induces a **wave function representation** for the space, i.e., it gives us a natural mapping of each abstract vector $|\psi\rangle$ in $S$ onto a complex valued wave function $\psi(\alpha)$, which give the expansion coefficients for the state in that continuously-indexed basis. Similarly, it maps each vector $\langle\psi|$ in $S^*$ onto a complex-valued wave function $\psi^*(\alpha)$. We speak, therefore, of $\psi(\alpha)$ as the wave function for the state $|\psi\rangle$ in the $\{|\alpha\rangle\}$ representation.

This similarity to what we see with the functional spaces works in the opposite direction as well. That is, the form of the inner product associated with functions on $R^3$

$$\langle\psi|\phi\rangle = \int d^3r\,\psi^*(\vec{r})\phi(\vec{r}) \tag{2.49}$$

suggests that the wave function $\psi(\vec{r})$ can be viewed as "merely" the function giving the expansion coefficients in *some* orthonormal basis of states $\{|\vec{r}\rangle\}$ labeled by the (continuously distributed) position vectors $\vec{r}$ in $R^3$. This leads to something called the position representation.

### *2.1.7 The Position Representation*

In the state space of a single quantum mechanical particle moving in three dimensions we wish to make an association

$$\psi(\vec{r}) \longleftrightarrow |\psi\rangle \tag{2.50}$$

of each function $\psi(\vec{r})$ with an underlying abstract vector in the space $|\psi\rangle$. We define this association more precisely by considering a *particular* set of functions, namely, the Dirac delta functions $\{\delta(\vec{r}-\vec{r}\,')\}$. For each point $\vec{r}\,'$ in space, there is a delta function centered at that point, which quantum mechanically would correspond to a particle which has all of its probability density located at $\vec{r}\,'$. Thus, we have an entire set of functions labeled by the points in $R^3$. With each of these position-localized wave functions we associate a state $|\vec{r}\,'\rangle$ of the state space. Thus we have the association

$$|\vec{r}\,'\rangle \longleftrightarrow \phi_{\vec{r}\,'}(\vec{r}) \equiv \delta(\vec{r}-\vec{r}\,') \tag{2.51}$$

where we have introduced a notation which suggests that, in this context, the delta function is to be considered a *function* of $\vec{r}$, which happens to be *labeled* by the point $\vec{r}\,'$. We claim that this set of kets $\{|\vec{r}\,'\rangle\}$ forms a continuous orthonormal basis for the underlying space. This is intuitively reasonable insofar as a particle located at $\vec{r}$ is incompatible with it being located at any other point. For the moment we will simply assume that this is true and see where it leads us. The assumed orthonormality of the states $\{|\vec{r}\,'\rangle\}$ leads us to the orthonormality relation, as defined for continuously indexed states,

$$\langle\vec{r}|\vec{r}\,'\rangle = \delta(\vec{r}-\vec{r}\,'). \tag{2.52}$$

In addition, the assumed completeness of the set implies that any other state in the space $|\psi\rangle$ must be expandable in this basis. Thus we can write

$$|\psi\rangle = \int d^3r\,'\,\psi(\vec{r}\,')|\vec{r}\,'\rangle \tag{2.53}$$

where we can write $\psi(\vec{r}\,') = \langle\vec{r}\,'|\psi\rangle$, in analogy to what we did with the continuous basis $|\alpha\rangle$. Since the integration variable is just a dummy, we can drop the prime and write this

$$|\psi\rangle = \int d^3r\;\psi(\vec{r})|\vec{r}\rangle. \tag{2.54}$$

It is now natural to assume that the function $\psi(\vec{r}) = \langle\vec{r}|\psi\rangle$, which here just gives the expansion coefficients for $|\psi\rangle$ in the $|\vec{r}\rangle$ basis, is precisely the wave function $\psi(\vec{r})$ that we wanted to associate with the state $|\psi\rangle$ to begin with. This turns out to be a consistent interpretation. In, particular, it predicts that the "wave function" associated with one of the basis states $|\vec{r}\,'\rangle$ should be given by the expression

$$\phi_{\vec{r}\,'}(\vec{r}) = \langle\vec{r}|\vec{r}\,'\rangle = \delta(\vec{r}-\vec{r}\,') \tag{2.55}$$

which *is* consistent with our original association of the state $\vec{r}\,'$ having all of its probability density located at $\vec{r}\,'$. Thus, the orthonormality relation for the states $\{|\vec{r}\,'\rangle\}$ also gives the form that their wave functions take in this representation, which we refer to as the *position representation*, or the $\vec{r}$ representation.

### *2.1.8 The Wavevector Representation*

Recall, that the set of plane waves

$$\phi_{\vec{k}}(\vec{r}) = \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}} \tag{2.56}$$

for all wavevectors $\vec{k}$ forms a complete set of functions for the space of transformable functions. We use this to define a set of underlying kets $\{|\vec{k}\rangle\}$ which are the states *represented* by the **plane** waves in the position representation just introduced. Thus, by definition,

$$|\vec{k}\rangle = \int d^3r\ \phi_{\vec{k}}(\vec{r})|\vec{r}\rangle = \int \frac{d^3r}{(2\pi)^{3/2}}\ e^{i\vec{k}\cdot\vec{r}}|\vec{r}\rangle \tag{2.57}$$

These states are the Fourier transforms, in this sense, of the position localized states $|\vec{r}\rangle$. If the states $\{|\vec{r}\rangle\}$ are complete and orthonormal, then it is straightforward to show that the states $\left\{|\vec{k}\rangle\right\}$ are as well. Orthonormality of the kets $|\vec{k}\rangle$ follows from the orthonormality the plane waves, i.e.,

$$\langle\vec{k}|\vec{k}'\rangle = \int d^3r\ \phi^*_{\vec{k}}(\vec{r})\phi_{\vec{k}'}(\vec{r}) = \int \frac{d^3r}{2\pi}\ e^{i(\vec{k}-\vec{k}')\cdot\vec{r}} = \delta(k-k'). \tag{2.58}$$

These states are also complete, since they can be used to expand an arbitrary member of the $\{|\vec{r}\rangle$ basis, and thereby be used to expand an arbitrary state $|\psi\rangle$ of the system. To determine this expansion we note that if, by assumption,

$$|\vec{r}\rangle = \int d^3k\ \phi_{\vec{r}}(\vec{k})|\vec{k}\rangle \tag{2.59}$$

then it must be that

$$\phi_{\vec{r}}(\vec{k}) = \langle\vec{k}|\vec{r}\rangle = (\langle\vec{r}|\vec{k}\rangle)^* = \phi^*_{\vec{k}}(r) = \frac{e^{-i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}} \tag{2.60}$$

Thus, we have the proposed expansion,

$$|\vec{r}\rangle = \int \frac{d^3k}{(2\pi)^{3/2}}\ e^{-i\vec{k}\cdot\vec{r}}|\vec{k}\rangle, \tag{2.61}$$

which is the counterpart to the definition

$$|\vec{k}\rangle = \int \frac{d^3r}{(2\pi)^{3/2}}\ e^{i\vec{k}\cdot\vec{r}}|\vec{r}\rangle \tag{2.62}$$

which is verified to be correct, by expressing the states $|\vec{k}\rangle$ in terms of the states $|\vec{r}\,'\rangle$, and collapsing the delta functions which develop. Having established the fact that the states $\{|\vec{k}\rangle\}$ also form an ONB for this space we can now expand arbitrary states

$$|\psi\rangle = \int d^3k\ \psi(\vec{k})|\vec{k}\rangle. \tag{2.63}$$

in this basis. This leads to interesting relations. For example, in the real space representation we have that

$$\psi(\vec{r}) = \langle \vec{r}|\psi\rangle = \int d^3k\ \psi(\vec{k})\langle \vec{r}|\vec{k}\rangle = \int \frac{d^3k}{(2\pi)^{3/2}}\ \psi(\vec{k})\, e^{i\vec{k}\cdot\vec{r}}, \tag{2.64}$$

which shows that $\psi(\vec{r})$ is the Fourier transform of $\psi(\vec{k})$. Similarly, we find that

$$\psi(\vec{k}) = \langle \vec{k}|\psi\rangle = \int d^3r\ \psi(\vec{r})\langle \vec{k}|\vec{r}\rangle = \int \frac{d^3r}{(2\pi)^{3/2}}\ \psi(r)\, e^{-i\vec{k}\cdot\vec{r}}, \tag{2.65}$$

so that $\psi(\vec{r})$ and $\psi(\vec{k})$ are Fourier transform pairs. Note where the plus and minus signs go in the exponentials in (2.64) and (2.65), particularly in comparison to (2.61) and (2.62). Also note a subtle shift in notation. In the last chapter we denoted the wavefunction by $\psi(\vec{r})$ and its Fourier transform by $\hat{\psi}(\vec{k})$. We now know, however, that there are as many possible wavefunctions representing the state $|\psi\rangle$ as there are continuous orthonormal bases for the space. In fact, we will see that if there exists *one* such continuous basis there will always be an infinite number of other ones that can be constructed. Thus, rather than coming up with a different diacritical mark $\psi, \hat{\psi}, \tilde{\psi}$, etc. to differentiate the wavefunction in each new representation that we introduce, we will agree to always include the argument of the wavefunction to indicate which representation we are working in at the moment. Thus, it will be understood that $\psi(\vec{r}), \psi(\vec{k})$,and, e.g., $\psi(\alpha)$ all represent *different* functions of their arguments, even though we use the same symbol $\psi$ for each, to indicate that they all provide a means of representing the same underlying state vector $|\psi\rangle$. We note in passing that we could avoid this problem all together by simply agreeing to use our identification of the expansion coefficient, or wave function, with the associated inner product. Thus, instead of writing $\psi(\vec{r}), \psi(\vec{k})$,and, $\psi(\alpha)$, we could always just write $\langle \vec{r}|\psi\rangle$, $\langle \vec{k}|\psi\rangle$, and $\langle \alpha|\psi\rangle$. In keeping with modern usage, we will use both interchangeably.

All of these ideas can be applied to a particle moving in lower dimensions, as well. In the space of a particle moving in one dimension we introduce a set of position localized states $\{|x\rangle\}$ labeled by the positions where the particle can be localized. We then have the following relations similar to those developed above

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

$$|\psi\rangle = \int dx\ \psi(x)|x\rangle$$

$$\psi(x) = \langle x|\psi\rangle$$

and a set of plane wave states $\{|k\rangle\}$ labeled by wave vector, which obey the following relations

$$\langle k|k'\rangle = \delta(k - k')$$

$$|\psi\rangle = \int dk\ \psi(k)|k\rangle$$

$$\psi(k) = \langle k|\psi\rangle$$

and which are related to the position localized states through the following relations

$$|k\rangle = \int dx\ |x\rangle\langle x|k\rangle = \int \frac{dx}{(2\pi)^{1/2}}\ e^{ikx}|x\rangle$$

$$|x\rangle = \int dk\ |k\rangle\langle k|x\rangle = \int \frac{dk}{(2\pi)^{1/2}}\ e^{-ikx}|k\rangle$$

$$\langle x|k\rangle = \frac{e^{ikx}}{(2\pi)^{1/2}} = \langle k|x\rangle^*$$

$$\psi(x) = \langle x|\psi\rangle = \int dk\ \psi(k)\langle x|k\rangle = \int \frac{dk}{(2\pi)^{1/2}}\ \psi(k)\, e^{ikx},$$

$$\psi(k) = \langle k|\psi\rangle = \int dx\ \psi(x)\langle k|x\rangle = \int \frac{dx}{(2\pi)^{1/2}}\ \psi(x)\, e^{-ikx}.$$

## 2.2 Postulate II: Observables of Quantum Mechanical Systems

In keeping with the general scheme introduced earlier, the second of the postulates that we will explore describes the nature of the observables of quantum mechanical systems.

> Every observable $\mathcal{A}$ of a quantum mechanical system is associated with a linear Hermitian operator $A$ whose eigenstates form a complete orthonormal basis for the quantum mechanical state space.

We are led, therefore, to investigate the nature of linear operators defined on linear vector spaces.

### *2.2.1 Operators and Their Properties*

An **operator** $A$ associated with a linear vector space $S$ acts on the elements $|\chi\rangle$ in $S$ and maps them onto (possibly) other elements $|\chi_A\rangle$ of the same space. We express this mapping of one vector onto another in the form

$$A|\chi\rangle = |\chi_A\rangle. \tag{2.66}$$

An operator $A$ is **linear** if it satisfies the following **linearity condition**

$$A(\lambda|\chi\rangle + \mu|\psi\rangle) = \lambda|\chi_A\rangle + \mu|\psi_A\rangle = \lambda A|\chi\rangle + \mu A|\psi\rangle, \tag{2.67}$$

for arbitrary states $|\chi\rangle, |\psi\rangle$, and arbitrary scalars $\lambda$ and $\mu$. In what follows we assume, unless otherwise stated, that all operators under consideration are linear. One of the useful properties of linear operators is that their action on arbitrary states is determined once their action on the elements of any ONB is specified. To see this, let $\{|i\rangle\}$ be an arbitrary ONB, and let the action

$$|\phi_i\rangle = A|i\rangle \tag{2.68}$$

of the linear operator $A$ on these states be known. When an arbitrary state $|\psi\rangle$ is acted upon by $A$ we can use the expansion of $|\psi\rangle$ in this basis to see that

$$A|\psi\rangle = A\sum_i \psi_i|i\rangle = \sum_i \psi_i\, A|i\rangle = \sum_i \psi_i|\phi_i\rangle \tag{2.69}$$

which uniquely determines the resulting vector. We describe below some of the common properties associated with linear operators.

The **sum and difference of operators** are defined through vector addition

$$(A+B)|\psi\rangle = A|\psi\rangle + B|\psi\rangle = |\psi_A\rangle + |\psi_B\rangle \tag{2.70}$$

$$(A - B)|\psi\rangle = A|\psi\rangle - B|\psi\rangle = |\psi_A\rangle - |\psi_B\rangle. \tag{2.71}$$

The **product of operators** is defined through the combined action of each. If $C = AB$, then

$$C|\psi\rangle = AB|\psi\rangle = A|\psi_B\rangle. \tag{2.72}$$

In general the operator product is not commutative, since reversing the order can give a different result, i.e., the vector

$$BA|\psi\rangle = B|\psi_A\rangle \tag{2.73}$$

need not have any relation to the vector $A|\psi_B\rangle$. It is useful, therefore, to define the **commutator** of two operators

$$[A, B] = AB - BA = -[B, A] \tag{2.74}$$

which is also an operator. If $[A, B] = 0$, then $AB = BA$, and the two operators commute. From the definition of the commutator it is straightforward to prove the following useful relations

$$\begin{gathered}
[A, A] = 0 \\
[A, B + C] = [A, B] + [A, C] \\
[A + B, C] = [A, C] + [B, C] \\
[A, BC] = B[A, C] + [A, B]C \\
[AB, C] = A[B, C] + [A, C]B \\
[A[B, C]] + [C[A, B]] + [B[C, A]] = 0.
\end{gathered}$$

The **null operator**, maps each vector in the space onto the null vector, i.e., $0|\psi\rangle = 0$.

The **identity operator**, maps each vector in the space onto itself, i.e., $\mathbf{1}|\psi\rangle = |\psi\rangle$.

The **inverse of an operator** $A$, if it exists, is denoted $A^{-1}$ and obeys the property

$$AA^{-1} = A^{-1}A = \mathbf{1}. \tag{2.75}$$

A nonzero vector $|\chi\rangle$ is said to be an **eigenvector** of an operator $A$ with eigenvalue $a$ (where generally, $a \in C$) if it satisfies the **eigenvalue equation**

$$A|\chi\rangle = a|\chi\rangle. \tag{2.76}$$

The set of eigenvalues $\{a\}$ for which solutions to this equation exist is referred to as the **spectrum** of $A$, and denoted spectrum($A$).

It is also possible to define operators that are, themselves, functions of other operators. This can be done in a number of ways. For example, from the product rule given above, it is clear that in general the $n$-fold product of an operator $A$ with itself is well-defined. Thus, we may always speak of positive integer powers $A^n$ of an operator. If the inverse $A^{-1}$ of an operator is also defined, then we can define negative powers through the relation $A^{-n} = \left(A^{-1}\right)^n$. Then, if $f(x) = \sum_n f_n x_n$ is any power series expandable function with a suitable radius of convergence, we can define the operator valued function $F(A) = \sum_n f_n A_n$ of the operator $A$. We will ignore for the moment a discussion of the conditions under which such series converge, because, as we will see, there are other ways of defining operator valued functions that are often more useful that allow this question to be avoided.

We now consider a variety of different operators. We begin by noting that if we multiply any state $|\psi\rangle$ in the space by a scalar $\lambda$ we generate a new vector $|\psi_\lambda\rangle = \lambda|\psi\rangle$. Thus, we can define a very simple type of operator which carries out this operation. To avoid any cumbersome notation we simply will denote by $\lambda$ that operator which multiplies a vector by the scalar $\lambda$. This allows us, e.g., to form operators of the form $\lambda + A$, where $A$ is an arbitrary operator and $\lambda$ a scalar, whose action is the obvious one, i.e., $(\lambda + A)\,|\psi\rangle = \lambda|\psi\rangle + A|\psi\rangle$. As special cases of operators formed from scalars, we have $\lambda = 1$ corresponding to the identity operator and $\lambda = 0$ corresponding to the null operator. Multiplication by a scalar is an operation that always commutes with any other linear operator, i.e., scalars always commute with everything. Thus, we can write $[\lambda, A] = 0$.

### *2.2.2 Multiplicative Operators*

In the space of a quantum particle moving in one dimension let us introduce an operator $X$ by defining its action on the one-dimensional position states $\{|x\rangle\}$ as follows:

$$X|x\rangle = x|x\rangle. \tag{2.77}$$

Thus, $X$ just multiplies the basis vector $|x\rangle$ by its label, i.e., by the point where the associated delta function is centered. (Note that this implies that the basis states $|x\rangle$ are all eigenstates of the operator $X$, and are in fact labeled by their associated eigenvalues). With this definition, we find that the action of $X$ on an arbitrary state $|\psi\rangle$ is rather simply expressed in the position representation, i.e.,

$$X|\psi\rangle = X\int dx\,\psi(x)|x\rangle = \int dx\,\psi(x)X|x\rangle = \int dx\,\psi(x)x|x\rangle. \tag{2.78}$$

Rearranging a little, we see that

$$X|\psi\rangle = \int dx\,[x\psi(x)]|x\rangle. \tag{2.79}$$

This shows that the wave function representing $X|\psi\rangle$ is just $x\psi(x)$. When we don't mind being a little imprecise we will say that "in the $x$-representation, $X\psi(x) = x\psi(x)$." or "$X$ multiplies the wave function by $x$". We have to be careful, though: in reality the operator $X$ does not actually act on the wave function (which is just a scalar-valued function), it acts on the kets in the expansion for the state, giving rise to this apparent effect.

This idea is easily extended to functions of $x$. For any function $f(x)$ we can define an operator $F$ which has the action

$$F|x\rangle = f(x)|x\rangle \tag{2.80}$$

of multiplying each basis vector $|x\rangle$ in the position representation by the function $f$ evaluated at the point $x$ labeling the basis vector. When $F$ acts on arbitrary vectors it leads to the result

$$F|\psi\rangle = \int dx\,\psi(x)\,F|x\rangle = \int dx\,\psi(x)f(x)|x\rangle \tag{2.81}$$

so that the wave function representing $F|\psi\rangle$ is just $f(x)\psi(x)$. "In the $x$ representation, $F$ multiplies the wave function by $f(x)$". The potential energy function $V(x)$ is associated with an operator of this type. In fact, all we are doing here is providing another way of defining a function of an operator, i.e., we can formally view $F$ as a function of the operator $X$, i.e., $F = F(X)$. In a representation in which $X$ just multiplies by $x$ the action of $F(X)$ is to multiply by $f(x)$. This definition is easily verified to agree with our former one regarding series expansions in those situations where the series converges, and extends it to allow for more general kinds of operator functions.

These ideas also extend to higher dimensions. We can define, e.g., the Cartesian operator components $X, Y$, and $Z$, of the **position operator** $\vec{R}$, through their action on the basis vectors $\{|\vec{r}\rangle\}$ of the position representation in three dimensions:

$$X|\vec{r}\rangle = X|x, y, z\rangle = x|x, y, z\rangle = x|\vec{r}\rangle, \tag{2.82}$$

$$Y|\vec{r}\rangle = Y|x, y, z\rangle = y|x, y, z\rangle = y|\vec{r}\rangle, \tag{2.83}$$

$$Z|\vec{r}\rangle = Z|x, y, z\rangle = z|x, y, z\rangle = z|\vec{r}\rangle. \tag{2.84}$$

Thus, $X, Y$, and $Z$ multiply the basis vectors of the position representation by the cartesian components of the points where they are centered. The position operator $\vec{R}$ is a **vector operator**, i.e., a collection of the three operators $X, Y$, and $Z$ which transform like the components of a vector in $R^3$. The effect of the operator $\vec{R}$ on the position states

$$\vec{R}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle \tag{2.85}$$

is to multiply them by the position vector with which they are labeled. In this representation, then,

$$\vec{R}|\psi\rangle = \int d^3r\ \psi(\vec{r})\,\vec{R}|\vec{r}\rangle = \int d^3r\ \psi(\vec{r})\,\vec{r}|\vec{r}\rangle = \int d^3r\ \left[\vec{r}\psi(\vec{r})\right]|\vec{r}\rangle. \tag{2.86}$$

Thus, the effect of $\vec{R}$ is to multiply the wave function $\psi(\vec{r})$ by $\vec{r}$. Finally, we can extend this to *functions* of $\vec{r}$, as in one-dimension. For each function $V(\vec{r})$ we can define an operator $V = V(\vec{R})$ such that

$$V|\vec{r}\rangle = V(\vec{r})|\vec{r}\rangle \tag{2.87}$$

and so

$$V|\psi\rangle = \int d^3r\ \psi(\vec{r})\,V|\vec{r}\rangle = \int d^3r\ V(\vec{r})\psi(\vec{r})\,|\vec{r}\rangle. \tag{2.88}$$

Thus the operator $V$ acts in the $\vec{r}$ representation to multiply the wave function by the function $V(\vec{r})$.

These kinds of **multiplicative operators** can be defined for any representation. If $\{|\alpha\rangle\}$ is an ONB for the space, we can define an operator $A$ such that

$$A|\alpha\rangle = \alpha|\alpha\rangle$$

for all basis vectors of this representation. Then, for any function $g(\alpha)$ we can define an operator $G = G(A)$ such that

$$G|\alpha\rangle = g(\alpha)|\alpha\rangle, \tag{2.89}$$

then

$$G|\psi\rangle = \int d\alpha\ \psi(\alpha)\,G|\alpha\rangle = \int d\alpha\ g(\alpha)\psi(\alpha)\,|\alpha\rangle, \tag{2.90}$$

so that $G$ acts in the $\alpha$ representation to multiply the wave function in that representation by the function $g(\alpha)$.

