

Chapter 2

Mathematical Apparatus

Abstract The mathematical tools of quantum mechanics are summarized. This overview, which makes no attempt to be mathematically complete and rigorous, is intended as an introduction for readers unfamiliar with the subject. We begin with some geometrical analogies of the basic concepts and techniques of the mathematical formalism used to treat the extended Hilbert space of the quantum-mechanical states, the abstract vector space spanned by the *state* vectors or the associated wave functions of the physical system of interest. Dirac's vector notation, which greatly simplifies manipulations on these mathematical objects, and the alternative representations of the singular delta "function" are given. The linear operators acting on the state vectors as well as their adjoints are defined and the basis set representations of vectors and operators are introduced. The eigenvalue problem of the linear *self*-adjoint (Hermitian) operators is examined in some detail and the complete set of the commuting observables is defined. The two most important (continuous) bases of vectors for representing quantum states of a single particle, defined by the eigenvectors of the particle *position* and *momentum* operators, respectively, are explored. In particular, the position representation of the momentum operator, as well as the momentum representation of the position operator, are examined in some detail. Next, the discrete energy representation is briefly examined and the unitary transformation of states and operators is discussed. Finally, the functional derivatives are introduced and the associated Taylor expansion of functionals is formulated. The localized displacements of the functional argument function are defined using Dirac's delta function and the rules of functional differentiation are outlined stressing analogies to familiar operations performed on functions of many variables. The *chain* rule transformations of functional derivatives are summarized.

2.1 Geometrical Analogies

The ordinary *three-dimensional physical space* R^3 is spanned by the orthonormal basis $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\} \equiv \mathbf{e}^{(3)}$ (a row vector of vector elements), consisting of three unit vectors $\{\mathbf{e}_i, i = 1 \equiv x, 2 \equiv y, 3 \equiv z\}$ along the mutually perpendicular axes $\{x, y, z\}$, respectively, in the Cartesian coordinate system. The orthogonality of different basis vectors, $i \neq j$, expressed by the vanishing scalar product $\mathbf{e}_i \cdot \mathbf{e}_j = 0$, and their unit length (normalization), $\mathbf{e}_i \cdot \mathbf{e}_i = |\mathbf{e}_i|^2 \equiv e_i^2 = 1$, can be combined into the *orthonormality* relations expressed in terms of Kronecker's delta,

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{i,j} = \{1, \text{ for } i = j; 0, \text{ for } i \neq j\}, \quad (2.1a)$$

defining the *three-dimensional, unit-metric tensor* represented by the identity matrix $\mathbf{I}^{(3)} = \{\delta_{i,j}\}$:

$$\mathbf{e}^{(3)} \cdot \mathbf{e}^{(3)} \equiv \mathbf{e}^{(3)T} \mathbf{e}^{(3)} = \mathbf{I}^{(3)}, \quad (2.1b)$$

where $\mathbf{e}^{(3)T}$ denotes the *transposed* (T), *column* vector of transposed vector elements.

Any vector in R^3 can be expanded in this reference system,

$$\mathbf{A} = \mathbf{A}_x + \mathbf{A}_y + \mathbf{A}_z \equiv \sum_{i=1}^3 \mathbf{A}_i = i\mathbf{a}_x + j\mathbf{a}_y + k\mathbf{a}_z \equiv \sum_{i=1}^3 \mathbf{e}_i a_i = \mathbf{e}^{(3)} \mathbf{a}^{(3)T}, \quad (2.2)$$

with the row vector of coordinates $\mathbf{a}^{(3)} = \{a_i = \mathbf{e}_i \cdot \mathbf{A}\} = [a_x, a_y, a_z]$, measuring the lengths $\{a_i = |\mathbf{A}_i|\}$ of projections $\{\mathbf{A}_i\}$ of \mathbf{A} onto the corresponding axes, providing the *matrix representation* of \mathbf{A} in the adopted basis set: $\mathbf{A} \leftrightarrow \mathbf{a}^{(3)}$.

It should be also observed that in the preceding equation the resolution of \mathbf{A} into its projections $\{\mathbf{A}_i\}$ along the directions of basic vectors $\mathbf{e}^{(3)}$ in this coordinate system can be also interpreted as a result of acting on \mathbf{A} with the *projection operator* $\hat{\mathbf{P}}(R^3)$ onto the whole R^3 space,

$$\hat{\mathbf{P}}(R^3) = \sum_{i=1}^3 (\mathbf{e}_i \mathbf{e}_i \cdot) \equiv \sum_{i=1}^3 \hat{\mathbf{P}}(\mathbf{e}_i), \quad (2.3)$$

defined by the sum of individual projectors $\{\hat{\mathbf{P}}(\mathbf{e}_i)\}$ onto the specified axes. Indeed, the following identity directly follows from (2.2):

$$\mathbf{A} = \sum_{i=1}^3 \mathbf{e}_i a_i = \left(\sum_{i=1}^3 \mathbf{e}_i \mathbf{e}_i \cdot \right) \mathbf{A} = \hat{\mathbf{P}}(R^3) \mathbf{A} = \sum_{i=1}^3 \hat{\mathbf{P}}(\mathbf{e}_i) \mathbf{A} \equiv \sum_{i=1}^3 \mathbf{A}_i. \quad (2.4)$$

The preceding relation also implies that the projection of any vector \mathbf{A} in R^3 , or $A(R^3)$ for short, amounts to multiplying it by the unity (identity) operation

$\hat{P}(R^3) = 1$: $\hat{P}(R^3)\mathbf{A}(R^3) = \mathbf{A}(R^3)$. Clearly, the sum of projections onto any two basis vectors $\hat{P}(\mathbf{e}_i, \mathbf{e}_j) = \hat{P}(\mathbf{e}_i) + \hat{P}(\mathbf{e}_j)$ defines the projection onto the plane defined by these two axes:

$$\hat{P}(\mathbf{e}_i, \mathbf{e}_j)\mathbf{A} = \hat{P}(\mathbf{e}_i)\mathbf{A} + \hat{P}(\mathbf{e}_j)\mathbf{A} = \mathbf{A}_i + \mathbf{A}_j \equiv \mathbf{A}_{(i,j)}. \quad (2.5)$$

This overall projection onto the whole physical space allows one to interpret the scalar product of two vectors \mathbf{A} and \mathbf{B} in R^3 in terms of their coordinates $\mathbf{a}^{(3)}$ and $\mathbf{b}^{(3)}$, respectively:

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{A} \cdot \hat{P}(R^3)\mathbf{B} = \sum_{i=1}^3 (\mathbf{A} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{B}) = \sum_{i=1}^3 a_i b_i = \mathbf{a}^{(3)} \mathbf{b}^{(3)T}. \quad (2.6)$$

As seen from this example, the coordinate-resolved expression results directly from placing the identity operator $\hat{P}(R^3) = 1$ between the two vectors in the scalar product. Obviously, this formal manipulation has no effect on the product value.

The characteristic property of projections is that the effect of a singular projection is identical to that of the subsequent repetition of the same projection. This immediately implies the *idempotency* property of the projection operators,

$$\hat{P}(R^3)\hat{P}(R^3) \equiv [\hat{P}(R^3)]^2 = \hat{P}(R^3), \quad [\hat{P}(\mathbf{e}_i, \mathbf{e}_j)]^2 = \hat{P}(\mathbf{e}_i, \mathbf{e}_j), \quad [\hat{P}(\mathbf{e}_i)]^2 = \hat{P}(\mathbf{e}_i), \quad (2.7)$$

where we have identified the square of an operator as a double execution of the operation it symbolizes. One can straightforwardly verify these identities using the orthonormality relations of (2.1a, 2.1b), which also imply that the product of projections into the mutually orthogonal subspaces identically vanishes, e.g.,

$$\hat{P}(\mathbf{i})\hat{P}(\mathbf{k}) = \hat{P}(\mathbf{i})\hat{P}(\mathbf{j}) = \hat{P}(\mathbf{j})\hat{P}(\mathbf{k}) = \hat{P}(\mathbf{i}, \mathbf{j})\hat{P}(\mathbf{k}) = 0. \quad (2.8)$$

These observations can be naturally generalized into the n -dimensional *Euclidean* space R^n , spanned by n orthonormal basic vectors $\mathbf{e}^{(n)} = \{\mathbf{e}_i, i = 1, 2, \dots, n\}$, $\mathbf{e}^{(n)T} \cdot \mathbf{e}^{(n)} = \mathbf{I}^{(n)}$, also including the $n \rightarrow \infty$ limit. In particular, the matrix representations of vectors and the coordinate-resolved expression for the scalar product of vectors $\mathbf{A}(R^n)$ and $\mathbf{B}(R^n)$ directly follow from applying the projector onto the whole space R^n ,

$$\hat{P}(R^n) = \sum_{i=1}^n (\mathbf{e}_i \mathbf{e}_i \cdot) \equiv \sum_{i=1}^n \hat{P}(\mathbf{e}_i), \quad (2.9)$$

$$\begin{aligned} \mathbf{A}(R^n) &= \hat{P}(R^n)\mathbf{A}(R^n) = \sum_{i=1}^n \hat{P}(\mathbf{e}_i)\mathbf{A}(R^n) = \sum_{i=1}^n \mathbf{e}_i [\mathbf{e}_i \cdot \mathbf{A}(R^n)] \\ &= \sum_{i=1}^n \mathbf{e}_i a_i = \sum_{i=1}^n \mathbf{A}_i = \mathbf{e}^{(n)} \mathbf{a}^{(n)T}, \end{aligned} \quad (2.10)$$

$$\begin{aligned}
\mathbf{A}(R^n) \cdot \mathbf{B}(R^n) &= \mathbf{A}(R^n) \cdot \hat{\mathbf{P}}(R^n) \mathbf{B}(R^n) = \sum_{i=1}^n (\mathbf{A} \cdot \mathbf{e}_i)(\mathbf{e}_i \cdot \mathbf{B}) \\
&= \sum_{i=1}^n a_i b_i = \mathbf{a}^{(n)} \mathbf{b}^{(n)T}.
\end{aligned} \tag{2.11}$$

In particular, for two identical vectors $\mathbf{A}(R^n) = \mathbf{B}(R^n)$ one obtains the following expression for the vector length (norm):

$$A = |\mathbf{A}| = \sqrt{A^2} = \left(\sum_{i=1}^n a_i^2 \right)^{1/2} \geq 0. \tag{2.12}$$

One similarly defines the projection operators into various subspaces in R^n , e.g., its complementary, mutually orthogonal parts $P^m = \{\mathbf{e}_i, i = 1, 2, \dots, m\} \equiv \mathcal{P}$ and $Q^{n-m} = \{\mathbf{e}_i, j = m+1, m+2, \dots, n\} \equiv \mathcal{Q}$:

$$\begin{aligned}
\hat{\mathbf{P}}(P^m) &\equiv \hat{\mathbf{P}}_{\mathcal{P}} = \sum_{i \in \mathcal{P}} \hat{\mathbf{P}}(\mathbf{e}_i), \quad \hat{\mathbf{P}}(Q^{n-m}) \equiv \hat{\mathbf{P}}_{\mathcal{Q}} = \sum_{j \in \mathcal{Q}} \hat{\mathbf{P}}(\mathbf{e}_j), \quad \hat{\mathbf{P}}_{\mathcal{P}} \hat{\mathbf{P}}_{\mathcal{Q}} = 0, \\
\mathbf{A}(R^n) &= (\hat{\mathbf{P}}_{\mathcal{P}} + \hat{\mathbf{P}}_{\mathcal{Q}}) \mathbf{A}(R^n) = \mathbf{A}_{\mathcal{P}} + \mathbf{A}_{\mathcal{Q}},
\end{aligned} \tag{2.13}$$

where $\mathbf{A}_{\mathcal{P}}$ and $\mathbf{A}_{\mathcal{Q}}$ stand for the projections of $\mathbf{A}(R^n)$ into the P^m and Q^{n-m} subspaces, respectively.

The scalar product of (2.11) can be also given the (linear) *functional* interpretation. In mathematics the linear functional $F[\varphi]$ of the argument φ , e.g. a function or vector, is a linear operation performed on the argument, which gives the scalar quantity F , $F[\varphi] = F$, e.g., the definite integral $I[f] = \int_{x_1}^{x_2} f(x) dx = I$. The same property can be associated with the (discrete) scalar product, say a projection of the argument vector $\mathbf{A} \equiv \vec{A}$ onto another vector $\mathbf{B} \equiv \vec{B}$:

$$\mathbf{B} \cdot \mathbf{A} = \vec{B} \cdot \vec{A} \equiv \vec{B}[\vec{A}], \tag{2.14}$$

where $\vec{B}[\vec{V}]$ denotes the functional of the vector argument \vec{V} giving the value of its scalar product with the vector \vec{B} . The latter thus defines the functional $\vec{B}[\mathbf{X}]$ itself, denoted as the “reversed” vector, by specifying the *direction* onto which the argument vector \mathbf{X} is to be projected.

It can be then demonstrated that these scalar product functionals also span the vector space, called the *dual space*, since any combination of such quantities represents another linear functional of the same type. Let us examine these reversed “vector” quantities (functionals) associated with the independent basis vectors $\{\mathbf{e}_i = \vec{e}_i\}$. They represent the *dual* basis “vectors” $\{\vec{e}_i[\vec{V}] \equiv \vec{e}_i\}$ of the

scalar product functionals. Indeed, any combination of them also belongs to this dual space, e.g.,

$$C_i \bar{e}_i[\vec{V}] + C_j \bar{e}_j[\vec{V}] = (C_i \bar{e}_i + C_j \bar{e}_j) \cdot \vec{V} \equiv \vec{W} \cdot \vec{V} = \vec{W}[\vec{V}], \quad (2.15)$$

and to every vector \vec{A} corresponds its functional analog \bar{A} in the dual space, since the vector is uniquely specified by the complete set of its scalar products (components) with all independent vectors $\mathbf{e}^{(n)}$:

$$\vec{A} = \sum_{i=1}^n a_i \bar{e}_i = \sum_{i=1}^n \bar{A}_i \Rightarrow \bar{A}[\vec{V}] = \sum_{i=1}^n a_i \bar{e}_i[\vec{V}] = \sum_{i=1}^n \bar{A}_i[\vec{V}]. \quad (2.16)$$

It also follows from these relations that in Euclidean space this correspondence is linear: the linear combination of vectors in R^n is represented in the associated dual space by the associated combination, with the same expansion coefficients, of the corresponding dual-space functionals.

