

THE QUANTUM ATOM

There can be no doubt but that in quantum mechanics one has the complete solution to the problems of chemistry.

G. N. Lewis (1933)

5.1 Chemistry and quantum mechanics

5.1.1 From Lewis to quantum mechanical models

Lewis introduced the concept of the electron pair into chemistry in 1916. In this paper (1916) he was able to rationalize the known geometries of carbon compounds by replacing the prevailing notion of a cubical arrangement of an octet of electrons by a tetrahedral arrangement of four pairs of electrons. What particularly served to later convince Lewis (1933) of the potential of the new quantum mechanics, as evidenced by the opening quotation, was its ability to account for observed molecular geometries and so provide a theoretical basis for his electron pair model. The early development of the valence bond theory of directed valence by Slater (1931) and Pauling (1931, 1960) provided the first quantum mechanical rationalization of molecular geometry. The valence bond approximation to the wave function, as expressed in terms of the overlap of hybridized atomic orbitals to describe the pairing of electrons, coupled with the notion of electron spin, was the natural theoretical extension of the pair concept of Lewis. The hybrid orbitals were spatially localized in directions coincident or nearly coincident with observed molecular geometries and this gave theoretical substance to the notion of picturing a bond as a spatially localized pair of electrons. Pauling's emphasis of the use of ionic-covalent resonance structures to account for the differing character of bonded interactions was in turn a restatement of Lewis's idea to represent the bonding between a pair of atoms in terms of tautomeric extremes of electron-sharing possibilities. It was a principal claim of Lewis's first paper (1916) that 'According to the theory which I am now presenting, it is not necessary to consider the two extreme types of chemical combination, corresponding to the very polar and the very nonpolar compounds, as different in kind, but only as different in degree.'

In the description of nature afforded by quantum mechanics, one classifies and characterizes the state of a total system in terms of the eigenvalues of a set of commuting observables acting on an element of the Hilbert space, the state vector. Molecular orbital theory in its canonical representation as originally

developed by Mulliken (1928*a-c*, 1932*a*, 1935) and Hund (1933, 1937) was, and is, the theory of *electronic structure* for the prediction and classification of quantum states of many-electron systems. Molecular orbital theory provides a model of the electronic structure of a many-electron system in terms of a set of coupled one-electron states or orbitals which greatly facilitates the prediction and classification of the electronic states of many-electron atoms and molecules. In a molecular system, the one-electron states themselves can be classified as bonding, non-bonding, or antibonding (Mulliken 1932*b*, 1939) and this classification correlates with the observed changes in bond lengths and dissociation energies of the new states generated as a consequence of a change in the occupation of a particular orbital. The delocalized nature of the canonical set of molecular orbitals provided a useful model for understanding the properties of unsaturated systems (Coulson and Longuet-Higgins 1947*a,b*, 1948*a-c*), but was unsatisfactory in providing models of directed bonds. Lennard-Jones (1949*a,b*, 1952) defined equivalent orbitals (which later evolved into localized orbitals) to provide the molecular orbital equivalent of directed valence orbitals and it was only after their introduction that molecular orbital theory found widespread application to structural problems in chemistry. Since the advent of suitable computational methods and machines in the 1960s, molecular orbitals calculated within the self-consistent field approximation and expanded in terms of a finite set of basis functions (Roothaan 1951; Hall 1951) have provided the principal means of obtaining approximate solutions to the wave functions of many-electron systems of chemical interest.

The orbital model has, however, been extended beyond its intended use of predicting and providing an understanding of the electronic structure of a system, by associating the forms of individual orbitals with the assumed spatially localized pairs of bonded or non-bonded electrons, and by attempting to define atomic properties in terms of coefficients of atomic-centred basis functions appearing in the expansion of molecular orbitals. These steps are admittedly arbitrary, as are attempts to define atoms through a partitioning of the Hamiltonian operator. This latter step violates the indistinguishability of the electrons from the outset.

One might imagine that, with the advent of quantum mechanics and its application to chemistry, Dalton's atomic theory would have been reinforced. This has not happened. Quantum mechanics has been shown to account for the properties of isolated atoms and for the total properties of a molecular system. The increased understanding that would result from the discovery of a firm theoretical basis for Dalton's theory has not been obtained because of the lack of a quantum definition of an atom in a molecule. This is not to say that the concepts of atoms and bonds do not appear in the quantum mechanical treatments of chemical systems. They do, but in the reverse manner to that described above. Rather than finding its quantum basis, the

atomic and bond concepts are built into an approximate theory to model a real system. Thus, in valence bond theory a molecular wave function is approximated in terms of products of atomic wave functions in appropriate valence states. Moffitt's (1951) 'atoms in molecules approach' to the calculation of molecular binding energies is a further example, as is the success of using atomic functions as a basis for the expansion of molecular orbitals. Properties of an ionic crystal strongly suggest that it is best regarded as a collection of interacting ions. Hence, one may successfully approximate the wave functions for such systems using free-ion wave functions with adjustments for the overlap of non-orthogonal functions as illustrated by the work of Löwdin (1948, 1956) in an early theoretical calculation of the properties of ionic crystals.

In 1929 Dirac wrote: 'The underlying physical laws necessary for the mathematical theory of . . . the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.' The stumbling block to obtaining useful approximate solutions to the equations of quantum mechanics for systems of chemical interest has been largely overcome. Overcoming this hurdle, however, has not in itself led to a mathematical theory of the whole of chemistry, at least not to one which is expressible in the language of chemistry. Given a state function, or a good approximation to it, how does one obtain from it a description of a system's properties in terms of its atoms and its bonds which, from a chemist's point of view, summarizes in a concise manner its important chemical properties? The varying successes of the quantum mechanical models of atoms and bonds have served to demonstrate the soundness of the atomic concept but, by building in the idea of atoms, they have not furthered Dalton's theory of atoms in molecules. The increased understanding that would result from the discovery of a firm theoretical basis for Dalton's theory requires a quantum definition of an atom in a molecule.

5.1.2 The role of the charge density in defining structure

* [Atoms and bonds have meaning in real space and are a reflection of the structure present in real space. This structure is not reflected in the properties of the infinite-dimensional Hilbert space of the state function.] [Instead, the physical basis for molecular structure should reside in the quantum mechanical function which provides a description of a system as it exists in real space. This function is a system's distribution of charge] as defined in eqn (1.3) for a stationary state or as defined now for the general time-dependent case,

$$\rho(\mathbf{r}, \mathbf{X}, t) = N \sum (\text{spins}) \int \left\{ \prod_{j \neq i} d\tau_j \right\} \Psi^*(\mathbf{x}, \mathbf{X}, t) \Psi(\mathbf{x}, \mathbf{X}, t) \quad (5.1)$$

where Ψ is a properly antisymmetrized solution to the general time-dependent Schrödinger equation,

$$i\hbar \partial\Psi/\partial t = \hat{H}\Psi \quad \text{and} \quad -i\hbar \partial\Psi^*/\partial t = \hat{H}\Psi^*. \quad (5.2) \quad *$$

As in Chapter 1, \mathbf{x} denotes the collection of electronic space and spin coordinates, \mathbf{X} the nuclear coordinates, and \mathbf{r} the space coordinates (x, y, z) of a single electron and $d\tau_j = dx_j dy_j dz_j$. [It should be noted that the electronic charge density contains the information needed to determine the distribution of nuclear charge as well, and thus it determines the total distribution of charge. When $\mathbf{r} = \mathbf{X}_\alpha$, a nuclear position coordinate, ρ exhibits a cusp and through the cusp condition, eqn (E2.2), it determines the nuclear charge Z_α . In addition, the positions of the nuclei are evident in the topology of the electronic charge density, as each nuclear cusp behaves like a local maximum in the charge distribution with $\rho(\mathbf{X}_\alpha) \approx Z_\alpha^3/2$.]

Molecular orbital theory has played the central role in the definition and understanding of problems of electronic structure. The charge density plays the corresponding role in the definition and understanding of the concepts associated with molecular structure. The previous chapters have shown that atoms, bonds, and structure are indeed consequences of the dominant topological property exhibited by a molecular charge distribution. What remains to be done is to demonstrate that the topological atom and its properties have a basis in quantum mechanics.

5.2 Need for a quantum definition of an atom

5.2.1 Observational basis for a quantum atom

A theory is only justified by its ability to account for observed behaviour. It is important, therefore, to note that the theory of atoms in molecules is a result of observations made on the properties of the charge density. These observations give rise to the realization that a quantum mechanical description of the properties of the topological atom is not only possible but is also necessary, for the observations are explicable only if the virial theorem applies to an atom in a molecule. The original observations are among the most important of the properties exhibited by the atoms of theory (Bader and Beddall 1972). For this reason and for the purpose of emphasizing the observational basis of the theory, these original observations are now summarized. They provide an introduction to the consequences of a quantum mechanical description of an atom in a molecule.

Figure 5.1 displays, in the form of contour plots, the ground-state charge distributions of LiF, LiO, and LiH, each at its equilibrium internuclear separation. Superimposed on each of these plots is the intersection of the

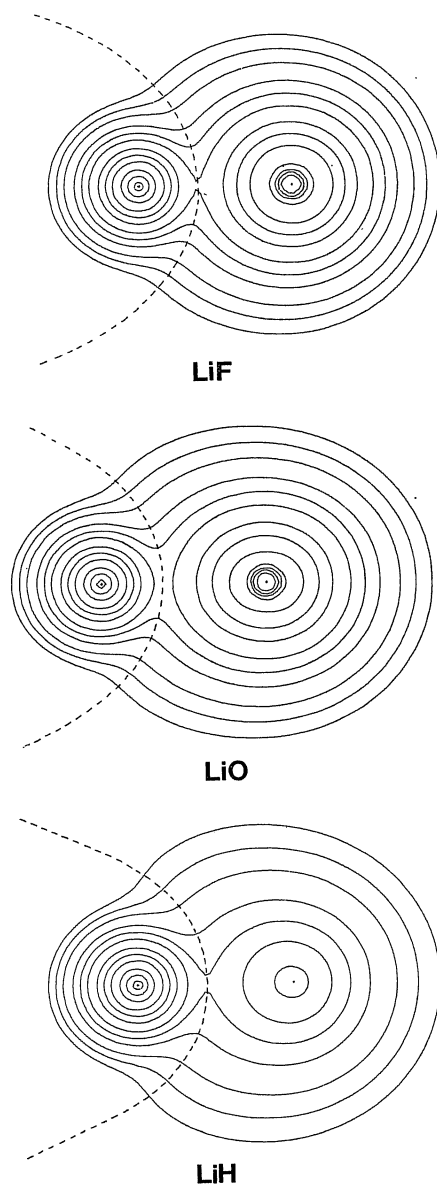


FIG. 5.1. Contour plots of the ground-state molecular charge distribution of LiF, LiO, and LiH. The intersection of the interatomic surface with the plane shown in the diagram is indicated by a dashed line.

zero-flux interatomic surface, as defined in eqn (2.9), with the plane of the diagram. The distribution of charge within the Li atom so defined is seen to be remarkably similar in all three of these molecules in spite of the very different natures of the neighbouring atom. The net charge of the Li atom (as determined by an integration of $\rho(\mathbf{r})$ over the atomic basin to obtain its average number of electrons followed by its subtraction from the nuclear charge) is very nearly the same for all three molecules (see Table 5.1), the observed variation following the trend anticipated on the basis of decreasing electronegativity of the ligand $F > O > H$. This near constancy in the charge distribution of Li correctly reflects the properties assigned to Li on the basis of the chemistry observed for compounds in which Li is bonded to a more electronegative element—those corresponding to a relatively small, singly-charged positive ion with a tightly bound distribution of electronic charge.

Figure 5.1 illustrates an elementary but important and general observation: the choice of the zero-flux surface for defining an atom maximizes the possibility of assigning an atomic identity to a given mononuclear region of a charge distribution. It is clear that any other choice of a partitioning surface would either include a portion of the neighbouring atom that is very different in all three cases, or omit a portion of the Li atom charge density that changes by only small amounts through the series of molecules. Since the partitioning must exhaust the space of a system, the latter possibility would assign the very similar omitted portions to the very different neighbouring atoms.

Coupled with the observation of near constancy in the charge distributions of the Li atoms in this series of molecules was the further observation that the kinetic energy density distributions exhibited a corresponding degree of constancy. (As demonstrated later, while there is no unique definition of a kinetic energy density, all definitions lead to the same average value for the kinetic energy when integrated over the basin of an atom up to its surface of zero flux. The average kinetic energy of an atom is a well-defined quantity as

Table 5.1
*Some properties of bound Li atoms**

Molecule	Net charge of Li atom	Average electronic kinetic energy of Li atom (au)
LiF ($X^1\Sigma^+$)	+ 0.937	7.354
LiO ($X^2\Pi$)	+ 0.932	7.356
LiH ($X^1\Sigma^+$)	+ 0.913	7.368

*From Bader and Beddall (1972); calculated from state functions close to the Hartree-Fock limit.

a consequence of the zero-flux surface condition.) Thus, it is an observation in this set of molecules and subsequently in others, that a constancy in the distribution of charge for a topologically defined atom leads to a corresponding degree of constancy in the average kinetic energy of the atom, the total spread in values for the Li atoms in the present examples being ~ 32 kJ/mol (see Table 5.1).

* [According to the virial theorem, the total energy for a system with inverse square forces is equal to minus the average kinetic energy. If one postulates the existence of an atomic statement of the virial theorem, then the above observation predicts that, when the charge distribution of an atom is identical in two different systems, the atom will contribute identical amounts to the total energies in both systems. *Equally important is the more general conclusion that the properties of such a topologically defined atom are directly determined by its charge distribution, the properties changing in direct response to changes in the charge density of the atom.*

The postulation of an atomic virial theorem for the topologically defined atoms leads to a number of important conclusions (Bader and Beddall 1972).

1. The total energy of a molecule is expressible as a sum of atomic energies.
2. The average potential energy of an atom is defined as the average of the virial of the forces exerted on it—as demonstrated later, this is the only non-arbitrary way of partitioning potential energies of interaction between systems.
3. A relationship must exist between the distribution of charge and the virial of the total force exerted on each element of the charge density, the virial field. The form of the atom must be independent of the individual contributions to the forces exerted on it, since these forces change radically between any pair of systems. The charge density of an atom must respond instead only to the sum of the local forces it experiences. The virials of the individual contributions to the forces exerted on the density of the Li ion by the hydride ion differ by thousands of kJ/mol from those exerted by the fluoride ion (the one-electron potential energies differ by 14×10^3 kJ/mol, for example) but the net force in each case corresponds to that emanating from a polarized singly-charged negative ion. The constancy in the charge distribution of Li must be viewed as remarkable when one contemplates these very different contributions to the force exerted on it by its neighbours. One field, however, changes by little—the virial of the total Ehrenfest force exerted on the electrons. This relationship between ρ and the virial field is the basis of the essential observation that, if the distribution of charge for an atom is identical in two different systems, then the atom will contribute identical amounts to the total energy in both systems.]