### *2.2.3 Differential Operators*

Another class of operators can also be defined through the position representation. Let us define operators $D_x, D_y$, and $D_z$ in such a way that if

$$|\psi\rangle = \int d^3r \; \psi(\vec{r}) \, |\vec{r}\rangle, \tag{2.91}$$

then

$$D_x|\psi\rangle = \int d^3r \; \frac{\partial \psi}{\partial x} \, |\vec{r}\rangle. \tag{2.92}$$

Thus, $D_x$ replaces the wave function in the position representation by its partial derivative with respect to $x$. Similar actions are implicitly defined for $D_y$ and $D_z$. Such operators are **differential** operators in this representation. These three operators form the components of the vector operator $\vec{D}$ which "takes the gradient in the $\vec{r}$ representation". That is to say,

$$\vec{D}|\psi\rangle = \int d^3r \; \vec{\nabla}\psi \, |\vec{r}\rangle. \tag{2.93}$$

A more useful variation of this operator is obtained by multiplying it by the square root of $-1$. We thus introduce the operator $\vec{K} = -i\vec{D}$ which we will refer to as the **wavevector operator**, and define through the expression

$$\vec{K}|\psi\rangle = \int d^3r \; [-i\vec{\nabla}\psi(\vec{r})] \, |\vec{r}\rangle. \tag{2.94}$$

It is instructive to consider the action of this operator in the $\vec{k}$ representation. By assumption, an arbitrary state can be expanded in the $\vec{k}$ representation in the form

$$|\psi\rangle = \int d^3r \; \psi(\vec{k}) \, |\vec{k}\rangle, \tag{2.95}$$

where

$$\psi(\vec{r}) = \int \frac{d^3k}{(2\pi)^{3/2}} \; e^{i\vec{k}\cdot\vec{r}}\psi(\vec{k}). \tag{2.96}$$

Thus,

$$-i\vec{\nabla}\psi(\vec{r}) = \int \frac{d^3k}{(2\pi)^{3/2}} \; [-i\vec{\nabla}e^{i\vec{k}\cdot\vec{r}}]\psi(\vec{k}). \tag{2.97}$$

The gradient operator, which acts only on the position variables, just "pulls down the wavevector" from the exponential, i.e.,

$$-i\vec{\nabla}\psi(\vec{r}) = \int \frac{d^3k}{(2\pi)^{3/2}} \; [\vec{k}e^{i\vec{k}\cdot\vec{r}}]\psi(\vec{k}). \tag{2.98}$$

Thus, we deduce that

$$\vec{K}|\psi\rangle = \int d^3r \; \left[-i\vec{\nabla}\psi(\vec{r})\right] |\vec{r}\rangle = \int d^3r \left[\int \frac{d^3k}{(2\pi)^{3/2}} \; \vec{k}e^{i\vec{k}\cdot\vec{r}} \, \psi(\vec{k})\right] |\vec{r}\rangle. \tag{2.99}$$

Interchanging the order of integration, this becomes

$$\vec{K}|\psi\rangle = \int d^3k \; \vec{k}\psi(\vec{k}) \int \frac{d^3r}{(2\pi)^{3/2}} \; e^{i\vec{k}\cdot\vec{r}}|\vec{r}\rangle. \tag{2.100}$$

In the last integral we recognize the definition of the basis states $|\vec{k}\rangle$. Thus we obtain the simple result

$$\vec{K}|\psi\rangle = \int d^3k \;\; \vec{k}\psi(\vec{k})|\vec{k}\rangle. \tag{2.101}$$

Thus, $\vec{K}$ acts "in the $\vec{k}$ representation" to multiply the wave function in that representation by $\vec{k}$. Since $\vec{K}$ really acts only on the kets $|\vec{k}\rangle$, we deduce the action

$$\vec{K}|\vec{k}\rangle = \vec{k}|\vec{k}\rangle. \tag{2.102}$$

Thus, the operator $\vec{K}$ plays the same role in the $\vec{k}$ representation that the operator $\vec{R}$ plays in the $\vec{r}$ representation, i.e., it simply multiplies the basis vectors by the value of the parameter $\vec{k}$ that labels them. (It is useful to think of this operation as one in which the operator "pulls out" the label.) Similarly the kinetic energy operator

$$H_0 = \frac{\hbar^2 K^2}{2m} = \frac{\hbar^2}{2m}\vec{K}\cdot\vec{K}$$

is a multiplicative operator in the $\vec{k}$ representation that multiplies the wave function by $\hbar^2 k^2/2m$, but is a differential operator in the $\vec{r}$ representation whose action is take $\psi(\vec{r})$ onto the function $-\left(\hbar^2/2m\right)\nabla^2\psi$.

Thus, whether an operator is a multiplicative operator or a differential operator is very much a *representation-dependent* statement. It is left as an exercise to show that in the $\vec{k}$ representation the position operator $\vec{R}$ actually acts as a differential operator, i.e., that

$$\vec{R}|\psi\rangle = \int d^3k \;\; [i\vec{\nabla}_k\psi(\vec{k})]|\vec{k}\rangle, \tag{2.103}$$

where $\vec{\nabla}_k$ means to take the gradient with respect to the variables $k_x, k_y$, and $k_z$, and that in the position representation the kinetic energy operator $H_0$ is a differential operator proportional to the Laplacian (as in Schrödinger's equation).

### *2.2.4 Ket-Bra Operators*

A very useful class of operator can be defined using any two vectors in the space. If $|\phi\rangle$ and $|\chi\rangle$ are vectors in $S$ then we can define an operator

$$A = |\phi\rangle\langle\chi| \tag{2.104}$$

whose action on any state $|\psi\rangle$ is as follows:

$$A|\psi\rangle = |\phi\rangle\langle\chi|\psi\rangle \tag{2.105}$$

which is just the vector $|\phi\rangle$ multiplied by the number $\langle\chi|\psi\rangle$. A linear sum of operators of this form is itself an operator. This "ket-bra" form is particularly useful for expressing what are referred to as projection operators.

### *2.2.5 Projection Operators: The completeness relation*

An operator $P$ is said to be a **projection operator**, or simply a **projector**, if it satisfies the **idempotency condition**

$$P^2 = P. \tag{2.106}$$

Note that this implies that $P^n = P$ for all integers $n \geq 1$. As an example, if $|\phi\rangle$ is square normalized to unity, so that $\langle\phi|\phi\rangle = 1$, then the operator

$$P_\phi = |\phi\rangle\langle\phi| \tag{2.107}$$

is a projector onto the direction of the vector $|\phi\rangle$. It is a projector because

$$P_\phi^2 = (|\phi\rangle\langle\phi|)\,(|\phi\rangle\langle\phi|) = |\phi\rangle\langle\phi|\phi\rangle\langle\phi| = |\phi\rangle\langle\phi| = P_\phi, \tag{2.108}$$

where the central inner product collapsed to unity because of the normalization of the state $|\phi\rangle$.Thus, the action of $P_\phi$ on an arbitrary state $|\psi\rangle$ is to take away those parts of $|\psi\rangle$ not lying along $|\phi\rangle$, and to leave the part lying along the direction of $|\phi\rangle$ alone.

As a simple extension of this idea we note that, a set of states $\{|i\rangle\}$ forms an orthonormal set of vectors so that $\langle i|j\rangle = \delta_{ij,}$ then the operator

$$P = \sum_{i=1}^{n} |i\rangle\langle i| \tag{2.109}$$

is also a projection operator because

$$P^2 = \left(\sum_{i=1}^{n} |i\rangle\langle i|\right)\left(\sum_{j=1}^{n} |j\rangle\langle j|\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} |i\rangle\langle i|j\rangle\langle j| = \sum_{i=1}^{n}\sum_{j=1}^{n} |i\rangle\delta_{ij}\langle j| = \sum_{i=1}^{n} |i\rangle\langle i| = P. \tag{2.110}$$

**Comment:** Projection operators always "project onto something". In this latter case, the operator $P$ projects onto the *subspace* spanned by the vectors in the orthonormal set. Recall the definition of a subspace:

> A set of vectors $S' \subset S$ which is a subset of a vector space $S$ is a **subspace** of $S$ if it is closed under the same operations that are defined in the parent space.

Any subset of vectors spans some subspace, namely the subspace of all vectors that can be produced by them by forming all possible linear combinations of vectors in the subset. In the example above, if the orthonormal states $\{|i\rangle\}$ were **complete**, so that they actually formed an ONB for the space, then the subspace that they project upon would be the entire space. Since, for such a basis, an expansion of the form

$$|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle \tag{2.111}$$

exists for any state $|\psi\rangle$ in the space, the action of the operator $P$ on such a state

$$P|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle = |\psi\rangle \tag{2.112}$$

is to just reproduce the state it acted on. We deduce that if the states $\{|i\rangle\}$ form an orthonormal basis, then

$$\sum_i |i\rangle\langle i| = \mathbf{1}. \tag{2.113}$$

This relation, which is of fundamental importance is referred to as a **decomposition of unity** in the basis $\{|i\rangle\}$, or as a statement of the **completeness relation** for the states $\{|i\rangle\}$.

These ideas are also extensible with some care to continuously-indexed states. If the states $\{|\alpha\rangle\}$ form an ONB for the space $S$ then the operator

$$\rho_\alpha = |\alpha\rangle\langle\alpha| \tag{2.114}$$

is *not* a projection operator, since the state $|\alpha\rangle$ is not square-normalized to unity. Indeed, it is of infinite norm, since $\langle\alpha|\alpha\rangle = \delta(\alpha - \alpha) = \delta(0)$. However, the integral of this operator over any region of the possible values taken on by the parameter $\alpha$ is a projector. That is, if we define

$$P_{ab} = \int_a^b d\alpha\ \rho_\alpha = \int_a^b d\alpha\ |\alpha\rangle\langle\alpha| \tag{2.115}$$

then

$$P_{ab}^2 = \left(\int_a^b d\alpha\ |\alpha\rangle\langle\alpha|\right)\left(\int_a^b d\alpha'\ |\alpha'\rangle\langle\alpha'|\right) = \int_a^b d\alpha \int_a^b d\alpha'\ |\alpha\rangle\langle\alpha|\alpha'\rangle\langle\alpha'|. \tag{2.116}$$

In this last expression, the normalization condition on the states $|\alpha\rangle$ yield a delta function, which means that

$$P_{ab}^2 = \int_a^b d\alpha \int_a^b d\alpha'\ |\alpha\rangle\delta(\alpha - \alpha')\langle\alpha'| = \int_a^b d\alpha\ |\alpha\rangle\langle\alpha| = P_{ab} \tag{2.117}$$

(Note, that in any expression in which one is integrating over a delta function, one simply removes the delta function and the integral sign and replaces the integration variable wherever it occurs with the value which makes the argument of the delta function vanish.). We will refer to an operator such as $\rho_\alpha$ as a **projector density**, since it's integral always gives a projector (in the same way that the integral of a charge density always gives a charge, etc.). The action of $\rho_\alpha$ on an arbitrary state $|\psi\rangle$ gives a vector

$$\rho_\alpha|\psi\rangle = |\alpha\rangle\langle\alpha|\psi\rangle = \psi(\alpha)|\alpha\rangle$$

along the direction of the state $|\alpha\rangle$ multiplied by the expansion coefficient $\psi(\alpha)$. Thus, we will continue to refer to $\rho_\alpha|\psi\rangle$ as the "part of $|\psi\rangle$ lying along the state $|\alpha\rangle$". At any rate, it is easy to see that $\rho_\alpha$ strips away any part of $|\psi\rangle$ not lying along that direction in state space.

As in the discrete case, if we now consider the projector which includes all the states in the basis, we project onto the entire space. Thus, since we can always write an arbitrary state in the form

$$|\psi\rangle = \int d\alpha\ |\alpha\rangle\langle\alpha|\psi\rangle \tag{2.118}$$

the action of the operator

$$P = \int d\alpha\ |\alpha\rangle\langle\alpha|, \tag{2.119}$$

which has no restrictions on the values of $\alpha$, is to reproduce whatever state it acts upon, i.e.,

$$P|\psi\rangle = \int d\alpha\ |\alpha\rangle\langle\alpha|\psi\rangle = |\psi\rangle \tag{2.120}$$

This being true for all $|\psi\rangle$, we deduce the completeness relation

$$\int d\alpha\ |\alpha\rangle\langle\alpha| = \mathbf{1} \tag{2.121}$$

for continuously-indexed states.

These decompositions of the identity operator facilitate the development of expansions for vectors and inner products, as demonstrated below:

$$|\chi\rangle = \mathbf{1}|\chi\rangle = \sum_i |i\rangle\langle i|\chi\rangle = \sum_i \chi_i \, |i\rangle \tag{2.122}$$

$$|\chi\rangle = \mathbf{1}|\chi\rangle = \int d\alpha \, |\alpha\rangle\langle\alpha|\chi\rangle = \int d\alpha \; \chi(\alpha) \, |\alpha\rangle \tag{2.123}$$

$$\langle\psi|\chi\rangle = \langle\psi| \left(\mathbf{1}|\chi\rangle\right) = \sum_i \langle\psi|i\rangle\langle i|\chi\rangle = \sum_i \psi_i^* \chi_i \tag{2.124}$$

$$\langle\psi|\chi\rangle = \langle\psi| \left(\mathbf{1}|\chi\rangle\right) = \int d\alpha \, \langle\psi|\alpha\rangle\langle\alpha|\chi\rangle = \int d\alpha \, \psi^*(\alpha)\chi(\alpha). \tag{2.125}$$

In this way representation independent expressions (on the left) are converted into representation dependent expressions (on the right) by inserting an appropriate "complete set of states". Expressions such as $\langle\psi|\chi\rangle$ are representation independent. Expressions such as

$$\sum_i \langle\psi|i\rangle\langle i|\chi\rangle = \sum_i \psi_i^* \chi_i \tag{2.126}$$

are representation dependent, because they depend upon a particular choice of representation, i.e., of basis. (Note that expressions such as $|\psi(\vec{r})\rangle$ or $|\hat{\psi}(\vec{k})\rangle$ are actually "ill-defined", have no official meaning, and are to be avoided. While we probably can guess what is intended, they are ambiguous and an abuse of standard usage. The view here is that the functions $\psi(\vec{r})$ and $\psi(\vec{k})$ give specific representations of a single underlying state vector $|\psi\rangle$, and not the other way around.)

### *2.2.6 Matrix Elements*

The **matrix element** of an operator $A$ between (or connecting) the states $|\chi\rangle$ and $|\psi\rangle$ is the scalar quantity

$$\langle\psi| \left(A|\chi\rangle\right) = \langle\psi|\chi_A\rangle, \tag{2.127}$$

where $|\chi_A\rangle = A|\chi\rangle$.

### *2.2.7 Action of Operators on Bras of $S^*$*

We have defined the action of operators in the space of kets $|\chi\rangle \in S$. We now extend the definition to allow them to operate in the space $S^*$ of bras $\langle\chi|$, by requiring that any matrix element be unchanged if the operator acts to the left, rather than to the right. Thus, for all $|\chi\rangle, |\psi\rangle \in S$ we require that

$$\langle\psi| \left(A|\chi\rangle\right) = \left(\langle\psi|A\right)|\chi\rangle = \langle\psi|A|\chi\rangle \tag{2.128}$$

where in the last expression we have removed the parentheses since now (by construction) $A$ can act in either direction. (Note that expressions like $A\langle\psi|$ and $|\chi\rangle A$ are undefined. Operators can only act on vectors when they are next to the vertical line appearing in the notation.)
If the state $|\phi\rangle$ is a unit vector, then the matrix element $\langle\phi|A|\phi\rangle$ is referred to as the **expectation value** of the operator $A$ taken with respect to the state $|\phi\rangle$.

An extremely important consequence of this definition of the action of an operator on bras is brought out by the answer to the following question: If $A|\chi\rangle = |\psi\rangle$, does it follow that $\langle\chi|A = \langle\psi|$? One might be tempted to think so, since the states $|\chi\rangle$ and $|\psi\rangle \in S$ are supposed to be in 1-1 correspondence with the states $\langle\chi|$ and $\langle\psi| \in S$. Nonetheless, the answer is "no". There is a relationship *similar* to this that we can write, but it involves what is referred to as the Hermitian adjoint of the operator $A$. The point is that we have to extend the 1-1 correspondence to include the relationship that exists between the operators that act in $S$ and those that act in $S^*$. We explore this below.

### *2.2.8 Hermitian Conjugation*

We have previously established a 1-1 correspondence between the ket's $|\chi\rangle$ of $S$ and the bra's $\langle\chi|$ of $S^*$ (which in any *representation* can be thought of as the correspondence between column vectors and the complex-conjugated row vectors). Corresponding *elements* of $S$ and $S^*$ are said to be **Hermitian Conjugates**, **Hermitian Adjoints**, or simply **Adjoints** of one another. Thus $S$ is the linear vector space adjoint to $S^*$. Similarly, the bra $\langle\chi|$ is the adjoint of the ket $|\chi\rangle$. It is convenient to use the notation $[\cdots]^+$ to denote the adjoint of $[\cdots]$. Thus we write

$$[|\chi\rangle]^+ = \langle\chi| \qquad\qquad [\langle\chi|]^+ = |\chi\rangle \tag{2.129}$$

which shows that $[[\cdots]^+]^+ = [\cdots]$. We can also refer, e.g., to the ONB of bra's $\{\langle i|\}$ as being adjoint to the ONB of ket's $\{|i\rangle\}$. This gives rise to the following point. Given the expansion

$$|\chi\rangle = \sum_i b_i\,|i\rangle \tag{2.130}$$

where, of course, $b_i = \langle i|\chi\rangle$, and the similar expansion

$$\langle\chi| = \sum_i c_i\,\langle i| \tag{2.131}$$

associated with the dual space, how are the expansion coefficients $b_i$ and $c_i$ related to one another? To find out, note that

$$\langle\chi|j\rangle = \sum_i c_i\,\langle i|j\rangle = \sum_i c_i\,\delta_{ij} = c_j = (\langle j|\chi\rangle)^* = b_j^*. \tag{2.132}$$

Thus we can write $c_i = b_i^*$, and therefore if

$$|\chi\rangle = \sum_i \chi_i\,|i\rangle = \sum_i |i\rangle\langle i|\chi\rangle \tag{2.133}$$

then

$$\langle\chi| = \sum_i \chi_i^*\,\langle i| = \sum_i \langle\chi|i\rangle\langle i|. \tag{2.134}$$

In general, this implies that if $\lambda$ is an element of the field associated with the space $S$, then $\lambda^*$ is the corresponding element of the field in $S^*$ (i.e., the scalar which plays the same role in $S^*$ that $\lambda$ plays in $S$). Thus, the rule for taking the *Hermitian conjugate* of any complex number is simply to take its *complex conjugate*. We write

$$[\lambda]^+ = \lambda^* \qquad\qquad [\lambda^*]^+ = \lambda. \tag{2.135}$$

We now extend this idea to operators. If the operator $A$ maps the ket $|\chi\rangle$, say, onto the ket $|\psi\rangle$ then the *adjoint* of the operator $A$ must have the corresponding effect in the adjoint space. Thus, the operator $A^+$ (which is read "$A$ adjoint" or "$A$ dagger") has the effect that if

$$A|\chi\rangle = |\psi\rangle \tag{2.136}$$

then

$$\langle\chi|A^+ = \langle\psi|. \tag{2.137}$$

This is the relationship we were looking for in the question posed above. It shows that when we "flip things around" we have to replace operators by their adjoints. Thus we can write

$$[A|\chi\rangle]^+ = \langle\chi|A^+ \tag{2.138}$$

A few moments of study of the adjoint process allows the following rules to be developed: To take the adjoint of any product of operators, numbers, bra's, ket's etc., (1) replace all elements by their adjoints (bra's are replaced by ket's, operators by their adjoints, numbers by their conjugates), and (2) reverse the order of all elements in the original product. Once this operation is performed, any numbers can be commuted past any operators or vectors to simplify the expression. As an example, note that

$$[\langle\phi|\xi\rangle]^+ = \langle\xi|\phi\rangle = \langle\phi|\xi\rangle^* \tag{2.139}$$

$$[\langle\psi|A|\phi\rangle]^+ = \langle\phi|A^+|\psi\rangle = \langle\psi|A|\phi\rangle^*$$

The first of identity in each of the last two equations follows from the rules for taking the adjoint, the second from the fact that both quantities are scalars, and the adjoints of which are just the complex conjugates. Finally, as a more complicated example we note that the operator

$$A = \lambda\,\langle\phi|B|\xi\rangle\,|\nu\rangle\langle\mu| \tag{2.140}$$

has as its adjoint

$$A^+ = |\mu\rangle\langle\nu|\,\langle\xi|B^+|\phi\rangle\lambda^*. \tag{2.141}$$

A short list of properties of the Hermitian adjoint are given below:

$$[A^+]^+ = A \tag{2.142}$$

$$[\lambda A]^+ = \lambda^* A^+ \tag{2.143}$$

$$[A+B]^+ = A^+ + B^+ \tag{2.144}$$

$$[AB]^+ = B^+A^+ \tag{2.145}$$

This last rule, which displays the reversal of order of the adjoint of a product, is easily proved. If $AB|\psi\rangle = A|\psi_B\rangle = |\phi\rangle$, then the adjoint is

$$\langle\phi| = \langle\psi_B|A^+ = [|\psi_B\rangle]^+A^+ = [B|\psi\rangle]^+A^+ = \langle\psi|B^+A^+ \tag{2.146}$$

from which we see that if $AB|\psi\rangle = |\phi\rangle$, then $\langle\psi|B^+A^+ = \langle\phi|$, which proves the result. We are now in a position to define some additional terms, one of which appears in the statement of the second postulate.

### *2.2.9 Hermitian, Anti-Hermitian, and Unitary Operators*

An operator $A$ is **Hermitian** or **self adjoint** if it is equal to its Hermitian adjoint, i.e., if

$$A = A^+. \tag{2.147}$$

In terms of matrix elements, the property

$$\langle\psi|A|\phi\rangle = \langle\phi|A^+|\psi\rangle^*, \tag{2.148}$$

which is true for any operator, reduces for Hermitian operators to the relation

$$\langle\psi|A|\phi\rangle = \langle\phi|A|\psi\rangle^*. \tag{2.149}$$

As a special case this implies that $\langle\phi|A|\phi\rangle = \langle\phi|A|\phi\rangle^*$, which implies that expectation values of a Hermitian operator are strictly real.

An operator $A$ is **anti-Hermitian** if it is equal to the negative of its adjoint, i.e., if

$$A = -A^+. \tag{2.150}$$

The matrix elements of an anti-Hermitian operator obey the easily-derivable condition $\langle\psi|A|\phi\rangle = -\langle\phi|A|\psi\rangle^*$,which implies that $\langle\phi|A|\phi\rangle = -\langle\phi|A|\phi\rangle^*$. Thus, expectation values of anti-Hermitian operators are strictly imaginary.
Note that if $A$ is any operator, it may be written in the form

$$\begin{aligned} A &= \frac{1}{2}\left(A + A^+\right) + \frac{1}{2}\left(A - A^+\right) \\ &= A_H + A_A \end{aligned} \tag{2.151}$$

where $A_H = \frac{1}{2}\left(A + A^+\right)$ is Hermitian (take its adjoint and see!) and $A_A = \frac{1}{2}\left(A - A^+\right)$ is anti-Hermitian (likewise!). Thus an arbitrary operator can be uniquely decomposed into a sum of Hermitian and anti-Hermitian operators.

An operator $U$ is **unitary** if its adjoint is equal to it inverse. Thus, for a unitary operator

$$U^+ = U^{-1}, \tag{2.152}$$

or equivalently,

$$UU^+ = U^+U = \mathbf{1}. \tag{2.153}$$

We will see that unitary operators (or the transformations they induce) play the same role in quantum mechanical Hilbert spaces that orthogonal transformations play in Cartesian vector spaces such as $R^3$.

