

DENSITY-FUNCTIONAL THEORY OF ATOMS AND MOLECULES

ROBERT G. PARR and WEITAO YANG

University of North Carolina



OXFORD UNIVERSITY PRESS · NEW YORK
CLARENDON PRESS · OXFORD

1

ELEMENTARY WAVE MECHANICS

1.1 The Schrödinger equation

The principles of density-functional theory are conveniently expounded by making reference to conventional wave-function theory. Therefore, this first chapter reviews elementary quantum theory (Levine 1983, Merzbacher 1970, Parr 1963, McWeeny and Sutcliffe 1969, Szabo and Ostlund 1982). The next chapter summarizes the more advanced techniques that we shall need, mainly having to do with density matrices.

Any problem in the electronic structure of matter is covered by Schrödinger's equation including the time. In most cases, however, one is concerned with atoms and molecules without time-dependent interactions, so we may focus on the time-independent Schrödinger equation. For an isolated N -electron atomic or molecular system in the Born-Oppenheimer nonrelativistic approximation, this is given by

$$\hat{H}\Psi = E\Psi \quad (1.1.1)$$

where E is the electronic energy, $\Psi = \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ is the wave function, and \hat{H} is the Hamiltonian operator,

$$\hat{H} = \sum_{i=1}^N \left(-\frac{1}{2} \nabla_i^2 \right) + \sum_{i=1}^N v(\mathbf{r}_i) + \sum_{i < j}^N \frac{1}{r_{ij}} \quad (1.1.2)$$

in which

$$v(\mathbf{r}_i) = - \sum_{\alpha} \frac{Z_{\alpha}}{r_{i\alpha}} \quad (1.1.3)$$

is the "external" potential acting on electron i , the potential due to nuclei of charges Z_{α} . The coordinates \mathbf{x}_i of electron i comprise space coordinates \mathbf{r}_i and spin coordinates s_i . Atomic units are employed here and throughout this book (unless otherwise specified): the length unit is the Bohr radius $a_0 (= 0.5292 \text{ \AA})$, the charge unit is the charge of the electron, e , and the mass unit is the mass of the electron, m_e . When additional fields are present, of course, (1.1.3) contains extra terms.

We may write (1.1.2) more compactly as

$$\hat{H} = \hat{T} + \hat{V}_{ne} + \hat{V}_{ee} \quad (1.1.4)$$

where

$$\hat{T} = \sum_{i=1}^N \left(-\frac{1}{2}\nabla_i^2\right) \quad (1.1.5)$$

is the kinetic energy operator,

$$\hat{V}_{ne} = \sum_{i=1}^N v(r_i) \quad (1.1.6)$$

is the electron–nucleus attraction energy operator, and

$$\hat{V}_{ee} = \sum_{i<j}^N \frac{1}{r_{ij}} \quad (1.1.7)$$

is the electron–electron repulsion energy operator. The total energy W is the electronic energy E plus the nucleus–nucleus repulsion energy

$$V_{nn} = \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \quad (1.1.8)$$

That is,

$$W = E + V_{nn} \quad (1.1.9)$$

It is immaterial whether one solves (1.1.1) for E and adds V_{nn} afterwards, or includes V_{nn} in the definition of \hat{H} and works with the Schrödinger equation in the form $\hat{H}\Psi = W\Psi$.

Equation (1.1.1) must be solved subject to appropriate boundary conditions. Ψ must be well-behaved everywhere, in particular decaying to zero at infinity for an atom or molecule or obeying appropriate periodic boundary conditions for a regular infinite solid. $|\Psi|^2$ is a probability distribution function in the sense that

$$|\Psi(\mathbf{r}^N, s^N)|^2 d\mathbf{r}^N = \text{probability of finding the system} \\ \text{with position coordinates between } \mathbf{r}^N \text{ and} \\ \mathbf{r}^N + d\mathbf{r}^N \text{ and spin coordinates} \\ \text{equal to } s^N \quad (1.1.10)$$

Here $d\mathbf{r}^N = d\mathbf{r}_1, d\mathbf{r}_2, \dots, d\mathbf{r}_N$; \mathbf{r}^N stands for the set $\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$, and s^N stands for the set s_1, s_2, \dots, s_N . The spatial coordinates are continuous, while the spin coordinates are discrete. Because electrons are fermions, Ψ also must be antisymmetric with respect to interchange of the coordinates (both space and spin) of any two electrons.

There are many acceptable independent solutions of (1.1.1) for a given system: the eigenfunctions Ψ_k , with corresponding energy eigenvalues E_k . The set Ψ_k is complete, and the Ψ_k may always be taken to be orthogonal and normalized [in accordance with (1.1.10)],

$$\int \Psi_k^* \Psi_l d\mathbf{x}^N = \langle \Psi_k | \Psi_l \rangle = \delta_{kl} \quad (1.1.11)$$

We denote the ground-state wave function and energy by Ψ_0 and E_0 . Here $\int d\mathbf{x}^N$ means integration over $3N$ spatial coordinates and summation over N spin coordinates.