It will be shown that each atom in a system makes an additive contribution to the average value of every system property. This is the principle underlying the cornerstone of chemistry—that atoms and functional groupings of atoms make recognizable contributions to the total properties of a system. In practice, we recognize a group and predict its effect upon the static and reactive properties of a system in terms of a set of properties assigned to the group. In those limiting situations wherein a group is essentially the same in two different systems, one obtains a so-called additivity scheme for the total properties, for in this case the atomic contributions, as well as being additive over each system, are transferable between the systems.

Examples are given later of near perfect transferability of groupings of atoms, cases where the group energies change by less than 4 kJ/mol. It is at this limit that one can determine the properties of atoms in molecules experimentally, as corresponding additive contributions to heats of formation. Such transferability is, of course, the exception rather than the rule but these limiting examples provide the touchstone for obtaining an understanding of the properties of the chemical atom—that these properties change in direct response to the extent of the changes in its form in real space. When this form remains constant, so too do the contributions which the atom makes to the properties of the total system in which it is found. *

[It is upon the strength of these observations that the theory of atoms in molecules rests, for they form the basis for the identification of the topological atom with the chemical atom of a molecular system. This chapter demonstrates that the postulated existence of the atomic virial theorem is confirmed, as is the complete quantum description of the properties of the topological atom. The development of the quantum mechanics of a subsystem requires careful consideration of the properties of observables and of the derivation of their equations of motion.] These fundamental aspects of the quantum description of a total system are considered next, with an emphasis on the changes anticipated in the description of a subsystem.

5.2.2 Observables and their properties for a total system

[It is a postulate of quantum mechanics that everything that can be known about a system is contained in the state function Ψ . The value of a physical property is obtained from the state function through the action of a corresponding operator on Ψ . Thus quantum mechanics is concerned with observables, the name given to the linear, Hermitian operators whose action on Ψ yields the values of a system's properties. The Hermitian character of an operator is illustrated for the Hamiltonian \hat{H}] an illustration which introduces the quantum mechanical current density and demonstrates its role in the description of the properties of subsystem averages of operators. If

$\Psi^*\Psi d\tau$ is to represent a probability, then the norm, N , of the state function must remain constant in time. Using Schrödinger's equation and its complex conjugate (eqn (5.2)), one has

$$\partial|\Psi|^2/\partial t = \Psi^*\dot{\Psi} + \dot{\Psi}^*\Psi = -(i/\hbar)\{\Psi^*(\hat{H}\Psi) - (\hat{H}\Psi)^*\Psi\} \quad (5.3)$$

* where $\dot{\Psi} = d\Psi/dt$. Integration of eqn (5.3) over all configuration space yields

$$dN/dt = -(i/\hbar) \int \{\Psi^*(\hat{H}\Psi) - (\hat{H}\Psi)^*\Psi\} d\tau. \quad (5.4)$$

In order that the norm remain constant in time it is necessary and sufficient that

HERMETIANO $\rightarrow \int \Psi^*(\hat{H}\Psi) d\tau = \int (\hat{H}\Psi)^*\Psi d\tau. \quad (5.5)$

An operator with the property exhibited in eqn (5.5) is said to be Hermitian if it satisfies this equation for all functions Ψ defined in the function space in which the operator is defined. The mathematical requirement for Hermiticity of \hat{H} expressed in eqn (5.5) places a corresponding physical requirement on the system—that there be a zero flux in the vector current through the surface S bounding the system. To illustrate this and other properties of the total system we shall assume, without loss of generality, a form for \hat{H} corresponding to a single particle moving under the influence of a scalar potential $\hat{V}(\mathbf{r})$

$$\hat{H} = -(\hbar^2/2m)\nabla^2 + \hat{V}(\mathbf{r}) = \hat{p}^2/2m + \hat{V}(\mathbf{r}). \quad (5.6)$$

It is worthwhile mentioning at this point that all properties of a subsystem defined in real space, including its energy, necessarily require the definition of corresponding three-dimensional density distribution functions. Thus, all the properties of an atom in a molecule are determined by averages over effective single-particle densities or 'dressed operators' and the one-electron picture is an appropriate one.

The potential energy operator in \hat{H} is a real quantity which does not involve derivatives and, in this case, eqn (5.5) can be rewritten as

$$\int \{\Psi^*(\nabla^2\Psi) - (\nabla^2\Psi)^*\Psi\} d\tau = 0. \quad (5.7)$$

The integral in eqn (5.7), like many to come, can be transformed into a surface integral using the three-dimensional analogue of integration by parts. In one-dimension, one has

$$\int (\phi^*\partial^2\phi/\partial x^2) dx = (\phi^*\partial\phi/\partial x)|_{x_1}^{x_2} - \int (\partial\phi^*/\partial x)(\partial\phi/\partial x) dx \quad (5.8)$$

* where x_1 and x_2 represent the limits of integration. In the extension of this result to three dimensions, the contributions arising from these limits combine to yield an integral over the surface S bounding the system. When applied to the complex conjugate term in eqn (5.7) as well, one obtains Green's theorem

$$\int \{\phi^*\nabla^2\phi - \phi\nabla^2\phi^*\} d\tau = \oint dS(\phi^*\nabla\phi - \phi\nabla\phi^*) \cdot \mathbf{n} \quad (5.9)$$

GAUSS THEOREM

$$\iiint_V dV \nabla \cdot \mathbf{V} = \iint_S dS \mathbf{V} \cdot \mathbf{n}$$

A INTERFERÊNCIA QUANTICA
EM UM SISTEMA QUANTICO
NA INTERFERÊNCIA QUANTICA
A DIFERENÇA DE FASE ENTRE OS
VETORES DE CAMPO

where \mathbf{n} is the vector normal to the surface S . Use of eqn (5.9) to re-express eqn (5.7) yields

$$\int \{\Psi^*\nabla^2\Psi - (\nabla^2\Psi)^*\Psi\} d\tau = \oint dS(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*) \cdot \mathbf{n} = (2mi/\hbar) \oint dS \mathbf{j} \cdot \mathbf{n} \quad (5.10)$$

where the quantum mechanical vector current density \mathbf{j} is defined as

$$\mathbf{j} = (\hbar/2mi)(\Psi^*\nabla\Psi - \Psi\nabla\Psi^*). \quad (5.11)$$

Eqn (5.10) can be expressed in terms of the Hamiltonian operator to yield

$$\int \Psi^*(\hat{H}\Psi) d\tau - \int (\hat{H}\Psi)^*\Psi d\tau = -i\hbar \oint dS \mathbf{j} \cdot \mathbf{n}. \quad (5.12a)$$

For a bound system, Ψ is square integrable and, hence, Ψ and its derivatives vanish on all elements $dS = dS\mathbf{n}$ of the surface when the surface is removed to infinity. Thus the right-hand side of eqn (5.10) vanishes for the total system with boundaries at infinity and \hat{H} is Hermitian. However, if the integration is limited to a subsystem Ω bounded by a surface $S(\Omega)$, part or all of which occurs for finite values of the integration variables, then one has

$$\int_{\Omega} \Psi^*(\hat{H}\Psi) d\tau - \int_{\Omega} (\hat{H}\Psi)^*\Psi d\tau = -i\hbar \oint dS(\Omega) \mathbf{j} \cdot \mathbf{n} \quad (5.12b)$$

where integration over a subsystem is indicated by the subscript Ω on the integral sign. In this case the flux in the current density through the surface will not, in general, vanish and one cannot assume that the Hamiltonian integral is equal to its Hermitian conjugate as is true for the total system.

Finally we note that eqn (5.3) can be re-expressed as

$$d\rho/dt + \nabla \cdot \mathbf{j} = 0, \quad (5.13)$$

which is a statement of the conservation law for a fluid. In its integrated form

$$dN(\Omega)/dt = - \int_{\Omega} \nabla \cdot \mathbf{j} d\tau = - \oint dS(\Omega) \mathbf{j} \cdot \mathbf{n} \quad (5.14)$$

where $N(\Omega)$, the average number of electrons in Ω , equals the integral of ρ over the region Ω . This law states that the change in the average number of particles in the volume Ω is given by the flux in the vector current density through its bounding surface $S(\Omega)$. The final equality given in eqn (5.14) results from the use of Gauss's theorem which states that the volume integral of the divergence of a vector is equal to the flux in the vector through the surface bounding the system. The divergence of a vector and Gauss's theorem are reviewed in Section E5.3.

Depending on the nature of the state function, some observables yield 'sharp' values— Ψ is an eigenfunction of the observable—while others yield

$$\nabla \cdot \mathbf{V} = \text{div } \mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z}$$
$$\nabla^2 \rho = \nabla \cdot \nabla \rho \therefore \iiint dV \nabla^2 \rho = \oint dS \nabla \rho \cdot \mathbf{n}$$

H E
HERMETIA-
NO PRA
* UM
SISTEMA
ABERTO,
MAS NAO
TEM UM
QUANTO

only 'average values'. [If two observables \hat{A} and \hat{B} commute, i.e.

$$[\hat{A}, \hat{B}] = (\hat{A}\hat{B} - \hat{B}\hat{A}) = 0, \quad (5.15)$$

then they possess a common, complete set of orthonormal eigenfunctions. For a system in a stationary state for which the state function is an eigenfunction of the Hamiltonian operator,

$$\hat{H}\psi = E\psi \quad \text{and} \quad \hat{H}\psi^* = E\psi^*, \quad (5.16)$$

all of the observables which form a pairwise commuting set with \hat{H} yield a set of sharp values and this set of numbers is used to classify the state of the system. The ground state of the water molecule, in addition to possessing a sharp value for the energy, is labelled 1A_1 as a consequence of the spin and symmetry operators forming such a commuting set with \hat{H} . The average value of an observable that is not a member of this commuting set of observables is given by

$$\langle \hat{A} \rangle = \langle \Psi, \hat{A}\Psi \rangle / \langle \Psi, \Psi \rangle = \int \Psi^* \hat{A}\Psi d\tau / \int \Psi^* \Psi d\tau \quad (5.17)$$

and, of course, for a system in a stationary state, the sharp value for a commuting observable equals its average value.

[The equation of motion for the average value of an observable \hat{A} can be obtained directly from Heisenberg's equation for $\hat{A}(t)$ (Messiah 1958, p. 319). The result is

$$d\langle \hat{A} \rangle / dt = (i/\hbar) \langle \Psi, [\hat{H}, \hat{A}]\Psi \rangle + \langle \partial \hat{A} / \partial t \rangle. \quad (5.18)$$

Equation (5.18) can also be obtained by differentiating the expression for the average value of \hat{A} (eqn (5.17)) assuming Ψ is normalized to unity,

$$d\langle \hat{A} \rangle / dt = \int \{ (\partial \Psi^* / \partial t) \hat{A} \Psi + \Psi^* \hat{A} (\partial \Psi / \partial t) + \Psi^* (\partial \hat{A} / \partial t) \Psi \} d\tau,$$

followed by the use of Schrödinger's equations (5.2), and subsequently taking into account the Hermiticity of the Hamiltonian operator,

$$\begin{aligned} d\langle \hat{A} \rangle / dt &= (i/\hbar) \{ \langle \hat{H}\Psi, \hat{A}\Psi \rangle - \langle \Psi, \hat{A}\hat{H}\Psi \rangle \} + \langle \partial \hat{A} / \partial t \rangle \\ &= (i/\hbar) \langle \Psi, (\hat{H}\hat{A} - \hat{A}\hat{H})\Psi \rangle + \langle \partial \hat{A} / \partial t \rangle \\ &= (i/\hbar) \langle \Psi, [\hat{H}, \hat{A}]\Psi \rangle + \langle \partial \hat{A} / \partial t \rangle. \end{aligned} \quad (5.19)$$

For an observable \hat{B} which commutes with the Hamiltonian and which does not possess an explicit time dependence one has

$$d\langle \hat{B} \rangle / dt = 0 \quad (5.20)$$

and its value is independent of time. In analogy with classical mechanics the observable \hat{B} is called a constant of the motion. [None of the observables considered here have an explicit time dependence and the final term on the right-hand side of eqns (5.18) or (5.19) will be omitted from this point on.]

IMPORTANTE: NOTAÇÃO DO LIVRO \equiv NOTAÇÃO USUAL

$$\langle \Psi, \hat{A}\Psi \rangle = \langle \Psi^* \hat{A} \Psi \rangle$$

[With Schrödinger's equation to describe how the state function changes with time and Heisenberg's equation to determine the corresponding change in the average value of each observable, one has a complete mechanical description of a system. Heisenberg's equation is important from another point of view, as it can be used to obtain relationships of great general importance for particular choices of the observable \hat{A} . Examples are the Ehrenfest theorems governing the time rate of change of the average values of an electronic position coordinate $\hat{\mathbf{r}}$ and momentum $\hat{\mathbf{p}} = -i\hbar\nabla$, and the virial theorem which is obtained when $\hat{A} = \hat{\mathbf{r}} \cdot \hat{\mathbf{p}}$. These theorems play important roles in the mechanical description of an atom in a molecule.]