### *2.2.10 Matrix Representation of Operators*

Let $\{|n\rangle\}$ be an ONB for the space $S$ and let $A$ be an operator acting in the space. From the trivial identity

$$A = \mathbf{1}A\mathbf{1} \tag{2.154}$$

we obtain a representation for $A$ by substituting a decomposition of unity in the $\{|n\rangle\}$ basis. Thus, we obtain

$$A = \left(\sum_n |n\rangle\langle n|\right) A \left(\sum_{n'} |n'\rangle\langle n'|\right) \tag{2.155}$$

With the different dummy indices we can now remove the parentheses to obtain

$$A = \sum_{n,n'} |n\rangle \langle n|A|n'\rangle \langle n'| \tag{2.156}$$

which we write in the form

$$A = \sum_{n,n'} |n\rangle A_{nn'} \langle n'|, \tag{2.157}$$

where

$$A_{nn'} = \langle n|A|n'\rangle \tag{2.158}$$

is the matrix element of $A$ connecting the basis states $|n\rangle$ and $|n'\rangle$. Thus, we obtain a decomposition of the operator $A$ in the "ket-bra" form, which makes its action on any state self-evident. The operator $A$, therefore, is completely determined by its matrix elements in any ONB. Thus, suppose that

$$|\phi\rangle = A|\psi\rangle \tag{2.159}$$

for some states $|\psi\rangle$ and $|\phi\rangle$. The expansion coefficients for the states $|\psi\rangle$ and $|\phi\rangle$ are clearly related. Note that if

$$|\phi\rangle = \sum_n \phi_n \, |n\rangle \qquad\qquad |\psi\rangle = \sum_n \psi_n \, |n\rangle \tag{2.160}$$

then

$$\phi_n = \langle n|\phi\rangle = \langle n|A\psi\rangle = \sum_{n'} \langle n|A|n'\rangle\langle n'|\psi\rangle \tag{2.161}$$

which can be written

$$\phi_n = \sum_{n'} A_{nn'}\psi_{n'}. \tag{2.162}$$

But this is precisely the form of a matrix multiplication

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \\ & & \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \end{pmatrix} \tag{2.163}$$

of a matrix having elements $A_{nn'}$ with a column vector having elements $\psi_n$, resulting in a column vector with elements $\phi_n$. Thus we see that in the row-vector-column vector representation induced by any discrete ONB, an operator is naturally represented by a **matrix** having entries which are just the matrix elements of that operator connecting the different members of the basis. Note that in producing the matrix of elements $A_{nn'} = \langle n|A|n'\rangle$ the bra corresponds to the row index, while the ket corresponds to the column index. Note also that this *expansion* of the operator

$$A = \sum_{n,n'} |n\rangle \, A_{nn'} \, \langle n'| = |1\rangle \, A_{11}\langle 1| + |1\rangle \, A_{12}\langle 2| + \cdots + |2\rangle \, A_{21}\langle 1| + \ldots \tag{2.164}$$

in ket-bra form, has the matrix interpretation

$$\begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} = \begin{pmatrix} A_{11} & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \begin{pmatrix} 0 & A_{12} & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \cdots + \begin{pmatrix} 0 & 0 & \cdots \\ A_{21} & 0 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} + \cdots \tag{2.165}$$

where we are simply filling up each slot of the matrix one element at a time. It is worthwhile looking at a few additional examples. Consider the matrix element $\langle\psi|A|\phi\rangle$ between arbitrary states $|\phi\rangle$ and $|\psi\rangle$. Inserting our expansion for $A$ this becomes

$$\langle\psi|A|\phi\rangle = \sum_{n,n'} \langle\psi|n\rangle \, A_{nn'} \, \langle n'|\phi\rangle \, , \tag{2.166}$$

in which we recognize $\phi_{n'} = \langle n'|\phi\rangle$ and $\psi_n^* = \langle\psi|n\rangle$. Thus, we obtain the result

$$\langle\psi|A|\phi\rangle = \sum_{n,n'} \psi_n^* \, A_{nn'} \, \phi_{n'} \, , \tag{2.167}$$

which has the matrix interpretation associated with the following operation

$$\langle\psi|A|\phi\rangle = \begin{pmatrix} \psi_1^* & \psi_2^* & \cdots \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \end{pmatrix}. \tag{2.168}$$

As another example, consider the operator product of

$$A = \sum_{n,n'} |n\rangle\, A_{nn'}\, \langle n'| \tag{2.169}$$

and

$$B = \sum_{n,n'} |n\rangle\, B_{nn'}\, \langle n'|. \tag{2.170}$$

The product operator $C = AB$ has a similar expansion, i.e.,

$$C = \sum_{n,n'} |n\rangle\, C_{nn'}\, \langle n'| \tag{2.171}$$

where

$$C_{nn'} = \langle n|C|n'\rangle = \langle n|AB|n'\rangle = \sum_{n''} \langle n|A|n''\rangle\langle n''|B|n'\rangle \tag{2.172}$$

Thus

$$C_{nn'} = \sum_{n''} A_{nn''} B_{n''n'}.$$

which is equivalent to the matrix multiplication

$$\begin{pmatrix} C_{11} & C_{12} & \cdots \\ C_{21} & C_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots \\ A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots \\ B_{21} & B_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix}. \tag{2.173}$$

As a final example, consider the matrix representing the adjoint of an operator. If

$$A = \sum_{n,n'} |n\rangle\, A_{nn'}\, \langle n'| \tag{2.174}$$

then by the two-part rule we developed for taking the adjoint, it follows that

$$A^+ = \sum_{n,n'} |n'\rangle\, A^*_{nn'}\, \langle n|. \tag{2.175}$$

Since $n$ and $n'$ are simply summation indices we can switch them to find that

$$A^+ = \sum_{n,n'} |n\rangle\, A^*_{n'n}\, \langle n'| = \sum_{n,n'} |n\rangle\, A^+_{nn'}\, \langle n'|, \tag{2.176}$$

from which we deduce that

$$A^+_{nn'} = A^*_{n'n}. \tag{2.177}$$

To interpret this properly a little care must be taken with this notation: the symbol $A^+_{nn'}$ means the $n, n'$ matrix element of the operator $A^+$, while the symbol $A^*_{n'n}$ means the complex conjugate of the $n', n$ matrix element of the operator $A$. Thus, in any ONB the matrix representing $A^+$ is the **complex-conjugate transpose** of the matrix representing $A$. A Hermitian operator is equal to its adjoint, so that the matrix elements representing such an operator obey the relation

$$\langle n|A|n'\rangle = A_{nn'} = \langle n'|A|n\rangle^* = A^*_{n'n}. \tag{2.178}$$

Thus, for a Hermitian operator

$$A_{nn'} = A^*_{n'n}$$

which implies, e.g., that the diagonal elements of the matrix representing a Hermitian operator are *real*. More generally, this shows that any matrix $\mathbf{A}$ representing a Hermitian operator is equal to its complex-conjugate transpose, i.e., $\mathbf{A} = \left(\mathbf{A}^{\mathrm{T}}\right)^*$. Any matrix obeying this relationship is a **Hermitian matrix**. The symmetry properties of an anti-Hermitian operators and matrices are left as an exercise.

Let us give some examples, suppose that in a 3-dimensional vectors space the operators $A$ and $B$ are represented in some orthonormal basis by the following matrices

$$\mathbf{A} = \begin{pmatrix} 0 & 2i & 7-3i \\ -2i & -3 & 4 \\ 7+3i & 4 & 8 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 4i & 4i & 14 \\ 0 & -6 & 2+6i \\ 8 & 8 & 8 \end{pmatrix}. \tag{2.179}$$

Note that we use non-italicized boldface symbols to represent the matrices in order to distinguish them from the operators themselves. The matrices $\mathbf{A}^+$ and $\mathbf{B}^+$ representing the adjoints of the operators $A$ and $B$ are

$$\mathbf{A}^+ = \left(\mathbf{A}^{\mathrm{T}}\right)^* = \begin{pmatrix} 0 & 2i & -3i+7 \\ -2i & -3 & 4 \\ 3i+7 & 4 & 8 \end{pmatrix} = \mathbf{A}$$

and

$$\mathbf{B}^+ = \left(\mathbf{B}^{\mathrm{T}}\right)^* = \begin{pmatrix} -4i & 0 & 8 \\ -4i & -6 & 8 \\ 14 & 2-6i & 8 \end{pmatrix},$$

from which we see that the operator $A$ is Hermitian, and is represented by a **Hermitian matrix**, while the operator $B$ is not Hermitian. The latter operator can be written as a sum of Hermitian and anti-Hermitian parts, however, as can the matrices representing it, i.e., we can write

$$\mathbf{B} = \mathbf{B}_H + \mathbf{B}_A$$

where

$$\mathbf{B}_H = \frac{1}{2}\left(\mathbf{B} + \mathbf{B}^+\right) = \begin{pmatrix} 0 & 2i & 11 \\ -2i & -6 & 5+3i \\ 11 & 5-3i & 8 \end{pmatrix}$$

and

$$\mathbf{B}_A = \frac{1}{2}\left(\mathbf{B} - \mathbf{B}^+\right) = \begin{pmatrix} 4i & 2i & 3 \\ 2i & 0 & -3+3i \\ -3 & 3+3i & 0 \end{pmatrix}.$$

It is an interesting fact that *neither* the transpose or the complex conjugate of an operator are, *by themselves*, well defined concepts; i.e., given an operator $A$, there is no operator that can be uniquely identified with the transpose of $A$. Although one can form the transpose $\mathbf{A}^{\mathrm{T}}$ of the *matrix* $\mathbf{A}$ representing $A$ in any basis, the operator associated with the transposed matrix will not generally correspond to the operator associated with the transpose of the matrix representing $A$ in any other basis. Thus, the act of transposition is a *representation dependent* operation. Similarly, complex conjugation can be performed on matrices, or on matrix elements, but it is not an operation that is uniquely defined for the operators themselves. It is somewhat surprising, therefore, that the Hermitian adjoint, which in a sense combines these two representation dependent operations, yields an operator that is independent of representation. This again emphasizes one of the basic themes, which is that bras, kets, and operators are *not* row vectors, column vectors, and matrices. The former may be represented by the latter, but the representation and that which is represented are two conceptually different things.

Matrix representations of this form were developed extensively by Heisenberg and gave rise to the term "matrix mechanics", in analogy to the "wave mechanics" developed by Schrödinger, which focuses on a wave function representation for the underlying space. Clearly, however, whether one has a wave mechanical or matrix mechanical representation depends simply upon the choice of basis (i.e., discrete or continuous) in which one is working.

**Extension to Continuous Representations** - Let $\{|\alpha\rangle\}$ be a continuous ONB for the space $S$ and let $A$ be an operator acting in the space. From the trivial identity

$$A = \mathbf{1}A\mathbf{1} \tag{2.180}$$

we obtain a representation for $A$ by substituting a decomposition of unity in the $\{|\alpha\rangle\}$ basis. Thus, we obtain

$$A = \left(\int d\alpha\, |\alpha\rangle\langle\alpha|\right) A \left(\int d\alpha'\, |\alpha'\rangle\langle\alpha'|\right). \tag{2.181}$$

With the different dummy indices safely in place we can now remove the parentheses to obtain

$$A = \int d\alpha \int d\alpha'\ |\alpha\rangle\langle\alpha|A\,|\alpha'\rangle\langle\alpha'| \tag{2.182}$$

which we write in the form

$$A = \int d\alpha \int d\alpha'\ |\alpha\rangle\, A(\alpha,\alpha')\, \langle\alpha'| \tag{2.183}$$

where the "kernel"

$$A(\alpha,\alpha') = \langle\alpha|A\,|\alpha'\rangle \tag{2.184}$$

of this integral relation is just the matrix element of $A$ connecting the basis states $|\alpha\rangle$ and $|\alpha'\rangle$. Thus, we obtain a decomposition of the operator $A$ in the "ket-bra" form, which makes its action on any state self-evident. The operator $A$ is, therefore, completely determined by its matrix elements in any continuous ONB. Thus, if

$$|\phi\rangle = A|\psi\rangle \tag{2.185}$$

for some states

$$|\phi\rangle = \int d\alpha\, \phi(\alpha)\, |\alpha\rangle \tag{2.186}$$

and

$$|\psi\rangle = \int d\alpha\, \psi(\alpha)\, |\alpha\rangle \tag{2.187}$$

then

$$\phi(\alpha) = \langle\alpha|\phi\rangle = \langle\alpha|A|\psi\rangle = \int d\alpha' \ \langle\alpha|A|\alpha'\rangle\langle\alpha'|\psi\rangle \tag{2.188}$$

or

$$\phi(\alpha) = \int d\alpha' \, A(\alpha,\alpha') \, \psi(\alpha'). \tag{2.189}$$

This is of the form of a "continuous" matrix multiplication, with integration replacing the summation process. It is worthwhile looking at a few additional examples. Consider the matrix element $\langle\psi|A|\phi\rangle$ between arbitrary states $|\phi\rangle$ and $|\psi\rangle$. Inserting our expansion for $A$ this becomes

$$\langle\psi|A|\phi\rangle = \int d\alpha \int d\alpha' \, \langle\psi\,|\alpha\rangle \, A(\alpha,\alpha') \, \langle\alpha'|\phi\rangle \, , \tag{2.190}$$

in which we recognize $\phi(\alpha') = \langle\alpha'|\phi\rangle$ and $\psi^*(\alpha) = \langle\psi|\alpha\rangle$. Thus, we obtain the result

$$\langle\psi|A|\phi\rangle = \int d\alpha \int d\alpha' \, \psi^*(\alpha) \, A(\alpha,\alpha') \, \phi(\alpha') \, , \tag{2.191}$$

which is the continuous analog of the matrix expression we wrote earlier. As a final example, consider the operator product of the operators

$$A = \int d\alpha \int d\alpha' \ |\alpha\rangle \, A(\alpha,\alpha') \, \langle\alpha'| \tag{2.192}$$

and

$$B = \int d\alpha \int d\alpha' \ |\alpha\rangle \, B(\alpha,\alpha') \, \langle\alpha'| \, . \tag{2.193}$$

The product $C = AB$ has an expansion

$$C = \int d\alpha \int d\alpha' \ |\alpha\rangle \, C(\alpha,\alpha') \, \langle\alpha'| \tag{2.194}$$

in which

$$C(\alpha,\alpha') = \langle\alpha|C|\alpha'\rangle = \langle\alpha|AB|\alpha'\rangle = \int d\alpha'' \, \langle\alpha|A|\alpha''\rangle\langle\alpha''|B|\alpha'\rangle. \tag{2.195}$$

Thus, we find that

$$C(\alpha,\alpha') \ = \int d\alpha'' \, A(\alpha,\alpha'') \, B(\alpha'',\alpha') \tag{2.196}$$

which is the continuous analog of a matrix multiplication.
The adjoint of an operator has a kernel which is the continuous analog of the "complex-conjugate transpose", i.e.,

$$A^{+}(\alpha,\alpha') = A^*(\alpha',\alpha). \tag{2.197}$$

For the kernel representing a Hermitian operator in a continuous basis we have the simpler relationship $A^*(\alpha,\alpha') = A(\alpha',\alpha)$.

**Examples:**
As an example, the operator $X$ has as its matrix elements in the position representation

$$\langle\vec{r}\,'|X|\vec{r}\rangle = \langle\vec{r}\,'|(x|\vec{r}\rangle) = x\langle\vec{r}\,'|\vec{r}\rangle = x\,\delta(\vec{r}\,' - \vec{r}) \tag{2.198}$$

This allows us to construct the expansion for this operator

$$X = \int d^3\vec{r}\,' \int d^3\vec{r} \ |\vec{r}\,'\rangle x\delta(\vec{r}\,' - \vec{r})\langle\vec{r}| = \int d^3\vec{r} \ |\vec{r}\rangle x\langle\vec{r}| \tag{2.199}$$

where the double integral has been reduced to a single integral because of the delta function. The operator $X$ is said to be **diagonal** in the position representation, because it has no nonzero elements connecting different states in this representation. This concept of diagonality extends beyond the position representation. In particular, an operator $A$ is said to be diagonal in the $\{|i\rangle\}$ representation if

$$A_{ij} = \langle i|A|j\rangle = A_i\delta_{ij} \tag{2.200}$$

so that

$$A = \sum_{i,j} |i\rangle\, A_{ij}\, \langle j| = \sum_{i,j} |i\rangle\, A_i\delta_{ij}\, \langle j| = \sum_i |i\rangle\, A_i \langle i|, \tag{2.201}$$

which only has one summation index, in contrast to the general form which requires two. In a representation in which an operator is diagonal, therefore, it is represented by a diagonal matrix

$$\mathbf{A} = \begin{pmatrix} A_1 & 0 & 0 & \cdots \\ 0 & A_2 & 0 & \cdots \\ 0 & 0 & A_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Similarly, in a continuous representation $\{|\alpha\rangle\}$, an operator $A$ is diagonal if

$$A(\alpha, \alpha') = \langle\alpha|A|\alpha'\rangle = A(\alpha)\delta(\alpha - \alpha'), \tag{2.202}$$

so that

$$A = \int d\alpha \int d\alpha'\ |\alpha\rangle A(\alpha, \alpha')\langle\alpha'| = \int d\alpha \int d\alpha'\ |\alpha\rangle A(\alpha)\delta(\alpha - \alpha')\langle\alpha'|, \tag{2.203}$$

or

$$A = \int d\alpha\ |\alpha\rangle A(\alpha)\langle\alpha|. \tag{2.204}$$

It is easy to show that in any basis in which an operator is diagonal, it is what we referred to earlier as a "multiplicative operator". That is, if

$$G = \int d\alpha\ |\alpha\rangle\, g(\alpha)\, \langle\alpha|, \tag{2.205}$$

is diagonal in the $\{|\alpha\rangle\}$ representation, and if $|\psi\rangle = \int d\alpha\ \psi(\alpha)|\alpha\rangle$, then

$$G|\psi\rangle = \int d\alpha\ |\alpha\rangle\, g(\alpha)\, \langle\alpha|\psi\rangle = \int d\alpha\ \ [g(\alpha)\, \psi(\alpha)]|\alpha\rangle, \tag{2.206}$$

which shows that a diagonal operator $G$ acts in the $\{|\alpha\rangle\}$ representation to multiply the wave function by $g(\alpha)$. We list below some additional operators and matrix elements in the basis in which they are diagonal. Derivation is straightforward and left as an exercise.

1. The position operator

$$\vec{R} = \int d^3\vec{r}\ |\vec{r}\rangle\, \vec{r}\, \langle\vec{r}| \qquad\qquad \langle\vec{r}\,'|\vec{R}\,|\vec{r}\rangle = \vec{r}\,\delta(\vec{r} - \vec{r}\,') \tag{2.207}$$

2. The potential energy operator

$$V = \int d^3\vec{r}\ |\vec{r}\rangle\, V(\vec{r})\, \langle\vec{r}| \qquad\qquad \langle\vec{r}\,'|V\,|\vec{r}\rangle = V(\vec{r})\,\delta(\vec{r} - \vec{r}\,'). \tag{2.208}$$

3. The wavevector operator

$$\vec{K} = \int d^3\vec{k}\ |\vec{k}\rangle\, \vec{k}\, \langle\vec{k}|, \qquad \langle\vec{k}\,'|\vec{K}\,|\vec{k}\rangle = \vec{k}\,\delta(\vec{k}-\vec{k}\,'), \tag{2.209}$$

4. The momentum operator

$$\vec{P} = \hbar\vec{K} = \int d^3\vec{k}\ |\vec{k}\rangle\, \hbar\vec{k}\, \langle\vec{k}| \qquad \langle\vec{k}\,'|\vec{P}\,|\vec{k}\rangle = \hbar\vec{k}\,\delta(\vec{k}-\vec{k}\,'). \tag{2.210}$$

5. The kinetic energy operator

$$T = \frac{\hbar^2 K^2}{2m} = \int d^3\vec{k}\ |\vec{k}\rangle\, \frac{\hbar^2 k^2}{2m}\, \langle\vec{k}| \qquad \langle\vec{k}\,'|T\,|\vec{k}\rangle = \frac{\hbar^2 k^2}{2m}\,\delta(\vec{k}-\vec{k}\,'). \tag{2.211}$$

It is important to point out that, although an operator may be diagonal in one representation, it is generally not diagonal in most others. As an additional example we work out below the matrix elements of the wavevector operator $\vec{K}$ in the position representation. Recall that the wavevector operator $\vec{K} = -i\vec{D}$ is a differential operator in the position representation. This means that for any state $|\psi\rangle = \int d^3r\ \psi(\vec{r})|\vec{r}\rangle$, the state $\vec{K}|\psi\rangle$ is given by the expansion

$$\vec{K}|\psi\rangle = \int d^3r\ \left[-i\vec{\nabla}\psi(\vec{r})\right]|\vec{r}\rangle. \tag{2.212}$$

On the other hand, we know that we can always write

$$\vec{K} = \int d^3r \int d^3r'\ |\vec{r}\rangle\, \vec{K}(\vec{r},\vec{r}\,')\, \langle\vec{r}\,'| \tag{2.213}$$

where $\vec{K}(\vec{r},\vec{r}\,') = \langle\vec{r}\,|\,\vec{K}\,|\,\vec{r}\,'\rangle$, so that

$$\vec{K}|\psi\rangle = \int d^3r \int d^3r'\ |\vec{r}\rangle\, \vec{K}(\vec{r},\vec{r}\,')\, \langle\vec{r}\,'|\psi\rangle = \int d^3r\ \left[\int d^3r'\ \vec{K}(\vec{r},\vec{r}\,')\,\psi(\vec{r}\,')\right]|\vec{r}\rangle. \tag{2.214}$$

Comparing the last equation to (2.212) we deduce that for any wavefunction $\psi(\vec{r})$,

$$\int d^3r'\ \vec{K}(\vec{r},\vec{r}\,')\,\psi(\vec{r}\,') = -i\vec{\nabla}\psi(\vec{r}).$$

Comparing this to the basic property associated with the *gradient of the delta function*, namely

$$\int d^3r'\ \vec{\nabla}\delta(\vec{r}-\vec{r}\,')\,f(\vec{r}\,') = f'(\vec{r}),$$

which holds for any function $f(\vec{r})$, we deduce that $\vec{K}(\vec{r},\vec{r}\,') = -i\vec{\nabla}\delta(\vec{r}-\vec{r}\,')$. Thus the matrix elements of the wavevector operator in the position representation take the form

$$\langle\vec{r}\,|\,\vec{K}\,|\,\vec{r}\,'\rangle = -i\vec{\nabla}\delta(\vec{r}-\vec{r}\,').$$

Although these matrix elements appear to be zero everywhere, *the wavevector operator is not diagonal in the position representation.* The reason for this is that the gradient of the delta function (or the derivatives of the delta function in general), represent a limiting process involving the difference between two values of a function infinitesimally displaced from the diagonal (e.g., $\vec{r} = \vec{r} \pm \delta\vec{r}$). The example above is useful in that it shows that even *differential* operators can legitimately be viewed as having matrix elements, although their matrix elements involve very odd distributional functions, such as the derivatives of delta functions. In a similar fashion, the following matrix elements are readily established

$$\langle \vec{k} \,|\, \vec{R} \,|\, \vec{k}\,' \rangle = i\vec{\nabla}_k \delta(\vec{k} - \vec{k}\,'), \tag{2.215}$$

$$\langle\, \vec{r}\,|\, T \,\,|\, \vec{r}\,'\rangle = -\frac{\hbar^2}{2m}\nabla^2\delta(\vec{r} - \vec{r}\,') \tag{2.216}$$

where in the last line it is the Laplacian of the delta function that appears in the expression for the matrix elements of the kinetic energy operator.

### *2.2.11 Canonical Commutation Relations*

It is clear that in the space of a single quantum mechanical particle, there is a very close relationship between the position operator $\vec{R}$ and the wavevector operator $\vec{K}$, or equivalently, the momentum operator $\vec{P} = \hbar\vec{K}$. This relationship is often expressed in terms of the commutation relations between the different cartesian components of these operators. These relations, which are referred to as **canonical commutation relations** are easy to derive. We note first that the Cartesian components of the position operator commute with one another, i.e., their action on the basis in which they are diagonal shows, e.g., that

$$XY|\vec{r}\rangle = xy|\vec{r}\rangle = YX|\vec{r}\rangle \tag{2.217}$$

since this is true for each element of an ONB we deduce an operator identity $XY = YX$, or $[X, Y] = 0$. This extends to the operator $Z$ as well, so that we can write, quite generally,

$$[X_i, X_j] = 0. \tag{2.218}$$

By an analogous argument it is found that the Cartesian components of the wavevector or momentum operator commute with one another, i.e.,

$$[K_i, K_j] = 0 = [P_i, P_j]. \tag{2.219}$$

On the other hand, the Cartesian components of position do not generally commute with the Cartesian components of wavevector or momentum. To see this it is useful to work in a specific representation (either one would suffice). In the position representation, we note that for an arbitrary state $|\psi\rangle$ represented by the wave function $\psi(\vec{r})$,

$$X_iK_j\psi(\vec{r}) = x_i\left(-i\frac{\partial\psi}{\partial x_j}\right) = -ix_i\frac{\partial\psi}{\partial x_j}. \tag{2.220}$$

On the other hand,

$$K_jX_i\psi(\vec{r}) = -i\frac{\partial}{\partial x_j}\{x_i\psi(\vec{r})\} = -i\{\delta_{ij}\psi(\vec{r}) + x_i\frac{\partial\psi}{\partial x_j}\} \tag{2.221}$$

where we have used the standard relation $\partial x_i/\partial x_j = \delta_{ij}$. Thus the action of the commutator $[X_i, K_{j,}] = X_iK_j - K_jX_i$ on such a state is given in the position representation by the expression

$$[X_i, K_j]\psi(\vec{r}) = -ix_i\frac{\partial\psi}{\partial x_j} + i\{\delta_{ij}\psi(\vec{r}) + x_i\frac{\partial\psi}{\partial x_j}\} = i\delta_{ij}\psi(\vec{r}). \tag{2.222}$$

This being true for all states $|\psi\rangle$ we deduce the operator identity

$$[X_i, K_j] = i\delta_{ij} \tag{2.223}$$

or the equivalent relation expressed in terms of the Cartesian components of the momentum operator

$$[X_i, P_j] = i\hbar\delta_{ij}. \tag{2.224}$$

Thus, the components of position and momentum along the same direction of space do not commute with one another. More generally, one can derive the relationship

$$[\vec{R} \cdot \hat{a}, \vec{K} \cdot \hat{b}] = i\hat{a} \cdot \hat{b} \tag{2.225}$$

for arbitrary components of the position and wavevector operator. These basic commutation relations can be used to develop more complicated commutators involving functions of the position and wavevector (or momentum) operators.

