From Schrödinger to atoms in molecules

R.F.W. Bader

Department of Chemistry, McMaster University, Hamilton, Ontario, Canada L8S 4M1

ABSTRACT

The properties of atoms in molecules can be measured experimentally and quantum mechanics predicts these atomic properties just as it predicts the properties of a total system. This paper presents a demonstration of this statement and an example of its application to the measured energy of an atom in a molecule.

INTRODUCTION

The quantum description of an atom in a molecule which exists in a stationary state can be obtained as a direct extension of Schrödinger's original derivation of the wave equation [1] through the introduction and use of operators which are the generators of infinitesimal unitary transformations. Such operators, as first pointed out by Dirac [2], are the quantum counterparts of the generators of infinitesimal canonical transformations in classical mechanics and the correspondence between the properties of these two sets of generators is responsible for the fundamental similarity in the structures of classical and quantum mechanics. Because of this correspondence, Schwinger [3] was able to state a single principle – the principle of stationary action – from which either classical or quantum mechanics can be derived in its entirety. In his first paper, Schrödinger demonstrated that the postulated 'quantum conditions' could be replaced by a variation principle from which the quantum numbers are obtained in a natural way. Schwinger's principle of stationary action is obtained as a generalization of the variation of the action integral through a relaxation of the constraint that the variations vanish at the time end points and their subsequent identification with the generators of infinitesimal unitary transformations. For a system in a stationary state, the variation of the action integral reduces to the variation of an energy integral which is identical to that first constructued by Schrödinger and used by him to derive the wave equation for a stationary state. We shall illustrate that for a time independent system, Schwinger's principle of stationary action generalizes Schrödinger's derivation of the wave equation to yield the hypervirial theorem as well. The further generalization of this result to a subsystem of a total system in a stationary state is obtained by retaining the variations in the state function on the (finite) boundaries of the subsystem and by a variation of the boundary itself. The boundary variations are again identified with the generators of infinitesimal unitary transformations. What is remarkable and physically important is that this final generalization is possible only for a particular class of subsystems, those which satisfy a variational constraint that is stated in terms of a property of a system's charge distribution [4,5,6].

We shall first review Schrödinger's derivation of the wave equation as given in the first of the four papers he published in 1926. This is followed by a discussion of Schwinger's principle of stationary action and of how it is obtained through the use of generators of infinitesimal transformations. The properties of these generators and their associated transformations are illustrated. They are then employed in the generalization of Schrödinger's derivation of quantum mechanics, first for a total system and then for a subsystem. The paper concludes by using the resulting quantum description of the properties of a subsystem to define the energy of an atom in a molecule. It is then demonstrated that this energy is the one determined experimentally, one which can be measured in those cases where there is an observed additivity in the heats of formation.

SCHRÖDINGER'S DERIVATION OF QUANTUM MECHANICS

Schrödinger concludes his first paper [1] by stating that 'the function Ψ be such as to make the "Hamilton integral"

$$\int dr \{ \hbar^2 T(q, \partial \Psi / \partial q) + \Psi^2 V \} \tag{1}$$

stationary, while fulfilling the normalising, accessory condition

$$\int \Psi^2 d\tau = 1 \quad' \tag{2}$$

$T(q, \partial \Psi / \partial q)$ is the kinetic energy expressed as a function of the coordinates and momenta and V is the potential energy. Schrödinger considered specifically the problem of the hydrogen atom for which the explicit form of the functional is

$$J[\Psi] = \int d\tau \{ (\hbar^2/2m) \nabla \Psi \cdot \nabla \Psi - (E - V) \Psi^2 \} \tag{3}$$

with $V = -e^2/r$. The quantity E in eqn (3) is the Lagrange multiplier introduced to ensure the normalisation constraint as expressed in eqn (2). Expressing the integrand in eqn (3) by the symbol f, the variation of J using the standard methods of the calculus of variations is given by the expression

$$\delta J[\Psi] = \int d\tau \{ (\partial f / \partial \Psi) \delta \Psi + (\partial f / \partial \nabla \Psi) \delta \nabla \Psi \} \tag{4}$$

The variation with respect to $\nabla \Psi$ yields a term containing $\delta \nabla \Psi$ and this must be replaced by an equivalent one expressed in terms of $\delta \Psi$ using integration by parts as outlined below

$$\nabla \cdot \{ (\nabla \Psi) \delta \Psi \} = \nabla^2 \Psi \delta \Psi + (\nabla \Psi) \cdot \delta \nabla \Psi \tag{5}$$

Requiring the resulting variation to be stationary gives

$$\delta J[\Psi] = 2 \int \{ -(\hbar^2/2m) \nabla^2 \Psi + (V - E) \Psi \} \delta \Psi d\tau + 2(\hbar^2/2m) \oint dS \nabla \Psi \cdot n \delta \Psi = 0 \tag{6}$$

Demanding that this requirement be met for all arbitrary variations $\delta \Psi$ yields

$$\hat{H}\Psi - E\Psi = 0 \tag{7}$$

as the Euler–Lagrange equation if, in addition, one requires the surface term to vanish

$$\oint dS \nabla \Psi \cdot n \delta \Psi = 0 \tag{8}$$

by demanding that $\delta \Psi$ vanish on the boundaries of the system which are taken to be at infinity. Schrödinger demonstrated that the variation problem had a discrete ($E < 0$) and a continuous ($E > 0$) spectrum of proper values. For the bound states of the hydrogen atom, eqn (7) had solutions only if E satisfied a particular condition. This condition introduced the principal quantum number labelled "ℓ" by Schrödinger, which together with the angular momentum quantum number which he labelled "n", demonstrated that for each value of "ℓ" there was a degeneracy of acceptable solutions equal to ℓ^2.

