Atomic And Molecular Electronic Structure

Horacio Grinberg

Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, and Consejo Nacional de Investigaciones Científicas y Técnicas

Agosto 2004

I. Schrödinger Equation and Operators Algebra

II. The Variational Method

III The Born Oppenheimer Approximation

IV Electron Spin and the Antisymmetry Principle

V The Independent Particle Model

I. SCHRÖDINGER EQUATION AND OPERATOR ALGEBRA

A. The Schrödinger Equation

For an N-particle system the many-fermion time-independent Schrödinger equation, may be written in the general form

$$\mathsf{H}\,\mathbf{\Psi}(\mathbf{x}_1, \cdots \mathbf{x}_N) \;=\; E\,\mathbf{\Psi}(\mathbf{x}_1, \cdots, \mathbf{x}_N) \tag{1.1}$$

The *wavefunction* $\mathbf{\Psi}(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ is a function of the spatial and spin (or other) coordinates of the N particles, with the combined coordinates of Particle i being represented by the single symbol $\mathbf{x}_i$. The wavefunction is required to be continuous and single-valued everywhere in the N-particle configuration space. The present discussion will be restricted to systems in which the particles are identical *fermions*, which by definition have the property that the many-particle wavefunction is antisymmetric under interchange of any two $\mathbf{x}_i$. The wavefunction must also be differentiable everywhere except where the potential becomes singular. In these places, if $\mathbf{\Psi}$ is nonzero it will exhibit discontinuities in its first spatial derivatives (cusps) to compensate the singularities then occurring in the potential energy. The wavefunction must also be normalizable, i.e., satisfying

$$\int \mathbf{\Psi}^*(\mathbf{x}_1, \cdots, \mathbf{x}_N)\; \mathbf{\Psi}(\mathbf{x}_1, \cdots, \mathbf{x}_N)\, d\mathbf{x}_1 \cdots d\mathbf{x}_N \;<\; \infty \tag{1.2}$$

This integration, and those in subsequent equations, are over the entire range of both spatial and other coordinates of all particles; for example, if there is a spin coordinate, then its integration will in practice involve a summation over the discrete set of spin states. If normalized to unity, $|\mathbf{\Psi}|^2$ may be directly identified with probability, but other normalizations are found useful for some discussions.

A solution to the Schrödinger equation consists of a $\mathbf{\Psi}$ satisfying the previously identified conditions, together with the corresponding value of E, which is the energy of the system. The solutions will normally include a discrete set of negative E values, with one (or a few) $\mathbf{\Psi}$ for a given E. When there is more than one $\mathbf{\Psi}$ for a particular E, the solutions are termed

degenerate; a unique Ψ is referred to as *nondegenerate*. In addition to the discrete set of E, there will be solutions for a continuous range of positive E. We restrict attention here to the discrete solutions and mostly to that of the lowest E, which is that of the ground state. It will also be assumed that this state is nondegenerate.

We will deal with a particular class of fermions, namely electrons. We therefore specialize from now on to a system containing, in addition to N electrons, S fixed atomic nuclei. In atomic units, the Hamiltonian operator H then assumes the form (ignoring magnetic and relativistic effects, and nuclear kinetic energy)

$$\mathsf{H} = \sum_{\mu=1}^{N} \left(-\frac{1}{2}\nabla_\mu^2 - \sum_{k=1}^{S} \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}_\mu|} \right) + \sum_{1\le\mu<\nu\le N} \frac{1}{|\mathbf{r}_\mu - \mathbf{r}_\nu|} + \sum_{1\le K<l\le S} \frac{Z_k Z_l}{|\mathbf{R}_k - \mathbf{R}_l|} \tag{1.3}$$

Here $\mathbf{R}_k$ is the position of Nucleus k, whose nuclear charge is Z_k, and $\mathbf{r}_\mu$ is the spatial part of $\mathbf{x}_\mu$. When we are concerned with the fact that H consists of one- and two-electron contributions but do no require their specific form, we alternatively write

$$\mathsf{H} = \mathsf{U} + \mathsf{V} \tag{1.4}$$

where

$$\mathsf{U} = \sum_{\mu} \mathsf{u}(\mathbf{x}_\mu) \tag{1.5}$$

$$\mathsf{V} = \sum_{\mu<\nu} \mathsf{v}(\mathbf{x}_\mu, \mathbf{x}_\nu) \tag{1.6}$$

with v symmetric under interchange of $\mathbf{x}_\mu$ and $\mathbf{x}_\nu$. For atomic and molecular systems it is frequently convenient to define H so that it excludes the nuclear-nuclear repulsion. Then U will describe the electronic kinetic energy plus the electron-nuclear electrostatic interaction, while V will consist of the electron-repulsion contributions.

B. Density Matrices

Physical quantities representable as averages over the particle distribution may be calculated as expectation values of corresponding operators. If Ψ is normalizable to unity, then

the expectation value $< \mathsf{B} >$ of the quantity corresponding to the operator B is given by

$$< \mathsf{B} > = \int \mathbf{\Psi}^* \, \mathsf{B} \, \mathbf{\Psi} \, d\mathbf{x}_1 \cdots d\mathbf{x}_N \tag{1.7}$$

If we take into account the fact that B will normally depend upon the particle coordinates one, or at most two at a time, then we may rewrite Eq. (1.7) in terms of reduced density matrices. The first-order density matrix of a unit-normalized N-particle wavefunction $\mathbf{\Psi}$, denoted

$\gamma(\mathbf{x}_1, \mathbf{x}'_1)$, is

$$\gamma(\mathbf{x}_1, \mathbf{x}'_1) \; = \; N \int d\mathbf{x}_2 \cdots d\mathbf{x}_N \, \mathbf{\Psi}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) \, \mathbf{\Psi}^*(\mathbf{x}'_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) \tag{1.8}$$

The second-order denstiy matrix of such an N-particle wavefunction, denoted $\Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2)$, is

$$\begin{aligned} &\Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) \\ &= \frac{N(N-1)}{2} \int d\mathbf{x}_3 \cdots d\mathbf{x}_N \, \mathbf{\Psi}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_N) \, \mathbf{\Psi}^*(\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}_3 \cdots, \mathbf{x}_N) \end{aligned} \tag{1.9}$$

Density matrices of higher order, through N, may be defined analogously. Comparison of Eqs. (1.8) and (1.9) indicates that the first-order density matrix may be obtained by integrating one set of coordinates in the second-order matrix:

$$\gamma(\mathbf{x}_1; \mathbf{x}'_1) \; = \; \frac{2}{N-1} \int d\mathbf{x}_2 \, \Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}_2) \tag{1.10}$$

If B is a one-particle operator i.e., is of the form

$$\mathsf{B} \; = \; \sum_i \mathsf{b}(\mathbf{x}_i) \tag{1.11}$$

its expectation value can be expressed in terms of the first-order matrix:

$$< \; \mathsf{B} \; > = \int d\mathbf{x}_1 \, d\mathbf{x}'_1 \, \delta(\mathbf{x}_1 - \mathbf{x}'_1) \, \mathsf{b}(\mathbf{x}_1) \, \gamma(\mathbf{x}_1, \mathbf{x}'_1) \tag{1.12}$$

Here $\delta(\mathbf{x})$ is the Dirac delta-function, with the property that for an arbitrary nonsingular function $f(\mathbf{x})$,

$$\int d\mathbf{x} \, \delta(\mathbf{x} - \mathbf{x}_0) \, f(\mathbf{x}) \; = \; f(\mathbf{x}_0) \tag{1.13}$$

if $\mathbf{x}_0$ is within the region of integration; otherwise, the integral vanishes. Equation (1.12) becomes obvious if we insert the explicit formula for γ and recognize that the delta-function and $\mathbf{x}_1'$ integration simply cause $\mathsf{b}(\mathbf{x}_1)$ to operate on $\mathbf{\Psi}$ but not on $\mathbf{\Psi}^*$. We have

$$< \mathsf{B} >= N \int d\mathbf{x}_1 \cdots d\mathbf{x}_N \, \mathbf{\Psi}^*(\mathbf{x}_1, \cdots, \mathbf{x}_N) \, \mathsf{b}(\mathbf{x}_1) \, \mathbf{\Psi}(\mathbf{x}_1, \cdots, \mathbf{x}_N) \tag{1.14}$$

Because $\mathbf{\Psi}$ is antisymmetric in the $\mathbf{x}_i$, $\mathsf{b}(\mathbf{x}_1)$ can be replaced by B/N, leading directly to Eq. (1.7).

In a similar fashion it can be seen that a two-particle operator of the form

$$\mathsf{B} = \sum_{\mu<\nu} \mathsf{b}(\mathbf{x}_\mu, \mathbf{x}_\nu) \tag{1.15}$$

has an expectation value expressible using the second-order density matrix

$$< \mathsf{B} >= \int d\mathbf{x}_1 \, d\mathbf{x}_2 \, d\mathbf{x}_1' \, d\mathbf{x}_2' \, \delta(\mathbf{x}_1 - \mathbf{x}_1') \, \delta(\mathbf{x}_2 - \mathbf{x}_2') \, \mathsf{b}(\mathbf{x}_1, \mathbf{x}_2) \, \Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}_1', \mathbf{x}_2') \tag{1.16}$$

Equation (1.16) shows that the Hamiltonian, a two-particle operator, has an expectation value that can be expressed in terms of the two-particle density matrix. Since this density matrix is a far less complicated object than the N-electron wavefunction $\mathbf{\Psi}$, it is nataural to ask whether it might be possible to reduce the full Scchrödinger equation to a two-particle to a two-particle equation that could then be solved directly for Γ. The thus-far insurmountable flaw in this proposal is that a Γ that solves the reduced Schrödinger equation cannot be guaranteed to be derivable from an N-fermion wavefunction. The search for tractable necessary and sufficient conditions that Γ correspond to a $\mathbf{\Psi}$ is known as the *N-Representability Problem.* This problem has not been satisfactorily resolved.

C. The Virial Theorem for a Many-Particle System

Consider a stationary state ψ satisfying the Schrödinger equation $\mathsf{H}\psi = E\psi$. Let F be any arbitrary Hermitian operator depending only on the coordinates and momenta of the system. It then turns out that the commutator of H and F has a zero expectation value in the state ψ, namely,

$$\frac{d < \mathsf{F} >}{dt} = \frac{i}{\hbar} < [\mathsf{H}, \mathsf{F}] >= 0 \tag{1.17}$$

In fact, expanding $< [\mathsf{H},\mathsf{F}] >$ and using $\mathsf{H}\psi = E\psi$ we get

$$< [\mathsf{H},\mathsf{F}] > = < \psi \,|\, \mathsf{H}\,\mathsf{F} - \mathsf{F}\,\mathsf{H} \,|\, \psi > = < \psi \,|\, \mathsf{H}\,\mathsf{F} \,|\, \psi > - < \psi \,|\, \mathsf{F}\,\mathsf{H} \,|\, \psi >$$

$$= < \mathsf{H}\,\psi \,|\, \mathsf{F} \,|\, \psi > - < \psi \,|\, \mathsf{F}\,\mathsf{H} \,|\, \psi > = E < \psi \,|\, \mathsf{F} \,|\, \psi > - E < \psi \,|\, \mathsf{F} \,|\, \psi > = 0$$

This relationship is known as the *hypervirial theorem*. It states that the expectation values of time-independent operators do not vary with time in stationary states.

Let us now consider the special case where the operators F and H are given by

$$\mathsf{F} = \mathsf{r}.\mathsf{p} \qquad\qquad \mathsf{H} = -\frac{\hbar^2}{2m}\nabla^2 + \mathsf{V}(r) \tag{1.18}$$

where r and p are the coordinate and momentum operators in the Schrödinger representation and may be constructed from

$$\mathsf{r} = \mathsf{i}\mathsf{x} + \mathsf{j}\mathsf{y} + \mathsf{k}\mathsf{z} \qquad\qquad \mathsf{p} = \frac{\hbar}{i}\nabla$$

The operator F then becomes

$$\mathsf{F} = \frac{\hbar}{i}\left(\mathsf{x}\frac{\partial}{\partial x} + \mathsf{y}\frac{\partial}{\partial y} + \mathsf{z}\frac{\partial}{\partial z}\right) \tag{1.19}$$

The commutator of H and F turns out to be

$$[\mathsf{H},\mathsf{F}] = \frac{\hbar}{i}\left[-\frac{\hbar^2}{m}\nabla^2 - \mathsf{r}.\nabla\mathsf{V}\right]$$

$$= \frac{\hbar}{i}\left(2\,\mathsf{T} - \mathsf{r}.\nabla\mathsf{V}\right) \tag{1.20}$$

The expectation value of this commutator then is

$$\frac{\hbar}{i}\left(2 < \mathsf{T} > - < \mathsf{r}.\nabla\mathsf{V} >\right) = 0$$

which may be written

$$2 < \mathsf{T} > = < \mathsf{r}.\nabla\mathsf{V} > \tag{1.21}$$

Equation (1.21) is an important relationship known as the quantum-mechanical *virial theorem*. An analogous relationship holds in classical mechanics and can be discussed with the

Bohr theory of the one-electron atom. If the potential energy is given by a relationship of the form

$$\mathsf{V}(r) \;=\; k\,\mathbf{r}^{n} \tag{1.22}$$

where k is a constant, then Eq. (1.21) becomes

$$2 \;<\mathsf{T}>=\; n \;<\mathsf{V}> \tag{1.23}$$

It is the form of the virial which is most useful for discussiing the electronic energies of atoms and molecules. In this case V represents electrostatic interactions among charged particles and $n = -1$. In general, we say that a potential energy defined by Eq. (1.22) is homogeneous in $\mathbf{r}$ of degree n.

For a system of N particles having position vectors $\mathbf{r}_1,\ \mathbf{r}_2, \cdots, \mathbf{r}_N$ and momenta $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N$, the virial theorem (1.21) becomes

$$2 \;<\mathsf{T}>=\; \sum_{i} <\mathbf{r}_i.\nabla_i \mathsf{V}> \tag{1.24}$$

The virial theorem is of great importance in classical mechanics. It can be used, for example, to derive the well-known virial equation of state often discussed in connection with nonideal gases. The theorem is also useful in connection with the construction of approximate solutions to Schrödinger's equation.

D. Matrix Representations

An arbitrary function φ can be expanded in terms of a complete orthonormal set of functions $\{\phi_\mu\}$; the expansion can be written

$$\varphi \;=\; \sum_{\mu} c_\mu\, \phi_\mu \tag{1.25}$$

with

$$c_\mu \;=< \phi_\mu \,|\, \varphi > \tag{1.26}$$

This relationship can be written in the Dirac bra-ket notation

$$|\varphi> = \sum_{\mu} |\phi_\mu><\phi_\mu|\,\varphi> = \left(\sum_{\mu} |\phi_\mu><\phi_\mu|\right)|\,\varphi> \tag{1.27}$$

where $|\phi>$ corresponds to ϕ, $<\phi|$ corresponds to ϕ^*, and complete brackets $<\phi|\,|\chi>$ are to be interpreted as scalar products, and are therefore usually written $<\phi|\chi>$. Equation (1.27) shows that $\sum_\mu |\phi_\mu><\phi_\mu|$ can be thought of as an operator whose $|\phi_\mu><\phi_\mu|$ term converts φ into $<\phi_\mu|\varphi>\phi_\mu$ (i.e., into $c_\mu\phi_\mu$). Since Eq. (1.27) must be valid for arbitrary φ, it is equivalent to the operator equation

$$\mathcal{F} = \sum_{\mu} |\phi_\mu><\phi_\mu| \tag{1.28}$$

Equation (1.28) is sometimes described as a *resolution of the identity* defined by the orthonormal functions ϕ_μ; its use decomposes φ into components, with c_μ the amplitude of the component of φ associated with ϕ_μ.

If we apply Eq. (1.25) to the expansion of the function $\mathsf{B}\varphi$, where B is an operator, we have

$$\mathsf{B}\,\varphi \;=\; \sum_{\nu} d_\nu\,\phi_\nu \tag{1.29}$$

with

$$d_\nu \;=<\phi_\nu\,|\,\mathsf{B}\,\varphi>=<\phi_\nu\,|\,\mathsf{B}\,|\,\varphi> \tag{1.30}$$

Alternatively, we could have first used Eq. (1.25) to expand φ, then expanding the resulting quantities $\mathsf{B}\varphi$ according to

$$\mathsf{B}\,\phi_\mu \;=\; \sum_{\nu} B_{\nu\mu}\,\phi_\nu \tag{1.31}$$

with

$$B_{\nu\mu} \;=<\phi_\nu\,|\,\mathsf{B}\,|\,\phi_\mu> \tag{1.32}$$

The result is

$$\mathsf{B}\,\varphi \;=\; \sum_{\mu} c_\mu\,\mathsf{B}\,\phi_\mu \;=\; \sum_{\mu\nu} c_\mu\,B_{\nu\mu}\,\phi_\nu \tag{1.33}$$

or, comparing with Eq. (1.29),

$$d_\nu = \sum_\mu B_{\nu\mu}\, c_\mu \tag{1.34}$$

Equations (1.29) and (1.34) correspond in Dirac notation to

$$\mathsf{B}\,\varphi = \sum_{\mu\nu} |\phi_\nu >< \phi_\nu \,|\, \mathsf{B} \,|\, \phi_\mu >< \phi_\mu \,|\, \varphi > \tag{1.35}$$

indicating that a representation of the operator B is provided by

$$\mathsf{B} = \sum_{\mu\nu} |\phi_\nu >< \phi_\nu \,|\, \mathsf{B} \,|\, \phi_\mu >< \phi_\mu \,| = \sum_{\mu\nu} |\phi_\nu > B_{\nu\mu} < \phi_\mu| \tag{1.36}$$

Equation (1.36) becomes obvious if we recognize that both summations are resolutions of the identity. We see that $B_{\nu\mu}$ can be thought of as the amplitude of a component of B associated with ϕ_ν and ϕ_μ in the manner shown. We also note that Eq. (1.34) shows that the amplitudes of $\mathsf{B}\varphi$ can be produced by a matrix multiplication of the amplitudes of B and those of φ.

The component amplitudes of functions φ and operators B will of course depend upon the set of functions that has been used for the expansions involved. However, since a different choice of expansion set only corresponds to a different component resolution of the same quantities, we would expect all operator equations and physical conclusions to be independent of the set of functions selected as an expansion basis. However, different expansion sets may differ greatly in the convenience of their use, and we therefore wish to analyze the effect of changing from one expansion set, $\{\phi_\mu\}$, to another, which we denote $\{\phi'_\mu\}$. Let U be the operator converting the successive functions ϕ_μ of the first expansion set into the respective functions ϕ'_μ of the second set. Since both sets are assumed orthonormal, U must be unitary.

The equality of ϕ'_μ and $\mathsf{U}\phi_\mu$ may be trivially expressed by the equation

$$\phi'_\mu = \mathsf{U}\,\phi_\mu \tag{1.37}$$

Regarding ϕ'_μ as expanded in terms of the original set of ϕ_μ, we have

$$\phi'_\mu = \sum_\nu U_{\nu\mu}\,\phi_\mu \tag{1.38}$$

with

$$U_{\nu\mu} = < \phi_\nu \,|\, \phi'_\mu > \tag{1.39}$$

that is,

$$U_{\nu\mu} = < \phi_\nu \,|\, \mathsf{U} \,|\, \phi_\mu > \tag{1.40}$$

The fact that U is unitary means that the inverse of Eq. (1.37)

$$\phi_\mu \,=\, \mathsf{U}^{-1}\, \phi'_\mu \tag{1.41}$$

leads to the expansion

$$\phi_\mu \,=\, \sum_\nu U^*_{\mu\nu}\, \phi'_\nu \tag{1.42}$$

Consider now the expansion of an operator B and a function φ in terms of the expansion set $\{\phi'_\mu\}$. Analogously to Eqs. (1.27) and (1.36),

$$|\varphi > = \sum_\mu |\phi'_\mu >< \phi'_\mu \,|\, \varphi > \tag{1.43}$$

$$\mathsf{B} \,=\, \sum_{\mu\nu} |\phi'_\mu >< \phi'_\nu \,|\, \mathsf{B} \,|\, \phi'_\mu >< \phi'_\mu| \tag{1.44}$$

Inserting the identity operator in the form given by Eq. (1.28) (in terms of the original expansion set $\{\phi_\mu\}$), Eqs. (1.43) and (1.44) can be written

$$\begin{aligned} |\varphi > &= \sum_\mu |\phi'_\mu >< \phi'_\mu| \left(\sum_\nu |\phi_\nu >< \phi_\nu| \right) |\varphi > \\ &= \sum_{\mu\nu} |\phi'_\mu >< \phi'_\mu \,|\, \phi_\nu >< \phi_\nu \,|\, \varphi > = \sum_{\mu\nu} |\phi'_\mu >\, U^*_{\nu\mu}\, c_\nu \end{aligned} \tag{1.45}$$

$$\begin{aligned} \mathsf{B} \,&=\, \sum_{\mu\nu} |\phi'_\nu >< \phi'_\nu| \left(\sum_\lambda \phi_\lambda >< \phi_\lambda| \right) \mathsf{B} \left(\sum_\sigma |\phi_\sigma >< \phi_\sigma| \right) |\phi'_\mu >< \phi'_\mu| \\ &= \sum_{\mu\nu\lambda\sigma} |\phi'_\nu >< \phi'_\nu| \,|\, \phi_\lambda >< \phi_\lambda \,|\, \mathsf{B} \,|\, \phi_\sigma >< \phi_\sigma \,|\, \phi'_\mu >< \phi'_\mu| \\ &= \sum_{\mu\nu\lambda\sigma} |\phi'_\nu >\, U^*_{\lambda\nu}\, B_{\lambda\sigma}\, U_{\sigma\mu}\, < \phi'_\mu| \end{aligned} \tag{1.46}$$

If we let $\mathbf{c}'$ and $\mathbf{B}'$ stand for the component amplitudes of φ and B in the $\{\phi'_\mu\}$ expansion set, so that

$$\varphi = \sum_\mu c'_\mu \, \phi'_\mu \tag{1.47}$$

$$\mathsf{B} = \sum_{\mu\nu} |\phi'_\nu > \, B'_{\nu\mu} \, < \phi'_\mu| \tag{1.48}$$

Eqs. (1.45) and (1.46) show that

$$c'_\mu = \sum_\nu U^*_{\nu\mu} \, c_\nu \tag{1.49}$$

$$B'_{\nu\mu} = \sum_{\lambda\sigma} U^*_{\lambda\nu} \, B_{\lambda\sigma} \, U_{\sigma\mu} \tag{1.50}$$

which in matrix form are

$$\mathbf{c}' = \mathbf{U}^\dagger \, \mathbf{c} \tag{1.51}$$

$$\mathbf{B}' = \mathbf{U}^\dagger \, \mathbf{B} \, \mathbf{U} \tag{1.52}$$

Eqs. (1.51) and (1.52) indicate that a unitary transformation between expansion sets is equivalent to the corresponding matrix transformation for all operators and functions involved. Since matrix equations are unaffected by unitary matrix transformations, we have now verified that operator equations will have solutions that are independent of the expansions introduced to obtain matrix equations.

Finally, consider the case that B is a Hermitian operator and that U has been chosen in such a way that $\mathsf{B}' = \mathsf{U}^\dagger \mathsf{B} \mathsf{U}$ is a diagonal matrix. This means that $B'_{\nu\mu} = < \phi'_\nu | \mathsf{B} | \phi'_\mu >$ vanishes unless $\mu = \nu$, which in turn means that the set $\{\phi'_\mu\}$ are the eigenfunctions of B. That is, an operator has a diagonal matrix if and only if its eigenfunctions are used as the expansion set.

E. Normal Operators

Since the Hermitian operators of quantum mechanics do not, in general, commute, it is evident that the product of two Hermitian operators is not necessarily Hermitian. Thus if A and B are two Hermitian operators, then

$$(\mathsf{A}\,\mathsf{B})^\dagger = \mathsf{B}\,\mathsf{A} \neq \mathsf{A}\,\mathsf{B} \tag{1.53}$$

However, if we form the *symmetrized sum* of the product by

$$\mathsf{A}\,\mathsf{B} + \mathsf{B}\,\mathsf{A} \tag{1.54}$$

then the result is easily verified to be Hermitian. This method provides a way of constructing Hermitian operators from products of Hermitian operators and thus enables one to compute expectation values of products of dynamical variables. In practice, a factor of $\frac{1}{2}$ is included in the expression (1.54).

Since the commutator of two Hermitian operators is an antihermitian operator, one may consider the commutator as an *antisymmetrized sum* of the product of two Hermitian operators.

A general property of operators is that it is always possible to resolve them into a Hermitian part and an antihermitian part. Consider now the arbitrary operator

$$\Lambda = \mathsf{A} + i\,\mathsf{B} \tag{1.55}$$

where both A and B are Hermitian operators. One can easily show that multiplication of an antihermitian operator by i (also an antihermitian operator) produces a Hermitian operator. In fact, the product of two commuting antihermitian operators is always Hermitian. One then sees that A is the Hermitian component of Λ and that $i\,\mathsf{B}$ is the antihermitian component. Taking the adjoint of Λ, adding it to Eq. (1.55), and solving for A leads to

$$\mathsf{A} = \frac{1}{2}\,(\Lambda + \Lambda^\dagger) \tag{1.56}$$

Similarly, we get

$$\mathsf{B} = \frac{1}{2\,i}\,(\Lambda - \Lambda^\dagger) \tag{1.57}$$

The operator Λ may then be written as

$$\Lambda = \frac{1}{2}(\Lambda + \Lambda^\dagger) + \frac{1}{2}(\Lambda - \Lambda^\dagger) \tag{1.58}$$

From Eq. (1.58) we see that if Λ is Hermitian, its antihermitian component is the zero operator (or null operator).

It is instructive to note that if one regards Hermitian operators as analogous to real numbers and antihermitian operators as analogous to imaginary numbers, Eq. (1.55) may be interpreted in terms of an analogy between operators space and complex-number space.

A *normal operator* is now defined as any operator which commutes with its adjoint operator; i.e., if $[\Lambda, \Lambda^\dagger] = 0$, we say Λ is a normal operator. Obviously this definition is a generalization of Hermitian operators, since every Hermitian operator is self-commuting and therefore normal. We shall now show that the converse is not true; i.e., the class of normal operators contains Hermitian operators as a subclass and includes other operators as well. The proof follows from the commutator of the operators Λ and $\Lambda^\dagger$ in terms of the Hermitian and antihermitian components as in Eq. (1.55), namely,

$$\Lambda\,\Lambda^\dagger = (\mathsf{A} + i\,\mathsf{B})(\mathsf{A} - i\,\mathsf{B}) = \mathsf{A}^2 + \mathsf{B}^2 + i\,[\mathsf{B}, \mathsf{A}]$$

$$\Lambda^\dagger\,\Lambda = (\mathsf{A} - i\,\mathsf{B})(\mathsf{A} + i\,\mathsf{B} = \mathsf{A}^2 + \mathsf{B}^2 - i\,[\mathsf{B}, \mathsf{A}]$$

Substracting the second expression from the first, we get

$$[\Lambda, \Lambda^\dagger] = 2\,i\,[\mathsf{B}, \mathsf{A}] \tag{1.59}$$

which is zero only if A and B themselves commute. Only for the special case $\mathsf{B} = 0$ is Λ also Hermitian. It is apparent from Eq. (1.59) that any pair of commuting Hermitian operators may be combined as in Eq. (1.55) to form a normal operator.

The concept of a normal operator is useful in the discussion of certain types of quantum-mechanical variables which have no classical analogs. Furthermore, certain properties of Hermitian operators are rather easily deducted by investigating the properties of the more general normal operators. Some of these properties are made evident in the following five theorems.

Theorem 1. If a normal operator Λ has an eigenfunction ψ_k with the eigenvalue λ_k, the adjoint operator $\Lambda^\dagger$ has an eigenvalue λ_k^* for the same eigenfunction ψ_k.

Let $\Lambda\psi_k = \lambda_k\psi_k$ and consider the integral

$$< (\Lambda^\dagger - \lambda_k^*)\,\psi_k \,|\, (\Lambda^\dagger - \lambda_k^*)\,\psi_k > \tag{1.60}$$

which can be rewritten as

$$< \psi_k \,|\, (\Lambda - \lambda_k)\,(\Lambda^\dagger - \lambda_k^*) \,|\, \psi_k > \tag{1.61}$$

Since $[\Lambda, \Lambda^\dagger] = 0$, the central factors may be interchanged, and one obtains

$$< \psi_k \,|\, (\Lambda^\dagger - \lambda_k^*)\,(\Lambda - \lambda_k) \,|\, \psi_k >= 0 \tag{1.62}$$

Then the integral (1.60) also vanishes, which implies

$$\Lambda^\dagger\,\psi_k \;=\; \lambda_k^*\,\psi_k \tag{1.63}$$

Theorem 2. The eigenvalues of Hermitian operators are real.

From Theorem 1 we have

$$\Lambda\,\psi_k \;=\; \lambda_k\,\psi_k \qquad\qquad \Lambda^\dagger\,\psi_k \;=\; \lambda_k^*\,\psi_k \tag{1.64}$$

Multiplying each expression on the right by ψ_k^*, integrating, and substracting the second expression from the first, we get

$$< \psi_k \,|\, \Lambda - \Lambda^\dagger \,|\, \psi_k >= \;(\lambda_k - \lambda_k^*) \;< \psi_k \,|\, \psi_k > \tag{1.65}$$

If $\Lambda = \Lambda^\dagger$, that is, Λ is Hermitian, and ψ_k is normalized to unity, we obtain

$$\lambda_k \;=\; \lambda_k^* \;=\; real \tag{1.66}$$

Theorem 3. If two Hermitian operators A and B commute, there exists a simultaneous set of eigenfunctions $\{\psi_k\}$ of A and B such that, for all k

$$\mathsf{A}\,\psi_k \;=\; a_k\,\psi_k$$

$$\mathsf{B}\,\psi_k \;=\; b_k\,\psi_k$$

In fact, since by hypothesis $[\mathsf{A}, \mathsf{B}] = 0$, the operators $\Lambda = \mathsf{A} + i\mathsf{B}$ and $\Lambda^\dagger = \mathsf{A} - i\mathsf{B}$ are normal. Thus

$$\Lambda\,\psi_k \;=\; (\mathsf{A} + i\,\mathsf{B})\,\psi_k \;=\; (a_k + i\,b_k)\,\psi_k \;=\; \lambda_k\,\psi_k$$

$$\Lambda^\dagger\,\psi_k \;=\; (\mathsf{A} - i\,\mathsf{B})\,\psi_k \;=\; (a_k - i\,b_k)\,\psi_k \;=\; \lambda_k^*\,\psi_k$$

Adding the above two equations yields

$$\mathsf{A}\,\psi_k \;=\; a_k\,\psi_k$$

Substracting the same equations gives

$$\mathsf{B}\,\psi_k \;=\; b_k\,\psi_k$$

A converse of Theorem 3 is also true; viz, if there exists a complete set of eigenfunctions common to both B and A, then $[\mathsf{A}, \mathsf{B}] = 0$. Theorem 3 should not be mistakenly interpreted as implying that since two operators commute, any eigenfunctions of one operator are automatically eigenfunctions of the other. The theorem merely states the conditions under which simultaneous eigenfunctions of two operators exist.

Theorem 4. If ψ_k and ψ_l are two different eigenfunctions of a normal operator Λ such that $\Lambda\psi_k = \lambda_k\psi_k$ and $\Lambda\psi_l = \lambda_l\psi_l$ with $\lambda_k \neq \lambda_l$, then ψ_k and ψ_l are orthogonal.
In fact

$$\lambda_k^* \;<\psi_k\,|\,\psi_l> \;=\; <\lambda_k\,\psi_k\,|\,\psi_l> \;=\; <\Lambda\,\psi_k\,|\,\psi_l> \;=\; <\psi_k\,|\,\Lambda^\dagger\,|\,\psi_l> \;=\; \lambda_l^*\;<\psi_k\,|\,\psi_l>$$

Substracting the first and last terms from each other, we get

$$(\lambda_k^* - \lambda_l^*)\;<\psi_k\,|\,\psi_l> \;=\; 0$$

By hypothesis $\lambda_k \neq \lambda_l$; hence $\lambda_k^* \neq \lambda_l^*$, and it follows that

$$<\psi_k\,|\,\psi_l> \;=\; 0$$

Theorem 5. If Λ is a normal operator which commutes with the operator A and ψ_k and ψ_l are two different eigenfunctions of Λ with different eigenvalues, then

$$<\psi_k\,|\,\mathsf{A}\,|\,\psi_l> \;=\; A_{kl} \;=\; 0$$

According to the hypothesis of the theorem we have

$$\lambda_l \, < \psi_k \,|\, \mathsf{A} \,|\, \psi_l > = < \psi_k \,|\, \mathsf{A} \,|\, \lambda_l \, \psi_l > = < \psi_k \,|\, \mathsf{A}\,\Lambda \,|\, \psi_l > = < \psi_k \,|\, \Lambda\,\mathsf{A} \,|\, \psi_l >$$

$$= < \Lambda^\dagger \, \psi_k \,|\, \mathsf{A} \,|\, \psi_l > = \, \lambda_k \, < \psi_k \,|\, \mathsf{A} \,|\, \psi_l >$$

Substracting the first and last terms from each other, we obtain

$$(\lambda_k - \lambda_l)\, A_{kl} \, = \, 0$$

Since, by hypothesis, $\lambda_k \neq .\lambda_l$, it follows that

$$A_{kl} \, = \, 0$$

F. Expectation Values

Let us now look at the expressions obtained for expectation values when an operator and wavefunction are expanded in an orthonormal set. Anticipating a discussion involving both many-particle and one-particle functions, let Ψ be an N-arbitrary wavefunction and B be an N-particle operator, and let $\{\Phi_\mu\}$ be an orthonormal set of N-electron functions. Then, writing

$$\mathsf{B} \, = \, \sum_{\mu\nu} |\Phi_\nu > \, B_{\nu\mu} \, < \Phi_\mu \,| \tag{1.67}$$

$$|\Psi > = \, \sum_\mu c_\mu \, |\Phi_\mu > \tag{1.68}$$

with $B_{\nu\mu} =< \Phi_\nu |\mathsf{B}| \Phi_\mu >$ and $c_\mu =< \Phi_\mu | \Psi >$ we find

$$< \mathsf{B} > = < \Psi \,|\, \mathsf{B} \,|\, \Psi > = \, \sum_{\nu\mu} < c_\nu \, \Phi_\nu \,|\, \mathsf{B} \,|\, c_\mu \, \Phi_\mu >$$

$$= \, \sum_{\nu\mu} c_\nu^* \, B_{\nu\mu} \, c_\mu \, = \, \mathbf{c}^\dagger \, \mathsf{B} \, \mathbf{c} \tag{1.69}$$

If we introduce the notion that B is a one- or two-particle operator, we can alternatively examine the expansion of $< \mathsf{B} >$ in terms of the reduced density matrices. Starting with a one-particle operator, for which

$$< \mathsf{B} > = \, \int d\mathbf{x}_1 \, d\mathbf{x}_1' \, \delta(\mathbf{x}_1 - \mathbf{x}_1') \, \mathsf{b}(\mathbf{x}_1) \, \gamma(\mathbf{x}_1, \mathbf{x}_1') \tag{1.70}$$

we expand b and γ. The expansion of b in an orthonormal set of one-particle functions $\{\phi_\mu\}$ is

$$\mathsf{b}(\mathbf{x}_1) \;=\; \sum_{\lambda\mu} |\phi_\lambda(\mathbf{x}_1) > \; b_{\lambda\mu} \; < \phi_\mu(\mathbf{x}_1)| \tag{1.71}$$

For $\gamma(\mathbf{x}_1, \mathbf{x}_1')$, we expand the $\mathbf{x}_1$ dependence in terms of the $\{\phi_\mu\}$ and the $\mathbf{x}_1'$ dependence in terms of the complex conjugate set $\{\phi_\mu^*\}$. This choice is natural because γ contains $\mathbf{x}_1$ as an argument of Ψ and $\mathbf{x}_1'$ is an argument of Ψ^*. The expansion is therefore of the form

$$\gamma(\mathbf{x}_1, \mathbf{x}_1') \;=\; \sum_{\mu\nu} |\phi_\mu(\mathbf{x}_1) > \; \gamma_{\mu\nu} \; < \phi_\nu(\mathbf{x}_1')| \tag{1.72}$$

with

$$\gamma_{\mu\nu} \;=\; \int d\mathbf{x}_1 \, d\mathbf{x}_1' \, \phi_\mu^*(\mathbf{x}_1) \, \gamma(\mathbf{x}_1, \mathbf{x}_1') \, \phi_\nu(\mathbf{x}_1') \tag{1.73}$$

Inserting Eqs. (1.71) and (1.72) into Eq. (1.70)

$$\begin{aligned} < \mathsf{B} > = & \int d\mathbf{x}_1 \, d\mathbf{x}_1' \, \delta(\mathbf{x}_1 - \mathbf{x}_1') \\ & \times \left[\sum_{\lambda\sigma} |\phi_\lambda(\mathbf{x}_1) > \; b_{\lambda\sigma} \; < \phi_\sigma(\mathbf{x}_1)| \right] \left[\sum_{\mu\nu} |\phi_\mu(\mathbf{x}_1) > \; \gamma_{\mu\nu} \; < \phi_\nu(\mathbf{x}_1')| \right] \\ & = \int d\mathbf{x}_1 \, d\mathbf{x}_1' \, \delta(\mathbf{x}_1, \mathbf{x}_1') \sum_{\lambda\mu\nu} \phi_\lambda(\mathbf{x}_1) \, b_{\lambda\mu} \gamma_{\mu\nu} \, \phi_\nu^*(\mathbf{x}_1') \\ & = \sum_{\mu\nu} b_{\nu\mu} \, \gamma_{\mu\nu} \end{aligned} \tag{1.74}$$

The last two lines in Eq. (1.74) are reached by invoking the orthonormality of the ϕ_μ. We conclude that the $\gamma_{\mu\nu}$ form a matrix representation of $\gamma(\mathbf{x}_1, \mathbf{x}_1')$ that can be combined according to the usual rules with that of b. To make the situation more explicit, we write

$$< \mathsf{B} > = \sum_\nu (\mathsf{b}\,\gamma)_{\nu\nu} \;=\; Tr(\mathsf{b}\,\gamma) \tag{1.75}$$

where "Tr" denotes the *trace* (sum of the diagonal elements) of its argument matrix. The fact that γ has a matrix representation similar to that of an operator gives new significance to the name "density matrix" and also explains why these quantities are sometimes called *reduced density operators. If* B is a two-particle operator, then the development analogous to that of the two preceding paragraphs leads to the following equations

$$\mathsf{b}(\mathbf{x}_1, \mathbf{x}_2) \;=\; \sum_{\rho\tau\,\mu\nu} |\phi_\rho(\mathbf{x}_1) \, \phi_\tau(\mathbf{x}_2) > \; b_{\rho\tau,\mu\nu} \; < \phi_\mu(\mathbf{x}_1) \, \phi_\nu(\mathbf{x}_2)| \tag{1.76}$$

$$b_{\rho\tau,\mu\nu} = < \phi_\rho(\mathbf{x}_1)\,\phi_\tau(\mathbf{x}_2) \,|\, \mathsf{b}(\mathbf{x}_1,\mathbf{x}_2) \,|\, \phi_\mu(\mathbf{x}_1)\,\phi_\nu(\mathbf{x}_2) > \tag{1.77}$$

$$\mathbf{\Gamma}(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1',\mathbf{x}_2') = \sum_{\mu\nu\lambda\sigma} |\phi_\mu(\mathbf{x}_1)\,\phi_\nu(\mathbf{x}_2)> \Gamma_{\mu\nu,\lambda\sigma} < \phi_\lambda(\mathbf{x}_1')\,\phi_\sigma(\mathbf{x}_2')| \tag{1.78}$$

$$\Gamma_{\mu\nu,\lambda\sigma} = \int d\mathbf{x}_1\, d\mathbf{x}_2\, d\mathbf{x}_1'\, d\mathbf{x}_2'\, \phi_\mu^*(\mathbf{x}_1)\,\phi_\nu^*(\mathbf{x}_2)\,\Gamma(\mathbf{x}_1,\mathbf{x}_2;\mathbf{x}_1',\mathbf{x}_2')\,\phi_\lambda(\mathbf{x}_1')\,\phi_\sigma(\mathbf{x}_2') \tag{1.79}$$

$$\begin{aligned} < \mathsf{B} > &= \sum_{\mu\nu\lambda\sigma} b_{\lambda\sigma,\mu\nu}\,\Gamma_{\mu\nu,\lambda\sigma} \\ &= \sum_{\lambda\sigma} (\mathbf{b}\,\mathbf{\Gamma})_{\lambda\sigma,\lambda\sigma} \end{aligned} \tag{1.80}$$

Regarding the pair $\lambda\sigma$ as a single composite index, the last line of Eq. (1.80) is seen to be the trace of $\mathbf{b\Gamma}$.

G. Projection Operators

A *projection operator* can be thought of as an operator that retains only certain of the components resulting when a function to which it is applied is subjected to a specified orthonormal expansion. The number of components to be retained may be finite (example: retain the first three components) or infinite (example: retain alternate components); the retained components may have certain symmetry properties (example: retain only even functions), or they may not.

One of the key properties of a projection operator is that its second consecutive application to the same function will not produce further change. This must be so because the first application removes all components of the function that are not to be retained; all reamining components will survive a second screening. The requirement we have identified is equivalent to the operator equation

$$\mathcal{P}^2 = \mathcal{P} \tag{1.81}$$

Projection operators are therefore *idempotent*.

A second requirement on a projection operator is that it gives the same result whether it is applied to the first, or to the second factor of a scalar product, so for arbitrary ϕ and χ,

$$< \phi \,|\, \mathcal{P}\, \chi > = < \mathcal{P}\, \phi \,|\, \chi > \tag{1.82}$$

To see that this must be so, consider ϕ and χ both to be expanded in the orthonormal set with respect to which $\mathcal{P}$ is defined. Both of the preceding scalar products will consist only of contributions from the components retained by $\mathcal{P}$. We conclude that $\mathcal{P}$ must be self-adjoint:

$$\mathcal{P}^\dagger \,=\, \mathcal{P} \tag{1.83}$$

Equations (1.81) and (1.83) may be taken as the formal conditions defining a projection operator.

Every projection operator $\mathcal{P}$ has an orthogonal complement $\mathcal{Q}$ whose effect is to retain only the components removed by $\mathcal{P}$. For any function Φ, $\Phi = \mathcal{P}\Phi + \mathcal{Q}\Phi$; this is equivalent to the operator equation

$$\mathcal{P} + \mathcal{Q} \,=\, \mathcal{F} \tag{1.84}$$

where $\mathcal{F}$ may be viewed as the limiting case of a projection operator which retains all components. The orthogonal complement of $\mathcal{F}$ is the null vector 0. The orthogonal complement of a projection operator is itself a projection operator. We show this formally:

$$\mathcal{Q}^\dagger \,=\, (\mathcal{F} - \mathcal{P})^\dagger \,=\, \mathcal{F} - \mathcal{P} \,=\, \mathcal{Q} \tag{1.85}$$

$$\mathcal{Q}^2 \,=\, (\mathcal{F} - \mathcal{P})\,(\mathcal{F} - \mathcal{P}) \,=\, \mathcal{F} - 2\,\mathcal{P} + \mathcal{P}^2 \,=\, \mathcal{F} - \mathcal{P} \,=\, \mathcal{Q} \tag{1.86}$$

The term *orthogonal complement* is aptly chosen, as

$$\mathcal{Q}\,\mathcal{P} \,=\, (\mathcal{F} - \mathcal{P})\,\mathcal{P} \,=\, \mathcal{P} - \mathcal{P}^2 \,=\, \mathcal{P} - \mathcal{P} \,=\, 0 \tag{1.87}$$

Since $\mathcal{P}$ is the orthogonal complement of $\mathcal{Q}$ as well as $\mathcal{Q}$ being that of $\mathcal{P}$, we also have $\mathcal{P}\mathcal{Q} = 0$.

If a projection operator $\mathcal{P}$ has been specified in terms of the particular orthonormal functions it retains, it may be written in bra-ket notation as

$$\mathcal{P} = \sum_{\mu;(\phi_\mu \in \mathcal{G}_P)} |\phi_\mu > < \phi_\mu| \tag{1.88}$$

where $\mathcal{G}_P$ is the space containing the functions retained by $\mathcal{P}$. The orthogonal complement of $\mathcal{P}$, denoted $\mathcal{Q}$, consists in bra-ket notation of the terms omitted from the $\mathcal{P}$ sum:

$$\mathcal{Q} = \sum_{\mu;(\phi_\mu \perp \mathcal{G}_P)} |\phi_\mu > < \phi_\mu| \tag{1.89}$$

where the restriction is to ϕ_μ orthogonal to every function in $\mathcal{G}_P$. It can be shown that the same $\mathcal{P}$ results if we transform from the $\{\phi_\mu\} \in \mathcal{G}_P$ to any other orthonormal set of functions within the same space, so that the projection operator $\mathcal{P}$ is determined by the *space* onto which it projects, rather than upon the individual functions used in Eq. (1.88). Similar remaks apply to $\mathcal{Q}$. Equations (1.88) and (1.89) show that a projection operator and its orthogonal complement can be interpreted as generating a partial resolution of the identity defined by the subspace $\mathcal{G}_P$.

It is sometimes inconvenient to specify the particular functions needed to express a projection operator in the form given by Eq. (1.88). It may instead be possible to devise an operator that can be shown to accomplish the desired projection on formal grounds. This approach is ordinarily practical when the projection subspace is defined by the symmetry of the functions it contains. In particular, we shall now concentrate on one type of symmetry that is always present in many-fermion systems, that with respect to permutation of the particle numbering. For an N-particle system, there are $N!$ permutation operators (including the identity operator); these operators form a group known as the *symmetry group*. A requirement of the Fermi-Dirac statistics is that many-particle wavefunctions be antisymmetric under permutationns of the particle numbering; this is equivalent to requiring that the wavefunction belong to the fully antisymmetric representation of the symmetry group, for which the representative of each permutation is its parity. The partity of a permutation is $+1$ if an even number of pairwise interchanges is needed to carry it out, and -1 if an

odd number of such interchanges is required. We shall write $(-)^P$ to denote the parity of a permutation P.

An antisymmetric function Φ can be generated from a suitable Φ_0 by the application of the operator

$$\mathcal{A} = \frac{1}{N!} \sum_{\mathsf{P}} (-)^P \mathsf{P} \tag{1.90}$$

where the sum is over all $N!$ permutations. The operator $\mathcal{A}$ is called the *antisymmetrizer*. The factor $1/N!$ has been inserted to give $\mathcal{A}$ what will prove to be a convenient scaling. As is well known, the construction of Φ from Φ_0 will fail if Φ_0 is symmetric under any permutational interchange.

It is instructive to verify that the antisymmetrizer indeed constructs a totally antisymmetric function. Consider application of a permutation Q to a function $\Phi = \mathcal{A}\Phi_0$:

$$\mathsf{Q}\,\Phi = \mathsf{Q}\,\frac{1}{N!} \sum_{\mathsf{P}} (-)^P \mathsf{P}\,\Phi_0 = (-)^Q \frac{1}{N!} \sum_{\mathsf{P}} (-)^Q (-)^P \mathsf{Q}\,\mathsf{P}\,\Phi_0 \tag{1.91}$$

where a factor of unity has been inserted into the last member of Eq. (1.91) in the form $(-)^Q(-)^Q$. This has been done so as to obtain the product $(-)^Q(-)^P$, which is the parity of the product QP. We next use the group-theoretical result that different operators P must yield different products when multiplied by Q. This means that a summation over P is equivalent to a summation over (QP), or setting $\mathsf{QP} = \mathsf{P}'$,

$$\frac{1}{N!} \sum_{\mathsf{P}} (-)^Q\,(-)^P\,\mathsf{Q}\,\mathsf{P} = \frac{1}{N!} \sum_{\mathsf{P}'} (-)^{P'} \mathsf{P}' \tag{1.92}$$

Eq. (1.91) thus reduces to $\mathsf{Q}\Phi = (-)^Q \mathcal{A}\Phi_0 = (-)^Q \Phi$, as required. In conclusion, the antisymmetrizer can be considered as the projection operator that retains the permutationally totally antisymmetric component of a many-particle wavefunction. We now verify that $\mathcal{A}$ is indeed a projection operator. First

$$\mathcal{A}^2 = \frac{1}{N!} \sum_{\mathsf{Q}} (-)^Q\,\mathsf{Q}\,\frac{1}{N!} \sum_{\mathsf{P}} (-)^P\,\mathsf{P} = \frac{1}{N!} \sum_{\mathsf{Q}} \frac{1}{N!} \sum_{\mathsf{P}} (-)^Q\,(-)^P\,\mathsf{Q}\,\mathsf{P} \tag{1.93}$$

As before, we identify QP as P', $(-)^Q(-)^P$ as $(-)^{P'}$, and the summation over P as equivalent to one over P'. We thus have

$$\mathcal{A}^2 = \frac{1}{N!}\sum_{\mathsf{Q}} \mathcal{A} = \mathcal{A} \tag{1.94}$$

the last equality resulting because the Q summation accumulates $N!$ terms, each equal to $\mathcal{A}$. Finally,

$$\mathcal{A}^\dagger = \frac{1}{N!}\sum_{\mathsf{P}}(-)^P \mathsf{P}^\dagger = \frac{1}{N!}\sum_{\mathsf{P}}(-)^P \mathsf{P}^{-1} \tag{1.95}$$

We have replaced $\mathsf{P}^\dagger$ by P^{-1} by making use of the fact that P is a symmetry operator and must be unitary. Then, noting that the parity of P is the same as that of P^{-1} and that the sum over P can be shown equivalent to a sum over P^{-1}, we can verify the other requirement that $\mathcal{A}$ must satisfy to be a projection operator: $\mathcal{A}^\dagger = \mathcal{A}$.