[With \hat{A} set equal to the product of the mass and the position coordinate for an electron, the commutator is

$$m[\hat{H}, \hat{\mathbf{r}}] = (\hbar/i) (-i\hbar\nabla) = (\hbar/i)\hat{\mathbf{p}} \quad (5.21) \text{ EHRENFEST}$$

and the first of Ehrenfest's relations is obtained,

$$m d\langle \hat{\mathbf{r}} \rangle / dt = \langle \hat{\mathbf{p}} \rangle. \quad (5.22) \text{ RELATIONS}$$

For $\hat{A} = \hat{\mathbf{p}}$, the commutator is

$$[\hat{H}, \hat{\mathbf{p}}] = i\hbar\nabla\hat{V}(\mathbf{r}) \quad (5.23) \text{ DE EQUAÇÃO}$$

and the second Ehrenfest relation is obtained

$$d\langle \hat{\mathbf{p}} \rangle / dt = \langle -\nabla\hat{V}(\mathbf{r}) \rangle = \langle \hat{\mathbf{F}}(\mathbf{r}) \rangle \quad (5.24) \text{ FORÇA}$$

where, in analogy with classical definitions, the force operator $\hat{\mathbf{F}}(\mathbf{r})$ is defined as the negative of the gradient of the potential. The operator $\hat{\mathbf{F}}(\mathbf{r})$ will be termed the Ehrenfest force. Equations (5.22) and (5.24) are both quantum analogues of classical relations but, in the quantum case, the relations hold only for the average values of the mechanical properties and not in a pointwise manner for each specific time as they do on a classical trajectory. The two expressions can be combined to yield the analogue of Newton's equation of motion]

$$m d^2\langle \hat{\mathbf{r}} \rangle / dt^2 = \langle \hat{\mathbf{F}}(\mathbf{r}) \rangle. \quad (5.25) \text{ EHRENFEST FORCE}$$

The left-hand side of eqns (5.22) and (5.24) may be explicitly evaluated using Schrödinger's equations (5.2) for Ψ and Ψ^* , and expressed in terms of the quantum vector current density, eqn (5.11). For the time derivative of $\langle \hat{\mathbf{r}} \rangle$ one has

$$\begin{aligned} d\langle \Psi^* \hat{\mathbf{r}} \Psi \rangle / dt &= (i/\hbar) \{ (\hat{H}\Psi^*) \hat{\mathbf{r}} \Psi - \Psi^* \mathbf{r} (\hat{H}\Psi) \}, \\ d\langle \Psi^* \hat{\mathbf{p}} \Psi \rangle / dt &= (-i\hbar/2m) \{ \hat{\mathbf{r}} (\nabla^2 \Psi^*) \Psi - \hat{\mathbf{r}} \Psi^* \nabla^2 \Psi \} \\ &= -\hat{\mathbf{r}} \nabla \cdot \mathbf{j}. \end{aligned}$$

In general, the notation \mathbf{AB} will be used to denote a multiplication of two vectors to yield a dyadic. A review of the properties of dyadics is given in

IMPORTANTE!
CIA DA
EQUAÇÃO
5.2 E
5.18

EHRENFEST
RELATIONS
DE EQUAÇÃO
FORÇA
5.25

EHRENFEST
FORCE

Section E5.2. In particular, the following identity holds

$$\nabla \cdot (\hat{\mathbf{j}}\mathbf{r}) = \mathbf{j} + \hat{\mathbf{r}}\nabla \cdot \mathbf{j}$$

and hence

$$d\langle \hat{\mathbf{r}} \rangle / dt = \int \mathbf{j} d\tau - \int \nabla \cdot (\hat{\mathbf{j}}\mathbf{r}) d\tau,$$

which, using Gauss's theorem followed by multiplication by m , becomes

$$m d\langle \hat{\mathbf{r}} \rangle / dt = m \int \mathbf{j} d\tau - m \oint dS (\mathbf{n} \cdot \mathbf{j}) \hat{\mathbf{r}}. \quad (5.26a)$$

The surface term vanishes for a system with boundaries at infinity and one obtains the result that the time derivative of the average position vector is given by the integral of the current density,

$$m d\langle \hat{\mathbf{r}} \rangle / dt = m \int \mathbf{j} d\tau = \frac{1}{2} \{ \langle \hat{\mathbf{p}}\Psi, \Psi \rangle + \langle \Psi, \hat{\mathbf{p}}\Psi \rangle \} = \overline{\langle \hat{\mathbf{p}} \rangle}. \quad (5.26b)$$

The current has the dimensions of a velocity and thus, when multiplied by the mass, it equals a momentum. Similarly, the time derivative of a velocity is an acceleration and hence one should obtain

$$d\overline{\langle \hat{\mathbf{p}} \rangle} / dt = m \int \{ \partial \mathbf{j}(\mathbf{r}) / \partial t \} d\tau = \langle \mathbf{F}(\mathbf{r}) \rangle. \quad (5.27)$$

The derivation of eqn (5.27) also entails the vanishing of a surface term, one involving the quantum stress tensor. The stress tensor was first introduced in its relativistic form in 1927 by Schrödinger and its properties were later discussed by Pauli (1933). The derivation of eqn (5.27) follows that given by Pauli (1958).

$$m \partial \mathbf{j} / \partial t = \frac{1}{2} \{ \hat{H}\Psi^* (\nabla\Psi) - \Psi^* \nabla(\hat{H}\Psi) + (\nabla\Psi^*)\hat{H}\Psi - \nabla(\hat{H}\Psi^*)\Psi \},$$

$$m \partial \mathbf{j} / \partial t = (\hbar^2/4m) \{ (-\nabla^2\Psi^*)\nabla\Psi + \Psi^*\nabla(\nabla^2\Psi) - (\nabla\Psi^*)\nabla^2\Psi + \nabla(\nabla^2\Psi^*)\Psi \} + 1/2 \{ \hat{V}(\nabla\Psi^*)\Psi - \nabla(\hat{V}\Psi^*)\Psi + \hat{V}\Psi^*\nabla\Psi - \Psi^*\nabla(\hat{V}\Psi) \},$$

$$m \partial \mathbf{j} / \partial t = -(\hbar^2/4m) \{ (\nabla^2\Psi^*)\nabla\Psi - \Psi^*\nabla(\nabla^2\Psi) + (\nabla\Psi^*)\nabla^2\Psi - \nabla(\nabla^2\Psi^*)\Psi \} - \Psi\Psi^*\nabla\hat{V}.$$

Define the stress tensor $\vec{\sigma}$ as

$$\vec{\sigma} = (\hbar^2/4m) \{ \Psi^*\nabla(\nabla\Psi) + \nabla(\nabla\Psi^*)\Psi - \nabla\Psi^*\nabla\Psi - \nabla\Psi\nabla\Psi^* \} \quad (5.28)$$

and the expression for the time derivative of \mathbf{j} may be expressed as

$$m \partial \mathbf{j} / \partial t = + \nabla \cdot \vec{\sigma} - \Psi^*\Psi \nabla \hat{V} \quad (5.29)$$

Hence,

$$m \int \{ \partial \mathbf{j}(\mathbf{r}) / \partial t \} d\tau = \int \Psi^* (-\nabla\hat{V})\Psi d\tau + \oint dS \vec{\sigma} \cdot \mathbf{n}$$

where Gauss's theorem has been used to transform the volume integral of $\nabla \cdot \vec{\sigma}$ into an integral over the surface, an integral which vanishes for a system with infinite boundaries to yield the equations listed under (5.27). While the current density does not appear in the description of the properties of a stationary state in the absence of a magnetic field, it and the quantum stress tensor $\vec{\sigma}$ play important roles in the expressions obtained for the average values of subspace properties and they are introduced here for that reason.

[One further theorem of paramount importance obtainable from Heisenberg's equation is the virial theorem as obtained from the operator $\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}$. In this case the commutator is

$$(i/\hbar) [\hat{H}, \hat{\mathbf{r}} \cdot \hat{\mathbf{p}}] = (-\hbar^2/m)\nabla^2 - \mathbf{r} \cdot \nabla \hat{V} = 2\hat{T} - \mathbf{r} \cdot \nabla \hat{V}$$

where \hat{T} denotes the kinetic energy operator. Hence,

$$d\langle \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} \rangle / dt = 2\langle \hat{T} \rangle + \langle -\hat{\mathbf{r}} \cdot \nabla \hat{V} \rangle. \quad (5.30)$$

The usual statement of the virial theorem for a stationary state is

$$2\langle \hat{T} \rangle + \langle -\hat{\mathbf{r}} \cdot \nabla \hat{V} \rangle = 0. \quad (5.31)$$

In the original derivation of the classical virial theorem given by Clausius, an expression corresponding to eqn (5.30) is also obtained. In the classical case one argues that the time average of $d(\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}) / dt$ vanishes over a sufficiently long period of time or that the motion is periodic to obtain the equivalent of eqn (5.31). *

Alternatively, the virial theorem may be derived directly from eqn (5.29), which expresses the forces in terms of the time derivative of the current density. Dotted the vector \mathbf{r} into eqn (5.29) to obtain the virial of these forces gives

$$m\mathbf{r} \cdot \{ \partial \mathbf{j} / \partial t \} = + \mathbf{r} \cdot \nabla \cdot \vec{\sigma} - \Psi^*\Psi (\mathbf{r} \cdot \nabla \hat{V}).$$

The term $\mathbf{r} \cdot \nabla \cdot \vec{\sigma}$ may be re-expressed using the identity

$$\nabla \cdot (\mathbf{r} \cdot \vec{\sigma}) = \text{Tr} \vec{\sigma} + \mathbf{r} \cdot \nabla \cdot \vec{\sigma}$$

where $\text{Tr} \vec{\sigma}$ denotes the trace of the tensor $\vec{\sigma}$,

$$\text{Tr} \vec{\sigma} = (\hbar^2/4m) \{ \Psi^*\nabla^2\Psi + (\nabla^2\Psi^*)\Psi - \nabla\Psi^* \cdot \nabla\Psi - \nabla\Psi \cdot \nabla\Psi^* \}. \quad (5.32)$$

Each term in eqn (5.32) gives minus one-half of the average kinetic energy when integrated over all space. As demonstrated later in eqn (5.47), the averages of these two ways of expressing the kinetic energy in eqn (5.32) differ by a surface integral and one has

$$- \int \text{Tr} \vec{\sigma} d\tau = 2\langle \hat{T} \rangle + (\hbar^2/4m) \oint dS \nabla \rho \cdot \mathbf{n}.$$

DO COMUTADOR:

$$[\hat{H}, \hat{x}\hat{p}_x] = i \left[-\hat{p}_x^2 + \hat{x} \frac{\partial \hat{V}}{\partial x} \right] \quad \text{EM TRÊS DIMENSÕES, TEMOS:}$$

$$[\hat{H}, \hat{\mathbf{r}} \cdot \hat{\mathbf{p}}] = i \left[-\hat{p}^2 + \hat{\mathbf{r}} \cdot \nabla \hat{V} \right]$$

QUANTUM
VIRIAL
THEOREM

Use of these results yields

$$m \int \mathbf{r} \cdot \{\partial \mathbf{j} / \partial t\} d\tau = 2 \langle \hat{T} \rangle + \langle -\mathbf{r} \cdot \nabla \hat{V} \rangle + (\hbar^2 / 4m) \oint dS \nabla \rho \cdot \mathbf{n} + \oint dS \mathbf{r} \cdot \vec{\sigma} \cdot \mathbf{n}$$

where the final surface term comes from applying Gauss's theorem to the integral of $\nabla \cdot (\mathbf{r} \cdot \vec{\sigma})$. The two surface integrals vanish for a system with boundaries at infinity. Hence the final result, a statement of the virial theorem for the general time-dependent case,

$$d \langle \hat{\mathbf{r}} \cdot \hat{\mathbf{p}} \rangle / dt = m \int \mathbf{r} \cdot \{\partial \mathbf{j} / \partial t\} d\tau = 2 \langle \hat{T} \rangle + \langle -\hat{\mathbf{r}} \cdot \nabla \hat{V} \rangle. \quad (5.33)$$

The virial of a system, denoted by the symbol \mathcal{V} , is defined to be the average

$$\mathcal{V} = \langle -\mathbf{r} \cdot \nabla \hat{V}(\mathbf{r}) \rangle = \langle \mathbf{r} \cdot \mathbf{F}(\mathbf{r}) \rangle.$$

It is the average of the virial of the forces acting on the particles which possess the kinetic energy $\langle \hat{T} \rangle$. When $\mathbf{F}(\mathbf{r})$ is the Ehrenfest force acting on the electrons, \mathcal{V} is the potential energy of the electrons. In those cases where \hat{V} , the potential energy operator of the system, is a homogeneous function of the coordinates of degree n , one can, by using Euler's theorem, re-express the virial in terms of the average of the potential energy operator

$$\mathcal{V} = -n \langle \hat{V}(\mathbf{r}) \rangle$$

and, consequently, for a stationary state one has

$$2 \langle \hat{T} \rangle = n \langle \hat{V} \rangle. \quad (5.34)$$

For an isolated atom, the potential energy operator for the Coulombic forces is a homogeneous function of degree -1 and in this case one has

$$2 \langle \hat{T} \rangle = - \langle \hat{V} \rangle.$$

HERMITIANO [For a system in a stationary state, the equation of motion for the mean value of \hat{A} , eqn (5.18), assumes the form

$$\langle [\hat{H}, \hat{A}] \rangle = 0.$$

$$\hat{H}\Psi^* = E\Psi^* \\ \hat{H}\Psi = E\Psi$$

Alternatively, this result can be derived directly by making use of the Hermitian property of \hat{H} and Schrödinger's equation for a stationary state (eqn (5.16)),

$$\langle [\hat{H}, \hat{A}] \rangle = \langle \psi, (\hat{H}\hat{A} - \hat{A}\hat{H})\psi \rangle = \langle \hat{H}\psi, \hat{A}\psi \rangle - \langle \psi, \hat{A}\hat{H}\psi \rangle = E \{ \langle \psi, \hat{A}\psi \rangle - \langle \psi, \hat{A}\psi \rangle \} = 0. \quad (5.35)$$

If \hat{H} contains a real parameter s such as a nuclear charge or a nuclear coordinate and \hat{A} corresponds to the operator $\partial / \partial s$, then the derivation of eqn (5.35) must be modified to include the term $\partial E / \partial s$, for E contains the

IMPORTANTE: HERMITIANO

$$\langle [\hat{H}, \hat{A}] \rangle = \langle \Psi^* (\hat{H}\hat{A} - \hat{A}\hat{H}) \Psi \rangle = \langle \Psi^* \hat{H} \hat{A} \Psi \rangle - \langle \Psi^* \hat{A} \hat{H} \Psi \rangle = \langle \hat{H} \Psi^* \hat{A} \Psi \rangle - \langle \Psi^* \hat{A} \hat{H} \Psi \rangle$$

same parameters as does \hat{H} . The result in this case is

$$\langle [\hat{H}, \hat{A}] \rangle - \langle [E, \hat{A}] \rangle = 0$$

and, when the commutators are evaluated, one obtains the Hellmann-Feynman theorem

$$\langle \psi, (\partial \hat{H} / \partial s) \psi \rangle = \partial E / \partial s. \quad (5.36)$$

When s is a nuclear coordinate X_α , eqn (5.36) gives the Hellmann-Feynman electrostatic theorem

$$\langle \psi, (-\nabla_\alpha \hat{V}) \psi \rangle = -\nabla_\alpha E = \mathbf{F}_\alpha. \quad (5.37)$$

[Equation (5.35) and the generalized statement of the Hellmann-Feynman theorem can also be derived through a variation of the state function (Epstein 1974b). In this case eqn (5.35) is known as the *hypervirial theorem* (Hirschfelder 1960). The variational derivations are important because they demonstrate that the theorems hold not only for the exact state function but also for approximations to these functions when the latter functions, such as Hartree-Fock, unrestricted Hartree-Fock, and multiconfigurational Hartree-Fock, are invariant to the variational parameters. The original such proof of the Hellmann-Feynman theorem was given by Hurley (1954)]

HYPER VIRIAL THEOREM PARA SISTEMA ALBERTO ED. 5.35

[It is clear from the use of the Hermiticity of \hat{H} in the derivation of eqn (5.35) that the hypervirial theorem will assume a different form for a subsystem. A derivation of the hypervirial theorem for a subsystem is given at this point to illustrate the consequence of the loss of the Hermitian property of \hat{H} . The result is general and applies to any subsystem regardless of its definition. To ensure that the subsystem average of an observable \hat{A} be a real number, one must work with the mean of the average value and its complex conjugate (cc),

HYPER VIRIAL THEOREM PARA UN SUBSYSTEMA

$$\frac{1}{2} \left\{ \int_\Omega \psi^* [\hat{H}, \hat{A}] \psi d\tau + \int_\Omega ([\hat{H}, \hat{A}] \psi)^* \psi d\tau \right\} = \frac{1}{2} \{ \langle [\hat{H}, \hat{A}] \rangle_\Omega + cc \}. \quad (5.38)$$