In the fourth paper Schrödinger [7] gave his time dependent equation for the state function

$$i\hbar\dot{\Psi} = \hat{H}\Psi \qquad\qquad -i\hbar\dot{\Psi}^* = \hat{H}\Psi^* \tag{9}$$

and noted that in general Ψ could be complex. Thus in the present day variation of the integral $J[\Psi]$, the factor of two in equation (6) would be gone and two separate and independent variations would be performed, one with respect to Ψ and the other with respect Ψ^*, variations which would in turn yield the Euler equations for both Ψ and Ψ^*.

In this same paper he discusses the physical significance of the field scalar Ψ and stresses 'that Ψ cannot and may not be interpreted directly in terms of three–dimensional space –'. Instead, he points out that the product $\Psi^*\Psi$ must determine the 'electric density'. The electric density for a N–electron system he defines as

$$\text{electric density at } r_1 = eN \int dr_2 \int dr_3 \ldots \int dr_N \Psi^*\Psi \tag{10}$$

which, aside from a summation over the electronic spin coordinates, is the definition of what is presently called the electronic charge density and denoted here by the symbol $\rho(r)$. It is most interesting to note that Schrödinger regarded $\Psi^*\Psi$ as providing a description of the actual spreading out of electronic charge in real space – as a description of a system's static charge distribution: 'As an obvious generalization of the procedure of spreading out the electronic charge according to a relative density function $\Psi^*\Psi$ (which furnished satisfactory results in the one–electron problem), the following view would present itself in the case of a general mechanical system: the real natural system does not behave like the picture which ordinary mechanics forms of it (e.g. a system of point–charges in a definite configuration), but rather behaves like what would be the result of spreading out the system, described by $q_1, \ldots, q_n$, throughout its configuration–space in accordance with a relative density function $\Psi^*\Psi$' [8]. This fourth paper also introduced the current density j and showed that $\rho(r)$ and $j(r)$ obey the conservation law for a fluid.

FROM SCHRÖDINGER TO SCHWINGER

The work of Schrödinger, Heisenberg, Dirac and others led to the Hamiltonian formulation of quantum mechanics. This approach, in addition to the general mathematical scheme of linear operators and state vectors with its associated probability interpretation,

contains two distinct postulates: the commutation relations between generalized coordinates and momenta and the equations of motion as introduced with the array of assumptions based on the classical Hamiltonian dynamics and the correspondence principle. In the development of quantum mechanics given by Schwinger [3], both postulates are replaced by a single dynamical principle based upon the Lagrangian formulation of mechanics.

In the Lagrangian approach, one obtains the equations of motion, classical or quantum, as the Euler equations resulting from the requirement that the action integral W_{12}, the action integral operator in the quantum case, be stationary with respect to variations in the classical trajectory or the quantum state function respectively, with the added constraint that these variations vanish at the time end points t_1 and t_2

$$\delta W_{12} = \delta \int_{t_1}^{t_2} \mathcal{L}[\Psi,t]dt = 0 \qquad (11)$$

where $\mathcal{L}[\Psi,t]$ is the Lagrangian. Schwinger demonstrated that one can obtain a complete description of mechanics through a generalization of the requirement that W_{12} be stationary. This generalization corresponds to the removal of the constraint that the variations δq or $\delta \Psi$ vanish at the time end points together with the introduction of a variation of the time end points themselves, followed by the demonstration that these end point variations may be identified with the generators of infinitesimal canonical or unitary transformations. The implementation of these steps leads to the principle of stationary action which is

$$\delta W_{12} = G(t_2) - G(t_1) \qquad (12)$$

where $G(t)$ denotes a generator of an infinitesimal canonical or unitary transformation. The equations of motion are still contained in this principle. The generator G has the form

$$G(t) = (\delta \mathcal{L}/\delta \dot{q})\delta q - H\delta t \qquad (13)$$

The two terms appearing in the generator encompass all possible dynamical changes that can occur in a mechanical system, both spatial and temporal. To appreciate the beauty and power of this principle requires an understanding of the properties of the generators G.

PROPERTIES OF INFINITESIMAL UNITARY TRANSFORMATIONS

A unitary transformation is one which when applied to both the state functions and the observables of a system, leaves the description of the system unchanged. Denote by $\hat{U}$ an operator with the property that its adjoint is equal to its inverse, that is

$$\hat{U}\hat{U}^+ = \hat{U}^+\hat{U} = 1 \qquad \text{or} \qquad \hat{U}^{-1} = \hat{U}^+ \qquad (14)$$

The operator $\hat{U}$ is used to perform simultaneous transformations on the state vector Ψ and on the observables $\hat{O}$ in the manner indicated in equations (15)

$$\Psi \rightarrow \Psi' = \hat{U}\Psi$$
$$\hat{O} \rightarrow \hat{O}' = \hat{U}\hat{O}\hat{U}^{-1} \qquad (15)$$

Such a transformation leaves the description of the system completely unchanged in that all properties of the transformed system are the same as those of the original one. One easily shows that an operator equation of the original system

$$\hat{O}\Psi = o\Psi$$

is transformed into a corresponding equation with the same eigenvalue

$$\hat{O}'\Psi' = o\Psi'$$

and the expectation value of an operator remains unchanged as well

$$\langle \Psi',\hat{O}'\Psi'\rangle = \langle \hat{U}\Psi\hat{U}\hat{O}\hat{U}^{-1}\hat{U}\Psi\rangle = \langle \Psi,\hat{O}\Psi\rangle$$

as does the normalization of Ψ.