### *2.2.12 Matrix Elements of Unitary Operators (Changing Representation)*

If $U$ is a unitary operator then it obeys the unitarity condition

$$U^+U = UU^+ = \mathbf{1}. \tag{2.226}$$

We can express this relationship in terms of the matrix elements of $U$ in any ONB $\{|\phi_i\rangle\}$ for the space, in the form

$$\sum_k U^+_{ik} U_{kj} = \delta_{ij}. \tag{2.227}$$

In this expression, $U_{kj} = \langle\phi_k|U|\phi_j\rangle$, while the matrix elements of $U^+$ satisfy the relation $U^+_{ik} = U^*_{ki}$. This leads to the result

$$\sum_k U^*_{ki} U_{kj} = \delta_{ij}, \tag{2.228}$$

which, we assert, looks something like the orthonormality relation for a set of vectors. To make this a little more clear, let us define a set of vectors

$$|u_i\rangle = \sum_k U_{ki}|\phi_k\rangle \tag{2.229}$$

and

$$|u_j\rangle = \sum_k U_{kj}|\phi_k\rangle \tag{2.230}$$

whose expansion coefficients in the $\{|\phi_i\rangle\}$ representation are the *columns* of the matrix representing the unitary operator $U$. The inner product of these vectors is

$$\langle u_i|u_j\rangle = \sum_k U^*_{ki} U_{kj} = \delta_{ij}. \tag{2.231}$$

Thus, the set of vectors $\{|u_i\rangle\}$ form an orthonormal set. Since they are also equal in number to the columns of the matrix representing $U$, and therefore to the number of basis vectors in the original basis, these vectors form another ONB for the same space. A unitary operator $U$, therefore, allows us to construct a new orthonormal basis from the original one. In fact, $U$ is precisely that operator which maps the original basis vectors onto the new ones. To see this, note that by construction,

$$|u_i\rangle = \sum_k U_{ki}|\phi_k\rangle = \sum_k \langle\phi_k|U|\phi_i\rangle|\phi_k\rangle. \tag{2.232}$$

Changing the order of the matrix element (a number) and the vector in the last expansion, we obtain

$$|u_i\rangle = \sum_k |\phi_k\rangle\langle\phi_k|U|\phi_i\rangle = \mathbf{1}U|\phi_i\rangle, \tag{2.233}$$

in which we have identified a decomposition of the unit operator in the original basis. Thus, we find that

$$|u_i\rangle = U|\phi_i\rangle. \tag{2.234}$$

Thus, any unitary operator $U$ maps an arbitrary ONB onto another ONB. The inverse of this is also true. Given any two ONB's for the space, there exists a unitary operator which connects them, and which can be used to *change representations* from one basis to the other. Let $\{|\phi_i\rangle\}$ and $\{|\psi_i\rangle\}$ be two arbitrary ONB's for a space. Let $U$ be the operator that maps the $i$th element of the set $\{|\phi_i\rangle\}$ onto the corresponding element of the set $\{|\psi_i\rangle\}$, i.e.,

$$U|\phi_i\rangle = |\psi_i\rangle \quad \text{for } i = 1, 2, \cdots \tag{2.235}$$

We will show that this operator is, in fact, unitary. To see this, note that the matrix elements of this operator in the $\{|\phi_i\rangle\}$ representation are given by the expression

$$U_{ij} = \langle\phi_i|U|\phi_j\rangle = \langle\phi_i|\psi_j\rangle, \tag{2.236}$$

so that we can expand $U$ in the form

$$U = \sum_{i,j} |\phi_i\rangle U_{ij}\langle\phi_j| = \sum_{i,j} |\phi_i\rangle\langle\phi_i|\psi_j\rangle\langle\phi_j| = \sum_j |\psi_j\rangle\langle\phi_j|. \tag{2.237}$$

Note that in the last expression we have again identified a decomposition of the unit operator, allowing for the simpler form. The adjoint of this relation is

$$U^+ = \sum_i |\phi_i\rangle\langle\psi_i|. \tag{2.238}$$

Thus, the product of $U$ and its adjoint gives

$$U^+U = \sum_j\sum_i |\phi_i\rangle\langle\psi_i|\psi_j\rangle\langle\phi_j| = \sum_j\sum_i |\phi_i\rangle\delta_{ij}\langle\phi_j| = \sum_i |\phi_i\rangle\langle\phi_i| = \mathbf{1}, \tag{2.239}$$

the identity operator. Taking the product in reverse order, on the other hand, yields a similar result

$$UU^+ = \sum_j\sum_i |\psi_j\rangle\langle\phi_j|\phi_i\rangle\langle\psi_i| = \sum_j\sum_i |\psi_j\rangle\delta_{ji}\langle\psi_i| = \sum_j |\psi_j\rangle\langle\psi_j| = \mathbf{1}. \tag{2.240}$$

Note that we have used the orthonormality relation associated with each ONB in simplifying these expressions. Hence, the operators $U$ and $U^+$ are unitary. The operator $U$ takes kets in $\{|\phi_i\rangle\}$ to kets in $\{|\psi_i\rangle\}$, while $U^+$ takes kets in $\{|\psi_i\rangle\}$ to kets in $\{|\phi_i\rangle\}$. That is, using the expansion for $U^+$,

$$U^+|\psi_i\rangle = \sum_j |\phi_j\rangle\langle\psi_j|\psi_i\rangle = |\phi_i\rangle, \tag{2.241}$$

In the $\{|\psi_i\rangle\}$ basis, $U^+$ has matrix elements

$$U^+_{ij} = \langle\psi_i|U^+|\psi_j\rangle = \langle\psi_i|\phi_j\rangle. \tag{2.242}$$

Thus, we see that

$$U_{ij} = \langle \phi_i | \psi_j \rangle \tag{2.243}$$

$$U^+_{ij} = \langle \psi_i | \phi_j \rangle. \tag{2.244}$$

The practical use of these matrix elements comes when we wish to transform from one representation to another. We first consider the transformation of vectors.

*a) Transformation of Kets* - Let $|\chi\rangle$ be an arbitrary ket in the space. It can be expanded in either of the two bases considered above, i.e., we can write

$$|\chi\rangle = \sum_i \chi_i \, |\phi_i\rangle \tag{2.245}$$

where $\chi_i = \langle \phi_i | \chi \rangle$, or

$$|\chi\rangle = \sum_i \chi'_i \, |\psi_i\rangle, \tag{2.246}$$

where $\chi'_i = \langle \psi_i | \chi \rangle$. The question is how are the expansion coefficients in these two bases related to one another. To find out we use an appropriate decomposition of the identity operator, i.e., we write

$$\chi_i = \langle \phi_i | \chi \rangle = \langle \phi_i | \mathbf{1} | \chi \rangle = \sum_j \langle \phi_i | \psi_j \rangle \langle \psi_j | \chi \rangle \tag{2.247}$$

But we have seen above, that the quantities $\langle \phi_i | \psi_j \rangle = U_{ij}$ are just the matrix elements of the unitary operator connecting these two bases, while $\langle \psi_j | \chi \rangle = \chi'_j$ is the expansion coefficient in the other basis. Thus we have the relation

$$\chi_i = \sum_j U_{ij} \chi'_j, \tag{2.248}$$

which is of the form of a matrix multiplication

$$\begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} U_{11} & U_{12} & \cdots \\ U_{21} & U_{22} & \cdots \\ \vdots & \vdots & \end{pmatrix} \begin{matrix} \chi'_1 \\ \chi'_2 \\ \vdots \end{matrix} . \tag{2.249}$$

By a similar approach it can be shown that the reverse transformation is effected by the matrix representing $U^+$. Thus, we have the relation

$$\chi'_i = \sum_j U^+_{ij} \chi_j. \tag{2.250}$$

*b) Transformation of Matrices* - If $A$ is an operator it has matrix elements in the two bases considered above of the form

$$A_{ij} = \langle \phi_i | A | \phi_j \rangle \tag{2.251}$$

and

$$A'_{ij} = \langle \psi_i | A | \psi_j \rangle. \tag{2.252}$$

To find the relationship between the matrices representing this operator in these two bases we write

$$A_{ij} = \langle \phi_i | \mathbf{1} A \mathbf{1} | \phi_j \rangle \tag{2.253}$$

and insert decompositions of unity in the $\{|\psi_j\rangle\}$. This yields

$$A_{ij} = \sum_{k,k'} \langle\phi_i|\psi_k\rangle\langle\psi_k|A|\psi_{k'}\rangle\langle\psi_{k'}|\phi_j\rangle \tag{2.254}$$

which we identify from above as

$$A_{ij} = \sum_{k,k'} U_{ik}\ A'_{kk'}\, U^{+}_{k'j} \tag{2.255}$$

which is of the form of a matrix multiplication $\mathbf{A} = \mathbf{U}\mathbf{A}'\mathbf{U}^{+}$, where $\mathbf{A}$ is the matrix with elements $A_{ij}$. The reverse transformation is found in the same way, and yields the result $\mathbf{A}' = \mathbf{U}^{+}\mathbf{A}\mathbf{U}$.

Before considering an example, it should be noted that unitary operators preserve the norm of any vector that they act upon. This is intuitively reasonable, since they have the simple effect of "changing the coordinate system", in the same way that orthogonal transformations do in real vector spaces, but is also quite easy to prove. If $|\chi\rangle$ is an arbitrary vector which is taken by a unitary operator $U$ onto the vector $|\phi\rangle = U|\chi\rangle$ then the squared norm of the transformed vector is given by the relation

$$\langle\phi|\phi\rangle = (\langle\chi|U^{+})(U|\chi\rangle) = \langle\chi|U^{+}U|\chi\rangle = \langle\chi|\chi\rangle \tag{2.256}$$

since $U^{+}U = \mathbf{1}$.

**Example (Extension to Continuous Representations)**
Let $|\psi\rangle$ be an arbitrary vector in the space of a quantum particle in three dimensions, i.e., the space spanned by the vectors $\{|\vec{r}\rangle\}$ of the position representation and by the vectors $\{|\vec{k}\rangle\}$ of the wavevector representation. We can expand the ket $|\psi\rangle$ in either of these two bases, i.e.,

$$|\psi\rangle = \int d^3r\, \psi(\vec{r})\, |\vec{r}\rangle \tag{2.257}$$

where $\psi(\vec{r}) = \langle\vec{r}|\psi\rangle$ and

$$|\psi\rangle = \int d^3k\, \psi(\vec{k})\, |\vec{k}\rangle \tag{2.258}$$

where $\psi(\vec{k}) = \langle\vec{k}|\psi\rangle$. How are the expansion coefficients $\psi(\vec{r})$ related to the expansion coefficients $\psi(\vec{k})$. We can find out in the same way as we just did for the discrete case, i.e., we write

$$\psi(\vec{r}) = \langle\vec{r}|\psi\rangle = \langle\vec{r}|\mathbf{1}|\psi\rangle = \int d^3k\ \langle\vec{r}|\vec{k}\rangle\langle\vec{k}|\psi\rangle \tag{2.259}$$

which we write as

$$\psi(\vec{r}) = \int d^3k\ U(\vec{r},\vec{k})\psi(\vec{k}) \tag{2.260}$$

where the (continuous) matrix elements of the unitary operator connecting these two bases are

$$U(\vec{r},\vec{k}) = \langle\vec{r}|\vec{k}\rangle = \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}. \tag{2.261}$$

Thus, we find that

$$\psi(\vec{r}) = \int d^3k\ \frac{e^{i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}\psi(\vec{k}) \tag{2.262}$$

which, of course, we already knew. Similarly, we find that

$$\psi(\vec{k}) = \int d^3r \; U^+(\vec{k},\vec{r})\psi(\vec{r}) = \int d^3r \; \frac{e^{-i\vec{k}\cdot\vec{r}}}{(2\pi)^{3/2}}\psi(\vec{r}). \tag{2.263}$$

Thus, the Fourier transform is just an example of a unitary transformation from one continuous basis to another.

It is also possible to use the unitary transformation represented by the Fourier transform to derive the matrix elements of some of the operators already encountered. As an example, consider the position operator $\vec{R}$, whose matrix elements in the position representation are given by the expression $\langle \vec{r} \,|\, \vec{R} \,|\, \vec{r}\,' \rangle = \vec{R}(\vec{r},\vec{r}\,') = \vec{r}\delta(\vec{r}-\vec{r}\,')$. The matrix elements in the wavevector representation can be obtained from this by a unitary transformation, i.e.,

$$\begin{aligned}
\vec{R}(\vec{k},\vec{k}\,') &= \int d^3r \int d^3\vec{r}\,' \; U^+(\vec{k},\vec{r})\,\vec{R}(\vec{r},\vec{r}\,')\,U(\vec{r}\,',\vec{k}) \\
&= \frac{1}{(2\pi)^3}\int d^3r \int d^3\vec{r}\,' \;\; e^{-i\vec{k}\cdot\vec{r}}\;\vec{r}\delta(\vec{r}-\vec{r}\,')\;e^{i\vec{k}'\cdot\vec{r}\,'} \\
&= \frac{1}{(2\pi)^3}\int d^3r \;\; \vec{r}\,e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} = i\vec{\nabla}_k\left[\int \frac{d^3r}{(2\pi)^3}\;\; e^{i(\vec{k}-\vec{k}')\cdot\vec{r}}\right]
\end{aligned} \tag{2.264}$$

which gives the result stated earlier without proof, namely, that

$$\langle \vec{k} \,|\, \vec{R} \,|\, \vec{k}\,' \rangle = \vec{R}(\vec{k},\vec{k}\,') = i\vec{\nabla}_k\delta(\vec{k}-k'). \tag{2.265}$$

### *2.2.13 Representation Independent Properties of Operators*

There are a number of properties associated with operators which are independent of the representation used to express them. These properties include:

The **trace** of an operator $A$, denoted by $\mathrm{Tr}(A)$, is the sum of the diagonal elements of any matrix representing the operator, i.e.,

$$\mathrm{Tr}(A) = \mathrm{Tr}(\mathbf{A}) = \sum_i A_{ii} = \sum_i \langle i|A|i\rangle \tag{2.266}$$

for any orthonormal basis of states $\{|i\rangle\}$. In a continuous basis, by definition,

$$\mathrm{Tr}(A) = \int d\alpha\, A(\alpha,\alpha). \tag{2.267}$$

The trace of an operator (or matrix) has many interesting properties. It is easily verified, e.g., that in any finite-dimensional space the trace of a product of matrices (or operators) is invariant under cyclic permutation of the elements in the product. That is,

$$\mathrm{Tr}(ABCD) = \mathrm{Tr}(BCDA) = \mathrm{Tr}(CDAB) = \mathrm{Tr}(DABC). \tag{2.268}$$

As an important consequence of this fact, it follows that

$$\mathrm{Tr}(UAU^+) = \mathrm{Tr}(UU^+A) = \mathrm{Tr}(A) \tag{2.269}$$

which shows that the trace of $A$ is invariant under a unitary transformation, and hence is independent of the specific representation used to evaluate it.

The **determinant** of an operator $A$, denoted $\det(A)$, is the determinant of any matrix representing the operator, i.e.,

$$\det(A) = \begin{vmatrix} A_{11} & A_{12} & \cdots & \\ A_{21} & A_{22} & \cdots & \\ \vdots & \vdots & & \\ & & & \end{vmatrix} \tag{2.270}$$

Basic familiarity with general properties of the determinant of a matrix will be assumed. For example, the determinant of a $2 \times 2$ matrix is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc, \tag{2.271}$$

while the determinant of a diagonal matrix is just the product

$$\begin{vmatrix} a_{11} & 0 & 0 & \cdots \\ 0 & a_{22} & 0 & \cdots \\ 0 & 0 & a_{33} & 0 \\ \vdots & \vdots & 0 & \ddots \end{vmatrix} = \prod_i a_i \tag{2.272}$$

of the diagonal elements. Thus, e.g., the identity operator has a determinant of unity, $\det(\mathbf{1}) = 1$. In addition, it turns out that the determinant of a product is equal to the product of the determinants, i.e.,

$$\det(ABC) = \det(A)\det(B)\det(C). \tag{2.273}$$

This last result implies that

$$\det(UAU^{+}) = \det(U)\det(A)\det(U^{+}) = \det(U^{+}U)\det(A) = \det(A), \tag{2.274}$$

in which we have used the result in both directions to recombine the product of the determinant into the determinant of the product $\det(UU^{+}) = \det(\mathbf{1}) = 1$. Thus, the determinant of an operator is also invariant with respect to a change of representation. Finally, you may recall that a necessary and sufficient condition for the inverse of a matrix to exist is that its determinant not vanish. This condition extends to any operator represented by such a matrix, i.e.

If $\det(A) = 0$, then $A$ is **non-invertible** or **singular**.

If $\det(A) \neq 0$, then there exists an inverse operator $A^{-1}$ such that $AA^{-1} = A^{-1}A = \mathbf{1}$.

### *2.2.14 Eigenvalues and Eigenvectors*

A nonzero vector $|\chi\rangle$ is said to be an **eigenvector** of the operator $A$ with eigenvalue $a$ (where generally, $a \in C$) if it satisfies the **eigenvalue equation**

$$A|\chi\rangle = a|\chi\rangle. \tag{2.275}$$

The set of eigenvalues $\{a\}$ for which solutions to this equation exist is referred to as the **spectrum** of the operator $A$, and we write $\text{spectrum}(A) = \{a\}$. The spectrum of an arbitrary operator can be real, complex, continuous, discrete, mixed, bounded, or unbounded.

A number of features follow from the eigenvalue equation. For example, it follows that if $|\chi\rangle$ is an eigenvector of $A$ then so is any multiple $\lambda|\chi\rangle$ of $|\chi\rangle$. This follows from the fact that $A$ is a linear operator so that

$$A(\lambda|\chi\rangle) = \lambda A|\chi\rangle = \lambda a|\chi\rangle = a(\lambda|\chi\rangle). \tag{2.276}$$

Thus, only the direction in Hilbert space of a given eigenvector is unique.

By taking the adjoint of the eigenvalue equation

$$A|\chi\rangle = a|\chi\rangle, \tag{2.277}$$

we see that if $|\chi\rangle$ is an **eigenket** with eigenvalue $a$ then

$$\langle\chi|A^{+} = \langle\chi|a^{*}, \tag{2.278}$$

which implies that $\langle\chi|$ is an **eigenbra** of $A^{+}$ with eigenvalue $a^{*}$.

An eigenvalue $a$ of an operator is **degenerate** if there exists more than one linearly independent eigenvector corresponding to that eigenvalue. The **degeneracy** $n_a$ of an eigenvalue $a$ is equal to the maximum number of linearly independent eigenvectors associated with it. We also say that an eigenvalue with degeneracy $n_a$ is $n_a$-fold degenerate. An eigenvalue with only one linearly independent eigenvector is said to be **nondegenerate**.

It should be clear, that *any* set of linearly independent vectors form a basis for a *subspace* of the original space (namely, the subspace formed from all possible linear combinations of those vectors). It follows, also, that any set of $n$ linearly independent vectors $|\chi_i\rangle$, each of which is an eigenvector of an operator $A$ associated with the *same* $n$-fold degenerate eigenvalue $a$, forms a basis for an entire subspace $S_a$, each vector of which is an eigenvector of $A$ with that same eigenvalue. Again, this follows from the assumption that we are dealing with linear operators, since if

$$A|\chi_i\rangle = a|\chi_i\rangle, \tag{2.279}$$

for $i = 1, 2, \cdots n$, then the action of $A$ on any linear combination

$$|\psi\rangle = \sum_{i=1}^{n} \lambda_i |\chi_i\rangle \tag{2.280}$$

of these vectors is

$$A|\psi\rangle = \sum_{i=1}^{n} \lambda_i A|\chi_i\rangle = \sum_{i=1}^{n} \lambda_i a|\chi_i\rangle = a\sum_{i=1}^{n} \lambda_i |\chi_i\rangle = a|\psi\rangle. \tag{2.281}$$

Thus, any vector $|\psi\rangle$ in $S_a$ is also an eigenvector with the same eigenvalue. Within this subspace we may form linear combinations of the linear independent vectors $|\chi_i\rangle$ using the Gram-Schmidt process to construct an orthonormal basis of eigenvectors for this **eigensubspace**.

From the definitions given above it is readily verified that the basis states of the position representation are eigenstates of the position operator, and are actually labeled by the associated eigenvalues. The position states are also eigenstates of the potential energy operator. Similarly, the basis vectors of the wavevector representation are eigenstates of the wavevector operator and are similarly labeled by their associated eigenvalues. They are also eigenstates of the momentum operator and of the kinetic energy operator.

### *2.2.15 Eigenproperties of Hermitian Operators*

The second postulate (which we introduced some time ago now) associates observables $\mathcal{A}$ with Hermitian operators $A$. The reason for this largely stems from the special properties associated with such operators. These properties include the following:

*Reality of the Eigenvalues* - If $A$ is a Hermitian operator, so that $A = A^+$, and $|\chi\rangle$ is one of its eigenvectors, so that A$|\chi\rangle = a|\chi\rangle$, then

$$\langle\chi|A|\chi\rangle = a\langle\chi|\chi\rangle. \tag{2.282}$$

Now for a Hermitian operator the adjoint of this equation is

$$\langle\chi|A|\chi\rangle = a^*\langle\chi|\chi\rangle. \tag{2.283}$$

Comparing the last two relations we deduce that

$$a^* = a. \tag{2.284}$$

Thus, we conclude that the eigenvalues of Hermitian operators are real. Formerly we showed that expectation values of Hermitian operators are real. The two statements are obviously closely related. The requirement that measurable quantities be real valued motivates the identification of observables with Hermitian operators. Note that, because of the reality of the eigenvalues, the adjoint of the eigenvalue equation for a Hermitian operator has the simple form

$$\langle\chi|A = \langle\chi|a. \tag{2.285}$$

*Orthogonality of Eigenvectors* - It is straightforward to show that eigenvectors of a Hermitian operator corresponding to different eigenvalues are necessarily orthogonal. Let $|\chi\rangle$ and $|\chi'\rangle$, be eigenvectors of a Hermitian operator $A$ corresponding to eigenvalues $a$ and $a'$, respectively. Thus, we can write

$$A|\chi\rangle = a|\chi\rangle, \tag{2.286}$$

and

$$A|\chi'\rangle = a'|\chi'\rangle. \tag{2.287}$$

Taking the inner product of the first of these with $|\chi'\rangle$ we find that

$$\langle\chi'|A|\chi\rangle = a\langle\chi'|\chi\rangle. \tag{2.288}$$

But the adjoint of the second expression shows that

$$\langle\chi'|A = \langle\chi'|a', \tag{2.289}$$

where we have used the reality of the eigenvalues deduced above. Taking the inner product of this equation on the right with $|\chi\rangle$, we find that

$$\langle\chi'|A|\chi\rangle = a'\langle\chi'|\chi\rangle. \tag{2.290}$$

Equating these two expressions for the matrix element $\langle\chi'|A|\chi\rangle$ we find that

$$a\langle\chi'|\chi\rangle = a'\langle\chi'|\chi\rangle, \tag{2.291}$$

or

$$(a - a')\langle\chi'|\chi\rangle = 0. \tag{2.292}$$

There are two ways in which this product can vanish. Either $a = a'$, in which case we haven't found out anything, or $a \neq a'$, in which case we deduce

$$\langle\chi'|\chi\rangle = 0 \tag{2.293}$$

showing that the eigenstates of a Hermitian operator corresponding to two different eigenvalues are always orthogonal.

### *2.2.16 Obtaining Eigenvectors and Eigenvalues*

The methods that one takes to actually solve the eigenvalue problem depend to some extent upon the size of the space that one is working in. For finite dimensional spaces the problem is reduced to a standard one of linear algebra. We seek *nontrivial* solutions to the eigenvalue equation

$$A|\chi\rangle = a|\chi\rangle, \tag{2.294}$$

which means, generally speaking, the two step process of finding the eigenvalues $a$ for which acceptable solutions exist, and then finding the associated eigenvectors. Note first that the *null vector* is always a solution to the eigenvalue equation for any value of $a$, but is of no interest since it does not represent a true dynamical state of the system, and so is referred to as a **trivial solution**. Thus, we seek nontrivial eigenvectors of nonzero length. To this end we rewrite the eigenvalue equation in the form

$$(A - a)|\chi\rangle = 0 \tag{2.295}$$

where $a = a\mathbf{1}$ is a scalar multiplicative operator that multiplies all vectors by the scalar $a$. We now rewrite this a second time, in the form

$$B|\chi\rangle = 0 \tag{2.296}$$

where the operator

$$B = B(a) = A - a \tag{2.297}$$

is an operator function of the parameter $a$. Now, *if* the inverse of $B$ existed, the solution to this equation could be obtained by multiplying both sides by $B^{-1}$:

$$|\chi\rangle = B^{-1}B|\chi\rangle = B^{-1}0 = 0 \tag{2.298}$$

This shows that if $B^{-1}$ exists, the only solution is the trivial one. It follows that for those values of $a$ for which *nontrivial* solutions exist, the operator $B(a)$ *cannot* possess an inverse. Since the inverse of $B$ *will* exist unless the determinant of $B$ vanishes, we conclude that the eigenvalues of $A$ are those values which make $\det(B) = 0$. Thus, we identify the eigenvalues of $A$ with the roots of the **characteristic or secular equation**

$$\det(A - a) = 0. \tag{2.299}$$

Since the determinant of an operator is representation independent, so will be the eigenvalues. Thus, the spectrum of an operator is representation independent, i.e., invariant under unitary transformations, and *any representation can be used to evaluate the determinant.* In a vector space of dimension $N$, the operator $A$ will be represented by an $N \times N$ matrix $\mathbf{A}$, and the characteristic equation will involve a polynomial of degree $N$

$$\det(A - a\mathbf{1}) = \begin{vmatrix} A_{11} - a & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} - a & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} - a \end{vmatrix} = c_0 + c_1 a + \cdots + c_N a^N = 0 \tag{2.300}$$

in the variable $a$, referred to as the **characteristic polynomial**. The fundamental theorem of algebra guarantees that any such polynomial will have, including possible multiplicities, exactly $N$ roots (which will generally be complex numbers). Thus, the characteristic polynomial can generally be factored into the form

$$\det(A - a\mathbf{1}) = (a - a_1)^{n_1}(a - a_2)^{n_2} \cdots (a - a_m)^{n_m} = 0 \tag{2.301}$$

where $a_{1,}a_{2,} \cdots a_m$ represent the $m \leq N$ distinct roots, which are assumed to have multiplicities $n_{1,}n_{2,} \cdots n_m$, such that

$$\sum_{i=1}^{m} n_i = N. \tag{2.302}$$

Once the eigenvalues are found by solving the characteristic equation, the eigenvectors are generally found one at a time by substituting the eigenvalues back into the eigenvalue equation and solving the linear system of equations obtained by expressing the eigenvectors and operators in any convenient representation. Thus, if the states $\{|i\rangle\}$ represent an ONB for the space, the eigenvalue equation $(A - a\mathbf{1})|\chi\rangle$ can be written

$$\sum_j (A_{ij} - a\delta_{ij})\chi_j = 0 \qquad\qquad i = 1, 2, \cdots N, \tag{2.303}$$

in terms of the matrix elements of the operator $A$ in this representation and the coefficients $\chi_i$ of the expansion for the eigenvector

$$|\chi\rangle = \sum_i \chi_i |i\rangle \tag{2.304}$$

of interest. Note that while this procedure gives a set of $N$ linear equation in $N$ unknowns which we can solve using standard techniques of linear algebra, the $N$ equations are *not linearly independent*. This follows from the fact that the matrix of coefficients has a vanishing determinant. Indeed, we *chose* the eigenvalue $a$ to make this determinant vanish. The point here is that the solution to these linear equations is *not unique*. Generally, we will find that the equations allow us to determine all of the coefficients in terms of one undetermined coefficient, which can then take any value. But this is equivalent to the observation that a scalar multiple of any eigenvector is also an eigenvector. Indeed, it is this freedom that allows us to produce a set of square-normalized eigenvectors.