Before leaving the subject of projection operators, let us consider them from the viewpoint of matrix representations. If we use a basis containing the functions of the projection space, placing these last, Eq. (1.88) shows that all elements of $\mathbf{Q}$, the matrix of a projection operator $\mathcal{Q}$, will be zero except for the diagonal elements referring to ϕ_ν in $\mathcal{G}_Q$, which will be unity. The matrix $\mathbf{Q}$ is therefore of the form

$$\mathbf{Q} = \begin{pmatrix} 0 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & 1 \end{pmatrix} \tag{1.96}$$

where the unit submatrix has the dimensionality of $\mathcal{G}_Q$. If a function ψ is represented by a vector $\mathbf{c}$, the action of $\mathcal{Q}$ will be to set all the components of ψ outside of $\mathcal{G}_Q$ to zero, and the vector representing $\mathcal{Q}\psi$ will be $\mathbf{Qc}$, of the block form

$$\mathbf{Q}\,\mathbf{c} = \begin{pmatrix} 0 \\ \vdots \\ \cdots\cdots \\ c_\mu \\ \vdots \end{pmatrix} \tag{1.97}$$

If we wish to restrict consideration to the components within $\mathcal{G}_Q$ that can be produced by action of an arbitrary operator A on functions themselves within $\mathcal{G}_Q$, we may form $\mathcal{Q}\mathsf{A}\mathcal{Q}$, whose matrix will be of the block form

$$\mathbf{Q}\,\mathbf{A}\,\mathbf{Q} = \begin{pmatrix} 0 & \vdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \vdots & \cdots A_{\mu\nu} \cdots \end{pmatrix} \tag{1.98}$$

It is clear from Eq. (1.96) that $\mathbf{Q}$ (and therefore also Q) will have eigenvalues 0 and 1 (except in the trivial case $\mathsf{Q} = \mathsf{F}$), and that projection operators (other than $\mathcal{F}$) are therefore singular. As is apparent from the form of Eq. (1.98) and from the fact that $\mathbf{QAQ}$ contains $\mathbf{Q}$ as factors, the projection of any operator must also be singular. It is sometimes useful to construct the matrix obtained therefrom by inverting the block of $A_{\mu\nu}$ from Eq. (1.98); note that this is not the same as $\mathbf{QA}^{-1}\mathbf{Q}$, because every element of $\mathbf{A}^{-1}$ contains contributions from all rows and columns of $\mathbf{A}$. One way of indicating this "inverse of the corner" of $\mathbf{A}$ is as $\mathbf{Q}[\mathbf{QAQ} + \mathbf{P}]^{-1}\mathbf{Q}$, where $\mathbf{P} = \mathbf{1} - \mathbf{Q}$; the matrix whose inverse is taken is of the form

$$\mathbf{Q}\,\mathbf{A}\,\mathbf{Q} + \mathbf{P} = \begin{pmatrix} 1 & \vdots & 0 \\ \cdots & \vdots & \cdots \\ 0 & \vdots & \cdots A_{\mu\nu} \cdots \end{pmatrix} \tag{1.99}$$

and the further applications of $\mathbf{Q}$ remove the unit matrix after the inversion. It is also possible to use the shorthand notation, suggested by Löwdin

$$\frac{\mathbf{Q}}{\mathbf{A}} = \mathbf{Q}\,[\mathbf{Q}\,\mathbf{A}\,\mathbf{Q} + \mathbf{P}]^{-1}\,\mathbf{Q} \tag{1.100}$$

with $\mathcal{Q}/\mathsf{A}$ for its operator equivalent. The foregoing analysis shows that $\mathcal{Q}/\mathsf{A}$ is a projected inverse of $\mathbf{A}$ in the sense that

$$\mathbf{Q}\,\mathbf{A}\,\mathbf{Q}\,\frac{\mathbf{Q}}{\mathbf{A}} = \frac{\mathbf{Q}}{\mathbf{A}}\,\mathbf{Q}\,\mathbf{A}\,\mathbf{Q} = \mathbf{Q} \tag{1.101}$$

The equivalent operator equation is

$$\mathcal{Q}\,\mathsf{A}\,\mathcal{Q}\,\frac{\mathcal{Q}}{\mathsf{A}} = \frac{\mathcal{Q}}{\mathsf{A}}\,\mathcal{Q}\,\mathsf{A}\,\mathcal{Q} = \mathcal{Q} \tag{1.102}$$

In what follows we discuss some useful theorems concerning the projection and permutation operators.

Theorem 1: The diagonal elements of a projection operator (in its matrix representation) obey the inequality

$$0 \leq \mathcal{P}_{ij} \leq 1 \tag{1.103}$$

The diagonal elements of the matrix representation of the projection operator $\mathcal{P}$ can be represented by

$$< \mathcal{P} > = < \phi \,|\, \mathcal{P} \,|\, \phi > = < \phi \,|\, \mathcal{P}^\dagger \mathcal{P} \,|\, \phi > \tag{1.104}$$

that is

$$< \mathcal{P} > = < \mathcal{P}\,\phi \,|\, \mathcal{P}\,\phi > \geq 0 \tag{1.105}$$

Now consider the operator

$$\mathcal{Q} = 1 - \mathcal{P} \tag{1.106}$$

This operator belongs to the remaining subspace not described by $\mathcal{P}$. Now if $\mathcal{P}^\dagger = \mathcal{P}$, then $\mathcal{Q}^\dagger = \mathcal{Q}$. Since $1 - \mathcal{P}$ is also a projection operator, it follows that

$$< 1 - \mathcal{P} > \geq 0 \tag{1.107}$$

which leads directly to

$$< \mathcal{P} > \leq 1 \tag{1.108}$$

Combining Eqs. (1.105) and (1.108) we get

$$0 \leq < \mathcal{P} > \leq 1 \tag{1.109}$$

In order to make better use of notational advantage of the antisymmetrization operator, it is necessary to reveal some of its mathematical properties. To do this we first prove the following theorem.

Theorem 2: The permutation operator P is unitary. Consider a generalized wavefunction of N electrons given by

$$\phi = \phi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) \tag{1.110}$$

The operation of P is to permute the electron coordinates to some new order $\mathbf{x}'_1, \mathbf{x}'_2, \cdots, \mathbf{x}'_N$, where each $\mathbf{x}'_i$ is one of the $\mathbf{x}_i$. We designate this permutation symbolically by

$$\mathsf{P}\,\phi(\mathbf{x}) \;=\; \phi(\mathbf{x}') \tag{1.111}$$

where $\mathbf{x}$ represents the collection of coordinates in the unpermuted order and $\mathbf{x}'$ the collection of coordinates in the permuted order. In terms of the inverse permutation this relationship becomes

$$\phi(\mathbf{x}) \;=\; \mathsf{P}^{-1}\,\phi(\mathbf{x}') \tag{1.112}$$

Now consider the definite integral

$$<\mathsf{P}\,\phi_1(\mathbf{x})\,|\,\phi_2(\mathbf{x})> \;=\; <\phi_1(\mathbf{x})\,|\,\mathsf{P}^{\dagger}\,|\,\phi_2(\mathbf{x})> \tag{1.113}$$

Another way to transform the left-hand integral is to use Eq. (1.111) for $\mathsf{P}\phi_1(\mathbf{x})$ and Eq. (1.112) for $\phi_2(\mathbf{x})$. This produces

$$<\mathsf{P}\,\phi_1(\mathbf{x})\,|\,\phi_2(\mathbf{x})> \;=\; <\phi_1(\mathbf{x}')\,|\,\mathsf{P}^{-1}\,|\,\phi_2(\mathbf{x}')> \tag{1.114}$$

Comparing the right-hand integral in Eqs. (1.113) and (1.114) and remembering that $\mathbf{x}$ and $\mathbf{x}'$ represent the same collection of dummy variables of integration, we obtain

$$\mathsf{P}^{\dagger} \;=\; \mathsf{P}^{-1} \tag{1.115}$$

i.e., P is a unitary operator.

Theorem 3: The expectation values of the two-electron transposition operator $\mathsf{P}_{\mu\nu}$ obey the inequality

$$-1 \;\leq <\mathsf{P}_{\mu\nu}> \leq\; 1 \tag{1.116}$$

Consider the operator

$$\frac{1}{2}\left(1 - \mathsf{P}_{\mu\nu}\right) \tag{1.117}$$

It is readily seen that $\mathsf{P}_{\mu\nu} = \mathsf{P}_{\mu\nu}^{-1}$; that is, $\mathsf{P}_{\mu\nu}$ is its own inverse. Furthermore, from Eq. (1.115) $\mathsf{P}_{\mu\nu}$ is also unitary, i.e., $\mathsf{P}^{\dagger} = \mathsf{P}^{-1}$. This means that $\mathsf{P}_{\mu\nu}^{\dagger}$ and therefore $\mathsf{P}_{\mu\nu}$ is also self-adjoint. The square of the operator (1.117) is

$$\begin{aligned}\left[\frac{1}{2}(1-\mathsf{P}_{\mu\nu})\right]^2 &= \left[\frac{1}{2}(1-\mathsf{P}_{\mu\nu})\right]\left[\frac{1}{2}(1-\mathsf{P}_{\mu\nu}^{-1})\right] \\ &= \frac{1}{4}\left(1+\mathsf{P}_{\mu\nu}\mathsf{P}_{\mu\nu}^{-1}-\mathsf{P}_{\mu\nu}-\mathsf{P}_{\mu\nu}^{-1}\right) = \frac{1}{2}\left(1-\mathsf{P}_{\mu\nu}\right) \qquad (1.118)\end{aligned}$$

which means that $\frac{1}{2}(1-\mathsf{P}_{\mu\nu})$ is also idempotent and therefore is a projection operator. Theorem 1 now gives Eq. (1.116) immediately.

Theorem 3 now leads to some useful inequalities involving integrals. The expectation value of $\mathsf{P}_{\mu\nu}$ for an arbitrary two-electron wavefunction $\xi(\mu,\nu)$ is given by

$$<\mathsf{P}_{\mu\nu}> = <\xi(\mu,\nu)\,|\,\mathsf{P}_{\mu\nu}\,|\,\xi(\mu,\nu)> = <\xi(\mu,\nu)\,|\,\xi(\nu,\mu)> \qquad (1.119)$$

where $\xi(\mu,\nu)$ and $\xi(\nu,\mu)$ have transposed electron coordinates. Then by Theorem 3,

$$-1 \leq <\xi(\mu,\nu)\,|\,\xi(\nu,\mu)> \leq 1 \qquad (1.120)$$

If $\xi(\mu,\nu)$ is normalized, we can write Eq. (1.120) as

$$-1 \leq <\xi(\mu,\nu)\,|\,\xi(\nu,\mu)> \leq <\xi(\mu,\nu)\,|\,\xi(\mu,\nu)> \qquad (1.121)$$

Now consider the special case when $\xi(\mu,\nu)$ is given as a product of two orbitals, namely,

$$\xi(\mu,\nu) = \phi_1(\mu)\,\phi_2(\nu) \qquad (1.122)$$

Equation (1.121) then gives

$$-1 \leq <\phi_1(\mu)\,|\,\phi_2(\mu)><\phi_1(\nu)\,|\,\phi_2(\nu)> \leq <\phi_1(\mu)\,|\,\phi_1(\mu)><\phi_2(\nu)\,|\,\phi_2(\nu)> \qquad (1.123)$$

which can also be written

$$|\,<\phi_1(\mu)\,|\,\phi_2(\mu)>\,|^2 \leq <\phi_1(\mu)\,|\,\phi_1(\mu)><\phi_2(\mu)\,|\,\phi_2(\mu)> \qquad (1.124)$$

If orbitals ϕ_1 and ϕ_2 are merely required to be quadratically integrable functions, Eq. (1.124) is known as the *Schwarz integral inequality*. This relationship tells us that the absolute

value of an overlap integral is always less than unity if the two overlapping functions are normalized.

Now let

$$\xi(\mu,\nu) \;=\; \phi_i(\mu)\,\phi_j(\nu)\,(r_{\mu\nu})^{-1} \tag{1.125}$$

where ϕ_i and ϕ_j are normalized one-electron functions. Eq. (1.121) then leads to

$$\left\langle \phi_i(\mu)\,\phi_j(\nu) \mid \frac{1}{r_{\mu\nu}} \mid \phi_j(\mu)\,\phi_i(\nu) \right\rangle \;\leq\; \left\langle \phi_i(\mu)\,\phi_j(\nu) \mid \frac{1}{r_{\mu\nu}} \mid \phi_i(\mu)\,\phi_j(\nu) \right\rangle \tag{1.126}$$

II. THE VARIATIONAL METHOD

A. Preliminary Remarks

The many-electron Schrödinger equation [Eq. (1.1)] cannot be separated into ordinary differential equations in terms of any known coordinate system. This fact is, of course, the consequence of the occurrence of the $r_{\mu\nu}^{-1}$ terms, associated with interactions among the electrons. Nonseparability of the Schrödinger equation for systems of physical interest turns out to be the general rule, and only a very few simple systems admit of exact solutions. Furthermore, many of these simple systems do not directly correspond to those exhibited by nature, e.g., the particle in a box, the harmonic oscillator, and the rigid rotator. In what follows we shall discuss in some detail the basic features of the variational method which has been found useful in the construction of approximate solutions to the Schrödinger equation. In fact, the variation method can be described as an alternative formulation of Schrödinger equation, a formulation which has the advantage of suggesting a route to approximate solutions of any desired degree of accuracy.

B. Basis of the Method

Let us consider the time-independent Schrödinger equation in the compact symbolic form

$$\mathcal{H}\,\psi \;=\; E\,\psi \tag{2.1}$$

where ψ and E are an exact eigenfunction and eigenenergy, respectively, of the Hamiltonian operator $\mathcal{H}$. In general, the mathematical form of the Hamiltonian will be such that Eq. (2.1) will not be separable in any known coordinate system. Thus one cannot hope to obtain an analytical closed-form expression for the eigenfunction ψ. Let us now examine the properties of the functional (whereas the domain of a function, say $f(x)$, is a region of coordinate space, the domain of a functional, say $F[f(x)]$, is a space of admissible functions;

thus, one may think of a functional as a function of a function, i.e., a *superfunction*) $\mathcal{E}[\xi]$ defined by

$$\mathcal{E}[\xi] \;=\; \frac{<\xi\,|\,\mathcal{H}\,|\,\xi>}{<\xi\,|\,\xi>} \tag{2.2}$$

where $\mathcal{H}$ is the exact Hamiltonian of the system defined by Eq. (2.1) and ξ is an arbitrary function of the system coordinates, subject only to the restriction that it is normalizable over the configuration space of the system whose exact wavefunction is ψ. It is apparent that if ξ is identical with ψ, the functional $\mathcal{E}[\xi]$ is the energy E (a constant). Let us regard ξ as a trial approximation to the exact wavefunction ψ and assume that ξ differs from ψ by no more than a first-order variation, namely,

$$\xi \;=\; \psi + \delta\psi \tag{2.3}$$

This means that ξ and ψ are normalizable in the same space. Now let us consider the expectation value of the operator $\mathcal{H} - E$ with respect to the trial function ξ. First, we note that application of the operator $\mathcal{H} - E$ to the trial function ξ leads to the result

$$(\mathcal{H} - E)\,\xi \;=\; (\mathcal{H} - E)\,(\psi + \delta\psi) \;=\; (\mathcal{H} - E)\,\delta\psi \tag{2.4}$$

i.e., the operator $\mathcal{H} - E$ annihilates the exact eigenfunction. The desired expectation value then is

$$\begin{aligned} <\mathcal{H} - E> = \frac{<\xi\,|\,\mathcal{H} - E\,|\,\xi>}{<\xi\,|\,\xi>} &= \frac{<\xi\,|\,\mathcal{H} - E\,|\,\delta\psi>}{<\xi\,|\,\xi>} = \frac{<(\mathcal{H} - E)\,\xi\,|\,\delta\psi>}{<\xi\,|\,\xi>} \\ &= \frac{<(\mathcal{H} - E)\,\delta\psi\,|\,\delta\psi>}{<\xi\,|\,\xi>} = \frac{<\delta\psi\,|\,\mathcal{H} - E\,|\,\delta\psi>}{<\xi\,|\,\xi>} \end{aligned} \tag{2.5}$$

where the turnover rule has been used twice. By simple rearrangement of the above equation, the function $\mathcal{E}[\xi]$ is seen to be

$$\mathcal{E}[\xi] \;=<\mathcal{H}>=\; E + \frac{<\delta\psi\,|\,\mathcal{H} - E\,|\,\delta\psi>}{<\xi\,|\,\xi>} \tag{2.6}$$

In this result we see that the functional $\mathcal{E}[\xi]$ is an approximation to the exact energy E and differs from E by only a second-order term in $\delta\psi$. In other words, although the trial

function ξ is in error to the first order, the total energy is in error to only the second order. Thus we can write

$$\delta\mathcal{E}[\xi] = \delta < \mathcal{H} >= 0 \tag{2.7}$$

Equation (2.7), which may be regarded as a statement of the variational principle, can now be shown to lead to to Schrödinger's equation; i.e., one can show that the condition (2.7) establishes that the trial function ξ is an eigenfunction of $\mathcal{H}$. This problem is most generally approached by the use of the calculus of variaitons, an area of mathematics concerned with a generalization of maxima and minima problems. It is desired to find the function ξ such that the functional

$$J[\xi] = < \xi \,|\, \mathcal{H} \,|\, \xi > \tag{2.8}$$

is stationary, i.e., has an extremal value with respect to arbitrary, small variations in ξ. To ensure consistency with probabilistic interpretations we impose on the functions ξ the normalization restriction

$$< \xi \,|\, \xi >= 1 \tag{2.9}$$

The variational problem can be cast in a more convenient form by defining the additional functional

$$K[\xi] = < \xi \,|\, \xi > -1 \tag{2.10}$$

and to find, instead, the functions ξ which make the new funcitonal

$$L[\xi] = J[\xi] - \lambda\, K[\xi] \tag{2.11}$$

stationary without any restrictions. The quantity λ is a real parameter known as a *lagrangian multiplier*. The stationary values of $L[\xi]$ will be those for which $\delta L[\xi]$ vanishes, and provided ξ is always normalized, this means that $\delta < \mathcal{H} >= 0$ is also satisfied.

The first-order variation in $L[\xi]$ is

$$\delta L[\xi] = < \delta\xi \,|\, \mathcal{H} \,|\, \xi > + < \xi \,|\, \mathcal{H} \,|\, \delta\xi > -\lambda\, < \delta\xi \,|\, \xi > -\lambda\, < \xi \,|\, \delta\xi > -\delta\lambda\, K[\xi] \tag{2.12}$$

It is convenient to introduce the notational simplification

$$< \delta\xi \,|\, \mathcal{H} \,|\, \xi > - \lambda \, < \delta\xi \,|\, \xi > = \, Q \tag{2.13}$$

Then, since

$$(< \delta\xi \,|\, \mathcal{H} \,|\, \xi > - \lambda \, < \delta\xi \,|\, \xi >)^{\dagger} \, = < \xi \,|\, \mathcal{H} \,|\, \delta\xi > - \lambda \, < \xi \,|\, \delta\xi > \tag{2.14}$$

we can write Eq. (2.12) as

$$\delta L[\xi] \, = \, Q + Q^{*} - \delta\lambda \, K[\xi] \tag{2.15}$$

If $< \xi | \xi > = 1$, the coefficient of $\delta\lambda$ vanishes, and the functional $L[\xi]$ will be stationary only if Q (and thus also its complex conjugate Q^{*}) vanishes, that is,

$$< \delta\xi \,|\, \mathcal{H} \,|\, \xi > - \lambda \, < \delta\xi \,|\, \xi > = \, 0 \tag{2.16}$$

If this equation is to be satisfied by any arbitrary variaiton $\delta\xi$, it follows at once that

$$\mathcal{H} \,|\xi > - \lambda \,|\xi > = \, 0 \tag{2.17}$$

which is just the Schrödinger equation. The condition that $\delta < \mathcal{H} >$ vanish is thus sufficient to make ξ an eigenfunction of the exact Hamiltonian $\mathcal{H}$. It is apparent from Eq. (2.2) that the lagrangian multiplier λ is to be identified with the functional $\mathcal{E}[\xi] = < \mathcal{H} >$.

Unfortunately, the condition $\delta < \mathcal{H} > = 0$ is necessary but not sufficient to establish that λ is a minimum. The problem of establishing true sufficiency of the conditions leading to true minima of functionals is a difficult one and has not been solved in general. In practice, however, λ is usually found to be a true minimum, and one accepts the fundamental theorem on faith in most physical applications.

As it now stands, the variation principle does not immediately disclose any particular usefulness. Now let $\{\psi_i\}$ be a complete orthonormal set of eigenfunctions of $\mathcal{H}$ and let $\{E_i\}$ be the corresponding eigenvalues. Even though we do not usually know the ψ_i explicitly, we can nevertheless express the trial function ξ in this set by the expansion

$$\xi \, = \, \sum_{i=1}^{\infty} \psi_i \, c_i + \int \psi_\lambda \, C_\lambda \, d\lambda \tag{2.18}$$

where the integral allows for the possibility of a continuum, i.e., a set of continuous eigenvalues. The complete set of eigenfunctions of the central-field Hamiltonian consists of a discrete set [the orbitals $R_{n;}(r)\Theta_{lm_l}(\theta)\Phi_{m_l}(\phi)$] and just such a continuum (associated with the radial equation). The discrete set corresponds physically to the bound states of an electron moving about the nucleus ($E < 0$), while the continuous set corresponds to a free electron scattered by the nuecleus ($E > 0$). For simplicity we shall assume that our set $\{\psi_i\}$ is discrete, so that the integral term in the expansion (2.18) is missing. Equation (2.2) now becomes

$$\mathcal{E} = \frac{<\xi \,|\, \mathcal{H} \,|\, \xi>}{<\xi \,|\, \xi>} = \frac{\sum_i \sum_j c_i^* \, c_j \, H_{ij}}{\sum_i \sum_j c_i^* \, c_j \, \Delta_{ij}} \tag{2.19}$$

where we have dropped the $[\xi]$ notation for the functional $\mathcal{E}[\xi]$, and where H_{ij} and Δ_{ij} are the matrix elements of the Hamiltonian $\mathcal{H}$ and of the overlap matrix Δ respectively, in the basis of the $\{\psi_i\}$. Since $\{\psi_i\}$ is a complete orthonormal set of eigenfunctions of $\mathcal{H}$ with eigenvalues E_i, Eq. (2.19) reduces to

$$\mathcal{E} = \frac{\sum_i |c_i|^2 \, E_i}{\sum_i |c_i|^2} \tag{2.20}$$

If the trial function ξ is normalized, it follows that

$$<\xi \,|\, \xi> = \sum_i |c_i|^2 = 1 \tag{2.21}$$

so that Eq. (2.20) becomes

$$\mathcal{E} = \sum_i |c_i|^2 \, E_i \tag{2.22}$$

Substracting E_0, the exact energy of the ground state, from both sides of Eq. (2.22) and using Eq. (2.21), we obtain

$$\mathcal{E} - E_0 = \sum_i |c_i|^2 \, (E_i - E_0) \tag{2.23}$$

Since, in general, $E_i \geq E_0$ and $|c_i|^2 \geq 0$ for all i, it follows that

$$\mathcal{E} \geq E_0 \tag{2.24}$$

This very important result is an upper-limit theorem for the energy. The theorem shows that any trial wavefunction ξ (which is normalizable) leads to a value of the energy which

is never lower than the true ground-state energy of the system. In this result resides the power (and also the weakness) of the variational method of approximating solutions to the Schrödinger wave equation. The power lies in the fact that one can choose the "best" wavefunction from several alternatives on the basis of the criterion of lowest energy; the weakness is that the energy turns out to be an insensitive criterion with respect to a "best" wavefunction for other physical properties. In fact, properties other than the energy are not variational, because only the Hamiltonian is used to obtain the wavefunction in the widely used computational chemistry methods. More precisely, Eq. (2.6), which shows that a first-order error in a trial wavefunction leads to no first-order in the expectation value of the Hamiltonian, does not hold for the expectation values of operators in general. The first-order in ψ may be averaged over the coordinates in such a way that even though a very good energy is obtained, the expectation values of other operators may be exceedingly poor. This may happen, for example, if the error is large for certain values of the coordinates upon which a particular operator depends. As a more specific example, suppose we are approximating the electron distribution of a many-electron atom. The first-order error may be distributed in such a way that it is very large in the neighborhood of the nucleus and small elsewhere. This error will affect the energy only to second order, but an operator sensitive to properties of the electron distribution next to the nucleus, e.g., an operator related to electron field gradients, may yield expectation values which are completely erroneous; e.g., the sign as well as the magnitude may be incorrect. One of the major problems of quantum mechanics is to find other and more sensitive criteria than the energy for the general acceptability of wavefunctions.

Since it is not possible in practice to expand trial wavefunctions in terms of all the infinite number of functions of a complete basis set, one has to attempt an expansion in terms of a finite number of such functions. Such a finite set of functions is referred to as a *truncated basis set*. In practice one seeks a finite number n of basis functions, attempting to find those n particular functions which lead to the minimum energy. The actual number of functions required to lead to a given energy depends upon the particular basis set and upon the system

of interest. For a given system of interest, some basis sets may be more efficient than others, in the sense that fewer basis functions are required to lead to a given energy.

Let us now expand the trial wavefunction ξ_k in some arbitrary basis set $\{\phi_i\}$ and truncate the expansion at the nth function, namely

$$\xi_k \;=\; \sum_{i=1}^{n} \phi_i \, c_{ik} \tag{2.25}$$

We assume that the functions $\{\phi_i\}$ are linearly independent but not necessarily orthonormal. If we define the matrices

$$\mathbf{\Phi}^\dagger \;=\; \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_n \end{pmatrix}$$

$$\mathbf{\Delta} \;=< \mathbf{\Phi} \,|\, \mathbf{\Phi} > \qquad\qquad \mathbf{C}_k \;=\; \begin{pmatrix} c_{ik} \\ c_{2k} \\ \vdots \\ c_{nk} \end{pmatrix}$$

$$\mathbf{H} \;=< \mathbf{\Phi} \,|\, \mathcal{H} \,|\, \mathbf{\Phi} > \tag{2.26}$$

then the energy $\mathcal{E}$ may be written

$$\mathcal{E} \;=\; \frac{\mathbf{C}_k^\dagger \, \mathbf{H} \, \mathbf{C}_k}{\mathbf{C}_k^\dagger \, \mathbf{\Delta} \, \mathbf{C}_k} \tag{2.27}$$

Applying the variational theorem (2.7), we obtain

$$\delta\mathcal{E} \;=\; 0 \;=\; \delta \, \frac{\mathbf{C}_k^\dagger \, \mathbf{H} \, \mathbf{C}_k}{\mathbf{C}_k^\dagger \, \mathbf{\Delta} \, \mathbf{C}_k} \;=\; \frac{1}{\mathbf{C}_k^\dagger \, \mathbf{\Delta} \, \mathbf{C}_k} \, (\delta \mathbf{C}_k^\dagger \, \mathbf{H} \, \mathbf{C}_k - \mathcal{E} \, \delta \mathbf{C}_k^\dagger \, \mathbf{\Delta} \, \mathbf{C}_k) + c.c. \tag{2.28}$$

Thus, the variational theorem is satisfied if

$$\delta \mathbf{C}_k^\dagger \, (\mathbf{H} \, \mathbf{C}_k - \mathcal{E} \, \mathbf{\Delta} \, \mathbf{C}_k) \;=\; 0 \tag{2.29}$$

In order for Eq. (2.29) to be satisfied for any arbitrary variation $\delta \mathbf{C}_k$ of the coefficients, it is necessary that the term in brackets vanish, namely,

$$(\mathbf{H} - \mathcal{E} \, \mathbf{\Delta}) \, \mathbf{C}_k \;=\; 0 \tag{2.30}$$

The conditions leading to Eq. (2.30) are equivalent to the n conditions

$$\frac{\partial \mathcal{E}}{\partial c_{ik}} = 0 \qquad \forall\, i = 1, 2, \cdots, n \tag{2.31}$$

i.e., the expansion coefficients $\{c_{ik}\}$ are variation paramters to be chosen so as to minimize $\mathcal{E}$. The nontrivial solutions of Eq. (2.30) are given by the n roots of the secular determinant

$$Det\,(\mathbf{H} - \mathcal{E}\,\mathbf{\Delta}) = 0 \tag{2.32}$$

Labelling the n roots $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$, we get the associated eigenvectors $\mathbf{C}_1, \mathbf{C}_2, \cdots, \mathbf{C}_n$, respectively. We now define the matrices

$$\mathbf{C} = \Big(\mathbf{C}_1\ \ \mathbf{C}_2 \cdots \mathbf{C}_n\Big) \qquad (\mathbf{\Xi})_{ij} = \epsilon_i\, \delta_{ij} \tag{2.33}$$

and write the Schrödinger equation in the matrix form

$$\mathbf{H}\,\mathbf{C} = \mathbf{\Delta}\,\mathbf{C}\,\mathbf{\Xi} \tag{2.34}$$

In the event that the basis $\{\phi_i\}$ is orthonormal, the above equation becomes

$$\mathbf{H}\,\mathbf{C} = \mathbf{C}\,\mathbf{\Xi} \tag{2.35}$$

One then obtains the n different trial functions

$$\xi = \mathbf{\Phi}\,\mathbf{C} \tag{2.36}$$

where

$$\xi = \Big(\xi_1\ \ \xi_2 \cdots \xi_n\Big) \tag{2.37}$$

Equation (2.36) is just the matrix form of the n equations given by Eq. (2.25) when $i = 1, 2, \cdots, n$.

If ϵ_1 is the lowest root, then by Eq. (2.24) we have

$$\epsilon_1 \geq E_1 \tag{2.38}$$

where E_1 is the exact energy of the lowest state of the system whose wavefunction is approximated by ξ_1.

It can be shown that if one extends the truncated basis set by one more function to the $n+1$ order, the roots of the $n \times n$ secular equation (2.32) will separate the roots of the $(n+1) \times (n+1)$ secular equation resulting from the extended basis. Denoting the roots of the extended basis by primes, this fact can be written

$$\epsilon'_1 < \epsilon_1 < \epsilon'_2 < \epsilon_2 < \epsilon'_3 < \epsilon_3 < \epsilon'_4 < \epsilon_4 < \cdots < \epsilon'_{n-2} < \epsilon_{n-2} < \epsilon'_{n-1}$$
$$< \epsilon_{n-1} < \epsilon'_n < \epsilon_n < \epsilon'_{n+1} \tag{2.39}$$

By extending this analysis to an infinite basis set one can readily show that the jth root ϵ_j of the secular equation of order $n \times n$ is an upper limit to the jth exact energy, that is

$$\epsilon_j \geq E_j \qquad\qquad j = 1, 2, \cdots, n \tag{2.40}$$

The roots $\{\epsilon_j\}$ are thus approximations to the energy of the ground state and the $n-1$ lowest excited states of the system. Generally, however, even if ϵ_1 is a very good approximation to the exact ground-state energy E_1, it turns out that the remaining roots are poorer approximations to the excited states. This is due largely to the fact that the n basis functions best suited to represent the ground state are generally not nearly as suitable for representation of other states.

The success of the above approxach obviously depends upon a judicious choice of a particular basis set, the number of basis functions retained, and the particular functions chosen. The criteria for the choice will depend upon the particular problem and cannot easily be generalized.

An alternative way of utilizing the variational principle is to construct the trial wavefunctions in terms of a set of variaiton parameters $\{\alpha_i\}$ which enter into the function in a nonlinear manner. In general we can write

$$\xi = \xi(\alpha_1, \alpha_2, \cdots, \alpha_n) \tag{2.41}$$

For example, the parameters $\{\alpha_i\}$ may enter the trial function ξ as exponential factors. One then chooses the $\{\alpha_i\}$ so as to minimize $\mathcal{E}$, that is

$$\frac{\partial \mathcal{E}}{\partial \alpha_i} \qquad\qquad \forall\, i = 1, 2, \cdots, n \tag{2.42}$$

In practice one often uses trial functions which contain both nonlinear and linear variation parameters. This freedom is useful in that it allows one to introduce the variational parameters in such a way that their effect can be interpreted in a simple physical manner.

C. Atomic Units

Before we begin a discussion of some ways of solving Schrödinger's equation in an approximate manner, it is convenient to introduce a system of dimensionless units commonly used in the quantum theory of atoms and molecules. In this system, mass, length, and time are expressed in units called *atomic units* (abbreviated a.u.).

The atomic unit of mass is the rest mass of the electron, namely, $m_e = 9.1091 \times 10^{-28}$ g. The atomic unit of length is the radius of the first Bohr orbit in hydrogen when the reduced mass μ is replaced with the electron rest mass m_e. Thus the atomic unit of length is

$$a_0 = \frac{\hbar^2}{m_e e^2} = 0.52917 \times 10^{-8} \text{ cm} \tag{2.43}$$

One then writes 1 a.u. (of length) instead of the above value in centimeters. All other lengths are then given in terms of a_0. For example, a length of $1.397\mathring{A}$ (the C-C bond length in benzene) is

$$\frac{1.397}{0.5292} = 2.640 \text{ a.u.} \quad (\text{of length}) \tag{2.44}$$

The atomic unit of time is defined as the time required for an electron to travel 1 a. u. of length in the first Bohr orbit. Since the velocity of the electron in the first Bohr orbit is $e^2/\hbar$, it is apparent that the atomic unit of time is $a_0\hbar/e^2$ (about 2.42×10^{-17} sec). This also means that the atomic unit of velocity is $e^2/\hbar$ (about 2.19×10^8 cm/sec). From the foregoing we see that the rest mass of the proton is 1836.1 a.u. and that the velocity of light (to three significant figures) is given by

$$\frac{3.00 \times 10^{10}}{e^2/\hbar} = 137 \text{ a.u.} \tag{2.45}$$

Relativistic effects on electrons in atoms are of the order of $(v/c)^2$, where v is the velocity of the electron and c is the speed of light. Since v is directly proportional to the atomic number Z, it follows that the relativistic effects are of the order of $(Z/137)^2$.

The atomic unit of energy is chosen as

$$\frac{e^2}{a_0} = 27.210 \text{ ev} = 1 \text{ a.u.} \quad (\text{of} \quad \text{energy}) \tag{2.46}$$

This is just twice the ionization potential of the hydrogen atom if the reduced mass of the electron is replaced by the rest mass. In terms of atomic units the experimental nonrelativistic energy of the helium atom in its ground state (-79.99 ev) is

$$\frac{79.99}{27.21} = -2.905 \text{ a.u.} \tag{2.47}$$

It follows from the above that the charge on the electron, $e = 4.80298 \times 10^{-10}$ esu, represents 1 atomic unit of charge. Also it follows that an atomic unit of angular momentum is given by $\hbar$ $(= 1.05450 \times 10^{-27}$ erg sec).

It is important to be aware of the fact that other conventions are sometimes used for expressing certain physical quantities. For example, some prefer to define the atomic unit of energy as half that given in (2.46), i.e., as $e^2/2a_0$ $(= 13.605$ ev). Sometimes the energy is given in terms of the *Rydberg constant for infinite mass* given by

$$R_\infty = \frac{e^2}{2a_0\hbar c} = 1.09737 \times 10^5 \text{ cm}^{-1} \tag{2.48}$$

Still others use $2R_\infty$ (the double Rydberg). The latter is consistent with the use of (2.46). The *infinite mass* refers to that of the nucleus, i.e.,

$$\lim_{M_N \to \infty} \mu = \frac{m_e M_N}{m_e + M_N} = m_e \tag{2.49}$$

When atomic units are used, one sets $e = \hbar = m_e = 1$ in quantum-mechanical expressions. For example, the kinetic-energy operator $-\hbar^2\nabla^2/2m$ becomes $-\frac{1}{2}\nabla^2$.

The advantage of atomic units is that if all calculations are directly expressed in such units, the results do not vary with subsequent revision of the numerical values of the fundamental constants. Furthermore, there is certain notational simplicity which is useful in writing down quantum-mechanical operators and functions.

D. Variational Treatment of the H_2^+ Ion

As an example of the procedure described in section II.B let us consider the hydrogen molecular ion H_2^+. In this molecule just one electron is moving in the potential field of the two nuclei. We may get a rough description of the lowest orbital of this molecule by considering it as a linear combination of the $1s$ orbitals of the two hydrogen atoms. Let us designate the nuclei by the letters a and b, the $1s$ orbital of an electron in the field of nucleus a alone by ϕ_a, and the $1s$ orbital of an electron in the field of nucleus b alone by ϕ_b. Analytically (in a.u.)

$$\phi_a \;=\; \frac{1}{\pi^{1/2}}\, e^{-r_a} \qquad\qquad \phi_b \;=\; \frac{1}{\pi^{1/2}}\, e^{-r_b} \tag{2.50}$$

If we put

$$\mathbf{\Delta}_{ab} \;=< \phi_a \,|\, \phi_b > \qquad\qquad \mathbf{H}_{aa} \;=< \phi_a \,|\, \mathsf{H} \,|\, \phi_a > \tag{2.51}$$

$$\mathbf{H}_{bb} \;=< \phi_b \,|\, \mathsf{H} \,|\, \phi_b > \qquad\qquad \mathbf{H}_{ab} \;=< \phi_a \,|\, \mathsf{H} \,|\, \phi_b >=\; \mathbf{H}_{ba} \tag{2.52}$$

the determination of the coefficients in the linear combination

$$\psi \;=\; c_a\, \phi_a + c_b\, \phi_b \tag{2.53}$$

leads to the following secular determinant for the energy

$$Det \begin{pmatrix} \mathbf{H}_{aa} - \mathcal{E} & \mathbf{H}_{ab} - \mathbf{\Delta}_{ab}\, \mathcal{E} \\ \mathbf{H}_{ba} - \mathbf{\Delta}_{ab}\, \mathcal{E} & \mathbf{H}_{bb} - \mathcal{E} \end{pmatrix} \;=\; 0 \tag{2.54}$$

From the symmetry of the problem it is evident that $\mathbf{H}_{aa} = \mathbf{H}_{bb}$. Using this relation, the roots of the determinant are found to be

$$\mathbf{H}_{aa} - \mathcal{E} \;=\; \mp\,(\mathbf{H}_{ab} - \mathbf{\Delta}_{ab}\, \mathcal{E})$$

or

$$\mathcal{E}_1 \;=\; \frac{\mathbf{H}_{aa} + \mathbf{H}_{ab}}{1 + \mathbf{\Delta}_{ab}}; \qquad\qquad \mathcal{E}_2 \;=\; \frac{\mathbf{H}_{aa} - \mathbf{H}_{ab}}{1 - \mathbf{\Delta}_{ab}} \tag{2.55}$$

The first root gives the following set of simulataneous equations for the coefficients in (2.53)

$$\pm\left[\left(\mathbf{H}_{aa}\,\mathbf{\Delta}_{ab}-\mathbf{H}_{ab}\right)c_a+\left(\mathbf{H}_{ab}-\mathbf{\Delta}_{ab}\,\mathbf{H}_{aa}\right)c_b\right]\;=\;0 \tag{2.56}$$

Equations (2.56) are satisfied only if $c_a = c_b$. In order that the eigenfunction (2.53) be normalized, we must have

$$c_a^2+c_b^2+2\,c_a\,c_b\,\mathbf{\Delta}_{ab}\;=\;1$$

so that

$$c_a\;=\;c_b\;=\;\left(2+2\,\mathbf{\Delta}_{ab}\right)^{-1/2} \tag{2.57}$$

Similarly, the second root gives

$$c_a\;=\;-c_b\;=\;\left(2-2\,\mathbf{\Delta}_{ab}\right)^{-1/2} \tag{2.58}$$

The wavefunctions and their associated energies are therefore

$$\psi_1\;=\;\frac{\phi_a+\phi_b}{\left(2+2\,\mathbf{\Delta}_{ab}\right)^{1/2}}\qquad\qquad \mathcal{E}_1\;=\;\frac{\mathbf{H}_{aa}+\mathbf{H}_{ab}}{1+\mathbf{\Delta}_{ab}} \tag{2.59}$$

$$\psi_2\;=\;\frac{\phi_a-\phi_b}{\left(2-2\,\mathbf{\Delta}_{ab}\right)^{1/2}}\qquad\qquad \mathcal{E}_2\;=\;\frac{\mathbf{H}_{aa}-\mathbf{H}_{ab}}{1-\mathbf{\Delta}_{ab}} \tag{2.60}$$

The integrals $\mathbf{\Delta}_{ab}$, $\mathbf{H}_{aa}$, and $\mathbf{H}_{ab}$ may all be evaluated exactly. In atomic units the nonrelativistic Hamiltonian is

$$\mathsf{H}\;=\;-\left(\frac{1}{2}\,\nabla^2+\frac{1}{r_a}+\frac{1}{r_b}-\frac{1}{R}\right) \tag{2.61}$$

where R is the internuclear distance in a.u. Since

$$\mathsf{H}\,\phi_a\;=\;\left(\mathcal{E}_H-\frac{1}{r_b}+\frac{1}{R}\right)\phi_a \tag{2.62}$$

where $\mathcal{E}_H$ is the energy of the ground state of hydrogen, the matrix elements for the energy become

$$\mathbf{H}_{aa}\;=\;\mathcal{E}_H+\frac{1}{R}-J\qquad\qquad J\;=<\phi_a\,|\,\frac{1}{r_b}\,|\,\phi_a> \tag{2.63}$$

$$\mathbf{H}_{ab} = \left(\mathcal{E}_H + \frac{1}{R}\right)\mathbf{\Delta}_{ab} - K \qquad\qquad K = < \phi_a \,|\, \frac{1}{r_a} \,|\, \phi_b > \tag{2.64}$$

so that the energy levels are

$$\mathcal{E}_1 = \mathcal{E}_H + \frac{1}{R} - \frac{J+K}{1+\mathbf{\Delta}_{ab}} \tag{2.65}$$

$$\mathcal{E}_2 = \mathcal{E}_H + \frac{1}{R} - \frac{J-K}{1-\mathbf{\Delta}_{ab}} \tag{2.66}$$

In order to evaluate the integrals involved in (2.63), (2.64) and (2.65), (2.66), it is convenient to transform to elliptical coordinates

$$\mu = \frac{r_a + r_b}{R}; \qquad \nu = \frac{r_a - r_b}{R}; \qquad \varphi$$
$$d\tau = \frac{R^3}{8}\,(\mu^2 - \nu^2)\,d\mu\,d\nu\,d\varphi$$
$$1 \le \mu \le \infty; \qquad -1 \le \nu \le 1; \qquad 0 \le \varphi \le 2\,\pi$$

For the "overlap" integral $\mathbf{\Delta}_{ab}$ we have

$$\begin{aligned}\mathbf{\Delta}_{ab} &= \int \phi_a\,\phi_b\,d\tau = \frac{1}{\pi}\int e^{-(r_a+r_b)}\,d\tau \\ &= \frac{R^3}{8\,\pi}\int_1^\infty e^{-R\mu}\,d\mu \int_{-1}^1 (\mu^2-\nu^2)\,d\nu \int_0^{2\pi} d\varphi \\ &= \frac{R^3}{2}\int_1^\infty \mu^2\,e^{-R\mu}\,d\mu - \frac{R^3}{6}\int_1^\infty e^{-R\mu}\,d\mu\end{aligned} \tag{2.67}$$

The integrals involved are special cases of the general integral

$$\int_1^\infty z^n\,e^{-az}\,dz = \frac{n!\,e^{-a}}{a^{n+1}}\sum_{k=0}^{n}\frac{a^k}{k!} \equiv A_n(a) \tag{2.68}$$

so that $\mathbf{\Delta}_{ab}$ is readily found to be

$$\mathbf{\Delta}_{ab} = e^{-R}\left(1 + R + \frac{R^2}{3}\right) \tag{2.69}$$

The integral J is

$$\begin{aligned}J &= \frac{1}{\pi}\int \frac{e^{-2r_a}}{r_b}\,d\tau = \frac{1}{\pi}\,\frac{2}{R}\int \frac{e^{-R(\mu+\nu)}}{\mu-\nu}\,d\tau \\ &= \frac{R^2}{2}\left[\int_1^\infty \mu\,e^{-R\mu}\,d\mu\int_{-1}^1 e^{-R\nu}\,d\nu + \int_1^\infty e^{-R\mu}\,d\mu\int_{-1}^1 \nu\,e^{-R\nu}\,d\nu\right]\end{aligned} \tag{2.70}$$

The integrals in ν are special cases of the integral

$$\int_{-1}^{1} z^n \, e^{-az} \, dz \; = \; (-1)^{n+1} \, A_n(-a) - A_n(a) \tag{2.71}$$

Inserting the proper values for the integrals gives

$$J \; = \; \frac{1}{R} \, [1 - e^{-2R} \, (1 + R)] \tag{2.72}$$

In the same way, we find for the K integral the expression

$$K \; = \; e^{-R} \, (1 + R) \tag{2.73}$$

For large values of R we see that $\mathbf{\Delta}_{ab} = 0$, $\mathbf{H}_{aa} = \mathcal{E}_H$, $\mathbf{H}_{ab} = 0$, so that $\mathcal{E}_1 = \mathcal{E}_2 = \mathcal{E}_H$, that is, just the energy of a normal hydrogen atom, as of course it should be. For $R = 0$, $\mathbf{\Delta}_{ab} = 1$, $\mathbf{H}_{aa} = \mathcal{E}_H - 1 + 1/R$, $\mathbf{H}_{ab} = \mathbf{H}_{aa}$. Neglecting the nuclear repulsion term $1/R$ for the time being, we see that the electronic energies are $\mathcal{E}'_1 = 3 \, \mathcal{E}_H$, $\mathcal{E}'_2 = 0$. When $R = 0$, the lowest molecular orbital should become the $1s$ atomic orbital of He, with an energy $4\mathcal{E}_H$. Our approximation is therefore in error by an amount $\mathcal{E}_H$ for $R = 0$, although it is correct for large R. The reason for this is clear. For large R the orbital ψ_1 is the correct orbital for a hydrogen atom, but for $R = 0$ it is again a hydrogen orbital, but surrounding a nucleus whose charge is two instead of one. In order to get a good approximation for small R we should take a number of orbitals for each atom.

Referring to (2.65) and (2.66), and for simplicity neglecting Δ_{ab} as compared with unity, we see that the difference in energy between the two states, to this approximation, is just $2 \, K$. Also, to this approximation, state ψ_2 is unstable with respect to a hydrogen atom and a proton by an amount $(1 + R) \, (e^{-R} + e^{-2 \, R})$ while state ψ_1 is stable by an amount $(1 + R) \, (e^{-R} - e^{-2 \, R})$. That this should be so may be seen qualitatively in the following manner. For state ψ_1, the electron density is

$$\rho_1 \; = \; \psi_1^* \, \psi_1 \; = \; \frac{1}{2 + 2 \, \Delta_{ab}} \, (\phi_a^2 + \phi_b^2 + 2 \, \phi_a \, \phi_b) \tag{2.74}$$

While for state ψ_2 it is

$$\rho_2 \; = \; \psi_2^* \, \psi_2 \; = \; \frac{1}{2 - 2 \, \Delta_{ab}} \, (\phi_a^2 + \phi_b^2 - 2 \, \phi_a \, \phi_b) \tag{2.75}$$

At a point midway between the two nuclei, se have

$$\rho_1 = \frac{2}{1+\Delta_{ab}}\,\phi_a^2; \qquad \rho_2 = 0 \tag{2.76}$$

State ψ_1 thus has a much greater accumulation of charge between the two nuclei than state ψ_2; the attraction between this accumulation of charge and the two protons may be considered as producing the stability of state ψ_1.

It is of some interest to look at this problem from the following viewpoint. The wave-functions, including the time-dependent term, are, if we neglect the overlap Δ_{ab} as compared to unity,

$$\Psi_1 = \psi_1\, e^{-i\mathcal{E}_1\hbar/t} = \frac{1}{2^{1/2}}(\phi_a+\phi_b)\, e^{-i\mathcal{E}_1\hbar/t}$$
$$\Psi_2 = \psi_2\, e^{-i\mathcal{E}_2\hbar/t} = \frac{1}{2^{1/2}}\,(\phi_a-\phi_b)\, e^{-i\mathcal{E}_2\hbar/t}$$

Any linear combination of these two solutions will represent some particular distribution of electron density. Let us consider the combination $\Psi = 2^{-1/2}\,(\Psi_1+\Psi_2)$. The electron density corresponding to this state is

$$\rho = \frac{1}{2}\left(\psi_1^2 + \psi_1\,\psi_2\left[e^{i(\mathcal{E}_1-\mathcal{E}_2)t/\hbar} + e^{-i(\mathcal{E}_1-\mathcal{E}_2)t/\hbar}\right] + \psi_2^2\right)$$

For $t = 0$, $\rho = 1/2\,(\psi_1+\psi_2)^2 = \phi_a^2$, so that the electron is on nucleus a. For $t = \pi\,\hbar\,(\mathcal{E}_1-\mathcal{E}_2)^{-1}$,

$$\rho = \frac{1}{2}\,(\psi_1-\psi_2)^2 = \phi_b^2$$

so that the electron is on nucleus b. From this viewpoint (which should not be taken too literally), the electron oscillates between a and b, the frequency of the oscillation being $\nu = (\mathcal{E}_1-\mathcal{E}_2)/h$. We thus have the result $\Delta\mathcal{E} = h\,\nu$, where ν is the frequency of the oscillation between the two states and $\Delta\mathcal{E}$ is the difference in energy between these two states.

In Fig. II.1 the energy levels for these two states are plotted as a function of the internuclear distance R, along with the experimental curve as determined from spectroscopic data. In Fig. II.2 the distribution of charge along the internuclear axis is shown. It is seen that

this approximation gives qualitatively correct results, although quantitatively the treatment is not very satisfactory.

The results can be somewhat improved by taking more complicated zero-order functions.

A simple method of improving the agreement would be to introduce a parameter into ϕ_a and ϕ_b; for example, we might take

$$\phi_a = \left(\frac{\alpha^3}{\pi}\right)^{1/2} e^{-\alpha r_a}; \qquad \phi_b = \left(\frac{\alpha^3}{\pi}\right)^{1/2} e^{-\alpha r_b}$$

and then vary α so that the energy is minimized. We should then get agreement at $R = 0$ as well as at $R = \infty$. The energy may be further improved by including in the secular equation the $2p$ orbitals of the hydrogen atoms. The inclusion of these terms partly takes into account the polarization of the hydrogen atom by the other nucleus.

For H_2^+ the band spectra indicate the values $D_e = 2.791$ e.v. and $R_e = 1.06\mathring{A}$. The simple theory as described above gives $D_e = 1.76$ e.v. and $R_e = 1.32\mathring{A}$. Introduction of the parameter α improves these results to $D_e = 2.25$ e.v. and $R_e = 1.06\mathring{A}$, and inclusion of the $2p$ orbitals gives $D_e = 2.71$ e.v., $R_e = 1.06\mathring{A}$. The value of D_e could, of course, be improved by adding more and more hydrogen orbitals.

III. THE BORN-OPPENHEIMER APPROXIMATION

A. Preliminary Remarks

Since molecules are very complicated arrangements of mutually nuclei and electrons, solution of the Schrödinger equation, even for the simplest case of diatomic molecules, cannot be accomplished without restoring to many approximations. In what follows we shall outline the procedure usually employed in molecular calculations. It is based on the empirical observations of molecular spectroscopy, which show that the total energy of a molecule can be viewed as the sum of several approximately noninteracting parts. Apart from the translational energy, a molecule has internal-energy contributions from electronic, vibrational, and rotational motions. In this section we show how the electronic motions can be approximately separated from the remaining motions.

B. Separation of the Electronic and Nuclear Motions

Changing slightly the notation of section I, the nonrelativistic Hamiltonian operator for a molecule of S nuclei and N electrons is

$$\mathsf{H} = -\frac{1}{2}\sum_{k=1}^{S}\frac{1}{M_k}\nabla_k^2 - \frac{1}{2}\sum_{\mu=1}^{N}\nabla_\mu^2 - \sum_{\mu}^{N}\sum_{k}^{S}\frac{Z_k}{r_{\mu k}} + \sum_{k<l}^{S}\frac{Z_k Z_l}{r_{kl}} + \sum_{\mu<\nu}^{N}\frac{1}{r_{\mu\nu}} \tag{3.1}$$

where atomic units are used and where it is assumed that all nuclear and electronic coordinates have been referred to the center of mass of the system. The Hamiltonian defined by Eq. (3.1) is then referred to as the *internal* Hamiltonian. The exact Schrödinger equation which this Hamiltonian satisfies can be written

$$\mathsf{H}\,\Psi(r;R) = E\,\Psi(r;R) \tag{3.2}$$

where E is the total internal energy of the molecule; that is, E is all the energy but that due to translation. The wavefunction $\Psi(r;R)$ depends upon both the electronic coordinates r and the nuclear coordinates R. In general, each electronic and nuclear coordinate includes both spatial and spin variables.