Our interest is in such a transformation which is infinitely close to unity as defined by the operator

$$\hat{U} = \hat{1} - i\epsilon\hat{G} \qquad (16)$$

where ϵ is a real infinitesimal quantity and $\hat{G}$ is a Hermitian operator. Since $\hat{G}$ is Hermitian, the inverse of $\hat{U}$ which equals $\hat{U}^+$ is

$$\hat{U}^{-1} = \hat{1} + i\epsilon\hat{G} \qquad (17)$$

and the product $\hat{U}\hat{U}^{-1}$ *is of order unity to first-order in* ϵ

$$\hat{U}\hat{U}^{-1} = \hat{1} + \epsilon^2\hat{G}\hat{G} \approx 1$$

The operator $\hat{U}$ induces an infinitesimal unitary transformation. An operator $\hat{O}$ is transformed in the following way under such a transformation

$$\hat{O} \rightarrow \hat{O}' = (\hat{1} - i\epsilon\hat{G})\hat{O}(\hat{1} + i\epsilon\hat{G}) = \hat{O} + i\epsilon[\hat{O},\hat{G}] \tag{18}$$

The infinitesimal change in $\hat{O}$, the variation in $\hat{O}$, is determined by its commutator with $\hat{G}$, the generator of the infinitesimal transformation according to

$$\delta\hat{O} = \hat{O}' - \hat{O} = i\epsilon[\hat{O},\hat{G}] \tag{19}$$

Similarly, for the state function one obtains

$$\Psi \rightarrow \Psi' = \hat{U}\Psi = \Psi - i\epsilon\hat{G}\Psi \tag{20}$$

and the infinitesimal change or variation in Ψ is given by action of the generator of the transformation on Ψ

$$\delta\Psi = \Psi' - \Psi = -i\epsilon\hat{G}\Psi \tag{21}$$

Thus the operator $\hat{G}$ can be considered to generate changes in the dynamical variables of a system. The infinitesimal change in the expectation value of an observable $\hat{A}$ is equivalently effected by the action of the generator $\hat{G}$ on Ψ or on the operator $\hat{A}$, for using the above expressions one finds

$$\langle\delta\Psi,\hat{A}\Psi\rangle + \langle\Psi,\hat{A}\delta\Psi\rangle = i\epsilon\langle\Psi,[\hat{G},\hat{A}]\Psi\rangle = -\langle\Psi,\delta\hat{A}\Psi\rangle \tag{22}$$

This general property of an infinitesimal unitary transformation will appear in the result to be obtained in the generalization of the variation of Schrödinger's energy functional.

Infinitesimal unitary transformations are important from another point of view: as first pointed out by Dirac they provide the means for establishing a fundamental correspondence between classical and quantum mechanics. Exploring this correspondence in its classical terms will provide a deeper understanding of the meaning of such transformations in the quantum case.

Corresponding to the unitary transformation in a quantum system is a canonical or contact transformation in classical mechanics. This is a transformation of the position and momentum coordinates such that Hamilton's equations of motion

$$\dot{q}_i = \partial H\backslash\partial p_i \qquad \text{and} \qquad \dot{p}_i = -\partial H\backslash\partial q_i \tag{23}$$

are satisfied using the transformed set of coordinates. The system and its properties are not changed by this transformation: all properties, while perhaps having a different functional form, have the same values when evaluated in the new coordinate system. A canonical transformation is therefore, like a unitary transformation in this regard. Canonical transformations are defined in terms of so-called generating functions which are functions of one of the original set of coordinates q and p and one of the new set, which we shall label as P and Q, as well as of the time [9]. There are therefore, four such generating functions and their individual properties are dertermined by a modification of Hamilton's principle. We are interested in the particular transformation generated by the function $F(q_i,P_i,t) = \Sigma_i q_i P_i$. The equations for obtaining the old momenta and the new coordinates from this function are

$$p_i = \partial F/\partial q_i \qquad \text{and} \qquad Q_i = \partial F/\partial P_i$$

and from these equations one sees that this particular function generates the identity transformation with $p_i = P_i$ and $Q_i = q_i$.

We are now in a position to define in analogy with the infinitesimal unitary transformation of quantum mechanics, the infinitesimal canonical transformation of classical mechanics. We consider the transformations

$$Q_i = q_i + \delta q_i$$
$$P_i = p_i + \delta p_i$$

for which the generating function F is given by

$$F = \sum_i q_i P_i + \epsilon G(q,P)$$

As in the quantum case, G is referred to as the generator of the transformation. Using the equations given above one finds

$$p_i = \partial F/\partial q_i = P_i + \epsilon\partial G/\partial q_i$$

$$Q_i = \partial F/\partial P_i = q_i + \epsilon \partial G/\partial P_i$$

In terms of these results, one finds that the infinitesimal changes in the coordinates to first-order are

$$\delta p_i = P_i - p_i = -\epsilon \partial G/\partial q_i \qquad (24)$$

$$\delta q_i = Q_i - q_i = \epsilon \partial G/\partial p_i \qquad (25)$$

where the partial derivative with respect to P_i has been replaced by one with respect to p_i in the latter equation as the results are to first-order only.

We shall use the classical expressions in equations (24) and (25) to illustrate that such infinitesimal transformations generate real changes in a system. They describe transformations which while leaving the functional form of a property unchanged, change its value by an infinitesimal amount. As a most important example of this property consider the results obtained when G is set equal to the Hamiltonian H and ϵ to dt. From equations (24) and (25) one finds that δq and δp are given by

$$\delta q_i = dt(\partial H/\partial p_i) = \dot{q}_i dt = dq_i \quad \text{and} \quad \delta p_i = -dt(\partial H/\partial q_i) = \dot{p}_i dt = dp_i$$

showing that the transformation changes the values that q and p have at time t to the values they have at time t + dt. Thus Hdt is the generator of an infinitesimal temporal change in a classical system. One may envisage the time evolution of a mechanical system as being the result of the successive application of the generator Hdt, so that its motion corresponds to the unfolding of a canonical transformation [2,9].