We now assert some basic properties of linear algebra without proof. First we define the concept of a **normal operator**. An operator is said to be *normal* if it commutes with its adjoint. Thus, if

$$[A, A^+] = 0 \tag{2.305}$$

then $A$ is normal. It follows that Hermitian operators and unitary operators are normal. It can be shown, that in a finite dimensional vector space the number of linearly independent eigenvectors of a normal operator is always equal to the dimension of the space, and that the multiplicity of any root in the characteristic equation is equal to the degeneracy of that eigenvalue. Thus, the eigenvectors of a normal operator form a basis for the space. Combining this with the orthogonality of the eigenvectors of Hermitian operators, we deduce an important fact: *In any finite dimensional space a Hermitian operator possesses an orthonormal basis of eigenvectors.*

In an orthonormal basis composed of its own eigenstates, an operator is clearly diagonal, since $\langle a_i|A|a_j\rangle = a_i\delta_{ij}$, hence

$$A = \sum_i |a_i\rangle a_i \langle a_i|. \tag{2.306}$$

It follows that the trace of any normal operator is just the sum

$$\mathrm{Tr}\,(A) = \sum_i a_i \tag{2.307}$$

of its eigenvalues (each eigenvalue being included in the sum $n_i$ times), while the determinant of any normal operator is just the product

$$\det\,(A) = \prod_i a_i \tag{2.308}$$

(including degeneracies) of its eigenvalues.

**Extended Example:** Consider a three dimensional vector space. Let the states $\{|1\rangle, |2\rangle, |3\rangle\}$ form an orthonormal basis for this space. Let $A$ be an operator on the space having the following four nonzero matrix elements $\langle 1|A|2\rangle = \langle 2|A|1\rangle = \langle 2|A|3\rangle = \langle 3|A|2\rangle = \sqrt{2}$, with all other matrix elements being equal to zero. In this representation, an arbitrary state $|\psi\rangle = \sum_{i=1}^{3} \psi_i |i\rangle$ can be represented by a column vector with components $\psi_i$ and the operator $A$ can be represented by the matrix

$$\mathbf{A} = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix}.$$

When $A$ acts on $|\psi\rangle$ it produces a state $|\phi\rangle$ which can be represented by the column vector

$$\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix} = \begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix} = \begin{pmatrix} \sqrt{2}\psi_2 \\ \sqrt{2}\psi_1 + \sqrt{2}\psi_3 \\ \sqrt{2}\psi_2 \end{pmatrix}.$$

To find the eigenvalues of $A$ we find the roots of the characteristic equation

$$\det(A - a) = \begin{vmatrix} -a & \sqrt{2} & 0 \\ \sqrt{2} & -a & \sqrt{2} \\ 0 & \sqrt{2} & -a \end{vmatrix} = 0.$$

Evaluating the determinant yields the condition $a(a^2 - 4) = 0$, which has three distinct roots. $a_1 = -2$, $a_2 = 0$, $a_3 = +2$. Thus, we can write spectrum$(A) = \{-2, 0, 2\}$. Let us denote the eigenvectors of $A$ by $\{|a_1\rangle, |a_2\rangle, |a_3\rangle\}$. To find $|a_1\rangle$ we substitute into the eigenvalue equation $(A - a_1)|a_1\rangle = (A + 2)|a_1\rangle = 0$, which in this representation can be written

$$\begin{pmatrix} 2 & \sqrt{2} & 0 \\ \sqrt{2} & 2 & \sqrt{2} \\ 0 & \sqrt{2} & 2 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} = 0 = \begin{pmatrix} 2a_{11} + \sqrt{2}a_{12} \\ \sqrt{2}a_{11} + 2a_{12} + \sqrt{2}a_{13} \\ \sqrt{2}a_{12} + 2a_{13} \end{pmatrix}.$$

From the first and last components of the column vector on the right we find that $a_{12} = -\sqrt{2}a_{11} = -\sqrt{2}a_{13}$. [Note that we can always ignore the equation from one of the components: it will just give us redundant information because $\det(A - a_1) = 0$]. Thus the column vector with components $\{1, -\sqrt{2}, 1\}$ is an eigenvector with eigenvalue $a_1 = -2$, but is not normalized. To normalize, we divide by the length to obtain the column vector

$$|a_1\rangle \rightarrow \begin{pmatrix} 1/2 \\ -1/\sqrt{2} \\ 1/2 \end{pmatrix}$$

representing the state $|a_1\rangle$. This means that

$$|a_1\rangle = \frac{1}{2}\left(|1\rangle - \sqrt{2}|2\rangle + |3\rangle\right)$$

is an eigenvector of $A$ with eigenvalue $a_1$. By repeating this process for $a_2$ and $a_3$ we find after a little work that

$$|a_2\rangle = \frac{1}{\sqrt{2}}\left(|1\rangle - |3\rangle\right) \qquad\qquad |a_3\rangle = \frac{1}{2}\left(|1\rangle + \sqrt{2}|2\rangle + |3\rangle\right).$$

Since $A$ is obviously Hermitian, these states are orthonormal, as is readily verified by computing the inner products. Since they are orthonormal, they also form a basis for the space. It readily verified that the original basis states can be rewritten in this new set as follows

$$|1\rangle = \frac{1}{2}\left(|a_1\rangle + \sqrt{2}|a_2\rangle + |a_3\rangle\right) \qquad |2\rangle = \frac{1}{\sqrt{2}}\left(-|a_1\rangle + |a_3\rangle\right) \qquad |3\rangle = \frac{1}{2}\left(|a_1\rangle - \sqrt{2}|2\rangle + |3\rangle\right)$$

The coefficients in these expansions are just the inner products of the new basis vectors with the old. The unitary matrices which transform between these two sets of orthonormal basis vectors are obtained by arranging in columns the coefficients expressing each set of vectors of one basis in terms of the other, i.e.,

$$\mathbf{U}^{+} = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \qquad\qquad \mathbf{U} = \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix}.$$

It is readily verified that the product of $\mathbf{U}$ and $\mathbf{U}^{+}$ yields the $3 \times 3$ identity matrix. The matrix $\mathbf{A}'$ representing the operator $A$ in the basis $\{|a_i\rangle\}$ of its eigenstates is obtained by transforming the matrix $\mathbf{A}$ using these unitary matrices, i.e., $\mathbf{A}' = \mathbf{U}^{+}\mathbf{A}\mathbf{U}$. We find that

$$\mathbf{A}' = \begin{pmatrix} 1/2 & -1/\sqrt{2} & 1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & 1/\sqrt{2} & 1/2 \end{pmatrix} \begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1/2 & 1/\sqrt{2} & 1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{pmatrix}$$

or

$$\mathbf{A}' = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

which is as we expect: In the basis of its own eigenstates, $A$ is diagonal and the diagonal elements are just the associated eigenvalues. We can see that $\mathrm{Tr}(A) = 0 = \det(A)$, the latter of which implies that this operator has no inverse.

**Extension to Infinite Dimensional Spaces**

The extension of these ideas to infinite dimensional spaces is often non-trivial. There is the obvious problem of taking the determinant of an infinite-dimensional matrix, but additional complications can also get in the way. One problem is that in contrast to what happens in finite dimensional spaces, it is not always true that the eigenvectors of a Hermitian operator span the space. Indeed, it is not all that unusual for a Hermitian operator to have no eigenvectors in the space at all. An example should help to make this clear. Consider the space $L^2(R^n)$ of square-integrable functions on $R^n$, which is a subspace of the larger set of Fourier transformable functions. The position operator $\vec{R}$ is Hermitian in this space, since for any two square-integrable functions $\phi(\vec{r})$ and $\psi(\vec{r})$ we have

$$\begin{aligned} \langle\phi|\vec{R}|\psi\rangle &= \int d^3r\,\phi^*(\vec{r})\,\vec{r}\,\psi(\vec{r}) = \langle\phi|\vec{R}|\psi\rangle \\ &= [\int d^3r\,\psi^*(\vec{r})\,\vec{r}\,\phi(\vec{r})]^* = \langle\psi|\vec{R}|\phi\rangle^*. \end{aligned} \tag{2.309}$$

But the eigenstates of the position operator are represented by the wave functions

$$\phi_{\vec{r}'}(\vec{r}) = \langle\vec{r}\,|\,\vec{r}^{\,'}\rangle = \delta(\vec{r}-\vec{r}') \tag{2.310}$$

which are *not* square integrable, since

$$\int d^3r\;\phi^*_{\vec{r}'}(\vec{r})\phi_{\vec{r}'}(\vec{r}) = \int d^3r\,\delta(\vec{r}-\vec{r}')\delta(\vec{r}-\vec{r}') = \delta(\vec{r}'-\vec{r}') = \infty. \tag{2.311}$$

Thus, the Hermitian operator $\vec{R}$ has *no* eigenfunctions in the space of square integrable functions. Neither does the wavevector operator $\vec{K}$. It is for this reason that we have chosen the space in which we are working to be the *space of Fourier transformable functions* (which includes the delta functions and plane waves). This does not entirely dispose of the problem, but it does allow us to confront it less frequently. It has become common to essentially define the problem away by mathematically introducing the concept of an *observable.* A Hermitian operator is said to be an **observable** for a space $S$ if it possesses a complete orthonormal basis of eigenvectors for the space. This mathematical definition makes close physical contact with our earlier intuition involving the principle of spectral decomposition. As examples of observables, we have in the space of Fourier transformable functions the operators $X, Y, Z, \vec{R}$, and $V(\vec{R})$, which are all diagonal in the position representation, and the operators $K_x, K_y, K_z, \vec{K}, \vec{P}$,and $T = \hbar^2K^2/2m$ which are all diagonal in the wavevector representation.

In many cases the solution to the eigenvector problem in infinite dimensional spaces takes the form of an *integro-differential equation.* To see how this arises, consider, for a single particle, the Hamiltonian operator

$$H = \frac{P^2}{2m} + V \tag{2.312}$$

the first term of which is diagonal in the wavevector representation, the second term $V = V(\vec{R})$ is diagonal in the position representation. The operator $H$, however, is generally not diagonal in either. To find solutions to the eigenvalue equation

$$H|\phi\rangle = E|\phi\rangle \tag{2.313}$$

we choose a representation in which to express them.

*Position Representation* - In the position representation, the eigenvector

$$|\phi\rangle = \int d^3r\,\phi(\vec{r})|\vec{r}\rangle \tag{2.314}$$

is represented by the eigenfunction $\phi(\vec{r})$. In this representation, the kinetic energy is a differential operator, while the potential energy is a multiplicative operator. Projecting the eigenvalue equation onto a basis vector of the position representation we find that

$$\begin{aligned}\langle\vec{r}|H|\phi\rangle &= \langle\vec{r}|\frac{P^2}{2m}|\phi\rangle + \langle\vec{r}|V|\phi\rangle \\ &= -\frac{\hbar^2}{2m}\nabla^2\phi(\vec{r}) + V(\vec{r})\phi(\vec{r}) = E\langle\vec{r}|\phi\rangle = E\phi(\vec{r}). \end{aligned} \tag{2.315}$$

Thus, we recover the energy eigenvalue equation

$$-\frac{\hbar^2}{2m}\nabla^2\phi(\vec{r}) + V(\vec{r})\phi(\vec{r}) = E\phi(\vec{r}) \tag{2.316}$$

as it arises in Schrödinger's mechanics. The eigenvalues of $H$ are then identified with those solutions which lie within the relevant space, which in our case, is the space of Fourier transformable functions. Thus, for example, in the limit that $V = 0$, this leads us to the plane waves, but it excludes exponentials of the form $\phi(\vec{r}) = \exp(\vec{\alpha}\cdot\vec{r})$, which for real vectors $\vec{\alpha}$ would diverge exponentially and not be Fourier transformable. Thus negative energies are excluded from the free particle spectrum.

*Wavevector Representation* - To find the form that the eigenvector equation takes in the wavevector representation, we project onto a basis vector of this representation. Thus, we consider

$$\langle\vec{k}|H|\phi\rangle = \langle\vec{k}|\frac{P^2}{2m}|\phi\rangle + \langle\vec{k}|V|\phi\rangle. \tag{2.317}$$

In this representation, $P^2$ is a multiplicative operator

$$\langle\vec{k}|\frac{P^2}{2m}|\phi\rangle = \frac{\hbar^2k^2}{2m}\phi(\vec{k}) \tag{2.318}$$

while the potential energy term can be written

$$\begin{aligned}\langle\vec{k}|V|\phi\rangle &= \int d^3k'\,\langle\vec{k}|V|\vec{k}'\rangle\langle\vec{k}'|\phi\rangle \\ &= \int d^3k'\,V(\vec{k},\vec{k}')\phi(\vec{k}'). \end{aligned} \tag{2.319}$$

Combining this with the result $\langle\vec{k}|E|\phi\rangle = E\phi(\vec{k})$, the eigenvalue equation in the wavevector representation takes the form of an integral equation

$$\frac{\hbar^2k^2}{2m}\phi(\vec{k}) + \int d^3k'\,V(\vec{k},\vec{k}')\phi(\vec{k}') = E\phi(\vec{k}). \tag{2.320}$$

As before, the eigenvalues are determined by the requirement that the solutions to this equation be in the space of interest. The question remains as to the form of the kernel $V(\vec{k},\vec{k}')$ representing the potential energy in the $\vec{k}$-representation. Inserting a decomposition of $V$ in the $\vec{r}$-representation in which it is diagonal we find that

$$\begin{aligned}V(\vec{k},\vec{k}') &= \langle\vec{k}|V|\vec{k}'\rangle = \int d^3r\,\langle\vec{k}|\vec{r}\rangle V(\vec{r})\langle\vec{r}|\vec{k}'\rangle \\ &= \int \frac{d^3r}{(2\pi)^3}\,V(\vec{r})\,e^{-i(\vec{k}-\vec{k}')\cdot\vec{r}} = \tilde{V}(\vec{k}-\vec{k}'), \end{aligned} \tag{2.321}$$

where $\tilde{V}(\vec{q})$ is the Fourier transform of $V(\vec{r})$.

### *2.2.17 Common Eigenstates of Commuting Observables*

Consider two observables $A$ and $B$. Each observable possesses a complete ONB of eigenstates, but the eigenstates of $A$ need have no relation to the eigenstates of $B$. Suppose, however, that the two observables commute with one another, so that $[A, B] = 0$, or $AB = BA$, and let $|a\rangle$ be an eigenstate of $A$ so that

$$A|a\rangle = a|a\rangle. \tag{2.322}$$

The action of $A$ on the state $B|a\rangle$ is then easily determined, i.e.,

$$A(B|a\rangle) = AB|a\rangle = BA|a\rangle = Ba|a\rangle = a(B|a\rangle) \tag{2.323}$$

or more simply,

$$A(B|a\rangle) = a(B|a\rangle). \tag{2.324}$$

From which we see that $B|a\rangle$ is also an eigenstate of $A$ with the same eigenvalue as the state $|a\rangle$. This leaves two possibilities.

*(i)* If $a$ is nondegenerate, then there is at most one linearly-independent eigenvector with this eigenvalue. This means that the states $|a\rangle$ and $B|a\rangle$ are linearly dependent, which for *two* vectors means that one is a scalar multiple of the other. Hence, there must exist a scalar $b$, say, such that

$$B|a\rangle = b|a\rangle. \tag{2.325}$$

But this shows that *if* $a$ is nondegenerate, and $[A, B] = 0$, then $|a\rangle$ is also an eigenvector of $B$. In this circumstance, we can label the state by the (real) eigenvalues $a$ and $b$, and write $|a\rangle = |a, b\rangle$. A nondegenerate eigenvector of an observable $A$ is also an eigenvector of any observable that commutes with $A$.

*(ii)* If $a$ is degenerate, with degeneracy $n_a$, then all we can say is that both $|a\rangle$ and $B|a\rangle$ are elements of the same $n_a$ dimensional eigensubspace $S_a$ of the operator $A$. But this implies that an operator $B$ which commutes with $A$ takes vectors in $S_a$ only onto other vectors in $S_a$. This is described by saying that $S_a$ is **globally-invariant** under the action of $B$. Recall that eigenvectors of $A$ with different eigenvalues are orthogonal to each other. This implies that if $|a'\rangle$ is an eigenvector of $A$ with eigenvalue $a' \neq a$, then

$$\langle a'|a\rangle = 0 \tag{2.326}$$

and that

$$\langle a'|B|a\rangle = 0. \tag{2.327}$$

Thus, if $B$ commutes with $A$ its matrix elements cannot connect states in different eigensubspaces. It is useful to see what all this looks like. In a basis of eigenstates of the operator $A$, the operator $A$ itself is represented by a diagonal matrix whose diagonal elements are the eigenvalues of $A$. In one such basis, e.g., $A$ may be represented by the matrix

$$A \longleftrightarrow \begin{bmatrix} a_1 & & & & & & & & \\ & a_1 & & & & & & & \\ & & a_1 & & & & & & \\ & & & a_2 & & & & & \\ & & & & \ddots & & & & \\ & & & & & a_2 & & & \\ & & & & & & a_3 & & \\ & & & & & & & \ddots & \\ & & & & & & & & a_3 \end{bmatrix} \tag{2.328}$$

where we have only shown the nonzero elements, and where, e.g., the eigenvalue $a_1$ is threefold degenerate, and so on. Each diagonal string of the same eigenvalue is associated with the associated eigensubspace of that dimension. In this basis, the operator $B$ only connects sites within each subspace. Thus, in this representation, $B$ must be represented by a matrix which is of **block diagonal** form

$$B \longleftrightarrow \begin{bmatrix} B_{11} & B_{12} & B_{13} & & & & & & \\ B_{21} & B_{22} & B_{23} & & & & & & \\ B_{31} & B_{32} & B_{33} & & & & & & \\ & & & B_{44} & \cdots & B_{4m} & & & \\ & & & \vdots & \ddots & \vdots & & & \\ & & & B_{m4} & \cdots & B_{mm} & & & \\ & & & & & & \cdots & \cdots & \cdots \\ & & & & & & \cdots & \cdots & \cdots \\ & & & & & & \cdots & \cdots & \cdots \end{bmatrix} \tag{2.329}$$

in which each block associated with the eigensubspaces of $A$ have nonzero matrix elements, which are arbitrary (except for the restriction that the diagonal subblocks be represented by Hermitian submatrices), while the matrix elements outside these blocks are all zero.

Given this block diagonal form, however, we can now form linear combinations of the basis vectors in each eigensubspace to find new basis vectors for each eigensubspace without affecting what happens in the other eigenspaces. This means that we can choose the basis vectors in each eigenspace so that they *diagonalize* the submatrix representing the operator $B$ in that eigensubspace. If we do this for each eigensubspace we generate a new basis of eigenvectors of $A$ in which the matrix representing the operator $A$ will be the same as in the original, but the matrix representing the operator $B$ will be represented by a diagonal matrix

$$\begin{bmatrix} b_{11} & & & & & & & & \\ & b_{12} & & & & & & & \\ & & b_{13} & & & & & & \\ & & & b_{21} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & & \ddots & \\ & & & & & & & & \ddots \end{bmatrix} \tag{2.330}$$

Thus, it is always possible to find a simultaneous ONB $\{|a, b, \tau\rangle\}$ of eigenstates associated with any two commuting observables. (We write the basis in this form, recognizing that there may be more than one linearly-independent basis vector labeled by the same pair of eigenvalues $a$ and $b$. The index $\tau$ distinguishes between such otherwise identically-labeled basis vectors.) Thus, even when the eigenvalues of two commuting operators are degenerate, *it always possible to find a complete orthonormal basis of eigenstates common to both.*

Let us consider an example intended to make these ideas a bit more concrete . The **parity** operator $\Pi$ in one dimension is defined through its action on the position eigenstates $|x\rangle$ as follows

$$\Pi|x\rangle = |-x\rangle. \tag{2.331}$$

Thus $\Pi$ takes a particle located at $x$ and moves it to $-x$. Equivalently, it reflects the particle through the origin, and it is straightforward to deduce that in the $x$ representation $\Pi\psi(x) = \psi(-x)$. It is then a straightforward exercise to show that this definition leads to a similar action

$$\Pi|k\rangle = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{ikx}\, \Pi|x\rangle = \Pi \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{ikx}\, |\text{-}\,x\rangle = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx}\, |x\rangle = |\text{-}\,k\rangle \quad (2.332)$$

on the eigenstates of the wavevector operator, i.e., $\Pi|k\rangle = |\text{-}\,k\rangle$. Thus, the parity operator also can be viewed as reversing the direction of motion of a particle. Consider, now, a particle moving in a symmetric potential $V(x)$. The total energy of such a particle is associated with the Hamiltonian operator

$$H = T + V = \frac{\hbar^2 K^2}{2m} + V. \quad (2.333)$$

It is readily verified that $\Pi$ commutes with the kinetic energy operator, by considering the action of $\Pi T$ and $T\Pi$ on the states $|k\rangle$, i.e.,

$$\Pi T|k\rangle - T\Pi|k\rangle = \Pi\frac{\hbar^2 k^2}{2m}|k\rangle - T|\text{-}\,k\rangle = \frac{\hbar^2 k^2}{2m}\left[|\text{-}\,k\rangle - |\text{-}\,k\rangle\right] = 0 \quad (2.334)$$

. If $V$ is symmetric about the origin, then $\Pi$ also commutes with $V$, since in the $x$ representation

$$\begin{aligned}[\Pi, V]\,\psi(x) &= \Pi V(x)\psi(x) - V(x)\Pi\psi(x) \\ &= V(-x)\psi(-x) - V(x)\psi(-x) \\ &= V(x)\psi(-x) - V(x)\psi(-x) = 0 \end{aligned} \quad (2.335)$$

Thus the Hamiltonian of a particle moving in a symmetric potential commutes with the parity operator, i.e., $[H, \Pi] = 0$. It follows that it is always *possible* to find a basis of energy eigenstates for this potential that are also states of good parity, i.e., that are simultaneously eigenstates of the parity operator. On the other hand, an eigenstate of $H$ will *necessarily* be an eigenstate of $\Pi$ *only* if it corresponds to a nondegenerate eigenvalue.

Getting back to our general development, we have shown that it is possible to find a simultaneous basis $\{|a, b, \tau\rangle\}$ of eigenstates of two operators $A$ and $B$ whenever they commute. When there are more than one linearly-independent basis vector $|a, b, \tau\rangle$ labeled by the same pair of eigenvalues $a$ and $b$, it may be possible to find a third operator $C$ which commutes with both $A$ and $B$ which we can diagonalize within each simultaneous eigensubspace $S_{ab}$ of $A$ and $B$ to further lift the degeneracy. This process can then be repeated until each basis vector is labeled by a unique set of eigenvalues associated with the mutually commuting operators introduced in order to distinguish each basis vector from the others.

A set $\{A, B, C\}$ of mutually commuting observables is said to form a **complete set of commuting observables** (CSCO) for a space if it possess a complete ONB of simultaneous eigenvectors $\{|a, b, c\rangle\}$ for the space, *each element of which is uniquely determined by the eigenvalues* $(a, b, c)$ *which label it.*

For example, the operators $\{X, Y, Z\}$ form a CSCO for the space of a single spinless particle in 3D, since there exists a complete set of basis vectors $\{|\vec{r}\rangle\} = \{|x, y, z\rangle\}$ each of which are uniquely identified by the eigenvalues $x, y$, and $z$ which label it. For similar reasons the operators $\{K_x, K_y, K_z\}$ form a CSCO, with associated basis vectors $\{|\vec{k}\rangle\} = \{|k_x, k_y, k_z\rangle\}$. It turns out that the one can mix and match Cartesian components of position and momentum to arrive at other CSCO's, provided that one does not include corresponding components of position and momentum along the same direction in space. Examples include sets of operators such as $\{X, K_y, Z\}, \{K_x, K_y, Z\}$, and $\{K_x, Y, Z\}$. The first of these, $\{X, K_y, Z\}$, has a set of basis vectors of the form $\{|x, k_y, z\rangle\}$ corresponding to a particle whose position in the $xz$ plane is known, but whose $y$ coordinate is unknown, and whose motion along the $y$ direction is characterized by wavevector $k_y$ (i.e., whose momentum along the $y$ axis is $p_y = \hbar k_y$). The set of operators $\{K_x, X, Z\}$ do not form a CSCO since they do not mutually commute.

