Roman F. Nalewajski

Perspectives in Electronic Structure Theory

Springer

Perspectives in Electronic Structure Theory

Roman F. Nalewajski

Perspectives in Electronic Structure Theory

Springer

Prof. Dr. Roman F. Nalewajski
Jagiellonian University
Department of Theoretical Chemistry
ul. R. Ingardena 3
30-060 Krakow
Poland
nalewajs@chemia.uj.edu.pl

ISBN 978-3-642-20179-0 e-ISBN 978-3-642-20180-6
DOI 10.1007/978-3-642-20180-6
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011934249

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This book is primarily intended as a textbook introducing to the reader the basic elements of the quantum theory of the electronic structure of molecular systems, including in its first two parts the basic axioms of the nonrelativistic quantum mechanics and rudiments of the wave function and density based theories. Its remaining two parts, of a more monographic character, contain the *Information Theory* (IT) description and some elements of the modern theory of chemical reactivity, respectively. The basic aim of this book is to present in a single text alternative outlooks on the molecular electronic structure, including the basic principles and techniques of the contemporary conceptual and computational quantum chemistry, covering also the insights provided by IT. Together these complementary perspectives enhance the depth of our understanding of the electronic/geometric structure of molecules and provide a full "vocabulary" to tackle diverse conditions, which influence their reactivity behavior. Indeed, only the insights from several different point of view amount to a real understanding of the problem. The emphasis is on the concepts involved and the key ideas encountered in these alternative approaches in the molecular quantum mechanics, and on the interpretation of calculated results in chemical terms: the bonded atoms and molecular fragments, the chemical bonds that connect these building blocks of molecules, and on their responses in a changing environment, which shape the reactivity preferences of reactants.

Explanation and understanding of chemical phenomena ultimately call for the quantum mechanical description provided by the modern quantum chemistry. The latter uses ideas and concepts that differ substantially from their classical analogs. A precise formulation of these generalized physical concepts, which requires some new mathematical tools, is the subject of Part I of this book. The depth and rigor of this physical/mathematical supplement have been dictated by the main didactic purpose of this text: to introduce all tools necessary for understanding the abstract ideas of the modern theory of the molecular structure and chemical reactivity. The foundations of quantum mechanics are covered using the familiar axiomatic approach, with only an introductory summary of the key experiments that led to their formulation.

The reader can familiarize himself with these novel ideas in the simplest problem of the stationary (bonded) states of the hydrogen-like atom presented in the part closing chapter.

The main theories of the molecular electronic structure are presented in Part II. In its opening chapter it examines available techniques of reducing the complexity of solving the molecular Schrödinger equation. In particular, the rudiments of the adiabatic separation of the electronic and nuclear motions are given and the elements of the approximate perturbational and variational approaches for determining the electronic quantum states are outlined. This brief overview also covers the basics of the orbital approximation and the idea of a pseudopotential, which effectively removes the chemically inactive electrons of the atomic inner shells from an explicit treatment in molecular calculations. The subsequent exposition of the principal *Wave Function Theories* (WFT), in which the system wave-function (probability amplitude) defines the quantum state of the molecule, covers the *Self-Consistent-Field Molecular Orbital* (SCF MO) theory, major *Configuration-Interaction* (CI) techniques for dealing with the Coulomb correlation problem, and rudiments of the *Valence Bond* (VB) treatment, which gives a more chemical understanding of molecules compared to its chief rival, the *Molecular Orbital* (MO) description and currently experiences a notable revival.

The following presentation of theoretical basis of the modern techniques of the *Density-Functional Theory* (DFT), in which the electron density or the density matrix constitute the system basic *state*-variables, covers the famous *Hohenberg–Kohn* (HK) theorems and some of their refinements/extensions, the basic elements of the ground-state *Kohn–Sham* (KS) theory and the associated ensemble approach to excited states. The theory of the density functional for the exchange-correlation energy is summarized, including the rudiments of the adiabatic connection and some more recent developments in the field of the density-matrix and orbital-dependent functionals, time-dependent DFT and alternative approaches to the molecular *van der Waals* (vdW) interactions. This short exposition also introduces the main concepts of the density-based reactivity theory: the hardness and softness responses of the electron distribution in molecules in the complementary *electron-following* (EF) and *electron-preceding* (EP) perspectives.

The additional insights from IT are presented in the monographic Part III of this textbook. Its dominating theme is the electron distribution as a source and carrier of information in molecules. First, the basic elements are summarized in the part opening chapter, to be followed by a brief exposition of the information principles in molecular quantum mechanics. The local IT probes of the presence of the direct chemical bonds are formulated and the importance of the nonadditive (interference) information tools is emphasized. In particular, the *Electron Localization Function* (ELF) and the *Contra-Gradience* (CG) bond criterion are used to explore the molecular electronic structure and the IT variational principles are used to derive the so called stockholder scheme for dividing the molecular electron density into the associated atomic pieces. Various *Charge Transfer* (CT) and *Polarization* (P) displacements accompanying the formation of chemical bonds in molecular systems are examined, including the equilibrium redistribution of electrons among the bonded *Atoms-in-Molecules* (AIM) and the molecular promotion of the latter.

Alternative bond-multiplicity descriptors and the electron localization criteria are introduced and molecules are interpreted as communication systems. This concept, developed within the standard SCF MO description, gives rise to the *Orbital Communication Theory* (OCT) of the chemical bond (Nalewajski 2010) an extension of the bond Communication Theory in atomic resolution (Nalewajski 2006). They both use the standard entropic descriptors of information channels in exploring patterns of the chemical bonds in molecules and their constituent parts, as well as the bond covalent/ionic composition.

The molecularly promoted AIM are only slightly modified, compared to their free (separated) analogs, mainly in the outer (valence) shell of electrons. These "external" electrons are responsible for the AIM chemical behavior and the equilibrium bonding pattern they exhibit in the field exerted by the framework of the practically unchanged atomic-cores. This bonding shell of the (delocalized) electrons is also crucial for the propagation of information in the molecule among the system constituent AIM and the *Atomic Orbitals* (AO) the latter contribute to the bonding subspace of the occupied MO, which ultimately determine the system network of chemical bonds. Using the standard tools of IT (summarized in the opening chapter of Part III) in treating these information scattering phenomena due to "communications" via the system chemical bonds provides a novel perspective on the origins and multiplicity of the system chemical bonds, as well as on the entropic nature of their covalent and ionic composition. In particular, the IT multiplicities of the localized chemical bonds are generated, the bond-coupling phenomena in molecular subsystems are discussed and the interference effects due to the multiple information scattering in molecules are examined. The new indirect (*through-bridge*) bonding mechanism is identified, which complements the familiar direct (*through-space*) chemical interactions in molecular systems, and its origins due to the implicit dependencies between AO in the molecular bonding subspace are explored.

The chemical concepts are discussed in a more depth in Part IV. It first provides a survey of alternative perspectives on diverse phenomena conditioning the chemical reactivity, stressing the importance of the conceptual approaches for a more chemical understanding of these bond-forming/bond-breaking processes. The distinction between the "horizontal" (involving displacements of the system electron density) and "vertical" (for the fixed electron distribution) changes in the molecular electronic structure is made and the responses of molecular fragments in the fragment-constrained equilibria are described in terms of the subsystem charge sensitivities. These perturbation–response relations are summarized for all admissible representations of the molecular/subsystem states, covering both the EF perspective of the Born–Oppenheimer approximation and the complementary EP picture, in the spirit of modern DFT. The illustrative case of the bimolecular reactive system is discussed in a more detail and alternative measures of the adiabatic coupling between the electronic and geometrical degrees-of-freedom of the molecular and reactive systems, including the novel compliant theoretical framework, are identified and modeled.

Finally, several qualitative approaches to reactivity phenomena are summarized. They cover recent IT probes of the elementary reaction mechanisms, chemical reactivity indices provided by the alternative hardness/softness (Fukui function) descriptors of molecules and their fragments, e.g., reactants in the *Donor–Acceptor* (DA) systems, as well as the associated equilibrium and stability criteria of molecules and the maximum hardness and the *Hard/Soft Acids and Bases* (HSAB) principles of chemistry. The importance of the complementary internal and external eigenvalue problem of quantum-mechanical observables for a compact description of the electronic processes in molecules and reactants is stressed and alternative hardness-decoupling schemes are examined.

This joint exposition of a variety of perspectives on the electronic structure of molecular systems, which are usually presented in separate texts, aims at comparing these diverse philosophies of treating the subject in the unifying language of the (nonrelativistic) molecular quantum mechanics and IT. Such presentation should help in uncovering the mutual relations between the specific concepts and techniques of these complementary approaches by extracting their common roots in the molecular quantum mechanics, in the frameworks of both the molecular states involved and the associated probability/density distributions.

The book may serve as both the classroom and reference text of the classical and modern ideas in the field of the chemical bond and reactivity theories. This text has evolved from teaching both the graduate and undergraduate courses in quantum chemistry, density-functional and reactivity theories, as well as the IT of molecular systems. It is intended for graduate and advanced undergraduate students and chemical researchers interested in the new ways of looking at the subject. It is hoped that a significant diversity of the student backgrounds have been accommodated in this textbook/monograph of the contemporary ways of thinking about classical issues in the theory of the electronic structure and reactivity behavior of molecules.

Cracow Roman F. Nalewajski
June 2011

References

Nalewajski RF (2010) Information origins of the chemical bond. Nova Science Publishers, New York

Nalewajski RF (2006) Information theory of molecular systems. Elsevier, Amsterdam

Contents

Part I Foundations of Quantum Mechanics

Part II Theories of Electronic Structure

Part III Insights from Information Theory

Part IV Chemical Concepts for Molecular Structure and Reactivity

Acronyms

A	Acidic, acceptor reactant, fragment
a.u.	Atomic units
AB	Acid–base, interaction, complex
AC	Adiabatic connection, of KS DFT
ACFDT	Adiabatic-connection fluctuation-dissipation theorem, DFT method
AIM	Atoms-in-molecules
AO	Atomic orbital(s)
APSG	Antisymmetrized product of strongly-orthogonal geminals, method
Av	Average value
B	Basic, reactant, fragment
BEBO	Bond-energy–bond-order, energy surface, reaction coordinate
BEF	Binary entropy function
BO	Born–Oppenheimer approximation
BOVB	Breathing orbital VB, theory
BSSE	Basis set superposition error
c	Coulomb correlation, hole, energy, potential
CAS	Complete active space, method
CBO	Charge-and-bond-order (1-density) matrix
CC	Coupled-cluster, approximation, method
CCCI	Correlation consistent CI, method, CI extension of GVB
CCSD(T)	CC theory including singly-, doubly-, and triply-excited configurations
CDFT	Conceptual DFT, approach
CEPA	Coupled electron pair approximation, method
CF	Coulson–Fischer, wave function
CG	Contra-gradience, criterion of bond localization
CI	Configuration interaction, theory, method
CID	Limited SR CI method involving *double*-excitations

CIDQ	Limited SR CI method involving *double*- and *quadruple*-excitations
CIDQH	Limited SR CI method involving *double*-, *quadruple*- and *hextuple*- excitations
CISD	Limited SR CI method involving *single*- and *double*-excitations
CISTQ	Limited SR CI method involving *single*-, *double*-, *triple*-, and *quadruple*-excitations
CM	Center of mass
CPKS	Coupled-perturbed KS, method
CPMET	Coupled-pair many electron theory
CS	Charge sensitivities
CSA	Charge sensitivity analysis
CT	Charge transfer, stage of bond formation or reaction, configurations
CTCB	Communication theory of the chemical bond
D	Donor, reactant, fragment
DA	Donor–acceptor, interaction, complex, reactive system
DCACP	Dispersion-corrected atom-centered potentials, DFT method
DFPT	Density functional perturbation theory, see GL theory
DFT	Density functional theory, method
DFT-SAPT	Combined DFT/SAPT, method
DHF	Dirac–Hartree–Fock, method
DIM	Diatomics-in-molecules, method, PES
DMFT	Density-matrix functional theory
E	Electrophilic, site in molecules, attacking agents
EE	Electronegativity-equalization, principle of Sanderson, equations
EF	Electron following, perspective
EHK	Ensemble HK, theory, functional
EKT	Extended Koopmans theorem
ELF	Electron-localization function
EP	Electron preceding, perspective
EPI	Extreme physical information, principle
ESP	Electrostatic potential
EXX	Exact-exchange, approach in DFT
F	Fisher information, intrinsic accuracy
FCI	*Full*-CI, scheme
FDT	Fluctuation-dissipation theorem
FE	Frontier electron, densities
FF	Fukui function, electronic descriptors
FO	Frontier orbitals, theory, electrons, densities
g	*Gerade*, even parity
GEA	Gradient expansion approximation, of *xc*-energy in DFT
GFF	Geometric Fukui Function