Expectation values of observables are given by formulas of the type

$$\langle \hat{A} \rangle = \frac{\int \Psi^* \hat{A} \Psi d\mathbf{x}}{\int \Psi^* \Psi d\mathbf{x}} = \frac{\langle \Psi | \hat{A} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (1.1.12)$$

where \hat{A} is the Hermitian linear operator for the observable A . Many measurements average to $\langle \hat{A} \rangle$; particular measurements give particular eigenvalues of \hat{A} . For example, if Ψ is normalized, expectation values of kinetic and potential energies are given by the formulas

$$T[\Psi] = \langle \hat{T} \rangle = \int \Psi^* \hat{T} \Psi d\mathbf{x} \quad (1.1.13)$$

and

$$V[\Psi] = \langle \hat{V} \rangle = \int \Psi^* \hat{V} \Psi d\mathbf{x} \quad (1.1.14)$$

The square brackets here denote that Ψ determines T and V ; we say that T and V are *functionals* of Ψ (see Appendix A).

1.2 Variational principle for the ground state

When a system is in the state Ψ , which may or may not satisfy (1.1.1), the average of many measurements of the energy is given by the formula

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (1.2.1)$$

where

$$\langle \Psi | \hat{H} | \Psi \rangle = \int \Psi^* \hat{H} \Psi d\mathbf{x} \quad (1.2.2)$$

Since, furthermore, each particular measurement of the energy gives one of the eigenvalues of \hat{H} , we immediately have

$$E[\Psi] \geq E_0 \quad (1.2.3)$$

The energy computed from a guessed Ψ is an upper bound to the true ground-state energy E_0 . Full minimization of the functional $E[\Psi]$ with respect to all allowed N -electron wave functions will give the true ground state Ψ_0 and energy $E[\Psi_0] = E_0$; that is,

$$E_0 = \min_{\Psi} E[\Psi] \quad (1.2.4)$$

Formal proof of the minimum-energy principle of (1.2.3) goes as follows. Expand Ψ in terms of the normalized eigenstates of \hat{H} , Ψ_k :

$$\Psi = \sum_k C_k \Psi_k \quad (1.2.5)$$

Then the energy becomes

$$E[\Psi] = \frac{\sum_k |C_k|^2 E_k}{\sum_k |C_k|^2} \quad (1.2.6)$$

where E_k is the energy for the k th eigenstate of \hat{H} . Note that the orthogonality of the Ψ_k has been used. Because $E_0 \leq E_1 \leq E_2 \leq \dots$, $E[\Psi]$ is always greater than or equal to E_0 , and it reaches its minimum E_0 if and only if $\Psi = C_0 \Psi_0$.

Every eigenstate Ψ is an extremum of the functional $E[\Psi]$. In other words, one may replace the Schrödinger equation (1.1.1) with the variational principle

$$\delta E[\Psi] = 0 \quad (1.2.7)$$

When (1.2.7) is satisfied, so is (1.1.1), and vice versa.

It is convenient to restate (1.2.7) in a way that will guarantee that the final Ψ will automatically be normalized. This can be done by the method of Lagrange undetermined multipliers (§17.6 of Arfken 1980, or Appendix A). Extremization of $\langle \Psi | \hat{H} | \Psi \rangle$ subject to the constraint $\langle \Psi | \Psi \rangle = 1$ is equivalent to making stationary the quantity $[\langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle]$ without constraint, with E the Lagrange multiplier. This gives

$$\delta[\langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle] = 0. \quad (1.2.8)$$

One must solve this equation for Ψ as a function of E , then adjust E until normalization is achieved. It is elementary to show the essential equivalence of (1.2.8) and (1.1.1). Solutions of (1.2.8) with forms of Ψ restricted to approximate forms $\tilde{\Psi}$ of a given type (that is, a subset of all allowable Ψ) will give well-defined best approximations $\tilde{\Psi}_0$ and \tilde{E}_0 to the correct Ψ_0 and E_0 . By (1.2.3), $\tilde{E}_0 \geq E_0$, and so convergence of the energy, from above, is assured as one uses more and more flexible $\tilde{\Psi}$. Most contemporary calculations on electronic structure are done with this variational procedure, in some linear algebraic implementation.

Excited-state eigenfunctions and eigenvalues also satisfy (1.2.8), but the corresponding methods for determining approximate $\tilde{\Psi}_k$ and \tilde{E}_k encounter orthogonality difficulties. For example, given $\tilde{\Psi}_1$, \tilde{E}_1 is not necessarily above E_1 , unless $\tilde{\Psi}_1$ is orthogonal to the exact Ψ_0 .

To summarize: For a system of N electrons and given nuclear potential $v(\mathbf{r})$, (1.2.8) defines a procedure for going from N and $v(\mathbf{r})$ to the

ground-state wave function Ψ and hence through (1.1.12) to the ground-state energy $E[N, v]$ and other properties of interest. Note that in this statement there is no mention of the kinetic-energy or electron-repulsion parts of \hat{H} , because these are universal in that they are determined by N . We say that E is a *functional* of N and $v(\mathbf{r})$.

1.3 The Hartree-Fock approximation

Suppose now that Ψ is approximated as an antisymmetrized product of N orthonormal spin orbitals $\psi_i(\mathbf{x})$, each a product of a spatial orbital $\phi_k(\mathbf{r})$ and a spin function $\sigma(s) = \alpha(s)$ or $\beta(s)$, the *Slater determinant*

$$\begin{aligned} \Psi_{\text{HF}} &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \\ &= \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 \cdots \psi_N] \end{aligned} \quad (1.3.1)$$

The *Hartree-Fock approximation* (Roothaan 1951) is the method whereby the orthonormal orbitals ψ_i are found that minimize (1.2.1) for this determinantal form of Ψ .