Having discussed the means of specifying the dynamical state of the system, and having described the basic properties associated with the observables of such a system, we now turn to the question of what happens when we actually go to perform a measurement.

## 2.3 Postulate III: The Measurement of Quantum Mechanical Systems

**Postulate III(a)** - (Values obtained during measurement) The only value which can be obtained as a result of an attempt to measure an observable $\mathcal{A}$ of a quantum mechanical system in a normalized state $|\psi\rangle$ is one of the eigenvalues in the spectrum of the Hermitian operator $A$ associated with it. Exactly which eigenvalue will be measured cannot generally be predicted. It is possible, however, to predict the *probability* for obtaining each eigenvalue. We consider separately the cases in which the eigenvalue obtained is (i) discrete and (ii) continuous.

*i*) The probability $P(a)$ that a measurement will yield one of the discrete eigenvalues $a$ of $A$ can be written

$$P(a) = \langle\psi|P_a|\psi\rangle, \tag{2.336}$$

where

$$P_a = \sum_{\tau=1}^{n_a} |a, \tau\rangle\langle a, \tau| \tag{2.337}$$

is the projector onto the eigensubspace $S_a$ of the observable $A$ associated with that eigenvalue. In this last expression the vectors $\{|a, \tau\rangle\}$ are the orthonormal basis vectors of $A$ associated with eigenvalue $a$, and so the operator $P_a$ is the sum of the orthogonal projectors for each such basis vector.

*ii*) The probability density $\rho(\alpha)$ that a measurement will yield one of the continuous eigenvalue $\alpha$ of $A$ can be written

$$\rho(\alpha) = \langle\psi|\rho_\alpha|\psi\rangle, \tag{2.338}$$

where

$$\rho_\alpha = \sum_{\tau=1}^{n_\alpha} |\alpha, \tau\rangle\langle \alpha, \tau|. \tag{2.339}$$

In this expression, the vectors $\{|\alpha, \tau\rangle\}$ are the orthonormal basis vectors of $A$ associated with the continuous eigenvalue $\alpha$. These basis vectors, by assumption, are labeled by the continuous index $\alpha$. They are not square-normalizable. They obey some form of Dirac normalization condition, e.g., $\langle\alpha, \tau|\alpha', \tau'\rangle = \delta(\alpha - \alpha')\delta_{\tau,\tau'}$ and so are of infinite length. Thus, the operator $\rho_a$ is not a projection operator, but a projector density. As we have seen in Sec. 2.2.5, the integral of $\rho_\alpha$ over any region of the allowed values of $\alpha$ does give a projector.

In writing the expression above, it is important to point out that we have assumed that the state $|\psi\rangle$ is square-normalized to unity; otherwise this expression just gives the relative probability of obtaining that value. Note that, in the discrete case, if the eigenvalue $a$ is nondegenerate, so that there is only one linearly independent basis eigenvector $|a\rangle$, then the projector $P_a = |a\rangle\langle a|$ onto the subspace $S_a$ is just the projector onto this one state. In this limit, our expression for the probability $P(a)$ reduces to the one we obtained within Schrödinger's postulates, namely,

$$P(a) = \langle\psi|a\rangle\langle a|\psi\rangle = \psi_a^*\psi_a = |\psi_a|^2, \tag{2.340}$$

where $\psi_a$ is the associated expansion coefficient for the state $|\psi\rangle$ in the basis of eigenstates of $A$. Thus, the probability reduces to the squared modulus of the associated amplitude, exactly as in Schrödinger's mechanics, except that we are now using $\psi_a$ to denote the expansion coefficient rather than $c_a$. Similarly if the continuous eigenvalue $\alpha$ of $A$ is nondegenerate, then the expression for the probability density reduces to

$$\rho(\alpha) = \langle\psi|\alpha\rangle\langle\alpha|\psi\rangle = \psi^*(\alpha)\psi(\alpha) = |\psi(\alpha)|^2. \tag{2.341}$$

Thus, for a single particle, the identification of $\rho(\vec{r}) = |\psi(\vec{r})|^2 = \langle\psi|\vec{r}\rangle\langle\vec{r}|\psi\rangle$ as the probability density to find the particle at $\vec{r}$ follows from the above prescription applied to a measurement of the position operator $\vec{R}$. Similarly, we have the identification of $\rho(\vec{k}) = |\psi(\vec{k})|^2$ for the probability density associated with measurements of wavevector or momentum.

For degenerate eigenvalues, the form that the probability takes actually depends upon whether the different linearly-independent basis vectors associated with the same eigenvalue are discretely or continuously indexed. For a discretely-indexed set the probability can be written using the result stated as above, i.e.,

$$P(a) = \langle\psi|P_a|\psi\rangle = \sum_{\tau=1}^{n_a}\langle\psi|a,\tau\rangle\langle a,\tau|\psi\rangle = \sum_{\tau=1}^{n_a}|\psi_{a,\tau}|^2 \tag{2.342}$$

which is similar to the nondegenerate case except that now we must sum over each basis vector of the associated eigensubspace. Similarly, in if the basis vectors associated with the continuous eigenvalue $\alpha$ are discretely-indexed, the probability density $\rho(\alpha)$ can be written

$$\rho(\alpha) = \langle\psi|\rho_\alpha|\psi\rangle = \sum_\tau\langle\psi|\alpha,\tau\rangle\langle\alpha,\tau|\psi\rangle = \sum_\tau\psi_\tau^*(\alpha)\psi_\tau(\alpha) = \sum_\tau|\psi_\tau(\alpha)|^2 \tag{2.343}$$

(recall our convention that discretely-indexed components are distinguished from one another with a subscript, while continuously-indexed ones use parentheses).

By contrast, if the basis vectors for an eigensubspace $S_a$ associated with an eigenvalue $a$ are continuously-indexed, then the sum over $\tau$ must be replaced by an integral over the continuous index, i.e.

$$P_a = \int d\tau\ |a,\tau\rangle\langle a,\tau| \tag{2.344}$$

so that

$$P(a) = \int d\tau\langle\psi|a,\tau\rangle\langle a,\tau|\psi\rangle = \int d\tau|\psi_a(\tau)|^2. \tag{2.345}$$

The corresponding expressions for the probability density $\rho(\alpha)$ to obtain a continuous eigenvalue $\alpha$ in a basis of continuously-indexed eigenvectors associated with that eigenvalue is

$$\rho_\alpha = \int d\tau\ |\alpha,\tau\rangle\langle\alpha,\tau| \tag{2.346}$$

and

$$\rho(\alpha) = \int d\tau \langle\psi|\alpha,\tau\rangle\langle\alpha,\tau|\psi\rangle = \int d\tau |\psi(\alpha,\tau)|^2. \tag{2.347}$$

To make this a little more concrete, consider as an example a measurement of just the $x$ coordinate of a particle moving in three dimensions. The position eigenstates $|\vec{r}\rangle = |x,y,z\rangle$ form a basis of eigenstates of the operator $X$, but are not completely labeled by the associated (continuous) eigenvalue $x$. Indeed, there are a continuous infinity of states having the same value of $x$, but different values of $y$ and $z$. Thus ,in this example the continuous index $\tau$ is actually a vector in $R^2$ denoting the coordinates $(y,z)$ in the $yz$ plane. Hence, in this example, the projector density associated with the value $x$ is

$$\rho_x = \int dy\, dz\, |x,y,z\rangle\langle x,y,z| \tag{2.348}$$

where the integral is over all points in the $yz$ plane. The associated probability density that a measurement of $X$ will yield the value $x$ is

$$\rho(x) = \int dy\, dz\, \langle\psi|x,y,z\rangle\langle x,y,z|\psi\rangle = \int dy\, dz\, \psi^*(x,y,z)\psi(x,y,z) = \int dy\, dz\, |\psi|^2. \tag{2.349}$$

In our formal development, since the only difference between these expressions is to replace a sum over $\tau$ by an integral we will, in what follows, adopt the practice of using the summation over $\tau$ to stand for either a summation or integration depending on whether the degeneracy of the state is discrete or continuous.

An important part of the traditional interpretation of the measurement process is that the value of an observable $A$ is really not defined unless the system is in an eigenstate associated with that observable. If it is in such an eigenstate, let us call it $|\psi_a\rangle$, then it lies entirely within the associated eigensubspace $S_a$. The act of a projector $P_{a'}$ on such a state will be to annihilate the state if $a \neq a'$, and to leave it alone if $a = a'$. Thus, under such circumstances, it is clear that

$$\begin{aligned} P(a) &= 1 \\ P(a') &= 0 \qquad \text{for } a' \neq a \end{aligned} \tag{2.350}$$

Hence the probability of obtaining the eigenvalue $a$ associated with such a state is equal to unity, which is operationally the only time that the value of the associated observable is well defined.

We now need to address the second part of the measurement postulate which describes what happens to a quantum mechanical system as a result of the measurement process. We assert that in an *ideal* measurement, which is one which perturbs the system the *least*, the state of the system immediately after the measurement is one that (1) is consistent with the particular eigenvalue obtained as a result of the measurement process, and (2) retains as much information about the state of the system immediately before the measurement as is consistent with (1). These ideas form the basis for the following:

**Postulate III(b)** (Collapse of the State Vector) - Immediately after a measurement of an observable $\mathcal{A}$ performed on a system in the state $|\psi\rangle$ that yields the value $a$, the state of the system is the normalized projection of $|\psi\rangle$ onto the eigensubspace $S_a$ associated with the eigenvalue measured, i.e., it is just that part of $|\psi\rangle$ lying within the subspace. We schematically indicate this change in the state vector as follows:

$$|\psi\rangle \quad \cdots \underset{a}{\overset{A}{\rightarrow}} \cdots \quad \frac{P_a|\psi\rangle}{||P_a|\psi\rangle||}. \tag{2.351}$$

Thus, simply speaking, nature throws away those parts of the state vector which are *not* consistent with the actual value obtained. It should be noted that this only indicates one possible *branch* of the evolution of the system during the course of the measurement process, namely, that one which occurs when the particular eigenvalue $a$ is obtained. As we have seen, it is not possible to predict which of these branches will actually be followed by any single quantum mechanical system. Thus, this change in the state vector during measurement is inherently *non-deterministic*.

The viewpoint usually taken is that the collapse of the state vector to the associated eigensubspace occurs as the result of an unspecified interaction of the system with the (classical) measurement device used to measure the observable. We will avoid the many interesting questions and apparent paradoxes which arise in the attempt to simultaneously treat the system-plus-measuring-device as a closed quantum system, as well as any discussion of whether the collapse of the state vector is a "physical" process or a statistical one. We will however make a few practical observations.

First, it is easy to show that the normalization factor $||P_a|\psi\rangle||$ which appears in the denominator of the reduced state vector is simply related to the probability that system would have taken that particular route during the measurement process, i.e.,

$$||P_a|\psi\rangle|| = \sqrt{\langle\psi|P_aP_a|\psi\rangle} = \sqrt{\langle\psi|P_a|\psi\rangle} = \sqrt{P(a)}, \tag{2.352}$$

where we have used the characteristic feature $P_a^2 = P_a$ associated with projection operators along with the fact that a projection operator is Hermitian. Thus, we can express the reduction process in the form

$$|\psi\rangle \quad \cdots \underset{a}{\overset{A}{\rightarrow}} \cdots \quad \frac{P_a|\psi\rangle}{\sqrt{\langle\psi|P_a|\psi\rangle}}. \tag{2.353}$$

Viewed from the other direction, this shows that the probability $P(a)$ is just the squared norm of that part of $|\psi\rangle$ lying within the eigensubspace.

Secondly, we should note that if $a$ is nondegenerate, then $P_a = |a\rangle\langle a|$, and

$$\frac{P_a|\psi\rangle}{||P_a|\psi\rangle||} = \frac{|a\rangle\langle a|\psi\rangle}{|\langle a|\psi\rangle|} = e^{i\phi}|a\rangle, \tag{2.354}$$

where all that is left of the original state is the phase information contained in the factor

$$e^{i\phi} = \frac{\langle a|\psi\rangle}{|\langle a|\psi\rangle|} \tag{2.355}$$

which is a complex number of unit modulus $\left|e^{i\phi}\right| = 1$, since we have divided the complex number $\langle a|\psi\rangle = |\langle a|\psi\rangle|\, e^{i\phi}$ by its magnitude. It should be noted that all expectation values of the states $e^{i\phi}|a\rangle$ and $|a\rangle$ are identical, and hence they are not physically distinct states. Thus, in the nondegenerate case the system is just left in the eigenstate associated with that eigenvalue.

**Extension to Continuous Spectra** - If the spectrum of $A$ is continuous, a measurement of $A$ with *infinite* precision would leave the system in an associated eigenvector, such that

$$|\psi\rangle \;\cdots \underset{\alpha}{\overset{A}{\rightarrow}} \cdots\; \frac{\rho_\alpha|\psi\rangle}{|\langle\psi|\rho_\alpha|\psi\rangle|}. \tag{2.356}$$

which is similar to the previous result except the projector has been replaced by the projector density. In the case of a nondegenerate spectrum this would imply the reduction

$$|\psi\rangle \;\cdots \underset{\alpha}{\overset{A}{\rightarrow}} \cdots\; \frac{|\alpha\rangle\langle\alpha|\psi\rangle}{|\langle\alpha|\psi\rangle|} = e^{i\phi}|\alpha\rangle. \tag{2.357}$$

This, of course leads to an apparent problem: if $\alpha$ is in the continuous part of the spectrum, then the basis vector $|\alpha\rangle$ is *not* square normalizable. Thus, even if $|\psi\rangle$ is square normalized, the state of the system immediately after the measurement is not square-normalizable, and hence would be considered "unphysical". One might imagine that this could lead to a real conceptual problem, since we are often interested in making measurements of operators with continuous spectra, such as $\vec{R}$ and $\vec{P}$, which in the classical theory are the primary dynamical observables. In fact, the difficulty is really only mathematical, since it arises as a practical problem only to the extent that we can actually perform a measurement with *infinite precision.* In practice, such measurements never take place, since there is always a finite resolution to the measuring apparatus, and any measurement with finite precision always leaves the system in a physically acceptable state.

To see how the assumption of a measuring device with finite precision resolves the difficulty, consider a hypothetical device which would register a "click" (eigenvalue = 1) if measurement of an observable $A$ would yield a value in the interval $I_\alpha = (\alpha, \alpha + \Delta\alpha)$ and would remain silent (eigenvalue = 0) otherwise. Here $\Delta\alpha$ clearly represents the finite precision of the device. An infinite string or collection of such devices centered on intervals separated by an amount $\Delta\alpha$ would then be a complete $A$-meter, since it would allow a complete measurement of $A$ with precision $\Delta\alpha$. (When applied to the system, only one of the elemental devices would click, telling us in which interval $I_\alpha$ the particle's value was located.) But in this circumstance, the observable actually measured by the one device that clicks is not $A$, but is an operator $P_\alpha$ that can be expanded in the $\alpha$ representation in terms of its eigenvalues, 1 and 0, in the form

$$P_\alpha = \int\limits_{\alpha'\in I_\alpha} d\alpha' \; |\alpha'\rangle 1\langle\alpha'| + \int\limits_{\alpha'\notin I_\alpha} d\alpha' \; |\alpha'\rangle 0\langle\alpha'|. \tag{2.358}$$

Obviously we can drop the part proportional to zero and write

$$P_\alpha = \int\limits_{\alpha'\in I_\alpha} d\alpha' \; |\alpha'\rangle\langle\alpha'|. \tag{2.359}$$

But this is a projection operator, being an integral over a projector density. Thus, measuring a "click", tells us that the system has a value $\alpha$ in the associated interval, and projects the state of the system onto a *normalizable* state vector having its components in this interval, i.e.,

$$|\psi\rangle \quad \begin{array}{c} P_\alpha = 1 \\ \cdots \rightarrow \cdots \\ \alpha \in I_\alpha \end{array} \quad \frac{P_\alpha|\psi\rangle}{||P_\alpha|\psi\rangle||}. \tag{2.360}$$

Thus, any *real* measurement (i.e., one made with finite precision) leaves the system in a normalized state vector. On the other hand, if the precision $\Delta\alpha$, although finite, is *very* small compared to the variation of $\psi(\alpha)$ over this interval, we can still properly speak of the probability density $\rho(\alpha) = |\psi(\alpha)|^2$ of obtaining a particular eigenvalue, without having to explicitly mention the actual precision of the device. Thus, in what follows we will refer to measurements of position, momentum, kinetic energy, etc., without worrying too much about the actual measuring apparatus.

We now consider some consequences of the two parts of the measurement postulate.

### *2.3.1 Sum of Probabilities*

First, we note that the probability of obtaining *some* value is guaranteed through the mathematical structure of the theory, since

$$\sum_a P(a) = \sum_a \sum_{\tau=1}^{n_a} \langle\psi|a,\tau\rangle\langle a,\tau|\psi\rangle = \langle\psi| \left( \sum_a \sum_{\tau=1}^{n_a} |a,\tau\rangle\langle a,\tau| \right) |\psi\rangle = \langle\psi|\psi\rangle = 1, \tag{2.361}$$

where we have identified the decomposition of unity implicit in the fact that the operator $A$ is an observable. It is this completeness property of course, which motivated our mathematical definition of the term "observable" in the first place.

### *2.3.2 Mean Values*

Given that the predictions of the postulates are statistical in nature, there are a number of *statistics* of the measurement process which are useful to evaluate. Consider, e.g., an **ensemble** of $N$ identically prepared systems (with $N \gg 1$), all in the same quantum mechanical state $|\psi\rangle$. If the same observable $A$ is measured on each member of this ensemble, the result will be a collection of values $\{a\}$ (all being eigenvalues of $A$). Each value $a$ will occur with a frequency $f_a = NP(a)$, related to the associated probability appearing in the postulates. We can then compute the **mean value** $\langle A\rangle$ as the arithmetic average associated with this series of measurements. This mean value, it should be emphasized, may not actually coincide with any of the measurements actually performed, but it does gives some information about the underlying distribution. To compute the mean value we proceed as follows:

$$\langle A\rangle = \frac{1}{N}\sum_a f_a a = \sum_a aP(a) = \sum_a \sum_{\tau=1}^{n_a} a\,\langle\psi|a,\tau\rangle\langle a,\tau|\psi\rangle. \tag{2.362}$$

Note, however, that if we move the eigenvalue $a$ between the elements of the projectors, we can identify the resulting expression as an expansion for the operator $A$ in the basis consisting of its own eigenstates, i.e.,

$$\sum_a \sum_{\tau=1}^{n_a} \langle\psi|a,\tau\rangle a \langle a,\tau|\psi\rangle = \langle\psi| \left( \sum_a \sum_{\tau=1}^{n_a} |a,\tau\rangle \; a \; \langle a,\tau| \right) |\psi\rangle = \langle\psi|A|\psi\rangle. \tag{2.363}$$

Thus we obtain the very simple result that the *mean value*

$$\langle A\rangle = \langle\psi|A|\psi\rangle \tag{2.364}$$

*is equivalent to the expectation value* of the operator with respect to the given state. The useful thing about this result is that we can compute it in any basis that we like: we do not actually have to know the eigenstates and eigenvalues of an operator to compute its expectation value. For example, for any ONB of states $\{|n\rangle\}$ for which we know the expansion coefficients $\psi_n = \langle n|\psi\rangle$ for the state $|\psi\rangle$ and the matrix elements $A_{nn'} = \langle n|A|n'\rangle$ we can compute the expression

$$\langle A\rangle = \langle\psi|A|\psi\rangle = \sum_{n,n'} \psi_n^* A_{nn'} \psi_{n'}, \tag{2.365}$$

obtained by inserting decompositions if unity on each side of the operator $A$.

For reference, we list below some of the more common expectation values associated with a single particle. In the position representation, we have

$$\langle\vec{R}\rangle = \langle\psi|\vec{R}|\psi\rangle = \int d^3r\ \langle\psi|\vec{r}\rangle\vec{r}\langle\vec{r}|\psi\rangle = \int d^3r\ \psi^*(\vec{r})\,\vec{r}\,\psi(\vec{r}) = \int d^3r\ \vec{r}\rho(\vec{r}), \tag{2.366}$$

$$\langle V\rangle = \langle\psi|V|\psi\rangle = \int d^3r\ \langle\psi|\vec{r}\rangle V(\vec{r})\langle\vec{r}|\psi\rangle = \int d^3r\ \psi^*(\vec{r})V(\vec{r})\psi(\vec{r}) = \int d^3r\ V(\vec{r})\rho(\vec{r}), \tag{2.367}$$

$$\langle\vec{K}\rangle = \langle\psi|\vec{K}|\psi\rangle = \int d^3r\ \langle\psi|\vec{r}\rangle\langle\vec{r}|\vec{K}|\psi\rangle = -i\int d^3r\ \psi^*(\vec{r})\vec{\nabla}\psi(\vec{r}), \tag{2.368}$$

$$\langle\vec{P}\rangle = \langle\psi|\vec{P}|\psi\rangle = \int d^3r\ \langle\psi|\vec{r}\rangle\langle\vec{r}|\vec{P}|\psi\rangle = -i\hbar\int d^3r\ \psi^*(\vec{r})\vec{\nabla}\psi(\vec{r}), \tag{2.369}$$

$$\langle T\rangle = \langle\psi|\frac{P^2}{2m}|\psi\rangle = \frac{1}{2m}\int d^3r\ \langle\psi|\vec{r}\rangle\langle\vec{r}|P^2|\psi\rangle = -\frac{\hbar^2}{2m}\int d^3r\ \psi^*(\vec{r})\vec{\nabla}^2\psi(\vec{r}), \tag{2.370}$$

while in the wavevector representation

$$\langle\vec{R}\rangle = \int d^3k\ \langle\psi|\vec{k}\rangle\langle\vec{k}|\vec{R}|\psi\rangle = i\int d^3k\ \psi^*(\vec{k})\vec{\nabla}_{\vec{k}}\,\psi(\vec{k}), \tag{2.371}$$

$$\langle V\rangle = \int d^3k\int d^3k'\ \langle\psi|\vec{k}\rangle\langle\vec{k}|V|\vec{k}\,'\rangle\langle\vec{k}\,'|\psi\rangle = \int d^3k\int d^3k'\ \psi^*(\vec{k})\tilde{V}(\vec{k}-\vec{k}\,')\psi(\vec{k}') \tag{2.372}$$

$$\langle\vec{P}\rangle = \int d^3k\ \langle\psi|\vec{k}\rangle\hbar\vec{k}\langle\vec{k}|\psi\rangle = \int d^3k\ \psi^*(\vec{k})\hbar\vec{k}\psi(\vec{k}) = \int d^3k\ \hbar\vec{k}\,\rho(\vec{k}), \tag{2.373}$$

$$\langle T\rangle = \frac{\hbar^2}{2m}\int d^3k\ \langle\psi|\vec{k}\rangle k^2\langle\vec{k}|\psi\rangle = \frac{\hbar^2}{2m}\int d^3k\ \psi^*(\vec{k})k^2\psi(\vec{k}) = \frac{\hbar^2}{2m}\int d^3k\ k^2\rho(\vec{k}). \tag{2.374}$$

### *2.3.3 Statistical Uncertainty*

The mean value of an observable tells us roughly where we can expect the majority of the values obtained in an ensemble of measurements to be clustered. It tells us nothing about how big of a region around the mean value we might expect to obtain these values. It is useful to have a measure of this statistical spread, which reflects the intrinsic quantum mechanical uncertainty associated with the measurement process. One useful measure of this dispersion is the **root-mean-square uncertainty** of a series of measurements, defined through the relation

$$\Delta A = \sqrt{\langle(A-\langle A\rangle)^2\rangle}, \tag{2.375}$$

which we can write in an equivalent and sometimes more useful form by expanding the quadratic

$$\langle(A-\langle A\rangle)^2\rangle = \langle A^2 - 2A\langle A\rangle + \langle A\rangle^2\rangle. \tag{2.376}$$

Since $\langle A\rangle$ is a constant, this reduces to

$$\langle(A-\langle A\rangle)^2\rangle = \langle A^2\rangle - 2\langle A\rangle^2 + \langle A\rangle^2 = \langle A^2\rangle - \langle A\rangle^2, \tag{2.377}$$

so that

$$\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}. \tag{2.378}$$

Note that if the system is in a normalized eigenstate $|a\rangle$ of $A$ then

$$\langle A^n\rangle = \langle a|A^n|a\rangle = a^n\langle a|a\rangle = a^n, \tag{2.379}$$

in which case

$$\Delta A = \sqrt{a^2 - a^2} = 0, \tag{2.380}$$

so that there is *no* uncertainty when the system is in an eigenstate of the operator, as we have repeatedly asserted.