The first term in the Hamiltonian (3.1) represent the kinetic energy of the nuclei (M_k is the mass in atomic units of the kth nucleus); the second term represents electron kinetic energy; the third term is the electron-nucleus-attraction potential energy; the fourth term is the nuclear mutual-repulsion potential energy; and the last term is the electron mutual-repulsion potential energy. Whereas the entire energy of an atom can be viewed as that of the electrons (relative to the energy of a stationary center of mass), that of a molecule arises from several different types of motions, viz, electronic motions, vibrational motions of the nucleus, and rotational motions of the nucleus. Moreover, all these motions are coupled in a rather subtle fashion, so that solution of the Schrödinger equation (3.2) can be solved approximately, even for fairly large molecules, by making certain assumptions and approximations concerning the magnitude of the coupling among the various types of motion.

According to Born and Oppenheimer, the solutions of the Schrödinger equation (3.2) can be expanded in a power series in the function $M^{-1/4}$ (where M is the average mass of the nuclei). If this function is very much smaller than unity, it turns out that one can approximate solutions to Eq. (3.2) by first solving the Schrödinger equation for a series of fixed nuclear positions, thus obtaining the electronic energy for particular arrangements of nuclei. This electronic energy (which depends on the nuclear positions) can then be used as the potential energy for the wavefunctions involving the nuclei alone. This approximation rests upon the physical picture of massive nuclei moving so slowly relative to the electronic motions that the electrons can be thought of as being in quasi-stationary states during the course of the nuclear vibrations and rotations. This implies that the forces acting between the nuclei can be calculated as gradients of the electronic energies and the electrostatic repulsions between the nuclei themselves. The situation is closely related to the *Ehrenfest adiabatic principle*, which states that if a system is perturbed slowly enough, it remains in definite stationary states. In this case it is the nuclear motions which are perturbing the system.

According to the Born-Oppenheimer approximation, the toal wavefunction of the molecule can be written

$$\Psi(r;R) \;=\; \psi_R(r)\,\phi(R) \tag{3.3}$$

The function $\psi_R(r)$ is called the *electronic wavefunction* and, for *fixed* nuclear positions, depends only upon the quantum states of the electrons. Thus, this function depends *parametrically* upon the nuclear coordinates; otherwise, it is independent of the nuclear quantum states. The function $\phi(R)$ is called the *nuclear wavefunction.* It describes the rotational and vibrational motions in a potential field supplied by the electrons.

It is now convenient to rewrite the internal Hamiltonian (3.1) in the simple symbolic form

$$\mathsf{H} \;=\; \mathsf{T}_R + \mathsf{U} + \mathsf{V} \tag{3.4}$$

where T_R is the kinetic-energy operator of the nuclei, U is a sum of monoelectronic operators each of the form

$$\mathsf{u}_\mu \;=\; -\frac{1}{2}\,\nabla_\mu^2 - \sum_k \frac{Z_k}{r_{\mu k}} \tag{3.5}$$

and V is the potential-energy operator for nuclear repulsions and electronic repulsions. The operator $\mathsf{U}+\mathsf{V}$ is customarily referred to as the *electronic Hamiltonian.* This operator is assumed to satisfy the Schrödinger equation

$$(\mathsf{U}+\mathsf{V})\,\psi_R(r) \;=\; E(R)\,\psi_R(r) \tag{3.6}$$

When the nuclear coordinates R are fixed, $E(R)$ is the energy of the N electrons moving in the field of S fixed nuclei *plus* the mutual-repulsion energy of the S fixed nuclei. $E(R)$ is usually referred to as the *molecular energy in the fixed-nuclei approximation.* The most stable configuration for a given electronic quantum state is defined by those values of R leading to a minimum value in the molecular energy.

The validity of this approximation can be examined by susbtituting the approximation (3.3) into the *exact* Schrödinger equation (3.2). In so doing, it is helpful to note the relationships

$$\nabla_k^2 \psi_R(r)\,\phi(R) \;=\; \psi_R(r)\,\nabla_k^2\,\phi(R) + \phi(R)\,\nabla_k^2\,\psi_R(r) + 2\,\nabla_k\,\psi_R(r).\nabla_k\,\phi(R)$$

$$\nabla_\mu^2 \psi_R(r)\,\phi(R) \;=\; \phi(R)\,\nabla_\mu^2\,\psi_R(r)$$

One then obtains

$$\begin{aligned}
&-\sum_k \frac{1}{M_k}\left[\nabla_k\,\psi_R(r).\nabla_k\,\phi(R) + \frac{1}{2}\,\phi(R)\,\nabla_k^2\,\psi_R(r)\right] \\
&-\sum_k \frac{1}{2\,M_k}\,\psi_R(r)\,\nabla_k^2\,\phi(R) + \sum_\mu \left[\left(-\,\frac{1}{2}\,\nabla_\mu^2 - \sum_k \frac{Z_k}{r_{\mu k}}\right)\psi_R(r)\right]\phi(R) \\
&+\left(\sum_{k<l}\frac{Z_k\,Z_l}{r_{kl}} + \sum_{\mu<\nu}\frac{1}{r_{\mu\nu}}\right)\psi_R(r)\,\phi(R) \;=\; E\,\psi_R(r)\,\phi(R) \qquad (3.7)
\end{aligned}$$

Assuming for the moment that the entire first summation can be ignored, Eq. (3.7) reduces to

$$[(\mathsf{T}_R + \mathsf{U} + \mathsf{V})\,\psi_R(r)]\,\phi(R) \;=\; \mathcal{E}\,\psi_R(r)\,\phi(R) \qquad (3.8)$$

where $\mathcal{E}$ is an approximation to the true energy E. Using Eq. (3.6) and dividing through by $\psi_R(r)$, one obtains

$$[\mathsf{T}_R + E(R)]\,\phi(R) \;=\; \mathcal{E}\,\phi(R) \qquad (3.9)$$

The above equation implies that one can describe the nuclear motions by using an effective Hamiltonian in which the potential energy is that provided by the electrons in the fixed-nucleus approximation. Equation (3.8) is a justification of the separation of nuclear and electronic variables assumed in (3.3) provided one can justify neglect of the first summation in Eq. (3.7). Without spelling out the rigorous details, we state only that the dot-product portion represents the rate of transition between different electronic states, a quantity which is usually small enough to be neglected. The remaining terms are of the form $\nabla_k^2\,\psi_R(r)$ divided by M_k. Since $\nabla_k^2\,\psi_R(r)$ is essentially $\nabla_\mu^2\,\psi_R(r)$, these remaining terms are more than three orders of magnitude smaller than the kinetic energies of single elctrons. These

terms may then be neglected. Thus, the Born-Oppenheimer approximation amounts to replacing the solution of the single equation (3.2) by the successive solutions of Eqs. (3.6) and (3.9).

In certain types of calculations, e.g., those in which interactions between the electronic motions and vibrational motions must be explicitly considered, one may treat the first summation in Eq. (3.7) as a small perturbation.

An interesting sidelight to the Born-Oppenheimer approximation is that the approximate energy $\mathcal{E}$ of Eq. (3.9) is a *lower bound* to the exact energy E of Eq. (3.2) (recall that the variation theorem provides an *upper bound* to the exact energy). In fact, the exact energy satisfies the relationship

$$E < \Psi \mid \Psi >_{r,R} = < \Psi \mid \mathsf{H} \mid \Psi >_{r,R} \tag{3.10}$$

where the subscripts on the inegrals are used as a remainder to integrate over both nuclear and electronic coordinates. Applying the variational principle to Eq. (3.6) and considering *fixed nuclear coordinates* allows one to write

$$< \Psi \mid \mathsf{U} + \mathsf{V} \mid \Psi >_r \geq E(R) < \Psi \mid \Psi >_r \tag{3.11}$$

Equations (3.4) and (3.10) then imply

$$E < \Psi \mid \Psi >_{r,R} \geq < \Psi \mid \mathsf{T}_R + E(R) \mid \Psi >_{r,R} \tag{3.12}$$

Now applying the variational principle to Eq. (3.9) and considering *fixed electronic coordinates* leads to

$$< \Psi \mid \mathsf{T}_R + E(R) \mid \Psi >_R \geq \mathcal{E} < \Psi \mid \Psi >_R \tag{3.13}$$

Therefore, Eq. (3.12) can be written

$$E \geq \mathcal{E} \tag{3.14}$$

C. The Electronic States of the H_2^+ Ion

This one-electron molecule-ion plays a role in the electronic structure of diatomic molecules which is analogous, in some respects, to the role of the hydrogen atom in atomic structure. Just as the nonrelativistic wave equation of the hydrogen atom can be solved exactly, so can the nonrelativistic wave equation of H_2^+, provided the Born-Oppenheimer approximation is assumed. The Hamiltonian operator of H_2^+ in the Born-Oppenheimer approximation is given by Eq. (2.61), where the coordinate system is as shown in Fig. III.1. The azimuthal angle φ is a rotation about the internuclear axis. The variable μ is defined in the interval 1 to ∞, and ν is defined in the interval -1 to 1. The variable μ describes confocal ellipsoids of revolution with the nuclei as foci, and ν describes confocal hyperboloids. Transformation to these elliptical coordinates introduced in Section II.D, the Laplacian operator ∇^2 is found to be

$$\nabla^2 = \frac{4}{R^2\,(\mu^2-\nu^2)} \times \left\{ \frac{\partial}{\partial\mu}\left[(\mu^2-1)\,\frac{\partial}{\partial\mu}\right] + \frac{\partial}{\partial\nu}\left[(1-\nu^2)\,\frac{\partial}{\partial\nu}\right] + \frac{\mu^2-\nu^2}{(\mu^2-1)\,(1-\nu^2)}\,\frac{\partial^2}{\partial\varphi^2}\right\} \tag{3.15}$$

On using the relationships

$$\frac{1}{r_a}+\frac{1}{r_b} = \frac{4\,\mu}{(\mu^2-\nu^2)\,R} \qquad\qquad E_{el} = E - \frac{1}{R} \tag{3.16}$$

The Schrödinger equation for H_2^+ then becomes

$$\left\{ \frac{\partial}{\partial\mu}\left[(\mu^2-1)\,\frac{\partial}{\partial\mu}\right] + \frac{\partial}{\partial\nu}\left[(1-\nu^2)\,\frac{\partial}{\partial\nu}\right] + \left(\frac{1}{\mu^2-1}+\frac{1}{1-\nu^2}\right)\frac{\partial^2}{\partial\varphi^2} + 2\,R^2\left[\frac{E_{el}}{4}\,(\mu^2-\nu^2)+\frac{\mu}{R}\right]\right\}\Psi(\mu,\nu,\varphi) = 0 \tag{3.17}$$

It is seen that on replacing $\Psi(\mu,\nu,\varphi)$ by the product function

$$\Psi(\mu,\nu,\varphi) = M(\mu)\,N(\nu)\,F(\varphi) \tag{3.18}$$

one can effect a separation of the variables to obtain the three ordinary differential equations

$$\left\{\frac{d}{d\mu}\left[(\mu^2-1)\,\frac{d}{d\mu}\right] + \epsilon\,\mu^2 + 2\,R\,\mu - \frac{\lambda^2}{\mu^2-1} + \kappa\right\} M(\mu) = 0 \tag{3.19}$$

$$\left\{ \frac{d}{d\nu}\left[(1-\nu^2)\,\frac{d}{d\nu}\right] + \epsilon\,\nu^2 + \frac{\lambda^2}{1-\nu^2} - \kappa \right\} N(\nu) \;=\; 0 \tag{3.20}$$

$$\left(\frac{d^2}{d\varphi^2} + \lambda^2 \right) F(\varphi) \;=\; 0 \tag{3.21}$$

The parameters λ and κ are separation constants, and the parameter ϵ is defined by

$$\epsilon \;=\; -\frac{1}{2}\,R^2\,E_{el} \tag{3.22}$$

Equations (3.19 - 3.21) have well-behaved solutions only if the parameters λ, κ, and ϵ have certain definite values. The equation in φ is of the same form as the φ portion of the spherical harmonics (in complex form) and is well behaved only if $\lambda = 0, \pm 1, \pm 2, \cdots$. Thus, λ is a quantum number associated with the component of the electronic orbital angular momentum along the axis of nuclear centers. The case of $\lambda = 0$ corresponds to no rotation about the bond axis; for $\lambda \neq 0$, the two possible integral values (differing only in sign) represent the fact that clockwise and counterclockdwise rotations about the bond axis are possible (even though we cannot distinguish the two). The solution of the equations in μ and ν is rather involved and must be carried out by numerical methods. The latter equations involve the energy parameter ϵ, which, in turn, invovles λ^2. Thus the energy depends only upon $|\lambda|$.

Furthermore, for $|\lambda| > 0$ the energy levels must be doubly degenerate. The quantum number λ can be seen to reduce to the atomic magnetic quantum number m_l as the internuclear distance R approaches zero. However, it turns out that $|\lambda|$ plays much the same role in H_2^+ as the azimuthal quantum number l does in a one-electron atom, for example, He^+, to which H_2^+ reduces as $R \to 0$. The various states of H_2^+ are usually symbolized by a notation analogous to the hydrogen-atom spectroscopic notation $1s, 2s, 2p, \cdots$, namely,

$$n|\lambda| = 1s\sigma_g, 2s\sigma_g, 2p\sigma_u, 2p\pi_u, \cdots$$

where nl refers to the one-electron atomic orbital (AO) to which the H_2^+ orbitals [called molecular orbitals (MO's)] reduce as $R \to 0$. The value of $|\lambda|$ is given by $\sigma, \pi, \delta, \phi, \cdots$ when $|\lambda| = 0, 1, 2, 3, \cdots$, respectively. The u and g refer to the inversion symmetry of the MO.

The ground state of H_2^+ ($|\lambda| = 0$) is designated by $1s\sigma_g$, which reduces to $1s$ of He^+ when $R \to 0$. The next state is designated $2p\sigma_u$. The representation of the states of H_2^+ by the $n|\lambda|$ notation is sometimes called the *united-atom correlation* since He^+ is formed by uniting the nuclei of H_2^+ to one nucleus at $R = 0$. In fact, one could regard H_2^+ as a perturbed united atom and obtain the wavefunction by a perturbation calculation.

IV. ELECTRON SPIN AND THE ANTISYMMETRY PRINCIPLE

A. Preliminary Remarks

Let us consider two identical particles which may be part of a many-body system and which, like the electron, possess a spin angular momentum, i.e., a fourth degree of freedom. We label the coordinates of the two particles (relative to some given origin) by the position vectors $\tau_1 = (\mathbf{r}_1, \sigma_1)$ and $\tau_2 = (\mathbf{r}_2, \sigma_2)$, where, in general, $\mathbf{r}_i$ represents the cartesian coordinates point (x_i, y_i, z_i) or the polar coordinates point $(r_i, \theta_i, \varphi_i)$ and σ_i is a spin coordinate. For example, a single electron could have two possible spin coordinates, σ_1 or σ_2 which represent the two possible projections of the spin angular momentum on an arbitrary axis. The probability of finding particle 1 at the point τ_1 and particle 2 at the point τ_2 is given by

$$|\,\psi[\tau_1(1), \tau_2(2)]\,|^2\, d\tau_1\, d\tau_2 \tag{4.1}$$

where ψ is a wavefunction describing the two-particle system. Since the two particles are identical, they must be physically indistinguishable. Therefore, the above probability could just as well be written

$$|\,\psi[\tau_2(1), \tau_1(2)]\,|^2\, d\tau_1\, d\tau_2 \tag{4.2}$$

which is the probability of finding the particle 1 at the point τ_2 and particle 2 at the point τ_1. In (4.1) and (4.2) it is understood that $\tau_i(j)$ is to interpreted as "an electron labeled j is located at a point labeled i". No functional dependence is implied by the notation. It is apparent that the two expressions, Eqs. (4.1) and (4.2) differ only in the exchange of the coordinates of two identical particles. In the absence of magnetic fields the wavefunction ψ can always be chosen as real, so that we can write

$$\psi^2[\tau_1(1), \tau_2(2)] \;=\; \psi^2[\tau_2(1), \tau_1(2)] \tag{4.3}$$

i.e.,

$$\psi[\tau_1(1), \tau_2(2)] \;=\; \epsilon_p\, \psi[\tau_2(1), \tau_1(2)] \qquad\qquad \epsilon_p \;=\; \pm 1 \tag{4.4}$$

Equation (4.3) expresses a basic symmetry law of quantum mechanics; viz, *if a physically measurable property depends on the coordinates (including the spin) of identical particles, the outcome of any measurement of that property must be independent of any attempt to label the particles of the system.* In other words, the measurable property (and hence its operator) must be a symmetric function of the coordinates. The quadratic expression (4.3) can be satisfied in two different ways with respect to the symmetry of the wavefunction ψ under an exchange of the coordinates of two indistinguishable particles. One possibility is to take $\epsilon_p = +1$ in Eq. (4.4), in which case ψ is said to be *symmetric* with respect to an interchange of the coordinates of a pair of identical particles. The remaining possibility is to take $\epsilon_p = -1$, in which case ψ is said to be *antisymmetric* with respect to an interchange of the coordinates of a pair of identical particles. We now postulate that all fundamental particles are described by wavefunctions which belong to one or the other of the above symmetry types and that particles never go from one symmetry type to another. Furthermore, we postulate that *all particles with half-integral spin are described by antisymmetric wavefunctions, and all particles with zero or integral spin are described by symmetric wavefunctions.*

Examples of particles described by antisymmetric wavefunctions are the electron, proton, neutron, and, in general, all particles with odd mass number. Such particles obey Fermi-Dirac statistics and are collectively referred to as *fermions.* Fermions tend to repel each other more than one would expect on the basis of otherwise similar particles. For example, a gas of fermions at a given temperature has a higher internal energy and pressure than an otherwise similar gas of identical particles.

Particles with zero or integral spin obey Bose-Einstein statistics and are called *bosons.* Such particles attract each other more than would expect on the basis of otherwise similar particles, and therefore a gas of bosons at a given temperature has a lower internal energy and pressure than an otherwise similar gas of identical particles. Examples of bosons are the photon (spin= 1), deuteron (spin= 1), and, in general, particles of even mass number. The spin is always zero if the atomic number is also even.

The postulate that electrons must be described by wavefunctions which are antisymmetric with respect to an interchange of the coordinates (including the spin) of a pair of electrons is known as the *Pauli principle* or the *antisymmetry principle*. The postulate is justified a posteriori by the fact that it leads to the exclusion principle, which is in accord with the experimental facts.

B. Spin Operators

We postulate that the spin operators obey commmutation relations of the same general form as the orbital-angular momentum operators, namely

$$[\mathsf{S}_x, \mathsf{S}_y] = i\,\mathsf{S}_z, \qquad [\mathsf{S}_y, \mathsf{S}_z] = i\,\mathsf{S}_x, \qquad [\mathsf{S}_x, \mathsf{S}_z] = i\,\mathsf{S}_y \tag{4.5}$$

where atomic units were used, i.e., $\hbar = 1$. Any operators satisfying commutation relations like the above would have eigenvalues of the form

$$0, \pm 1, \pm 2, \pm 3, \cdots$$

or

$$\pm 1/2, \pm 3/2, \pm 5/2, \cdots$$

This is certainly consistent with the $\pm 1/2$ eigenvalues associated with the spin components. Since there are only two components of the projection of the spin, we know that the spin eigenfunctions must be representable as two-component column matrices, i.e., two-component spinors. If these are required to be orthonormal, one possibility is

$$\mathbf{\Omega}\left(\frac{1}{2}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \mathbf{\Omega}\left(-\frac{1}{2}\right) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.6}$$

If we regard S_z as the operator of the spin component which is quantized, the matrix representation of this operator must satisfy the matrix eigenvalue equations

$$\mathbf{S}_z\,\mathbf{\Omega}\left(\frac{1}{2}\right) = \frac{1}{2}\,\mathbf{\Omega}\left(\frac{1}{2}\right), \qquad \mathbf{S}_z\,\mathbf{\Omega}\left(-\frac{1}{2}\right) = -\frac{1}{2}\,\mathbf{\Omega}\left(-\frac{1}{2}\right) \tag{4.7}$$

where $\mathbf{S}_z$ is the matarix representation of S_z,

$$\mathbf{S}_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{4.8}$$

The remaining matrix representations are

$$\mathbf{S}_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathbf{S}_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{4.9}$$

It is readily verified that each of these matrices is self-adjoint.

The matarix representation of S^2 is most conveniently found by use of the step-up and step-down operaors whose matrix representations are

$$\mathbf{S}_+ = \mathbf{S}_x + i\,\mathbf{S}_y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{S}_- = \mathbf{S}_x - i\,\mathbf{S}_y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{4.10}$$

and by analogy to the orbital angular momentum we write

$$\mathbf{S}^2 = \mathbf{S}_-\mathbf{S}_+ + \mathbf{S}_z\,(\mathbf{S}_z + 1) = \mathbf{S}_+\mathbf{S}_- + \mathbf{S}_z\,(\mathbf{S}_z - 1) = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.11}$$

if H is a Hamiltonian which contains no spin coordinates, e.g., the nonrelativistic or spin-free Hamiltonian, we have the commutation relations

$$[\mathsf{S}_z, \mathsf{H}] = 0 = [\mathsf{S}^2, \mathsf{H}] \tag{4.12}$$

Furthermore, analogous to the orbital angular momentum

$$[\mathsf{S}_z, \mathsf{S}^2] = 0 \tag{4.13}$$

The fact that S_z and S^2 commute with the Hamiltonian means that we can make use of Theorem 5 (see Sect. I.E) in the simplification of matrix elements of H.

The spin eigenfunctions whose matrix representations are given by Eq. (4.6) are conveniently defined as follows

$$\omega_1, \omega_2 = \begin{cases} \alpha = \begin{cases} 1 & spin\ +1/2 \\ 0 & spin\ -1/2 \end{cases} \\ \\ \beta = \begin{cases} 0 & spin\ +1/2 \\ 1 & spin\ -1/2 \end{cases} \end{cases}$$

For example, $\omega_i = \alpha$ ($i = 1$ or 2) means that the probability of finding an electron with spin 1/2 is unity if the electron has that spin and zero otherwise. This means that α is the eigefunction of S_z with the eigenvalue $+1/2$ and that β is the eigenfunction of S_z with the eigenvalue $-1/2$. These eigenfunctions satisfy the orthonormality condition

$$\sum_{m_s} \omega_1^* \, \omega_2 \; = \; \delta_{12} \tag{4.14}$$

where the summation is over the two possible values of m_s. More explicitly, this can be written

$$\sum_{m_s} \omega_1^* \, \omega_2 \; = \; \omega_1^* \, \omega_2 \Big|_{m_s = 1/2} + \omega_1^* \, \omega_2 \Big|_{m_s = -1/2} \tag{4.15}$$

If $\omega_1 = \omega_2 = \alpha$ or $\omega_1 = \omega_2 = \beta$, the above leads to $1 + 0 = 1$ or $0 + 1 = 1$, respectively. For $\omega_1 \neq \omega_2$, one obtains $0 + 0 = 0$. An alternative notation to (4.14) is

$$\int \omega_1^* \, \omega_2 \, d\sigma \; = \; \delta_{12} \tag{4.16}$$

where integration is interpreted as summmation over m_s.

Using the spin function notation, we may summarize the following important relations:

$$\begin{aligned}
\mathsf{S}_x \, \alpha \; &= \; \frac{1}{2} \, \beta & \qquad \mathsf{S}_+ \, \alpha \; &= \; 0 \\
\mathsf{S}_x \, \beta \; &= \; \frac{1}{2} \, \alpha & \qquad \mathsf{S}_+ \, \beta \; &= \; \alpha \\
\mathsf{S}_y \, \alpha \; &= \; \frac{1}{2} \, i \, \beta & \qquad \mathsf{S}_- \, \alpha \; &= \; \beta \\
\mathsf{S}_y \, \beta \; &= \; -\frac{1}{2} \, i \, \alpha & \qquad \mathsf{S}_- \, \beta \; &= \; 0 \\
\mathsf{S}_z \, \alpha \; &= \; \frac{1}{2} \, \alpha & \qquad \mathsf{S}^2 \, \alpha \; &= \; \frac{3}{4} \, \alpha \\
\mathsf{S}_z \, \beta \; &= \; -\frac{1}{2} \, \beta & \qquad \mathsf{S}^2 \, \beta \; &= \; \frac{3}{4} \, \beta
\end{aligned}$$

If ω is a many-electron spin function, we obtain the relationships

$$\begin{aligned}
\mathsf{S}_z \; &= \; \sum_i \mathsf{S}_{zi} \\
\mathsf{S}_z \, \omega \; &= \; M_S \, \omega \\
\mathsf{S}^2 \, \omega \; &= \; S \, (S + 1) \, \omega
\end{aligned}$$

$$-S \leq M_S \leq S$$

$$M_S = \sum_i m_{si}$$

$$S = \sum_i S_i, \sum_i S_i - 1, \cdots \geq 0$$

V. THE INDEPENDENT-PARTICLE MODEL

A. Preliminary Remarks

The independent-particle model has played a central role in electronic-structure calculations. While this model can be motivated by the observation that it becomes exact with the suppression of inter-electron interactions, it is in actuality far more adequate than that observation might suggest, and in addition has the great virtue of ease of visualization and interpretation.

As currently understood, the term *independent-particle model* refers to any formalism in which a many-particle wavefunction is based on a product of single-particle functions. Each single-particle function, often called an *orbital* (or if it is desired to distinguish spatial functions from those including spin a *spinorbital*) is usually determined by methods that take cognizance of the nature of the other single-particle functions to be used, so that the "independence" is formal, rather than physical. Moreover, the energy and other properties associated with the many-particle wavefunction are ordinarily not the sums of orbital contributions, and the requirement of antisymmetry on particle interchange even causes the probability distributions of the various particles to be correlated. However, these departures from a truly independent-particle description do not preclude the use of the intuitively appealing and relatively satisfactory single-particle concepts.

B. The Hartree Product and Slater Determinants

We discuss a system containing N electrons and one or more fixed nuclei, and assume the system to be characterized by an electronic Hamiltonian H of the form

$$\mathsf{H} = \mathsf{U} + \mathsf{V} \tag{5.1}$$

where U and V denote the one-electron and two-electron contributions to H respectively, and may therefore be written in the forms

$$\mathsf{U} = \sum_{\mu} \mathsf{u}(\mu) \tag{5.2}$$

$$\mathsf{V} = \sum_{\mu<\nu} \mathsf{v}(\mu,\nu) \tag{5.3}$$

where μ and ν refer to electrons. Ordinarily $\mathsf{u}(\mu)$ will consist of the kinetic energy operator for Electron μ, plus terms describing its potential energy in the field of the nuclei, i.e., for a system containing in addition to N electrons, S fixed atomic nuclei

$$\mathsf{u}(\mu) = -\frac{1}{2}\nabla_{\mu}^{2} - \sum_{i=1}^{S} \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}_{\mu}|} \tag{5.4}$$

$\mathsf{v}(\mu,\nu)$ in Eq. (5.3) describes the interaction between Electrons μ and ν (i.e., $\mathsf{v}(\mu,\nu) = r_{\mu\nu}^{-1}$). The Hamiltonian expressed in Eqs. (5.1) - (5.3) is in the Born-Oppenheimer approximation, in which the coupling between electronic and nuclear motions is neglected. Except where explicitly indicated otherwise, the potential energy of internuclear interactions will also be omitted.

The Schrödinger equation that this N-electron system satisfies may be written symbolically as

$$\mathsf{H}\,\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) = E\,\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) \tag{5.5}$$

It is found that there is no known coordinate system which will permit one to separate the variables by writing (Hartree product)

$$\Psi(\mathbf{r}_1,\mathbf{r}_2,\cdots,\mathbf{r}_N) = \psi_1(\mathbf{r}_1)\,\psi_2(\mathbf{r}_2),\cdots\psi_N(\mathbf{r}_N) \tag{5.6}$$

This unfortunate result is due to the presence of the coulombic term for the mutual repulsions of the electrons. This fact presents one of the major mathematical difficulties in the solution of Schrödinger's equation for electronic systems more complicated than the hydrogen atom. One possible approach to an approximate solution of the Schrödinger equation is to ignore the electron-electron-repulsion terms in the Hamiltonian. Although one would hardly expect

this to lead to generally good results, nevertheless it suggests certain refinements which are much more satisfactory.

If the electron-repulsion terms in the Hamiltonian are ignored, the Schrödinger equation is separable in the coordinates of individual electrons, and one may apply Eq. (5.6) to obtain N independent equations. In fact, inserting Eq. (5.6) into

$$\mathsf{U}\,\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) \;=\; E\,\Psi(\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N) \tag{5.7}$$

and dividing by Ψ gives

$$\sum_{\mu=1}^{N} \frac{1}{\psi_\mu(\mathbf{r}_\mu)}\,\mathsf{u}(\mu)\,\psi_\mu(\mathbf{r}_\mu) \;=\; E \tag{5.8}$$

Each term in the sum on the left-hand side depends upon the coordinates of one electron only, so that the equation can be separated into the N equations

$$\mathsf{u}(\mu)\,\psi_\mu(\mathbf{r}_\mu) \;=\; \epsilon_\mu\,\psi_\mu(\mathbf{r}_\mu) \qquad\qquad \mu \;=\; 1, 2, \cdots, N \tag{5.9}$$

where ϵ_μ is the energy of the μth electron. The total energy is now given by the simple expression

$$\sum_{\mu=1}^{N} \epsilon_\mu \;=\; E \tag{5.10}$$

Althogh it is a solution of Eq. (5.7) the function Ψ in Eq. (5.6) is not a correct wave-function for two reasons. First, it should include the spin coordinates of the electrons. If spin-orbit interaction is neglected, we may include spin by writing, instead of $\psi_\mu(\mathbf{r})$

$$\phi_\mu(\mathbf{x}) \;=\; \psi_\mu(\mathbf{r})\,\chi_\mu(\xi) \tag{5.11}$$

where ξ is the spin coordinate, $\mathbf{x}$ standing for both $\mathbf{r}$ and ξ, and the spin function χ_μ is either α or β, defined by

$$\begin{aligned} \alpha(1) \;&=\; 1, & \alpha(-1) \;&=\; 0 \\ \beta(1) \;&=\; 0, & \beta(-1) \;&=\; 1 \end{aligned} \tag{5.12}$$

The presence of the spin functions doew not affect Eqs. (5.9) and (5.10).

Secondly, the wavefunction of any system of particles must be either symmetric or antisymmetric in the coordinates of the particles. Particles whose wavefunctions are symmetric are called *bosons*, and those whose wavefunctions are antisymmetric are called *fermions*. Photons, for exaple, are bosons, while electrons are fermions. Instead of a single product like (5.6) we must therefore take an antisymmetric sum of such products with the electronic coordinates permuted in all possible ways. The result, including spin, is a normalized N-electron independent-particle wavefunction, denoted $\mathbf{\Phi}$, assumed to be of the form

$$\mathbf{\Phi} \;=\; (N!)^{1/2}\mathcal{A}\,[\phi_\alpha(\mathbf{x}_1)\,\phi_\beta(\mathbf{x}_2)\cdots\phi_\omega(\mathbf{x}_N)] \tag{5.13}$$

where the ϕ_α are single-particle functions and $\mathcal{A}$ denotes the antisymmetrizer whose definition in terms of permutations P of the coordinates of the N electrons is

$$\mathcal{A} \;=\; \frac{1}{N!}\sum_{\mathsf{P}}(-)^P\mathsf{P} \tag{5.14}$$

where P is a permutation operator upon the electronic coordinates, P is the number of interchanges in P, and the sum is over the $N!$ different permutations. As shown in Sect. I.G $\mathcal{A}$ is a projection operator. This is just a compact way of writing the determinant (Slater determinant)

$$\mathbf{\Phi} \;=\; \frac{1}{(N!)^{1/2}}\,Det\begin{pmatrix} \phi_1(\mathbf{x}_1) & \phi_1(\mathbf{x}_2) & \cdots & \phi_1(\mathbf{x}_N) \\ \phi_2(\mathbf{x}_1) & \phi_2(\mathbf{x}_2) & \cdots & \phi_2(\mathbf{x}_N) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_N(\mathbf{x}_1) & \phi_N(\mathbf{x}_2) & \cdots & \phi_N(\mathbf{x}_N) \end{pmatrix} \tag{5.15}$$

The factor $1/(N!)^{1/2}$ (arising from the $N!$ spinorbital products resulting in the expansion of the determinant) normalizes $\mathbf{\Phi}$, provided the one-electron functions are normalized. In fact, integration with respect to $\mathbf{x}$ implies a sum over the two values ±1 of the spin variable ξ; thus

$$\int|\phi_\mu(\mathbf{x})|^2\,d\mathbf{x} \;\equiv\; \sum_{\xi=\pm1}\int|\psi_\mu(\mathbf{r})|^2\,|\chi_\mu(\xi)|^2\,d\mathbf{r} \;=\; \int|\psi_\mu(\mathbf{r})|^2\,d\mathbf{r} \;=\; 1 \tag{5.16}$$

The functions ψ_μ are, of course, orthogonal and satisfy the completeness relation, since they are the eigenfunctions of (5.9), i.e.,

$$\int\phi^*_\mu(\mathbf{x}\,\phi_\nu(\mathbf{x})\,d\mathbf{x} \;=\; \delta_{\mu\nu} \tag{5.17}$$

$$\sum_{\mu} \phi_{\mu}^{*}(\mathbf{x}')\,\phi_{\mu}(\mathbf{x}) \;=\; \delta(\mathbf{x}-\mathbf{x}') \tag{5.18}$$

where $\delta(\mathbf{x}'-\mathbf{x})$ stands for $\delta(\mathbf{r}-\mathbf{r}')\,\delta_{\xi\xi'}$

The product of the diagonal elements in Eq. (5.15) is just the Hartree product wavefunction (with spin included) (5.6); the remaining $N!-1$ spinorbital products represent all possible permutations of the electron coordinates among the spinorbitals. That this determinantal form is really an antisymmetric function follows from the fact that a determinant is antisymmetric under an exchange of any two rows, an operation equivalent to exchanging the coordinates of a pair of electrons. In the special case that each spinorbital is associated with four quantum numbers (say, n, l, m_l, and m_s), the determinant will vanish if any two electrons have all four quantum numbers the same. This result follows from the fact that two electrons with the same set of quantum numbers are described by the same spinorbital, i.e., two of the columns of the determinant are identical, and thus the determinant vanishes. In general, whether each spinorbital is associated with definite quantum numbers or not, the determinant will vanish whenever two electrons are represented by the same spinorbital. This result is known as the *Pauli exclusion principle.* The vanishing of the wavefunction whenever two electrons occupy the same point of space (coordinate space and spin space) may be thought of as representing a repulsion between electrons of the same spin. However, the Slater determinantal form of the many-electron wavefunction allows for too high a probability of two electrons of different spin occupying the same point of coordinate space.

The Hartree product is an uncorrelated or independent-electron wavefunction because

$$\prod_{\mu} |\,\phi_{\mu}(\mathbf{x}_{\mu})\,|^{2}\, d\mathbf{x}_1 \cdots d\mathbf{x}_N \tag{5.19}$$

which is the simulataneous probability of finding electron-one in the volume element $d\mathbf{x}_1$, centered at $\mathbf{x}_1$, electron-two in $d\mathbf{x}_2$, etc., is just equal to the product of probabilities

$$|\,\phi_1(\mathbf{x}_1)\,|^{2}\, d\mathbf{x}_1\, |\,\phi_2(\mathbf{x}_2)\,|^{2}\, d\mathbf{x}_2 \cdots |\,\phi_N(\mathbf{x}_N)\,|^{2}\, d\mathbf{x}_N \tag{5.20}$$

that electron-one is in $d\mathbf{x}_1$ times the probability that electron-two is in $d\mathbf{x}_2$, etc.

Antisymmetrizing a Hartree product to obtain a Slater determinant introduces *exchange effects*, so-called because they arise from the requirement that $|\mathbf{\Phi}|^2$ be invariant to the exchange of the space and spin coordinates of any two electrons. In particular, a Slater determinant incorporates *exchange correlation*, which means that the motion of two electrons with parallel spins is correlated. Since the motion of electrons with opposite spins remains uncorrelated, it is customary to refer to a single determinantl wavefunction as an uncorrelated wavefunction.

To see how exchange correlation arise, let us investigate the effect of antisymmetrizing a Hartree product on the electron density. Consider a two-electron Slate determinant in which spinorbitals ϕ_1 and ϕ_2 are occupied

$$\mathbf{\Phi}(\mathbf{x}_1, \mathbf{x}_2) \;=\; |\, \phi_1(\mathbf{x}_1)\, \phi_2(\mathbf{x}_2) \;> \tag{5.21}$$

If the two electrons have opposite spins and occupy different spatial orbitals

$$\phi_1(\mathbf{x}_1) \;=\; \psi_1(\mathbf{r}_1)\, \alpha(\omega_1) \tag{5.22}$$

$$\phi_2(\mathbf{x}_2) \;=\; \psi_2(\mathbf{r}_2)\, \beta(\omega_2) \tag{5.23}$$

Then by expanding the determinant, one obtains

$$\begin{aligned} &|\mathbf{\Phi}|^2\, d\mathbf{x}_1\, d\mathbf{x}_2 \\ &= \frac{1}{2}\, |\, \psi_1(\mathbf{r}_1)\, \alpha(\omega_1)\, \psi_2(\mathbf{r}_2)\, \beta(\omega_2) - \psi_1(\mathbf{r}_2)\, \alpha(\omega_2)\, \psi_2(\mathbf{r}_1)\, \beta(\omega_1)\, |^2\, d\mathbf{x}_1\, d\mathbf{x}_2 \end{aligned} \tag{5.24}$$

for the simultaneous probability of electron-one being in $d\mathbf{x}_1$ and electron-two being in $d\mathbf{x}_2$. Let $P(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2$ be the probability of finding electron-one in $d\mathbf{r}_1$ at $\mathbf{r}_1$ and simultaneously electron-two in $d\mathbf{r}_2$ at $\mathbf{r}_2$. This probability is obtained by integrating (averaging) Eq. (5.24) over the spin coordinates of the two electrons, i.e.,

$$\begin{aligned} P(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 \;&=\; \int d\omega_1\, d\omega_2\, |\, \mathbf{\Phi}\, |^2\, d\mathbf{r}_1\, d\mathbf{r}_2 \\ &= \frac{1}{2} \Big[\, |\, \psi_1(\mathbf{r}_1)\, |^2\, |\, \psi_2(\mathbf{r}_2)\, |^2 + |\, \psi_1(\mathbf{r}_2)\, |^2\, |\, \psi_2(\mathbf{r}_1)\, |^2 \Big]\, d\mathbf{r}_1\, d\mathbf{r}_2 \end{aligned} \tag{5.25}$$

The first term in Eq. (5.25) is the product of the probability of finding electron-one in $d\mathbf{r}_1$ at $\mathbf{r}_1$ times the probability of finding electron-two in $d\mathbf{r}_2$ at $\mathbf{r}_2$, if electron-one occupies ψ_1 and electron-two occupies ψ_2. The second term has electron-one occupying ψ_2 and electron-two occupying ψ_1. Since electrons are indistinguishable, the correct probability is the average of the two terms as shown. Thus, the motions of the two electrons is uncorrelated. This is particularly obvious if $\psi_1 = \psi_2$, for in that case

$$P(\mathbf{r}_1, \mathbf{r}_2) \;=\; |\,\psi_1(\mathbf{r}_1)\,|^2\,|\,\psi_1(\mathbf{r}_2)\,|^2 \tag{5.26}$$

Note that $P(\mathbf{r}_1, \mathbf{r}_1) \neq 0$ so that there is a finite probability of finding two electrons with opposite spins at the same point in space.

If the two electrons have the same spin (say β), we have

$$\phi_1(\mathbf{x}_1) \;=\; \psi_1(\mathbf{r}_1)\,\beta(\omega_1) \tag{5.27}$$

$$\phi_2(\mathbf{x}_2) \;=\; \psi_2(\mathbf{r}_2)\,\beta(\omega_2) \tag{5.28}$$

then, by steps identical to the above, we obtain

$$\begin{aligned} P(\mathbf{r}_1, \mathbf{r}_2) \;=\; \frac{1}{2}\Big\{ & |\,\psi_1(\mathbf{r}_1)\,|^2\,|\,\psi_2(\mathbf{r}_2)|^2 + |\,\psi_1(\mathbf{r}_2)\,|^2|\,\psi_2(\mathbf{r}_1)\,|^2 \\ & - \Big[\,\psi_1^*(\mathbf{r}_1)\,\psi_2(\mathbf{r}_1)\,\psi_2^*(\mathbf{r}_2)\,\psi_1(\mathbf{r}_2) + \psi_1(\mathbf{r}_1)\,\psi_2^*(\mathbf{r}_1)\,\psi_2(\mathbf{r}_2)\,\psi_1^*(\mathbf{r}_2)\,\Big]\Big\} \end{aligned} \tag{5.29}$$

where we now have an extra cross term, making the probabilities correlated. This is exchange correlation between electrons of parallel spin. Note that $P(\mathbf{r}_1, \mathbf{r}_1) = 0$, and thus the probability of finding two electrons with parallel spins at the same point in space is zero. A *Fermii hole* is said to exist around an electron. In summary, within the single Slater determinant description, the motion of electrons with parallel spins is correlated but the motion of electrons with opposite spins is not.

For the non-interacting system we have been considering, whose Hamiltoniain is U, the determinantal wavefunction $\mathbf{\Phi}$ gives the same energy E, Eq. (5.10), as the single product function (5.6). The wavefunction of the ground state is just a determinant whose elements

are the one-electron functions corresponding to the lowest energy levels ϵ_μ (two functions with opposite spin factors to each orbital state).

When the interaction term V is included in the Hamiltonian, so that we have the Schrödinger equation (5.5), this equation can no longer be separated, and the wavefunction cannot be expressed as a single determinant of one-electron functions. However, it would seem possible to use the determinantal solutions of Eq. (5.7) as the basis of a perturbation treatment of V - in other words, to take the non-interacting system as the unperturbed system, with V as a perturbation.

Although it is possible to use the one-electron functions ϕ_μ, whose orbital factors are solutions of Eq. (5.9), in constructing the solutions of Eq. (5.5), these may not be the best choice -it may, for example, be better to use the one-electron functions obtained by the Hartree or Hartree-Fock methods. Determinants formed from the latter functions would in general no longer be solutions of Eq. (5.7), but solutions of (5.5) could still be expressed in terms of them.

C. Expectation Values of Operators in a Basis of Determinantal Wavefunctions

If G is any operator which is symmetric in the system coordinates, it will commute with every permutation, thereby also commuting with the sum of such permutations. Therefore

$$[\mathsf{G}, \mathcal{A}] = 0 \tag{5.30}$$

The spin-free Hamiltonian of an N-electron system is just such an operator.

Now consider any quantum-mechanical operator G which commutes with the projection operator $\mathcal{A}$. Using the notation

$$\mathbf{\Phi}^{HP} = [\,\phi_\alpha(\mathbf{x}_1)\,\phi_\beta(\mathbf{x}_2)\cdots\phi_\omega(\mathbf{x}_N)\,] \tag{5.31}$$

for the Hartree product, the expectation value of this operator for the normalized function $\mathbf{\Phi}$ is given by

$$<\mathsf{G}> = <\mathbf{\Phi}\,|\,\mathsf{G}\,|\,\mathbf{\Phi}> = N!\,<\mathcal{A}\,\mathbf{\Phi}^{HP}\,|\,\mathsf{G}\,|\,\mathcal{A}\,\mathbf{\Phi}^{HP}> \tag{5.32}$$

The commutation relation (5.30) and the projection properties of $\mathcal{A}$ allows this expression to be written as

$$< \mathsf{G} >= N! \; < \mathbf{\Phi}^{HP} \,|\, \mathsf{G} \,|\, \mathcal{A}\, \mathbf{\Phi}^{HP} >= \sum_{\mathsf{P}} (-)^P \; < \mathbf{\Phi}^{HP} \,|\, \mathsf{G} \,|\, \mathsf{P}\, \mathbf{\Phi}^{HP} > \tag{5.33}$$

This relationship is very useful in the evaluation of expectation values of oprators with respect to determinantal wavefunctions. It is particularly useful for many-electron wavefunctins for which no factorization into space and spin functions is possible. We shall now illustrate the use of this relationship by obtaining a general expression for the total energy of a system with a closed-shell configuration, i.e., a configuration in which the N electrons may thought of as occupying $N/2$ doubly filled spatial orbitals.

The interaction of the μth electron with the different nuclei is given by the second term on the right-hand side of the monoelectronic operator (5.4). The expectation value of this operator represents the independent-particle energy of the μth electron.

In the following we shall find it convenient to use lowercase Greek letters $\mu, \nu, \cdots$ to represent electron coordinates and lowercase Roman letters $i, j, \cdots$ to designate the orbitals.

We shall assume that the spatial orbitals $\{\psi_i\}$ are orthonormal; i.e., for the μth electron

$$< \psi_i(\mu) \,|\, \psi_j(\mu) \; >= \delta_{ij} \tag{5.34}$$

It then follows at once that the spinorbitals will also be orthonormal, either through the orthonormality of their spatial factors or their spin factors. Equation (5.33) for the Hamiltonian (5.1) and a wavefunction constructed as a single determinant of $N/2$ doubly occupied spatial orbitals becomes

$$< \mathsf{H} >= \sum_{\mathsf{P}} (-)^P \; < \mathbf{\Phi}^{HP} \,|\, \sum_{\mu} \mathsf{u}(\mu) + \sum_{\mu<\nu} \mathsf{v}(\mu,\nu) \,|\, \mathsf{P}\mathbf{\Phi}^{HP} > \tag{5.35}$$

First, let us investigate the monoelectronic portion of $< \mathsf{H} >$. The summation over the identity permutations can be written

$$\sum_{\mu}^{N} < \phi_\mu(\mu) \,|\, \mathsf{u}(\mu) \,|\, \phi_\mu(\mu) > \tag{5.36}$$

Because of the assumed form of the wavefunction, the successive odd- and even- subscripted spinorbitals have the forms

$$\phi_\mu(\mu) \;=\; \psi_i(\mu)\,\alpha(\mu) \qquad\qquad \phi_{\mu+1}(\mu+1) \;=\; \psi_i(\mu+1)\,\beta(\mu+1) \tag{5.37}$$

Thus the summation in (5.36) can be more simply written [after integrating over the spin coordinates by use of Eq. (4.16)]

$$2\sum_{i=1}^{N/2} < \psi_i(\mu) \,|\, \mathsf{u}(\mu) \,|\, \psi_i(\mu) > \tag{5.38}$$

The subscript μ is now superfluous as long as one keeps in mind that the above are one-electrons integrals and that u is a monoelectronic operator. It is convenient to introduce the symbol

$$\epsilon_i^{(0)} \;=< \psi_i \,|\, \mathsf{u} \,|\, \psi_i > \tag{5.39}$$

For all other permutations, the monoelectronic contributions to $< \mathsf{H} >$ vanish because of spin orthogonality. Thus we obtain

$$\sum_\mu < \mathsf{u}(\mu) >= \; 2\sum_{i=1}^{N/2} \epsilon_i^{(0)} \tag{5.40}$$

In the special case that the $\{\psi_i\}$ are eigenfunctions of the monoelectronic operator u, Eq. (5.40) represents the zeroth-order energy of the system, i.e., the energy predicted by the independent-particle model. In general, however, we shall not assume that the $\{\phi_i\}$ are eigenfunctions of u, so that the interpretation of (5.40) as a zeroth-order energy is not strictly true. Nevertheless, it is still convenient to regard (5.40) as something near to the zeroth-order energy of the system. This simplifies interpretation of certain calculations within the format of perturbation theory.

Now let us look at the two-electron portion of $< \mathsf{H} >$. For identity permutations we obtain

$$\sum_{\mu<\nu} < \phi_\mu(\mu)\,\phi_\nu(\nu) \,|\, \mathsf{v}(\mu,\nu) \,|\, \phi_\mu(\mu)\,\phi_\nu(\nu) > \tag{5.41}$$

Now whenever μ is odd and $\nu = \mu + 1$, we have

$$\phi_\mu(\mu) \;=\; \psi_i(\mu)\,\alpha(\mu) \qquad\qquad \phi_\nu(\nu) \;=\; \psi_i(\nu)\,\beta(\nu) \tag{5.42}$$

There are $N/2$ such occurrences, so that this leads to a contribution (after integration over spin)

$$\sum_{i=1}^{N/2} < \psi_i(\mu)\,\psi_i(\nu) \,|\, \mathsf{v}(\mu,\nu) \,|\, \psi_i(\mu)\,\psi_i(\nu) > = \sum_{i=1}^{N/2} J_{ii} \tag{5.43}$$

where J_{ii} is a coulombic integral, defined in general by

$$J_{ij} \;=\; \left\langle \psi_i(\mu)\,\psi_j(\nu) \left| \frac{1}{r_{\mu\nu}} \right| \psi_i(\mu)\psi_j(\nu) \right\rangle \tag{5.44}$$

Such an integral represents the energy due to electrostatic repulsion between a pair of electrons having charge distributions $|\psi_i(\mu)|^2$ and $|\psi_j(\nu)|^2$, respectively. For all other values of $\mu < \nu$ there are four ways one can obtain a given coulombic integral J_{ij} from (5.41). These four ways occur whenever

Identity permutations then lead to a two-electron contribution to $< \mathsf{H} >$ of

$$\sum_{i=1}^{N/2} J_{ii} + 4 \sum_{i<j}^{N/2} J_{ij} \tag{5.45}$$

In the case of two-electron permutations (odd parity) one obtains

$$-\sum_{\mu<\nu} < \phi_\mu(\mu)\,\phi_\nu(\nu) \,|\, \mathsf{v}(\mu,\nu) \,|\, \phi_\nu(\mu)\,\phi_\mu(\nu) > \tag{5.46}$$

When μ is odd and $\nu = \mu + 1$, this leads to zero integrals as a result of spin orthogonality. For other values $\mu < \nu$ one obtains integrals of the form

$$\left\langle \psi_i(\mu)\,\psi_j(\nu) \left| \frac{1}{r_{\mu\nu}} \right| \psi_j(\mu)\,\psi_i(\nu) \right\rangle \;=\; K_{ij} \tag{5.47}$$

known as *exchange integrals*. Unlike the coulombic integrals, the exchange integrals have no simple classical interpretation, since they arise solely as a consequence of the nonclassical antisymmetry principle. Of the four $\mu < \nu$ combinations leading to a particular K_{ij}, only two (those involving only one type of spin function) lead to nonzero integrals for two-electron permutations. Thus, the total contribution due to tw-electron permutations is

$$2 \sum_{i<j}^{N/2} K_{ij} \tag{5.48}$$

It is readily verified that permutations involving more than two electrons lead to no nonzero integrals. This, however, would not be true if we had not assumed orthonormality of the spatial orbitals.