The normalization integral $\langle \Psi_{\text{HF}} | \Psi_{\text{HF}} \rangle$ is equal to 1, and the energy expectation value is found to be given by the formula (for example, see Parr 1963)

$$E_{\text{HF}} = \langle \Psi_{\text{HF}} | \hat{H} | \Psi_{\text{HF}} \rangle = \sum_{i=1}^N H_i + \frac{1}{2} \sum_{i,j=1}^N (J_{ij} - K_{ij}) \quad (1.3.2)$$

where

$$H_i = \int \psi_i^*(\mathbf{x}) \left[-\frac{1}{2} \nabla^2 + v(\mathbf{x}) \right] \psi_i(\mathbf{x}) d\mathbf{x} \quad (1.3.3)$$

$$J_{ij} = \iint \psi_i(\mathbf{x}_1) \psi_i^*(\mathbf{x}_1) \frac{1}{r_{12}} \psi_j^*(\mathbf{x}_2) \psi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \quad (1.3.4)$$

$$K_{ij} = \iint \psi_i^*(\mathbf{x}_1) \psi_j(\mathbf{x}_1) \frac{1}{r_{12}} \psi_i(\mathbf{x}_2) \psi_j^*(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \quad (1.3.5)$$

These integrals are all real, and $J_{ij} \geq K_{ij} \geq 0$. The J_{ij} are called *Coulomb integrals*, the K_{ij} are called *exchange integrals*. We have the important equality

$$J_{ii} = K_{ii} \quad (1.3.6)$$

This is the reason the double summation in (1.3.2) can include the $i = j$ terms.

Minimization of (1.3.2) subject to the orthonormalization conditions

$$\int \psi_i^*(\mathbf{x}) \psi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij} \quad (1.3.7)$$

now gives the Hartree–Fock differential equations

$$\hat{F} \psi_i(\mathbf{x}) = \sum_{j=1}^N \varepsilon_{ij} \psi_j(\mathbf{x}) \quad (1.3.8)$$

where

$$\hat{F} = -\frac{1}{2} \nabla^2 + v + \hat{g} \quad (1.3.9)$$

in which the Coulomb-exchange operator $\hat{g}(x_1)$ is given by

$$\hat{g} = \hat{j} - \hat{k} \quad (1.3.10)$$

Here

$$\hat{j}(\mathbf{x}_1)f(\mathbf{x}_1) = \sum_{k=1}^N \int \psi_k^*(\mathbf{x}_2) \psi_k(\mathbf{x}_2) \frac{1}{r_{12}} f(\mathbf{x}_1) d\mathbf{x}_2 \quad (1.3.11)$$

and

$$\hat{k}(\mathbf{x}_1)f(\mathbf{x}_1) = \sum_{k=1}^N \int \psi_k^*(\mathbf{x}_2) f(\mathbf{x}_2) \frac{1}{r_{12}} \psi_k(\mathbf{x}_1) d\mathbf{x}_2 \quad (1.3.12)$$

with $f(\mathbf{x}_1)$ an arbitrary function. The matrix ε consists of Lagrange multipliers (in general complex) associated with the constraints of (1.3.7). Also,

$$\varepsilon_{ji}^* = \varepsilon_{ij} \quad (1.3.13)$$

so that ε is Hermitian (Roothaan 1951).

Multiplying (1.3.8) by ψ_i^* and integrating, one obtains the formula for “orbital energies,”

$$\varepsilon_i = \varepsilon_{ii} = \langle \psi_i | \hat{F} | \psi_i \rangle = H_i + \sum_{j=1}^N (J_{ij} - K_{ij}) \quad (1.3.14)$$

Summing over i and comparing with (1.3.2), we find

$$E_{\text{HF}} = \sum_{i=1}^N \varepsilon_i - V_{ee} \quad (1.3.15)$$

where the symbol V_{ee} stands for the total electron–electron repulsion energy

$$\begin{aligned} V_{ee} &= \int \Psi_{\text{HF}}^*(\mathbf{x}^N) \left(\sum_{i < j} \frac{1}{r_{ij}} \right) \Psi_{\text{HF}}(\mathbf{x}^N) d\mathbf{x}^N \\ &= \frac{1}{2} \sum_{i,j=1}^N (J_{ij} - K_{ij}) \end{aligned} \quad (1.3.16)$$

For the total molecular energy including nuclear–nuclear repulsion, one has from (1.1.9),

$$W_{\text{HF}} = \sum_{i=1}^N \varepsilon_i - V_{ee} + V_{nn} \quad (1.3.17)$$

$$= \sum_{i=1}^N H_i + V_{ee} + V_{nn} \quad (1.3.18)$$

Note that neither E_{HF} nor W_{HF} is equal to the sum of orbital energies.

Solution of (1.3.8) must proceed iteratively, since the orbitals ψ_i that solve the problem appear in the operator \hat{F} . Consequently, the Hartree–Fock method is a nonlinear “self-consistent-field” method.