GGA	Generalized gradient approximation, of correlation functionals in DFT
GL	Görling–Levy theory, see DFPT
GTO	Gaussian-type orbitals
GVB	Generalized VB, method
GVB-PPA	GVB method using PPA
H	Hartree, theory, method
H	Heisenberg picture of time evolution
HF	Hartree–Fock, theory, method
HFl	HF limit, energy reference
HGGA	*Hyper*-GGA, hybrid gradient functionals in DFT
HK	Hohenberg–Kohn, theorems, method, principle
Hl	Hartree limit, energy reference
HL	Heitler–London, theory, wave-function
Hn	Hellmann, pseudopotential
HOMO	Highest occupied molecular orbital, frontier orbital
HSAB	Hard-soft-acids-and-bases, principle
I	Interaction picture of time evolution
IEPA	Independent electron pair approximation, method
INM	Internal normal modes
IRC	Intrinsic reaction coordinate
IRM	Interreactant modes, see also ISM
ISM	Intersubsystem Modes
IT	Information theory, Information theoretic
K	Kullback symmetrized cross-entropy, divergence
KL	Kullback–Leibler cross-entropy, entropy deficiency, missing information, information distance, directed divergence
KLI	Krieger–Li–Yafrate, method
KS	Kohn–Sham, method, orbitals
KW	Kołos–Wolniewicz, results for H_2
LCAO MO	Linear combinations of AO representing MO
LCGTO	Linear combinations of GTO representing MO
LCSTO	Linear combinations of STO representing MO
LDA	Local density approximation, of functionals
LEPS	London–Eyring–Polanyi–Sato, PES
LO	Localized orbitals
LR	Linear-response, approximation, function
LSDA	Local spin-density approximation of DFT
LUMO	Lowest unoccupied molecular orbital
MBPT	*Many*-body PT
ME	Maximum entropy, rule
MEC	Minimum energy coordinates, in the compliant approach
MED	Minimum entropy deficiency, principle
MEP	Minimum energy path

MGGA	*Meta*-GGA, gradient functionals in DFT
MH	Maximum hardness, principle
MIM	Molecules-in-molecules, method
MO	Molecular orbitals, theory
MP	Møller–Plesset, method
MR (MC) SCF	Multireference (configuration) SCF, method
MRCI	Multireference CI, method
MRVB	Multireference VB, theories
N	Nucleophilic, sites in molecules, attacking agents
NFF	Nuclear FF, geometric descriptors
NG	Natural geminals
NLDA	Nonlocal density approximation, of functionals
NO	Natural orbitals
NOCV	Natural orbitals for chemical valence
NREL	Nonrelativistic limit, energy reference
NSO	Natural spin-orbitals
OAO	Symmetrically (Löwdin) orthogonalized AO
OCT	Orbital communication theory, of the chemical bond
OEP	Optimized effective potential, method in DFT
OEP_x	x-only OEP approach
OPM	Optimized potential method, in orbitally-dependent DFT
OPW	Orthogonalized plane wave, method
P	Polarization (promotion), stage of bond formation or reaction
PES	Potential energy surface
PK	Phillips–Kleinman, local pseudo-potential
PNM	Populational normal modes
PNO	Pseudonatural orbitals
PPA	Perfect pairing approximation, in VB theory
PPP	Pariser–Parr–Pople, semiempirical SCF MO theory
PT	Perturbation theory
QMC	Quantum Monte Carlo, method, results
QSPR	Quantitative structure–property relationships
R	Radical, sites in molecules, attacking agents
RGVB	Resonating GVB, theory
RHF	Spin-restricted HF, theory
RPA	Random phase approximation
RPA DFT	RPA-ACFDT method in DFT
RS	Range separation, in electron interactions
S	Schrödinger picture of time evolution
S	Shannon entropy
SAL	Separated atoms limit, dissociation energy reference, wave-function
SAPT	Symmetry adapted perturbation theory
SBC	Symmetric binary channel

SCF MO	Self-consistent-field MO method
SCVB	Spin-coupled extension of VB, method
SE	Schrödinger equation
SI	Self-interaction, error, hole
SO	Spin orbital(s)
SPA	Single-pole approximation, of TDDFT
SR CI	*Single*-reference CI, method
SR SCF	*Single*-reference SCF, HF method
SRL	Separated reactant limit, reference
STO	Slater-type orbitals
TDDFT	Time dependent DFT, theory, method
TF	Thomas–Fermi, theory
TFD	Thomas–Fermi–Dirac, theory
TS	Transition state, concept of Slater, complex in absolute rate theory
TST	Transition state theory, of reaction rates
u	*ungerade*, odd parity
UHF	Spin-unrestricted HF theory
VB	Valence-bond, theory, structures
VB CI	Valence bond CI, method
VB CIS	Limited CI involving *single*-excitations from VBSCF reference, method
VB CISD	Limited CI involving *single*- and *double*-excitations from VBSCF reference, method
VBSCF	Valence bond SCF, method
vdW	van der Waals, interactions
vW	von Weizsäcker, nonhomegeneity correction
WFT	Wave-function theory
x	Exchange (Fermi) correlation, hole, energy, potential
xc	Exchange-correlation, hole, energy, potential

Part I
Foundations of Quantum Mechanics

Chapter 1
Sources

Abstract A brief presentation of the experimental origins of quantum mechanics is given. The key experiments leading to contradictions with accepted physical theories of matter and radiation, signaling a need for a thorough revision of classical mechanics and electrodynamics, are surveyed. The early attempts to resolve these controversies, formulated at the beginning of twentieth century and often named as the Old Quantum Theory, which mark the genesis of the modern quantum mechanics, are summarized. The specificity of the classical description of physical processes is briefly outlined and main suggestions addressed to a more general mechanics describing the elementary particles, atoms, and molecules are enumerated. The particle diffraction experiment is examined in some detail to pinpoint the essence of the wave–particle duality and to identify the key elements of the quantum description: the initial and final experiments, as well as the free evolution of the system dynamic state which separates them, without any interference from the measuring apparatus. The internal angular momentum of an elementary particle, called spin, is introduced. The emphasis in this historical background is on the development of the classical concepts into their more general quantum counterparts, rather than on their discontinuity in the two theories. On one hand, the classical (approximate) mechanics, in which some very small quantities such as the quantum of the physical action – measured by the Planck constant – are approximated by zero, provides the *geometric* optics limit of the quantum (exact) mechanics. On the other hand, the quantum description has to use the classical concepts due to a macroscopic character of the measuring devices, which adds to the intimate relationship between the two formulations.

1.1 Experimental Origins and Old Quantum Theory

At the current state of our understanding of matter the modern quantum mechanics plays a fundamental role in describing phenomena and processes in the surrounding world, particularly at the *microscopic* level of photons, elementary particles, atoms,

R.F. Nalewajski, *Perspectives in Electronic Structure Theory*,
DOI 10.1007/978-3-642-20180-6_1,

and molecules. It should be emphasized, however, that the complete theory of *macroscopic* objects, of dimensions perceived by our senses, also requires the quantum mechanical description of interactions between their constituent atoms and molecules since the quantum nature of these microscopic particles can be manifested also at the macroscopic level. Clearly, in the limit of very large masses and energies of macroscopic objects the predictions of quantum mechanics must be identical with those resulting from its classical analog. Thus, when supplemented by the laws of statistical thermodynamics the quantum mechanics gives rise to the complete description of the natural world.

It was born in the atmosphere of severe confusion at the beginning of twentieth century, when the accepted physical theories were challenged by numerous dilemmas resulting from a series of remarkable new experimental observations, which could not be explained by the classical mechanics and electrodynamics. The physics at the end of nineteenth century distinguished the categories of matter and radiation, and used separate laws to describe them: Newton's mechanics, to predict motions of material bodies, and the Maxwell equations of the electromagnetic theory of radiation, which unites the electric, magnetic, and optical phenomena. We recall at this point that the so-called *wave* optics becomes the *geometric* optics in the limit of infinitely small wavelength, $\lambda \to 0$, i.e., for infinitely large frequency, $v \to \infty$, of the monochromatic radiation.

Let us now briefly summarize the key stages of the development of quantum ideas in physics (see, e.g., van der Waerden 1968) with the experiment and intuitive insight ultimately leading to a new philosophy of science (Heisenberg 1949, 1958; Yourgrau and van der Merve 1979; Bohm 1980) with the *exact* determinism of classical predictions being replaced by the *statistical* determinism of quantum laws. This "revolution" has also led to a dramatically different way of thinking about the process of measurement, to a discovery of the universal character of the *particle–wave dualism* of both the radiation and matter, and a new definition of the mechanical state of microscopic systems. The crisis of classical physics was indeed observed first on the subatomic and atomic/molecular scales, in processes involving interactions of such objects with electromagnetic radiation, a diffraction of radiation and elementary particles, etc.

We begin this short survey with the problem of the black-body radiation, at equilibrium in the given temperature T, which could not be explained by the classical electrodynamics and eventually led to formulation in 1901 of the famous Planck's hypothesis of the *energy quantization*. The question was this: how much energy is present as radiation in the given volume of an empty space of a cavity in an object held at the definite temperature T, and how it is distributed as a function of the radiation frequency? The quantity describing such a distribution is called the radiation energy density $u(v, T)$, which measures the energy of the monochromatic radiation of frequency v per unit volume of the cavity, in thermal equilibrium at absolute temperature T. The Rayleigh–Jeans law of 1900, $u(v, T) \propto v^2 T$, derived using the classical electrodynamics and statistical thermodynamics, is correct only for low frequencies (in the infrared region of the electromagnetic radiation spectrum) and it dramatically fails for high frequencies (in the ultraviolet region), where the experimental data show a sharp drop in the energy distribution, with $u \to 0$ in the *geometric* optics limit of $v \to \infty$. This classical distribution has been obtained by first calculating the number of elementary oscillators (cavity standing waves) of the electromagnetic field, each corresponding to a particular frequency of radiation, and then ascribing them an average energy $k_B T$, where the Boltzmann constant $k_B = 1.381 \times 10^{-23}$ [J K^{-1}], in accordance with the classical energy *equi*-partition principle.

In order to overcome this discrepancy, also known as the *ultraviolet catastrophe*, which could not be explained by classical means, Planck has proposed that the energy of the elementary radiation oscillator of frequency v, is restricted to integral multiples of the *energy quantum*, $hv \equiv \hbar\omega$, where the new universal constant h has a dimension of the mechanical action [energy × time]; here, the radiation angular frequency $\omega = 2\pi v$ [radians/s] and the symbol $\hbar = h/2\pi$. In other words, this finite "grain" of the oscillator energy constitutes the smallest amount by which the oscillator energy can be increased or lowered. Hence, the energy absorbed by the elementary oscillators of the surrounding cavity can also be absorbed or emitted in integral multiples of such energy quanta, for all frequencies allowed by the cavity standing-wave boundary conditions, as implied by the condition of a thermal equilibrium in the black-body radiation problem: $\Delta E = hv$. This quantum (non-classical) assumption gives rise to the celebrated *Planck*'*s distribution law*:

$$u(v,T) \propto v^3[\exp(hv/k_B T) - 1]^{-1}, \tag{1.1}$$

which is in perfect agreement with experimental observations for the Planck constant (quantum portion of the physical action) $h = 6.626 \times 10^{-34}$ [Js] or $\hbar = h/2\pi = 1.055 \times 10^{-34}$ [Js].

It should be emphasized that this assumption was incompatible with the principles of classical physics. Thus, the agreement with experiment has been achieved only by introducing into the framework of the contemporary physics, in which the oscillator energy and mechanical action constitute the continuous dynamical quantities, the artificial "discrete" quantum condition, incompatible with the basic principles of the classical theory.

This energy quantization has been generalized in 1905 by Einstein into hypothesis of the elementary, localized (indivisible) portions of the electromagnetic energy, defining the radiation particles called *photons*, each containing Planck's portion of the energy: $E = hv$. This assumption provides the complete explanation of the *photoelectric effect* discovered by Hertz in 1886 and 1887. Photoelectrons are produced instantaneously, when the light of a frequency higher than some threshold value v_0 strikes any substance. This phenomenon is governed by the two laws formulated by Lenard in 1899–1902: (1) the number of photoelectrons is proportional to the intensity of the incident radiation; (2) their maximum velocity v and hence also the kinetic energy are affected only by the radiation frequency, and not by its intensity as predicted by the classical, wave theory of radiation. In Einstein's hypothesis the photoelectron energy of motion originates entirely from a single photon, representing a localized corpuscle of the energy, and satisfies the energy conservation

$$\frac{1}{2}m_e v^2 = h\nu - h\nu_0, \tag{1.2}$$

where m_e denotes the mass of an electron and the threshold energy $\Phi = h\nu_0$ measures the so-called *work function* of the irradiated substance.