Working with just the first of the averages, one obtains

$$\langle \psi, [\hat{H}, \hat{A}] \psi \rangle_\Omega = \langle \psi, \hat{H} \hat{A} \psi \rangle_\Omega - \langle \hat{H} \psi, \hat{A} \psi \rangle_\Omega + \langle \hat{H} \psi, \hat{A} \psi \rangle_\Omega - \langle \psi, \hat{A} \hat{H} \psi \rangle_\Omega \quad (5.39)$$

where the term $\langle \hat{H} \psi, \hat{A} \psi \rangle_\Omega$ has been both added and subtracted to maximize the correspondence with eqn (5.35). Making use of Schrödinger's equation as before to get rid of the last two terms on the right-hand side,

$$\langle \psi, [\hat{H}, \hat{A}] \psi \rangle_\Omega = (\hbar^2 / 2m) \int_\Omega \{ (\nabla^2 \psi^*) \hat{A} \psi - \psi^* \nabla^2 (\hat{A} \psi) \} d\tau \quad (5.40)$$

where, as in the derivation of eqn (5.7), the terms involving \hat{V} have cancelled

(SCHRÖDINGER): $\hat{H}\Psi^* = E\Psi^*$

$$\begin{aligned} & \langle \Psi^* \hat{H} \hat{A} \Psi \rangle_\Omega - \langle \hat{H} \Psi^* \hat{A} \Psi \rangle_\Omega = \\ & = \langle E \Psi^* \hat{A} \Psi \rangle - \langle \Psi^* \hat{A} E \Psi \rangle = \\ & = E \langle \Psi^* \hat{A} \Psi \rangle - \langle \Psi^* \hat{A} \Psi \rangle E \end{aligned}$$

FIRST TWO TERMS OF 5.39

out. Using eqn (5.9) with $\phi^* = \psi^*$ and $\phi = \hat{A}\psi$, gives

$$\langle \psi, [\hat{H}, \hat{A}]\psi \rangle_{\Omega} = -i\hbar \oint dS(\Omega) \mathbf{j}_A \cdot \mathbf{n} \quad (5.41)$$

where \mathbf{j}_A , the current density for an observable \hat{A} , is defined in a manner analogous to that for the vector current \mathbf{j} itself,

$$\mathbf{j}_A = (\hbar/2mi) \{ \psi^* \nabla(\hat{A}\psi) - (\nabla\psi^*)\hat{A}\psi \}. \quad (5.42)$$

The final statement of the subsystem hypervirial theorem is

$$\frac{1}{2} \{ \langle [\hat{H}, \hat{A}] \rangle_{\Omega} + cc \} = -\frac{1}{2} \{ i\hbar \oint dS(\Omega) \mathbf{j}_A \cdot \mathbf{n} + cc \}, \quad (5.43)$$

a result first obtained by Epstein (1974); (see also Srebrenik and Bader 1974).

By performing corresponding substitutions on the subsystem analogue of eqn (5.19), the expression for the time derivative of a subsystem expectation value is

$$\begin{aligned} & \frac{1}{2} d \{ \langle \Psi, \hat{A}\Psi \rangle_{\Omega} + \langle \hat{A}\Psi, \Psi \rangle_{\Omega} \} / dt \\ &= \frac{1}{2} \{ (i/\hbar) \langle [\hat{H}, \hat{A}] \rangle_{\Omega} + cc \} - \frac{1}{2} \{ \oint dS(\Omega) \mathbf{j}_A \cdot \mathbf{n} + cc \} \\ &+ \frac{1}{2} \{ \oint dS(\Omega) (\partial S / \partial t) \Psi^* \hat{A}\Psi + cc \}. \end{aligned} \quad (5.44)$$

The final term accounts for the dependence of the surface S bounding the region Ω on the time.

The importance of Heisenberg's equation and of the corresponding hypervirial theorem for a stationary state in the description of the properties of the total system is maintained in the description of a subsystem. Indeed, the generalized action principle that is employed to establish the quantum mechanics of a subsystem can be expressed in the form of a variational statement of the Heisenberg equation of motion. For a stationary state, the same principle reduces to a variation of the energy and the resulting theorem is a variational statement of the hypervirial theorem applicable to both the total system and a subsystem. It is to be emphasized, however, that the variational derivations of these statements as obtained through the action principle apply only to a particular class of subsystems, those bounded by a surface of zero flux in the gradient vector of the charge density—the definition of a topological atom. The use of these theorems in the description of a subsystem leads to novel results because of the general non-vanishing of the surface integral in eqns (5.43) and (5.44).

5.3 Need for a subsystem variation principle

An understanding of chemistry requires a regional description of a system. The notion that a molecule can be viewed as a collection of atoms linked by a network of bonds, a notion that has already been shown to be rooted in the topological properties of the charge distribution, is the operational principle

USANDO O TEOREMA DE GAUSS PARA A EQ 5.40
 $(\hbar^2/2m) \int_{\Omega} ((\nabla^2 \psi^*)\hat{A}\psi - \psi^* \nabla^2(\hat{A}\psi)) d\tau =$
 $= (\hbar^2/2m) \oint dS \{ \nabla \psi^* \hat{A}\psi - \psi^* \nabla \hat{A}\psi \} \cdot \mathbf{n} \rightarrow \text{Eq 5.41}$

underlying our classification and understanding of chemical behaviour. It would appear, therefore, that to find chemistry within the framework of quantum mechanics one must determine the values of observables for pieces of a total system, i.e. for a subsystem. But how is one to choose these pieces? Is there but one, or are there many ways of dividing a system into 'atoms' and its properties into corresponding atomic contributions? If there is an answer to this problem, then the necessary information must be contained in the state function, for Ψ tells us everything we can know about a system.

It is clear from the examples given in the preceding section that one can write down expressions for the expectation values of observables for a subsystem with an arbitrarily defined boundary. One has no guarantee, however, that the results obtained for such an arbitrarily defined region of space will be of physical significance. Indeed, one can show very simply and as one important example that the kinetic energy of an arbitrarily defined region of space is not well-defined.

Consider the expression

$$\begin{aligned} \nabla^2(\Psi^*\Psi) &= \nabla \cdot \{ (\nabla\Psi^*)\Psi + \Psi^*(\nabla\Psi) \} \\ &= (\nabla^2\Psi^*)\Psi + \Psi^*(\nabla^2\Psi) + 2\nabla\Psi^* \cdot \nabla\Psi. \end{aligned} \quad (5.45)$$

Thus, in the general many-electron case one has

$$\begin{aligned} -(\hbar^2/4m) \sum_i \{ \Psi^* \nabla_i^2 \Psi + \Psi \nabla_i^2 \Psi^* \} &= (\hbar^2/2m) \sum_i \nabla_i \Psi^* \cdot \nabla_i \Psi \\ &- (\hbar^2/4m) \sum_i \nabla_i^2 (\Psi^* \Psi). \end{aligned} \quad (5.46)$$

Since electrons are indistinguishable and Ψ is antisymmetrized, the average of a sum of N one-electron operators can be replaced by N times the average of one of the operators. If the sum of the operators is replaced by N times a single operator in each term in the above equation and is then summed over all the spin coordinates and integrated over the space coordinates of all electrons *but one* (operations denoted by the symbol $\int d\tau'$, see eqn (1.4)), the result is

$$\begin{aligned} -(\hbar^2/4m) N \int d\tau' \{ \Psi^* \nabla^2 \Psi + \Psi \nabla^2 \Psi^* \} &= (\hbar^2/2m) N \int d\tau' \nabla \Psi^* \cdot \nabla \Psi \\ &- (\hbar^2/4m) N \int d\tau' \nabla^2 (\Psi^* \Psi). \end{aligned} \quad (5.47)$$

This is readily restated in terms of the one-density matrix, $\Gamma^{(1)}(\mathbf{r}, \mathbf{r}')$ (see eqn (E1.1)), and its trace, the charge density, $\rho(\mathbf{r})$ (eqn (1.4))

$$\begin{aligned} -(\hbar^2/4m) \{ \nabla^2 + \nabla'^2 \} \Gamma^{(1)}(\mathbf{r}, \mathbf{r}')|_{\mathbf{r}=\mathbf{r}'} &= (\hbar^2/2m) (\nabla \cdot \nabla') \Gamma^{(1)}(\mathbf{r}, \mathbf{r}')|_{\mathbf{r}=\mathbf{r}'} \\ &- (\hbar^2/4m) \nabla^2 \rho(\mathbf{r}). \end{aligned} \quad (5.48)$$

This equation or, equivalently, eqn (5.47) can be used to define, respectively, the quantities (Bader and Preston 1969)

$$K(\mathbf{r}) = G(\mathbf{r}) + L(\mathbf{r}). \quad (5.49)$$

KINETIC ENERGY OPERATOR FOR ANY MANY-ELECTRON SYSTEM

* $K(\mathbf{r})$ and $G(\mathbf{r})$ are kinetic energy densities and $L(\mathbf{r})$, a function of the Laplacian of the charge density. Integration of the final coordinate \mathbf{r} in eqn (5.49) over a region of space Ω yields

$$\int_{\Omega} K(\mathbf{r}) d\tau = \int_{\Omega} G(\mathbf{r}) d\tau - (\hbar^2/4m) \int_{\Omega} \nabla \cdot \nabla \rho(\mathbf{r}) d\tau, \quad (5.50)$$

$$K(\Omega) = G(\Omega) - (\hbar^2/4m) \oint dS(\Omega) \nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = G(\Omega) + L(\Omega)$$

where Gauss's theorem has been used to replace the final integral on the right-hand side with an integral over the surface bounding the region Ω . The surface integral is seen to be a measure of the flux in the gradient vectors of the charge density through the surface bounding the region. From the discussion of the gradient vector field of the charge density given in Chapter 3, it is clear that the surface integral will not vanish for an arbitrary surface. There are two cases for which the surface integral will always vanish yielding an equality in the kinetic energy expectation values: (1) when the integral is taken over all space and one has the result first pointed out by Schrödinger, that the average kinetic energy T may be calculated using either of the above forms for the kinetic energy operator,

$$\langle \hat{T} \rangle = T = K = G, \quad (5.51)$$

and (2) when the region Ω is bounded by a surface of zero flux in the gradient vector of the charge density (eqn (2.9) and repeated here as eqn (5.52)),

$$\nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \text{for all } \mathbf{r} \text{ on the surface } S(\mathbf{r}). \quad (5.52)$$

Thus, the kinetic energy of the topologically defined atom $T(\Omega)$ is a well-defined quantity and one has

$$T(\Omega) = K(\Omega) = G(\Omega) \quad \text{if } \Omega \text{ is bounded by a zero-flux surface} \quad (5.53)$$

in correspondence with the total system result, eqn (5.51). The expression for the kinetic energy

$$(1/2m) \langle \hat{p}\psi, \hat{p}\psi \rangle = (\hbar^2/2m) \langle \nabla\psi \cdot \nabla\psi \rangle$$

is the quantum analogue of the classical expression $|p|^2/2m$. It corresponds to the use of the first term on the right-hand side of eqn (5.46) or the kinetic energy density $G(\mathbf{r})$ in eqn (5.49) in the many-electron case.

Cohen (1979) has shown that all kinetic energy densities, when defined in terms of the generalizations made possible through the use of the Wigner distribution function, yield identical average values when integrated over a region bounded by a surface of zero flux, as in eqn (5.52). It is possible to have a surface through which the net integrated flux is zero, the flux being of opposite sign in different regions of the surface. Such a situation will not, in general, persist for changes in the system's surface, whether the change is real, as caused by nuclear displacements, or virtual, as caused by a mathematical

variation of the surface elements. A surface of zero flux in the charge density is maintained at all times through real motions and through virtual changes as brought about by a variation of the charge density. It will be demonstrated later that the virial, as well as the kinetic energy, is not uniquely defined for a subsystem with an arbitrary boundary when either eqn (5.43) or (5.44) is evaluated for the operator $\hat{\mathbf{r}} \cdot \hat{\mathbf{p}}$. Thus, the average total energy of the electrons in a region of space with arbitrary boundaries is not defined. These results have a related and most important consequence: the possible energy functionals, whose variation yields Schrödinger's equation for a total system exhibit different properties with respect to the variations when the functionals are evaluated over regions with arbitrary boundaries. As demonstrated below, a unique variational result is obtained only for a subsystem satisfying eqn (5.52).

It was stated in the introductory chapter that the question 'are there atoms in molecules' is equivalent to asking two equally necessary questions of quantum mechanics: (1) does the state function contain the necessary information to predict a unique partitioning of a molecule into subsystems, and (2) does quantum mechanics provide a complete description of the subsystems so defined? The only method at one's disposal to answer questions of this nature are extrema principles based upon a variational principle. Schrödinger (1926) used such a method in his original derivation of the equation which bears his name. In relatively recent times, two new but equivalent formulations of quantum mechanics, one by Feynman (1948), the other by Schwinger (1951), have appeared, both of them based upon a single principle, the quantum analogue of the principle of least action, or more simply the action principle. The action principle in its generalized form as developed by Schwinger offers an alternative variational approach to the laws of physics, opening the door to new questions and new answers, and herein lies the power of the method. We shall use it to answer the question 'are there atoms in molecules?'

For a system in a stationary state, the variation of the action integral reduces to the variation of an energy integral which is identical to that first constructed by Schrödinger and used by him to derive the wave equation for a stationary state. We shall introduce this subject in the following section by reviewing Schrödinger's derivation of his wave equation and showing how this derivation can be generalized to yield a quantum description of a particular class of quantum subsystems—the topological atom (Srebrenik and Bader 1975; Bader 1988). This demonstration is sufficient to provide an operational understanding of the quantum mechanics of an atom in a molecule.

Chapter 8 gives a full account of Schwinger's principle of stationary action and of how, through its generalization, one obtains a definition of a quantum subsystem and a description of its properties.

5.3.1 Schrödinger's derivation of wave mechanics

Schrödinger concludes his first paper (1926) by stating that 'the function ψ be such as to make the 'Hamilton integral'

$$\int d\tau \{ \hbar^2 T(q, \partial\psi/\partial q) + \psi^2 V \}$$

stationary while fulfilling the normalizing, accessory condition

$$\int \psi^2 d\tau = 1.$$

The quantity $T(q, \partial\psi/\partial q)$ is the kinetic energy expressed in terms of the momentum as in the first term on the right-hand side of eqn (5.47) that is, as $|\hat{p}^2|/2m$, and V is the potential energy. Schrödinger considered specifically the problem of the hydrogen atom for which the explicit form of the functional is

$$\mathcal{J}[\psi] = \int d\tau \{ (\hbar^2/2m) \nabla\psi \cdot \nabla\psi + V\psi\psi \} \quad (5.54)$$

with $V = -e^2/r$. In his first paper Schrödinger did not allow ψ to be complex. The problem is to find the function ψ such that $\mathcal{J}[\psi]$ attains its minimum value subject to the constraint that ψ remain normalized to unity. In ordinary extrema problems a function is a maximum or a minimum at a single point in the space of its variables and these problems are handled by the methods of differential calculus. In contrast, $\mathcal{J}[\psi]$ is an integral and its stationarity requires that its value be a minimum with respect to the averaging of the integrand over *all* points in configuration space. This is a problem requiring the calculus of variations.