For a system having an even number of electrons, in what is called the *restricted Hartree–Fock method* (RHF), the N orbitals ψ_i are taken to comprise $N/2$ orbitals of form $\phi_k(\mathbf{r})\alpha(s)$ and $N/2$ orbitals of form $\phi_k(\mathbf{r})\beta(s)$. The energy formula (1.3.2) becomes

$$E_{\text{HF}} = 2 \sum_{k=1}^{N/2} H_k + \sum_{k,l=1}^{N/2} (2J_{kl} - K_{kl}) \quad (1.3.19)$$

where

$$H_k = \int \phi_k^*(\mathbf{r}) \left[-\frac{1}{2} \nabla^2 + v(\mathbf{r}) \right] \phi_k(\mathbf{r}) d\mathbf{r} \quad (1.3.20)$$

$$J_{kl} = \iint |\phi_k(\mathbf{r}_1)|^2 \frac{1}{r_{12}} |\phi_l(\mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \quad (1.3.21)$$

$$K_{kl} = \iint \phi_k^*(\mathbf{r}_1) \phi_l(\mathbf{r}_1) \frac{1}{r_{12}} \phi_k(\mathbf{r}_2) \phi_l^*(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (1.3.22)$$

while the Hartree–Fock equations (1.3.8) now read

$$\hat{F} \phi_k(\mathbf{r}) = \sum_{l=1}^{N/2} \varepsilon_{kl} \phi_l(\mathbf{r}) \quad (1.3.23)$$

with the operator \hat{F} given by (1.3.9) and (1.3.10), with (1.3.11) and (1.3.12) replaced by

$$\hat{j}(\mathbf{r}_1)f(\mathbf{r}_1) = 2 \sum_{m=1}^{N/2} \int |\phi_m(\mathbf{r}_2)|^2 \frac{1}{r_{12}} d\mathbf{r}_2 f(\mathbf{r}_1) \quad (1.3.24)$$

$$\hat{k}(\mathbf{r}_1)f(\mathbf{r}_1) = \sum_{m=1}^{N/2} \int \phi_m^*(\mathbf{r}_2) f(\mathbf{r}_2) \frac{1}{r_{12}} d\mathbf{r}_2 \phi_m(\mathbf{r}_1) \quad (1.3.25)$$

The determinantal wave function (1.3.1) for this “closed-shell” case is

explicitly

$$\Psi_{\text{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{r}_1)\alpha(s_1) & \phi_1(\mathbf{r}_1)\beta(s_1) & \cdots & \phi_{N/2}(\mathbf{r}_1)\beta(s_1) \\ \phi_1(\mathbf{r}_2)\alpha(s_2) & \phi_1(\mathbf{r}_2)\beta(s_2) & \cdots & \phi_{N/2}(\mathbf{r}_2)\beta(s_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{r}_N)\alpha(s_N) & \phi_1(\mathbf{r}_N)\beta(s_N) & \cdots & \phi_{N/2}(\mathbf{r}_N)\beta(s_N) \end{vmatrix} \quad (1.3.26)$$

An important property of this wave function [and also of the more general (1.3.1)] is that a unitary transformation of the occupied orbitals ϕ_k (or ψ_i) to another set of orbitals η_m leaves the wave function unchanged except possibly by an inconsequential phase factor. The operators \hat{j} , \hat{k} , and \hat{F} of (1.3.23) through (1.3.25) [or of (1.3.9) through (1.3.12)] are also invariant to such a transformation (Roothaan 1951, Szabo and Ostlund 1982, page 120). That is to say, if we let

$$\eta_m = \sum_k U_{mk} \psi_k \quad (1.3.27)$$

where \mathbf{U} is a unitary matrix,

$$\mathbf{U}^+ \mathbf{U} = \mathbf{1} \quad (1.3.28)$$

then (1.3.23) becomes

$$\hat{F}\eta_m = \sum_{n=1}^{N/2} \epsilon_{mn}^{\eta} \eta_n \quad (1.3.29)$$

where

$$\epsilon^{\eta} = \mathbf{U} \epsilon \mathbf{U}^+ \quad (1.3.30)$$

This exhibits the considerable freedom that exists in the choice of the matrix ϵ .

Since the matrix ϵ is Hermitian, one may choose the matrix \mathbf{U} to diagonalize it. The corresponding orbitals λ_m , called the *canonical Hartree-Fock orbitals*, satisfy the *canonical Hartree-Fock equations*,

$$\hat{F}\lambda_m(\mathbf{r}) = \epsilon_m^{\lambda} \lambda_m(\mathbf{r}) \quad (1.3.31)$$

Equation (1.3.31) is considerably more convenient for calculation than (1.3.23). Furthermore, the orbitals that are solutions of (1.3.31) are uniquely appropriate for describing removal of electrons from the system in question. There is a theorem due to Koopmans (1934) that if one assumes no reorganization (change of orbitals) on ionization, the best (lowest-energy) single-determinantal description for the ion is the determinant built from the canonical Hartree-Fock orbitals of (1.3.31). One then finds, approximately,

$$\epsilon_m^{\lambda} = -I_m \quad (1.3.32)$$

where I_m is the ionization energy associated with removal of an electron from the orbital λ_m . This equation is in error, because it ignores both

reorganization and errors in the Hartree-Fock description (called correlation energy: see the next section); fortunately these tend to cancel.