The electromagnetic radiation thus exhibits a dual character. On one hand, in the diffraction (interference) experiments, it behaves as a wave characterized by the frequency ν $[s^{-1}]$ or wave length $\lambda = c/\nu$, where c stands for the velocity of light in vacuum. On the other hand, as the localized particle of energy, it should be characterized by the linear momentum p. Using the relativistic expression for the energy, $E = m_f c^2 = p_f c = h\nu$, where m_f stands for the photon mass of motion (its rest mass vanishes), one obtains the relativistic expression for the photon momentum:

$$p_f = h\nu/c = h/\lambda \quad \text{or} \quad p_f = \hbar(2\pi/\lambda) \equiv \hbar k, \tag{1.3}$$

where k $[\mathrm{m}^{-1}]$ stands for the photon wave number.

In 1922 this corpuscular nature of radiation has been confirmed experimentally by Compton in the X-ray photon scattering by electrons. The collisions between photon (particle of radiation) and electron (particle of matter) have been shown to be governed by the conservation of the system energy and linear momentum, the two laws that govern any perfectly elastic collisions, e.g., of the billiard balls in the macroscopic world. It also follows from this experiment that any measurement of the particle position, effected by a scattering of light, influences the particle linear momentum; the more precise is this experiment, i.e., the shorter the wave of the incident radiation, the more perturbed is the particle motion after collision with the photon. This implies that in the microscopic world the measuring device and the object of measurement are not absolutely separable as it is implicitly assumed in the classical theory.

A second challenge to the established theory came from the atomic physics. In 1911 Rutherford had demonstrated, by scattering the α-radiation particles (nuclei of the helium atoms) on thin layers of heavy metals, that each atom contains the positively charged, heavy nucleus, with the estimated diameter of the order 10^{-15} [m], surrounded by light, negatively charged electrons, with the estimated diameter of the atom as a whole of the order 10^{-10} [m]. He also guessed that electrons are moving along the circular or elliptic trajectories around the nuclear attractor. This "planetary" model of an atom was in an obvious conflict with the accepted classical electrodynamics, which predicted that electrons moving on a circular orbit, thus being accelerated, should radiate electromagnetic energy and ultimately collapse onto the nucleus. Therefore, the very stability of such a "classical" atomic model has been put in doubt.

To remove this troubling inconsistency, in 1913 Bohr has followed the Planck approach of incorporating in the classical theory subsidiary quantum conditions which contradicted it. He has achieved an excellent agreement with the available experimental data for the hydrogen atom by assuming that in the circular motion of an electron allowed are only specific, *stationary* orbits, on which the particle energy remains fixed. These *stationary* energy levels $\{E_n\}$ and corresponding radii $\{r_n\}$ are identified by the orbit *quantum number* $n = 1, 2, \ldots$. The energy is emitted/absorbed in the discrete manner, not continuously as predicted by the classical electrodynamics, only when electron makes a transition between the two stationary orbits. Emission takes place when electron "jumps" from an outer orbit, exhibiting larger radius, to an inner orbit of smaller radius, identified by the higher and lower values of n, respectively. Accordingly, the inner → outer transitions are possible only after absorbing the energy from an incident radiation. Bohr has used Planck's relation between the transition energy and frequency of the emitted/absorbed radiation:

$$\Delta E_{n\to n'} = E_{n'} - E_n = h\nu_{n\to n'}. \tag{1.4}$$

Bohr's quantum conditions, which determine the stationary orbits, can be formulated as those for the allowed, discrete values of the length of the electron angular momentum $\boldsymbol{l}_n = \boldsymbol{r}_n \times \boldsymbol{p}_n$,

$$l_n = |\boldsymbol{l}_n| = m_{\mathrm{e}} v_n r_n = n\hbar, \tag{1.5}$$

where $\boldsymbol{r}_n$ denotes the electron position vector on nth orbit, and $\boldsymbol{p}_n = m_{\mathrm{e}}\boldsymbol{v}_n$ stands for its linear momentum.

This model has been subsequently developed in 1915 and 1916 by Sommerfeld and Planck, who introduced the elliptic orbits and the spatial quantization of the angular momentum. This generalized planetary model still gave wrong predictions already for helium atom (*two*-electron system), which signaled that this *Old Quantum Theory* was far from the final formulation of the new, generalized mechanics of microscopic objects. It should be realized, however, that new physical ideas are always arrived at by understanding the novel in terms of the familiar. Clearly, Bohr's quantization rules, successful as they were, entail assumptions which are in conflict with the classical physics. For example, the latter predicts that an electron on the circular orbit should emit radiation and this contradicts the assumed stationary character of such a trajectory. Although it was clear already at the time of its invention that this *ad hoc* synthesis of the quantum elements with the classical theory has hardly any future as the consistent physical theory, Bohr's planetary model has turned out to be quite successful in explaining the observed series of spectral lines emitted by hydrogen. The predictive power of the model was quite limited, however, since – despite later improvements – it dramatically failed to explain the spectral data of *many* electron atoms.

Since the *micro*-objects escape perception by human sense organs, their observation always requires the measurement devices, the *macro*-objects which translate their interactions with the *micro*-objects in terms of macroscopic quantities. This points out to a subtle relationship between the quantum mechanics and classical physics. In his celebrated *Correspondence Principle* Bohr has recognized that quantum mechanics must be consistent with classical mechanics. The classical limit corresponds to very large energies (quantum numbers), when such minute quantities as the Planck constant can be formally treated as zeros, in the $h \to 0$ limit.

In 1924 the quantum condition (1.5) of Bohr's model has gained a convincing interpretation in de Broglie's hypothesis of the *universal* character of the *particle–wave dualism*, which was first observed in the electromagnetic radiation. He suggested that the relations between corpuscular (E, p) and wave (ν, λ) attributes of material particles, which exhibit a nonzero rest mass, are the same as for photons, for which the rest mass vanishes (1.3). Therefore, there should also be a new, wave facet of electrons, linked to their more familiar corpuscular aspect by the associated relations:

$$E_\mathrm{e} = h\nu_\mathrm{e}, \quad p_\mathrm{e} = h\nu_\mathrm{e}/c = h/\lambda_\mathrm{e}. \tag{1.6}$$

The existence of such *matter waves* has been confirmed experimentally in 1927 by Davisson and Germer, who diffracted the electron beam on a crystal. This development has quantitatively verified the preceding relations thus demonstrating that the particle–wave duality constitutes a universal characteristic of nature, i.e., of all objects in the microworld, or the *micro-objects* for short, rather than being a monopoly of light. Apparently, in this scale of the linear dimensions 10^{-8}–10^{-15} [m], the differences between the material and radiation particles are significantly blurred. The hope was that in the final version of the quantum theory this important discovery will find a consistent synthesis and a more explicit dynamical expression. At this time it has not been understood yet as to how de Broglie's waves propagate and how they influence the motion of individual particles. They do offer, however, a solid basis for explaining Bohr's quantum condition of (1.5). More specifically, rewriting it in terms of the electron de Broglie's wavelength of an electron moving on nth stationary orbit, $\lambda_n = h/p_n$ (1.6), gives: $2\pi r_n = n\lambda_n$. This condition thus represents the classical criterion for the standing wave along the whole perimeter of the electron circular orbit. In other words, only on the stationary orbits of Bohr the constructive interference of de Broglie's (traveling) waves explains the stability of the electron distribution. Accordingly, the destructive interference of the de Broglie waves in an atom disallows any orbit which fails to satisfy this quantum condition.

Since science is concerned only with observable things one has to let the micro-particle to respond to some outside influence, in order to observe it. As we have already argued above, when examining the implications of the Compton experiment, the measurement process inadvertently modifies the state of the micro-object. A careful examination of the limitations imposed by this influence on the accuracies Δx and Δp_x of the simultaneous determination of the particle position (Cartesian) coordinate x and its conjugate linear momentum p_x, respectively, has led Heisenberg to formulate in 1926 and 1927 his famous *Uncertainty Principle*, also known as the *Principle of Indeterminacy*, which states that the limiting value of the product of these two indeterminacies has a very small but finite value of the order of Planck's constant:

$$\Delta x \; \Delta p_x \geq \hbar. \tag{1.7}$$

The specific multiple of $\hbar$ in r.h.s. of the preceding inequality depends on the adopted measure of the measurement precision. For example, the *standard deviation* σ_A of physical quantity A, $\Delta A \cong \sigma_A = \left\langle (A - \langle A \rangle)^2 \right\rangle^{\frac{1}{2}} = \left(\left\langle A^2 \right\rangle - \langle A \rangle^2 \right)^{\frac{1}{2}}$, where $\langle A \rangle$ is the average, statistical expectation value of A and $\left\langle A^2 \right\rangle$ denotes the average value of its square, can be used to quantify the accuracy of such measurements. We shall use this familiar descriptor of a random variable later in this book, when formulating the Uncertainty Principle in terms of concepts of the molecular wave mechanics.

This limit to the fineness of our power to observe the atomic objects and the smallness of their accompanying disturbance in an act of measurement introduces the *absoluteness* to the distinction between the micro- and macro-objects. This limit can never be surpassed by an improved technique or increased skill of an observer, since a fraction of a photon is never observed. It is inherent in natural world and the dual particle–wave behavior, "anomalous" from the classical perspective, is not peculiar to light, but it is universally present in all material particles as well.

1.2 Classical–Mechanical Description and a Need for Its Revision in Generalized Mechanics

A necessity for a departure from the classical mechanics and its causality is thus clearly demonstrated by the experimental observations. The classical concepts have been proved to be inadequate to describe the molecular, atomic, and subatomic events. The uncertainty principle denies an observer the ability to simultaneously measure the conjugate components of the position and momentum vectors of micro-objects with arbitrary high precision. This contradicts the basic assumption of the classical mechanics, in the canonical formulation of the Hamilton equations of motion, where the exact knowledge of such quantities is required for the very definition of the particle dynamic state. According to the Heisenberg principle of indeterminacy such simultaneously (sharply) unobserved quantities are *unknowable*. Therefore, one is forced to resign from the classical concept of the particle trajectory, e.g., Bohr's orbit, which is unobservable thus belonging in the micro-world to a "metaphysical" rather than physical category.

Hence, the precise description of the time evolution of a micro-object, which requires an exact knowledge of its position and momentum at the given time, is unavailable in the quantum theory. This restriction does not reflect our technical inability of a precise measurement, but rather it signifies the incompatibility of the two observations involved. Such physical quantities, which cannot be sharply defined simultaneously, are called the *complementary observables*. As we shall see later in the book, besides the complementary pair of the particle position and momentum, (x, p_x), there is a number of such relations in quantum physics: energy and time, (E, t), any two Cartesian components of the angular momentum, e.g., (l_x, l_y), etc.

The uncertainty relations give rise to *statistical* predictions of the quantum theory, in contrast to the *deterministic* predictions of the classical physics. In the macroscale of objects perceived by our senses, the statistical distribution of the alternative outcomes of a measurement, represented by the normal (Gaussian) distribution, can be made infinitely sharp in the limit of the Dirac *delta* function (Dirac 1967), which can be thought of as representing the ordinary Gauss curve of the probability theory in the limit of its vanishing variance. Therefore, the *statistical* (*multiple*-valued) determinism of quantum mechanics constitutes a natural extension of its limiting form in the *strict* (*single*-valued) determinism of the classical theory.

According to Bohr's *Complementarity Principle* both coexisting wave and particle aspects of all objects in the microworld are essential for their full description. However, the precise specification of one complementary observable rules out any specification of the other. Should the particle momentum be known exactly, $\Delta p_x \rightarrow 0$, one would then have no knowledge of its position whatsoever, $\Delta x \rightarrow \infty$; accordingly, when the object position is sharply defined, $\Delta x \rightarrow 0$, one looses all the knowledge about its momentum: $\Delta p_x \rightarrow \infty$. The principle operates not only in these limiting cases, but it also covers all intermediate, finite precisions of specifying the pairs of complementary observables. The more the precise localization of an electron (or photon) in space, when its momentum is not well specified, the more the particle-like behavior. Accordingly, the wave-like character is uncovered, when the particle localization is not well specified, i.e., when its momentum is determined more precisely.

As further articulated by Bohr and his Copenhagen School, all physical quantities such as position, momentum, angular momentum, energy, etc., have to be specified by measurement, which conveys information to our senses. It has to contain amplification mechanisms by which microscopic effects are translated into macroscopic effects accessible to our understanding. Indeed, all experiments in the atomic, nuclear, and subnuclear scales in the final analysis are described in classical terms, related to attributes of the macroscopic measuring apparatus. This emphasizes a unique, intimate relationship between the quantum mechanics and its classical limit, with the former being destined to use the concepts of the latter to describe the behavior of the micro-objects.

