TOPOLOGICAL AND QUANTUM MECHANICAL STUDY OF MOLECULAR CHARGE DISTRIBUTIONS

By

© THANH-TUNG NGUYEN-DANG, Lic.es Sc.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
July 1980

TOPOLOGICAL AND QUANTUM MECHANICAL STUDY OF MOLECULAR CHARGE DISTRIBUTIONS

DOCTOR OF PHILOSOPHY (1980)
(Chemistry)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Topological and Quantum Mechanical Study of Molecular Charge Distributions

AUTHOR: Thanh-Tung Nguyen-Dang, Lic. ès Sc. (Université de Liège, Liège, Belgium)

SUPERVISOR: Professor R.F.W. Bader

NUMBER OF PAGES: xi, 213

ACKNOWLEDGEMENTS

I wish to thank all the members of the Theoretical Chemistry Laboratory for providing me with the opportunity of numerous constructive discussions. Special thanks are addressed to Dr. Steven G. Anderson for his assistance in the computational work, to Dr. Yoram Tal for many suggestions and stimulating discussions, as well as for his frequent encouragement.

During the course of this work, I have had the chance to appreciate the teachings of Dr. R.F.W. Bader, on scientific matters, as well as on the value of conscientious and honest labour. I wish to express my gratitude towards Dr. Bader, for his guidance, his encouragement, and for the confidence he often expresses regarding my scientific judgment.

The typing of this work has been done superbly, and with expertise, by Mrs. Jan Gallo. Many thanks go to her, in particular for her patience and her minutious care in handling this manuscript.

I thank all my friends, for their love and care.

I dedicate this work to my parents, my brother, my sisters, and my beloved Bich-Thuy.

ABSTRACT

This thesis is concerned with the rigorous definitions of the two central concepts embodied in the notion of molecular structure, namely, the concepts of atoms and bonds. The basis for the present approach is provided by the topological properties of the charge distribution in a given molecular system. The essential observation is that the only local maxima of a ground state distribution occur at the positions of the nuclei. The nuclei are, therefore, identified as point attractors of the gradient vector field of the charge density. The associated basins partition the molecule into atomic fragments. Each atom is a stable structural unit defined as the union of an attractor and its basin. The common boundary of two neighbouring atomic fragments, the interatomic surface, contains a particular critical point, which generates a pair of gradient paths linking the two neighbouring attractors. The union of this pair of gradient paths and their endpoints is called a bond path. The network of bond paths defines a molecular graph of the system.

Having defined a unique molecular graph for any molecular geometry, the total configuration space is partitioned into a finite number of regions. Each region is associated with a particular structure defined as an equivalence class of molecular graphs. A chemical reaction in which chemical bonds are broken and/or formed is, therefore, a trajectory

in configuration space which must cross one of the boundaries between two neighbouring structural regions. These boundaries form the catastrophe set of the system which, like a phase diagram in thermodynamics, denotes the points of "balance" between neighbouring structures. A general analysis of the structural changes in an ABC type system is given in detail, together with specific examples of all possible structural elements in a molecular system.

The topological definition of an atom as the union of a nuclear attractor and its basin is equivalent to its definition in terms of its boundary. An atomic boundary consists of the union of a number of surfaces through which the flux of the gradient vector of the charge density is zero. Such a surface is called a zero-flux surface. In general, the zero-flux surfaces define subsystems of a molecular system which represent atomic fragments or functional groupings of such fragments. The properties of these subsystems are delineated through a variational approach to the quantum mechanics of molecular subsystems. In both the time-dependent and the stationary state case, the variational equations for a subsystem are generalizations of the corresponding expressions for a total system. The stationary-state variational principle relates the first-order change in the subsystem energy functional to the flux, through the boundary of the subsystem, of the current density generated by an infinitesimal arbitrary variation in the state function. From this principle, the generalizations of the important hypervirial and virial theorems are obtained, which apply equally to a subsystem as to the total system. In particular, a definition of the subsystem energy, which exhibits the important additivity property, is obtained through the subsystem virial theorem. The time-dependent variational equations generalize the stationary state variational principle; they are shown to follow from Schwinger's quantum action principle and principle of stationary action. The application of these principles to a subsystem requires that the subsystem quantum action integral be stationary in the sense that its first variation, ensuing from arbitrary variations in the state function between two given time end-points, contains only contributions from these time end-points and from the spatial boundary of the subsystem, at intermediate times. When the aforementioned variations are produced by an infinitesimal unitary transformation, the subsystem variational principle relates the change in the subsystem Lagrangian to the commutator of the generator of the transformation with the Hamiltonian of the system. From this statement of the time-dependent variational principle, a force law is obtained, which generalizes Ehrenfest's theorem, and can be regarded as the quantum analogue of Newton's equation of motion. The variational principle also leads to a generalization of the continuity equation, which governs the time change of an atomic population, and an extension of the virial theorem to describe a subsystem - an atom, or a functional grouping of atoms - in a time-dependent molecular environment.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

INTRODUCTION

Feynman once wrote:[1] "If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generations of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis (or the atomic fact, or whatever you wish to call it) that all things are made of atoms - little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another." The atomic hypothesis referred to above by Feynman has its origin in the speculations of the greek philosophers of the antiquity. However, it was not until the end of the eighteenth century and the beginning of the nineteenth century, that the atomic hypothesis was revived to allow for the rationalization of the body of chemical facts known at that time. Dalton proposed his atomic theory during the period 1803-1807; this theory played an important role in the later development of Chemistry, and stimulated the emergence of concepts which nowaday constitute the working language of the chemist. Thus, the structural hypotheses advanced by Kékulé, Vant'Hoff and Le Bel before the turn of the nineteenth century complement Dalton's atomic theory by postulating the existence of particularly strong interactions - bonded interactions - between atoms, leading to the formation of molecules. The synthesis of Dalton's atomic theory, modified by the ideas of Avogadro and Cannizaro, and the structural ideas of Kékulé, Vant'Hoff and Le Bel leads to the description of the constitution of matter in terms of the concept of molecular structure: The atoms of a molecule are viewed to be linked to one another by lines which denote the bonds, whose union represents what a chemist calls the structure of the molecule. From these considerations grew the notion of valence, as an important atomic property. Valence is some measure of the ability of an atom to form bonds with other atoms. In general, the maximum number of bonds a given atom can form with other atoms is fixed and corresponds to the saturation of its valency. Thus, it is necessary to assume that certain pairs of atoms in a molecule are bonded to one another while other pairs are not. Bonds, therefore, possess directional properties, as inferred by Vant'Hoff and Le Bel, and it is through these directional properties that observed geometrical arrangements of atoms in molecules later found their explanation.

The models of the bonds linking atoms were initially purely descriptive: electrovalent bonds for ionic or polar substances, covalent bonds for non-polar substances. The forces leading to binding in the former substances could be rationalized in terms of the then understood electrical make-up of matter, while the nature of the bonding in the latter substances was a mystery. The next significant advance in the development of a model of molecular structure was made by Lewis.[2] He proposed that the electron pair (rather than the octet of electrons) should be regarded as the fundamental structural unit. By introducing the idea that a single chemical bond was the result of sharing a pair of electrons between two atomic kernels, Lewis was led to the conclusion that the difference between the two extreme types of chemical combination, corresponding to the very polar and the non-polar substances, respectively, is a difference in degree, and not one in type. Lewis' ideas had been so successful in accounting for known chemical combinations and predicting new ones, that they still dominate our present thinking.

The first half of the twentieth century witnessed the birth and development of quantum mechanics. The application of quantum mechanics to chemistry, resulting in what is now known as quantum chemistry, however, inevitably leads to new questions concerning the validity of the Lewis electron pair model, and, more generally, the concept of molecular structure. The description of a molecular system by a wavefunction satisfying a Schrödinger equation introduces a radical change in our thinking: the wavefunction describes all particles as delocalized, in the sense that their instantaneous positions cannot be determined, all that is accessible to us being their probability distribution in space. As a consequence, the picture of the electron pairs in some geometric arrangement, tetrahedral arrangement as used by Lewis to rationalize the known structure of carbon compounds, breaks down. The essential idea of Dalton's atomic theory now appears only in the construction of the Hamiltonian operator of an isolated molecular system, that is, as an input to the non-relativistic theoretical study of the molecule: the Hamiltonian operator, which determines the dynamics of the system through Schrödinger's time-dependent equation, is known only if the constituting particles of the system, and the potential to which they are subjected, are completely specified. The picture of the atoms in a molecule as having some extension in space, and exhibiting a set of characteristic properties, is no longer evident. Instead, in the new description, at first sight, it appears that once in chemical combination, an atom loses its identity, since its electrons are indistinguishable from the electrons of its neighbours, and are delocalized over the entirety of the volume to which the molecular system is confined. The only invariant property which may still be associated with the atom is then its nucleus. In spite of this difficulty, much efforts have been invested, since the first application of quantum mechanics to chemistry, to salvage the notion of bonds. As an example, the Valence Bond theory, first developed by Heitler, London,[3] Pauling,[4] and Slater,[5] reflects an attempt to preserve Lewis' pair model. However, the explanation of bonding in terms of electron pairing provided by this theory suffers from its being a consequence of the à priori introduction of Lewis' idea at the outset of all Valence Bond calculations, in constructing the appropriate trial wavefunctions to be used in the application of the variation principle. Another noteworthy approach to chemical binding, in which one finds an interesting attempt to recover the concept of bonds, is the Localized Molecular Orbital theory, first introduced and developed by Lennard-Jones[6,7] and Pople.[8] In this theory, one constructs molecular orbitals (MO's) whose associated orbital charge densities are localized in some regions of real space, and minimize the total interorbital electronic repulsion.[9] By this procedure, one generally obtains a number of orbitals whose orbital densities exhibit directional properties which are reminiscent of those of the chemical bonds. The chemical bonds in a molecular system are thus identified with some localized orbitals and therefore are imaginary objects. An orthogonal transformation of the localized MO's gives an equivalent description, within the general Molecular Orbital theory, in which, however, there need not be a localized orbital corresponding to a bond. In other words, and more generally, the possibility of representing the electronic structure of a system in terms of localized molecular orbitals does not itself imply the existence of spatially localized pairs of electrons. In his original papers,[6] Lennard-Jones noted that the degree to which pairs of electrons are spatially localized is determined by the quantum mechanical pair distribution function and not by a particular orbital representation of the charge density. Thus, the ability to express the total charge density in terms of a sum of localized orbital densities does not imply the existence of localized pairs of electrons. In spite of this remark concerning the reality of localized electron pairs, the works of Lennard-Jones and Pople have stimulated the development of models of electronic structure in which the charge densities of localized molecular orbitals are exhibited as the physical embodiment of the Lewis electron pair.[9-11] The failure of these models in obtaining the definitions of chemical bonds as real - perhaps someday, directly observable - properties of a molecular system is due to the fact that the charge densities of the orbitals, taken individually, are not unitary-invariant properties of the system, as opposed to the total charge distribution.

This thesis is concerned with the study of the total charge distribution of molecular systems, and its role in defining the two important elements of the atomic hypothesis, namely, the concepts of atoms and bonds. Recent work [12-19] in this laboratory on the total charge distribution of molecular systems have revealed special topological features of this distribution which may be regarded as the spatial reflections of chemical bonding in molecules. These topological features are some particular paths traced out by the gradient vector of the charge distribution. They are called bond-paths, and are uniquely associated with a set of critical points exhibited by the charge distribution. The same critical points also give rise to surfaces with special properties, which partition a molecular system into fragments identified as the atoms of the molecule. Chapter I is devoted to the study of the critical points in the molecular charge distribution, and their role in determining structural elements of the molecule. For each given nuclear configuration (or geometry) of the molecular system, the complete identification of the critical points of the associated charge distribution determines the network of bond-paths linking pairs of nuclei of the molecule. This network of bond-paths, the molecular graph, and its change with the geometry of the system, determine the structure(s) of the molecule. The concepts of molecular structure and its changes are elaborated in Chapter II, where the notion of structural stability is introduced and utilized in the study of the changes in the gradient vector field of the charge distribution. We are naturally led to employ the language of Catastrophe Theory and the theory of Qualitative Dynamics throughout the first two chapters; in particular, Thom's Elementary Catastrophe Theory finds its application in Chapter II, in the study of structural changes in trinuclear systems. The partitioning of a molecular system by the interatomic surfaces defined in Chapter I yields fragments which can be regarded as atomic fragments or functional groupings of such fragments. These are examples of subsystems of a molecular system, and are all defined by the zero-flux boundary condition - the requirement that the subsystem be bounded by a surface through which there is no net (local) flux in the gradient vector of the charge distribution. The quantum mechanical properties of molecular subsystems, as defined by the zero-flux boundary condition, are developed and discussed in Chapter III; previous work [20] in this laboratory on the quantum mechanical properties of atomic fragments have led to the observations that in systems where one finds additivity of properties, the charge distribution of the individual atoms remain essentially unchanged. Moreover, it was also observed that when the distribution of charge of an atom is nearly unchanged on transfer between different systems, the local kinetic energy density of the atom exhibits the same degree of invariance as does the charge density. Together with the experimentally inferred additivity of atomic properties, the above observations suggest the existence of a subsystem virial theorem. Generalized expressions of the virial theorem, applicable both to a subsystem or to a total molecular system, are obtained in Chapter III. Therein, the quantum mechanical properties of a subsystem in a molecule are delineated, through the considerations of particular statements of the variational principle, generalized so as to properly describe a subsystem as well as the total system. In addition to the subsystem virial theorem, these considerations also lead to the variational derivation of a force law for the subsystem. The mentioned force law can be considered an equation of motion governing the evolution of the subsystem in time. The derivations of the generalized variational principle, the virial theorem and the force law, presented in Chapter III, require a single condition to be satisfied by the subsystem, which is precisely the zero-flux boundary condition. Thus, the formal properties exhibited by the subsystems considered in Chapter III lend further support to our proposition, made in Chapter I and used throughout Chapter II, that the fragments obtained in the topological study of the molecular charge distribution represent the constituting atoms of a molecule, as viewed in the atomic hypothesis.

CHAPTER I

STRUCTURAL ELEMENTS IN THE MOLECULAR CHARGE DISTRIBUTION

A. Construction of the Charge Distribution:

An isolated molecular system consisting of N nuclei, labelled $\alpha, \beta, \ldots$, and n electrons, labelled $i, j, \ldots$, is described in non-relativistic quantum mechanics by its hamiltonian, $\hat{H}$, and a wavefunction, ψ, satisfying the Schrödinger time-independent equation[21]

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

where E is a real constant. The hamiltonian, $\hat{H}$, is a self-adjoint operator on the Hilbert space $L^2(R^{3(N+n)})$ of square-integrable complex-valued functions of all the constituting particle coordinates. Its explicit expression in the coordinates representation is given in eq. (1.2).

$$\hat{H}(\underset{\sim}{r},\underset{\sim}{X},ih\nabla_{\underset{\sim}{r}},ih\nabla_{\underset{\sim}{X}}) = \sum_{\alpha} - \frac{\hbar^2}{2M_{\alpha}}\Delta_{\alpha} - \frac{\hbar^2}{2m}\sum_{i}\Delta_i + \sum_{\alpha<\beta}\frac{Z_{\alpha}Z_{\beta}e^2}{R_{\alpha\beta}} - \sum_{\alpha,i}\frac{Z_{\alpha}e^2}{r_{i\alpha}} + \sum_{i<j}\frac{e^2}{r_{ij}} \tag{1.2}$$

Above, M_{α}, Z_{α} denote the mass and the charge of nucleus α, respectively; $\Delta_{\alpha}(\Delta_i)$ the Laplacian operator with respect to the coordinates of the indicated particle (nucleus α or electron i, respectively). Also, the symbols $R_{\alpha\beta}$, $r_{i\alpha}$, r_{ij} denote the Euclidian distances between the respective particles.

A wavefunction, Ψ, obeying eq. (1.1) describes a stationary state of the system, a state which depends on the time variable only through a phase factor, giving rise to constant expectation values of all time-independent observables. In fact, the state of the system is not uniquely defined by eq. (1.1). There always exists a set of properties whose corresponding operators mutually commute, and commute with $\hat{H}$. These properties can serve together with the energy, E, to uniquely define a state of the system.[21,22] The requirement that the associated operators commute with $\hat{H}$ also defines these properties as the generators of coordinates transformations which leave the system invariant.[22] Noteworthy are the permutation operators $\hat{P}_{ij}$ whose action is to interchange the identical particles i and j. The invariance of the system upon such permutations is a quantal property, to be contrasted with the invariance with respect to spatial translations and rotations of the system as a whole, which exists in classical mechanics as well. A direct consequence of the relations $[\hat{H},\hat{P}_{ij}] = 0$, is that the wavefunction of the system must be an eigenfunction of the operator $\hat{P}_{ij}$, that is, it must not distinguish identical particles. In the present considerations, the solutions of eq. (1.1) are also required to be at least antisymmetric with respect to the permutations of the electrons. To an appropriately antisymmetrized wavefunction, Ψ, one constructs the density matrices,[23,24]

$$\Gamma^{(n)}(\underset{\sim}{X},1,2,\ldots,n|\underset{\sim}{X}',1',2',\ldots,n') = \Psi(\underset{\sim}{X},1,2,\ldots,n)\Psi^*(\underset{\sim}{X}',1',2',\ldots,n')$$

$$\vdots$$

$$\Gamma^{(p)}(\underset{\sim}{X},1,2,\ldots,p|\underset{\sim}{X}',1',2',\ldots,p') = L_n^p\Gamma^{(n)} = \qquad (1.3)$$
$$\binom{n}{p}\int\Gamma^{(n)}(X,1,2,\ldots,p,p+1,\ldots,n|X',1',2',\ldots,p',p+1,\ldots,n)d(p+1)\ldots dn$$

$$\vdots$$

$$\Gamma^{(1)}(\underset{\sim}{X},1|\underset{\sim}{X}',1') = L_n^1\Gamma^{(n)} = n\int\Gamma^{(n)}(\underset{\sim}{X},1,2,\ldots,n|\underset{\sim}{X}',1',2,\ldots,n)$$
$$d2d3\ldots dn$$

In equations (1.3), the electronic spatial and spin coordinates are indicated by the particles' indices $i = 1,2,\ldots,n$, whereas the nuclear coordinates are designated by the symbol $\underset{\sim}{X}$. The contraction operator L_n^p acts on $\Gamma^{(n)}$ by taking its partial trace with respect to $(n-p)$ electronic indices;[25] this partial trace is then multiplied by the normalization factor $\binom{n}{p}$.[23] A density matrix $\Gamma^{(p)}$ is thus a function of $(N+p)$ sets of variables, namely the primed and unprimed spin and spatial coordinates of N nuclei and p of the n electrons. Of particular interest are the second-order density matrix $\Gamma^{(2)} = L_n^2\Gamma^{(n)}$ and the one matrix $\Gamma^{(1)} = L_2^1\Gamma^{(2)}$. The two-matrix owes its importance to the fact that the energy of the system in the state Ψ is linear in $\Gamma^{(2)}$.[23,24] As to the one-matrix, it has been the object of intensive investigation[26,27,28] in the last decade, since the discovery of the Höhenberg and Köhn theorem,[26] in 1964. These researches reflect the conception that much of the physics of molecular systems can be understood without the full knowledge of the exhaustive information content of the wavefunction. In the present context, we only mention these works to point out that they share the same philosophical point of view as our work. However, the analogy does not extend further: we do not aim at reformulating the laws of quantum mechanics in terms of the one-matrix, and even do not believe that the Schrödinger equation could be equivalent to some variational equation governing $\Gamma^{(1)}$. However, we wish to extract the maximum of information from $\Gamma^{(1)}$, and particularly from its (spinless) diagonal elements

$$\rho(\underset{\sim}{r};\underset{\sim}{X}) = \Gamma^{(1)}(\underset{\sim}{X},\underset{\sim}{r}|\underset{\sim}{X},\underset{\sim}{r}) = n\int\Psi^*(\underset{\sim}{X},\underset{\sim}{r},2,..,n)\Psi(\underset{\sim}{X},\underset{\sim}{r},2,..,n)d\sigma_1 d^N\sigma d2..dn \quad (1.4$$

where the integral $n\int d\sigma_1 d^N\sigma d2..dn$ involves the summation over the spin states of one electron and those of the nuclei, in addition to the contraction L_n^1.

We shall attempt to map this information onto chemical concepts. Equation (1.4) defines what is termed the charge distribution. In fact, $\rho(\underset{\sim}{r},\underset{\sim}{X})$ affords the following probabilistic interpretation: (ρ/N) is the probability density of finding one (any one) electron at $\underset{\sim}{r}$, and the nuclei in the configuration $\underset{\sim}{X}$. It is thus the density of a joint probability. In our calculations, we follow the usual practice of quantum chemistry, and consider the wavefunction in the Born-Oppenheimer approximation,[21,22,29] where the nuclear and electronic motions are decoupled. The wavefunction is then written in the form

$$\Psi_\ell(\underset{\sim}{X},\underset{\sim}{r}) = \chi_\ell(\underset{\sim}{X})\psi_\ell(\underset{\sim}{r};\underset{\sim}{X}) \qquad (1.5)$$

The electronic part $\psi(\underset{\sim}{r};\underset{\sim}{X})$ of the wavefunction given in eq. (1.5) is customarily required to satisfy the electronic Schrödinger equation

$$\hat{H}_{e\ell}\psi_\ell(\underset{\sim}{r};\underset{\sim}{X}) = E_\ell(\underset{\sim}{X})\psi(\underset{\sim}{r};\underset{\sim}{X}) \qquad (1.6a)$$

where

$$\hat{H}_{e\ell} = \hat{H} - \{\sum_\alpha(-\frac{\hbar^2}{2M_\alpha})\Delta_\alpha + \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta e^2}{R_{\alpha\beta}}\} = \hat{H} - \hat{H}_{Nuc} \qquad (1.6b)$$

is the electronic hamiltonian. The nuclear hamiltonian, $\hat{H}_{Nuc}$, is obviously defined by

$$\hat{H}_{Nuc} = \hat{T}_{Nuc} + \hat{V}_{nn}(\underset{\sim}{X}) \qquad (1.7a)$$

where

$$\hat{T}_{Nuc} = \sum_\alpha(-\frac{\hbar^2}{2M_\alpha})\Delta_\alpha \qquad (1.7b)$$

is the kinetic-energy operator of the nuclei, and

$$\hat{V}_{nn}(\underset{\sim}{X}) = \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta e^2}{R_{\alpha\beta}}$$

is the nuclear-nuclear repulsion potential. The electronic hamiltonian, $\hat{H}_{e\ell}$, depends on the nuclear configuration $\underset{\sim}{X}$ only through the multiplicative potential term

$$\hat{V}_{en} = -\sum_{\alpha,i}\frac{Z_\alpha e^2}{r_{i\alpha}}$$

and therefore is responsible for the parametric dependence of the solutions to eq. (1.6a) on $\underset{\sim}{X}$. In fact, the most general solution to eq. (1.1) can be written as an expansion over a basis constituted of the products of the eigenfunctions ψ_ℓ and $\hat{H}_{e\ell}$ and a complete basis set $\{\eta_m(\underset{\sim}{X})\}$ of the nuclear Hilbert space, yielding

$$\Psi(\underset{\sim}{r},\underset{\sim}{X}) = \sum_{m,\ell} c_{m\ell}\eta_m(\underset{\sim}{X})\psi_\ell(\underset{\sim}{r};\underset{\sim}{X}) \qquad (1.8)$$

Using eq. (1.8) in the Schrödinger equation (1.1), one arrives at

$$[\hat{H}_{nuc} + E_\ell(\underset{\sim}{X})]\chi_\ell(\underset{\sim}{X}) + \sum_{\ell'}\{<\chi_\ell|\hat{T}_{nuc}\psi_{\ell'}>\chi_{\ell'}(\underset{\sim}{X}) - \sum_\alpha \frac{\hbar^2}{M_\alpha} <\psi_\ell|\nabla_\alpha\psi_{\ell'}>$$

$$\cdot\nabla_\alpha\chi_{\ell'}(\underset{\sim}{X})\} = E\chi_\ell(\underset{\sim}{X}) \qquad (1.9)$$

where $\chi_\ell(\underset{\sim}{X}) = \sum_m c_{m\ell}\eta_m(\underset{\sim}{X})$.

The Born-Oppenheimer approximation, eq. (1.5), is obtained if one neglects the coupling terms on the ℓhs of eq. (1.9). In the situations where such procedure is justified, eq. (1.9) indeed reduces to

$$[\hat{H}_{nuc} + E_\ell(\underset{\sim}{X})]\chi_\ell(\underset{\sim}{X}) = E\chi_\ell(\underset{\sim}{X}) \qquad (1.10)$$

Equation (1.10) describes the nuclei as moving in an average external potential field whose value for a given configuration $\underset{\sim}{X}$ is determined by the electronic Schrödinger equation (1.6a). Multiplying both sides of equation (1.10) by $\psi_\ell(\underset{\sim}{r};\underset{\sim}{X})$ and using eq. (1.6a), we obtain

$$[\hat{H}_{nuc} + \hat{H}_{e\ell}](\chi_\ell(\underset{\sim}{X})\psi_\ell(\underset{\sim}{r};\underset{\sim}{X})) = \hat{H}(\chi_\ell\psi_\ell) = E(\chi_\ell\psi_\ell) \qquad (1.11)$$

that is, the product $\Psi_\ell(\underset{\sim}{r},\underset{\sim}{X}) = \chi_\ell(\underset{\sim}{X})\psi_\ell(\underset{\sim}{r};\underset{\sim}{X})$ is an eigenfunction of $\hat{H}$, the total hamiltonian. The neglect of the coupling terms in eq. (1.9) can be justified by the large ratio between the nuclear masses and the electronic mass;[29] it is also justified when the electronic states $\psi_\ell(\underset{\sim}{r};\underset{\sim}{X})$, $\ell = 0,1..$, are well separated from one another for all values of the nuclear coordinate vector $\underset{\sim}{X}$.[30] With the Born-Oppenheimer separation, the charge distribution defined by eq. (1.4) factorizes into a product of a nuclear distribution function $P_\ell(\underset{\sim}{X})$ with an electronic distribution $\rho^\ell_{\underset{\sim}{X}}(\underset{\sim}{r})$:

$$\rho_\ell(\underset{\sim}{r},\underset{\sim}{X}) = P_\ell(\underset{\sim}{X})\rho^\ell_{\underset{\sim}{X}}(\underset{\sim}{r};\underset{\sim}{X}) \tag{1.12}$$

For a given state ℓ, the factor

$$P_\ell(\underset{\sim}{X}) = \chi^*_\ell(\underset{\sim}{X})\chi_\ell(\underset{\sim}{X}) \tag{1.13}$$

describes the probability density of finding the nuclei in the configuration $\underset{\sim}{X}$, while

$$\rho^\ell_{\underset{\sim}{X}}(\underset{\sim}{r};\underset{\sim}{X}) = n\int\psi^*_\ell(\underset{\sim}{r},2,\ldots,n;\underset{\sim}{X})\psi_\ell(\underset{\sim}{r},2,\ldots,n;\underset{\sim}{X})d\sigma_1 d^N\sigma d2..dn \tag{1.14}$$

is the proper electronic charge distribution, when the nuclei are in the given configuration $\underset{\sim}{X}$. Its product with (1/n) is the probability density of finding an electron at $\underset{\sim}{r}$, conditional upon the nuclei being in the configuration $\underset{\sim}{X}$. The general scheme that we shall present in the following involves the study of the topological properties of the distribution $\rho_\ell(\underset{\sim}{r},\underset{\sim}{X})$ for various fixed configurations $\underset{\sim}{X}$, and a given state ℓ.

When the Born-Oppenheimer approximation is used, the field studied is the charge distribution $\rho^{\ell}_{\underset{\sim}{X}}(\underset{\sim}{r})$ associated with the nuclear configuration $\underset{\sim}{X}$, which now enters the definition of the field as a set of parameters. For each configuration $\underset{\sim}{X}$, the topological study of $\rho^{\ell}_{\underset{\sim}{X}}(\underset{\sim}{r})$ enables us to associate with $\underset{\sim}{X}$ a set of structural elements. This set thus shares the same probability density $P_{\ell}(\underset{\sim}{X})$ as the configuration $\underset{\sim}{X}$. Beyond the Born-Oppenheimer separation, one can still consider the non-adiabatic charge distribution $\rho_{\ell}(\underset{\sim}{r},\underset{\sim}{X})$ and associate to each configuration $\underset{\sim}{X}$ a set of structural elements, which shares the same marginal probability density as $\underset{\sim}{X}$. However, this marginal probability density can only be obtained from $\rho_{\ell}(\underset{\sim}{r},\underset{\sim}{X})$ by integration over the electronic coordinates, and cannot be used in reconstructing the joint probability density $\rho_{\ell}(\underset{\sim}{r},\underset{\sim}{X})$. Nevertheless, the association of structural elements to each nuclear configuration is a generally valid and physically meaningful procedure. The subsequent definition of molecular structure is thus permissible both within and/or without the Born-Oppenheimer approximation.[31] In our applications, we always consider molecular systems in their ground states. The following formulation of the topological study of the molecular charge distributions will, therefore, concern a field $\rho_{\ell}(\underset{\sim}{r},\underset{\sim}{X})$, over the space $R^{3(N+1)}$ of the nuclear coordinates and the coordinates of an electron, for a given ℓ (say $\ell = 0$). Henceforth, we shall delete the subscript ℓ, and call the nuclear configuration space R^{3N} by the term "control space".

The space R^3 of the electron's coordinates will be called, accordingly, the "behaviour space".[32]

B. Elements of Molecular Structure in the Charge Distribution

The topological study of the charge distribution $\rho(\utilde{r},\utilde{X})$ of a molecular system in a given nuclear configuration $\utilde{X}$ was customarily carried out by plotting contours of constant ρ,[13-16] or by utilizing three dimensional displays of the distribution.[33] We propose to complement the contours of ρ by the orthogonal trajectories[34] traced out by the gradient vector field $\nabla\rho(\utilde{r},\utilde{X})$, the gradient being taken with respect to the behaviour variables $\utilde{r} \in R^3$. We shall show that the complementary picture so obtained is capable of revealing the essential topological properties of the charge distribution for any given configuration $\utilde{X}$, and, moreover, affords concise description of the distribution by the specification of a set of critical points with characteristic properties.
In the following, unless otherwise stated, ρ will stand for $\rho(\utilde{r},\utilde{X})$ for a given $\utilde{X}$. Also, all molecular systems considered for illustration in this section will be taken in their ground state and at the equilibrium geometry.

Because of the Kato cusp conditions,[35,36] the gradient of the charge distribution is not defined at the nuclear positions, where it presents a discontinuity. Thus, ρ can be considered as a smooth function of $\utilde{r}$, only if one excludes the nuclear positions. In other words, it is an almost everywhere smooth function.[37] However, in the neighbourhood of a nuclear cusp position, the charge distribution can be replaced by a smooth function through a homeomorphism. That is, one can define a bijective and bicontinuous mapping between ρ and some smooth integrable function ρ_h which coincides with ρ everywhere, except within open balls of infinitesimal radii centered at the nuclear positions. The homeomorphism described above, defines ρ_h as topologically equivalent[38] to the actual distribution ρ, and enables one to discuss the topological properties of ρ via the corresponding properties of the equivalent field ρ_h. Bearing this remark in mind, we now consider the properties of the vector field $\nabla\rho$. We note that the gradient vectors of ρ are (almost everywhere) orthogonal to contours of constant ρ. The trajectories generated by these vectors, the <u>gradient paths</u> of ρ, thus provide an alternative description of the charge distribution. Mathematically, the <u>gradient path</u> g_m, through a given point $\underset{\sim}{r}_m \in R^3$, is defined by the differential equation[34,18,19]

$$\frac{d\underset{\sim}{r}(s)}{ds} = \nabla\rho(\underset{\sim}{r}(s),\underset{\sim}{X}); \qquad \underset{\sim}{r}(0) = \underset{\sim}{r}_m, s \in R \tag{1.15}$$

The assignment of the continuous parameter s to the points of the gradient path g_m, as described in eq. (1.15), can be regarded as the limit of the following process:

Suppose that at $\underset{\sim}{r}_m$, $\nabla\rho(\underset{\sim}{r}_m,\underset{\sim}{X}) \neq \underset{\sim}{0}$. There exists a neighbourhood $B(\underset{\sim}{r}_m)$ of $\underset{\sim}{r}_m$ whose intersection V_m with g_m is non-empty, and such that

$$\forall \underset{\sim}{r} \varepsilon V_m,\ \exists\ \theta(\underset{\sim}{r}) \varepsilon [-\theta_m, \theta_m],\ \underset{\sim}{r} - \underset{\sim}{r}_m = \theta(\underset{\sim}{r}) \nabla\rho(\underset{\sim}{r}_m, \underset{\sim}{X}) \tag{1.16}$$

where $\theta_m > 0$ is related to the radius $\varepsilon(\underset{\sim}{r}_m)$ of the largest open ball contained in $B(\underset{\sim}{r}_m)$ by

$$\theta_m = \frac{\varepsilon(\underset{\sim}{r}_m)}{|\nabla\rho(\underset{\sim}{r}_m)|}$$

Equation (1.16) is obtained since, by definition, the gradient path g_m is tangent to the vector $\nabla\rho$ at any of its points $\underset{\sim}{r}$, and in particular, at $\underset{\sim}{r}_m$ itself. Choose a point $\underset{\sim}{r}_{m+1} \varepsilon V_m$, by eq. (1.16), there exists a parameter $\theta(\underset{\sim}{r}_{m+1})$ such that

$$\underset{\sim}{r}_{m+1} = \underset{\sim}{r}_m + \theta(\underset{\sim}{r}_{m+1}) \nabla\rho(\underset{\sim}{r}_m, \underset{\sim}{X}).$$

If $\underset{\sim}{r}_{m+1} = \underset{\sim}{r}_m$, one obviously has $\theta(\underset{\sim}{r}_m) = 0$. For $\underset{\sim}{r}_{m+1} \neq \underset{\sim}{r}_m$, since $\underset{\sim}{r}_{m+1} \varepsilon g_m$, one can assign a V_{m+1} to $\underset{\sim}{r}_{m+1}$ in the same manner as V_m was associated with $\underset{\sim}{r}_m$, and eq. (1.16) holds for $\underset{\sim}{r}_{m+1}$, upon the replacements of V_m by V_{m+1} and of θ_m by $\theta_{m+1} = \varepsilon(\underset{\sim}{r}_{m+1})/|\nabla\rho(\underset{\sim}{r}_{m+1})|$. Again, one defines $\varepsilon(\underset{\sim}{r}_{m+1})$ as the radius of the maximal open ball centered at $\underset{\sim}{r}_{m+1}$ within which g_m can be approximated by its tangent. Another point $\underset{\sim}{r}_{m+2} \neq \underset{\sim}{r}_{m+1}$ or $\underset{\sim}{r}_m$ of the gradient path g_m can be chosen in V_{m+1} and is related to $\underset{\sim}{r}_m$ by

$$\underset{\sim}{r}_{m+2} = \underset{\sim}{r}_m + \theta(\underset{\sim}{r}_{m+1}) \nabla\rho(\underset{\sim}{r}_m, X) + \theta(\underset{\sim}{r}_{m+2}) \nabla\rho(\underset{\sim}{r}_{m+1}, \underset{\sim}{X})$$

The procedure described above can be repeated for each new point, and, with $\theta(\underset{\sim}{r}_{m+k+1}) = \delta_k$, yields the recursive formula

$$\underset{\sim}{r}_{m+i} = \underset{\sim}{r}_m + \sum_{j=0}^{i} \delta_j \nabla\rho(\underset{\sim}{r}_{m+j}) \qquad (1.17)$$

Equation (1.17) describes the realization of the gradient paths by a discrete countable set of its points. Just as eq. (1.16) furnishes a good local approximation to the gradient paths in the neighbourhood of $\underset{\sim}{r}_m$, eq. (1.17) provides a good global approximation to the path in its entirety. The smaller the δ_j's are, the better will this approximation be. In the limit where all δ_j's approach zero, the summation in the r.h.s. of eq. (1.17) can be replaced by an integral in which a continuous parameter s replaces the index j, and its differential ds, the local parameters δ_j. One thus obtains:

$$\underset{\sim}{r}(s') = \underset{\sim}{r}(0) + \int_{s=0}^{s=s'} \nabla\rho(\underset{\sim}{r}(s),\underset{\sim}{X})ds; \quad \underset{\sim}{r}(0) = \underset{\sim}{r}_m \qquad (1.18)$$

which is the integral version of eq. (1.15).

Above, we have considered only points at which $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}) \neq 0$, and have seen that for each such point $\underset{\sim}{r}_m$, there exists a maximal neighbourhood within which the gradient path through $\underset{\sim}{r}_m$ is a straight line, given by the tangent to the path at $\underset{\sim}{r}_m$. Points $\underset{\sim}{r}_c$ of the real space at which $\nabla\rho(\underset{\sim}{r}_c,\underset{\sim}{X}) = 0$ do not possess this property. Any neighbourhood of $\underset{\sim}{r}_c$ intersects the gradient path through $\underset{\sim}{r}_c$ at a point, the point $\underset{\sim}{r}_c$ itself. In other words, the gradient path through $\underset{\sim}{r}_c$ is the singleton $g_c = \{\underset{\sim}{r}_c\}$. This result can also be obtained by using eq. (1.15). A point $\underset{\sim}{r}_c$ at which $\nabla\rho(\underset{\sim}{r}_c,\underset{\sim}{X}) = 0$ is called a <u>critical point</u> of $\rho(\underset{\sim}{r},\underset{\sim}{X})$.[39] In the neighbourhood of a critical point $\underset{\sim}{r}_c$, there are in general an infinite number of points at which $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}) \neq 0$. Each of these points belongs to a gradient path which may come as close as possible to $\underset{\sim}{r}_c$, without ever attaining it. We wish to classify these paths, and through the so obtained description of all the paths in its neighbourhood, characterize the critical point. For this purpose, we need make use of the concepts of α and ω-limit sets.[40,41] Consider the gradient path through a given point $\underset{\sim}{r}_m$, with the parametrization described in eqs. (1.15) and (1.18). The α-limit set of the path is defined as the collection of points $\underset{\sim}{r}(s)$ obtained in the limit $s \to -\infty$:

$$\alpha(\underset{\sim}{r}_m) = \lim_{s\to-\infty} \underset{\sim}{r}(s) \tag{1.19a}$$

Likewise, the ω-limit set is obtained when $s \to +\infty$

$$\omega(\underset{\sim}{r}_m) = \lim_{s\to+\infty} \underset{\sim}{r}(s) \tag{1.19b}$$

Pictorially, the α-limit set is the origin of the path, whereas the ω-limit set is the terminus of the path. The notions of α and ω-limit sets were introduced to deal with general vector fields.[40,41] In our specific applications to the gradient vector fields of the charge distribution, these limit sets are generally singletons, constituted of critical points of $\rho(\underset{\sim}{r},\underset{\sim}{X})$. In some neighbourhood of a critical point $\underset{\sim}{r}_c$, ρ can be written in a Taylor series expansion up to second order

$$\rho(\underset{\sim}{r},\underset{\sim}{X}) = \rho(\underset{\sim}{r}_c;\underset{\sim}{X}) + (\underset{\sim}{r} - \underset{\sim}{r}_c)^T\cdot(\nabla\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}))\Big|_{\underset{\sim}{r}=\underset{\sim}{r}_c} \cdot(\underset{\sim}{r} - \underset{\sim}{r}_c) \tag{1.20}$$

where $\nabla\nabla\rho(\underset{\sim}{r},\underset{\sim}{x})|_{\underset{\sim}{r}=\underset{\sim}{r}_c} = \{e\}^T\underset{\approx}{A}\{e\}$ is the dyadic of second partial derivatives of ρ with respect to the three components of $\underset{\sim}{r}$, evaluated at the critical point $\underset{\sim}{r}_c$. The corresponding matrix $\underset{\approx}{A} = (\partial^2\rho/\partial r_i\partial r_j)_{\underset{\sim}{r}=\underset{\sim}{r}_c}$ is called the <u>Hessian matrix</u> of the critical point $\underset{\sim}{r}_c$.[34] Above we have also used the notation $\{e\}$ to denote the matrix $(\underset{\sim}{e}_1,\underset{\sim}{e}_2,\underset{\sim}{e}_3)$ and indicated the transpose of any (3 x 3) matrix by the superscript T. Using eq. (1.20), the differential equation (1.15) becomes

$$\frac{d\underset{\sim}{r}(s)}{ds} = \underset{\approx}{A}(\underset{\sim}{r}(s) - \underset{\sim}{r}_c) \tag{1.21}$$

in some neighbourhood of $\underset{\sim}{r}_c$.

The Hessian matrix $\underset{\approx}{A}$ is a hermitian real matrix and therefore possesses only real eigenvalues $\lambda_1,\lambda_2,\lambda_3$. When $\lambda_i \neq 0$ for all $i \in \{1,2,3\}$, eq. (1.21) admits solutions of the form[34,42,43]

$$\underset{\sim}{r}(s) - \underset{\sim}{r}_c = \alpha\underset{\sim}{v}_1 e^{\lambda_1 s} + \beta\underset{\sim}{v}_2 e^{\lambda_2 s} + \gamma\underset{\sim}{v}_3 e^{\lambda_3 s} \tag{1.22}$$

where $\underset{\sim}{v}_1,\underset{\sim}{v}_2,\underset{\sim}{v}_3$ are the eigenvectors of $\underset{\approx}{A}$ associated with $\lambda_1,\lambda_2,\lambda_3$, respectively. In eq. (1.22), the constants α,β,γ depends on the initial conditions imposed on eq. (1.21), that is, on the particular gradient path this equation is to describe. A critical point in ρ can be characterized by two properties of its Hessian matrix:[34,17-19] τ, <u>the rank of the critical point</u> is the number of non-zero eigenvalues of $\underset{\approx}{A}$, and ζ, <u>the signature of $\underset{\sim}{r}_c$</u>, is the excess number of positive over negative eigenvalues of $\underset{\approx}{A}$. In this section, we consider cases where only rank three critical points in ρ arise, i.e., $\tau = 3$. Only four types of critical points (3,s) are thus to be analyzed: (3,-3), (3,-1), (3,+1) and (3,+3) critical points.

At a (3,-3) critical point, $\rho(\underset{\sim}{r},\underset{\sim}{X})$ is a local maximum: all gradient paths in some neighbourhood of the critical point terminate at that point. In other words, a (3,-3) critical point constitutes the ω-limit set of all paths in its neighbourhood. No path originates from the critical point. We have observed that only nuclear positions exhibit such properties,[17,18] within a homeomorphism (on ρ) as discussed before. A local maximum in ρ satisfies the mathematical conditions of an attractor of the vector field $\nabla\rho$.[40,41] These conditions are stated as follows: Let $\underset{\sim}{v}(m)$ be a vector field over a manifold M ($m \in M$). A closed subset F of M is an <u>attractor</u> of the vector field $\underset{\sim}{v}$ if and only if

i) F is invariant under $\underset{\sim}{v}$, that is the trajectory of $\underset{\sim}{v}$ through any point of F is contained entirely in F

ii) Any trajectory of $\underset{\sim}{v}$ whose α-limit set contains a point of F is contained in F

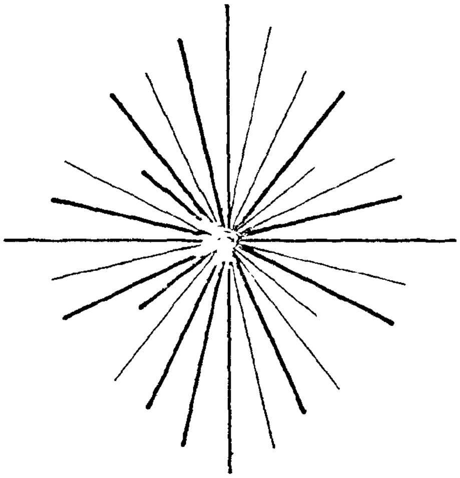
iii) There exists an open invariant neighbourhood U of F, such that every trajectory starting from a point of U has F as its ω-limit set. The maximal open set $U_{max} \subseteq M$ satisfying this condition is called the <u>basin</u> of the attractor F

iv) F is indecomposable, that is, almost every trajectory of $\underset{\sim}{v}$ in F, is dense in F.

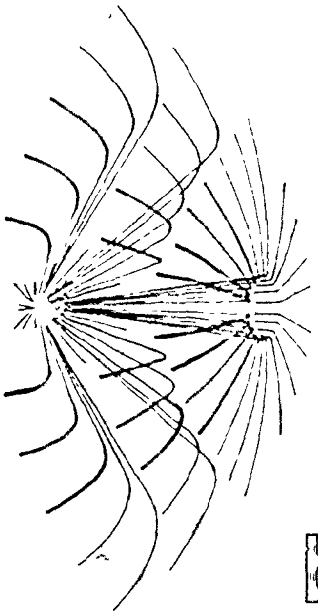
In our particular applications, where $M = R^3$, $\underset{\sim}{v} = \nabla\rho$, the singleton $\{\underset{\sim}{r}_c\}$ of any critical point satisfies conditions (i) and (iv), since the gradient paths through $\underset{\sim}{r}_c$ reduce to the point $\underset{\sim}{r}_c$ itself. However, conditions (ii) and (iii) are met only by the local maxima in ρ. As will be illustrated later, these conditions are also satisfied by the saddle points in ρ (either (3,-1) or (3,+1) critical points) only with respect to the gradient vector field over a subset of R^3 of dimension less than three. We shall occasionally refer to these as one- or two-dimensional pseudo-attractors. Unless otherwise stated, attractor or, more precisely, nuclear attractor will designate a subset (a singleton) satisfying all four conditions stated above, with respect to the field $\nabla\rho$ over R^3.

When only one nuclear attractor is present, the system has spherical symmetry: all gradient paths originate at infinity and terminate at the attractor. One, then, has a representation of an isolated atom, in its ground state, in terms of its charge distribution: the atom is the union of a nuclear attractor and its basin, which, in this case, extends over all the behaviour space. Pictorially, an isolated attractor is characterized by its phase portrait,[17-19,41-44] the pattern of gradient paths in the neighbourhood of the critical point, which is shown in Fig. 1.1a. When two nuclear attractors are present, the behaviour space is partitioned into two regions, say A and B. The interior of each region, e.g., of region A,

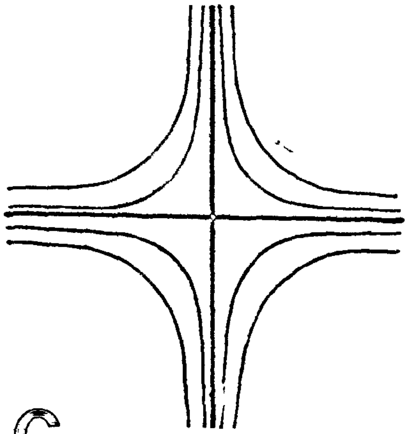
Figure 1.1. Patterns of the trajectories of $\nabla\rho$ in the neighbourhood of rank-three critical points. (a) Phase portrait of an *isolated* attractor. All trajectories *terminate* at the attractor, which is found coincident with the nucleus of an isolated atom. (b) Phase portrait of an attractor with a single neighbour. The trajectories which *terminate* at the attractor fill an *open subset* of real space, the *basin of the attractor*. The union of the attractor and its basin defines an atom with a single neighbour. (c) Phase portrait of a (3,-1) critical point, in a plane containing the *unique* pair of gradient-paths which *originate* at the critical point. In the plane normal to these two paths and containing the critical point, *locally*, all paths terminate at the critical point. All other paths avoid the critical points and terminate at nuclear attractors. The phase portrait of a (3,-1) critical point is thus of the *hyperbolic type*. (d) Phase portrait of a (3,-1) critical point, showing the trajectories which *terminate* at the critical point. *Globally*, these gradient paths *span an interatomic surface*, partitioning the molecular system into *distinct atoms*. Also shown is the unique pair of gradient paths originating at the critical point (compare with (c)). (e) Phase portrait of a (3,+1) critical point, showing the *unique* pair of gradient paths *terminating* at the critical point, together with the paths which *originate* at this point. These paths span *the ring surface*, and the (3,+1) critical point denotes the presence of a *bonded ring*.



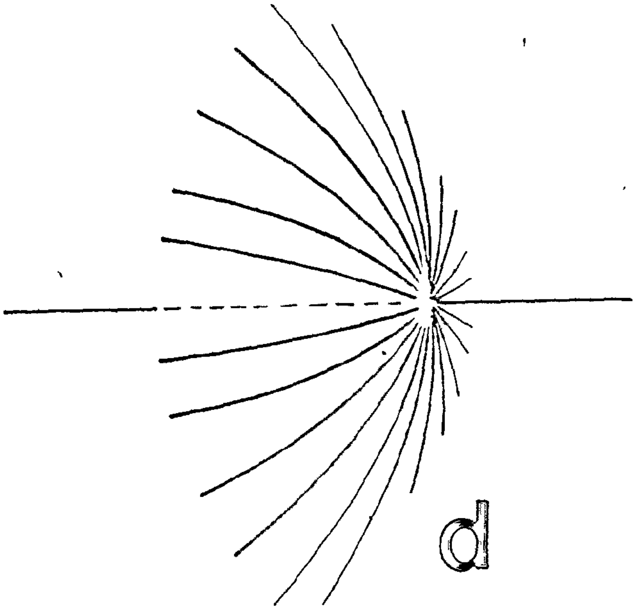
a



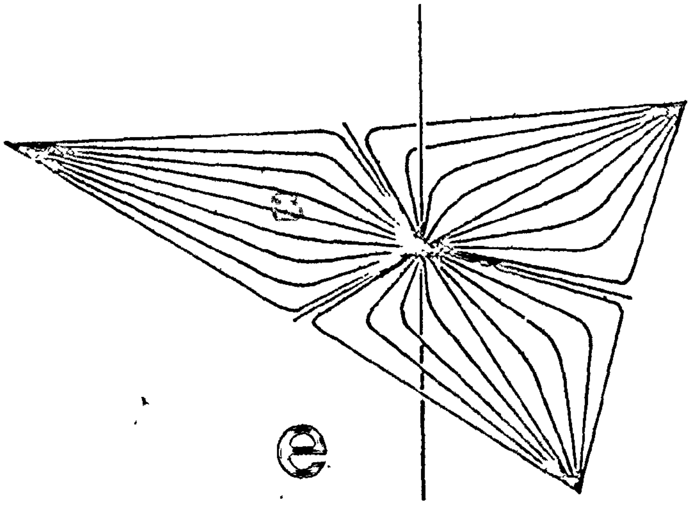
b



c



d



e

is the basin U_A of a corresponding attractor. The phase portrait associated with one of the attractors is illustrated by Fig. 1.1b. The partitioning surface S_{AB} is the common boundary of the two neighbouring basins U_A, U_B. The gradient paths through the points of S_{AB} are all contained in the surface. Originating at infinity, they all terminate at a critical point $\underset{\sim}{r}_S$ contained in S_{AB}. The critical point constitutes the ω-limit set of all the gradient paths of the surface, and therefore, is an attractor with respect to the vector field $\nabla\rho$ over S_{AB}. Consider a three-dimensional neighbourhood of $\underset{\sim}{r}_S \in S_{AB}$. By definition of S_{AB}, such a neighbourhood has non-empty intersections with the neighbouring basins U_A and U_B. Any point of this intersection is traversed by a gradient path which terminates at one of the two nuclear attractors, avoiding the critical point $\underset{\sim}{r}_S$ or otherwise originating from $\underset{\sim}{r}_S$. In fact, there are only two gradient paths whose α-limit sets are identically the singleton $\{\underset{\sim}{r}_S\}$, while all other paths in the neighbourhood of $\underset{\sim}{r}_S$ avoid the critical point. The phase portrait of the critical point $\underset{\sim}{r}_S$ is thus of the hyperbolic type illustrated in Figs. 1.1c and 1.1d. The hyperbolic phase portrait is characteristic of a (3,-1) critical point. The eigenvectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2$ associated with the two negative eigenvalues λ_1, λ_2 of the Hessian matrix $\underset{=}{A}$ of $\underset{\sim}{r}_S$ span the surface S_{AB}. The remaining eigenvector $\underset{\sim}{v}_3$, corresponding to the unique positive eigenvalue λ_3, generates the two gradient paths whose α-limit set is the singleton $\{\underset{\sim}{r}_S\}$. The union of the closures of these paths defines a line along which the charge distribution is maximal with respect to all displacements perpendicular to the line. This line is called a bond-path.[17,18,19] We now define[18,19] an atom generally as the union of an attractor of $\nabla\rho$ and its basin. Then, two atoms are neighbouring when they share a common two-dimensional boundary, an interatomic surface, which contains a (3,-1) critical point. Alternatively, one may say: two neighbouring atoms are linked by a bond path. The bond-paths are the topological antecedents of the chemical bonds: we have observed[17,18,19] that in all the systems for which the bonding schemes are thought to be well understood, the bond-paths obtained in the study of the charge distribution coincide with the lines that a chemist would draw according to his observations to represent chemical bonds. While this correspondance is still an observation, we conjecture that: two atoms are bonded if and only if they are linked by a bond-path. We shall therefore speak of a bonded pair of atoms to designate a pair of atoms between which the charge distribution exhibits a (3,-1) critical point giving rise to a bond-path, which links the respective nuclear attractors. The description of the relationship between the topologically defined atoms afforded by the bond paths is complementary to the one provided by the interatomic surfaces:

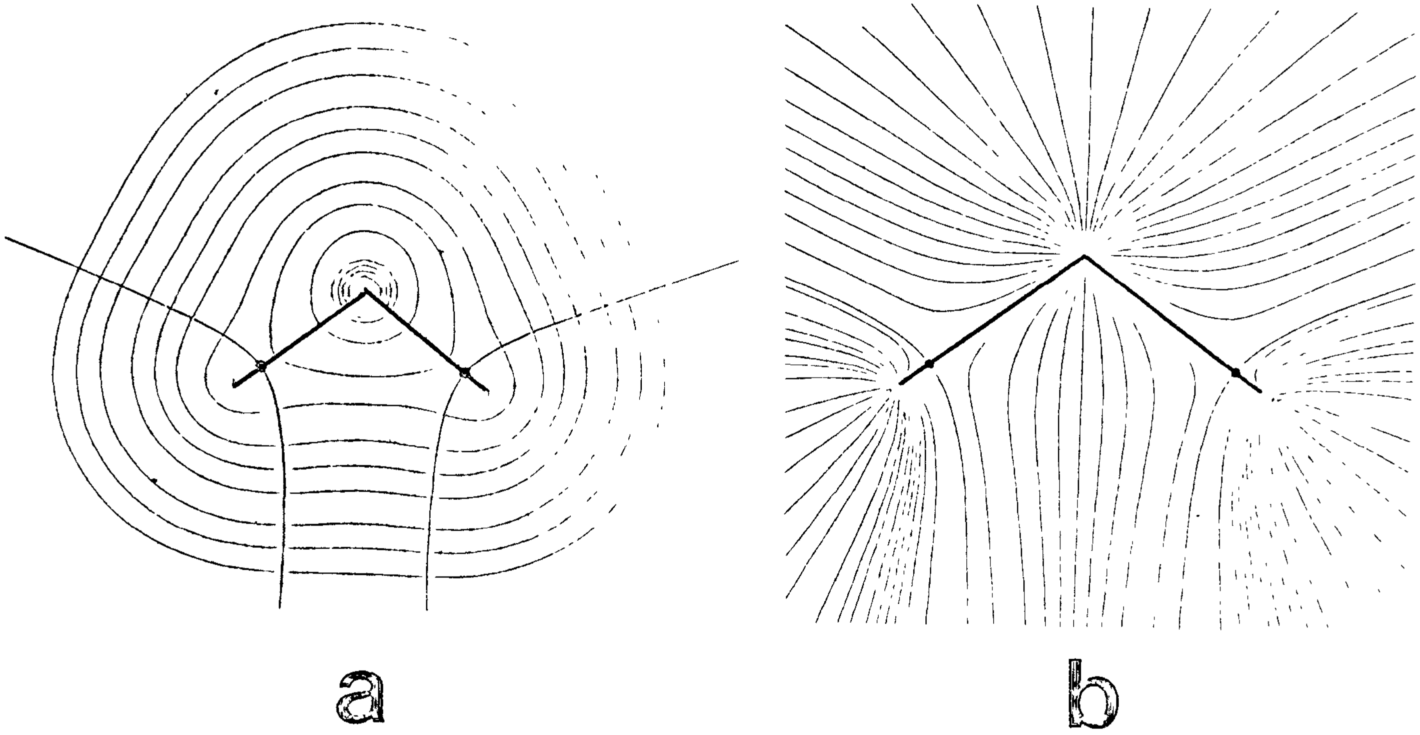
The interatomic surfaces delineate, hence define the atoms. An atom A is the portion of the electronic and nuclear charge distribution contained in the region bounded by all interatomic surfaces S_{AB}. The boundary S_A of A thus satisfies

$$\underset{B}{\cup} S_{AB} \subseteq S_A .$$

On the other hand, the bond-paths determine the bonding of the atoms. We thus see that atoms and bonds can only be defined and considered simultaneously, and interdependently. Once the attractors of a system, and the (3-1) saddle points are identified, one obtains a network of bond-paths. We shall later show that this network of bond-paths summarizes all the essential topological properties of the charge distribution when the nuclei are in a given configuration $\underset{\sim}{X}$. We define the <u>molecular graph</u> of the system in the configuration $\underset{\sim}{X}$, as this network of bond-paths. Mathematically, a <u>molecular graph</u> is, thus, the union of the closures of all the bond paths exhibited by $\rho(\underset{\sim}{r},\underset{\sim}{X})$. The definitions advanced above are illustrated in Figs. 1.2, through the example of the water molecule.[45] Fig. 1.2a shows the contours of ρ in the plane of the nuclei, to be compared with Fig. 1.2b which shows the trajectories of $\nabla\rho$ in the same plane. The hydrogen atoms are not bonded to each other since there is no bond-path linking the respective attractors. Each of them is bonded to the oxygen atom. Thus, there are two bond-paths, meeting each other at the attractor of the oxygen atom, thereby yielding the molecular graph, also shown in Fig. 1.2a.