Thus, in a statistical sense, the uncertainty in an observable associated with a given quantum state $|\psi\rangle$ is a measure of the extent to which the state can be said to actually possess a value of the associated observable. It is interesting, in this context, to ask about the simultaneous possibility of reducing the uncertainty associated with two *different* observables. We know, for example, that if $B$ is an observable which commutes with $A$, then it is possible to find simultaneous eigenstates $|a,b\rangle$ of *both* observables. For such a state the uncertainty in both observables will vanish.

If, however, $B$ is an operator which does not commute with $A$, then there need be *no* simultaneous eigenstates (although a few *may* exist, there will not generally exist a common *basis* of eigenstates). Under these circumstances, it is not always possible to reduce the simultaneous statistical uncertainty associated with the measurement of both observables on a given quantum state. There turns out to be a precise statement which can be made about the so-called **uncertainty product**

$$\Delta A\Delta B = \sqrt{\langle(A-\langle A\rangle)^2\rangle}\sqrt{\langle(B-\langle B\rangle)^2\rangle} \tag{2.381}$$

associated with any given state of the system. This product is clearly a measure of the joint uncertainty associated with these two observables. In particular, we prove below the well-known **uncertainty principle**, the statement of which follows:

### *2.3.4 The Uncertainty Principle*

For any quantum state $|\psi\rangle$, the joint uncertainty in the values of two observables $A$ and $B$ as measured through the uncertainty product $\Delta A\Delta B$ is bounded from below through the relation

$$\Delta A\Delta B \geq \frac{1}{2}|\langle[A,B]\rangle|, \tag{2.382}$$

it being understood that all expectation values are to be taken with respect to the same quantum state $|\psi\rangle$.

To prove the uncertainty principle, we need first to prove a simple but useful theorem known as **Schwarz's inequality**, which states that if $|x\rangle$ and $|y\rangle$ are any two states in the space, then $\langle x|x\rangle\langle y|y\rangle \geq \langle x|y\rangle\langle y|x\rangle$ or, equivalently, $||x||^2||y||^2 \geq |\langle x|y\rangle|^2$, or more simply,

$$||x||\,||y|| \geq |\langle x|y\rangle|. \tag{2.383}$$

To prove this relation, set $|\alpha\rangle = |x\rangle + \lambda|y\rangle$, for some constant $\lambda$ to be chosen later. It then follows that the length of this vector is positive so that

$$\langle\alpha|\alpha\rangle = [\langle x| + \lambda^*\langle y|][|x\rangle + \lambda|y\rangle] \geq 0. \tag{2.384}$$

Expanding, we find that

$$\langle x|x\rangle + \lambda\langle y|x\rangle + \lambda^*\langle y|x\rangle + \lambda\lambda^*\langle y|y\rangle \geq 0. \tag{2.385}$$

This statement is true for arbitrary $\lambda$, so we can set

$$\lambda = -\frac{\langle x|y\rangle}{\langle y|y\rangle} \qquad \lambda^* = -\frac{\langle y|x\rangle}{\langle y|y\rangle}, \tag{2.386}$$

which leads to the result

$$\langle x|x\rangle - \frac{\langle x|y\rangle\langle y|x\rangle}{\langle y|y\rangle} - \frac{\langle y|x\rangle\langle x|y\rangle}{\langle y|y\rangle} + \frac{\langle x|y\rangle}{\langle y|y\rangle}\frac{\langle y|x\rangle}{\langle y|y\rangle}\langle y|y\rangle \geq 0. \tag{2.387}$$

This reduces to

$$\langle x|x\rangle - \frac{\langle x|x\rangle\langle y|x\rangle}{\langle y|y\rangle} \geq 0 \tag{2.388}$$

so

$$\langle x|x\rangle \geq \frac{\langle x|y\rangle\langle y|x\rangle}{\langle y|y\rangle}, \tag{2.389}$$

and hence

$$\langle x|x\rangle\langle y|y\rangle \geq \langle x|y\rangle\langle y|x\rangle, \tag{2.390}$$

which proves Schwarz's inequality.

We now are in a position to prove the uncertainty relation. First, we introduce shifted operators

$$\hat{A} = A - \langle A\rangle \tag{2.391}$$

$$\hat{B} = B - \langle A\rangle \tag{2.392}$$

which are just like the originals, except that they have zero mean value with respect to the state $|\psi\rangle$. These operators obey the following relationships, as is readily verified:

$$\Delta\hat{A} = \sqrt{\langle\hat{A}^2\rangle} = \Delta A \tag{2.393}$$

$$\Delta\hat{B} = \sqrt{\langle\hat{B}^2\rangle} = \Delta B, \tag{2.394}$$

and

$$[\hat{A}, \hat{B}] = [A, B], \tag{2.395}$$

so that if we prove the uncertainty relation for the shifted operators $\hat{A}$ and $\hat{B}$ we also prove it for the unshifted operators $A$ and $B$. Now set $|x\rangle = \hat{A}|\psi\rangle$, and $|y\rangle = \hat{B}|\psi\rangle$ and apply Schwarz's inequality to find that

$$\langle\psi|\hat{A}^2|\psi\rangle\langle\psi|\hat{B}^2|\psi\rangle \geq \langle\psi|\hat{A}\hat{B}|\psi\rangle\langle\psi|\hat{B}\hat{A}|\psi\rangle = |\langle\psi|\hat{A}\hat{B}|\psi\rangle|^2. \tag{2.396}$$

This is already a useful inequality, but to put it in the standard form, we can observe that the quantity on the right is the squared modulus of the complex number $\langle\psi|\hat{B}\hat{A}|\psi\rangle$ and so is larger in magnitude than the square of just its imaginary part. The latter we can obtain by taking one-half the difference of this number with its complex conjugate, i.e.,

$$|\langle\psi|\hat{A}\hat{B}|\psi\rangle| \geq |\mathbf{Im}\langle\psi|\hat{A}\hat{B}|\psi\rangle| = \frac{1}{2}|\langle\psi|\hat{A}\hat{B}|\psi\rangle - \langle\psi|\hat{B}\hat{A}|\psi\rangle| = \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle| \tag{2.397}$$

Combining this with the result above and taking the square root we obtain

$$\sqrt{\langle\hat{A}^2\rangle}\sqrt{\langle\hat{B}^2\rangle} \geq \frac{1}{2}|\langle[\hat{A},\hat{B}]\rangle|, \tag{2.398}$$

which is equivalent to the statement of the uncertainty principle.

Perhaps the most common application of the uncertainty principle is to the Cartesian components of the position and momentum operator along the same direction, for which the canonical commutation relations $[X_i, P_j] = i\hbar\delta_{ij}$ and the result above imply that

$$\Delta X_i \Delta P_i \geq \frac{\hbar}{2} \tag{2.399}$$

or equivalently

$$\Delta X_i \Delta K_i \geq \frac{1}{2}. \tag{2.400}$$

In this form, the uncertainty relation shows that, past a certain point, we can increase our knowledge of a particle's position along a certain direction only if we are willing to put up with a concomitant loss of information about its momentum along the same direction, and vice versa. More generally, we can increase our knowledge of an operator $A$ at the expense of decreasing our knowledge of observables $B$ with which $A$ does not commute. On the other hand, the uncertainty principle is also consistent with our observation that there is no limit to the precision with which we may simultaneously specify the value of *commuting* observables. Commuting observables are, therefore, often referred to as being **compatible observables.**

### *2.3.5 Preparation of a State Using a CSCO*

It follows from the last discussion and the statement of the second part of the measurement postulate, that if a system is initially in an unknown state, it should be possible to "collapse" it into a known state through a series of nearly instantaneous measurements performed using the operators in a complete set of commuting observables $\{A, B, C\}$. Suppose, for example, the system is in the state $|\psi\rangle$, which we may or may not know. This state can be expanded in the basis vectors associated with the simultaneous eigenstates of the operators $A, B$, and $C$ in the usual way, i.e.,

$$|\psi\rangle = \sum_{a',b',c'} \psi_{a',b',c'}|a',b',c'\rangle. \tag{2.401}$$

If we now measure these three compatible observables in a very short interval of time, we will see the subsequent reduction of the state vector onto one of these basis vectors, as represented by the diagram

$$|\psi\rangle \quad \cdots \xrightarrow[a]{A} \cdots \quad \sum_{b',c'}\psi_{a,b',c'}|a,b',c'\rangle \quad \cdots \xrightarrow[b]{B} \cdots \quad \sum_{c'}\psi_{a,b,c'}|a,b,c'\rangle \quad \cdots \xrightarrow[c]{C} \cdots \quad \frac{\psi_{a,b,c}}{|\psi_{a,b,c}|}|a,b,c\rangle. \tag{2.402}$$

This conceptual ability to prepare a system in a well characterized state is extremely important for the theory, for it allows us to test the statistical predictions of the theory, which rely on the idea of an ensemble of similarly prepared systems upon which to perform a subsequent measurement. Thus, after performing such a complete series of measurements on an ensemble of arbitrary initial state vectors, we can extract those which end up in a particular quantum state $|a, b, c\rangle$ to produce a subensemble of systems upon which to perform further experiments.

We are now ready to finish up the set of postulates that we have been developing to describe the formalism of quantum mechanics. The last postulate describes the way a quantum mechanical system behaves in between the times during which measurements are being made. As we have seen, *during* a measurement process, a quantum mechanical system, in contact with a classical measuring device, evolves non-deterministically as the state vector collapses into one of the eigensubspaces of the particular observable being measured. In between these measurement events, evolution is governed by the fourth postulate.

## 2.4 Postulate IV : Evolution

> Between measurements the state vector $|\psi(t)\rangle$ of a quantum system evolves deterministically according to **Schrödinger's equation of motion**
>
> $$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle, \tag{2.403}$$
>
> in which the Hamiltonian operator $H$ is the observable associated with the total energy of the system at time $t$.

In practice, to use Schrödinger's equation we project it onto the basis vectors of an appropriate representation. Thus, if the vectors $\{|n\rangle\}$ form an ONB for the space of interest, then we can write

$$\langle n|i\hbar\frac{d}{dt}|\psi(t)\rangle = \langle n|H|\psi(t)\rangle. \tag{2.404}$$

Sliding the bra $\langle n|$ past the time derivative and inserting a complete set of states to the right of the Hamiltonian, we obtain

$$i\hbar\frac{d}{dt}\langle n|\psi(t)\rangle = \sum_{n'}\langle n|H|n'\rangle\langle n'|\psi(t)\rangle \tag{2.405}$$

in which we recognize coefficients for the expansion

$$|\psi(t)\rangle = \sum_n |n\rangle\langle n|\psi(t)\rangle = \sum_n \psi_n(t)\,|n\rangle. \tag{2.406}$$

Thus, in this representation, Schrödinger's equation takes the form

$$i\hbar\frac{d\psi_n}{dt} = \sum_m H_{nn'}\psi_{n'}(t) \tag{2.407}$$

of a set of first-order coupled differential equations for the time-dependent expansion coefficients for the state $|\psi(t)\rangle$ in this basis.

In a continuous representation $|\alpha\rangle$, the state of the system is represented by the wavefunction $\psi(\alpha)$ and the Hamiltonian becomes an integro-differential operator acting on this function. In the most general case, the matrix elements of $H$ between the continuous basis states $|\alpha\rangle$ are defined by some kernel $H(\alpha,\alpha') = \langle\alpha|H|\alpha'\rangle$. Projection of the Schrödinger equation onto the basis states of that representation then leads to the expression

$$\langle\alpha|i\hbar\frac{d}{dt}|\psi(t)\rangle = \langle\alpha|H|\psi(t)\rangle. \tag{2.408}$$

As before, taking the derivative with respect to time is a linear operation, so we can write

$$\langle\alpha|i\hbar\frac{d}{dt}|\psi(t)\rangle = i\hbar\frac{d}{dt}\langle\alpha|\psi(t)\rangle = i\hbar\frac{\partial\psi(\alpha,t)}{\partial t} \tag{2.409}$$

where the exact differential for the vector $|\psi(t)\rangle$ (which only depends parametrically on time, not on $\alpha$ or anything else) turns into a partial derivative when it acts on the function $\psi(\alpha,t) = \langle\alpha|\psi(t)\rangle$, which is formally a function of two variables. Making this substitution and inserting a complete set of states between $H$ and $\psi(t)$ on the right we obtain an integral equation

$$i\hbar\frac{\partial\psi}{\partial t} = \int d\alpha'\ H(\alpha,\alpha')\psi(\alpha',t). \tag{2.410}$$

for the wave function $\psi(\alpha,t)$.

Under certain special situations (which occur rather often) the matrix elements of $H$ will involve derivatives of delta functions, and the integral equation will reduce to a differential equation, as we have seen occur with the energy eigenvalue equation in the position representation.

Thus, for a single particle in $3D$ moving under the influence of a potential $V(\vec{r},t)$, the Hamiltonian is simply the sum of the kinetic and potential energy operators

$$H = \frac{P^2}{2m} + V(\vec{R},t). \tag{2.411}$$

Under these circumstances, the Schrödinger equation can be written in the position representation by projecting it onto the basis vectors of that representation, i.e.,

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle = \frac{P^2}{2m}|\psi(t)\rangle + V|\psi(t)\rangle, \tag{2.412}$$

$$i\hbar\frac{d}{dt}\langle\vec{r}|\psi(t)\rangle = \frac{1}{2m}\langle\vec{r}|P^2|\psi(t)\rangle + \langle\vec{r}|V|\psi(t)\rangle, \tag{2.413}$$

which we recognize as

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{r},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r},t) + V(\vec{r},t)\psi(t), \tag{2.414}$$

which is Schrödinger's equation in its original form. Alternatively, we can choose to work in the momentum or wavevector representation:

$$i\hbar\frac{d}{dt}\langle\vec{k}|\psi(t)\rangle = \frac{1}{2m}\langle\vec{k}|P^2|\psi(t)\rangle + \langle\vec{k}|V|\psi(t)\rangle, \tag{2.415}$$

which we can write in the form

$$i\hbar\frac{\partial}{\partial t}\psi(\vec{k},t) = \frac{\hbar^2k^2}{2m}\psi(\vec{k},t) + \int d^3k'\ \tilde{V}(\vec{k}-\vec{k}\,',t)\psi(\vec{k}\,',t). \tag{2.416}$$

Unless $\tilde{V}(\vec{k}-\vec{k}\,',t)$ has special properties which enable a simplification, Schrödinger's equation for a single particle in the wavevector representation is an integrodifferential equation.

### *2.4.1 Construction of the Hamiltonian and Other Observables*

In principle, the evolution of a quantum mechanical system is reduced to the solution of a set of coupled first order differential equations once the Hamiltonian is known. As in classical mechanics, therefore, the first step in solving the dynamical problem is the construction of a suitable Hamiltonian. In many cases a Hamiltonian *operator* can be obtained from the Hamiltonian *function* of an associated classical system. For the paradigmatic case of a single spinless particle moving in three dimensions, for example, we move from the classical description, which is based upon the dynamical variables $\vec{r}$ and $\vec{p}$, to the quantum mechanical one, by replacing the dynamical variables by the Hermitian operators $\vec{R}$ and $\vec{P}$, whose components obey the canonical commutation relations

$$[X_i, X_j] = [P_i, P_j] = 0 \tag{2.417}$$

$$[X_i, P_j] = i\hbar\delta_{ij}. \tag{2.418}$$

In a similar fashion, it seems reasonable to associate with any classical observable $A(\vec{r}, \vec{p}, t)$ an operator $A(\vec{R}, \vec{P}, t)$ obtained by replacing the dynamical variables appearing in the function with the corresponding operators. Unfortunately, while this procedure works a good deal of the time, there are circumstances where it can give ambiguous results and/or suffer the drawback that the operator which is produced is not Hermitian. To illustrate the basic difficulties that arise, consider the following classical observable

$$xp_x = p_x x \tag{2.419}$$

which we can write in either of these two ways, since classical variables always commute with one another. Quantum mechanically, however, the operators obtained by replacing the dynamical variables with associated operators

$$XP_x \neq P_x X \tag{2.420}$$

are *not* equal because these operators do not commute. A moments reflection will reveal that there are actually an infinite number of classically equivalent expressions that each generate a different quantum mechanical operator (consider, e.g., $\sqrt[n]{x^n p^n} = xp_x$) Which of these operators should be used to represent the classical observable? Most of them are patently unusable because they are not Hermitian. In the simple example above, e.g., neither $XP_x$ or $P_x X$ is a Hermitian operator,

$$[XP_x]^+ = P_x^+ X^+ = P_x X. \tag{2.421}$$

and so cannot represent an observable.

To resolve this problem we use the idea of **Hermitian symmetrization** of an operator. Recall that any operator $A$ can be written in the form

$$A = A_H + A_A, \tag{2.422}$$

where the operator $A_H = \frac{1}{2}(A + A^+)$ is Hermitian and $A_A = \frac{1}{2}(A + A^+)$ is anti-Hermitian. Thus, an arbitrary operator can be decomposed in a unique way into Hermitian and anti-Hermitian parts in a way analogous to the manner in which an arbitrary complex number can be decomposed into real and imaginary parts. We refer to $A_H$, therefore, as the Hermitian part of $A$, and stipulate that the observable corresponding to any classical quantity be associated with the Hermitian part of the operator obtained through the replacement discussed above, i.e.,

$$A(\vec{r},\vec{p},t) \to \frac{A(\vec{R},\vec{P},t)+A^{+}(\vec{R},\vec{P},t)}{2}. \tag{2.423}$$

In the particular example considered, this implies that

$$xp_x = p_x x \to \frac{1}{2}(XP_x + P_x X) \tag{2.424}$$

so that both of the operators in question are treated on an equal footing. An additional example would be

$$\vec{r}\cdot\vec{p} \to \frac{1}{2}(\vec{R}\cdot\vec{P}+\vec{P}\cdot\vec{R}). \tag{2.425}$$

Unfortunately, this procedure only *reduces* the underlying problem, it does not eliminate it. This can be seen by considering the following equivalent classical expressions

$$x^2 p_x = p_x x^2 = x p_x x, \tag{2.426}$$

the first two of which are both associated under the above prescription with the operator

$$\frac{1}{2}(X^2 P_x + P_x X^2), \tag{2.427}$$

while the third of which is associated with the operator

$$XP_xX. \tag{2.428}$$

This difficulty is, in a sense, only mathematical, and presumably arises because we are attempting to go from a *less complete* description (i.e., classical mechanics) to a *more complete* description (i.e., quantum mechanics) of the physical universe. It is not unreasonable to expect that in any domain where classical behavior is observed, the differences between the predictions associated with any of these operators will become unimportant. In the quantum domain, however, this underscores the fact that the association of a measuring device with an observable can sometimes involve subtle distinctions.

It should also be pointed out that this problem really only arises in operators involving *products* of non-commuting observables. In the most common situation, namely that of a particles moving in response to a classical potential function $V(r,t)$, the problem never arises because such products don't appear in the Hamiltonian.

### *2.4.2 Some Features of Quantum Mechanical Evolution*

*Determinism* - Note that the differential equation governing the evolution of the state vector is first order in time. This means that the solution depends only on the initial state of the system, and not, e.g., on its initial rate of change. Thus, any initial state $|\psi(t_0)\rangle$ of the system at time $t_0$ will evolve into a single unique vector $|\psi(t)\rangle$ at time $t > t_0$. We note that this implicitly defines a mapping of the space onto itself, and thus implies the existence of an operator $U$, or a family of operators $U(t,t_0)$, that map an arbitrary state at time $t_0$ onto the state into which it evolve in time $t$. This **evolution operator** is defined through the relation

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle. \tag{2.429}$$

*Linearity* - The linearity of the equations of motion imply a superposition principle for the solutions of the Schrödinger equation. That is if $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ are two possible solutions to the Schrödinger equation (which have, e.g., evolved from two different initial state vectors $|\psi_1(t_0)\rangle$ and $|\psi_2(t_0)\rangle$), then the time-dependent vector

$$|\psi(t)\rangle = \alpha|\psi_1(t)\rangle + \beta|\psi_2(t)\rangle \tag{2.430}$$

is also a solution to the Schrödinger equation for any complex constants $\alpha$ and $\beta$, since

$$\begin{aligned} i\hbar\frac{d}{dt}|\psi(t)\rangle &= \alpha(i\hbar\frac{d}{dt}|\psi_1(t)\rangle) + \beta(i\hbar\frac{d}{dt}|\psi_2(t)\rangle) &(2.431)\\ &= \alpha H|\psi_1(t)\rangle + \beta H|\psi_1(t)\rangle = H(\alpha|\psi_1(t)\rangle + \beta|\psi_2(t)\rangle), &(2.432)\end{aligned}$$

so that

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle. \tag{2.433}$$

This implies, as a consequence, that if we find out how the basis vectors of any ONB evolve under the Schrödinger equation, we can determine the evolution of any other vector in the system. It also implies that the evolution operator $U(t, t_0)$ introduced above is a linear operator.

*Conservation of the Norm* - It is also relatively easy to show that quantum mechanical evolution preserves the norm of the state vector, a condition which is obviously important if we wish the total sum of probabilities to be conserved. Thus, we consider the rate of change of the (squared) length of a vector $|\psi(t)\rangle$ evolving under the Schrödinger equation

$$\frac{d}{dt}\langle\psi(t)|\psi(t)\rangle = \langle\psi|\left(\frac{d}{dt}|\psi\rangle\right) + \left(\frac{d}{dt}\langle\psi|\right)|\psi\rangle, \tag{2.434}$$

where we have simply used the chain rule on the right hand side. From the Schrödinger equation itself we deduce that

$$\frac{d}{dt}|\psi\rangle = -\frac{i}{\hbar}H|\psi\rangle \tag{2.435}$$

the adjoint of which gives

$$\frac{d}{dt}\langle\psi| = \frac{i}{\hbar}\,\langle\psi|H. \tag{2.436}$$

Substituting these in above we find that

$$\frac{d}{dt}\langle\psi|\psi\rangle = -\frac{i}{\hbar}\langle\psi|H|\psi\rangle + \frac{i}{\hbar}\langle\psi|H|\psi\rangle = 0, \tag{2.437}$$

so that $\langle\psi(t)|\psi(t)\rangle = \langle\psi(t_0)|\psi(t_0)\rangle$ is constant. This implies that the evolution operator $U = U(t, t_0)$ that we defined earlier is *unitary*, since

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(t_0)|U^+U|\psi(t_0)\rangle = \langle\psi(t_0)|\psi(t_0)\rangle. \tag{2.438}$$

Since this must be true for arbitrary states $|\psi(t_0)\rangle$, it follows that $U^+U = \mathbf{1}$. This feature is also described by saying that the Schrödinger equation leads to a **unitary evolution**.

### *2.4.3 Evolution of Mean Values*

Let us now consider how the mean value associated with an arbitrary observable $A(t)$ evolves in time. In general, operators can have an intrinsic time dependence. As an example consider the potential associated with the application of a spatially uniform sinusoidally-varying electric field, i.e.,

$$V_{ext}(t) = -e\vec{E} \cdot \vec{R}\cos(\omega t). \tag{2.439}$$

The mean value of this operator

$$\langle V_{ext}(t)\rangle = \langle\psi(t)|V_{ext}(t)|\psi(t)\rangle \tag{2.440}$$

evolves in time, since both the state and the operator itself is changing. At any instant of time, this mean value gives a measure of the interaction of the system with the external field. In general, the mean value of an arbitrary observable

$$\langle A(t)\rangle = \langle\psi(t)|A(t)|\psi(t)\rangle \tag{2.441}$$

may have two sources of time dependence: a part due to the operator itself, and a part due to the evolution of the system. We can, however, use the chain rule to write

$$\frac{d}{dt}\langle A(t)\rangle = \frac{d}{dt}\langle\psi(t)|A(t)|\psi(t)\rangle \tag{2.442}$$

$$= \left(\frac{d}{dt}\langle\psi|\right) A|\psi\rangle + \langle\psi|\frac{\partial A}{\partial t}|\psi\rangle + \langle\psi|A\left(\frac{d}{dt}|\psi\rangle\right). \tag{2.443}$$

From our earlier manipulations this can be written

$$\frac{d}{dt}\langle A(t)\rangle = \frac{i}{\hbar}\langle\psi|HA|\psi\rangle + \langle\frac{\partial A}{\partial t}\rangle - \frac{i}{\hbar}\langle\psi|AH|\psi\rangle. \tag{2.444}$$

$$= \langle\frac{\partial A}{\partial t}\rangle - \frac{i}{\hbar}\langle\psi|AH - HA|\psi\rangle \tag{2.445}$$

in which we recognize the commutator of $A$ and $H$. Thus, we have the equation of motion

$$\frac{d}{dt}\langle A(t)\rangle = \langle\frac{\partial A}{\partial t}\rangle - \frac{i}{\hbar}\langle[A, H]\rangle. \tag{2.446}$$

This form may be familiar to the student of classical mechanics, in that it resembles the equation of motion

$$\frac{dA}{dt} = \frac{\partial A}{\partial t} - \{A, H\}_{\mathrm{PB}}. \tag{2.447}$$

for a classical observable $A(q, p, t)$, where the bracketed quantity represents the *Poisson bracket* of the two functions $A(q, p, t)$ and $H(q, p, t)$, defined through the relation

$$\{f, g\} = \sum_i \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i}\frac{\partial f}{\partial p_i}. \tag{2.448}$$

As a consequence of this equation of motion, we note that any time-independent operator that commutes with the Hamiltonian has a mean value that remains constant in time, since the equation of motion then predicts that $d\langle A(t)\rangle/dt = 0$.