The indeterminacy principle also implies a relativity of the quantum description with respect to the adopted method of measurement, since the specific experimental device uncovers its own "projection" of the observed "reality." This also constitutes a natural extension of the classical relativity of the description of physical phenomena with respect to the adopted reference frame. This feature signifies a deeper, fully objective approach, which resigns from the subjective classical idealization of the exact separability of the observed object and the measuring device. It is implicitly assumed in the classical theory that the progress of a physical process is independent of the experimental observations, which monitor its current stage. In other words, classical theory claims a lack of interference of the measuring device into the state of the probed mechanical system, i.e., the absolute separability of these two subsystems of an experimental arrangement.

Clearly, the physical objects evolve freely when undisturbed by an act of measurement, but finally we have to bring them into contact with the experimental apparatus to monitor their current (final) state. The progress of classical process is assumed to be independent whether they are observed experimentally or not, but in the realm of quantum mechanics the experimental monitoring is not without an influence, sometimes decisive, on the behavior of the observed micro-object. In the macroworld this influence can be practically neglected. For example, the perturbation of the airplane trajectory created by the photons of the illuminating radar radiation is nonexistent for all practical reasons. To summarize, the impression of the unequivocal determinism in the Newtonian mechanics is created by the very high masses and energies of the classical objects. It hardly implies the universality of this limiting macroconjecture of the absolute separability of the object and measuring device, to also cover the microworld where such small perturbations do matter.

The classical description also assumes the possibility of limitless gathering of simultaneous measurement information, i.e., the availability of the precise values of all mechanical properties of all constituent particles at the given time. In other words, this approach assumes that in principle at a given time all objects can be absolutely localized in space and their momenta can be determined with arbitrary precision, as can be any physical property of the dynamical system under consideration. Clearly, for practical reasons only, we are unable to reach this level of the precise specification of the mechanical microstate of all atoms/molecules in a macroscopic amount of matter. However, as claimed in the classical statistical thermodynamics, such detailed data are in principle knowable with arbitrary precision. Only due to the obvious "technical" difficulties of reaching this goal, and in view of the implications of the *Law of Large Numbers*, which renders such information irrelevant, we resort to familiar methods of the statistical mechanics in predicting the *average* descriptors of the system macrostate.

Let us briefly summarize the main suggestions addressed to the generalized mechanics capable of describing the behavior of micro-objects. As we have already argued in the preceding section, the relation between this, yet unknown, new mechanics and its classical analog should be similar to the relation between the *wave*- and *geometrical* optics; the former becomes the latter in the formal short-wave limit of $\lambda \to 0$ ($v \to \infty$), which is a characteristic of de Broglie's wave of a macro-object, when the free particle would not be diffracted but going along a straight rectilinear path, just as we expect classically. The new mechanics should thus include the classical mechanics as its limiting case for very large energies and hence also large values of its quantum numbers $-$ or equivalently $-$ in the formal limit of the vanishing quantum of the physical action: $h \to 0$. This can be argued more precisely by observing that the wave aspect of matter will be hidden from our sight, if de Broglie's wavelength λ is much lower than a characteristic length d involved in describing the motion of a body of momentum p: $\lambda/d = h/(dp) \ll 1$.

Thus, the $\lambda \rightarrow 0$ and $h \rightarrow 0$ limits are equivalent in identifying the range of applications of the classical mechanics. This postulate is known as Bohr's Correspondence Principle.

In contrast to old quantum theories, the general quantum theory must be internally consistent, i.e., all its experimental consequences must follow from the same axiomatic basis. It has to be capable of explaining all known experimental facts, rather than a narrow selection of such data. In the new mechanics we have to refrain from the classical definition of the system dynamic state, which uses the complementary observables. The new definition must instead be based only on the strictly knowable state parameters, which can be simultaneously determined with utmost precision. Clearly, such a positivistic attitude is a prerequisite of any sound physical theory.

The new definition of the mechanical state must be complete so that the results of all possible experiments performed on the microsystem can be extracted from it. In particular, it must offer means to predict the possible outcomes (spectrum) $\{a_i\}$ of any single measurement of quantity A, as well as the frequencies m_i (or probabilities) $\{p_i = m_i/m\}$ of these experimentally allowed values of the measured physical quantity in many repetitions $m = \sum_i m_i$ of the given experiment, performed on systems in the same dynamical state. This information on a multitude of measurements performed on replicas of the system then suffices to determine the statistical expectation value of the measured physical quantity:

$$\langle A \rangle = \sum\nolimits_i p_i\, a_i. \tag{1.8}$$

1.3 Implications from the Particle Diffraction Experiment

Let us consider the double-slit interference of photons or electrons, in analogy with Young's optical experiment. In this experimental arrangement the monochromatic stream of quantum particles falls on the opaque diaphragm with two slits O_1 and O_2. This experiment is crucial for distinguishing whether a perturbation traveling in space is of the particle or wave character.

The intensities $I_1(x)$ and $I_2(x)$ of two streams of the noninteracting particles passing through the openings O_1 and O_2, respectively, when the other slit is closed, upon reaching the screen $\mathcal{E}$ would produce the sum of such individual intensities (probabilities), $I_1(x) + I_2(x)$. The superposition of the corresponding waves $\psi_1(x) = |\psi_1(x)| \exp[\mathrm{i}\phi_1(x)]$ and $\psi_2(x) = |\psi_2(x)| \exp[\mathrm{i}\phi_2(x)]$,

$$\psi(x) = \psi_1(x) + \psi_2(x), \tag{1.9}$$

gives rise to the screen intensity distribution exhibiting the interference effects,

$$\begin{aligned} I(x) &= |\psi(x)|^2 = \psi(x)\psi^*(x) = |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2|\psi_1(x)||\psi_2(x)|\cos[\phi_1(x) - \phi_2(x)] \\ &\equiv [I_1(x) + I_2(x)] + 2[I_1(x)I_2(x)]^{\frac{1}{2}}\cos[\phi_1(x) - \phi_2(x)] \equiv I^{add}(x) + I^{nadd}(x), \end{aligned} \tag{1.10}$$

because of the last, nonadditive (oscillatory) term $I^{nadd}(x)$. Above, we have identified the intensity of wave by the squared modulus of the scalar wave field $\psi(x)$, by analogy to the intensities of the electric, $E(x)$, or magnetic, $H(x)$, fields.

It has been established experimentally that the interference fringes are the statistical result of a very large number of independent particles hitting the screen, when each particle retains its individuality being finally deposited on a single grain of the photographic plate of the screen, at apparently random positions, hitting also the regions no classical particle could reach. The same interference pattern appears when a beam of particles goes through the slits simultaneously, and when single particles are scattered, one at a time, with the impact locations being observed in seemingly random fashion, now here, now there, over a length of time. The statistical determinism in this scattering of micro-objects, which give the impression of being truly indeterminable and chaotic, is only revealed after very many repetitions of such elementary, single-particle experiments, when the interference pattern finally emerges.

The appearance of interference depends critically on *both* slits being open, and it vanishes when one of them is closed, i.e., when a single particle goes definitely through one slit or the other, giving after many repetitions the separate distributions $I_1(x)$ or $I_2(x)$ on the screen. One thus concludes that the observance of interference denies us the determination of the slit through which the particle has actually passed. The interference pattern cannot be explained in the corpuscular representation, as a result of some collective effect of interactions between the beam particles. More specifically, by diminishing the density of the incident stream of particles, and hence also the number of particles passing through the slits in unit time, one changes such interactions, and this should affect the interference pattern on the screen. However, the experiment does not exhibit any influence of this kind; the diffraction pattern remains the same even in the limit of a single particle passing the slits at a time. The attempts to explain this phenomenon in the wave representation alone also fail, as the interference intensities, i.e., the wave determinism of the particle distribution is uncovered only after many repetitions of the single-particle scatterings performed at the specified dynamical conditions of the incident beam.

These apparent contradictions illustrate the wave–particle dualism of the micro-objects. Indeed, in accordance with the Heisenberg indeterminacy principle, it is impossible to simultaneously, sharply specify the particle momentum $p = h/\lambda$, which implies the knowledge of the interference pattern, and its position, which presupposes the knowledge of the slit, through which the particle has passed, when the other slit remains closed.

Therefore, there is a distinct *wave* causality in this at first glance "random" scattering of independent particles so that de Broglie's wave $\psi(x, t)$, or the *wave* (*state*) *function* for short, indeed describes in a statistical sense a movement of a single particle, with the wave intensity $I(x, t) = |\psi(x, t)|^2$ (1.10) measuring the chance of finding it hitting the screen at location x at time t. This probabilistic interpretation of the waves of matter is due to Born, who proposed in 1927 to call the intensity $I(x, t)$ the *probability density* of observing the particle at specified localization at the given time. As we shall see later in the book, in the modern quantum mechanics this identification forms a basis for interpreting the system wave function, which carries the complete information about the dynamic state of the micro-object. It should also be emphasized that this function itself, the solution of the Schrödinger wave equation formulated in 1926, which governs the dynamics of microsystems, cannot be treated as a measure of the likelihood of finding a particle at the given position, since for that it should be positive everywhere, being then incapable of the destructive interference, which is the observed fact.

The *double*-slit diffraction of microparticles identifies two types of experiments involved in establishing the classical attributes of quantum systems. Let us examine the consecutive stages of a general setup in a thought experiment shown in Fig. 1.1. We denote the initial and final states (wave functions) of the quantum system, at time $t_0 \equiv 0$ and $t > 0$, respectively, by $\psi(x, t_0)$ and $\psi(x, t)$. The classical attributes of the initial state are determined by performing the so-called *initial experiment*, which in fact creates $\psi(x, t_0)$, e.g., the monochromatic beam of particles of the specified momentum. Thus, this first category of experiment in quantum mechanics always refers to the *future*, by preparing the quantum state the time evolution of which we intend to study.

In the period $t_0 \to t$ the system evolves freely, $\psi(x, t_0) \to \psi(x, t)$, without any perturbing influence from measuring devices. This wave deterministic process will be described by the Schrödinger equation of motion, which in the modern quantum mechanics replaces the Newton (Hamilton) equations of motion of the classical theory. As we shall see later in the book, this evolution of the state function in the

$$\begin{bmatrix}\textit{initial}\\ \textit{experiment}\end{bmatrix} \Rightarrow \begin{bmatrix}\psi(t_0)\\ \textit{initial}\\ \textit{state}\end{bmatrix} \overset{\textit{free evolution}}{\Rightarrow} \begin{bmatrix}\psi(t)\\ \textit{final}\\ \textit{state}\end{bmatrix} \Rightarrow \begin{bmatrix} & \textit{final experiments}: & \\ \textit{measurements of } A & & \textit{spectrum, probabilities}\\ \{\hat{A}\psi(t)\} & \Rightarrow & \{a_i\}, \{p_i\}\end{bmatrix}$$

Fig. 1.1 Qualitative diagram of the *initial* and *final* experiments involved in preparing the initial state $\psi(t_0)$ and extracting the classical attributes of the final state $\psi(t)$ reached after free (undisturbed by measurement) evolution in the time interval $t_0 \to t$. The initial experiment arrangement, including the particle collimating slits and an appropriate velocity selector, transforms the polychromatic electron beam into its monochromatic component, thus preparing the initial state $\psi(t_0)$. In the time interval $t_0 \to t$ the system evolves freely, without any intervention from the measuring devices, in the specified dynamical conditions, e.g., when the particle motion is influenced by the force field generated by the external potential $v(x)$, in accordance with the *strictly* deterministic laws of quantum dynamics: $\psi(t_0) \to \psi(t)$. The statistically distributed classical attributes of the final state $\psi(t)$ are then extracted by performing the final experiment, using, e.g., the *double*-slit arrangement or a crystal as the measuring apparatus, which diffracts electrons to the movable detector or a photographic plate. This position-extraction experiment is an illustrative example of a general measurement-event of any physical observable A. The process of extracting the observed values $\{a_i\}$ (spectrum) of A in the *single*-particle experiments performed on the final state $\psi(t)$ has been symbolically depicted in the diagram as performance of the relevant mathematical operation $\hat{A}$ on $\psi(t)$, $\hat{A}\psi(t)$, with the operator $\hat{A}$ being specific for the measured quantity A. The observed spectrum $\{a_i\}$ of A and the associated probabilities $\{p_i = m_i/m\}$ can be determined only after many $m = (\sum_i m_i) \to \infty$ repetitions of the *single*-electron scatterings, with m_i denoting the frequency of observing a_i

specified dynamical conditions is strictly deterministic, with the given initial state $\psi(x, t_0)$ giving rise to a single final state $\psi(x, t)$.

The aim of the *final experiment* is to determine the classical descriptors of the quantum system in state $\psi(x, t)$. It should be stressed that after the particle has been localized on the screen, by using the photographic plate or some clever monitoring device, its dynamical state has been inadvertently and irreversibly destroyed as a result of the interaction with such an apparatus. Indeed the particle's precise localization denies us of any knowledge about the particle momentum. Thus, the final experiment can have implications only to the very *past* event, when the micro-object reaches the screen.