Some bond-paths in a polyatomic molecular system may form a ring. Consider the atoms linked by these bond-paths.

Figure 1.2. The charge distribution of ground-state water, at the equilibrium nuclear configuration, and in the plane of the nuclei. (a) Contours of $\rho(\underset{\sim}{r})$. Superimposed on the contours are also shown the *molecular graph* associated with this configuration, and the interatomic surfaces which intersect the molecular graphs at the pair of (3,-1) critical point. (b) Trajectories of $\nabla\rho(\underset{\sim}{r})$. The behaviour of the trajectories in the neighbourhood of the (3,-1) critical points are to be compared with Fig. 1.1c.

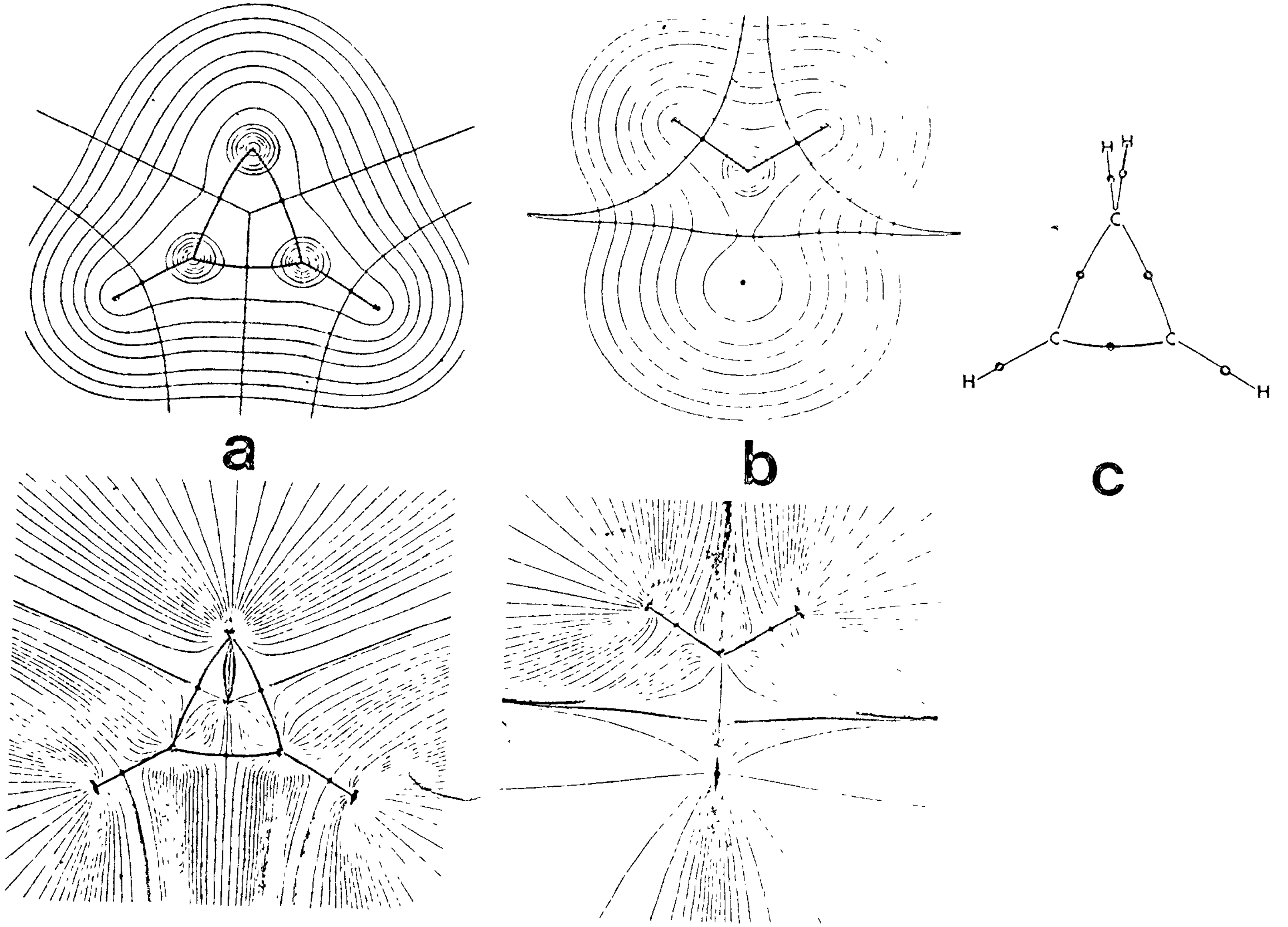

a
b

Their interatomic surfaces may meet each other in a line d. Any point along d is traversed by a gradient path whose ω-limit set is a singleton, constituted of a critical point, $\underset{\sim}{r}_{S'}$. Let M be the union of all the above interatomic surfaces:

$$M = \bigcup_{\substack{A,B\\ \varepsilon\ \text{ring}}} S_{AB}$$

As discussed above, the (3,-1) saddle points of the ring are attractors of the vector field $\nabla\rho$ over M. Their basins share a common boundary, the line d. Therefore, any gradient path which intersects d non-trivially is contained in d. In other words, d is the union of two gradient paths which terminate at $\underset{\sim}{r}_{S'}$. Any three dimensional neighbourhood of $\underset{\sim}{r}_{S'}$ has non-empty intersections with all the basins of the pseudo-attractors in M as well as with all the atoms of the ring. Thus, any point of such a neighbourhood, not on the line d, is traversed by a gradient path which terminates either at the (3,-1) or at the (3,-3) critical points on the ring. This gradient path either avoids the critical point $\underset{\sim}{r}_{S'}$, or otherwise it originates from $\underset{\sim}{r}_{S'}$. The collection of all the paths which originate from $\underset{\sim}{r}_{S'}$ is a two-dimensional submanifold of R^3, called the ring surface.[18] The phase portrait associated with the ring critical point $\underset{\sim}{r}_{S'}$, shown in Fig. 1.le, is characteristic of a (3,+1) critical point. The ring axis d is generated by the eigenvector $\underset{\sim}{v}_1$, associated with the unique negative eigenvalue λ_1 of the Hessian matrix of $\underset{\sim}{r}_{S'}$. The remaining eigenvectors $\underset{\sim}{v}_2, \underset{\sim}{v}_3$ generate the ring surface. In the ring surface, the charge distribution is a local minimum at $\underset{\sim}{r}_{S'}$, whereas at the same point, it attains a local maximum along the ring axis, d. In the case where a (3,+1) critical point is found within it, the ring is called a <u>bonded ring</u>. A bonded ring is thus the boundary of a ring surface. These definitions are illustrated in Fig. 1.3 by the example[17,45] of cyclopropene (C_3H_4). Fig. 1.3a shows the gradient map in the symmetry plane containing the carbon nuclei. Each carbon nucleus is joined to the others by a bond path. The molecular graph thus exhibits a ring. Two carbon atoms are each bonded to a hydrogen atom whose attractor also lies in this plane. In a classical structural assignment, these two carbon atoms are joined by a double bond. In the present description, only one bond-path links their respective attractors. However, the value of ρ at the corresponding (3,-1) saddle point is higher than that observed for the two other C-C bonds. All bond paths of the ring are slightly curved outward, probably denoting the presence of ring strains.[16,17,19] The ring formed by the three carbon atoms in this example is a bonded ring. Owing to the symmetry of the system, the ring surface lies in the plane of the carbon nuclei. Perpendicular to the C-C bond-paths, at the corresponding (3,-1) saddle points, are also shown the intersections of the interatomic surfaces S_{CC} with this plane. These meet at the (3,+1) ring saddle point. Another view of the ring saddle point is given in Fig. 1.3b: in the symmetry plane

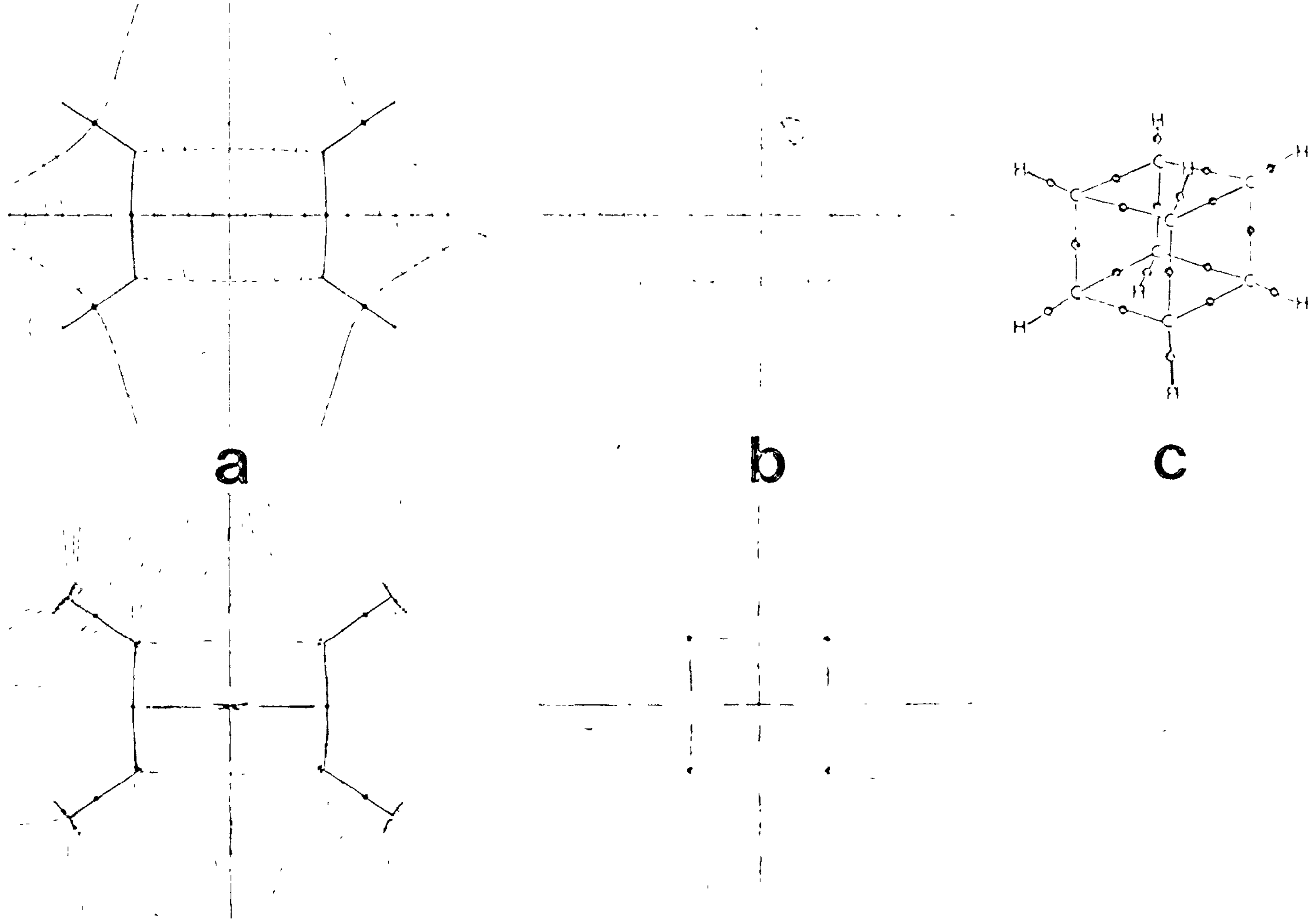
Figure 1.3. The charge distribution of cyclopropene (C_3H_4), in the ground-state and at the equilibrium geometry. (a) Contours of ρ and trajectories of $\nabla\rho$ in the symmetry plane containing all three carbon nuclei. The bonded ring exhibited by the molecular graph is clearly shown on both maps. (b) Contours of ρ and trajectories of $\nabla\rho$ in the symmetry plane perpendicular to the plane of the carbon nuclei. The intersection of the molecular graph with this plane is shown on both maps by two heavy lines; each of these is the bond-path associated with a (3,-1) critical point which is indicated by a dot, and denotes a CH bond. The (3,+1) ring critical point behaves as a (2,0) saddle point in this plane, and is characterized by a phase portrait similar to that of Fig. 1.1c. However, unlike a (3,-1) critical point, this ring critical point is the α-limit set of an infinite number of paths. (c) Molecular graph of cyclopropene in its ground-state equilibrium geometry.


a
b
H
H
C
C
C
H
H
c

perpendicular to the plane of the carbon nuclei, the (3,+1) critical point seems to play the same role as a (3,-1) critical point. In fact, it is, in this plane, a (2,0) critical point with respect to the topology of the plane inherited from the topology of R^3. The ring axis can thus be interpreted as the common boundary of the two dimensional basins in this plane, just as an interatomic surface separates a pair of three dimensional basins in R^3. A pair of gradient paths originate from the ring critical point and terminate at the two dimensional attractors, forming a line suggestive of a bond-path which would link the pair of pseudo-attractors. In reality, this line merely indicates the intersection of the ring surface with the symmetry plane considered. By symmetry, the line must itself be the union of two gradient paths belonging to the ring surface. It, thus, belongs to an infinite countable set of lines, all of them being capable to accomplish the same role: to link the same pairs of two-dimensional attractors. However, none of them is a ridge of maximum charge distribution as a true bond-path would be. The molecular graph for cyclopropene is shown in Fig. 1.3c. It may happen that the molecular graph of a system exhibits four or more (adjacent) rings whose ring surfaces together bound a three-dimensional region of the behaviour space. The corresponding ring axes then meet at a point within this region. This point is a (3,+3) critical point in ρ. At such a point, ρ is a local minimum. A (3,+3) critical point constitutes the α-limit set of all paths in some of its neighbourhoods. No paths terminate at a (3,+3) critical point. The paths which originate at a (3,+3) critical point span the bounded region of R^3 mentioned above, which we call the cage volume. The smallest part of the molecular graph which defines a cage volume is called a cage.[17,19] An example of a cage is provided by the molecular graph of cubane,[17,19,45] C_8H_8. In its ground state equilibrium geometry, the system has the symmetry of a cube: the eight carbon nuclei are located at the corners of a cube, at the centre of which is found a (3,+3) critical point. Fig. 1.4 illustrates the phase portrait of the central (3,+3) critical point by showing its cross-sections, through a diagonal symmetry plane, σ_d (Fig. 1.4a), and through a symmetry plane perpendicular to a four-fold symmetry axis (Fig. 1.4b). All paths within the cage originate from the (3,+3) critical point. Each such path terminates either at a nuclear attractor, or at a (3,-1) bond critical point, as seen in Fig. 1.4a. Otherwise, it terminates at a (3,+1) ring critical point. In Fig. 1.4b, four such paths can be observed, while all others terminate at (3,-1) critical points. Fig. 1.4c gives the molecular graph of cubane.

The above analysis emphasizes the importance of two types of critical points: the nuclear attractors, or (3,-3) critical points, which ensure the existence of atoms and bonds. The (3,-1) critical points, on the other hand, precisely define both the extensions and the bondings of atoms, via their

Figure 1.4. Phase portrait of a (3,+3) critical point, illustrated by the charge distribution of cubane (C_8H_8), at its equilibrium geometry and in the ground state. The system has the symmetry of a cube. The three-dimensional phase portrait of the (3,+3) critical point, the local minimum, found at the inversion centre of the cube, can be inferred from its cross-sections through (a) a diagonal symmetry plane, σ_d, of the cube, and (b) a symmetry plane perpendicular to a four-fold symmetry axis. In (a), the intersection of the molecular graph with the diagonal plane illustrated is shown by heavy lines carrying dots which represent (3,-1) bond critical points. The molecular graph intersects the symmetry plane, illustrated in (b), at the positions of the two-dimensional attractors found in this plane. These 2D attractors, in fact, are the (3,-1) critical points whose associated bond-paths constitute the curved edges of the cage. In (c), the resulting molecular graph of cubane is shown.


a
b
c
H
C
H

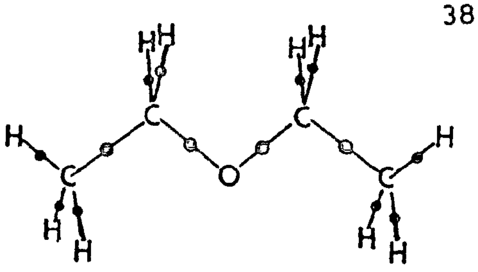
interatomic surfaces and the bond-paths linking their attractors, respectively. The two other types of critical points arise as the consequences of some particular relative configurations of the (3,-3) and (3,-1) critical points: A (3,+1) ring critical point is observed when some bond-paths form a bonded ring, which bounds a ring surface. A (3,+3) critical point denotes the existence of a cage formed by the union of a number of bond-paths. Thus, for a given configuration $\underset{\sim}{X}$, the essential topological properties of the charge distribution $\rho(\underset{\sim}{r},\underset{\sim}{X})$ are summarized in the specification of the nuclear attractors, and the bond-paths linking them, in other words, in the molecular graph. Fig. 1.5 gives the molecular graphs of a few systems of chemical interest.

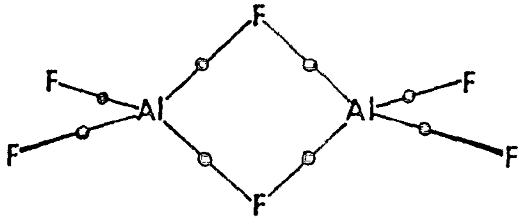
The number and types of rank 3 critical points which can coexist in a system with a finite number of nuclei is governed by the Poincaré-Hopf theorem.[17-19,34] Denoting by n, b, r, c, the number of nuclei ((3,-3) critical points), bonds ((3,-1) critical points), bonded rings ((3,+1) critical points) and cages ((3,+3) critical points), respectively, this theorem requires that these numbers satisfy the simple relationship

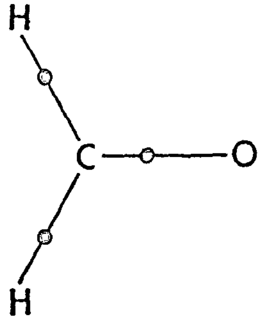
$$n - b + r - c = 1 \tag{1.23}$$

Finally, in Table 1.1, we summarize the classification of rank 3 critical points, according to their mathematical and physical properties, as detailed in this section.

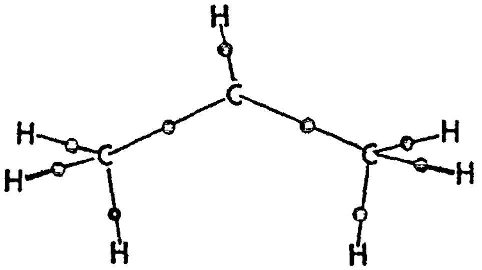
Figure 1.5. Molecular graphs as defined by the unique axes of the (3,-1) critical points found in the ground-state charge distribution of $(C_2H_5)_2O$ (diethyl ether), Al_2F_6, HCOH (formaldehyde), C_2H_2 (acetylene), $C_3H_7^+$ (for which graphs representative of two different structures are shown (cf. Chapter II, section 2)), B_2H_6 and B_5H_9. Except for $C_3H_7^+$, all molecules are considered to be at the equilibrium nuclear configuration. The graphs associated with HCOH and C_2H_2 serve to illustrate the observation that multiple bonds do not arise in the present picture. The strength of a bond may instead be inferred from the value of ρ at the bond critical point. The examples of Al_2F_6, B_2H_6 and B_5H_9 illustrate how the analysis of the charge distribution unambiguougly determines the bondings in those molecules that classical valence theory fails to describe with a unique bond structure.

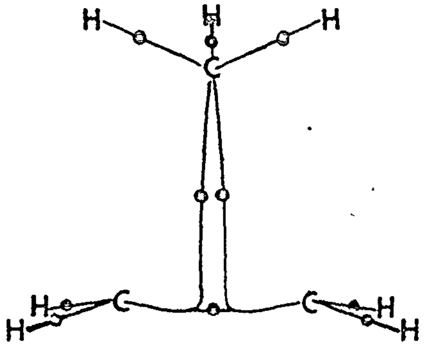

H H
H H
H
C
C
H
O
C
C
H
H
H
H H

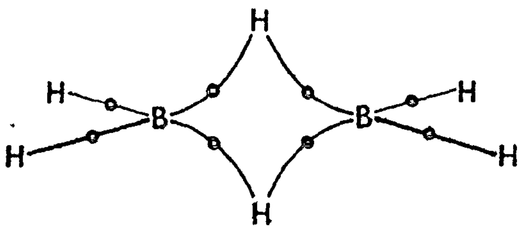

F
F
Al
Al
F
F
F
F


H
C
O
H


H
C
C
H


H
C
H
C
C
H
H
H
H
H
H


H
H
H
C
H
C
C
H
H
H


H
H
B
B
H
H
H
H
H

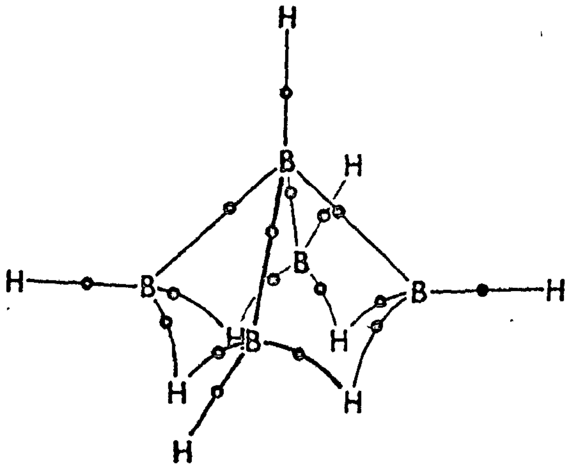

H
B
H
B
H
B
B
H
H
HB
H
H
H

TABLE 1.1

Classification of Rank 3 Critical Points

ζ = signature = (number of positive eigenvalues - number of negative eigenvalues)
n_α = number of paths originating from critical point
n_ω = number of paths terminating at critical point

ζ	number	n_α	n_ω	Type of Attractor	Physical Interpretations
-3	n	0	∞	3D	nuclear attractor, centre of atom (atom = union of attractor and basin)
-1	b	2	∞	2D	denotes possibly bonding interaction between two neighbouring atoms (A,B). In S_{AB}, the (3,-1) critical point is a pseudo-attractor (2D).
+1	r	∞	2	1D	denotes the presence of a bonded ring. Along the ring axis, the (3,+1) critical point is a 1D-pseudo-attractor.
+3	c	∞	0	non-attractor (repellor)	denotes the presence of a cage.

Poincaré-Hopf Theorem: $n - b + r - c = +1$

CHAPTER II

MOLECULAR STRUCTURE AND ITS CHANGES

A. Molecular structure and structural stability

In the preceding chapter, we have discussed and illustrated the structural elements of a molecular system at isolated points in the associated nuclear configuration space. In particular, we have emphasized the topological definition of atoms and bonds in terms of the molecular charge distribution. Indeed, in the chemists' conception of the constitution of matter, atoms and bonds are the central concepts which embody the notion of molecular structure. As chemists, we conceive molecular structure as the result of particularly strong interactions - bonded interactions - between various pairs of atoms in a molecule. In Chapter I, we have proposed that the bond-paths exhibited by the molecular charge distribution indicates the directions along which such interactions are to be found. A molecular graph associated with a given nuclear configuration, e.g., the equilibrium configuration $\underset{\sim}{X}_e$, thus summarizes the bonded interactions which would be operative if the molecule is instantaneously in the configuration $\underset{\sim}{X}_e$. This situation can never be observed, however, since, even in a classical description, the atoms of a molecule are in constant motion under the influence of thermal excitations. Moreover, in quantum theory, we expect the molecule to have an appreciable probability to be found in nuclear configurations other than $\underset{\sim}{X}_e$, within some neighbourhood of the equilibrium configuration. Thus, if molecular structure has ever been evidenced experimentally, it can only be either an average property of the system, or a property which is associated with some open subset of the nuclear configuration space. This remark constitutes the guideline for the definition of molecular structure to be found in the present chapter. Through an analysis based on the concepts of structural stability,[40,41,46] we shall define an equivalence relation of molecular graphs. This allows us to define a molecular structure as an equivalence class of molecular graphs. By the very nature of the above equivalence relation, this equivalence class is associated with an open subset of the nuclear configuration space, called a structural region. Thus, the use of the concept of structural stability leads to a partitioning of the nuclear configuration space into structural regions, to each of which is assigned a definite structure. This partitioning also depicts the possible mechanisms of structural changes in a molecule. Molecular structure and its change are thus delineated simultaneously through the analysis of structural stability.[18,19]

In Chapter I, the structural elements of a molecular system, which are summarized in the molecular graph, have been defined in terms of the gradient vector field of the charge distribution. To determine their changes, we thus need to study and characterize the changes in this vector field, $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$, which accompany the variations in its nuclear control parameters $\underset{\sim}{X}$. The description of these changes necessitates the introduction and use of a mathematical device which allows the comparison of vector fields over R^3. This mathematical device is an equivalence relation, defined as follows:[40,41] two vector fields $\underset{\sim}{v},\underset{\sim}{v}'$ over R^3 are said to be <u>equivalent</u> if and only if there exists a homeomorphism which maps the trajectories of $\underset{\sim}{v}$ into the trajectories of $\underset{\sim}{v}'$. We denote the equivalence between $\underset{\sim}{v}$ and $\underset{\sim}{v}'$ symbolically by

$$\underset{\sim}{v} \sim \underset{\sim}{v}'$$

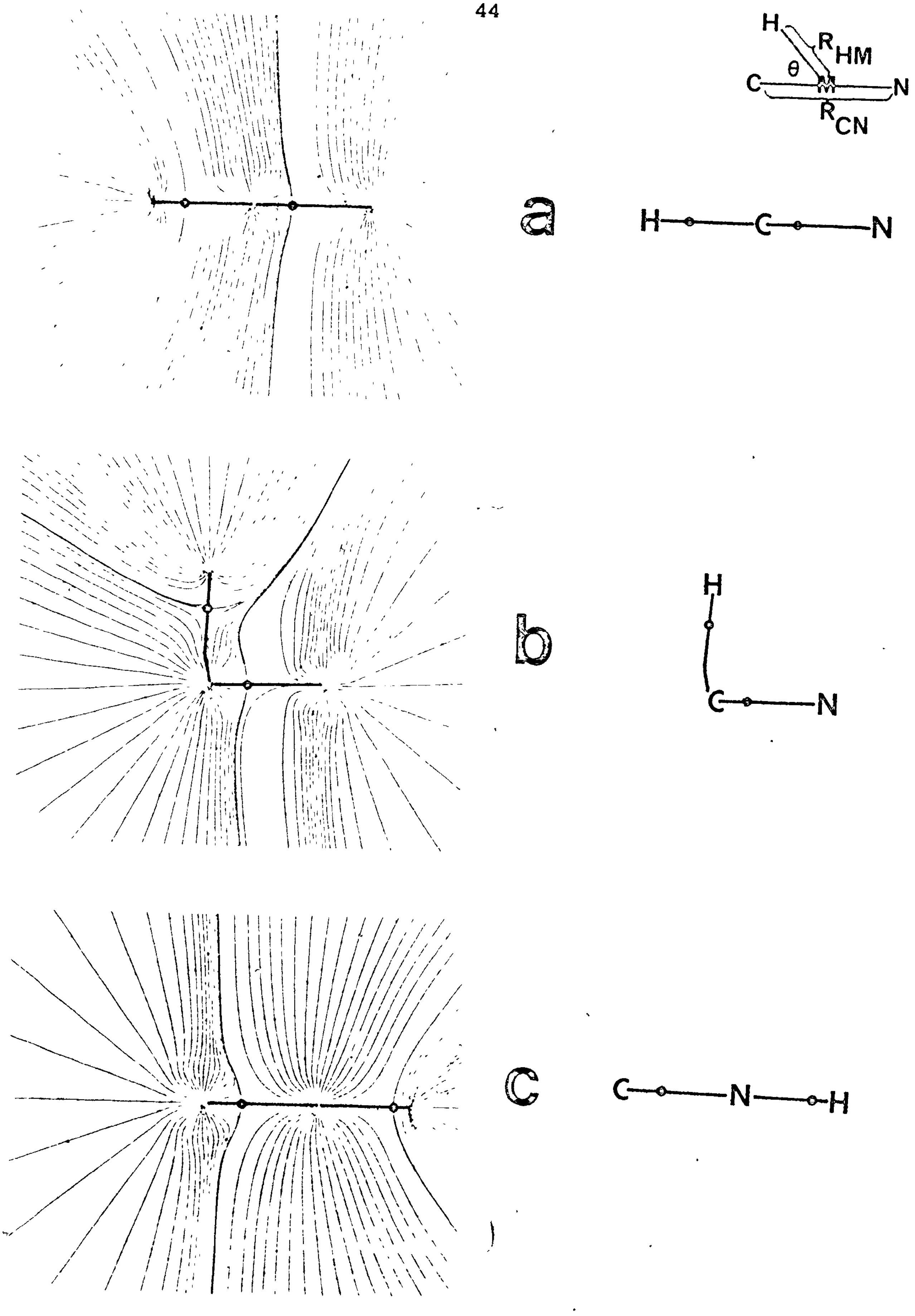
Consider two equivalent vector fields $\underset{\sim}{v}$ and $\underset{\sim}{v}'$, and let h be the homeomorphism mentioned above. Since h is, by definition, bijective and bicontinuous, it maps an $\omega(\alpha)$-limit set of a path g_m of $\underset{\sim}{v}$ into the corresponding $\omega(\alpha)$-limit set of a path g_m' of $\underset{\sim}{v}'$. In particular, if $\underset{\sim}{v}$ and $\underset{\sim}{v}'$ are the gradient vector fields $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$ and $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$ associated with the configurations $\underset{\sim}{X}$ and $\underset{\sim}{X}'$, respectively, the homeomorphism h maps a (τ,ζ) critical point of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ into a (τ,ζ) critical point of $\rho(\underset{\sim}{r},\underset{\sim}{X}')$. Moreover, a path connecting a pair of critical points $(\underset{\sim}{r}_c,\underset{\sim}{r}_c')$ of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ is transformed by h into a corresponding path of $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$ which connects the images of $\underset{\sim}{r}_c$ and $\underset{\sim}{r}_c'$. In other words, the molecular graph of the system in the configuration $\underset{\sim}{X}'$ will qualitatively be similar to the molecular graph arising from $\rho(\underset{\sim}{r},\underset{\sim}{X})$. However, these graphs may correspond to different bonding schemes, since the equivalence of the associated vector fields defined above only ensures that atoms correspond to each other in bonded pairs, regardless of their chemical nature. The equivalence of gradient vector fields of the charge distribution is illustrated in Fig. 2.1 on the example of the HCN system.[45,47] The three vector fields in Fig. 2.1 are equivalent, according to the above definition. One notes, however, that while the fields in Figs. 2.1a and 2.1b correspond to the same bonding scheme, represented by a molecular graph of the type H-C-N, the field in Fig. 2.1c corresponds to the molecular graph C-N-H. In order to topologically distinguish between two different bonding schemes which nevertheless arise from two equivalent vector fields $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$ and $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$, we make use of the following concepts of *structural stability*:[40,41,46,18,19] A configuration $\underset{\sim}{X} \in R^{3N}$ is *structurally stable* iff (if and only if) there exists a neighbourhood $V_{\underset{\sim}{X}}$ of $\underset{\sim}{X}$ in R^{3N} such that

$$\forall \underset{\sim}{X}' \in V_{\underset{\sim}{X}} \;,\quad \nabla\rho(\underset{\sim}{r},\underset{\sim}{X}') \sim \nabla\rho(\underset{\sim}{r},\underset{\sim}{X}) \tag{2.1}$$

An attractor F of the vector field $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$ is *structurally stable* iff, for every point $\underset{\sim}{X}'$ sufficiently close to $\underset{\sim}{X}$, there is an attractor F' of the associated field $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$ and a *local* homeomorphism h_{loc}: $R^3 \to R^3$ such that $h_{loc}(F) = F'$ and h_{loc} maps trajectories of $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$ into trajectories of $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$; moreover, h_{loc} is required to be close to the identity.

According to the above definition, a structurally stable con-

Figure 2.1. Charge distributions of the HCN system, at three different nuclear configurations, along the minimum energy path of the ground-state potential energy surface.[47] The charge distributions are here displayed via their gradient vector fields. Inset at the top of the figure is the coordinate system used to geometrically characterize the configurations of the minimum energy path: for a given internuclear distance R_{CN}, M is the centre of mass of the two-nuclei system (C,N). The position of the Hydrogen is uniquely defined by the values of the parameters R_{HM} (distance between the proton and M) and θ. The minimum energy path is obtained by minimizing the ground-state electronic energy E_o (R_{CN}, R_{HM}, θ) with respect to R_{CN} and R_{HM} for each given value of θ (P.K. Pearso, H.F. Schaeffer III and W. Wahlgren, J. Chem. Phys. 62, 350 (1972)). In the above coordinate system, (a) corresponds to $\theta = 0.0$ deg., (b) corresponds to $\theta = 60.0$ deg., and (c) to 180.0 deg. The gradient vector fields exhibited in (a)-(c) are equivalent to one another. However, (a) and (b) give rise to molecular graphs (shown opposite to the respective gradient maps) which correspond to the structure H-C-N while (c) corresponds to the structure C-N-H. Only through an analysis of structural stability does the distinction between (c) and (a) (or (b)) becomes clear.


H
R_HM
θ
C
N
R_CN
a
H
C
N
b
H
C
N
c
C
N
H

figuration, $\underset{\sim}{X}_0 \in R^{3N}$ admits a maximal neighbourhood, W, which satisfies eq. (2.1). The set W is simply the union of all $V_{\underset{\sim}{X}_0}$ of the type defined in eq. (2.1). One readily verifies that all the points of W are structurally stable, and that W plays the same role with respect to any of its points as it does with $\underset{\sim}{X}_0$. That is, for any $\underset{\sim}{X} \in W$, W is the largest $V_{\underset{\sim}{X}}$ satisfying eq. (2.1). Thus, since $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}) \sim \nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$, for any pair $\underset{\sim}{X},\underset{\sim}{X}' \in W$, one can find a homeomorphism $h_{\underset{\sim}{X},\underset{\sim}{X}'}$, which maps the gradient paths of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ into those of $\rho(\underset{\sim}{r},\underset{\sim}{X}')$. We have observed that, in general, the homeomorphisms $h_{\underset{\sim}{X},\underset{\sim}{X}'}$ also ensure the structural stability of all the nuclear attractors of the molecular system over W, in the sense of the definition given above. That is, to an attractor F_i of $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X})$ corresponds an attractor F_i' of $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}')$ such that

$$F_i' = h_{\underset{\sim}{X},\underset{\sim}{X}'}(F_i)$$

Furthermore, in some neighbourhood U_i of F_i, the homeomorphism $h_{\underset{\sim}{X},\underset{\sim}{X}'}$ approaches the identity transformation, as $|\underset{\sim}{X}' - \underset{\sim}{X}| \to 0$:

$$\forall\, \underset{\sim}{r} \in U_i, \qquad \lim_{|\underset{\sim}{X}-\underset{\sim}{X}'|\to 0} h_{\underset{\sim}{X},\underset{\sim}{X}'}(\underset{\sim}{r}) = \underset{\sim}{r}$$

Consequently, the molecular graphs associated with $\underset{\sim}{X}$ and $\underset{\sim}{X}'$ contain the same number of bond-paths, linking the same nuclear attractors. They are said to be *equivalent.*[18,19] Anticipating the definition of molecular structure, we call the open set W considered above, a structural region. Thus, a *structural* region, W, is the maximal neighbourhood of some structurally stable configuration $\underset{\sim}{X}$, which satisfies eq. (2.1). Then, two molecular graphs are equivalent iff they are associated with two configurations belonging to the same structural region. Two equivalent molecular graphs represent the same bonding scheme, which we shall refer to as a structure of the molecular system. In other words, a molecular structure is defined as an equivalence class of molecular graphs.[18,19] From the definition of the equivalence relation of molecular graphs, it is seen that a unique structure is associated with all the configurations of a structural region. In contrast, to each point of the nuclear configuration space, there corresponds a distinct molecular graph. A property which holds true for all the points of an open, dense subset, of a set E, is said to be generic in E.[40,41,46] Thus, a structure is generic in the closure, $\bar{W}$, of the associated structural region W, whereas a molecular graph is nowhere generic. Another example of a non-generic property is the molecular geometry, defined as the set of internuclear distances and angles.

Through the above definition of structure, which employs the notions of structural stability, the nuclear configuration space of a molecular system is partitioned into a finite number, ℓ, of structural regions, W_i, $i = 1,2,\ldots,\ell$. We have observed that, in all systems studied, the union of these structural regions, $\cup_i W_i$, is dense in R^{3N}, i.e., its closure equals R^{3N},

$$\overline{\left(\bigcup_{i=1}^{\ell} W_i\right)} = R^{3N}$$

A point which belongs to this union, belongs to some structural region, and is called a regular point. A configuration, $\underset{\sim}{Y}$, belonging to the complementary, C, of the set of regular points, is called a catastrophe point. The catastrophe set, C, is thus the collection of all the structurally unstable, catastrophe points of the nuclear configuration space. Let ∂W_i denote the boundary of the structural region W_i. Since $\cup_i W_i$ is dense in R^{3N},

$$C = \bigcup_{i=1}^{\ell} (\partial W_i) \qquad (2.2)$$

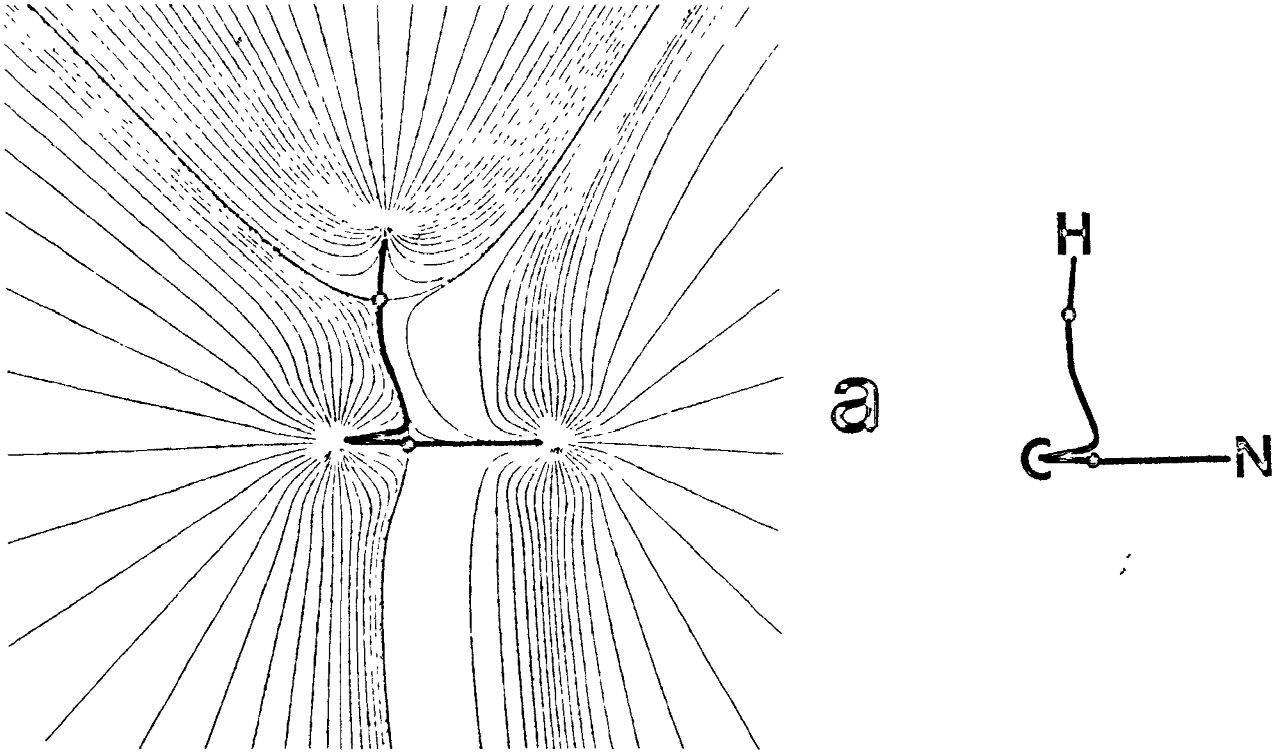
viz. the catastrophe set is the union of the boundaries of all the ℓ structural regions W_i. The result of eq. (2.2) depicts the catastrophe set C as the loci of structural changes. Consider two points, $\underset{\sim}{X}_1, \underset{\sim}{X}_2$, belonging to two neighbouring structural regions W_1, W_2. A continuous deformation of the system from the nuclear configuration $\underset{\sim}{X}_1$, leading to the configuration $\underset{\sim}{X}_2$, will necessitate the passage of the system through a catastrophe configuration, $\underset{\sim}{X}_C$, at which the molecular graph is neither characteristic of W_1, nor is it representative of the structure associated with W_2. It represents an intermediate, or the transition, between the structures associated with the structural regions W_1, W_2. We may use the notions of structural stability in which a subset C' of C replaces R^{3N}, C' being topologized as a submanifold of R^{3N} to define structures over C, just as molecular structures were defined over the nuclear configuration space, and assigned to regular points of this space. We thereby obtain an equivalence relation of the molecular graphs associated with the points of C, which gives a partitioning of C into structural regions. Each of these structural regions is associated with an equivalence class of molecular graphs, that is a structure, over C. Again the boundaries of these structural regions together form the loci of structural changes in C, its points being structurally unstable (both in C and R^{3N}). In principle, the above analysis can be repeated indefinitely, until no further differentiation based on structural stability (over some submanifold of R^{3N}) is obtained. Thus, structure is definable over the entire nuclear configuration space of a molecular system and is generic in some closed non-empty subset of this space. We shall henceforth call *regular structure* a structure which is found at a regular point. A regular structure is thus uniquely attached to a (3N dimensional) open subset of the nuclear control space. In contrast, a *catastrophe structure*, observed at a catastrophe point is associated with a closed subset of the control space. Finally, we distinguish two types of catastrophe points:[18,40,41] a point $\underset{\sim}{X}$ of C is a *bifurcation (catastrophe) point* iff the associated (scalar) field $\rho(\underset{\sim}{r},\underset{\sim}{X})$ exhibits a *singularity*, i.e., a critical point of rank less than three. The study of bifurcation catastrophes, will form the subject matter of the next section. A point of C which is not a bifurcation point is called a conflict catastrophe point. Thus, at a conflict catastrophe point $\underset{\sim}{Y}$, the charge distribution does not exhibit any singularity: all the critical points of $\rho(\underset{\sim}{r},\underset{\sim}{Y})$ are of rank three and can be identified, according to their signature, as described in Chapter I. The configuration $\underset{\sim}{Y}$ is unstable, however: while a slight displacement of the system from $\underset{\sim}{Y}$ preserves the number and types of critical points, it causes a change in the structure. In general, it is found that a catastrophe structure of the conflict type involves a bond-path which terminates at a saddle point, i.e., a pseudo-attractor, instead of a nuclear attractor. Upon a slight displacement from the catastrophe point, this bond-path is subjected to a change in one of its ω-limit sets, which, previously a pseudo-attractor, is now found coincident with a nuclear attractor. Thus, the mechanism of the structural change encountered at a conflict catastrophe point is the switching of the ω-limit set of a bond-path from one attractor to another, a mechanism which we shall call switching of attractors.[18,19]

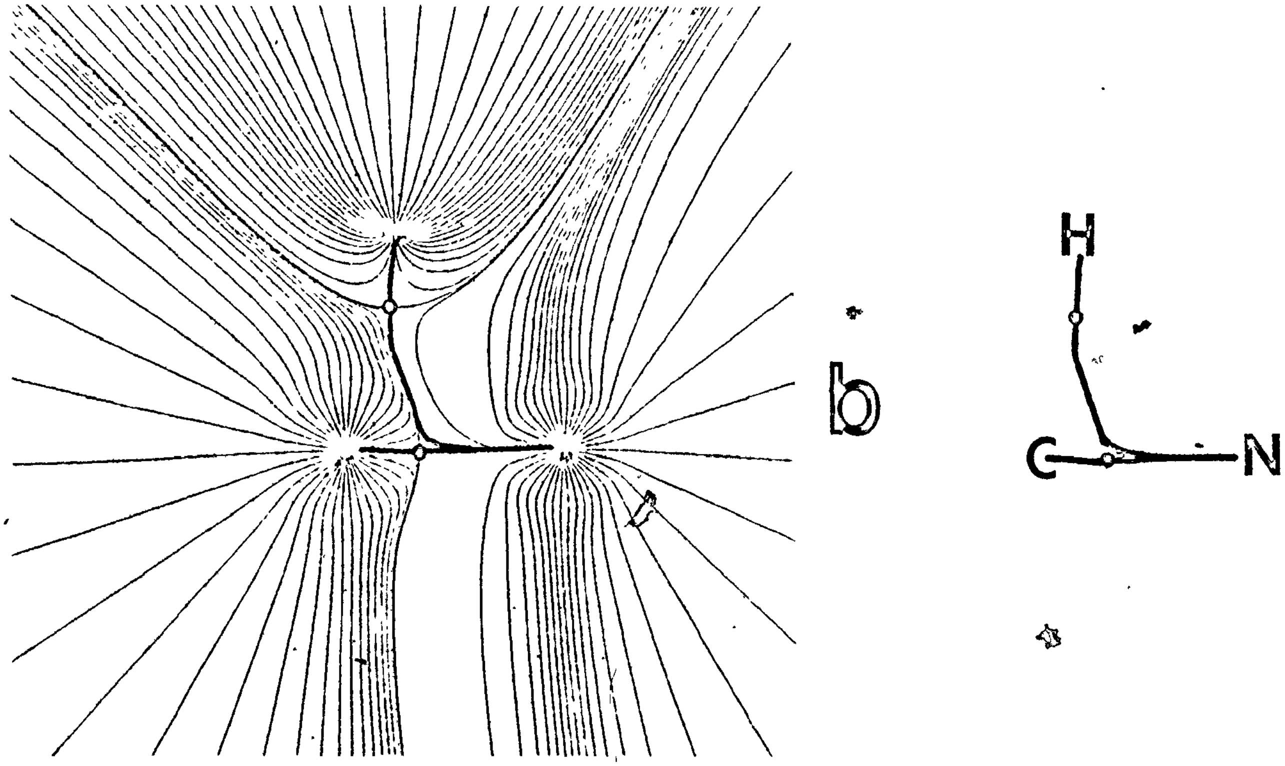
To illustrate the definitions and ideas advanced above, we again consider the example of the HCN system, and refer to Fig. 2.1. The molecular graphs of Fig. 2.1a and 2.1b are equivalent. They represent a unique structure, which is designated by the graph, H-C-N. The associated nuclear configurations thus belong to the same structural region, which does not contain the configuration illustrated in Fig. 2.1c. The molecular graph associated with this configuration is not equivalent to the molecular graphs of Figs. 2.1a and 2.1b, and represents a different structure, in which the nitrogen atom is bonded to the proton, a structure designated by C-N-H. To obtain this structure from the molecular structure of Figs. 2.1a and 2.1b, a passage through a catastrophe configuration is necessitated. Detailed numerical calculations have indicated that a catastrophe point of the conflict type separates the configurations of Figs. 2.1a, 2.1b from the configuration of Fig. 2.1c, along the path of minimum energy for the rearrangement reaction I[47]

$$\text{H-C-N} \rightarrow \text{C-N-H} \qquad \text{I}$$

Because of its inherent structural instability, a catastrophe point can seldom be located precisely in a numerical study, but its existence and its nature can be established unambiguously by the observations made within a sufficiently small neighbourhood of the catastrophe point. Figs. 2.2 illustrate this procedure, for the identification of the conflict catastrophe point encountered in the reaction I. The two gradient maps and molecular graphs of Fig. 2.2 refer to two geometries lying in a small interval of the minimum energy path for this reaction.[47] The field of Fig. 2.2a exhibits the structure common to Figs. 2.1a, 2.1b. The proton is still linked to

Figure 2.2. Charge distribution of the HCN system, at two nuclear configurations along the minimum energy path of the ground-state potential energy surface.[47] The configurations illustrated lie within a small neighbourhood of the conflict catastrophe configuration found along the above path, which is the reaction path of the rearrangement reaction (I). As in Fig. 2.1, the charge distribution is exhibited via its gradient vector field. With respect to the coordinate system defined in Fig. 2.1, (a) refers to θ = 72.1 deg, and (b) refers to 72.4 deg. A switching of attractors is observed in going from (a) to (b). The bond path involved in this process is the one linking the proton to the rest of the system. The conflict catastrophe configuration thus corresponds to a value of θ (in deg.) in the interval [72.1, 72.4]. The transition state was determined to correspond to θ = 70.2 deg. when, along the reaction path, all wavefunctions are approximated by single-determinantal SCF wavefunctions, and θ = 73.7 deg. from a set of CI calculations. It thus appears that, to the accuracy of the determination of the minimum energy path, the conflict catastrophe occurs in the region of the transition state.


a
H
C
N

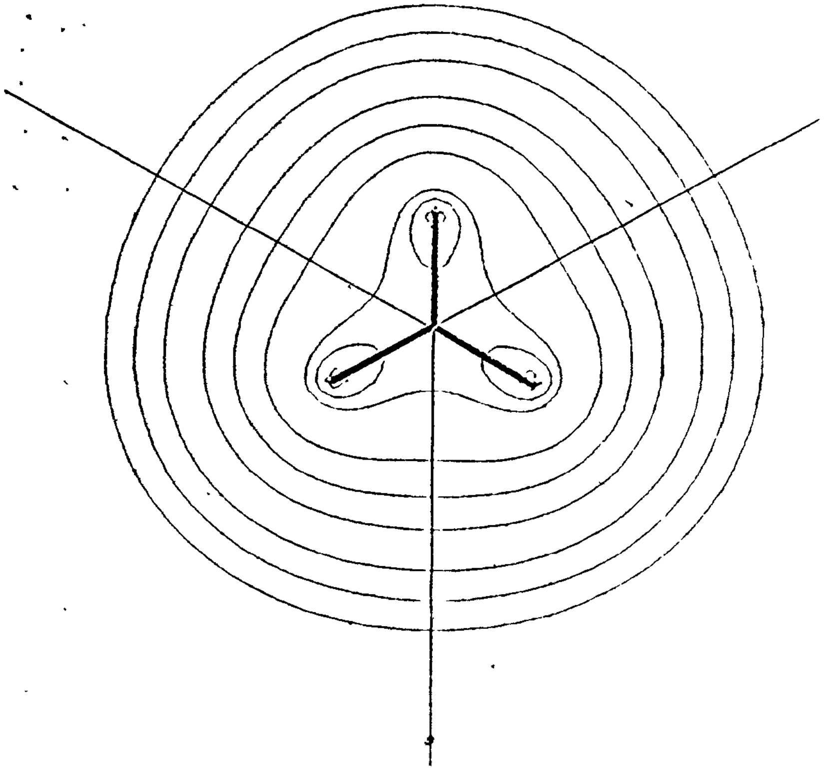

b
H
C
N

the carbon nucleus by a bond-path. However, near the carbon nuclear attractor, this bond-path is strongly curved, and closely approaches the (C-N) bond. Upon a slight displacement of the proton towards the nitrogen atom, the field of Fig. 2.2b is obtained, in which the proton is bonded to the nitrogen atom, via a bond-path which almost merges with the (C-N) bond, near its terminus at the nitrogen nuclear attractor. The molecular graph of Fig. 2.2b represents the same structure as that of Fig. 2.1c. The observed abrupt change in the structure, as the configuration of Fig. 2.2a is deformed into that of Fig. 2.2b, and vice-versa, only involves a switching of attractors, thus indicating the passage of the system through a conflict catastrophe configuration. The particular appearance of the bond-path which connects the proton to the rest of the system, in both Figs. 2.2a and 2.2b, suggests that in the conflict catastrophe structure, this bond-path terminates at the (3,-1) saddle point of the C-N bond.