The total energy can now be written as a summation over the $N/2$ spatial orbitals, namely,

$$< \mathsf{H} >= 2 \sum_{i=1}^{N/2} \epsilon_i^{(0)} + \sum_{i<j}^{N/2} (4\, J_{ij} - 2\, K_{ij}) + \sum_{i=1}^{N/2} J_{ii} \tag{5.49}$$

Equation (5.49) can be put into a more convenient form by noting that

$$J_{ij} \;=\; J_{ji} \qquad\qquad K_{ij} \;=\; K_{ji} \qquad\qquad J_{ii} \;\equiv\; K_{ii} \tag{5.50}$$

One can then write

$$\begin{aligned} < \mathsf{H} >&= 2 \sum_{i=1}^{N/2} \epsilon_i^{(0)} + \sum_{i=1}^{N/2} J_{ii} + 2 \sum_{i<j}^{N/2} (2\, J_{ij} - K_{ij}) \\ &= 2 \sum_{i=1}^{N/2} \epsilon_i^{(0)} + \sum_{i,j}^{N/2} (2\, J_{ij} - K_{ij}) \end{aligned} \tag{5.51}$$

We note that the restriction $i < j$ is removed in the last expression in (5.51).

Equation (5.51) is valid only for electronic systems with closed-shell configurations in which the total wavefunction is approximated as a single determinant of doubly occupied spatial orbitals. For other types of configurations and wavefunctions one must use a different expression.

Had one not taken the indistinguishability of electrons into account by requiring the total wavefunction to be antisymmetric, the energy expression (5.51) would not contain the exchange integrals K_{ij}. The inclusion of the exchange integral K_{ij} always lowers the total energy in the case of singlet states described by a determinantal wavefunctions of doubly occupied orbitals. The first summation term in Eq. (5.51) is always negative, since it represents the energy of bound electrons in an independent-particle model. The coulombic integrals represent electrostatic-repulsion energies and are always positive. The exchange integrals, however, have no classical interpretation, and thus one cannot tell at first sight

just what the sign of the exchange integral is. As an example, let us consider the exchange integral given by

$$K_{12} = \left\langle \psi_1(\mu)\,\psi_2(\nu) \left| \frac{1}{r_{\mu\nu}} \right| \psi_2(\mu)\,\psi_1(\nu) \right\rangle \tag{5.52}$$

where ψ_1 and ψ_2 are two orthonormalized functions. We see that if the $1/r_{\mu\nu}$ term is omitted from K_{12}, that is, if $r_{\mu\nu} = 1$, the integral vanishes as a result of the orthogonality of ψ_1 and ψ_2. In general, such an orthogonality arises because the regions of space where $\psi_1^*(\mu)\psi_2(\mu)$ is positive are just canceled by regions of space where it is negative. Now considering both electrons, the integrand will be positive whenever both electrons are in regions of space where $\psi_1^*(\mu)\psi_2(\mu)$ and $\psi_2^*(\nu)\psi_1(\nu)$ are of the same sign. However, if the electrons are in regions of space where $\psi_1^*(\mu)\psi_2(\mu)$ and $\psi_2^*(\nu)\psi_1(\nu)$ are of opposite sign, the integrand will be negative. Since, on the average, $r_{\mu\nu}$ is much smaller in the former case, then, compared to the case $r_{\mu\nu} = 1$, the positive portions of the integral with $r_{\nu\nu} \neq 1$ contribute more and the negative portions contribute less. This means, then, that the exchange integral must always be positive. Since neglect of the antisymmetry principle would lead to an energy (5.51) without the exchange term, we see that antisymmetrization of the closed-shell wavefunction leads to a lowering of the total energy.

It is also possible to show that the exchange integral K_{12} is always smaller than the corresponding coulombic integral J_{12}. In fact, from Eq. (1.126) one has

$$0 \leq K_{ij} \leq J_{ij} \tag{5.53}$$

D. Ground-State Energy of the Helium Atom

Let us calculate the ground-state energy of the helium atom using the independent-particle model. The two electrons will be described by the orbitals (in atomic units)

$$\psi_{1s}(1) = \pi^{-1/2} Z^{3/2} e^{-Zr_1} \qquad \psi_{1s}(2) = \pi^{-1/2} Z^{3/2} e^{-Zr_2} \tag{5.54}$$

The total wavefunction (normalized) is the simple product

$$\psi(1,2) = \psi_{1s}(1)\psi_{1s}(2) = \frac{Z^3}{\pi} e^{-Z(r_1+r_2)} \tag{5.55}$$

Since

$$\epsilon_i = -\frac{Z^2}{2n^2} \quad \text{a.u.} \tag{5.56}$$

and $Z = 2$ for helium, the total electronic energy is

$$E = \epsilon_1 + \epsilon_2 = -\frac{1}{2}Z^2 - \frac{1}{2}Z^2 = -Z^2 = -4.00 \quad \text{a.u.} \tag{5.57}$$

which is just eight times the energy of the hydrogen atom in its ground state. The experimental value of the ground state of the helium atom is -2.905 a.u. The fact that the calculated energy is too low is partially accounted for on the basis of neglecting the energy due to electron repulsion. The latter is obviously a positive quantity, so that its inclusion would raise the total energy.

One might argue that the use of $Z = 2$ cannot be strictly correct since each electron would act as a sort of screen over the nucleus with respect to the other electron. Thus although the electrons are not assumed to interact directly, they may be assumed to interact in the sense of screening the nucleus. This factor may be taken into account by defining the modified nuclear charge

$$\eta = Z - S \tag{5.58}$$

where S is a parameter called the *screening constant.* The total electronic energy of the helium atom is now

$$E = -\eta^2 = -(Z - S)^2 \tag{5.59}$$

If we choose $\eta = 2.905^{1/2} = 1.70$, then $S = 0.30$, and the correct experimental energy is obtained. However, it is easy to show that this device cannot be a proper way of avoiding the explicit introduction of electron repulstion. The ionization potential of the helium atom on the basis of this model is

$$\frac{1}{2}\eta^2 = 1.453 \quad \text{a.u.} = 39.5 \quad \text{ev} \tag{5.60}$$

whereas the experimental value is very close to 0.9 a.u., or 24.5 ev. The fact that the ionization potential turns out to be too high means that the model predicts an electron more

tightly bound than is actually the case; i.e., the amount of electron repulsion is underestimated in spite of the agreement with the total energy. It is evident that the model must be physically incorrect.

One could continue such calculations for heavier atoms, but one would find that even though screening constants could be chosen empirically so as to reproduce the total electronic energy, the ionization potentials would become worse and worse. For the lithium atom with the configuration $1s^2 2s$, oue would need to use two different screening constants, one for the $1s$ electron and another for the $2s$ electron.

The quantitative aspects of the independent-particle model notwithstanding, one might ask how the model fares in a strictly qualitative fashion; e.g., is it adequate for the development of a theory of the periodic table? According to Eq. (5.6) and the Pauli exclusion principle, the order of energy levels should follow the hydrogen-atom order

$$1s < 2s = 2p < 3s = 3p = 3d < 4s = 4p = 4d = 4f \cdots \tag{5.61}$$

Experimentally, one finds that this is not the case. For example, the configuration $1s^2 2s^2$ leads to a lower energy than the configuration $1s^2 2p^2$ does. It turns out that one effect of the electron repulsions is to remove the degeneracies associated with the azimuthal quantum number. The modified order of levels then becomes

$$1s < 2s < 2p < 3s < 3p < 3d < 4s < 4p < 4d < 4f \cdots \tag{5.62}$$

However, there is yet another effect due to electron repulsion which the above does not account for. Experimentally it is found that the $1s^2 2s^2 2p^6 3s^2 3p^6 4s$ configuration leads to a lower total energy than the $1s^2 2s^2 2p^6 3s^2 3p^6 3d$ configuration does, suggesting an apparent reversal of the $3d$ and $4s$ orbital energies. In the following subsection we shall show that this apparent reversal of certain orbital energies is the result of details of the electron repulsion.

E. The Effect of Electron Repulsions on Atomic Energies

Let us correct the independent-particle model of the helium atom in a different way, i.e., by retaining the independent-particle wavefunction (5.55) but using the physically correct Hamiltonian

$$\mathsf{H} = -\frac{1}{2}(\nabla_1^2 + \nabla_2^2) - Z\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{1}{r_{12}} \tag{5.63}$$

The total electronic energy is now approximated by taking the expectation value of the correct Hamiltonian with respect to the wavefunction obtained from the physically incorrect independent-particle model. The total electronic energy is now given by

$$\begin{aligned}
&< \mathsf{H} > = < \psi(1,2) \,|\, \mathsf{H} \,|\, \psi(1,2) > = -\frac{1}{2} < \psi_{1s}(1)\,\psi_{1s}(2) \,|\, \nabla_1^2 + \nabla_2^2 \,|\, \psi_{1s}(1)\,\psi_{1s}(2) > \\
&-Z\left\langle \psi_{1s}(1)\,\psi_{1s}(2) \left| \frac{1}{r_1} + \frac{1}{r_2} \right| \psi_{1s}(1)\,\psi_{1s}(2) \right\rangle \\
&+\left\langle \psi_{1s}(1)\,\psi_{1s}(2) \left| \frac{1}{r_{12}} \right| \psi_{1s}(1)\,\psi_{1s}(2) \right\rangle
\end{aligned} \tag{5.64}$$

The first two integrals represent just the energy of the independent-particle model, $-Z^2$, but we shall show formally just how the integrals are separately evaluated. For the kinetic-energy integral we have

$$\begin{aligned}
&-\frac{1}{2} < \psi_{1s}(1)\,\psi_{1s}(2) \,|\, \nabla_1^2 + \nabla_2^2 \,|\, \psi_{1s}(1)\,\psi_{1s}(2) > = -< \psi_{1s}(1) \,|\, \nabla_1^2 \,|\, \psi_{1s}(1) > \\
&= -< \psi_{1s}(2) \,|\, \nabla_2^2 \,|\, \psi_{1s}(2) > = -\frac{Z^2}{\pi} < e^{-Zr} \,|\, \nabla^2 \,|\, e^{-Zr} > \\
&= -\frac{Z^3}{\pi} \int_0^\infty \int_0^{2\pi} \int_0^\pi e^{-Zr} \left(\frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} e^{-Zr} \right) r^2 \,\sin\theta \, d\theta \, d\phi \, dr \\
&= -4\,Z^3 \int_0^\infty (Z^2\,r^2 - 2\,Z\,r)\, e^{-2Zr}\, dr \;=\; Z^2
\end{aligned} \tag{5.65}$$

In the last step we have used the standard integral

$$\int_0^\infty x^n\, e^{-ax}\, dx \;=\; \frac{n!}{a^{n+1}} \qquad\qquad a > 0 \qquad\qquad n = 0, 1, 2, \cdots \tag{5.66}$$

The nucleus-electron attraction integral is

$$-Z\left\langle \psi_{1s}(1)\,\psi_{1s}(2) \left| \frac{1}{r_1} + \frac{1}{r_2} \right| \psi_{1s}(1)\,\psi_{1s}(2) \right\rangle \;=\; -2\,Z \left\langle \psi_{1s}(1) \left| \frac{1}{r_1} \right| \psi_{1s}(1) \right\rangle$$

$$= -2\,Z\left\langle \psi_{1s}(2)\left|\frac{1}{r_2}\right|\psi_{1s}(2)\right\rangle = -\frac{2\,Z^4}{\pi}\left\langle e^{-Zr}\left|\frac{1}{r}\right|e^{-Zr}\right\rangle$$

$$= -\frac{2\,Z^4}{\pi}\int_0^{\infty}\int_0^{2\pi}\int_0^{\pi} r\,e^{-2Zr}\,\sin\theta\,d\theta\,d\phi\,dr = -8\,Z^4\int_0^{\infty} r\,e^{-2Zr}\,dr = -2\,Z^2 \quad (5.67)$$

Combining the results of Eqs. (5.65) and (5.67), we obtain the independent-particle model energy $Z^2-2Z^2=-Z^2$. The remaining integral in Eq. (5.64) accounts for electron repulsion and is a bit more difficult to evaluate since it contains the interelectronic coordinate r_{12}. Three somewhat different approaches are commonly used to evaluate this electron-repulsion integral. One way is to interpret the integral in terms of the electrostatic repulsion of two charged spheres whose charge distribution is represented by the $1s$ orbitals. A second approach involves expansion of $1/r_{12}$ in terms of associated Legendre polynomials. A third way is perhaps the most useful in general and is the method we shall illustrate here.

The volume element associated with the two-electron repulsion integral in terms of spherical polar coordinates is

$$r_1^2\,\sin\theta_1\,d\theta_1\,d\phi_1\,dr_1\,r_2^2\,\sin\theta_2\,d\theta_2\,d\phi_2\,dr_2 \quad (5.68)$$

Insert on this page Fig. V.1

It is possible to express the coordinates of the second particle relative to those of the first particle in terms of the interelectronic distance r_{12} and the angles χ and ω as shown in Fig. V.1. One then obtains for the volume element

$$r_1^2 \sin\theta_1 \, d\theta_1 \, d\phi_1 \, dr_1 \, r_{12}^2 \sin\omega \, d\omega \, d\chi \, dr_{12} \tag{5.69}$$

Using the law of cosines, we obtain

$$r_2^2 = r_1^2 + r_{12}^2 - 2\, r_1 \, r_{12} \cos\omega \tag{5.70}$$

If r_1 and r_{12} are held fixed, then

$$dr_2 = r_1 \, r_{12} \sin\omega \, d\omega \tag{5.71}$$

Substituting Eq. (5.71) into Eq. (5.69) for $\sin\omega \, d\omega$, the volume element becomes

$$r_1 \sin\theta_1 \, d\theta_1 \, d\phi_1 \, dr_1 \, r_2 \, r_{12} \, d\chi \, dr_2 \, dr_{12} \tag{5.72}$$

The last integral in Eq. (5.64) then becomes

$$\begin{aligned} I &= \left\langle \psi_{1s}(1)\,\psi_{1s}(2) \left| \frac{1}{r_{12}} \right| \psi_{1s}(1)\,\psi_{1s}(2) \right\rangle \\ &= \frac{Z^6}{\pi^2} \int_0^{2\pi} \int_0^{2\pi} \int_0^{\pi} \int_{|r_1 - r_2|}^{r_1 + r_2} \int_0^{\infty} \int_0^{\infty} e^{-2\,Z(r_1 + r_2)} \, r_1 \, r_2 \, dr_1 \, dr_2 \, dr_{12} \sin\theta_1 \, d\theta_1 \, d\phi_1 \, d\chi \end{aligned} \tag{5.73}$$

where we note that r_{12}^{-1} has been eliminated. Integrating over the three angles θ_1, ϕ_1, and χ, we get

$$I = 8\, Z^6 \int_0^{\infty} e^{-2\,Z\,r_1} \, r_1 \, dr_1 \int_0^{\infty} e^{-2\,Z\,r_2} \, r_2 \, dr_2 \int_{|r_1 - r_2|}^{r_1 + r_2} dr_{12} \tag{5.74}$$

Because of the presence of the lower limit $|r_1 - r_2|$ on r_{12}, the integration must be carried out in two parts. When $r_2 > r_1$, the variables r_1, r_2, and r_{12} are defined in the intervals

$$\begin{aligned} &r_1 \le r_2 \le \infty \\ &0 \le r_1 \le \infty \\ &r_2 - r_1 \le r_{12} \le r_1 + r_2 \end{aligned} \tag{5.75}$$

When $r_1 > r_2$, we have

$$r_2 \leq r_1 \leq \infty$$

$$0 \leq r_2 \leq \infty$$

$$r_1 - r_2 \leq r_{12} \leq r_1 + r_2 \tag{5.76}$$

The integrals occurring in each part of the integration are symmetrical in r_1 and r_2 and thus are equal. The integral I then becomes

$$I \;=\; 16\,Z^6 \int_0^\infty e^{-2\,Z\,r_1}\, r_1\, dr_1 \int_{r_1}^\infty e^{-2\,Z\,r_2}\, r_2\, dr_2 \int_{r_2-r_1}^{r_1+r_2} dr_{12} \tag{5.77}$$

Integrating over r_{12} yields $2r_1$. Then integration over r_2 by parts and use of Eq. (5.66) leads to the final result

$$I \;=\; \frac{5}{8}\,Z \tag{5.78}$$

The total energy of the helium atom is the sum of Eqs. (5.65), (5.67), and (5.78), namely,

$$<\mathsf{H}> = \; Z^2 - 2\,Z^2 + \frac{5}{8}\,Z \;=\; -Z^2 + \frac{5}{8}\,Z \;=\; -2.75 \;\;\text{a.u.} \tag{5.79}$$

This is only about 0.15 a.u. higher than the experimental value of -2.905 a.u. In general, one can write the total energy of any atom in the form

$$E \;=\; \sum_{i=1}^{N} \epsilon_i + \sum_{i<j}^{N} \left\langle \frac{1}{r_{ij}} \right\rangle \tag{5.80}$$

where the ϵ_i are obtained from the independent-particle model and the $< r_{ij}^{-1} >$ are computed from the independent-particle wavefunction. This would give unsatisfactory results from a quantitative standpoint but would lead to some qualitative results useful in the understanding of the periodic table. The form of Eq. (5.80) suggests that the apparent reversal of the $3d$ and $4s$ levels is not due to a reversal in the relative energies of the $3d$ and $4s$ orbitals but rather to an increased electron repulsion in the configuration $1s^2 2s^2 2p^6 3s^2 3p^6 3d$ relative to the $1s^2 2s^2 2p^6 3s^2 3p^6 4s$ configuration. Thus if we are to retain our original notion

of putting electrons into orbitals on by one and yet be able to predict relative electronic energies correctly, we must replace the order of orbitals in (5.62) by the empirical order

$$1s < 2s < 2p < 3s < 3p < 4s < 3d < 4p < 5s < 4d < 5p$$
$$< 6s < 4d < 5d < 6p < 7s \tag{5.81}$$

We do not imply that this represents the atual order of orbital energies. The above sequence is merely an empirical order which enables one to predict ground-state configurations of atoms and thus to qualitatively account for the periodic table. Experimentally determined orders of orbital energies may be obtained from x-ray terms values or from optical spectroscopic term values and are found to agree with the order given in Eq. (5.62). As an example, the ground-state configuration of the copper atom is consistent with $1s^2 2s^2 2p^6 2s^2 3p^4 4s 3d^9$, but that of the ion Cu^+ is $1s^2 2s^2 2p^6 3s^2 3p^6 3d^9$; that is, it is the $4s$ electron which is lost in ionization.

Unfortunately (5.81) does not provide a complete answer to the empirically determined configuration order. For a detailed explanation of such deviations one needs to take electron spin into account.

F. Scaling and the Virial Theorem

We shall now show how the virial theorem introduced in Sec. I.C, can be used to improve the treatment of the helium atom discussed in the previous subsection. We shall also show how the virial theorem can be used to analyze such a treatment in terms of errors in the kinetic and potential energies taken separately.

For a system of electrons and nuclei in which the potential energy is of coulombic form, the virial theorem states that the exact wavefunction leads to

$$< \mathsf{T} > = -\frac{1}{2} < \mathsf{V} > \tag{5.82}$$

Introducing the fact that the total energy $< \mathsf{H} >$ is conserved, we may rewrite the above as

$$< \mathsf{H} > = - < \mathsf{T} > = \frac{1}{2} < \mathsf{V} > \tag{5.83}$$

Let us consider a system of q particles (N electrons and p nuclei) having charges $e_1, e_2, \cdots, e_q$; masses $m_1, m_2, \cdots, m_q$; position vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_q$; and a nonrelativistic Hamiltonian operator of the general form ($\hbar = 1$)

$$\mathsf{H} = -\frac{1}{2}\sum_{i=1}^{q}\frac{\nabla_i^2}{m_i} + \sum_{i<j}^{q}\frac{e_i e_j}{r_{ij}} = \mathsf{T} + \mathsf{V} \tag{5.84}$$

Now let a stationary state of the system be described by the Schrödinger

$$\mathsf{H}\,\Psi = E\,\Psi \tag{5.85}$$

where Ψ is the *exact* wavefunction and E is the *exact* energy. We now let $\Phi(\mathbf{v})$, an approximation to Ψ, be given by

$$\Phi(\mathbf{v}) = \Phi(\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_q) \tag{5.86}$$

which leads to an approximate energy ϵ_i such that the virial theorem (5.83) is not necessarily satisfied. We shall now show how we can ensure that such a function always leads to satisfaction of the virial theorem. Let us replace $\Phi(\mathbf{v})$ by a new function in which each position vector $\mathbf{v}_i$ is multiplied by some arbitrary scale factor η. This means that each coordinate x_i, y_i, and z_i in v_i is multiplied by a scale factor. The new function is called a *scaled* wavefunction and may be written

$$\Phi_\eta = \eta^{3/2q}\,\Phi(\eta\mathbf{v}) \tag{5.87}$$

such that for $\eta = 1$ this reduces to Eq. (5.86). The factor $\eta^{3/2q}$ merely ensures that Φ_η is normalized whenever $\Phi(\mathbf{v})$ is normalized. The action of the scale factor may be thought of as a transformation which uniformly stretches all the coordinates from the origin by a factor η.

Now let $< \mathsf{T} >^{(\eta)}$ and $< \mathsf{V} >^{(\eta)}$ represent the expectation values of the kinetic-energy and potential-energy operators, respectively, with respect to the scaled wavefunction (5.87). For the kinetic energy we get

$$< \mathsf{T} >^{(\eta)} = < \Phi_\eta \,|\, \mathsf{T} \,|\, \Phi_\eta > \tag{5.88}$$

where Φ_η is assumed to be normalized. We now introduce the change of notation

$$\eta^3 \, \mathbf{v}_i \; = \; \mathbf{v}'_i \tag{5.89}$$

We note that

$$dv \; = \; dv_1 \, dv_2 \cdots dv_q \; = \; dx_1 \, dy_1 \, dz_1 \cdots dx_q \, dy_q \, dz_q \tag{5.90}$$

so that $\eta x_i = x'_i$, $\eta y_i = y'_i$, and $\eta z_i = z'_i$. Equation (5.88) then becomes

$$\begin{aligned} < \mathsf{T} >^{(\eta)} &= \left\langle \Phi(\mathbf{v}') \, \middle| \, -\frac{1}{2} \, \eta^2 \sum_{i=1}^{q} \frac{(\nabla'_i)^2}{m_i} \, \middle| \, \Phi(\mathbf{v}') \right\rangle \\ &= \eta^2 < \Phi(\mathbf{v}') \, | \, \mathsf{T}' \, | \, \Phi(\mathbf{v}') > = \; \eta^2 \; < \mathsf{T} >^{(1)} \end{aligned} \tag{5.91}$$

This result follows from the fact that all functions and operators in $\mathbf{v}'$ have exactly the same form as those in $\mathbf{v}_i$, that is, $\mathbf{v}'$ is a dummy variable, and the two are therefore equivalent. The quantity $< \mathsf{T} >^{(1)}$ is the expectation value of the kinetic-energy operator with respect to the unscaled wavefunction $\Phi(\mathbf{v})$. Similarly one can show that

$$< \mathsf{V} >^{(\eta)} = \; \eta \; < \mathsf{V} >^{(1)} \tag{5.92}$$

The energy associated with the scaled wavefunction, which we shall denote by ϵ_η is now given in terms of the expectation values of the unscaled function ($\eta = 1$) as

$$< \mathsf{H} > = \; \epsilon_\eta \; = \; \eta^2 \; < \mathsf{T} >^{(1)} + \eta \; < \mathsf{V}^{(1)} \tag{5.93}$$

We now wish to choose the scale function η such that the quantity $|E - \epsilon_\eta|$ is a minimum. Since the exact energy E must be independent of η, this leads to

$$\frac{\partial \epsilon_\eta}{\partial \eta} \; = \; 2 \, \eta \; < \mathsf{T} >^{(1)} + < \mathsf{V} >^{(1)} = \; 0 \tag{5.94}$$

Solving for the scale factor gives

$$\eta \; = - \, \frac{1}{2} \, \frac{< \mathsf{V} >^{(1)}}{< \mathsf{T} >^{(1)}} \tag{5.95}$$

Substituting Eq. (5.95) into (5.93), we see that

$$|E - \epsilon_\eta|_{min} \; = \; \left| E + \frac{[< \mathsf{V} >^{(1)}]^2}{4 \; < \mathsf{T} >^{(1)}} \right| \tag{5.96}$$

If $\Phi(\mathbf{v})$ happens to be the exact wavefunction, then $\eta = 1$, and Eq. (5.95) leads to the virial theorem (5.82). For any scaled wavefunction Φ_η in which η is given by (5.95), substitution of this value of η into Eqs. (5.91) and (5.92) leads to

$$< \mathsf{T} >^{(\eta)} = -\frac{1}{2} < \mathsf{V} >^{(\eta)} \tag{5.97}$$

which is a form of the virial theorem first derived by Fock. This equation states that the virial equation can always be satisfied by a proper choice of scale factor; i.e., we can always stretch the position vectors $\mathbf{v}_i$ by some amount η such that the average value of the kinetic energy is equal to one-half the negative of the average value of the potential energy.

Let us now consider the model in which the independent-particle wavefunction was used along with the correct nonrelativistic Hamiltonian. The average values of the kinetic and potential energies were

$$< \mathsf{T} >^{(Z)} = Z^2 \qquad\qquad < \mathsf{V} >^{(Z)} = -2\,Z^2 + \frac{5}{8}\,Z \tag{5.98}$$

which clearly *do not* satisfy the virial theorem. If we use the scaled wavefuncion

$$\psi_\eta(1,2) \;=\; \frac{\eta^3}{\pi}\, e^{-\eta(r_1+r_2)} \tag{5.99}$$

the kinetic and potential enrgies turn out to be

$$< \mathsf{T} >^{(\eta)} = \eta^2 \qquad\qquad < \mathsf{V} >^{(\eta)} = -2\,\eta\,Z + \frac{5}{8}\,\eta \tag{5.100}$$

The value of η which will lead to satisfaction of the virial theorem is obtained by setting $\eta = 1$ in Eq. (5.100) and using Eq. (5.95), namely,

$$\eta \;=\; -\frac{1}{2}\frac{< \mathsf{V} >^{(1)}}{< \mathsf{T} >^{(1)}} \;=\; \frac{1}{2}\,(4 - \frac{5}{8}) \;=\; \frac{27}{16} \;=\; 1.69 \tag{5.101}$$

Substituting this value of η into Eqs. (5.100) and adding to get the total energy, we find

$$\epsilon_\eta \;=\; -2.85 \text{ a.u.} \tag{5.102}$$

which is an improvement of 0.1 a.u. over the previous result using $Z = \eta = 2$. This represents the closest to the exact energy one can get using a wavefunction of the form of Eq. (5.55).

The value of η may be physically interpreted as representing an effective nuclear charge, just as used in the independent-particle model. In the present case, however, the interpretation is based on the correct Hamiltonian. The ionization potential of the helium atom based on the properly scaled wavefunction is now given by

$$-[-2.85 - (-\frac{1}{2}Z^2)] = 0.85 \text{ a.u.} \tag{5.103}$$

This result is low by only 0.05 a.u.

Let us now define the minimum error resulting from the wavefunction (5.55) by

$$\Delta E = E - \epsilon_\eta = -0.05 \text{ a.u.} \tag{5.104}$$

Using the virial theorem in the form (5.83), we see that the errors in the kinetic and potential energies are

$$\Delta \mathsf{T} = 0.05 \text{ a.u.} \qquad \Delta \mathsf{V} = -0.10 \text{ a.u.} \tag{5.105}$$

This means that the method *underestimates* the kinetic energy by 0.05 a.u. This is due to the fact that we have assumed that the electrons undergo simpler motions than they do in actuality; i.e., the electrons resort to more complicated motions in avoiding each other than is implicit in the wavefunction (5.99). Similarly, the method *overestimates* the potential energy, i.e., the electrons do not repel each other as strongly as the wavefunction (5.99) implies. From the low value of the ionization potential in (5.103), we see that this is indeed the case.

G. Scaling and the Virial Theorem for Diatomic Molecules

We now extend the discussion of the virial theorem and scaling of the previous subsection to diatomic molecules in the Born-Oppenheimer approximation, i.e., for fixed nuclei. If the nuclei are fixed to a distance R, the external force holding the nuclei in this position is given by

$$F = -\,grad\ E(R) = -\frac{dE(R)}{dR} \tag{5.106}$$

where $E(R)$ is the total energy at the internuclear separation R. The virial theorem then becomes

$$2 < \mathsf{T} > = - < \mathsf{V} > - R \frac{dE(R)}{dR} \tag{5.107}$$

where $< \mathsf{T} >$ is the average kinetic energy of the electrons (the nuclei are stationary and have no kinetic energy) and $< \mathsf{V} >$ is the average electrostatic energy.

When $R = R_e$ or ∞, the last term in Eq. (5.107) vanishes, and we obtain the ordinary virial theorem (5.82). We can thus apply the ordinary virial theorem to the variational or perturbation calculation of Sec. II.D. The expectation value of the kinetic energy in the ground state, using the same notation as in Sec. II.D, is

$$\begin{aligned} < \mathsf{T} > &= < \psi_1 \mid -\frac{1}{2}\nabla^2 \mid \psi_1 > = [2(1 + \mathbf{\Delta}_{ab})]^{-1} \; < \phi_a + \phi_b \mid -\frac{1}{2}\nabla^2 \mid \phi_a + \phi_b > \\ &= (1 + \mathbf{\Delta}_{ab})^{-1} [< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_a > + < \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_b >] \end{aligned} \tag{5.108}$$

Using the integral values (in atomic units)

$$< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_a > = \frac{1}{2} \tag{5.109}$$

$$< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_b > = -\frac{1}{2}[\mathbf{\Delta}_{ab} - 2(1 + R)\, e^{-R}] \tag{5.110}$$

and Eq. (2.69) for $\mathbf{\Delta}_{ab}$, one obtains (at $R = 2.50$ a.u.)

$$< \mathsf{T} > = 0.3831 \text{ a.u.} \tag{5.111}$$

Since the total energy at $R = 2.50$ a.u. is -0.5654 a.u., the average potential energy is

$$< \mathsf{V} > = -0.5654 - 0.3831 = -0.9485 \text{ a.u.} \tag{5.112}$$

Clearly, the virial theorem is not satisfied, since

$$\frac{< \mathsf{T} >}{-\frac{1}{2} < \mathsf{V} >} = \frac{0.3831}{0.4743} = 0.8077 \tag{5.113}$$

The calculated dissociation energy D_e can be written

$$D_e = -(\Delta \mathsf{T} + \Delta \mathsf{V}) = 0.0654 \text{ a.u.} \tag{5.114}$$

where $\Delta\mathsf{T}$ and $\Delta\mathsf{V}$ are the changes in kinetic and potential energies, respectively, in going from a hydrogen atom and a bare proton to H_2^+, that is,

$$\Delta\mathsf{T} = <\mathsf{T}> -0.5000 = -0.1169 \text{ a.u.} \tag{5.115}$$

$$\Delta\mathsf{V} = <\mathsf{V}> -(-1.0000) = 0.0515 \text{ a.u.} \tag{5.116}$$

Equations (5.115) and (5.116) imply that the process of forming H_2^+ involves a decrease in the kinetic energy and an increase in the potential energy. However, exact solution led to a total energy of -0.6026 a.u., so that $<\mathsf{T}>= 0.6026$ a.u., and $<\mathsf{V}>= -1.2052$ a.u. Thus, the true values of $\Delta\mathsf{T}$ and $\Delta\mathsf{V}$ are given by

$$\Delta\mathsf{T} = 0.6026 - 0.5000 = 0.1026 \text{ a.u.} \tag{5.117}$$

$$\Delta\mathsf{V} = -1.2052 - (-1.0000) = -0.2052 \text{ a.u.} \tag{5.118}$$

which shows just the opposite behavior to that predicted from the variational or perturbation calculation. We therefore conclude that it is dangerous to interpret the perturbation calculation too literally in terms of the chemical bonding, since the treatment partitions the total energy incorrectly.

We now discuss the role of scaling when not all atoms are in their equilibrium positions. Let us represent the unscaled and scaled wavefunctions of an N- electron diatomic molecule by

$$\phi(\tau, R) = \phi(\tau_1, \tau_2, \cdots, \tau_N, R) \tag{5.119}$$

$$\phi_\eta = \eta^{3N/2}\, \phi(\eta\tau, \rho) \tag{5.120}$$

where $\rho = \eta R$. Thus the internuclear distance R is stretched by a scale factor in the same way as the electron position vectors. Instead of Eqs. (5.91) and (5.92), we now obtain

$$<\mathsf{T}>^{(\eta,R)} = <\phi_\eta \,|\, \mathsf{T} \,|\, \phi_\eta> = \eta^2 \; <\mathsf{T}>^{(1,\rho)} \tag{5.121}$$

$$<\mathsf{V}>^{(\eta,R)} = <\phi_\eta \,|\, \mathsf{V} \,|\, \phi_\eta> = \eta \; <\mathsf{V}>^{(1,\rho)} \tag{5.122}$$

The total energy associated with the scaled wavefunction is

$$E(\eta, R) \;=\; \eta^2 \;<\mathsf{T}>^{(1,\rho)} + \eta \;<\mathsf{V}>^{(1,\rho)} \tag{5.123}$$

Letting R be a fixed parameter and applying the variation principle with η as a variation parameter, we get

$$\begin{aligned} &\frac{\partial E(\eta, R)}{\partial \eta} \\ &= \; 2\,\eta \;<\mathsf{T}>^{(1,\rho)} + <\mathsf{V}>^{(1,\rho)} + \eta^2\, R\, \frac{\partial <\mathsf{T}^{(1,\rho)}}{\partial \rho} + \eta\, R\, \frac{\partial <\mathsf{V}>^{(1,\rho)}}{\partial \rho} \;=\; 0 \end{aligned} \tag{5.124}$$

The last two terms in the above arise because the expectation values depend on η through ρ. The form of these terms arises when we employ the substitution

$$\left(\frac{\partial \rho}{\partial \eta}\right)_R \;=\; R \tag{5.125}$$

When $\phi(\tau, R)$ is the exact solution, then $\eta = 1$, and $\rho = R$, so that Eq. (5.124) reduces to

$$2 \;<\mathsf{T}>^{(1,R)} + <\mathsf{V}>^{(1,R)} + R\, \frac{dE(1,R)}{dR} \;=\; 0 \tag{5.126}$$

which is just the virial theorem (5.107). The last term in Eq. (5.126) arises from Eq. (5.124) when $\eta = 1$ and $\rho = R$, since then $E(1, R)$ is the exact energy and does not depend on η. The last two terms in Eq. (5.124) then become

$$R\, \frac{d}{dR}(<\mathsf{T}>^{(1,R)} + <\mathsf{V}>^{(1,R)} = \; R\, \frac{dE(1,R)}{dR} \tag{5.127}$$

If one uses the relationship (5.123) for $\eta = 1$ and $\rho = R$ along with the virial theorem (5.126), one can obtain the two useful relationships

$$<\mathsf{T}>^{(1,R)} = \; -E(1,R) - R\, \frac{dE(1,R)}{dR} \tag{5.128}$$

$$<\mathsf{V}>^{(1,R)} = \; 2\, E(1,R) + R\, \frac{dE(1,R)}{dR} \tag{5.129}$$

which are analogous to Eq. (5.83). These relationships allow one to calculate potential energies and kinetic energies individually and in a unique manner if the total energy is

known as a function of R. Figure V.2 is a plot of total kinetic and potential energies for a diatomic molecule whose total energy is closely approximated by the Morse function

$$E(R) \;=\; D_e\,(1 - e^{-\beta(R-R_0)})^2 \tag{5.130}$$

Next we investigate the case when $\phi(\tau, R)$ is not the exact wavefunction, so that the virial theorem is not automatically satisfied. We let $\rho = \eta R$, as before, but now we regard ρ as an auxiliary basic parameter in terms of which R and η may be expressed, that is,

$$R \;=\; R(\rho) \qquad\qquad \eta \;=\; \eta(\rho) \tag{5.131}$$

Solving Eq. (5.124) for η, we obtain

$$\eta \;=\; \eta(\rho) \;=\; -\frac{<\mathsf{V}>^{(1,\rho)} + \rho\,\mathsf{V}_\rho}{2\,<\mathsf{T}>^{(1,\rho)} + \rho\,\mathsf{T}_\rho} \tag{5.132}$$

where we introduce the convenient notation

$$\mathsf{V}_\rho \;=\; \frac{\partial <\mathsf{V}>^{(1,\rho)}}{\partial\rho} \qquad\qquad \mathsf{T}_\rho \;=\; \frac{\partial <\mathsf{T}>^{(1,\rho)}}{\partial\rho} \tag{5.133}$$

The internuclear distance R is given in terms of ρ as

$$R \;=\; \rho\,\eta^{-1} \;=\; \rho\left(-\frac{2\,<\mathsf{T}>^{(1,\rho)} + \rho\,\mathsf{T}_\rho}{<\mathsf{V}>^{(1,\rho)} + \rho\,\mathsf{V}_\rho}\right) \tag{5.134}$$

Insert on this page Fig. V.2

From Eq. (5.134) we can obtain the inverse relationship $\rho = \rho(R)$, so that η and R can be related through Eq. (5.132) as well as through $\rho = \eta R$. This inverse relationship is most readily found by using Eq. (5.134) to plot a graph of R versus ρ.

Multiplying Eq. (5.124) by η, we obtain

$$2\,\eta^2 < \mathsf{T} >^{(1,\rho)} + \eta\ < \mathsf{V} >^{(1,\rho)} + R\,(\eta^3\,\mathsf{T}_\rho + \eta^2\,\mathsf{V}_\rho)\ =\ 0 \tag{5.135}$$

Now using Eqs. (5.121) and (5.122), we see that the first two terms in (5.135) are

$$2\ < \mathsf{T} >^{(\eta,R)} + < \mathsf{V} >^{(\eta,R)} \tag{5.136}$$

From Eq. (5.123) we see that the term in parentheses in Eq. (5.135) is

$$\eta \left(\frac{\partial E(\eta,R)}{\partial \rho}\right)_\eta \ = \ \left(\frac{\partial E(\eta,R)}{\partial R}\right)_\eta \tag{5.137}$$

Since $E(\eta, R)$ depends upon η and R, we can write

$$\frac{dE(\eta,R)}{dR} \ = \ \left(\frac{\partial E(\eta,R)}{\partial R}\right)_\eta + \left(\frac{\partial E(\eta,R)}{\partial \eta}\right)_R \frac{d\eta}{dR} \tag{5.138}$$

But since $\eta R = \rho$, the second term in Eq. (5.138) vanishes, and we get

$$\frac{dE(\eta,R)}{dR} \ = \ \left(\frac{\partial E(\eta,R)}{\partial R}\right)_\eta \tag{5.139}$$

Combining Eqs. (5.136) and (5.139), Eq. (5.135) can be rewritten as

$$2\ < \mathsf{T} >^{(\eta,R)} + < \mathsf{V} >^{(\eta,R)} + R\,\frac{dE(\eta,R)}{dR} \ = \ 0 \tag{5.140}$$

which shows that the trial wavefunction can always be made to satisfy the virial theoremm at any internuclear separation R provided η and R are related by Eq. (5.132) and $\rho = \eta R$.

If we are interested only in obtaining the minimum energy (for which $R = R_e$), then from Eqs. (5.135) and (5.139)

$$\frac{dE(\eta,R)}{dR} \ = \ 0 \ = \ \eta^3\,\mathsf{T}_\rho + \eta^2\,\mathsf{V}_\rho \tag{5.141}$$

Using the above result in Eq. (5.124) (after dividing by η) and solving for η, we obtain

$$\eta \ = \ -\frac{< \mathsf{V} >^{(1,\rho)}}{2\ < \mathsf{T} >^{(1,\rho)}} \tag{5.142}$$

which is analogous to Eq. (5.95). The value of η for the minimum in the energy curve (which we shall call η_e) may be obtained from (5.142) by finding the value of $\rho(=\rho_e)$ which minimizes $E(\eta, R)$. From Eqs. (5.121) and (5.122) it follows that the energy can be written as

$$E(\eta, R) = \eta^2 < \mathsf{T} >^{(1,\rho)} + \eta < \mathsf{V} >^{(1,\rho)} \tag{5.143}$$

and from Eq. (5.142) it is obtained

$$E(\eta, R) = -\frac{(< \mathsf{V} >^{(1,\rho)})^2}{4 < \mathsf{T} >^{(1,\rho)}} \tag{5.144}$$

which is analogous to Eq. (5.96). An alternative and equivalent procedure is to obtain ρ_e by minimizing Eq. (5.144). From this ρ_e the minimum energy and η_e can be computed by using Eqs. (5.142) and (5.141), respectively. The internuclear distance at which $E(\eta, R)$ is a minimum is then given by $R_e = \rho_e \eta_e^{-1}$.

As an example of the method just discussed, we consider the simple MO treatment of H_2^+ obtained by scaling the wavefunction. The results for the equilibrium separation can be obtained as follows. Using Eq. (5.142), we obtain the scale factor

$$\eta_e = \frac{0.9485}{2\,(0.3831)} = 1.238 \tag{5.145}$$

where the values given by Eqs. (5.111) and (5.112) have been used, i.e., this value of η_e is valid for $\rho_e = 2.50$ a.u. Thus $R_e = \rho_e/\eta_e = 2.50/1.24 = 2.02$ a.u. The minimum energy is given by Eq. (5.144) as -0.5871 a.u. and leads to a dissociation energy of $[-0.5 - (-0.5871)] = 0.0871$ a.u. (2.37 ev), over 0.6 ev higher than that found from the unscaled function [Eq. (5.114)].

To close this subsection it remains to calculate the kinetic energy operator matrix elements given in Eqs. (5.109) and (5.110). Using the atomic functions given in Eq. (2.50) and the elliptical coordinates introduced in Sec. II.D, we write

$$< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_a > = -\frac{1}{2\,\pi} < e^{-(\mu+\nu)\,R/2} \mid \nabla^2 \mid e^{-(\mu+\nu)\,R/2} > \tag{5.146}$$

that is,

$$< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_a > = -\frac{1}{2\,\pi}\int_1^\infty d\mu \int_{-1}^1 d\nu \int_0^{2\pi} d\varphi\, e^{-(\mu+\nu)\,R/2}\,\nabla^2\, e^{-(\mu+\nu)\,R/2}\, d\tau \quad (5.147)$$

Application of the Laplacian operator given in Eq. (3.15) to the exponential function $e^{-(\mu+\nu)R/2}$ yields

$$\nabla^2\, e^{-(\mu+\nu)\,R/2} = \frac{4}{R^2\,(\mu^2-\nu^2)}\left[\frac{R^2}{4}\,(\mu^2-\nu^2) + R\,(\mu-\nu)\right] e^{-(\mu+\nu)\,R/2} \quad (5.148)$$

Substitution of this expression in Eq. (5.147) yields, after integrating over φ,

$$< \phi_a \mid -\frac{1}{2}\nabla^2 \mid \phi_a > = -\frac{R}{2}\int_{-1}^1 d\nu \int_1^\infty d\mu \left[\frac{R^2}{4}\,(\mu^2-\nu^2) + R\,(\mu-\nu)\right] e^{-(\mu+\nu)\,R} \quad (5.149)$$

The integral over μ, with the help of Eq. (2.68), can be expressed as

$$\begin{aligned}&\int_1^\infty d\mu \left[\frac{R^2}{4}\,(\mu^2-\nu^2) + R\,(\nu-\mu)\right] e^{-\mu\,R} \\ &= \frac{R^2}{4}\,A_2(R) - \frac{R^2}{4}\,\nu^2\,A_0(R) + R\,\nu\,A_0(R) - R\,A_1(R)\end{aligned} \quad (5.150)$$

Inserting this expression in Eq. (5.149), the resulting integral over ν is calculated from Eq. (2.71). After some cancelations it yields

$$\begin{aligned}&\int_{-1}^1 d\nu \left[\frac{R^2}{4}\,A_2(R) - \frac{R^2}{4}\,\nu^2\,A_0(R) + R\,\nu\,A_0(R) - R\,A_1(R)\right] e^{-\nu\,R} \\ &= \frac{R^2}{4}\,(1-\mathsf{C})\,A_0(R)\,A_2(-R) + R\,(1+\mathsf{C})\,A_0(R)\,A_1(-R)\end{aligned} \quad (5.151)$$

where C is an operator that interchanges the sign of the argument in the product functions $A_i(R)A_j(-R)$ $(i \neq j = 0,1,2)$, i.e., $\mathsf{C}A_i(R)A_j(-R) = A_i(-R)A_j(R)$. Using the explicit expressions of these $A_i(R)$ as given in Eq. (2.68) one gets Eq. (5.109). The off diagonal matrix element (5.110) can be obtained similarly.

H. Forces in Molecules: The Generalized Hellmann-Feynman Theorem

The Born-Oppenheimer approximation states that one can view the nuclear motions of a molecule as occurring in a potential field $E(R)$ provided by the electrons. This, then, implies that the forces acting upon the nuclei are expressible as gradients of $E(R)$. In the

following we shall investigate the calculation of such forces from both the exact wavefunction and approximate wavefunctions.

Let $\xi = \xi(\alpha)$ be an arbitrary normalizable wavefunction which may or may not be an exact eigenfunction of the system Hamiltonian $\mathsf{H} = \mathsf{H}(\alpha)$. We assume that H is the electronic Hamiltonian of a molecule in the nonrelativistic and Born-Oppenheimer approximation, and we let α represent some parameter present in H; for example, α may be a nuclear coordinate. The energy asociated with the function ξ is given by

$$\mathcal{E} = \frac{<\xi\,|\,\mathsf{H}\,|\,\xi>}{<\xi\,|\,\xi>} \tag{5.152}$$

Differentiating $\mathcal{E}$ with respect to the parameter α and rearranging, we obtain

$$\frac{d\mathcal{E}}{d\alpha} =<\xi\,|\,\xi>^{-1}\,\left(2\left\langle\frac{\partial\xi}{\partial\alpha}\right|\mathsf{H}-\mathcal{E}\left|\xi\right\rangle+\left\langle\xi\left|\frac{\partial\mathsf{H}}{\partial\alpha}\right|\xi\right\rangle\right) \tag{5.153}$$

where it is assumed that ξ is a real function. Now in the special case that ξ is the exact wavefunction ψ satisfying $\mathsf{H}\psi = E\psi$ and $<\psi|\psi>= 1$, the above reduces to

$$\frac{dE}{d\alpha} = \left\langle\frac{\partial\mathsf{H}}{\partial\alpha}\right\rangle \qquad\qquad \left(\equiv\left\langle\psi\left|\frac{\partial\mathsf{H}}{\partial\alpha}\right|\psi\right\rangle\right) \tag{5.154}$$

This relationship will be referred to as the *generalized Hellmann-Feynman theorem.* The theorem states that the slope of the curve $E(\alpha)$ versus α can be calculated as the expectation value of the operator $\partial\mathsf{H}/\partial\alpha$. The form of the theorem is reminiscent of perturbation theory in that changes in the energy of a system are related to changes in the Hamiltonian.

For a given choice of the parameter α, the physical consequences of the generalized Hellmann-Feynman theorem depend upon which electronic coordinates are held stationary during the differentiation. To illustrate this very important fact we consider the special case of a diatomic molecule AB with an internuclear separation R. We then choose the parameter α as the internuclear distance R.

First, we consider fixed electron coordinates measured along a space-fixed axis, viz, measured along the molecular axis. For simplicity we let the nucleus B serve as the origin of

the electronic coordinates (see Fig.III.1). The μth electron can then be described in terms of $r_{A\mu}$, $\theta_{A\mu}$, and $\varphi_{A\mu}$. Since T is independent of R, we can write

$$\left\langle \frac{\partial \mathsf{H}}{\partial R} \right\rangle = \left\langle \frac{\partial \mathsf{V}}{\partial R} \right\rangle \tag{5.155}$$

The potential-energy operator is

$$\mathsf{V} = -\sum_{\mu} \left(\frac{Z_A}{r_{A\mu}} + \frac{Z_B}{r_{B\mu}} \right) + \sum_{\mu<\nu} \frac{1}{r_{\mu\nu}} + \frac{Z_A Z_B}{R} \tag{5.156}$$

Since $r_{A\mu}$ and $r_{\mu\nu}$ do not depend upon R, we obtain

$$\left\langle \frac{\partial \mathsf{V}}{\partial R} \right\rangle = -Z_B \left(\frac{Z_A}{R^2} + \left\langle \psi \left| \sum_{\mu} \frac{\partial r_{B\mu}^{-1}}{\partial R} \right| \psi \right\rangle \right) = -Z_B \left(\frac{Z_A}{R^2} + \int \gamma \frac{\partial r_B^{-1}}{\partial R} \, dv_1 \right) \tag{5.157}$$

where γ is the first-order density matrix of the system. We can then write

$$\frac{dE}{dR} = -Z_B \left(\frac{Z_A}{R^2} + \int \gamma \frac{\partial r_B^{-1}}{\partial R} \, dv_1 \right) \tag{5.158}$$

Equation (5.158) is usually called the *electrostatic theorem*. The slope dE/dR is just the negative of the z component of the force acting upon the nucleus B (the z direction is chosen to coincide with the internuclear axis). A similar expression can be written down for the z component of the force acting upon nucleus A. The electrostatic theorem states that forces on nuclei can be calculated by classical electrostatics *provided* one describes the electronic distribution by the correct quantum-mechanical distribution function.

Insert on this page Fig. V.3

Next, we consider scaled electronic coordinates, i.e., any coordinates which have the angles $\theta_{A\mu}$ and $\theta_{B\mu}$ invariant under the rotation d/dR (see Fig. V.3). The confocal elliptical coordinates of Sec. II.D are just such coordinates. The use of scaled electronic coordinates is equivalent to using $1/R$ as a scale factor. By analogy to Eqs. (5.121) and (5.122) we now obtain

$$\mathsf{T} = \mathsf{T}(\eta, R) = \frac{1}{R^2}\,\mathsf{T}(1,1) \tag{5.159}$$

$$\mathsf{V} = \mathsf{V}(\eta, R) = \frac{1}{R}\,\mathsf{V}(1,1) \tag{5.160}$$

$$\eta = \frac{1}{R} \tag{5.161}$$

Thus

$$\frac{\partial \mathsf{T}}{\partial R} = -\frac{2}{R^3}\,\mathsf{T}(1,1) = -\frac{2}{R}\,\mathsf{T} \tag{5.162}$$

$$\frac{\partial \mathsf{V}}{\partial R} = -\frac{1}{R^2}\,\mathsf{V}(1,1) = -\frac{1}{R}\,\mathsf{V} \tag{5.163}$$

$$\frac{\partial \mathsf{H}}{\partial R} = -\frac{1}{R}\,(2\,\mathsf{T} + \mathsf{V}) \tag{5.164}$$

The slope dE/dr then becomes

$$\frac{dE}{dR} = -\frac{1}{R}\,(2 < \mathsf{T} > + < \mathsf{V} >) \tag{5.165}$$

which is just the virial theorem (5.107). Equation (5.165) represents another way of calculating the force on a nucleus.