Due to the particle–wave duality, the link between $\psi(x, t)$ and possible outcomes of the final experiment is generally of the "*one-to-many*" type, thus giving rise to statistical predictions of specific values of classical descriptors of the system final state. Indeed, we cannot a priori predict, where the scattered electron hits the screen, but the final interference pattern, obtained after numerous repetitions of the single-electron diffractions, uniquely identifies the probability distribution $|\psi(x, t)|^2$ of the final state. It should be emphasized that only very numerous repetitions of the single-particle "experiment" together constitute the complete final experiment in quantum mechanics.

The preceding discussion prompts us to revise our ideas of causality (Heisenberg 1949, 1958; Born 1964; Bohm 1980; see also: Penrose 1989). Causality applies only to the micro-objects which are left undisturbed. Therefore, only the free-evolution in the chain of events depicted in Fig. 1.1 represents the causal stage, while the final measurement produces a disturbance in the state of the object serious enough to destroy any causal connection between the *separate* results of observations monitoring the object final state.

The statistical predictions and the indeterminism of quantum laws are a property inherent in nature, and should not be regarded as resulting from our temporary ignorance, which could be removed by some future theory, better and more complete. Although the modern quantum theory provides a thoroughly rational, coherent, and extremely successful description of micro-objects of the subatomic and atomic/molecular levels, one should not dogmatically rule out its future improvements and extensions, e.g., on the subnuclear level. However, as much as the quantum mechanics was forced upon the modern science by the physical rather than metaphysical necessity, these developments have to address future experimental findings, which could not be explained by the quantum theory. Indeed, as history teaches us, no matter how complete the description of the dynamical state may seem today, sooner or later new experimental facts will require us to improve the theoretical model and arrive at an even more general description, more detailed and usually more complex.

For example, all empirical evidence, including the *Stern–Gerlach experiment* and atomic spectra, points to the need for attributing to many elementary particles, notably electrons, protons, and neutrons, the intrinsic angular momentum, or *spin*, and the associated magnetic moment. Therefore, such particles can hardly be treated as mass points without any internal structure. Hence, for the complete specification of their dynamic states one has to provide the relevant spin quantum numbers, which fix these internal degrees-of-freedom of such micro-objects. These new dynamical variables of entirely nonclassical origin have to be specified besides the remaining simultaneously measurable observables.

1.4 Particle Spin

In 1925 Uhlenbeck and Goudsmit hypothesized the existence of yet another internal attribute of atoms and elementary particles, called *spin* angular momentum and the associated intrinsic magnetic dipole moment, which complement such properties of these micro-objects as mass, electric dipole moment, moment of inertia, electric charge, etc. This internal state variable has been originally introduced to simplify the classification of atomic spectra. This goal has been achieved, when one envisaged the existence of the internal angular momentum $\boldsymbol{s}$ of an electron, called the spin, the length of which is quantized by the *half* integral quantum number $\mathscr{s} = \frac{1}{2} : s = |\boldsymbol{s}| = [\mathscr{s}(\mathscr{s}+1)]^{\frac{1}{2}}\hbar$ (Fig. 1.2).

Confirmation of this experimental conjecture came in 1928 from the relativistic quantum theory of Dirac. The existence of the electronic spin also transpires from

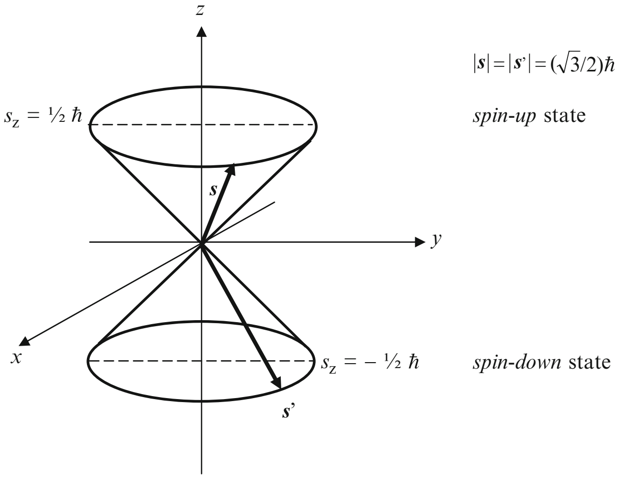

Fig. 1.2 The electron spin $\boldsymbol{s}$ can be characterized in quantum mechanics by two simultaneously observable attributes: its length $s = |\boldsymbol{s}| = [\mathscr{s}(\mathscr{s}+1)]^{\frac{1}{2}}\hbar = (\sqrt{3}/2)\hbar$, for the half-integral spin quantum number $\mathscr{s} = \frac{1}{2}$, and its projection on the specified axis, say axis "z" of the Cartesian coordinate system: $s_z = \sigma\hbar$, where $\sigma = \pm\,\mathscr{s}$. These two observables do not strictly specify the spin vector, but rather they define the whole family of admissible vector directions determining the cone surfaces shown in the diagram. The length and a single projection exhaust the complete list of simultaneously observed properties of any angular momentum in quantum mechanics. In other words, the direction of the angular momentum of the microparticle is not an observable

the earlier Stern–Gerlach experiment of 1921 in which a beam of silver atoms, containing a single, outermost spin-unpaired electron, produce two traces corresponding to the spin-up ($s_z = {}^1\!/_2\hbar$) and spin-down ($s_z = -{}^1\!/_2\hbar$) states (Fig. 1.2) of their valence electron, after being deflected in a nonuniform magnetic field.

These two spin states of a single electron can be uniquely specified by the quantum numbers determining the two simultaneously measurable attributes of the spin vector: $\mathscr{s}$, for its length, and $\sigma = \pm\mathscr{s}$, for its projection along the specified direction, say the "z" axis in Fig. 1.2: $s_z = \sigma\hbar$. They can be symbolically represented as the following "state vectors," in which one provides an explicit or symbolic specification of the state spin quantum numbers within the arrow-like symbol of Dirac:

$$\textit{spin-up}\ \text{state:}\qquad |\,\alpha\,\rangle = |\mathscr{s}, \sigma = +\mathscr{s}\rangle = |{}^1\!/_2, +{}^1\!/_2\rangle = |\,+\,\rangle,$$

$$\textit{spin-down}\ \text{state:}\quad |\,\beta\,\rangle = |\mathscr{s}, \sigma = -\mathscr{s}\rangle = |{}^1\!/_2, -{}^1\!/_2\rangle = |\,-\,\rangle.$$

1.5 Birth of Modern Quantum Mechanics

The consistent quantum mechanics (see, e.g., Messiah 1961; Davydov 1965; Dirac 1967; Merzbacher 1967; Cohen-Tannoudji et al. 1977; Fock 1986), which explains the origins of the quantization of the physical observables and introduces the generalized dynamics of quantum states, has emerged in 1926–1927 in two equivalent forms: the *Matrix Mechanics* of Heisenberg and the *Wave Mechanics* of Schrödinger. Although using quite different mathematical apparatuses, the matrix algebra and differential equations, respectively, these two rival theories gave rise to identical physical predictions, in complete agreement with all experimental data. It was clear, therefore, that these two approaches represent the same physical theory, as indeed demonstrated later by Schrödinger and Dirac (see, e.g., Buckley and Peat 1979).

Heisenberg discovered the need for a generally *noncommutative* multiplication of physical quantities in quantum mechanics, which gives rise to the position–momentum indeterminacy. The analogies with systems in classical mechanics, which are governed by the linear equations of motion, a consequence of the superposition relationships between states of vibrating strings or membranes, have led Schrödinger to establish the basic equations of the Wave Mechanics. The resulting equation of state is also linear in the unknowns, because of the assumption of the quantum superposition principle. In Heisenberg's approach the quantum states and physical observables are represented by the matrix vectors and square matrices, respectively, while in Schrödinger's treatment they are accordingly associated with functions and differential operators. The important contributions to the final form of the modern quantum theory have also been made by other members of the Göttingen School, Born and Jordan, and by Dirac and Pauli, who invented the relativistic version of the quantum theory.

These revolutionary departures from principles of the classical theory, and particularly in the form of the quantum superposition of states demanding indeterminacy in the results of observations, are necessary to provide a sensible physical interpretation and to explain all known experimental facts. These new ideas find their expression through the introduction of a new mathematical formalisms as well as novel axioms and rules of manipulation. The two original formulations of the modern quantum mechanics can be united in a more general and abstract form of the quantum theory, which includes both the wave mechanics and matrix theory as its special cases. This "geometric" formulation requires the complex linear vector space, called the *Hilbert space*, in which vectors represent state functions. Both n-dimensional and $n \to \infty$ spaces are invoked, including the indenumerably infinite case of vectors corresponding to continuous variables. The matrix and wave function theories then appear as corresponding to different choices of the basic vectors in the Hilbert-space, which define the chosen reference frame for concepts and equations of quantum mechanics. This is similar to the relationship between the form of equations in classical physics and the adopted coordinate system in which they are formulated. With the increased elegance and mathematical abstractness of this unifying geometric formulation one also gains a great deal of understanding.

The geometric approach using Dirac's vector notation is the method chosen in the present short presentation of the principles of quantum mechanics. Its relation to the two original formulations will be briefly explored, emphasizing their equivalence in predicting the possible outcomes of experiments and the dynamical equation of motion. Since the wave mechanics appears to be conceptually simpler in chemical applications and directly connecting to the particle–wave dualism, a stronger emphasis will be made on this (nonrelativistic) version of the quantum theory. However, for reasons of convenience, in specific problems covered by the book the matrix theory will also be applied. In this study an emphasis is put on the conceptual developments rather than specific applications. For the solvable problems in quantum mechanics and quantum chemistry the reader is referred to specific textbooks and monographs (e.g., Flügge 1974; Szabo and Ostlund 1982; Atkins 1983; Levine 1983; McQuarrie 1983; Johnson and Pedersen 1986).

References

Atkins PW (1983) Molecular quantum mechanics. Oxford University Press, Oxford

Bohm D (1980) Causality and chance in modern physics. University of Pennsylvania Press, Philadelphia

Born M (1964) Natural philosophy of cause and chance. Dover, New York

Buckley P, Peat FD (eds) (1979) A question of physics: conversations in physics and biology. University of Toronto Press, Toronto

Cohen-Tannoudji C, Diu B, Laloë F (1977) Quantum mechanics. Hermann, Wiley, Paris

Davydov AS (1965) Quantum mechanics. Pergamon, Oxford

Dirac PAM (1967) The principles of quantum mechanics. Clarendon, Oxford
Flügge S (1974) Practical quantum mechanics. Springer, New York
Fock VA (1986) Fundamentals of quantum mechanics. Mir, Moscow
Heisenberg W (1949) The physical principles of the quantum theory. Dover, New York
Heisenberg W (1958) Physics and philosophy: the revolution in modern science. Dover, New York
Johnson CS, Pedersen LG (1986) Problems and solutions in quantum chemistry and physics. Dover, New York
Levine IN (1983) Quantum chemistry. Allyn and Bacon, Boston
McQuarrie DA (1983) Quantum chemistry. University Science Books and Oxford University Press, London
Merzbacher E (1967) Quantum mechanics. Wiley, New York
Messiah A (1961) Quantum mechanics. North Holland, Amsterdam
Penrose R (1989) The emperor's new mind: concerning computers, minds, and the laws of physics. Oxford University Press, New York, p 225
Szabo A, Ostlund NS (1982) Modern quantum chemistry: introduction to advanced electronic structure theory. Macmillan, New York
van der Waerden BL (ed) (1968) Sources of quantum mechanics. Dover, New York
Yourgrau W, van der Merve A (eds) (1979) Perspectives in quantum theory. Dover, New York

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.