In general, the change in a property A(p,q) caused by an infinitesimal canonical transformation is

$$\delta A = A(q_i + \delta q_i, \ p_i + \delta p_i) - A(q_i, p_i)$$

Performing a Taylor series expansion and keeping terms to first-order only, one finds

$$\delta A = \sum_i \{(\partial A/\partial q_i)\delta q_i + (\partial A/\partial p_i)\delta p_i\}$$

which, using the defining equations (24) and (25) may be re-expressed as

$$\delta A = \epsilon \sum_i \{(\partial A/\partial q_i)(\partial G/\partial p_i) - (\partial A/\partial p_i)(\partial G/\partial q_i)\} \qquad (26)$$

Eqn (26) is of course the definition of the Poisson bracket for the quantities A and G, a quantity denoted by the symbol {A,G} as in the equation

$$\delta A = \epsilon\{A,G\} \qquad (27)$$

This latter equation is in direct correspondence with the quantum result given in eqn (19) for the variation of the observable $\hat{O}$ in terms of the commutator of $\hat{O}$ and the generator $\hat{G}$. The root of this correspondence, as noted by Dirac, derives from the similarity in the properties of infinitesimal unitary and canonical transformations. This similarity is a reflection of a structure which is common to both mechanics [2]. The quantity Hdt is the generator of an infinitesimal temporal transformation in both a classical and a quantum system. Using the Heisenberg representation for the quantum system, both mechanics yield the important result that if

$$\delta H = \epsilon\{H,G\} = (i/\hbar)\epsilon[\hat{H},\hat{G}] = 0$$

then the property G is a constant of the motion. In this view, constants of motion are generating functions of infinitesimal canonical transformations which leave H invariant.

The generator $\hat{G}$ defined in the principle of stationary action may act on the state function or on the observable corresponding to some property. Staying within the operator representation we shall derive the Heisenberg equations of motion from this principle by choosing a variation in which only the time is changed by an amount δt and δq is zero. The change in an operator $\hat{O}$ is $-(d\hat{O}/dt)\delta t$ and according to eqn (19) which describes the effect of the generator $\hat{H}\delta t$ one has

$$d\hat{O}/dt = -(i/\hbar)[\hat{O},\hat{H}] \qquad (28)$$

Again using eqn (19), this time setting the operator $\hat{O}$ equal to the position operator $\hat{q}_s$, and considering a variation in which only the coordinate $\hat{q}_r$ is changed at a given time, one obtains

$$\delta \hat{q}_r = (i/\hbar)[\hat{q}_s, (\partial \mathscr{L}/\partial \dot{\hat{q}}_r)\delta \hat{q}_r] \qquad (29)$$

The RHS of eqn (29) reduces to $\delta \hat{q}_s$ if and only if

$$[\hat{q}_s, \partial \mathscr{L}/\partial \dot{\hat{q}}_r] = i\hbar\delta_{rs} \qquad (30)$$

Thus the canonical commutation rules are obtained as a consequence of the principle of stationary action.

The principle of stationary action yields the equations of motion and the commutation relationships for a quantum system. It introduces the observables of quantum mechanics by identifying the variations of the action integral with the generators of infinitesimal transformations and yields as a consequence of this identification, the Heisenberg equation of motion for an observable.

An operational statement of the principle of stationary action can be obtained from eqn (12) in the Schrödinger representation by considering the change in action for an infinitesimal time interval and by restricting the action of the generator to the state function (which by eqn (21) gives $\delta\Psi = -(i/\hbar)\epsilon\hat{G}\Psi$). The resulting statement is given in terms of the variation of the Lagrangian integral as

$$\delta\mathcal{L}[\hat{G}\Psi] = (\epsilon/2)(i/\hbar)\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle \tag{31}$$

where $\mathcal{L}[\Psi,t] = \int L(\Psi,\nabla\Psi,\dot{\Psi})d\tau$ and the many-particle Lagrangian density is the invariant hermitian function

$$L(\Psi,\nabla\Psi,\dot{\Psi}) = (i\hbar/2)\{\Psi^*\dot{\Psi} - \dot{\Psi}^*\Psi\} - \{(\hbar^2/2m)\sum_i\nabla_i\Psi^*\cdot\nabla_i\Psi + \hat{V}\Psi^*\Psi\} \tag{32}$$

Equation (31) is a variational derivation of the Heisenberg equation of motion for the observable $\hat{G}$

$$d\langle\hat{G}\rangle/dt = (i/\hbar)\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle \tag{33}$$

Schrödinger's choice of the form for the functional $J[\Psi]$ was based on analogy with the Hamilton-Jacobi differential equation of classical mechanics. Its form with respect to its dependence on only first derivatives of Ψ with respect to a coordinate is important for this is the restriction placed on the Lagrangian. For a stationary state at the point of variation the state function satisfies

$$(i\hbar)\partial\Psi/\partial t = \hat{H}\Psi = E\Psi$$

and $\Psi(r,t) = \exp\{(i/h)Et\}\Psi(r)$. In this situation it is clear that the Lagrangian will reduce to the negative of the integral previously given for the energy functional $J[\Psi]$ which is rewritten here for the many-electron case

$$J[\Psi] = \int d\tau\{(\hbar^2/2m)\sum_i\nabla_i\Psi^*\cdot\nabla_i\Psi + (\hat{V}-E)\Psi^*\Psi\} = -\mathcal{L}[\Psi,t] = -\int d\tau L(\Psi,\nabla\Psi,\dot{\Psi}) \tag{34}$$

where $\int d\tau$ implies integration over all electronic coordinates. For a stationary state the variations in Ψ will be of the type

$$\delta\Psi(r,t) = \exp\{(i/\hbar)\lambda t\}\delta\Psi(r) \tag{35}$$

thereby preserving the separation of the temporal and spatial variables observed at the point of variation. The Lagrangian will be equal to the functional J at every stage of the variation and the two variations are simply related to give

$$-\delta\mathcal{L}[\Psi,t] = \delta J[\Psi] = -(\epsilon/2)(i/\hbar)\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle \tag{36}$$

For a stationary state, $\delta E = \delta J = 0$, and in this case the principle of stationary action yields the hypervirial theorem [10]

$$\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle = 0 \tag{37}$$

Infinitesimal unitary transformations can be used to obtain eqn (37) directly through a variation of the usual stationary state energy functional where $\hat{H}$ is the Hamiltonian

$$\langle\delta\Psi|(\hat{H} - E)|\Psi\rangle + \langle\Psi|(\hat{H} - E)|\delta\Psi\rangle = 0 \tag{38}$$