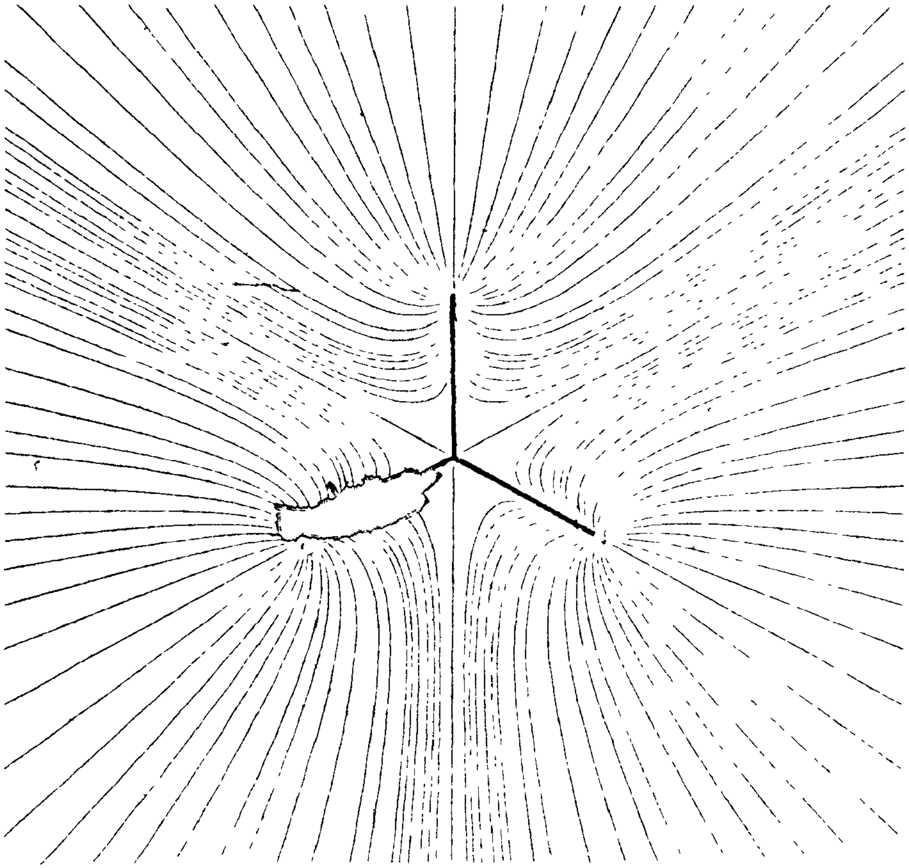
It has been found that the interval of the reaction path in which the two configurations illustrated by Fig. 2.2 are located, also corresponds to the transition state separating the reactant and the product of reaction I.[47] Thus, in this specific example, the topological intermediate, the conflict catastrophe configuration, is observed to coincide with an energetic instability, the transition state of the rearrangement reaction I. The structure of the reactant, denoted

H-C-N, is observed over a large neighbourhood of the equilibrium geometry, which, along the reaction path, extends from the conflict catastrophe point to the dissociation asymptote corresponding to the dissociated system H + CN. This observation justifies the assignment of the above structure to the HCN molecule. An exceptional case arises when the equilibrium geometry of a system corresponds to a catastrophe point. An example is the H_3^+ system, which has only two electrons to bind three nuclei.[18,45] In the ground state, and at its equilibrium geometry, the charge distribution of this system exhibits a singularity. The phase portrait associated with this singularity is shown in Fig. 2.3b. It is characteristic of a critical point of rank one, which will be studied in detail in the next section. The ground state equilibrium nuclear configuration of H_3^+ is thus a bifurcation point, whose associated catastrophe structure is shown in Fig. 2.3a. This structure is found only at the equilibrium geometry, as well as at neighbouring nuclear configuration of D_{3h} symmetry, and thus is not observable. To determine the other structures accessible to the system in the neighbourhood of the equilibrium geometry, we subject one of the protons to various displacements from its equilibrium position. For a displacement in the dissociative direction which preserves a two-fold symmetry axis, thus yielding a nuclear configuration of C_{2v} symmetry, a gradient map and a molecular graph of the

<u>Figure 2.3</u>. Charge distribution of the H_3^+ ground-state system, at the equilibrium geometry and in the plane of the nuclei. The ground-state equilibrium geometry illustrated has D_{3h} symmetry and is characterized by $d_1 = d_2 = d_3 = d = 1.6575$ a.u., where d_1, d_2, d_3 denote the distances between the pairs (H1,H2), (H2,H3), (H3,H1), respectively. (a) Contours of ρ. Superimposed on the contours are also shown the molecular graph and the interatomic surfaces of the system. (b) Trajectories of $\nabla\rho$. The charge distribution exhibits a <u>singularity</u> at the intersection of the two-fold symmetry axes of the system. The singularity is of corank two, codimension three, and its phase portrait is characteristic of the <u>elliptic umbilic</u> singularity (see Section B of this Chapter). The same type of singularity and the same molecular graph are observed at neighbouring D_{3h} configuration. Thus, the molecular graph shown in (a) is representative of a <u>bifurcation catastrophe</u> structure, found at all configurations of D_{3h} symmetry in the neighbourhood of the equilibrium geometry.



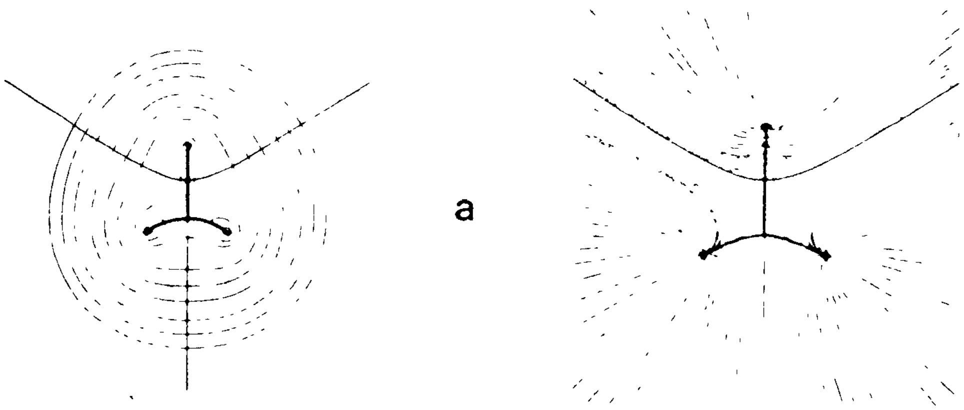
a

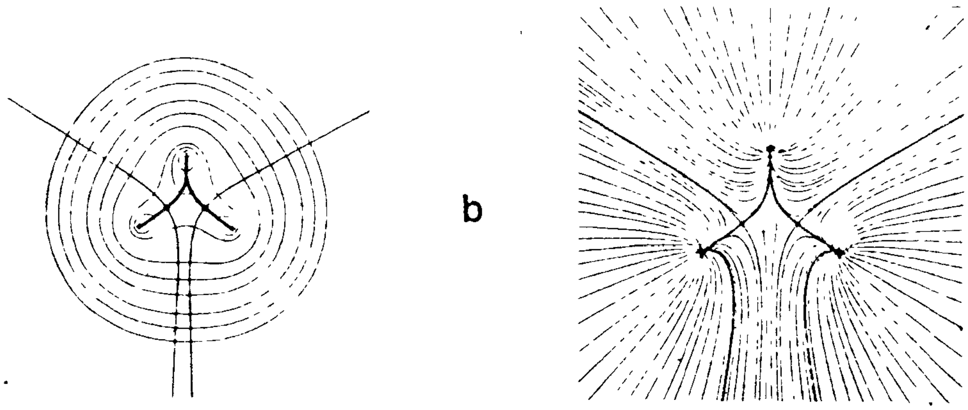


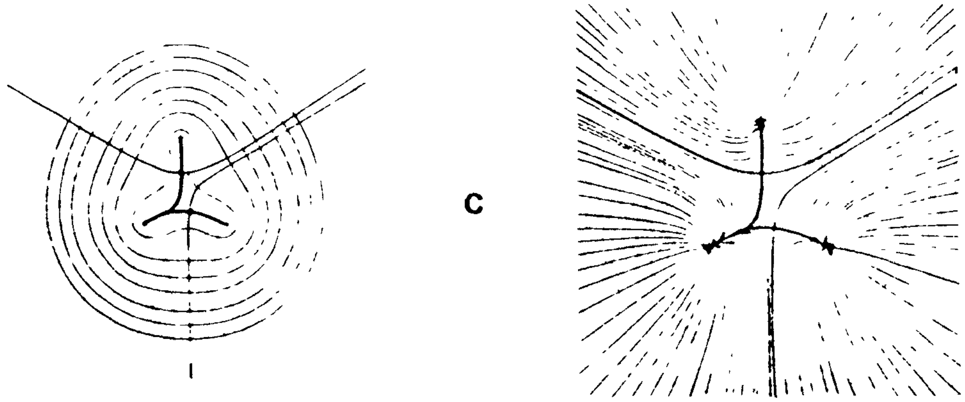
b

type shown in Fig. 2.4a are obtained. The displaced proton is now bonded to a hydrogen molecule, H_2, via its (3,-1) bond critical point. The corresponding structure will be indicated by H-(H-H). If a two-fold symmetry axis, of the original equilibrium geometry, is not strictly preserved in the above displacement, thus yielding a configuration of C_s symmetry, a molecular graph of the type shown in Fig. 2.4c would result. In other words, the nuclear configuration illustrated by Fig. 2.4a is a conflict catastrophe configuration, since the associated charge distribution does not exhibit any singularity, and the configuration is structurally unstable. The structure corresponding to Fig. 2.4c, denoted by H-H-H, is obtained for all configurations of C_s symmetry. It is also observed for configurations of C_{2v} symmetry resulting from a displacement of one proton towards the two other protons, Fig. 2.4b. It therefore is a regular structure. Above, we have noted that the configurations of D_{3h} symmetry in the neighbourhood of the equilibrium geometry, $\underset{\sim}{X}_e$, are also bifurcation configurations of the same type as $\underset{\sim}{X}_e$. We can repeat the above study, to determine the structural changes encountered upon a deformation of the system from each of these bifurcation configurations. It is found that the same structural changes are observed when one of the protons is subjected to the displacements described above, from its position in any of these bifurcation configurations. This allows us to characterize the catastrophe set in some neighbourhood of

Figure 2.4. Charge distributions of H_3^+ ground-state system, at three nuclear configurations lying in a neighbourhood of the equilibrium configuration. In each of the three sets of diagrams, the charge distribution is displayed via its contours and the trajectories of its gradient vector, in the plane of the nuclei. Superimposed on the contours of ρ are also shown the molecular graph and the interatomic surfaces. With d_1, d_2, d_3 as defined in Fig. 2.3, diagrams (a) correspond to $d_1 = d_3 = 1.8$ a.u., $d_2 = 1.6$ a.u. The molecular graph obtained is representative of the *conflict catastrophe* structure H-(H-H). The same structure is observed when, starting from its position in the equilibrium geometry, a proton is subjected to a displacement in the dissociative direction, which preserves a two-fold symmetry axis. When the direction of this displacement is reversed, the *regular structure* H-H-H represented by the molecular graph of diagrams (b), is obtained. The diagrams (b) were obtained for $d_1 = d_3 = 1.6$ a.u., $d_2 = 1.8$ a.u. The regular structure H-H-H is also observed when the uppermost proton in (a) is further subjected to a lateral displacement of magnitude 0.1 a.u., resulting in diagrams (c). This illustrates the inherent instability of the conflict structure of (a).


a


b


c

$\underset{\sim}{X}_e$ in a simple manner, when the nuclear configuration space is coordinatized by the internuclear distances. Let d_1, d_2, d_3 denote the (Euclidian) distances between the pairs (H1,H2), (H2,H3) and (H3,H1), respectively. The nuclear configurations of D_{3h} symmetry then constitute the line Δ, defined by

$$\Delta: \quad d_1 = d_2 = d_3$$

Let d be the value of $d_1 = d_2 = d_3$ in the equilibrium configuration $\underset{\sim}{X}_e$. A displacement of H1, from its position in $\underset{\sim}{X}_e$, in the dissociative direction, and which preserves a two-fold symmetry axis, gives a point along the semi-axis δ_1, defined by

$$\delta_1: \quad d_1 = d_3 > d_2, \; d_2 = d$$

Thus, along δ_1, the system assumes the conflict structure H1-(H2-H3). Likewise, similar displacements of H2,H3, respectively, define the semi-axes δ_2, δ_3

$$\delta_2: \quad d_2 = d_1 > d_3, \; d_3 = d$$

$$\delta_3: \quad d_3 = d_2 > d_1, \; d_1 = d$$

which are associated with the conflict structures H2-(H1-H3), and H3-(H1-H2). The rectangular system of semi-axes $(\delta_1, \delta_2, \delta_3)$ thus denotes the loci of all the conflict catastrophe structures accessible from the bifurcation point $\underset{\sim}{X}_e$ upon the displacement of a single proton. The above construction can be repeated for each point $\underset{\sim}{X}_b \in \Delta$, in some neighbourhood of $\underset{\sim}{X}_e$, yielding in each case, a rectangular system of semi-axes of the same type as the system $(\delta_1, \delta_2, \delta_3)$ found above, but with a different d, and which denotes the conflict catastrophes encountered as the geometry of the system is deformed from the one associated with $\underset{\sim}{X}_b$, by displacing a single proton appropriately. Thus, the catastrophe set is seen to be the union of three half-planes disposed symmetrically with respect to and meeting each other at the line Δ, Fig. 2.5a. Consider a plane, P, orthogonal to Δ at $\underset{\sim}{X}_e$. Its intersection with the catastrophe set is shown in Fig. 2.5b. It is constituted of three semi-axes, $\delta_1', \delta_2', \delta_3'$, which can be regarded as the orthogonal projections of $\delta_1, \delta_2, \delta_3$ respectively, onto the plane P. Thus, δ_i', the projection of δ_i, denotes the loci in P, of the conflict catastrophe structure Hi-(Hj-Hk), which is seen to be obtained from the bifurcation structure associated with $\underset{\sim}{X}_e$ by a simultaneous displacement of the three protons, such that

$$d_{(ij)} = d_{(ik)} > d_{(jk)}$$

and

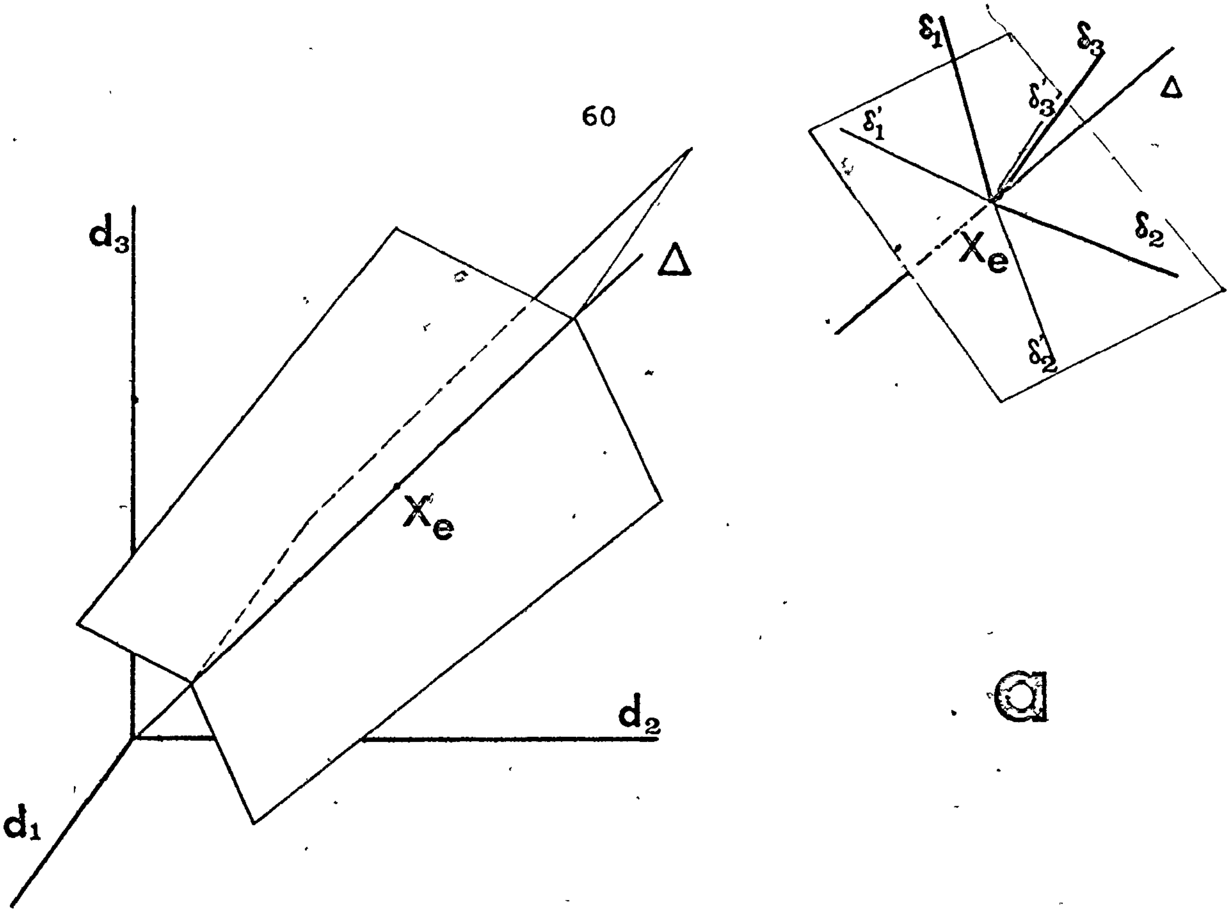
$$d_{(ij)} + d_{(ik)} + d_{(jk)} = 3d.$$

The last condition has been imposed to ensure that the nuclear configuration resulting from the above displacement lies in P.[48] We have also used the parentheses (ij) to denote the unique index which is associated with the pair (i,j), according to the rule

$$(ij) = (ji) = \begin{cases} 1 & \text{if} \quad i,j = 1,2 \\ 2 & \text{if} \quad i,j = 2,3 \\ 3 & \text{if} \quad i,j = 1,3 \end{cases}$$

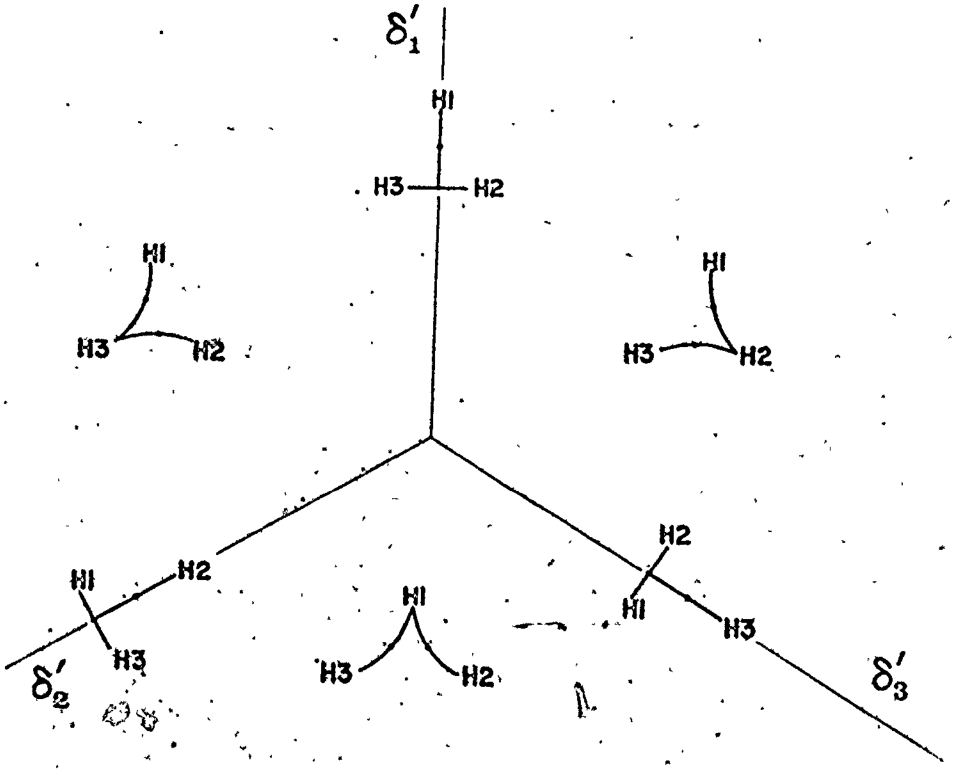
The three semi-axes, $\delta_1^{'}, \delta_2^{'}, \delta_3^{'}$, partition the plane P into three structural regions associated with the regular structures Hi-Hj-Hk, i,j,k = 1,2,3. These regular structures and the conflict structures observed along $\delta_1^{'}$, $\delta_2^{'}$ and $\delta_3^{'}$ are also represented in Fig. 2.5b. Fig. 2.5b is an example of what is called a structure diagram. It is defined generally as the pictorial representation of the catastrophe set in some subset W of the nuclear configuration space, on which is also superimposed the representation of all the structures defined over W. In the above example, W is chosen to be the plane P, which is orthogonal to the line Δ at $\underset{\sim}{X}_e$. A structure diagram is the portrait of what is termed a process in Catastrophe theory. In other applications of Catastrophe theory, the catastrophe set is completely defined by the process.[40,41] In the present context, we adopt the reverse point of view and instead regard the structure diagram as defining the process. Corresponding to the notion of isomorphic processes in Catastrophe theory, we define: two structure diagrams over the respective subsets W,W' of the nuclear configuration space are isomorphic iff there exists a homeomorphism, h:W → W', such that (i) h maps the catastrophe set K in W onto the catastrophe set K' in W' and (ii) restricted to regular points of W, h is a diffeomorphism[49] onto the regular points of W'. To illustrate this

Figure 2.5. The *catastrophe set* of the H_3^+ ground-state system, in a neighbourhood of the equilibrium nuclear configuration, $\underset{\sim}{X}_e$. (a) When the nuclear configuration space is coordinatized by d_1, d_2, d_3 (cf. Figs. 2.3 and 2.4), *locally*, the catastrophe set is found to be the union of three half-planes meeting at, and disposed symmetrically with respect to the line Δ: $d_1 = d_2 = d_3$. Also shown, in the inset at the top of the diagram, is the system of semi-axes (δ_1, δ_2, δ_3) used in the text, and its projection, (δ_1', δ_2', δ_3'), onto the plane P perpendicular to Δ at $\underset{\sim}{X}_e$. (b) The cross-section of the catastrophe set through the plane P is the union of the three semi-axes δ_1', δ_2', δ_3'. These denote the loci of conflict catastrophes encountered when the protons are displaced in such a manner as to preserve the value of the sum $d_1 + d_2 + d_3$. Superimposing, on this cross-section, the structures observed in P, we obtain a *structure diagram*, which summarizes all structural changes observable in the neighbourhood of the equilibrium configuration.


d3
Δ
Xe
d2
d1
δ1
δ3
δ'3
Δ
δ'1
Xe
δ2
δ'2

a

b


δ'1
H1
H3
H2
H1
H3
H2
H1
H3
H2
H1
H2
H3
H1
H3
H2
H2
H1
H3
δ'2
δ'3

notion, consider the structure diagram of Fig. 2.5b, which, as noted above, is defined over W = P. Let W' be the union of the planar sectors defined by the pairs (δ_1,δ_2), (δ_1,δ_3), (δ_2,δ_3). We obtain a structure diagram over W' by assigning to these planar sectors and the semi-axes δ_i, i = 1,2,3, the appropriate regular and conflict structures. We have noted that the catastrophe set in $W, K = \bigcup_{i=1}^{3} \delta_i'$, is just the orthogonal projection of the catastrophe set in W', $K' = \bigcup_{i=1}^{3} \delta_i$, onto W. Regarded as a mapping between W' and W, this projection is a homeomorphism which sends K' into K and vice-versa, and which is a diffeomorphism over the set of regular points in W'. Thus, the structure diagram over W', whose construction is outlined above, is isomorphic with the diagram of Fig. 2.5b. One can easily verify that the structure diagram over any plane P' through $\underset{\sim}{X}_e$, which is not parallel to Δ, is isomorphic with the structure diagram over P, Fig. 2.5b. The isomorphism of structure diagrams defined above is an equivalence relation. An equivalence class arising from this equivalence relation is represented by a typical structure diagram. Thus, all the structures and structural changes which may be observed for a given molecular system are determined by the complete specification of the typical structure diagrams associated with the system. In most cases, we are interested in the structures and structural changes which occur in some limited part of the nuclear configuration space, and often a single structure diagram can be used, which provides all the relevant information.

Thus, in the case of the H_3^+ system, the diagram of Fig. 2.5b determines all the structures and structural changes which can be observed in some neighbourhood of the equilibrium geometry, $\underset{\sim}{X}_e$. Three regular structures of the type Hi-Hj-Hk, separated from each other by the conflict structures Hi-(Hj-Hk), have been obtained, in the hypothesis that the protons are distinguishable. Recognizing that the protons are identical, hence indistinguishable, we conclude that the regular structure H-H-H is obtained over a locally dense open subset of the nuclear configuration space, and is thus generic. We, therefore, expect that this regular structure will likely be observed, if an attempt is made to experimentally determine the structure of the H_3^+ system.[50]

The examples of the HCN and H_3^+ systems considered above illustrate the utility and importance of the concepts of structural stability in defining molecular structure and delineating the possible mechanisms of structural changes in a molecular system. Molecular structure has been defined as an equivalence class of molecular graphs. By the very nature of the definition of the equivalence relation of molecular graphs, an equivalence class, hence a structure, is associated with an open region of the nuclear configuration space of the system. It thus satisfies the prerequisite for its being observable. While for HCN, it is expected that the observed structure will be the one associated with the equilibrium geometry of the system, for H_3^+, structural stability considerations anticipate that the observed structure will likely be the one associated with a regular configuration which results from a deformation of the equilibrium geometry. The use of the notions of structural stability also leads to the conclusion that a structural change can only occur via the establishment of a catastrophe point. Corresponding to the distinction between a catastrophe point of the conflict type and one of the bifurcation type, we distinguish two general mechanisms of structural changes: Structure can be changed either through the switching of attractors of some bond-paths, or via a mechanism which involves the creation of a singularity in the molecular charge distribution, at a bifurcation (nuclear) configuration. The determination of the structures and structural changes which are to be found in the neighbourhood of a bifurcation configuration is often facilitated by the application of Thom's Theory of Elementary Catastrophes.[40,41] In the following section, we shall review the statement of the central result of this theory, Thom's classification theorem, and illustrate its utility in the study of molecular structure and its change.

B. Bifurcation Catastrophes

In the previous section, we have defined a bifurcation catastrophe point as a nuclear configuration $\underset{\sim}{X}_b \in R^{3N}$, whose associated charge distribution, $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$, exhibits a singularity, i.e., a critical point $\underset{\sim}{r}_c$ whose Hessian matrix $\underline{\underline{A}}(\underset{\sim}{r}_c)$ satisfies

$$\det(\underline{A}(\underset{\sim}{r}_c)) = 0 \tag{2.3}$$

Thus, the Hessian matrix $\underline{A}(\underset{\sim}{r}_c)$ of a singularity $\underset{\sim}{r}_c$ possesses at least one zero eigenvalue. Let τ be the rank of $\underline{A}(\underset{\sim}{r}_c)$. The number $\sigma = (3 - \tau)$ is called the <u>corank</u> of the singularity. In some neighbourhood $\eta(\underset{\sim}{r}_c)$ of a singularity $\underset{\sim}{r}_c$ of corank σ, there exists a coordinate system centred at $\underset{\sim}{r}_c$, with respect to which ρ assumes the Taylor series expansion

$$\rho(\underset{\sim}{r},\underset{\sim}{X}_b) = \rho(\underset{\sim}{0}) + Q(\underset{\sim}{r}) + P(\underset{\sim}{r}), \quad \underset{\sim}{r} \in \eta(\underset{\sim}{r}_c) \tag{2.4}$$

where $Q(\underset{\sim}{r})$ is a quadratic form in the first τ components of $\underset{\sim}{r} \in \eta(\underset{\sim}{r}_c)$ and $P(\underset{\sim}{r})$ contains only terms of third or higher orders in $\underset{\sim}{r}$. In the above coordinates system, the definition of the singularity, $\underset{\sim}{r}_c = 0$, becomes

$$\forall\, i,j \in \{\tau,\ldots,3\}, \quad \left.\frac{\partial^2\rho}{\partial r_i \partial r_j}\right|_{\underset{\sim}{r}=\underset{\sim}{0}} = 0 \tag{2.5}$$

Equation (2.5) implies that the singularity is an unstable attractor of the vector field $\nabla_\sigma\rho$ over the intersection, $\eta(\underset{\sim}{r}_c) \cap R^\sigma$, of the neighbourhood $\eta(\underset{\sim}{r}_c)$ with R^σ, where $\nabla_\sigma\rho$ is the gradient vector of ρ with respect to the σ <u>essential variables</u> $(r_i, r_j, ..) \in R^\sigma$, $i = \tau,\ldots,3$. Thus, we can alternatively define a bifurcation point as a configuration $\underset{\sim}{X}_b \in R^{3N}$ for which $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ exhibits an unstable pseudo-attractor in some submanifold M of R^3. The unstable pseudo-attractor in question is then a singularity in $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ of corank $\sigma = \dim M$. When the configuration $\underset{\sim}{X}$ of the system is allowed to vary in some neighbourhood of $\underset{\sim}{X}_b$, a number of different structures, which are all distinct from the catastrophe structure found at $\underset{\sim}{X}_b$, will be observed. Henceforth, we shall use the name "morphogenesis" to designate the generation of new structures by varying the configuration of the system from a catastrophe point $\underset{\sim}{X}_c$ and within a neighbourhood of $\underset{\sim}{X}_c$. Thom has shown an important result which will be used in the following. Applied to our case, this result states that when $\sigma < 3$, the morphogenesis from a bifurcation point $\underset{\sim}{X}_b$ can locally be described by a simple algebraic function of the σ essential variables, those which were missing in the quadratic term $Q(\underset{\sim}{r})$ of equation (2.4). This function also depends on a finite number of control parameters. The minimum number q of control parameters needed is a characteristic of the singularity, and is called the <u>codimension</u> of the singularity. The simple algebraic function mentioned above is called <u>the universal unfolding</u> of the singularity.[40,41,46,51] It can be brought into one of seven canonical forms, the unfoldings of the seven "elementary catastrophes" in Thom's classification, through a diffeomorphism, i.e., a bijective smooth mapping of the behaviour manifold M into itself. An elementary exposition of some of the results embodied by Thom's classification theorem, and an extended list of the elementary catastrophes and their unfoldings can be found in the recently published book by T. Poston and I. Stewart.[52] In the two following remarks, we briefly describe how Thom's theorem works when applied to the analysis of structural stability in molecular systems:

i) The unfolding provides a local mathematical model of the morphogenesis, in the sense that it suitably describes the field $\rho(\underset{\sim}{r},\underset{\sim}{X})$ only for configurations $\underset{\sim}{X}$ lying in some neighbourhood of the bifurcation point $\underset{\sim}{X}_b$, and for $\underset{\sim}{r}$ belonging to some small neighbourhood, $\eta_M(\underset{\sim}{r}_c) \subseteq M$, of the singularity $\underset{\sim}{r}_c$. Let $f(\underset{\sim}{x};\underset{\sim}{\mu})$ denote the universal unfolding of the singularity $\underset{\sim}{r}_c$, in its canonical form, $\underset{\sim}{x}$ being the coordinates of a point sufficiently close to $\underset{\sim}{r}_c$ in M, with respect to a coordinates system centred at $\underset{\sim}{r}_c$, and $\underset{\sim}{\mu} \in B \subset R^q$ is the set of q control parameters, such that $\underset{\sim}{\mu} = \underset{\sim}{0}$ (the origin of R^q) corresponds to the bifurcation point $\underset{\sim}{X}_b$. Then, $f(\underset{\sim}{x};\underset{\sim}{\mu} = \underset{\sim}{0})$ is diffeomorphic to the restriction of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ to $\eta_M(\underset{\sim}{r}_c)$, and possesses a single critical point at $\underset{\sim}{x} = \underset{\sim}{0}$. We have noted above that in $\eta_M(\underset{\sim}{r}_c)$, this critical point is, in fact, an unstable attractor of the associated gradient vector field $\nabla_\sigma\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$. Since $f(\underset{\sim}{x};\underset{\sim}{0})$ is diffeomorphic to $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)\big|_{\eta_M(\underset{\sim}{r}_c)}$, the corresponding vector field $\nabla_x f(\underset{\sim}{x};\underset{\sim}{0})$ is equivalent to $\nabla_\sigma\rho(\underset{\sim}{r},\underset{\sim}{X}_b)\big|_{\eta_M(\underset{\sim}{r}_c)}$, and $\underset{\sim}{x} = \underset{\sim}{0}$ is an unstable attractor of $\nabla_x f(\underset{\sim}{x};\underset{\sim}{0})$. As $\underset{\sim}{\mu}$ is made to vary about the origin in R^q, the unstable attractor either disappears or bifurcates into a number of critical points whose locations, for each given $\underset{\sim}{\mu}$, can be found by solving the equation

$$\nabla_x f(\underset{\sim}{x};\underset{\sim}{\mu}) = \underset{\sim}{0}.$$

Through a local diffeomorphism, the set B of the control parameters $\underset{\sim}{\mu}$ can be identified with a neighbourhood, W, of $\underset{\sim}{X}_b$ in a q-dimensional submanifold of the control space. Then, for each $\underset{\sim}{\mu} \in B$, there exists a configuration $\underset{\sim}{X} \in W$ such that $f(\underset{\sim}{x};\underset{\sim}{\mu})$ is diffeomorphic to the restriction, $\rho(\underset{\sim}{r};\underset{\sim}{X})\big|_{\eta_M}$, of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ to $\eta_M(\underset{\sim}{r}_c)$. Thus, the critical points of $f(\underset{\sim}{x};\underset{\sim}{\mu})$ are also critical points of $\rho(\underset{\sim}{r},\underset{\sim}{X})\big|_{\eta_M}$. Now, the unfolding $f(\underset{\sim}{x};\underset{\sim}{\mu})$ only describes the topological properties of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ with respect to the σ essential variables which locally span the submanifold M. Referring to eq. (2.4), we recall that the absence of these σ variables in $Q(\underset{\sim}{r})$ is responsible for the establishment of the singularity in $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$. While $f(\underset{\sim}{x};\underset{\sim}{0})$ appropriately describes the restriction of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ to $\eta_M(\underset{\sim}{r}_c)$, the quadratic expression $Q(\underset{\sim}{r})$ governs the local topological properties of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ with respect to the τ directions orthogonal to $\eta_M(\underset{\sim}{r}_c)$. Noteworthy is the remark that all points of $\eta_M(\underset{\sim}{r}_c)$ are regular critical points of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ with respect to these directions. As $\underset{\sim}{X}$ varies in W (from $\underset{\sim}{X}_b$), $f(\underset{\sim}{x};0)$ evolves into $f(\underset{\sim}{x};\underset{\sim}{\mu})$ for some corresponding $\underset{\sim}{\mu} \in B$, $\underset{\sim}{\mu} \neq \underset{\sim}{0}$, and $Q(\underset{\sim}{r})$ into some function, denoted $Q_{\underset{\sim}{X}}(\underset{\sim}{r})$. Now $Q(\underset{\sim}{r})$, is structurally stable, in the sense that, for any $\underset{\sim}{X}$ sufficiently close to $\underset{\sim}{X}_b$, $Q_{\underset{\sim}{X}}(\underset{\sim}{r})$ is diffeomorphic to $Q(\underset{\sim}{r})$, and hence gives rise to an equivalent gradient vector field to that of $Q(\underset{\sim}{r})$. This equivalence ensures that essential topological properties of $Q(\underset{\sim}{r})$ are conserved upon variations of the configuration $\underset{\sim}{X}$ around the bifurcation point $\underset{\sim}{X}_b$. In particular, all points of $\eta_M(\underset{\sim}{r}_c)$

are critical points of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ with respect to the τ variables of $Q(\underset{\sim}{r})$, and, along the directions spanned by these variables, the points of $\eta_M(\underset{\sim}{r}_c)$ are critical points of the same type as the original singularity $\underset{\sim}{r}_c$ of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$. Thus, the identification of each critical point of $f(\underset{\sim}{x};\underset{\sim}{\mu})$, by its rank and signature in M, completely specifies a corresponding critical point of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ in R^3.

ii) In the foregoing exposition, we have interpreted eq. (2.5) as describing an unstable attractor of the restriction of the vector field $\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ to some submanifold M of the behaviour space. It is easy to see that a singularity $\underset{\sim}{r}_c$ of $\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ can also be regarded as an unstable attractor of the field $-\nabla\rho(\underset{\sim}{r},\underset{\sim}{X}_b)$ over M. As shown in Section B of Chapter I, a rank three critical point of ρ is either a nuclear attractor or a pseudo-attractor of $\nabla\rho$. Otherwise, it is a local minimum in ρ, thus a three-dimensional attractor of $-\nabla\rho$. Thus, at a regular point, or in a conflict catastrophe configuration, any critical point can be regarded as an attractor of some corresponding vector field derived from $\nabla\rho$, and which can become unstable as the configuration of the system is made to change. As a consequence, in studying structural changes in a given molecular system, it is crucial to follow simultaneously the evolution of all critical points which accompanies a smooth variation in the configuration of the system.[41]

In the previous section, we have noted that the ground state charge distribution of the H_3^+ system in its equilibrium geometry $\underset{\sim}{X}_e$ exhibits a singularity which is of corank two (rank one). In a neighbourhood of the equilibrium geometry, all nuclear attractors are structurally stable and only (3,-1) critical points are involved in the formation of the singularity observed at $\underset{\sim}{X}_b = \underset{\sim}{X}_e$. Because of the symmetry of the system, these critical points always lie in the plane, M, of the three nuclei. Thus, Thom's theorem assesses the existence of a neighbourhood, $n_M \subset M$, of the singularity within which $\rho(\underset{\sim}{r},\underset{\sim}{X}_e)$ is diffeomorphic to the catastrophe germ, $f(\underset{\sim}{x};\underset{\sim}{0})$, of an unfolding, $f(\underset{\sim}{x};\underset{\sim}{\mu})$, which describes an umbilic catastrophe.[52] Indeed, the phase portrait associated with this singularity, Fig. 2.3a, is characteristic of the elliptic umbilic, whose catastrophe germ is[40,41,46,51]

$$f(\underset{\sim}{x};\underset{\sim}{0}) = x_1^2x_2 - \frac{1}{3}x_2^3, \tag{2.6}$$

where x_1,x_2 are the components of $\underset{\sim}{x} \in n_M$, with respect to a coordinates system centred at the singularity. The universal unfolding of the elliptic umbilic is given in eq. (2.7)

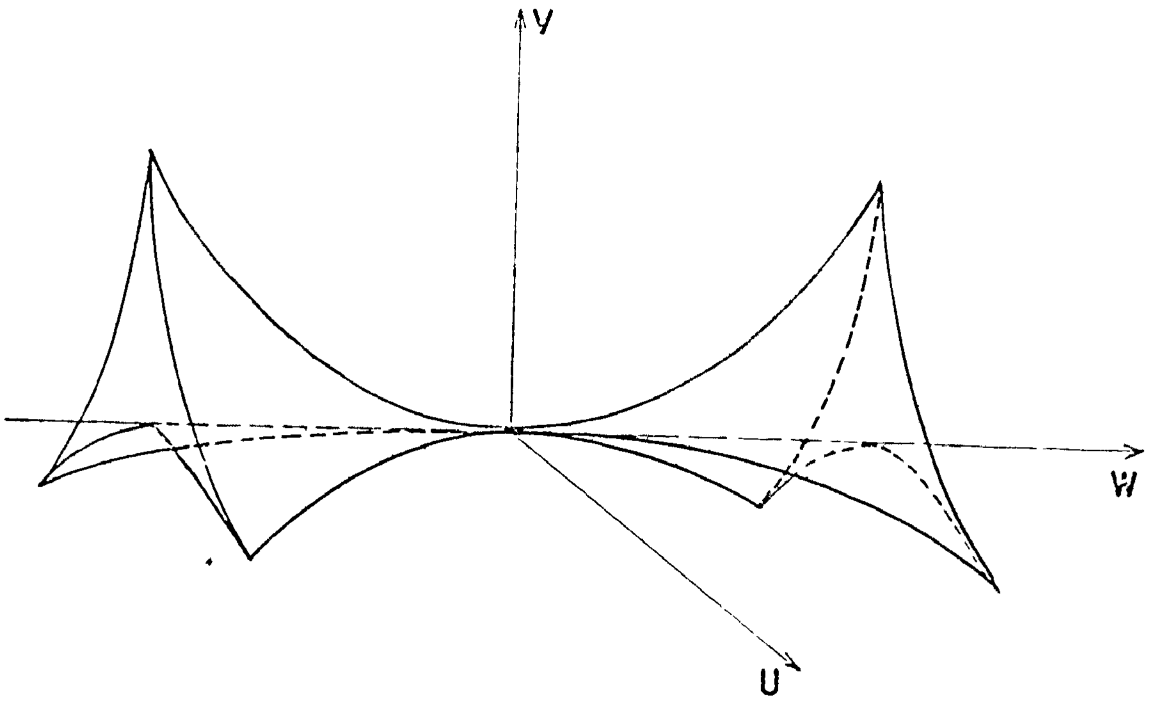
$$f(\underset{\sim}{x};\underset{\sim}{\mu}) = x_1^2x_2 - \frac{1}{3}x_2^3 + wx_1^2 + ux_1 + vx_2 \tag{2.7}$$

It depends on three control parameters $\underset{\sim}{\mu} = (w,u,v)$, and thus the singularity in $\rho(\underset{\sim}{r},\underset{\sim}{X}_e)$ is of codimension three. The control parameters $\underset{\sim}{\mu}$ might be defined only over a subset of R^3, but for the sake of simplicity we consider them to span all of the space R^3. The bifurcation set of eq. (2.7) is then the set of all $\underset{\sim}{\mu} \in R^3$, such that[18,19,53]

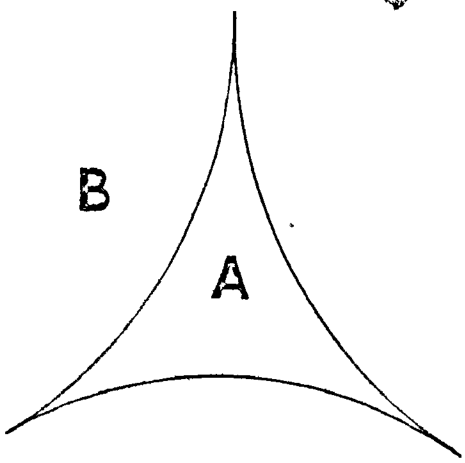
$$\frac{\partial f(\underset{\sim}{x};\underset{\sim}{\mu})}{\partial x_i} = 0 \qquad \text{and} \qquad \frac{\partial^2 f(\underset{\sim}{x};\underset{\sim}{\mu})}{\partial x_i \partial x_j} = 0 \qquad i,j \in \{1,2\}$$

This set is sketched in Fig. 2.6a.[18,19,40,51,53] It comprises two tapered cones meeting each other at the origin $\underset{\sim}{\mu} = \underset{\sim}{0}$. Its cross-section through a plane $w = w_0 \neq 0$ is a hypocycloid, shown in Fig. 2.6b. The hypocycloid partitions the plane $w = w_0$ into two open regions, denoted A and B, respectively. For $\underset{\sim}{\mu} \in A$, that is inside the hypocycloid, $f(\underset{\sim}{x};\underset{\sim}{\mu})$ always exhibits three saddle points - which are (2,0) critical points with respect to the topology of the plane - and a local minimum or a local maximum depending on whether w_0 is positive or negative.[18,53] Since the charge distribution always decreases along any line perpendicular to M, a local maximum of f in M would correspond to a local maximum of ρ in R^3, that is an attractor of the vector field $\nabla\rho$. Recalling the remark made in section B of Chapter I that the only (three dimensional) attractors of $\nabla\rho$ are the nuclear positions, we can safely discard the cases $w_0 < 0$, since they give rise to an additional (unobserved) attractor which would become unstable when the system is in the bifurcation configuration $\underset{\sim}{X}_b$. On the other hand, for $w_0 > 0$, $f(\underset{\sim}{x};\underset{\sim}{\mu})$ exhibits a local minimum when $\underset{\sim}{\mu} \in A$, which corresponds to a (3,+1) critical point of some locally diffeomorphic distribution $\rho(\underset{\sim}{r},\underset{\sim}{X})$. As discussed in section B of Chapter I, a (3,+1) critical point indicates the presence of a bonded ring. However, we have seen that a ring structure

Figure 2.6. The *bifurcation set* predicted by the *elliptic umbilic unfolding* $f(\underset{\sim}{x},\underset{\sim}{\mu})$, eq. (2.7). (a) In the space R^3 of the three control parameters $\underset{\sim}{\mu} = (u,v,w)$, this bifurcation set comprises two tapered cones which are joined at the origin of the control space. (b) The cross-section of the bifurcation set through the plane $w = w_o$ is a triangular hypocycloid, which partitions this plane into two regions. For a point $\underset{\sim}{\mu}$ contained in region A, the function $f(\underset{\sim}{x},\underset{\sim}{\mu})$ exhibits four critical points: three saddle points and a minimum if $w_o > 0$, or a maximum if $w_o < 0$. For a point $\underset{\sim}{\mu}$ contained in region B, the function $f(\underset{\sim}{x},\underset{\sim}{\mu})$ possesses only two saddle points.



a



b

is not observed in the neighbourhood of the equilibrium geometry of the ground state H_3^+ system. Therefore, $w = 0$, and the morphogenesis in the neighbourhood of $\underset{\sim}{X}_e$ would be described by[18,19]

$$f(\underset{\sim}{x};\underset{\sim}{\mu}') = x_1^2 x_2 - \frac{1}{3}x_2^3 + ux_1 + vx_2 \qquad (2.8)$$

where the control parameters $\underset{\sim}{\mu}' = (u,v)$ span a portion of a plane through $\underset{\sim}{X}_e$. For all $\underset{\sim}{\mu}' \neq \underset{\sim}{0}$, $f(\underset{\sim}{x};\underset{\sim}{\mu}')$ exhibits two saddle points, which correspond to two (3,-1) bond saddle points in the diffeomorphic field $\rho(\underset{\sim}{r};\underset{\sim}{X})$, thereby yielding a regular structure of the type (H-H-H), or a conflict catastrophe structure of the type [H-(H-H)]. The unfolding $f(\underset{\sim}{x};\underset{\sim}{\mu}')$ of eq. (2.8) is only able to predict the bifurcation point $\underset{\sim}{X}_e$ at the origin $\underset{\sim}{\mu}' = \underset{\sim}{0}$ of the (u,v) plane, and its associated bifurcation structure. For all other points of this plane, it can predict the number and type of critical points of signature > -3, but is unable to establish the associated molecular graphs. Hence, the complete identification of conflict catastrophe points, in general, requires a global analysis of structural stability in addition to the information obtained from the unfolding. This programme has already been carried out in the previous section. In particular, we have obtained the structure diagram of Fig. 2.5b by considering deformations of the system from its equilibrium geometry $\underset{\sim}{X}_e$, such that the sum of the internuclear distances remains constant. These deformations are appropriately described when the following identifications are made: i) a point $\underset{\sim}{\mu}' = (u = 0, v > 0)$ denotes the simultaneous displacement of the three protons such that $d_1 = d_3 > d_2$, and $d_1 + d_2 + d_3 = 3d$, using the notations of section A. Thus, the semi-axis associated with the positive values of the control parameter v coincides with the semi-axis δ_i' defined in Section A. ii) The axis of the control parameter u is perpendicular to the axis of the v's, and is contained in the plane P defined by the relation $d_1 + d_2 + d_3 = 3d$. Thus, in a configuration denoted by $\underset{\sim}{\mu}' = (u, v = 0)$, the internuclear distances d_1, d_2, d_3 satisfy $d_1 + d_3 = 2d$, $d_2 = d$. Denoting by Δd_i, $i = 1,2,3$, the variation in the internuclear distance d_i from its value at $\underset{\sim}{X}_e$, i.e., $\Delta d_i = d_i - d$, we recognize that a displacement of the system along the v-axis corresponds to $\Delta d_1 = \Delta d_3 = -\Delta d_2$, whereas along the u-axis, the protons are displaced in such a manner as to satisfy the relations $\Delta d_1 = -\Delta d_3$, $\Delta d_2 = 0$. We also make an identification of the behaviour variables (x_1, x_2) such that the x_2 axis of the behaviour plane coincides with the C_2 symmetry axis on which the proton H1 is found at the equilibrium geometry. With these identifications, the following observations result directly from the application of the unfolding $f(\underset{\sim}{x}; \underset{\sim}{\mu}')$ of eq. (2.8):

i) For a displacement denoted by $\underset{\sim}{\mu}' = (u = 0, v > 0)$, the two saddle points of $f(\underset{\sim}{x}; \underset{\sim}{\mu}')$ both lie on the x_2-axis, i.e., on the C_2 symmetry axis of the corresponding nuclear configuration.

ii) When $u = \sqrt{3}v$, and $v < 0$, i.e., for a point lying along the semi-axis δ_2' of Fig. 2.5b, we again obtain a nuclear configuration of C_{2v} symmetry, but with a different C_2 symmetry axis, namely the one associated with the proton H2 in $\underset{\sim}{X}_e$. The associated field $f(\underset{\sim}{x};\underset{\sim}{\mu}')$ then exhibits two saddle points, which are both found on this C_2 axis. The same observations hold for a displacement denoted by $\underset{\sim}{\mu}' = (u = -\sqrt{3}v, v < 0)$.

iii) For any other displacement $\underset{\sim}{\mu}'$ which preserves a twofold symmetry axis of the equilibrium configuration, the two saddle points of the associated $f(\underset{\sim}{x};\underset{\sim}{\mu}')$ are found to be located symmetrically with respect to this symmetry axis.

The above results are summarized in Table II.1. Because of the symmetry of the nuclear configurations, $\underset{\sim}{\mu}'$, obtained in i) and ii), the associated fields $f(\underset{\sim}{x},\underset{\sim}{\mu}')$ give rise to molecular graphs which characterize a conflict structure Hi-(Hj-Hk). Thus, together with the above identifications of the control parameters, (u,v), and the behaviour variables, (x_1,x_2), the unfolding of eq. (2.8) leads to the association of the semi-axes δ_1', δ_2', δ_3' with the conflict structures H1-(H2-H3), H2-(H1-H3) and H3-(H1-H2), respectively.

However, to ascertain that these semi-axes are the only loci of conflict catastrophes, we need to consider the charge distribution at all other nuclear configurations lying in the plane P and verify that these configurations are structurally stable. We then obtain the results summarized in Fig. 2.5b, which, in terms of the control parameters u, v, w,

TABLE II.1

Critical points of the elliptic umbilic unfolding, with w = 0

The particular expression of the elliptic umbilic unfolding,with $w = 0$, is given in eq. 2.8. For a given value of $\underset{\sim}{\mu}' = (u,v)$, the critical points of $f(\underset{\sim}{x};\underset{\sim}{\mu}')$ are found by solving the pair of equation $(\partial f(x;\underset{\sim}{\mu}')/\partial x_i) = 0$, $i = 1,2$. Listed below are the general solutions to these equations, their identifications as critical points of f, in the behaviour plane R^2, and as critical points of ρ, in R^3.

Values of Control Parameters	Number	Location	Type of Critical Point in R^2	Type of Critical Point in R^3
$\begin{cases} u = 0 \\ v > 0 \end{cases}$	2	$\begin{cases} x_1 = 0, \\ x_2 = \pm \sqrt{v} \end{cases}$	(2,0)	(3,-1)
$\begin{cases} u = 0 \\ v < 0 \end{cases}$	2	$\begin{cases} x_1 = \pm \sqrt{-v} \\ x_2 = 0 \end{cases}$	(2,0)	(3,-1)
$u \neq 0$ v arbitrary	2	$\begin{cases} x_1 = \pm \frac{1}{\sqrt{2}}[v - (v^2+u^2)^{\frac{1}{2}}]^{\frac{1}{2}} \\ x_2 = - \frac{u}{2x_1} \end{cases}$	(2,0)	(3,-1)
$\begin{cases} u = 0 \\ v = 0 \end{cases}$	1	$\begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$	singularity of corank two .	

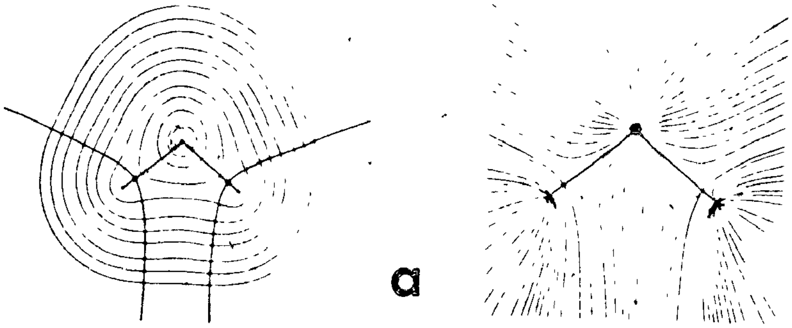
state: the semi-axis ($v > 0$), and its image through a rotation bv $\pm \frac{2\pi}{3}$ about the w-axis, denote the loci of conflict catastrophe structures, and partition the (u,v) plane into three structural regions, each associated with a regular structure of the type Hi-Hj-Hk. Because of the indistinguishability of the protons, these structures, in fact, represent a unique regular structure, denoted H-H-H, which is then seen to be generic.