It should be emphasized that the dual-space elements, the “reversed” vectors, represent mathematical quantities (functionals of vectors) quite different from the original (argument) vectors on which they act.

2.2 Dirac's Vector Notation and Delta Function

In accordance with the *Superposition Principle* of quantum mechanics (Dirac 1967), any combination of states represents an admissible quantum state of the given molecular or atomic system. This property is also typical of ordinary vectors, $C_A \mathbf{A} + C_B \mathbf{B} = \mathbf{C}$, where the numerical coefficients C_A and C_B determine the relative participation of both vectors in the combination. We shall use this analogy in the vector notation of Dirac, in which the quantum states Ψ and Φ are denoted as arrowed “ket” symbols $|\Psi\rangle$, $|\Phi\rangle$, \dots , called *state vectors*. Their linear combination $C_\Psi |\Psi\rangle + C_\Phi |\Phi\rangle = |\Theta\rangle$ determines another state $|\Theta\rangle$. When these states are functions of the continuous parameter $x \in [\xi, \zeta]$, $|\Psi\rangle = |\Psi(x)\rangle \equiv |x\rangle$, this summation of vector states is generalized into its continuous (integral) analog:

$$|\Theta\rangle = \int_{\xi}^{\zeta} c(x) |x\rangle dx. \quad (2.17)$$

Here, the combination coefficients $\{c(x), c(x'), \dots\}$ are in general complex since the quantum states are complex entities. The resultant state $|\Theta\rangle$ of the given combination is said to be *dependent* upon the component states $\{|x\rangle, |x'\rangle, \dots\}$. These *independent* state vectors cannot be expressed as combinations, with nonvanishing coefficients, of the remaining states in this basis set.

In the quantum kinematics it is the *direction* of the state vector $|\Psi\rangle$ that matters and uniquely identifies the quantum state Ψ . Therefore, the opposite state vectors along the same direction, e.g., $|\Psi\rangle$ and $-|\Psi\rangle$, in fact represent the same state Ψ , and any combination of the state with itself, $C_1|\Psi\rangle + C_2|\Psi\rangle = (C_1 + C_2)|\Psi\rangle = C|\Psi\rangle \equiv Me^{i\phi}|\Psi\rangle$, where M and ϕ stand for the modulus and phase of the complex coefficient C , also denote the same state Ψ . As we shall see later in this chapter, the length (norm) of the state vectors in quantum mechanics will be fixed by the appropriate *normalization* requirement resulting from the probabilistic interpretation of quantum states. In case of the square integrable wave functions it calls for $M = 1$, but the phase ϕ will be left undetermined as immaterial and having no physical meaning.

This property of the quantum superposition rule distinguishes it from the corresponding classical principle, e.g., that for combining vibrations of a string or a membrane, in which the combination of a state with itself gives another state exhibiting different amplitude. There is also another important distinction between the quantum and classical kinematics: in quantum mechanics the state vector of the vanishing norm (length), which thus has no specified direction in the vector space of quantum states, does not exist and thus has no physical meaning, while the classical vibration of the vanishing amplitude everywhere does in fact represent the real physical state of rest of a string or a membrane.

It was shown in the preceding section that to any vector space the dual space of the “reversed” vectors, the entities of quite different mathematical variety (functionals), can be ascribed through the concept of the scalar product (projection) of the vectors themselves. The dual space to the *ket*-space of state vectors $\{|\Psi_i\rangle\}$ is called the *bra*-space of the reversed “vectors” (functionals) $\{\langle\Psi_i|\}$, with the one-to-one (antilinear) correspondence: $\langle\Psi_i| \leftrightarrow |\Psi_i\rangle$, $(\langle\Psi_i| + \langle\Psi_j|) \leftrightarrow (|\Psi_i\rangle + |\Psi_j\rangle)$, $C^*\langle\Psi| \leftrightarrow C|\Psi\rangle$, etc., where C^* denotes the *complex* conjugate of C . In the original terminology of Dirac the *bra*-“vector” $\langle\Psi|$ represents the *conjugate-imaginary* of the associated *ket*-vector $|\Psi\rangle$. Again, the basic difference between the elements of the two vector spaces, with the “*bras*” in fact representing the functionals acting on “*kets*,” it is improper to regard the *bra*-“vectors” as the *complex conjugates* of the corresponding *ket*-vectors.

In Dirac’s notation the bra $\langle\Phi|$ and ket $|\Psi\rangle$ symbols are examples of an *incomplete* “bracket,” while the result of $\langle\Phi|$ acting on $|\Psi\rangle$ gives the *complete* bracket of the *scalar product* of $|\Psi\rangle$ and $|\Phi\rangle$, $\langle\Phi|\Psi\rangle \equiv \Phi[|\Psi\rangle]$, which measures the projection of $|\Psi\rangle$ on $|\Phi\rangle$. The complete bracket generates the complex number. This association also explains the English nomenclature of the “bra” and “ket” symbols. This definition also implies that in contrast to the Euclidean space the complex numbers of the projections of $|\Psi\rangle$ on $|\Phi\rangle$ and of $|\Phi\rangle$ on $|\Psi\rangle$, respectively, are not equal in general, one representing the complex conjugate of the other:

$$\langle\Phi|\Psi\rangle = \Phi[|\Psi\rangle] \equiv \langle\Psi|\Phi\rangle^* = \Psi[|\Phi\rangle]^*. \quad (2.18)$$

One also observes that this linear functional of the ket vector:

$$\Phi[C_1\Psi_1 + C_2\Psi_2] = C_1\Phi[\Psi_1] + C_2\Phi[\Psi_2], \quad (2.19)$$

is antilinear with respect to the bra vector, which determines the direction on which the projection is made:

$$\langle C_1\Phi_1 + C_2\Phi_2 | \Psi \rangle = C_1^* \Phi_1[|\Psi\rangle] + C_2^* \Phi_2[|\Psi\rangle]. \quad (2.20)$$

Any vector in the ket space has its unique analog in the dual space of the bra “vectors” (functionals). There is a close analogy with the Euclidean space, in which the scalar product functional has also been used to define the dual “vector”. Indeed the vector is uniquely defined by its projections on all (independent, orthonormal) vectors $\{|X_i\rangle = |i\rangle\}$, possibly including indenumerable vectors $\{|X(x)\rangle \equiv |x\rangle\}$ labeled by the continuous parameter(s) x . The set of projections $\{\langle\Phi|X_i\rangle = \langle X_i|\Phi\rangle^*\}$ thus uniquely determines the original ket $|\Phi\rangle$ associated with the functional $\Phi[] = \langle\Phi|$.

The “orthonormality” relations for the continuous basis vectors $\{|x\rangle\}$ are expressed in terms of the continuous analog of the Koronecker delta $\delta_{ij} = \langle i|j\rangle$, called the *Dirac delta* “function” $\delta(x' - x) = \langle x|x'\rangle$. For any function $f(x)$ of the continuous argument(s) x this kernel satisfies the following “projection” identity:

$$f(x) = \int \delta(x' - x) f(x') dx'. \quad (2.21)$$

This equation indicates that this singular function represents the kernel of the integral operator $\int dx' \delta(x' - x)$, which acting on function $f(x')$ generates $f(x)$. Moreover, since the integral of the preceding equation formally expresses the functional $f(x) = f[f(x')]$, Dirac's delta can also be interpreted as the functional derivative (see Sect. 2.7):

$$\delta(x' - x) = \frac{\delta f(x)}{\delta f(x')}. \quad (2.22)$$

We shall discuss other properties of this mathematical entity later in this section.

The Dirac delta function $\delta(x' - x)$ of (2.21) represents the *unity*-normalized, $\int \delta(x' - x) dx' = 1$, infinitely sharp distribution centered at $x' = x$, exhibiting vanishing values at any finite distance from this point. It can be thus envisaged as the limiting form of the ordinary Gaussian (normal) distribution of the probability theory in the limit of the vanishing variance:

$$\delta(x' - x) = \lim_{\sigma \rightarrow 0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x' - x)^2}{2\sigma^2}\right). \quad (2.23)$$

Alternatively, one can use any complete, say discrete, set of orthonormal basis functions $\{\chi_i(x)\}$, $\int \chi_i^*(x) \chi_j(x) dx = \delta_{ij}$, to generate the analytical representation of this singular function. Indeed, expanding $f(x)$ in terms of the complete (orthonormal) basis set $\{\chi_i(x)\}$ gives:

$$\begin{aligned}
 f(x) &= \sum_i \chi_i(x) c_i = \sum_i \chi_i(x) \left[\int \chi_i^*(x') f(x') dx' \right] \\
 &= \int \left\{ \sum_i \chi_i^*(x') \chi_i(x) \right\} f(x') dx'.
 \end{aligned} \tag{2.24}$$

Hence, comparing the last equation with (2.21) gives the *closure* relation:

$$\delta(x' - x) = \sum_i \chi_i(x) \chi_i^*(x'). \tag{2.25a}$$

For the continuous orthonormal basis set $\{u_\alpha(x)\}$ labeled by the continuous index α , $\int u_\alpha^*(x) u_{\alpha'}(x) dx = \delta(\alpha' - \alpha)$, one similarly finds

$$\delta(x' - x) = \int u_\alpha(x) u_\alpha^*(x') d\alpha. \tag{2.25b}$$

When the complete basis set is “mixed,” containing the discrete and continuous parts, $\{\chi_i(x), u_\alpha(x)\}$, with $\int u_\alpha^*(x) \chi_i(x) dx = 0$, this closure relation reads

$$\delta(x' - x) = \sum_i \chi_i(x) \chi_i^*(x') + \int u_\alpha(x) u_\alpha^*(x') d\alpha. \tag{2.25c}$$

Another important example of the continuous analytical representation of Dirac’s delta originates from the Fourier-transform relations, e.g., between the wave function in the position and momentum representations of quantum mechanics (see Sect. 2.6),

$$\begin{aligned}
 \Phi(k) &= \frac{1}{\sqrt{2\pi}} \int \exp(-ikx) f(x) dx \quad \text{and} \quad f(x) = \frac{1}{\sqrt{2\pi}} \int \exp(ik'x) \Phi(k') dk', \\
 i &= \sqrt{-1}.
 \end{aligned} \tag{2.26}$$

Substituting the second, inverse transformation into the first one then gives

$$\Phi(k) = \frac{1}{2\pi} \int \Phi(k') \left\{ \int \exp[ix(k' - k)] dx \right\} dk' \tag{2.27}$$

and hence

$$\delta(k' - k) = \frac{1}{2\pi} \int \exp[ix(k' - k)] dx. \tag{2.28}$$

The singular Dirac delta function $\delta(x' - x) \equiv \delta(z)$ satisfies the following identities:

$$\begin{aligned}
\delta(z) &= \delta(-z), \quad z\delta(z) = 0, \\
\delta(az) &= |a|^{-1}\delta(z), \quad f(x')\delta(x' - x) = f(x)\delta(x' - x), \\
\int \delta(x' - x) \delta(x - x'') dx &= \delta(x' - x''), \\
\delta(x^2 - a^2) &= (2|a|)^{-1}[\delta(x - a) + \delta(x + a)]. \tag{2.29}
\end{aligned}$$

Of interest also are the related properties of the derivative of Dirac's delta "function," $\delta'(z) \equiv d\delta(z)/dz$,

$$\int f(z)\delta'(z)dz = -f'(0) \quad \text{or} \quad \int f(z)\delta'(-z)dz = f'(0), \quad z\delta'(z) = -\delta(z). \tag{2.30}$$

2.3 Linear Operators and Their Adjoints

The complex number resulting from the scalar product between two state vectors is the result of applying the functional represented by its *bra* factor to its *ket* argument. When the linear action of a mathematical object on ket results in another ket, i.e., when it attributes in the linear fashion the uniquely specified *result*-vector $|\Psi'\rangle$ to the given *argument*-vector $|\Psi\rangle$, it is said to define the linear *operator* \hat{A} :

$$\hat{A}|\Psi\rangle = |\hat{A}\Psi\rangle \equiv |\Psi'\rangle, \quad \hat{A}(C_1\Psi_1 + C_2\Psi_2) = C_1\hat{A}|\Psi_1\rangle + C_2\hat{A}|\Psi_2\rangle. \tag{2.31}$$

The operator is defined when its action on every ket is determined; it becomes zero, $\hat{A} = 0$, when its action on every ket $|\Psi\rangle$ gives zero. Thus, two operators are equal when they produce equal results when applied to every ket.

The linear operators can be added and multiplied:

$$(\hat{A} + \hat{B})|\Psi\rangle = \hat{A}|\Psi\rangle + \hat{B}|\Psi\rangle, \quad (\hat{A}\hat{B})|\Psi\rangle = \hat{A}(\hat{B}|\Psi\rangle) \equiv \hat{A}\hat{B}|\Psi\rangle. \tag{2.32}$$

In general, they do not commute, giving rise to nonvanishing *commutator*

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A} \neq 0. \tag{2.33}$$

A multiplication by a number is a trivial case of a linear operation, which commutes with all linear operators. It can be easily verified that commutators satisfy the following identities:

$$[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}], \quad [\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}],$$

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}], \quad [\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0. \quad (2.34)$$

Linear operators can also act on the bra vectors, with the latter always put to the left of the operator, giving other bras. Indeed, the symbol $\hat{A}\langle\Phi|$ has no meaning of the bra vector (functional), since its action on the ket vector $|\Psi\rangle$ gives another operator, $(\hat{A}\langle\Phi|)|\Psi\rangle = \hat{A}\langle\Phi|\Psi\rangle = \langle\Phi|\Psi\rangle\hat{A}$, thus representing an alien object in the present mathematical formalism. However, it can be straightforwardly demonstrated, again using the scalar product functional as the link to the definition of (2.31), that $\langle\Phi|\hat{A} = \langle\Phi|$. Indeed, since \hat{A} is linear and the scalar product depends linearly on the ket, the scalar products $\Phi[\hat{A}|\Psi] = \langle\Phi|(\hat{A}|\Psi\rangle)$ for the specified $\langle\Phi|$ and \hat{A} , associate with every ket $|\Psi\rangle$ in the vector space a number which depends linearly on $|\Psi\rangle$. This new linear functional thus defines a new bra vector $\langle\Phi|$, which can be regarded as a result of \hat{A} acting on $\langle\Phi|$:

$$\langle\Phi|(\hat{A}|\Psi\rangle) = (\langle\Phi|\hat{A})|\Psi\rangle = \langle\Phi|\Psi\rangle. \quad (2.35)$$

Therefore, the linear operators act either on bras to their left or on kets to their right. In other words, the position of parentheses in the above matrix element of \hat{A} is of no importance:

$$\langle\Phi|(\hat{A}|\Psi\rangle) = (\langle\Phi|\hat{A})|\Psi\rangle = \langle\Phi|\hat{A}|\Psi\rangle. \quad (2.36)$$

The operation $\langle\Phi|\hat{A} = \langle\Phi|$ is linear, because for arbitrary $|\Psi\rangle$ and $\langle\Omega| = C_1\langle\Phi_1| + C_2\langle\Phi_2|$ one obtains:

$$\begin{aligned} \langle\Omega|\hat{A}|\Psi\rangle &= \langle\Omega|(\hat{A}|\Psi\rangle) = C_1\langle\Phi_1|(\hat{A}|\Psi\rangle) + C_2\langle\Phi_2|(\hat{A}|\Psi\rangle) \\ &= C_1(\langle\Phi_1|\hat{A})|\Psi\rangle + C_2(\langle\Phi_2|\hat{A})|\Psi\rangle, \end{aligned} \quad (2.37)$$

and hence $\langle\Omega|\hat{A} = C_1\langle\Phi_1|\hat{A} + C_2\langle\Phi_2|\hat{A}$.