With the variation $\delta\Psi = -(i/h)\epsilon\hat{G}\Psi$ and eqn (22), one obtains directly the hypervirial theorem, eqn (37), [11]. When $\hat{H}$ contains a real parameter σ such as a nuclear coordinate and $\hat{G}$ corresponds to the operator $\partial/\partial\sigma$, then the above derivation must be modified to include the term $\partial E/\partial\sigma$ for E contains the same parameters as does $\hat{H}$. The result in this case is

$$\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle - \langle\Psi|[E,\hat{G}]|\Psi\rangle = 0 \tag{39}$$

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

$$\langle\Psi|(\partial\hat{H}/\partial\sigma)|\Psi\rangle = \partial E/\partial\sigma \tag{40}$$

Epstein [11] has shown that unrestricted and spin unrestricted Hartree-Fock solutions are invariant to all unitary transformations with one-electron operators $\hat{G}$. Thus the hypervirial theorems for all such $\hat{G}$'s will be satisfied by UHF and SUHF functions. Hurley

[12] first demonstrated that when the set of trial functions is invariant to changes in a parameter σ, the generalized Hellmann-Feynman theorem will be satisfied for σ.

In summary, the principle of stationary action yields the equations of motion and it introduces the observables of quantum mechanics by identifying the variations of the action integral with the generators of infinitesimal unitary transformations. Through the use of these generators, the principle gives the canonical commutation relations and the Heisenberg equations of motion. For a stationary state the principle, as well as yielding the wave equation, gives the hypervirial theorem from which the properties of the observables for a stationary state can be determined.

ATOMIC STATEMENT OF THE PRINCIPLE OF STATIONARY ACTION

It has been demonstrated that Schwinger's principle of stationary action can be extended to a particular class of subsystems of a total system [5,6]. Before demonstrating that the simpler variation of Schrödinger's stationary state functional $J[\Psi]$ can be extended to define an atom in a molecule, we first outline the general case which involves the variation of a subsystem action integral. This is accomplished by first defining a subsystem Lagrangian $L[\Psi,\Omega]$, where Ω denotes the subsystem, in terms of the many-particle Lagrangian function given in eqn (32) as

$$\mathscr{L}[\Psi,\Omega] = \int_\Omega dr \int d\tau' L(\Psi, \nabla\Psi, \dot{\Psi}) \tag{41}$$

The symbol $\int d\tau'$ in eqn (41) denotes a summation over all spin coordinates and the integration over the spatial coordinates of all electrons but one. Since Ψ is antisymmetrized, it matters not which set of electronic coordinates is excluded. The symbol $\int_\Omega dr$ denotes the integration of the remaining electronic coordinates over the subsystem Ω. The variation of the corresponding subsystem action integral $W_{12}[\Omega]$ is generalized to include a variation of the surface bounding the subsystem, a step which in turn requires the retention of the variations in Ψ on the portions of this surface with finite coordinates, in addition to retaining the variations in Ψ at the time end points, the generalization introduced by Schwinger. It is important to note that the latter terms necessarily appear in the variation of the atomic action integral and thus the generalization of the variation of the action integral to a subsystem automatically transforms the variation into Schwinger's principle. The variations in Ψ at the time end points and on the surface of Ω are again identified with the generators of infinitesimal unitary transformations. The result of this very general variation is shown to reduce to the principle of stationary action only if the constraint given in eqn (42) is satisfied at every stage of the variation

$$\delta\left\{\int_\Omega \nabla^2\rho(r)d\tau\right\} = 0 \tag{42}$$

Eqn (42) is satisfied by a subsystem which is bounded by a surface of zero flux in the gradient vector of the charge density, the condition given in eqn (43)

$$\nabla\rho(r)\cdot n(r) = 0 \qquad \text{for every point on the surface } S(\Omega) \tag{43}$$

The atoms of chemistry are bounded by surfaces which satisfy the zero flux condition [13,14], a condition which demands that the integral of the Laplacian of the charge density over the atom vanish as stated in eqn (44)

$$\int_\Omega \nabla^2\rho(r)d\tau = \oint dS(\Omega)\nabla\rho(r)\cdot n(r) = 0 \tag{44}$$

This condition is satisfied by an atom free or bound. The surface of an atom is constantly and continuously deformable in such a way that eqn (44) is always satisfied, even for the virtual changes encountered in the variation of the action integral. The integral of the Lagrangian density of a many-particle quantum system, eqn (32) reduces to the integral given in eqn (44) when the system is described by Schrödinger's equation. Thus the Lagrangian integral of an atom, like that of a total quantum system, always vanishes.

The operational statement of the atomic principle of stationary action corresponding to that given in eqn (31) for a total system is

$$\delta\mathscr{L}[\hat{G}\Psi,\Omega] = (\epsilon/2)\{(i/\hbar)\langle\Psi|[\hat{H},\hat{G}]|\Psi\rangle_\Omega + \text{complex conjugate}\} \tag{45}$$

where the subscript Ω on the commutator average implies the same mode of integration as given in eqn (41). This statement of the principle applies to any region bounded by a surface of zero flux in the gradient vector of the charge density, a condition satisfied by the total system as well as by the atoms of which it is comprised. Eqn (45) thus represents a generalization of quantum mechanics to a subsystem of a total system.