The orbital energies for the canonical Hartree-Fock orbitals also control the long-range behavior of the orbitals. Naively, one would expect, from the one-electron nature of (1.3.31), $\lambda_m \sim \exp[-(-2\epsilon_m^{\lambda})^{1/2}r]$ for large r . This is correct for atoms with s electrons only, but not in general. Instead, in general the maximum (least-negative) of all of the occupied ϵ_m^{λ} determines the long-range behavior of *all* of the orbitals:

$$\lambda_m \sim \exp[-(-2\epsilon_{\text{max}})^{1/2}r] \quad \text{for large } r \quad (1.3.33)$$

The long-range properties of the exchange part of \hat{F} are responsible for this remarkable behavior (Handy, Marron, and Silverstone 1969). The operator \hat{F} is *not* a Sturm-Liouville operator.

For the closed-shell case, entirely equivalent to the canonical Hartree-Fock description are the circulant Hartree-Fock description and the localized Hartree-Fock description. *Circulant Hartree-Fock orbitals* (Parr and Chen 1981, Nyden and Parr 1983) are orbitals the absolute squares of which are as close to each other as possible in a certain sense; for them, the matrix ϵ of (1.3.29) is a circulant matrix (diagonal elements all equal, every row a cyclic permutation of every other). *Localized Hartree-Fock orbitals* (Edmiston and Ruedenberg 1963) are orbitals with maximum self-repulsion or minimum interorbital exchange interaction. The electron repulsion part of (1.3.19) is, from (1.3.6),

$$V_{ee} = J - K \quad (1.3.34)$$

where

$$J = \sum_{k,l=1}^{N/2} 2J_{kl} = \sum_k J_{kk} + \left[\sum_k J_{kk} + \sum_{k \neq l} J_{kl} \right] \quad (1.3.35)$$

and

$$K = \sum_{k,l=1}^{N/2} K_{kl} = \sum_k J_{kk} + \left[\sum_{k \neq l} K_{kl} \right] \quad (1.3.36)$$

J and K are each invariant to unitary transformation, but the terms in square brackets in these equations are not; the unitary transformation to localized orbitals can therefore be effected by maximizing

$$J(\text{self}) = \sum_k J_{kk} = K(\text{self}) \quad (1.3.37)$$

or, equivalently, by minimizing

$$\sum_{k \neq l} K_{kl}$$

Circulant orbitals are important because they are orbitals that have equivalent mathematical and physical footing, each very close to the

square root of the electron density per particle. Localized orbitals are important because their existence reconciles molecular-orbital theory with the more traditional descriptions of molecules as held together by localized chemical bonds.

If from the beginning one neglects all interorbital exchange terms in the Hartree–Fock method, which corresponds to using a product of orbitals as the wave function in place of the antisymmetrized product of (1.3.1) or (1.3.26), one gets the *orthogonalized Hartree method*. The closed-shell equation (1.3.23) is replaced by

$$\hat{F}_k^H \phi_k^H = \sum_l \varepsilon_{kl}^H \phi_l^H \quad (1.3.38)$$

where

$$\hat{F}_k^H = -\frac{1}{2}\nabla^2 + v + \hat{j}_k \quad (1.3.39)$$

in which

$$\hat{j}_k(\mathbf{x}_1) = \int \left[|\psi_k(\mathbf{x}_2)|^2 + 2 \sum_{m \neq k} |\psi_m(\mathbf{x}_2)|^2 \right] \frac{1}{r_{12}} d\mathbf{x}_2 \quad (1.3.40)$$

This method gives orbitals even more localized than the localized Hartree–Fock orbitals; these are useful for some purposes (Levy, Nee, and Parr 1975).

When the number of electrons is not even, the standard Hartree–Fock scheme is what is called the *unrestricted open-shell Hartree–Fock method* (UHF) (Szabo and Ostlund 1982, pages 205–229). Spatial parts of spin orbitals with α spin are allowed to be different from spatial parts of spin orbitals with β spin, even within a single “pair” of electrons. Noting that orthogonality between all α -spin spin orbitals and all β -spin spin orbitals is still preserved, we see that the only problem in implementation is the complication associated with handling all N orbitals in the Hartree–Fock equations. The mathematical apparatus is (1.3.8) to (1.3.12). The UHF method can also be used for an even number of electrons. Often, indeed usually, the UHF method then gives no energy lowering over the restricted HF method. But there are important cases in which energy lowering is found. For example, the UHF description of bond breaking in H_2 gives the proper dissociation products, while the RHF description of H_2 gives unrealistic ones.

Many physical properties of most molecules in their ground states are well accounted for by use of Hartree–Fock wave functions (Schaefer 1972).

In actual implementation of Hartree–Fock theory (and also in calculations of wave functions to an accuracy higher than those of Hartree–Fock), one usually (though not always) employs some set of fixed, one-electron basis functions, in terms of which orbitals are expanded and many-electron wave functions are expressed. This transforms the mathe-

matical problem into one (or more) matrix eigenvalue problems of high dimension, in which the matrix elements are calculated from arrays of integrals evaluated for the basis functions. If we call the basis functions $\chi_p(\mathbf{r})$, one can see from (1.1.2) what the necessary integrals will be: *overlap integrals*,

$$S_{pq} = \int \chi_p^*(\mathbf{r}) \chi_q(\mathbf{r}) d\mathbf{r} \quad (1.3.41)$$

kinetic energy integrals,

$$T_{pq} = \int \chi_p^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \chi_q(\mathbf{r}) d\mathbf{r} \quad (1.3.42)$$

electron–nucleus attraction integrals,

$$(A | pq) = \int \chi_p^*(\mathbf{r}_1) \frac{1}{r_{1A}} \chi_q(\mathbf{r}_1) d\mathbf{r}_1 \quad (1.3.43)$$

and *electron–electron repulsion integrals*,

$$(pq | rs) = \iint \chi_p^*(\mathbf{r}_1) \chi_q(\mathbf{r}_1) \frac{1}{r_{12}} \chi_r(\mathbf{r}_2) \chi_s^*(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \quad (1.3.44)$$

Sometimes these are all computed exactly, in which case one says that one has an *ab initio method*. (For reviews, see Schaefer 1977 and Lawley 1987. The term “*ab initio*” was used first, though intended to have a different meaning, in Parr, Craig, and Ross 1950.) Sometimes these are determined by some recourse to experimental data, in which case one has a *semiempirical method* (Parr 1963, Segal 1977). Such details are of course vital, but here they will not be of much concern to us in the present exposition.