It can be directly verified that the product of the ket and bra vectors, $|\Psi\rangle\langle\Phi|$, represents an *operator*. When acting on ket $|\Xi\rangle$ it generates another ket vector along $|\Psi\rangle$, $|\Psi\rangle\langle\Phi|\Xi\rangle = \langle\Phi|\Xi\rangle|\Psi\rangle$, while the result of its action on bra $\langle\Omega|$ produces another bra vector (functional), proportional to $\langle\Phi|$: $\langle\Omega|\Psi\rangle\langle\Phi|$. It thus defines the linear operator:

$$\begin{aligned} |\Psi\rangle\langle\Phi|C_1\Xi_1 + C_2\Xi_2 &= C_1\langle\Phi|\Xi_1\rangle|\Psi\rangle + C_2\langle\Phi|\Xi_2\rangle|\Psi\rangle, \\ (C_1\langle\Omega_1| + C_2\langle\Omega_2|)|\Psi\rangle\langle\Phi| &= C_1\langle\Omega_1|\Psi\rangle\langle\Phi| + C_2\langle\Omega_2|\Psi\rangle\langle\Phi|. \end{aligned} \quad (2.38)$$

In particular, the operator $|X_i\rangle\langle X_i|$ defined by the normalized vector $|X_i\rangle \equiv |i\rangle$ and its *bra* conjugate amounts to the projection onto the $|i\rangle$ direction:

$$|i\rangle\langle i|\Psi\rangle \equiv \hat{P}_i|\Psi\rangle = \langle i|\Psi\rangle|i\rangle \equiv \Psi_i|i\rangle, \quad (2.39)$$

where Ψ_i stands for i th component of $|\Psi\rangle$ in the $|\mathbf{i}\rangle = \{|\mathbf{i}\rangle\}$ representation (row vector). The projector idempotency then directly follows:

$$\hat{P}_i^2 = |\mathbf{i}\rangle\langle\mathbf{i}|\mathbf{i}\rangle\langle\mathbf{i}| = |\mathbf{i}\rangle\langle\mathbf{i}| = \hat{P}_i. \quad (2.40)$$

When this discrete (countable) basis set spans the complete space, the sum of all such projectors, i.e., the projection on the whole space, amounts to the identity operation,

$$\hat{P} = |\mathbf{i}\rangle\langle\mathbf{i}| = \sum_i \hat{P}_i = 1, \quad (2.41a)$$

where $\langle\mathbf{i}|$ stands for the column vector of bras associated with the row vector of the basis kets $|\mathbf{i}\rangle$, because then $\hat{P}|\Psi\rangle = |\Psi\rangle$. Similarly, when the complete basis set $|\mathbf{x}\rangle = \{|\mathbf{x}\rangle\}$ is noncountable in character, with the orthonormality relations expressed by Dirac's delta "function" of (2.21), the summation is replaced by the integral over the continuous parameter(s),

$$\hat{P} \equiv |\mathbf{x}\rangle\langle\mathbf{x}| = \int |\mathbf{x}\rangle\langle\mathbf{x}| dx = \int \hat{P}(x) dx = 1, \quad (2.41b)$$

where we have again interpreted $|\mathbf{x}\rangle$ and $\langle\mathbf{x}|$ as the (*continuous*) row and column vectors, respectively. Finally, when the complete (mixed) basis contains both the discrete part $|\boldsymbol{\alpha}\rangle = \{|\boldsymbol{\alpha}\rangle\}$ and the indenumerable subspace $|\mathbf{y}\rangle = \{|\mathbf{y}\rangle\}$, $|\mathbf{m}\rangle = [|\boldsymbol{\alpha}\rangle, |\mathbf{y}\rangle]$ the identity operator of the complete overall projection operator includes both the discrete and continuous projections:

$$\hat{P} \equiv |\mathbf{m}\rangle\langle\mathbf{m}| = |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| + |\mathbf{y}\rangle\langle\mathbf{y}| = \sum_{\alpha} \hat{P}_{\alpha} + \int \hat{P}(y) dy = 1. \quad (2.41c)$$

The (antilinear) one-to-one correspondence between kets and bras associates with every linear operator \hat{A} its *adjoint* (linear) operator \hat{A}^\dagger by the requirement that the bra associated with the ket $\hat{A}|\Psi\rangle = |\hat{A}\Psi\rangle \equiv |\Psi'\rangle$ is given by the result of action of \hat{A}^\dagger on the bra associated with $|\Psi\rangle$:

$$\langle\Psi'| = \langle\hat{A}\Psi| \equiv \langle\Psi|\hat{A}^\dagger. \quad (2.42)$$

Hence, since $\langle\Phi|\hat{A}\Psi\rangle = \langle\hat{A}\Psi|\Phi\rangle^*$ one obtains:

$$\langle\Phi|\hat{A}\Psi\rangle \equiv \langle\Phi|\hat{A}|\Psi\rangle = \langle\hat{A}\Psi|\Phi\rangle^* \equiv \langle\Psi|\hat{A}^\dagger|\Phi\rangle^*. \quad (2.43)$$

Moreover, because $(\hat{A}^\dagger)^\dagger = \hat{A}$ and hence $\langle\hat{A}^\dagger\Phi| = \langle\Phi|\hat{A}$, the adjoint operators can be alternatively defined by the identity:

$$\langle\hat{A}^\dagger\Phi|\Psi\rangle = \langle\Phi|\hat{A}|\Psi\rangle = \langle\Phi|\hat{A}\Psi\rangle. \quad (2.44)$$

Next, it is easy to show that $(\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger$ and $(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$. To determine the adjoint of the product of two operators one observes that the ket $|\Omega\rangle = \hat{A}\hat{B}|\Psi\rangle \equiv \hat{A}|\Theta\rangle$ is associated with the bra

$$\langle\Omega| = \langle\Psi|(\hat{A}\hat{B})^\dagger = \langle\Theta|\hat{A}^\dagger = \langle\Psi|\hat{B}^\dagger\hat{A}^\dagger, \quad (2.45)$$

where we have realized that the bra associated with $|\Theta\rangle$, $\langle\Theta| = \langle\Psi|\hat{B}^\dagger$. Hence, $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$. This change of order, when one takes the adjoint of a product of operators, can be generalized to an arbitrary number of them: $((\hat{A}\hat{B}\dots\hat{C}))^\dagger = \hat{C}^\dagger\dots\hat{B}^\dagger\hat{A}^\dagger$. One also observes that the following identity is satisfied for commutators: $[\hat{A}, \hat{B}]^\dagger = [\hat{B}^\dagger, \hat{A}^\dagger]$.

We can now summarize the mutual relations between the mathematical entities hitherto introduced in terms of the general *Hermitian* conjugation denoted by the adjoint symbol “ \dagger ”. In the Dirac notation the ket $|\Psi\rangle$ and its associated bra $\langle\Psi|$ are said to be Hermitian conjugates of each other: $\langle\Psi| = |\Psi\rangle^\dagger$ and $|\Psi\rangle = \langle\Psi|^\dagger$. Moreover, the operators \hat{A} and \hat{A}^\dagger are also related by the Hermitian conjugation. As we have observed in the preceding equation the hermitian conjugation of the product of operator factors changes the order in the product of the adjoint operators. This rule holds for other entities as well. For example, the Hermitian conjugate of $\hat{A}|\Psi\rangle$ gives:

$$(\hat{A}|\Psi\rangle)^\dagger = |\hat{A}\Psi\rangle^\dagger = |\Psi\rangle^\dagger\hat{A}^\dagger = \langle\Psi|\hat{A}^\dagger. \quad (2.46)$$

Similarly,

$$\begin{aligned} (|\Psi\rangle\langle\Phi|)^\dagger &= (\langle\Phi|^\dagger)(|\Psi\rangle^\dagger) = |\Phi\rangle\langle\Psi|, \quad (\langle\Phi|\Psi\rangle)^\dagger = (|\Psi\rangle^\dagger)(\langle\Phi|^\dagger) = \langle\Psi|\Phi\rangle, \\ (\lambda\langle\Phi|\Psi\rangle|\Psi\rangle\langle\Phi|)^\dagger &= |\Phi\rangle\langle\Psi|\langle\Psi|\Phi\rangle\lambda^* = \lambda^*\langle\Psi|\Phi\rangle|\Phi\rangle\langle\Psi|, \text{ etc.} \end{aligned} \quad (2.47)$$

Thus, to obtain the adjoint (Hermitian conjugate) of any expression composed of constants, kets, bras and linear operators, one replaces the constants by their complex conjugates, kets by the associated bras, bras by the associated kets, operators by their adjoints and reverses the order of factors in the products. However, as we have observed in the last line of (2.47), the position of constants, λ^* , $\langle\Psi|\Phi\rangle$, etc., is of no importance.

2.4 Basis Set Representations of Vectors and Operators

Selection of the complete (orthonormal) basis of the reference ket vectors in the vector space of the system quantum-mechanical states, either discrete $|i\rangle = \{|i\rangle\}$, $\langle i|j\rangle = \delta_{i,j}$, or the continuous infinity of vectors $|x\rangle = \{|x\rangle\}$, $\langle x|x'\rangle = \delta(x'-x)$, defines the specific *representation* in which both the vectors and operators can be expressed. By convention the basis vectors $|i\rangle$ and $|x\rangle$ are arranged as the *row*

vectors. Accordingly, their Hermitian conjugates define the respective column vectors of the bra basis: $|\mathbf{i}\rangle^\dagger = \langle \mathbf{i}|$ and $|\mathbf{x}\rangle^\dagger = \langle \mathbf{x}|$.

Using the closure relations of (2.41a), (2.41b) and the above orthonormality relations for these basis vectors gives the associated expansions of any ket $|\Psi\rangle$:

$$\begin{aligned} |\Psi\rangle &= \sum_i |\mathbf{i}\rangle \langle \mathbf{i}|\Psi\rangle = \sum_i |\mathbf{i}\rangle \Psi_i = |\mathbf{i}\rangle \langle \mathbf{i}|\Psi\rangle \equiv |\mathbf{i}\rangle \boldsymbol{\Psi}^{(i)}, \\ |\Psi\rangle &= \int |\mathbf{x}\rangle \langle \mathbf{x}|\Psi\rangle d\mathbf{x} = \int |\mathbf{x}\rangle \Psi(\mathbf{x}) d\mathbf{x} \equiv |\mathbf{x}\rangle \langle \mathbf{x}|\Psi\rangle \equiv |\mathbf{x}\rangle \boldsymbol{\Psi}^{(\mathbf{x})}. \end{aligned} \quad (2.48a)$$

The components $\{\Psi_i\}$ or $\{\Psi(\mathbf{x}), \Psi(\mathbf{x}'), \dots\}$, by convention arranged vertically as the *column* vectors, $\boldsymbol{\Psi}^{(i)} = \langle \mathbf{i}|\Psi\rangle$ and $\boldsymbol{\Psi}^{(\mathbf{x})} = \langle \mathbf{x}|\Psi\rangle$, provide the representations of the ket $|\Psi\rangle$ in the basis sets $|\mathbf{i}\rangle$ and $|\mathbf{x}\rangle$, respectively. In the mixed basis set case of (2.41c) the expansion of ket $|\Psi\rangle$ in $|\mathbf{m}\rangle$ will contain both the discrete and continuous components:

$$\begin{aligned} |\Psi\rangle &= |\mathbf{m}\rangle \langle \mathbf{m}|\Psi\rangle \equiv |\mathbf{m}\rangle \boldsymbol{\Psi}^{(\mathbf{m})} = |\boldsymbol{\alpha}\rangle \langle \boldsymbol{\alpha}|\Psi\rangle + |\mathbf{y}\rangle \langle \mathbf{y}|\Psi\rangle = |\boldsymbol{\alpha}\rangle \boldsymbol{\Psi}^{(\boldsymbol{\alpha})} + |\mathbf{y}\rangle \boldsymbol{\Psi}^{(\mathbf{y})} \\ &= \sum_{\boldsymbol{\alpha}} |\boldsymbol{\alpha}\rangle \langle \boldsymbol{\alpha}|\Psi\rangle + \int |\mathbf{y}\rangle \langle \mathbf{y}|\Psi\rangle d\mathbf{y}. \end{aligned} \quad (2.48b)$$