Above, we are led to consider the universal unfolding of the elliptic umbilic with the parameter $w = 0$ on the basis of two observations. In particular, we have used the observation that nowhere near the equilibrium geometry, does the H_3^+ system exhibit a ring structure, to discard the possibility of having $w > 0$. We now consider another three-attractor system which, however, exhibits a ring structure over a region of its nuclear control space, and illustrate the utility of the unfolding of the elliptic umbilic, eq. (2.7), in this case. Figs. 2.7 show the gradient paths maps and molecular graphs (superimposed on the contours of ρ) exhibited by the charge distribution of the H_2O system in various configurations.[18,45] The configurations, of C_{2v} symmetry, illustrated are located along the minimum energy path of the ground state potential energy surface, which is the reaction path for the separation of $H_2O(X\ ^1A_1)$ into $O(^1D)$ and $H_2(^1\Sigma_g^+)$. Henceforth, we shall call this path abbreviately as the C_{2v} path. Fig. 2.7a cor-

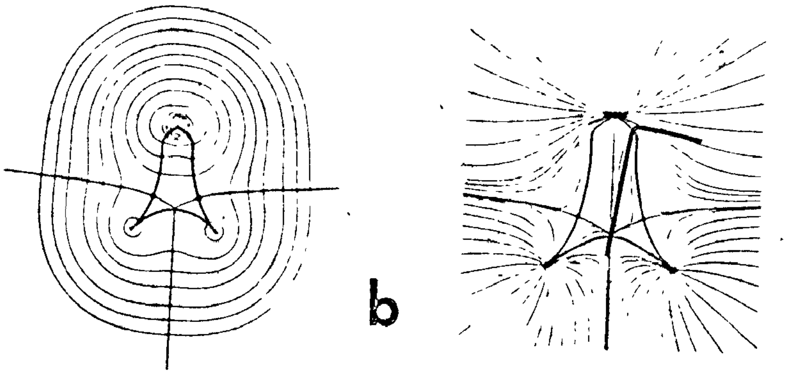
Figure 2.7. Charge distributions of ground-state water, at four nuclear configurations along the minimum energy C_{2v} path for the separation of $H_2O(X\ ^1A_1)$, into $O(^1D)$ and $H_2(^1\Sigma_g^+)$. In each configuration, the charge distribution is displayed in the plane of the nuclei, via its contours and the trajectories of its gradient vector. Superimposed on the contours are also shown the molecular graph and the interatomic surfaces. The geometric parameters quoted in the following are appropriate for configurations of C_{2v} symmetry and are defined as follows

$$R \equiv O\text{-}(H_2) \text{ separation}, \qquad R_H = H\text{-}H \text{ separation}$$

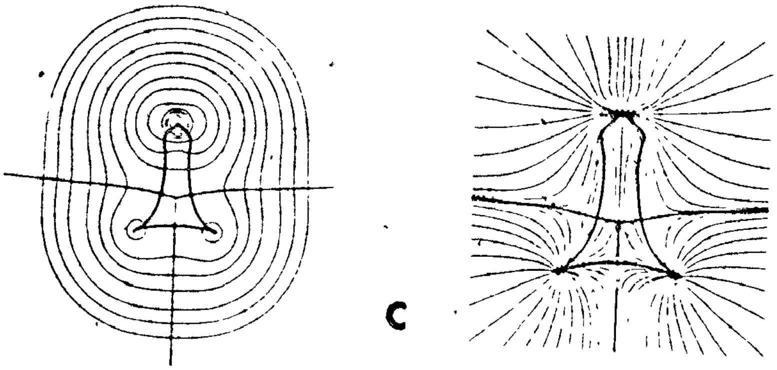
Diagrams (a) correspond to the equilibrium geometry ($R = 1.1089$ a.u., $R_H = 2.8677$ a.u.). Diagrams (b) correspond to the bifurcation configuration ($R = 2.39$ a.u., $R_H = 2.10$ a.u.), where a singularity in ρ appears between the two protons. Diagrams (c) correspond to the configuration ($R = 2.50$ a.u., $R_H = 2.00$ a.u.), where a ring structure is observed. At later stage, this ring structure collapses to give the conflict catastrophe structure $O\text{-}(H_2)$. This is illustrated in diagrams (d), for the configuration ($R = 2.75$ a.u., $R_H = 1.80$ a.u.).



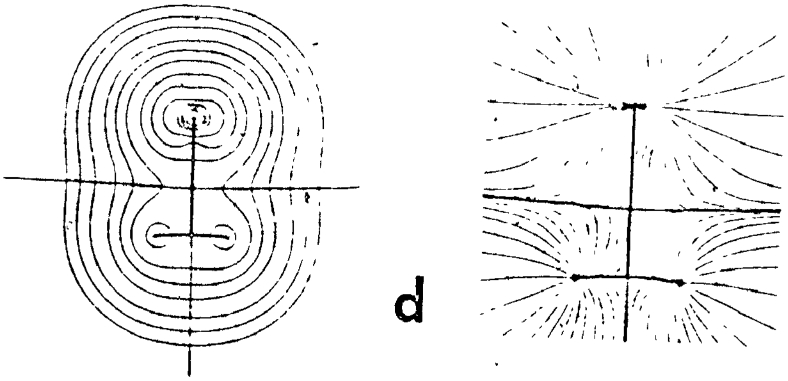
a



b

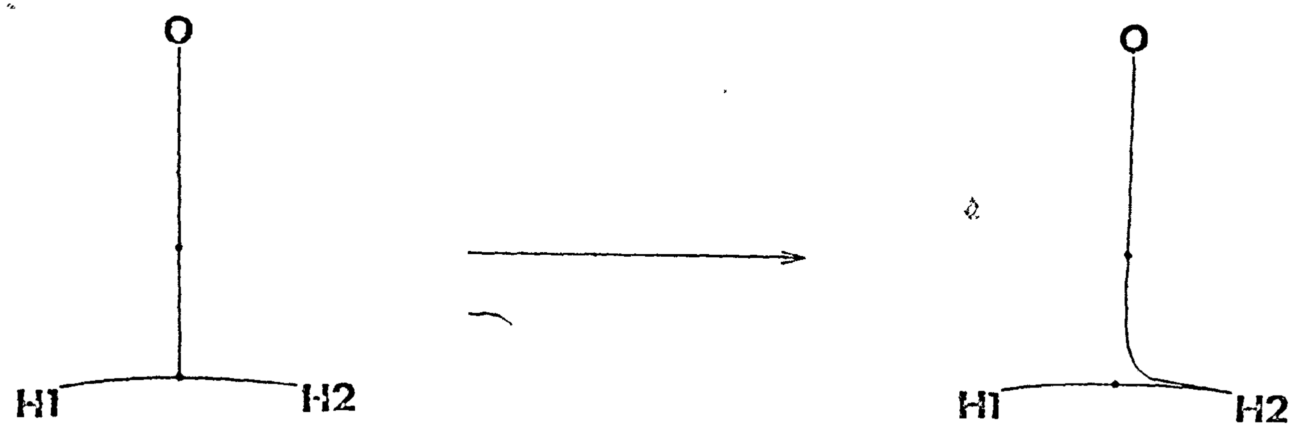


c



d

Figure 2.8. The instability of the conflict catastrophe structure $O-(H_2)$. A switching of attractors is observed when the Oxygen atom is subjected to a lateral displacement of 0.05 a.u. from its original position, shown in Fig. 2.7d.


O
H1
H2
O
H1
H2

responds to the equilibrium geometry of the water molecule. As we follow the C_{2v} path in the dissociative direction, the system traverses a bifurcation point, which is illustrated in Fig. 2.7b, to assume a ring structure, Fig. 2.7c, over a bounded interval of the path. Asymptotically in this direction, the structure of the system is as shown in Fig. 2.7d. This final structure, denoted $O-(H_2)$, in which the outgoing oxygen atom is bonded to the H_2 molecule via its (3,-1) bond saddle point, is reminiscent of the structure $[H-(H_2)]$ observed for the H_3^+ system, Fig. 2.3a. It is also a conflict catastrophe structure, since the vector field illustrated in Fig. 2.7d does not possess any singularity and the associated configuration is structurally unstable: a slight displacement of the oxygen nucleus, from the C_2 symmetry axis of this configuration, gives rise to a structure of the type O-H-H, Fig. 2.8. Separating the ring structure of Fig. 2.7c, and this conflict catastrophe structure, there must be a bifurcation structure: as will be shown later, to form a ring structure from the structure of Fig. 2.7d, the (3,-1) critical points associated with the bond $O-(H_2)$ must become, at some stage, a singularity. However, we have not been able to isolate the corresponding bifurcation point.

Consider the bifurcation configuration illustrated in Fig. 2.7b. Along its C_2 symmetry axis, and in a neighbourhood of the singularity which lies on this axis, the charge distribution exhibits a behaviour illustrated by the profile (a) of

Fig. 2.9. On the other hand, the bifurcation point admits a neighbourhood along the minimum energy path in which the morphogenesis only involves structural changes along the C_2 symmetry axis. Thus, the singularity mentioned above is of corank one. Furthermore, the profile (a) of Fig. 2.9 is typical of the function

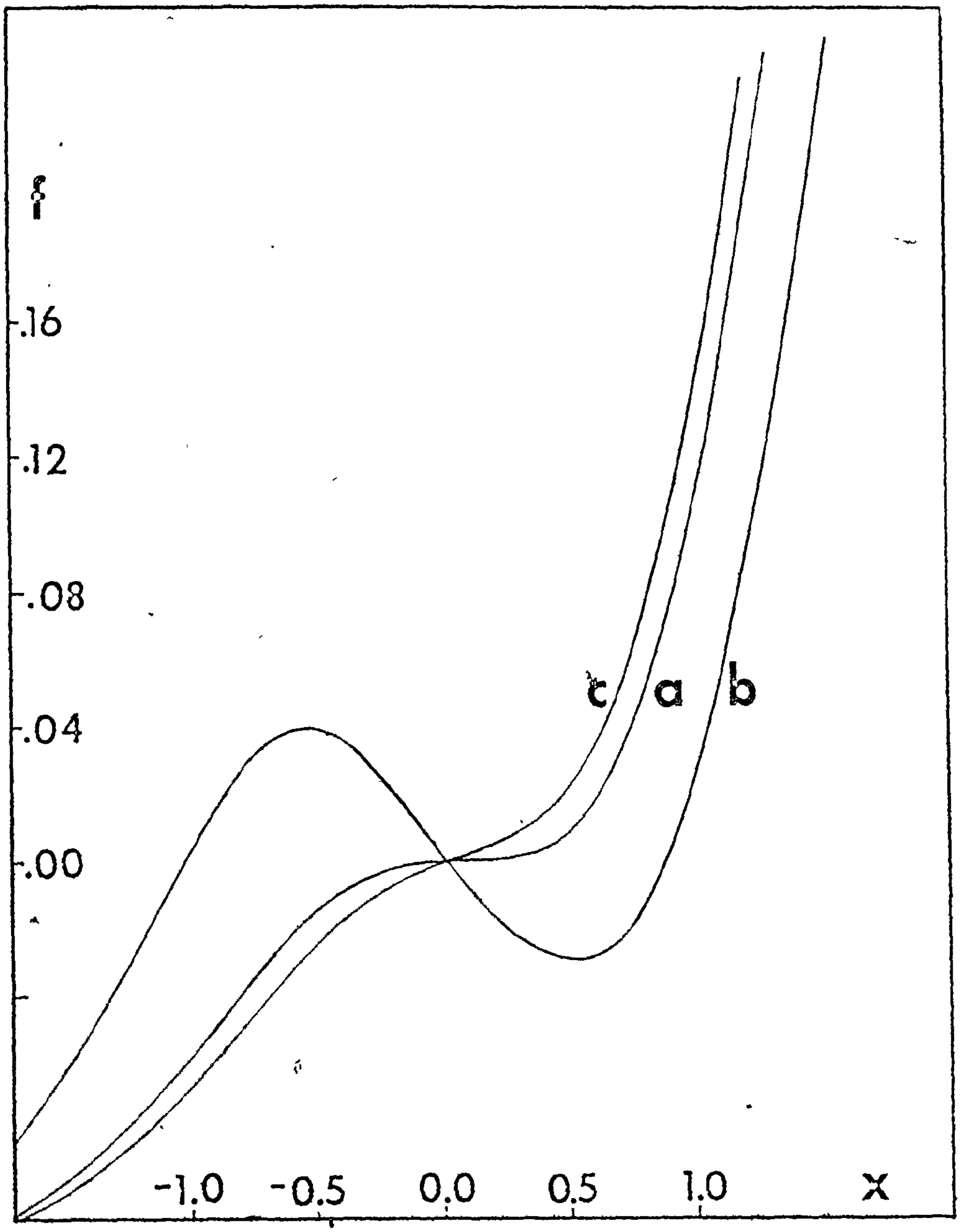
$$f(x;0) = \frac{1}{3}x^3 \qquad (2.9)$$

which describes a "fold catastrophe". Thus, along the C_2 symmetry axis, now identified as M, and in some neighbourhood of the singularity, the charge distribution associated with the bifurcation point of Fig. 2.7b is diffeomorphic to the germ, f(x;0), of the fold, eq. (2.9). The singularity is of codimension one, and admits the universal unfolding[40,41,46,51-53]

$$f(x;t) = \frac{1}{3}x^3 + tx, \quad t \in W \subset R \qquad (2.10)$$

In both equations (2.9) and (2.10), we readily identify the behaviour variable x as the coordinate of a point of the submanifold M, with the origin at the singularity of Fig. 2.7b. As t varies in eq. (2.10), two sets of curves are obtained. The curves associated with $t > 0$, generically exhibit two critical points, one local maximum and one local minimum, as shown in Fig. 2.9, curve (b). For all $t < 0$, f(x;t) possesses no critical point, and describes a profile similar to curve (c)

Figure 2.9. The fold catastrophe in H_2O. The three curves shown in this figure are the profiles of ρ along the C_2 symmetry axis of three C_{2v} configurations. Curve (a) corresponds to the bifurcation configuration of Fig. 2.7b (R = 2.39 a.u., R_H = 2.10 a.u.). The position of the singularity which is observed in this configuration defines the origin of the coordinate x along the C_2 symmetry axis. Curve (b) corresponds to the configuration of Fig. 2.7d (R = 2.75 a.u., R_H = 1.80 a.u.). The same type of profile is observed for any configuration lying in the ring structural region (e.g., that of Fig. 2.7c). Curve (c) corresponds to the configuration (R = 2.30 a.u., R_H = 2.16 a.u.) belonging to the same structural region as the equilibrium geometry. The three profiles characterize the unfolding of the fold, eq. (2.10).


f
.16
.12
.08
.04
.00
c
a
b
-1.0
-0.5
0.0
0.5
1.0
x

of Fig. 2.9. For $t < 0$, the singularity has thus disappeared. This situation can be assigned to the configurations along the minimum energy path in which the oxygen atom has been brought closer to form the open structure of water in its equilibrium geometry (Fig. 2.7a). Referring to Fig. 2.7b, we note that in the plane containing the two directions orthogonal to the C_2 symmetry axis, locally, the singularity behaves as a (2,0) saddle point. Thus, when $t > 0$, the local maximum predicted by eq. (2.10) corresponds to a (3,-1) critical point of some $\rho(\underset{\sim}{r},\underset{\sim}{X})$ which is locally diffeomorphic to $f(x;t)$ along the axis M. Likewise, the local minimum of $f(x;t)$, for $t > 0$ corresponds to a (3,+1) critical point of $\rho(\underset{\sim}{r},\underset{\sim}{X})$. Thus, the function $f(x;t)$, with $t > 0$, topologically represents the restriction of $\rho(\underset{\sim}{r},\underset{\sim}{X})$ to M, where the configuration $\underset{\sim}{X}$ lies within the interval of the C_{2v} path over which a ring structure is observed. One such configuration has already been illustrated in Fig. 2.7c. We thus see that eq. (2.10) correctly accounts for a part of the morphogenesis depicted in Figs. 2.7, if we identify the parameter t with some measure of the displacement of the system, along the C_{2v} path, from the bifurcation configuration of Fig. 2.7b. However, eq. (2.10) fails in predicting a bifurcation point at which the ring structure collapses to give the asymptotic structure of Fig. 2.7d. The charge distribution associated with this bifurcation point exhibits a singularity which arises from the merging of the

(3,+1) ring critical point with the (3,-1) bond saddle points of the two OH bonds. Along the C_2 axis of this configuration, the singularity still corresponds to a local minimum in ρ, as the profile (b) of Fig. 2.9 is still observed along this symmetry axis. On the other hand, the singularity is a local maximum along an axis perpendicular to the molecular plane. These two facts indicate that the singularity is of corank one.

This singularity does not admit an unfolding, however, since the structural changes observed in the neighbourhood of the associated bifurcation configuration involve a set of critical points which essentially are not colinear; on the other hand, the unfolding of a singularity of corank one would predict a set of colinear critical points to be involved in these structural changes. This remark suggests the use of the unfolding of some singularity of corank two, which is capable of predicting at least the ring structure of Fig. 2.7c and the conflict catastrophe structure $O-(H_2)$.[18,19] We have noted above that the elliptic umbilic unfolding, eq. (2.7), predicts a ring structure when the parameter w assumes a positive value. Consider eq. (2.7) again, with a given $w > 0$. The bifurcation set associated with $f(\underset{\sim}{x};\underset{\sim}{\mu})$ has already been illustrated in Fig. 2.6b. In the region labelled A, $f(\underset{\sim}{x};\underset{\sim}{\mu})$ exhibits a local minimum which would correspond to a (3,+1) ring critical point of ρ if the behaviour plane spanned by the variables $\underset{\sim}{x} = (x_1,x_2)$ is identified with the molecular plane of a three-attractor system, denoted (ABC). In fact, insofar as a local minimum in $f(\underset{\sim}{x};\underset{\sim}{\mu})$ can be identified with a ring critical point, one can always choose the parameter $\underset{\sim}{\mu}$ so that $f(\underset{\sim}{x};\underset{\sim}{\mu})$ appropriately describes a process in which a ring is formed and/or disrupted. In particular, as will be discussed later, (ABC) can denote a three-attractor subsystem of a many attractor system, for example, a part of some hydrocarbon molecule which exhibits a ring structure. All the points on the hypocycloid shown in Fig. 2.5b are bifurcation points. However, they fall into two classes: we distinguish between the bifurcation points at the three corners and those lying along the edges of the hypocycloid. The distinction is best illustrated by the evolution of $f(\underset{\sim}{x};\underset{\sim}{\mu})$ along the v-axis, $(u = 0)$. (We refer the reader to Table II.2.) It can be shown that at the bifurcation point $\underset{\sim}{\mu}_1 = (u = 0,\ v = 0)$, $f(\underset{\sim}{x};\underset{\sim}{\mu}_1)$ possesses a singularity of corank one and codimension one. Thus, the morphogenesis in some neighbourhood of $\underset{\sim}{\mu}_1$ can be described by the unfolding of the fold, eq. (2.10). On the other hand, at the bifurcation point $\underset{\sim}{\mu}_2 = (u = 0,\ v = w^2)$, i.e., at a corner of the hypocycloid, $f(\underset{\sim}{x};\underset{\sim}{\mu}_2)$ exhibits a singularity of corank one but of infinite codimension; it therefore does not admit an unfolding. Thus, the bifurcation points $\underset{\sim}{\mu}_1$ and $\underset{\sim}{\mu}_2$ are of the types of those observed in the above example of the H_2O system. At both bifurcation points, the singularity of $f(\underset{\sim}{x};\underset{\sim}{\mu})$ lies on the axis of the x_2 variable. For $0 < v < w^2$

TABLE II.2

Critical points of the elliptic umbilic unfolding, with $w = w_o > 0$

For given values of u,v, the critical points of $f(\underset{\sim}{x};\underset{\sim}{\mu})$ are obtained as the solutions to the pair of equations $(\partial f/\partial x_i) = 0$, $i = 1,2$, where $\underset{\sim}{\mu} = (u,v,w_o)$, and $f(\underset{\sim}{x};\underset{\sim}{\mu})$ is given by eq. (2.7). Listed below are the general solutions to the equations mentioned above for $u = 0$, their identifications as critical points of f, in the behaviour plane R^2, and as critical points of ρ, in R^3.

Values of Control Parameter v	Number	Location	Type of Critical Point in R^2	Type of Critical Point in R^3
$0<v<w_o^2$	3	$x_1=0,\ x_2=\sqrt{v}$ and $x_1=\pm\sqrt{w_o^2-v}, x_2=-w_o$	(2,0)	(3,-1)
	1	$x_o=0, x_2=-\sqrt{v}$	(2,+2)	(3,+1)
$v > w_o^2$	2	$x_1=0, x_2=\pm\sqrt{v}$	(2,0)	(3,-1)
$v < 0$	2	$x_1=\pm\sqrt{w_o^2-v}$ $x_2=-w_o$	(2,0)	(3,-1)
$v=w_o^2$	1	$x_1=0, x_2=+\sqrt{v}=w_o$	(2,0)	(3,-1)
	1	$x_1=0, x_2=-w_o$	singularity of corank one and infinite codimension†	
$v = 0$	2	$x_1=\pm w_o, x_2=-w_o$	(2,0)	(3,-1)
	1	$x_1=0, x_2=0$	singularity of corank one, codimension one††	

TABLE II.2 (Cont.)

† The Hessian matrix of $f(x;u=0,v=w_o^2,w=w_o)$ at the point $\underset{\sim}{x}_C = (x_1 = 0, x_2 = -w_o)$ has only one non-zero element, namely, the element $(\partial^2 f/\partial x_2^2)\big|_{\underset{\sim}{x}_C} = -2w_o$. The singularity is, therefore, of corank one, with x_1 being the essential variable. It is easily verified that $(\partial^n f/\partial x_1^n)\big|_{\underset{\sim}{x}_C} = 0$, for all $n \geq 3$. Thus, if f is expanded in a Taylor series in the neighbourhood of $\underset{\sim}{x}_C$, the variable x_1 will not appear to any finite order. A theorem of the general Catastrophe Theory then establishes that $\underset{\sim}{x}_C$ is a singularity with infinite codimension. It does not possess an unfolding.

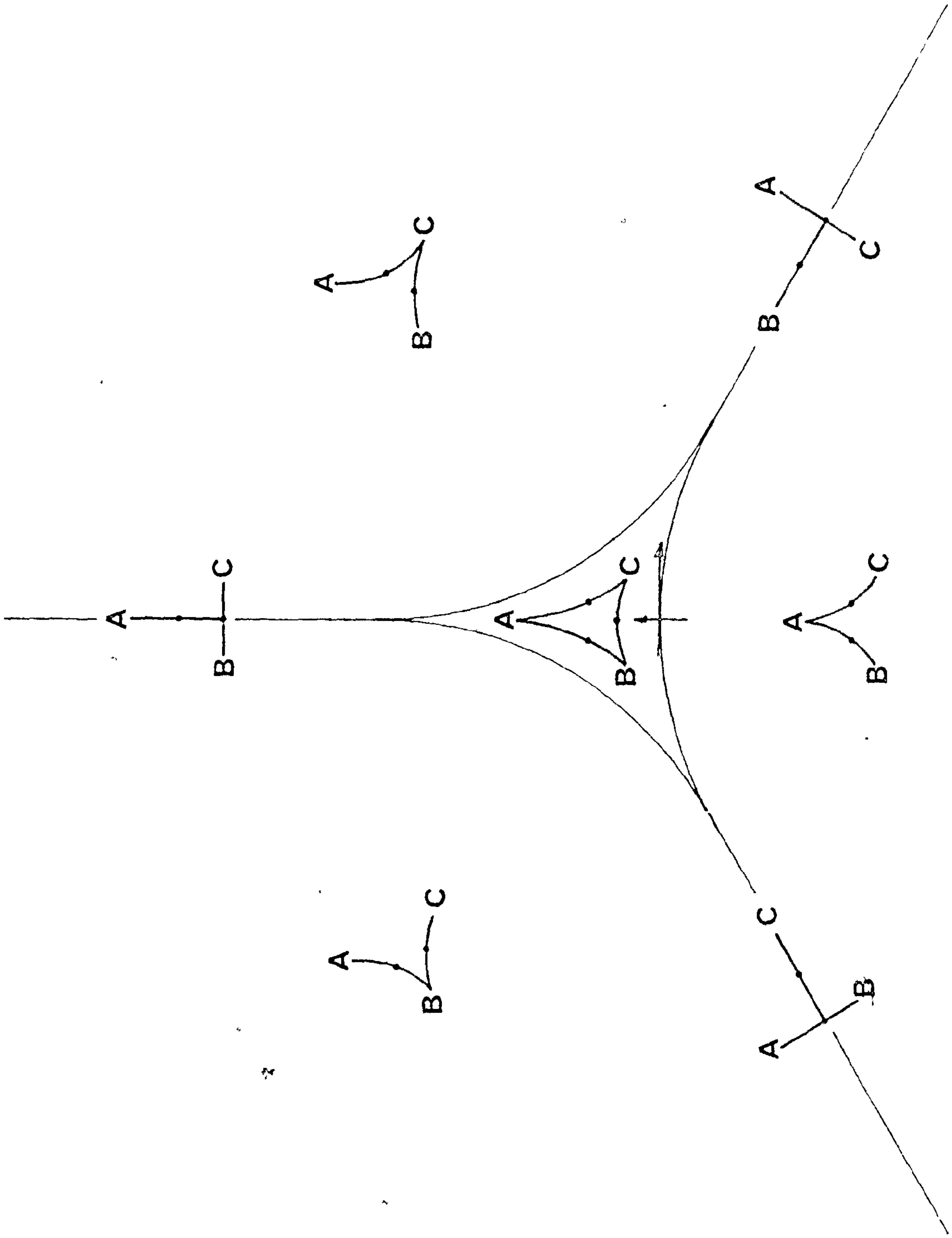
†† The only non-zero element of the Hessian matrix of $f(\underset{\sim}{x};u=0,v=0,w_o)$ at $x'_C = (0,0)$ is the element $(\partial^2 f/\partial x_1^2)\big|_{\underset{\sim}{x}'_C} = 2w_o$. The point $\underset{\sim}{x}'_C$ is thus a singularity of corank one, with essential variable x_2. Since $(\partial^3 f/\partial x_2^3)\big|_{\underset{\sim}{x}'_C} = -2$, a Taylor series of f at $\underset{\sim}{x}'_C$ will contain a third-power of x_2. A Theorem of Catastrophe Theory - the Reduction Lemma[51,52] - asserts that in some neighbourhood of $\underset{\sim}{x}'_C$, f is diffeomorphic to $g(x_1,x_2) = q(x_1) + \frac{1}{3}x_2^3$, where $q(x_1)$ is a quadratic expression in x_1. Thus, with respect to x_2, f is locally diffeomorphic to the catastrophe germ of the fold, eq. (2.9). The singularity $\underset{\sim}{x}'_C$ is, therefore, of codimension one.

and $u = 0$, $f(\underset{\sim}{x};\underset{\sim}{\mu})$ possesses a minimum on the x_2-axis, and which corresponds to a (3,+1) ring critical point. In addition, it also exhibits three saddle points, one of them being on the x_2-axis. These three critical points can be identified with the (3,-1) bond saddle points of the ring. When $v = w^2$, i.e., at $\underset{\sim}{\mu}_2$, the (3,+1) ring critical point merges with the two (3,-1) critical points which were not on the x_2-axis, to give the singularity of corank one and infinite codimension mentioned above. When $v > w^2$, this singularity becomes a saddle point in $f(\underset{\sim}{x};\underset{\sim}{\mu})$, i.e., a regular (3,-1) critical point of the associated ρ. Thus, the field $f(\underset{\sim}{x};\underset{\sim}{\mu})$ corresponding to $u = 0$ and $v > w^2$ exhibits two saddle points, both on the x_2-axis. For $u = 0$, $v < 0$, $f(\underset{\sim}{x};\underset{\sim}{\mu})$ also has two saddle points which, however, do not lie on the x_2-axis. Thus, the unfolding $f(\underset{\sim}{x};\underset{\sim}{\mu})$, given by eq. (2.7), with $w > 0$ and $u = 0$ properly describes the observations of Fig. 2.7 if we make the following identifications:[18,19] i) v measures the displacement of the system along the C_{2v} minimum energy path, with respect to the bifurcation configuration of Fig. 2.7b; ii) $\underset{\sim}{x} = (x_1,x_2)$ denotes the position of a point in the molecular plane with respect to an orthogonal coordinates system centred on the singularity of the above bifurcation configuration, such that: the axis of the x_2 variable coincides with the C_2 symmetry axis of this configuration. We finally identify the parameter u as describing a displacement of the system perpendicular to the minimum energy path. Such a displacement reduces the symmetry of the system from that of the C_{2v} point group to the symmetry of the C_s point group. Then, the unfolding of the elliptic umbilic, eq. (2.7), with $w > 0$, completely describes the morphogenesis in some neighbourhood of the bifurcation point illustrated by Fig. 2.7b. It appears that this neighbourhood extends over quite large a portion of the nuclear configuration space, since it encompasses an interval of the C_{2v} path containing the equilibrium geometry as well as the dissociation asymptote. Three noteworthy features are predicted by the model afforded by the elliptic umbilic unfolding. First, the ring structure exists not only over an interval of the dissociative path, it also can be obtained at configurations of C_s symmetry in some neighbourhood of this interval. Second, there must be altogether three conflict catastrophe structures and three regular open structures. To see this, we note that the bifurcation point $\underset{\sim}{\mu_2}$ separates the ring structural region A from the semi-axis $(u = 0,\ v > w^2)$ which represents the conflict catastrophe structure O-(H_2). Thus, $\underset{\sim}{\mu_2}$ belongs to the boundary of a region associated with a conflict structure. By symmetry, this result must also be true for the remaining corners of the hypocycloid. Through a rotation by $\pm \frac{2\pi}{3}$ about the three-fold symmetry axis of the hypocycloid, the semi-axis $(u = 0,\ v > w^2)$ is transformed into a line which represents another conflict catastrophe structure.

The three conflict structures are readily identified as corresponding to the bonding scheme A-(BC), B-(AC), C-(AB), respectively, where A = 0, B = H1, C = H2, and the hydrogen nuclei are considered distinguishable. The semi-axes associated with the three conflict structures found above now partition the region B of Fig. 2.6b into three open subsets. Each of these represents a regular open structure of the type A-B-C. Thus, the use of the elliptic umbilic unfolding leads to the structure diagram shown in Fig. 2.10. Finally, eq. (2.7) predicts the existence of a region of the nuclear configuration space in which the morphogenesis would be accounted for by the elliptic umbilic unfolding, but with w = 0, i.e., by eq. (2.8). In this region, there would exist a configuration whose associated charge distribution exhibits a singularity of corank two and codimension three, as was found for the H_3^+ system in its equilibrium geometry. Finding this region, and its geometrical relationship with the one for which Fig. 2.10 was obtained, will establish the physical interpretation of the parameter w in eq. (2.7), which we have left unidentified in the above analysis. While this region has not been found yet, for the moment, we anticipate, on the basis of observations made on the H_2O and H_3^+ systems that w depends on the nuclear configuration, the nuclear charges, the number of electrons of the system, as well as on its quantum state.[18,19]

To close this section, we now use the structure dia-

Figure 2.10. Structure diagram for a three-attractor (ABC) system which exhibits a ring structure, over a bounded region of the nuclear configuration space. The origin of the u,v coordinate system is indicated by a pair of crossed arrows. (The axis of the u's is parallel to the horizontal arrow, the axis of the v's, to the vertical arrow.)


A
B
C
A
B
C
A
B
C
A
B
C
A
B
C
A
B
C
A
B
C

gram of Fig. 2.10 to discuss the mechanism of a rearrangement reaction.[19] We have noted above that the model afforded by the elliptic umbilic unfolding is applicable to any case where a ring involving three attractors is formed or destroyed. The three attractors and their associated basins can form a subsystem of a larger system. A reaction in which every part of this larger system remains unaltered except the part containing the three nuclei A, B and C which undergoes the rearrangement

$$A - B - C \rightarrow B - C - A,$$

is called a rearrangement reaction. A rearrangement reaction can thus be conducted through either of two mechanisms which are predicted by eq. (2.7). If the reaction path intersects the catastrophe set of Fig. 2.10 at a conflict catastrophe point, or when the ring structural region shrinks to a point ($w = 0$, the case of H_3^+), the system passes through an intermediate whose structure is A-(BC). In the previous section, we have illustrated this mechanism in the case of the reaction I

$$H - C - N \rightarrow C - N - H \qquad \text{(I)}$$

In particular, we have noted that the intermediate conflict structure H-(CN) was found in the neighbourhood of an energetic instability, i.e., the configuration corresponding to the top of a potential barrier. On the other hand, if the reaction path crosses a ring structural region, the system will traverse two bifurcation points, each of them being located on an edge of the hypocycloid of Fig. 2.10. Thus, the reaction will follow the mechanism illustrated in Fig. 2.11. At the first bifurcation point, a ring structure begins to emerge from the open structure A-B-C, through the establishment of a singularity between the terminal atoms A and C. In the ring structural region, the singularity splits into a (3,-1) critical point associated with the newly formed A-C bond, and a (3,+1) ring critical point. As the system approaches the second bifurcation point, the (3,+1) ring critical point moves toward the (3,-1) critical point of the A-B bond, with which it merges to give a singularity at this second bifurcation point. The reaction proceeds to its completion as the system moves away from this bifurcation point, and assumes the open structure B-C-A of the product. This mechanism has been found to apply in the case of the carbonium-ion rearrangement reaction

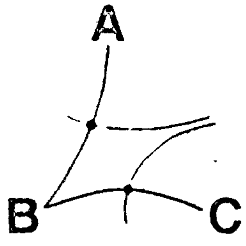
$$(CH_3\text{-}CH_2\text{-}C^*H_2)^+ \rightarrow (CH_2\text{-}C^*H_2\text{-}CH_3)^+,$$

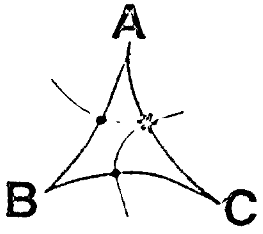
(where C* denotes a labelled carbon nucleus), which proceeds through an intermediate ring structure. This intermediate ring structure of the $C_3H_7^+$ system has already been illustrated in Fig. 1.5. In the previous section, we have noted that all the structures and structural changes which can be associated with a given molecular system are completely determined

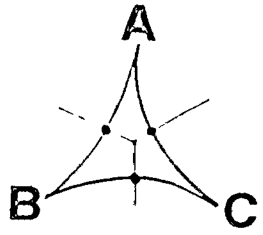
Figure 2.11. Sequence of structures encountered in a rearrangement reaction

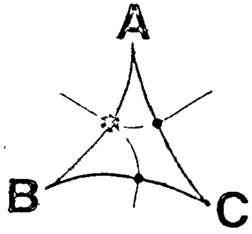
$$A\text{-}B\text{-}C \rightarrow B\text{-}C\text{-}A,$$

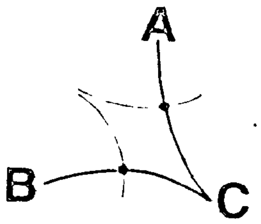
with a ring-structure intermediate. The position of a singularity in ρ is denoted by an asterisk. In this mechanism, the system passes through two bifurcation configurations corresponding to the formation and the destruction of the ring structure.


A
B
C


A
B
C


A
B
C


A
B
C


A
B
C

by the specification of a set of (typical) structure diagrams. The above discussions specifically illustrate the utilisation of the structure diagram in establishing the possible mechanisms of a rearrangement reaction. The partitioning of the nuclear configuration space by the catastrophe set, as indicated by a structure diagram, is reminiscent of a phase diagram in thermodynamics. In a classification of the local topological structure of phase diagrams, Coleman and O'Shea[41] have established that the phase boundary in a (P,T) thermodynamic control plane represents a set of conflict catastrophe points terminating in a bifurcation catastrophe at the critical point. Thus, a phase diagram depicts the various phases that a thermodynamic system can assume and the loci of phase transition in the associated thermodynamics control space. Correspondingly, a structure diagram defines the different structures that a molecular system may exhibit, and the structural changes that can be observed in some region of the molecular control space, the nuclear configuration space.

According to the definition of section A, molecular structure is not a property attached to a single point in the nuclear configuration space. Instead, a structure is associated with an open region in this space. Each structure thus has a probability of occurrence which is determined by the probability of finding the nuclear configuration **within** the associated structural region. In general, most of the probability distribution over the nuclear configuration space is concentrated in some neighbourhoods of the local minima of the potential energy surface which is described by the function $E_\ell(\underset{\sim}{X})$ of eq. (1.6a).[31] In particular, the equilibrium configuration, $\underset{\sim}{X}_e$, by definition delivers the absolute minimum in $E_\ell(\underset{\sim}{X})$ and corresponds to an absolute maximum in the above **probability** distribution. When $\underset{\sim}{X}_e$ is not a catastrophe point, we may consider the structure associated with $\underset{\sim}{X}_e$ as structurally representing the molecular system. By this procedure, we explicitly assign a unique definite structure to the molecular system. Alternatively, we may define <u>the</u> structure of the molecular system as the average over all the structures which are definable over the nuclear configuration space. Any of the two definitions suggested above leads to the conclusion that the assignment of a definite structure to an isolated molecular system is made possible by the interpretation of the wave function as a probability amplitude, in quantum mechanics. This conclusion is in sharp contrast with the one quoted by Woolley,[54] which states that the description of a molecular system with a definite structure is valid only in the classical limit. According to this author, the hypothesis of a definite molecular structure must form a part of the input to the theoretical study of a molecule, and cannot be considered a derived concept of quantum mechanics. The discrepancy between Woolley's statement and our conclusion, in fact, stems from a divergence of views and concerns: Molecular structure appears to be identified with molecular geometry in Woolley's work,[54] which aims at delineating the foundation of the model of a molecule which is provided by the Born-Oppenheimer approximation, and forms the basis of the theory of molecular spectra. On the other hand, we distinguish the structure of a molecule from its geometry, which is specified, for instance, by the internuclear distances.[19] We have already indicated an important difference between a molecular geometry and a molecular structure: the latter in general is generic, while the former never is. The importance of genericity in the mathematical definition of molecular structure, if this structure is to be an observable property of the molecular system, has been discussed in the introduction to this chapter. Conceptually, there is another fundamental difference between molecular geometry - Woolley's structure - and molecular structure as defined in this work: A given geometry specifies the relative position of the nuclei constituting the molecule, without defining the atoms - as subsystems of the molecule - and their bonding interactions. In contrast, a molecular structure is represented by a typical molecular graph, which unambiguously defines the atomic constituents of the molecular system, and depicts those interactions between the atoms which correspond to chemical bonds. As illustrated by the structure diagrams considered in this chapter, molecular structure and its change, corresponding to a change in bonding, are intimately related concepts, which are defined simultaneously in the application of the notions of structural stability to the topological study of molecular charge distribution.

CHAPTER III

QUANTUM MECHANICS OF A SUBSYSTEM

A. Quantum Subsystems

In Chapter I, we have defined an atom as the region of space traversed by the gradient paths of ρ which terminate at a nuclear attractor. Alternatively, the atom can be defined in terms of its boundary. Generally, this boundary comprises the union of a number of interatomic surfaces and some portions which are infinitely distant from the attractor. The interatomic surfaces as well as the surfaces found at infinity are the only closed surfaces S of R^3 which satisfy the equation

$$\nabla\rho\cdot\underset{\sim}{n} = 0, \quad \forall\, r \in S \qquad (3.1)$$

A surface which satisfies eq. (3.1) is called a zero-flux surface.[13-19] Thus, an atom is a region of space which contains a single nuclear attractor and which is bounded by a zero-flux surface. The union of two or more adjacent atoms is again a connected region bounded by a zero-flux surface. Such a union may represent a functional group and therefore is also of interest to a chemist. We thus consider generally a connected region $\Omega \subseteq R^3$ which is bounded by a zero-flux surface as representing a subsystem of the molecular system. In this chapter, the quantum mechanical properties of these subsystems will be derived and discussed. Before considering these properties, we first address ourselves to two questions, which by their very nature, can be given only somewhat speculative answers, at least at the present stage of development of the theory. The first question concerns the reality of the atoms and subsystem defined above. Referring to the considerations of Chapter II, we note that in (almost) all molecular systems studied, the nuclear attractors are structurally stable, even at catastrophe points, where the gradient vector fields of the associated charge distributions undergo global changes. In this sense, atoms are the indestructible units of matter. In contrast, a functional group or a molecule in a molecular environment is made up from the linkage of a number of atomic units, which can be disrupted at a catastrophe point. Such a subsystem thus exists only over an (open) domain of the nuclear configuration space, of the total system. When its internal structure, as represented by the network of bond-paths linking its constituent nuclei, is generic with respect to the nuclear configuration space, the domain of existence of the species extends over the entirety of this space. The existence of the subsystem (a functional group or a molecule in a molecular matrix) with a definite structure would then be obtained in spite of the interactions with the environment, rather than as a result of these interactions.[19]

The second question which we now consider concerns the nature of our subsystems: the partitioning of the total system by the zero-flux surfaces is a partitioning in real space, whereby we assign an open region of R^3 to a subsystem. A subsystem is thus not defined by an à priori partitioning of the Hamiltonian in the vector space of linear hermitian operators acting on the abstract Hilbert space of the total system. Such a partitioning of the Hamiltonian would violate the indistinguishability of the electrons at the outset, and moreover, is arbitrary. This arbitrariness can only be removed if the partitioning is decided on the basis of some physical considerations which would amount to building in the physics which is to be discovered. Thus, a subsystem in the present context is not defined as a collection of a given number of electrons to which one assigns a Hamiltonian, whose eigenfunctions would describe physical states of the subsystem. Instead, it is a portion of a charge distribution ρ, which contains an average number of electrons obtained by integrating ρ over the region of space it occupies. One may consider this average number of electrons as a statistical average of instantaneous electronic populations that the subsystem may assume. In this sense, the subsystem is open: it exchanges energy, momentum as well as matter with its environment. In general, such an open subsystem can be formally described only by a density matrix, and not by a state function.[55] The quantum theory of open systems is still in its early stages of development. An alternative to this formal approach is provided by the specification of the average values of important subsystem properties. We shall only be concerned with such a description, which uses the subsystem average properties. Notable among these are the subsystem energy and the average of the forces acting on the subsystem. The relationship between these two properties constitute the virial theorem. The average of the forces on the subsystem can also be related to the subsystem average of the linear momentum of an electron, through a force law which parallels Newton's equation of motion of classical mechanics. In the following, we shall obtain the definitions of these subsystem properties and derive their interrelationships - the subsystem virial theorem and force law - in the context of a general variational characterization of the subsystem.

The laws of nature are often expressible as variational principles.[56] These principles, which state that the correct description of a system is achieved when some associated quantity attains a minimum or a maximum value or is simply stationary, are both conceptually and practically appealing to physicists. They have the merit of being simple and general, and often constitute the starting points for formal developments or practical calculations. Fermat's principle in optic is an example of such a variational principle. In classical mechanics, one can cite two well-known variational principles which are d'Alembert's principle and Hamilton's principle.[56,57] Two variational principles exist in Quantum mechanics which will be reviewed in the following, and generalized so as to properly describe a subsystem as well as a total molecular system.[58]

B. Stationary-State Variational Principle

The first of these quantum mechanical variational principles identifies the bound energy levels of a molecular system, described by a Hamiltonian $\hat{H}$ as in eq. (1.1), as the stationary points of the functional

$$E[\varphi] = \frac{\int \varphi^*(\underset{\sim}{\tau})\hat{H}\varphi(\underset{\sim}{\tau})d\underset{\sim}{\tau}}{\int \varphi^*(\underset{\sim}{\tau})\varphi(\underset{\sim}{\tau})d\underset{\sim}{\tau}} = \frac{\langle\varphi|\hat{H}|\varphi\rangle}{\langle\varphi|\varphi\rangle} \tag{3.2a}$$

or equivalently of the functional

$$G[\varphi] = \int d\underset{\sim}{\tau}\varphi^*(\underset{\sim}{\tau})\hat{H}\varphi(\underset{\sim}{\tau}) = \langle\varphi|\hat{H}|\varphi\rangle \tag{3.2b}$$

subject to the normalization condition

$$\int \varphi^*(\tau)\varphi(\tau)d\tau = \text{constant} \tag{3.2c}$$

In both equations (3.2a) and (3.2b), $\underset{\sim}{\tau}$, $d\underset{\sim}{\tau}$ denote a position vector in the space of the coordinates and spins of all the particles, and a volume element in that space, respectively. The continuous, twice-differentiable trial function $\varphi(\underset{\sim}{\tau})$ is chosen to be square-integrable and is required to satisfy the boundary conditions

$$\lim_{|\underset{\sim}{r}_j|\to+\infty} \varphi^{(\ell)}(\underset{\sim}{\tau}) = 0, \quad \forall\ell = 0,1,2, \quad \forall j \in \{1,\ldots,N\} \text{ and}$$
$$\underset{\sim}{\tau} = \{\underset{\sim}{r}_1 s_1, \underset{\sim}{r}_2 s_2, \ldots, \underset{\sim}{r}_j s_j, \ldots, \underset{\sim}{r}_N s_N\} \tag{3.3}$$

where $\varphi^{(\ell)}$ denotes the ℓth derivatives of φ with respect to the components of $\underset{\sim}{r}_j$, (N is the total number of particles). To fix the ideas, and to simplify the following discussions, we shall be concerned specifically with the variational properties of a subsystem with respect to the electronic energy functional $E_e[\varphi]$ or $G_e[\varphi]$, constructed according to eqs. (3.2a and b) with $\hat{H}$ replaced by $\hat{H}_{e\ell}$, eq. (1.6b).[59] The trial function $\varphi(\underset{\sim}{\tau})$ then depends parametrically on the nuclear configuration. For simplicity, in the following it is considered a function of the electronic spatial and spin coordinates only, which are represented by the vector $\underset{\sim}{\tau}$. (Thus, N represents the total number of electrons in the molecule.) The stationary state variational principle then states that[58]

$$\frac{\delta E_e[\varphi]}{\delta\varphi(\underset{\sim}{\tau})} = 0 \qquad (3.4a)$$

determines the electronic stationary states of the system, where $[\delta E_e/\delta\varphi(\underset{\sim}{\tau})]$ denotes the functional derivative of $E_e[\varphi]$ with respect to $\varphi(\underset{\sim}{\tau})$. Equivalently, using $G_e[\varphi] = E_e[\varphi]\langle\varphi|\varphi\rangle$, eq. (3.4a) can be written as

$$\frac{\delta(G_e[\varphi] - \lambda\langle\varphi|\varphi\rangle)}{\delta\varphi(\tau)} = 0 \qquad (3.4b)$$

where the Lagrangian multiplier λ, associated with the normalization constraint, eq. (3.2c), is readily identified with the value of $E_e[\psi]$ for a given solution of eq. (3.4a), that is, an eigenvalue of $\hat{H}_{e\ell}$.

The well-established equivalence between the solutions of the time-independent electronic Schrödinger equation (1.6a) and those of the variational principle, eqs. (3.4), is most often used in the design of tractable practical methods of solving eq. (1.6a).[58] In this work, we adopt the complementary position, and regard eqs. (3.4) as the generating principle of the quantum mechanical laws which govern a time-independent isolated molecular system. Indeed, not only does the variational problem posed in eqs. (3.4) lead to an Euler equation which is precisely Schrödinger's equation (1.6a), it also yields directly the hypervirial theorem and all other related theorems, notably the virial theorem.[58] This property of eqs. (3.4) will be shown in the context of the following derivations, as we consider the variational properties of a subsystem. We construct an appropriate variational functional for a subsystem Ω by first defining the regional projection of $G_e[\psi]$, denoted by $G_e[\Omega,\psi]$:[60,61,62]

$$
\begin{aligned}
G_e[\Omega,\psi] &= \tfrac{1}{2}\{\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*(\underset{\sim}{r},\underset{\sim}{\tau}')\hat{H}_{e\ell}\psi(\underset{\sim}{r},\underset{\sim}{\tau}') + \text{c.c.}\} \\
&= \int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\tfrac{1}{2}[\psi^*(\underset{\sim}{r},\underset{\sim}{\tau}')\sum_i(-\frac{\hbar^2}{2m})\ \nabla_i^2\psi(\underset{\sim}{r},\underset{\sim}{\tau}') + \\
&\quad \sum_i(-\frac{\hbar^2}{2m})\ (\nabla_i^2\psi^*(\underset{\sim}{r},\underset{\sim}{\tau}'))\psi(\underset{\sim}{r},\underset{\sim}{\tau}')] + \psi^*(\underset{\sim}{r},\underset{\sim}{\tau}')\hat{v}(\underset{\sim}{r},\underset{\sim}{\tau}')\psi(\underset{\sim}{r},\underset{\sim}{\tau}')\}
\end{aligned}
\qquad (3.5)
$$

In eq. (3.5), the primed integral is taken over (N-1) electronic coordinates, and the spins of all electrons, denoted collectively by $\underset{\sim}{\tau}'$. This integral yields a function of the cartesian coordinates of one electron, which, when $\psi(\mathbf{r},\boldsymbol{\tau}')$ is required to be antisymmetric, can be any of the N electrons of the system. We also denote the sum of the electron-nuclear attraction, and electron-electron repulsion terms in $\hat{H}_{e\ell}$ by $\hat{v}(\mathbf{r},\boldsymbol{\tau}')$, i.e.,

$$\hat{v}(\mathbf{r},\boldsymbol{\tau}') = -\sum_{\alpha} \frac{Z_\alpha e^2}{|\mathbf{r}-\mathbf{R}_\alpha|} - \sum_{\substack{\alpha,i\\ \mathbf{r}_i\in\boldsymbol{\tau}'}} \frac{Z_\alpha e^2}{|\mathbf{r}_i-\mathbf{R}_\alpha|} + \sum_{\mathbf{r}_i\in\tau'} \frac{e^2}{|\mathbf{r}-\mathbf{r}_i|} + \sum_{\substack{i<j\\ \mathbf{r}_i,\mathbf{r}_j\in\tau'}} \frac{e^2}{|\mathbf{r}_i-\mathbf{r}_j|} \tag{3.6}$$

When $\Omega = R^3$, eq. (3.5) generates $G_e[\psi]$, which is seen to satisfy

$$\begin{aligned} G_e[\psi] = G_e[R^3,\psi] &= \int d\mathbf{r}\int' d\boldsymbol{\tau}'\psi^*[(-\frac{\hbar^2}{2m})\sum_i \nabla_i^2 + \hat{v}]\psi \\ &= \int d\mathbf{r}\int' d\boldsymbol{\tau}'\{\sum_i (\frac{\hbar^2}{2m})\ \nabla_i\psi^*\cdot\nabla_i\psi + \psi^*\hat{v}\psi\} \end{aligned} \tag{3.7}$$

where the boundary conditions at infinity, eq. (3.3), has been used in the application of the identity

$$\tfrac{1}{4}\nabla_i^2(\psi^*\psi) = \tfrac{1}{4}(\psi^*\nabla_i^2\psi + (\nabla_i^2\psi^*)\psi) + \tfrac{1}{2}\nabla_i\psi^*\cdot\nabla_i\psi \tag{3.8}$$

to rearrange the kinetic energy terms. On the other hand, the use of eq. (3.8) in eq. (3.5) for an arbitrary region $\Omega \subset R^3$ leads to a surface term, in addition to the regional projection of eq. (3.7). This surface term necessarily vanishes if Ω is bounded by a surface of zero-flux in $\nabla\rho_\psi(\mathbf{r})$, where ρ_ψ is the charge distribution associated with the trial function $\psi(\boldsymbol{\tau})$ since

$$G_e[\Omega,\psi] = \int_\Omega d\mathbf{r}\int' d\boldsymbol{\tau}'\{(\frac{\hbar^2}{2m})\ \sum_i\nabla_i\psi^*\cdot\nabla_i\psi + \psi^*\hat{v}\psi\} + (-\frac{\hbar^2}{4m})\oint ds\nabla\rho_\psi'\cdot\mathbf{n} \tag{3.9}$$

where $\rho'_{\varphi} = (\frac{1}{N})\rho_{\varphi}$. The zero-flux surface condition, eq. (3.1), thus ensures that $G_e[\Omega,\varphi]$ satisfies the same identity as $G_e[\varphi]$, eq. (3.7). We shall require $G_e[\Omega,\varphi]$ to possess the above property at any stage of the variation. That is, in some small neighbourhood of an eigenfunction $\psi(\underset{\sim}{\tau})$ of $\hat{H}_{e\ell}$ (a solution to eqs. (3.4)), the admissible trial function $\varphi(\underset{\sim}{\tau})$ must satisfy sufficient regularity conditions to ensure the existence of a surface of zero-flux in $\nabla\rho_{\varphi}(\underset{\sim}{r})$, bounding a region $\Omega(\varphi)$ which is continuously deformable into a given subsystem $\Omega = \Omega(\psi)$. The region $\Omega(\varphi)$ represents the subsystem in the varied total system which is described by $\varphi(\underset{\sim}{\tau})$, just as Ω represents the subsystem when the total system is in the stationary state $\psi(\underset{\sim}{\tau})$. Then, corresponding to $G_e[\varphi]$, we obtain the subsystem functional $G_e[\Omega,\varphi]$ which is given uniquely by

$$G_e[\Omega,\varphi] = \int_{\Omega(\varphi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{(-\frac{\hbar^2}{4m})\sum_i(\varphi^*\nabla_i^2\varphi + (\nabla_i^2\varphi^*)\varphi) + \varphi^*\hat{\nu}\varphi\}$$

$$= \int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{(\frac{\hbar^2}{2m})\sum_i\nabla_i\varphi^*\cdot\nabla_i\varphi + \varphi^*\hat{\nu}\varphi\} \tag{3.10}$$

We complete the construction of the variational functional for a subsystem $\Omega(\varphi)$ by adjoining to $G_e[\Omega,\varphi]$ the normalization condition

$$\int_{\Omega(\varphi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\varphi^*(\underset{\sim}{r},\underset{\sim}{\tau}')\varphi(\underset{\sim}{r},\underset{\sim}{\tau}') = \int_{\Omega(\varphi)} d\underset{\sim}{r}\rho'_{\varphi}(\underset{\sim}{r}) = \text{constant} \tag{3.11}$$

This can be achieved either by dividing $G_e[\Omega,\varphi]$ by the above normalization integral, to obtain $G'_e[\Omega,\varphi]$

$$G_e'[\Omega,\psi] = \frac{\int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{(\frac{\hbar^2}{2m})\sum_i \nabla_i\psi^*\cdot\nabla_i\psi + \psi^*\hat{v}\psi\}}{\int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*\psi}$$

$$= \tfrac{1}{2}\{\frac{\langle\psi^*,\hat{H}_{e\ell}\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} + \text{c.c.}\} \qquad (3.12a)$$

or by the use of a Lagrangian multiplier, λ, thus defining the variational functional $F_\lambda[\Omega,\psi]$

$$F_\lambda[\Omega,\psi] = \int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{(\frac{\hbar^2}{2m})\sum_i \nabla_i\psi^*\cdot\nabla_i\psi + (\hat{v}-\lambda)\psi^*\psi\}$$

$$= \tfrac{1}{2}\{\langle\psi^*,(\hat{H}_{e\ell}-\lambda)\psi\rangle_\Omega + \text{c.c.}\} \qquad (3.12b)$$

In eqs. (3.12), we have introduced the notation $\langle\psi^*,\hat{A}\psi\rangle_\Omega$ to denote the integral

$$\int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*(\underset{\sim}{r},\underset{\sim}{\tau}')\hat{A}(\underset{\sim}{r},\underset{\sim}{\tau}')\psi(\underset{\sim}{r},\underset{\sim}{\tau}').$$