R.F. Nalewajski, *Perspectives in Electronic Structure Theory*,
DOI 10.1007/978-3-642-20180-6_2, © Springer-Verlag Berlin Heidelberg 2012

2.1 Geometrical Analogies

The ordinary *three*-dimensional *physical* space R^3 is spanned by the orthonormal *basis* $\{\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\} \equiv \boldsymbol{e}^{(3)}$ (a *row* vector of vector elements), consisting of three unit vectors $\{\boldsymbol{e}_i, i = 1 \equiv x, 2 \equiv y, 3 \equiv z\}$ 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 $\boldsymbol{e}_i \cdot \boldsymbol{e}_j = 0$, and their unit length (normalization), $\boldsymbol{e}_i \cdot \boldsymbol{e}_i = |\boldsymbol{e}_i|^2 \equiv e_i^{\,2} = 1$, can be combined into the *orthonormality* relations expressed in terms of Kronecker's delta,

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{i,j} = \{1,\ \text{for}\ i = j; 0,\ \text{for}\ i \neq j\}, \tag{2.1a}$$

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

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

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

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

$$\boldsymbol{A} = \boldsymbol{A}_x + \boldsymbol{A}_y + \boldsymbol{A}_z \equiv \sum_{i=1}^{3} \boldsymbol{A}_i = \boldsymbol{i}a_x + \boldsymbol{j}a_y + \boldsymbol{k}a_z \equiv \sum_{i=1}^{3} \boldsymbol{e}_i a_i = \boldsymbol{e}^{(3)} \boldsymbol{a}^{(3)\mathrm{T}}, \tag{2.2}$$

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

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

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

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

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

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

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

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

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

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

As seen from this example, the coordinate-resolved expression results directly from placing the identity operator $\hat{\mathrm{P}}(R^3) = 1$ 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,

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

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.,

$$\hat{\mathrm{P}}(\boldsymbol{i})\hat{\mathrm{P}}(\boldsymbol{k}) = \hat{\mathrm{P}}(\boldsymbol{i})\hat{\mathrm{P}}(\boldsymbol{j}) = \hat{\mathrm{P}}(\boldsymbol{j})\hat{\mathrm{P}}(\boldsymbol{k}) = \hat{\mathrm{P}}(\boldsymbol{i},\boldsymbol{j})\hat{\mathrm{P}}(\boldsymbol{k}) = 0. \tag{2.8}$$

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

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

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

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

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

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

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

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

where $\boldsymbol{A}_{\mathcal{P}}$ and $\boldsymbol{A}_{\mathcal{Q}}$ stand for the projections of $\boldsymbol{A}(R^n)$ 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 $\boldsymbol{A} \equiv \vec{A}$ onto another vector $\boldsymbol{B} \equiv \vec{B}$:

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

where $\overleftarrow{B}[\vec{V}]$ 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 $\overleftarrow{B}[\boldsymbol{X}]$ itself, denoted as the "reversed" vector, by specifying the *direction* onto which the argument vector $\boldsymbol{X}$ 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 $\{\boldsymbol{e}_i = \vec{e}_i\}$. They represent the *dual* basis "vectors" $\{\overleftarrow{e}_i[\vec{V}] \equiv \overleftarrow{e}_i\}$ of the scalar product functionals. Indeed, any combination of them also belongs to this dual space, e.g.,

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

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

$$\vec{A} = \sum_{i=1}^{n} a_i\vec{e}_i = \sum_{i=1}^{n} \vec{A}_i \Rightarrow \overleftarrow{A}[\vec{V}] = \sum_{i=1}^{n} a_i\overleftarrow{e}_i[\vec{V}] = \sum_{i=1}^{n} \overleftarrow{A}_i[\vec{V}]. \tag{2.16}$$

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, $C_A\,\boldsymbol{A} + C_B\,\boldsymbol{B} = \boldsymbol{C}$, 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 $|\Psi\rangle$, $|\Phi\rangle, \ldots$, 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:

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

Here, the combination coefficients $\{c(x), c(x'), \ldots\}$ 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 $\{|x\rangle,\ |x'\rangle,\ \ldots\}$. 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, $C_1|\Psi\rangle + C_2|\Psi\rangle = (C_1 + C_2)\,|\Psi\rangle = C|\Psi\rangle \equiv Me^{\mathrm{i}\phi}\ |\Psi\rangle$, 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:

$$\langle\Phi|\Psi\rangle = \Phi[|\Psi\rangle] \equiv \langle\Psi|\Phi\rangle^* = \Psi[|\Phi\rangle]^*. \tag{2.18}$$

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

$$\Phi[|C_1\Psi_1 + C_2\Psi_2\rangle] = C_1\Phi[|\Psi_1\rangle] + C_2\Phi[|\Psi_2\rangle], \tag{2.19}$$

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

$$\langle C_1\Phi_1 + C_2\Phi_2|\Psi\rangle = C_1{}^*\Phi_1[|\Psi\rangle] + C_2{}^*\Phi_2[|\Psi\rangle]. \tag{2.20}$$

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 $\delta_{i,j} = \langle i|j\rangle$, 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:

$$f(x) = \int \delta(x' - x) f(x')\, dx'. \tag{2.21}$$

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):

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

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:

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

Alternatively, one can use any complete, say discrete, set of orthonormal basis functions $\{\chi_i(x)\}$, $\int\chi_i{}^*(x)\chi_j(x)dx = \delta_{i,j}$, 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:

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

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),

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

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

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

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:

$$\delta(z) = \delta(-z), \quad z\delta(z) = 0,$$

$$\delta(az) = |a|^{-1}\delta(z), \quad f(x')\delta(x' - x) = f(x)\delta(x' - x),$$

$$\int \delta(x' - x)\,\delta(x - x'')dx = \delta(x' - x''),$$

$$\delta(x^2 - a^2) = (2|a|)^{-1}[\delta(x - a) + \delta(x + a)]. \tag{2.29}$$

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

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

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* $\hat{A}$:

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

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:

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

$$[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}], \qquad [\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0. \tag{2.34}$$

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 $\Phi[\hat{A}|\Psi\rangle] = \langle\Phi|(\hat{A}|\Psi\rangle)$ 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 $\langle\Phi|$:

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

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:

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

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:

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

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:

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

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:

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

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

$$\hat{\mathrm{P}}_i^{\,2} = |i\rangle\langle i|i\rangle\langle i| = |i\rangle\langle i| = \hat{\mathrm{P}}_i. \tag{2.40}$$

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,

$$\hat{\mathrm{P}} = |\boldsymbol{i}\rangle\langle\boldsymbol{i}| = \sum_i \hat{\mathrm{P}}_i = 1, \tag{2.41a}$$

where $\langle\boldsymbol{i}|$ stands for the column vector of bras associated with the row vector of the basis kets $|\boldsymbol{i}\rangle$, because then $\hat{\mathrm{P}}|\Psi\rangle = |\Psi\rangle$. Similarly, when the complete basis set $|\boldsymbol{x}\rangle = \{|x\rangle\}$ 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),

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

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

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

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

$$\langle\Psi'| = \langle\hat{\mathrm{A}}\Psi| \equiv \langle\Psi|\hat{\mathrm{A}}^\dagger. \tag{2.42}$$

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

$$\langle\Phi|\hat{\mathrm{A}}\Psi\rangle \equiv \langle\Phi|\hat{\mathrm{A}}|\Psi\rangle = \langle\hat{\mathrm{A}}\Psi|\Phi\rangle^* \equiv \langle\Psi|\hat{\mathrm{A}}^\dagger|\Phi\rangle^*. \tag{2.43}$$

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

$$\langle\hat{\mathrm{A}}^\dagger\Phi|\Psi\rangle = \langle\Phi|\hat{\mathrm{A}}|\Psi\rangle = \langle\Phi|\hat{\mathrm{A}}\Psi\rangle. \tag{2.44}$$

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

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

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

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 "†". 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 $\langle\Psi|^\dagger = |\Psi\rangle$. Moreover, the operators $\hat{\mathrm{A}}$ and $\hat{\mathrm{A}}^\dagger$ 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 $\hat{\mathrm{A}}|\Psi\rangle$ gives:

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

Similarly,

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

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 $|\boldsymbol{i}\rangle = \{|i\rangle\}$, $\langle i|j\rangle = \delta_{i,j}$, or the continuous infinity of vectors $|\boldsymbol{x}\rangle = \{|x\rangle\}$, $\langle x|x'\rangle = \delta(x'-x)$, defines the specific *representation* in which both the vectors and operators can be expressed. By convention the basis vectors $|\boldsymbol{i}\rangle$ and $|\boldsymbol{x}\rangle$ are arranged as the *row* vectors. Accordingly, their Hermitian conjugates define the respective column vectors of the bra basis: $|\boldsymbol{i}\rangle^{\dagger} = \langle \boldsymbol{i}|$ and $|\boldsymbol{x}\rangle^{\dagger} = \langle \boldsymbol{x}|$.

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 $|\Psi\rangle$:

$$\begin{aligned} |\Psi\rangle &= \sum_i |i\rangle\,\langle i|\Psi\rangle = \sum_i |i\rangle \mathit{\Psi}_i = |\boldsymbol{i}\rangle\,\langle \boldsymbol{i}|\Psi\rangle \equiv |\boldsymbol{i}\rangle \boldsymbol{\Psi}^{(\boldsymbol{i})}, \\ |\Psi\rangle &= \int |x\rangle\,\langle x|\Psi\rangle\,dx = \int |x\rangle \mathit{\Psi}(x)\,dx \equiv |\boldsymbol{x}\rangle\,\langle \boldsymbol{x}|\Psi\rangle \equiv |\boldsymbol{x}\rangle \boldsymbol{\Psi}^{(\boldsymbol{x})}. \end{aligned} \tag{2.48a}$$

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

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

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

$$\begin{aligned} \langle\Phi| &= \sum_i \langle\Phi|i\rangle\,\langle i| = \sum_i \mathit{\Phi}_i^{*}\langle i| = \langle\Phi|\boldsymbol{i}\rangle\langle \boldsymbol{i}| \equiv \boldsymbol{\Phi}^{(\boldsymbol{i})\dagger}\langle \boldsymbol{i}|, \\ \langle\Phi| &= \int \langle\Phi|x\rangle\,\langle x|\,dx = \int \mathit{\Phi}^{*}(x)\ \langle x|\,dx = \langle\Phi|\boldsymbol{x}\rangle\,\langle \boldsymbol{x}| \equiv \boldsymbol{\Phi}^{(\boldsymbol{x})\dagger}\,\langle \boldsymbol{x}|, \\ \langle\Phi| &= \langle\Phi|\boldsymbol{m}\rangle\,\langle \boldsymbol{m}| \equiv \boldsymbol{\Phi}^{(\boldsymbol{m})\dagger}\langle \boldsymbol{m}| = \langle\Phi|\boldsymbol{\alpha}\rangle\,\langle \boldsymbol{\alpha}| + \langle\Phi|\boldsymbol{y}\rangle\,\langle \boldsymbol{y}| \equiv \boldsymbol{\Phi}^{(\boldsymbol{\alpha})\dagger}\langle \boldsymbol{\alpha}| + \boldsymbol{\Phi}^{(\boldsymbol{y})\dagger}\langle \boldsymbol{y}|. \end{aligned} \tag{2.49}$$

Therefore, the vector components $\boldsymbol{\Phi}^{(\boldsymbol{i})\dagger} = \langle\Phi|\boldsymbol{i}\rangle$, $\boldsymbol{\Phi}^{(\boldsymbol{x})\dagger} = \langle\Phi|\boldsymbol{x}\rangle$ and $[\{\mathit{\Phi}_\alpha{}^{*}\},\ \mathit{\Phi}^{*}(y)] = [\langle\Phi|\boldsymbol{\alpha}\rangle,\ \langle\Phi|\boldsymbol{y}\rangle]$, 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 $\mathit{\Phi}^{*}(x) = \langle\Phi|x\rangle$, can also be regarded as the continuous *row* vector with components $[\mathit{\Phi}^{*}(x),\ \mathit{\Phi}^{*}(x'),\ \ldots]$.

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

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

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

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

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

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:

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

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

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

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, $x = \boldsymbol{r}$, or its momentum, $x = \boldsymbol{p}$, the associated representations $\boldsymbol{\Psi}^{(\boldsymbol{r})} \equiv \Psi(\boldsymbol{r}) = \langle \boldsymbol{r}|\Psi\rangle$ and $\boldsymbol{\Psi}^{(\boldsymbol{p})} \equiv \Psi(\boldsymbol{p}) = \langle \boldsymbol{p}|\Psi\rangle$, called the wave functions in the position ($\boldsymbol{r}$) and momentum ($\boldsymbol{p}$) 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: $\hat{\mathrm{A}}^\dagger = \hat{\mathrm{A}}$. Only such *Hermitian* operators can be associated with the measurable quantities of physics. They satisfy the following scalar product identity [see (2.43)]:

$$\langle \Phi|\hat{\mathrm{A}}|\Psi\rangle = \langle \Psi|\hat{\mathrm{A}}|\Phi\rangle^* = \langle \Phi|\hat{\mathrm{A}}|\Psi\rangle^\dagger. \tag{2.54}$$

The projector $\hat{\mathrm{P}}_\Psi = |\Psi\rangle\langle\Psi|$ provides an example of the Hermitian operator: $\hat{\mathrm{P}}_\Psi^\dagger = \hat{\mathrm{P}}_\Psi$. 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 $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] = 0$, $(\hat{\mathrm{A}}\hat{\mathrm{B}})^\dagger = \hat{\mathrm{B}}^\dagger\hat{\mathrm{A}}^\dagger = \hat{\mathrm{B}}\hat{\mathrm{A}} = \hat{\mathrm{A}}\hat{\mathrm{B}}$.

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

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

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 $\hat{\mathrm{A}}$ 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 $\hat{\mathrm{A}}$, 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 $\hat{\mathrm{P}}_\Psi$ is an example of the quantum-mechanical observable, which exhibits only two eigenvectors. Indeed, for an arbitrary ket $|\Phi\rangle$ the two functions $|1\rangle \equiv \hat{\mathrm{P}}_\Psi|\Phi\rangle$ and $|0\rangle \equiv (1 - \hat{\mathrm{P}}_\Psi)|\Phi\rangle$ can be shown to satisfy the eigenvalue problem of $\hat{\mathrm{P}}_\Psi$:

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

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 $\hat{\mathrm{P}}_\Psi$ 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 $\hat{\mathrm{P}}_\Psi$ is an observable.