Let us now return to Eq. 5.154) and consider the case when ξ is *not* the exact wavefunction ψ. In general, the generalized Hellmann-Feynman theorem (5.154) no longer follows. Nevertheless, we may ask what it means to say that there exist approximate wavefunctions $\xi = \xi(\alpha)$ which satisfy the generalized Hellmann-Feynman theorem, i.e., which satisfy

$$\frac{d\mathcal{E}}{d\alpha} = \left\langle \frac{\partial \mathsf{H}}{\partial \alpha} \right\rangle \qquad \left(\equiv \left\langle \xi \left| \frac{\partial \mathsf{H}}{\partial \alpha} \right| \xi \right\rangle \right) \tag{5.166}$$

or, alternatively, which satisfy

$$\left\langle \frac{\partial \xi}{\partial \alpha} \middle| \mathsf{H} - \mathcal{E} \middle| \xi \right\rangle = 0 \tag{5.167}$$

The meaning of the statement that a given approximate wavefunction $\xi(\alpha)$ satisfies (5.166) is that the force calculated from the expectation value of $\partial \mathsf{H}/\partial \alpha$ agrees with the force obtained from the slope of the approximate energy $\mathcal{E}(\alpha)$ as a function of α. In general, one will not obtain such an agreement; furthermore, neither the force $d\mathcal{E}/d\alpha$ nor the force $< \partial \mathsf{H}/\partial \alpha >$ will agree with the exact force calculated from the exact wavefunction. One would expect the force obtained from $d\mathcal{E}/d\alpha$ to be closer to the true force $dE/d\alpha$ than the force obtained from $< \xi | \partial \mathsf{H}/\partial \alpha | \xi >$, since the former contains only second-order errors in the energy, whereas the latter *may* contain first-order errors. Nevertheless, it is much easier to evaluate forces from the latter. For this reason it is convenient to investigate the conditions under which Eq. (5.167) is satisfied, so that we can use $< \partial \mathsf{H}/\partial \alpha >$ to compute approximate forces which are equivalent to those obtained from $d\mathcal{E}/d\alpha$. To carry out such an analysis it is convenient to define the functional

$$\mathcal{E}[\alpha, \lambda] = < \xi(\lambda) \,|\, \mathsf{H}(\alpha) \,|\, \xi(\lambda) > \tag{5.168}$$

such that $\mathcal{E}[\alpha, \alpha] = \mathcal{E}$, the approximate energy defined by (5.152). The total differential of $\mathcal{E}[\alpha, \lambda]$ is

$$d\mathcal{E}[\alpha, \lambda] = \frac{\partial \mathcal{E}[\alpha, \lambda]}{\partial \alpha}\, d\alpha + \frac{\partial \mathcal{E}[\alpha, \lambda]}{\partial \lambda}\, d\lambda \tag{5.169}$$

Dividing the above by $d\alpha$ and setting $\lambda = \alpha$, we get

$$\frac{d\mathcal{E}}{d\alpha} = \left. \frac{\partial \mathcal{E}[\alpha, \lambda]}{\partial \alpha} \right|_{\lambda=\alpha} + \left. \frac{\partial \mathcal{E}[\alpha, \lambda]}{\partial \lambda} \right|_{\lambda=\alpha} \tag{5.170}$$

where it is understood that we set $\lambda = \alpha$ *after* carrying out the partial differentiation. We see at once that

$$\left. \frac{\partial \mathcal{E}[\alpha, \lambda]}{\partial \alpha} \right|_{\lambda=\alpha} = \left\langle \xi \middle| \frac{\partial \mathsf{H}}{\partial \alpha} \middle| \xi \right\rangle \tag{5.171}$$

so that ξ satisfies the Hellmann-Feynman theorem if

$$\left.\frac{\partial \mathcal{E}[\alpha,\lambda]}{\partial \lambda}\right|_{\lambda=\alpha} = 0 \tag{5.172}$$

Equation (5.172) states that of all possible approximate wavefunctions $\xi\lambda$, the best (by the variation criterion) is $\xi(\alpha)$. However, of greater interest is the case when the Hellmann-Feynman theorem is satisfied for $\lambda \neq \alpha$. We choose some variational parameter $\lambda = \lambda(\alpha)$ and associate $\mathsf{H}(\alpha)$ with the wavefunction $\xi[\lambda(\alpha)]$ such that

$$\mathcal{E} = \mathcal{E}[\alpha, \lambda(\alpha)] \tag{5.173}$$

Dividing (5.169) by $d\alpha$ but now setting $\lambda = \lambda(\alpha)$, we obtain

$$\frac{d\mathcal{E}}{d\alpha} = \left.\frac{\partial \mathcal{E}[\alpha,\lambda]}{d\alpha}\right|_{\lambda=\lambda(\alpha)} + \left.\frac{\partial \mathcal{E}[\alpha,\lambda]}{\partial \lambda}\right|_{\lambda=\lambda(\alpha)} \frac{d\lambda(\alpha)}{d\alpha} \tag{5.174}$$

Since

$$\left.\frac{\partial \mathsf{H}[\alpha,\lambda]}{\partial \alpha}\right|_{\lambda=\lambda(\alpha)} = \left\langle \xi \left| \frac{\partial \mathsf{H}}{\alpha} \right| \xi \right\rangle \tag{5.175}$$

$$\frac{d\lambda(\alpha)}{d\alpha} \neq 0 \tag{5.176}$$

we see that the Hellmann-Ferynman theorem is satisfied if

$$\left.\frac{\partial \mathsf{H}[\alpha,\lambda]}{\partial \lambda}\right|_{\lambda=\lambda(\alpha)} = 0 \tag{5.177}$$

i.e., an approximate wafefunction ξ will satisfy the Hellmann-Feynman theorem if one includes a variational parameter $\lambda = \lambda(\alpha)$. The specific choice of λ depends upon the coordinates which are held fixed during the variation. In the case of diatomic molecules, if fixed electronic coordinates are used and the wavefunction is constructed from AO's, one must choose a variation parameter $\zeta = \zeta(R)$ which describes a new coordinate origin shifted by the amount ζ from one of the nuclei (see Fig. V.4). AO's defined with reference to such an off-nuclear origin are called *floating orbitals*. Use of the floating parameter ζ as a variational parameter leads to an approximate wavefunction ξ which satisfies the electrostatic theorem (5.158); i.e., the slope of the curve of $\mathcal{E}(R)$ versus R can be calculated from $< \partial \mathsf{V}/\partial R >$.

When scaled coordinates are used, one must choose $\lambda(\alpha)$ as a scale factor $\eta = \eta(R)$. Use of this parameter as a variationa parameter leads to an approximate wavefunction ξ which satisfies the virial theorem, a situation discussed in the previous Subsection. It is important to realize that unless both $\zeta = \zeta(R)$ and $\eta = \eta(R)$ are used as variational parameters, the force calculated by the electrostatic theorem will not necessarily agree with the force calculated from the virial theorem. If the approximate wavefunction is fully optimized, i.e., if the energy is minimized with respect to all possible variational parameters, not only is the Hellmann-Feynman theorem satisfied, but also forces calculated by the electrostatic theorem and the virial theorem agree. Such a situation obtains for exact Hartree-Fock wavefunctions. However, this would no longer be true if the Hartree-Fock calculation were improved (from the point of view of energy) by an incomplete CI treatment, in spite of the fact that such a treatment lowers the energy.

Insert on this page Fig. V.4

VI. SPIN-ADAPTED CONFIGURATIONS

A. Preliminary Considerations

We have described the spin of a single electron by the two spin functions $\alpha(\omega) \equiv \alpha$ and $\beta(\omega) \equiv \beta$. In this Sect. we will discuss spin in more detail and consider the spin states of many-electron systems. We will describe *restricted* Slater determinants that are formed from spinorbitals whose spatial partis are restricted to be the same for α and β spins (i.e., $\{\phi_i\} = \{\psi_i\alpha, \psi_i\beta\}$). Restricted determinants, except in special cases, are not eigenfunctions of the total electron spin operator. However, by taking appropriate linear combinations of such determinants we can form *spin-adapted configurations*, which are proper eigenfunctions. Finally, we will describe *unrestricted* determinants, which are formed from spinorbitals that have different spatial parts for different spins (i.e., $\{\phi_i\} = \{\psi_i^\alpha\alpha, \psi_i^\beta\beta\}$).

In the usual nonrelativistic treatment, such as considered here, the Hamiltonian does not contain any spin coordinates and hence both S^2 and S_z commute with the Hamiltonian

$$[\mathsf{H}, \mathsf{S}^2] \;=\; 0 \;=\; [\mathsf{H}, \mathsf{S}_z] \tag{6.1}$$

Consequently, the exact eigenfunctions of the Hamiltonian are also eigenfunctions of the two spin operators

$$\mathsf{S}^2 \,|\, \mathbf{\Phi} \,> = \, S\,(S+1)\,|\, \mathbf{\Phi} \,> \tag{6.2}$$

$$\mathsf{S}_z \,|\, \mathbf{\Phi} \,> = \, M_S \,|\, \mathbf{\Phi} \,> \tag{6.3}$$

where S and M_S are the spin quantum numbers describing the total spin and its z component of an N-electron $|\mathbf{\Phi}>$. States with $S = 0, 1/2, 1, 3/2, \cdots$ have multiplicity $2S+1 = 1, 2, 3, 4, \cdots$ and are called singlets, doublets, triplets, quartets, etc. Approximate solutions of the Schrödinger equation are not necessarily pure spin states. However, it is often convenient to constrain approximate wavefunctions to be pure singlets, doublets, triplets, etc.

Any single determinant is an eigenfunction of S_z. In particular

$$\begin{aligned} \mathsf{S}_z \,|\, \phi_i \, \phi_j \cdots \phi_k \,> &= \frac{1}{2} \, (N_\alpha - N_\beta) \,|\, \phi_i \, \phi_j \cdots \phi_k \,> \\ &= M_S \,|\, \phi_i \, \phi_j \cdots \phi_k \,> \end{aligned} \tag{6.4}$$

where N_α is the number of spinorbitals with α spin and N_β is the number of spinorbitals with β spin. However, single determinants are not necessarily eigenfunctions of S^2. As we will discuss in the next Subsection, by combining a small number of single determinants it is possible to form spin-adapted configurations that are correct eigenfunctions of S^2.

B. Restricted Determinants and Spin-Adapted Configurations

As we have seen, given a set of K orthonormal spatial orbitals $\{\psi_i \,|\, i = 1, 2, \cdots, K\}$ we can form a set of $2K$ spinorbitals $\{\phi_i \,|\, i = 1, 2, \cdots, 2K\}$ by multiplying each spatial orbital by either the α or β spin function

$$\begin{aligned} \phi_{2i-1}(\mathbf{x}) \;&=\; \psi_i(\mathbf{r}) \, \alpha(\omega) \\ & \qquad\qquad\qquad\qquad\qquad\qquad i \;=\; 1, 2, \cdots, K \\ \phi_{2i}(\mathbf{x}) \;&=\; \psi_i(\mathbf{r}) \, \beta(\omega) \end{aligned}$$

Such spinorbitals are called *restricted* spinorbitals and determinants formed from them are restricted determinants. In such a determinant a given spatial orbital ψ_i can be occupied either by a single electron (spin up or down) or by two electrons (one with up and the other with spin down). It is convenient to classify types of restricted determinants according to the number of spatial orbitals that are singly occupied. A determinant in which each spatial orbital is doubly occupied is called a *closed-shell* determinant (see Fig. VI.1). An *open shell* referes to a spatial orbital that contains a single electron. One refers to determinants by the number of open shells they contain.

$$
\begin{array}{ll}
\vdots & \\
\vdots & \\
\cdots\cdots\cdots & \psi_6 \\
\cdots\cdots\cdots & \psi_5 \\
\cdots\uparrow\downarrow\cdots & \psi_4 \\
\cdots\cdots\cdots & \psi_3 \\
\cdots\uparrow\downarrow\cdots & \psi_2 \\
\cdots\uparrow\downarrow\cdots & \psi_1 \\
|\,^1\mathbf{\Phi}\,> = |\,\phi_1\,\overline{\phi}_1\,\phi_2\,\overline{\phi}_2\,\phi_4\,\overline{\phi}_4\,> &
\end{array}
$$

Fig.VI.1 A singlet closed − shell determinant

All the electron spins are paired in a closed-shell determinant, and it is not surprising that a closed-shell determinant is a pure singlet. That is, it is an eigenfunction of S^2 with eigenvalue zero,

$$\mathsf{S}^2\,|\,\phi_i\,\overline{\phi}_i\,\phi_j\,\overline{\phi}_j\cdots> = 0\,(0+1)\,|\,\phi_i\,\overline{\phi}_i\,\phi_j\,\overline{\phi}_j\cdots> = 0 \tag{6.5}$$

Let us consider the construction of a wavefunction of a two-electron system in which the two spinorbitals in the Hartree product are to be fabricated from two different space orbitals ψ_1 and ψ_2 which are assumed to be separately normalized and mutually orthogonal; for example, in the case of a two-electron atom, ψ_1 and ψ_2 may be two different hydrogenlike atomic orbitals. Since we have two possible spin functions, α and β, we can form the four different spinorbitals $\psi_1\alpha$, $\psi_1\beta$, $\psi_2\alpha$, and $\psi_2\beta$. In general, when one has $2N$ spinorbitals to be used n at a time ($n \leq 2N$), the total number of different combinations is given by

$$\eta = \binom{2\,N}{n} = \frac{(2\,N)!}{n!\,(2\,N-n)!} \tag{6.6}$$

In the two-electron function we have four spinorbitals to be used two at a time, so that we can obtain six different simple products, namely,

$$\begin{aligned}
\mathbf{\Phi}_{12}^{HP\,\alpha\alpha} &= \psi_1(1)\,\alpha(1)\,\psi_2(2)\,\alpha(2) && 1 \\
\mathbf{\Phi}_{12}^{HP\,\beta\beta} &= \psi_1(1)\,\beta(1)\,\psi_2(2)\,\beta(2) && -1 \\
\mathbf{\Phi}_{12}^{HP\,\alpha\beta} &= \psi_1(1)\,\alpha(1)\,\psi_2(2)\,\beta(2) && 0 \\
\mathbf{\Phi}_{12}^{HP\,\beta\alpha} &= \psi_1(1)\,\beta(1)\,\psi_2(2)\,\alpha(2) && 0 \\
\mathbf{\Phi}_{11}^{HP\,\alpha\beta} &= \psi_1(1)\,\alpha(1)\,\psi_1(2)\,\beta(2) && 0 \\
\mathbf{\Phi}_{22}^{HP\,\alpha\beta} &= \psi_2(1)\,\alpha(1)\,\psi_2(2)\,\beta(2) && 0
\end{aligned} \tag{6.7}$$

Products such as $\mathbf{\Phi}_{11}^{HP\,\alpha\alpha}$ are excluded on the basis of the exclusion principle. The numbers to the right of each product function signify the S_z eigenvalues, i.e., M_S. These eigenvalues are readily verified by the direct application of S_z to a particular function using the appropriate relations, for example

$$\begin{aligned}
\mathsf{S}_z\,\mathbf{\Phi}_{12}^{HP\,\alpha\alpha} &= (\mathsf{S}_{z1} + \mathsf{S}_{z2})\,\psi_1(1)\,\alpha(1)\,\psi_2(2)\,\alpha(2) \\
&= \psi_1(1)\,[\,\mathsf{S}_{z1}\,\alpha(1)\,]\,\psi_2(2)\,\alpha(2) + \psi_1(1)\,\alpha(1)\,\psi_2(2)\,[\,\mathsf{S}_{z2}\,\alpha(2)\,] \\
&= \frac{1}{2}\,\psi_1(1)\,\alpha(1)\,\psi_2(2)\,\alpha(2) + \frac{1}{2}\,\psi_1(1)\,\alpha(1)\,\psi_2(2)\,\alpha(2) \;=\; \mathbf{\Phi}_{12}^{HP\,\alpha\alpha}
\end{aligned} \tag{6.8}$$

i.e., the eigenvalue M_S of $\mathbf{\Phi}_{12}^{HP\,\alpha\alpha}$ is 1. In general, the M_S eigenvalue of a product wavefunction (antisymmetrized or not) is given by the simple expression

$$M_S \;=\; \frac{1}{2}\,(N_\alpha - N_\beta) \tag{6.9}$$

The product functions (6.7), as they now stand, are not antisymmetric as required by the Pauli principle. For example, we could just as well represent the first function in (6.7) by

$$\mathbf{\Phi}_{21}^{HP\,\alpha\alpha} \;=\; \psi_2(1)\,\alpha(1)\,\psi_1(2)\,\alpha(2) \tag{6.10}$$

where the coordinates of the two electrons have been interchanged. Now consider the function obtained by taking the difference of $\mathbf{\Phi}_{12}^{HP\,\alpha\alpha}$ and $\mathbf{\Phi}_{21}^{HP\,\alpha\alpha}$, namely,

$$
\begin{aligned}
2^{-1/2}\,(\mathbf{\Phi}_{12}^{HP\,\alpha\alpha} - \mathbf{\Phi}_{21}^{HP\,\alpha\alpha}) \;&=\; 2^{-1/2}\,[\psi_1(1)\,\alpha(1)\,\psi_2(2)\,\alpha(2) - \psi_2(1)\,\alpha(1)\,\psi_1(2)\,\alpha(2)] \\
&=\; 2^{-1/2}\,[\psi_1(1)\,\psi_2(2) - \psi_2(1)\,\psi_1(2)]\,\alpha(1)\,\alpha(2) \\
&=\; -2^{-1/2}\,(\mathbf{\Phi}_{21}^{HP\,\alpha\alpha} - \mathbf{\Phi}_{12}^{HP\,\alpha\alpha}) \qquad (6.11)
\end{aligned}
$$

where $2^{-1/2}$ is the normalization factor. It is evident that the function (6.11) is antisymmetric, as a result, in this case, of the spatial functions. It is readily verified that the antisymmetric function (6.11) is also representable by the determinant

$$
\mathbf{\Phi}_{12}^{\alpha\alpha} \;=\; \frac{1}{2!^{1/2}}\,Det\,\begin{pmatrix} \psi_1(1)\,\alpha(1) & \psi_2(1)\,\alpha(1) \\ \psi_1(2)\,\alpha(2) & \psi_2(2)\,\alpha(2) \end{pmatrix} \qquad (6.12)
$$

in which $\mathbf{\Phi}_{12}^{HP\,\alpha\alpha}$ itself is the product of the diagonal elements and $\mathbf{\Phi}_{21}^{HP\,\alpha\alpha}$ is the product of the remaining elements. It is often convenient to adopt a simplified notation for determinantal wavefunctions and to write (6.12) as

$$
\mathbf{\Phi}_{12}^{\alpha\alpha} \;=\; 2^{-1/2}\,(\psi_1\,\psi_2 - \psi_2\,\psi_1)\,\alpha\,\alpha \qquad (6.13)
$$

where it is understood that electrons 1 and 2 are associated with each product in the natural order 1, 2 from left to right. An even simpler notation, which we shall use quite frequently, is

$$
\mathbf{\Phi}_{12}^{\alpha\alpha} \;=\; |\,\phi_1\,\phi_2\,| \qquad (6.14)
$$

where the vertical bars imply a determinant (including the normalization factor) and ϕ_1 and ϕ_2 imply spinorbitals formed from the spatial orbitals ψ_1 and ψ_2 along with spin functions. For spinorbitals formed with β spin functions, one would write

$$
\mathbf{\Phi}_{12}^{\beta\beta} \;=\; 2^{-1/2}\,(\psi_1\,\psi_2 - \psi_2\,\psi_1)\,\beta\,\beta = |\,\overline{\phi}_1\,\overline{\phi}_2\,| \qquad (6.15)
$$

The horizontal bar over a spatial orbital indicates that a β spin function is to be associated with that spatial orbital in forming the spinorbital.

Exactly the same trick suffices to antisymmetrize each of the remaining product functions of Eq. (6.7). We thus obtain

$$\begin{aligned}
\mathbf{\Phi}_{12}^{\alpha\beta} &= 2^{-1/2}\,(\psi_1\,\psi_2\,\alpha\,\beta - \psi_2\,\psi_1\,\beta\,\alpha) \;=\; |\,\phi_1\,\overline{\phi}_2\,| \\
\mathbf{\Phi}_{12}^{\beta\alpha} &= 2^{-1/2}\,(\psi_1\,\psi_2\,\beta\,\alpha - \psi_2\,\psi_1\,\alpha\,\beta) \;=\; |\,\overline{\phi}_1\,\phi_2\,| \\
\mathbf{\Phi}_{11}^{\alpha\beta} &= 2^{-1/2}\,\psi_1\,\psi_1\,(\alpha\,\beta - \beta\,\alpha) \;=\; |\,\phi_1\,\overline{\phi}_1\,| \\
\mathbf{\Phi}_{22}^{\alpha\beta} &= 2^{-1/2}\,\psi_2\,\psi_2\,(\alpha\,\beta - \beta\,\alpha) \;=\; |\,\phi_2\,\overline{\phi}_2\,|
\end{aligned} \tag{6.16}$$

These determinantal forms arise from the special choice of representing approximate wavefunctions as antisymmetrized products of orbitals.

Examination of an expanded determinantal wavefunction reveals that it is not possible, in general, to speak of an electron as occupying a definite orbital, since the antisymmetrization has the formal effect of distributing each electron over more than one space orbital.

By use of the relations of spin operators we can investigate the behavior of the six antisymmetrized functions given by Eqs. (6.14) to (6.16) with respect to the operator S^2. It is found (after some rather tedious algebra) that all except $|\,\phi_1\,\overline{\phi}_2\,|$ and $|\,\overline{\phi}_1\,\phi_2\,|$ are eigenfunctions of S^2. However, if we form the linear combinations

$$2^{-1/2}\,(|\,\phi_1\,\overline{\phi}_2\,| \pm |\,\overline{\phi}_1\,\phi_2\,|) \;=\; 2^{-1}\,(\psi_1\,\psi_2 \mp \psi_2\,\psi_1)\,(\alpha\,\beta \pm \beta\,\alpha) \tag{6.17}$$

we obtain two different eigenfunctions of S^2. The determinantal wavefunctions formed from the products (6.7) along with their S_z and S^2 eigenvalues are given by

Eigenfunctions	M_S	$S\,(S+1)$
$\|\,\phi_1\,\phi_2\,\|$	1	2
$\|\,\overline{\phi}_1\,\overline{\phi}_2\,\|$	-1	2
$2^{-1/2}\,(\|\,\phi_1\,\overline{\phi}_2\,\| + \|\,\overline{\phi}_1\,\phi_2\,\|)$	0	2
$2^{-1/2}\,(\|\,\phi_1\,\overline{\phi}_2\,\| - \|\,\overline{\phi}_1\,\phi_2\,\|)$	0	0
$\|\,\phi_1\,\overline{\phi}_1\,\|$	0	0
$\|\,\phi_2\,\overline{\phi}_2\,\|$	0	0

In the simple two-electron case considered here, the final wavefunctions all factor into a spatial wavefunction and a spin wavefunction. This behavior does not carry over into wavefunctions involving more than two electrons.

The first three functions have the same spatial function, that is, $\psi_1\,\psi_2 - \psi_2\,\psi_1$, but each has a different spin function. Since each of these has the same value of $S\,(S+1)$, we say that these functions form the three components of a triplet state, i.e., a state which has threefold spin degeneracy. This degeneracy follows from the fact that the energy associated with a wavefunction depends only upon the spatial functions. The remaining functions all have $S\,(S+1)$ values of zero, and since they have different spatial portions, they represent three different nondegenerate states, called *singlet states*. It should be noted that the triplet-state spatial function is antisymmetric with symmetric spin functions and that the singlet-state spatial functions are all symmetric with antisymmetric spin functions.

Either of the last two functions can be used to describe the ground state of a two-electron system. This if ϕ_1 (or ϕ_2) is ϕ_{1s} (a hydrogenlike $1s$ orbital), we obtain the familiar ground-state wavefunction for the helium atom, namely,

$$\xi \;=\; |\,\phi_{1s}\,\overline{\phi}_{1s}\,| \;=\; 2^{-1/2}\,\phi_{1s}(1)\,\phi_{1s}(2)\,(\alpha\,\beta - \beta\,\alpha) \tag{6.18}$$

except that there is now a spin wavefunction present. It is easy to demonstrate that this spin wavefunction is immaterial as far as expectation values of spin-free operators are concerned. If G is a spin-free operator, its expectation value for a system described by the wavefunction (6.18) is

$$\begin{aligned} <\mathsf{G}> &= <\xi\,|\,\mathsf{G}\,|\,\xi> = \int\int |\phi_{1s}\,\overline{\phi}_{1s}\,|\,\mathsf{G}\,|\,\phi_{1s}\,\overline{\phi}_{1s}\,|\,d\mathbf{r}\,d\sigma \\ &= \int \phi_{1s}^*(1)\,\phi_{1s}^*(2)\,\mathsf{G}\,\phi_{1s}(1)\,\phi_{1s}(2)\,d\mathbf{r}\int \frac{|\,\alpha\,\beta - \beta\,\alpha\,|^2}{2}\,d\sigma \end{aligned} \tag{6.19}$$

By use of Eq. (4.16) we see that the right-hand integral becomes

$$\begin{aligned} &\frac{1}{2}\int |\,\alpha\,\beta - \beta\,\alpha\,|^2\,d\sigma \;=\; \frac{1}{2}\left(\int \alpha^*\,\alpha\,d\sigma_1\int \beta^*\,\beta\,d\sigma_2\right. \\ &\left.-2\int \alpha^*\,\beta\,d\sigma_1\int \beta^*\,\alpha\,d\sigma_2 + \int \beta^*\,\beta\,d\sigma_1\int \alpha^*\,\alpha\,d\sigma_2\right) \;=\; \frac{1}{2}\,(1-0+1) \;=\; 1 \end{aligned} \tag{6.20}$$

Thus, the spin functions integrate out to unity, and the expectation value of G is simply

$$< \mathsf{G} > = < \phi_{1s}(1)\,\phi_{1s}(2) \mid \mathsf{G} \mid \phi_{1s}(1)\,\phi_{1s}(2) > \tag{6.21}$$

In the case of systems with more than two electrons, it turns out that the spin functions also integrate to unity even though one cannot factor the total electronic wavefunction into a spatial part and a spin part. In such a case it is necessary to take the spin functions into account in setting up the total wavefunction in order to obtain a spatial wavefunction of the correct symmetry. Once this is done, it is always possible to set up an expression for $< \mathsf{G} >$ which does not involve the spin functions (see Sect. V.B).

C. The Excited States of the Helium Atom

The lowest excited state of the helium atom is represented to zeroth order by the configuration $1s2s$. In Sect. VI.B we saw that this configuration leads to singlet and triplet states. The zeroth-order approximation to the singlet-state wavefunction is

$${}^1\xi \;=\; 2^{-1/2}\,(|1s\,\overline{2s}| - |\overline{1s}\,2s|)$$

where we let $1s$ and $2s$ represent normalized hydrogenlike atomic orbitals. The zeroth-order approximation to the triplet-state wavefunctions is given by one of the functions

$${}^3\xi \;=\; \begin{pmatrix} 2^{-1/2}\,(|1s\,\overline{2s}| + |\overline{1s}\,2s|) \\ |1s\,2s| \\ |\overline{1s}\,\overline{2s}| \end{pmatrix}$$

We recall that the singlet function has the spatial portion $\psi_{1s}\psi_{2s} + \psi_{2s}\psi_{1s}$ with an antisymmetric spin function, whereas the triplet function has the spatial portion $\psi_{1s}\psi_{2s} - \psi_{2s}\psi_{1s}$ with one of three symmetric spin functions. The symmetric and antisymmetric spatial functions may be said to arise as a result of the double degeneracy of the unperturbed state (the independent-particle model); i.e., the two electrons are indistinguishable. The antisymmetrization is irrelevant to the unperturbed state (since the zeroth-order Hamiltonian is a sum of two monoelectronic operators) but is required for the perturbed state since the

complete Hamiltonian contains the two-electron term $1/r_{12}$. The lowest singlet and triplet states are usually designated by $2\,^1S_0$ and $2\,^3S_1$, the number preceding the term symbol representing an effective principal quantum number. This effective principal quantum number arises from the Rydberg series of the atomic spectrum.

The first-order perturbation energies of helium in the $1s2s$ configuration (and in the absence of spin-orbit interaction) is given by the roots of the determinant

$$Det \begin{pmatrix} \mathbf{H}_{11}^{(1)} - \epsilon & \mathbf{H}_{12}^{(1)} \\ \mathbf{H}_{21}^{(1)} & \mathbf{H}_{22}^{(1)} - \epsilon \end{pmatrix} = 0$$

where

$$\mathbf{H}_{11}^{(1)} = \left\langle {}^1\xi \left| \frac{1}{r_{12}} \right| {}^1\xi \right\rangle$$

$$\mathbf{H}_{22}^{(1)} = \left\langle {}^3\xi \left| \frac{1}{r_{12}} \right| {}^3\xi \right\rangle$$

$$\mathbf{H}_{12}^{(1)} = \mathbf{H}_{21}^{(1)} = \left\langle {}^3\xi \left| \frac{1}{r_{12}} \right| {}^1\xi \right\rangle$$

Since S^2 commutes with $1/r_{12}$ and ${}^3\xi$ and ${}^1\xi$ have different S^2 eigenvalues, the off-diagonal matrix elements $\mathbf{H}_{12}$ and $\mathbf{H}_{21}$ vanish. The energies of the singlet and triplet states to first order then are

$${}^1E = \epsilon_{1s}^{(0)} + \epsilon_{2s}^{(0)} + \mathbf{H}_{11}^{(1)} \qquad\qquad {}^3E = \epsilon_{1s}^{(0)} + \epsilon_{2s}^{(0)} + \mathbf{H}_{22}^{(1)}$$

where the zeroth-order energy is given by ($\epsilon_i^{(0)} = -Z^2/2n^2$ in a.u.)

$$\epsilon_{1s}^{(0)} + \epsilon_{2s}^{(0)} = -\frac{4}{2} - \frac{4}{8} = -\frac{5}{2}\ a.u.$$

Since two-electron wavefunctions (in the orbital approximation) factor into space and spin function, the matrix elements $\mathbf{H}_{11}^{(1)}$ and $\mathbf{H}_{22}^{(1)}$ can be evaluated from the spatial portions of ${}^1\xi$ and ${}^3\xi$, respectively; i.e., the spin integrates out to unit. One obtains, with the notation $1s \equiv \phi_{1s} = \psi_{1s}\alpha$, $2s \equiv \phi_{2s} = \psi_{2s}\alpha$, $\overline{1s} \equiv \overline{\phi}_{1s} = \psi_{1s}\beta$, $\overline{2s} \equiv \overline{\phi}_{2s} = \psi_{2s}\beta$,

$$\begin{aligned} \mathbf{H}_{11}^{(1)} &= \frac{1}{2}\left\langle |\phi_{1s}\,\overline{\phi}_{2s}| - |\overline{\phi}_{1s}\,\phi_{2s}| \left| \frac{1}{r_{12}} \right| |\phi_{1s}\,\overline{\phi}_{2s}| - |\overline{\phi}_{1s}\,\phi_{2s}| \right\rangle \\ &= \left\langle \psi_{1s}\,\psi_{2s} \left| \frac{1}{r_{12}} \right| \psi_{1s}\,\psi_{2s} \right\rangle + \left\langle \psi_{1s}\,\psi_{2s} \left| \frac{1}{r_{12}} \right| \psi_{2s}\,\psi_{1s} \right\rangle = J_{12} + K_{12} \end{aligned}$$

$$\mathbf{H}_{22}^{(1)} = \frac{1}{2}\left\langle |\phi_{1s}\,\overline{\phi}_{2s}| + |\overline{\phi}_{1s}\,\phi_{2s}| \left| \frac{1}{r_{12}} \right| |\phi_{1s}\,\overline{\phi}_{2s}| + |\overline{\phi}_{1s}\,\phi_{2s}| \right\rangle$$
$$= \left\langle \psi_{1s}\,\psi_{2s} \left| \frac{1}{r_{12}} \right| \psi_{1s}\,\psi_{2s} \right\rangle - \left\langle \psi_{1s}\,\psi_{2s} \left| \frac{1}{r_{12}} \right| \psi_{2s}\,\psi_{1s} \right\rangle = J_{12} - K_{12}$$

Thus, the energies of the 2^1S_0 and 2^3S_1 states are

$$^1E = -\frac{5}{2} + J_{12} + K_{12} \qquad\qquad ^3E = -\frac{5}{2} + J_{12} - K_{12}$$

The energy difference between the singlet and triplet states is

$$^1E - ^3E = 2\,K_{12}$$

From Sect. we recall that $K_{12} > 0$, so that the triplet state is the lower in energy. This is in accord with Hund's rule of maximum multiplicity.

It is important to note that if one had used a simple product wavefunction $\phi_{1s}\phi_{2s}$, the singlet and triplet states would have the same energy, namely, $-\frac{1}{2} + J_{12}$. It is clear that the antisymmetry principle accounts for the separation of different spin states. *The spin functions, although not affecting the total energy directly, nevertheless influence the total energy by determining the form of the spatial portion of the wavefunction.*

The values of the coulombic and exchange integrals are 0.419 and 0.044 a.u., respectively. The singlet- and triplet-state energies then are

$$^1E = -2.037\ a.u. \qquad (exp.\ -2.147\ a.u.)$$
$$^3E = -2.125\ a.u. \qquad (exp.\ -2.176\ a.u.)$$

The singlet state is in error by 0.110 a.u., and the triplet state is in error by 0.051 a.u. The fact that the triplet-state error is lower is due largely to the fact that the screening effect is not so important for electrons with parallel spins as for electrons with antiparallel spins. This, in turn, is due to the fact that the antisymmetry principle tends to keep electrons with the same spin farther apart. On the other hand, the antisymmetry principle implies that electrons with different spins come closer together than is actually the case.

The tendency for electrons with like spins to avoid each other is often referred to as a *spin correlation* or *exchange correlation* effect. Such an effect is in addition to the radial and angular correlation. In general, the use of single-configuratioanl antisymmetrized wavefunctions accounts for correlation effects in triplet states much better than in singlet states.

D. Construction of Determinantal Eigenfunctions of S^2

As mentioned in the previous Subsection, wavefunctions which are eigenfunctions of S^2 are said to describe pure spin states, i.e., states which are characterized by a definite relative alignment of electron spins. Since one is almost always concerned with atoms and molecules which are in pure spin states, and since determinantal wavefunctions are generally not automatically eigenfunctions of S^2, it is convenient to describe a systematic procedure for constructing wavefunctions for pure spin states.

Consider the eigenvalue equation

$$\mathsf{S}_z\,\omega \;=\; M_S\,\omega \tag{6.22}$$

where ω is an N-electron function containing spin coordinates. For $N = 3$ we that

$$\begin{aligned}
M_S &= m_{s1} + m_{s2} + m_{s3} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2} \\
M_S &= m_{s1} + m_{s2} + m_{s3} = \frac{1}{2} + \frac{1}{2} - \frac{1}{2} = \frac{1}{2} \\
M_S &= m_{s1} + m_{s2} + m_{s3} = \frac{1}{2} - \frac{1}{2} - \frac{1}{2} = -\frac{1}{2} \\
M_S &= m_{s1} + m_{s2} + m_{s3} = -\frac{1}{2} - \frac{1}{2} - \frac{1}{2} = -\frac{3}{2}
\end{aligned}$$

i.e., there are four different possible spin alignments. Below we show that there are $2^3 = 8$ different ways in which these four M_S values can arise. These eight possibilities can be grouped as follows: the group $-\frac{3}{2} \le M_S \le \frac{3}{2}$, which corresponds to $S = \frac{3}{2}$ and thus represents a state with multiplicity $2\,S + 1 = 4$, and two groups $-\frac{1}{2} \le M_S \le \frac{1}{2}$, which corresponds to $S = \frac{1}{2}$ and thus represent states of multiplicity $2\,S + 1 = 2$. This means that

an atom or molecule having three electrons outside a closed shell may exist in a quadruplet state or in one of two different doublet states.

Possible values of the S_z *eigenvalues for a three − electron system*

m_{s1}	m_{s2}	m_{s3}	M_S	$\mathbf{\Phi}_i$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$	$\lvert\, \phi_1\, \phi_2\, \phi_3 \,\rvert$
$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\lvert\, \phi_1\, \phi_2\, \overline{\phi}_3 \,\rvert$
$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\lvert\, \phi_1\, \overline{\phi}_2\, \overline{\phi}_3 \,\rvert$
$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\lvert\, \overline{\phi}_1\, \overline{\phi}_2\, \overline{\phi}_3 \,\rvert$
$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\lvert\, \overline{\phi}_1\, \overline{\phi}_2\, \phi_3 \,\rvert$
$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\lvert\, \overline{\phi}_1\, \phi_2\, \phi_3 \,\rvert$
$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\lvert\, \overline{\phi}_1\, \phi_2\, \overline{\phi}_3 \,\rvert$
$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\lvert\, \phi_1\, \overline{\phi}_2\, \phi_3 \,\rvert$

For $N > 3$ the above procedure is rather tedious to carry out, since there will be 2^N different spin couplings to write down. Thus a systematic procedure is required.

We now develop general methods which enable us to obtain:

(i) the number of independent spin states characterized by the quantum number S which exists for a system of N electrons when the space orbitals occupied by each electron are different. We denote this number $f(N,S)$.

Consider a many-electron spin function which is a simple product of N one-electron spin wavefunctions. Since each one-electron spin wavefunction is restricted to either α or β type, it follows that 2^N different many-electron spin wavefunctions may be generated. Each of these, of necessity, is a eigenfunction of S_z.

For a trhee-electron system, the possible values of S are 3/2 and 1/2. The values of $f(N,S)$ for this system are readily seen to be

$$f(3,3/2) = \frac{3!}{3!\,0!} = 1 \tag{6.23}$$

$$f(3,1/2) = \frac{3!}{2!\,1!} - 1 = 2 \tag{6.24}$$

In Eq. (6.23) we have in essence calculated the number of possible spin orientations which yield $M_S = 3/2$; the state $S = 3/2$ also contains three other spin components with $M_S = 1/2$, $-1/2$, and $-3/2$. Thus, in Eq. (6.24) where we calculate the number of possible spin orientations with $M_S = 1/2$, before we associate the number 3 with the number of independent spin states $S = 1/2$ we must substract out that one spin orientation with $M_S = 1/2$ belonging to $S = 3/2$ - hence Eq. (6.24). Thus, the three-electron system possesses one quartet spin wavefunction and two doublet spin wavefunctions. The total number of wavefunctions is $8 = 2^3 = 1 \times 4 + 2 \times 2$.

For a four-electron system, the values of S are 2, 1, and 0. The values of $f(N,S)$ are

$$f(4,2) = \frac{4!}{4!\,0!} = 1 \qquad \text{1 quintet} \tag{6.25}$$

$$f(4,1) = \frac{4!}{3!\,1!} - 1 = 3 \qquad \text{3 triplets} \tag{6.26}$$

$$f(4,0) = \frac{4!}{2!\,2!} - 3 - 1 = 2 \qquad \text{2 singlets} \tag{6.27}$$

The total number of wavefunctions is $16 = 2^4 = 1 \times 5 + 3 \times 3 + 2 \times 1$.

The results for the three- and four-electrons systems are now generalized in analogical fashion to yield

$$f(N,S) = \frac{N!}{(N/2-S)!\,(N/2+S)!} - \frac{N!}{(N/2-S-1)!\,(N/2+S+1)!}$$
$$= \frac{(2\,S+1)\,N!}{(N/2+S+1)!\,(N/2-S)!} \tag{6.28}$$

Multiplication of this expression by $2\,S+1$ and performing a summation over S gives the total number of wavefunctions

$$\sum_S (2\,S+1)\,f(N,S) = 2^N \tag{6.29}$$

Thus, by insertion of the appropriate values of N and S into $f(N,S)$ we can evaluate the number of independent spin states of a given multiplicity available for an N-electron system. The results of such an evaluation are often presented in graphic form, in a construct such as Fig. VI.2 termed *branching diagram*, from which it is possible to carry out the analysis with very little labor. This diagram shows the number of states of different multiplicities obtainable for a given number of independent electrons. The diagram is very easy to construct beginning with a single electron and successively coupling other electron spins to it in all possible algebraic ways. In the diagram the number of states of a given multiplicity is indicated within a circle whose abscissa is the number of electrons and whose ordinate is the multiplicity. The diagram is constructed in such a way that each encircled number is the sum of the two adjacent encircled numbers to the left. We see that for four electrons one would have two singlet states ($M_S = 0$), three triplet states ($M_S = -1, 0,$ and 1), and one quintet state ($M_S = -2, -1, 0, 1,$ and 2). Thus for four electrons not in closed shells one could write down 16 linearly independent wavefunctions (not all of which would be automatically orthogonal) leading to six different energies.

(ii) A readily usable form of the operator S^2. We may expand this operator as

$$\mathsf{S}^2 = \mathsf{S}_x^2 + \mathsf{S}_y^2 + \mathsf{S}_z^2 \tag{6.30}$$

If we now define

$$\mathsf{S}_+ \equiv \mathsf{S}_x + i\,\mathsf{S}_y \qquad\qquad \mathsf{S}_- \equiv \mathsf{S}_x - i\,\mathsf{S}_y \tag{6.31}$$

Insert on this page Fig. VI.2

we obtain

$$\mathsf{S}_-\mathsf{S}_+ \;=\; \mathsf{S}_x^2 + \mathsf{S}_y^2 - \hbar\,\mathsf{S}_z \tag{6.32}$$

where the commutator relationship $[\mathsf{S}_x, \mathsf{S}_y] \;=\; i\,\hbar\,\mathsf{S}_z$ has been used. Insertion of Eq. (6.32) into (6.30) yields

$$\mathsf{S}^2 \;=\; \mathsf{S}_z^2 + \hbar\,\mathsf{S}_z + \mathsf{S}_-\,\mathsf{S}_+ \tag{6.33}$$

An equivalent form readily verifiable, is

$$\mathsf{S}^2 \;=\; \mathsf{S}_z^2 - \hbar\,\mathsf{S}_z + \mathsf{S}_+\,\mathsf{S}_- \tag{6.34}$$

At this point we note that

$$\mathsf{S}_+ \;\equiv\; \sum_j \mathsf{s}_+(j) \qquad\qquad \mathsf{S}_- \;\equiv\; \sum_j \mathsf{s}_-(j) \tag{6.35}$$

$$\mathsf{s}_+(j) \;\equiv\; \mathsf{s}_x(j) + i\,\mathsf{s}_y(j) \qquad\qquad \mathsf{s}_-(j) \;\equiv\; \mathsf{s}_x(j) - i\,\mathsf{s}_y(j) \tag{6.36}$$

where j is an electron numbering index and s is a one-electron operator. The effects of s_+ and s_- on the spin functions are

$$\mathsf{s}_-(j)\begin{pmatrix}\alpha(j)\\ \beta(j)\end{pmatrix} \;=\; \hbar\begin{pmatrix}\beta(j)\\ 0\end{pmatrix} \qquad\qquad \mathsf{s}_+(j)\begin{pmatrix}\alpha(j)\\ \beta(j)\end{pmatrix} \;=\; \hbar\begin{pmatrix}0\\ \alpha(j)\end{pmatrix} \tag{6.37}$$

The operator s_- steps down α to β but annihilates β whereas s_+ annihilates α but steps up β to α. If the operands of S^2 are limited to Slater determinants or to simple products of spinorbitals (these will be always assumed to be the cases), the Eqs. (6.33) and (6.34) can be simplified further.

Consider the determinant

$$\mathbf{\Phi} \;=\; |\,\psi_1\,\alpha(1)\,\psi_2\,\beta(2)\,\psi_3\,\alpha(3)\cdots\psi_N\,\alpha(N)\,| \tag{6.38}$$

in which each electron is associated with a different space function. Since the spin operators do not affect the orbital parts and since no restriction need be placed on the spin function

associated with a given space function (i.e., the spin function may be either α or β) it follows that the determinantal wavefunction of Eq. (6.38) may be abbreviated to

$$\mathbf{\Phi} \;=\; |\,\alpha(1)\,\beta(2)\,\alpha(3)\cdots\alpha(N)\,| \tag{6.39}$$

Indeed, with no loss of information, we can use the condensation

$$\mathbf{\Phi} \;=\; \alpha(1)\,\beta(2)\,\alpha(3)\cdots \;\equiv\; \alpha\,\beta\,\alpha\cdots \tag{6.40}$$

as long as we remember that all of these constitute a *short hand* for Eq. (6.38).

The results of operating on $\mathbf{\Phi}$ of Eqs. (6.38), (6.39), or (6.40) with S_z and S_z^2 are

$$\mathsf{S}_z\,\mathbf{\Phi} \;=\; \frac{1}{2}\,(N_\alpha - N_\beta)\,\hbar\,\mathbf{\Phi} \tag{6.41}$$

$$\mathsf{S}_z^2\,\mathbf{\Phi} \;=\; \frac{1}{4}\,(N_\alpha - N_\beta)^2\,\hbar^2\,\mathbf{\Phi} \tag{6.42}$$

We may rewrite $\mathsf{S}_-\mathsf{S}_+$ in the form

$$\mathsf{S}_-\,\mathsf{S}_+ \;=\; \sum_i^N \mathsf{s}_-(i)\,\mathsf{s}_+(i) + \sum_i^N \sum_{i\neq j}^N \mathsf{s}_-(i)\,\mathsf{s}_+(j) \tag{6.43}$$

and consider each of the parts on the right-hand side of this equation with respect to their effects on $\mathbf{\Phi}$. Consider first

$$\sum_i^N \mathsf{s}_-(i)\,\mathsf{s}_+(i)\,\mathbf{\Phi} \;=\; \sum_i^N \mathsf{s}_-(i)\,\mathsf{s}_+(i)\,|\,\alpha(1)\,\beta(2)\,\cdots\,\alpha(N)\,| \tag{6.44}$$

By virtue of the one-electron nature of the $\mathsf{s}_-(i)\,\mathsf{s}_+(i)$ operator, Eq. (6.44) reduces to

$$\sum_i^N \mathsf{s}_-(i)\,\mathsf{s}_+(i)\,\mathbf{\Phi} \;=\; N_\beta\,\hbar^2\,\mathbf{\Phi} \tag{6.45}$$

where the last equality follows from the noncommutativity of s_- and s_+ and from

$$\mathsf{s}_-(i)\,\mathsf{s}_+(i)\,\begin{pmatrix}\alpha(i)\\ \\ \beta(i)\end{pmatrix} \;=\; \hbar^2\,\begin{pmatrix}0\\ \\ \beta(i)\end{pmatrix} \tag{6.46}$$

Consider next the two-electron operator part $\sum_i \sum_{i\neq j} \mathsf{s}_-(i)\mathsf{s}_+(j)$ and note that the result

$$\mathsf{s}_-(i)\,\mathsf{s}_+(j)\,\alpha(i)\,\beta(j) \;=\; \hbar^2\,\beta(i)\,\alpha(j) \tag{6.47}$$

implies a formal identity of the operators

$$\sum_i \sum_{i \neq j} \mathsf{s}_-(i)\, \mathsf{s}_+(j) \;=\; \hbar^2 \sum_{\mathsf{P}} \mathsf{P}_{\alpha\beta} \tag{6.48}$$

where $\mathsf{P}_{\alpha\beta}$ is an operator which exchanges α and β functions in the original determinant (i.e., $\mathsf{P}_{\alpha\beta}$ is that subclass of all electron permutation operators which leads to an interchange of α and β spins), and where the sum is taken over all possible interchanges. For the two-electron permutations, no terms survive. The effect of the operator $\mathsf{S}_-\,\mathsf{S}_+$ on $\mathbf{\Phi}$ is then summarized in the expression

$$\mathsf{S}_-\,\mathsf{S}_+\,\mathbf{\Phi} \;=\; \left(N_\beta + \sum_{\mathsf{P}} \mathsf{P}_{\alpha\beta}\right) \hbar^2\,\mathbf{\Phi} \tag{6.49}$$

If we not collect all the terms of Eqs. (6.41), (6.42), and (6.49) and insert them into Eq. (6.33), we obtain after a small amount of rearrangement

$$\mathsf{S}^2\,\mathbf{\Phi} \;=\; \left[\sum_{\mathsf{P}} \mathsf{P}_{\alpha\beta} + \frac{1}{4}\left[(N_\alpha - N_\beta)^2 + 2\,N\right]\right] \hbar^2\,\mathbf{\Phi} \tag{6.50}$$

where $N = N_\alpha + N_\beta$. Thus, given $\mathbf{\Phi}$ as an eigenfunction of S^2, it is a relatively simple matter to determine S.

(iii) A general method for construction of eigenfunctions with given S and M_S characteristics. For this purpose, we use projection operator techniques.

We now turn out attention to the generation of eigenfunctions of S^2. Suppose we start with a trial wavefunction $\mathbf{\Phi}_t$ which satisfies one condition - namely, that it be resolvable into a linear combination of eigenfunctions of S^2

$$\mathbf{\Phi}_t \;=\; \sum_k c_{ik}\,\mathbf{\Phi}(S_k) \tag{6.51}$$

where the summation runs over all possible values of S_k consistent with the specific number of electrons under consideration and where

$$\mathsf{S}^2\,\mathbf{\Phi}(S_k) \;=\; S_k\,(S_k+1)\,\hbar^2\,\mathbf{\Phi}(S_k) \tag{6.52}$$

The coefficients c_{ik} need not be known; all that is required is that $\mathbf{\Phi}_t$ be a function in the domain spanned by the basis set $\mathbf{\Phi}(S_k)$. It then follows that

$$[\mathsf{S}^2 - S_i\,(S_i+1)\,\hbar^2]\,\mathbf{\Phi}_t \;=\; \sum_{k \neq i} d_k\,\mathbf{\Phi}(S_k) \tag{6.53}$$

We may consider $[\mathsf{S}^2 - S_i(S_i+1)]$ to be an operator which eliminates any admixture of an S_i wavefunction from $\mathbf{\Phi}_t$. Successive application of such *annihilators* finally yields one pure spin state of specified multiplicity - if such a spin state exists in $\mathbf{\Phi}_t$. The required operator, with $\lambda_j = S_j(S_j+1)\hbar^2$, can be written as

$$\mathcal{O}(S_k) = \prod_{i \neq k} \left(\frac{\mathsf{S}^2 - \lambda_i}{\lambda_k - \lambda_i} \right) \tag{6.54}$$

This operator projects a wavefunction of the required spin quantum number (i.e., S_k) out of the general spin space; $\mathcal{O}(S_k)$ is a spin-projection operator.

E. Spin Eigenfunctions of a Three-Electron System: An Example Calculation

The eigenfunctions of S_z may be written immediately. In a simple short-hand form, using the condensate notation of Eq. (6.40) and from the results previously obtained for a three-electron system, they are

$$\begin{aligned}
&\alpha\,\alpha\,\alpha && M_S = \frac{3}{2} \\
&\alpha\,\alpha\,\beta && M_S = \frac{1}{2} \\
&\alpha\,\beta\,\alpha && M_S = \frac{1}{2} \\
&\beta\,\alpha\,\alpha && M_S = \frac{1}{2} \\
&\alpha\,\beta\,\beta && M_S = -\frac{1}{2} \\
&\beta\,\alpha\,\beta && M_S = -\frac{1}{2} \\
&\beta\,\beta\,\alpha && M_S = -\frac{1}{2} \\
&\beta\,\beta\,\beta && M_S = -\frac{3}{2}
\end{aligned}$$

where, for example, $\alpha\,\beta\,\alpha$ denotes $|\,\psi_1\,\alpha(1)\,\psi_2\,\beta(2)\,\psi_3\,\alpha(3)\,| \equiv |\,\phi_1\,\overline{\phi}_2\,\phi_3\,|$ with $\psi_1 \neq \psi_2 \neq \psi_3$. Inspection of the branching diagram of Fig. VI.2 indicates that we should obtain one quartet spin state ($S = \frac{3}{2}$) and two doublet spin states ($S = \frac{1}{2}$). The projection operators are

$$\mathcal{O}[3/2] = \frac{\mathsf{S}^2 - 3/4}{15/4 - 3/4} \tag{6.55}$$

$$\mathcal{O}[1/2] \;=\; \frac{\mathsf{S}^2 - 15/4}{3/4 - 15/4} \tag{6.56}$$

where, for brevity, we have eliminated the factor $\hbar^2$ from the operator S^2. Since we have made no attempt to normalize the spin-projection operators, we fully expect that the spin-projected wavefunctions require normalization.