The atomic analogue of the energy functional $J[\Psi]$ is defined as

$$J[\Psi,\Omega] = \int_\Omega dr \int d\tau' \{(\hbar^2/2m)\sum_i \nabla_i\Psi^* \cdot \nabla_i\Psi + (\hat{V} - E)\Psi^*\Psi\} \tag{46}$$

The variation of this integral, including a variation of the surface $S(r)$ bounding the region Ω, yields

$$\delta J[\Psi,\Omega] = \int_\Omega dr \int d\tau' \{\hat{H}\Psi - E\Psi\}\delta\Psi^* + \oint dS(r) \int d\tau' \{(\hbar^2/2m)\nabla\Psi\cdot n(r)\delta\Psi^* + \delta S(r)f(\Psi,\nabla\Psi)\}$$
$$+ \text{ complex conjugate} \tag{47}$$

where $f(\Psi,\nabla\Psi)$ again represents the integrand of the functional J. The variations in Ψ are assumed to vanish when any of the coordinates assume the infinite values associated with the boundary of the total system. Thus the only surface terms to survive are those for the surface bounding the region Ω. One may use the argument from Courant and Hilbert [15] regarding the stationarity of J with respect to the vanishing of $\delta\Psi$ on the boundary, to again obtain Schrödinger's equations for Ψ and Ψ^* as the Euler equations for the variation of $J[\Psi,\Omega]$. One may consider contained in this variation the special case where $\Omega = R^3$ with $\delta\Psi$ thus vanishing on all the boundaries. This gives Schrödinger's original derivation as a particular case of the variation in eqn (47) and one obtains the wave equations for Ψ and Ψ^* as the Euler equations.

The variation of $J[\Psi,\Omega]$ is thus reduced to

$$\delta J[\Psi,\Omega] = \oint dS(r) \int d\tau' \{(\hbar^2/2m)\nabla\Psi\cdot n(r)\delta\Psi^* + \delta S(r)f(\Psi,\nabla\Psi)\} + cc \tag{48}$$

and further progress towards obtaining a general physical result can be made only by removal of the term involving a variation of the surface. Consider, towards this goal, the identity connecting the integrand $f(\Psi,\nabla\Psi)$ of Schrödinger's functional and the energy functional expressed in terms of the Hamiltonian operator $\hat{H}$ as expressed in eqn (49)

$$f(\Psi,\nabla\Psi) + E = (1/2)\{\Psi^*\hat{H}\Psi + (\hat{\Psi H})^*\Psi\} + (\hbar^2/4m)\sum_i \nabla_i^2(\Psi^*\Psi) \tag{49}$$

The subsystem integration of the final term on the RHS of this equation yields

$$(\hbar^2/4m)\int_\Omega dr \int d\tau' \sum_i \nabla_i^2(\Psi^*\Psi) = (\hbar^2/4m)\int_\Omega \nabla^2\rho'(r)dr$$

(50) where $\rho'(r)$ is the electron density divided by N. This result follows from the fact that the volume integral of $\nabla^2(\Psi^*\Psi)$ vanishes for a system with boundaries at infinity because of Green's theorem

$$\ldots \int dr_1 \ldots \{\nabla_i^2(\Psi^*\Psi)\} = \ldots \oint dS(r_i) \ldots \nabla_i(\Psi^*\Psi)\cdot n(r_i) = 0 \tag{50}$$

since both Ψ and $\nabla_i\Psi$, and their complex conjugates, vanish when any electronic coordinate becomes infinite. We also point out here the related identity connecting the two expressions for the kinetic energy

$$-(\hbar^2/4m)\sum_i \{\Psi^*\nabla_i^2\Psi + \Psi\nabla_i^2\Psi^*\} = (\hbar^2/2m)\sum_i \nabla_i\Psi^*\cdot\nabla_i\Psi - (\hbar^2/4m)\nabla^2\rho(r) \tag{51}$$

an expression which is conveniently re-stated in terms of kinetic energy densities using the one-matrix as

$$-(\hbar^2/4m)\{(\nabla^2 + \nabla'^2)\Gamma^{(1)}(r,r')\} = (\hbar^2/2m)\{\nabla\cdot\nabla'\Gamma^{(1)}(r,r')\} - (\hbar^2/4m)\nabla^2\rho(r) \tag{52}$$

Because of eqn (44), it is clear that the integration of either kinetic energy density over a region satisfying the zero flux surface condition as given in eqn (43) will yield the same value for its average kinetic energy.

Substitution of the identity given in eqn (49) into eqn (48) and recalling that at the point of variation Schrödinger's equation applies, the variation of $J[\Psi,\Omega]$ reduces to

$$\delta J[\Psi,\Omega] = (\hbar^2/4m)\oint dS(r)\{\delta S(r)\nabla^2\rho'(r) + 2\int d\tau'\nabla\Psi\cdot n(r)\delta\Psi\} + cc \tag{53}$$

The imposition of the variational constraint given in eqn (42) now enables one to replace the term involving the surface variation with the volume integral of the variation of $\nabla^2\rho(r)$ for one has

$$\delta\{\int_\Omega \nabla^2\rho(r)dr\} = 0 = \int_\Omega \delta\{\nabla^2\rho(r)\}dr + \oint dS(r)\delta S(r)\nabla^2\rho(r) \tag{54}$$

The required variation of the Laplacian of ρ yields only surface terms for the coordinate r and is given by

$$\int_\Omega \delta\{d\rho'(r)\}dr = \oint dS(r)\int d\tau' \{(\nabla\Psi^*)\delta\Psi + \Psi^*\delta(\nabla\Psi)\}\cdot n(r) + cc \tag{55}$$

It is important that no new volume contributions to the variation result from the impositionn of the variational constraint or one would no longer obtain Schrödinger's equation as the Euler equation of the variation. Substitution of this result into eqn (53) yields

$$\delta J[\Psi,\Omega] = (\hbar^2/4m)\oint dS(r)\int d\tau' \{(\nabla\Psi^*)\delta\Psi - \Psi^*\delta(\nabla\Psi)\}\cdot n(r) + cc \tag{56}$$

The single-particle current density is defined as

$$j(r) = (\hbar/2mi)\int d\tau' \{\Psi^*\nabla\Psi - (\nabla\Psi^*)\Psi\} \tag{57}$$

and in terms of this quantity the variation in the atomic functional $J[\Psi,\Omega]$ is seen to be

given by the flux in the infinitesimal quantum current density through the surface bounding the atom

$$\delta J[\Psi,\Omega] = -(i\hbar/2)\oint dS(r)\delta_\Psi j(r) \cdot n(r) + cc \qquad (58)$$