The integral $\mathcal{J}[\psi]$ is extremized in the sense that any arbitrary change in the function $\psi(\mathbf{r})$ from its true value causes no first-order change in the value of the integral. One can think of varying the integral by assuming the existence of the 'correct function $\psi(\mathbf{r})$ ' and then changing the integral by changing $\psi(\mathbf{r})$ at each point \mathbf{r} into the 'trial function $\phi(\mathbf{r})$ ' by adding to it the small amount $\eta(\mathbf{r})$

$$\phi(\mathbf{r}) = \psi(\mathbf{r}) + \eta(\mathbf{r}) \quad (5.55)$$

where $\eta(\mathbf{r})$ is assumed to vanish at the boundaries of the system but is otherwise arbitrary. This varied expression for each value of the coordinate \mathbf{r} is substituted into the integral to yield $\mathcal{J}[\phi]$ and all terms of order η^2 and higher are discarded. The difference between $\mathcal{J}[\phi]$ and $\mathcal{J}[\psi]$ is the first-order difference in \mathcal{J} and is labelled $\delta\mathcal{J}[\psi]$, the variation of \mathcal{J} . Since $\psi(\mathbf{r})$ makes $\mathcal{J}[\psi]$ a minimum, $\delta\mathcal{J}[\psi]$ must vanish. Feynman *et al.* (1964, Section 19.1) give a useful and engaging introduction to the calculus of variations in a derivation of Newton's equation by a minimization of the classical action integral.

We shall allow for the possibility of ψ being complex in our review of Schrödinger's work. The functions ψ and ψ^* are treated as independent

variables and are varied separately. Proceeding with the variation of ψ as outlined above, one obtains

$$\delta\mathcal{J}[\psi] = \int \{ (\hbar^2/2m) \nabla\psi^* \cdot \nabla\eta + V\psi^*\eta \} d\tau.$$

To this must be added the effect of the constraint on the normalization of ψ which is written as

$$\int \psi^*\psi d\tau - 1 = 0.$$

Substitution of the trial function for ψ yields

$$\int \psi^*\psi d\tau + \int \psi^*\eta d\tau - 1 = 0$$

or

$$\langle \psi, \eta \rangle = 0. \quad (5.56)$$

Eqn (5.56) is multiplied by a factor λ and then added to the variation of \mathcal{J} . Requiring the resultant variation, which is labelled $\delta\mathcal{G}[\psi]$, to be stationary yields

$$\delta\mathcal{G}[\psi] = \delta\mathcal{J}[\psi] + \lambda \langle \psi, \eta \rangle = 0$$

or

$$\delta\mathcal{G}[\psi] = \int \{ (\hbar^2/2m) \nabla\psi^* \cdot \nabla\eta + (V + \lambda)\psi^*\eta \} d\tau = 0. \quad (5.57)$$

This method of handling the constraint is known as Lagrange's method of undetermined multipliers. One may introduce the constraint in this manner from the very beginning by defining a new functional $\mathcal{G}[\psi]$ as

$$\begin{aligned} \mathcal{G}[\psi] &= \mathcal{J}[\psi] + \lambda \langle \psi, \psi \rangle \\ &= \int \{ (\hbar^2/2m) \nabla\psi^* \cdot \nabla\psi + (V + \lambda)\psi^*\psi \} d\tau; \end{aligned} \quad (5.58)$$

the variation of $\mathcal{G}[\psi]$ then yields the expression $\delta\mathcal{G}[\psi]$.

We desire an expression for $\delta\mathcal{G}[\psi]$ in which no derivatives of η appear so that the complete integrand is multiplied by $\eta(\mathbf{r})$, the small arbitrary change we make in $\psi(\mathbf{r})$. This is easily done using the equivalent of an integration by parts. One notes that

$$\nabla \cdot (\nabla\psi^*\eta) = \nabla^2\psi^*\eta + \nabla\psi^* \cdot \nabla\eta \quad \text{or} \quad \nabla\psi^* \cdot \nabla\eta = \nabla \cdot (\nabla\psi^*\eta) - \nabla^2\psi^*\eta$$

and, using Gauss's theorem to transform the volume integral of $\nabla \cdot (\nabla\psi^*\eta)$, we obtain

$$\begin{aligned} \delta\mathcal{G}[\psi] &= \int \{ -(\hbar^2/2m) \nabla^2\psi^* + (V + \lambda)\psi^* \} \eta d\tau \\ &\quad + (\hbar^2/2m) \oint dS \nabla\psi^* \cdot \mathbf{n} \eta = 0. \end{aligned} \quad (5.59)$$

The surface integral is over the boundary of the system as found for $\mathbf{r} = \infty$

and η is required to vanish on this boundary. This yields

$$\begin{aligned} \delta \mathcal{G}[\psi] &= \int \{ -(\hbar^2/2m)\nabla^2\psi^* + (V - E)\psi^* \} \eta \, d\tau \\ &= \int \{ \hat{H}\psi^* - E\psi^* \} \eta \, d\tau = 0 \end{aligned} \quad (5.60)$$

where the constant λ has been identified with $-E$, the negative of the total energy. This result is to be true for all arbitrary functions $\eta(\mathbf{r})$, including, for example, the case where $\eta(\mathbf{r}) = 0$ for all \mathbf{r} but one arbitrarily chosen narrow range of values. The only way in which the integral in eqn (5.60) can vanish for all such variations η is for the quantity multiplied by η to vanish. This yields

$$\hat{H}\psi^* - E\psi^* = 0. \quad (5.61a)$$

A similar variation of ψ^* yields the complex conjugate of eqn (5.61)

$$H\psi - E\psi = 0. \quad (5.61b)$$

Equation (5.61a or b) is called the Euler-Lagrange or simply the Euler equation of the variation and, as first demonstrated by Schrödinger in 1926, variation of $\mathcal{G}[\psi]$ subject to the normalization constraint yields the wave equation. Because of the equivalence of the two kinetic energy densities when integrated over all space (eqn (5.51)) $\mathcal{G}[\psi] = E$ and thus the function ψ which satisfies the wave equation also minimizes the energy of the system. Schrödinger demonstrated that the 'quantum conditions' then postulated for the hydrogen atom could be replaced by this variation principle since the resulting Euler equation yields the quantum numbers n and l in a natural way.

One may obtain eqn (5.61a or b) as the Euler equation in the variation of $\mathcal{G}[\psi]$ without imposing any prescribed boundary conditions on $\eta(\mathbf{r})$, the change or variation of $\psi(\mathbf{r})$. This is accomplished by introducing 'the natural boundary conditions' (Courant and Hilbert 1953). The necessary condition for $\mathcal{G}[\psi]$ to be stationary is that its first variation $\delta \mathcal{G}[\psi]$ as given in eqn (5.59) vanish. If $\mathcal{G}[\psi]$ is stationary with respect to variations which do not have prescribed boundary values, then it is certainly stationary with respect to the smaller class of variations for which $\eta(\mathbf{r})$ vanishes on the boundary which implies Euler's equation. Thus we need consider only that part of the variation which depends on the boundary, a step which yields the 'natural boundary conditions',

$$\nabla\psi^* \cdot \mathbf{n} = 0 \quad \text{or} \quad \nabla\psi \cdot \mathbf{n} = 0 \quad \text{for all points on the boundary.} \quad (5.62)$$

This is the necessary boundary condition to obtain Schrödinger's equation as the Euler equation in the variation of $\mathcal{G}[\psi]$. For a bound system, one imposes the further condition that the state functions ψ and ψ^* themselves vanish on the boundary at infinity. This is a further, necessary restriction on the state

function if, as stated in the Born postulate, $\psi^*\psi \, d\tau$ is to represent a probability distribution. This probability must necessarily vanish on the boundaries infinitely far from the attractor which binds the system.

The normalization requirement on ψ can be handled differently by using the functional,

$$\mathcal{E}[\psi] = \int \{ (\hbar^2/2m)\nabla\psi^* \cdot \nabla\psi + V\psi^*\psi \} \, d\tau / \langle \psi, \psi \rangle. \quad (5.63)$$

A variation of $\mathcal{E}[\psi]$ through a variation of ψ as given in eqn (5.55) yields through first-order in η ,

$$\delta \mathcal{E}[\psi] = \int \{ (\hbar^2/2m)\nabla\psi^* \cdot \nabla\eta + (V - \mathcal{E}[\psi])\psi^*\eta \} \, d\tau / \langle \psi, \psi \rangle. \quad (5.64)$$

The expansion of $(\langle \psi, \psi \rangle + \langle \psi, \eta \rangle)^{-1}$ followed by that for $(1 + \langle \psi, \eta \rangle / \langle \psi, \psi \rangle)^{-1}$ up to first-order in η have been used to obtain eqn (5.64). Once again, ridding the expression of $\nabla\eta$ and demanding that η vanish at the boundaries of the system yields

$$\delta \mathcal{E}[\psi] = \int \{ \hat{H}\psi^* - \mathcal{E}[\psi]\psi^* \} \eta \, d\tau / \langle \psi, \psi \rangle = 0. \quad (5.65)$$

Since $\mathcal{E}[\psi] = E$ at the point of variation, one again obtains the wave equation as the Euler equation.

One can put the variation problem in a form where the usual ideas of differential calculus can be used to obtain an extremum in some function. Simply label all possible changes in ψ with a parameter α , such that for some value of α , say $\alpha = 0$, the varied function will coincide with the function which extremizes the integral. One possible parametric form, for example, is

$$\psi(\mathbf{r}, \alpha) = \psi(\mathbf{r}, 0) + \alpha\eta(\mathbf{r}) \quad (5.66)$$

where, as before, $\eta(\mathbf{r})$ is any arbitrary function which vanishes at infinity. If we denote by f the integrand of some integral $I[\psi]$ to be extremized

$$f = f(\psi, \nabla\psi),$$

then the parametrization given in eqn (5.66) gives $I[\psi]$ as a function of α ,

$$I[\psi, \alpha] = \int f(\psi(\alpha), \nabla\psi(\alpha)) \, d\tau.$$

The condition for an extremum is now given by the usual condition from differential calculus that $(\partial I / \partial \alpha)_{\alpha=0} = 0$. To obtain the extremum condition we multiply by $d\alpha$ and evaluate all derivatives at $\alpha = 0$,

$$\begin{aligned} (\partial I / \partial \alpha)_{\alpha=0} \, d\alpha &= \int \{ (\partial f / \partial \psi) (\partial \psi / \partial \alpha)_0 \, d\alpha \\ &\quad + (\partial f / \partial \nabla\psi) (\partial \nabla\psi / \partial \alpha)_0 \, d\alpha \} \, d\tau = 0. \end{aligned} \quad (5.67)$$

If one makes the following identifications at $\alpha = 0$,

$$(\partial I / \partial \alpha)_0 \, d\alpha = \delta I, \quad (\partial \psi / \partial \alpha)_0 \, d\alpha = \delta \psi, \quad (\partial \nabla\psi / \partial \alpha)_0 \, d\alpha = \delta \nabla\psi,$$

then one can write down the general expression for the variation of $I[\psi, \alpha]$ as

$$\delta I = \int \{ (\partial f / \partial \psi) \delta \psi + (\partial f / \partial \nabla \psi) \delta \nabla \psi \} d\tau. \quad (5.68)$$

Correspondingly, the expression for a trial function is amended to read

$$\phi(\mathbf{r}) = \psi(\mathbf{r}) + \delta\psi(\mathbf{r}). \quad (5.69)$$

To proceed beyond eqn (5.68) the procedure is always the same: one rides the expression of $\delta \nabla \psi$ (or $\delta \dot{\psi}$ for a time-dependent system) using an integration by parts to transform the integrand into a quantity multiplied only by $\delta \psi$. Setting this quantity equal to zero yields the Euler equation.

5.3.2 The variational definition of a subsystem and its properties

It will be demonstrated that the generalization of Schrödinger's variation of the functional $\mathcal{G}[\psi]$ to a subsystem has two important consequences: (1) The variation of the subsystem energy functional $\mathcal{G}[\psi, \Omega]$ yields the hypervirial theorem for a subsystem,

$$\delta \mathcal{G}[\psi, \Omega] = (-\varepsilon/2) \{ (i/\hbar) \langle [\hat{H}, \hat{G}] \rangle_{\Omega} + cc \} \quad (5.70)$$

where ε denotes an infinitesimal and the action of $\varepsilon \hat{G}$ on ψ causes the variation in ψ . Equation (5.70) is the stationary-state analogue of Schwinger's principle of stationary action and it will be shown to be a variational statement of the hypervirial theorem; (2) the generalization represented by eqn (5.70) is possible only if the subsystem is bounded by a surface of zero flux in the gradient vectors of the charge density—the topological atom (Srebrenik and Bader 1975; Bader 1988).

The extension of Schrödinger's energy functional to the many-electron case, including the Lagrange multiplier λ is (in analogy with eqn (5.58))

$$\mathcal{G}[\psi] = \int d\tau \{ (\hbar^2/2m) \sum_i \nabla_i \psi^* \cdot \nabla_i \psi + (\hat{V} + \lambda) \psi^* \psi \} \quad (5.71)$$

where, for brevity, the symbol $\int d\tau$ is used here to denote a sum over the spin coordinates and an integration over the spatial coordinates of all N electrons. The definition of the corresponding functional for a subsystem Ω is

$$\mathcal{G}[\psi, \Omega] = \int_{\Omega} d\tau_1 \int d\tau' \{ (\hbar^2/2m) \sum_i \nabla_i \psi^* \cdot \nabla_i \psi + (\hat{V} + \lambda) \psi^* \psi \} \quad (5.72)$$

where \hat{V} denotes the full many-electron potential energy operator. The reader is reminded that the symbol $\int d\tau'$ implies a summation over all spins and the integration over the spatial coordinates of all electrons but one. The symbol $\int_{\Omega} d\tau_1$ implies that the coordinates of electron '1' are integrated over the subsystem Ω . Since ψ is antisymmetrized with respect to the electronic coordinates, all electrons are given equivalent descriptions and it matters not

which set of electronic coordinates is integrated over Ω , just as it matters not which set of electronic coordinates is chosen to define the charge density in eqn (5.1). There is in this procedure no partitioning of kinetic or potential energy operators into sets for different regions, a step which does violate the indistinguishability of the electrons. All electrons occupy the total space of the system, as the limits on the electronic coordinates are $\pm \infty$ for each degree of freedom. The symbol $\langle \psi, \psi \rangle_{\Omega}$ is correspondingly defined as

$$\langle \psi, \psi \rangle_{\Omega} = \int_{\Omega} d\tau_1 \int d\tau' \psi^* \psi. \quad (5.73)$$

The trial functions ϕ representing variations in ψ are given by eqn (5.69) and substitution of $\phi(\mathbf{r})$ for $\psi(\mathbf{r})$ into $\mathcal{G}[\psi, \Omega]$ yields $\mathcal{G}[\phi, \Omega]$. At the point of variation, $\phi = \psi$ and $\mathcal{G}[\phi, \Omega]$ equals $\mathcal{G}[\psi, \Omega]$. The variations $\delta \psi$ and $\delta \psi^*$ are not given prescribed values on any of the boundaries, including the boundary of the subsystem. Instead only the natural boundary condition, that $\nabla_i \psi \cdot \mathbf{n}_i$ and $\nabla_i \psi^* \cdot \mathbf{n}_i$, together with ψ and ψ^* , vanish on all infinite boundaries, will be invoked. The functional $\mathcal{G}[\phi, \Omega]$ is to be varied not only with respect to ϕ , however, but also with respect to the surface defining the subsystem Ω . Only by having the surface itself considered to be a function of ϕ can the definition of the subsystem be determined entirely in a non-arbitrary way by the variational procedure.