We shall often use this notation in what follows to simplify the exposition. The value of the Lagrangian multiplier λ can again be identified as the value of $G_e'[\Omega,\psi]$ where ψ corresponds to a stationary state, through the requirement that the variations of $G_e'[\Omega,\psi]$ and $F_\lambda[\Omega,\psi]$ be formally equivalent. From eq. (3.12a), the variation of $G_e'[\Omega,\psi]$, in a small neighbourhood of a stationary state ψ, satisfies

$$\langle\psi^*,\psi\rangle_\Omega\delta G_e'[\Omega,\psi] = \delta(\tfrac{1}{2}\{\langle\psi^*,\hat{H}_{e\ell}\psi\rangle_\Omega + \text{c.c.}\}) - G_e'[\Omega,\psi]\delta\langle\psi^*,\psi\rangle_\Omega$$

$$= \delta(\tfrac{1}{2}\{\langle\psi^*,(\hat{H}_{e\ell} - G_e'[\Omega,\psi])\psi\rangle_\Omega + \text{c.c.}\}) \qquad (3.13)$$

The varied functional of the r.h.s. of eq. (3.13) is exactly $F_\lambda[\Omega,\varphi]$, if $\lambda = G_e'[\Omega,\psi] = E$, E being the energy of the system in the stationary state $\psi(\underset{\sim}{\tau})$. Thus, the identification $\lambda = E$ ensures that the results of the variations in $G_e'[\Omega,\varphi]$ and $F_\lambda[\Omega,\varphi]$ are identical, to a multiplicative constant factor equal to $\langle\psi^*,\psi\rangle_\Omega$. Having defined the variational functionals, we now consider their variations.[63] Using eq. (3.12b), with $\lambda = E$, to first order in $\delta\varphi = (\varphi-\psi)$, one has

$$\begin{aligned}
\delta F_E[\Omega,\varphi] = {} & \int_{\Omega(\varphi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{(\tfrac{\hbar^2}{2m})\sum_i \nabla_i\varphi^*\cdot\nabla_i\delta\varphi + (\hat{\nu}-E)\varphi^*\delta\varphi\} + \text{c.c.} \\
& + \oint dS\delta S\{\int' d\underset{\sim}{\tau}'[(\tfrac{\hbar^2}{2m})\sum_i \nabla_i\varphi^*\cdot\nabla_i\varphi + (\hat{\nu}-E)\varphi^*\varphi]\} \\
= {} & \int_{\Omega(\psi)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\hat{H}_{e\ell}\varphi^* - E\varphi^*\}\delta\varphi + (\tfrac{\hbar^2}{2m})\oint dS\int' d\underset{\sim}{\tau}'(\nabla\varphi^*\cdot\underset{\sim}{n})\delta\varphi \\
& + \text{c.c.} + \tfrac{1}{2}\oint dS\delta S\{\int' d\underset{\sim}{\tau}'\varphi^*(\hat{H}_{e\ell} - E)\varphi + \text{c.c.}\} \\
& + (\tfrac{\hbar^2}{2m})\oint dS\delta S\nabla^2\rho_\varphi'
\end{aligned} \tag{3.14}$$

The second equality in eq. (3.14) was obtained by using eq. (3.8) and the identity

$$\nabla_i\varphi^*\cdot\nabla_i\delta\varphi = \nabla_i\cdot(\nabla_i\varphi^*\delta\varphi) - (\nabla_i^2\varphi^*)\delta\varphi. \tag{3.14a}$$

At any point $\underset{\sim}{r}$ of the surface of $\Omega(\varphi)$, $\delta S(\underset{\sim}{r})$ denotes the normal shift of a surface element centred on $\underset{\sim}{r}$, which is conditioned by the continuous variation in the admissible trial function

$\varphi(\underset{\sim}{\tau})$. We note that the last term in the r.h.s. of eq. (3.14) is the contribution to the total variation of the integral

$$I[\Omega,\varphi] = (\frac{\hbar^2}{4m}) \int_{\Omega(\varphi)} \nabla^2 \rho_{\varphi}' d\underset{\sim}{r} \tag{3.15}$$

from the above surface variation. Explicitly, the total variation in $I[\Omega,\varphi]$ is

$$\begin{aligned} \delta I[\Omega,\varphi] &= (\frac{\hbar^2}{4m}) \{ \oint dS \delta(\nabla\rho_{\varphi}' \cdot \underset{\sim}{n}) + \oint dS \delta S \nabla^2 \rho_{\varphi}' \} \\ &= (\frac{\hbar^2}{4m}) \{ \oint dS \int' d\underset{\sim}{\tau}' (\nabla\delta\varphi^*\varphi + \delta\varphi^*\nabla\varphi + \nabla\varphi^*\delta\varphi + \varphi^*\nabla\delta\varphi) \cdot \underset{\sim}{n} \\ &+ \oint dS \delta S \nabla^2 \rho_{\varphi}' \} \end{aligned} \tag{3.16}$$

Using eq. (3.16), we can rearrange eq. (3.14) to give

$$\begin{aligned} \delta F_E[\Omega,\varphi] = & \int_{\Omega(\varphi)} d\underset{\sim}{r} \int' d\underset{\sim}{\tau}' \{ \hat{H}_{e\ell}\varphi^* - E\varphi^* \} \delta\varphi + (\frac{\hbar^2}{4m}) \oint dS \int' d\tau' \{ \delta\varphi\nabla\varphi^* - \varphi^*\nabla\delta\varphi \} \cdot \underset{\sim}{n} \\ & + \text{c.c.} + \tfrac{1}{2} \oint dS \delta S \{ \int' d\underset{\sim}{\tau}' \varphi^* (\hat{H}_{e\ell} - E)\varphi + \text{c.c.} \} + \delta I[\Omega,\varphi] \end{aligned} \tag{3.17}$$

Because of the definition of $\Omega(\varphi)$, $I[\Omega,\varphi]$ identically vanishes at all stages of the variation, and so does $\delta I[\Omega,\varphi]$.[60,62] In the limit $\delta\varphi \to 0$, i.e. $\varphi \to \psi$, the first and third terms on the r.h.s. of eq. (3.17) also vanish, leaving the result

$$\delta F_E[\Omega,\psi] = (\frac{\hbar^2}{4m}) \oint dS \int' d\underset{\sim}{\tau}' \{ \delta\psi\nabla\psi^* - \psi^*\nabla\delta\psi \} \cdot \underset{\sim}{n} + \text{c.c.} \tag{3.18a}$$

where we have replaced $\delta\varphi$ by $\delta\psi$, as $\varphi \to \psi$. By virtue of eq. (3.13), the variation in $G_e'[\Omega,\varphi]$, from $G_e'[\Omega,\psi]$, to first order, is simply,

$$\delta G_e'[\Omega,\psi] = \frac{(\hbar^2/4m)\oint dS\int' d\underset{\sim}{\tau}'\{\delta\psi\nabla\psi^* - \psi^*\nabla\delta\psi\}\cdot\underset{\sim}{n}}{N\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*\psi} + \text{c.c.} \qquad (3.18b)$$

Equations (3.18) are the particular statements of the stationary state variational principle, which variationally characterize a subsystem, as defined by its boundary, a surface of zero-flux in $\nabla\rho$.[19,60,62] They are the consequences of the all-space principles, eqs. (3.4), since the functionals involved, $G_e'[\Omega,\psi]$ and $F_E[\Omega,\psi]$, are the subsystem projections of $E_e[\psi]$ and $(G_e[\psi] - E\langle\psi|\psi\rangle)$, respectively, and the extremal properties for the latters have been used, via the Schrödinger equation (1.6a). On the other hand, requiring eqs. (3.18) to be satisfied by any region Ω, which is bounded by a zero-flux surface, one obtains eqs. (3.4) since $\Omega = R^3$ is such a region, which in addition to the zero-flux boundary condition, is characterized by the boundary conditions of eq. (3.3b). The above remarks indicate that eqs. (3.18) are equivalent to eqs. (3.4). We regard the subsystem variational principle, eqs. (3.18)as the generalization of the all-space (or total system) principle, eqs. (3.4), in the sense that the former allows for formal discussions which refer both to a subsystem and/or to the total system. This variational principle for a subsystem no longer requires that some quantity, specifically $G_e'[\Omega,\psi]$ or $F_E[\Omega,\psi]$, be stationary, in the classical sense of the vanishing of its first variation. Instead, it demands that this first variation only has contributions from the bounding surface of the subsystem. Following Schwinger,[64] and anticipating the forthcoming discussions of time-dependent subsystems which will be based on the principle of stationary action,[62,64,65] we may re-interpret the property of a subsystem functional of being stationary as demanding the absence of interior contribution in its first variation. In this sense, eqs.(3.18) still correspond to $F_E[\Omega,\psi]$ or $G_e'[\Omega,\psi]$ being stationary with respect to the admissible variations $\delta\psi$ defined above. More precisely, the variations in these energy functionals are determined by the flux in the current, through the bounding surface of the subsystem, as generated by the variation $\delta\psi$.[62] The above physical interpretation is obtained by rewriting eqs.(3.18) as

$$\delta F_E[\Omega,\psi] = -\left(\frac{\hbar^2}{2N}\right)\oint dS\,\delta_\psi\{\mathbf{J}\}\cdot\mathbf{n} + \text{c.c.}, \tag{3.19a}$$

and

$$\delta G_e'[\Omega,\psi] = -\left(\frac{\hbar^2}{2N}\right)\langle\psi^*,\psi\rangle_\Omega^{-1}\oint dS\,\delta_\psi\{\mathbf{J}\}\cdot\mathbf{n} + \text{c.c.} \tag{3.19b}$$

where

$$\mathbf{J}(\mathbf{r}) = \frac{\hbar N}{2mi}\int' d\boldsymbol{\tau}'\{\psi^*\nabla\psi - (\nabla\psi^*)\psi\} \tag{3.19c}$$

is the (single electron probability) current density,[22] and δ_ψ denotes a variation which does not alter ψ^*. Since ψ is a stationary state, the flux of the associated current density $\mathbf{J}(\mathbf{r})$ through any closed surface vanishes, as a consequence of the continuity equation[22]

$$\frac{d\rho}{dt} = \frac{\partial\rho}{\partial t} + \nabla\cdot\underset{\sim}{J} = 0 \qquad (3.20)$$

and the separate vanishing of $\frac{\partial\rho}{\partial t}$. However, the change in $\underset{\sim}{J}(\underset{\sim}{r})$, $\delta_\psi \underset{\sim}{J}(\underset{\sim}{r})$, which accompanies an infinitesimal variation $\delta\psi$, does not possess necessarily a vanishing divergence, and it is precisely its flux which governs the corresponding changes in the subsystem functionals, eqs. (3.19). To gain a better understanding of the above physical statement of the subsystem variational principle, we consider variations in ψ which are generated by the application of some infinitesimal operator $\epsilon\hat{A}$, i.e. $\delta\psi = \epsilon\hat{A}\psi$.[60,62,65] In particular, $\hat{A}$ may be related to the generator, $\hat{A}'$, of an (infinitesimal) unitary transformation by

$$\hat{A} = i\hat{A}'$$

The r.h.s. of eqs. (3.19) can then be written as the flux of the current associated with the density defined in eq. (3.21)

$$\rho_{iA}(\underset{\sim}{r}) = N\int{}' d\underset{\sim}{\tau}'\{\psi^*(i\hat{A})\psi + \text{c.c.}\} \qquad (3.21)$$

Differentiating both sides of eq. (3.21) with respect to the time variable t, we get

$$\frac{\partial\rho_{iA}(\underset{\sim}{r})}{\partial t} = -i\nabla\cdot\underset{\sim}{J}_A(\underset{\sim}{r}) - \left(\frac{N}{\hbar}\right)\int{}' d\underset{\sim}{\tau}'\psi^*[\hat{H}_{e\ell},\hat{A}]\psi + \text{c.c.} \qquad (3.22a)$$

where

$$\mathbf{J}_A(\mathbf{r}) = \frac{N\hbar}{2mi}\int' d\boldsymbol{\tau}'[\psi^*\nabla(\hat{A}\psi) - (\nabla\psi^*)(\hat{A}\psi)] \qquad (3.22b)$$

In deriving eq. (3.22a), we have made use of the Schrödinger time-dependent equation[21,22]

$$i\hbar\,\frac{\partial\psi(\mathbf{r},t)}{\partial t} = \hat{H}_{e\ell}\,\psi(\mathbf{r},t), \qquad (3.23)$$

and assumed that $\hat{A}$ does not depend explicitly on time. When ψ is an eigenfunction of $\hat{H}_{e\ell}$, by eq. (3.23), it can be shown[21,22] that it depends on the time variable only through the phase factor $e^{(-i/\hbar)Et}$. In this instance, the l.h.s. of eq. (3.22a) vanishes,[66] yielding

$$-i\hbar\nabla\cdot\mathbf{J}_A(\mathbf{r}) + \text{c.c.} = N\int' d\boldsymbol{\tau}'\psi^*[\hat{H}_{e\ell},\hat{A}]\psi + \text{c.c.} \qquad (3.24)$$

Thus, the value of the divergence of $\{i\mathbf{J}_A + \text{c.c.}\}$, which is the current associated with $\rho_{iA}(\mathbf{r})$, is determined by the commutator of the operator $\hat{A}$ with the Hamiltonian $\hat{H}_{e\ell}$. By eqs. (3.19), it is seen that the flux of $\{i\mathbf{J}_A + \text{c.c.}\}$ through the boundary of Ω governs the variations in the subsystem energy functionals, which result from the variations $\delta\psi = \epsilon\hat{A}\psi$ throughout the entire molecular system. Thus,

$$\delta F_E[\Omega,\psi] = -\left(\frac{\epsilon\hbar}{2N}\right)\oint dS\{i\mathbf{J}_A + \text{c.c.}\}\cdot\mathbf{n} = \langle\psi^*,\psi\rangle_\Omega\,\delta G_e'[\Omega,\psi] \qquad (3.25)$$

Applying Gauss' divergence theorem, and using (3.24), one has

$$\delta F_E[\Omega,\psi] = \frac{\epsilon}{2}\{\langle\psi^*,[\hat{H}_{e\ell},\hat{A}]\psi\rangle_\Omega + \text{c.c.}\} \qquad (3.26a)$$

and

$$\delta G_e'[\Omega,\psi] = \frac{\varepsilon}{2}\left\{\frac{\langle\psi^*,[\hat{H}_{e\ell},\hat{A}]\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} + \text{c.c.}\right\} \tag{3.26b}$$

We wish to discuss two important aspects of eqs. (3.26). First, we note that the change in the energy functional $G_e'[\Omega,\psi]$, as generated by an infinitesimal unitary transformation $\psi \rightarrow (1+i\varepsilon\hat{A}')\psi$, is related to the generator $\hat{A} = i\hat{A}'$ of the transformation in a formally similar manner to the way the change of the Hamiltonian function of a classical system is related to the generator of an infinitesimal canonical transformation.[57,62] Let $H(\underset{\sim}{p},\underset{\sim}{q})$ be the Hamiltonian function associated with a classical system of n degrees of freedom, described by the generalized coordinates, $\underset{\sim}{q}$, whose conjugate momenta are denoted by $\underset{\sim}{p}$. If $A(\underset{\sim}{p},\underset{\sim}{q})$ is the generator of a canonical transformation, which induces changes $\delta\underset{\sim}{q} = \nabla_{\underset{\sim}{p}} A(\underset{\sim}{p},\underset{\sim}{q})$, $\delta\underset{\sim}{p} = -\nabla_{\underset{\sim}{q}} A(\underset{\sim}{p},\underset{\sim}{q})$, where $\nabla_{\underset{\sim}{q}(\underset{\sim}{p})}$ denotes the gradient operator with respect to $\underset{\sim}{q}(\underset{\sim}{p})$, then A induces a change in the Hamiltonian function given by

$$\delta H(\underset{\sim}{p},\underset{\sim}{q}) = \{H,A\}_p \tag{3.27a}$$

where $\{f(\underset{\sim}{p},\underset{\sim}{q}),g(\underset{\sim}{p},\underset{\sim}{q})\}_p$ denotes the Poisson bracket between f and g, i.e.,

$$\{f(\underset{\sim}{p},\underset{\sim}{q}),g(\underset{\sim}{p},\underset{\sim}{q})\}_p = \nabla_{\underset{\sim}{q}} f(\underset{\sim}{p},\underset{\sim}{q})\cdot\nabla_{\underset{\sim}{p}} g(\underset{\sim}{p},\underset{\sim}{q}) - \nabla_{\underset{\sim}{q}} g(\underset{\sim}{p},\underset{\sim}{q})\cdot\nabla_{\underset{\sim}{p}} f(\underset{\sim}{p},\underset{\sim}{q}) \tag{3.27b}$$

The correspondence between eq. (3.27a) and eq. (3.26b) is established by replacing H and A by $\langle\psi^*,\hat{H}\psi\rangle_\Omega$ and $\langle\psi^*,\hat{A}\psi\rangle_\Omega$ and the Poisson bracket $\{H,A\}_p$ by $\langle\psi^*,[\hat{H},\hat{A}]\psi\rangle_\Omega$. Thus, we can regard eqs. (3.26) as the quantal counterpart of the classical equation (3.27a). In fact, it can be shown that the average expressions of eqs. (3.26) correspond to an operator equation which is the quantized form of eq. (3.27a).[65] In obtaining this result, we need to construct a projection operator, $\hat{\Pi}(\Omega)$, such that

$$\langle\psi^*,\psi\rangle_\Omega = \langle\psi|\hat{\Pi}(\Omega)|\psi\rangle \qquad (3.28a)$$

and

$$\langle\psi^*,\hat{H}_{e\ell}\psi\rangle_\Omega = \langle\psi|\hat{\Pi}(\Omega)\hat{H}_{e\ell}|\psi\rangle \qquad (3.28b)$$

Thus,

$$\hat{\Pi}(\Omega) = \int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\,|\underset{\sim}{r},\underset{\sim}{\tau}'\rangle\langle\underset{\sim}{r},\underset{\sim}{\tau}'| \qquad (3.28c)$$

In terms of $\hat{\Pi}(\Omega)$, we define the subsystem Hamiltonian

$$\hat{H}(\Omega) = \tfrac{1}{2}\{\hat{\Pi}(\Omega)\hat{H}_{e\ell} + \hat{H}_{e\ell}\hat{\Pi}(\Omega)\} \qquad (3.28d)$$

A unitary transformation $\hat{U} = \hat{1} + i\varepsilon\hat{A}'$ then induces a change in $\hat{H}(\Omega)$ given by

$$\hat{H}(\Omega) = i\varepsilon[\hat{H}(\Omega),\hat{A}'] + \delta_S\hat{H}(\Omega) \qquad (3.29)$$

where $\delta_S\hat{H}(\Omega)$ accounts for the infinitesimal shifts in the surface boundary Ω. Eq. (3.29) is the quantum analogue of eq. (3.27a), generalized to include the above variation in the boundary of the subsystem. The average of eq. (3.29) in the eigenstate $|\psi\rangle$ of $\hat{H}_{e\ell}$ yields eqs. (3.26), using the following results:

$$[\hat{H}(\Omega),\hat{A}'] = \tfrac{1}{2}\{\hat{\Pi}(\Omega)[\hat{H}_{e\ell},\hat{A}'] + [\hat{H}_{e\ell},\hat{A}']\hat{\Pi}(\Omega)\} + \tfrac{1}{2}\{[\hat{\Pi}(\Omega),\hat{A}']\hat{H}_{e\ell} + \hat{H}_{e\ell}[\hat{\Pi}(\Omega),\hat{A}']\} \tag{3.30a}$$

and

$$\delta\langle\psi^*,\psi\rangle_\Omega = i\varepsilon\langle\psi|[\hat{\Pi}(\Omega),\hat{A}']|\psi\rangle + \langle\psi|\delta_S\hat{\Pi}(\Omega)|\psi\rangle = 0 \tag{3.30b}$$

$$\delta_S\hat{H}(\Omega) = \tfrac{1}{2}\{\delta_S\hat{\Pi}(\Omega)\hat{H} + \hat{H}\delta_S\hat{\Pi}(\Omega)\} \tag{3.30c}$$

Note that equation (3.30b) is the consequence of imposing the normalization constraints, $\langle\psi^*,\psi\rangle_\Omega$ = constant, on the variations $\delta\psi = i\varepsilon\hat{A}'\psi$.

As a second remark concerning eqs. (3.26), we note that the variations in the subsystem energy functional vanish identically in two instances. Because of their relationship with the commutator $[\hat{H}_{e\ell},\hat{A}]$ depicted in eqs. (3.26), these variations are necessarily zero when the operator $\hat{A}$ corresponds to a constant of the motion of the total electronic system.^{62} This is hardly a surprise when one notes that in this instance, the density $\rho_{iA}(\underset{\sim}{r})$ satisfies a continuity equation of the same form as the probability density $\rho(\underset{\sim}{r})$. The current $\{i\underset{\sim}{J}_A(\underset{\sim}{r}) + \text{c.c.}\}$ arising from the variations $\delta\psi = \varepsilon\hat{A}\psi$ thus contributes a zero-flux through the boundary of Ω, since its divergence identically vanishes at all points in real space, just as does $\nabla\cdot\underset{\sim}{J}(\underset{\sim}{r})$. Another situation in which $\delta G_e'[\Omega,\psi]$ (and/or $\delta F_E[\Omega,\psi]$) vanishes corresponds to $\Omega = R^3$, that is when the subsystem is the total molecular system. We have already obtained this property, by using eqs. (3.18) and the boundary conditions of eq. (3.3). In the present context, where the variations in ψ are generated by the operator A, the vanishing of $\delta G_e'[\Omega,\psi] = \langle\psi^*,\psi\rangle_\Omega^{-1}\delta F_E[\Omega,\psi]$ follows from eqs. (3.26) and the hypervirial theorem[21,58]

$$\frac{\langle\psi^*,[\hat{H}_{e\ell},\hat{A}]\psi\rangle_{R^3}}{\langle\psi^*,\psi\rangle_{R^3}} = 0 \qquad (3.31)$$

Equation (3.31) is customarily derived by exploiting the hermiticity of the Hamiltonian $\hat{H}_{e\ell}$, and the electronic Schrödinger equation obeyed by the state function $\psi(\underset{\sim}{\tau})$. Alternatively, we may regard eqs. (3.26), together with eq. (3.25) and the boundary conditions at infinity, eq. (3.3), as providing a variational derivation of the hypervirial theorem, eq. (3.31). That is, the property denoted in eq. (3.31), that the all space average of the commutator $[\hat{H}_{e\ell},\hat{A}]$ - in an eigenstate of $\hat{H}_{e\ell}$ - vanishes, is a consequence of the generalized variational principle, eqs. (3.26). This remark suggests that the hypervirial theorem be generalized to read[62,67]

$$\frac{N}{2}\left\{\frac{\langle\psi^*,[\hat{H}_{e\ell},\hat{A}]\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} + \text{c.c.}\right\} \doteq -\frac{\hbar}{2\langle\psi^*,\psi\rangle_\Omega}\oint dS\{i\underset{\sim}{J}_A + \text{c.c.}\}\cdot\underset{\sim}{n} \qquad (3.32)$$

so as to apply to both a subsystem and the total system. The subsystem hypervirial theorem, eq. (3.32), is then related to the subsystem variational principle in the same manner as eq. (3.31) is related to the all space principle, eqs. (3.4).

An important consequence of the hypervirial theorem, eq. (3.31), is the virial theorem, obtained by setting[58,60-62]

$$\hat{A} = \sum_i \underset{\sim}{r}_i \cdot \nabla_i = \frac{i}{\hbar}\sum_i \underset{\sim}{r}_i \cdot \hat{p}_i , \tag{3.33}$$

the summation in the r.h.s. being over the electronic indices i. The commutator of $\hat{A}$ and $\hat{H}_{e\ell}$ is readily determined as

$$[\hat{H}_{e\ell},\hat{A}] = \frac{i}{\hbar}\sum_i\{[\hat{H}_{e\ell},\underset{\sim}{r}_i]\cdot\hat{p}_i + \underset{\sim}{r}_i\cdot[\hat{H}_{e\ell},\hat{p}_i]\} = \sum_i\{\frac{\hat{p}_i^2}{m} - \underset{\sim}{r}_i\cdot\nabla_i\hat{\nu}\} \tag{3.34}$$

where $\hat{\nu}$ is given by eq. (3.6). In deriving eq. (3.34), the explicit expression of $\hat{H}_{e\ell}$,

$$\hat{H}_{e\ell} = \frac{1}{2m}\sum_i \hat{p}_i^2 + \hat{\nu}(\underset{\sim}{\tau})$$

has been used, together with the commutation rules

$$[r_i^k, p_j^\ell] = i\hbar\delta_{ij}\delta_{k\ell}$$

$$[r_i^k, r_j^\ell] = 0 = [p_i^k, p_j^\ell]$$

$$\forall i,j \in \{1,2,\ldots,n\}$$

$$\forall\ k,\ell = 1,2,3$$

where u^k denotes the k^{th} component of the vector $\underset{\sim}{u}$. Substituting $[\hat{H}_{e\ell},\hat{A}]$ in eq. (3.31) by its explicit expression, eq. (3.34), one arrives at the result

$$2\langle\psi|\hat{T}_{e\ell}|\psi\rangle = \langle\psi|\sum_i \underset{\sim}{r}_i\cdot\nabla_i\hat{\nu}|\psi\rangle \tag{3.35a}$$

or

$$2N\langle\psi^*,(-\frac{\hbar^2}{2m})\nabla^2\psi\rangle_{R^3} = N\langle\psi^*,\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_{R^3} \quad , \tag{3.35b}$$

Above, the antisymmetry of ψ, viz. the indistinguishability of the electrons, has been used to rewrite the integrals of eq. (3.35a) in the forms shown in eq. (3.35b). Eqs. (3.35) embody what is called the virial theorem.[21,58,60-62] Classically, the gradient of the potential energy $\hat{v}(\underset{\sim}{r},\underset{\sim}{\tau}',\{\underset{\sim}{R}_\alpha\})$ with respect to the electronic coordinates $\underset{\sim}{r}$ gives the (negative of the) total force on the electron whose instantaneous position vector is $\underset{\sim}{r}$, when the other electrons are in the configuration described by $\underset{\sim}{\tau}'$, and the nuclei assume the instantaneous configuration $\{\underset{\sim}{R}_\alpha\}$. The scalar product of $-\nabla\hat{v}$ with $\underset{\sim}{r}$ is the virial of this force. Thus, eq. (3.35a) states that minus twice the average kinetic energy of the electronic system equals the average of the virial of the forces exerted on the electrons. Because the electrons are indistinguishable, the same relationship is obtained between the average kinetic energy and the virial of the force associated with a single electron, eq. (3.35b). The above relationship, eq. (3.35b), can also be obtained by using the single electron generator $\hat{A} = \underset{\sim}{r}\cdot\nabla$ in eq. (3.31).[60,62]

It is well known that the transformation generated by the operator $\varepsilon\sum_i \underset{\sim}{r}_i\cdot\nabla_i$ is the scaling of all the electronic coordinates of ψ[21,58,60-62]

$$(1+\varepsilon\sum_i \underset{\sim}{r}_i\cdot\nabla_i)\psi(\underset{\sim}{\tau}) = \psi(\zeta\underset{\sim}{\tau}) \tag{3.36a}$$

where

$$\zeta = (1 + \varepsilon)$$

In contrast, the operator $\varepsilon\underset{\sim}{r}\cdot\nabla$ corresponds to the scaling of the coordinates of one electron at a time, the electron involved being arbitrary,[60,62]

$$(1 + \varepsilon\underset{\sim}{r}\cdot\nabla)\psi(\underset{\sim}{r},\underset{\sim}{\tau}') = \psi(\zeta\underset{\sim}{r},\underset{\sim}{\tau}'). \qquad (3.36b)$$

To evaluate the commutator between $\hat{H}_{e\ell}$ and the single particle generator $\hat{A} = \underset{\sim}{r}\cdot\nabla$, we rewrite the Hamiltonian accordingly so as to make the dependence on the single electronic coordinates $\underset{\sim}{r}$ and its conjugate momentum, $\hat{\underset{\sim}{p}} = -i\hbar\nabla$, appear explicitly

$$\hat{H}_{e\ell} = (-\frac{\hbar^2}{2m})\nabla^2 + (-\frac{\hbar^2}{2m})\sum_{i\varepsilon\tau'}\nabla_i^2 + \hat{v}(\underset{\sim}{r},\underset{\sim}{\tau}') \qquad (3.37)$$

The commutator between $\hat{H}_{e\ell}$ and $\hat{A}$ is then readily found to be

$$[\hat{H}_{e\ell},\hat{A}] = -\frac{\hbar^2}{2m}\nabla^2 - \underset{\sim}{r}\cdot\nabla\hat{v} \qquad (3.38)$$

whereupon, the application of the subsystem hypervirial theorem, eq. (3.32), gives

$$\delta G_e'[\Omega,\psi] = \frac{\varepsilon}{2}\{\frac{\langle\psi^*,\{(-\frac{\hbar^2}{2m})\nabla^2 - \underset{\sim}{r}\cdot\nabla\hat{v}\}\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} + \text{c.c.}$$

$$= \frac{-\varepsilon\hbar}{2N\langle\psi^*,\psi\rangle_\Omega}\oint dS\{iJ_{\underset{\sim}{r}\cdot\underset{\sim}{\nabla}} + \text{c.c.}\}\cdot\underset{\sim}{n} \qquad (3.39)$$

Equation (3.39) can be rearranged to give the subsystem virial theorem, eq. (3.40).

$$\frac{2\langle\psi^*,(-\frac{\hbar^2}{2m})\nabla^2\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} + \text{c.c.} = \frac{\langle\psi^*,\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_\Omega}{\langle\psi^*,\psi\rangle_\Omega} - \frac{i\hbar}{2N\langle\psi^*,\psi\rangle_\Omega}\oint dSJ_{\underset{\sim}{r}\cdot\underset{\sim}{\nabla}}\cdot\underset{\sim}{n} + \text{c.c.} \qquad (3.4$$

The l.h.s. of eq. (3.40) is readily recognized as representing the kinetic energy of the subsystem, while the first term in the r.h.s. represents the subsystem average of the virial of the total force exerted on an electron. The surface integral in the r.h.s. of eq. (3.40) can be rewritten as

$$\frac{i\hbar}{2}\oint dS\underset{\sim}{J}_{\underset{\sim}{r}\cdot\nabla}\cdot\underset{\sim}{n} + \text{c.c.} = \frac{N\hbar^2}{4m}\oint dS\oint' d\underset{\sim}{\tau}'\{\psi^*\nabla(\underset{\sim}{r}\cdot\nabla\psi) - \nabla\psi^*(\underset{\sim}{r}\cdot\nabla\psi) +$$

$$\nabla(\underset{\sim}{r}\cdot\nabla\psi^*)\psi - (\underset{\sim}{r}\cdot\nabla\psi^*)\nabla\psi\}\cdot\underset{\sim}{n}$$

$$= \frac{N\hbar^2}{4m}\oint dS\{\int' d\tau'\underset{\sim}{r}\cdot[\psi^*\nabla(\nabla\psi) - (\nabla\psi^*)\nabla\psi]\cdot\underset{\sim}{n} + \text{c.c.} +$$

$$\nabla\rho'\cdot\underset{\sim}{n}\}$$

$$= \frac{i\hbar}{2}\oint dS(\underset{\sim}{r}\cdot\underset{\sim}{J}_{\nabla})\cdot\underset{\sim}{n} + \text{c.c.} \tag{3.41a}$$

or, since $\hat{\underset{\sim}{p}} = -i\hbar\nabla$,

$$\frac{i\hbar}{2}\oint dS\underset{\sim}{J}_{\underset{\sim}{r}\cdot\nabla}\cdot\underset{\sim}{n} + \text{c.c.} = -\tfrac{1}{2}\oint dS(\underset{\sim}{r}\cdot\underset{\sim}{J}_{\hat{p}}\cdot\underset{\sim}{n}) + \text{c.c.} \tag{3.41b}$$

In the form of eq. (3.41b), the surface term which appears in the statement of the subsystem virial theorem, eq. (3.40), acquires the significance of a "surface flux virial" - as called by J.A. Schweitz[68,69] - which is caused by the exchange of particles through the boundary of an open system. Schweitz has considered a generalization of the classical virial theorem, to include the case where the classical system is open, and showed that Clausius' theorem should be generalized to read[68]

$$2\overline{T} + Z_p + Z_s = 0 \tag{3.42a}$$

where $\bar{T}$ is the time average of the kinetic energy of the system, which is associated with a volume Ω, Z_p is the time average of the virial of the forces on the particles of the system

$$Z_p = \overline{(\sum_{i\epsilon\Omega} \underset{\sim}{r}_i \cdot \underset{\sim}{F}_i)} \tag{3.42b}$$

The last term, Z_s, of eq. (3.42a) is given by

$$Z_s = \oint dS \underset{\sim}{r} \cdot (\overleftarrow{p}(\underset{\sim}{r}) - \vec{p}(\underset{\sim}{r})) \tag{3.42c}$$

where $\overleftarrow{p}(\underset{\sim}{r})$ and $\vec{p}(\underset{\sim}{r})$ are the "mean flux densities of momentum" into and out of the volume Ω. More precisely, they are defined as the time average of the surface momentum density of particles, respectively, entering or leaving the volume Ω, through the point $\underset{\sim}{r}$ of its boundary. On the basis of the classical result, eq. (3.42a), Schweitz anticipates that an additional surface term, as given in eq. (3.41b), will also arise in the quantal counterpart of eq. (3.42a).[69] In the above, we have obtained the anticipated result, by explicitly using the zero-flux boundary condition for the (open) subsystem Ω, as illustrated by the rearrangements in eq. (3.41a). One can also regard the surface term of eq. (3.40) as the virial of the forces exerted on the surfaces of Ω.[70] In this interpretation, the resultant force on the portion of the charge distribution which constitutes the subsystem is equal to the negative of the resultant force acting on its boundary.

The result is obtained by noting that eq. (3.41a) can also be written as[70]

$$-\frac{i\hbar}{2}\oint dS J_{\underset{\approx}{r}:\nabla}\cdot\underset{\sim}{n} + \text{c.c.} = -\oint dS\underset{\sim}{r}\cdot\overleftrightarrow{P}(\underset{\sim}{r})\cdot\underset{\sim}{n}(\underset{\sim}{r}) \qquad (3.43a)$$

where

$$\overleftrightarrow{P}(\underset{\sim}{r}) = \frac{N\hbar^2}{4m}\int' d\underset{\sim}{\tau}'\{\nabla(\nabla\psi^*)\psi + \psi^*\nabla(\nabla\psi) - \nabla\psi^*\nabla\psi - \nabla\psi\nabla\psi^*\} \qquad (3.43b)$$

is a symmetric dyadic, referred to as the pressure tensor. Its "scalar" product with $\underset{\sim}{n}(\underset{\sim}{r})$ gives the surface density of the force exerted on the surface element which is specified by its outwardly directed normal, $\underset{\sim}{n}(\underset{\sim}{r})$. The pressure tensor $\overleftrightarrow{P}(\underset{\sim}{r})$, also appears in a particular expression of a force law, baptized "the differential force law" by Epstein.[71] When referring to a stationary state, and when integrated over (N - 1) electronic coordinates, Epstein's differential force law writes

$$\underset{\sim}{F}(\underset{\sim}{r}) = N\int' d\underset{\sim}{\tau}'\psi^*(-\nabla\hat{v})\psi = -\nabla\cdot\overleftrightarrow{P}(\underset{\sim}{r}) \qquad (3.44)$$

stating that the force density $\underset{\sim}{F}(\underset{\sim}{r})$ is proportional to the divergence of $\overleftrightarrow{P}(\underset{\sim}{r})$. Thus, the resultant force on the subsystem Ω is[70]

$$\underset{\sim}{F}(\Omega) = \int_\Omega d\underset{\sim}{r}\underset{\sim}{F}(\underset{\sim}{r}) = -\oint dS\overleftrightarrow{P}(\underset{\sim}{r})\cdot\underset{\sim}{n}(\underset{\sim}{r}) \qquad (3.45)$$

establishing the anticipated result.

In summary, the subsystem virial theorem can be written in the following equivalent forms:

$$-2\bar{T}(\Omega) = \bar{V}(\Omega) + \bar{V}_S(\Omega) \qquad (3.46a)$$

$$-2\bar{T}(\Omega) = \bar{V}(\Omega) + \{-\tfrac{1}{2}\oint dS(\underset{\sim}{r}\cdot\underset{\sim}{J}_{\hat{p}}\cdot\underset{\sim}{n}) + \text{c.c.}\} \qquad (3.46b)$$

$$-2\bar{T}(\Omega) = \bar{V}(\Omega) + \{\frac{i\hbar}{2}\oint dS\underset{\sim}{J}_{\underset{\sim}{r}\cdot\nabla}\cdot\underset{\sim}{n} + \text{c.c.}\} \qquad (3.46c)$$

In all three equations (3.46), $\bar{V}(\Omega)$ denotes the subsystem average of the operator $-\underset{\sim}{r}\cdot\nabla\hat{v}$, and $\bar{T}(\Omega)$ denotes the subsystem average kinetic energy. The term $\bar{V}_S(\Omega)$ in the r.h.s. of eq. (3.46a) is the virial of the forces exerted on the surface bounding Ω, and, as discussed above, is related to the pressure tensor, $\overleftrightarrow{P}(\underset{\sim}{r})$, by

$$\bar{V}_S(\Omega) = \oint dS\underset{\sim}{r}\cdot\overleftrightarrow{P}(\underset{\sim}{r})\cdot\underset{\sim}{n}(\underset{\sim}{r})$$

Equation (3.46b) is the particular statement of the virial theorem which parallels Schweitz's result of eq. (3.42a). This equation, as well as the physical statement of eq. (3.46a), are obtained only if Ω is bounded by a zero-flux surface. Thus, the zero-flux boundary condition is fundamental in two respects: first, it ensures the equivalence of the two subsystem energy functionals $F_E[\Omega,\psi]$ and $G_e'[\Omega,\psi]$, by establishing the identity between the two possible expressions for the subsystem kinetic energy, $\bar{T}(\Omega)$:[13-16,20,60,62]

$$\bar{T}(\Omega) = \frac{N}{2}\{\langle\psi^*,(-\frac{\hbar^2}{2m})\nabla^2\psi\rangle_\Omega + \text{c.c.}\} = N\langle(\frac{\hbar^2}{2m})\nabla\psi^*,\nabla\psi\rangle_\Omega \qquad (3.47)$$

Secondly, the zero-flux boundary condition leads to the physical expressions, eqs. (3.46a,b), of the subsystem virial theorem. This theorem has been obtained via the subsystem variational principle, eqs. (3.18) and eqs. (3.26), where the variations in $\psi(\underset{\sim}{\tau})$ are generated by the infinitesimal scaling operator $\varepsilon\hat{A} = \varepsilon\underset{\sim}{r}\cdot\nabla$. The all-space result, eq. (3.35b), is trivially obtained, as the surface terms in the r.h.s. of eqs. (3.46) vanish, when $\Omega = R^3$. In this case, the term $\bar{V}(\Omega)$ can simply be related to the average potential energy and the virial of the forces acting on the nuclei. One first verifies that, because $\hat{\nu}(\underset{\sim}{r},\underset{\sim}{\tau}')$ is a homogeneous function of the electronic and nuclear coordinates of degree -1, it satisfies[60,62,65]

$$-\underset{\sim}{r}\cdot\nabla\hat{\nu} - \sum_{\underset{\sim}{r}_i\in\underset{\sim}{\tau}'} \underset{\sim}{r}_i\cdot\nabla_i\hat{\nu} = \hat{\nu} + \sum_\alpha R_\alpha\cdot\nabla_\alpha\hat{\nu} \tag{3.48a}$$

One then notes the identity

$$\langle\psi^*,\underset{\sim}{r}_i\cdot\nabla_i\hat{\nu}\psi\rangle_{R^3} = \langle\psi^*,\underset{\sim}{r}\cdot\nabla\hat{\nu}\psi\rangle_{R^3}, \quad \forall i = 1,2,..,n \tag{3.48b}$$

to write

$$-N\langle\psi^*,\underset{\sim}{r}\cdot\nabla\hat{\nu}\psi\rangle_{R^3} = \langle\psi^*,\hat{\nu}\psi\rangle_{R^3} + \sum_\alpha \underset{\sim}{R}_\alpha\cdot\langle\psi^*,\nabla_\alpha\hat{\nu}\psi\rangle_{R^3} \tag{3.49}$$

By virtue of the Hellmann-Feynman theorem,[21,58,72] we have

$$\langle\psi^*,\psi\rangle_{R^3}\nabla_\alpha E = -\langle\psi^*,\psi\rangle_{R^3}\underset{\sim}{F}_\alpha = \langle\psi^*,\nabla_\alpha\hat{\nu}\psi\rangle_{R^3}$$

thus

$$\bar{V}(\Omega=R^3) = -N\langle\psi^*,\underset{\sim}{r}\cdot\nabla\hat{\nu}\psi\rangle_{R^3} = \langle\psi^*,\hat{\nu}\psi\rangle_{R^3} - \langle\psi^*,\psi\rangle_{R^3}\sum_\alpha \underset{\sim}{R}_\alpha\cdot\underset{\sim}{F}_\alpha \tag{3.50}$$

where $\underset{\sim}{F}_\alpha$ is the force exerted on the nucleus α by the electronic system. We can also use eq. (3.6) to further decompose the first term in the r.h.s. of eq. (3.50) into electron-nuclear and electron-electron interaction terms:

$$\langle\psi^*,\hat{\nu}\psi\rangle_{R^3} = N\langle\psi^*,\nu_{ne}\psi\rangle_{R^3} + \frac{N(N-1)}{2}\langle\psi^*,\nu_{ee}\psi\rangle_{R^3}$$

$$= \int d\underset{\sim}{r}\,\nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r}) + \int d\underset{\sim}{r}\int d\underset{\sim}{r}'\,\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') \quad (3.51a)$$

where the 2-matrix has been defined in eq. (1.3), and

$$\nu_{ne}(\underset{\sim}{r}) = -\frac{Z_\alpha e^2}{|\underset{\sim}{r} - \underset{\sim}{R}_\alpha|}$$

$$\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}') = \frac{e^2}{|\underset{\sim}{r} - \underset{\sim}{r}'|} \quad . \qquad (3.51b)$$

Thus, considering the wavefunction ψ normalized to unity, we rewrite eq. (3.50) as

$$\overline{V}(\Omega=R^3) = \overline{V}_{ne}(\Omega=R^3) + \overline{V}_{ee}(\Omega=R^3) - \sum_\alpha \underset{\sim}{R}_\alpha\cdot\underset{\sim}{F}_\alpha \qquad (3.52a)$$

where[20,60]

$$\overline{V}_{ne}(\Omega) = \int_\Omega d\underset{\sim}{r}\,\nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r}) \qquad (3.52b)$$

$$\overline{V}_{ee}(\Omega) = \int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\,\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}'). \qquad (3.52c)$$

The virial theorem for the total system is thus

$$-2\overline{T} = \overline{V}_{ne} + \overline{V}_{ee} - \sum_\alpha \underset{\sim}{R}_\alpha\cdot\underset{\sim}{F}_\alpha \quad , \qquad (3.53)$$

where the volume dependence of the various terms has been deleted, as $\Omega = R^3$. When Ω refers to a proper subsystem (i.e., $\Omega \neq R^3$), eq. (3.48b) is no longer applicable, while using eq. (3.48a) we have

$$\bar{V}(\Omega) = -N\langle\psi^*, \underset{\sim}{r}\cdot\nabla\hat{\nu}\psi\rangle_\Omega = N\langle\psi^*,\hat{\nu}\psi\rangle_\Omega + 2\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'(\underset{\sim}{r}'\cdot\nabla'\hat{\nu})\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}')$$

$$+ N\sum_\alpha \underset{\sim}{R}_\alpha\cdot\langle\psi^*,\nabla_\alpha\hat{\nu}\psi\rangle_\Omega \;.$$

Using eq. (3.6), the definitions of eq. (3.51b), and those of eq. (1.3), the r.h.s. of the above equation can be developed to give

$$\begin{aligned}\bar{V}(\Omega) = {} & \int_\Omega \nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r})d\underset{\sim}{r} + \sum_\alpha \underset{\sim}{R}_\alpha\cdot\int_\Omega\nabla_\alpha\nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r})d\underset{\sim}{r} \\ & + 2\{\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\nu_{ne}(\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') + \int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\underset{\sim}{r}'\cdot\nabla'\nu_{ne}(\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') \\ & + \sum_\alpha \underset{\sim}{R}_\alpha\cdot\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\nabla_\alpha\nu_{ne}(\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}')\} + 2\{\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') \\ & + \int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\underset{\sim}{r}'\cdot\nabla'\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}')\} + 3\{\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\int d\underset{\sim}{r}''\nu_{ee}(\underset{\sim}{r}',\underset{\sim}{r}'')\times \\ & \Gamma^{(3)}(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'') + 2\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\int d\underset{\sim}{r}''\underset{\sim}{r}'\cdot\nabla'\nu_{ee}(\underset{\sim}{r}',\underset{\sim}{r}'')\Gamma^{(3)}(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'')\}\end{aligned} \tag{3.54}$$

The sum of the three terms in the first curly brackets of eq. (3.53) gives exactly zero, because we have

$$\underset{\sim}{r}'\cdot\nabla'\nu_{ne}(\underset{\sim}{r}') + \sum_\alpha \underset{\sim}{R}_\alpha\cdot\nabla_\alpha\nu_{ne}(\underset{\sim}{r}') = -\nu_{ne}(\underset{\sim}{r}')$$

identically. Likewise, the two terms in the last curly brackets of eq. (3.53) sum up to zero, since

$$2\int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\int d\underset{\sim}{r}''\underset{\sim}{r}'\cdot\nabla'\nu_{ee}(\underset{\sim}{r}',\underset{\sim}{r}'')\Gamma^{(3)}(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'') = \int_\Omega d\underset{\sim}{r}\int d\underset{\sim}{r}'\int d\underset{\sim}{r}''(\underset{\sim}{r}'\cdot\nabla' + \underset{\sim}{r}''\cdot\nabla'')\nu_{ee}\Gamma^{(3)}$$

follows from the symmetry of the triple integral with respect to the permutation of the primed and double-primed coordinates. Thus, eq. (3.53) reduces to

$$\bar{V}(\Omega) = \int_{\Omega} \nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r})d\underset{\sim}{r} + \int_{\Omega} d\underset{\sim}{r}\int d\underset{\sim}{r}'\,\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') + \sum_{\alpha} \underset{\sim}{R}_{\alpha}\cdot\int_{\Omega} \nabla_{\alpha}\nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r})d\underset{\sim}{r}$$

$$+ \int_{\Omega} d\underset{\sim}{r}\int_{\Omega'} d\underset{\sim}{r}'(\underset{\sim}{r}'\cdot\nabla' - \underset{\sim}{r}\cdot\nabla)\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') \cdot \qquad (3.55)$$

In obtaining eq. (3.54), we have rearranged the second curly brackets of eq. (3.53), using the relation[60,62,70]

$$2\int_{\Omega} d\underset{\sim}{r}\int d\underset{\sim}{r}'\,\underset{\sim}{r}'\cdot\nabla'\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') + \int_{\Omega} d\underset{\sim}{r}\int d\underset{\sim}{r}'\,\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') =$$

$$\int_{\Omega} d\underset{\sim}{r}\int_{\Omega'} d\underset{\sim}{r}'(\underset{\sim}{r}'\cdot\nabla' - \underset{\sim}{r}\cdot\nabla)\nu_{ee}(\underset{\sim}{r},\underset{\sim}{r}')\Gamma^{(2)}(\underset{\sim}{r},\underset{\sim}{r}') = \bar{V}(\Omega,\Omega')\cdot$$

where $\Omega' = R^3 - \Omega$. Define

$$F_{\alpha}(\Omega) = -\int_{\Omega} \nabla_{\alpha}\nu_{ne}(\underset{\sim}{r})\rho(\underset{\sim}{r})d\underset{\sim}{r} \;,$$

the Hellmann-Feynman force on nucleus α due to the charge distribution in Ω. The virial of the forces operating in the subsystem now assumes the simple expression[60]

$$\bar{V}(\Omega) = \bar{V}_{ne}(\Omega) + \bar{V}_{ee}(\Omega) - \sum_{\alpha} \underset{\sim}{R}_{\alpha}\cdot\underset{\sim}{F}_{\alpha}(\Omega) + \bar{V}''(\Omega,\Omega') \;, \qquad (3.56)$$

which is to be compared with eq. (3.52a). We see that in addition to the subsystem projections of the terms in eq. (3.52a), $\bar{V}(\Omega)$ also contains the interfragment term $\bar{V}''(\Omega,\Omega')$, whose value depends on the origin of the molecular coordinates system. Its vanishing for some particular choice of this origin would represent a balance between the virial of the forces which the electronic distribution in Ω exerts on the distribution in Ω', and the virial of the forces exerted on Ω by the distribution in the complementary space.[60]

Numerical results do exist which suggest that the choice of origin which causes $\overline{V}''(\Omega,\Omega')$ to vanish also ensures the vanishing of the surface terms in eqs. (3.46).[60] If this property is indeed true in general, it will imply that for some particular choice of origin, the statement of the virial theorem for a subsystem is exactly of the same form as for the total system. That is, the subsystem kinetic energy is related to the various potential energy components, and to the virial of the nuclear forces by

$$-2\overline{T}(\Omega) = \overline{V}_{ne}(\Omega) + \overline{V}_{ee}(\Omega) - \sum_{\alpha} \mathbf{R}_{\alpha} \cdot \mathbf{F}_{\alpha}(\Omega) \, , \qquad (3.57a)$$

when the origin of the coordinates system in R^3 is chosen so as to make $\overline{V}''(\Omega,\Omega')$ and $\overline{V}_S(\Omega)$ both vanish. If $\overline{V}''(\Omega,\Omega')$ only is made to vanish, the surface term $\overline{V}_S(\Omega)$ will appear as an additional term in the r.h.s. of eq. (3.57). Even in this instance, the subsystem appears to be self-contained: the surface term $\overline{V}_S(\Omega)$ accounts for all the interactions of the subsystem with its surrounding, and the virial relationship obtained by adding $\overline{V}_S(\Omega)$ to the r.h.s. of eq. (3.57a) leads to the definition of the subsystem energy, $\overline{E}(\Omega)$, which satisfies

$$-\overline{T}(\Omega) = \overline{E}(\Omega) - \sum_{\alpha} \mathbf{R}_{\alpha} \cdot \mathbf{F}_{\alpha}(\Omega) + \overline{V}_S(\Omega) \, . \qquad (3.57b)$$

Thus,

$$\bar{E}(\Omega) = \bar{T}(\Omega) + \bar{V}_{ne}(\Omega) + \bar{V}_{ee}(\Omega), \tag{3.58a}$$

that is, $\bar{E}(\Omega)$ is given by the sum of the subsystem average kinetic and potential energies. Eq. (3.58a) is obtained from the statement of the virial theorem of eq. (3.57b), only when the origin of the molecular coordinates system is chosen so that $\bar{V}''(\Omega,\Omega') = 0$. In general, one obtains, from eq. (3.57b)

$$\begin{aligned}\bar{E}(\Omega) &= \bar{T}(\Omega) + \bar{V}_{ne}(\Omega) + \tfrac{1}{2}\langle -\underset{\sim}{r}\cdot\nabla\hat{v}_{ee}\rangle_{\Omega} \\ &= N\langle \psi^*,\hat{H}_1\psi\rangle_{\Omega}\,. \end{aligned} \tag{3.58b}$$

We define

$$\hat{H}_i = -\frac{\hbar^2}{2m}\nabla_i + \hat{v}_{ne}(\underset{\sim}{r}_i) - \tfrac{1}{2}\underset{\sim}{r}_i\cdot\nabla_i\Big(\sum_{j\neq i}\hat{v}_{ee}(\underset{\sim}{r}_i,\underset{\sim}{r}_j)\Big),$$

for all $i = 1,2,..,N$. Since $\hat{v}_{ee}(\underset{\sim}{r}_i,\underset{\sim}{r}_j)$ is a homogeneous function of degree -1 in $\underset{\sim}{r}_i$ and $\underset{\sim}{r}_j$, we have

$$\hat{H}_{e\ell} = \sum_{i=\ell}^{N}\hat{H}_i\,.$$

Defining the projection operator $\hat{\Pi}(\Omega,i)$ by

$$\hat{\Pi}(\Omega,i) = \int_{\Omega} d\underset{\sim}{r}_i \int \prod_{j\neq i} d\underset{\sim}{r}_j\,|..\underset{\sim}{r}_i..\rangle\langle..\underset{\sim}{r}_i|\,,$$

we further have

$$\hat{H}_{e\ell} = \sum_{i\Omega}\tfrac{1}{2}(\hat{\Pi}(\Omega,i)\hat{H}_i + \hat{H}_i\hat{\Pi}(\Omega,i)),$$

so that

$$\begin{aligned} E_e &= \langle\psi^*,\hat{H}_{e\ell}\psi\rangle_{R^3} = \tfrac{1}{2}\sum_{i\Omega}\sum\{\langle\psi|\hat{\Pi}(\Omega,i)\hat{H}_i|\psi\rangle + \text{c.c.}\} \\ &= \sum_{\Omega} N\langle\psi^*,\hat{H},\psi\rangle_{\Omega} \\ &= \sum_{\Omega}\bar{E}(\Omega) \ . \end{aligned} \tag{3.59}$$

The result of eq. (3.59) denotes the important additivity property exhibited by the subsystem energy, $\bar{E}(\Omega)$, defined in eq. (3.58b). The above construction clearly indicates that this additivity property is a consequence of the inverse-square force law governing the electronic interactions, and the equivalence of the electrons of the molecular system.[60,62,70]

Throughout this section, we have derived the general statements of the stationary-state variational principle which apply equally to a subsystem and to the total system. As expressed in eqs. (3.18), the generalized variational principle requires that the first-order changes in the subsystem energy functionals $G_e'[\Omega,\psi]$ and $F_E[\Omega,\psi]$ be governed by the flux of the single-particle vector current through the subsystem boundary, as generated by the variations $\delta\psi(\underset{\sim}{\tau})$. When these variations are expressible in the form $\delta\psi(\underset{\sim}{\tau}) = \varepsilon\hat{A}\psi(\underset{\sim}{\tau})$, the flux of the current generated by the variations, through the surface of the subsystem, can be related to the subsystem average of the commutator between the generator $\hat{A}$ and the electronic Hamiltonian $\hat{H}_e$. We note that when $\varepsilon\hat{A}$ describes an infinitesimal unitary transformation, the variational equations (3.26) are the quantal counterparts of the classical equation (3.27a). Equations (3.26) can also be regarded as the variational expressions of the hypervirial theorem, generalized to refer both to a subsystem and to a total system. An important consequence of the subsystem hypervirial theorem is obtained by considering $\hat{A} = \underset{\sim}{r}\cdot\nabla$, the generator of a scaling transformation. The relationship so obtained is the subsystem virial theorem. It relates the subsystem average kinetic energy to the virial of the forces acting within the subsystem and on its bounding surface.