1.4 Correlation energy

When one is interested in higher accuracy, there are straightforward extensions of the single-determinantal description to simple “multiconfiguration” descriptions involving few determinants (for example, Section 4.5 of Szabo and Ostlund 1982).

The exact wave function for a system of many interacting electrons is never a single determinant or a simple combination of a few determinants, however. The calculation of the error in energy, called *correlation energy*, here defined to be negative,

$$E_{\text{corr}}^{\text{HF}} = E - E_{\text{HF}} \quad (1.4.1)$$

is a major problem in many-body theory on which there has been a vast amount of work and much progress has been made. The methods

employed include the linear mixing of many determinants (millions!), called configuration interaction (Chapter 4 of Szabo and Ostlund 1982), and many-body perturbation techniques (Chapter 6 of Szabo and Ostlund 1982). For comprehensive reviews, see Sinanoğlu and Brueckner (1970), Hurley (1976), and Wilson (1984).

Correlation energy tends to remain constant for atomic and molecular changes that conserve the numbers and types of chemical bonds, but it can change drastically and become determinative when bonds change. Its magnitude can vary from 20 or 30 to thousands of kilocalories per mole, from a few hundredths of an atomic unit on up. Exchange energies are an order of magnitude or more bigger, even if the self-exchange term is omitted.

1.5 Electron density

In an electronic system, the number of electrons per unit volume in a given state is the *electron density* for that state. This quantity will be of great importance in this book; we designate it by $\rho(\mathbf{r})$. Its formula in terms of Ψ is

$$\rho(\mathbf{r}_1) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\mathbf{s}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N \quad (1.5.1)$$

This is a nonnegative simple function of three variables, x , y , and z , integrating to the total number of electrons,

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \quad (1.5.2)$$

There has been much attention paid to the electron density over the years (Smith and Absar 1977). Maps of electron densities are available in many places (for example, Bader 1970). For an atom in its ground state, the density decreases monotonically away from the nucleus (Weinstein, Politzer, and Srebrenik 1975), in approximately piecewise exponential fashion (Wang and Parr 1977). For molecules, at first sight, densities look like superposed atomic densities; on closer inspection (experimental or theoretical), modest (but still quite small in absolute terms) buildups of density are seen in bonding regions.

At any atomic nucleus in an atom, molecule, or solid, the electron density has a finite value; for an atom we designate this $\rho(0)$. In the neighborhood of a nucleus there always is a cusp in the density owing to the necessity for Hamiltonian terms $-\frac{1}{2}\nabla^2 - (Z_\alpha/r_\alpha)$ not to cause blowups in $\hat{H}\Psi$ there. The specific cusp condition is (for example, see Davidson

1976, page 44)

$$\frac{\partial}{\partial r_\alpha} \bar{\rho}(r_\alpha)|_{r_\alpha=0} = -2Z_\alpha \bar{\rho}(0) \quad (1.5.3)$$

where $\bar{\rho}(r_\alpha)$ is the spherical average of $\rho(r_\alpha)$.

Another important result is the long-range law for electron density,

$$\rho \sim \exp[-2(I_{\min})^{1/2}r] \quad (1.5.4)$$

where I_{\min} is the exact first ionization potential (Morrell, Parr, and Levy 1975; this paper also contains a generalization of Koopmans' theorem). The corresponding Hartree-Fock result will be, from (1.3.33),

$$\rho_{\text{HF}} \sim \exp[-2(-\epsilon_{\max})^{1/2}r] \quad (1.5.5)$$

where ϵ_{\max} approximates I_{\min} by (1.3.32).

Finally, we record here certain results about electron density from the standard first-order perturbation theory for a nondegenerate state. Suppose the state Ψ_k^0 is perturbed to the state $\Psi_k = \Psi_k^0 + \Psi_k^1$ by the one-electron perturbation $\Delta V = \sum_i \Delta v(\mathbf{r}_i)$. The energy change to first order is

$$E_k^{(1)} = \int \Psi_k^{0*} \Delta V \Psi_k^0 dx^N = \int \rho_k(\mathbf{r}_1) \Delta v(\mathbf{r}_1) d\mathbf{r}_1, \quad (1.5.6)$$

while the perturbed wavefunction is, to first order,

$$\Psi_k = \Psi_k^0 + \sum_{j \neq k} \frac{\langle \Psi_j^0 | \Delta V | \Psi_k^0 \rangle}{E_k^0 - E_j^0} \Psi_j \quad (1.5.7)$$