The general expression for the variation of an integral including a variation of its surface for variations with respect to ψ is given by

$$\delta_{\psi} \left\{ \int_{\Omega} f(\psi, \nabla \psi) d\tau \right\} = \int_{\Omega} \{ (\partial f / \partial \psi) \delta \psi + (\partial f / \partial \nabla \psi) \delta \nabla \psi \} d\tau + \oint dS(\Omega, \mathbf{r}) f(\psi, \nabla \psi) \delta_{\psi} S(\Omega, \mathbf{r}). \quad (5.74)$$

where $\delta_{\psi} S$ denotes a variation of the surface through variations in ψ . The variation of the surface appears in an integral of f over the surface and this term contributes to the first-order change in the functional (Courant and Hilbert 1953).

No particular difficulties arise in passing from the one- to the many-electron case in the variation of $\mathcal{G}[\psi, \Omega]$. Referring to eqns (5.72) and (5.74), one has

$$\delta \mathcal{G}[\psi, \Omega] = \int_{\Omega} d\tau_1 \int d\tau' \{ (\partial f / \partial \psi) \delta \psi + \sum_i (\partial f / \partial \nabla_i \psi) \delta \nabla_i \psi \} + \oint dS(\Omega, \mathbf{r}_1) \int d\tau' f(\psi, \nabla \psi) \delta S(\Omega, \mathbf{r}_1) + cc. \quad (5.75)$$

In this case $f(\psi, \nabla \psi) \equiv f$ refers to the integrand of eqn (5.72). Only the integral over the surface of the subsystem Ω survives because of the vanishing of ψ and ψ^* on the boundaries at infinity. One again uses an integration by parts based

on the identity

$$\nabla_i \cdot (\nabla_i \psi^* \delta \psi) = \nabla_i^2 \psi^* \delta \psi + \nabla_i \psi^* \cdot \delta \nabla_i \psi$$

to rid the expression of the terms $\delta \nabla_i \psi = \nabla_i \delta \psi$. Each of the surface integrals obtained in this manner vanishes except for the integral over the surface of Ω . This step is illustrated in eqn (5.76) which details the variation for one such term,

$$\begin{aligned} & \int_{\Omega} d\tau_1 \int d\tau' (\partial f / \partial \nabla_i \psi) \delta \nabla_i \psi \\ &= (\hbar^2 / 2m) \int_{\Omega} d\tau_1 \int d\tau' \nabla_i \psi^* \cdot \delta \nabla_i \psi \\ &= -(\hbar^2 / 2m) \int_{\Omega} d\mathbf{r}_1 \int d\tau' \nabla_i^2 \psi^* \delta \psi \\ &+ (\hbar^2 / 2m) \int_{\Omega} d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \oint dS(\mathbf{r}_i) \dots \int d\mathbf{r}_N \nabla_i \psi^* \cdot \mathbf{n}_i \delta \psi. \end{aligned} \quad (5.76)$$

The surface integral in eqn (5.76) comes from the application of Gauss's theorem to the term involving $\nabla_i \cdot (\nabla_i \psi^* \delta \psi)$. As before, all such surface integrals vanish except for $\mathbf{r}_i = \mathbf{r}_1$ because of the vanishing of $\nabla_i \psi^*$ on the boundaries at infinity. From this point on, the coordinate \mathbf{r}_1 and the volume element $d\tau_1$ will be set equal to \mathbf{r} and $d\tau$, respectively, and ∇_1 and ∇_1^2 to their corresponding unscripted quantities.

The expression for the variation of $\mathcal{G}[\psi, \Omega]$ including a variation of the surface $S(\Omega, \mathbf{r})$ is, therefore,

$$\begin{aligned} \delta \mathcal{G}[\psi, \Omega] &= \int_{\Omega} d\tau \int d\tau' \{ \hat{H} \psi^* + \lambda \psi^* \} \delta \psi \\ &+ \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ (\hbar^2 / 2m) \nabla \psi^* \cdot \mathbf{n}(\mathbf{r}) \delta \psi \\ &+ \delta_{\psi} S(\Omega, \mathbf{r}) f(\psi, \nabla \psi) \} + cc \end{aligned} \quad (5.77)$$

where

$$\hat{H} = -(\hbar^2 / 2m) \sum_i \nabla_i^2 + \hat{V}. \quad (5.78)$$

If eqn (5.77) is to be obtained for any variations $\delta \psi$ and $\delta \psi^*$ and any region Ω , it applies to the case where $\Omega = R^3$, i.e. the total system. In this case, all surface terms vanish and $\mathcal{G}[\psi, \Omega]$ and $\delta \mathcal{G}[\psi, \Omega]$ are identical with $\mathcal{G}[\psi]$ and $\delta \mathcal{G}[\psi]$, respectively, and the Euler equations obtained from the variation are Schrödinger's equations

$$\hat{H} \psi = E \psi \quad \text{and} \quad \hat{H} \psi^* = E \psi^* \quad (5.79)$$

with \hat{H} defined as in eqn (5.78) and where $\lambda = -E$. Using eqns (5.79) the expression for $\delta \mathcal{G}[\psi, \Omega]$ reduces at the point of variation to

$$\begin{aligned} \delta \mathcal{G}[\psi, \Omega] &= \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ (\hbar^2 / 2m) \nabla \psi^* \cdot \mathbf{n}(\mathbf{r}) \delta \psi \\ &+ \delta_{\psi} S(\Omega, \mathbf{r}) f(\psi, \nabla \psi) \} + cc. \end{aligned} \quad (5.80)$$

Further progress towards obtaining a general physical result can be made only by removal of the term involving the variation of the surface of the subsystem. Consider, towards this goal, an alternative expression for the integrand $f(\psi, \nabla \psi)$, one involving the Hamiltonian operator \hat{H} . Including the complex conjugate term in eqn (5.80), the integrand $f(\psi, \nabla \psi)$ appears twice and, using eqn (5.46) which relates two forms of the kinetic energy density, one has

$$\begin{aligned} 2f(\psi, \nabla \psi) &= \{ \psi^* \hat{H} \psi + (\hat{H} \psi)^* \psi \} - 2E \psi^* \psi \\ &+ 2(\hbar^2 / 4m) \sum_i \nabla_i^2 (\psi^* \psi). \end{aligned} \quad (5.81)$$

Since Schrödinger's equations are assumed to apply, eqn (5.81) simplifies to.

$$2f(\psi, \nabla \psi) = 2(\hbar^2 / 4m) \sum_i \nabla_i^2 (\psi^* \psi). \quad (5.82)$$

Integration of the right-hand side of eqn (5.82) in the manner indicated in eqn (5.80) transforms it into an integral of the Laplacian of the charge density. A typical term in this integration can be transformed using Gauss's theorem to yield

$$\begin{aligned} & \oint dS(\Omega, \mathbf{r}) \int d\tau_2 \dots \int d\tau_i \nabla_i^2 (\psi^* \psi) \dots \int d\tau_N \\ &= \oint dS(\Omega, \mathbf{r}) \int d\tau_2 \dots \oint dS(\mathbf{r}_i) \nabla_i (\psi^* \psi) \cdot \mathbf{n} \dots \int d\tau_N, \end{aligned} \quad (5.83)$$

and each such term vanishes for $\mathbf{r}_i \neq \mathbf{r}$ since $\nabla_i \psi^*$ and $\nabla_i \psi$ vanish along with ψ and ψ^* on the infinite boundaries. Thus only the term involving the coordinate \mathbf{r} survives the integration over the sum of operators ∇_i^2 . Defining $\rho'(\mathbf{r})$ as the charge density per electron or $\rho(\mathbf{r})/N$,

$$\rho'(\mathbf{r}) = \int d\tau' \psi^* \psi, \quad (5.84)$$

one obtains for the multiple integration of the sum of terms appearing in eqn (5.82) the result

$$2(\hbar^2 / 4m) \oint dS(\Omega, \mathbf{r}) \nabla^2 \rho'(\mathbf{r}).$$

Dividing this contribution equally between the appropriate term appearing explicitly in eqn (5.80) and its complex conjugate yields

$$\begin{aligned} \delta \mathcal{G}[\psi, \Omega] &= \oint dS(\Omega, \mathbf{r}) (\hbar^2 / 2m) \{ \int d\tau' \nabla \psi^* \cdot \mathbf{n}(\mathbf{r}) \delta \psi \\ &+ \frac{1}{2} \delta_{\psi} S(\Omega, \mathbf{r}) \nabla^2 \rho'(\mathbf{r}) \} + cc. \end{aligned} \quad (5.85)$$

Further progress in obtaining a general principle from this expression for $\delta\mathcal{G}[\psi, \Omega]$ is not possible for an arbitrary region of space. We now demonstrate that this result is transformed into the atomic statement of the hypervirial theorem when the subsystem Ω is restricted to one which satisfies a particular variational constraint. We shall introduce the variational constraint in terms of a trial function ϕ and show how the constraint, which limits the variations to a particular class of subsystems, can be applied from the beginning of the variation process, before the point $\phi = \psi$ is attained. The following conditions are to be fulfilled in the constrained variation:

A region $\Omega(\phi)$ is defined in terms of the trial function ϕ that is bound by a zero-flux surface in $\nabla\rho_\phi$

$$\nabla\rho_\phi(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \quad \text{for all } \mathbf{r} \text{ in surface } S(\Omega, \mathbf{r}) \quad (5.86)$$

where the trial density is defined as

$$\rho_\phi(\mathbf{r}) = \int d\tau' \phi^*(\mathbf{x}, \tau') \phi(\mathbf{x}, \tau'). \quad (5.87)$$

Recalling that the topological definition of an atom implies the zero-flux surface condition, a region of space bounded by a surface satisfying eqn (5.86) at the point of variation is henceforth called an atom. It is required that, as ϕ tends to ψ , $\Omega(\phi)$ is continuously deformable into the region $\Omega(\psi)$ associated with the atom. The region $\Omega(\phi)$ thus represents the atom in the varied total system, which is described by the trial function ϕ , just as $\Omega(\psi)$ represents the atom when the total system is in the state described by ψ .

Requiring the fulfillment of these conditions amounts to imposing the variational constraint that the divergence of $\nabla\rho_\phi$ integrates to zero at all stages of the variation, i.e. that

$$L(\phi, \Omega) = -(\hbar^2/4m) \int_{\Omega(\phi)} \nabla^2 \rho_\phi(\mathbf{r}) d\tau = 0 \quad (5.88a)$$

for all admissible ϕ , which implies

$$\delta L(\psi, \Omega) = \delta \left\{ -(\hbar^2/4m) \int_{\Omega} \nabla^2 \rho'(\mathbf{r}) d\tau \right\} = 0. \quad (5.88b)$$

Recalling the expression for the variation of an integral which includes a variation of its surface as given in eqn (5.74) with $f = \nabla^2 \rho$, the variational constraint given in eqn (5.88) leads immediately to a result enabling one to eliminate the term involving the variation of the surface which appears in the general expression for the variation of $\mathcal{G}[\psi, \Omega]$. From eqn (5.74) one has

$$\delta_\psi \left\{ \int_{\Omega} \nabla^2 \rho'(\mathbf{r}) d\tau \right\} = \int \delta_\psi \{ \nabla^2 \rho'(\mathbf{r}) \} d\tau + \oint dS(\Omega, \mathbf{r}) \delta_\psi S(\Omega, \mathbf{r}) \nabla^2 \rho'(\mathbf{r}),$$

and applying the constraint that $\delta L(\psi, \Omega) = 0$ to this result yields

$$(\hbar^2/4m) \oint dS \delta_\psi S(\Omega, \mathbf{r}) \nabla^2 \rho'(\mathbf{r}) = -(\hbar^2/4m) \int_{\Omega} \delta_\psi \{ \nabla^2 \rho'(\mathbf{r}) \} d\tau. \quad (5.89)$$

The variation of $\nabla^2 \rho'$ is easily determined using the recipe given in eqn (5.68). Re-expressing $\nabla^2 \rho'$ in terms of ψ and ψ^* , one has

$$\nabla^2 \rho' = \int d\tau' \nabla^2 (\psi^* \psi) = \int d\tau' \{ (\nabla^2 \psi^*) \psi + 2\nabla\psi^* \cdot \nabla\psi + \psi^* \nabla^2 \psi \} \quad (5.90)$$

Hence the variation is with respect to ψ , $\nabla\psi$, and $\nabla^2 \psi$ with the result

$$\begin{aligned} \int_{\Omega} \delta_\psi \{ \nabla^2 \rho'(\mathbf{r}) \} d\tau &= \int_{\Omega} d\tau \int d\tau' \{ \nabla^2 \psi^* \delta\psi \\ &\quad + 2\nabla\psi^* \cdot \delta\nabla\psi + \psi^* \delta\nabla^2 \psi \}. \end{aligned} \quad (5.91)$$

The Laplacian of ρ is what is termed a divergence expression—its variation yields only surface terms. The integral in eqn (5.91) can be expressed as the divergence of a sum of terms

$$\int_{\Omega} \delta_\psi \{ \nabla^2 \rho'(\mathbf{r}) \} d\tau = \int_{\Omega} d\tau \int d\tau' \nabla \cdot \{ \nabla\psi^* \delta\psi + \psi^* \nabla\delta\psi \},$$

and by Gauss's theorem this can be re-expressed as a surface integral. Thus, the result of imposing the variational constant given in eqn (5.88b) is to replace the integral of the surface variation appearing in the expression for $\delta\mathcal{G}[\psi, \Omega]$ (eqn (5.85)) with the result appearing on the right-hand side of eqn (5.92),

$$\begin{aligned} (\hbar^2/4m) \oint dS \delta_\psi S(\Omega, \mathbf{r}) \nabla^2 \rho'(\mathbf{r}) \\ = -(\hbar^2/4m) \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ (\nabla\psi^*) \delta\psi + \psi^* \delta\nabla\psi \} \cdot \mathbf{n}(\mathbf{r}). \end{aligned} \quad (5.92)$$

It is important that no new volume contributions to the variation be obtained by the imposition of the constraint given in eqn (5.88b), or one would no longer obtain Schrödinger's equations as the Euler-Lagrange equation of the variation.