C. Time Dependent Variational Principles

In the previous section, the variational properties of an open subsystem in a time-independent molecular environment were studied and delineated. The total system, consisting of the subsystem and its environment, was described by a wavefunction satisfying the electronic time-independent Schrödinger equation (1.6a). The use of the subsystem stationary-state variational principle, eqs. (3.18) or eqs. (3.26), has been shown to be formally equivalent to utilizing a solution of eq. (1.6a). Moreover, it has the merit of offering a pathway towards the understanding of the mechanics of molecular fragments - atoms or functional groups - and of open systems in general, when the total system is time-independent. In this section, we consider the case in which the total system is described by a time-dependent state function $\psi(\underset{\sim}{\tau},t)$, which satisfies the Schrödinger equation

$$i\hbar \frac{\partial\psi(\underset{\sim}{\tau},t)}{\partial t} = \hat{H}\psi(\underset{\sim}{\tau},t) \; . \qquad (3.60)$$

Just as eq. (1.6a) is the Euler equation associated with the stationary state variational principle, $\delta E_e[\psi] = 0$, the time dependent Schrödinger equation (3.60) can be obtained as the Euler equation of a variational problem which we now formulate. We first define the Lagrangian density[62,70,73]

$$L[\phi(t);\underset{\sim}{r}] = \int' d\underset{\sim}{\tau}'\{\frac{i\hbar}{2}(\phi^*\dot{\phi} - \dot{\phi}^*\phi) - \frac{\hbar^2}{2m}\sum_i \nabla_i\phi^*\cdot\nabla_i\phi - \hat{v}\phi^*\phi\} \; , \qquad (3.61)$$

where the trial functions $\phi(\underset{\sim}{\tau},t)$ are again required to satisfy the same boundary conditions as $\psi(\underset{\sim}{\tau})$ in eq. (3.3). That is

$$\lim_{|\underset{\sim}{r}_j|\to+\infty} \phi^{(n)}(\underset{\sim}{\tau},t) = 0, \quad \forall\, j = 1,2,..,N, \; \forall\, n = 0,1,2, \forall\, t. \qquad (3.62)$$

The potential energy $\hat{v}(\underset{\sim}{r},\underset{\sim}{\tau}',t)$ is defined as in eq. (3.6), with possibly an additional time-dependent term. As in the previous section $\underset{\sim}{\tau} = (\underset{\sim}{r},\underset{\sim}{\tau}')$ denotes the coordinates and spin of all the electrons, and $\underset{\sim}{r}$ designates the spatial coordinates of an arbitrary electron. The Lagrangian density defined in eq. (3.61) is a functional of $\phi, \nabla_i\phi$, $i = 1,2,..,N$ and $\dot{\phi}$ as well as of the respective complex conjugate quantities. Its all-space integral defines the Lagrangian, $\mathcal{L}[R^3,\phi(t)]$, associated with the total system. Thus,[62,70]

$$\mathcal{L}[R^3,\phi(t)] = \int d\underset{\sim}{r} L[\phi(t);\underset{\sim}{r}] \; . \qquad (3.62a)$$

In terms of the total system Lagrangian $\mathcal{L}[R^3,\phi(t)]$, we construct the action integral

$$W_{12}[R^3,\phi] = \int_{t_1}^{t_2} dt\, \mathcal{L}[R^3,\phi(t)] \ , \tag{3.62b}$$

where the time end-points, t_1, t_2, are arbitrary. We now show that a solution of the Schrödinger equation (3.60), which we shall denote by $\psi(\underset{\sim}{\tau},t)$, corresponds to a stationary point of W_{12}, viz. it is a solution of the variational problem

$$\delta W_{12}\Big|_{\phi=\psi} = 0 \ , \tag{3.63}$$

where the variations, $\delta\psi = (\phi - \psi)$, are required to vanish at the time end-points of W_{12}. Indeed, the first-order change in W_{12}, which results from the variations $\delta\phi$ in the trial function, is

$$\begin{aligned}
\delta W_{12} &= \int_{t_1}^{t_2} dt \int d\underset{\sim}{\tau} \{ (\frac{\partial L}{\partial \phi})\delta\phi + \sum_i (\frac{\partial L}{\partial(\nabla_i\phi)})\cdot\delta(\nabla_i\phi) + (\frac{\partial L}{\partial\dot{\phi}})\delta(\dot{\phi}) + \text{c.c.}\} \\
&= \int_{t_1}^{t_2} dt \int d\underset{\sim}{\tau} \{ (-\frac{i\hbar}{2}\dot{\phi}^* - \hat{v}\phi^*)\delta\phi + \sum_i (-\frac{\hbar^2}{2m})\nabla_i\phi^*\cdot\nabla_i(\delta\phi) \\
&\quad + \frac{i\hbar}{2}\phi^* \frac{\partial}{\partial t}(\delta\phi) + \text{c.c.}\} \\
&= \int_{t_1}^{t_2} dt \int d\underset{\sim}{\tau} \{ -i\hbar\dot{\phi}^* - (\frac{-\hbar^2}{2m})\sum_i \nabla_i^2\phi^* - \hat{v}\phi^* \}\delta\phi + \text{c.c.}
\end{aligned} \tag{3.64}$$

In deriving eq. (3.64), we have used the following identities

$$\frac{\partial L}{\partial\phi} = -\frac{i\hbar}{2}\dot{\phi}^* - \hat{v}\phi^* \ , \tag{3.65a}$$

$$\frac{\partial L}{\partial(\nabla_i\phi)} = (-\frac{\hbar^2}{2m})\nabla_i\phi^* \ , \tag{3.65b}$$

$$\frac{\partial L}{\partial\dot{\phi}} = \frac{i\hbar}{2}\phi^* \ , \tag{3.65c}$$

which follow directly from the definition of $L[\phi(t);\underset{\sim}{r}]$, eq. (3.61). The vector whose components are the partial derivatives of L with respect to the components of $\nabla_i\phi$, $(\partial L/\partial(\partial\phi/\partial x_i))$, is denoted by $\partial L/\partial(\nabla_i\phi)$. The last identity given in eq. (3.64) results from a rearrangement which uses the identities

$$\nabla_i\phi^*\cdot\phi_i(\delta\phi) = \nabla_i\cdot(\nabla_i\phi^*\delta\phi) - (\nabla_i^2\phi^*)\delta\phi \ , \tag{3.65d}$$

and

$$\phi^*\frac{\partial}{\partial t}(\delta\phi) = \frac{\partial}{\partial t}(\phi^*\delta\phi) - \dot\phi^*\delta\phi \ , \tag{3.65e}$$

together with the boundary conditions on the trial function, ϕ, and its variations, $\delta\phi$. When $\phi = \psi$, a solution to the Schrödinger equation (3.60), the variation of W_{12} vanishes identically, since eq. (3.64) gives

$$\delta W_{12}\Big|_{\phi=\psi} = \int_{t_1}^{t_2} dt\int d\underset{\sim}{\tau}\{-i\hbar\dot\psi^* - \hat{H}\psi^*\}\delta\psi + c.c. = 0 \ . \tag{3.66}$$

Conversely, requiring δW_{12} to vanish for any variation $\delta\psi$ satisfying the boundary conditions of eq. (3.60), and which vanishes at the time end-points, one obtains the Euler equations

$$-i\hbar\dot\psi^* - \hat{H}\psi^* = 0 \ , \tag{3.67a}$$

and

$$i\hbar\dot\psi - \hat{H}\psi = 0 \ , \tag{3.67b}$$

which are seen to be the time-dependent Schrödinger equation and its complex conjugate expression. Thus, we have established the equivalence between the solutions of eq. (3.63) and those of eq. (3.60). The condition of the vanishing of $\delta\psi$ at the time end-points, $\delta\psi(\underset{\sim}{\tau},t_1) = \delta\psi(\underset{\sim}{\tau},t_2) = 0$, has been used in obtaining the above result. If this condition is relaxed, the variation in W_{12} is

$$\delta W_{12}\Big|_{\phi=\psi} = \int_{t_1}^{t_2} dt \int d\underset{\sim}{\tau}\,(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\,\delta\psi + \left[\frac{i\hbar}{2}\int d\underset{\sim}{\tau}\,\psi^*\delta\psi\right]\Bigg|_{t_1}^{t_2} + \text{c.c.} \qquad (3.68)$$

A comparison of eq. (3.68) with eq. (3.66) suggests that the principle of stationary action, eq. (3.63), can be replaced by the more general principle stated in eq. (3.69)

$$\delta W_{12}\Big|_{\phi=\psi} = \left[\frac{i\hbar}{2}\int d\underset{\sim}{\tau}\,\psi^*\delta\psi\right]\Bigg|_{t_1}^{t_2} + \text{c.c.} \qquad (3.69)$$

Indeed, if eq. (3.69) is to be satisfied by any variation, $\delta\psi$, of the function ψ, without any restriction on its value at the time end-points, it must be satisfied by those particular variations which vanish at t_1 and t_2. Thus, the Euler equations (3.67) are again obtained. The results of eqs. (3.68) and (3.69) are also obtained when the condition of fixed time end-points is relaxed. The variations in the time end-points, denoted $\delta t_1, \delta t_2$, then contribute an additional term, $\delta^{(a)}W_{12}$, to the variation in W_{12}. This additional term is given by

$$\delta^{(a)}W_{12}\Big|_{\phi=\psi} = \Sigma[R^3,\psi(t)]\,\delta t\,\Bigg|_{t_1}^{t_2} . \qquad (3.70)$$

We verify that this term vanishes identically as a consequence of the Schrödinger equation, and the definition of $\Sigma[R^3,\psi(t)]$, eqs. (3.61), (3.62a). Indeed, using the identity

$$\tfrac{1}{2}\nabla_i\phi^*\cdot\nabla_i\phi = \tfrac{1}{2}\{-\tfrac{1}{2}\phi^*\nabla_i^2\phi + \text{c.c.}\} + \tfrac{1}{4}\nabla_i^2(\phi^*\phi)\,, \qquad (3.71)$$

eqs. (3.62a), (3.61) gives

$$\mathcal{L}[R^3,\dot{\psi}(t)] = \tfrac{1}{2}\int d\underset{\sim}{\tau}\{(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\psi + \text{c.c.}\} + (-\frac{\hbar^2}{4m})\int d\underset{\sim}{\tau}\sum_i\nabla_i^2(\psi^*\psi) = 0\,, \qquad (3.7$$

since the state function, ψ, and its first and second derivatives vanish at the spatial boundary of the total system.

In this general case, where the time end-points are varied, we may define the total variation in ψ by

$$\bar{\delta}\psi = \delta\psi + \dot{\psi}\delta t\,, \qquad (3.73)$$

and rewrite eq. (3.70) as follows[62]

$$\delta W_{12}\Big|_{\phi=\psi} = (\frac{i\hbar}{2}\int d\underset{\sim}{\tau}\psi^*\bar{\delta}\psi)\Big|_{t_1}^{t_2} + \text{c.c.} - \{\frac{i\hbar}{2}\int d\underset{\sim}{\tau}(\psi^*\dot{\psi} - \dot{\psi}^*\psi) - \mathcal{L}[R^3,\psi]\}\delta t\,\Big|_{t_1}^{t_2}$$

$$= \{\frac{i\hbar}{2}\int d\underset{\sim}{\tau}\psi^*\bar{\delta}\psi + \text{c.c.} - (\int d\underset{\sim}{\tau}\psi^*\hat{H}\psi)\,\delta t\}\Big|_{t_1}^{t_2}\,. \qquad (3.74)$$

In this form, the variational principle presents an analogy with a result in classical mechanics.[57,73] In Appendix II, we review Hamilton's principle and its generalization in classical mechanics. Therein, we have arrived at eq. (AII.16), which states that the change in the classical action integral

$$W_{12}^{Cl} = \int_{t_1}^{t_2} dt\,L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t)\,, \qquad (3.75)$$

ensuing from general variations, $\delta\underset{\sim}{q}$, $\delta\dot{\underset{\sim}{q}}$ and δt, of the variables of the Lagrangian $L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t)$, equals the difference in the values assumed by the generator, F, of an associated infinitesimal canonical transformation, at the time end-points, t_1, t_2. Thus

$$\delta W_{12}^{Cl} = F(\underset{\sim}{p}, \underset{\sim}{q}, t)\Big|_{t_1}^{t_2} , \qquad (3.76a)$$

where

$$\underset{\sim}{p} = \nabla_{\underset{\sim}{\dot{q}}} L(\underset{\sim}{q}, \underset{\sim}{\dot{q}}, t) , \qquad (3.76b)$$

$$F = \underset{\sim}{p}\bar{\delta}\underset{\sim}{q} - H(\underset{\sim}{p}, \underset{\sim}{q})\delta t, \quad \bar{\delta}\underset{\sim}{q} = \delta\underset{\sim}{q} + \underset{\sim}{\dot{q}}\delta t , \qquad (3.76c)$$

and $H(\underset{\sim}{p}, \underset{\sim}{q})$ is the Hamiltonian of the system, constructed from the generalized coordinates $\underset{\sim}{q}$ and their conjugate momenta $\underset{\sim}{p}$. The canonical transformation described by F has the effect of simultaneously changing the time parameter, t, of an instantaneous state of the system, by δt, and the instantaneous coordinates of the system by $\delta\underset{\sim}{q}$, (cf. eqs. (AII.18)). Taken separately, each term on the r.h.s. of the definition of F, eq. (3.76c), is the generator of a separate infinitesimal canonical transformation:[57] $\underset{\sim}{p}\cdot\bar{\delta}\underset{\sim}{q}$ generates a spatial translation by the infinitesimal vector $\bar{\delta}\underset{\sim}{q}$, whereas $H\delta t$ generates a displacement in time by the infinitesimal δt. We note that the variational result of eq. (3.74), which refers to a total quantum molecular system, is of the same form as eq. (3.76a), with

$$F = \frac{i\hbar}{2}\int d\underset{\sim}{\tau}\,\psi^*\bar{\delta}\psi + \text{c.c.} - \left(\int d\underset{\sim}{\tau}\,\psi^*\hat{H}\psi\right)\delta t . \qquad (3.77)$$

The correspondence between classical dynamical quantities and quantum mechanical averages is thus obtained, in comparing eq. (3.77) with eq. (3.76c). Corresponding to the classical Hamiltonian, $H(\underset{\sim}{p},\underset{\sim}{q})$, of eq. (3.76c), is the average of the Hamiltonian operator, $\hat{H}$, of the molecular system. When $\delta\psi = i\hat{A}\psi$, and $\hat{A}$ is the hermitian, the first term on the r.h.s. of eq. (3.77) is simply the average of the operator $\hat{A}$, which, in this case, is also the generator of an infinitesimal unitary transformation. Since $\hat{H}$ is the generator of time evolution in quantum mechanics, the quantity $F(t)$ represents the quantum mechanical average of the generator of an infinitesimal unitary transformation. Thus, the identification of eq. (3.74) as the quantal counterpart of eq. (3.76a) establishes the well known correspondence between unitary transformations in quantum mechanics and canonical transformations in classical mechanics. This correspondence is also obtained as a consequence of Schwinger's quantum action principle and its corollary, the principle of stationary action.[64] Schwinger's ideas are reviewed in Appendix III. In coordinate representation, the quantum action principle states that the transformation function $\langle\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle$, which connects the coordinate representation basis vectors $|\underset{\sim}{q}',t_1\rangle$, $|\underset{\sim}{q}'',t_2\rangle$, at the times t_1,t_2, is variationally characterized by[64]

$$\delta\langle\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle = \frac{i}{\hbar}\langle\underset{\sim}{q}',t_1|\delta\hat{W}_{12}^{(q)}|\underset{\sim}{q}'',t_2\rangle \quad , \qquad (3.78)$$

where the quantum action integral $\hat{W}_{12}^{(q)}$ is constructed in exactly the same manner as the classical action integral, W_{12}^{Cl} (cf. eq. (3.75)), from the quantized dynamical variables $\underset{\sim}{q}(t),\underset{\sim}{p}(t)$.

Thus,

$$\hat{W}_{12}^{(q)} = \int_{t_2}^{t_1} \hat{L}^{(q)} dt, \qquad (3.79a)$$

where the quantum Lagrangian $\hat{L}^{(q)}$ is defined through

$$\hat{L}^{(q)} dt = \hat{p} \cdot d\hat{q} - \hat{H} dt \ . \qquad (3.79b)$$

When the variations involved in eq. (3.78) do not alter the structure of the Hamiltonian of the system, the ensuing change in the action integral, $\hat{W}_{12}^{(q)}$ is further given by the difference in the values, at t_1 and t_2, of the generator, $\hat{F}(t)$, of the infinitesimal unitary transformation which is responsible for the variations considered

$$\delta \hat{W}_{12}^{(q)} = \hat{F}(t_1) - \hat{F}(t_2) \ . \qquad (3.80)$$

This last result, which is the mathematical statement of the principle of stationary quantum action, is exactly of the same form as the classical result, eq. (3.76a). As shown and discussed in Appendix III, eqs. (3.78) and (3.80) can be obtained from general considerations of unitary transformations in quantum theory. On the other hand, assuming their validity as generating principles, these equations lead to the variational derivations of the fundamental laws of quantum mechanics. In particular, through the principle of stationary action, eq. (3.80), one can identify the generator $\hat{F}(t)$ in the explicit expression of $\delta \hat{W}_{12}^{(q)}$, which is obtained by applying the calculus of variation to $\hat{W}_{12}^{(\underset{\sim}{q})}$. Subjecting $\hat{W}_{12}^{(\underset{\sim}{q})}$, given by eqs. (3.79), to general variations in $\hat{\underset{\sim}{q}}(t)$, $\hat{\underset{\sim}{p}}(t)$ and the time end-points, t_1, t_2, one arrives at the following identification of $\hat{F}(t)$

$$\hat{F}(t) = \hat{\underset{\sim}{p}}(t)\cdot\delta\underset{\sim}{q}(t) - \hat{H}(t)\delta t \ , \tag{3.81}$$

which is to be compared with eq. (3.76c). In eq. (3.81), the variations, $\delta\underset{\sim}{q}(t)$, are considered to be c-numbers, i.e., they commute with $\hat{\underset{\sim}{q}}(t)$ and $\hat{\underset{\sim}{p}}(t)$. Considering the particular case where $\delta\underset{\sim}{q}(t) = 0$, at all t, in eq. (3.81), we recognize the role of the Hamiltonian, $\hat{H}$, as the generator of the time-development of the system, thereby obtaining the Heisenberg equations of motion

$$\nabla_{\underset{\sim}{p}}\hat{H} = \frac{d\hat{\underset{\sim}{q}}(t)}{dt} = \frac{1}{i\hbar}[\hat{\underset{\sim}{q}}(t),\hat{H}] \ , \tag{3.82a}$$

$$-\nabla_{\underset{\sim}{q}}\hat{H} = \frac{d\hat{\underset{\sim}{p}}(t)}{dt} = \frac{1}{i\hbar}[\hat{\underset{\sim}{p}}(t),\hat{H}] \ . \tag{3.82b}$$

We note that in eqs. (3.82), only the second set of equality signs are obtained as the consequences of the identification of eq. (3.81). The first identities, which give the equations of motion for $\hat{\underset{\sim}{q}}$ and $\hat{\underset{\sim}{p}}$ in the form of quantized Hamilton's equations (cf. eqs. (AII.6) of Appendix II) constitute the Euler equations of the variational problem, eq. (3.80), together with the result

$$\frac{d\hat{H}}{dt} = \frac{\partial\hat{H}}{\partial t}$$

(See Appendix III, eqs. (AIII.28), (AIII.29).) Finally, setting $\delta t = 0$ in eq. (3.81), we obtain the generator for a purely spatial variation, $\delta \underset{\sim}{q}$. Applying the general formalism of unitary transformations to this generator, one can derive the commutation rules governing the dynamical variables, $\underset{\sim}{\hat{q}}$, $\underset{\sim}{\hat{p}}$.[64]

Since the principle of stationary action leads to the identification of $\hat{H}$ as the generator of time development, Schrödinger's equations (3.67) can also be regarded as a consequence of Schwinger's quantum action principle. On the other hand, we have derived these equations as the Euler equations of the variational problem stated in eqs. (3.74), (3.69). Henceforth, the results of eqs. (3.69), (3.74) will be referred to as the Schrödinger variational principle. We now show that Schrödinger's variational principle follows from the Schwinger quantum action principle. In view of the derivation of the latter in Appendix III, the anticipated result will automatically establish the equivalence between the two principles. To fix the ideas, we first let the generalized coordinates $\underset{\sim}{q}$ be the 3n electronic coordinates of a molecular system. As a further simplification, we assume that the Hamiltonian, $\hat{H}$, is spin-independent, so that we can ignore the spin in the following derivations. Consider a state vector, $|\psi,t\rangle$, of the system, and the transformation function $\langle\psi,t_1|\psi,t_2\rangle$, where the time end-points t_1,t_2 are arbitrary.

Using the completeness relation[22]

$$\int d\underset{\sim}{q}'|\underset{\sim}{q}',t_1><\underset{\sim}{q}',t_1| = \hat{1} , \qquad (3.83)$$

at t_1 and t_2, we can write

$$<\psi,t_1|\psi,t_2> = \int d\underset{\sim}{q}'d\underset{\sim}{q}''<\psi,t_1|\underset{\sim}{q}',t_1><\underset{\sim}{q}',t_2|\underset{\sim}{q}'',t_2><\underset{\sim}{q}'',t_2|\psi,t_2> , \qquad (3.84)$$

which amounts to expanding the state vector at t_1 and t_2 in terms of the coordinates eigenstates at the same time:

$$|\psi,t_1> = \int d\underset{\sim}{q}'\psi(\underset{\sim}{q}',t_1)|\underset{\sim}{q}',t_1> , \qquad (3.85a)$$

$$|\psi,t_2> = \int d\underset{\sim}{q}''\psi(\underset{\sim}{q}'',t_2)|\underset{\sim}{q}'',t_2> \cdot \qquad (3.85b)$$

We now consider the change in $<\psi,t_1|\psi,t_2>$ ensuing from arbitrary variations in the coordinates eigenvectors, keeping the expansion coefficients $\psi(\underset{\sim}{q}',t_1)$, $\psi(\underset{\sim}{q}'',t_2)$ fixed. By eq. (3.84), this change is given by

$$\delta^{(\underset{\sim}{q})}<\psi,t_1|\psi,t_2> = \int d\underset{\sim}{q}'d\underset{\sim}{q}''<\psi,t_1|q',t_2>\delta(<\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2>)<q'',t_2|\psi,t_2> \cdot (3.86)$$

Using eq. (3.78), and the completeness relation, eq. (3.83), we find

$$\delta^{(\underset{\sim}{q})}<\psi,t_1|\psi,t_2> = \frac{i}{\hbar}<\psi,t_1|\delta\hat{W}_{12}^{(\underset{\sim}{q})}|\psi,t_2> , \qquad (3.87a)$$

where the action integral $\hat{W}_{12}^{(\underset{\sim}{q})}$ is defined in eqs. (3.79). As indicated by the superscript $(\underset{\sim}{q})$, the above variational result, eq. (3.87a), is obtained when the state vector $|\psi,t>$ is expanded in terms of the basis vectors of the coordinates representation, which then are the only objects of the variation.

The variation in $|\psi,t\rangle$ is thus representation dependent, and so is the change in $\langle\psi,t_1|\psi,t_2\rangle$. We may enlarge the class of admissible variations in $|\psi,t\rangle$ by including changes in the momentum representation. In Appendix III, we note that in this representation, Schwinger's quantum action principle also holds, and reads

$$\delta\langle \underset{\sim}{p}',t_1|\underset{\sim}{p}'',t_2\rangle = \frac{i}{\hbar}\langle \underset{\sim}{p}',t_1|\delta\hat{W}_{12}^{(\underset{\sim}{p})}|\underset{\sim}{p}'',t_2\rangle \;, \tag{3.88a}$$

where

$$\hat{W}_{12}^{(p)} = \int_2^1 [-d\hat{\underset{\sim}{p}}\cdot\hat{\underset{\sim}{q}} - \hat{H}dt] \;. \tag{3.88b}$$

Following the same procedure as above, which has led to eq. (3.87a), with the coordinates eigenvectors $|\underset{\sim}{q}',t\rangle$ replaced by the momentum eigenvectors $|\underset{\sim}{p}',t\rangle$, we obtain

$$\delta^{(\underset{\sim}{p})}\langle\psi,t_1|\psi,t_2\rangle = \frac{i}{\hbar}\langle\psi,t_1|\delta\hat{W}_{12}^{(p)}|\psi,t_2\rangle \;; \tag{3.87b}$$

where $\delta^{(\underset{\sim}{p})}$ denotes a change which is wholly accounted for by variations in the basis vectors of the momentum representation. A general, arbitrary variation in $|\psi,t\rangle$ may be one of the type $\delta^{(\underset{\sim}{q})}$, or one of the type $\delta^{(\underset{\sim}{p})}$, or a combination of both types. Thus, we may write

$$\delta\langle\psi,t_1|\psi,t_2\rangle = \frac{i}{\hbar}\langle\psi,t_1|\delta\hat{W}_{12}|\psi,t_2\rangle \;, \tag{3.87c}$$

where $\hat{W}_{12}$ is a linear combination of $\hat{W}_{12}^{(q)}$ and $\hat{W}_{12}^{(p)}$. For variations which leave the Hamiltonian invariant in its structure, the principle of stationary action applies to both $\hat{W}_{12}^{(q)}$ and $\hat{W}_{12}^{(p)}$, hence to $\hat{W}_{12}$. Thus

$$\delta\hat{W}_{12} = \hat{A}(t_1) - \hat{A}(t_2) \tag{3.89}$$

where $\hat{A}(t)$ is the generator of a unitary transformation which is responsible for the variations at time t. We now consider the case $t_1 = t$, $t_2 = t + dt$, where t is arbitrary, and dt is an infinitesimal time interval. By virtue of eq. (3.87c),

$$\delta\langle\psi,t|\psi,t+dt\rangle = \frac{i}{\hbar}\langle\psi,t|\delta\hat{W}_{t,t+dt}|\psi,t+dt\rangle \cdot \tag{3.90}$$

Since the variations involved in Schrödinger's variational principle do not alter the structure of the Hamiltonian, eq. (3.89) is applicable and gives

$$\delta\langle\psi,t|\psi,t+dt\rangle = -\frac{i}{\hbar}\langle\psi,t|\hat{A}(t+dt) - \hat{A}(t)|\psi,t+dt\rangle \cdot \tag{3.91}$$

On the l.h.s. of eq. (3.91), we can write

$$|\psi,t+dt\rangle = |\psi,t\rangle + dt\,\frac{\partial|\psi,t\rangle}{\partial t}\ , \tag{3.92a}$$

and use the result

$$\delta\langle\psi,t|\psi,t\rangle = 0 = \delta\langle\psi,t+dt|\psi,t+dt\rangle = \delta\langle\psi,t|\hat{H}dt|\psi,t+dt\rangle\ , \tag{3.92b}$$

to obtain

$$\delta\langle\psi,t|(i\hbar\frac{\partial}{\partial t} - \hat{H})dt|\psi,t\rangle = \langle\psi,t|\frac{d\hat{A}(t)}{dt}dt|\psi,t\rangle \tag{3.93}$$

to first order in all infinitesimal quantities. The first two identities of eq. (3.92b), which state the vanishing of the variations of $<\psi,t|\psi,t>$, $<\psi,t+dt|\psi,t+dt>$, follow directly from the quantum action principle, eqs. (3.87c) and (3.89). The last identity of eq. (3.92b) implicitly makes use of the interpretation of $\hat{H}$ as the time-development generator,[22] or, equivalently, of the Schrödinger equation. Using the Heisenberg equation of motion for $\hat{A}(t)$,

$$\frac{d\hat{A}(t)}{dt} = \frac{1}{i\hbar}[\hat{A},\hat{H}] + \frac{\partial\hat{A}}{\partial t} \quad ,$$

we rewrite eq. (3.93) as

$$\begin{aligned}\delta<\psi,t|(i\hbar\frac{\partial}{\partial t} - \hat{H})dt|\psi,t> &= <\psi,t|\frac{1}{i\hbar}[\hat{A},\hat{H}] + \frac{\partial\hat{A}}{\partial t}|\psi,t>dt \\ &= \frac{d}{dt}(<\psi,t|\hat{A}(t)|\psi,t>)dt \; . \qquad (3.94)\end{aligned}$$

Adding the complex conjugate expressions to both sides of eq. (3.94), and noting that $\delta(dt) \to 0$ when $dt \to 0$, we obtain

$$\begin{aligned}\delta[\tfrac{1}{2}<\psi,t|(i\hbar\frac{\partial}{\partial t} - \hat{H})|\psi,t> + c.c.] &= \frac{d}{dt}(\tfrac{1}{2}<\psi,t|\hat{A}(t)|\psi,t> + c.c.) \\ &= \tfrac{1}{2}<\{\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\}> + c.c. \qquad (3.95)\end{aligned}$$

Using eq. (3.71), we verify that the varied expression on the l.h.s. of eq. (3.95) is equal to $\mathcal{L}[R^3,\psi]$:

$$\begin{aligned}\tfrac{1}{2}<\psi,t|(i\hbar\frac{\partial}{\partial t} - \hat{H})|\psi,t> + c.c. &= \int d\underset{\sim}{\tau}\{\frac{i\hbar}{2}(\psi^*\dot{\psi} - \dot{\psi}^*\psi) - \tfrac{1}{2}(\psi^*\hat{H}\psi + \hat{H}\psi^*\psi)\} \\ &= \int d\underset{\sim}{\tau}\{\frac{i\hbar}{2}(\psi^*\dot{\psi} - \dot{\psi}^*\psi) - \frac{\hbar^2}{2m}\sum_i \nabla_i\psi^*\cdot\nabla_i\psi \\ &\quad - \hat{V}\psi^*\psi\} \equiv \mathcal{L}[R^3,\psi(t)] \; . \qquad (3.96)\end{aligned}$$

On the other hand, the r.h.s. of eq. (3.95) is simply the time derivative of the average of the operator $\hat{A}(t)$; this term can also be written as the average of the operator $\{\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\}$, which in the Heisenberg representation is the time derivative of the generator $\hat{A}$. Thus, the use of the Schwinger quantum action principle and the principle of stationary action leads to the following equation (3.97), which variationally characterizes the total system Lagrangian $\mathcal{L}[R^3,\psi]$:[62,70]

$$\delta\mathcal{L}[R^3,\psi] = \frac{d}{dt}\{\tfrac{1}{2}\langle\hat{A}(t)\rangle + c.c.\}$$

$$= \tfrac{1}{2}\langle\{\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\}\rangle + c.c. \qquad (3.97)$$

Since $\hat{A}$ is the generator of the unitary transformation responsible for the variations in eq. (3.97), we have

$$-\frac{i}{\hbar}\hat{A}(t)|\psi,t\rangle = \delta|\psi,t\rangle \quad ,$$

so that eq. (3.95) can also be written as

$$\delta\mathcal{L}[R^3,\psi] = \frac{d}{dt}\{\frac{i\hbar}{2}\int\psi^*\delta\psi d\tau + c.c.\} \cdot \qquad (3.98)$$

Integrating eq. (3.98) over the time, from t_1 to t_2, we obtain

$$\int_{t_1}^{t_2}[\delta\mathcal{L}[R^3,\psi]]dt = \{\frac{i\hbar}{2}\int\psi^*\delta\psi d\tau + c.c.\}\Big|_{t_1}^{t_2} . \qquad (3.99)$$

Adding to both sides of eq. (3.99), the terms $\mathcal{L}[R^3,\psi(t)]\delta t\Big|_{t_1}^{t_2}$, which account for the variations in the time end-points t_1,t_2, and which vanish identically, as shown in eq. (3.72), we derive the Schrödinger variational principle, eqs. (3.69), and (3.74). We note that eq. (3.97) also follows from eq. (3.69) upon the replacements $t_1 = t$, $t_2 = t + dt$, and $\delta\psi = -\frac{i}{\hbar}\hat{A}\psi$.

The above derivations clearly demonstrate the equivalence of Schwinger's quantum action principle and Schrödinger's variational principle. The latter can be expressed either in terms of the action integral, eqs. (3.69) and (3.74), or in terms of the Lagrangian, eq. (3.97). Both these functionals, $W_{12}[R^3,\phi(t)]$ and $\mathcal{L}[R^3,\phi(t)]$, refer to the total system as they both involve the all-space integral of the Lagrangian density $L[\phi(t);\underset{\sim}{r}]$ of eq. (3.61). Integrating $L[\phi;\underset{\sim}{r}]$ over the volume $\Omega(t)$ occupied by a subsystem at time t, we obtain the subsystem Lagrangian $\mathcal{L}[\Omega,\phi(t)]$,

$$\mathcal{L}[\Omega,\phi(t)] = \int_{\Omega(t)} d\underset{\sim}{r}L[\phi(t);\underset{\sim}{r}] \ , \qquad (3.100a)$$

from which an action integral, $W_{12}[\Omega,\phi]$, associated with the subsystem can be constructed, in the same manner as $W_{12}[R^3,\phi]$ was constructed from $\mathcal{L}[R^3,\phi(t)]$, i.e.,

$$W_{12}[\Omega,\phi] = \int_{t_1}^{t_2} dt\mathcal{L}[\Omega,\phi(t)] \ . \qquad (3.100b)$$

Explicitly, we have[62,70]

$$\mathcal{L}[\Omega,\phi(t)] = \int_{\Omega(t)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\frac{i\hbar}{2}(\phi^*\dot{\phi} - \dot{\phi}^*\phi) - \frac{\hbar^2}{2m}\sum_i \nabla_i\phi^*\cdot\nabla_i\phi - \hat{v}\phi^*\phi\} \ , \qquad (3.101a$$

and

$$W_{12}[\Omega,\phi] = \int_{t_1}^{t_2} dt \int_{\Omega(t)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\frac{i\hbar}{2}(\phi^*\dot{\phi} - \dot{\phi}^*\phi) - \frac{\hbar^2}{2m}\sum_i\nabla_i\phi^*\cdot\nabla_i\phi - \hat{v}\phi^*\phi\} \ . \qquad (3.101b$$

Using eq. (3.71), we can also write $\mathcal{L}[\Omega,\phi]$ in the form

$$\mathcal{L}[\Omega,\phi(t)] = \tfrac{1}{2}\int_{\Omega(t)} d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\phi^*(i\hbar\tfrac{\partial}{\partial t} - \hat{H})\phi + c.c.\} - \frac{\hbar^2}{4m}\oint dS(t)\nabla\rho_\phi'\cdot\underset{\sim}{n} \ , \qquad (3.102$$

where

$$\rho_\phi'(\underset{\sim}{r},t) = \frac{1}{N}\rho_\phi(\underset{\sim}{r},t) = \int' d\underset{\sim}{\tau}'\phi^*(\underset{\sim}{\tau},t)\phi(\underset{\sim}{\tau},t) \ .$$

Through eq. (3.102), it is seen that when $\phi(t) = \psi(t)$, the subsystem Lagrangian vanishes as a consequence of the Schrödinger equations (3.67), and the zero-flux boundary condition, eq. (3.1), which defines the subsystem at any time t.

We now consider the variational properties which the subsystem functionals, $W_{12}[\Omega,\phi]$ and $\mathcal{L}[\Omega,\phi(t)]$, exhibit as a consequence of the Schrödinger variational principle. Thus, we consider the changes in $W_{12}[\Omega,\phi]$ and $\mathcal{L}[\Omega,\phi(t)]$ resulting from the variations in the trial function $\phi(t)$ about the state function $\psi(t)$. Since the boundary of the subsystem is defined in terms of ρ, it will also be subjected to the variations.[62] Thus, as in section A, we require the trial function ϕ to satisfy sufficient analytical conditions so as to ensure the existence of a surface of zero-flux in $\nabla\rho_\phi(\underset{\sim}{r},t)$, at all times t. This surface defines a region $\Omega(\phi(t))$ which, at any time t, is continuously deformable into the region $\Omega(t) = \Omega(\psi(t))$ associated with the subsystem. In other words, we impose the constraint[62,65,70]

$$\delta(\int_\Omega \nabla^2\rho d\underset{\sim}{r}) = 0 \ , \qquad (3.103a)$$

at all times between t_1 and t_2, along with the condition

$$\int_\Omega \nabla^2 \rho d\underset{\sim}{r} = 0 \cdot \qquad (3.103b)$$

Explicitly, the constraint of eq. (3.103a) reads

$$\oint dS \int' d\tau' \{(\nabla\psi^*)\delta\psi + \psi^*\nabla(\delta\psi) + \nabla(\delta\psi^*)\psi + \delta\psi^*\nabla\psi\}\cdot\underset{\sim}{n} + \oint dS\delta S\nabla^2\rho' = 0. \quad (3.104$$

The change in $W_{12}[\Omega,\psi]$ resulting from the variations, $\delta\psi$, in ψ, and δt_i, in the time end-points t_1, t_2, is evaluated by directly applying the calculus of variations[63] to eq. (3.101b).

We thus find

$$\begin{aligned}
\delta W_{12}[\Omega,\phi]\Big|_{\phi=\psi} = &\{\int_{t_1}^{t_2} dt \int_\Omega d\underset{\sim}{r} \int' d\underset{\sim}{\tau}'(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\delta\psi + \text{c.c.}\} \\
&+ \{\int_{t_1}^{t_2} dt \oint dS \int' d\underset{\sim}{\tau}'\}-\frac{\hbar^2}{2m})\nabla\psi^*\cdot\underset{\sim}{n}\delta\psi + \text{c.c.}\} \\
&+ \int_{t_1}^{t_2} dt \oint dS \int' d\underset{\sim}{\tau}'\{\tfrac{1}{2}(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\psi + \text{c.c.}\}\delta S \\
&- \frac{\hbar^2}{4m}\int_{t_1}^{t_2} dt \oint dS\delta S(\nabla^2\rho') + \int_{t_1}^{t_2} dt \oint dS \int' d\underset{\sim}{\tau}' \frac{\partial S}{\partial t}\{-\frac{i\hbar}{2}\psi^*\delta\psi + \text{c.c.}\} \\
&+ \{\int_\Omega d\underset{\sim}{r} \int' d\underset{\sim}{\tau}'(\frac{i\hbar}{2})\psi^*\delta\psi + \text{c.c.} + \mathcal{L}[\Omega,\psi]\delta t\}\Big|_{t_1}^{t_2} .
\end{aligned}$$

Using the result of eq. (3.104), we obtain

$$\begin{aligned}
\delta W_{12}[\Omega,\phi]\Big|_{\phi=\psi} = &\{\int_{t_1}^{t_2} dt \int_\Omega d\underset{\sim}{r} \int' d\underset{\sim}{\tau}'(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\delta\psi + \text{c.c.}\} \\
&+ \int_{t_1}^{t_2} dt \oint dS \int' d\underset{\sim}{\tau}'\{\tfrac{1}{2}(-i\hbar\dot{\psi}^* - \hat{H}\psi^*)\psi + \text{c.c.}\}\delta S
\end{aligned}$$

$$+\left(-\frac{h^2}{4m}\right)\left\{\int_{t_1}^{t_2}dt\oint dS\int' d\underset{\sim}{\tau}'[(\nabla\psi^*)\delta\psi - \psi^*\nabla(\delta\psi)]\cdot\underset{\sim}{n} + \text{c.c.}\right\}$$

$$+\left\{-\frac{i\hbar}{2}\int_{t_1}^{t_2}dt\oint dS\int' d\underset{\sim}{\tau}'\ \frac{\partial S}{\partial t}\psi^*\delta\psi + \text{c.c.}\right\}$$

$$+\left\{\frac{i\hbar}{2}\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*\delta\psi + \text{c.c.} + \mathcal{L}[\Omega,\psi]\delta t\right\}\Bigg|_{t_1}^{t_2} . \qquad (3.105)$$

Since ψ and ψ^* satisfy the Schrödinger equations (3.67), the first two terms on the r.h.s. of eq. (3.105) vanish, as well as the terms containing $\mathcal{L}[\Omega,\psi]\delta t$ at the time end-points. Thus,

$$\delta W_{12}[\Omega,\phi]\Big|_{\phi=\psi} = \left(-\frac{\hbar^2}{4m}\right)\left\{\int_{t_1}^{t_2}dt\oint dS\int' d\underset{\sim}{\tau}'[(\nabla\psi^*)\delta\psi - \psi^*\nabla(\delta\psi)]\cdot\underset{\sim}{n} + \text{c.c.}\right\}$$

$$+\left\{-\frac{i\hbar}{2}\int_{t_1}^{t_2}dt\oint dS\int' d\underset{\sim}{\tau}'\ \frac{\partial S}{\partial t}\psi^*\delta\psi + \text{c.c.}\right\}$$

$$+\left\{\frac{i\hbar}{2}\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\psi^*\delta\psi + \text{c.c.}\right\}\Bigg|_{t_1}^{t_2} . \qquad (3.106)$$

In the above derivation of eq. (3.106), we have used the Schrödinger variational principle via its Euler equations (3.67). The result stated in eq. (3.106) can also be considered a generalization of Schrödinger's variational principle. Indeed, requiring eq. (3.106) to be satisfied for any subsystem bounded by a zero-flux surface, we obtain eqs. (3.69) and (3.74), since $\Omega = R^3$ can be regarded as constituting such a subsystem. The Schrödinger variational principle, generalized to refer to a subsystem or to a total system, thus states that the action integral $W_{12}[\Omega,\phi]$ is stationary, with respect to variations about a state function $\phi = \psi$, in the sense that the change in

$W_{12}[\Omega,\phi]$ resulting from these variations contains only contributions from its spatial and temporal boundaries.[65] Making the substitution $\delta\psi = -\frac{i}{\hbar}\hat{A}\psi$ on the r.h.s. of eq. (3.102), we find

$$\delta W_{12}[\Omega,\phi]\Big|_{\phi=\psi} = \frac{1}{N}\{\tfrac{1}{2}\int_{t_1}^{t_2}dt\oint dS\underset{\sim}{J}_A(\underset{\sim}{r})\cdot\underset{\sim}{n} + \text{c.c.} - \int_{t_1}^{t_2}dt\oint dS\,\frac{\partial S}{\partial t}\rho_A(\underset{\sim}{r})$$

$$+ \int_\Omega d\underset{\sim}{r}\rho_A(\underset{\sim}{r})\Big|_{t_1}^{t_2} \quad , \qquad (3.107a)$$

where

$$\rho_A(\underset{\sim}{r}) = \frac{N}{2}\{\int' d\underset{\sim}{\tau}'\psi^*\hat{A}\psi + \text{c.c.}\} \quad , \qquad (3.107b)$$

and

$$\underset{\sim}{J}_A(\underset{\sim}{r}) = \frac{N\hbar}{2mi}\int' d\underset{\sim}{\tau}'[\psi^*\nabla(\hat{A}\psi) - (\nabla\psi^*)\hat{A}\psi]\cdot\underset{\sim}{n} \quad , \qquad (3.107c)$$

are, respectively, the density and vector current density associated with the property described by the generator $\hat{A}$ (cf. eqs. (3.21) and (3.22)). In the form given in eq. (3.107a), the variation in $\delta W_{12}[\Omega,\phi]$ evaluated at $\phi = \psi$ is seen to be governed by a change in the generator $\hat{A}$, due to the net flux of its vector current density across the spatial boundary of the subsystem and the change in this boundary with time, from t_1 to t_2, in addition to the difference in the values, at t_1 and t_2, of the subsystem average of $\hat{A}$.[65] We may write the latter in the form given by the following eq. (3.108),

$$\int_\Omega d\underset{\sim}{r}\rho_A(\underset{\sim}{r})\Big|_{t_1}^{t_2} = \int_{t_1}^{t_2}dt\,\frac{d}{dt}\int_\Omega d\underset{\sim}{r}\rho_A(\underset{\sim}{r}) = -\tfrac{1}{2}\int_{t_1}^{t_2}dt\oint dS\underset{\sim}{J}_A\cdot\underset{\sim}{n} + \text{c.c.} + \int_{t_1}^{t_2}dt\oint dS\frac{\partial S}{\partial t}\rho_A(\underset{\sim}{r})$$

$$+\frac{N}{2}\int_{t_1}^{t_2}dt\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\,\psi^*\left(\frac{i}{\hbar}[\hat{H},\hat{A}]\right)\psi + \text{c.c.}$$

$$+\frac{N}{2}\int_{t_1}^{t_2}dt\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\,\psi^*\,\frac{\partial\hat{A}}{\partial t}\psi + \text{c.c.}\quad, \tag{3.108}$$

and reduce eq. (3.107a) to

$$\delta W_{12}[\Omega,\phi]\Big|_{\phi=\psi} = \tfrac{1}{2}\Big\{\int_{t_1}^{t_2}dt\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\,\psi^*\left(\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\right)\psi + \text{c.c.}\Big\}$$

$$= \frac{1}{N}\int_{t_1}^{t_2}dt\int_{\Omega}d\underset{\sim}{r}\,\rho_{\dot{A}}(\underset{\sim}{r})\ . \tag{3.109a}$$

The density, $\rho_{\dot{A}}(\underset{\sim}{r})$, associated with the time derivative of $\hat{A}$ is defined analogously with $\rho_{\hat{A}}$, eq. (3.107b), i.e.,

$$\rho_{\dot{A}}(\underset{\sim}{r}) = \frac{N}{2}\Big\{\int' d\underset{\sim}{\tau}'\,\psi^*\dot{\hat{A}}\psi + \text{c.c.}\Big\}$$

$$= \frac{N}{2}\Big\{\int' d\underset{\sim}{\tau}'\,\psi^*\left(\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\right)\psi + \text{c.c.}\Big\}\quad, \tag{3.109b}$$

where the Heisenberg equation of motion for $\hat{A}$ has been used. Consider the case $t_1 = t$, $t_2 = t + dt$, where t is arbitrary, and dt, an arbitrary infinitesimal time interval. Equation (3.109a) gives the result[62,65,70]

$$\delta\mathcal{L}[\Omega,\dot{\phi}(t)]\Big|_{\phi=\psi} = \frac{1}{N}\int_{\Omega}d\underset{\sim}{r}\,\rho_{\dot{A}}(\underset{\sim}{r})$$

$$= \tfrac{1}{2}\Big\{\Big\langle\Big(\frac{i}{\hbar}[\hat{H},\hat{A}] + \frac{\partial\hat{A}}{\partial t}\Big\rangle_{\Omega} + \text{c.c.}\Big\}\quad, \tag{3.110}$$

which constitutes the generalization of eq. (3.97). In words, eq. (3.110) states that the variation in the subsystem Lagrangian resulting from an infinitesimal unitary transformation with generator $\hat{A}$ is equal to the subsystem average of the time derivative of $\hat{A}$. This statement of the variational principle applies to a subsystem as well as to a total molecular system, in which case, $\Omega = R^3$, and eq. (3.110) is seen to be identical with eq. (3.97). Eq. (3.110) also embodies the stationary state variational principle, since in a stationary state,

$$i\hbar\dot{\psi} = \hat{H}\psi = E\psi \ ,$$

so that $\mathcal{L}[\Omega,\psi(t)] = -F_E[\Omega,\psi]$, where the energy functional F_E has been defined in the preceding section, eq. (3.12b). Defining $\varepsilon\hat{A} = -\frac{i}{\hbar}\hat{A}$, and considering a generator $\hat{A}$ which does not depend on the time explicitly, we verify that eq. (3.110) is identical with eq. (3.26a), when ψ describes a stationary state[62],

$$-\delta\mathcal{L}[\Omega,\phi(t)]\Big|_{\phi=\psi} = \delta F_E[\Omega,\psi]\Big|_{\phi=\psi} = \frac{\varepsilon}{2}\{\langle[\hat{H},\hat{A}]\rangle_\Omega + c.c.\} \ .$$

Thus, the stationary state variational principle, used in the preceding section, appears to be a particular statement of the Schrödinger variational principle, which, in its most general form, can be expressed either in terms of the action integral $W_{12}[\Omega,\phi]$, eqs. (3.106), (3.107), or in terms of the subsystem Lagrangian $\mathcal{L}[\Omega,\phi(t)]$, eq. (3.110). We have thus completed the variational characterization of a subsystem, defined as a portion of the molecular charge distribution enclosed by a zero-flux surface, for a system which is in a stationary state as well as for a time-dependent system. The variational principle demands that the change in the subsystem Lagrangian $\mathcal{L}[\Omega,\psi]$, which results from an infinitesimal unitary transformation of generator $\hat{A}$, be given by the projection over the region occupied by the subsystem of the all-space result. We now consider a few examples which will illustrate the use of the above statement of the variational principle. First, let $\hat{A} = \varepsilon\hat{1}N$, where $\hat{1}$ is the unit operator and ε an infinitesimal real number. Then, the subsystem average of $\hat{A}$, $\langle\hat{A}\rangle_\Omega$, is ε times the average electronic population of the subsystem. On the other hand, the infinitesimal unitary transformation described by the generator $\hat{A}$ is just a phase transformation which, when operating on the state function ψ, leaves the Lagrangian $\mathcal{L}[\Omega,\psi]$ invariant. This result can be obtained either by using eq. (3.110) or by directly evaluating $\delta\mathcal{L}[\Omega,\psi]$. From eq. (3.107a), it is also seen that

$$\delta\mathcal{L}[\Omega,\psi] = \tfrac{1}{2}\{\frac{d}{dt}\langle\hat{A}\rangle_\Omega + \frac{1}{n}\oint dS\,\mathbf{J}_A\cdot\mathbf{n} - \oint dS\frac{\partial S}{\partial t}\int' d\boldsymbol{\tau}'\,\psi^*\hat{A}\psi + \text{c.c.}\} \,. \quad (3.111)$$

Specializing to $\hat{A} = \varepsilon\hat{1}N$, we thus have

$$\dot{N}(\Omega) = -\oint dS\,\mathbf{J}\cdot\mathbf{n} + \oint dS\frac{\partial S}{\partial t}\rho \,, \quad (3.112a)$$

where

$$N(\Omega) = N\int_\Omega d\mathbf{r}\int' d\boldsymbol{\tau}'\,\psi^*\psi \,. \quad (3.112b)$$

Eqs. (3.112) constitute a generalization of the usual equation of continuity. This generalization describes the time evolution of the portion of the charge distribution contained in the subsystem, whose boundary also changes with time.[62] This time dependence of the spatial boundary of the subsystem is accounted for by the second term on the r.h.s. of eq. (3.112a). While this result, eq. (3.112a), can be derived in a much simpler manner, namely by directly differentiating $N(\Omega)$, eq. (3.112b), and using the Schrödinger equations (3.67), we use the present derivation, which utilizes the variational principle, to illustrate a general property observed whenever the generator $\hat{A}$ leaves the subsystem Lagrangian invariant. From eq. (3.107), it is seen that, when $\delta\mathcal{L}[\Omega,\psi] = 0$,

$$\tfrac{1}{2}\frac{d}{dt}\{\langle\hat{A}\rangle_{\Omega} + \text{c.c.}\} = \frac{1}{N}\{\tfrac{1}{2}\oint dS\,\mathbf{J}_{A}\cdot\mathbf{n} + \text{c.c.} - \oint dS\,\frac{\partial S}{\partial t}\,\rho_{A}\}\ , \qquad (3.113)$$

viz. the total change in the subsystem average of $\hat{A}$ with time is restricted to changes in the property at the boundary of the subsystem. The average value of $\hat{A}$ is independent of time only when $\Omega = \mathbf{R}^3$, or when the r.h.s. of eq. (3.113) vanishes.[62] We now consider the case $\hat{A} = \boldsymbol{\varepsilon}\cdot\hat{\mathbf{p}}$, where $\hat{\mathbf{p}}$ denotes the conjugate momenta associated with the electron whose coordinates, $\mathbf{r}$, are integrated over the subsystem in the definition of $\mathcal{L}[\Omega,\psi(t)]$. Each of the components of the vector $\boldsymbol{\varepsilon}$ is an arbitrary, infinitesimal real number, fixed for all $\mathbf{r}$. For each given $\boldsymbol{\varepsilon}$, the generator $\hat{A}$ then describes a uniform translation of the electronic coordinates, $\mathbf{r}$, by $-\boldsymbol{\varepsilon}$:

$$\underset{\sim}{r}' = \underset{\sim}{r} - \underset{\sim}{\varepsilon} .$$

Denoting by $\nabla_{\underset{\sim}{\varepsilon}}$ the gradient operator with respect to $\underset{\sim}{\varepsilon}$, the variational principle states that

$$\underset{\sim}{\varepsilon}\cdot\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}(t)]\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}} = \frac{\underset{\sim}{\varepsilon}}{2}\{<\frac{i}{\hbar}[\hat{H},\hat{\underset{\sim}{p}}]>_{\Omega} + \text{c.c.}\} . \qquad (3.114)$$

On the l.h.s. of eq. (3.114), we have introduced the notation ψ_{ε} to denote the image of the state function, ψ, in the infinitesimal unitary transformation described by $\hat{A}$:

$$\psi_{\varepsilon}(\underset{\sim}{r},\underset{\sim}{\tau}',t) = (\hat{1} - \frac{i}{\hbar}\underset{\sim}{\varepsilon}\cdot\hat{\underset{\sim}{p}})\psi(\underset{\sim}{r},\underset{\sim}{\tau}',t) = \psi(\underset{\sim}{r} - \underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t) = \psi(\underset{\sim}{r}',\underset{\sim}{\tau}',t). \quad (3.115$$

Thus, for a given $\underset{\sim}{\varepsilon}$, and to first order

$$\begin{aligned}\mathcal{L}[\Omega,\psi_{\varepsilon}] = \int_{\Omega}d\underset{\sim}{r}\int d\underset{\sim}{\tau}'\{&\frac{i\hbar}{2}(\psi^*(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t)\dot{\psi}(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t) - \dot{\psi}^*(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t)\ \psi(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t))\\ &- \frac{\hbar^2}{2m}\sum_i\nabla_i\psi^*(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t)\cdot\nabla_i\psi(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t) - \hat{v}(\underset{\sim}{r},\underset{\sim}{\tau}',t)\psi^*(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t)\\ &\psi(\underset{\sim}{r}-\underset{\sim}{\varepsilon},\underset{\sim}{\tau}',t)\} . \end{aligned} \qquad (3.116)$$

Differentiating both sides of eq. (3.116), with respect to $\underset{\sim}{\varepsilon}$, then setting $\underset{\sim}{\varepsilon} = \underset{\sim}{0}$, we find

$$\begin{aligned}\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}]\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}} = -\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{&\frac{i\hbar}{2}((\nabla\psi^*)\dot{\psi} + \psi^*\nabla\dot{\psi} - (\nabla\dot{\psi}^*)\psi - \dot{\psi}^*\nabla\psi)\\ &- \frac{\hbar^2}{2m}\sum_i((\nabla\nabla_i\psi^*)\cdot\nabla_i\psi) + (\nabla\nabla_i\psi)\cdot\nabla_i\psi^*) - \hat{v}(\nabla\psi^*\psi + \psi^*\nabla\psi)\} .\end{aligned} \qquad (3.117)$$

In eq. (3.117), we have used the dyadic notations to express the contributions to $\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}]$ arising from the kinetic energy part of the Lagrangian, $\mathcal{L}[\Omega,\psi_{\varepsilon}]$. Explicitly, e.g., the term

$(\nabla\nabla_i\psi^*)\cdot\nabla_i\psi$ reads

$$(\nabla\nabla_i\psi^*)\cdot\nabla_i\psi = \sum_i \underset{\sim}{e}_k\Big(\sum_\ell \frac{\partial^2\psi^*}{\partial r_k \partial r_\ell^{(i)}}\cdot\frac{\partial\psi}{\partial r_\ell^{(i)}}\Big) \quad , \tag{3.118}$$

where $r_k(r_k^{(i)})$ is the component of $\underset{\sim}{r}(\underset{\sim}{r}_i)$ along the direction of the unit vector $\underset{\sim}{e}_k$, $k = 1,2,3$. The r.h.s. of eq. (3.118) can also be written as

$$\sum_k \underset{\sim}{e}_k\Big(\sum_\ell \frac{\partial^2\psi^*}{\partial r_k\partial r_\ell^{(i)}}\cdot\frac{\partial\psi}{\partial r_\ell^{(i)}}\Big) = \sum_i \underset{\sim}{e}_k\Big(\sum_\ell\Big[\frac{\partial}{\partial r_\ell^{(i)}}\Big(\frac{\partial\psi^*}{\partial r_k}\frac{\partial\psi}{\partial r_\ell^{(i)}}\Big) - \frac{\partial\psi^*}{\partial r_k}\frac{\partial^2\psi}{\partial(r_\ell^{(i)})^2}\Big]\Big)$$

$$= \sum_\ell \frac{\partial}{\partial r_\ell^{(i)}}\Big[\Big(\sum_k \underset{\sim}{e}_k\frac{\partial\psi^*}{\partial r_k}\Big)\frac{\partial\psi}{\partial r_\ell^{(i)}}\Big] - \Big(\sum_k \underset{\sim}{e}_k\frac{\partial\psi^*}{\partial r_k}\Big)\Big(\sum_\ell\frac{\partial^2\psi}{\partial(r_\ell^{(i)})^2}\Big)$$

$$= \nabla_i\cdot(\nabla_i\psi\nabla\psi^*) - \nabla\psi^*\nabla_i^2\psi \;,$$

where the dyadic $\nabla_i\psi\nabla\psi^*$ is given by

$$\nabla_i\psi\nabla\psi^* = \sum_{\ell,k}\underset{\sim}{e}_\ell \frac{\partial\psi}{\partial r_\ell^{(i)}}\frac{\partial\psi^*}{\partial r_k}\underset{\sim}{e}_k \;.$$

Thus,

$$(\nabla\nabla_i\psi^*)\cdot\nabla_i\psi = \nabla_i\cdot(\nabla_i\psi\nabla\psi^*) - \nabla\psi^*\nabla_i^2\psi \;. \tag{3.119a}$$

Likewise, we can write the term $(\nabla\nabla_i\psi)\cdot\nabla_i\psi^*$ as

$$(\nabla\nabla_i\psi)\cdot\nabla_i\psi^* = \nabla_i\cdot(\nabla_i\psi^*\nabla\psi) - \nabla_i^2\psi^*\nabla\psi \;. \tag{3.119b}$$

Using eqs. (3.119b) and Schrödinger's equations (3.67), we infer, from eq. (3.117)

$$\underset{\sim}{\varepsilon}\cdot\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}]\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}} = -\underset{\sim}{\varepsilon}\cdot\Big[\int_{\Omega}d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\frac{i\hbar}{2}(\dot{\psi}^{*}\nabla\psi + \psi^{*}\nabla\dot{\psi} - (\nabla\psi^{*})\dot{\psi} - (\nabla\dot{\psi}^{*})\psi)\}$$

$$+ \frac{\hbar^{2}}{2m}\oint dS\int' d\underset{\sim}{\tau}'\{\nabla\psi^{*}\nabla\psi + \nabla\psi\nabla\psi^{*}\}\cdot\underset{\sim}{n}\Big],$$

or, equivalently, using eq. (3.19c)

$$\underset{\sim}{\varepsilon}\cdot\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}]\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}} = \underset{\sim}{\varepsilon}\cdot\{\frac{m}{N}\int_{\Omega}d\underset{\sim}{r}(\partial\underset{\sim}{J}/\partial t) + \frac{\hbar^{2}}{2m}\oint dS\int' d\underset{\sim}{\tau}'(\nabla\psi^{*}\nabla\psi + \nabla\psi\nabla\psi^{*})\cdot\underset{\sim}{n}\}. \quad (3.120)$$

Since the subsystem is bounded by a zero-flux surface at all stages of the variations generated by $\hat{\mathcal{A}}$, eq. (3.103a) applies, and can be written in the form

$$\underset{\sim}{\varepsilon}\cdot\{\frac{\hbar^{2}}{4m}\nabla_{\varepsilon}(\int_{\Omega}\nabla^{2}\rho_{\psi_{\varepsilon}}d\underset{\sim}{r})\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}}\} = \underset{\sim}{\varepsilon}\cdot\{-\frac{\hbar^{2}}{4m}\oint dS\nabla\nabla\rho\cdot\underset{\sim}{n}\}$$

$$= \underset{\sim}{\varepsilon}\cdot\{-\frac{\hbar^{2}}{4m}\oint dS\int' d\underset{\sim}{\tau}'[(\nabla\nabla\psi^{*})\psi + \nabla\psi^{*}\nabla\psi$$

$$+ \nabla\psi\nabla\psi^{*} + \psi^{*}(\nabla\nabla\psi)]\cdot\underset{\sim}{n}\}$$

$$= 0 \quad (3.121)$$

Eq. (3.121) denotes $\underset{\sim}{\varepsilon}$ as a vector arbitrarily chosen in the plane perpendicular to the resultant vector of the surface integral of $\nabla\nabla\rho\cdot\underset{\sim}{n}$. Using the identity of eq. (3.121), we rewrite eq. (3.120) as

$$\underset{\sim}{\varepsilon}\cdot\nabla_{\underset{\sim}{\varepsilon}}\mathcal{L}[\Omega,\psi_{\varepsilon}]\Big|_{\underset{\sim}{\varepsilon}=\underset{\sim}{0}} = \underset{\sim}{\varepsilon}\cdot\{\frac{m}{N}\int_{\Omega}d\underset{\sim}{r}(\partial\underset{\sim}{J}/\partial t) - \frac{1}{N}\oint dS\overleftrightarrow{P}\cdot\underset{\sim}{n}\} \quad (3.122)$$

where the pressure tensor $\overleftrightarrow{P}$ has been defined in eq. (3.43b).