(a) Spin Eigenfunctions with $S = \frac{3}{2}$, $M_S = \frac{3}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{3}{2}$

We first let $\mathcal{O}[3/2]$ operate on $\alpha\,\alpha\,\alpha$. The result is

$$\mathcal{O}[3/2]\,\alpha\,\alpha\,\alpha \;=\; \frac{[0 + \frac{1}{4}\,(3^2 + 2\times 3)]\,\alpha\,\alpha\,\alpha - \frac{3}{4}\,\alpha\,\alpha\,\alpha}{15/4 - 3/4} \;=\; \alpha\,\alpha\,\alpha \tag{6.57}$$

Thus, it is clear that $\alpha\,\alpha\,\alpha$ is a satisfactory eigenfunction with $S = \frac{3}{2}$, $M_S = \frac{3}{2}$. Using the notation ${}^{2S+1}\mathbf{\Phi}_{M_S}$ to denote the eigenfunctions of a given multiplicity and a given M_S we can write

$${}^4\mathbf{\Phi}_{3/2} \;=\; \alpha\,\alpha\,\alpha \tag{6.58}$$

In order to exemplify the manner in which unwanted functions are eliminated, consider the result of $\mathcal{O}[1/2$ operating on $\alpha\,\alpha\,\alpha$. The result is

$$\mathcal{O}[1/2]\,\alpha\,\alpha\,\alpha \;=\; \frac{0 + \frac{1}{4}\,(3^2 + 2\times 3)]\,\alpha\,\alpha\,\alpha - \frac{15}{4}\,\alpha\,\alpha\,\alpha}{3/4 - 15/4} \;=\; 0 \tag{6.59}$$

Thus, as evidenced by our inability to project a doublet spin eigenfunction out of $\alpha\,\alpha\,\alpha$, there is no doublet character in $\alpha\,\alpha\,\alpha$. This, of course, had to be the case - $\alpha\,\alpha\,\alpha$ is, after all, a quartet eigenfunction.

We next construct the spin wavefunction with $S = \frac{3}{2}$, $M_S = \frac{1}{2}$. Clearly, this eigenfunction may be a linear combination of $\alpha\,\alpha\,\beta$, $\alpha\,\beta\,\alpha$, and $\beta\,\alpha\,\alpha$ only. We start with $\alpha\,\alpha\,\beta$ to find

$$\begin{aligned}
\mathcal{O}[3/2]\,\alpha\,\alpha\,\beta \;&=\; \frac{\left(\sum_{\mathsf{P}} \mathsf{P}_{\alpha\beta} + \frac{1}{4}\,[(N_\alpha - N_\beta)^2 + 2\,N_\alpha + 2\,N_\beta]\right)\alpha\,\alpha\,\beta - \frac{3}{4}\,\alpha\,\alpha\,\beta}{15/4 - 3/4} \\
&=\; \frac{\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha + \frac{1}{4}\,[1^2 + 2\times 2 + 2\times 1]\,\alpha\,\alpha\,\beta - \frac{3}{4}\,\alpha\,\alpha\,\beta}{15/4 - 3/4} \\
&=\; \frac{1}{3}\,(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha + \alpha\,\alpha\,\beta)
\end{aligned} \tag{6.60}$$

Normalization yields

$$^4\mathbf{\Phi}_{1/2} = \frac{1}{3^{1/2}}(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha + \alpha\,\alpha\,\beta) \tag{6.61}$$

The spin wavefunction with $S = \frac{3}{2}$, $M_S = -\frac{1}{2}$ is constructed in like manner from $\alpha\,\beta\,\beta$, $\beta\,\alpha\,\beta$, $\beta\,\beta\,\alpha$. It is found to be

$$^4\mathbf{\Phi}_{-1/2} = \frac{1}{3^{1/2}}(\alpha\,\beta\,\beta + \beta\,\alpha\,\beta + \beta\,\beta\,\alpha) \tag{6.62}$$

The last of these four spin wavefunctions is

$$^4\mathbf{\Phi}_{-3/2} = \beta\,\beta\,\beta \tag{6.63}$$

(b) Spin Eigenfunctions with $S = \frac{1}{2}$, $M_S = \frac{1}{2}, -\frac{1}{2}$

There are two doublet spin wavefunctions. Those with $M_S = 1/2$ are constructed from $\alpha\,\alpha\,\beta$, $\alpha\,\beta\,\alpha$, and $\beta\,\alpha\,\alpha$. Letting $\mathcal{O}[1/2]$ operate on each of these in turn, we find

$$\begin{aligned}\mathcal{O}[1/2]\,\alpha\,\alpha\,\beta &= \frac{\left(\sum_{\mathsf{P}} \mathsf{P}_{\alpha\beta} + \frac{1}{4}\left[(N_\alpha - N_\beta)^2 + 2\,N_\alpha + 2\,N_\beta\right]\right)\alpha\,\alpha\,\beta - \frac{15}{4}\,\alpha\,\alpha\,\beta}{3/4 - 15/4} \\ &= \frac{\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha + \frac{1}{4}\left[1^2 + 2\times 2 + 2\times 1\right]\alpha\,\alpha\,\beta - \frac{15}{4}\,\alpha\,\alpha\,\beta}{3/4 - 15/4} \\ &= -\frac{1}{3}(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha - 2\,\alpha\,\alpha\,\beta)\end{aligned} \tag{6.64}$$

Analogously, it is found

$$\mathcal{O}[1/2]\,\alpha\,\beta\,\alpha = -\frac{1}{3}(\alpha\,\alpha\,\beta + \beta\,\alpha\,\alpha - 2\,\alpha\,\beta\,\alpha) \tag{6.65}$$

$$\mathcal{O}[1/2]\,\beta\,\alpha\,\alpha = -\frac{1}{3}(\alpha\,\beta\,\alpha + \alpha\,\alpha\,\beta - 2\,\beta\,\alpha\,\alpha) \tag{6.66}$$

Normalization of these three functions yields

$$^2\mathbf{\Phi}_{1/2} = (1/6)^{1/2}(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha - 2\,\alpha\,\alpha\,\beta)$$

$$^2\mathbf{\Phi}_{1/2} = (1/6)^{1/2}(\alpha\,\alpha\,\beta + \beta\,\alpha\,\alpha - 2\,\alpha\,\beta\,\alpha)$$

$$^2\mathbf{\Phi}_{1/2} = (1/6)^{1/2}(\alpha\,\beta\,\alpha + \alpha\,\alpha\,\beta - 2\,\beta\,\alpha\,\alpha)$$

Thus three wavefunctions with $S = \frac{1}{2}$, $M_S = \frac{1}{2}$ have been constructed. However, one of them is linearly dependent on the other two and they are not orthogonal. Redundancy can

be eliminated and orthogonality produced if we note the cyclic symmetry exhibited by these three wavefunctions and make maximum use of it. It is clear that these three wavefunctions can be represented as three in-plane vectors separated by 120°. Thus, the difference of any two functions is orthogonal to the remaining one. A possible set of orthonormal wavefunctions, then, is

$$^2\mathbf{\Phi}_{1/2} = (1/6)^{1/2}\,(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha - 2\,\alpha\,\alpha\,\beta) \tag{6.67}$$

$$^2\mathbf{\Phi}_{1/2} = (1/2)^{1/2}\,(\beta\,\alpha\,\alpha - \alpha\,\beta\,\alpha) \tag{6.68}$$

The generation of the two wavefunctions $^2\mathbf{\Phi}_{-1/2}$ proceeds in a similar way. One gets

$$^2\mathbf{\Phi}_{-1/2} = (1/6)^{1/2}\,(\beta\,\alpha\,\beta + \beta\,\beta\,\alpha - 2\,\alpha\,\beta\,\beta) \tag{6.69}$$

$$^2\mathbf{\Phi}_{-1/2} = (1/2)^{1/2}\,(\beta\,\beta\,\alpha - \beta\,\alpha\,\beta) \tag{6.70}$$

These results are summarized as follows:

S	M_S	*Spin Adapted Configuration*
3/2	+ 3/2	$^4\mathbf{\Phi}_{3/2} = \alpha\,\alpha\,\alpha$
3/2	+ 1/2	$^4\mathbf{\Phi}_{1/2} = 1/3^{1/2}\,(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha + \alpha\,\alpha\,\beta)$
3/2	− 1/2	$^4\mathbf{\Phi}_{-1/2} = 1/3^{1/2}\,(\alpha\,\beta\,\beta + \beta\,\alpha\,\beta + \beta\,\beta\,\alpha)$
3/2	− 3/2	$^4\mathbf{\Phi}_{-3/2} = \beta\,\beta\,\beta$
1/2	+ 1/2	$^2\mathbf{\Phi}_{1/2} = 1/6^{1/2}\,(\alpha\,\beta\,\alpha + \beta\,\alpha\,\alpha - 2\,\alpha\,\alpha\,\beta)$
1/2	+ 1/2	$^2\mathbf{\Phi}_{1/2} = 1/2^{1/2}\,(\beta\,\alpha\alpha - \alpha\,\beta\,\alpha)$
1/2	− 1/2	$^2\mathbf{\Phi}_{-1/2} = 1/6^{1/2}\,(\beta\,\alpha\,\beta + \beta\,\beta\,\alpha - 2\,\alpha\,\beta\,\beta)$
1/2	− 1/2	$^2\mathbf{\Phi}_{-1/2} = 1/2^{1/2}\,(\beta\,\beta\,\alpha - \beta\,\alpha\,\beta)$

VII. OCCUPATION-NUMBER FORMALISM

A. Occupation-Number Wavefunction

Many-fermion wavefunctions are often expressed in terms of determinants formed from a basis of one-particle functions. These determinants may be described by their *occupancy*, (i.e., by listing the basis functions that are used therein). Much of the discussion of many-fermion systems can be systematized and thereby simplified, by adopting a suitable set of notations for representing determinantal functions, the relationships between them, and the effect of various operators thereon. The purpose of this Section is to develop such a notation and the accompanying algebra.

Let $\phi_\mu, \phi_\nu, ...$ denote an orthonormal basis of one-particle states to be used in constructing many-fermion wavefunctions. The set of one-particle functions is assigned a lexical ordering, which may be arbitrary but must remain unchanged during all of the discussion to follow. For an N-fermion system, an *occupation-number* determinantal wavefunction can be identified by specifying the N one-particle states to be occupied, then forming the determinantal function based on these states *in their lexical order*. For example, if $\phi_1, \phi_2, ...$ denote the one-particle states in their lexical ordering, the normalized occupation-number wavefunction Ψ formed from ϕ_7, ϕ_3, ϕ_8 will be

$$\Psi = (3!)^{1/2}\mathcal{A}[\phi_3(1)\phi_7(2)\phi_8(3)] \equiv [\phi_3, \phi_7, \phi_8] \tag{7.1}$$

where $\mathcal{A}$ is the antisymmetrizer, given in general, for an N-particle system, by

$$\mathcal{A} = \frac{1}{N!}\sum_{\mathsf{P}}(-)^P\mathsf{P} \tag{7.2}$$

and where the sum is over all $N!$ permutations.

Note that the definition (7.1) uniquely determines the sign to be associated with each occupation-number wafefunction. We shall use the definition of Eq. (7.1) even when it does not indicate a lexically ordered occupation-number wavefunction. For example, $[\phi_7, \phi_3, \phi_8]$ will denote a function equal to $-[\phi_3, \phi_7, \phi_8]$.

It is convenient to have a shorthand for characterizing the relationships between different occupation-number wavefunctions. For this purpose we introduce the so-called *annihilation operator* $\mathbf{a}_\mu$, whose effect is to remove the function ϕ_μ from the occupied states list. Its adjoint, the operator $\mathbf{a}_\mu^\dagger$, is called a *creation operator*, with the effect that it adds the function ϕ_μ to the occupied states list. If a function to be annihilated is already absent, or if a function to be added is already present, the operation yields vanishing results. Further discussions are simplified by making a superficially complex choice of the signs to be associated with the application of $\mathbf{a}_\mu$ and $\mathbf{a}_\mu^\dagger$; the precise definition of these operators is as follows:

$$\begin{aligned}
&\mathbf{a}_\mu^\dagger[\cdots,\phi_\lambda,\phi_\nu,\cdots] = (-)^{n_\mu}[\cdots\phi_\lambda,\phi_\mu,\phi_\nu,\cdots] \\
&\mathbf{a}_\mu^\dagger[\cdots,\phi_\mu,\cdots] = 0 \\
&\mathbf{a}_\mu[\cdots,\phi_\lambda,\phi_\mu,\phi_\nu,\cdots] = (-)^{n_\mu}[\cdots,\phi_\lambda,\phi_\nu,\cdots] \\
&\mathbf{a}_\mu[\cdots,\phi_\lambda,\phi_\nu,\cdots] = 0
\end{aligned} \tag{7.3}$$

where the one-particle states in each [ ] are in lexical order, and n_μ is the number of occupied one-particle states that lexically precede ϕ_μ in the adjacent [ ].

We should verify that the definitions of Eqs. (7.3) are consistent with the mathematical meaning of *adjoint*. It is sufficient to have

$$< \mathbf{a}_\mu^\dagger[\,]|[\,]' >=< [\,]|\mathbf{a}_\mu[\,]' > \tag{7.4}$$

for all lexically ordered determinantal function [ ] and [ ]′ such that [ ]′ contains one more one-particle state than [ ]. Eq. (7.4) is clearly satisfied, as both its members vanish unless $\mathbf{a}_\mu^\dagger[\,] = (-)^{n_\mu}[\,]'$, in which case we also have $\mathbf{a}_\mu[\,]' = (-)^{n_\mu}[\,]$, and each member reduces to $(-)^{n_\mu}$.

B. Commutation Rules for Creation and Annihilation Operators

Since the result of applying $\mathbf{a}_\mu$ or $\mathbf{a}_\mu^\dagger$ has a sign dependent upon the number of occupied states preceding ϕ_μ in lexical order, creation-annihilation operator products produce results

whose signs depend upon the ordering of the factors. In particular, consider the sccessive application of $\mathbf{a}_\mu$ and $\mathbf{a}_\nu$ to an occupation-number wavefunction Ψ containing both ϕ_μ and ϕ_ν. We assume without loss of generality that ϕ_μ precedes ϕ_ν, and write $\Psi = [\cdots, \phi_\mu, \cdots, \phi_\nu, \cdots]$. Letting n_μ and n_ν be the numbers of occupied one-particle states respectively preceding ϕ_μ and ϕ_ν in Ψ, we have

$$\begin{aligned}\mathbf{a}_\nu\mathbf{a}_\mu\Psi &= \mathbf{a}_\nu\mathbf{a}_\mu[\cdots, \phi_\mu, \cdots, \phi_\nu, \cdots] = (-)^{n_\mu}\mathbf{a}_\nu[\cdots, \phi_\nu, \cdots] \\ &= (-)^{n_\mu}(-)^{n_\nu-1}[\cdots\cdots] \end{aligned} \tag{7.5}$$

The factor $(-)^{n_\nu-1}$ arises because the prior removal of ϕ_μ causes only $n_\nu - 1$ one-particle states to precede ϕ_ν in the function to which $\mathbf{a}_\nu$ is applied. The situation is different if $\mathbf{a}_\nu$ is applied first:

$$\begin{aligned}\mathbf{a}_\mu\mathbf{a}_\nu\Psi &= \mathbf{a}_\mu\mathbf{a}_\nu[\cdots, \phi_\mu, \cdots, \phi_\nu, \cdots] = (-)^{n_\nu}\mathbf{a}_\mu[\cdots, \phi_\mu, \cdots] \\ &= (-)^{n_\nu}(-)^{n_\mu}[\cdots\cdots] \end{aligned} \tag{7.6}$$

Adding Eqs. (7.5) and (7.6), we have

$$(\mathbf{a}_\mu\mathbf{a}_\nu + \mathbf{a}_\nu\mathbf{a}_\mu)\Psi = 0 \tag{7.7}$$

for all occupation-number wavefunctions Ψ containing ϕ_μ and ϕ_ν. Since Eq. (7.7) is trivially satisfied if ϕ_μ or ϕ_ν is absent from Ψ, we have the operator identity

$$\mathbf{a}_\mu\mathbf{a}_\nu + \mathbf{a}_\nu\mathbf{a}_\mu = 0 \tag{7.8}$$

Since Eq. (7.8) is symmetric in μ and ν and is clearly satisfied if $\mu = \nu$, the restriction that ϕ_μ precedes ϕ_ν may now be dropped.

An entirely similar argument leads to the analogous result for creation operators:

$$\mathbf{a}_\mu^\dagger\mathbf{a}_\nu^\dagger + \mathbf{a}_\nu^\dagger\mathbf{a}_\mu^\dagger = 0 \tag{7.9}$$

Eq. (7.9) may also be derived simply by taking the adjoint of Eq. (7.8).

Consider next the operator products $\mathbf{a}_\mu^\dagger \mathbf{a}_\nu$, and $\mathbf{a}_\nu \mathbf{a}_\mu^\dagger$. Suppose again the ϕ_μ precedes ϕ_ν in lexical order. Let Ψ be an occupation-number wavefunction containing ϕ_ν but not ϕ_μ, with η_ν and η_μ the numbers of occupied states respectively preceding ϕ_ν and ϕ_μ. Then

$$\mathbf{a}_\mu^\dagger \mathbf{a}_\nu \Psi = \mathbf{a}_\mu^\dagger \mathbf{a}_\nu [\cdots, \phi_\nu, \cdots] = (-)^{n_\nu} \mathbf{a}_\mu^\dagger [\cdots\cdots]$$
$$= (-)^{n_\mu} (-)^{n_\nu} [\cdots, \phi_\mu, \cdots] \tag{7.10}$$

$$\mathbf{a}_\nu \mathbf{a}_\mu^\dagger \Psi = \mathbf{a}_\nu \mathbf{a}_\mu^\dagger [\cdots, \phi_\nu, \cdots] = (-)^{n_\mu} \mathbf{a}_\nu [\cdots, \phi_\mu, \cdots, \phi_\nu, \cdots]$$
$$= (-)^{n_\mu} (-)^{n_\nu + 1} [\cdots, \phi_\mu, \cdots] \tag{7.11}$$

Eqs. (7.10) and (7.11) may be added to reach

$$(\mathbf{a}_\mu^\dagger \mathbf{a}_\nu + \mathbf{a}_\nu \mathbf{a}_\mu^\dagger) \Psi = 0 \tag{7.12}$$

Eq. (7.12) can also be satisfied if ϕ_ν precedes ϕ_μ in lexical order. Thus,

$$\mathbf{a}_\mu^\dagger \mathbf{a}_\nu + \mathbf{a}_\nu \mathbf{a}_\mu^\dagger = 0 \qquad\qquad (\nu \neq \mu) \tag{7.13}$$

Eq. (7.13) in contrast to Eqs. (7.8) and (7.9), is *not* satisfied for $\mu = \nu$. However, for an occupation-numnber wafunction Ψ containing ϕ_μ:

$$\mathbf{a}_\mu^\dagger \mathbf{a}_\mu \Psi = \mathbf{a}_\mu^\dagger \mathbf{a}_\mu [\cdots, \phi_\mu, \cdots] = (-)^{n_\mu} \mathbf{a}_\mu^\dagger [\cdots\cdots]$$
$$= (-)^{n_\mu} (-)^{n_\mu} [\cdots, \phi_\mu, \cdots] = \Psi \tag{7.14}$$

$$\mathbf{a}_\mu \mathbf{a}_\mu^\dagger \Psi = 0 \tag{7.15}$$

where n_μ is the number of occupied states preceding ϕ_μ. If Ψ does not contain ϕ_μ,

$$\mathbf{a}_\mu^\dagger \mathbf{a}_\mu \Psi = 0 \tag{7.16}$$

$$\mathbf{a}_\mu \mathbf{a}_\mu^\dagger \Psi = \mathbf{a}_\mu \mathbf{a}_\mu^\dagger [\cdots\cdots] = (-)^{n_\mu} \mathbf{a}_\mu [\cdots, \phi_\mu, \cdots]$$
$$= (-)^{n_\mu} (-)^{n_\mu} [\cdots\cdots] = \Psi \tag{7.17}$$

Thus, whether or not Ψ contains ϕ_μ,

$$(\mathbf{a}_\mu \mathbf{a}_\mu^\dagger + \mathbf{a}_\mu^\dagger \mathbf{a}_\mu) \Psi = \Psi \tag{7.18}$$

Eq. (7.18) implies

$$\mathbf{a}_\mu \mathbf{a}_\mu^\dagger + \mathbf{a}_\mu^\dagger \mathbf{a}_\mu = 1 \tag{7.19}$$

Eqs. (7.8), (7.9), (7.13), and (7.19) all have left sides of the form of *anticommutators*. These equations may therefore be written

$$\begin{aligned} [\mathbf{a}_\mu, \mathbf{a}_\nu]_+ &= 0 \\ [\mathbf{a}_\mu^\dagger, \mathbf{a}_\nu^\dagger]_+ &= 0 \\ [\mathbf{a}_\mu, \mathbf{a}_\nu^\dagger]_+ &= \delta_{\mu\nu} \end{aligned} \tag{7.20}$$

C. Reference States and Antisymmetry

The most convenient way to use creation and annihilation operators in specifying determinantal wavefunctions is by using them to generate such wavefunctions from a reference state. The simplest such state is the *vacuum state* $|0>$, which corresponds formally to an occupation-number wavefunction containing no occupied one-particle states. The state $|0>$ itself has no observational meaning, but completely meainingful results may be obtained by applying creation or annihilation operators thereto, using the rules contained in Eq. (7.20). For example, taking $\phi_1, \phi_2, \cdots$ to be the one-particle states in lexical order, $\mathbf{a}_4^\dagger|0> = [\phi_4]$, $\mathbf{a}_6^\dagger \mathbf{a}_4^\dagger|0> = -[\phi_4, \phi_6]$, $\mathbf{a}_3|0> = 0$, and so on.

An arbitrary determinant may be generated from $|0>$ by operation with a string of creation operators for the one-particle states to be occupied therein. If the creation operators are arranged in ascending lexical order from left to right, their operation on $|0>$ will generate the appropriate occupation-number wavefunction with a positive sign. For example:

$$\mathbf{a}_2^\dagger \mathbf{a}_3^\dagger \mathbf{a}_5^\dagger \mathbf{a}_6^\dagger|0> = +[\phi_2, \phi_3, \phi_5, \phi_6] \tag{7.21}$$

This statement becomes obvious if we recognize that at the time of application of each $\mathbf{a}_\mu^\dagger$, it operates on a function with no occupancy preceding ϕ_μ. If, however, the $\mathbf{a}_\mu^\dagger$ are not arranged in ascending order, we can permute them to that order with the aid of the anticommutator

listed in Eq. (7.20). We see that each interchange of adjacent $\mathbf{a}_\mu^\dagger$ and $\mathbf{a}_\nu^\dagger$ will introduce a minus sign, so that the ordering of the $\mathbf{a}_\mu^\dagger$ will be accompanied by the sign factor $(-)^P$ associated with the parity of the ordering permutation. Since this is the same sign factor as is associated with the reordering of the rows or columns of a determinant, we may make an appropriate permutation of the one-particle functions in [ ] instead of including a factor $(-)^P$; the result of this observation is the equation

$$\mathbf{a}_\mu^\dagger \cdots \mathbf{a}_\nu^\dagger |0> = +[\phi_\mu, \cdots, \phi_\nu] \tag{7.22}$$

which is valid whether or not $\phi_\mu, \cdots, \phi_\nu$ are in lexical order. Of course, if these functions are not in lexical order, the right-hand side of Eq. (7.22) will not be in the standard form of an occupation-number wavefunction.

Another, even more useful reference state for an N-fermion problem is an N-particle determinantal state Φ obtained as its independent-particle zero-order solution. Letting ϕ_α, $\phi_\beta, \cdots$ stand for the one-particle states occupied in Φ, in lexical order, we write

$$\begin{aligned} \Phi &= (N!)^{1/2} \mathcal{A}[\phi_\alpha(1)\phi_\beta(2), \cdots, \phi_\omega(N)] \\ &= \mathbf{a}_\alpha^\dagger \mathbf{a}_\beta^\dagger \cdots \mathbf{a}_\omega^\dagger |0> \end{aligned} \tag{7.23}$$

Other determinantal states may now be obtained by applying annihilation operators to remove unwanted one-particle states in Φ, and applying creation operators to add new occupied one-particle states. If a new determinantal state is to describe the same number of particles as Φ, the operator string applied to Φ must contain equal numbers of creation and annihilation operators.

Let us consider more closely the determinantal states we may produce from Φ. We may classify such states by their degree of excitation, defined as the number of one-particle states occupied in Φ but not in the new wavefunction. Letting ϕ_r, $\phi_s, \cdots$ denote one-particle states unoccupied in Φ, the singly, doubly, ... excited determinants relative to Φ are of the forms

$$\Phi_\alpha^r = \mathbf{a}_r^\dagger \mathbf{a}_\alpha \Phi$$

$$\Phi_{\alpha\beta}^{rs} = \mathbf{a}_r^\dagger \mathbf{a}_s^\dagger \mathbf{a}_\beta \mathbf{a}_\alpha \Phi$$

$$\cdots\cdots \tag{7.24}$$

We shall use the notation of Eq. (7.24) irrespective of the lexical ordering of α, β, r, $s, \cdots$. For describing determinants with arbitrary degrees of excitation, we introduce collective indices R and Δ to stand for r, $s, \cdots$ and α, $\beta, \cdots$, respectively, and write A^R_Δ to stand for the operator sequence $\mathbf{a}^\dagger_r \mathbf{a}^\dagger_s \cdots \mathbf{a}^\dagger_\beta \mathbf{a}^\dagger_\alpha$. Then the general form of Eq. (7.24) may be expressed

$$\Phi^R_\Delta = \mathsf{A}^R_\Delta \Phi \tag{7.25}$$

It should be obvious that the remainder of a complete set of N-particle determinants can be formed from Φ according to Eq. (7.25) (i.e., using creation (annihilation) operators only for the one-particle states unoccupied (occupied) in Φ). For logical completeness, the possible index sets denoted by R or Δ must include the empty set.

One way of describing a wavefunction like Φ^r_α is by comparison with the reference state Φ; we may say that Φ^r_α has a *hole* in state ϕ_α and a *particle* in state ϕ_r. Following this terminology, $\phi_\alpha, \phi_\beta, \cdots$ are in more general contexts called *hole states*, while $\phi_r, \phi_s, \cdots$ are known as *particle states*. The index sets $\alpha, \beta, \cdots$ and $r, s, \cdots$ are respectively called *hole* and *particle indices*. The transfer of a fermion from ϕ_α to ϕ_r may therefore be characterized as the creation of a hole and a particle, or a *particle-hole pair*. The reverse transfer from ϕ_r to ϕ_α, annihilates both members of the pair. The operator A^r_α is thus a particle-hole (pair) creation operator. Wavefunctions of higher degrees of excitation may be thought of as the result of the creation of larger numbers of particles and holes. For example, $\mathsf{A}^{rs}_{\alpha\beta}$ create two particles and two holes.

Because creation and annihilation operators anticommute, wavefunctions such as $\Phi^{rs}_{\alpha\beta}$ have signs that can be changed by index permutation. In particular, we have

$$\Phi^{rs}_{\alpha\beta} = -\Phi^{sr}_{\alpha\beta} = -\Phi^{rs}_{\beta\alpha} = +\Phi^{sr}_{\beta\alpha} \tag{7.26}$$

In general, we may apply arbitrary permutations to the upper or lower indices of any Φ^R_Δ, changing the wavefunction by $(-)^P$, where P is the parity of the permutation applied; if permutations are applied to both upper and lower indices, their respective sign factors must

both be included. This means that Φ^R_Δ is invariant when R and Δ are both subjected to the same permutation.

We look next at the sign associated with the determinantal function Φ^R_α. Remembering that Φ is lexically ordered, and letting η_α and η_r in Φ respectively stand for the numbers of occupied states lexically preceding ϕ_α and ϕ_r in Φ, and tentatively assuming that ϕ_α precedes ϕ_r,

$$\Phi^r_\alpha = \mathbf{a}^\dagger_r \mathbf{a}_\alpha [\cdots \phi_\alpha \cdots] = (-)^{\eta_\alpha} \mathbf{a}^\dagger_r [\cdots] = (-)^{\eta_\alpha} (-)^{\eta_r - 1} [\cdots \phi_r \cdots] \tag{7.27}$$

The sign factor is exactly that which could be produced by moving ϕ_r in [ ] from its lexical position to the position originally occupied by ϕ_α. We thus may interpret Φ^r_α as the determinantal function obtained by direct replacement of ϕ_α by ϕ_r in [ ]. The same conclusion may be reached if ϕ_r precedes ϕ_α. We illustrate the foregoing using the ordered basis ϕ_1, $\phi_2, \cdots$, the reference state $\Phi = [\phi_1, \phi_2, \phi_3, \phi_4]$ and the production of Φ^5_3:

$$\begin{aligned} \Phi^5_3 &= \mathbf{a}^\dagger_5 \mathbf{a}_3 \Phi = \mathbf{a}^\dagger_5 \mathbf{a}_3 [\phi_1, \phi_2, \phi_3, \phi_4] = (-)^2 \mathbf{a}^\dagger_5 [\phi_1, \phi_2, \phi_4] \\ &= (-)^2 (-)^3 [\phi_1, \phi_2, \phi_4, \phi_5] = +[\phi_1, \phi_2, \phi_5, \phi_4] \end{aligned} \tag{7.28}$$

It is important to note that this analysis is not restricted to the reference state; the application of any product $\mathbf{a}^\dagger_\mu \mathbf{a}_\nu$ to any [ ] will produce a function in which ϕ_ν has been directly replaced by ϕ_μ.

In $\Phi^{rs}_{\alpha\beta}$ we encounter a string containing two creation and two annihilation operators. Using the anticommutator property, we may make the following rearrangement:

$$\Phi^{rs}_{\alpha\beta} = \mathbf{a}^\dagger_r \mathbf{a}^\dagger_s \mathbf{a}_\beta \mathbf{a}_\alpha \Phi = (\mathbf{a}^\dagger_s \mathbf{a}_\beta)(\mathbf{a}^\dagger_r \mathbf{a}_\alpha) \Phi \tag{7.29}$$

the sign is positive because an even number of adjacent-operator interchanges is required to move $\mathbf{a}^\dagger_r$ to its final position. Drawing on the result of the preceding paragraph, we see that $\Phi^{rs}_{\alpha\beta}$ is the determinant reached from Φ by direct replacement of ϕ_α by ϕ_r, followed by direct replacement of ϕ_β by ϕ_s. A similar analysis shows that in general, Φ^R_Δ (with R and Δ index strings of the same length) has the sign corresponding to the direct replacements of ϕ_α by ϕ_r, ϕ_β by $\phi_s, \cdots$

D. Representation of Dynamical Operators

The effect of any dynamical operator on a determinantal function can be represented using creation and annihilation operators to describe the linear combination of determinants thereby produced. This representation is possible because dynamical operators are symmetric in the coordinates of identical particles. We start with a discussion of one-particle operators, i.e., operators U that can be written in the form

$$\mathsf{U} = \sum_i \mathsf{u}(i) \tag{7.30}$$

where $\mathsf{u}(i)$ operates on Particle i. In the basis $\phi_\mu, \phi_\nu, \cdots$, the action of u may be described by

$$\mathsf{u}\,\phi_\nu = \sum_\lambda < \phi_\lambda|\mathsf{u}|\phi_\nu > \phi_\lambda \tag{7.31}$$

Consider now the application of U to a determinantal function:

$$\begin{aligned}
\mathsf{U}[\phi_\mu, \phi_\nu, \cdots, \phi_\omega] &= \left(\sum_i \mathsf{u}(i)\right)(N!)^{-1/2}\sum_P (-)^P \mathsf{P}[\phi_\mu(1)\phi_\nu(2)\cdots\phi_\omega(N)] \\
&= (N!)^{-1/2}\sum_P (-)^P \mathsf{P}\left[\left(\sum_i \mathsf{u}(i)\right)\phi_\mu(1)\phi_\nu(2)\cdots\phi_\omega(N)\right] \\
&= [\mathsf{u}\phi_\mu, \phi_\nu, \cdots, \phi_\omega] + [\phi_\mu, \mathsf{u}\phi_\nu, \cdots, \phi_\omega] + \cdots + [\phi_\mu, \phi_\nu, \cdots, \mathsf{u}\phi_\omega]
\end{aligned} \tag{7.32}$$

Because U is symmetric in the particle numbering, it is invariant under permutation, and for that reason could be commuted through the antisymmetrizer. Then, because it is a sum of one-particle contributions, it produces the sum of determinants shown in the last line of Eq. (7.32). Next, we insert expansions for $\mathsf{u}\phi_\mu$, $\mathsf{u}\phi_\nu, \cdots, \mathsf{u}\phi_\omega$ from Eq. (7.31), obtaining

$$\begin{aligned}
\mathsf{U}[\phi_\mu, \phi_\nu, \cdots, \phi_\omega] &= \sum_\lambda < \phi_\lambda|\mathsf{u}|\phi_\mu > [\phi_\lambda, \phi_\nu, \cdots, \phi_\omega] \\
&+ \sum_\lambda < \phi_\lambda|\mathsf{u}|\phi_\nu > [\phi_\mu, \phi_\lambda, \cdots, \phi_\omega] + \cdots \\
&+ \sum_\lambda < \phi_\lambda|\mathsf{u}|\phi_\omega > [\phi_\mu, \phi_\nu, \cdots, \phi_\lambda]
\end{aligned} \tag{7.33}$$

Eq. (7.33) is simpified by using the result from the preceding section that the operator $\mathsf{a}_\lambda^\dagger\mathsf{a}_\mu$ causes direct replacement of ϕ_μ by ϕ_λ in a determinantal function, irrespective of the

location of ϕ_μ therein. Thus

$$[\phi_\lambda, \phi_\nu, \cdots, \phi_\omega] = \mathbf{a}_\lambda^\dagger \mathbf{a}_\mu [\phi_\mu, \phi_\nu, \cdots, \phi_\omega],$$

$$[\phi_\mu, \phi_\lambda, \cdots, \phi_\omega] = \mathbf{a}_\lambda^\dagger \mathbf{a}_\nu [\phi_\mu, \phi_\nu, \cdots, \phi_\omega], \cdots,$$

$$[\phi_\mu, \phi_\nu, \cdots, \phi_\lambda] = \mathbf{a}_\lambda^\dagger \mathbf{a}_\omega [\phi_\mu, \phi_\nu, \cdots, \phi_\omega],$$

so

$$\mathsf{U}[\phi_\mu, \phi_\nu, \cdots, \phi_\omega] = \Bigg[\sum_\lambda < \phi_\lambda | \mathsf{u} | \phi_\mu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\mu + \sum_\lambda < \phi_\lambda | \mathsf{u} | \phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\nu + \cdots$$

$$+ \sum_\lambda < \phi_\lambda | \mathsf{u} | \phi_\omega > \mathbf{a}_\lambda^\dagger \mathbf{a}_\omega \Bigg] [\phi_\mu, \phi_\nu, \cdots, \phi_\omega] \tag{7.34}$$

The operator on the right-hand side of Eq. (7.34) is of the form $\sum_{\lambda,\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma$, with the λ sum unrestricted but with σ summed only over the occupied states of [ ]. However, the σ summation can be extended to cover all one-particle states without error, as the added terms have a vanishing effect on [ ]. This extension has the virtue that the right-hand side operator then has a form independent of the occupancy of [ ], and let us write

$$\mathsf{U}[\phi_\mu, \phi_\nu, \cdots, \phi_\omega] = \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma [\phi_\mu, \phi_\nu, \cdots, \phi_\omega] \tag{7.35}$$

In line with the immediately preceding discussion, both sums are now unrestricted. Since the determinantal function [ ] is arbitrary, Eq. (7.35) is equivalent to the operator equation

$$\mathsf{U} = \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma \tag{7.36}$$

An important illustration of Eq. (7.36) is its application to the reference function Φ. Using the convention of Section VII.C we have

$$\mathsf{U}\Phi = \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma \Phi = \sum_\alpha < \phi_\alpha | \mathsf{u} | \phi_\alpha > \Phi + \sum_{r\alpha} < \phi_r | \mathsf{u} | \phi_\alpha > \Phi_\alpha^r \tag{7.37}$$

In obtaining Eq. (7.37), we used the fact that the σ summation from Eq. (7.36) yields nonzero results only if σ referes to a *hole* state (occupied in Φ). The λ summation then contributes both for the value assigned to σ and for all *particle* states.

In general, as Eqs. (7.36) and (7.37) illustrate, a one-electron U can convert a determinant into a linear combination of determinants, each differing from the original determinant by

at most one occupied single-particle state. However, in the special case that the ϕ_μ are eigenfunctions of u, with $\mathsf{u}\,\phi_\mu = \epsilon_\mu \phi_\mu$, the only nonzero terms in Eq. (7.36) are those with $\lambda = \sigma$, and we have

$$\mathsf{U} = \sum_\lambda \epsilon_\lambda \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda \tag{7.38}$$

The operator product $\mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda$ is sometimes called the *number operator* for the one-particle state λ, as

$$\begin{aligned} \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda [\,] &= 1.[\,], \qquad && if\ \phi_\lambda\ occupied\ in\ [\,] \\ &= 0, && otherwise \end{aligned} \tag{7.39}$$

In other words, the eigenvalues of $\mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda$ are the possible numbers of fermions occupying ϕ_λ. Also, because $\mathbf{a}_\lambda \mathbf{a}_\lambda^\dagger = 1 - \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda$, the operator $\mathbf{a}_\lambda \mathbf{a}_\lambda^\dagger$ has eigenvalue unity for an unoccupied state, and zero for an occupied state:

$$\begin{aligned} \mathbf{a}_\lambda \mathbf{a}_\lambda^\dagger [\,] &= 1.[\,], \qquad && if\ \phi_\lambda\ occupied\ in\ [\,] \\ &= 0, && otherwise \end{aligned} \tag{7.40}$$

In view of Eq. (7.39), we see that if U is diagonal [i.e., satisfying Eq. (7.38)], then

$$\mathsf{U}[\,] = \sum_{(\lambda\, occ)} \epsilon_\lambda [\,] \tag{7.41}$$

where the sum is over the one-particle states occupied in $[\,]$. Eq. (7.41) shows that $[\,]$ is an eigenfunction of U, with an eigenvalue that is the sum of the occupied-state eigenvalues of u.

One of the obvious features of Eqs. (7.36) - (7.41) is that their forms are independent of N, the number of particles described. This fact is suggestive of possibilities for convenient formulation of problems in which the number of particles changes (e.g., electron attachment or ionization). If we should somehow become sufficiently confused that we need to ask how many particles are being represented by a given determinant, Eq. (7.39) even indicates an operator that can do that job - the total number operator

$$\mathsf{N} = \sum_\lambda \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda \tag{7.42}$$

Application of N to a determinantal function will multiply it by the number of occupied one-particle states contained therein:

$$\mathsf{N}[\,] = \sum_{\lambda} \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda [\,] = \mathsf{N}[\,] \tag{7.43}$$

We proceed now to a discussion of two-particle operators V of the form

$$\mathsf{V} = \sum_{i<j} \mathsf{v}(i,j) \tag{7.44}$$

where

$$\mathsf{v}(i,j)\phi_\mu(i)\phi_\nu(j) = \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\nu > \phi_\lambda(i)\phi_\sigma(j) \tag{7.45}$$

V is completely symmetric under permutations of the particle numbering, and therefore, like U, may be commuted through the antisymmetrizer:

$$\begin{aligned}
&\mathsf{V}[\phi_\mu, \phi_\nu, \phi_\rho, \cdots, \phi_\tau, \phi_\omega] \\
&= (N!)^{-1/2} \sum_P (-)^P \mathsf{P}\left[\sum_{i<j} \mathsf{v}(i,j)\phi_\mu(1)\phi_\nu(2)\phi_\rho(3)\cdots\phi_\tau(N-1)\phi_\omega(N)\right]
\end{aligned} \tag{7.46}$$

Using Eq. (7.45) and the previously proved result that $\mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu$ causes direct replacement of ϕ_μ by ϕ_λ and ϕ_ν by ϕ_σ,

$$\begin{aligned}
&\mathsf{v}[\phi_\mu, \phi_\nu, \phi_\rho, \cdots, \phi_\tau, \phi_\omega] = \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\nu > [\phi_\lambda, \phi_\sigma, \phi_\rho, \cdots, \phi_\tau, \phi_\omega] \\
&+ \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\rho > [\phi_\lambda, \phi_\nu, \phi_\sigma, \cdots, \phi_\tau, \phi_\omega] \\
&+ \cdots \\
&+ \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\tau\phi_\omega > [\phi_\mu, \phi_\nu, \phi_\rho, \cdots, \phi_\lambda, \phi_\sigma] \\
&= \Bigg[\sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu \\
&+ \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\rho > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\rho \mathbf{a}_\mu \\
&+ \cdots + \sum_{\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\tau\phi_\omega > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\omega \mathbf{a}_\tau\Bigg] \\
&\times [\phi_\mu, \phi_\nu, \phi_\rho, \cdots, \phi_\tau, \phi_\omega]
\end{aligned} \tag{7.47}$$

The operator on the right-hand side of Eq. (7.47) is of the form $\sum_{\mu\nu\lambda\sigma} < \phi_\lambda\phi_\sigma|\mathsf{v}|\phi_\mu\phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu$, with λ and σ unrestricted, but with μ and ν restricted to occupied states of $[\,]$

with ϕ_μ therein preceding ϕ_ν. The restriction that ϕ_μ precedes ϕ_ν can be removed if we insert a factor "$\frac{1}{2}$", and the restriction to occupied states can, as for U, be dropped without error. We thus have

$$\mathsf{V} = \frac{1}{2} \sum_{\mu\nu\lambda\sigma} < \phi_\lambda \phi_\sigma | \mathsf{v} | \phi_\mu \phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu \tag{7.48}$$

Eq. (7.48) is consistent with our expectation that "direct" and "exchange" contributions should have opposite signs. For example, look at two terms of V containing the same indices, but in a different order:

$$\mathsf{V} = \cdots + \frac{1}{2} < \phi_\lambda \phi_\sigma | \mathsf{v} | \phi_\mu \phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu + \frac{1}{2} < \phi_\lambda \phi_\sigma | \mathsf{v} | \phi_\nu \phi_\mu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\mu \mathbf{a}_\nu + \cdots \tag{7.49}$$

If we interchange the order of $\mathbf{a}_\mu$ and $\mathbf{a}_\nu$ in the second explicit term and introduce a corresponding minus sign, we have

$$\mathsf{V} = \cdots + \frac{1}{2} < \phi_\lambda \phi_\sigma | \mathsf{v} | \phi_\mu \phi_\nu - \phi_\nu \phi_\mu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu + \cdots \tag{7.50}$$

Then, defining

$$< \phi_\lambda \phi_\sigma | \tilde{\mathsf{v}} | \phi_\mu \phi_\nu > = < \phi_\lambda \phi_\sigma | \mathsf{v} | \phi_\mu \phi_\nu - \phi_\nu \phi_\mu > \tag{7.51}$$

we may write

$$\mathsf{V} = \frac{1}{4} \sum_{\mu\nu\lambda\sigma} < \phi_\lambda \phi_\sigma | \tilde{\mathsf{v}} | \phi_\mu \phi_\nu > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu \tag{7.52}$$

We illustrate the use of Eq. (7.52) by calculating $\mathsf{V}\Phi$. Nonzero contributions only arise if ν and μ refer to *holes*, λ and σ may refer either to *particles* or to the holes created by $\mathbf{a}_\nu$ or $\mathbf{a}_\mu$. Using the conventions of Section VII.C, we have

$$\begin{aligned} \mathsf{V}\Phi = & \sum_{(\alpha\beta)} < \phi_\alpha \phi_\beta | \tilde{\mathsf{v}} | \phi_\alpha \phi_\beta > \Phi + \sum_{r\alpha\beta} < \phi_r \phi_\beta | \tilde{\mathsf{v}} | \phi_\alpha \phi_\beta > \Phi_\alpha^r \\ & + \sum_{(rs)(\alpha\beta)} < \phi_r \phi_s | \tilde{\mathsf{v}} | \phi_\alpha \phi_\beta > \Phi_{\alpha\beta}^{rs} \end{aligned} \tag{7.53}$$

The factor "$\frac{1}{4}$" in Eq. (7.52) has been cancelled by the elimination of identical terms and the restriction of index sums to inequivalent sets.

E. Example-Hartree-Fock State

As an initial example of the use of the occupation-numner formalism, let us consider the equations of the Hartree-Fock method. We take the basis one-particle states to be the eigenstates of the Hartree-Fock Hamiltonian $\mathsf{H}_0 = \mathsf{U} + \mathsf{U}_{HF}$, with $\phi_\alpha, \phi_\beta, \cdots$ representing occupied single-particle states and $\phi_r, \phi_s, \cdots$ representing unoccupied single-particle states. We take the reference function Φ to be the Hartree-Fock wavefunction:

$$\Phi = \mathbf{a}_\alpha^\dagger \mathbf{a}_\beta^\dagger \cdots |0> \tag{7.54}$$

where $\alpha, \beta, \cdots$ are in lexical order. In accord with Eq. (7.38), the occupation-number representation of H_0 is

$$\mathsf{H}_0 = \sum_\lambda \epsilon_\lambda \mathbf{a}_\lambda^\dagger \mathbf{a}_\lambda \tag{7.55}$$

where the sum over λ is unrestricted. If we apply H_0 to Φ, the only non-vanishing terms will be those for which λ describes an occupied single-particle state. Each such term is simply $\epsilon_\lambda \Phi$ and we therefore get

$$\mathsf{H}_0 \Phi = \left(\sum_\alpha \epsilon_\alpha \right) \Phi \tag{7.56}$$

where the sum is now restricted to the occupied single-particle states.

The terms comprising H_0, namely U and U_{HF}, may also be written in occupation-number language. From Eq. (7.36)

$$\mathsf{U} = \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma \tag{7.57}$$

and using Eq. (7.36) again, this time in conjunction with the equation

$$< \phi_\mu | \mathsf{u}_{HF} | \phi_\nu > = \sum_\beta (< \phi_\mu \phi_\beta | \mathsf{v} | \phi_\nu \phi_\beta > - < \phi_\mu \phi_\beta | \mathsf{v} | \phi_\beta \phi_\nu >) \tag{7.58}$$

for the matrix elements of u_{HF},

$$\mathsf{U}_{HF} = \sum_{\lambda\sigma} \sum_\beta < \phi_\lambda \phi_\beta | \tilde{\mathsf{v}} | \phi_\sigma \phi_\beta > \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma \tag{7.59}$$

If we now form the expectation values $< \Phi|\mathsf{U}|\Phi >$ and $< \Phi|\mathsf{U}_{HF}|\Phi >$, we can see that we get nonvanishing contributions only if σ refers to a single-particle state occupied in Φ, and if $\mathbf{a}^\dagger_\lambda$ restores the state annihilated by $\mathbf{a}_\sigma$. Thus,

$$< \mathsf{U} >= \sum_\alpha < \phi_\alpha|\mathsf{u}|\phi_\alpha > \tag{7.60}$$

$$< \mathsf{U}_{HF} >= \sum_\alpha \sum_\beta < \phi_\alpha\phi_\beta|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta > \tag{7.61}$$

The two-particle part of H may be handled analogously. Starting from Eq. (7.52),

$$\mathsf{V} = \frac{1}{4}\sum_{\mu\nu\lambda\sigma} < \phi_\lambda\phi_\sigma|\tilde{\mathsf{v}}|\phi_\mu\phi_\nu > \mathbf{a}^\dagger_\lambda\mathbf{a}^\dagger_\sigma\mathbf{a}_\nu\mathbf{a}_\mu \tag{7.62}$$

We note that the $< \Phi|\mathsf{V}|\Phi >$ contains nonvanishing terms only when μ and ν refer to occupied single-particle states, and then only when $\mathbf{a}^\dagger_\lambda$ and $\mathbf{a}^\dagger_\sigma$ restore the same states as were annihilated by $\mathbf{a}_\mu$ and $\mathbf{a}_\nu$. There are two ways in which this can happen: either $\lambda = \mu$ and $\sigma = \nu$, or $\sigma = \mu$ and $\lambda = \nu$. In the first alternative, the operator sequence $\mathbf{a}^\dagger_\mu\mathbf{a}^\dagger_\nu\mathbf{a}_\nu\mathbf{a}_\mu$ restores the original wavefunction; in the second, the sequence $\mathbf{a}^\dagger_\nu\mathbf{a}^\dagger_\mu\mathbf{a}_\nu\mathbf{a}_\mu$ produces a final result differing from Φ by the interchange in position of ϕ_μ and ϕ_ν (i.e., by a minus sign). Thus,

$$< \mathsf{V} >= \frac{1}{4}\sum_{\alpha\beta}[< \phi_\alpha\phi_\beta|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta > - < \phi_\beta\phi_\alpha|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta >] \tag{7.63}$$

Referring to the definition of $\tilde{\mathsf{v}}$ given in Eq. (7.51), we see that
$< \phi_\beta\phi_\alpha|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta >= - < \phi_\alpha\phi_\beta|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta >$, so that Eq. (7.62) reduces to

$$< \mathsf{V} >= \frac{1}{2}\sum_{\alpha\beta} < \phi_\alpha\phi_\beta|\tilde{\mathsf{v}}|\phi_\alpha\phi_\beta > \tag{7.64}$$

Remembering now that $\mathsf{H} = \mathsf{H}_0 + \mathsf{H}'$, with $\mathsf{H}' = \mathsf{V} - \mathsf{U}_{HF}$, Eqs. (7.56), (7.61), and (7.64) yield the expected results:

$$\begin{aligned} < \mathsf{H} >&= E_0 + < \mathsf{V} > - < \mathsf{U}_{HF} > \\ &= E_0 - < \mathsf{V} > \end{aligned} \tag{7.65}$$

with

$$E_0 = \sum_\alpha \epsilon_\alpha \tag{7.66}$$

F. Matrix Elements Between Determinantal Wavefunctions

As a second and most important example of the use of the occupation-number formalism, we consider next the evaluation of matrix elements of various operators between determinantal wavefunctions. As in Section VII.C, we use Φ to denote a reference determinant and $\Phi^{rs\cdots}_{\alpha\beta\cdots}$ to denote a determinant in which one-particle states α, $\beta, \cdots$ (collectively denoted Δ) have been respectively replaced by states r, $s, \cdots$ (collectively denoted R). Letting R' and Δ' stand for index sets as well, the matrix elements of concern here are of the general forms $< \Phi^R_\Delta|\mathsf{U}|\Phi^{R'}_{\Delta'} >$ and $< \Phi^R_\Delta|\mathsf{V}|\Phi^{R'}_{\Delta'} >$, where U and V are one- and two-particle operators, respectively. We will also need the overlap integral $< \Phi^R_\Delta|\Phi^{R'}_{\Delta'} >$.