If the variations in Ψ are now identified with the generators of infinitesimal unitary transformations as $\delta\Psi = -(i/h)\epsilon\hat{G}\Psi$, then the expression for the variation of $J[\Psi,\Omega]$ becomes equal to the infinitesimal flux in the current density for the observable $\hat{G}$

$$\delta J[\hat{G}\Psi,\Omega] = -(\epsilon/2)\oint dS(r) j_G(r) \cdot n(r) + cc \qquad (59)$$

where $j_G(r)$ is given by

$$j_G(r) = (\hbar/2mi)\int d\tau' \{\Psi^*\nabla(\hat{G}\Psi) - (\nabla\Psi^*)\hat{G}\Psi\} \qquad (60)$$

The result for the variation of $J[\Psi,\Omega]$ is put into its final form by using the subsystem statement of the Heisenberg equation for an observable $\hat{G}$ for the case of a stationary state which is

$$\{\langle[\hat{H},\hat{G}]\rangle_\Omega + cc\} = -\{i\hbar\oint dS(r) j_G(r).n(r) + cc\} \qquad (61)$$

a result which may also be easily obtained starting from the LHS of eqn (61) using Schrödinger's equation and recalling that $\hat{H}$ is not Hermitian over a subsystem. Eqn (61) is the hypervirial theorem for a subsystem with an arbitrary boundary. With the use of eqn (61), the variation in the atomic functional becomes

$$\delta J[\hat{G}\Psi,\Omega] = -(\epsilon/2)\{(i/\hbar)\langle[\hat{H},\hat{G}]\rangle_\Omega + cc\} \qquad (62)$$

which is the principle of stationary action for a stationary state as given in eqn (36), generalized to a subsystem bounded by a surface of zero flux in the gradient vector of the charge density.

Eqn (62) together with Schrödinger's equation, are obtained through a generalization of the variation of Schrödinger's energy functional. The same description of the properties of a total system afforded by quantum mechanics applies to a particular class of subsystem, one which is bounded by a surface of zero flux in the gradient vector of the charge density – a quantum subsystem.

Eq. (62) is the variational derivation of the hypervirial theorem for a quantum subsystem. Evaluation of $\delta J[\hat{G}\Psi,\Omega]$ for a given $\hat{G}$ yields the surface integral of the flux in j_G through the surface of the subsystem, and equating this to the commutator average, yields the hypervirial theorem for a subsystem, eqn (61). The variational derivation of this theorem as determined by the principle of stationary action is restricted to subsystems which satisfy the variational constraint of a zero flux boundary condition, eqn (43). From eqn (62) one can, as in the quantum description of a total system, derive the Ehrenfest momentum and force theorems, the Hellmann–Feynman and the virial theorems. In general there is a non-vanishing fluctuation in the subsystem average value of an observable. Thus unlike a total system, the subsystem average of the commutator of $\hat{H}$ and an observable $\hat{G}$ in general does not vanish and new relationships are obtained for a subsystem which relate its properties to virtual fluxes in corresponding quantum currents through its bounding surface [16].

We are interested in defining the average energy of a subsystem and for this purpose we need the subsystem statement of the virial theorem. This theorem is obtained [4,6] by setting the generator $\hat{G}$ in eqn (62) equal to the virial operator $\hat{r} \cdot \hat{p}$, the product of the position and momentum coordinates for an electron. This yields

$$2T(\Omega) = -\mathcal{V}(\Omega) \qquad (63)$$

where $T(\Omega)$, the kinetic energy of Ω is obtained by integration of either of the kinetic energy densities given in eqn (52) over Ω. The quantity $\mathcal{V}(\Omega)$ is the virial of the Ehrenfest force exerted on the subsystem. It is expressible in terms of the quantum mechanical stress tensor, which in turn is a functional of the one-matrix $r^{(1)}(r,r')$. The energy of Ω is defined as

$$E(\Omega) = T(\Omega) + \mathcal{V}(\Omega) \qquad (64)$$

The reader is referred to reference [6] for a complete discussion of this topic. It is however, important to note that only for a quantum subsystem can the virial theorem be derived <u>variationally</u> from the principle of stationary action. While one can partially integrate Heisenberg's equation of motion for the operator $\hat{r} \cdot \hat{p}$ over any arbitrarily defined region of space this does not yield an expression which is the analogue of the virial theorem for the total system for two reasons: the kinetic energy is not uniquely defined for a subsystem with arbitrarily defined boundaries and the virial for an arbitrary subsystem contains a spurious contribution, either positive or negative in sign,

arising from the nonvanishing of the integral of $(\hbar^2/4m)\nabla^2\rho$. The vanishing of this integral and the unique definition of a kinetic energy for an atom are both consequences of the quantum boundary condition of zero flux, eqn (43). The difficulty in partitioning a total energy enters in the division of the potential energy of interaction between the subsystems. A force however, is a local quantity and is determined by the divergence of the quantum stress tensor, $-\nabla\bullet\sigma(r)$. The virial of this quantity yields the corresponding potential energy of the electronic charge located at r and in this manner the potential energy is transformed into a local quantity which may be separately averaged over each subsystem and thus, partitioned into separate atomic contributions. Any property M of a total system is expressible as a sum of atomic contributions

$$M = \sum_{\Omega} M(\Omega) \tag{65}$$

TRANSFERABILITY AND ADDITIVITY OF ATOMIC PROPERTIES

The knowledge of chemistry is ordered, classified, and understood by assigning properties to atoms and functional groups and then relating the properties of the total system to those of its constituent atoms. The atoms defined by theory provide the physical basis for this hypothesis in the following way: (1) They are the most transferable pieces of a system that one can define in real space and they thus maximize the transfer of information between molecules at the level of the charge density. (2) The average value of a property for the total system is obtained by summing the atomic averages for the same property over all the atoms in the molecule, eqn (65). (3) The most important characteristic of an atom is that the constancy in its properties, including its contribution to the total energy of a system, is observed to be directly determined by the constancy in its distribution of charge. When the distribution of charge over an atom is the same in two different molecules, i.e., when the atom or some functional grouping of atoms is the same in the real space of two systems, then it makes the same contribution to the total energy in both systems. It is because of the direct relationship betwen the spatial form of an atom and its properties that we are able to identify them in different systems. This relationship has its basis in the observation that the atoms of theory respond only to changes in the total force exerted on their charge distributions and not to changes in the individual contributions to this force, changes which are large even between closely related systems, be they members of a homologous series or chemically similar in structure. If it were not for this property of responding only to some net field (actually the virial of the total Ehrenfest force) rather than to individual potential contributions, there would be no chemically recognizable atoms or functional groups.