By virtue of eq. (3.82b) and the definition of the Hamiltonian $\hat{H}$, the r.h.s. of eq. (3.114) is seen to give the subsystem average of the total force acting on the electron described by $\underset{\sim}{r}$

$$\tfrac{1}{2}\underset{\sim}{\varepsilon}\cdot\{\langle\frac{i}{\hbar}[\hat{H},\hat{p}]\rangle_{\Omega} + \text{c.c.}\} = \underset{\sim}{\varepsilon}\cdot\langle-\nabla\hat{v}\rangle_{\Omega} = \underset{\sim}{\varepsilon}\cdot\{\frac{1}{N}\underset{\sim}{F}(\Omega,t)\} \qquad (3.123)$$

Through the requirement that eq. (3.114) be obtained for all admissible $\underset{\sim}{\varepsilon}$, the variational principle yields the force law[70]

$$\underset{\sim}{F}(\Omega,t) = m\int_{\Omega} d\underset{\sim}{r}(\partial\underset{\sim}{J}/\partial t) - \oint dS\,\overset{\leftrightarrow}{P}\cdot\underset{\sim}{n} \qquad (3.124)$$

Eq. (3.124) is the generalization of Ehrenfest's theorem:[21,22,74] It relates the forces acting on the subsystem to the time derivative of the momentum density $\underset{\sim}{J}$. It constitutes the quantum analogue of Newton's equations of motion in classical mechanics.

A particular expression for the above force law has been used in the preceding section, eq. (3.45), which is seen to result from the vanishing of $(\partial\underset{\sim}{J}/\partial t)$ in a stationary state. Above, we have derived eq. (3.124) as a consequence of the variational principle. It can be shown that the same result also follows from Epstein's differential force law, which, in its original form, states that[71]

$$\psi^*\nabla_j \hat{\nu}\psi = -\frac{i\hbar}{2}\frac{\partial}{\partial t}[(\nabla_j\psi^*)\psi - \psi^*\nabla_j\psi] + \frac{\hbar^2}{4m}\sum_k \nabla_k\cdot[(\nabla_k\nabla_j\psi^*)\psi$$

$$+ \psi^*(\nabla_k\nabla_j\psi) - \nabla_k\psi^*\nabla_j\psi - \nabla_k\psi\nabla_j\psi^*] \,. \qquad (3.125)$$

Indeed, we verify that eq. (3.124) follows directly from eq. (3.125), upon replacing ∇_j by ∇ and carrying out the integration denoted by $\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'$. The derivation of eq. (3.124) from Epstein's differential force law constitutes an alternative to the one presented above which utilized the variational principle, in the form of eq. (3.114).[70] Thus, we can regard eq. (3.124) as an independent result, which we may require to be satisfied by the subsystem averaged force $\underset{\sim}{F}(\Omega)$, together with the variational result of eq. (3.114). A comparison of the right hand sides of eqs. (3.120), (3.124), then shows that the simultaneous satisfaction of both eqs. (3.114) and (3.124) necessitates the result of eq. (3.121a). Thus, to meet the requirements that i) the subsystem satisfies the variational principle, eq. (3.110), and hence, eq. (3.114); for variations generated by $\hat{A} = \underset{\sim}{\varepsilon}\cdot\hat{\underset{\sim}{p}}$;

ii) it obeys the force law of eq. (3.124), and finally; iii) its Lagrangian vanishes at the point of variation, we must impose the zero-flux boundary condition on the subsystem. In other words, the variational principle, eq. (3.110), defines the subsystem unambiguously as one bounded by a zero-flux surface, if the subsystem is to satisfy the physical force law of eq. (3.124).^{70}

A comparable result is also obtained in the case $\hat{A} = -\varepsilon \underset{\sim}{r}\cdot\underset{\sim}{\hat{p}}$, where ε is an infinitesimal real number. The operator $\hat{A}$ then generates the scaling in the electronic coordinates $\underset{\sim}{r}$ by the factor $\zeta = 1 + \varepsilon$. That is

$$\psi_\varepsilon(\underset{\sim}{r},\underset{\sim}{\tau}',t) = (\hat{1} + \frac{i}{\hbar}\varepsilon\underset{\sim}{r}\cdot\underset{\sim}{\hat{p}})\psi(\underset{\sim}{r},\underset{\sim}{\tau}',t) = \psi(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\,.$$

The variational principle now states that

$$\frac{\partial}{\partial\varepsilon}\mathcal{L}[\Omega,\psi_\varepsilon]\Big|_{\varepsilon=0} = -\tfrac{1}{2}\{\langle\frac{i}{\hbar}[\hat{H},\underset{\sim}{\hat{r}}\cdot\underset{\sim}{\hat{p}}]\rangle_\Omega + \text{c.c.}\}\,. \qquad (3.126)$$

Writing the Lagrangian $\mathcal{L}[\Omega,\psi_\varepsilon]$ explicitly in terms of ε, we have (to first order in ε):

$$\mathcal{L}[\Omega,\psi_\varepsilon] = \int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\{\frac{i\hbar}{2}[\psi^*(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\dot{\psi}(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t) - \dot{\psi}^*(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\psi(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)]$$
$$- \frac{\hbar^2}{2m}\sum_i\nabla_i\psi^*(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\cdot\nabla_i\psi(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t) - \hat{v}\psi^*(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\psi(\zeta\underset{\sim}{r},\underset{\sim}{\tau}',t)\}. \quad (3.127$$

Differentiating both sides of eq. (3.127) with respect to ε, and noting that $(\partial/\partial\varepsilon) = \underset{\sim}{r}\cdot\nabla$ as $\varepsilon \to 0$, we find

$$\frac{\partial}{\partial\varepsilon}\mathcal{L}[\Omega,\psi_\varepsilon]\Big|_{\varepsilon=0} = \int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\underset{\sim}{r}\cdot[\frac{i\hbar}{2}(\nabla\psi^*\dot{\psi} + \psi^*\nabla\dot{\psi} - (\nabla\dot{\psi}^*)\psi - \dot{\psi}^*\nabla\psi)$$
$$- \frac{\hbar^2}{2m}\sum_i((\nabla\nabla_i\psi^*)\cdot\nabla_i\psi + (\nabla\nabla_i\psi)\cdot\nabla_i\psi^*) - \hat{v}(\nabla\psi^*\psi + \psi^*\nabla\psi)]$$
$$+ \int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'(-\frac{\hbar^2}{2m})\nabla\psi^*\cdot\nabla\psi\quad.$$

The last result can again be rearranged by using eqs. (3.119), and the Schrödinger equation, to give

$$\frac{\partial}{\partial\varepsilon}\mathcal{S}[\Omega,\psi_\varepsilon]\Big|_{\varepsilon=0} = -\frac{m}{N}\int_\Omega d\underset{\sim}{r}\,(\underset{\sim}{r}\cdot\frac{\partial \underset{\sim}{J}}{\partial t}) - \frac{\hbar^2}{2m}\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\sum_i \nabla_i\cdot(\nabla_i\psi^*\nabla_i\psi + \nabla_i\psi\nabla_i\psi^*)\cdot\underset{\sim}{r}$$
$$-(\frac{\hbar^2}{2m})\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'\nabla\psi^*\cdot\nabla\psi \;. \qquad (3.128)$$

Using the identities

$$\nabla_i\cdot(\nabla_i\psi^*\nabla\psi)\cdot\underset{\sim}{r} = \nabla_i\cdot[(\nabla_i\psi^*\nabla\psi)\cdot\underset{\sim}{r}] - (\nabla_i\psi^*\nabla\psi)\cdot\nabla_i\underset{\sim}{r} \;,$$

$$\nabla_i\cdot(\nabla_i\psi\nabla\psi^*)\cdot\underset{\sim}{r} = \nabla_i\cdot[(\nabla_i\psi\nabla\psi^*)\cdot\underset{\sim}{r}] - (\nabla_i\psi\nabla\psi^*)\cdot\nabla_i\underset{\sim}{r} \;,$$

we obtain, from eq. (3.128)

$$\frac{\partial}{\partial\varepsilon}\mathcal{S}[\Omega,\psi_\varepsilon]\Big|_{\varepsilon=0} = -\frac{m}{N}\int_\Omega d\underset{\sim}{r}\,(\underset{\sim}{r}\cdot\frac{\partial \underset{\sim}{J}}{\partial t}) - \frac{\hbar^2}{2m}\oint dS\int' d\underset{\sim}{\tau}'\underset{\sim}{r}\cdot[\nabla\psi^*\nabla\psi + \nabla\psi\nabla\psi^*]\cdot\underset{\sim}{n} \;. \quad (3.129)$$

Evaluating the commutator of the r.h.s. of eq. (3.126), we find

$$\langle\frac{i}{\hbar}[\hat{H},\underset{\sim}{\hat{r}}\cdot\underset{\sim}{\hat{p}}]\rangle_\Omega = 2\langle\psi^*,(-\frac{\hbar^2}{2m})\nabla^2\psi\rangle_\Omega + \langle\psi^*,-\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_\Omega \;,$$

so that, combined with eq. (3.129), eq. (3.126) becomes

$$\langle\psi^*,-\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_\Omega = -\tfrac{1}{2}\{2\langle\psi^*,(-\frac{\hbar^2}{2m})\nabla^2\psi\rangle_\Omega + \text{c.c.}\} + \frac{m}{N}\int_\Omega d\underset{\sim}{r}\,(\underset{\sim}{r}\cdot\frac{\partial \underset{\sim}{J}}{\partial t})$$
$$+\frac{\hbar^2}{2m}\oint dS\int' d\underset{\sim}{\tau}'\underset{\sim}{r}\cdot[\nabla\psi^*\nabla\psi + \nabla\psi\nabla\psi^*]\cdot\underset{\sim}{n} \qquad (3.130)$$

We can compare the last result with an expression derived from Epstein's differential force law. For this, we set $\nabla_j = \nabla$ in eq. (3.125), take the scalar product of both sides of the resulting equation with $-\underset{\sim}{r}$, then apply the integration indicated by $\int_\Omega d\underset{\sim}{r}\int' d\underset{\sim}{\tau}'$, and rearrange terms to get

$$\langle\psi^*,-\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_\Omega = -\tfrac{1}{2}\{2\langle\psi^*,-\frac{\hbar^2}{2m}\nabla^2\psi\rangle_\Omega + \text{c.c.}\} + \frac{m}{N}\int_\Omega d\underset{\sim}{r}(\underset{\sim}{r}\cdot\frac{\partial\underset{\sim}{J}}{\partial t})$$

$$- \frac{\hbar^2}{4m}\oint dS\int' d\underset{\sim}{\tau}'\underset{\sim}{r}\cdot[(\nabla\nabla\psi^*)\psi + \psi^*(\nabla\nabla\psi) - \nabla\psi^*\nabla\psi - \nabla\psi\nabla\psi^*]\cdot\underset{\sim}{n}$$

$$- \frac{\hbar^2}{4m}\int_\Omega d\underset{\sim}{r}\nabla^2\rho(\underset{\sim}{r}) \ . \tag{3.131}$$

Comparing eqs. (3.131) and (3.130), we see that the two expressions differ by the term

$$\frac{\hbar^2}{4m}\frac{\partial}{\partial\varepsilon}(\int_\Omega\nabla^2\rho_{\psi_\varepsilon}d\underset{\sim}{r})\Big|_{\varepsilon=0} = \frac{\hbar^2}{4m}\{\oint dS\int' d\underset{\sim}{\tau}'\underset{\sim}{r}\cdot[(\nabla\nabla\psi^*)\psi + \psi^*(\nabla\nabla\psi) + \nabla\psi^*\nabla\psi$$

$$+ \nabla\psi\nabla\psi^*]\cdot\underset{\sim}{n} + \int_\Omega d\underset{\sim}{r}\nabla^2\rho(\underset{\sim}{r})\} \ , \tag{3.132}$$

which corresponds to the variation in $(\frac{\hbar^2}{4m}\int_\Omega\nabla^2\rho d\underset{\sim}{r})$ in the limit $\epsilon \to 0$. The consistency requirement on the two expressions of the quantity $\langle\psi^*,-\underset{\sim}{r}\cdot\nabla\hat{v}\psi\rangle_\Omega$, eqs. (3.130) and (3.131), thus demands that this variation in $(\frac{\hbar^2}{4m})\int_\Omega\nabla^2\rho d\underset{\sim}{r}$ vanishes, i.e., that the constraint of eq. (3.103a) be satisfied. Referring to eq. (3.132), we recognize that this condition can be fulfilled only if the two terms on the r.h.s. of this equation separately vanish since the first time is origin-dependent while the value of the last term does not depend on any particular choice of the origin. To determine the origin for which the first term vanishes amounts to delineating the class of admissible trial functions ψ_ε to be used in the application of the variational principle. Combined with eq. (3.103a), which, as discussed above, is to be observed to ensure the consistency between eqs. (3.130) and (3.131), the vanishing of the second term on the r.h.s. of eq. (3.132) demands that the subsystem be bounded by a zero-flux surface. Only in this case, from eqs. (3.130), (3.131), can a time-dependent virial theorem be uniquely associated with the subsystem, and reads

$$-2\bar{T}(\Omega) = \bar{V}(\Omega) - m\int_{\Omega} d\underset{\sim}{r}\,\Big(\underset{\sim}{r}\cdot\frac{\partial \underset{\sim}{J}}{\partial t}\Big) + \oint dS\,\underset{\sim}{r}\cdot\overset{\leftrightarrow}{P}\cdot\underset{\sim}{n}\,. \tag{3.133}$$

We note that eq. (3.133) properly reduces to the virial theorem for a stationary state, which was derived in the previous section using the stationary state variational principle (cf. eqs. (3.46) and eq. (3.43a)). The generalized virial theorem, eq. (3.133), does not depend on the choice of the origin of the vector $\underset{\sim}{r}$ - i.e., it is form-invariant in a change of origin - since it follows from Epstein's differential force law with $\underset{\sim}{r}$ referring to an arbitrary origin. Only its derivation from the variational principle is origin-dependent: eq. (3.133) can be regarded as the consequence of the variational principle, eq. (3.110), only when the variations $\delta\psi$ are restricted to be of the type $\delta\psi = -\frac{i}{\hbar}\varepsilon\hat{\underset{\sim}{r}}\cdot\hat{\underset{\sim}{p}}\psi$, with ε being an arbitrary infinitesimal number, and the vector $\underset{\sim}{r}$ measured from some particular origin so as to ensure the vanishing of eq. (3.132).

The time-dependent virial theorem stated in eq. (3.133) differs from the statement for a stationary state, eqs. (3.46), by just the term $-m\int_{\Omega} d\underset{\sim}{r}\,(\underset{\sim}{r}\cdot\partial\underset{\sim}{J}/\partial t)$. This term can be interpreted as the virial of the forces which arise from the change in the momentum density $\underset{\sim}{J}(\underset{\sim}{r},t)$ with time, while $\overline{V}(\Omega)$ gives the virial of the forces on an electron in Ω, resulting from its instantaneous interaction with the other electrons of the system. Through the three examples considered above, corresponding respectively to $\hat{A} = \varepsilon\hat{1}N$, $\underset{\sim}{\varepsilon}\cdot\underset{\sim}{p}$, and $\varepsilon\underset{\sim}{r}\cdot\underset{\sim}{p}$, we have illustrated the use of the general variational principle, eq. (3.110), in obtaining physical laws governing a subsystem in a time-dependent molecular system. When $\hat{A} = \varepsilon\hat{1}N$, describing a simple infinitesimal phase transformation, the generalized equation of continuity, eq. (3.112a), is obtained, which governs the time change of the electronic population of the subsystem. In the case $\hat{A} = \underset{\sim}{\varepsilon}\cdot\underset{\sim}{p}$, the generator of an infinitesimal uniform translation of an electronic coordinates vector $\underset{\sim}{r}$, we have derived the force law of eq. (3.124), which is seen to be the quantum analogue of Newton's equation of motion in classical mechanics, and a generalization of the well-known Ehrenfest theorem. Finally, with $\hat{A} = \varepsilon\underset{\sim}{\hat{r}}\cdot\underset{\sim}{\hat{p}}$, which generates the scaling of the electronic coordinates vector $\underset{\sim}{r}$ by the factor $\zeta = (1+\varepsilon)$, we have arrived at the time-dependent virial theorem, eq. (3.133). In the last two cases, we have also demonstrated that the variational principle, together with the physical statements of the above force law and virial theorem, provides a characterization of the subsystem which is completely equivalent to its definition via the zero-flux boundary condition, eq. (3.1).[70] Previously, we showed that eq. (3.110) is but a particular statement of the Schrödinger variational principle, eqs. (3.106) and (3.107). The expression of the latter for the particular case where the subsystem occupies all the electronic space, the total system case, is found to be equivalent to the fundamental Schwinger quantum action principle and principle of stationary action.[75] This connection depicts the significance of, and establishes the role played by the Schrödinger variational principle in the quantum mechanical study of molecular subsystems.

CONCLUSION

The derivations and discussions of the last chapter clearly demonstrate the fundamental role played by the zero-flux condition, eq. (3.1), in defining subsystems of a molecular system, with unique physical properties. In both the time-dependent case and the case of a total system in a stationary state, these subsystems can be characterized by various expressions of the variational principle which constitute the generalizations of the corresponding statements pertaining to a total isolated system. Noteworthy are the results of eqs. (3.26a) and eq. (3.110), which relate, in a simple manner, the change in the subsystem variational functionals $F_E[\Omega,\psi]$ and $\mathcal{L}[\Omega,\psi]$, respectively, to the generator $\hat{A}$ of an infinitesimal unitary transformation. By an appropriate choice of this generator, these statements of the variational principle lead to the simultaneous identification of important subsystem properties and the physical laws which they obey. Thus, the subsystem total energy, $\bar{E}(\Omega)$, and kinetic energy, $\bar{T}(\Omega)$, together with the subsystem virial theorem, eqs. (3.46), (3.57) and (3.133), are obtained by considering the effect of an infinitesimal scaling transformation on the subsystem functionals $F_E[\Omega,\psi]$ and $\mathcal{L}[\Omega,\psi]$. Similarly, the relationship between the subsystem average momentum, $\underset{\sim}{p}(\Omega)$, and the average of the forces on the subsystem, $\mathbf{F}(\Omega)$, follows from the variational principle, eq. (3.110), with the variation in $\mathcal{L}[\Omega,\psi]$ caused by an infinitesimal translation of the coordinates of an electron. This relationship, the subsystem force law, is given in eq. (3.124), and can be rewritten as[70]

$$\mathbf{F}(\Omega) = \frac{d}{dt}(\mathbf{p}(\Omega)) - m\oint dS\frac{\partial S}{\partial t}\mathbf{J} - \oint dS\,\overleftrightarrow{P}\cdot\mathbf{n} \quad , \tag{4.1a}$$

where

$$\mathbf{p}(\Omega) = m\int_\Omega d\mathbf{r}\mathbf{J}(\mathbf{r},t) = \int_\Omega d\mathbf{r}\int' d\boldsymbol{\tau}'\{\psi^*\hat{\mathbf{p}}\psi + \text{c.c.}\} \tag{4.1b}$$

denotes the subsystem average linear momentum.

The subsystem properties, $\overline{E}(\Omega)$, $\overline{T}(\Omega)$, $\mathbf{p}(\Omega)$ and $\mathbf{F}(\Omega)$, are all defined by the integral of the associated densities over the region Ω occupied by the subsystem. They can also be written in the form

$$\begin{aligned} A(\Omega) &= \langle\psi|\tfrac{1}{2}\sum_i(\hat{\Pi}(\Omega,i)\hat{a}(i) + \hat{a}(i)\hat{\Pi}(\Omega,i))|\psi\rangle \\ &= N\langle\psi|\tfrac{1}{2}(\hat{\Pi}(\Omega)\hat{a}(1) + \hat{a}(1)\hat{\Pi}(\Omega))|\psi\rangle \quad , \end{aligned} \tag{4.2a}$$

where

$$\begin{aligned} \hat{\Pi}(\Omega,i) &= \int_\Omega d\mathbf{r}_i\int\prod_{j\neq i}d\mathbf{r}_j|..\mathbf{r}_i..\rangle\langle..\mathbf{r}_i..| \quad , \\ \hat{\Pi}(\Omega) &\equiv \hat{\Pi}(\Omega,1) \quad , \end{aligned} \tag{4.2b}$$

and $\hat{a}(i)$ is an operator associated with a single electron, obtained in the symmetric partitioning of an N-electron operator $\hat{A}$, i.e.,

$$\hat{A}(1,2,\ldots,N) = \sum_{i=1}^{N}\hat{a}(i) \ . \tag{4.2c}$$

This definition of $A(\Omega)$ then imparts to the subsystem property associated with the operator $\hat{A}$, the additivity property[70]

$$\sum_{\Omega} A(\Omega) = \langle \hat{A} \rangle \quad , \tag{4.3}$$

as a consequence of eqs. (4.2b,c), and the obvious result

$$\sum_{\Omega} \hat{\Pi}(\Omega,i) = \hat{1}, \quad \forall i = 1,2,\ldots,N \ .$$

While this additivity property can be obtained for any partitioning of the molecular system into non-overlapping regions Ω, only the partitioning by the zero-flux surfaces defines subsystems with well-defined kinetic energies, and which satisfy the virial theorem, eqs. (3.133), (3.46), as well as the force law of eqs. (3.124) and (4.1).

We have derived Schrödinger's time independent and time dependent equations as the Euler equations of the variational problems posed in eqs. (3.18) and eq. (3.106), respectively. The solutions of these Euler equations thus describe the total system instead of a subsystem only. In other words, the variational principle does not lead to the association of a state-function with a subsystem. This is a reflection of the fact that a subsystem is open, and, in general, cannot be separated from its environment. Formally, the state of an open system can only be described by a density matrix which determines the average value of the physical properties of the system, and represents a statistical ensemble of identical open systems.[55]

This remark suggests that future developments in the quantum mechanics of molecular subsystems should focus on the definition of the subsystem density matrix and the study of its time evolution.

The definition of the subsystem by its boundary, a zero-flux surface, makes use of a topological property exhibited by the charge distribution of the molecular system. A zero-flux surface is the union of a number of interatomic surfaces defined in Chapter I, and possibly some portions of a surface found at infinity. When the subsystem contains only one nucleus, it represents an atomic fragment of the molecular system. The mentioned nucleus is the only attractor of the gradient vector field of the charge distribution within the region occupied by the subsystem. Thus, the same partitioning of the molecular system is obtained by the zero-flux boundary condition, eq. (3.1), which, as discussed above, uniquely defines subsystems with recognizable physical properties, and by the identification of the attractors of the gradient vector field of the molecular charge distribution. Together with their associated basins, these attractors constitute the atomic fragments of the molecular system. The phenomenological study of the topological properties of the charge distribution, detailed in Chapters I and II, also leads to the identification of a chemical bond with a bond path, the union of the pair of gradient paths originating from a (3,-1) critical point. Since an interatomic surface can also be regarded as a property - complementary to the bond path - associated with a (3,-1) critical point, the concepts of atoms and bonds provide two equivalent descriptions of the constitution of molecular matter. Atoms and bonds are the basic ingredients in the definition of molecular structure, and the determination of structural changes, which were discussed and illustrated in Chapter II. Structure and structural changes, as defined therein, are essentially topological concepts. Their quantum-mechanical significance are still to be discovered. To that effect, much effort is being invested in this laboratory to elucidate the followings: i) the relationship between the external nuclear potential and the charge distribution. A comparative study of the topological properties of the nuclear potential to which the electrons of a molecular system are subjected, and those of the charge distribution, has revealed a close relationship between the two fields mentioned above.[76] This observation is in accordance with the Höhenberg and Köhn theorem which asserts a functional dependence of ρ on the external potential and vice-versa, for a system in a non-degenerate ground state.[26]

ii) The relationship between the topological properties of the charge distribution $\rho(\underset{\sim}{r})$ and those of the first-order spinless density matrix $\Gamma^{(1)}(\underset{\sim}{r},\underset{\sim}{r}')$, in the neighbourhood of the submanifold $(\underset{\sim}{r} = \underset{\sim}{r}')$ of the six-dimensional space spanned by the variables $\underset{\sim}{r}$ and $\underset{\sim}{r}'$.[70] This study is motivated by the fact that, because of the virial theorem, and Epstein's differential force law, all molecular properties of interest, and their associated densities are expressible in terms of $\Gamma^{(1)}$ and ρ. Only in a variational derivation, does one need the second-order density matrix and encounter the n-representability problem. It is thus hoped that, utilizing the one-matrix, a finer description will emerge, which will allow for a direct identification of the topological features of ρ with quantum mechanical properties of the molecular system. In particular, emphasis is being placed upon the role of the pressure tensor defined in Chapter III, since its components along the normal vectors to the atomic surface determine the total force on a given atom. It thus appears to play a central role in any attempt to delineate the quantum mechanical signification of the bonds which are represented by the bond paths exhibited by a molecular charge distribution.

APPENDIX I

SOURCES AND BASIS SETS OF WAVEFUNCTIONS

Listed in the following are the sources and basis sets of the wavefunctions of the molecules used for the illustrations in Chapters I and II. The list follows the order of appearance of the chemical systems discussed in these two chapters, except for H_2O, which, in this list, appears after the H_3^+ system. All wavefunctions are single determinantal SCF wavefunctions of near Hartree-Fock quality. A bracket $[\ldots n_i t_i \ldots | \ldots n_i' t_i \ldots]$ denotes a set of basis functions in which there are n_i' atomic orbitals of type t_i, obtained by the contraction of n_i atomic orbitals (the primitive functions) of the same type.

1) $\underline{C_3H_4}$ (Cyclopropene): This laboratory[19] - GTO (Gaussian Type Orbitals) basis set consisting of [10s,5p,2d|5s,3p,2d] on C and [4s,1p|2s,1p] on H. This basis set was originally used by Neumann and Moskowitz for *ab initio* calculations on formaldehyde (HCHO) and carbon monoxide (CO) (D. Neumann and J.W. Moskowitz, J. Chem. Phys. **50**, 2216 (1969)).

2) $\underline{C_8H_8}$ (Cubane): This laboratory,[19] STO-3G (Slater-type Orbitals, each expanded on three Gaussian primitive functions) minimal basis set.

3) <u>HCN</u>: This laboratory[47] - GTO basis set consisting of

[9s,5p,1d|4s,2p,1d] on C and N, and [4s,1p|2s,1p] on H. The potential energy surface for the rearrangement reaction I

$$HCN \rightarrow CNH \quad (I)$$

has been investigated by Schaeffer III *et al.* (P.K. Pearso, H.F. Schaeffer III and W. Wahlgren, J. Chem. Phys. 62, 350 (1972)). The discussion found in Chapter II, concerning the coincidence of the topological and quantum mechanical instabilities in this system, is based on the results obtained by these authors.

4) H_3^+: This laboratory[18,19] - GTO basis set consisting of [9s,3p,1d|4s,3p,1d] on each H. This basis set was originally used by Siegbahn and Lin (P. Siegbahn and B. Lin, J. Chem. Phys. 68, 2457 (1978)). Geometries corresponding to the minimum energy C_{2v} path were interpolated from the results of Conroy (H. Conroy, J. Chem. Phys. 51, 3979 (1969)). It has been shown that an SCF charge density for H_3^+ in the neighbourhood of the equilibrium geometry, of near Hartree-Fock quality, is in error by no more than 2% compared to a density calculated from a CI wavefunction using single and double configurations (A.J. Duben and J.P. Lowe, J. Chem. Phys. 55, 4276 (1971)).

5) H_2O: This laboratory[18,19] - GTO basis set consisting of [10s,5p,2d|5s,3p,2d] on O, and [4s,1p|2s,1p] on H, as originally used by Neumann and Moskowitz (D. Neumann and J.W. Moskowitz, J. Chem. Phys. 49, 2056 (1969)). The C_{2v}

nuclear configurations illustrated in Fig. 2.7 are obtained by interpolation from the results of previous studies on the ground state potential energy surface of this system (R.A. Gangi and R.F.W. Bader, J. Chem. Phys. 55, 5369 (1971); R.F.W. Bader and R.A. Gangi, J. Am. Chem. Soc. 93, 1831 (1971)).

APPENDIX II

HAMILTON'S PRINCIPLE AND ITS GENERALIZATIONS IN CLASSICAL MECHANICS

Consider a classical conservative system of n-degrees of freedom, described by the generalized coordinates $\underset{\sim}{q}$ and conjugate momenta $\underset{\sim}{p}$. In terms of $\underset{\sim}{q}$ and $\underset{\sim}{p}$, one constructs the action integral

$$W^{Cl} = \int_{t_1}^{t_2} dt(\underset{\sim}{p}\cdot\dot{\underset{\sim}{q}} - H(\underset{\sim}{p},\underset{\sim}{q})) \quad , \qquad \text{(AII.1)}$$

where $H(\underset{\sim}{p},\underset{\sim}{q})$ is the Hamiltonian of the system, which is the Legendre transform of the Lagrangian $L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t)$

$$H(\underset{\sim}{p},\underset{\sim}{q}) = \underset{\sim}{p}\cdot\dot{\underset{\sim}{q}} - L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t); \quad \underset{\sim}{p} \underset{\text{def}}{=} \nabla_{\dot{\underset{\sim}{q}}} L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t) \, . \qquad \text{(AII.2)}$$

In eq. (AII.2), we use the notation $\nabla_{\dot{q}} L$ to denote the n-vector whose components are the partial derivatives of L with respect to the n components of the generalized velocity vector $\dot{\underset{\sim}{q}}$. We will henceforth adopt this notation, and denote by $\nabla_{\underset{\sim}{u}} f(\underset{\sim}{u})$, the n-gradient vector of $f(\underset{\sim}{u})$ with respect to the n-vector $\underset{\sim}{u} = (u_1,u_2,\ldots,u_n)$. Accordingly, the scalar product $\underset{\sim}{u}\cdot\underset{\sim}{v}$ denotes the sum of all $u_i v_i$, $i = 1,\ldots,n$, for any pair of n-vectors $(\underset{\sim}{u},\underset{\sim}{v})$. Using eq. (AII.2), the action integral can also be expressed in terms of the Lagrangian as

$$W^{Cl} = \int_{t_1}^{t_2} dt L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t) \, . \qquad \text{(AII.3)}$$

Hamilton's principle states that the path along which the system actually evolves between t_1 and t_2 is determined by the condition

$$\delta W^{Cl} = 0 \quad , \tag{AII.4}$$

where δW^{Cl} denotes the first order change in W^{Cl} ensuing from small variations in the pair $(\underset{\sim}{q},\dot{\underset{\sim}{q}})$, or $(\underset{\sim}{q},\underset{\sim}{p})$, which vanish at the time end-points. In other words, given the states of the system at two fixed instants of time, t_1,t_2 - these states being specified either by the pair $(\underset{\sim}{q}(t_i),\dot{\underset{\sim}{q}}(t_i))$ or by $(\underset{\sim}{q}(t_i),\underset{\sim}{p}(t_i))$, $i = 1,2$ - the actual path followed by the system between t_1 and t_2, in real coordinates space or in the phase space, is selected among all paths connecting the two given states at the time end-points, so as to give a stationary value to W^{Cl}.[57,73] Applying Hamilton's principle, eq. (AII.4), to the expression of W^{Cl} given in eq. (AII.3), one obtains the Lagrange equation of motion

$$\frac{d}{dt}(\nabla_{\dot{\underset{\sim}{q}}}L) = \nabla_{\underset{\sim}{q}}L \quad , \tag{AII.5}$$

whereas using eq. (AII.1), Hamilton's equations of motion are obtained

$$\begin{aligned} \dot{\underset{\sim}{q}} &= \nabla_{\underset{\sim}{p}}H(\underset{\sim}{p},\underset{\sim}{q}) \quad , \\ \dot{\underset{\sim}{p}} &= -\nabla_{\underset{\sim}{q}}H(\underset{\sim}{p},\underset{\sim}{q}) \quad . \end{aligned} \tag{AII.6a}$$

Since both the Lagrange equation of motion, eq. (AII.5), and the Hamilton equations (AII.6a) were derived as the Euler equations of the same variational problem, they are equivalent. Lagrange's equation of motion gives the path of the system, in the real coordinate space (a n-dimensional space), as the integral curve of a second-order differential equation in $\underset{\sim}{q}$. In contrast, Hamilton's equations of motion utilize the 2n-dimensional phase space, the cartesian product of the coordinates space and the space of the momenta $\underset{\sim}{p}$, in which the actual path of the system is the integral curve of a set of 2n first-order differential equations.

Hamilton's equations of motion are preserved in a canonical transformation. The form invariance of these equations in fact, constitutes the requirement which defines canonical transformations in classical mechanics.[57,73] A canonical transformation is thus one of the type

$$(\underset{\sim}{p},\underset{\sim}{q}) \rightarrow (\underset{\sim}{P},\underset{\sim}{Q}); \qquad \underset{\sim}{P} = \underset{\sim}{P}(\underset{\sim}{p},\underset{\sim}{q}), \qquad \underset{\sim}{Q} = \underset{\sim}{Q}(\underset{\sim}{p},\underset{\sim}{q}) \ ,$$

such that the transformed generalized coordinates and momenta $\underset{\sim}{Q},\underset{\sim}{P}$ satisfy

$$\begin{aligned} \dot{\underset{\sim}{Q}} &= \nabla_{\underset{\sim}{P}} K(\underset{\sim}{P},\underset{\sim}{Q}) \ , \\ \dot{\underset{\sim}{P}} &= -\nabla_{\underset{\sim}{Q}} K(\underset{\sim}{P},\underset{\sim}{Q}) \ , \end{aligned} \qquad \text{(AII.6b)}$$

for some function $K(\underset{\sim}{P},\underset{\sim}{Q})$ which plays the same role, with respect to $\underset{\sim}{Q},\underset{\sim}{P}$ as $H(\underset{\sim}{p},\underset{\sim}{q})$ does with respect to the original coordinates and momenta $\underset{\sim}{q},\underset{\sim}{p}$. Again, eqs. (AIII.6b) can be regarded as the consequence of Hamilton's principle, i.e., it is the Euler equation of the variational problem

$$\delta W_1^{Cl} = 0 \quad , \tag{AII.7a}$$

where

$$W_1^{Cl} = \int_{t_1}^{t_2} dt\{\underset{\sim}{P}\cdot\dot{\underset{\sim}{Q}} - K(\underset{\sim}{P},\underset{\sim}{Q})\} \quad , \tag{AII.7b}$$

is the transformed action integral. The variations considered in eq. (AII.7a) are of the same type as those involved in eq. (AII.4), viz. they are required to vanish at the time end-points t_1 and t_2. A necessary condition for the simultaneous satisfaction of eq. (AII.7a) and eq. (AII.4) is that the change in the action integral induced by the canonical transformation be given by the difference of the values that some function, of $\underset{\sim}{q},\underset{\sim}{Q},\underset{\sim}{p},\underset{\sim}{P}$ and the time, assumes at the two time end-points, since the variation of such a term automatically vanishes. Thus,

$$\delta_c W^{Cl} = W_1^{Cl} - W^{Cl} = F(\underset{\sim}{q},\underset{\sim}{p},\underset{\sim}{Q},\underset{\sim}{P},t)\Big|_{t_1}^{t_2} . \tag{AII.8a}$$

The function F arising on the r.h.s. of eq. (AII.8a) can be related to the generating function of the canonical transformation. The 4n variables $(\underset{\sim}{q},\underset{\sim}{p},\underset{\sim}{Q},\underset{\sim}{P})$ of F are not all independent, and the function F, in fact, depends only on 2n of these variables.[57] For instance, F may be considered a function of $\underset{\sim}{q}$ and $\underset{\sim}{Q}$, in which case, we write

$$F = F_1(\underset{\sim}{q},\underset{\sim}{Q},t) \cdot$$

In differential form, the condition of eq. (AII.8a) reads

$$\underset{\sim}{p}\cdot\dot{\underset{\sim}{q}} - H(\underset{\sim}{p},\underset{\sim}{q}) - \underset{\sim}{P}\cdot\dot{\underset{\sim}{Q}} + K(\underset{\sim}{P},\dot{\underset{\sim}{Q}}) - \frac{d}{dt} F_1(\underset{\sim}{q},\underset{\sim}{Q},t) = 0 \cdot \qquad \text{(AII.8b)}$$

Expressing the total time derivative of F_1 as

$$\frac{d}{dt}F_1(\underset{\sim}{q},\underset{\sim}{Q},t) = \nabla_{\underset{\sim}{q}}F_1\cdot\dot{\underset{\sim}{q}} + \nabla_{\underset{\sim}{Q}}F_1\cdot\dot{\underset{\sim}{Q}} + \frac{\partial F_1}{\partial t} \quad , \qquad \text{(AII.9)}$$

and using the independence of $\underset{\sim}{q},\underset{\sim}{Q}$, hence of $\dot{\underset{\sim}{q}}$ and $\dot{\underset{\sim}{Q}}$, eq. (AII.8b) leads to

$$\underset{\sim}{p} = \nabla_{\underset{\sim}{q}}F_1(\underset{\sim}{q},\underset{\sim}{Q},t) \quad , \qquad \text{(AII.10a)}$$

$$\underset{\sim}{P} = -\nabla_{\underset{\sim}{Q}}F_1(\underset{\sim}{q},\underset{\sim}{Q},t) \quad , \qquad \text{(AII.10b)}$$

$$K(\underset{\sim}{P}(\underset{\sim}{q},\underset{\sim}{p}),\underset{\sim}{Q}(\underset{\sim}{q},\underset{\sim}{p})) = H(\underset{\sim}{p},\underset{\sim}{q}) + \frac{\partial F_1}{\partial t} \cdot \qquad \text{(AII.10c)}$$

Eq. (AII.10a) permits $\underset{\sim}{Q}$ to be determined as a function of $\underset{\sim}{p},\underset{\sim}{q}$ and t. Once $\underset{\sim}{Q}$ is obtained, eq. (AII.10b) directly gives $\underset{\sim}{P} = \underset{\sim}{P}(\underset{\sim}{q},\underset{\sim}{p})$. Finally, eq. (AII.10c) provides the expression of K as a function of $\underset{\sim}{q},\underset{\sim}{p}$. Thus, in this case, F is the generating function of the canonical transformation, since it completely determines the transformation $(\underset{\sim}{q},\underset{\sim}{p}) \to (\underset{\sim}{Q},\underset{\sim}{P})$, via eqs. (AII.10a,b). A generating function depending on $(\underset{\sim}{q},\underset{\sim}{P})$ can be constructed as the Legendre transform of F_1,[57]

$$F_2(\underset{\sim}{q},\underset{\sim}{P},t) = F_1(\underset{\sim}{q},\underset{\sim}{Q},t) + \underset{\sim}{P}\cdot\underset{\sim}{Q} \quad ,$$

in terms of which eq. (AII.8b) can be written as

$$\underset{\sim}{p}\cdot\dot{\underset{\sim}{q}} - H(\underset{\sim}{p},\underset{\sim}{q}) - \underset{\sim}{P}\cdot\dot{\underset{\sim}{Q}} + K(\underset{\sim}{P},\underset{\sim}{Q}) - \frac{d}{dt}\{F_2(\underset{\sim}{q},\underset{\sim}{P},t) - \underset{\sim}{P}\cdot\underset{\sim}{Q}\} = 0 \quad ; \qquad \text{(AII.8c)}$$

using

$$\frac{d}{dt}F_2 = \nabla_{\underset{\sim}{q}}F_2\cdot\dot{\underset{\sim}{q}} + \nabla_{\underset{\sim}{P}}F_2\cdot\dot{\underset{\sim}{P}} + \frac{\partial F_2}{\partial t} \quad ,$$

and the linear independence of $\dot{\underset{\sim}{q}}$ and $\dot{\underset{\sim}{P}}$, we now obtain

$$\underset{\sim}{P} = \nabla_{\underset{\sim}{q}}F_2 \quad , \qquad \text{(AII.11a)}$$

$$\underset{\sim}{Q} = \nabla_{\underset{\sim}{P}}F_2 \quad , \qquad \text{(AII.11b)}$$

$$K = H + \frac{\partial F_2}{\partial t} \quad . \qquad \text{(AII.11c)}$$

Eq. (AII.11a) determines $\underset{\sim}{P}$ as a function of $(\underset{\sim}{q},\underset{\sim}{p})$, the knowledge of which gives $\underset{\sim}{Q}$ by eq. (AII.11b) Repeating the above procedure, one can generate two other types of canonical transformations, whose generating functions depend on the pairs $(\underset{\sim}{p},\underset{\sim}{Q})$ and $(\underset{\sim}{p},\underset{\sim}{P})$, respectively, and can be obtained from F_1 or F_2 by appropriate Legendre transformations.[57] In terms of the generating function F_2, eq. (AII.8a) assumes the form

$$\delta_C W^{Cl} = \{F_2(\underset{\sim}{q},\underset{\sim}{P},t) - \underset{\sim}{P}\cdot\underset{\sim}{Q}\}\Big|_{t_1}^{t_2} = F\Big|_{t_1}^{t_2} \quad . \qquad \text{(AII.12)}$$

A particular F_2 is given by

$$F_2 = \underset{\sim}{q}\cdot\underset{\sim}{P} \quad ,$$

which simply generates the identity transformation, since, in accordance with eqs. (AII.11)

$$\underset{\sim}{p} = \underset{\sim}{P}; \qquad \underset{\sim}{Q} = \underset{\sim}{q}; \qquad K = H \ . \qquad \text{(AII.13)}$$

Thus, the function F appearing on the r.h.s. of eq. (AII.12) describes that part of the generating function which is responsible for a non-trivial (canonical) transformation, if any, i.e., one which is not the identity transformation. We call F the generator of the canonical transformation described by eqs. (AII.11). In equation (AII.8a), (AII.12), we have used the notation δ_C to temporarily differentiate the variations induced by the canonical transformation from those involved in eqs. (AII.4), (AII.7a) which were indicated by the symbol δ. The former are physically realized, whereas the latter are merely mathematical constructs, which were restricted to be vanishing at the time end-points, and thus never represent physical situations. We now relax this boundary condition to incorporate physically realizable variations, as induced by an infinitesimal canonical transformation whose generator is

$$F(\underset{\sim}{p},\underset{\sim}{q},t) = \underset{\sim}{p}\cdot\delta\underset{\sim}{q} \ . \qquad \text{(AII.13)}$$

Using eqs. (AII.11), it is straightforward to verify that $F(\underset{\sim}{p},\underset{\sim}{q},t)$, as given in eq. (AII.13), generates only variations in $\underset{\sim}{q}$, denoted $\delta\underset{\sim}{q}$. The change in W^{Cl} ensuing from these variations is readily evaluated using eq. (AII.3), and is found to be

$$\delta W^{Cl} = \int_{t_1}^{t_2} dt\left[\frac{d}{dt}(\nabla_{\dot{\underset{\sim}{q}}} L) - \nabla_{\underset{\sim}{q}} L\right]\cdot\delta\underset{\sim}{q} + \underset{\sim}{p}\cdot\delta\underset{\sim}{q}\Big|_{t_1}^{t_2} \quad . \qquad \text{(AII.14)}$$

Using Lagrange's equation of motion, eq. (AII.5), the above result leads to

$$\delta W^{Cl} = \underset{\sim}{p}\cdot\delta\underset{\sim}{q}\Big|_{t_1}^{t_2} = F\Big|_{t_1}^{t_2} \quad , \qquad \text{(AII.15)}$$

which is identical with eq. (AII.12). Conversely, requiring eq. (AII.15) to be obtained as a consequence of arbitrary variations, $\delta\underset{\sim}{q}$, in the components of $\underset{\sim}{q}$, one is led to an Euler equation which is Lagrange's equation of motion, by virtue of eq. (AII.14). Thus, eq. (AII.15) appears to be the generalization of Hamilton's principle which allows for the use of variations which do not necessarily vanish at the time end-points. These variations may represent actual changes in the path followed by the system between t_1 and t_2, which result from the action of an infinitesimal canonical transformation on the system, whose generator $F(\underset{\sim}{p},\underset{\sim}{q},t) = \underset{\sim}{p}\cdot\delta\underset{\sim}{q}$. Finally, we may also relax the condition of fixed time end-points, and let t_1,t_2 to be varied in W^{Cl}. We then obtain an additional term to the r.h.s. of eq. (AII.14), which equals the difference of the values of $L(\underset{\sim}{q},\dot{\underset{\sim}{q}},t)\delta t = (\underset{\sim}{p}\cdot\dot{\underset{\sim}{q}} - H(\underset{\sim}{p},\underset{\sim}{q}))\delta t$ taken at the time end-points. Accordingly, eq. (AII.15) is generalized to read

$$\delta W^{Cl} = [\underset{\sim}{p}\cdot\bar{\delta}\underset{\sim}{q} - H(\underset{\sim}{p},\underset{\sim}{q})\delta t]\Big|_{t_1}^{t_2} ; \quad \bar{\delta}\underset{\sim}{q} = \delta\underset{\sim}{q} + \dot{\underset{\sim}{q}}\delta t \quad . \qquad \text{(AII.16)}$$

Since the Hamiltonian $H(\underset{\sim}{p},\underset{\sim}{q})$ is the generator of time evolution, we again observe that the generalized variational statement, eq. (AII.16), is of the form of eqs. (AII.8a), (AII.12), with

$$F(\underset{\sim}{p},\underset{\sim}{q},t) = \underset{\sim}{p}\cdot\delta\underset{\sim}{q} - H(\underset{\sim}{p},\underset{\sim}{q})\delta t \ . \qquad \text{(AII.17)}$$

The canonical transformation described by F then has the effect of simultaneously changing the time variable t, by δt, and the coordinates of the system, by $\delta\underset{\sim}{q}$, while the canonical momenta $\underset{\sim}{p}$ are varied by

$$\delta\underset{\sim}{p} = \dot{\underset{\sim}{p}}\delta t \ .$$

Indeed, using eqs. (AII.11a,b), with $F_2 = \underset{\sim}{p}\cdot\underset{\sim}{q} + F$, we obtain

$$\underset{\sim}{Q} - \underset{\sim}{q} = \nabla_{\underset{\sim}{p}}F = \delta\underset{\sim}{q} - (\nabla_{\underset{\sim}{p}}H)\delta t = \delta\underset{\sim}{q} - \dot{\underset{\sim}{q}}\delta t = \delta\underset{\sim}{q} \ , \qquad \text{(AII.18a)}$$

and

$$\underset{\sim}{P} - \underset{\sim}{p} = \delta\underset{\sim}{p} = \nabla_{\underset{\sim}{q}}F = (-\nabla_{\underset{\sim}{q}}H)\delta t = \dot{\underset{\sim}{p}}\delta t \ . \qquad \text{(AII.18b)}$$

In deriving eqs. (AII.18), we have used the Hamilton equations of motion (AII.6). We have also noted that since F describes an infinitesimal transformation, its dependence on $\underset{\sim}{P}$ can be replaced by a dependence on the original momenta $\underset{\sim}{p}$. Thus, in obtaining eq. (AII.18a) via eq. (AII.11b), we have replaced the gradient operator $\nabla_{\underset{\sim}{P}}$ by $\nabla_{\underset{\sim}{p}}$, the gradient with respect to the components of the original momentum vector $\underset{\sim}{p}$.