### *2.4.4 Eherenfest's Theorem*

As an interesting application of the use of the equations of motion for the mean value of an observable, consider the motion of a particle under the influence of a force $\vec{F}(\vec{r}) = -\nabla V(\vec{r})$ derivable from a scalar potential $V(\vec{r})$. Quantum mechanically, this corresponds to the usual Hamiltonian

$$H = \frac{P^2}{2m} + V(\vec{R}). \tag{2.449}$$

The classical dynamical variables $\vec{r}(t)$ and $\vec{p}(t)$ associated with such a Hamiltonian evolve according to Hamilton's equations

$$\frac{d\vec{r}(t)}{dt} = \frac{p(t)}{m}, \tag{2.450}$$

$$\frac{d\vec{p}(t)}{dt} = -\nabla V(r(t)) = \vec{F}(\vec{r}(t)). \tag{2.451}$$

which are equivalent to Newton's second law. Let us now consider how the *mean values* $\langle\vec{R}(t)\rangle$ and $\langle\vec{P}(t)\rangle$ associated with the corresponding quantum mechanical observables change in time. First, we examine the equation of motion for the position operator $\vec{R}$, which being independent of time ($\partial\vec{R}/\partial t = 0$) leads to the equation of motion

$$\frac{d}{dt}\langle\vec{R}(t)\rangle = -\frac{i}{\hbar}\langle[\vec{R}, H]\rangle. \tag{2.452}$$

This leads us to evaluate

$$[\vec{R}, H] = \frac{1}{2m}[\vec{R}, P^2] + [\vec{R}, V]. \tag{2.453}$$

Since $V$ is a function of $\vec{R}$, the second commutator vanishes. The $x$ component of the first commutator is

$$[X, P^2] = [X, P_x^2] + [X, P_y^2] + [X, P_z^2] = [X, P_x^2], \tag{2.454}$$

where we have recognized that the only non-commuting part involves position and momentum operators along the same direction. Using the standard trick for evaluating the commutator of a product we find that

$$[X, P_x^2] = P_x[X, P_x] + [X, P_x]P_x = 2i\hbar P_x, \tag{2.455}$$

and similarly for the other two cartesian components of the commutator in question. As a vector operator relation, therefore, we have the result that

$$[\vec{R}, P^2] = 2i\hbar\vec{P}, \tag{2.456}$$

which we can put back into the equation of motion for $\langle\vec{R}(t)\rangle$ to obtain

$$\frac{d}{dt}\langle\vec{R}(t)\rangle = -\frac{i}{\hbar}\frac{1}{2m}\langle 2i\hbar\vec{P}\rangle = \frac{1}{m}\langle\vec{P}\rangle. \tag{2.457}$$

Thus, we find that

$$\frac{d\langle\vec{R}\rangle}{dt} = \left\langle\frac{\vec{P}}{m}\right\rangle, \tag{2.458}$$

which, brackets aside, looks like its classical counterpart. Thus, as in classical mechanics the mean velocity equals the mean momentum divided by the mass.

In a similar fashion we can compute the equation of motion for the mean momentum,

$$\frac{d}{dt}\langle \vec{P}(t)\rangle = -\frac{i}{\hbar}\langle[\vec{P}, H]\rangle. \tag{2.459}$$

which leads us to evaluate

$$[\vec{P}, H] = \frac{1}{2m}[\vec{P}, P^2] + [\vec{P}, V]. \tag{2.460}$$

Now the kinetic energy term disappears, but the potential energy term does not, since it is a function of the operator $\vec{R}$, which does not commute with $\vec{P}$. Since we have not specified the exact functional form of the potential energy function, it is convenient to work in a representation in which $V$ is diagonal, namely the position representation. In the position representation we can write

$$\langle\vec{r}|[\vec{P}, V]|\psi\rangle = -i\hbar[\vec{\nabla}V(\vec{r})\psi(r) - V(\vec{r})\vec{\nabla}\psi(\vec{r})] \tag{2.461}$$

$$= -i\hbar[\psi(r)\vec{\nabla}V(\vec{r}) + V(\vec{r})\vec{\nabla}\psi(r) - V(\vec{r})\vec{\nabla}\psi(\vec{r})] = -i\hbar[\vec{\nabla}V(\vec{r})]\psi(r), \tag{2.462}$$

so that in the position representation, $[\vec{P}, V]$ acts to multiply the wave function by the function $-i\hbar\vec{\nabla}V(\vec{r}) = i\hbar\vec{F}(\vec{r})$, where $\vec{F}(\vec{r})$ is the classical force function. Thus, we make the identification

$$[\vec{P}, V] = i\hbar\vec{F}(\vec{R}), \tag{2.463}$$

where

$$\vec{F}(\vec{R}) = \int d^3r\ |\vec{r}\rangle\vec{F}(\vec{r})\langle\vec{r}|$$

is a vector observable associated with the force on the particle, i.e., the *force operator*. Using this result in the equation of motion we find that

$$\frac{d}{dt}\langle\vec{P}(t)\rangle = -\frac{i}{\hbar}\langle i\hbar\vec{F}(\vec{R})\rangle = \langle\vec{F}\rangle. \tag{2.464}$$

Thus, the equations of motion for the position and momentum operators can be written

$$\frac{d\langle\vec{R}\rangle}{dt} = \langle\vec{P}\rangle, \tag{2.465}$$

$$\frac{d}{dt}\langle\vec{P}\rangle = \langle\vec{F}\rangle. \tag{2.466}$$

which looks like Newton's equations, aside from the taking of expectation values. These classically familiar-looking expressions are referred to as **Ehrenfest's equations of motion** for the mean values. Their interpretation requires a little care. It might be expected, for example, that these equations imply that if the initial mean values were equal to those of some hypothetical classical system with the same potential, so that $\langle\vec{R}(0)\rangle = \vec{r}(0)$ and $\langle\vec{P}(0)\rangle = \vec{p}(0)$, then as both systems evolved the mean values $\langle\vec{R}(t)\rangle$ and $\langle\vec{P}(t)\rangle$ for the quantum particle would simply follow the corresponding classical trajectory $\vec{r}(t)$ and $\vec{p}(t)$. This is, however, *not* generally the case. To see this, we note that the classical equation of motion

$$\frac{d\vec{p}}{dt} = \vec{F} = \vec{F}(\vec{r}(t)), \tag{2.467}$$

equates the derivative of $\vec{p}$ to the force function *evaluated at the particle's instantaneous position.* If the mean value $\langle \vec{P} \rangle$ *were* to obey this same equation, it would have to satisfy the following relation

$$\frac{d\langle \vec{P} \rangle}{dt} = \vec{F}(\langle \vec{R} \rangle), \tag{2.468}$$

which involves the force function $\vec{F}(\vec{r})$ evaluated at the mean value of the particle's position. But this is not the equation we derived, which contains $\langle \vec{F}(\vec{R}) \rangle$ on the right hand side, not $\vec{F}(\langle \vec{R} \rangle)$. Thus, the only situation in which the quantum mean values will follow classical trajectories is when, for all instants of time,

$$\vec{F}(\langle \vec{R} \rangle) = \langle \vec{F}(\vec{R}) \rangle. \tag{2.469}$$

In general, of course these are not the same. It is straightforward to show, however, that if the potential function can be written as a polynomial of degree two or less in the position of the particle, then this condition is satisfied. This means that the mean value of position and momentum for a particle subject to no force, a constant force, or a linear (e.g., Hooke's law) force will always follow the corresponding classical trajectory.

### *2.4.5 Evolution of Systems with Time Independent Hamiltonians*

We now consider the evolution of quantum mechanical systems in which the Hamiltonian operator is independent of time, so that $\partial H/\partial t = 0$. Classically, in such a system the total energy is conserved. Quantum mechanically, this implies that the *mean value* of the energy will be conserved since, under these circumstances

$$\frac{d}{dt}\langle H(t) \rangle = \langle \frac{\partial H}{\partial t} \rangle - \frac{i}{\hbar}\langle [H, H] \rangle = 0. \tag{2.470}$$

It is important to realize, of course, that the energy of a quantum system is still generally undefined *unless* the system is actually in an eigenstate of the energy operator. Thus, the mean value only predicts the statistical outcome associated with many measurements of energy performed on an ensemble of identically-prepared quantum mechanical systems. Nonetheless, when the Hamiltonian is time independent, the evolution of the system is most easily expressed in terms of the ONB of energy eigenstates. In the case of a discrete system, we can express the energy eigenstates in terms of a discrete index $n$ so that

$$H|n\rangle = E_n|n\rangle, \tag{2.471}$$

with

$$\langle n|n' \rangle = \delta_{n,n'}. \tag{2.472}$$

Under such circumstances, the instantaneous dynamical state can be expanded in the form

$$|\psi(t)\rangle = \sum_n |n\rangle\langle n|\psi(t)\rangle = \sum_n \psi_n(t)|n\rangle. \tag{2.473}$$

For a system with a continuous energy spectrum, the energy eigenstates can be indexed by a continuous index, $\nu$, such that

$$H|v\rangle = E_v|v\rangle, \tag{2.474}$$

with

$$\langle v|v' \rangle = \delta(v - v'), \tag{2.475}$$

and

$$|\psi(t)\rangle = \int dv\ |v\rangle\langle v|\psi(t)\rangle = \int dv\ \psi(v,t)|v\rangle. \tag{2.476}$$

In general, the Hamiltonian can have both a discrete and a continuous part to its spectrum, with both discrete and continuously distributed eigenvectors $\{|n\rangle, |v\rangle\}$, with

$$\langle n|n'\rangle = \delta_{n,n'}. \tag{2.477}$$

$$\langle v|v'\rangle = \delta(v-v'), \tag{2.478}$$

$$\langle n|v\rangle = 0, \tag{2.479}$$

and

$$|\psi(t)\rangle = \sum_n |n\rangle\langle n|\psi(t)\rangle + \int dv\ |v\rangle\langle v|\psi(t)\rangle = \sum_n \psi_n(t)|n\rangle + \int dv\ \psi(v,t)|v\rangle. \tag{2.480}$$

In what follows, we will, for simplicity, write expressions in the form of a discrete index, but corresponding expressions for the general case should be straightforward to generate.

If we project the Schrödinger equation onto the basis vectors of the energy operator, we obtain

$$i\hbar\frac{d\psi_n(t)}{dt} = \sum_{n'} H_{nn'}\psi_{n'}(t), \tag{2.481}$$

where, by assumption, $H_{nn'} = \langle n|H|n'\rangle = E_n\langle n|n'\rangle = E_n\delta_{nn'}$. Making this substitution we find that the equations of motion in the energy representation

$$i\hbar\frac{d}{dt}\psi_n = E_n\psi_n = \hbar\omega_n\psi_n, \tag{2.482}$$

are uncoupled. In this last expression we have introduced the notation $\omega_n = E_n/\hbar$. This equation is readily manipulated into the integral

$$\int_{\psi_n(t_0)}^{\psi_n(t)} \frac{d\psi_n}{\psi_n} = -i\omega_n\int_{t_0}^{t} dt, \tag{2.483}$$

which gives

$$\psi_n(t) = \psi_n(t_0)e^{-i\omega_n(t-t_0)}, \tag{2.484}$$

or if we wish to make $t_0 = 0$,

$$\psi_n(t) = \psi_n(0)e^{-i\omega_n t}, \tag{2.485}$$

so that

$$|\psi(t)\rangle = \sum_n \psi_n(0)\ e^{-i\omega_n t}|n\rangle = \sum_n \psi_n(t_0)\ e^{-i\omega_n(t-t_0)}|n\rangle$$

Thus, in the energy representation the coefficients which determine the state vector at an initial instant of time each acquire a simple time-dependent phase factor that depends upon the energy of the associated basis state.

If at $t = 0$ the system is in a single energy eigenstate, so that

$$|\psi(0)\rangle = |n\rangle, \tag{2.486}$$

then $\psi_{n'}(0) = \delta_{n,n'}$. The system will then simply stay in that eigenstate, but will acquire an oscillating phase factor, i.e.,

$$|\psi(t)\rangle = e^{-i\omega_n t}|n\rangle. \tag{2.487}$$

In such a simply evolving state all physical properties are static, or stationary. Consider, e.g., any time-independent observable $B$, which may or may not commute with the Hamiltonian. With the system initially in an energy eigenstate, the mean value of such an observable will be

$$\langle B(t)\rangle = \langle\psi(t)|B|\psi(t)\rangle = e^{+i\omega_n t}\langle n|B|n\rangle e^{-i\omega_n t} = \langle n|B|n\rangle = \langle B(0)\rangle, \tag{2.488}$$

which is independent of time. For this reason, the energy eigenstates are referred to as **stationary states**.

In general, of course, the system will be in a linear superposition of energy eigenstates of different energies, and observables of the system will therefore evolve in time. Thus, an initial state

$$|\psi(0)\rangle = \sum_n \psi_n(0)|n\rangle, \tag{2.489}$$

will evolve into the state

$$|\psi(t)\rangle = \sum_n \psi_n(0)e^{-i\omega_n t}|n\rangle. \tag{2.490}$$

Once we know the expansion in the energy basis we can view the evolution in other representations as well. Thus, e.g., the real space wave function $\psi(\vec{r},t)$ for such a state will evolve in a manner that depends upon the projection of the state vector onto one of the basis states of the position representation, i.e.,

$$\psi(\vec{r},t) = \langle\vec{r}|\psi(t)\rangle = \sum_n \psi_n(0)\; e^{-i\omega_n t}\langle\vec{r}|n\rangle = \sum_n \psi_n(0)\; e^{-i\omega_n t}\phi_n(\vec{r}). \tag{2.491}$$

where $\phi_n(\vec{r}) = \langle\vec{r}|n\rangle$ are the energy eigenfunction in the position representation, and

$$\psi_n(0) = \langle n|\psi(0)\rangle = \int d^3r\; \langle n|\vec{r}\rangle\langle\vec{r}|\psi(0)\rangle = \int d^3r\; \phi_n^*(\vec{r})\psi(\vec{r},0)$$

can be computed from the initial real space wavefunction. Thus, we obtain a natural decomposition of $\psi(\vec{r},t)$ in the orthonormal functions $\phi_n(\vec{r})$ associated with the energy eigenstates.

In such a superposition state, the mean value of an observable will also evolve, as we have seen. This evolution, whose equation of motion we have already explored can also be expressed in the energy representation

$$\langle A(t)\rangle = \langle\psi(t)|A|\psi(t)\rangle = \sum_n\sum_{n'}\psi_n^*(t)A_{nn'}\psi_{n'}(t) \tag{2.492}$$

$$= \sum_n\sum_{n'}\psi_n^*(0)A_{nn'}\psi_{n'}(0)\; e^{+i(\omega_n-\omega_{n'})t}, \tag{2.493}$$

in which we have inserted the expansion for $|\psi(t)\rangle$ and its adjoint

$$\langle\psi(t)| = \sum_n \psi_n^*(0)\; e^{+i\omega_n t}\langle n|. \tag{2.494}$$

From this expression it is easy to see that the mean value of all time-independent observables will have components which oscillate in time at the so-called **Bohr frequencies** of the system,

$$\Omega_{nn'} = \omega_n - \omega_{n'} = \frac{E_n - E_{n'}}{\hbar}, \tag{2.495}$$

which are simply related to the energy differences between the different eigenstates of $H$.

### *2.4.6 The Evolution Operator*

We finish up our formal discussion of the consequences of the evolution postulate by discussing the explicit form of the unitary operator $U(t, t_0)$ which evolves the state vector from an initial state at time $t_0$ to the state of the system at some time $t$ later. We treat separately the case of a time-dependent and a time-independent Hamiltonian.

*Time-Independent Hamiltonian* -For the case in which $H$ is independent of time it is possible to explicitly construct the evolution operator by considering the expansion of the evolving state vector in the energy basis. The evolution operator is defined through the relation

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle, \tag{2.496}$$

while the equations of motion imply the expansion

$$|\psi(t)\rangle = \sum_n \psi_n(t_0)e^{-iE_n(t-t_0)/\hbar}|n\rangle, \tag{2.497}$$

in which

$$\psi_n(t_0) = \langle n|\psi(t_0)\rangle. \tag{2.498}$$

Making this substitution into the expansion and doing some judicious re-arranging of terms, we find that

$$|\psi(t)\rangle = \sum_n \langle n|\psi(t_0)\rangle e^{-iE_n(t-t_0)/\hbar}|n\rangle = \sum_n |n\rangle e^{-iE_n(t-t_0)/\hbar}\langle n|\psi(t_0)\rangle, \tag{2.499}$$

in which we can, by comparison with the definition of the evolution operator, make the identification

$$U(t, t_0) = \sum_n |n\rangle e^{-iE_n(t-t_0)/\hbar}\langle n|. \tag{2.500}$$

Thus, the evolution operator is diagonal in the energy representation, and its diagonal elements are a simple function of the associated eigenvalues of the energy operator. We can therefore write the evolution operator as the corresponding function of the energy operator itself, i.e.,

$$U(t, t_0) = e^{-iH(t-t_0)/\hbar}. \tag{2.501}$$

In this form is clear that when $H$ is independent of time the evolution operator $U(t, t_0) = U(t - t_0)$ only depends on the length of the time interval over which the system is being evolved. Thus, we can write $U(t) = e^{-iHt/\hbar}$ as the operator which evolves the system for a time $t$. We also note that in this form the evolution operator is explicitly unitary, since for any Hermitian operator $A$, the adjoint of the operator $U = e^{iA}$ is the operator $U^+ = e^{-iA}$, and so

$$UU^+ = U^+U = e^{i(A-A)} = \mathbf{1}. \tag{2.502}$$

This form resembles another operator of this type, namely the **spatial translation operator**

$$T(\vec{L}) = e^{-i\vec{L}\cdot\vec{K}} = e^{-i\vec{L}\cdot\vec{P}/\hbar}, \tag{2.503}$$

which is a multiplicative operator in the momentum representation, but which in the position representation has the effect of shifting the wave function through a displacement $\vec{L}$, i.e.,

$$T(\vec{L})\psi(\vec{r}) = \psi(\vec{r} - \vec{L}). \tag{2.504}$$

Thus, if originally the peak of the wave function $\psi(\vec{r})$ was located at the origin (at $\vec{r}=0$), it will now be shifted to the point where the argument of $\psi(\vec{r}-\vec{L})$ is zero, i.e., to the point $\vec{r}=\vec{L}$. To see how this comes about, consider the corresponding one-dimensional version of this operator

$$T(L)=e^{-iLK_x}=\mathbf{1}+(-iLK_x)+\frac{(-iLK_x)^2}{2!}+\frac{(-iLK_x)^3}{3!}+\cdots \tag{2.505}$$

If we let this act on a state described by the wave function $\psi(x)$, and use the result appropriate to the position representation that $iK_x=d/dx$, we find that

$$T(L)\psi(x)=[\mathbf{1}-L\frac{d}{dx}+\frac{L^2}{2!}\frac{d^2}{dx^2}-\frac{L^3}{3!}\frac{d^3}{dx^3}+\cdots]\psi(x) \tag{2.506}$$

$$=\psi(x)-L\frac{d\psi}{dx}+\frac{L^2}{2!}\frac{d^2\psi}{dx^2}-\frac{L^3}{3!}\frac{d^3\psi(x)}{dx^3}+\cdots]=\psi(x-L), \tag{2.507}$$

where we see that the expansion of the exponential operator automatically generates the Taylor series for the function $\psi(x-L)$ expanded about the point $x$. In terms of the basis vectors of the position and momentum representation it is straightforward to show that $T(\vec{L})|\vec{r}\rangle=|\vec{r}+\vec{L}\rangle$ and that $T(\vec{L})|\vec{k}\rangle=e^{-i\vec{L}\cdot\vec{k}}|\vec{k}\rangle$.

Thus, by analogy, the evolution operator $U(t)$ is sometimes referred to as the **time-translation operator**. Translation operators for other observables are also easily defined as operator exponentials of this type. If the variable being shifted is an angle, then the "translation" is actually a rotation. The Hermitian operators which appear in exponents of these translation operators are referred to as the **generators** of the associated translation. Thus, $H$ is the generator of time translations, while $\vec{P}$ is the generator of spatial translations. Through similar analyses, one can establish the fact that $\vec{R}$ is the generator of translations in *momentum*, and that any component of the angular momentum operator $\vec{L}$ generates rotations about that axis.

*Time-Dependent Hamiltonian* - When the Hamiltonian is not independent of time no simple closed form expression for the evolution operator exists. It is possible, however, to develop an integral equation for the evolution operator that is sometimes quite useful. To this end we combine the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle=H(t)|\psi(t)\rangle \tag{2.508}$$

with the defining equation

$$|\psi(t)\rangle=U(t,t_0)|\psi(t_0)\rangle \tag{2.509}$$

for the evolution operator to obtain

$$i\hbar\frac{d}{dt}U(t,t_0)|\psi(t_0)\rangle=H(t)U(t,t_0)|\psi(t_0)\rangle. \tag{2.510}$$

Since this is valid for all initial state vectors $|\psi(t_0)\rangle$, we deduce that

$$i\hbar\frac{d}{dt}U(t,t_0)=H(t)U(t,t_0), \tag{2.511}$$

showing that the evolution operator itself obeys an operator form of the Schrödinger equation. Unlike the Schrödinger equation for the state vector, however, the operator $U$ has a well-defined initial condition, namely,

$$U(t_0,t_0)=\mathbf{1}, \tag{2.512}$$

which stems from the fact that $\lim_{t\to t_0} U(t,t_0)|\psi(t_0)\rangle = |\psi(t_0)\rangle$. This boundary condition on $U$ is sometimes described by saying that it is "smoothly connected to the unit operator". We can use this initial condition to formally integrate the Schrödinger equation for $U$ as follows. We write (treating $t_0$ simply as a parameter, not a variable of integration)

$$\frac{d}{dt'}U(t',t_0)dt' = -\frac{i}{\hbar}H(t')U(t',t_0)dt', \tag{2.513}$$

or

$$\int_{U(t_0,t_0)}^{U((t,t_0)} dU = -\frac{i}{\hbar}\int_{t_0}^{t} H(t')U(t',t_0)dt'. \tag{2.514}$$

Integrating and using the initial condition this becomes

$$U(t,t_0) = \mathbf{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')U(t',t_0). \tag{2.515}$$

This is only a formal solution because the right hand side contains the evolution operator itself; however it does have the initial condition already built in, and it can be *iterated* to obtain an expansion for $U$ in "powers of $H$". To do this, we repeatedly insert the whole integral for $U$ into the integral in which it appears

$$\begin{aligned} U(t,t_0) &= \mathbf{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')\left[\mathbf{1} - \frac{i}{\hbar}\int_{t_0}^{t'} dt^{''} H(t'')U(t'',t_0)\right] \\ &= \mathbf{1} + \left(\frac{-i}{\hbar}\right)\int_{t_0}^{t} dt' H(t') + \left(\frac{-i}{\hbar}\right)^2 \int_{t_0}^{t} dt' \int_{t_0}^{t'} dt^{''} H(t')H(t'') + \cdots \quad (2.516) \\ &= \sum_{n=0}^{\infty} U^{(n)}(t,t_0) \quad (2.517) \end{aligned}$$

where the $n$th term of the expansion has the general form

$$U^{(n)}(t,t_0) = \left(\frac{-i}{\hbar}\right)^n \int_{t_0}^{t} dt_n \int_{t_0}^{t_n} dt_{n-1}\cdots \int_{t_0}^{t_2} dt_1\, H(t_n)H(t_{n-1})\cdots H(t_1). \tag{2.518}$$

This and similar expansions for the evolution operator provide a useful starting point for the development of time-dependent perturbation theories. As a final note, we observe that this expansion for $U$ implies a similar integral equation and expansion for the state vector, i.e.,

$$\begin{aligned} |\psi(t)\rangle &= U(t,t_0)|\psi(t_0)\rangle \\ &= \left[\mathbf{1} + \left(\frac{-i}{\hbar}\right)\int_{t_0}^{t} dt' H(t')U(t',t_0)\right]|\psi(t_0)\rangle \quad (2.519) \end{aligned}$$

or

$$|\psi(t)\rangle = |\psi(t_0)\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')|\psi(t')\rangle, \tag{2.520}$$

where the second term on the right-hand-side of the last expression represents the change that has occurred in the state vector between times $t_0$ and $t$.