Considering first the overlap integral, we note that it must vanish unless the index sets R and R' contain the same individual indices, although not necessarily in the same order. For a nonvanishing result, it is also necessary that Δ and Δ' contain the same indices. For R, R', Δ, Δ' satisfying these conditions, we write

$$\begin{aligned} &< \Phi^R_\Delta|\Phi^{R'}_{\Delta'} >=< \mathsf{A}^R_\Delta\Phi|\mathsf{A}^{R'}_{\Delta'}\Phi > \\ &=< \Phi|(\mathsf{A}^R_\Delta)^\dagger\mathsf{A}^{R'}_{\Delta'}|\Phi > \\ &=< \Phi|\mathbf{a}^\dagger_\alpha\mathbf{a}^\dagger_\beta\cdots\mathbf{a}_s\mathbf{a}_r\mathbf{a}^\dagger_{r'}\mathbf{a}^\dagger_{s'}\cdots\mathbf{a}_{\beta'}\mathbf{a}_{\alpha'}|\Phi > \end{aligned} \tag{7.67}$$

after which we permute the primed-index operators to the unprimed orderings r, $s, \cdots$ and $\cdots, \beta$, α. This permutation introduces a factor $(-)^P$ that is determined by its parity, and the operator string thereby produced leaves Φ unchanged. Specifically, we have

$$\begin{aligned} &< \Phi^R_\Delta|\Phi^{R'}_{\Delta'} >= (-)^P < \Phi|\mathbf{a}^\dagger_\alpha\mathbf{a}^\dagger_\beta\cdots\mathbf{a}_s\mathbf{a}_r\mathbf{a}^\dagger_r\mathbf{a}^\dagger_s\cdots\mathbf{a}_\beta\mathbf{a}_\alpha|\Phi > \\ &= (-)^P < \Phi|(\mathbf{a}_r\mathbf{a}^\dagger_r)(\mathbf{a}_s\mathbf{a}^\dagger_s)\cdots(\mathbf{a}^\dagger_\beta\mathbf{a}_\beta)(\mathbf{a}^\dagger_\alpha\mathbf{a}_\alpha)|\Phi > \end{aligned} \tag{7.68}$$

where the operator permutation carrying the first line of Eq. (7.68) into the second is even and does not produce a sign change. As indicated in Eqs. (7.39) and (7.40), each parenthesized pair of operators multiplies Φ by unity, so all may be dropped. Thus,

$$< \Phi^R_\Delta|\Phi^{R'}_{\Delta'} >= (-)^P < \Phi|\Phi >= (-)^P \tag{7.69}$$

where P is the permutation bringing R', Δ' to the same order as R, Δ. We will encounter this overlap integral so often that we condense its definition to

$$\left\langle \frac{R}{R'} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \tag{7.70}$$

This symbol has the value of $+1$ (-1) if R, $\Delta \to R', \Delta'$ can be produced by an even (odd) permutation; it is equal to 0 if R, Δ and R', Δ' differ by more than a permutation. It is also defined so as to apply (with value zero) even when the strings R and R' (or Δ and Δ') are of different lengths.

Next, we proceed to matrix elements of one-particle operators. As pointed out in the discussion after Eq. (7.37), a one-particle operator can change the single-particle state of at most one fermion, so that a matrix element $< \Phi_\Delta^R | \mathsf{U} | \Phi_{\Delta'}^{R'} >$ must vanish unless $\Phi_{\Delta'}^{R'}$ and Φ_Δ^R differ by at most one single-particle state. We may distinguish four cases, of which the first is the possibility that R, Δ and R', Δ' represent the same occupancy. For this case, permuting the operator string $\mathsf{A}_{\Delta'}^{R'}$ to the ordering A_Δ^R, we have

$$\begin{aligned} < \Phi_\Delta^R | \mathsf{U} | \Phi_{\Delta'}^{R'} > &= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \Phi_\Delta^R | \mathsf{U} | \Phi_\Delta^R > \\ &= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > < \Phi_\Delta^R | \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma | \Phi_\Delta^R > \end{aligned} \tag{7.71}$$

where we have inserted Eq. (7.36) for U. The final matrix element of Eq. (7.71) vanishes unless $\sigma = \lambda$, with λ occupied in Φ_Δ^R. When these conditions are met the matrix element has the value unity, and

$$< \Phi_\Delta^R | \mathsf{U} | \Phi_{\Delta'}^{R'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \left[\sum_{r \in R} < \phi_r | \mathsf{u} | \phi_r > + \sum_{\alpha \tilde{\in} \Delta} < \phi_\alpha | \mathsf{u} | \phi_\alpha > \right] \tag{7.72}$$

Remembering our convention that Greek indices $\alpha, \beta, \cdots$ denote occupied states in Φ, we see that the last sum in Eq. (7.72) is over the occupied states of Φ that are *not* listed in Δ.

In the next case for one-particle-operator matrix elements, the two determinantal functions differ by one orbital not occupied in Φ. Denoting these wavefunctions Φ_Δ^{rR} and $\Phi_{\Delta'}^{sR'}$, where $r \neq s$ and R, Δ and R', Δ' differ only by a permutation, and using the relationship

$\Phi_\Delta^{rR} = \mathbf{a}_r^\dagger \Phi_\Delta^R$,

$$< \Phi_\Delta^{rR} | \mathsf{U} | \Phi_{\Delta'}^{sR'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \Phi_\Delta^{rR} | \mathsf{U} | \Phi_\Delta^{sR} >$$

$$= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > < \Phi_\Delta^R | \mathbf{a}_r \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma \mathbf{a}_s^\dagger | \Phi_\Delta^R > \tag{7.73}$$

The only nonzero term of Eq. (7.73) is that with $\lambda = r$ and $\sigma = s$, and Eq. (7.73) reduces to

$$< \Phi_\Delta^{rR} | \mathsf{U} | \Phi_{\Delta'}^{sR'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r | \mathsf{u} | \phi_s > \tag{7.74}$$

A third one-particle-operator case is of the form $< \Phi_{\alpha\Delta}^R | \mathsf{U} | \Phi_{\beta\Delta'}^{R'} >$. Because $\Phi_{\alpha\Delta}^R = \mathsf{A}_\Delta^R \mathbf{a}_\alpha \Phi$, and A_Δ^R is a string containing an odd number of operators, we have $\mathsf{A}_\Delta^R \mathbf{a}_\alpha = -\mathbf{a}_\alpha \mathsf{A}_\Delta^R$, so $\Phi_{\alpha\Delta}^R = -\mathbf{a}_\alpha \Phi_\Delta^R$, and the equation analogous to Eq. (7.73) is

$$< \Phi_{\alpha\Delta}^R | \mathsf{U} | \Phi_{\beta\Delta'}^{R'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{u} | \phi_\sigma > < \Phi_\Delta^R | (-\mathbf{a}_\alpha^\dagger) \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma (-\mathbf{a}_\beta) | \Phi_\Delta^R > \tag{7.75}$$

Since by hypothesis $\alpha \neq \beta$, the only nonzero term of Eq. (7.75) occurs when $\sigma = \alpha$ and $\lambda = \beta$; because $< \Phi_\Delta^R | (-\mathbf{a}_\alpha^\dagger) \mathbf{a}_\beta^\dagger \mathbf{a}_\alpha (-\mathbf{a}_\beta) | \Phi_\Delta^R > = -1$

$$< \Phi_{\alpha\Delta}^R | \mathsf{U} | \Phi_{\beta\Delta'}^{R'} > = -\left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_\beta | \mathsf{u} | \phi_\alpha > \tag{7.76}$$

The fourth one-particle-operator case is $< \Phi_{\alpha\Delta}^{rR} | \mathsf{U} | \Phi_{\Delta'}^{R'} >$, where R, Δ and R', Δ' are related by permutation. Using $\Phi_{\alpha\Delta}^{rR} = +\mathbf{a}_r^\dagger \mathbf{a}_\alpha \Phi_\Delta^R$, we obtain by the methods already illustrated

$$< \Phi_{\alpha\Delta}^{rR} | \mathsf{U} | \Phi_{\Delta'}^{R'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r | \mathsf{u} | \phi_\alpha > \tag{7.77}$$

Analogous methods may be applied for two-particle operators. We use Eq. (7.52) for the two-particle operator V, noting that the summands of that equation are symmetric under index interchanges $\mu \leftrightarrow \nu$ and $\lambda \leftrightarrow \sigma$. We illustrate for one case:

$$< \Phi_\Delta^{rsR} | \mathsf{V} | \Phi_{\Delta'}^{uvR'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \Phi_\Delta^R | \mathbf{a}_s \mathbf{a}_r \mathsf{V} \mathbf{a}_u^\dagger \mathbf{a}_v^\dagger | \Phi_\Delta^R >$$

$$= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \frac{1}{4} \sum_{\mu\nu\lambda\sigma} < \phi_\lambda \phi_\sigma | \tilde{\mathsf{v}} | \phi_\mu \phi_\nu > < \Phi_\Delta^R | \mathbf{a}_s \mathbf{a}_r \mathbf{a}_\lambda^\dagger \mathbf{a}_\sigma^\dagger \mathbf{a}_\nu \mathbf{a}_\mu \mathbf{a}_u^\dagger \mathbf{a}_v^\dagger | \Phi_\Delta^R > \tag{7.78}$$

Nonzero terms result only if indices μ and ν are equal to u and v (in either order), and if λ and σ are equal to r and s (in either order). There are four index choices that satisfy

these conditions; each makes the same contribution to Eq. (7.78). We therefore take any one index choice, multiply by 4, and thereby cancel the factor "$\frac{1}{4}$" that came as part of V. We choose $\lambda = r$, $\sigma = s$, $\mu = u$, $\nu = v$; the final matrix element of Eq. (7.78) is then $+1$, and

$$< \Phi_{\Delta}^{rsR} | \mathsf{V} | \Phi_{\Delta'}^{uvR'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_s | \tilde{\mathsf{v}} | \phi_u \phi_v > \tag{7.79}$$

Since the matrix element formulas are useful in discussions of some rather sophisticated quantum formulations of many-particle systems, such as the configuration-interaction method, many-body diagrammatic perturbation theory, coupled-cluster method, we recapitulate those already derived and add the others needed to provide a complete collection. The only matrix elements omitted from the following lists are those that can be obtained by transposition of the formulas given.

One-particle operator:

$$< \Phi_{\Delta}^{R} | \mathsf{U} | \Phi_{\Delta'}^{R'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \left[\sum_{r \in R} < \phi_r | \mathsf{u} | \phi_r > + \sum_{\alpha \tilde{\in} \Delta} < \phi_\alpha | \mathsf{u} | \phi_\alpha > \right] \tag{7.80}$$

$$< \Phi_{\Delta}^{rR} | \mathsf{U} | \Phi_{\Delta'}^{sR'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r | \mathsf{u} | \phi_s > \tag{7.81}$$

$$< \Phi_{\alpha\Delta}^{R} | \mathsf{U} | \Phi_{\beta\Delta'}^{R'} >= -\left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_\beta | \mathsf{u} | \phi_\alpha > \tag{7.82}$$

$$< \Phi_{\alpha\Delta}^{rR} | \mathsf{U} | \Phi_{\Delta'}^{R'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r | \mathsf{u} | \phi_\alpha > \tag{7.83}$$

Two-particle operator:

$$\begin{aligned} < \Phi_{\Delta}^{R} | \mathsf{V} | \Phi_{\Delta'}^{R'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \Bigg[& \frac{1}{2} \sum_{rs;r,s \in R} < \phi_r \phi_s | \tilde{\mathsf{v}} | \phi_r \phi_s > \\ & + \sum_{r \in R} \sum_{\alpha \tilde{\in} \Delta} < \phi_r \phi_\alpha | \tilde{\mathsf{v}} | \phi_r \phi_\alpha > + \frac{1}{2} \sum_{\alpha\beta;\alpha,\beta \tilde{\in} \Delta} < \phi_\alpha \phi_\beta | \tilde{\mathsf{v}} | \phi_\alpha \phi_\beta > \Bigg] \end{aligned} \tag{7.84}$$

$$< \Phi_{\Delta}^{rR} | \mathsf{V} | \Phi_{\Delta'}^{sR'} >= \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \left[\sum_{u \in R} < \phi_r \phi_u | \tilde{\mathsf{v}} | \phi_s \phi_u > + \sum_{\alpha \tilde{\in} \Delta} < \phi_r \phi_\alpha | \tilde{\mathsf{v}} | \phi_s \phi_\alpha > \right] \tag{7.85}$$

$$< \Phi^R_{\alpha\Delta} | \mathsf{V} | \Phi^{R'}_{\beta\Delta'} >$$
$$= -\left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \left[\sum_{r\in R} < \phi_\beta \phi_r | \tilde{\mathbf{v}} | \phi_\alpha \phi_r > + \sum_{\gamma\tilde{\in}\Delta} < \phi_\beta \phi_\gamma | \tilde{\mathbf{v}} | \phi_\alpha \phi_\gamma > \right] \tag{7.86}$$

$$< \Phi^{rR}_{\alpha\Delta} | \mathsf{V} | \Phi^{R'}_{\Delta'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle \left[\sum_{s\in R} < \phi_r \phi_s | \tilde{\mathbf{v}} | \phi_\alpha \phi_s > + \sum_{\beta\tilde{\in}\Delta} < \phi_r \phi_\beta | \tilde{\mathbf{v}} | \phi_\alpha \phi_\beta > \right] \tag{7.87}$$

$$< \Phi^{rsR}_{\Delta} | \mathsf{V} | \Phi^{uvR'}_{\Delta'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_s | \tilde{\mathbf{v}} | \phi_u \phi_v > \tag{7.88}$$

$$< \Phi^{R}_{\alpha\beta\Delta} | \mathsf{V} | \Phi^{R'}_{\gamma\delta\Delta'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_\gamma \phi_\delta | \tilde{\mathbf{v}} | \phi_\alpha \phi_\beta > \tag{7.89}$$

$$< \Phi^{rR}_{\alpha\Delta} | \mathsf{V} | \Phi^{sR'}_{\beta\Delta'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_\beta | \tilde{\mathbf{v}} | \phi_s \phi_\alpha > \tag{7.90}$$

$$< \Phi^{rsR}_{\alpha\Delta} | \mathsf{V} | \Phi^{uR'}_{\Delta'} > = -\left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_s | \tilde{\mathbf{v}} | \phi_u \phi_\alpha > \tag{7.91}$$

$$< \Phi^{rR}_{\alpha\beta\Delta} | \mathsf{V} | \Phi^{R'}_{\gamma\Delta'} > = -\left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_\gamma | \tilde{\mathbf{v}} | \phi_\alpha \phi_\beta > \tag{7.92}$$

$$< \Phi^{rsR}_{\alpha\beta\Delta} | \mathsf{V} | \Phi^{R'}_{\Delta'} > = \left\langle \frac{R}{\Delta} \;\middle|\; \frac{R'}{\Delta'} \right\rangle < \phi_r \phi_s | \tilde{\mathbf{v}} | \phi_\alpha \phi_\beta > \tag{7.93}$$

G. Density Matrix Elements

The matrix elements of reduced density operators can be cast in a form that permits their evaluation by the methods outlined in the preceding section. To find such a formulation, we start by writing an N-particle wavefunction $[\phi_\mu, \phi_\nu, \phi_\lambda, \cdots]$ in a way that explicates its dependence on the coordinates $\mathbf{x}_1$. Taking account of the N-dependent factors implicit in the definition of [ ], we have

$$[\phi_\mu, \phi_\nu, \phi_\lambda, \cdots] = \left[\frac{(N-1)!}{N!} \right]^{1/2} (\phi_\mu(\mathbf{x_1})[\phi_\nu, \phi_\lambda, \cdots]$$
$$-\phi_\nu(\mathbf{x_1})[\phi_\mu, \phi_\lambda, \cdots] + \phi_\lambda(\mathbf{x_1})[\phi_\mu, \phi_\nu, \cdots] - \cdots) \tag{7.94}$$

In Eq. (7.94), the [ ] on the left has arguments $\mathbf{x}_1, \cdots, \mathbf{x}_n$, while the arguments of the [ ] on the right are $\mathbf{x}_2, \cdots, \mathbf{x}_n$. If we assume ϕ_μ, $\phi_\nu, \cdots$ are in lexical order, Eq. (7.94) can be written

$$[\phi_\mu, \phi_\nu, \phi_\lambda, \cdots] = \left[\frac{(N-1)!}{N!} \right]^{1/2} \sum_\sigma \phi_\sigma(\mathbf{x}_1) \mathbf{a}_\sigma [\phi_\mu, \phi_\nu, \phi_\lambda, \cdots] \tag{7.95}$$

The alternating signs in Eq. (7.94) are just needed for use of $\mathbf{a}_\sigma$, and the arguments of $\mathbf{a}_\sigma[\,]$ are $\mathbf{x}_2, \cdots, \mathbf{x}_n$. The σ sum may be regarded as unrestricted, as the σ values not represented in $[\phi_\mu, \phi_\nu, \phi_\lambda, \cdots]$ lead to vanishing contributions.

Since the form of the operator in Eq. (7.95) is independent of the wavefunction to which it is applied, we consider applying it to a general wafefunction Ψ consisting of an arbitrary linear combination of lexically ordered $[\,]$. We get

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) = \left[\frac{(N-1)!}{N!}\right]^{1/2} \sum_\mu \phi_\mu(\mathbf{x}_1)\mathbf{a}_\mu\Psi \tag{7.96}$$

where the arguments of $\mathbf{a}_\mu\Psi$ are $\mathbf{x}_2, \cdots, \mathbf{x}_n$. It is obvious, then, that

$$\Psi(\mathbf{x}_1', \mathbf{x}_2, \cdots, \mathbf{x}_n) = \left[\frac{(N-1)!}{N!}\right]^{1/2} \sum_\nu \phi_\nu(\mathbf{x}_1')\mathbf{a}_\nu\Psi \tag{7.97}$$

The first-order density matrix of a unit-normalized N-particle wavefunction Ψ, denoted $\gamma(\mathbf{x}_1, \mathbf{x}_1')$, is

$$\gamma(\mathbf{x}_1, \mathbf{x}_1') = N \int d\mathbf{x}_2 \cdots d\mathbf{x}_n \Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)\Psi^*(\mathbf{x}_1', \mathbf{x}_2, \cdots, \mathbf{x}_n) \tag{7.98}$$

Introducing Eqs. (7.96) and (7.97) into the definition (7.98) we therefore obtain

$$\begin{aligned}\gamma(\mathbf{x}_1, \mathbf{x}_1') &= N\left[\frac{(N-1)!}{N!}\right] \sum_{\mu\nu} \phi_\mu(\mathbf{x}_1)\phi_\nu^*(\mathbf{x}_1') < \mathbf{a}_\nu\Psi|\mathbf{a}_\mu\Psi > \\ &= \sum_{\mu\nu} \phi_\mu(\mathbf{x}_1)\phi_\nu^*(\mathbf{x}_1') < \Psi|\mathbf{a}_\nu^\dagger\mathbf{a}_\mu|\Psi >\end{aligned} \tag{7.99}$$

Comparing Eq. (7.6) with the expansion of $\gamma(\mathbf{x}_1, \mathbf{x}_1')$

$$\gamma(\mathbf{x}_1, \mathbf{x}_1') = \sum_{\mu\nu} |\phi_\mu(\mathbf{x}_1) > \gamma_{\mu\nu} < \phi_\nu(\mathbf{x}_1')| \tag{7.100}$$

with

$$\gamma_{\mu\nu} = \int d\mathbf{x}_1 d\mathbf{x}_1' \phi_\mu^*(\mathbf{x}_1)\gamma(\mathbf{x}_1, \mathbf{x}_1')\phi_\nu(\mathbf{x}_1') \tag{7.101}$$

we see that

$$\gamma_{\mu\nu} =< \Psi|\mathbf{a}_\nu^\dagger\mathbf{a}_\mu|\Psi > \tag{7.102}$$

The formula for the second-order density matrix corresponding to Eq. (7.102) is

$$\Gamma_{\mu\nu,\lambda\sigma} = \frac{1}{2} < \Psi|\mathbf{a}_\lambda^\dagger\mathbf{a}_\sigma^\dagger\mathbf{a}_\nu\mathbf{a}_\mu|\Psi > \tag{7.103}$$

H. Basis Transformation

In this section we consider the way in which expressions involving annihilation or creation operators are affected by transformations of the single-particle basis. Since the formalism we are using is based on orthonormal single-particle states, we limit the discussion to transformations that preserve orthonormality, namely *unitary transformations*. Accordingly, we consider the transformation from an orthonormal set $\{\phi_\mu\}$ to another such set $\{\phi'_\mu\}$, with the two sets related by a transformation operator U:

$$\phi'_\mu = \mathsf{U}\phi_\mu \tag{7.104}$$

The matarix representation of Eq. (8.1) is

$$\phi'_\mu = \sum_\nu U_{\nu\mu}\phi_\nu \tag{7.105}$$

The matrix $\mathbf{U}$ is unitary, so that $(\mathbf{U}^{-1})_{\mu\nu} = (\mathbf{U}^\dagger)_{\mu\nu} = U^*_{\nu\mu}$.

The matrix elements of a one-particle operator B transform in such a way that $\mathbf{B'}$, the matrix of B in the $\{\phi'\}$ basis, is related to the transformed matrix $\mathbf{B}$ by

$$\mathbf{B'} = \mathbf{U}^\dagger\mathbf{B}\mathbf{U} \tag{7.106}$$

$$< \phi'_\mu|\mathsf{B}|\phi'_\nu >= \sum_{\lambda\sigma}(\mathbf{U}^\dagger)_{\mu\lambda} < \phi_\lambda|\mathsf{B}|\phi_\sigma > U_{\sigma\nu} \tag{7.107}$$

This result was obtained by substituting the transformation equations for ϕ'_μ and ϕ'_ν. The corresponding result for two-particle operators is

$$< \phi'_\mu\phi'_\nu|\mathsf{B}|\phi'_\lambda\phi'_\sigma >= \sum_{\rho\tau\eta\zeta}(\mathbf{U}^\dagger)_{\mu\rho}(\mathbf{U}^\dagger)_{\nu\tau} < \phi_\rho\phi_\tau|\mathsf{B}|\phi_\eta\phi_\zeta > U_{\eta\lambda}U_{\zeta\sigma} \tag{7.108}$$

We are also concerned here with the relations between the operators $\mathbf{a}^\dagger_\mu$, $\mathbf{a}_\mu$ in the original basis and their counterparts in the transformed basis, which we shall denote $\mathbf{a}'^\dagger_\mu$, $\mathbf{a}'_\mu$. Since the operator $\mathbf{a}'^\dagger_\mu$ is supposed to add a particle in the state ϕ'_μ, it must be equivalent to a linear combination of additions to the original $\{\phi_\mu\}$ with appropriate coefficients. That is,

$$\mathbf{a}'^\dagger_\mu = \sum_\nu U_{\nu\mu}\mathbf{a}^\dagger_\nu \tag{7.109}$$

Taking the adjoints of both sides of Eq. (7.109), we also have

$$\mathbf{a}'_\mu = \sum_\nu (\mathbf{U}^\dagger)_{\mu\nu} \mathbf{a}_\nu \tag{7.110}$$

Since Eq. (7.110) may seem to some extent nonintuitive, let us digress briefly to verify its consistency. For this purpose, consider the anticommutation relation for the tranformed operators:

$$\mathbf{a}'^\dagger_\mu \mathbf{a}'_\nu + \mathbf{a}'_\nu \mathbf{a}'^\dagger_\mu = \delta_{\mu\nu} \tag{7.111}$$

Our objective is to verify Eq. (7.111) by rewriting it in the original basis. Using Eqs. (7.109) and (7.110), we have

$$\begin{aligned}
&\mathbf{a}'^\dagger_\mu \mathbf{a}'_\nu + \mathbf{a}'_\nu \mathbf{a}'^\dagger_\mu = \sum_{\lambda\sigma} U_{\lambda\mu} (\mathbf{U}^\dagger)_{\nu\sigma} [\mathbf{a}^\dagger_\lambda \mathbf{a}_\sigma + \mathbf{a}_\sigma \mathbf{a}^\dagger_\lambda] \\
&= \sum_{\lambda\sigma} U_{\lambda\mu} (\mathbf{U}^\dagger)_{\nu\sigma} \delta_{\lambda\sigma} \\
&= \sum_\lambda (\mathbf{U}^\dagger)_{\nu\lambda} U_{\lambda\mu} = \sum_\lambda (\mathbf{U}^{-1})_{\nu\lambda} U_{\lambda\mu} = \delta_{\nu\mu}
\end{aligned} \tag{7.112}$$

as required.

We are now ready to verify that the occupation-number expressions for operators are invariant with respect to basis transformation; for a one-particle operator B we start from

$$\mathsf{B} = \sum_{\mu\nu} < \phi'_\mu | \mathsf{b} | \phi'_\nu > \mathbf{a}'^\dagger_\mu \mathbf{a}'_\nu \tag{7.113}$$

Applying Eqs. (8.2), (8.6), and (8.7), we find

$$\begin{aligned}
\mathsf{B} &= \sum_{\mu\nu} \sum_{\lambda\sigma\rho\tau} [(\mathbf{U}^\dagger)_{\mu\lambda} U_{\sigma\nu} < \phi_\lambda | \mathsf{b} | \phi_\sigma >] [U_{\rho\mu} \mathbf{a}^\dagger_\rho] [(\mathbf{U}^\dagger)_{\nu\tau} \mathbf{a}_\tau] \\
&= \sum_{\lambda\sigma\rho\tau} \left[\sum_\mu U_{\rho\mu} (\mathbf{U}^\dagger)_{\mu\lambda} \right] \left[\sum_\nu U_{\sigma\nu} (\mathbf{U}^\dagger)_{\nu\tau} \right] < \phi_\lambda | \mathsf{b} | \phi_\sigma > \mathbf{a}^\dagger_\rho \mathbf{a}_\tau \\
&= \sum_{\lambda\sigma\rho\tau} \delta_{\rho\lambda} \delta_{\sigma\tau} < \phi_\lambda | \mathsf{b} | \phi_\sigma > \mathbf{a}^\dagger_\rho \mathbf{a}_\tau \\
&= \sum_{\lambda\sigma} < \phi_\lambda | \mathsf{b} | \phi_\sigma > \mathbf{a}^\dagger_\lambda \mathbf{a}_\sigma
\end{aligned} \tag{7.114}$$

This is the desired result. A corresponding result can be obtained for two-particle operators.

Given any trial function $\tilde{\mathbf{\Phi}}$, the expectation value $E[\tilde{\mathbf{\Phi}}]$ of the Hamiltonian operator H is a number given by

$$E[\tilde{\mathbf{\Phi}}] = < \tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \tilde{\mathbf{\Phi}} > \tag{8.1}$$

We say that $E[\tilde{\mathbf{\Phi}}]$ is a functional of $\tilde{\mathbf{\Phi}}$ since its value depends on the form of a function, i.e., the function $\tilde{\mathbf{\Phi}}$, rather than any single independent variable. Suppose we vary $\tilde{\mathbf{\Phi}}$ by an arbitrarily small amount, by changing the parameters upon which it depends, for example. That is,

$$\tilde{\mathbf{\Phi}} \rightarrow \tilde{\mathbf{\Phi}} + \delta\,\tilde{\mathbf{\Phi}} \tag{8.2}$$

The energy then becomes

$$\begin{aligned} E[\tilde{\mathbf{\Phi}} + \delta\,\tilde{\mathbf{\Phi}}] &= < \tilde{\mathbf{\Phi}} + \delta\,\tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \tilde{\mathbf{\Phi}} + \delta\,\tilde{\mathbf{\Phi}} > \\ &= E[\tilde{\mathbf{\Phi}}] + \left[< \delta\,\tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \tilde{\mathbf{\Phi}} > + < \tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \delta\,\tilde{\mathbf{\Phi}} > \right] + \cdots \\ &= E[\tilde{\mathbf{\Phi}}] + \delta\,E \end{aligned} \tag{8.3}$$

where $\delta\,E$, which is called the first variation in E, includes all terms that are linear, i.e., first order, in the variation $\delta\,\tilde{\mathbf{\Phi}}$. Notice that we can treat "δ" just like a differential operator, i.e.,

$$\delta\ < \tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \tilde{\mathbf{\Phi}} > = < \delta\,\tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \tilde{\mathbf{\Phi}} > + < \tilde{\mathbf{\Phi}} \,|\, \mathsf{H} \,|\, \delta\,\tilde{\mathbf{\Phi}} >$$

In the variational method, we are looking for that $\tilde{\mathbf{\Phi}}$ for which $E[\tilde{\mathbf{\Phi}}]$ is a minimum. In other words, we wish to find that $\tilde{\mathbf{\Phi}}$ for which the first variation in $E[\tilde{\mathbf{\Phi}}]$ is zero, i.e.,

$$\delta\,E\ =\ 0 \tag{8.4}$$

A. Minimization of the Energy of a Single Determinant

The definition of the ϕ_α in Eq. (5.13) is completed by the requirement that $\mathbf{\Phi}$ be an approximation to $\mathbf{\Psi}$, the exact solution to the Scrödinger equation (1.1), i.e.,

$$\mathsf{H}\,\mathbf{\Psi}\ =\ E\,\mathbf{\Psi} \tag{8.5}$$

Different criteria for the determination of $\mathbf{\Phi}$ lead to different sets of equations for the ϕ_α.

In terms of $\mathbf{\Phi}$, an approximation $< \mathsf{H} >$ to the exact energy E is given by the relation

$$< \mathsf{H} > = < \mathbf{\Phi} \,|\, \mathsf{H} \,|\, \mathbf{\Phi} > \tag{8.6}$$

Inserting the definition of $\mathbf{\Phi}$ from Eq. (5.13) and using the fact that $\mathcal{A}$ is self-adjoint, commutes with H, and is idempotent [cf. Eq. (1.94)], we may write

$$\begin{aligned} < \mathsf{H} > &= N! < \phi_\alpha(1) \cdots \phi_\omega(N) \,|\, \mathsf{H} \,|\, \mathcal{A}\, \phi_\alpha(1) \cdots \phi_\omega(N) > \\ &= \sum_{\mathsf{P}}^{N!} (-)^{\mathsf{P}} < \phi_\alpha(1) \cdots \phi_\omega(N) \,|\, \mathsf{H}\, \mathsf{P} \,|\, \phi_\alpha(1) \cdots \phi_\omega(N) > \end{aligned} \tag{8.7}$$

Now, introducing the definition of H as given in Eqs. (5.1)-(5.3), and taking note of the orthonormality of the ϕ_α, Eq. (8.7) is reduced to the form

$$< \mathsf{H} > = \sum_{\alpha}^{N} < \phi_\alpha \,|\, \mathsf{u} \,|\, \phi_\alpha > + \sum_{\alpha<\beta}^{N} \Big[< \phi_\alpha\, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\alpha\, \phi_\beta > - < \phi_\alpha\, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\beta\, \phi_\alpha > \Big] \tag{8.8}$$

In accordance with our general notational conventions, the α and β summations run over the orbitals included in $\mathbf{\Phi}$. Noticing that if we include the term $\alpha = \beta$ in the last sum, it vanishes, we may remove the restriction $\alpha < \beta$ in favor of a factor $\frac{1}{2}$, giving

$$\begin{aligned} < \mathsf{H} > = & \sum_{\alpha}^{N} < \phi_\alpha \,|\, \mathsf{u} \,|\, \phi_\alpha > \\ & + \frac{1}{2} \sum_{\alpha\beta}^{N} \Big[< \phi_\alpha\, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\alpha\, \phi_\beta > - < \phi_\alpha\, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\beta\, \phi_\alpha > \Big] \end{aligned} \tag{8.9}$$

Given the single determinant

$$|\mathbf{\Phi} > = \, |\, \phi_1\, \phi_2 \cdots \phi_a\, \phi_b \cdots \phi_N >$$

the energy

$$E_0 = < \mathbf{\Phi} \,|\, \mathsf{H} \,|\, \mathbf{\Phi} >$$

is a functional of the spinorbitals $\{\phi_a\}$. To derive the Hartree-Fock equations we need to minimize $E_o[\{\phi_a\}]$ with respect to the spinorbitals, subject to the constraint that the spinorbitals remain orthonormal,

$$< \phi_a \,|\, \phi_b > = \delta_{ab} \tag{8.10}$$

that is, the constraints are of the form

$$< \phi_a \,|\, \phi_b > - \delta_{ab} \;=\; 0 \tag{8.11}$$

We therefore consider the functional $L[\{\phi_a\}]$ of the spinorbitals

$$L[\{\phi_a\}] \;=\; E_o[\{\phi_a\}] - \sum_{a=1}^{N}\sum_{b=1}^{N} \lambda_{ab}\,(< \phi_a \,|\, \phi_b > - \delta_{ab}) \tag{8.12}$$

where $E_o[\{\phi_a\}]$ is the expectation value of the single determinant $|\mathbf{\Phi}>$, written here using a slightly modified notation, in the form

$$E_o[\{\phi_a\}] \;=\; \sum_{a=1}^{N} < a \,|\, \mathsf{u} \,|\, a > + \frac{1}{2}\sum_{a=1}^{N}\sum_{b=1}^{N}\Big(< a\,b \,|\, a\,b > - < a\,b \,|\, b\,a > \Big) \tag{8.13}$$

and the λ_{ab} in Eq. (8.12) constitute a set of Lagrangian multipliers. Minimization of E_o subject to the constraints, is thus obtained by minimizing L, i.e., we now wish to find the conditions which the orbitals $\{\phi_a\}$ must satisfy in order for the functional L to have a minimum value; thus, we require that $\delta\, L \;=\; 0$ for small variations $\delta\,\phi_a$ of these optimum orbitals. We therefore vary the spinorbitals an arbitrary, infinitesimal amount, i.e.,

$$\phi_a \;\rightarrow\; \phi_a + \delta\,\phi_a \tag{8.14}$$

In the following treatment it is convenient to define Coulomb and exchange operators by

$$\mathsf{J}_b(\mu)\,\phi_a(\mu) \;=< \phi_b(\nu) \,|\, r_{\mu\nu}^{-1} \,|\, \phi_b(\nu) > \;\phi_a(\mu) \tag{8.15}$$

$$\mathsf{K}_b(\mu)\,\phi_a(\mu) \;=< \phi_b(\nu) \,|\, r_{\mu\nu}^{-1} \,|\, \phi_a(\nu) > \;\phi_b(\mu) \tag{8.16}$$

Using these definitions, the coulomb and exchange integrals occurring in Eq. (8.13) can be written

$$\mathsf{J}_{ab} \;=< a\,b \,|\, a\,b > =< \phi_a(\mu) \,|\, \mathsf{J}_b(\mu) \,|\, \phi_a(\mu) > =< \phi_b(\nu) \,|\, \mathsf{J}_a(\nu) \,|\, \phi_b(\nu) > \tag{8.17}$$

$$\mathsf{K}_{ab} \;=< a\,b \,|\, b\,a > =< \phi_a(\mu) \,|\, \mathsf{K}_b(\mu) \,|\, \phi_a(\mu) \;=< \phi_b(\nu) \,|\, \mathsf{K}_a(\nu) \,|\, \phi_b(\nu) > \tag{8.18}$$

It is seen that the coulomb operator $\mathsf{J}_b(\mu)$ is just the operator for the potential energy which would arise from an electron distribution in space with a density $\phi_b^*\,\phi_b$. Such operators

correspond to the effective potentials occurring in the Hartree SCF mehtod. The exchange operator, on the other hand, has no classical analogy, since it arises from the antisymmetry principle.

The first-order variation in the functional L is

$$\begin{aligned}\delta L \;=\; & \sum_{a=1}^{N}\Big(<\delta\,\phi_a\,|\,\mathsf{u}\,|\,\phi_a> + <\phi_a\,|\,\mathsf{u}\,|\,\delta\,\phi_a>\Big)\\ & +\frac{1}{2}\sum_{a=1}^{N}\sum_{b=1}^{N}\Big(<\delta\,\phi_a\,|\,\mathsf{J}_b-\mathsf{K}_b\,|\,\phi_a> + <\phi_a\,|\,\mathsf{J}_b-\mathsf{K}_b\,|\,\delta\,\phi_a>\Big)\\ & +\frac{1}{2}\sum_{a=1}^{N}\sum_{b=1}^{N}\Big(<\delta\,\phi_b\,|\,\mathsf{J}_a-\mathsf{K}_a\,|\,\phi_b> + <\phi_b\,|\,\mathsf{J}_a-\mathsf{K}_a\,|\,\delta\,\phi_b>\Big)\\ & -\sum_{ab}^{N}\Big(\lambda_{ab}\;<\delta\,\phi_a\,|\,\phi_b> +\lambda_{ab}\;<\phi_a\,|\,\delta\,\phi_b>\Big)\end{aligned} \tag{8.19}$$

The first and second double summations are symmetric in their indices and lead to the same final sums. Thus, Eq. (8.19) can be written

$$\begin{aligned}\delta L \;=\; & \sum_{a=1}^{N}\left[\left\langle\delta\,\phi_a\right|\mathsf{u}+\sum_{b=1}^{N}(\mathsf{J}_b-\mathsf{K}_b)\left|\,\phi_a\right\rangle\right]\\ & +\sum_{a=1}^{N}\left[\left\langle\phi_a\right|\mathsf{u}+\sum_{b=1}^{N}(\mathsf{J}_b-\mathsf{K}_b)\left|\,\delta\,\phi_a\right\rangle\right]\\ & -\sum_{ab}^{N}\Big(\lambda_{ab}\;<\delta\,\phi_a\,|\,\phi_b> +\lambda_{ab}\;<\phi_a\,|\,\delta\,\phi_b>\Big)\end{aligned} \tag{8.20}$$

Since u, J_b, and K_b are hermitian, we see that the first and second summations are just the adjoints of each other. Furthermore, we can interchange summations indices in the last term of the double summation. Then using the fact that $<\phi_b|\delta\phi_a>$ and $<\delta\phi_a|\phi_b>$ are adjoints of each other, we can write

$$\sum_{ab}^{N}\lambda_{ab}\;<\phi_a\,|\,\delta\,\phi_b> = \sum_{ab}\lambda_{ba}\;<\delta\,\phi_a\,|\,\phi_b>^{*} \tag{8.21}$$

Eq. (8.20) then becomes

$$\begin{aligned}\delta L \;=\; & \sum_{a=1}^{N}\left[\left\langle\delta\,\phi_a\right|\mathsf{u}+\sum_{b=1}^{N}(\mathsf{J}_b-\mathsf{K}_b)\left|\,\phi_a\right\rangle-\sum_{b=1}^{N}\lambda_{ab}\;<\delta\,\phi_a\,|\,\phi_b>\right]\\ & +\sum_{a=1}^{N}\left[\left\langle\delta\,\phi_a\right|\mathsf{u}+\sum_{b=1}^{N}(\mathsf{J}_b-\mathsf{K}_b)\left|\,\phi_a\right\rangle^{*}-\sum_{b=1}^{N}\lambda_{ba}\;<\delta\,\phi_a\,|\,\phi_b>^{*}\right]\end{aligned} \tag{8.22}$$

The vanishing of δL for an arbitrary variation $\delta\phi_a$ is now satisfied by the conditions

$$\left[\mathsf{u} + \sum_{b=1}^{N}(\mathsf{J}_b - \mathsf{K}_b)\right] |\phi_a> = \sum_{b=1}^{N} \lambda_{ab} \,|\, \phi_b > \tag{8.23}$$

$$\left[\mathsf{u} + \sum_{b=1}^{N}(\mathsf{J}_b - \mathsf{K}_b)\right] |\phi_a>^* = \sum_{b=1}^{N} \lambda_{ba} \,|\, \phi_b >^* \tag{8.24}$$

Taking the complex conjugate of Eq. (8.24) and substracting from Eq. (8.23), we obtain

$$\sum_{b=1}^{N}(\lambda_{ab} - \lambda_{ba}^*) \,|\, \phi_b > = \, 0 \tag{8.25}$$

Since the orbitals $\{\phi_b\}$ are linearly independent, it follows that $\lambda_{ab} = \lambda_{ba}^*$, i.e., the lagrangian multipliers are the elements of a hermitian matrix. Thus, Eqs. (8.23) and (8.24) are complex conjugates of each other and are equivalent. These equations are known as the *Hartree-Focck equations*. Defining the Fock operator as

$$\mathsf{F} \;=\; \mathsf{u} + \sum_{b=1}^{N}(\mathsf{J}_b - \mathsf{K}_b) \;=\; \mathsf{u} + \mathsf{u}_{HF} \tag{8.26}$$

the equation for the spinorbitals takes the form

$$\mathsf{F} \,|\, \phi_a > = \sum_{b=1}^{N} \lambda_{ab} \,|\, \phi_b > \qquad\qquad a \;=\; 1, 2, \cdots, N \tag{8.27}$$

This result is perhaps surprising at first sight since it is not in the canonical (standard) eigenvalue form of the equation

$$\mathsf{F} \,|\, \phi_a > = \, \epsilon_a \,|\, \phi_a > \tag{8.28}$$

The reason is that any single determinant wavefunction $\mathbf{\Phi}$ formed from a set of spinorbitals $\{\phi_a\}$ retains a certain degree of flexibility in the spinorbitals; the spinorbitals can be mixed among themselves without changing the expectation values $E_o \;=< \mathbf{\Phi} \,|\, \mathsf{H} \,|\, \mathbf{\Phi} >$. Before obtaining the canonical form of the Hartree-Fock equations, we need to consider unitary transformations of the spinorbitals among themselves.

Starting from Eq. (8.9), we may conveniently demonstrate an important property of $< \mathsf{H} >$, namely that its value is invariant with respect to unitary transformations among the single-particle functions ϕ_a. This invariance means that we cannot attach unique physical significance to the individual ϕ_a, but only to the N-particle space spanned by the set $\{\phi_a\}$ occurring in the many-particle wavefunction $\mathbf{\Phi}$. To prove the invariance, consider a transformation connecting the set $\{\phi_a\}$ to another set of single-particle functions denoted $\{\phi'_a\}$. We write

$$\phi_\beta \;=\; \sum_\alpha U_{\alpha\beta}\,\phi'_\alpha \tag{8.29}$$

Since the transformation is assumed unitary, the $U_{\alpha\beta}$ are the elements of a unitary matrix $\mathbf{U}$, and the ϕ'_α, like the ϕ_α, are orthonormal. Substituting Eq. (8.29) into Eq. (8.9), we obtain

$$\begin{aligned} < \mathsf{H} > = & \sum_{\alpha\beta\gamma} U^*_{\beta\alpha}\, U_{\gamma\alpha}\; < \phi'_\beta \,|\, \mathsf{u} \,|\, \phi'_\gamma > \\ & + \frac{1}{2} \sum_{\alpha\beta} \sum_{\gamma\lambda\sigma\tau} U^*_{\gamma\alpha}\, U^*_{\lambda\beta}\, U_{\sigma\alpha}\, U_{\tau\beta}\, (< \phi'_\gamma\, \phi'_\lambda \,|\, \mathsf{v} \,|\, \phi'_\sigma\, \phi'_\tau > - < \phi'_\gamma\, \phi'_\lambda \,|\, \mathsf{v} \,|\, \phi'_\tau\, \phi'_\sigma >) \end{aligned} \tag{8.30}$$

Because $\mathbf{U}$ is unitary, $U^*_{\beta\alpha} \;=\; (\mathbf{U}^{-1})_{\alpha\beta}$, and we have the relation

$$\sum_\alpha U^*_{\alpha\beta}\, U_{\gamma\alpha} \;=\; \delta_{\beta\gamma} \tag{8.31}$$

Taking similar steps for the last term of Eq. (8.30), we find

$$\begin{aligned} < \mathsf{H} > = & \sum_{\beta\gamma} \delta_{\beta\gamma}\; < \phi'_\beta \,|\, \mathsf{u} \,|\, \phi'_\gamma > \\ & + \frac{1}{2} \sum_{\gamma\lambda\sigma\tau} \delta_{\gamma\sigma}\, \delta_{\lambda\tau}\, (< \phi'_\gamma\, \phi'_\lambda \,|\, \mathsf{v} \,|\, \phi'_\sigma\, \phi'_\tau > - < \phi'_\gamma\, \phi'_\lambda \,|\, \mathsf{v} \,|\, \phi'_\tau\, \phi'_\sigma >) \end{aligned} \tag{8.32}$$

Equation (8.32) reduces to the same form as Eq. (8.9), but with the functions ϕ'_α in place of the ϕ_α. We thus see that the same value will be obtained for $< \mathsf{H} >$ irrespective of whether $\{\phi_\alpha\}$ or $\{\phi'_\alpha\}$ are considered to be the single-particle functions comprising $\mathbf{\Phi}$. In fact, it can be shown that the many-particle function built from the ϕ'_α is identical to $\mathbf{\Phi}$ except for a possible phase factor.

The Hartree-Fock equations can be written in the matrix form

$$\mathsf{F}\,\phi \;=\; \phi\,\mathbf{\Lambda} \tag{8.33}$$

where ϕ and $\mathbf{\Lambda}$ are given by

$$\phi \;=\; (\phi_1 \quad \phi_2 \quad \cdots \quad \phi_N) \qquad\qquad \mathbf{\Lambda} \;=\; \begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1N} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ \lambda_{N1} & \lambda_{N2} & \cdots & \lambda_{NN} \end{pmatrix} \tag{8.34}$$

It is evident from Eq. (8.26) that the Fock operator is a monoelectronic operator. The summation term over the coulomb and exchange operators represents a one-electron approximation to the behavior of one electron in the field of the others. Now, let ϕ represent a matrix of Hartree-Fock spinorbitals. It is not difficult to show that the orthonormality of the Hartree-Fock orbitals is invariant under a unitary transformation. Moreover, a Slater determinant is invariant under such transformation. Thus, one finds that the Hartree-Fock wavefunction is determined only within a unitary transformation among the orbitals; i.e., the orbitals themselves are not unique. Given a particular set of Hartree-Fock orbitals, there are many different other sets obtainable from them by different unitary transformations. It turns out that the Hartree-Fock orbitals defined by Eqs. (8.23) and (8.24) are not particularly suitable for the description of many-electron systems in terms of orbitals associated with definite energy values. It is then convenient to choose a unitary transformation such that

$$\mathbf{U}^{\dagger}\,\mathbf{\Lambda}\,\mathbf{U} \;=\; \mathbf{\Xi} \tag{8.35}$$

where $\mathbf{\Xi}$ is a diagonal matrix, namely,

$$\mathbf{\Xi}_{ij} \;=\; \epsilon_i\,\delta_{ij} \tag{8.36}$$

Now multiplying Eq. (8.33) from the right by $\mathbf{U}$, using Eq. (8.35), and the unitarity of $\mathbf{U}$, we obtain

$$\mathsf{F}\,\phi' \;=\; \phi'\,\mathbf{\Xi} \tag{8.37}$$

where

$$\phi' \;=\; \phi\,\mathbf{U} \tag{8.38}$$

However, the operator F is defined in terms of the original Hartree-Fock orbitals$\{\phi_a\}$ and not in terms of the transformed orbitals $\{\phi'_a\}$, so that Eqs. (8.33) and (8.37) are not of the same form. We now show that the operator (8.26) is invariant under the unitary transformation. Since u does not depend on the Hartree-Fock orbitals, we need consider only the coulomb and exchange operators. Considering the coulomb operator first, we write

$$\mathsf{J}'_b(\mu)\,\phi_a(\mu) \;=< \phi'_b(\nu)\,|\,r_{\mu\nu}^{-1}\,|\,\phi'_b(\nu) >\; \phi_a(\mu) \tag{8.39}$$

Letting U_{ab} be an element of the unitary transformation matrix, we obtain

$$\begin{aligned}
\sum_b \mathsf{J}'_b(\mu)\,\phi_a(\mu) \;&=\; \sum_{kl} < \phi_k(\nu)\,|\,r_{\mu\nu}^{-1}\,|\,\phi_l(\nu) >\; \sum_b U^*_{kb}\,U_{lb}\,\phi_a(\mu) \\
&= \sum_{kl} < \phi_k(\nu)\,|\,r_{\mu\nu}^{-1}\,|\,\phi_l(\nu) >\; \delta_{kl}\,\phi_a(\mu) \\
&= \sum_{k} < \phi_k(\nu)\,|\,r_{\mu\nu}^{-1}\,|\,\phi_k(\nu) >\; \phi_a(\mu) \\
&= \sum_{b} < \phi_b(\nu)\,|\,r_{\mu\nu}^{-1}\,|\,\phi_b(\nu) >\; \phi_a(\mu) \\
&= \sum_b \mathsf{J}_b(\mu)\,\phi_a(\mu)
\end{aligned} \tag{8.40}$$

A similar result obtains for the exchange portion of the Hartree-Fock operator, so that

$$\sum_{b=1}^{N}(\mathsf{J}'_b - \mathsf{K}'_b) \;=\; \sum_{b=1}^{N}(\mathsf{J}_b - \mathsf{K}_b) \tag{8.41}$$

The advantage of the particular unitary transformation (8.35) is that the Hartree-Fock equations now assume the pseudoeigenvalue form (8.28)

$$\mathsf{F}\,|\,\phi_a >=\; \epsilon_a\,|\,\phi_a > \tag{8.42}$$

where we dropped the primes on the transformed orbitals for notational convenience. The $\{\epsilon_a\}$ are called the Hartree-Fock eigenvalues and are the energies associated with the Hartree-Fock orbitals. Eq. (8.42), plus the equations of constraint, Eq. (8.11), define the ϵ_α and ϕ_α. The constraint equations are important, because in their absence Eq. (8.42) would have

solutions for all values of the ϵ_α. However, the requirement that the ϕ_α be normalized makes Eq. (8.42) a *boundary-value* equation, with solutions only for certain values of ϵ_α. We also note that $\mathsf{F} = \mathsf{u} + \mathsf{u}_{HF}$ is Hermitian, confirming our expectation that the ϵ_α are real and verifying that different ϕ_α are orthogonal.

Since $\mathbf{\Lambda}$ is a hermitian matrix, it is always possible to find a unitary matrix $\mathbf{U}$ such that the transformation (8.35) diagonalizes $\mathbf{\Lambda}$. We are not concerned with how to obtain such a matrix, only that such a matrix exists and is unique. There must exist then a set of spinorbitals $\{\phi'_a\}$ for which the matrix of Lagrange multipliers is diagonal

$$\mathsf{F}\,|\,\phi'_a> = \epsilon_a\,|\,\phi'_a> \tag{8.43}$$

The unique set of spinorbitals $\{\phi'_a\}$ obtained from a solution of this eigenvalue equation is called the set of *canonical spinorbitals*. We henceforth drop the primes and write the Hartree-Fock equations as (8.42).