The electron-density change is then, to first order in Δv ,

$$\begin{aligned} \Delta \rho_k(\mathbf{r}_1) &= N \iint (\Psi_k^* \Psi_k - \Psi_k^{0*} \Psi_k^0) d\mathbf{s}_1 d\mathbf{x}_2 d\mathbf{x}_3 \cdots d\mathbf{x}_N \\ &= 2N \operatorname{Re} \left\{ \sum_{j \neq k} \frac{\langle \Psi_j^0 | \Delta V | \Psi_k^0 \rangle}{E_k^0 - E_j^0} \int \cdots \int \Psi_k^{0*} \Psi_j^0 d\mathbf{s}_1 d\mathbf{x}_2 d\mathbf{x}_3 \cdots d\mathbf{x}_N \right\} \\ &= \int \frac{\delta \rho_k(\mathbf{r}_1)}{\delta v(\mathbf{r}_2)} \Delta v(\mathbf{r}_2) d\mathbf{r}_2 \end{aligned} \quad (1.5.8)$$

where the "functional derivative" $\delta \rho / \delta v$ is defined by

$$\frac{\delta \rho_k(\mathbf{r}_1)}{\delta v(\mathbf{r}_2)} = \frac{\delta \rho_k(\mathbf{r}_2)}{\delta v(\mathbf{r}_1)} = 2N^2 \times \sum_{j \neq k} \frac{\left[\int \cdots \int \Psi_j^{0*} \Psi_k^0 d\mathbf{s}_2 d\mathbf{x}_1 d\mathbf{x}_3 \cdots d\mathbf{x}_N \right] \left[\int \cdots \int \Psi_k^{0*} \Psi_j^0 d\mathbf{s}_1 d\mathbf{x}_2 d\mathbf{x}_3 \cdots d\mathbf{x}_N \right]}{(E_k^0 - E_j^0)} \quad (1.5.9)$$

This quantity is called the *linear response function*. The symmetry represented in (1.5.9) is important. If a perturbation at point 1 produces a density change at point 2, then the same perturbation at point 2 will produce at point 1 precisely the same density change. Note that

$$\int \frac{\delta \rho_k(\mathbf{r}_1)}{\delta v(\mathbf{r}_2)} d\mathbf{r}_1 = 0 \quad (1.5.10)$$

All of these formulas assume that the number of electrons is fixed. For a general discussion of functional derivatives, see Appendix A.

1.6 Hellmann–Feynman theorems and virial theorem

Let λ be a parameter in the Hamiltonian and $\Psi(\lambda)$ be an eigenfunction of \hat{H} . Then

$$\frac{dE}{d\lambda} = \frac{\langle \Psi | \partial \hat{H} / \partial \lambda | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (1.6.1)$$

$$E(\lambda_2) - E(\lambda_1) = \frac{\langle \Psi_2 | \hat{H}(\lambda_2) - \hat{H}(\lambda_1) | \Psi_1 \rangle}{\langle \Psi_2 | \Psi_1 \rangle} \quad (1.6.2)$$

and

$$E(\lambda_2) - E(\lambda_1) = \int_{\lambda_1}^{\lambda_2} \frac{\langle \Psi | \partial \hat{H} / \partial \lambda | \Psi \rangle}{\langle \Psi | \Psi \rangle} d\lambda \quad (1.6.3)$$

These identities are the *differential Hellmann–Feynman theorem* (formula), the *integral Hellmann–Feynman theorem* (formula), and the *integrated Hellmann–Feynman theorem* (formula) (Epstein, Hurley, Wyatt, and Parr 1967). The derivative $\partial \hat{H} / \partial \lambda$ is written as a partial derivative to emphasize that the integral $\langle \Psi | \partial \hat{H} / \partial \lambda | \Psi \rangle$ can depend on the coordinate system chosen to describe a particular situation.

The equation (1.6.1) is a direct result of the first-order perturbation formula for energy, (1.5.6) above. Integrating (1.6.1) from λ_1 to λ_2 gives (1.6.3). Theorem (1.6.2) can be put in a general form,

$$E_A - E_B = \frac{\langle \Psi_B | \hat{H}_A - \hat{H}_B | \Psi_A \rangle}{\langle \Psi_B | \Psi_A \rangle} \quad (1.6.4)$$

where \hat{H}_A and \hat{H}_B are different Hamiltonians acting on the same N -electron wave-function space, but they need not be related to each other by a parameter λ . Since Ψ_B and Ψ_A are eigenfunctions of \hat{H}_B and \hat{H}_A , then

$$\hat{H}_A \Psi_A = E_A \Psi_A \quad \text{and} \quad \hat{H}_B \Psi_B = E_B \Psi_B \quad (1.6.5)$$

Multiply these equations respectively by Ψ_B^* and Ψ_A^* and integrate. Take

the complex conjugate of the second result. Subtraction then gives

$$(E_A - E_B) \langle \Psi_B | \Psi_A \rangle = \langle \Psi_B | \hat{H}_A - \hat{H}_B | \Psi_A \rangle$$

Provided $\langle \Psi_B | \Psi_A \rangle \neq 0$, this is equivalent to (1.6.4) [and also (1.6.1) follows as the special case when the change is small].