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

$$\langle\Psi_i|\hat{\mathrm{A}} = \langle\Psi_i|a_i{}^*. \tag{2.55b}$$

It then follows from the Hermitian character of $\hat{\mathrm{A}}$ 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:

$$0 = (a_i - a_i{}^*)\,\langle\Psi_i|\Psi_i\rangle \Rightarrow a_i = a_i{}^*. \tag{2.57}$$

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

$$\hat{\mathrm{A}}|i_1\rangle = a_i|i_1\rangle,\ \hat{\mathrm{A}}|i_2\rangle = a_i|i_2\rangle,\ \ldots,\ \hat{\mathrm{A}}|i_g\rangle = a_i|i_g\rangle; \tag{2.58}$$

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 $\hat{\mathrm{A}}$ that any combination of such states, say $|\Psi\rangle = C_1|i_1\rangle + C_2|i_2\rangle + \ldots C_g|i_g\rangle$, also represents the eigenvector of $\hat{\mathrm{A}}$ belonging to this eigenvalue:

$$\hat{\mathrm{A}}|\Psi\rangle = C_1\hat{\mathrm{A}}|i_1\rangle + C_2\hat{\mathrm{A}}|i_2\rangle + \cdots + C_g\hat{\mathrm{A}}|i_g\rangle = a_i|\Psi\rangle. \tag{2.59}$$

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, $\hat{\mathrm{A}}|i\rangle = a_i|i\rangle$ and $\langle j|\hat{\mathrm{A}} = \langle j|a_j$, 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,

$$0 = (a_i - a_i)\,\langle j|i\rangle \Rightarrow \langle j|i\rangle = 0. \tag{2.60}$$

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 othonormal, $\langle i_k|i_l\rangle = \delta_{k,l}$, 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 $|\boldsymbol{i}\rangle$, 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)]:

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

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

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

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

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

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

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 $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$, $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] = 0$, 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 $\hat{\mathrm{A}}$ and $\hat{\mathrm{B}}$, which satisfy the simultaneous eigenvalue problems of these two operators:

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

Indeed, when $|a_i\rangle$ denotes the eigenvector of $\hat{\mathrm{A}}, \hat{\mathrm{A}}|a_i\rangle = a_i|a_i\rangle$, and $[\hat{\mathrm{A}}, \hat{\mathrm{B}}] = 0$, applying $\hat{\mathrm{B}}$ to both sides of this eigenvalue equation gives: $\hat{\mathrm{B}}\hat{\mathrm{A}}|a_i\rangle = \hat{\mathrm{A}}(\hat{\mathrm{B}}|a_i\rangle) = a_i(\hat{\mathrm{B}}|a_i\rangle)$. Therefore, $\hat{\mathrm{B}}|a_i\rangle$ is also the eigenvector of $\hat{\mathrm{A}}$ belonging to the same eigenvalue a_i. Hence, for the nondegenerate eigenvalue $a_i, \hat{\mathrm{B}}|a_i\rangle$ 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, $\hat{\mathrm{B}}|a_i\rangle$ is then proportional to $|a_i\rangle$, thus also satisfying the eigenvalue equation of $\hat{\mathrm{B}}$,

$$\hat{\mathrm{B}}|a_i\rangle = b_j|a_i\rangle \Rightarrow |a_i\rangle = |a_i, b_j\rangle. \tag{2.65}$$

For the degenerate eigenvalue $a_i, \hat{\mathrm{B}}|a_i\rangle$ gives a vector belonging to the subspace $\{|a_i\rangle_k\}$ of a_i, so that such eigenvalue subspace of $\hat{\mathrm{A}}$ remains globally invariant under the action of $\hat{\mathrm{B}}$. 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 $\hat{\mathrm{A}}, a_i \neq a_j$, give the vanishing matrix element of the other operator: $\langle a_i|\hat{\mathrm{B}}|a_j\rangle = 0$. This directly follows from their vanishing commutator which implies

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

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

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

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

Therefore:

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

The minimum set of the mutually commuting observables $\{\hat{\mathrm{A}}, \hat{\mathrm{B}}, \ldots, \hat{\mathrm{C}}\}$, 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, ..., c_k) provide the complete specification of the state under consideration: $|\Psi\rangle = |a_i, b_j, \ldots, c_k\rangle$. 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 $\boldsymbol{r} = (x, y, z)$ or momentum $\boldsymbol{p} = (p_x, p_y, p_z)$. These states, $\{|\boldsymbol{r}\rangle\}$ and $\{|\boldsymbol{p}\rangle\}$, respectively, labeled by the respective three continuous coordinates are the eigenvectors of the particle position and momentum operators, $\hat{\mathbf{r}} = (\hat{\mathrm{x}}, \hat{\mathrm{y}}, \hat{\mathrm{z}})$ and $\hat{\mathbf{p}} = (\hat{\mathrm{p}}_x, \hat{\mathrm{p}}_y, \hat{\mathrm{p}}_z)$,

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

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

$$\Psi(\boldsymbol{r}') = \int \langle\boldsymbol{r}'|\boldsymbol{r}\rangle\langle\boldsymbol{r}|\Psi\rangle\, d\boldsymbol{r} = \int u_{\boldsymbol{r}'}{}^*(\boldsymbol{r})\ \Psi(\boldsymbol{r})\ d\boldsymbol{r},$$
$$\Psi(\boldsymbol{p}') = \int \langle\boldsymbol{p}'|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\Psi\rangle\, d\boldsymbol{p} = \int u_{\boldsymbol{p}'}{}^*(\boldsymbol{p})\ \Psi(\boldsymbol{p})\ d\boldsymbol{p}. \tag{2.70}$$

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

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

They also identify the function "coordinates" as the corresponding projections in the function space spanned by the bases $\{u_{\boldsymbol{r}'}(\boldsymbol{r})\}$ and $\{u_{\boldsymbol{p}'}(\boldsymbol{p})\}$, 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:

$$\langle\Psi|\Phi\rangle = \int \langle\Psi|\boldsymbol{r}\rangle\langle\boldsymbol{r}|\Phi\rangle\, d\boldsymbol{r} = \int \Psi^*(\boldsymbol{r})\,\Phi(\boldsymbol{r})\, d\boldsymbol{r}$$
$$= \int \langle\Psi|\boldsymbol{p}\rangle\langle\boldsymbol{p}|\Phi\rangle\, d\boldsymbol{p} = \int \Psi^*(\boldsymbol{p})\,\Phi(\boldsymbol{p})\, d\boldsymbol{p} = 0. \tag{2.71}$$

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

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

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 $\boldsymbol{k} = \boldsymbol{p}/\hbar$,

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

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

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

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

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

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

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

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)]:

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

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

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

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

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

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 $\hat{\mathbf{p}}|\Psi\rangle$,

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

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

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

where the differential vector operator $\nabla_{\boldsymbol{r}} = \boldsymbol{i}\partial/\partial x + \boldsymbol{j}\partial/\partial y + \boldsymbol{k}\partial/\partial z \equiv \partial/\partial\boldsymbol{r}$ stands for the *position* gradient. Therefore, the action of the momentum operator in the position representation amounts to performing the differential operation $\hat{\mathbf{p}}(\boldsymbol{r}) = -i\hbar\nabla_{\boldsymbol{r}}$ on the wave function $\Psi(\boldsymbol{r})$. 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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\langle\boldsymbol{p}|\hat{\mathrm{r}}|\Psi\rangle = i\hbar\nabla_{\boldsymbol{p}}\langle\boldsymbol{p}|\Psi\rangle = \hat{\mathrm{r}}(\boldsymbol{p})\Psi(\boldsymbol{p}). \tag{2.89}$$

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 $\hat{\mathrm{E}} \equiv \hat{\mathrm{H}}$,

$$\hat{\mathrm{H}}|E_n\rangle = E_n|E_n\rangle. \tag{2.90}$$

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

$$\int \langle\xi|\hat{\mathrm{H}}|\xi'\rangle\langle\xi'|E_n\rangle\, d\xi' = E_n\langle\xi|E_n\rangle \tag{2.91}$$

or

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

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

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

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

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

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,

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

their mutual reciprocity relations imply:

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

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:

$$\begin{gathered}u_{\boldsymbol{r}}(\boldsymbol{p}) = \langle\boldsymbol{p}|\boldsymbol{r}\rangle \equiv \mathbf{t}(\boldsymbol{p},\boldsymbol{r}) \quad \text{and} \quad u_{\boldsymbol{p}}(\boldsymbol{r}) = \langle\boldsymbol{r}|\boldsymbol{p}\rangle \equiv \mathbf{t}(\boldsymbol{r},\boldsymbol{p}),\\ \int \mathbf{t}(\boldsymbol{p},\boldsymbol{r})\,\mathbf{t}(\boldsymbol{r},\boldsymbol{p}')d\boldsymbol{r} = \delta(\boldsymbol{p}-\boldsymbol{p}'), \quad \int \mathbf{t}(\boldsymbol{r},\boldsymbol{p})\;\mathbf{t}(\boldsymbol{p},\boldsymbol{r}')d\boldsymbol{p} = \delta(\boldsymbol{r}-\boldsymbol{r}').\end{gathered} \tag{2.98}$$

They also define the unitary kernels,

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

of the integral transformations between the position and momentum representations:

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

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

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

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

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

$$\begin{aligned}\widehat{\mathrm{T}}(\boldsymbol{r},\boldsymbol{p})\widehat{\mathrm{T}}^{\dagger}(\boldsymbol{r},\boldsymbol{p}) &= \widehat{\mathrm{T}}(\boldsymbol{r},\boldsymbol{p})\widehat{\mathrm{T}}(\boldsymbol{p},\boldsymbol{r}) = 1 \quad \text{and}\\ \widehat{\mathrm{T}}(\boldsymbol{p},\boldsymbol{r})\widehat{\mathrm{T}}^{\dagger}(\boldsymbol{p},\boldsymbol{r}) &= \widehat{\mathrm{T}}(\boldsymbol{p},\boldsymbol{r})\widehat{\mathrm{T}}(\boldsymbol{r},\boldsymbol{p}) = 1,\end{aligned} \qquad (2.102)$$

because the double exchange of variables $\boldsymbol{p} \to \boldsymbol{r} \to \boldsymbol{p}$ amounts to identity operation on the wave function $\Psi(\boldsymbol{p})$ and the *double* exchange $\boldsymbol{r} \to \boldsymbol{p} \to \boldsymbol{r}$ operation performed of $\Psi(\boldsymbol{r})$ 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),

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

and using the identity (2.102) one obtains

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

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

$$\widehat{\mathrm{T}}(\boldsymbol{r},\boldsymbol{p})\hat{\mathrm{A}}(\boldsymbol{p},\boldsymbol{p}')\widehat{\mathrm{T}}^{\dagger}(\boldsymbol{r}',\boldsymbol{p}') = \widehat{\mathrm{T}}(\boldsymbol{r},\boldsymbol{p})\hat{\mathrm{A}}(\boldsymbol{p},\boldsymbol{p}')\widehat{\mathrm{T}}^{-1}(\boldsymbol{p}',\boldsymbol{r}') = \hat{\mathrm{A}}(\boldsymbol{r},\boldsymbol{r}') \qquad (2.104)$$

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 $\hat{\mathrm{S}}(\boldsymbol{x}) = \exp[i\hat{\alpha}(\boldsymbol{x})]$. It involves the linear Hermitian operator $\hat{\alpha}(\boldsymbol{x})$, the function of the same list of variables as those of the wave function itself.

This transformation of $\Psi(\boldsymbol{x})$ 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*): $\Psi(\boldsymbol{x}) = \Psi(\boldsymbol{x},\ t = 0)$. 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*:

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

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,

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

This functional attributes to the argument function f the scalar $F = F[f]$. It is defined by the functional density $\mathcal{L}[x, f(x), f'(x), \ldots]$, 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:

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

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: $d^{(1)}f(x_1, x_2, \ldots) = \sum_i (\partial f/\partial x_i)\, dx_i$.

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:

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

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):

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

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

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

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: $d^{(2)}f(x_1, x_2, \ldots) = ½\sum_i\sum_j (\partial^2 f/\partial x_i \partial x_j)\, dx_i\, dx_j$.