Substitution of the identity in eqn (5.92) into eqn (5.85) and combining terms yields the final result for the variation of the functional $\mathcal{G}[\psi, \Omega]$,

$$\begin{aligned} \delta\mathcal{G}[\psi, \Omega] &= (\hbar^2/4m) \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ \nabla\psi^* \delta\psi \\ &\quad - \psi^* \delta\nabla\psi \} \cdot \mathbf{n}(\mathbf{r}) + \text{cc.} \end{aligned} \quad (5.93)$$

Equation (5.93) is a statement of a physical principle as it can be restated in terms of the flux in the infinitesimal change in vector current density through the surface bounding the atom. In analogy with the definition of the charge density, the single-particle vector current density, $\mathbf{j}(\mathbf{r})$, of a many-particle

system is defined as

$$\mathbf{j}(\mathbf{r}) = (\hbar/2mi) \int d\tau' \{ \psi^* \nabla \psi - (\nabla \psi^*) \psi \}. \quad (5.94)$$

The infinitesimal change in $\mathbf{j}(\mathbf{r})$ caused by a variation in ψ is

$$\delta_\psi \mathbf{j}(\mathbf{r}) = (\hbar/2mi) \int d\tau' \{ \psi^* \delta \nabla \psi - (\nabla \psi^*) \delta \psi \} \quad (5.95)$$

or, equivalently,

$$\delta_\psi \mathbf{j}(\mathbf{r}) = (\hbar/2mi) \int d\tau' \{ \psi^* \nabla (\delta \psi) - (\nabla \psi^*) \delta \psi \},$$

and the expression for the variation of the subsystem energy functional constrained to a region bounded by a zero-flux surface becomes

$$\delta \mathcal{G}[\psi, \Omega] = - (i\hbar/2) \oint dS(\Omega, \mathbf{r}) \delta_\psi \mathbf{j}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) + cc. \quad (5.96)$$

To demonstrate that the expression for $\delta \mathcal{G}[\psi, \Omega]$ is equivalent to a variational derivation of the hypervirial theorem for an atom in a molecule, it is necessary to take one more very important step which is to consider the variations in ψ to be generated by the action of an operator on ψ . That is, one makes the identifications

$$\delta \psi = - \varepsilon (i/\hbar) \hat{G} \psi \quad \text{and} \quad \delta \psi^* = \varepsilon (i/\hbar) \hat{G} \psi^* \quad (5.97)$$

where \hat{G} is a linear Hermitian operator, an observable, and ε denotes an infinitesimal. At this point we shall use these expressions to substitute for $\delta \psi$ and $\delta \psi^*$ in the expression for $\delta \mathcal{G}[\psi, \Omega]$ and simply state that their use corresponds to identifying the variations in the state function with the action of generators of infinitesimal unitary transformations. Such operators, when acting on a state function, cause changes in the dynamical properties of a system. It is this identification which enabled Schwinger to transform the variation of the action integral into the more powerful principle of stationary action. These ideas are fully developed in Chapter 8 where it will also be shown that the final result to be obtained for $\delta \mathcal{G}[\psi, \Omega]$, using eqn (5.97), is a statement of Schwinger's principle for a stationary state as it applies to an atom in a molecule.

We proceed by defining a current density for the property associated with the operator \hat{G} in analogy with the current \mathbf{j}_A defined in eqn (5.42)

$$\mathbf{j}_G(\mathbf{r}) = (\hbar/2mi) \int d\tau' \{ \psi^* \nabla (\hat{G} \psi) - (\nabla \psi^*) (\hat{G} \psi) \}. \quad (5.98)$$

In terms of this vector current the variation in the atomic energy functional is given by

$$\delta \mathcal{G}[\psi, \Omega] = - (\varepsilon/2) \{ \oint dS(\Omega, \mathbf{r}) \mathbf{j}_G(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) + cc \}. \quad (5.99)$$

Using the expression derived previously for a subsystem statement of the hypervirial theorem, eqn (5.43), we arrive at the atomic statement of the

principle of stationary action for a stationary state

$$\delta \mathcal{G}[\psi, \Omega] = - (\varepsilon/2) \{ (i/\hbar) \langle \psi, [\hat{H}, \hat{G}] \psi \rangle_\Omega + cc \}. \quad (5.100)$$

Equation (5.100) is a variational result and it applies only to regions of space bounded by surfaces which satisfy the variational constraint of exhibiting a zero-flux in the gradient vector field of the charge density. It applies, therefore, to an atom, to any linked grouping of atoms, and to the total system. Equation (5.100) is a generalization of Schrödinger's original derivation of his equations to yield the same equations together with the variational statement of the hypervirial theorem. When Ω equals the total space of the system, one has

$$\delta \mathcal{G}[\psi, \Omega] = \delta \mathcal{G}[\psi] = - (\varepsilon/2) \{ (i/\hbar) \langle \psi, [\hat{H}, \hat{G}] \psi \rangle + cc \} = 0, \quad (5.101)$$

a result which yields not only Schrödinger's equations but, in addition, the hypervirial theorem

$$\langle \psi, [H, A] \psi \rangle = 0 \quad (5.102)$$

and hence the variational development of all the mechanical theorems derivable from it. For an atom, the variation in $\mathcal{G}[\psi, \Omega]$ is proportional to the flux in the vector current density of the generator \hat{G} (eqn (5.99)) and eqn (5.100) is consequently a variational derivation of the subsystem hypervirial theorem. Thus, for an atom, or any functional grouping of atoms, one obtains, in analogy with the total system, Schrödinger's equations and a *variational statement* of the hypervirial theorem, a statement that applies uniquely to systems bounded by zero-flux surfaces, which for any operator \hat{G} is

$$\{ (i/\hbar) \langle \psi, [\hat{H}, \hat{G}] \psi \rangle_\Omega + cc \} = \{ \int dS(\Omega, \mathbf{r}) \mathbf{j}_G(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) + cc \}. \quad (5.103)$$

Through requisite choices for the operator \hat{G} , eqn (5.103) determines the force acting on an atom in a molecule and, through the atomic statement of the virial theorem, its energy. It establishes the mechanics of an atom in a molecule, as is demonstrated in the next chapter.

E5.1 The variation of Hamiltonian-based functionals

Schrödinger's use of the expression $(1/2m) \langle \hat{\mathbf{p}} \psi, \hat{\mathbf{p}} \psi \rangle$ to express the kinetic energy in his functional $\mathcal{S}[\psi]$ was based on an analogy with the Hamilton-Jacobi equation of classical mechanics. This is also the form used for the kinetic energy as it appears in the Lagrangian of the action integral, a consequence of the Lagrangian, either quantum or classical, being a function of the field variables and their first derivatives only. Since the Lagrangian degenerates into a functional of the form $\mathcal{S}[\psi]$ for a stationary state, these are termed Lagrangian-based functionals. One may, however, construct an

energy functional based upon the Hamiltonian operator which differs from the above by expressing the kinetic energy in terms of the operator $-(\hbar^2/2m)\nabla^2$. We investigate the variational properties of such a functional here. In this section the normalization constant will be handled by division with $\langle\psi, \psi\rangle$ as in eqn (5.63) for $\mathcal{E}[\psi]$.

In analogy with $\mathcal{E}[\psi]$ and its many-electron subsystem counterpart, the atomic form of the energy functional $\mathcal{E}[\psi, \Omega]$ is defined as

$$\mathcal{E}[\psi, \Omega] = \int_{\Omega} d\tau \int d\tau' \{ (\hbar^2/2m) \sum_i \nabla_i \psi^* \cdot \nabla_i \psi + \hat{V} \psi^* \psi \} / \langle \psi, \psi \rangle_{\Omega} \quad (\text{E5.1})$$

where

$$\langle \psi, \psi \rangle_{\Omega} = \int_{\Omega} d\tau \int d\tau' \psi^* \psi. \quad (\text{E5.2})$$

The Hamiltonian-based functional $\mathcal{E}'[\psi, \Omega]$ is defined as

$$\mathcal{E}'[\psi, \Omega] = \frac{1}{2} \{ \mathcal{H}[\psi, \Omega] + \mathcal{H}^*[\psi, \Omega] \} \quad (\text{E5.3})$$

where

$$\mathcal{H}[\psi, \Omega] = \int_{\Omega} d\tau \int d\tau' \{ -(\hbar^2/2m) \sum_i \psi^* \nabla_i^2 \psi + \hat{V} \psi^* \psi \} / \langle \psi, \psi \rangle_{\Omega}. \quad (\text{E5.4})$$

The Lagrangian-based functionals employ the form $(1/2m) \langle \hat{\mathbf{p}}\psi, \hat{\mathbf{p}}\psi \rangle$ for the kinetic energy and are, therefore, necessarily real. The functional $\mathcal{H}[\psi, \Omega]$ can, however, be complex, as \hat{H} is not necessarily Hermitian over a subsystem as shown in eqn (5.12). Since one cannot have a complex quantity stationary with respect to a variation, one must vary $\mathcal{H}[\psi, \Omega]$ together with its complex conjugate as defined in eqn (E5.3).

The two types of functionals can be related using eqn (5.46)

$$\mathcal{E}[\psi, \Omega] = \mathcal{E}'[\psi, \Omega] + (\hbar^2/4m) \int_{\Omega} d\tau \nabla^2 \rho(\mathbf{r}) / \langle \psi, \psi \rangle_{\Omega} \quad (\text{E5.5})$$

where only the term involving the coordinate \mathbf{r} survives the integration over the sum of terms $\sum_i \nabla_i^2 (\psi^* \psi)$. For a total system, or for a subsystem satisfying the zero-flux surface condition, one has

$$\mathcal{E}[\psi, \Omega] = \mathcal{E}'[\psi, \Omega] = \mathcal{E} \quad (\text{E5.6a})$$

where \mathcal{E} is the energy eigenvalue for a stationary state. Thus the zero-flux surface condition insures that the real part of $\mathcal{H}[\psi, \Omega]$ equals $\mathcal{E}[\psi, \Omega]$. Because of the constraint imposed on the variation of $\mathcal{E}[\psi, \Omega]$ (eqn (5.88b)),

one also has

$$\delta \mathcal{E}[\psi, \Omega] = \delta \mathcal{E}'[\psi, \Omega] \quad (\text{E5.6b})$$

and the variations of both functionals yield identical results. Since the function $\sum_i \nabla_i^2 (\psi^* \psi)$ is a divergence expression, its addition to the functional $\mathcal{E}[\psi, \Omega]$ to obtain $\mathcal{E}'[\psi, \Omega]$ does not change the Euler equation obtained from the variation. However, as shown below, while the variation of $\mathcal{E}'[\psi, \Omega]$ yields the same results as does $\delta \mathcal{E}[\psi, \Omega]$, it does so without placing restrictions on the surface of Ω .

The variation of $\mathcal{E}[\psi, \Omega]$ with the imposition of the variational constants given in eqn (5.88) at every stage of the variation and including a variation of the surface yields

$$\delta \mathcal{E}[\psi, \Omega] = (\hbar^2/4m) \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ \nabla \psi^* \delta \psi - \psi^* \delta \nabla \psi \} \cdot \mathbf{n} / \langle \psi, \psi \rangle_{\Omega} + \text{cc} \quad (\text{E5.7})$$

in analogy with the result obtained for the variation of $\mathcal{G}[\psi, \Omega]$ given in eqn (5.93).

The variation of $\mathcal{E}'[\psi, \Omega]$ involves variation with regard to $\delta \psi$ and $\delta \nabla^2 \psi$. It is necessary to perform an integration by parts twice in succession to rid the expression of the variations of gradients of ψ . This procedure is illustrated below for a single term in $\mathcal{H}[\psi, \Omega]$

$$\begin{aligned} & -(\hbar^2/2m) \{ \partial(\psi^* \nabla_i^2 \psi) / \partial \nabla_i^2 \psi \} \delta \nabla_i^2 \psi \\ & = -(\hbar^2/2m) \psi^* \delta \nabla_i^2 \psi \\ & = -(\hbar^2/2m) (-\nabla \psi^* \cdot \delta \nabla \psi + \nabla \cdot (\psi^* \delta \nabla \psi)) \\ & = -(\hbar^2/2m) (+\nabla^2 \psi^* \delta \psi + \nabla \cdot (\psi^* \delta \nabla \psi - \nabla \psi^* \delta \psi)). \end{aligned} \quad (\text{E5.8})$$

Using this result, the variation of $\mathcal{E}'[\psi, \Omega]$ including a variation of the surface of Ω is found to be

$$\begin{aligned} & \langle \psi, \psi \rangle_{\Omega} \delta \mathcal{E}'[\psi, \Omega] \\ & = \frac{1}{2} \int_{\Omega} d\mathbf{r} \int d\tau' \{ \hat{H} \psi^* - \mathcal{H}^*[\psi, \Omega] \psi^* \} \delta \psi \\ & \quad + \left(\frac{1}{2} \right) \oint dS(\Omega, \mathbf{r}) \int d\tau' \delta_{\psi} S(\Omega, \mathbf{r}) \{ \psi^* \hat{H} \psi - \mathcal{H}^*[\psi, \Omega] \psi^* \} \\ & \quad + (\hbar^2/4m) \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ \nabla \psi^* \delta \psi - \psi^* \delta \nabla \psi \} \cdot \mathbf{n} + \text{cc}. \end{aligned} \quad (\text{E5.9})$$

At the point of variation with $\mathcal{H}[\psi, \Omega] = E$, one obtains Schrödinger's

equations as the Euler equations and

$$\delta \mathcal{E}'[\psi, \Omega] = (\hbar^2/4m) \oint dS(\Omega, \mathbf{r}) \int d\tau' \{ \nabla \psi^* \delta \psi - \psi^* \delta \nabla \psi \} \cdot \mathbf{n} / \langle \psi, \psi \rangle_{\Omega} + \text{cc.} \quad (\text{E5.10})$$

This result, which is obtained without any restrictions on Ω is identical to that obtained from the variation of $\mathcal{E}[\psi, \Omega]$ by constraining the subsystem to be bounded by a zero-flux surface. It is clear from eqn (E5.9) that, unlike the variation of $\mathcal{G}[\psi, \Omega]$ or $\mathcal{E}[\psi, \Omega]$, one obtains the same result for the variation of $\mathcal{E}'[\psi, \Omega]$ whether or not the surface is varied.

The Lagrangian-based functional $\mathcal{G}[\psi, \Omega]$ or $\mathcal{E}[\psi, \Omega]$ derives directly from the Lagrangian as employed in the quantum action principle. For a total system, both the Lagrangian- and Hamiltonian-based functionals yield identical variational results. This equivalence in variational behaviour is maintained for the corresponding subsystem functionals only if the subsystem is bounded by a zero-flux surface. Only an atomic region ensures an equivalence in both the values and the variational properties of the two types of functionals (eqns (E5.6a,b)) thereby preserving the properties obtained for a total system.