APPENDIX III

SCHWINGER'S QUANTUM ACTION PRINCIPLE

A variational principle exists in quantum mechanics which parallels Hamilton's principle of classical mechanics. The principle was first derived by Schwinger who baptized it the quantum action principle.[64] Consider a system of n degrees of freedom, which classically is described by the coordinate vector $\underset{\sim}{q}$ and the conjugate momenta $\underset{\sim}{p}$, as in Appendix II. The Schwinger quantum action principle provides a differential characterization of the transformation function

$$\langle q',t_1|q'',t_2\rangle = \langle q',t_2|\hat{U}(t_1,t_2)|q'',t_2\rangle \qquad \text{(AIII.1)}$$

which is the probability amplitude for the transition from the coordinate representation, the representation associated with the observable $\underset{\sim}{q}(t)$, at the time t_1 to the same representation at the time t_2. In eq. (AIII.1), the unitary operator $\hat{U}(t_1,t_2)$ represents the time evolution of the system from t_1 to t_2. Its explicit expression will be given later, (eq. (AIII.4c)). In the coordinate representation, Schrödinger's equation,

$$i\hbar \frac{\partial}{\partial t}\psi(\underset{\sim}{q}',t) = \hat{H}(\underset{\sim}{q}',\frac{\hbar}{i}\nabla_{\underset{\sim}{q}'},t)\,\psi(\underset{\sim}{q}',t), \qquad \text{(AIII.2)}$$

is formally equivalent to the following set of equations

$$i\hbar\frac{\partial}{\partial t}\langle \underset{\sim}{q}',t| = \langle \underset{\sim}{q}',t|\hat{H}(\hat{\underset{\sim}{q}}(t),\hat{\underset{\sim}{p}}(t),t)\;, \qquad \text{(AIII.3a)}$$

$$\langle \underset{\sim}{q}',t|(\hat{\underset{\sim}{q}}(t) - \underset{\sim}{q}') = 0 \quad , \qquad \text{(AIII.3b)}$$

$$\langle \underset{\sim}{q}',t|\hat{\underset{\sim}{p}}(t) = \frac{\hbar}{i}\nabla_{\underset{\sim}{q}'}\langle \underset{\sim}{q}',t| \quad , \qquad \text{(AIII.3c)}$$

since the state function $\psi(\underset{\sim}{q}',t)$, can be regarded as the projection of a vector $|\psi\rangle$, in the abstract Hilbert space, on the basis vector $|\underset{\sim}{q}',t\rangle$:

$$\psi(\underset{\sim}{q}',t) = \langle \underset{\sim}{q}',t|\psi\rangle \quad .$$

Equation (AIII.3a) identifies the Hamiltonian operator, $\hat{H}$, as the generator of time displacement. Indeed, eq. (AIII.3a) can be written in the form

$$\langle \underset{\sim}{q}',t + dt| = \langle \underset{\sim}{q}',t|\hat{U}(t,t + dt) \quad , \qquad \text{(AIII.4a)}$$

where the time evolution operator for the development of the system between t and t + dt is[2,64]

$$\hat{U}(t,t + dt) = \hat{1} - \frac{i}{\hbar}dt\hat{H}(\hat{\underset{\sim}{q}},(t),\hat{\underset{\sim}{p}}(t),t) \quad . \qquad \text{(AIII.4b)}$$

If $\hat{H}$ is explicitly time-independent, eq. (AIII.4b) may be integrated to give the operator $\hat{U}(t_1,t_2)$ of eq. (AIII.1)

$$\hat{U}(t_1,t_2) = e^{(i/\hbar)(t_2-t_1)\hat{H}} \quad . \qquad \text{(AIII.4c)}$$

Likewise, we identify the momentum operator $\hat{\underset{\sim}{p}}$ as the generator of spatial translation, by rewriting eq. (AIII.3c) in the form

$$\langle \underset{\sim}{q}' + \delta\underset{\sim}{q},t| = \langle \underset{\sim}{q}',t|\hat{U}(\delta\underset{\sim}{q}) \quad , \qquad \text{(AIII.5a)}$$

where[2,64]

$$\hat{U}(\delta\underset{\sim}{q}) = \hat{1} + \frac{i}{\hbar}\delta\underset{\sim}{q}\cdot\hat{\underset{\sim}{p}}(t) = \hat{1} + \frac{i}{\hbar}\hat{G}_{\delta\underset{\sim}{q}(t)} \ , \qquad \text{(AIII.5b)}$$

and $\delta\underset{\sim}{q}$ is a vector of n arbitrary, small, real numbers. The unitary transformation of eqs. (AIII.4b) determines the Heisenberg equation of motion

$$\frac{d\hat{F}}{dt} = \frac{\partial\hat{F}}{\partial t} + \frac{1}{i\hbar}[\hat{F},\hat{H}] \ , \qquad \text{(AIII.6a)}$$

while eq. (AIII.5b) leads to

$$\nabla_{\underset{\sim}{q}}\hat{F} = \frac{1}{i\hbar}[\hat{F},\hat{\underset{\sim}{p}}] \ , \qquad \text{(AIII.6b)}$$

where $\hat{F}$ is an operator function of the operators $\hat{\underset{\sim}{q}}(t)$, $\hat{\underset{\sim}{p}}(t)$, which may also depend explicitly on the time. In particular, when $\hat{F} = \hat{H}$, eqs. (AIII.6) give

$$\frac{d\hat{H}}{dt} = \frac{\partial\hat{H}}{\partial t} \ , \qquad \text{(AIII.7a)}$$

$$-\nabla_{\underset{\sim}{q}}\hat{H} = \frac{1}{i\hbar}[\hat{\underset{\sim}{p}},\hat{H}] = \frac{d\underset{\sim}{p}(t)}{dt} \ . \qquad \text{(AIII.7b)}$$

In the form of eq. (AIII.7b) the Heisenberg equation of motion for $\hat{\underset{\sim}{p}}(t)$ is seen to be the quantized expression of Hamilton's equation of motion, eq. (AIII.6).[64] Using the momentum representation and the same procedure as above, it can also be shown that the unitary operator

$$\hat{U}(\delta\hat{\underset{\sim}{p}}) = \hat{1} - \frac{i}{\hbar}\delta\underset{\sim}{p}\cdot\hat{\underset{\sim}{q}}(t)$$

generates the infinitesimal displacements $\delta\underset{\sim}{p}$ in the momentum space, whereby one obtains

$$\nabla_{\underset{\sim}{p}}\hat{F} = -\frac{1}{i\hbar}[\hat{F},\hat{\underset{\sim}{q}}(t)] \quad . \tag{AIII.6c}$$

When $\hat{F} = \hat{H}$, eqs. (AIII.6a) and (AIII.6c) give the quantized Hamilton equation of motion for $\hat{\underset{\sim}{q}}(t)$:

$$\nabla_{\underset{\sim}{p}}\hat{H} = \frac{1}{i\hbar}[\hat{\underset{\sim}{q}}(t),\hat{H}] = \frac{d\hat{\underset{\sim}{q}}(t)}{dt} \quad . \tag{AIII.7c}$$

Consider the transformation function connecting two eigenstates of the coordinate operator $\hat{\underset{\sim}{q}}$, at t and t + dt, respectively. By eqs. (AIII.1) and (AIII.4a)

$$<\underset{\sim}{q}',t + dt|\underset{\sim}{q}'',t> = <\underset{\sim}{q}',t|[\hat{1} - \frac{i}{\hbar}\hat{H}dt]|\underset{\sim}{q}'',t> \tag{AIII.8}$$

Schwinger's quantum action principle states that the variation of the transformation function $<\underset{\sim}{q}',t + dt|\underset{\sim}{q}'',t>$ is determined by the variation of an operator, $\hat{W}_{t,t+dt}$, called the action operator, which we shall shortly define:[64]

$$\delta<\underset{\sim}{q}',t+dt|\underset{\sim}{q}'',t> = \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|\delta\hat{W}_{t,t+dt}|\underset{\sim}{q}'',t> \tag{AIII.9}$$

The variations involved in eq. (AIII.9) may be of two types, which Schwinger classified as dynamical and kinematical changes, respectively.[64] A dynamical change, denoted by δ'', involves a variation of the time end-points t,t+dt or an alteration of the structure of the Hamiltonian, an instance arising in, e.g., a perturbation problem. In contrast, a kinematical change, indicated by the symbol δ', is one in which only the operators $\hat{\underset{\sim}{q}}(t)$, $\hat{\underset{\sim}{q}}(t+dt)$ are subjected to variations. These variations induce the unitary transformations of generators $\hat{G}_{\delta\underset{\sim}{q}(t)}$, $\hat{G}_{\delta\underset{\sim}{q}(t+dt)}$, respectively, where $\hat{G}_{\delta\underset{\sim}{q}(t)}$ has been defined in eq. (AIII.6b). They, in turn, cause the basis vector $|\underset{\sim}{q}",t>$, and the dual vector $|\underset{\sim}{q}',t+dt>$ to vary by

$$\delta'(|\underset{\sim}{q}",t>) = -\frac{i}{\hbar}\delta\underset{\sim}{q}(t)\cdot\hat{\underset{\sim}{p}}(t)|q",t> , \qquad \text{(AIII.10a)}$$

$$\delta'(<\underset{\sim}{q}',t+dt|) = \frac{i}{\hbar}\delta\underset{\sim}{q}(t+dt)\cdot<\underset{\sim}{q}',t+dt|\hat{\underset{\sim}{p}}(t+dt) .$$

Thus, to first order, the change in the transformation function, $<\underset{\sim}{q}',t+dt|\underset{\sim}{q}",t>$, is

$$\begin{aligned}\delta'<\underset{\sim}{q}',t+dt|\underset{\sim}{q}",t> &= \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|[\hat{\underset{\sim}{p}}(t+dt)\cdot\delta\underset{\sim}{q}(t+dt) - \hat{\underset{\sim}{p}}(t)\delta\underset{\sim}{q}(t)]|\underset{\sim}{q}",t> \\ &= \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|[\hat{G}_{\delta\underset{\sim}{q}(t+dt)} - \hat{G}_{\delta\underset{\sim}{q}(t)}]|\underset{\sim}{q}",t> . \qquad \text{(AIII.11)}\end{aligned}$$

Using the Heisenberg equation of motion, eq. (AIII.7b), we write

$$\hat{\underset{\sim}{p}}(t+dt) = \hat{\underset{\sim}{p}}(t) + \frac{1}{i\hbar}dt[\hat{\underset{\sim}{p}},\hat{H}] , \qquad \text{(AIII.12)}$$

so that eq. (AIII.11) gives

$$\begin{aligned}\delta'<\underset{\sim}{q}',t+dt|\underset{\sim}{q}",t> &= \frac{i}{\hbar}[<\underset{\sim}{q}',t+dt|\{\hat{\underset{\sim}{p}}(t)\cdot(\delta\underset{\sim}{q}(t+dt) - \delta\underset{\sim}{q}(t))|\underset{\sim}{q}",t> \\ &\qquad + <\underset{\sim}{q}',t+dt|\frac{1}{i\hbar}[\hat{\underset{\sim}{p}}\cdot\delta\underset{\sim}{q},\hat{H}dt]|\underset{\sim}{q}",t>] \\ &= \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|\delta'\{\hat{\underset{\sim}{p}}(t)\cdot d\underset{\sim}{q}(t) - dt\hat{H}\}|\underset{\sim}{q}",t> \\ &= \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|\delta'\hat{W}_{t,t+dt}|\underset{\sim}{q}",t> . \qquad \text{(AIII.13)}\end{aligned}$$

In obtaining the second identity of eq. (AIII.13), use has been made of the definition of the variation δ', as a kinematical change, which implies

$$\delta'\hat{\underset{\sim}{p}}(t) = 0 \; , \qquad \text{(AIII.14a)}$$

$$\delta'\hat{H} = \frac{1}{i\hbar}[\hat{H},\underset{\sim}{p}\cdot\delta\underset{\sim}{q}] \; . \qquad \text{(AIII.14b)}$$

Also, we have introduced the differential of $\underset{\sim}{q}(t)$, defined as

$$d\underset{\sim}{q}(t) = \underset{\sim}{q}(t+dt) - \underset{\sim}{q}(t) = \frac{d\underset{\sim}{q}(t)}{dt}dt \; , \qquad \text{(AIII.14c)}$$

and defined the action operator $\hat{W}_{t,t+dt}$ by

$$\hat{W}_{t,t+dt} = \hat{\underset{\sim}{p}}d\underset{\sim}{q} - dt\hat{H} \; . \qquad \text{(AIII.14d)}$$

Extending the variation δ' to include variations in $\underset{\sim}{p}$ as well, we obtain an additional term which is the matrix element of the operator expression

$$\delta\underset{\sim}{p}(t)[\hat{\underset{\sim}{q}}(t+dt) - \hat{\underset{\sim}{q}}(t)] - dt\nabla_{\underset{\sim}{p}}\hat{H}\cdot\delta\underset{\sim}{p}(t) \; , \qquad \text{(AIII.15)}$$

between the state vectors $|\underset{\sim}{q}'',t\rangle$, $|\underset{\sim}{q}',t+dt\rangle$. In eq. (AIII.15) the variation $\delta\underset{\sim}{p}(t)$ is a vector of n numbers, just as $\delta\underset{\sim}{q}(t)$ in eqs. (AIII.10)-(AIII.14). Using equations (AIII.7c) and (AIII.14c) the operator expression of eq. (AIII.15), which is readily identified as the variation of $\hat{W}_{t,t+dt}$ ensuing from changes in $\underset{\sim}{p}$ alone, is seen to vanish identically. Thus, eq. (AIII.13) holds for all variations, (δ'), of the dynamical variables $\underset{\sim}{q}(t)$ and their conjugate momenta, $\underset{\sim}{p}(t)$. We now verify that this result is also obtained for variations of the dynamical type. By definition, such a variation involves either changes in the time end-point t,t+dt, and/or a change in the structure of the Hamiltonian. We first consider the case where t and t' = t + dt are subject to variations δt and δt', respectively, the Hamiltonian being left invariant. By virtue of eqs. (AIII.4) these variations induce the following changes in $<\underset{\sim}{q}',t'|$ and $|\underset{\sim}{q}",t>$:

$$\delta"(<\underset{\sim}{q}',t'|) = -\frac{i}{\hbar}\delta t'<\underset{\sim}{q}',t'|\hat{H}(\underset{\sim}{\hat{q}}(t'),\hat{p}(t'),t') , \quad \text{(AIII.16a)}$$

$$\delta"(|\underset{\sim}{q}",t>) = \frac{i}{\hbar}\delta t\hat{H}(\hat{q}(t),\hat{p}(t),t)|\underset{\sim}{q}",t> . \quad \text{(AIII.16b)}$$

Therefore, the ensuing variation in $<\underset{\sim}{q}',t'|\underset{\sim}{q}",t>$ is

$$\delta"<\underset{\sim}{q}',t'|\underset{\sim}{q}",t> = \frac{i}{\hbar}<\underset{\sim}{q}',t'|[\hat{H}(t)\delta t - \hat{H}(t')\delta t']|\underset{\sim}{q}",t> \cdot \quad \text{(AIII.16c)}$$

Since t' = t + dt, and dt is an infinitesimal, eq. (AIII.16c) can also be written as

$$\begin{aligned}\delta"<\underset{\sim}{q}',t+dt|\underset{\sim}{q}",t> &= -\frac{i}{\hbar}<\underset{\sim}{q}',t+dt|\hat{H}(t)\delta"(dt)|\underset{\sim}{q}",t> \\ &= \frac{i}{\hbar}<\underset{\sim}{q}',t+dt|\delta"[-dt\hat{H}]|\underset{\sim}{q}",t> , \quad \text{(AIII.17)}\end{aligned}$$

where δ"(dt) = δt' − δt. The last equality sign in eq. (AIII.17) was obtained since, by hypothesis, the variation δ" does not affect the Hamiltonian. On the other hand, consider a dynamical change in which only the structure of the Hamiltonian is altered. The accompanying variation in the transformation function $<\underset{\sim}{q}',t+dt|\underset{\sim}{q}",t>$ is readily evaluated, using eq. (AIII.8)

$$\delta''\langle \underset{\sim}{q}',t+dt|\underset{\sim}{q}'',t\rangle = -\frac{i}{\hbar}\langle \underset{\sim}{q}',t|(\delta''\hat{H})dt|\underset{\sim}{q}'',t\rangle = \frac{i}{\hbar}\langle \underset{\sim}{q}',t+dt|\delta''(-dt\hat{H})|\underset{\sim}{q}'',t\rangle ,$$

where, in the second identity, we have replaced $|\underset{\sim}{q}',t\rangle$ by $\langle \underset{\sim}{q}',t+dt|$ thereby introducing a term of second-order, which is inconsequential in the present variational context. Thus, eq. (AIII.17) holds for both types of dynamical changes, and is the general result for a variation depicted by δ''. Since in these variations, the term $\underset{\sim}{p}(t)d\underset{\sim}{q}(t)$ is invariant, i.e., $\delta''(\underset{\sim}{p}d\underset{\sim}{q}) = 0$, eq. (AIII.17) may also be written in the form of eq. (AIII.13):

$$\delta''\langle \underset{\sim}{q}',t+dt|\underset{\sim}{q}'',t\rangle = \frac{i}{\hbar}\langle \underset{\sim}{q}',t+dt|\delta''[\underset{\sim}{p}\cdot d\underset{\sim}{q} - dt\hat{H}]|\underset{\sim}{q}'',t\rangle$$

$$= \frac{i}{\hbar}\langle \underset{\sim}{q}',t+dt|\delta''\hat{W}_{t,t+dt}|\underset{\sim}{q}'',t\rangle \ . \qquad \text{(AIII.18)}$$

Combining eqs. (AIII.13) and (AIII.18), we obtain Schwinger's quantum action principle, eq. (AIII.9), where the variation, δ, includes both dynamical and kinematical changes.

The action operator

$$\hat{W}_{t,t+dt} = \hat{\underset{\sim}{p}}\cdot d\underset{\sim}{q} - dt\hat{H} = \hat{L}dt$$

defines the quantal Lagrangian operator $\hat{L}$, which is of the same form as the classical Lagrangian given in eq. (AII.2). We now show that for a finite time interval $[t_1,t_2]$, the transformation function $\langle \underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle$ also obeys the quantum action principle,[64] i.e.

$$\delta\langle \underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle = \frac{i}{\hbar}\langle \underset{\sim}{q}',t_1|\delta\hat{W}_{12}|\underset{\sim}{q}'',t_2\rangle , \qquad \text{(AIII.19a)}$$

with the quantum action integral defined by

$$\hat{W}_{12} = \int_{t_2}^{t_1} dt\hat{L} \quad . \qquad \text{(AIII.19b)}$$

For this purpose, we decompose the time interval $[t_1,t_2]$ into n intervals of equal lengths, dt. We shall choose n to be sufficiently large so that dt is an infinitesimal. To simplify the notations, we shall also denote the endpoints of the k^{th} interval by $t^{(k-1)}$, and $t^{(k)}$ $(t^{(k)} > t^{(k-1)})$. Thus $t_1 = t^{(0)}$, $t_2 = t^{(n)}$, and

$$[t_1,t_2] = [t^{(0)},t^{(1)}[U[t^{(1)},t^{(2)}[U....U[t^{(n-1)},t^{(n)}] \, .$$

For each k = 1,...,n-1, we use the completeness relation[2,64]

$$\int d\underset{\sim}{q}^{(k)} |\underset{\sim}{q}^{(k)},t^{(k)}><\underset{\sim}{q}^{(k)},t^{(k)}| = \hat{1} \quad , \qquad \text{(AIII.20)}$$

to write the transformation function $<\underset{\sim}{q}',t_1|\underset{\sim}{q}",t_2>$ in the form

$$<\underset{\sim}{q}',t_1|\underset{\sim}{q}",t_2> = <\underset{\sim}{q}',t_1|\int d\underset{\sim}{q}^{(1)}|\underset{\sim}{q}^{(1)},t^{(1)}><\underset{\sim}{q}^{(1)},t^{(1)}|\int d\underset{\sim}{q}^{(2)}|\underset{\sim}{q}^{(2)},t^{(2)}>$$

$$<\underset{\sim}{q}^{(2)},t_2^{(2)}|x \ldots \int d\underset{\sim}{q}^{(n-1)}|\underset{\sim}{q}^{(n-1)},t^{(n-1)}><\underset{\sim}{q}^{(n-1)},t^{(n-1)}|\underset{\sim}{q}",t_2>$$

$$= \int d\underset{\sim}{q}^{(1)}d\underset{\sim}{q}^{(2)} \,..\, d\underset{\sim}{q}^{(n-1)}<\underset{\sim}{q}',t^{(0)}|\underset{\sim}{q}^{(1)},t^{(1)}>$$

$$<\underset{\sim}{q}^{(1)},t^{(1)}|\underset{\sim}{q}^{(2)},t^{(2)}>x \ldots <\underset{\sim}{q}^{(n-1)},t^{(n-1)}|\underset{\sim}{q}",t^{(n)}> \, . \qquad \text{(AIII.21)}$$

In differential form, the multiplicative law of composition of the transformation functions, eq. (AIII.21), becomes

$$\delta<\underset{\sim}{q}',t_1|\underset{\sim}{q}",t_2> = \int d\underset{\sim}{q}^{(1)}\delta(<\underset{\sim}{q}',t^{(0)}|\underset{\sim}{q}^{(1)},t^{(1)}>)<\underset{\sim}{q}^{(1)},t^{(1)}|\underset{\sim}{q}",t^{(n)}>$$

$$+\sum_{k=1}^{n-2}\int d\underset{\sim}{q}^{(k)}d\underset{\sim}{q}^{(k+1)}\langle\underset{\sim}{q}',t^{(0)}|\underset{\sim}{q}^{(k)},t^{(k)}\rangle\;\delta(\langle\underset{\sim}{q}^{(k)},t^{(k)}|\underset{\sim}{q}^{(k+1)},t^{(k+1)}\rangle)\times\langle\underset{\sim}{q}^{(k+1)},t^{(k+1)}|\underset{\sim}{q}'',t^{(n)}\rangle$$

$$+\int d\underset{\sim}{q}^{(n-1)}\langle\underset{\sim}{q}',t^{(0)}|\underset{\sim}{q}^{(n-1)},t^{(n-1)}\rangle\times\delta(\langle\underset{\sim}{q}^{(n-1)},t^{(n-1)}|\underset{\sim}{q}'',t^{(n)}\rangle). \quad \text{(AIII.22}$$

Equation (AIII.22) gives the change in $\langle\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle$ explicitly in terms of the variations in the transformation functions $\langle\underset{\sim}{q}^{(k)},t^{(k)}|\underset{\sim}{q}^{(k+1)},t^{(k+1)}\rangle$, $k = 0,\ldots,n-1$. Since the time interval $[t^{(k)},t^{(k+1)}]$ is an interval of infinitesimal length, dt, eq. (AIII.9) is applicable to each of these transformation functions. Thus, we have

$$\delta\langle\underset{\sim}{q}^{(k)},t^{(k)}|\underset{\sim}{q}^{(k+1)},t^{(k+1)}\rangle=\frac{i}{\hbar}\langle\underset{\sim}{q}^{(k)},t^{(k)}|\delta\hat{W}_{t^{(k)},t^{(k+1)}}|\underset{\sim}{q}^{(k+1)},t^{(k+1)}\rangle. \quad \text{(AIII.23)}$$

Using the last result and the completeness relations, eq. (AIII.20), in eq. (AIII.22), we obtain

$$\begin{aligned}\langle\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle &=\frac{i}{\hbar}\langle\underset{\sim}{q}',t_1|\sum_{k=0}^{n-1}\delta\hat{W}_{t^{(k)},t^{(k+1)}}|\underset{\sim}{q}'',t_2\rangle\\ &=\frac{i}{\hbar}\langle\underset{\sim}{q}',t_1|\delta[\sum_{k=0}^{n-1}\hat{W}_{t^{(k)},t^{(k+1)}}]|\underset{\sim}{q}'',t_2\rangle\,. \quad \text{(AIII.24)}\end{aligned}$$

In the limit dt → 0, the sum $\sum\hat{W}_{t^{(k)},t^{(k+1)}}$ becomes the time integral of $\hat{L}$, thus yielding the result of eqs. (AIII.19). We have thus completed the derivation of Schwinger's action principle. We now consider a corollary of this principle; this corollary is called the principle of stationary action.[64] Referring to eqs. (AIII.11) and (AIII.16), we note that, when the changes involved in the variation of $<\underset{\sim}{q}^{(k)},t^{(k)}|\underset{\sim}{q}^{(k+1)},t^{(k+1)}>$ leave the Hamiltonian form-invariant, the variation in $\hat{W}_{t^{(k)},t^{(k+1)}}$ equals the difference in the values at $t^{(k)}$ and $t^{(k+1)}$ of an operator function, $\hat{G}(t)$, which is the generator of a unitary transformation

$$\delta\hat{W}_{t^{(k)},t^{(k+1)}} = \hat{G}(t^{(k)}) - \hat{G}(t^{(k+1)}) \quad ; \qquad \text{(AIII.25a)}$$

$$\hat{G}(t) = \hat{\underset{\sim}{p}}(t)\delta\underset{\sim}{q}(t) - \hat{H}(\underset{\sim}{q}(t),\underset{\sim}{p}(t),t)\delta t \ . \qquad \text{(AIII.25b)}$$

Thus,

$$\delta\hat{W}_{12} = \delta\left(\sum_{k=0}^{n-1}\hat{W}_{t^{(k)},t^{(k+1)}}\right) = \sum_{k=0}^{n-1}\delta\hat{W}_{t^{(k)},t^{(k+1)}}$$

$$= \hat{G}(t_1) - \hat{G}(t_2) \quad ; \qquad \text{(AIII.26)}$$

Eq. (AIV.26) is the mathematical expression of Schwinger's principle of stationary action: the variation of the action integral is unaffected by the details of the changes occurring at intermediate times; it is wholly determined by the values of the generator of the unitary transformation, which causes these changes, at the time end-points t_1,t_2. From the above derivation, this statement of the principle of stationary action is seen to be the consequence of the more general quantum action principle, which is obtained only if the variations do not affect the Hamiltonian, i.e., the physical nature, of the system. Eq. (AIII.26) is to be compared with its classical counterpart, eq. (AII.16). It is seen that $\hat{W}_{12}$ exhibits the same variational properties at W^{Cl}. Corresponding to the generator of the canonical transformation defined by δW^{Cl}, the variation in $\hat{W}_{12}$ determines the generator of a unitary transformation, which governs the changes in the quantized dynamical variables, $\underset{\sim}{\hat{q}}(t)$, and their conjugate momenta, $\underset{\sim}{\hat{p}}(t)$, via the commutator relationships

$$\delta\underset{\sim}{\hat{q}}(t) = \frac{1}{i\hbar}[\underset{\sim}{\hat{q}}(t),\hat{G}(t)] \quad , \qquad \text{(AIII.27a)}$$

$$\delta\underset{\sim}{\hat{p}}(t) = \frac{1}{i\hbar}[\underset{\sim}{\hat{p}}(t),\hat{G}(t)] \quad . \qquad \text{(AIII.27b)}$$

Schwinger has also shown that the principle of stationary action embodies all the laws of the quantum mechanics of an unperturbed system.[64] Indeed, using the calculus of variations, one may obtain the variation in $\hat{W}_{12}$ directly from its definition, equation (AIII.19b)

$$\delta\hat{W}_{12} = \int_{t_2}^{t_1}\{\delta\underset{\sim}{p}\cdot[d\underset{\sim}{q} - \nabla_{\underset{\sim}{p}}\hat{H}dt] + [-d\underset{\sim}{p} - \nabla_{\underset{\sim}{q}}\hat{H}dt]\cdot\delta\underset{\sim}{q} - \delta t[\frac{\partial\hat{H}}{\partial t}dt - \frac{d\hat{H}}{dt}dt]\}$$
$$+ \{\underset{\sim}{p}\cdot\delta\underset{\sim}{q} - \hat{H}\delta t\}\Big|_{t_2}^{t_1} \quad , \qquad \text{(AIII.28)}$$

where $\underset{\sim}{p}$, $\underset{\sim}{q}$ and t are considered to be varied independently. Requiring $\hat{W}_{12}$ to be stationary, in the sense of eq. (AIII.26) we obtain the Euler equations

$$\frac{d\underset{\sim}{q}}{dt} = \nabla_{\underset{\sim}{p}}\hat{H}; \qquad \frac{d\underset{\sim}{p}}{dt} = -\nabla_{\underset{\sim}{q}}\hat{H}; \qquad \frac{d\hat{H}}{dt} = \frac{\partial\hat{H}}{\partial t} \quad , \qquad \text{(AIII.29)}$$

which we can compare with eqs. (AIII.7c), (AIII.7b) and (AIII.7a) respectively. Thus, the (quantized) Hamilton equations of motion are obtained as the Euler equations of the variational problem posed by eq. (AIII.26). On the other hand, we also identify the generator $\hat{G}(t)$ as

$$\hat{G}(t) = \hat{\underset{\sim}{p}}(t)\delta\underset{\sim}{q}(t) - \hat{H}(t)\delta t \ .$$

Considering purely temporal changes, we set $\delta\underset{\sim}{q}(t) = 0$, for all t, thereby obtaining the generator of time displacement $\hat{G}_{\delta t}(t) = -\hat{H}(t)\delta t$. Eqs. (AIII.27) (or eq. (AIII.6a), with $\hat{F} = \hat{\underset{\sim}{q}}(t)$ and $\hat{F} = \hat{\underset{\sim}{p}}(t)$, respectively) then identify the Hamilton equations of motion, (AIII.29) with the Heisenberg equations of motion, as has been done in eqs. (AIII.7b and c). Finally, for purely kinematical changes, we set $\delta t = 0$ to obtain the generator $\hat{G}_{\delta\underset{\sim}{q}(t)}$, previously defined in eq. (AIII.5b). Using $\hat{G} = \hat{G}_{\delta\underset{\sim}{q}(t)}$ in eqs. (AIII.27) we get the well-known commutator relationships,

$$\begin{aligned} &[\hat{q}_i(t),\hat{p}_i(t)] = i\hbar; \\ &[\hat{p}_i(t),\hat{p}_i(t)] = 0. \end{aligned} \tag{AIII.30}$$

The above derivations of the Heisenberg equations of motion and the commutation rules illustrate the power of the principle of stationary action, and by extension, of the quantum action principle.

Above, the quantum action principle has been derived by explicitly writing the wavefunction in the coordinate representation. The principle then gives a differential characterization of the transformation functions $\langle\underset{\sim}{q}',t_1|\underset{\sim}{q}'',t_2\rangle$

connecting basis vectors of the coordinate representation. Exactly the same considerations can be given regarding the variations of the transformation functions in the momentum representation. We simply replace $\hat{\underset{\sim}{q}}(t)$, $\underset{\sim}{q}'$ by $\hat{\underset{\sim}{p}}(t)$, $\underset{\sim}{p}'$ in eqs. (AIII.3). The operator $\nabla_{\underset{\sim}{q}'}$ is accordingly replaced by $-\nabla_{\underset{\sim}{p}'}$, so that eq. (AIII.3c) is replaced by

$$<\underset{\sim}{p}',t|\underset{\sim}{q}(t) = -\frac{\hbar}{i}\nabla_{\underset{\sim}{p}'}<\underset{\sim}{p}',t| \ . \qquad \text{(AIII.31)}$$

Corresponding to eq.(AIII.11), we have

$$\delta_p'<\underset{\sim}{p}',t+dt|\underset{\sim}{p}'',t> = \frac{i}{\hbar}<\underset{\sim}{p}',t+dt|-[\hat{\underset{\sim}{q}}(t+dt)\delta\underset{\sim}{p}(t+dt) - \hat{\underset{\sim}{q}}(t)\delta\underset{\sim}{p}(t)]|\underset{\sim}{p}'',t>$$

$$= \frac{i}{\hbar}<\underset{\sim}{p}',t+dt|\hat{G}_{\delta\underset{\sim}{p}(t+dt)} - \hat{G}_{\delta\underset{\sim}{p}(t)}|\underset{\sim}{p}'',t> \ , \qquad \text{(AIII.32a)}$$

where

$$\hat{G}_{\delta\underset{\sim}{p}(t)} = -\delta\underset{\sim}{p}(t)\hat{\underset{\sim}{q}}(t) \qquad \text{(AIII.32b)}$$

is the generator of the uniform translations $\delta\underset{\sim}{p}(t)$ in momentum space. In eq. (AIII.32a) δ_p' denotes a kinematical variation in the momentum space, a variation which involves neither a change in the time end-points of the transformation function $<\underset{\sim}{p}',t+dt|\underset{\sim}{p}'',t>$, nor an alteration of the dynamics of the system through a change in its Hamiltonian. Using

$$\hat{\underset{\sim}{q}}(t+dt) = \hat{\underset{\sim}{q}}(t) + \frac{1}{i\hbar}dt[\hat{\underset{\sim}{q}}(t),\hat{H}]$$

$$\frac{1}{i\hbar}[-\hat{\underset{\sim}{q}}\cdot\delta\underset{\sim}{p},\hat{H}dt] = \delta_p'[-dt\hat{H}]$$

and

$$\delta \underset{\sim}{p}(t+dt) - \delta \underset{\sim}{p}(t) = \delta'_p (d\underset{\sim}{p}(t)),$$

we obtain, from eq. (AIII.32a)

$$\delta'_p \langle \underset{\sim}{p}', t+dt | \underset{\sim}{p}'', t \rangle = \frac{i}{\hbar} \langle \underset{\sim}{p}', t+dt | \delta'_p \hat{W}^{(p)}_{t,t+dt} | \underset{\sim}{p}'', t \rangle , \tag{AIII.33a}$$

where the action operator $\hat{W}^{(p)}_{t,t+dt}$ is defined by

$$\hat{W}^{(p)}_{t,t+dt} = -\hat{\underset{\sim}{q}}(t) \cdot d\underset{\sim}{p}(t) - dt\hat{H} . \tag{AIII.33b}$$

For a variation of the dynamical type, the same arguments as those used previously in deriving eqs. (AIII.17), (AIII.18), can be repeated, leading to

$$\delta'' \langle \underset{\sim}{p}', t+dt | \underset{\sim}{p}'', t \rangle = \frac{i}{\hbar} \langle \underset{\sim}{p}', t+dt | \delta'' \hat{W}^{(p)}_{t,t+dt} | \underset{\sim}{p}'', t \rangle , \tag{AIII.34}$$

so that, for a general variation, $\delta^{(p)}$, which includes both dynamical and kinematical changes, we have

$$\delta^{(p)} \langle \underset{\sim}{p}', t+dt | \underset{\sim}{p}'', t \rangle = \frac{i}{\hbar} \langle \underset{\sim}{p}', t+dt | \delta^{(p)} \hat{W}^{(p)}_{t,t+dt} | \underset{\sim}{p}'', t \rangle . \tag{AIII.35}$$

Furthermore, if $\delta^{(p)}$ does not alter the structure of the Hamiltonian,

$$\delta^{(p)} \hat{W}^{(p)}_{t,t+dt} = \hat{F}(t+dt) - \hat{F}(t) , \tag{AIII.36}$$

where $\hat{F}(t)$ is the generator of an infinitesimal unitary transformation responsible for the variation denotes by $\delta^{(p)}$.

Using the completeness relation

$$\int d\underset{\sim}{p}^{(k)} | \underset{\sim}{p}^{(k)}, t^{(k)} \rangle \langle \underset{\sim}{p}^{(k)}, t^{(k)} | = \hat{1} \tag{AIII.37}$$

at the time $t^{(k)}$ arising in a partitioning of a finite time interval $[t_1, t_2]$ into infinitesimal intervals, we generalize the quantum action principle, eq. (AIII.35), to write

$$\delta^{(p)}\langle \underset{\sim}{p}', t_1 | \underset{\sim}{p}'', t_2 \rangle = \frac{i}{\hbar}\langle \underset{\sim}{p}', t_1 | \delta^{(p)} \hat{W}^{(p)}_{12} | \underset{\sim}{p}'', t_2 \rangle , \qquad \text{(AIII.38a)}$$

where

$$\hat{W}^{(p)}_{12} = \int_{t_2}^{t_1} [-\hat{q}\,d\underset{\sim}{p} - \hat{H}dt] . \qquad \text{(AIII.38b)}$$

Accordingly, the principle of stationary action, eq. (AIII.36), is generalized to read

$$\delta^{(p)} \hat{W}^{(p)}_{12} = \hat{F}(t_1) - \hat{F}(t_2) . \qquad \text{(AIII.39)}$$

The generator $\hat{F}(t)$ appearing in eqs. (AIII.36) and (AIII.39) can be either $-\hat{\underset{\sim}{q}}\delta\underset{\sim}{p}$ or $\hat{H}\delta t$; it may also include both types of transformation.

Through these considerations, it appears that Schwinger's quantum action principle and the principle of stationary action, applicable whenever the variation $\delta^{(p)}$ does not alter the structure of the Hamiltonian, can be derived in any representation; only the explicit expression of the quantum action integral changes with the representation. This remark is exploited in Section C of Chapter III, where we connect Schwinger's action principle with the Schrödinger variational principle.

REFERENCES AND FOOTNOTES

1. R.P. Feynman, R.B. Leighton, M. Sands, "The Feynman Lectures on Physics", Vol. 1, Addison-Wesley, Reading, Massachusetts (1964), p. I-2.
2. G.N. Lewis, J. Am. Chem. Soc., 38, 762 (1916).
3. W. Heitler and F. London, Z. Physik, 44, 455 (1927).
4. L. Pauling, "The Nature of the Chemical Bonds", 3rd Ed., Cornell University Press, Ithaca, N.Y. (1960).
5. J.C. Slater, Phys. Rev., 37, 481 (1931); ibid., 38, 1109 (1931); ibid. 41, 255 (1932).
6. J.E. Lennard-Jones, J. Chem. Phys., 20, 1024 (1952).
7. J.E. Lennard-Jones, Proc. Roy. Soc. (London), A198, 1, 14 (1949).
8. J.E. Lennard-Jones and J.A. Pople, Proc. Roy. Soc. (London), A202, 166 (1950); J.A. Pople, Quart. Rev. London, 11, 273 (1957).
9. C. Edmiston and K. Ruedenberg, Rev. Mod. Phys., 35, 457 (1960); W. England, L.S. Salmon and K. Ruedenberg, in "Topics in Current Chemistry", Springer-Verlag, New York, N.Y. (1971), p. 31.
10. S.F. Boys, Rev. Mod. Phys., 32, 296 (1960); J.M. Foster and S.F. Boys, Rev. Mod. Phys., 32, 300 (1960).
11. H. Weinstein, R. Pauncz and M. Cohen, Adv. Atomic Mol. Phys., 7, 97 (1971).

12. Initial efforts in this laboratory were devoted to the partitioning of a molecular system by surfaces of zero-flux in the gradient vector of the charge distribution (Refs. 13-15). The bond-paths were introduced later (Ref. 16). The unified description of molecular systems in terms of both the virial partitioning and the bond-paths is achieved through the detailed topological study of the charge density (Refs. 17-19).

13. R.F.W. Bader and P.M. Beddall, J. Am. Chem. Soc., 95, 305 (1973); R.F.W. Bader and R.R. Messer, Can. J. Chem., 52, 2268 (1974); R.F.W. Bader and G.R. Runtz, Mol. Phys., 30, 117 (1975); G.R. Runtz and R.F.W. Bader, Mol. Phys., 30, 129 (1975); R.F.W. Bader, Acc. Chem. Res., 8, 34 (1975).

14. G.R. Runtz, "A Partitioning Scheme for Molecular Electronic Charge Distribution", Ph.D. Thesis, McMaster University (1973).

15. R.R. Messer, "Virial Partitioning: Theory and Applications", Ph.D. Thesis, McMaster University (1977).

16. G.R. Runtz, R.F.W. Bader and R.R. Messer, Can. J. Chem., 16, 3040 (1977).

17. R.F.W. Bader, S.G. Anderson and A.J. Duke, J. Am. Chem. Soc., 101, 1389 (1979).

18. R.F.W. Bader, T.T. Nguyen-Dang and Y. Tal, J. Chem. Phys., 70, 4316 (1979).

19. R.F.W. Bader, Y. Tal, S.G. Anderson and T.T. Nguyen-Dang, Isr. J. Chem., 19, 8 (1980).

20. R.F.W. Bader and P.M. Beddall, J. Chem. Phys., 56, 3320 (1972); R.F.W. Bader, P.M. Beddall and J. Peslak, Jr., J. Chem. Phys., 58, 557 (1973).

21. The major part of this section is a review of elementary notions of Quantum Mechanics, as they are currently being applied to Chemistry. These notions can be found in any textbook on Quantum Chemistry, for instance in, I.N. Levine, "Quantum Chemistry", 2nd Edition, Allyn and Bacon, Inc., Boston, Mass. (1970). Also, see ref. 22.

22. A. Messiah, "Quantum Mechanics", Vol. II, North-Holland, Amsterdam (1958); P. Roman, "Advanced Quantum Theory", Addison-Wesley, Reading, Mass. (1965).

23. P.O. Löwdin, Phys. Rev., 97, 1474 (1955). A density matrix of order p is defined in this paper with the normalization factor $\binom{n}{p}$. Elsewhere, the normalization $\binom{n}{p}p!$ is also adopted, e.g., in R. McWeeny, Rev. Mod. Phys., 32, 335 (1960).

24. A thorough discussion of the mathematical properties of density matrices can be found in A.J. Coleman, Rev. Mod. Phys., 35, 668 (1963). A complete review of the role of density matrices in Chemistry and an extensive bibliography on the subject can be found in E.R. Davidson, "Reduced Density Matrices in Quantum Chemistry", Academic Press, New York, N.Y. (1976).

25. H. Kummer, J. Math. Phys., 8, 2063 (1967).

26. P. Hohenberg and W. Kohn, Phys. Rev. B, 136, 864 (1964).

27. W. Kohn and L.J. Sham, Phys. Rev. A, 140, 1133 (1965); B.Y. Tong and L.J. Sham, Phys. Rev. A, 144, 1 (1966); R.G. Gordon and Y.S. Kim, J. Chem. Phys., 56, 3122 (1972).

28. T.L. Gilbert, Phys. Rev. B, 12, 2111 (1975); M. Berrondo and O. Goscinski, Int. J. Quant. Chem., Symp. No. 9, 67 (1975); R.A. Donnelly and R.G. Parr, J. Chem. Phys., 69, 4431 (1978); R.G. Parr, R.A. Donnelly, M. Levy and W.E. Palke, J. Chem. Phys., 68, 3802 (1978); R.G. Parr, S.R. Gadre and L. Bartolotti, Proc. Natl. Acad. Sci. USA, 76, 2522 (1979); M. Levy, Proc. Natl. Acad. Sci. USA, 76, 6062 (1979).

29. M. Born and J.R. Oppenheimer, Ann. Phys., 84, 457 (1927); Also see M. Born and K. Huang, "Dynamical Theory of Crystal Lattices", Oxford University Press, Oxford, England (1954), Appendix VIII.

30. W.D. Hobey and A.D. McLachlan, J. Chem. Phys., 33, 1695 (1960); J.C. Tully and R.K. Preston, J. Chem. Phys., 54, 4297 (1971); R.F.W. Bader and R.A. Gangi, in "Theoretical Chemistry", Vol. 2, R.N. Dixon and C. Thomson, Editors, The Chemical Society, London, England (1975), pp. 1-65.

31. Although the topological analysis developed in Chapters I and II can be applied to an exact (non-adiabatic) charge distribution, its illustrations by specific examples implicitly assume the Born-Oppenheimer separation. Also, we shall occasionally use various notions derived from

the Born-Oppenheimer Approximation, such as the notions of potential energy surface, minimum energy path of a reaction, and equilibrium (nuclear) configuration - or geometry - of a system, even in discussions pertaining to a non-adiabatic charge distribution.

32. This terminology originates from Catastrophe Theory (see refs. 40, 41, 46, 51-53).

33. J.R. Van Wazer and I. Absar, "Electron Densities in Molecules and Molecular Orbitals", Academic Press, New York, N.Y. (1975).

34. K. Collard and G.G. Hall, Int. J. Quantum Chem., <u>12</u>, 623 (1977).

35. T. Kato, Commun. Pure Appl. Math., <u>10</u>, 151 (1957).

36. W.A. Bingel, Z. Naturforsch. Teil <u>A18</u>, 1249 (1963); E. Steiner, J. Chem. Phys., <u>39</u>, 2365 (1963).

37. A function $f:M \rightarrow N$, where M and N are differentiable manifolds, is smooth if it is continuous and infinitely differentiable at all points of M (that is, with respect to a local coordinates system at $m \in M$, the partial derivatives of f exist and are continuous to any order, at m).[46] In our applications, in practice we need only require differentiability to a finite order larger than two.

38. Two functions $f,g:M \rightarrow R$ are topologically equivalent if there exist homeomorphisms $h:M \rightarrow M$ and $k:R \rightarrow R$, such that $fh = kg$.

39. We use the terminology of ref. 34. Elsewhere, what is presently called a critical point is also termed "stationary

points", "rest points", or "singularity". We shall later employ the name "singularity" to denote a critical point of a particular type.

40. R. Thom, "Structural Stability and Morphogenesis", Benjamin, Reading, Mass. (1975).

41. D.B. O'Shea, "An Exposition of Catastrophe Theory and its Applications to Phase Transitions", Queen's Papers in Pure and Applied Mathematics No. 47, Queen's University, Kingston, Ontario (1978); A.J. Coleman and D.B. O'Shea, Phys. Rev. B, to be published.

42. The mathematical analysis of differential equations through the local qualitative properties of their solutions, as adopted in this section, is detailed in refs. 43, 44.

43. M. Hirsch and S. Smale, "Differential Equations, Dynamical Systems, and Linear Algebra", Academic Press, New York, N.Y. (1974).

44. S. Lefschetz, "Differential Equations: Geometric Theory", 2nd Ed., Dover Publ., New York, N.Y. (1977).

45. Details of the determination of the wavefunctions for the systems illustrated in Chapters I and II are given in Appendix I.

46. Y.C. Lu, "Singularity Theory and an Introduction to Catastrophe Theory", Springer-Verlag, New York, N.Y. (1976).

47. M. Ojha, Senior Thesis, Department of Chemistry, McMaster University (1979).

48. We note that in addition to the condition $d_{(ij)} + d_{(ik)} + d_{(jk)} = 3d$, which defines the plane P, the distances $d_{(ij)}$, $i,j = 1,2,3$, must also satisfy the triangular inequality

$$d_{(ij)} + d_{(ik)} \geq d_{(jk)}.$$

Thus, only the portion of P enclosed by the triangle whose corners are at $(d_1 = \frac{3d}{2} = d_2,\ d_3 = 0)$, $(d_1 = \frac{3d}{2} = d_3,\ d_2 = 0)$ and $(d_1 = 0,\ d_2 = \frac{3d}{2} = d_3)$ represents physically realizable situations.

49. A function $h: W \to W'$ is a diffeomorphism iff i) h is bijective, and ii) h and h^{-1} are smooth (infinitely differentiable).

50. The equilibrium geometry of the H_3^+ system has been determined experimentally by Gaillard et al to be equilaterally triangular in shape: M.J. Gaillard et al, Phys.Rev. A17, 1797 (1978). Recently, experimental spectral data for this molecular ion have also been reported: T. Oka, Phys.Rev.Lett., 45, 531 (1980); J.T.Shy et al, Phys.Rev.Lett., 45, 535 (1980). The equilateral nuclear configuration illustrated in Fig. 2.3a has been found to deliver the minimum in the ground-state electronic energy function $E_o(X)$ by theoretical calculations. Christoffersen has proposed that the structure of this molecule at its equilibrium geometry is of the type illustrated in Fig. 2.3b: R.E. Christoffersen, J. Chem. Phys. 41, 960 (1964); J. Chem. Phys., 48, 1790 (1968).

51. A thorough account of Thom's classification theorem can be found in ref. 46, and in E.C. Zeeman and D.J.A. Trotman,

Lecture notes in Math. No. 525, Springer-Verlag, New York, N.Y. (1967), p. 263.

52. T. Poston and I. Stewart, "Catastrophe Theory and Its Applications", Pitman, London (1978).

53. T. Poston and A.E.R. Woodcock, Lecture notes in Math. No. 373, Springer-Verlag, New York, N.Y. (1974).

54. R.G. Woolley, Israel J. Chem., 19, 30 (1980); J. Am. Chem. Soc., 100, 1073 (1978); Chem. Phys. Lett., 55, 443 (1978); Int. J. Quant. Chem., XII, S1, 307 (1977); Adv. Phys., 25, 27 (1976).

55. E.B. Davies, "Quantum Theory of Open Systems", Academic Press, New York, N.Y. (1976); W. Peier, Physica, 57, 565 (1972).

56. B.L. Moiseiwitsch, "Variational Principles", Interscience Publ., New York, N.Y. (1966).

57. H. Goldstein, "Classical Mechanics", Addison-Wesley, Reading, Mass. (1965).

58. S.T. Epstein, "The Variation Method in Quantum Chemistry", Academic Press, New York, N.Y. (1974).

59. Generalization to a non-adiabatic wavefunction (i.e., going beyond the Born-Oppenheimer approximation) is straightforward: in this case, the primed integral appearing in the definition of any subsystem functional ($G_e'[\Omega,\psi]$ or $F_E[\Omega,\psi]$, or $\mathcal{L}[\Omega,\psi]$ of section C) also includes the integration over the nuclear coordinates and spins.

60. S. Srebrenik and R.F.W. Bader, J. Chem. Phys., 63, 3945 (1975).

61. A. Mazziotti, R.G. Parr and G. Simons, J. Chem. Phys., 59, 939 (1973).

62. S. Srebrenik, R.F.W. Bader and T.T. Nguyen-Dang, J. Chem. Phys., 68, 3667 (1978).

63. R. Courant and D. Hilbert, "Methods of Mathematical Physics", Vol. 1, Interscience Publ., Inc., New York, N.Y. (1953), pp. 164-274.

64. J. Schwinger, in "Brandeis University 1960 Summer Institute in Theoretical Physics, Lecture Notes", (1960), pp. 223-372. The quantum action principle was initially introduced and employed by Schwinger in his discussions on quantized field: J. Schwinger, Phys. Rev., 82, 914 (1951); Phys. Rev., 91, 713 (1953).

65. R.F.W. Bader, S. Srebrenik and T.T. Nguyen-Dang, J. Chem. Phys., 68, 3680 (1978).

66. This statement is true only if the operator $\hat{A}$ does not involve $\hat{\nabla}_{\underset{\sim}{X}}$, the momentum operators associated with the nuclei.

67. The regional hypervirial theorem was first obtained by Epstein, in the form of the vanishing of the regional average of $[\hat{H},\hat{A}]$. The condition for such a regional hypervirial theorem to be obtained is precisely the vanishing of the r.h.s. of eq. (3.32). See S.T. Epstein,

J. Chem. Phys., 60, 3351 (1974).

68. J.A. Schweitz, J. Phys. A, 10, 507 (1977).

69. J.A. Schweitz, J. Phys. A, 10, 517 (1977).

70. R.F.W. Bader, accepted for publication in J. Chem. Phys.

71. S.T. Epstein, J. Chem. Phys., 63, 3573 (1975).

72. H. Hellmann, "Einfuhrung in die Quantenchemie", Deuticke, Leipzig, Germany (1937), p. 285; R.P. Feynman, Phys. Rev., 56, 340 (1939).

73. P.M. Morse and H. Feshbach, "Methods of Theoretical Physics", Part 1, McGraw-Hill, New York, N.Y. (1953), Chapter 2.

74. P. Ehrenfest, Z. Physik, 45, 455 (1927).

75. We have also derived variational equations for a subsystem described in field theory (see ref. 65). In this description, the electronic structure of a molecular system is expressed in terms of the field operators $\hat{\psi}(\underset{\sim}{r},t)$, $\hat{\psi}^{+}(\underset{\sim}{r},t)$. These operators can be written in terms of the creation and annihilation operators $\hat{a}_i$, $\hat{a}_i^{+}$ associated with single-particle basis states ϕ_i (spin orbitals) as follows[77-79]

$$\hat{\psi}(\underset{\sim}{r},t) = \sum_i \hat{a}_i(t)\phi_i(\underset{\sim}{r},t), \qquad \hat{\psi}^{+}(\underset{\sim}{r},t) = \sum_i \hat{a}_i^{+}(t)\phi_i^{*}(\underset{\sim}{r},t) \tag{1}$$

To appropriately describe a system of identical fermions by this approach, customarily referred to as the "second-quantization formalism", the operators $\hat{a}_i$, $\hat{a}_i^{+}$ must satisfy the anticommutation rules at equal times

$$\{\hat{a}_i(t),\hat{a}_j^+(t)\} = \hat{a}_i\hat{a}_j^+ + \hat{a}_j^+\hat{a}_i = i\hbar\delta_{ij}, \quad \forall\, i,j$$

$$\{\hat{a}_i(t),\hat{a}_j(t)\} = 0 = \{\hat{a}_i^+,\hat{a}_j^+\}, \quad \forall\, i,j \qquad (2)$$

A single-determinantal state of a given system (with a given number of electrons) is then identified as a particular excitation state of the field, obtained by letting an appropriate combination of the operators $\hat{a}_i$, $\hat{a}_i^+$ act on the vacuum state of the field. By the transformation (1), this state can also be written in terms of the field operators $\hat{\psi},\hat{\psi}^+$. Corresponding to the Hartree-Fock equations governing the basis functions ϕ_i, one then has the field equation

$$-i\hbar\dot{\hat{\psi}}(\underset{\sim}{r},t)-\frac{\hbar^2}{2m}\nabla^2\hat{\psi}(\underset{\sim}{r},t)+\hat{v}\hat{\psi}(\underset{\sim}{r},t)+\left(\int\hat{\psi}^+(\underset{\sim}{r}',t)\hat{u}(\underset{\sim}{r},\underset{\sim}{r}')\hat{\psi}(\underset{\sim}{r}',t)d\underset{\sim}{r}'\right)\hat{\psi}(\underset{\sim}{r},t) = 0 \qquad (3)$$

where $\hat{v}(\underset{\sim}{r},t)$ denotes the interaction of the field with an external source (in here, the nuclear point charges), and $\hat{u}(\underset{\sim}{r},\underset{\sim}{r}')$ the interaction of the field with itself, corresponding to the two-body interactions between the excitations of the field. This field equation can be regarded as the Euler equation of the variational problem derived from Schwinger's quantum action principle,[64] when it is applied to this Schrödinger's field system. In ref. 65, we have considered the restriction of the field action integral to a subsystem, which is defined by the second-quantized version of the zero-flux boundary condition:

$$\forall r \in S, \qquad \nabla\hat{\rho}\cdot\underset{\sim}{n} = 0, \qquad \hat{\rho} = \psi^{+}\psi$$

We have obtained the field equation (3), together with the field-theoretic Hamiltonian operator associated with the subsystem, upon the requirement that the subsystem action operator satisfies the second-quantized version of the variational principle, which was derived in Chapter III (section C) of this Thesis.

76. Y. Tal, R.F.W. Bader and J. Erkku, Phys. Rev. A, 21, 1 (1980).

77. P. Roman, (see ref. 22)

78. G. Baym, "Lectures on Quantum Mechanics" , W.A.Benjamin, Inc., Reading, Massachusetts (1977).

79. H.C. Longuet-Higgins, in "Quantum Theory of Atoms, Molecules and the Solid State", P.O.Lowdin, Editor, Academic Press, New-York, N.Y. (1966), pp 105-121.