The relationship between the form and the properties of an atom are most evident in the limit of an atom being transferable between systems without change. The resulting constancy in its properties, together with the fact that the value of a property for the total system is given by the sum of the atomic contributions (point 2 above), then leads to the existence of atomic or group additivity schemes. The most important and fundamental of these is additivity of the energy, and one of the earliest examples of this was observed for the homologous series of saturated hydrocarbons [17].

It has been shown that the methyl and methylene groups defined by the theory of atoms in molecules account for the additivity of the energy observed in the normal hydrocarbons and for the deviations in this additivity found for small cyclic and bicyclic molecules, deviations which are used as the basis for the experimental definition of a strain energy [18,19,20]. The observed incremental differences in the heats of formation of the hydrocarbons yield measured values for the energies of the methyl and methylene groups. The energies assigned to these groups are independently determined by the theory of atoms in molecules. That this assignment leads to energies for the transferable methyl and methylene groups equal to their incremental values as determined by the total energies, confirms that these are the atoms of chemistry.

The charge distribution of the transferable methylene group of the normal hydrocarbons is found to undergo a small perturbation, corresponding to a shift in charge from the hydrogens to the carbon, when the group is placed in a small ring system with geometric strain. In cyclopropane and cyclobutane for example, these small changes in the charge distribution lead to small increases in the energies of the methylene groups in these two molecules relative to the standard value. The calculated increases correspond to the observed strain energies for these molecules providing further evidence that the atoms of theory recover the measureable properties of the atoms of chemistry. As a final, related example, it is found that the energy of the methylene group in cyclohexane is predicted by theory to be equal to that found for the standard group, a result again in accord with the experimental heats of formation which demonstrate that this molecule does not possess a measureable strain energy. Other properties parallel the constancy of the charge distribution of the transferable methyl and methylene groups as has been demonstrated for

the populations, group dipole (and other) moments, correlation energies and atomic volumes [20]. An atomic volume is another property that has been shown experimentally to be additive in the hydrocarbons.

The properties of atoms in molecules can be measured experimentally and quantum mechanics predicts these properties as it does the properties of the total molecule.

Acknowledgment

REFERENCES

1. E. Schrödinger, Ann. d. Phys. 79, 361 (1926). Referenced from "Collected Papers on Wave Mechanics", E. Schrodinger, together with his four lectures on wave mechanics. Chelsea Pub. Co., New York, N.Y., third ed. 1982.

2. P.A.M. Dirac, "The Principles of Quantum Mechanics", fourth ed., Oxford University Press, Oxford, 1958. pp. 103 to 107; pp. 125 to 130.

3. J. Schwinger, Phys. Rev. 82, 914 (1951).

4. S. Srebrenik and R.F.W. Bader, J. Chem. Phys. 63, 3945 (1975).

5. S. Srebrenik, R.F.W. Bader and T.T. Nguyen-Dang, J. Chem. Phys. 68, 3667 (1978). R.F.W. Bader, S. Srebrenik and T.T. Nguyen-Dang, J. Chem. Phys. 68, 3680 (1978).

6. R.F.W. Bader and T.T. Nguyen-Dang, Ad. Quantum Chem. 14, 63 (1981).

7. E. Schrödinger, Ann. d. Phys. 81, 109 (1926).

8. E. Schrödinger from fourth lecture, p. 205 of collected works in reference 1.

9. H. Goldstein, "Classical Mechanics", Addison-Wesley Pub. Co., Reading, Mass., 1950.

10. J.O. Hirschfelder, J. Chem. phys. 33, 1462 (1960). J.O. Hirschfelder and C.A. Coulson, J. Chem. Phys. 36, 941 (1962).

11. S.T. Epstein, "The Variationn Method in Quantum Chemistry", Academic Press, New York, 1974.

12. A.C. Hurley, Proc. Roy. Soc. A 226, 179 (1954).

13. R.F.W. Bader and P.M. Beddall, J. Chem. Phys. 56, 3320 (1972).

14. R.F.W. Bader, T.T. Nugyen-Dang and Y. Tal, Rep. Prog. Phys. 44, 893 (1981).

15. R. Courant and D. Hilbert, "Methods of Mathematical Physics", vol. I, Interscience Publishers, Inc., New York, 1953, p. 208.

16. R.F.W. Bader and K. Wiberg in "Density Matrices and Density Functionals", Edited by R. Erdahl and V.H. Smith, D. Reidel Pub. Co., Dordrecht, 1987, p. 677.

17. E.J. Prosen, W.H. Johnson and F.D. Rossini, J. Res. Natl. Bur. Stand. 37, 51 (1946); J.L. Franklin, Ind. Eng. Chem. 41, 1070 (1949); S.W. Benson, "Thermochemical Kinetics", Wiley, New York (1968); T.L. Allen, J. Chem. Phys. 31, 1039 (1959).

18. R.F.W. Bader, Can. J. Chem. 64, 1036 (1986).

19. K. Wiberg, R.F.W. Bader and C.D.H. Lau, J. Am. Chem. Soc. 109, 985, 1001 (1987).

20. R.F.W. Bader, A. Larouche, C. Gatti, M.T. Carroll and P.J. MacDougall, J. Chem. Phys., July, (1987).