E5.2 Vectors, tensors, and dyadics

A vector in three-dimensional Cartesian space is characterized by three components

$$\mathbf{A} = iA_x + jA_y + kA_z \quad (\text{E5.11})$$

where \mathbf{i} , \mathbf{j} , and \mathbf{k} are unit vectors along x , y , and z , respectively. Equation (E5.11) can be expressed as a product of a row and a column matrix as

$$\mathbf{A} = [\mathbf{i} \quad \mathbf{j} \quad \mathbf{k}] \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} = iA_x + jA_y + kA_z. \quad (\text{E5.12})$$

Because of the orthogonality of the unit vectors in a scalar or dot product, the scalar product of two vectors \mathbf{A} and \mathbf{B} can be expressed as

$$\mathbf{A} \cdot \mathbf{B} = [A_x \quad A_y \quad A_z] \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} = A_x B_x + A_y B_y + A_z B_z. \quad (\text{E5.13})$$

A dot product of two vectors yields a single number, a scalar. It is possible to generate a three-by-three matrix by the multiplication of two vectors. This

procedure is denoted by the symbol \mathbf{AB} which defines the operation

$$\mathbf{AB} = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} [B_x \quad B_y \quad B_z] = \begin{pmatrix} A_x B_x & A_x B_y & A_x B_z \\ A_y B_x & A_y B_y & A_y B_z \\ A_z B_x & A_z B_y & A_z B_z \end{pmatrix}. \quad (\text{E5.14})$$

The result of this operation is called a dyadic or tensor

$$\vec{\mathbf{D}} = \mathbf{AB}.$$

A dyadic behaves like a two-headed vector — the scalar product of $\vec{\mathbf{D}}$ with a vector \mathbf{A} yields a vector

$$\begin{aligned} \vec{\mathbf{D}} \cdot \mathbf{A} &= \mathbf{C}, \\ C_x &= D_{xx}A_x + D_{xy}A_y + D_{xz}A_z, \\ C_y &= D_{yx}A_x + D_{yy}A_y + D_{yz}A_z, \\ C_z &= D_{zx}A_x + D_{zy}A_y + D_{zz}A_z. \end{aligned} \quad (\text{E5.15})$$

A scalar is obtained only by taking a scalar product of $\vec{\mathbf{D}}$ from both sides

$$\mathbf{B} \cdot \vec{\mathbf{D}} \cdot \mathbf{A} = \mathbf{B} \cdot \mathbf{C}. \quad (\text{E5.16})$$

A dyadic is required to describe those directed (vectorial) properties of a system which result from the application of a force or field along directions orthogonal to the observed resultant. The polarizability of a molecule is, for example, described by the polarizability tensor α . The dipole μ induced by an applied field \mathbf{E} is given by

$$\mu = \alpha \cdot \mathbf{E}. \quad (\text{E5.17})$$

Even if \mathbf{E} is directed along the z -axis, $\mathbf{E} = kE_z$, the x - and y -components of μ can differ from zero. Using eqn (E5.15), the three components of μ are given by

$$\begin{aligned} \mu_x &= \alpha_{xz}E_z, \\ \mu_y &= \alpha_{yz}E_z, \\ \mu_z &= \alpha_{zz}E_z. \end{aligned}$$

If a dyadic is symmetric, then

$$\mathbf{A} \cdot \vec{\mathbf{D}} = \vec{\mathbf{D}} \cdot \mathbf{A},$$

a property possessed by the stress tensor $\vec{\sigma}$ (eqn (5.28)). An example of an unsymmetrical dyadic is provided by the product \mathbf{jr} as encountered in the derivation of Ehrenfest's first relationship, eqn (5.26). The identity used in the derivation of this relation

$$\mathbf{V} \cdot (\mathbf{jr}) = \mathbf{j} + \mathbf{rV} \cdot \mathbf{j}$$

is obtained from the matrix multiplication

$$\nabla \cdot (\mathbf{j}\mathbf{r}) = \begin{bmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \end{bmatrix} \begin{pmatrix} j_x x & j_x y & j_x z \\ j_y x & j_y y & j_y z \\ j_z x & j_z y & j_z z \end{pmatrix}.$$

In tensor notation, an element of the tensor $\mathbf{j}\mathbf{r}$ is

$$T_{k1} = j_k x_1$$

and

$$\frac{\partial T_{k1}}{\partial x_k} = j_{kk} x_1 + j_k x_{1k}.$$

Two types of terms appear in the stress tensor $\vec{\sigma}$ in eqn (5.28), terms of the form $\psi^* \nabla \nabla \psi$ and $\nabla \psi^* \nabla \psi$. The divergence of this tensor determines the force $-\nabla \cdot \vec{\sigma}$. Taking the divergence of $\vec{\sigma}$ yields terms of the form

$$\nabla \cdot (\psi^* \nabla \nabla \psi) = \nabla \psi^* \cdot \nabla \nabla \psi + \psi^* \nabla (\nabla^2 \psi)$$

and

$$\nabla \cdot (\nabla \psi^* \nabla \psi) = \nabla^2 \psi^* \nabla \psi + \nabla \psi^* \cdot \nabla \nabla \psi.$$

The subtraction of these two contributions as they appear in $\nabla \cdot \vec{\sigma}$ yields the terms appearing in the equation for the force preceding eqn (5.28).

E5.3 Divergence of a vector and Gauss's theorem

Consider an element of surface area dS and define \mathbf{n} to be the outwardly directed unit vector normal to dS . Then $\mathbf{j} \cdot \mathbf{n}$ is the component of the vector current density normal to the surface element dS and the current through this element is $\mathbf{j} \cdot \mathbf{n} dS$. The total current out of some region Ω bounded by a surface $S(\Omega)$ is obtained by summing the contributions from all its surface elements. The resulting surface integral is called 'the flux \mathbf{j} through the surface $S(\Omega)$ ', that is

$$\text{Flux of } \mathbf{j} \text{ through the surface } S(\Omega) = \oint dS(\Omega, \mathbf{r}) \mathbf{j}(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}). \quad (\text{E5.18})$$

This expression is applied to the corresponding surface integral of any vector quantity, even if there is no physical flow through the surface. Thus references in the text to a 'flux in the gradient vector of ρ ' do not imply any physical flow.

One can imagine the region Ω bounded by $S(\Omega)$ to be divided into an arbitrary number of smaller volumes. The total flux through $S(\Omega)$ is still given by eqn (E5.18) since this total flux is equal to the sum of the fluxes out of all the interior volumes. Consider, for example, a surface S_{ik} between two interior regions Ω_i and Ω_k . The flux out of Ω_i through S_{ik} is $\oint dS_{ik} \mathbf{j} \cdot \mathbf{n}_i$ which must equal the negative of the flux out of Ω_k , $\oint dS_{ik} \mathbf{j} \cdot \mathbf{n}_k$, since $\mathbf{n}_i = -\mathbf{n}_k$. Thus the flux

out of one region equals the flux into the neighbouring region through their common surface.

Gauss's theorem is proved by considering the flux out of an infinitesimally small cube. Consider a cube with edges of lengths Δx , Δy , and Δz and denote the coordinates of the corner nearest the origin as (x, y, z) . The flux of a vector, say \mathbf{j} , through the surface of the cube is equal to the fluxes through each of the six faces. The outward flux through the face $\Delta y \Delta z$ which is perpendicular to the x -axis in the negative x direction, called face (1), is $-j_x(x) \Delta y \Delta z$ where $j_x(x)$ is the x -component of \mathbf{j} . The flux out of the corresponding face at $x + \Delta x$, face (2), is $j_x(x + \Delta x) \Delta y \Delta z$. The component of j_x at $x + \Delta x$ is slightly different from j_x at x and to first-order one has

$$j_x(x + \Delta x) = j_x(x) + \frac{\partial j_x}{\partial x} \Delta x.$$

The sum of the fluxes through faces (1) and (2) is

$$\text{Flux out of (1) and (2)} = \frac{\partial j_x}{\partial x} \Delta x \Delta y \Delta z.$$

The total flux through all six faces of the cube is, therefore,

$$\text{Total flux in } \mathbf{j} \text{ out of cube} = \left(\frac{\partial j_x}{\partial x} + \frac{\partial j_y}{\partial y} + \frac{\partial j_z}{\partial z} \right) \Delta x \Delta y \Delta z.$$

This total flux is also given by the surface integral of $\mathbf{j} \cdot \mathbf{n}$ and, hence,

$$\oint dS \mathbf{j} \cdot \mathbf{n} = \nabla \cdot \mathbf{j} \Delta V$$

where ΔV is the volume of the cube. Thus, the divergence of a vector at a given point in space is the net outflow or flux of \mathbf{j} per unit volume in the neighbourhood of the point. From the above demonstration that the total flux from a volume is the sum of the fluxes out of each of its parts, we see that the flux through a surface $S(\Omega)$ bounding a finite volume Ω is given by the sum of the fluxes out of each of its infinitesimal elements dV , or

$$\oint dS(\Omega) \mathbf{j} \cdot \mathbf{n} = \int_{\Omega} \nabla \cdot \mathbf{j} dV \quad (\text{E5.19})$$

which is Gauss's theorem.

References

- Bader, R. F. W. (1988). *Pure appl. Chem.* **60**, 145.
 Bader, R. F. W. and Beddall, P. M. (1972). *J. Chem. Phys.* **56**, 3320.
 Bader, R. F. W. and Preston, H. J. T. (1969). *Int. J. quantum Chem.* **3**, 327.
 Cohen, L. (1979). *J. Chem. Phys.* **70**, 788.
 Coulson, C. A. and Longuet-Higgins, H. C. (1947a). *Proc. r. Soc., London* **A191**, 39.

- Coulson, C. A. and Longuet-Higgins, H. C. (1947b). *Proc. r. Soc., London* **A192**, 16.
- Coulson, C. A. and Longuet-Higgins, H. C. (1948a). *Proc. r. Soc., London* **A193**, 447.
- Coulson, C. A. and Longuet-Higgins, H. C. (1948b). *Proc. r. Soc., London* **A193**, 456.
- Coulson, C. A. and Longuet-Higgins, H. C. (1948c). *Proc. r. Soc., London* **A195**, 188.
- Courant, R. and Hilbert, D. (1953). *Methods of mathematical physics*, Vol. 1. Wiley Interscience, New York.
- Dirac, P. A. M. (1929). *Proc. r. Soc., London* **A123**, 714.
- Epstein, S. T. (1974a) *The variation method in quantum chemistry*. Academic Press, New York.
- Epstein, S. T. (1974b). *J. Chem. Phys.* **60**, 3351.
- Feynman, R. P. (1948). *Rev. mod. Phys.* **20**, 367.
- Feynman, R. P., Leighton, R. B., and Sands, M. (1964). *The Feynman lectures on physics*, Vol. II. Addison-Wesley, Reading, Massachusetts.
- Hall, G. G. (1951). *Proc. r. Soc., London* **A205**, 511.
- Hirschfelder, J. O. (1960). *J. Chem. Phys.* **33**, 1462.
- Hund, F. (1933). *Physik* **1**, 163.
- Hund, F. (1937). *Physik* **5**, 1.
- Hurley, A. C. (1954). *Proc. r. Soc., London* **A226**, 179.
- Lennard-Jones, J. E. (1949a). *Proc. r. Soc., London* **A198**, 1.
- Lennard-Jones, J. E. (1949b). *Proc. r. Soc., London* **A198**, 14.
- Lennard-Jones, J. E. (1952). *J. Chem. Phys.* **20**, 1024.
- Lewis, G. N. (1916). *J. Am. chem. Soc.* **38**, 762.
- Lewis, G. N. (1933). *J. Chem. Phys.* **1**, 17.
- Löwdin, P.-O. (1948). *A theoretical investigation into some properties of ionic crystals*. Almqvist and Wiksell, Stockholm.
- Löwdin, P.-O. (1956). *Phil. Mag. Suppl.* **5**, 1.
- Messiah, A. (1958). *Quantum mechanics*, Vol. I. Wiley, New York.
- Moffitt, W. (1951). *Proc. r. Soc., London* **A210**, 245.
- Mulliken, R. S. (1928a). *Phys. Rev.* **32**, 186.
- Mulliken, R. S. (1928b). *Phys. Rev.* **32**, 388.
- Mulliken, R. S. (1928c). *Phys. Rev.* **32**, 761.
- Mulliken, R. S. (1932a). *Phys. Rev.* **40**, 55.
- Mulliken, R. S. (1932b). *Rev. mod. Phys.* **4**, 1.
- Mulliken, R. S. (1935). *J. Chem. Phys.* **3**, 375.
- Mulliken, R. S. (1939). *Phys. Rev.* **56**, 778.
- Pauli, W. (1958) In *Encyclopedia of physics*, Vol. 5, part 1 (ed. S. F. Lügge). Springer, Berlin.
- Pauling, L. (1931). *J. Am. chem. Soc.* **53**, 1367.
- Pauling, L. (1960). *The nature of the chemical bond* (3rd edn). Cornell University Press, Ithaca, New York.
- Roothaan, C. C. J. (1951). *Rev. mod. Phys.* **23**, 69.
- Schrödinger, E. (1926). *Ann. d. Phys.* **79**, 361.
- Schrödinger, E. (1927). *Ann. d. Phys.* **82**, 265.
- Schwinger, J. (1951). *Phys. Rev.* **82**, 914.
- Slater, J. C. (1931). *Phys. Rev.* **37**, 481.
- Srebrenik, S. and Bader, R. F. W. (1974). *J. Chem. Phys.* **61**, 2536.
- Srebrenik, S. and Bader, R. F. W. (1975). *J. Chem. Phys.* **63**, 3945.

THE MECHANICS OF AN ATOM IN A MOLECULE

I hope and believe that the present statements will prove useful in the elucidation of the magnetic properties of atoms and molecules, and further for explaining the flow of electricity in solid bodies.

— Erwin Schrödinger (1926). The statement following his introduction of the charge and current densities and the quantum equation of continuity in his fourth paper on 'wave mechanics'.

6.1 An atomic view of the properties of matter

6.1.1 The charge and current densities

Schrödinger's belief in the ability of his 'wave mechanics' to provide a description of the properties of matter as demonstrated in the opening quotation was well founded and his mechanics has been used to understand phenomena unknown in 1926. The roles of the charge and current densities in quantum mechanics go beyond their use in the description of the electrical and magnetic properties of matter. The charge density provides a description of the distribution of charge throughout real space and is the bridge between the concept of state functions in Hilbert space and the physical model of matter in real space. Schrödinger subscribed to this same view, pointing out in his fourth paper wherein he defines the electric density at a point in real space, 'that the ψ function itself cannot be and may not be directly interpreted in terms of three-dimensional space—because it is in general a function in configuration space, and not real space.' The zero-flux boundary condition, as stated in terms of the gradient vector field of the charge density, defines a single atom and objects made from collections of atoms. As demonstrated in this chapter, the form and ultimately the properties of an object, microscopic or macroscopic, are determined by its distribution of charge.

A subsystem is an open system, free to exchange charge and momentum with its environment. Thus the current density \mathbf{j}_G for any observable \hat{G} is of particular importance in the mechanics of a subsystem, since a non-vanishing flux in this current implies a fluctuation in the subsystem average value of the property G .

Variation of the energy functional for a subsystem is equal to an infinitesimal times the flux in the current density of the generator \hat{G} through the