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:

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

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 $f(x) = f[g; x]$,

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

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

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

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

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

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 $F[f,\, g] = const.$ implies the associated functional relation between them, e.g., $g = g[f]_F$. The vanishing first differential,

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

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

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

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:

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

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

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

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

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

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

then gives:

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

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):

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

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

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

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: $f^{(i)}(x) = d^i f(x)/dx^i$, $i = 1, 2, \ldots, n$:

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

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

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

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

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

Similarly, for

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

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

Chapter 3
Basic Concepts and Axioms

Abstract The postulates of quantum mechanics are formulated using the mathematical tools of the preceding chapter. First, the axioms related to the quantum kinematics are summarized, dealing with a variety and physical meaning of quantum states at the specified time. They include alternative definitions and interpretations of the wave functions of microobjects as amplitudes of the particle probability distributions in the configuration or momentum spaces. As an illustrative example the electron densities are then discussed. The superposition principle is formulated, and the symmetry implications of indistinguishability of identical particles in quantum mechanics are examined. The links between the quantum states and outcomes of the physical measurements are then surveyed and the physical observables are attributed to quantum mechanical operators, linear and Hermitian, and their specific forms in the position and momentum representations are introduced. The eigenvalues of the quantum mechanical operator are postulated to determine a variety of all possible results of a single experiment measuring the physical property the operator represents, while the operator expectation value represents the average value of this quantity in a very large number of repeated measurements performed on the system in the same quantum state. The eigenstates of the quantum mechanical operator are shown to correspond to the sharply specified value of the physical property under consideration, while other quantum states exhibit distributions of its allowed eigenvalues. The statistical mixtures of quantum states are defined in terms of the density operator and the ensemble averages of physical observables in such mixed states are examined. The simultaneous sharp measurement of several physical observables is linked to the mutual commutation of their operators and the quantum mechanical formulation of the general Principle of Indeterminacy is given. Properties of the electron angular momentum and spin operators are examined.

In the dynamical development, the pictures of time evolution in quantum mechanics are introduced through the alternative time-dependent unitary transformations of the state vectors/operators. The Schrödinger equation is explored in some detail, with the emphasis placed upon the stationary states, time dependence of expectation values, conservation laws, the probability current, and continuity equation. The

R.F. Nalewajski, *Perspectives in Electronic Structure Theory*,
DOI 10.1007/978-3-642-20180-6_3,

correspondence between the quantum and classical dynamics is established through the Ehrenfest principle. Finally, the rudiments of the Heisenberg and interaction pictures of quantum dynamics are briefly summarized.

3.1 *N*-Electron Wave Functions and Their Probabilistic Interpretation

In the canonical formulation of classical mechanics, the system dynamics is formulated in terms of the Hamilton function $E = H(\boldsymbol{Q}, \boldsymbol{P})$ expressing the system energy E in terms of its generalized coordinates $\boldsymbol{Q} = \{Q_\alpha\}$ and their conjugated momenta $\boldsymbol{P} = \{P_\alpha\}$, $\alpha = 1, 2, \ldots, f$, with f denoting the system number of dynamical *degrees of freedom*. Together these conjugate dynamical variables uniquely specify the system mechanical state. Indeed, the knowledge of $\boldsymbol{Q}(t)$ and $\boldsymbol{P}(t)$ at the specified time $t = t_0$ allows one to determine the exact time evolution of these state parameters, via the Hamilton equations of motion:

$$\dot{Q}_\alpha = \frac{dQ_\alpha}{dt} = \frac{\partial H}{\partial P_\alpha}, \quad \dot{P}_\alpha = \frac{dP_\alpha}{dt} = -\frac{\partial H}{\partial Q_\alpha}, \quad \alpha = 1, 2, \ldots, f. \tag{3.1}$$

Since these are the *first*-order differential equations, their solutions $\{\boldsymbol{Q}(t), \boldsymbol{P}(t)\}$ are uniquely specified when the values of these classical state variables are fixed at $t = t_0$. Thus, knowing the state $\{\boldsymbol{Q}(t_0), \boldsymbol{P}(t_0)\}$ of the classical system at this time, one can in principle predict with certainty the system mechanical state at $t \neq t_0$, i.e., precisely determine the outcome of any measurement at an earlier or later stage of the system time evolution.

As we have argued in Chap. 1, this classical specification of the mechanical state is inapplicable in the quantum theory, due to the simultaneous indeterminacy of coordinates and momenta of microobjects (the Heisenberg principle). Indeed, since the state variables must be precisely specified, either the position coordinates or the components of the canonically conjugated momenta of the system particles should be used to uniquely characterize its quantum state. Therefore, at the given time t, which in the simplest (nonrelativistic) formulation of the quantum theory plays the role of a parameter, the quantum state corresponding to the state vector $|\Psi(t)\rangle$ is represented by the wave functions in either the position or momentum representations,

$$\Psi(\boldsymbol{Q}; t) = \langle \boldsymbol{Q} | \Psi(t) \rangle \quad \text{or} \quad \Psi(\boldsymbol{P}; t) = \langle \boldsymbol{P} | \Psi(t) \rangle; \tag{3.2a}$$

here, the representation basis sets $\{|\boldsymbol{Q}\rangle\}$ and $\{|\boldsymbol{P}\rangle\}$ correspond to the position and momentum eigenstates, respectively, in which these molecular variables are known precisely. For quantum particles these classical state "coordinates" should be also supplemented with all nonclassical, internal (*spin*) degrees of freedom for each particle, $\boldsymbol{\Sigma} = \{\Sigma_k\}$. Therefore, the full specification of the mechanical state of the given quantum system, in either the *position-spin* or *momentum-spin* representations, is embodied in the corresponding wave functions:

$$\Psi(\boldsymbol{Q},\boldsymbol{\Sigma};\, t) = \langle \boldsymbol{Q},\boldsymbol{\Sigma}|\Psi(t)\rangle \quad \text{or} \quad \Psi(\boldsymbol{P},\boldsymbol{\Sigma};\, t) = \langle \boldsymbol{P},\boldsymbol{\Sigma}|\Psi(t)\rangle. \tag{3.2b}$$

Since the theoretical description of the electronic structure of molecules is the main objective of this book, in what follows we shall focus on a general (atomic or molecular) *N*-electron system, with the list of the (coordinate/momenta)-spin variables in the Cartesian coordinates:

$$\begin{aligned} |\boldsymbol{Q},\boldsymbol{\Sigma}\rangle &\equiv |\boldsymbol{q}^N\rangle = |\{\boldsymbol{q}_k\}\rangle \equiv |\mathcal{Q}^N\rangle, \quad \boldsymbol{q}_k = (\boldsymbol{r}_k,\, \sigma_k), \\ |\boldsymbol{P},\boldsymbol{\Sigma}\rangle &\equiv |\boldsymbol{u}^N\rangle = |\{\boldsymbol{u}_k\}\rangle \equiv |\mathcal{P}^N\rangle, \quad \boldsymbol{u}_k = (\boldsymbol{p}_k,\, \sigma_k), \quad k = 1,2,\ldots,N; \end{aligned} \tag{3.3}$$

here $\boldsymbol{r}_k = (x_k, y_k, z_k), \boldsymbol{p}_k = (p_{x_k}, p_{y_k}, p_{z_k})$ and σ_k, respectively, denote the continuous position, momentum vectors of *k*th electron and its discrete spin orientation variable $\sigma_k \in (+½, -½)$ (see Fig. 1.2).

Therefore, the vector space of the *N*-electron system is spanned by all basis vectors in either the position $\{|\mathcal{Q}^N\rangle\}$ or momentum $\{|\mathcal{P}^N\rangle\}$ representations. In what follows we shall call this vector space the *molecular Hilbert space*. The specific state of such an *N*-electron system in time t will be denoted by the ket $|\Psi^N(t)\rangle$. Since each basis vector is specified by the three position/momentum coordinates and one spin variable for each electron, the overall dimensionality of either the *position*-spin or *momentum*-spin spaces is $4N$. The basis vectors $|\mathcal{Q}^N\rangle$ and $|\mathcal{P}^N\rangle$ are then identified by corresponding points in these configurational spaces. It should be observed that in the classical mechanics the system state was uniquely specified at the given time by selecting the point in the $6N$-dimensional position–momentum *phase space* of N particles.

Moreover, the corresponding position-spin or momentum-spin data for the atomic nuclei are also required for the complete specification of the molecular state. However, as we shall argue in Part II of this book, due to a huge difference in masses between the (light) electrons and (heavy) nuclei, the dynamics of the former can be to a good approximation described by examining their fast movements in the effective potential generated by the "frozen" nuclear framework, with the fixed positions of nuclei playing the role of parameters in the electronic structure theory. In this *adiabatic approximation* of Born and Oppenheimer the nuclei, sources of the external potential in which electrons move, thus determine the assumed molecular geometry.

After these short preliminaries, we are now in a position to formulate the important postulate of quantum mechanics, due to Born, which provides the physical interpretation of the wave functions of (3.2a) and (3.2b):

Postulate I: The (normalized) quantum mechanical state of the molecular system containing *N*-electrons in time *t*, $\langle\Psi^N(t)|\Psi^N(t)\rangle \equiv \left\|\Psi^N(t)\right\|^2 = 1,$ where $\left\|\Psi^N(t)\right\|$

stands for the norm ("length") of the state vector, is uniquely specified by the orientation of the state-vector $|\Psi^N(t)\rangle$ in the molecular Hilbert space or equivalently by its equivalent representations (wave functions) in the position or momentum basis sets, respectively,

$$\Psi(\mathcal{Q}^N;\, t) = \langle \mathcal{Q}^N | \Psi^N(t) \rangle \quad \text{or} \quad \Psi(\mathcal{P}^N; t) = \langle \mathcal{P}^N | \Psi^N(t) \rangle. \tag{3.2c}$$

These in general complex-valued functions determine the probability amplitudes of simultaneously observing at this time the specified positions/momenta and spin orientations of all N electrons, with the corresponding probability densities being determined by the squares of the wave function moduli:

$$\begin{aligned} p(\mathcal{Q}^N; t) &= \left|\langle \mathcal{Q}^N | \Psi^N(t) \rangle\right|^2 = \left|\Psi(\mathcal{Q}^N; t)\right|^2 \equiv P[\mathcal{Q}^N | \Psi^N(t)], \\ \int p(\mathcal{Q}^N; t)\, d\mathcal{Q}^N &= \int \langle \Psi^N(t) | \mathcal{Q}^N \rangle \langle \mathcal{Q}^N | \Psi^N(t) \rangle\, d\mathcal{Q}^N = \langle \Psi^N(t)\, | \Psi^N(t) \rangle = 1; \\ \pi(\mathcal{P}^N; t) &= \left|\langle \mathcal{P}^N | \Psi^N(t) \rangle\right|^2 = \left|\Psi(\mathcal{P}^N; t)\right|^2 \equiv P[\mathcal{P}^N | \Psi^N(t)], \\ \int \pi(\mathcal{P}^N; t)\, d\mathcal{P}^N &= \int \langle \Psi^N(t) | \mathcal{P}^N \rangle \langle \mathcal{P}^N | \Psi^N(t) \rangle\, d\mathcal{P}^N = \langle \Psi^N(t)\, | \Psi^N(t) \rangle = 1. \end{aligned} \tag{3.4}$$

Here, the generalized "integration" symbol $\int d\mathcal{Q}^N$ actually denotes the definite integrations over the position coordinates and summations over the spin variables of all electrons:

$$\int d\mathcal{Q}^N \equiv \int d\boldsymbol{q}_1 \ldots d\boldsymbol{q}_N \equiv \int d\boldsymbol{r}_1 \ldots d\boldsymbol{r}_N \sum_{\sigma_1} \cdots \sum_{\sigma_N}, \tag{3.5a}$$

The related operation in the momentum-spin space similarly reads:

$$\int d\mathcal{P}^N \equiv \int d\boldsymbol{u}_1 \ldots d\boldsymbol{u}_N \equiv \int d\boldsymbol{p}_1 \ldots d\boldsymbol{p}_N \sum_{\sigma_1} \cdots \sum_{\sigma_N}. \tag{3.5b}$$

In fact, the normalization conditions of this postulate, for the position-spin and momentum-spin probability densities $p(\mathcal{Q}^N; t)$ and $\pi(\mathcal{P}^N; t)$, respectively, express the unit probability of the sure event that at the specified time t all electrons are located somewhere in the physical or momentum spaces, and assume one of its allowed spin orientations. We have also indicated in (3.4) that the probability densities $P[\mathcal{Q}^N | \Psi^N(t)]$ and $P[\mathcal{P}^N | \Psi^N(t)]$ of the particle positions and momenta, respectively, are *conditional* upon the specified quantum state. Indeed, these densities represent the conditional probabilities of observing the basis set events corresponding to the wave function arguments $\mathcal{Q}^N$ or $\mathcal{P}^N$ (*variables*), given the molecular state $|\Psi^N(t)\rangle$ (the *parameter*): $p(\mathcal{Q}^N; t) = P[\mathcal{Q}^N | \Psi^N(t)]$ and $\pi(\mathcal{P}^N; t) = P[\mathcal{P}^N | \Psi^N(t)]$. The normalization relations thus involve the integrations/summations of these conditional