The associated expansions of the bra vector $\langle \Phi|$ in terms of the reference bra vectors $\langle \mathbf{i}|$, $\langle \mathbf{x}|$, and $\langle \mathbf{m}|$, respectively, directly follow from applying the corresponding *unity*-projections of (2.41a)–(2.41c) to $\langle \Phi|$ (from the right):

$$\begin{aligned} \langle \Phi| &= \sum_i \langle \Phi|\mathbf{i}\rangle \langle \mathbf{i}| = \sum_i \Phi_i^* \langle \mathbf{i}| = \langle \Phi|\mathbf{i}\rangle \langle \mathbf{i}| \equiv \boldsymbol{\Phi}^{(i)\dagger} \langle \mathbf{i}|, \\ \langle \Phi| &= \int \langle \Phi|\mathbf{x}\rangle \langle \mathbf{x}| d\mathbf{x} = \int \Phi^*(\mathbf{x}) \langle \mathbf{x}| d\mathbf{x} = \langle \Phi|\mathbf{x}\rangle \langle \mathbf{x}| \equiv \boldsymbol{\Phi}^{(\mathbf{x})\dagger} \langle \mathbf{x}|, \\ \langle \Phi| &= \langle \Phi|\mathbf{m}\rangle \langle \mathbf{m}| \equiv \boldsymbol{\Phi}^{(\mathbf{m})\dagger} \langle \mathbf{m}| = \langle \Phi|\boldsymbol{\alpha}\rangle \langle \boldsymbol{\alpha}| + \langle \Phi|\mathbf{y}\rangle \langle \mathbf{y}| \equiv \boldsymbol{\Phi}^{(\boldsymbol{\alpha})\dagger} \langle \boldsymbol{\alpha}| + \boldsymbol{\Phi}^{(\mathbf{y})\dagger} \langle \mathbf{y}|. \end{aligned} \quad (2.49)$$

Therefore, the vector components $\boldsymbol{\Phi}^{(i)\dagger} = \langle \Phi|\mathbf{i}\rangle$, $\boldsymbol{\Phi}^{(\mathbf{x})\dagger} = \langle \Phi|\mathbf{x}\rangle$ and $[\{\Phi_{\boldsymbol{\alpha}}^*\}, \Phi^*(\mathbf{y})] = [\langle \Phi|\boldsymbol{\alpha}\rangle, \langle \Phi|\mathbf{y}\rangle]$, when arranged horizontally as the associated *row* vectors, constitute the corresponding representations of $\langle \Phi|$ in these three types of the basis set. Again, the continuous representation of the bra vector, e.g., the complex conjugate wave function $\Phi^*(\mathbf{x}) = \langle \Phi|\mathbf{x}\rangle$, can also be regarded as the continuous *row* vector with components $[\Phi^*(\mathbf{x}), \Phi^*(\mathbf{x}'), \dots]$.

In these three types of the basis sets, the linear operator \hat{A} is accordingly represented by the square matrix and/or the continuous kernel, respectively,

$$\begin{aligned}
\mathbf{A}(\mathbf{i}, \mathbf{i}') &\equiv \langle \mathbf{i} | \hat{\mathbf{A}} | \mathbf{i}' \rangle = \{A_{i, i'} = \langle i | \hat{\mathbf{A}} | i' \rangle\} \equiv \mathbf{A}^{(i)}, \\
\mathbf{A}(\mathbf{x}, \mathbf{x}') &\equiv \langle \mathbf{x} | \hat{\mathbf{A}} | \mathbf{x}' \rangle = \{A(x, x') = \langle x | \hat{\mathbf{A}} | x' \rangle\} \equiv \mathbf{A}^{(x)}, \\
\mathbf{A}(\mathbf{m}, \mathbf{m}') &\equiv \langle \mathbf{m} | \hat{\mathbf{A}} | \mathbf{m}' \rangle = \begin{bmatrix} \mathbf{A}(\boldsymbol{\alpha}, \boldsymbol{\alpha}') = \langle \boldsymbol{\alpha} | \hat{\mathbf{A}} | \boldsymbol{\alpha}' \rangle & \mathbf{A}(\boldsymbol{\alpha}, \mathbf{y}') = \langle \boldsymbol{\alpha} | \hat{\mathbf{A}} | \mathbf{y}' \rangle \\ \mathbf{A}(\mathbf{y}, \boldsymbol{\alpha}') = \langle \mathbf{y} | \hat{\mathbf{A}} | \boldsymbol{\alpha}' \rangle & \mathbf{A}(\mathbf{y}, \mathbf{y}') = \langle \mathbf{y} | \hat{\mathbf{A}} | \mathbf{y}' \rangle \end{bmatrix} \equiv \mathbf{A}^{(m)}.
\end{aligned} \tag{2.50}$$

The adjoint operator $\hat{\mathbf{A}}^\dagger$ is similarly represented by the corresponding Hermitian conjugates of these “matrices,”

$$\begin{aligned}
\langle \mathbf{i} | \hat{\mathbf{A}}^\dagger | \mathbf{i}' \rangle &= \langle \hat{\mathbf{A}} \mathbf{i} | \mathbf{i}' \rangle = \langle \mathbf{i}' | \hat{\mathbf{A}} \mathbf{i} \rangle^* = \mathbf{A}(\mathbf{i}', \mathbf{i})^* = [\mathbf{A}(\mathbf{i}, \mathbf{i}')^*]^T = \mathbf{A}(\mathbf{i}, \mathbf{i}')^\dagger \equiv \mathbf{A}^{\dagger(i)}, \\
\langle \mathbf{x} | \hat{\mathbf{A}}^\dagger | \mathbf{x}' \rangle &= \langle \mathbf{x}' | \hat{\mathbf{A}} | \mathbf{x} \rangle^* = \langle \mathbf{x} | \hat{\mathbf{A}} | \mathbf{x}' \rangle^\dagger = \mathbf{A}(\mathbf{x}, \mathbf{x}')^\dagger = [\mathbf{A}(\mathbf{x}, \mathbf{x}')^*]^T \equiv \mathbf{A}^{\dagger(x)}, \\
\langle \mathbf{m} | \hat{\mathbf{A}}^\dagger | \mathbf{m}' \rangle &= \langle \mathbf{m}' | \hat{\mathbf{A}} | \mathbf{m} \rangle^* = \langle \mathbf{m} | \hat{\mathbf{A}} | \mathbf{m}' \rangle^\dagger = \mathbf{A}(\mathbf{m}, \mathbf{m}')^\dagger = [\mathbf{A}(\mathbf{m}, \mathbf{m}')^*]^T \\
&= \begin{bmatrix} \mathbf{A}(\boldsymbol{\alpha}, \boldsymbol{\alpha}')^\dagger = \langle \boldsymbol{\alpha}' | \hat{\mathbf{A}} | \boldsymbol{\alpha} \rangle^* & \mathbf{A}(\boldsymbol{\alpha}, \mathbf{y}')^\dagger = \langle \mathbf{y}' | \hat{\mathbf{A}} | \boldsymbol{\alpha} \rangle^* \\ \mathbf{A}(\mathbf{y}, \boldsymbol{\alpha}')^\dagger = \langle \boldsymbol{\alpha}' | \hat{\mathbf{A}} | \mathbf{y} \rangle^* & \mathbf{A}(\mathbf{y}, \mathbf{y}')^\dagger = \langle \mathbf{y}' | \hat{\mathbf{A}} | \mathbf{y} \rangle^* \end{bmatrix} \equiv \mathbf{A}^{\dagger(m)}.
\end{aligned} \tag{2.51}$$

Hence, the Hermitian (*self*-adjoint) operator $\hat{\mathbf{A}}$ of the physical observable A , for which $\hat{\mathbf{A}}^\dagger = \hat{\mathbf{A}}$, is represented by the Hermitian matrix/kernel: $\mathbf{A}^{\dagger(b)} = \mathbf{A}^{(b)}$, $\mathbf{b} = \mathbf{i}, \mathbf{x}, \mathbf{m}$.

The relations between vectors of (2.31) and (2.42) are thus transformed into the corresponding equations in terms of the basis set components. For example, (2.31) then reads:

$$\begin{aligned}
\hat{\mathbf{A}}|\Psi\rangle &= |\Psi'\rangle \leftrightarrow \mathbf{A}^{(b)} \boldsymbol{\Psi}^{(b)} = \boldsymbol{\Psi}'^{(b)}, \quad \mathbf{b} = \mathbf{i}, \mathbf{x}, \mathbf{m}, \quad \text{i.e.,} \\
\sum_{i'} A_{i, i'} \Psi_{i'} &= \Psi_i, \quad \int A(x, x') \Psi(x') dx' = \Psi'(x), \\
\sum_{\alpha'} A_{\alpha, \alpha'} \Psi_{\alpha'} + \int A(\alpha, y') \Psi(y') dy' &= \Psi_{\alpha}' \quad \text{and} \\
\sum_{\alpha'} A(\mathbf{y}, \alpha') \Psi_{\alpha'} + \int A(\mathbf{y}, y') \Psi(y') dy' &= \Psi'(\mathbf{y}).
\end{aligned} \tag{2.52}$$

The corresponding basis set transcriptions of (2.42) similarly give:

$$\begin{aligned}
\langle \Psi' | &= \langle \Psi | \hat{\mathbf{A}}^\dagger \Leftrightarrow (\hat{\mathbf{A}} |\Psi\rangle = |\Psi'\rangle)^\dagger \Leftrightarrow \boldsymbol{\Psi}^{(b)\dagger} \mathbf{A}^{\dagger(b)} = \boldsymbol{\Psi}'^{\dagger(b)}, \quad \mathbf{b} = \mathbf{i}, \mathbf{x}, \mathbf{m}, \quad \text{i.e.,} \\
\sum_{i'} \Psi_{i'}^* A_{i', i}^* &= \Psi_i'^*, \quad \int \Psi(x')^* A(x', x)^* dx' = \Psi'(x)^*, \\
\sum_{\alpha'} \Psi_{\alpha'}^* (A_{\alpha', \alpha})^* + \int \Psi(y')^* A(y', \alpha)^* dy' &= (\Psi_{\alpha}')^* \quad \text{and} \\
\sum_{\alpha'} \Psi_{\alpha'}^* A(\alpha', y)^* + \int \Psi(y')^* A(y', y)^* dy' &= \Psi'(y)^*.
\end{aligned} \tag{2.53}$$

It should be emphasized that the basis set representations of the state vector are fully equivalent to the state specification by the vector itself. For example (see Sect. 2.6), when the continuous basis set is labeled by the position of a particle in space, $x = \mathbf{r}$, or its momentum, $x = \mathbf{p}$, the associated representations $\Psi^{(r)} \equiv \Psi(\mathbf{r}) = \langle \mathbf{r} | \Psi \rangle$ and $\Psi^{(p)} \equiv \Psi(\mathbf{p}) = \langle \mathbf{p} | \Psi \rangle$, called the wave functions in the position (\mathbf{r}) and momentum (\mathbf{p}) representations, respectively, provide alternative specifications of the quantum state of the particle, which uniquely establish the direction of the ket $|\Psi\rangle$ in the system Hilbert space.

2.5 Eigenvalue Problem of Linear Hermitian Operators

For the linear operator to represent the physically observable quantity in quantum mechanics it has to be *self-adjoint*, i.e., its hermitian conjugate (adjoint) must be identical with the operator itself: $\hat{A}^\dagger = \hat{A}$. Only such *Hermitian* operators can be associated with the measurable quantities of physics. They satisfy the following scalar product identity [see (2.43)]:

$$\langle \Phi | \hat{A} | \Psi \rangle = \langle \Psi | \hat{A} | \Phi \rangle^* = \langle \Phi | \hat{A} | \Psi \rangle^\dagger. \quad (2.54)$$

The projector $\hat{P}_\Psi = |\Psi\rangle\langle\Psi|$ provides an example of the Hermitian operator: $\hat{P}_\Psi^\dagger = \hat{P}_\Psi$. One also observes that the change of order of the adjoint factors in the Hermitian conjugate of the product of two operators implies that the product of the *commuting* Hermitian operators also represents the Hermitian operator. Indeed, when $[\hat{A}, \hat{B}] = 0$, $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger = \hat{B}\hat{A} = \hat{A}\hat{B}$.

In quantum mechanics the *eigenvalue problem* of the linear Hermitian operator \hat{A} corresponding to the physical quantity A is of paramount importance in determining the outcomes of its measurement. It is defined by the following equation:

$$\hat{A}|\Psi_i\rangle = a_i|\Psi_i\rangle \quad \text{or} \quad \langle\Psi_i|\hat{A}^\dagger = \langle\Psi_i|a_i^* = \langle\Psi_i|\hat{A} = \langle\Psi_i|a_i, \quad (2.55a)$$

where a_i denotes *ith eigenvalue* (a number) and $|\Psi_i\rangle \equiv |a_i\rangle$ and $\langle\Psi_i| \equiv \langle a_i|$ stand for the associated *eigen-ket(bra)*, i.e., the eigenvector belonging to a_i . Therefore, the action of \hat{A} on its eigenvector does not affect the direction of the latter, with only its length being multiplied by the corresponding eigenvalue.

A trivial example is the multiplication by a number a . This operator has just one eigenvalue, this number itself: any ket is an eigenket and any bra is an eigenbra corresponding to this eigenvalue. One observes that this number has to be real for such a *number operator* to be self-adjoint [see (2.55a)].

In quantum theory the Hermitian operator \hat{A} , the eigenvectors of which form a basis in the state space, is called an *observable*. The projections onto all such eigenstates amounts to the identity operations of (2.41a)–(2.41c). The projector \hat{P}_Ψ is an example of the quantum-mechanical observable, which exhibits only two

eigenvectors. Indeed, for an arbitrary ket $|\Phi\rangle$ the two functions $|1\rangle \equiv \hat{P}_\Psi|\Phi\rangle$ and $|0\rangle \equiv (1 - \hat{P}_\Psi)|\Phi\rangle$ can be shown to satisfy the eigenvalue problem of \hat{P}_Ψ :

$$\hat{P}_\Psi|1\rangle = \hat{P}_\Psi^2|\Phi\rangle = \hat{P}_\Psi|\Phi\rangle = |1\rangle, \quad \hat{P}_\Psi|0\rangle = (\hat{P}_\Psi - \hat{P}_\Psi^2)|\Phi\rangle = 0|0\rangle, \quad (2.56)$$

where we have used the idempotency property of projection operators [(2.40)]. Therefore, the two state vectors $\{|1\rangle, |0\rangle\}$ are the eigenvectors of \hat{P}_Ψ corresponding to eigenvalues $\{1, 0\}$. Since every ket in the state space can be expanded in these two eigenstates, $|\Phi\rangle = |1\rangle + |0\rangle$, they form the basis in the state space, $|1\rangle\langle 1| + |0\rangle\langle 0| = 1$, thus confirming that \hat{P}_Ψ is an observable.