Use Cartesian coordinates and in (1.6.1) let λ be the coordinate X_α of the position of nucleus α . Suppose that no fields are present except those due to the nuclei; i.e., that there are no extra terms in (1.1.3). Then the only terms in \hat{H} that depend on X_α are v and V_{nn} , and (1.6.1), yields, using (1.1.9),

$$\frac{\partial W}{\partial X_\alpha} = - \sum_{\beta \neq \alpha} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}^3} (X_\alpha - X_\beta) - Z_\alpha \int \rho(\mathbf{r}_1) \frac{(x_1 - X_\alpha)}{r_{1\alpha}^3} d\mathbf{r}_1 \quad (1.6.6)$$

This is a purely classical expression. What it shows is that the force on nucleus α due to the other nuclei and the electrons, in some particular Born–Oppenheimer nuclear configuration, is just what would be computed from classical electrostatics from the locations of the other nuclei and the electronic charge density (see Deb 1981). This is the famous *electrostatic theorem* of Feynman (1939).

An application of (1.6.1) that we are interested in is the formula obtained if one replaces Z_α by λZ_α everywhere it appears in \hat{H} and then computes $W(1) - W(0)$ for a ground state. Note that the ground state of N electrons in the absence of any nuclei has zero energy: $W(0) = 0$. Hence, for a ground state (Wilson 1962, Politzer and Parr 1974)

$$W = \sum_{\alpha < \beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} - \sum_{\alpha} Z_\alpha \int_0^1 d\lambda \int \frac{\rho(\mathbf{r}_1, \lambda)}{r_{1\alpha}} d\mathbf{r}_1 \quad (1.6.7)$$

Here $\rho(\mathbf{r}, \lambda)$ is the density associated with the eigenfunction $\Psi(\mathbf{x}, \lambda)$ for the N -electron problem with scaled nuclear charges.

Note that the Hellmann–Feynman theorems (1.6.1) through (1.6.3) hold for any eigenstate, while (1.6.7) is only true for a ground state. Equation (1.6.7), as well as the electrostatic theorem (1.6.6), can be thought of as foreshadowing what we will be demonstrating at length in this book: For a ground state, the electron density suffices for the determination of *all* the properties. Another essential point is that these various theorems may or may not hold for approximate eigenfunctions. For example, (1.6.1) holds for exact Hartree–Fock wave functions (Stanton 1962).

Another important theorem is the *virial theorem*, which relates the kinetic energy and potential energy components of the energy in certain circumstances. This theorem results from homogeneity properties that

may or may not be present for a particular problem in the kinetic and potential energy components of \hat{H} . The kinetic energy component,

$$\hat{T} = \sum_i \left(-\frac{1}{2}\nabla_i^2\right) \quad (1.6.8)$$

is homogeneous of degree -2 in particle coordinates. The total potential energy component,

$$\hat{V} = -\sum_{i,\alpha} \frac{Z_\alpha}{r_{i\alpha}} + \sum_{i<j} \frac{1}{r_{ij}} + \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta}{R_{\alpha\beta}} \quad (1.6.9)$$

is homogeneous of degree -1 in all particle coordinates. Assuming no additional forces are acting, we then find for any eigenstate of an atom,

$$E = -\langle T \rangle = \frac{1}{2}\langle V \rangle \quad (1.6.10)$$

and for any eigenstate of a molecule or solid with a particular sufficient set of internuclear distances $R_j = |\mathbf{R}_{\alpha\beta}|$,

$$\langle T \rangle = -W - \sum_j R_j \left(\frac{\partial W}{\partial R_j} \right)_{R_k} \quad (1.6.11)$$

and

$$\langle V \rangle = 2W + \sum_j R_j \left(\frac{\partial W}{\partial R_j} \right)_{R_k} \quad (1.6.12)$$

Proofs are elementary (Löwdin 1959). Given a normalized eigenstate Ψ , it makes stationary the $E[\Psi]$ of (1.2.1). Take a normalized scaled version of this Ψ ,

$$\Psi_\zeta \equiv \zeta^{3N/2} \Psi(\zeta \mathbf{r}_1, \zeta \mathbf{r}_2, \dots, \zeta R_1, \zeta R_2, \dots) \quad (1.6.13)$$

and calculate $E[\Psi_\zeta]$. This is stationary for $\zeta = 1$, which gives (1.6.10) through (1.6.12). The scaling properties of the individual components of $E[\Psi_\zeta]$ are important. Using (1.1.13) and (1.1.14), these are found to be

$$T[\Psi_\zeta] = \zeta^2 T[\Psi, \zeta R] \quad (1.6.14)$$

and

$$V[\Psi_\zeta] = \zeta V[\Psi, \zeta R] \quad (1.6.15)$$

respectively. The dependences on ζR are parametric. For a comprehensive review of the virial theorem, see Marc and McMillan (1985).

Note that in (1.6.13) both electronic and nuclear coordinates are scaled. Another type of scaling, of electronic coordinates only, is important for the purposes of this book. Let

$$\Psi_\lambda \equiv \lambda^{3N/2} \Psi(\lambda \mathbf{r}_1, \lambda \mathbf{r}_2, \dots; R_1, R_2, \dots) \quad (1.6.16)$$

Then we find

$$T[\Psi_\lambda] = \lambda^2 T[\Psi_1] \quad (1.6.17)$$

and

$$V_{ee}[\Psi_\lambda] = \lambda V_{ee}[\Psi_1] \quad (1.6.18)$$

though this time V_{ne} for molecules does not scale simply. The scaling of (1.6.16) produces a simple dilation of the electronic cloud without changing its normalization; the scaling of (1.6.13) changes nuclear positions as well. For atoms, the two scalings are equivalent.