The eigenbra problem is similarly defined by the Hermitian conjugate of (2.55a):

$$\langle\Psi_i|\hat{A} = \langle\Psi_i|a_i^*. \quad (2.55b)$$

It then follows from the Hermitian character of \hat{A} that all its eigenvalues are real numbers. It suffices to multiply (2.55a) by $\langle\Psi_i|$ (from the left) and (2.55b) by $|\Psi_i\rangle$ (from the right), subtract the resulting equations and use (2.54) (for $\Phi = \Psi = \Psi_i$) to obtain the identity:

$$0 = (a_i - a_i^*) \langle\Psi_i|\Psi_i\rangle \Rightarrow a_i = a_i^*. \quad (2.57)$$

The eigenvalues can be degenerate, when several independent eigenvectors $\{|\Psi_{i,j}\rangle\} = \{|\Psi_{i,1}\rangle, |\Psi_{i,2}\rangle, \dots, |\Psi_{i,g}\rangle\} \equiv \{|i_j\rangle, j = 1, 2, \dots, g\}$ belong to the same eigenvalue a_i :

$$\hat{A}|i_1\rangle = a_i|i_1\rangle, \hat{A}|i_2\rangle = a_i|i_2\rangle, \dots, \hat{A}|i_g\rangle = a_i|i_g\rangle; \quad (2.58)$$

here the number g of such linearly independent (mutually orthogonal) components determines the *multiplicity* of such degenerate eigenvalue. It then directly follows from the linear character of \hat{A} that any combination of such states, say $|\Psi\rangle = C_1|i_1\rangle + C_2|i_2\rangle + \dots + C_g|i_g\rangle$, also represents the eigenvector of \hat{A} belonging to this eigenvalue:

$$\hat{A}|\Psi\rangle = C_1\hat{A}|i_1\rangle + C_2\hat{A}|i_2\rangle + \dots + C_g\hat{A}|i_g\rangle = a_i|\Psi\rangle. \quad (2.59)$$

The Hermitian character of the linear operator also implies that eigenvectors $|\Psi_i\rangle \equiv |i\rangle$ and $|\Psi_j\rangle \equiv |j\rangle$, which belong to different eigenvalues $a_i \neq a_j$, respectively, are automatically orthogonal. Indeed, the associated eigenvalue equations, $\hat{A}|i\rangle = a_i|i\rangle$ and $\langle j|\hat{A} = \langle j|a_j$, give by an analogous manipulation involving a multiplication of the former by $\langle j|$, of the latter by $|i\rangle$, and a subtraction of the resulting equations,

$$0 = (a_i - a_j) \langle j|i\rangle \Rightarrow \langle j|i\rangle = 0. \quad (2.60)$$

In the degenerate case, each vector belonging to the subspace $\{|i_k\rangle\}$ of eigenvalue a_i is thus orthogonal to every vector belonging to the subspace $\{|j_l\rangle\}$ of

eigenvalue a_j : $\langle i_k | j_l \rangle = 0$. Inside each degenerate subspace the vectors can always be constructed as orthonormal, $\langle i_k | i_l \rangle = \delta_{k,l}$, by choosing appropriate combinations of the initial independent (normalized but nonorthogonal) state vectors.

For the given representation in the Hilbert space, say, specified by the discrete orthonormal basis $|i\rangle$, the eigenvalue equation (2.55a) assumes the form of the separate systems of algebraic equations for each eigenvalue, which can be summarized in the joint matrix form [see (2.52)]:

$$\mathbf{A}^{(i)} \boldsymbol{\Psi}^{(i)} = \mathbf{a} \boldsymbol{\Psi}^{(i)}, \quad (2.61)$$

with the operator represented by the square matrix $\mathbf{A}^{(i)} = \{\langle i | \hat{A} | i' \rangle\}$, the diagonal matrix $\mathbf{a} = \{a_s \delta_{s,s'} = \langle \Psi_s | \hat{A} | \Psi_s \rangle \delta_{s,s'}\}$ grouping the eigenvalues $\{a_s\}$ corresponding to eigenvectors $|\boldsymbol{\Psi}\rangle = \{|\Psi_s\rangle \equiv |s\rangle\}$ determined by the corresponding columns $\boldsymbol{\Psi}_s^{(i)} = \langle i | s \rangle$ of the rectangular matrix $\boldsymbol{\Psi}^{(i)} = \{\boldsymbol{\Psi}_s^{(i)}\} = \langle i | \boldsymbol{\Psi} \rangle = \{\langle i | s \rangle\}$ grouping the relevant expansion coefficients (projections).

Moreover, since both $|\boldsymbol{\Psi}\rangle$ and $|i\rangle$ form bases in the Hilbert space, the overall projection $|\boldsymbol{\Psi}\rangle \langle \boldsymbol{\Psi}| = |i\rangle \langle i| = 1$ and hence

$$\begin{aligned} \boldsymbol{\Psi}^{(i)} \boldsymbol{\Psi}^{(i)\dagger} &= \langle i | \boldsymbol{\Psi} \rangle \langle \boldsymbol{\Psi} | i \rangle = \langle i | i \rangle = \{\delta_{i,i'}\} = \mathbf{I}^{(i)} \quad \text{and} \\ \boldsymbol{\Psi}^{(i)\dagger} \boldsymbol{\Psi}^{(i)} &= \langle \boldsymbol{\Psi} | i \rangle \langle i | \boldsymbol{\Psi} \rangle = \langle \boldsymbol{\Psi} | \boldsymbol{\Psi} \rangle = \{\delta_{s,s'}\} = \mathbf{I}^{(\boldsymbol{\Psi})}. \end{aligned} \quad (2.62)$$

Thus, the basis set components of eigenvectors, $\boldsymbol{\Psi}^{(i)}$, define the unitary matrix: $(\boldsymbol{\Psi}^{(i)})^\dagger = (\boldsymbol{\Psi}^{(i)})^{-1}$. Hence, the multiplication, from the right, of both sides of (2.61) by $\boldsymbol{\Psi}^{(i)\dagger}$ allows one to rewrite this matrix equation as the unitary (similarity) transformation (“rotation”), which diagonalizes the Hermitian matrix $\mathbf{A}^{(i)}$, the basis set representation of the Hermitian operator \hat{A} , to its eigenvector representation $\mathbf{a} = \langle \boldsymbol{\Psi} | \hat{A} | \boldsymbol{\Psi} \rangle \equiv \mathbf{A}^{(\boldsymbol{\Psi})}$:

$$\boldsymbol{\Psi}^{(i)\dagger} \mathbf{A}^{(i)} \boldsymbol{\Psi}^{(i)} = (\boldsymbol{\Psi}^{(i)})^{-1} \mathbf{A}^{(i)} \boldsymbol{\Psi}^{(i)} = \mathbf{a}. \quad (2.63)$$

This is the standard numerical procedure, which is routinely applied in the computer programs for the finite basis set determination of eigenvalues of Hermitian matrices.

When dealing with problems of the simultaneous measurements of physical quantities in quantum mechanics, one encounters the *common* eigenvalue problem of several mutually commuting observables. It can be straightforwardly demonstrated that the commutation of operators \hat{A} and \hat{B} , $[\hat{A}, \hat{B}] = 0$, implies the existence of their common eigenvectors, which form the basis in the space of state vectors. In other words, for the case of the discrete spectrum of eigenvalues $\{a_i\}$ and $\{b_j\}$ of these two operators, there exist the common eigenvectors $\{|a_i, b_j\rangle\}$ of \hat{A} and \hat{B} , which satisfy the simultaneous eigenvalue problems of these two operators:

$$\hat{A}|a_i, b_j\rangle = a_i|a_i, b_j\rangle \quad \text{and} \quad \hat{B}|a_i, b_j\rangle = b_j|a_i, b_j\rangle. \quad (2.64)$$

Indeed, when $|a_i\rangle$ denotes the eigenvector of \hat{A} , $\hat{A}|a_i\rangle = a_i|a_i\rangle$, and $[\hat{A}, \hat{B}] = 0$, applying \hat{B} to both sides of this eigenvalue equation gives: $\hat{B}\hat{A}|a_i\rangle = \hat{A}(\hat{B}|a_i\rangle) = a_i(\hat{B}|a_i\rangle)$. Therefore, $\hat{B}|a_i\rangle$ is also the eigenvector of \hat{A} belonging to the same eigenvalue a_i . Hence, for the nondegenerate eigenvalue a_i , $\hat{B}|a_i\rangle$ must be collinear with $|a_i\rangle$, since there is only one independent eigenstate corresponding to a_i , identified by the direction of $|a_i\rangle$. Hence, $\hat{B}|a_i\rangle$ is then proportional to $|a_i\rangle$, thus also satisfying the eigenvalue equation of \hat{B} ,

$$\hat{B}|a_i\rangle = b_i|a_i\rangle \Rightarrow |a_i\rangle = |a_i, b_i\rangle. \quad (2.65)$$

For the degenerate eigenvalue a_i , $\hat{B}|a_i\rangle$ gives a vector belonging to the subspace $\{|a_i\rangle_k\}$ of a_i , so that such eigenvalue subspace of \hat{A} remains globally invariant under the action of \hat{B} . One also observes that for such a pair of commuting operators, the two eigenvectors for different eigenvalues of one operator, say $|a_i\rangle$ and $|a_j\rangle$ of \hat{A} , $a_i \neq a_j$, give the vanishing matrix element of the other operator: $\langle a_i|\hat{B}|a_j\rangle = 0$. This directly follows from their vanishing commutator which implies

$$\langle a_i|[\hat{A}, \hat{B}]|a_j\rangle = (a_i - a_j)\langle a_i|\hat{B}|a_j\rangle = 0 \Rightarrow \langle a_i|\hat{B}|a_j\rangle = 0, \quad (2.66)$$

where we have recognized the Hermitian character of \hat{A} .

In fact the commutation of two operators constitutes both the necessary and sufficient condition for the two operators to have the common eigenvectors. The above demonstration of the sufficient criterion can be supplemented by the inverse theorem of the necessary condition that the existence of the common eigenvalue problem of the two operators implies that they commute. Since the common eigenvectors $\{|a_i, b_j\rangle\}$ constitute the basis (complete) set one can expand any ket

$$|\Psi\rangle = \sum_{i,j} |a_i, b_j\rangle \langle a_i, b_j|\Psi\rangle \equiv \sum_{i,j} |a_i, b_j\rangle C_{i,j}. \quad (2.67)$$

Therefore:

$$\begin{aligned} [\hat{A}, \hat{B}]|\Psi\rangle &= \sum_{i,j} C_{i,j} [\hat{A}\hat{B} - \hat{B}\hat{A}]|a_i, b_j\rangle = \sum_{i,j} C_{i,j} (a_i b_j - b_j a_i) |a_i, b_j\rangle = 0 \\ \Rightarrow [\hat{A}, \hat{B}] &= 0. \end{aligned} \quad (2.68)$$

The minimum set of the mutually commuting observables $\{\hat{A}, \hat{B}, \dots, \hat{C}\}$, which uniquely specify the direction of the state vector $|\Psi\rangle$, is called the complete set of commuting observables. Hence, there exists a unique orthonormal basis of their common eigenvectors and the corresponding eigenvalues (a_i, b_j, \dots, c_k) provide the complete specification of the state under consideration: $|\Psi\rangle = |a_i, b_j, \dots, c_k\rangle$. One should realize, however, that for a given molecular system there exist several such sets of observables. We shall encounter their examples in the next section.

2.6 Position and Momentum Representations

Two important cases of the continuous basis sets in the vector space of quantum states of a single (spinless) particle combine all state vectors corresponding to its sharply specified position $\mathbf{r} = (x, y, z)$ or momentum $\mathbf{p} = (p_x, p_y, p_z)$. These states, $\{|\mathbf{r}\rangle\}$ and $\{|\mathbf{p}\rangle\}$, respectively, labeled by the respective three continuous coordinates are the eigenvectors of the particle position and momentum operators, $\hat{\mathbf{r}} = (\hat{x}, \hat{y}, \hat{z})$ and $\hat{\mathbf{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$,

$$\begin{aligned}\hat{\mathbf{r}}|\mathbf{r}'\rangle &= \mathbf{r}'|\mathbf{r}'\rangle, & \langle \mathbf{r}'|\mathbf{r}'\rangle &= \delta(\mathbf{r}' - \mathbf{r}) = u_{\mathbf{r}'}(\mathbf{r}), & \int |\mathbf{r}\rangle \langle \mathbf{r}| d\mathbf{r} &= 1; \\ \hat{\mathbf{p}}|\mathbf{p}'\rangle &= \mathbf{p}'|\mathbf{p}'\rangle, & \langle \mathbf{p}'|\mathbf{p}'\rangle &= \delta(\mathbf{p}' - \mathbf{p}) = u_{\mathbf{p}'}(\mathbf{p}), & \int |\mathbf{p}\rangle \langle \mathbf{p}| d\mathbf{p} &= 1.\end{aligned}\quad (2.69)$$

The Dirac deltas $\{\delta(\mathbf{r}' - \mathbf{r})\}$ and $\{\delta(\mathbf{p}' - \mathbf{p})\}$ in these equations define the continuous basis functions $\{u_{\mathbf{r}'}(\mathbf{r})\}$ and $\{u_{\mathbf{p}'}(\mathbf{p})\}$ for expanding the particle wave functions $\Psi(\mathbf{r}') = \langle \mathbf{r}'|\Psi\rangle$ and $\Psi(\mathbf{p}') = \langle \mathbf{p}'|\Psi\rangle$ in these two bases:

$$\begin{aligned}\Psi(\mathbf{r}') &= \int \langle \mathbf{r}'|\mathbf{r}\rangle \langle \mathbf{r}|\Psi\rangle d\mathbf{r} = \int u_{\mathbf{r}'}^*(\mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r}, \\ \Psi(\mathbf{p}') &= \int \langle \mathbf{p}'|\mathbf{p}\rangle \langle \mathbf{p}|\Psi\rangle d\mathbf{p} = \int u_{\mathbf{p}'}^*(\mathbf{p}) \Psi(\mathbf{p}) d\mathbf{p}.\end{aligned}\quad (2.70)$$

Indeed, these two equations express the basic integral property of Dirac's delta function [(2.21)]:

$$\Psi(\mathbf{r}') = \int \delta(\mathbf{r} - \mathbf{r}') \Psi(\mathbf{r}) d\mathbf{r} \quad \text{and} \quad \Psi(\mathbf{p}') = \int \delta(\mathbf{p} - \mathbf{p}') \Psi(\mathbf{p}) d\mathbf{p}.$$

They also identify the function “coordinates” as the corresponding projections in the function space spanned by the bases $\{u_{\mathbf{r}'}(\mathbf{r})\}$ and $\{u_{\mathbf{p}'}(\mathbf{p})\}$, respectively.

The orthogonality relation between quantum states $|\Psi\rangle$ and $|\Phi\rangle$ can thus be expressed as the isomorphic relations between the corresponding wave functions:

$$\begin{aligned}\langle \Psi|\Phi\rangle &= \int \langle \Psi|\mathbf{r}\rangle \langle \mathbf{r}|\Phi\rangle d\mathbf{r} = \int \Psi^*(\mathbf{r}) \Phi(\mathbf{r}) d\mathbf{r} \\ &= \int \langle \Psi|\mathbf{p}\rangle \langle \mathbf{p}|\Phi\rangle d\mathbf{p} = \int \Psi^*(\mathbf{p}) \Phi(\mathbf{p}) d\mathbf{p} = 0.\end{aligned}\quad (2.71)$$

It also follows from (2.69) that the basis functions $u_{\mathbf{r}'}(\mathbf{r})$ and $u_{\mathbf{p}'}(\mathbf{p})$ are themselves wave functions of quantum states with the sharply defined position $\mathbf{r} = \mathbf{r}'$ and momentum $\mathbf{p} = \mathbf{p}'$, respectively. There is one-to-one correspondence between wave functions and the associated state vectors they represent, e.g.,

$$u_{\mathbf{r}'}(\mathbf{r}) \Leftrightarrow |\mathbf{r}'\rangle, \quad u_{\mathbf{p}'}(\mathbf{p}) \Leftrightarrow |\mathbf{p}'\rangle, \quad \Psi(\mathbf{r}) \Leftrightarrow |\Psi\rangle, \quad \Psi(\mathbf{p}) \Leftrightarrow |\Psi\rangle. \quad (2.72)$$

Of interest also are the relations between the wave functions in the momentum and position representations. They are summarized by the Fourier transformations of (2.26), which in three dimensions read in terms of the wave vector $\mathbf{k} = \mathbf{p}/\hbar$,

$$\begin{aligned}\Psi(\mathbf{k}) &= (2\pi)^{-3/2} \int \exp(-i\mathbf{k} \cdot \mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r} \text{ or } \Psi(\mathbf{p}) = (2\pi\hbar)^{-3/2} \int \exp(-\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r}, \\ \Psi(\mathbf{r}) &= (2\pi)^{-3/2} \int \exp(i\mathbf{k} \cdot \mathbf{r}) \Psi(\mathbf{k}) d\mathbf{k} = (2\pi\hbar)^{-3/2} \int \exp(\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}) \Psi(\mathbf{p}) d\mathbf{p}.\end{aligned}\tag{2.73}$$

Substituting one transform into the other then generates the following analytical representations of the Dirac deltas [see (2.28)]:

$$\begin{aligned}\delta(\mathbf{r}' - \mathbf{r}) &= (2\pi\hbar)^{-3} \int \exp\left[\frac{i}{\hbar}\mathbf{p} \cdot (\mathbf{r}' - \mathbf{r})\right] d\mathbf{p}, \\ \delta(\mathbf{p}' - \mathbf{p}) &= (2\pi\hbar)^{-3} \int \exp\left[\frac{i}{\hbar}(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{r}\right] d\mathbf{r}.\end{aligned}\tag{2.74}$$

Hence, by transcribing (2.73) in terms of corresponding state vectors,

$$\begin{aligned}\Psi(\mathbf{p}) &= \langle \mathbf{p} | \Psi \rangle = \int \langle \mathbf{p} | \mathbf{r} \rangle \langle \mathbf{r} | \Psi \rangle d\mathbf{r} = \int u_{\mathbf{p}}^*(\mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r}, \\ \Psi(\mathbf{r}) &= \langle \mathbf{r} | \Psi \rangle = \int \langle \mathbf{r} | \mathbf{p} \rangle \langle \mathbf{p} | \Psi \rangle d\mathbf{p} = \int u_{\mathbf{r}}^*(\mathbf{p}) \Psi(\mathbf{p}) d\mathbf{p},\end{aligned}\tag{2.75}$$

one identifies the following representation of basis vectors of one representation in terms of vectors comprised in the other basis set:

$$\begin{aligned}u_{\mathbf{p}}(\mathbf{r}) &= \langle \mathbf{r} | \mathbf{p} \rangle = (2\pi\hbar)^{-3/2} \exp\left(\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}\right) \text{ and} \\ u_{\mathbf{r}}(\mathbf{p}) &= \langle \mathbf{p} | \mathbf{r} \rangle = u_{\mathbf{p}}(\mathbf{r})^* = (2\pi\hbar)^{-3/2} \exp\left(-\frac{i}{\hbar}\mathbf{p} \cdot \mathbf{r}\right).\end{aligned}\tag{2.76}$$

Let us now examine the associated representations of the position and momentum operators in these two continuous basis sets. We first observe that these operators are the continuous diagonal when represented in the basis set of their own eigenvectors [see (2.69)]:

$$\langle \mathbf{r}'' | \hat{\mathbf{r}} | \mathbf{r}' \rangle = \mathbf{r}' \langle \mathbf{r}'' | \mathbf{r}' \rangle = \mathbf{r}' \delta(\mathbf{r}' - \mathbf{r}''), \quad \langle \mathbf{p}'' | \hat{\mathbf{p}} | \mathbf{p}' \rangle = \mathbf{p}' \langle \mathbf{p}'' | \mathbf{p}' \rangle = \mathbf{p}' \delta(\mathbf{p}' - \mathbf{p}''). \tag{2.77}$$

Therefore, the action of the position operator on the wave function in the position representation amounts to a straightforward multiplication by the position vector:

$$\int \langle \mathbf{r}'' | \hat{\mathbf{r}} | \mathbf{r}' \rangle \langle \mathbf{r}' | \Psi \rangle d\mathbf{r}' = \int \mathbf{r}' \delta(\mathbf{r}' - \mathbf{r}'') \Psi(\mathbf{r}') d\mathbf{r}' = \mathbf{r}'' \Psi(\mathbf{r}''). \quad (2.78)$$

The action of the momentum operator on the wave function in the momentum representation similarly represents the multiplication by the momentum vector:

$$\int \langle \mathbf{p}'' | \hat{\mathbf{p}} | \mathbf{p}' \rangle \langle \mathbf{p}' | \Psi \rangle d\mathbf{p}' = \int \mathbf{p}' \delta(\mathbf{p}' - \mathbf{p}'') \Psi(\mathbf{p}') d\mathbf{p}' = \mathbf{p}'' \Psi(\mathbf{p}''). \quad (2.79)$$

Next, let us establish the form of the *momentum* operator in the *position* representation. It can be recognized by examining the position representation of the ket $\hat{\mathbf{p}}|\Psi\rangle$,

$$\langle \mathbf{r} | \hat{\mathbf{p}} | \Psi \rangle = \int \langle \mathbf{r} | \mathbf{p} \rangle \langle \mathbf{p} | \hat{\mathbf{p}} | \Psi \rangle d\mathbf{p} = (2\pi\hbar)^{-3/2} \int \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) \mathbf{p} \Psi(\mathbf{p}) d\mathbf{p}. \quad (2.80)$$

Hence, by comparing the previous equation with the last (2.73) gives:

$$\langle \mathbf{r} | \hat{\mathbf{p}} | \Psi \rangle = -i\hbar \nabla_{\mathbf{r}} \langle \mathbf{r} | \Psi \rangle \equiv \hat{\mathbf{p}}(\mathbf{r}) \Psi(\mathbf{r}), \quad (2.81)$$

where the differential vector operator $\nabla_{\mathbf{r}} = i\partial/\partial x + j\partial/\partial y + k\partial/\partial z \equiv \partial/\partial \mathbf{r}$ stands for the *position* gradient. Therefore, the action of the momentum operator in the position representation amounts to performing the differential operation $\hat{\mathbf{p}}(\mathbf{r}) = -i\hbar \nabla_{\mathbf{r}}$ on the wave function $\Psi(\mathbf{r})$. Hence, the matrix element $\langle \Phi | \hat{\mathbf{p}} | \Psi \rangle$ in this representation is determined by the associated integral in terms of the position wave functions:

$$\langle \Phi | \hat{\mathbf{p}} | \Psi \rangle = \int \langle \Phi | \mathbf{r} \rangle \langle \mathbf{r} | \hat{\mathbf{p}} | \Psi \rangle d\mathbf{r} = -i\hbar \int \Phi^*(\mathbf{r}) \nabla_{\mathbf{r}} \Psi(\mathbf{r}) d\mathbf{r}. \quad (2.82)$$

One could alternatively calculate the kernel $\hat{\mathbf{p}}(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | \hat{\mathbf{p}} | \mathbf{r}' \rangle$ (the continuous matrix element) of the momentum operator, in terms of which the operation of (2.81) reads:

$$\langle \mathbf{r} | \hat{\mathbf{p}} | \Psi \rangle = \int \langle \mathbf{r} | \hat{\mathbf{p}} | \mathbf{r}' \rangle \langle \mathbf{r}' | \Psi \rangle d\mathbf{r}' = \int \hat{\mathbf{p}}(\mathbf{r}, \mathbf{r}') \Psi(\mathbf{r}') d\mathbf{r}'. \quad (2.83)$$

By twice inserting the closure relation into this matrix element, and using the analytical expression for the Dirac delta (2.74) one then finds:

$$\begin{aligned} \langle \mathbf{r} | \hat{\mathbf{p}} | \mathbf{r}' \rangle &= \iint \langle \mathbf{r} | \mathbf{p} \rangle \langle \mathbf{p} | \hat{\mathbf{p}} | \mathbf{p}' \rangle \langle \mathbf{p}' | \mathbf{r}' \rangle d\mathbf{p} d\mathbf{p}' \\ &= \iint u_{\mathbf{r}}^*(\mathbf{p}) \mathbf{p} \delta(\mathbf{p}' - \mathbf{p}) u_{\mathbf{r}'}(\mathbf{p}') d\mathbf{p}' d\mathbf{p} \\ &= (2\pi\hbar)^{-3} \int \exp\left[\frac{i}{\hbar} \mathbf{p} \cdot (\mathbf{r}' - \mathbf{r})\right] \mathbf{p} d\mathbf{p} = i\hbar \nabla_{\mathbf{r}'} \delta(\mathbf{r}' - \mathbf{r}). \end{aligned} \quad (2.84)$$

Substituting this result to (2.83), after integration by parts [see (2.30)], gives the same result as in (2.82):

$$\begin{aligned} \int \hat{\mathbf{p}}(\mathbf{r}, \mathbf{r}') \Psi(\mathbf{r}') d\mathbf{r}' &= i\hbar \int \nabla_{\mathbf{r}'} \delta(\mathbf{r}' - \mathbf{r}) \Psi(\mathbf{r}') d\mathbf{r}' \\ &= -i\hbar \int \delta(\mathbf{r}' - \mathbf{r}) \nabla_{\mathbf{r}'} \Psi(\mathbf{r}') d\mathbf{r}' = -i\hbar \nabla_{\mathbf{r}} \Psi(\mathbf{r}). \end{aligned} \quad (2.85)$$

One similarly derives the remaining kernel providing the momentum representation of the position operator,

$$\begin{aligned} \hat{\mathbf{r}}(\mathbf{p}, \mathbf{p}') &= \langle \mathbf{p} | \hat{\mathbf{r}} | \mathbf{p}' \rangle = \iint \langle \mathbf{p} | \mathbf{r} \rangle \langle \mathbf{r} | \hat{\mathbf{r}} | \mathbf{r}' \rangle \langle \mathbf{r}' | \mathbf{p}' \rangle d\mathbf{r} d\mathbf{r}' = \iint u_{\mathbf{p}}^*(\mathbf{r}) \mathbf{r} \delta(\mathbf{r}' - \mathbf{r}) u_{\mathbf{p}'}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ &= (2\pi\hbar)^{-3} \int \exp\left[\frac{i}{\hbar}(\mathbf{p}' - \mathbf{p}) \cdot \mathbf{r}\right] \mathbf{r} d\mathbf{r} = -i\hbar \nabla_{\mathbf{p}'} \delta(\mathbf{p}' - \mathbf{p}), \end{aligned} \quad (2.86)$$

where the *momentum* gradient $\nabla_{\mathbf{p}} = \mathbf{i}\partial/\partial p_x + \mathbf{j}\partial/\partial p_y + \mathbf{k}\partial/\partial p_z \equiv \partial/\partial \mathbf{p}$. It gives rise to the following momentum representation of the ket $\hat{\mathbf{r}}|\Psi\rangle$:

$$\begin{aligned} \langle \mathbf{p} | \hat{\mathbf{r}} | \Psi \rangle &= \int \langle \mathbf{p} | \hat{\mathbf{r}} | \mathbf{p}' \rangle \langle \mathbf{p}' | \Psi \rangle d\mathbf{p}' = \int \hat{\mathbf{r}}(\mathbf{p}, \mathbf{p}') \Psi(\mathbf{p}') d\mathbf{p}' \\ &= -i\hbar \int \nabla_{\mathbf{p}'} \delta(\mathbf{p}' - \mathbf{p}) \Psi(\mathbf{p}') d\mathbf{p}' = i\hbar \int \delta(\mathbf{p}' - \mathbf{p}) \nabla_{\mathbf{p}'} \Psi(\mathbf{p}') d\mathbf{p}' \\ &= i\hbar \nabla_{\mathbf{p}} \Psi(\mathbf{p}) \equiv \hat{\mathbf{r}}(\mathbf{p}) \Psi(\mathbf{p}). \end{aligned} \quad (2.87)$$

Therefore, the action of the position operator in the momentum space coincides with the differential operation $\hat{\mathbf{r}}(\mathbf{p}) = i\hbar \nabla_{\mathbf{p}}$ performed on the wave function $\Psi(\mathbf{p})$.

The same result directly follows from inserting the closure identity into the initial scalar product of the preceding equation:

$$\langle \mathbf{p} | \hat{\mathbf{r}} | \Psi \rangle = \int \langle \mathbf{p} | \mathbf{r} \rangle \langle \mathbf{r} | \hat{\mathbf{r}} | \Psi \rangle d\mathbf{r} = (2\pi\hbar)^{-3/2} \int \exp(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}) \mathbf{r} \Psi(\mathbf{r}) d\mathbf{r}. \quad (2.88)$$

Hence, by comparing this expression with the second (2.73) again gives:

$$\langle \mathbf{p} | \hat{\mathbf{r}} | \Psi \rangle = i\hbar \nabla_{\mathbf{p}} \langle \mathbf{p} | \Psi \rangle = \hat{\mathbf{r}}(\mathbf{p}) \Psi(\mathbf{p}). \quad (2.89)$$

2.7 Energy Representation and Unitary Transformations

The energy representation of quantum states and operators is defined by the basis set of the (orthonormal) eigenvectors $\{|E_n\rangle\}$ of the system energy operator, the Hamiltonian $\hat{E} \equiv \hat{H}$,

$$\hat{H}|E_n\rangle = E_n|E_n\rangle. \quad (2.90)$$

They represent the stationary states, with the sharply specified energy. Here, for the sake of simplicity we have assumed the discrete spectrum of the allowed energy levels $\{E_n\}$.

In the position representation $\xi = \{|\xi\rangle, |\xi'\rangle, \dots\} = \{|\xi\rangle\}$, $\langle\xi'|\xi\rangle = \delta(\xi - \xi')$, where ξ groups the system coordinates, the eigenkets $\{|E_n\rangle\}$ of the energy basis set are represented by the associated wave functions $\{\varphi_{E_n}(\xi) = \langle\xi|E_n\rangle\} = \langle\xi|E_n\rangle$ of the continuous *column* vector, while the corresponding eigenbras define the associated continuous *row* vector: $\{\varphi_{E_n}^*(\xi) = \langle\xi|E_n\rangle^* = \langle E_n|\xi\rangle\} = \langle E_n|\xi\rangle$. In this position basis the Hamiltonian \hat{H} is similarly represented by the continuous (diagonal) matrix: $\hat{H} \Rightarrow \{\hat{H}(\xi, \xi') = \langle\xi|\hat{H}|\xi'\rangle = \hat{H}(\xi)\delta(\xi' - \xi)\}$. In the position representation the energy eigenvalue problem of (2.90) reads:

$$\int \langle\xi|\hat{H}|\xi'\rangle \langle\xi'|E_n\rangle d\xi' = E_n \langle\xi|E_n\rangle \quad (2.91)$$

or

$$\int \hat{H}(\xi, \xi') \varphi_{E_n}(\xi') d\xi' = \int \hat{H}(\xi') \delta(\xi' - \xi) \varphi_{E_n}(\xi') d\xi' = \hat{H}(\xi) \varphi_{E_n}(\xi) = E_n \varphi_{E_n}(\xi). \quad (2.92)$$

The orthonormality of the energy eigenvectors (discrete spectrum), $\langle E_m|E_n\rangle = \delta_{E_m, E_n}$, can be also expressed in terms of the associated wave functions:

$$\langle E_m|E_n\rangle = \int \langle E_m|\xi\rangle \langle\xi|E_n\rangle d\xi = \int \varphi_{E_m}^*(\xi) \varphi_{E_n}(\xi) d\xi = \delta_{E_m, E_n}. \quad (2.93)$$

Any state vector $|\Psi\rangle$ is thus equivalently represented either by the components $\{\Psi_{E_n} = \langle E_n|\Psi\rangle = \int \varphi_{E_n}^*(\xi) \Psi(\xi) d\xi\} \equiv \mathbf{\Psi}^{(E)}$ in the energy representation or by the wave function $\Psi(\xi) = \langle\xi|\Psi\rangle \equiv \mathbf{\Psi}^{(\xi)}$ in the position representation. They are related *via* the following transformations:

$$\begin{aligned} \Psi(\xi) &= \langle\xi|\Psi\rangle = \sum_m \langle\xi|E_m\rangle \langle E_m|\Psi\rangle = \sum_m \varphi_{E_m}(\xi) \Psi_{E_m} \text{ or } \mathbf{\Psi}^{(\xi)} = \mathbf{T}(\xi, E) \mathbf{\Psi}^{(E)}, \\ \Psi_{E_n} &= \langle E_n|\Psi\rangle = \int \langle E_n|\xi\rangle \langle\xi|\Psi\rangle d\xi = \int \varphi_{E_n}^*(\xi) \Psi(\xi) d\xi \text{ or } \mathbf{\Psi}^{(E)} = \mathbf{T}(E, \xi) \mathbf{\Psi}^{(\xi)}. \end{aligned} \quad (2.94)$$

Thus, the energy eigenfunctions $\{\varphi_{E_m}(\xi)\}$, with the continuous (discrete) position (energy) labels, transform the energy representation of the state vector to its associated position representation. Accordingly, the complex conjugate functions $\{\varphi_{E_m}^*(\xi)\}$ are seen to define the reverse transformation of the state vector, from its position representation to the energy representation.

Therefore, should one regard the coefficients of these mutually reverse transformations as elements of the corresponding transformation matrices identified by the discrete $\{E_n\}$ and continuous $\{\xi\}$ indices,

$$\begin{aligned}\{\varphi_{E_n}(\xi)\} &\equiv \mathbf{T}(\xi, E) = \langle \xi | E \rangle, \\ \{\varphi_{E_n}^*(\xi)\} &\equiv \mathbf{T}(E, \xi) = \langle E | \xi \rangle = \mathbf{T}(\xi, E)^\dagger,\end{aligned}\tag{2.95}$$

their mutual reciprocity relations imply:

$$\begin{aligned}\mathbf{T}(\xi', E) \mathbf{T}(E, \xi) &= \langle \xi' | E \rangle \langle E | \xi \rangle = \langle \xi' | \xi \rangle = \delta(\xi - \xi') \\ &\Rightarrow \mathbf{T}(E, \xi) = \mathbf{T}(\xi, E)^\dagger = \mathbf{T}(\xi, E)^{-1}; \\ \mathbf{T}(E, \xi) \mathbf{T}(\xi, E') &= \langle E | \xi \rangle \langle \xi | E' \rangle = \langle E | E' \rangle = \delta_{E, E'} = \mathbf{I}. \\ &\Rightarrow \mathbf{T}(\xi, E) = \mathbf{T}(E, \xi)^\dagger = \mathbf{T}(E, \xi)^{-1}.\end{aligned}\tag{2.96}$$

One thus concludes that each of these mutually inverse transformation matrices is the Hermitian conjugate of the other thus defining the *unitary* transformations (“rotations”) of one orthonormal basis set into another.

To summarize, the system energy, with discrete (or continuous/mixed) set of eigenvalues, constitutes the independent variable of the energy representation. The square of the modulus of the component $\Psi_{E_n} = \langle E_n | \Psi \rangle$ measures the (conditional) probability $W(E_n | \Psi)$ of observing the system in state $|\Psi\rangle$ at the specified energy:

$$\begin{aligned}W(E_n | \Psi) &= |\langle E_n | \Psi \rangle|^2 = \langle \Psi | E_n \rangle \langle E_n | \Psi \rangle, \\ \sum_n W(E_n | \Psi) &= \sum_n \langle \Psi | E_n \rangle \langle E_n | \Psi \rangle = \langle \Psi | \Psi \rangle = 1.\end{aligned}\tag{2.97}$$

As we have already observed in (2.75) of the preceding section, the wave functions (2.76) define another pair of such mutually reverse transformations:

$$\begin{aligned}u_r(\mathbf{p}) &= \langle \mathbf{p} | \mathbf{r} \rangle \equiv \mathbf{t}(\mathbf{p}, \mathbf{r}) \quad \text{and} \quad u_p(\mathbf{r}) = \langle \mathbf{r} | \mathbf{p} \rangle \equiv \mathbf{t}(\mathbf{r}, \mathbf{p}), \\ \int \mathbf{t}(\mathbf{p}, \mathbf{r}) \mathbf{t}(\mathbf{r}, \mathbf{p}') d\mathbf{r} &= \delta(\mathbf{p} - \mathbf{p}'), \quad \int \mathbf{t}(\mathbf{r}, \mathbf{p}) \mathbf{t}(\mathbf{p}, \mathbf{r}') d\mathbf{p} = \delta(\mathbf{r} - \mathbf{r}').\end{aligned}\tag{2.98}$$

They also define the unitary kernels,

$$\mathbf{t}(\mathbf{p}, \mathbf{r}) = \mathbf{t}(\mathbf{r}, \mathbf{p})^\dagger = \mathbf{t}(\mathbf{r}, \mathbf{p})^{-1} \quad \text{and} \quad \mathbf{t}(\mathbf{r}, \mathbf{p}) = \mathbf{t}(\mathbf{p}, \mathbf{r})^\dagger = \mathbf{t}(\mathbf{p}, \mathbf{r})^{-1},\tag{2.99}$$

of the integral transformations between the position and momentum representations:

$$\begin{aligned}\int \mathbf{t}(\mathbf{p}, \mathbf{r}) \Psi(\mathbf{r}) d\mathbf{r} &\equiv \mathbf{T}(\mathbf{p}, \mathbf{r}) \Psi(\mathbf{r}) = \Psi(\mathbf{p}) \quad \text{or} \quad \widehat{\mathbf{T}}(\mathbf{p}, \mathbf{r}) \Psi(\mathbf{r}) = \Psi(\mathbf{p}), \\ \int \mathbf{t}(\mathbf{r}, \mathbf{p}) \Psi(\mathbf{p}) d\mathbf{p} &\equiv \mathbf{T}(\mathbf{r}, \mathbf{p}) \Psi(\mathbf{p}) = \Psi(\mathbf{r}) \quad \text{or} \quad \widehat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \Psi(\mathbf{p}) = \Psi(\mathbf{r}).\end{aligned}\tag{2.100}$$

Above, $\mathbf{T}(\mathbf{p}, \mathbf{r})$ represents the integral operator $\hat{\mathbf{T}}(\mathbf{p}, \mathbf{r})$ defined by the kernel $\mathbf{t}(\mathbf{p}, \mathbf{r})$, which replaces the arguments of the wave function: $\mathbf{r} \rightarrow \mathbf{p}$, etc.

It follows from the preceding equations that these transformations are unitary:

$$\mathbf{T}(\mathbf{r}, \mathbf{p}) = \mathbf{T}(\mathbf{p}, \mathbf{r})^{-1} = \mathbf{T}(\mathbf{p}, \mathbf{r})^\dagger \quad \text{or} \quad \hat{\mathbf{T}}^{-1}(\mathbf{p}, \mathbf{r}) = \hat{\mathbf{T}}^\dagger(\mathbf{p}, \mathbf{r}) = \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}), \quad (2.101)$$

where the inverse operator $\hat{\mathbf{T}}^{-1}(\mathbf{p}, \mathbf{r})$ replaces the variables in the wave function it acts upon in the inverse order: $\mathbf{p} \rightarrow \mathbf{r}$. Therefore, the reciprocity relations of (2.98) in fact express the unitary character of the above (integral) transformation operators,

$$\begin{aligned} \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \hat{\mathbf{T}}^\dagger(\mathbf{r}, \mathbf{p}) &= \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \hat{\mathbf{T}}(\mathbf{p}, \mathbf{r}) = 1 \quad \text{and} \\ \hat{\mathbf{T}}(\mathbf{p}, \mathbf{r}) \hat{\mathbf{T}}^\dagger(\mathbf{p}, \mathbf{r}) &= \hat{\mathbf{T}}(\mathbf{p}, \mathbf{r}) \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) = 1, \end{aligned} \quad (2.102)$$

because the double exchange of variables $\mathbf{p} \rightarrow \mathbf{r} \rightarrow \mathbf{p}$ amounts to identity operation on the wave function $\Psi(\mathbf{p})$ and the *double* exchange $\mathbf{r} \rightarrow \mathbf{p} \rightarrow \mathbf{r}$ operation performed of $\Psi(\mathbf{r})$ leaves it unchanged.

Transitions from one set of independent variables to another are called the *canonical* transformations. They have been shown to correspond to unitary operators, which also transform the matrix representations of the quantum-mechanical operators to a new set of variables. Indeed, by unitary transforming both sides of the momentum representation of (2.31),

$$\int \hat{\mathbf{A}}(\mathbf{p}, \mathbf{p}') \Psi(\mathbf{p}') d\mathbf{p}' = \Psi'(\mathbf{p}),$$

and using the identity (2.102) one obtains

$$\begin{aligned} [\hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \hat{\mathbf{A}}(\mathbf{p}, \mathbf{p}') \hat{\mathbf{T}}^\dagger(\mathbf{r}', \mathbf{p}')] [\hat{\mathbf{T}}(\mathbf{r}', \mathbf{p}') \Psi(\mathbf{p}')] &\equiv \hat{\mathbf{A}}(\mathbf{r}, \mathbf{r}') \Psi(\mathbf{r}') \\ &= \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \Psi'(\mathbf{p}) \equiv \Psi'(\mathbf{r}). \end{aligned} \quad (2.103)$$

Hence, the canonically transformed resultant vector $\Psi'(\mathbf{p})$ in the new variables becomes: $\hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \Psi'(\mathbf{p}) = \Psi'(\mathbf{r})$. It results from the transformed vector $\hat{\mathbf{T}}(\mathbf{r}', \mathbf{p}')$ $\Psi(\mathbf{p}') = \Psi(\mathbf{r}')$ by the action of the transformed operator

$$\hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \hat{\mathbf{A}}(\mathbf{p}, \mathbf{p}') \hat{\mathbf{T}}^\dagger(\mathbf{r}', \mathbf{p}') = \hat{\mathbf{T}}(\mathbf{r}, \mathbf{p}) \hat{\mathbf{A}}(\mathbf{p}, \mathbf{p}') \hat{\mathbf{T}}^{-1}(\mathbf{p}', \mathbf{r}') = \hat{\mathbf{A}}(\mathbf{r}, \mathbf{r}') \quad (2.104)$$

with the preceding equation thus expressing a general transformation law for changing representations of linear operators.

Another important type of the unitary operators is represented by the *phase* transformation $\hat{\mathbf{S}}(\mathbf{x}) = \exp[i\hat{\alpha}(\mathbf{x})]$. It involves the linear Hermitian operator $\hat{\alpha}(\mathbf{x})$, the function of the same list of variables as those of the wave function itself.

This transformation of $\Psi(\mathbf{x})$ modifies the wave function without affecting its set of the independent state variables.

All physical predictions of quantum mechanics can be shown to remain unaffected by the unitary transformations of states and operators, since they are related to specific *invariants* of such operations. These invariant properties include the linear and Hermitian character of quantum-mechanical observables, all algebraic relations between them, e.g., the commutation rules, spectrum of eigenvalues and the matrix elements of operators.

The diversity of unitary transformations is not limited to those changing a description of the system quantum-mechanical states at the given time (quantum *kinematics*): $\Psi(\mathbf{x}) = \Psi(\mathbf{x}, t = 0)$. In the next chapter, we shall examine other examples of unitary transformations of wave functions and operators, which generate different *pictures* of the quantum-mechanical *dynamics*, e.g., the evolution of quantum states with time in the *Schrödinger picture*:

$$\Psi(\mathbf{x}, t) = \hat{U}(t)\Psi(\mathbf{x}), \quad \hat{U}^\dagger(t)\hat{U}(t) = 1. \quad (2.105)$$

2.8 Functional Derivatives

The functional of the state vector argument or of its continuous basis representation – the wave function – gives the scalar. The representative example of such a mathematical entity is the definite integral, e.g., the scalar product of two wave functions. It may additionally involve various derivatives of the function argument. For simplicity, let us assume the functional F of a single function $f(x)$ of the continuous variable x ,

$$F[f] = \int \mathcal{L}[x, f(x), f'(x), \dots] dx. \quad (2.106)$$

This functional attributes to the argument function f the scalar $F = F[f]$. It is defined by the functional density $\mathcal{L}[x, f(x), f'(x), \dots]$, which in a general case depends on the current value of x , the argument function itself, $f(x)$, and its derivatives: $f'(x) = df(x)/dx$, etc.

An important problem, which we shall often encounter in this book, is to find the functional variation $\delta F = F[f + \delta f] - F[f]$ due to a small modification of the argument function, $\delta f = \varepsilon h$, where ε is a small parameter and h stands for the displacement function (perturbation). The first differential of the functional is the component of δF that depends on δf linearly:

$$\delta^{(1)}F = \int \frac{\delta F}{\delta f(x)} \delta f(x) dx, \quad (2.107)$$

with the (local) coefficient before $\delta f(x)$ in the integrand defining the *first functional derivative* of F with respect to f at point x . It is seen to transform the local displacements of the argument function into the first differential of the functional. This expression can be viewed as the continuous generalization of the familiar differential of the function of several variables: $d^{(1)}f(x_1, x_2, \dots) = \sum_i (\partial f / \partial x_i) dx_i$.

The global shift δf in the functional argument can be viewed as composed of local manipulations on f which are conveniently expressed in terms of the Dirac delta function: $\delta f(x) = \int \delta f(x' - x) dx'$, where $\delta f(x' - x) = \delta f(x') \delta(x' - x) = \varepsilon h(x') \delta(x' - x) \equiv \varepsilon h(x' - x)$. Here, $\delta f(x' - x)$ stands for the localized displacement of the argument function, centered around x , in terms of which the first functional derivative, itself the functional of f , reads:

$$\frac{\delta F}{\delta f(x)} = \lim_{\varepsilon \rightarrow 0} \frac{F[f(x') + \varepsilon h(x' - x)]_{x'} - F[f(x')]_{x'}}{\varepsilon} \equiv g[f; x], \quad (2.108)$$

where subscript x' in the functional symbol symbolizes integration over this argument [see (2.106)].

One similarly introduces *higher functional derivatives*, which define the consecutive terms in the functional Taylor–Volterra expansion (Volterra 1959; Gelfand and Fomin 1963):

$$\begin{aligned} \delta F[f] &= \int \frac{\delta F}{\delta f(x)} \delta f(x) dx + \frac{1}{2} \iint \delta f(x) \frac{\delta^2 F}{\delta f(x) \delta f(x')} \delta f(x') dx dx' + \dots \\ &\equiv \delta^{(1)} F[f] + \delta^{(2)} F[f] + \dots \end{aligned} \quad (2.109)$$

For example, in the localized perspective on modifying the argument function of the functional, one interprets its second functional derivative as the limiting ratio:

$$\frac{\delta^2 F}{\delta f(x) \delta f(x')} = \frac{\delta g[f; x]}{\delta f(x')} = \lim_{\varepsilon \rightarrow 0} \frac{g[f(x''') + \varepsilon h(x''' - x'); x]_{x'''} - g[f; x]}{\varepsilon}. \quad (2.110)$$

In (2.109) it determines the continuous transformation of the *two-point* displacements of the argument function, $\delta f(x'' - x) \delta f(x''' - x')$, centered around x and x' , respectively, into the second differential $\delta^{(2)} F[f]$. The latter again parallels the familiar expression for the second differential of a function of several variables: $d^{(2)}f(x_1, x_2, \dots) = \frac{1}{2} \sum_i \sum_j (\partial^2 f / \partial x_i \partial x_j) dx_i dx_j$.

The rules of the functional differentiation thus represent the local, function generalization of those characterizing the differentiation of functions. The functional derivatives of the sum and product of two functionals, respectively, read:

$$\begin{aligned} \frac{\delta}{\delta f(x)} \{aF[f] + bG[f]\} &= a \frac{\delta F}{\delta f(x)} + b \frac{\delta G}{\delta f(x)}, \\ \frac{\delta}{\delta f(x)} \{F[f] G[f]\} &= G \frac{\delta F}{\delta f(x)} + F \frac{\delta G}{\delta f(x)}. \end{aligned} \quad (2.111)$$

The *chain rule* transformation of functional derivatives also holds. Consider the *composite* functional $F[f] = F[f[g]] \equiv \bar{F}[g]$. Substituting the first differential of $f(x) = f[g; x]$,

$$\delta^{(1)}f[g; x] = \int \frac{\delta f(x)}{\delta g(x')} \delta g(x') dx', \quad (2.112)$$

into $\delta^{(1)}\bar{F}[g]$ of (2.108) gives:

$$\delta^{(1)}\bar{F}[g] = \int \frac{\delta \bar{F}}{\delta g(x')} \delta g(x') dx' = \int \frac{\delta F}{\delta f(x)} \left[\int \frac{\delta f(x)}{\delta g(x')} \delta g(x') dx' \right] dx. \quad (2.113)$$

Hence, the functional derivative of the composite functional follows from the chain rule

$$\frac{\delta \bar{F}}{\delta g(x')} = \int \frac{\delta F}{\delta f(x)} \frac{\delta f(x)}{\delta g(x')} dx. \quad (2.114)$$

One similarly derives the chain rules for *implicit* functionals. When functional $F[f, g]$ is held constant, the variations of the two argument functions are not independent, since the relation $F[f, g] = \text{const.}$ implies the associated functional relation between them, e.g., $g = g[f]_F$. The vanishing first differential,

$$\begin{aligned} \delta^{(1)}F[f, g] &= \int \left[\left(\frac{\partial F}{\partial f(x)} \right)_g [\delta f(x)]_F + \left(\frac{\partial F}{\partial g(x)} \right)_f [\delta g(x)]_F \right] dx = 0, \quad \text{or} \\ \int \left(\frac{\partial F}{\partial f(x)} \right)_g [\delta f(x)]_F dx &= - \int \left(\frac{\partial F}{\partial g(x')} \right)_f [\delta g(x')]_F dx', \end{aligned} \quad (2.115)$$

is determined by the *partial* functional derivatives, a natural local extension of the ordinary partial derivatives of a function of several variables, e.g.,

$$\left(\frac{\partial F}{\partial f(x)} \right)_g = \lim_{\varepsilon \rightarrow 0} \frac{F[f(x') + \varepsilon h(x' - x), g]_{x'} - F[f, g]}{\varepsilon}. \quad (2.116)$$

Finally, differentiating both sides of Eq. (2.115) with respect to one of the argument functions for constant F gives the following implicit chain rules:

$$\begin{aligned} \left(\frac{\partial F}{\partial f(x)} \right)_g &= - \int \left(\frac{\partial F}{\partial g(x')} \right)_f \left(\frac{\partial g(x')}{\partial f(x)} \right)_F dx', \\ \left(\frac{\partial F}{\partial g(x')} \right)_f &= - \int \left(\frac{\partial F}{\partial f(x)} \right)_g \left(\frac{\partial f(x)}{\partial g(x')} \right)_F dx. \end{aligned} \quad (2.117)$$

These relations parallel familiar manipulations of derivatives in the classical thermodynamics.

For the fixed value of the composite functional $F[f[u], g[u]] = \tilde{F}[u] = \text{const.}$ one similarly finds:

$$\begin{aligned} \left(\frac{\partial g(x')}{\partial f(x)} \right)_{\bar{F}} &= \int \left(\frac{\partial g(x')}{\partial u(x'')} \right)_{\bar{F}} \left(\frac{\partial u(x'')}{\partial f(x)} \right)_{\bar{F}} dx'', \\ \left(\frac{\partial f(x)}{\partial g(x')} \right)_{\bar{F}} &= \int \left(\frac{\partial f(x)}{\partial u(x'')} \right)_{\bar{F}} \left(\frac{\partial u(x'')}{\partial g(x')} \right)_{\bar{F}} dx''. \end{aligned} \quad (2.118)$$

Let us further assume that functions $f(x)$ and $g(x)$ are unique functionals of each other, $f(x) = f[g; x]$ and $g(x') = g[f; x']$. Substitution of (2.112) into

$$\delta^{(1)} g[f; x''] = \int \frac{\delta g(x'')}{\delta f(x)} \delta f(x) dx, \quad (2.119)$$

then gives:

$$\delta^{(1)} g[f; x''] = \int \frac{\delta g(x'')}{\delta f(x)} \delta f(x) dx = \iint \frac{\delta g(x'')}{\delta f(x)} \frac{\delta f(x)}{\delta g(x')} \delta g(x') dx' dx. \quad (2.120)$$

This equation identifies the Dirac delta function as the functional derivative of the function at one point with respect to its value at another point, as also implied by (2.107):

$$\int \frac{\delta g(x'')}{\delta f(x)} \frac{\delta f(x)}{\delta g(x')} dx = \frac{\delta g(x'')}{\delta g(x')} = \delta(x'' - x'), \quad (2.121)$$

where we have applied the functional chain rule. The preceding equation also defines the mutually inverse functional derivatives:

$$\frac{\delta g(x')}{\delta f(x)} = \left(\frac{\delta f(x)}{\delta g(x')} \right)^{-1}. \quad (2.122)$$

Let us assume the functional (2.106) in the typical form including the dependence of its density on the argument function itself and its first n derivatives: $f^{(i)}(x) = d^i f(x)/dx^i$, $i = 1, 2, \dots, n$:

$$\mathcal{L}(x) = \mathcal{L}(x, f(x), f^{(1)}(x), f^{(2)}(x), \dots, f^{(n)}(x)). \quad (2.123)$$

The functional derivative of $F[f]$ is then given by the following general expression:

$$\frac{\delta F}{\delta f(x)} = \frac{\partial \mathcal{L}(x)}{\partial f(x)} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}(x)}{\partial f^{(1)}(x)} \right) + \frac{d^2}{dx^2} \left(\frac{\partial \mathcal{L}(x)}{\partial f^{(2)}(x)} \right) - \cdots + (-1)^n \frac{d^n}{dx^n} \left(\frac{\partial \mathcal{L}(x)}{\partial f^{(n)}(x)} \right). \quad (2.124)$$

The first term in the r.h.s. of the preceding equation defines the so-called *variational derivative*. It determines the functional derivative of the *local* functionals, the densities of which depend solely upon the argument function itself.

This development can be straightforwardly generalized to cover functionals of functions in three dimensions. Consider, e.g., the functional of $f(\mathbf{r})$ depending on the position vector in the physical space: $\mathbf{r} = (x, y, z)$. Equation (2.124) can be then extended to cover the $f = f(\mathbf{r})$ case by replacing the operator d/dx by its three-dimensional analog – the gradient $\nabla \equiv \partial/\partial \mathbf{r}$. For example, for $\mathcal{L}(\mathbf{r}, f(\mathbf{r}), |\nabla f(\mathbf{r})|)$ the functional derivative of $F[f]$ is given by the expression:

$$\frac{\delta F}{\delta f(\mathbf{r})} = \frac{\partial \mathcal{L}(\mathbf{r})}{\partial f(\mathbf{r})} - \nabla \left(\frac{\partial \mathcal{L}(\mathbf{r})}{\partial |\nabla f(\mathbf{r})|} \right). \quad (2.125)$$

Similarly, for

$$\begin{aligned} \tilde{F}[f] &= F[f] + \int \mathcal{L}(\Delta f(\mathbf{r})) d\mathbf{r} \equiv \int \tilde{\mathcal{L}}(\mathbf{r}, f(\mathbf{r}), |\nabla f(\mathbf{r})|, \Delta f(\mathbf{r})) d\mathbf{r}, \quad \Delta = \nabla^2, \\ \frac{\delta \tilde{F}}{\delta f(\mathbf{r})} &= \frac{\delta F}{\delta f(\mathbf{r})} + \Delta \left(\frac{\partial \tilde{\mathcal{L}}(\mathbf{r})}{\partial \Delta f(\mathbf{r})} \right). \end{aligned} \quad (2.126)$$

References

- Dirac PAM (1967) The principles of quantum mechanics. Clarendon, Oxford
 Gelfand IM, Fomin SV (1963) Calculus of variations. Englewood Cliffs, Prentice-Hall
 Volterra V (1959) Theory of functionals. Dover, New York

Perspectives in Electronic Structure Theory

Nalewajski, R.F.

2012, XX, 720 p., Hardcover

ISBN: 978-3-642-20179-0