The canonical spinorbitals, which are solutions to $\mathsf{F}\,|\phi_a> = \epsilon_a\,|\phi_a>$, will generally be delocalized and form a basis for an irreducible representation of the point group of the molecule, i.e., they will have certain symmetry properties characteristic of the symmetry of the molecule or, equivalently, of the Fock operator. Once the canonical spinorbitals have been obtained it would be possible to obtain an infinite number of equivalent sets by a unitary transformation of the canonical set.

It is not difficult to show that the Hartree-Fock wavefunction is invariant under the unitary transformation (8.35), which transforms the N spinorbitals according to (8.38). Consider the linear transformation of the spinorbitals (8.29). Then the theorem is that

$$|\,\phi_1\,\phi_2\,\cdots\phi_N\,| \;=\; |\,\mathbf{U}\,|\,|\,\phi'_1\,\phi'_2\cdots\phi'_N\,| \tag{8.44}$$

where $|\mathbf{U}|$ is the determinant of the matrix of the $U_{\alpha\beta}$. To see this, we observe that in detail the left hand side of Eq. (8.44) is

$$Det\begin{pmatrix} \sum_{\beta=1}^{N} U_{1\beta}\,\phi'_\beta(1) & \cdots & \sum_{\beta=1}^{N} U_{1\beta}\,\phi'_\beta(N) \\ \vdots & \vdots & \vdots \\ \sum_{\beta=1}^{N} U_{N\beta}^{N}\,\phi'_\beta(1) & \cdots & \sum_{\beta=1}^{N} U_{N\beta}\,\phi'_\beta(N) \end{pmatrix} \tag{8.45}$$

which will be recognized as the determinant of the product of the two "matrices" $U_{\alpha\beta}$ and $\phi_\beta(s)$. Equation (8.44) then follows from the standard theorem that the determinant of a product of matrices is equal to the product of the determinants of the separate matrices.

One then concludes that ($Det\,\mathbf{U} = e^{i\varphi}$)

$$|\mathbf{\Phi}'| = |\mathbf{\Phi}| \tag{8.46}$$

that is, $\mathbf{\Phi}$ is invariant to a unitary transformation of the basis orbitals. Because any observable property depends on $|\mathbf{\Phi}|^2$, for all intents and purposes, the original wavefunction in terms of the spinorbitals $\{\phi_a\}$ and the transformed wavefunction in terms of the spinorbitals $\{\phi'_a\}$ are identical. For a single determinant wavefunction, any expectation value is therefore invariant to an arbitrary unitary transformation of the spinorbitals. Thus, the spinorbitals that make the total energy stationary are not unique, and no particular physical significance can be given to a particular set of spinorbitals. Localized spinorbitals, for example, are not more "physical" than delocalized spinorbitals.

One can even show that any transformation $\mathbf{\Phi}' = \mathbf{\Phi}\,\mathbf{T}$ ($\mathbf{T}$ nonsingular but not necessarily unitary) leaves $\mathbf{\Phi}$ invariant (except, possibly, for an immaterial normalization constant). This result follows from the fact that $|Det\ \mathbf{T}|$ is equal to a constant numerical quantity. However, if $\mathbf{\Phi}$ is not of the form assumed here (determinant of doubly occupied spatial orbitals), the preceding conclusions are not necessarily valid.

Turning now to the question of assuming the ϕ_a orthonormal, we first note that there are many linear transformations of the type (8.44); for example, the well-known Schmidt procedure, which, starting from a given linearly independent set of ϕ_a (and the ϕ_a must be linearly independent to start with or else $\mathbf{\Phi} \equiv 0$) will produce an orthonormal set of ϕ'_a. From (8.44), then, we see that $\mathbf{\Phi}$ is proproportional to a Slater determinant formed from the ϕ'_a, and therefore since the proportionality constant $|\mathbf{T}|$ will simply cancel out in calculating E, we have the desired result that we can, without loss of generality, restrict attention to determinants formed from orthonormal spinorbitals.

In summary, even with the requirement of orthonormality, the ϕ'_a are certainly not unique. One important consequence of this observation is that we must expect that our variational equations will not determine the optimal spinorbitals uniquely. Instead, at any stationary point, they will be arbitrary up to a unitary transformation.

Actually we earlier encountered an example of nonuniqueness of the variational parameters. Namely in the linear variation method only the ratios of the coefficients were determined, while the overall scale could be chosen arbitrarily. Thus lack of uniqueness of this variational parameters is not uncommon.

Alternatively, one may try to eliminate ambiguities by reparametrizing. Thus in the UHF approximation, it will turn out that the individual ϕ_a appear in E only in the form of the so called one-body density matrix

$$\rho^{UHF}(\mathbf{x}_1, \mathbf{x}'_1) \equiv \sum_{i=1}^{N} \phi_i(\mathbf{x}_1)\, \phi_i^*(\mathbf{x}'_1) \tag{8.47}$$

which is, as we will show, invariant to unitary transformations. Therefore one can bypass the ambiguity by basing the whole theory of the UHF approximation on ρ^{UHF} rather than on the ϕ_i individually.

To see that ρ^{UHF} is invariant we write

$$\phi_i = \sum_{j=1}^{N} T_{ij}\, \phi'_i \tag{8.48}$$

Then

$$\rho^{UHF}(\mathbf{x}_1, \mathbf{x}'_1) = \sum_{k=1}^{N} \sum_{j=1}^{N} \sum_{i=1}^{N} T_{ik}^*\, T_{ij}\, \phi'_j(\mathbf{x}_1)\, \phi_k'^*(\mathbf{x}'_1) \tag{8.49}$$

which in turn does become

$$\rho^{UHF}(\mathbf{x}_1, \mathbf{x}'_1) = \sum_{i=1}^{N} \phi'_k(\mathbf{x}'_1)\, \phi_k'^*(\mathbf{x}'_1) \equiv \rho'^{UHF}(\mathbf{x}_1, \mathbf{x}'_1) \tag{8.50}$$

if the transformation is unitary, that is, if

$$\sum_{l=1}^{N} T_{lk}^*\, T_{lj} = \delta_{kj} \tag{8.51}$$

The orbitals ϕ_α obtained from Eq. (8.42) with the Fock operator given by Eq. (8.26) have a natural physical interpretation. The operator $\mathsf{u} + \mathsf{u}_{HF}$ is of the form of the Hamiltonian for a single electron moving in the field of the nuclei, and subject to interaction with the Hartree-Fock charge distributions of the other electrons. This interaction consists of two parts, of which the first, or *Coulomb* part, is simply the electrostatic interaction of the electron with the charge density of all other electrons. The remaining, or *exchange* part, has no classical analog and is due to the antisymmetry requirement on the many-particle wavefunction. Even though the sums in Eq. (8.42) for u_{HF} are over all electrons, they do not include the interaction of an electron with its own distribution. In this case the Coulomb and exchange contributions are of equal magnitude, but opposite sign, and cancel each other.

We can now see that the iterative process leading to the solution of Eq. (8.42) involves the determination of individual electron distributions that are consistent with the Coulomb and exchange interactions produced by the distribution found for all other electrons. It is for this reason that the Hartree-Fock method is also known as the *self-consistent field* method (SCF method).

It is important to recognize that even though each Hartree-Fock ϕ_α is determined from an equation that includes interaction with all other electrons, the function Φ cannot be an exact solution to the many-electron Schrödinger equation. In brief, ϕ is deficient in that it ignores the fact that the probability distribution of each electron must be a function of the positions of all other electrons, and not merely dependent upon their distributions as a whole. In other words, the true probability of electron i being in any particular volume element must *correlate* with the probabilities that other electrons are individually in particular volume elements, and such a situation cannot be represented by a Φ that is an antisymmetric product of functions ϕ_α. It has become customary to refer to the deficiency of the Hartree-Fock wavefunction as a lack of *electron correlation*, and to identify the difference between

the exact energy and the Hartree-Fock $< \mathsf{H} >$ as the *correlation energy* that can be defined precisely as follows:

The correlation energy for a certain state with respect to a specified Hamiltonian is the difference between the exact eigenvalue of the Hamiltonian and its expectation value in the Hartree-Fock approximation for the state under consideration

The correlation energy so defined depends on the Hamiltonian and may be expressed mathematically as

$$E_{corr} = < \mathsf{H} > (exact) - < \mathsf{H} > (Hartree - Fock) \tag{8.52}$$

From now on, unless otherwise stated, we shall assume that H is the nonrelativistic Hamiltonian, and we shall use the term *correlation energy* to mean *nonrelativistic* correlation energy. It must be emphasized that the correlation energy defined in this way is *not* the difference between the exact experimental energy and the Hartree-Fock energy since the former contains contributions from relativistic effects. This is the reason of why the correlation energy is usually defined as the difference between the exact nonrelativistic energy and the restricted Hartree-Fock energy.

The correlation energy must be regarded as a strictly mathematical quantity which serves as a useful criterion for the acceptability of wavefunctions over Hartree-Fock wavefunctions. Good wavefunctions (on the basis of the variational principle) will be those accounting for as much of the correlation energy as possible. This is an important point in view of the fact that the Hartree-Fock energy is usually within about 1 percent of the experimental value. Although this may strike one as very good agreement, it must be pointed out that total energies per se are not of much significance to the problem of physics and chemistry. Rather, one is usually interested in energy *differences*, e.g., the energy difference between two spectroscopic states. Unfortunately, these energy differences themselves are often no larger than about 1 percent of the total energy of either state. Thus small absolute errors in the total energies may easily lead to large relative errors in their differences. For this reason there is a tremendous amount of interest in quantum-mechanical calculations which give better

energies than the Hartree-Fock method does. Since the Hartree-Fock method has been rather widely applied, and since one is beginning to understand just what its inadequacies are, it is helpful to use this method as a sort of reference point for more accurate calculations.

The above definition of correlation and correlation energy are unambiguous and well accepted, but unfortunately deviate to some extent from the usual meanings of these words. In a strict mathematical sense, there is in fact a correlation between individual electron distributions in the Hartree-Fock Φ. This mathematical correlation is produced by the antisymmetrization, and is most evident when we consider what happens when two electrons are at the same space-spin coordinate $\mathbf{x}$ - the antisymmetrization causes Φ to vanish, even if no $\phi_\alpha(\mathbf{x})$ vanishes. In fact, the so-called correlation error in the Hartree-Fock method reflects the fact that coulombic interaction between pairs of electrons, especially electrons with antiparallel spins, is not properly accounted for. Electrons of parallel spins are kept apart by the antisymmetry principle, an effect which overrides the coulombic repulstion, and are thus described somewhat better than electrons of antiparallel spins. Wavefunctions which lead to lower energies than the Hartree-Fock energy must somehow account for details of electronic motions in a more sophisticated manner than by just allowing each electron to move in a "smeared-out" field of the others. It is this assumption of an average potential field which allows electrons with antiparallel spins to avoid each other somewhat less assiduously than would be expected on the basis of electrostatic effects alone. According to the variational principle, it is evident that the correlation defined by Eq. (8.52) is a negative quantity, since the Hartree-Fock energy lies above the exact energy. From the virial theorem, which the Hartree-Fock energy satisfies, we see that

$$T_{corr} = -\frac{1}{2} V_{corr} \tag{8.53}$$

so that $T_{corr} > 0$, and $V_{corr} < 0$. As we shall soon discuss, the correlation energy for a two-electron atom is around -1.1 ev, so that $T_{corr} = 1.1$ ev. This means that the kinetic energy calculated by the Hartree-Fock method is always too low. This may be interpreted to mean that the electrons actually undergo more complicated movions to avoid each other

than is implicit in the mehtod. The potential-energy error of -2.2 ev results primarily from allowing two electrons with different spins to occupy the same spatial region, leading to a higher potential energy than is actually the case. The simple Hartree SCF method, which was discussed in Sect. V, is very similar to the Hartree-Fock SCF method except that exchange effects are entirely absent in the former. For this very reason the Hartree method (which does not include electron spin explicitly) is sometimes superior to the Hartree-Fock method. This comes about from the fact that the Hartree method gives an equally poor description of paired electrons and unpaired electrons, but the errors are of different sign and tend to cancel. The introduction of antisymmetry improves the description of parallel spins but does nothing to improve the description of antiparallel spins; hence cancellation of errors no longer occurs. In general the simple Hartree SCF method often leads to more sensible results than the Hartree-Fock method, e.g., in theories involving ferromagnetism, antiferromagnetism, and general magnetic properties.

The exact nonrelativistic energy of an atom is not always known with great accuracy, so that the correlation energy as defined by Eq. (8.52) is often known only approximately. In the case of two-electron atoms, accurate calculations are used obtain the exact nonrelativistic energies. For other atoms the exact nonrelativistic energies have been estimated using the relationship

$$E_{NR} = E_{exp} - E_R \tag{8.54}$$

(NR = nonrelativistic, R = relativistic). The experimental energy is obtained from summation of the successive ionization potentials of the atom and its ions. The relativistic correction to the energy is usually estimated by the use of perturbation theory.

The Hartree-Fock energy of the ground state of the helium atom is -2.861673 a.u. Using the value of -2.903724 a.u. as the exact nonrelativistic energy, we obtain a correlation energy of -0.042051 a.u. (about -1.14 ev). This represents an error of 1.14 ev in the kinetic energy and -2.8 ev in the potential energy. In the following Table are listed the theoretical and semiempirical correlation energies obtained for several two-electron atoms. It is seen from

this Table that the correlation energy is remarkably constant for the two-electron systems. The slight discrepancy between the theoretical and semiempirical values is within the limits of experimental error but may possibly indicate the existence of small, not yet understood effect. For atoms and ions with more than two electrons, the correlation energy is on the order of -2 ev ($= 46$ kcal mole^{-1}) per droubly filled orbital. This is rather a large error in view of the fact that it is of the same order of magnitude as chemical binding energies.

Correlation energies in electron volts of some heliumlike systems

		Theoretical	*Semiempirical*		
System	Z	E_{corr}	E_{corr}	$E_{exp} - EHF$	E_R
H^-	1	-1.08			
He	2	-1.14	-1.142	-1.145	-0.003
Li^+	3	-1.18	-1.182	-1.197	-0.015
Be^{++}	4	-1.20	-1.194	-1.250	-0.056
B^{3+}	5	-1.22	-1.196	-1.345	-0.149
C^{4+}	6	-1.23	-1.197	-1.521	-0.324
	∞	-1.28			

For heavier ions, for example, Al^{3+}, the correlation energy is on the order of -11 ev ($= 253$ kcal mole^{-1}). This is also, just about the magnitude of the relativistic energy. As seen from the Table, the relativistic energy varies with Z even as the number of electrons stays the same, increasing roughly as Z^4. It has been estimated that the relativistic correction for Al^{11+} is 9 ev, showing that the major portion of this correction comes from the inner shells. For $Z = 92$ it has been estimated that the relativistic correction is about 15 percent of the total energy. It is apparent that the relativistic effects contribute an important amount to

the total energy. This leads one to wonder whether or not relativistic effects are somehow experimentally observable in chemistry.

We now turn to the total electronic energy associated with the Hartree-Fock wavefunction. While this quantity is formally given by Eq. (8.9), now that the ϕ_α are presumed known, it is instructive to relate it to the orbital energies ϵ_α. Note that becuase ϵ_α includes the interaction of each other ϕ_β with ϕ_α, the sum of the ϵ_α cannot yield the correct value of the total energy, because it would count these interactions twice.

A convenient procedure for relating $< \mathsf{H} >$ to the sum of the ϵ_α may be developed from the partitioning of H into the two parts H_0 and H',

$$\mathsf{H} = \mathsf{H}_0 + \mathsf{H}' \tag{8.55}$$

defined by

$$\mathsf{H}_0 = \mathsf{U} + \mathsf{U}_{HF} \tag{8.56}$$

$$\mathsf{H}' = \mathsf{V} - \mathsf{U}_{HF} \tag{8.57}$$

with

$$\mathsf{U}_{HF} = \sum_i \mathsf{u}_{HF}(i) \tag{8.58}$$

U and V are given by Eqs. (5.2) and (5.3) respectively. This partitioning causes H_0 to be the sum of one-electron operators $\mathsf{F}(i)$

$$\mathsf{H}_0 = \sum_i \mathsf{F}(i) \tag{8.59}$$

where F, the Fock operator, is given by Eq. (8.26), i.e.,

$$\mathsf{F} = \mathsf{u} + \mathsf{u}_{HF} \tag{8.60}$$

and is the operator occurring in the Hartree-Fock equation (8.42). We thus have

$$\mathsf{H}_0 \Phi = E_0 \Phi \tag{8.61}$$

with

$$E_0 = \sum_\alpha \epsilon_\alpha \tag{8.62}$$

The previously described situation makes it natural to call H_0 the Hartree-Fock Hamiltonian.

We now write $< \mathsf{H} >$ as $< \mathsf{H}_0 > + < \mathsf{H}' >$, obtaining

$$< \mathsf{H} > = E_0 + < \mathsf{V} - \mathsf{U}_{HF} > \tag{8.63}$$

Eq. (8.63) shows that $< \mathsf{V} - \mathsf{U}_{HF} >$ is a measure of the exent to which E_0 overcounts the effect of the electron-electron interactions. Remembering that U_{HF} is a sum of one-electron operators, we write its expectation value in the form

$$< \mathsf{U}_{HF} > = \sum_\alpha < \phi_\alpha \,|\, u_{HF} \,|\, \phi_\alpha > \tag{8.64}$$

Using Eq. (8.26) for u_{HF} and from the definitions of the Coulomb and exchange operators given in Eqs. (8.15) and (8.16), we find

$$< \phi_\alpha \,|\, \mathsf{u}_{HF} \,|\, \phi_\alpha > = \sum_\beta (< \phi_\alpha \, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\alpha \, \phi_\beta > - < \phi_\alpha \, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\beta \, \phi_\alpha >) \tag{8.65}$$

which implies

$$< \mathsf{U}_{HF} > = 2 \, < \mathsf{V} > \tag{8.66}$$

We thus have

$$< \mathsf{H}' > = - < \mathsf{V} > \tag{8.67}$$

or

$$< \mathsf{H} > = E_0 - < \mathsf{V} > \tag{8.68}$$

confirming our earlier observation that the electron-electron interactions are counted twice in E_0. Since we also have $\mathsf{H} > = < \mathsf{U} > + < \mathsf{V} >$, we may convert Eq. (8.68) to the form

$$< \mathsf{H} > = \frac{1}{2} (E_0 + < \mathsf{U} >) \tag{8.69}$$

which clarifies the fact that once E_0 has been determined, no further terms explicitly involving two-electron operators need to be computed to get $< \mathsf{H} >$.

We close the discussion of the Hartree-Fock Hamiltonian with a more complete examination of the matrix elements of u_{HF}. These matrix elements occur whenever we partition H as in Eqs. (8.55) -(8.59), and enter into formulations of the correlation energy in particular. From Eq. (8.26), we find for any ϕ_μ and ϕ_ν, whether occupied or unoccupied in Φ,

$$< \phi_\mu \,|\, \mathsf{u}_{HF} \,|\, \phi_\nu > = \sum_\beta (< \phi_\mu \, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\nu \, \phi_\beta > - < \phi_\mu \, \phi_\beta \,|\, \mathsf{v} \,|\, \phi_\beta \, \phi_\nu >) \tag{8.70}$$

In fact, Eq. (8.70) holds for any functions ϕ_μ and ϕ_ν, whether or not they are solutions to the Hartree-Fock equations.

D. Koopmans' Theorem

In order to solve the Hartree-Fock equations it is necessary to introduce a basis set and solve a set of matrix equations. Before doing so, however, there are certain aspects of the eigenvalue equation and its solutions that are independent of any basis, and it is appropriate to discuss them at this point.

We have already seen that ϵ_α is an energy associated with a fermion in the orbital ϕ_α, interacting with all the other fermions. More significance can be given ϵ_α by comparing the energy expectation value of an N-electron system, assigned Hartree-Fock wavefunction Φ, with the energy expectation value of an $(N-1)$-electron system that is identical except for the removal of one electron from the single-particle ϕ_α, so that the $(N-1)$-electron system is assigned a wavefunction Φ', which is an antisymmetrized product of all the single-particle states in Φ except ϕ_α. The Φ' constructed in this way may be regarded as an approximation to the state reached from Φ by an ionization that removes the electron in ϕ_α.

Using Eq. (8.9) for $< \mathsf{H} >$, we find for the N-electron system the previously derived result

$$< \Phi \,|\, \mathsf{H} \,|\, \Phi > = \sum_\beta < \phi_\beta \,|\, \mathsf{u} \,|\, \phi_\beta > + \frac{1}{2} \sum_{\beta\gamma} < \phi_\beta \, \phi_\gamma \,||\, (\phi_\beta \, \phi_\gamma > \tag{8.71}$$

where the two-electron antisymmetric integral $< \cdots || \cdots >$ is given by

$$< \phi_\beta \, \phi_\gamma \, || \, \phi_\beta \, \phi_\gamma > = < \phi_\beta \, \phi_\gamma \, | \, \mathsf{v} \, | \, \phi_\beta \, \phi_\gamma > - < \phi_\beta \, \phi_\gamma \, | \, \mathsf{v} \, | \, \phi_\gamma \, \phi_\beta > \tag{8.72}$$

The corresponding result for the $(N-1)$-electron system (for which we let H_{N-1} denote the Hamiltonian) is

$$< \Phi' \, | \, \mathsf{H}_{N-1} \, | \, \Phi' > = \sum_{\beta(\beta \neq \alpha)} < \phi_\beta \, | \, \mathsf{u} \, | \, \phi_\beta > + \frac{1}{2} \sum_{\beta\gamma(\beta,\gamma \neq \alpha)} < \phi_\beta \, \phi_\gamma \, || \, \phi_\beta \, \phi_\gamma > \tag{8.73}$$

The difference between these expectation values is

$$< \Phi \, | \, \mathsf{H} \, | \, \Phi > - < \Phi' \, | \, \mathsf{H}_{N-1} \, | \, \Phi' > = < \phi_\alpha \, | \, \mathsf{u} \, | \, \phi_\alpha > + \sum_\beta < \phi_\alpha \, \phi_\beta \, || \, \phi_\alpha \, \phi_\beta > \tag{8.74}$$

A factor "$\frac{1}{2}$" is not needed in the last term of Eq. (8.74) because ϕ_α appears in Eq. (8.71) both when $\beta = \alpha$ and when $\gamma = \alpha$. The right-hand of Eq. (8.74) is seen to be

$$< \phi_\alpha \, | \, \mathsf{u} + \mathsf{u}_{HF} \, | \, \phi_\alpha > = \epsilon_\alpha \tag{8.75}$$

so we have

$$< \Phi \, | \, \mathsf{H} \, | \, \Phi > - < \Phi' \, | \, \mathsf{H}_{N-1} \, | \, \Phi' > = \epsilon_\alpha \tag{8.76}$$

This result, known as Koopmans' theorem shows that $-\epsilon_\alpha$ can be identified as an approximation to the ionization energy for an electron in orbital ϕ_α.

The identification of $-\epsilon_\alpha$ as an ionization energy is approximate, not only because the Hartree-Fock wavefunction of the N-electron system is inexact, but also because it was assumed that the single-particle states originally found for the N-electron system could be used without modification for the $(N-1)$-electron system. Actually, the removal of one electron will change the forces determining the optimum distributions of the other electrons, with the result that a better description for the $(N-1)$-electron system would be obtained if readjustments were permitted in all the remaining occupied ϕ_β. These readjustments, sometimes referred to as *relaxation*, can be significant and are often a major source of error in using $-\epsilon_\alpha$ as a measure of ionization energy. It is also obvious that because neither Φ nor Φ' permit a description of electron correlation effects, any differences in correlation energy

between the N-electron and $(N-1)$-electron systems will contribute to cause $-\epsilon_\alpha$ to deviate from a true ionization energy. Normally, these correlation energies will be significantly different, except when ϕ_α is but weakly coupled to the remaining ϕ_β (as, for example, when ϕ_α describes a lone electron in an outer-atomic shell).

Koopmmans' theorem can also be applied to the $(N+1)$-electron states formed by the addition of an electron in an unoccupied ϕ_r to the N-electron state Φ. Letting Φ" be the antisymmetrized product formed by appending ϕ_r to the ϕ_α in Φ, and letting H_{N+1} be the $(N+1)$-electron Hamiltonian, we can use Eq. (8.9) to show

$$< \Phi" \mid \mathsf{H}_{N+1} \mid \Phi" > - < \Phi \mid \mathsf{H} \mid \Phi > = < \phi_r \mid \mathsf{u} \mid \phi_r > + \sum_\beta < \phi_r \, \phi_\beta \mid\mid \phi_r \, \phi_\beta > \qquad (8.77)$$

The right hand side of Eq. (8.77) is the energy ϵ_r obtained for solution ϕ_r to the Hartree-Fock equations, and we see that $-\epsilon_r$ can be identified as an approximation to the electron affinity for addition of an electron to Φ in state ϕ_r. The earlier remarks regarding relaxation and correlation errors also apply here. For neutral systems, the errors in Koopmans's-theorem electron affinity estimates will usually be larger than for ionization energies, becuase the electron repulstions in the negative ion are often sufficient to make the optimum wavefunction far more diffuse than for the corresponding neutral system.

The foregoing discussion indicates that the ϵ_α can be identified with energies required to remove an alectron from ϕ_α or to add one to ϕ_r. However, it dones *not* follow that $\epsilon_r - \epsilon_\alpha$ is the energy required to move (excite) an electron from ϕ_α to ϕ_r, as can be seen by considering the two-step process of (1) removal from ϕ_α, (2) addition to ϕ_r. The first step corresponds to $-\epsilon_\alpha$, but the second involves addition to an $(N-1)$-electron system, not the N-electron system for which ϵ_r is appropriate. If we let Φ_α^r stand for the N-electron state in which ϕ_α is replaced by ϕ_r, we may calculate directly from Eq. (8.9)

$$\begin{aligned} < \Phi_\alpha^r \mid \mathsf{H} \mid \Phi_\alpha^r > = & \sum_{\beta(\beta\neq\alpha)} < \phi_\beta \mid \mathsf{u} \mid \phi_\beta > + < \phi_r \mid \mathsf{u} \mid \phi_r > \\ & + \frac{1}{2} \sum_{\beta\gamma(\beta,\gamma\neq\alpha)} < \phi_\beta \, \phi_\gamma \mid\mid \phi_\beta \, \phi_\gamma > + \sum_{\beta(\beta\neq\alpha)} < \phi_\beta \, \phi_r \mid\mid \phi_\beta \, \phi_r > \end{aligned} \qquad (8.78)$$

It is thus clear that

$$< \Phi_\alpha^r \,|\, \mathsf{H} \,|\, \Phi_\alpha^r > - < \Phi \,|\, \mathsf{H} \,|\, \Phi > = < \phi_r \,|\, \mathsf{u} \,|\, \phi_r > - < \phi_\alpha \,|\, \mathsf{u} \,|\, \phi_\alpha > + \sum_{\beta(\beta \neq \alpha)} < \phi_\beta \, \phi_r \,||\, \phi_\beta \, \phi_r > - \sum_\beta < \phi_\alpha \, \phi_\beta \,||\, \phi_\alpha \, \phi_\beta > \tag{8.79}$$

Introducing ϵ_r and ϵ_α, Eq. (8.79) can be rewritten

$$< \Phi_\alpha^r \,|\, \mathsf{H} \,|\, \Phi_\alpha^r > - < \Phi \,|\, \mathsf{H} \,|\, \Phi > = \epsilon_r - \epsilon_\alpha - < \phi_\alpha \, \phi_r \,||\, \phi_\alpha \, \phi_r > \tag{8.80}$$

Equation (8.80) shows that ϵ_r overestimates the energy for addition of an electron in ϕ_r by including the nonexistent interaction with an electron in ϕ_α. This formula for excitation energies is of course subject to the remarks already made regarding relaxation and correlation.

E. Basis-set Expansion

The equations for the Hartree-Fock orbitals, Eq. (8.42), are coupled integrodifferential equations, as u contains the diferential operator descriptive of the kinetic energy. We may also note that v, though algebraic, cannot be divided into factors dependent separately upon μ and ν. In spite of the difficulties associated with this mathematical structure, the first solutions to the Hartree-Fock equations were obtained by numerical integration methods. However, it soon became apparent that acceptable approximate solutions could be obtained more readily through the introduction of a basis set in which solutions might be expanded. This procedure has been particularly advantageous in molecular and crystalline systems, because of the complexity associated with their nonsphericity. The Hartree-Fock equations then give the set of conditions determining the coefficients $c_{\mu i}$ relating to each ϕ_i to the basis functions χ_μ of which it is composed.

To derive the Hartree-Fock equations appropriate to a basis expansion, sometimes referred to as the *matrix* Hartree-Fock equations, each Hartree-Fock orbital is expanded in terms of this basis set by

$$\phi_i = \sum_\mu \chi_\mu \, c_{\mu \, i} = \chi \, \mathbf{C}_i \tag{8.81}$$

If the complete basis set $\{\chi_\mu\}$ is used, the $\{\phi_i\}$ are given exactly. In practice, however, one truncates the expansion to m finite members of the basis set. $\mathbf{C}_i$ is then a column matrix of m rows. In order to be able to construct at least N linearly independent solutions, it is necessary that $m \geq N$.

Before deriving the matrix Hartree-Fock equations it is useful to express $< \mathsf{H} >$ in terms of integrals over the basis functions $\{\chi_\mu\}$. Direct substitution of Eq. (8.81) into (8.9) leads to

$$< \mathsf{H} > = \sum_{\alpha\mu\nu} c^*_{\mu\alpha}\, c_{\nu\alpha}\ < \chi_\mu \,|\, \mathsf{u} \,|\, \chi_\nu >$$
$$+\frac{1}{2} \sum_{\alpha\beta} \sum_{\mu\nu\lambda\sigma} c^*_{\mu\alpha}\, c^*_{\lambda\beta}\, c_{\nu\alpha}\, c_{\sigma\beta}\, (< \chi_\mu\, \chi_\lambda \,|\, \mathsf{v} \,|\, \chi_\nu\, \chi_\sigma > - < \chi_\mu\, \chi_\lambda \,|\, \mathsf{v} \,|\, \chi_\sigma\, \chi_\nu >) \tag{8.82}$$

Equation (8.82) may be cast in a simpler form if we introduce the reduced density matrices discussed in Section I. Starting from the definition in Eq. (1.8), the first-order density matrix corresponding to a normalized $\mathbf{\Phi}$ built from orthonormal orbitals [as given in Eq. (5.13)] simplifies to

$$\gamma(\mathbf{x}_1, \mathbf{x}'_1) \ = \sum_\alpha \phi_\alpha(\mathbf{x}_1)\, \phi_\alpha(\mathbf{x}'_1) \tag{8.83}$$

In the manipulations leading to Eq. (8.83), we may not combine the antisymmetrizers and appeal to the idempotence of $\mathcal{A}$, because we do not have a scalar product with respect to the variables $\mathbf{x}_1$ or $\mathbf{x}'_1$. But we can observe that we will get no contribution to γ except from terms such that the orbital ordering is identical in $\mathbf{\Psi}$ and $\mathbf{\Psi}^*$, and that there will be $(N-1)!$ such terms with a given orbital associated with $\mathbf{x}_1$ and $\mathbf{x}'_1$.

An important feature of Eq. (8.83) is that it shows γ to be the sum of the contributions from the individual occupied orbitals. If we now insert Eq. (8.81), thus replacing the set $\{\phi_\alpha\}$ with the set $\{\chi_\mu\}$, we reach

$$\gamma(\mathbf{x}_1, \mathbf{x}'_1) \ = \sum_{\alpha\mu\nu} c_{\mu\alpha}\, c^*_{\nu\alpha}\, \chi_\mu(\mathbf{x}_1)\, \chi^*_\nu(\mathbf{x}'_1) \tag{8.84}$$

Comparing Eq. (8.84) with the general expansion formula for γ given in Eq. (1.72), we see that if $\{\chi_\mu\}$ is regarded as the expansion for γ, its matrix elements arge given by

$$\gamma_{\mu\nu} \ = \sum_\alpha c_{\mu\alpha}\, c^*_{\nu\alpha} \tag{8.85}$$

The second-order density matrix corresponding to $\mathbf{\Phi}$ can also be simplified. The final result contains two kinds of terms: those that arise from contributions in which the orbitals of $\mathbf{\Psi}$ and those of $\mathbf{\Psi}^*$ occur in the same order, and those in which the nonintegrated orbitals (for Electrons 1 and 2) are reversed in $\mathbf{\Psi}^*$ relative to their ordering in $\mathbf{\Psi}$. Substituting Eq. (5.13) into the expression for Γ given in Eq. (1.9), we find

$$\Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = \frac{1}{2} \sum_{\alpha\beta} \phi_\alpha(\mathbf{x}_1)\, \phi_\beta(\mathbf{x}_2)\, [\phi_\alpha(\mathbf{x}'_1)\, \phi_\beta(\mathbf{x}'_2) - \phi^*_\beta(\mathbf{x}'_1)\, \phi^*_\alpha(\mathbf{x}'_2)] \tag{8.86}$$

Converting Eq. (8.86) to the $\{\chi_\mu\}$ basis,

$$\begin{aligned} &\Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) \\ &= \frac{1}{2} \sum_{\alpha\beta} \sum_{\mu\nu\lambda\sigma} c_{\mu\alpha}\, c_{\lambda\beta}\, (c^*_{\nu\alpha}\, c^*_{\sigma\beta} - c^*_{\nu\beta}\, c^*_{\sigma\alpha})\, \phi_\mu(\mathbf{x}_1)\, \phi_\lambda(\mathbf{x}_2)\, \phi^*_\nu(\mathbf{x}'_1)\, \phi^*_\sigma(\mathbf{x}'_2) \end{aligned} \tag{8.87}$$

Comparison with Eq. (1.78) yields

$$\Gamma_{\mu\lambda,\nu\sigma} = \frac{1}{2} \sum_{\alpha\beta} c_{\mu\alpha}\, c_{\lambda\beta}\, (c^*_{\nu\alpha}\, c^*_{\sigma\beta} - c^*_{\nu\beta}\, c^*_{\sigma\alpha}) \tag{8.88}$$

We note that Eq. (8.88) may be further simplified by substitution of the $\gamma_{\mu\nu}$ from Eq. (8.85). We obtain

$$\Gamma_{\mu\lambda,\nu\sigma} = \frac{1}{2} (\gamma_{\mu\nu}\, \gamma_{\lambda\sigma} - \gamma_{\mu\sigma}\, \gamma_{\lambda\nu}) \tag{8.89}$$

Equation (8.89) shows that for an independent-particle wavefunction the second-order density matrix may be written in terms of that of first order. This is an important simplification for independent-particle calculations, but one that will not be possible for wavefunctions that are not of the independent-particle form.

We are now ready to obtain an alternate form for $< \mathsf{H} >$. One possible route is to use the general expressions given in Eqs. (1.74) and (1.80). Using Eq. (8.89), this approach leads to

$$\begin{aligned} < \mathsf{H} > &= \sum_{\mu\nu} \gamma_{\nu\mu}\, < \chi_\mu \,|\, \mathsf{u} \,|\, \chi_\nu > + \frac{1}{2} \sum_{\mu\nu\lambda\sigma} (\gamma_{\nu\mu}\, \gamma_{\sigma\lambda} - \gamma_{\nu\lambda}\, \gamma_{\sigma\mu})\, < \chi_\mu\, \chi_\lambda \,|\, \mathsf{v} \,|\, \chi_\nu\, \chi_\sigma > \\ &= \sum_{\mu\nu} \gamma_{\nu\mu} \Big[< \chi_\mu \,|\, \mathsf{u} \,|\, \chi_\nu > \\ &\quad + \frac{1}{2} \sum_{\lambda\sigma} \gamma_{\sigma\lambda} (< \chi_\mu\, \chi_\lambda \,|\, \mathsf{v} \,|\, \chi_\nu\, \chi_\sigma > - < \chi_\mu\, \chi_\lambda \,|\, \mathsf{v} \,|\, \chi_\sigma\, \chi_\nu >) \Big] \end{aligned} \tag{8.90}$$

In obtaining the last line of Eq. (8.90), we have renamed the summation indices to permit extraction of a factor $\gamma_{\nu\mu}$ from all terms of $< \mathsf{H} >$. Equation (8.90) could alternatively have been derived by direct substitution of Eq. (8.85) into Eq. (8.82).

We now define the following matrices and matrix elements

$$\chi = (\chi_1 \quad \chi_2 \quad \cdots \quad \chi_m)$$

$$\mathbf{\Delta} = < \chi \,|\, \chi > \qquad\qquad \Delta_{\mu\nu} = < \chi_\mu \,|\, \chi_\nu >$$

$$< \phi_i \,|\, \mathsf{u} \,|\, \phi_i > = \mathbf{C}_i^\dagger \, < \chi^\dagger \,|\, \mathsf{u} \,|\, \chi > \, \mathbf{C}_i = \mathbf{C}_i^\dagger \, \mathbf{u} \, \mathbf{C}_i$$

$$(\mathbf{J}_i)_{\mu\nu} = < \chi_\mu \,|\, \mathsf{J}_i \,|\, \chi_\nu > \qquad (\mathbf{K}_i)_{\mu\nu} = < \chi_\mu \,|\, \mathsf{K}_i \,|\, \chi_\nu >$$

$$J_{ij} = \mathbf{C}_i^\dagger \, \mathbf{J}_j \, \mathbf{C}_i = \mathbf{C}_j^\dagger \, \mathbf{J}_i \, \mathbf{C}_j$$

$$K_{ij} = \mathbf{C}_i^\dagger \, \mathbf{K}_j \, \mathbf{C}_i = \mathbf{C}_j^\dagger \, \mathbf{K}_i \, \mathbf{C}_j$$

Using these definitions, the total energy associated with $\mathbf{\Phi}$ becomes, after projecting the single particle functions in the basis set (8.81),

$$E_0 = \sum_i^N \mathbf{C}_i^\dagger \, \mathbf{u} \, \mathbf{C}_i + \sum_{ij}^N \mathbf{C}_i^\dagger \, (\mathbf{J}_j - \mathbf{K}_j) \, \mathbf{C}_i \tag{8.91}$$

We now consider the functional

$$L = E_0 - \sum_{ij}^N \lambda_{ij} (\mathbf{C}_i^\dagger \, \mathbf{\Delta} \, \mathbf{C}_j - \delta_{ij}) \tag{8.92}$$

The first-order variation in this functional is

$$\begin{aligned} \delta L = & \sum_{i=1}^N \delta \, \mathbf{C}_i^\dagger \left[\left(\mathbf{u} + \sum_{j=1}^N (\mathbf{J}_j - \mathbf{K}_j) \right) \mathbf{C}_i - \sum_{j=1}^N \lambda_{ij} \, \mathbf{\Delta} \, \mathbf{C}_j \right] \\ & + \sum_{j=1}^N \left[\mathbf{C}_i^\dagger \left(\mathbf{u} + \sum_{i=1}^N (\mathbf{J}_j - \mathbf{K}_j) \right) - \sum_{j=1}^N \lambda_{ji} \, \mathbf{C}_j^\dagger \, \mathbf{\Delta} \right] \delta \, \mathbf{C}_i \end{aligned} \tag{8.93}$$

The vanishing of δL for an arbitrary variation $\delta \mathbf{C}_i$ is now satisfied by the conditions

$$\mathbf{F} \, \mathbf{C}_i - \sum_{j=1}^N \lambda_{ij} \, \mathbf{\Delta} \, \mathbf{C}_j = 0 \tag{8.94}$$

$$\mathbf{C}_i^\dagger \, \mathbf{F} - \sum_{j=1}^N \lambda_{ji} \, \mathbf{C}_j^\dagger \, \mathbf{\Delta} = 0 \tag{8.95}$$

where

$$\mathbf{F} = < \chi \,|\, \mathsf{F} \,|\, \chi > \tag{8.96}$$

Equations (8.94) and (8.95) are known as *Roothaan's equations*. By use of a properly chosen unitary transformation on the $\{\phi_i\}$, one can write Roothaan's equations in the pseudoeigenvalue form

$$\mathbf{F}\,\mathbf{C}_i \;=\; \epsilon_i\,\mathbf{\Delta}\,\mathbf{C}_i \tag{8.97}$$

or, more generally,

$$\mathbf{F}\,\mathbf{C} \;=\; \mathbf{\Delta}\,\mathbf{C}\,\epsilon \tag{8.98}$$

where

$$\mathbf{C} \;=\; (\,\mathbf{C}_1 \quad \mathbf{C}_2 \cdots \mathbf{C}_m\,) \qquad\qquad \mathbf{C}_i \;=\; \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{mi} \end{pmatrix} \tag{8.99}$$

The m eigenvectors $\{\mathbf{C}_i\}$ are orthogonal in the sense that $\mathbf{C}_i^{\dagger}\,\mathbf{\Delta}\,\mathbf{C}_j \;=\; \delta_{ij}$. The nontrivial solutions of Eq. (8.98) are obtained by solving for the m roots of the secular determinant

$$Det\,(\mathbf{F} - \epsilon\,\mathbf{\Delta}) \;=\; 0 \tag{8.100}$$

where each matrix element is of the form

$$F_{\mu\nu} - \epsilon\,\Delta_{\mu\nu} \;=< \chi_\mu \,|\, \mathsf{F} \,|\, \chi_\nu > - \epsilon \;< \chi_\mu \,|\, \chi_\nu > \tag{8.101}$$

and the Fock matrix elements (i.e., elements of $\mathbf{F}$) are explicitly given by

$$F_{\mu\nu} \;=< \chi_\mu \,|\, \mathsf{u} \,|\, \chi_\nu > + \sum_{\lambda\sigma} \gamma_{\sigma\lambda} \;< \chi_\mu\,\chi_\lambda \,||\, \chi_\nu\,\chi_\sigma > \tag{8.102}$$

and $\gamma_{\lambda\sigma}$ is an element of the first-order density matrix as given by Eq. (8.85). The quantities $< \chi_\mu|\mathsf{u}|\chi_\nu >$ are manifestly matrix elements of u in the bassis set $\{\chi_\mu\}$, and u_{HF} has matrix elements

$$(\mathsf{u}_{HF})_{\mu\nu} \;=\; \sum_{\lambda\sigma} \gamma_{\sigma\lambda} \;< \chi_\mu\,\chi_\lambda \,||\, \chi_\nu\,\chi_\sigma > \tag{8.103}$$

Equation (8.103) is an obvious analog of Eq. (8.70).

The whole discussion of this Section implicitly assumes that exact solutions are to be obtained for the Hartree-Fock equations. Of course, such solutions can only be realized within the numerical accuracy of the procedures used for solving the equations, and, if a basis set is introduced, to the limitations imposed by its lack of completeness. It therefore becomes relevant to inquire as to the extent to which the formulas given for $< \mathsf{H} >$ and for other quantities may be affected by these unavoidable inaccuracies. First of all, since the matrix $\mathbf{F}$ depends on the $\{\mathbf{C}_i\}$, the secular equations are nonlinear and must be solved by an iterative process. One approach is to assume some matrix $\mathbf{C}^{(1)}$ as first approximation, use it to construct $\mathbf{F}^{(1)}$, and then solve the Eqs. (8.100) to obtain a first-improved matrix $\mathbf{C}^{(2)}$. The process is repeated until an assumed matrix $\mathbf{C}^{(n)}$ and the calculated matrix $\mathbf{C}^{(n+1)}$ agree within previously specified limits. Since the basis orbitals $\{\chi_p\}$ are not varied, this method does not lead to the true Hartree-Fock solutions unless the basis used for the expansion (8.81) is rather large. The solutions are generally said to be *self-consistent* only with respect to the elements of the matrix $\mathbf{C}$.

Although it is not necessary to use an orthonormal basis set (linear independence is sufficient), we shall find it convenient to do so (under this condition $\mathbf{\Delta}$ represents the identity operator). We recall that it is always possible to tranform any m linearly independent functions to a new, orthonormal set so that the assumption of an orthonormal basis set leads to no loss of generality. Roothaan's equations now become

$$\mathbf{F}\,\mathbf{C} \;=\; \mathbf{C}\,\epsilon \tag{8.104}$$

It is clear that inaccuracies in solution of the Hartree-Fock equations will cause the ϕ_α thereby determined to deviate from the set of single-particle functions that actually minimizes $< \mathsf{H} >$. But we may take the viewpoint that the ϕ_α, however determined, are to be regarded as precisely defined quantities. Assuming only that they are orthonormal, Eqs. (8.82) and (8.90) remain valid. If we also accept Eqs. (8.20) and (8.70) as precise definitions of u_{HF}, it will then continue to be true that $< \mathsf{U}_{HF} >= 2 < \mathsf{V} >$ [cf. Eq. (8.66)]. However,

unless ϕ_α are in fact exact, they will not be eigenfunctions of $\mathsf{u} + \mathsf{u}_{HF}$.

Even though the ϕ_α are not the exact Hartree-Fock orbitals, they and their corresponding orbital energies ϵ_α will frequently be obtained by a process that

$$< \phi_\alpha \,|\, \mathsf{u} + \mathsf{u}_{HF} \,|\, \phi_\alpha > = \epsilon_\alpha \tag{8.105}$$

Equation (8.105) will, for example, be satisfied if ϕ_α is a solution to the matrix Hartree-Fock equations, even for a severely limited basis. Given Eq. (8.105), it is apparent that

$$< \mathsf{H} > = \sum_\alpha \epsilon_\alpha - < \mathsf{V} > \tag{8.106}$$

An approximate operator F consistent with this discussion is

$$\mathsf{F} = \sum_\alpha |\, \phi_\alpha > \epsilon_\alpha < \phi_\alpha \,| + \sum_r |\, \phi_r > \epsilon_r < \phi_r \,| \tag{8.107}$$

where the unoccupied orbitals ϕ_r are an arbitrary set orthogonal to the $\{\phi_\alpha\}$, and the ϵ_r are also arbitrary. With this F, Eq. (8.42) as well as Eqs. (8.61) - (8.69) will retain their validity.

In the case that the ϕ_α are solutions to the matrix Hartree-Fock equations for a (necessarily) limited basis, $(\mathsf{u} + \mathsf{u}_{HF} - \epsilon_\alpha)\,\phi_\alpha$ will have no component within the space spanned by the basis, and $\mathsf{V} - \mathsf{U}_{HF}$ will be an exact representation of H' within the corresponding many-particle space. We may thus use the partitioning implied by Eqs. (8.56) and (8.57) as a starting point from which to consider basis-set expansions of electron correlation effects not included in the independent-particle model.

F. Restrictions in the Hartree-Fock Method

Even for relatively simple systems it turns out to be both difficult and time-consuming to implement the Hartree-Fock method in the totally unrestricted form in which it has been presented in the preceding part of this section. Moreover, as will become clearer later, even if an exact solution to the Hartree-Fock equations was obtained, it would generally not be found to describe a many-particle wavefunction Φ possessing the spatial or spin symmetries

of the Hamiltonian, nor would its single-particle states ϕ_μ exhibit these symmetries. The true Hartree-Fock optimization would thus have been obtained at a considerable cost in ease of visualization and interpretation.

There are two alternatives for remedying the symmetry difficulties of the Hartree-Fock function (referred to as a symmetry dilemma), of which the first is simply to restrict the determinantal wavefunction $\mathbf{\Phi}$ so that it possesses a symmetry appropriate to the Hamiltonian. The second alternative, to be discussed in more detail later, is to generalize the independent-particle formulation to include the symmetry projections of determinantal wavefunctions. This approach constitutes a generalization because such projections usually yield wavefunctions that cannot be written as single Slater determinants.

In this subsection we consider restrictions that may usefully be applied to $\mathbf{\Phi}$. Those corresponding to one-particle operators can be effectuated by applying them to the single-particle states ϕ_μ. Restrictions of this class include some spatial symmetries, and for spin-independent electronic Hamiltonians, one (usually the z) component of the electron spin. Thus, we might require that the spatial part of each ϕ_μ transform as a basis function of some irreducible representation of the spatial symmetry group of the Hamiltonian, and that the spin part of each ϕ_μ be an eigenfunction of the z component of the spin operator, s_z. Because the single-particle spin space is spanned by two s_z eigenfunctions (denoted as in sections IV and VI, α and β) with different eigenvalues (respectively $+\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$), each ϕ_μ can be written as a *spinorbital* consisting of a spatial orbital multiplied by α or β.

For spin-independent electronic Hamiltonians, there is one important symmetry condition that involves a two-particle operator, namely that the wavefunction be an eigenfunction of the square of the total spin. Letting $\mathsf{s}(i)$ be the spin operator for Electron i, the two-particle nature of S^2 is evident from the formula

$$\mathsf{S}^2 = \sum_i \mathsf{s}(i)^2 + \sum_{ij(i\neq j)} \mathsf{s}(i).\mathsf{s}(j) \tag{8.108}$$

Unless the spinorbitals ϕ_α satisfy certain restrictive conditions, the wavefunction $\mathbf{\Phi}$ built from them cannot be an eigenfunction of S^2. There are two such restrictions, one or the

other of which must apply to each occupied ϕ_γ. The first alternative is that a spatial orbital be *doubly occupied* (i.e., that it occur twice among the occupied ϕ_γ, once with spin eigenfunction α and once with eigenfunction β). As is well known, a pair of electrons doubly occupying a spatial orbital will make no contribution to the total spin. The second restriction, applicable to the single-occupied spatial orbitals, is that they all be associated with the same s_z eigenfunction. This condition causes these spins to couple to an eigenfunction of maximum total spin.

Where the spatial symmetry group of the Hamiltonian possesses irreducible representations of dimension greater than unity, further restrictions may be necessary if $\mathbf{\Phi}$ is to transform appropriately under spatial symmetry operations. A general discussion is cumbersome and will not be presented here. Suffice it to say, $\mathbf{\Phi}$ must not contain different elements of two different bases of the same irreducible representation unless it contains at least one of the complete bases. This condition is referred to as an *equivalence restriction.* A typical example is provided by the p orbitals of an atom; the equivalence restriction demands that they occur in Φ in sets, each set having its own radial dependence.

Use of the symmetry, occupancy, and equivalence restrictions described in the preceding three paragraphs defines a standard procedure known as the *Restricted Hartree-Fock* (RHF) method. The RHF method is at its simplest form when $\mathbf{\Phi}$ can be assumed to consist completely of doubly-occupied orbitals and when there are no partially occupied bases for irreducible representations of the spatial symmetry group. A state well represented by such an RHF $\mathbf{\Phi}$ is termed a *closed-shell* state, and $\mathbf{\Phi}$ is referred to as a *closed-shell wavefunction.* The ground states of most molecules with an even number of electrons and without high symmetry fall into this class, as do the ground states of the noble gas atoms and other atoms with complete subshells (such as Be, configuration $1s^2 2s^2$). States requiring an RHF $\mathbf{\Phi}$ containing singly-occupied orbitals or partially-occupied irreducible-representation bases are termed *open shell*; odd-electron molecules and many atomic states fall into this category. It should be noted that there exist many molecular states (e.g., so-called nonclassical structures) and atomic states (e.g., C $1s^2 2s^2 2p^2$, ^1S) that are not well represented by any

single-determinant wavefunction and that are therefore inappropriate to the RHFmethod. Such systems should be handled either by generalizations of the Hartree-Fock scheme or by methods that include a suitable portion of the correlation energy.

G. Restricted Hartree-Fock Method - Closed Shell States

In this subsection we consider the simplifications resulting from the Hartree-Fock method if we limit consideration to a spin-independent electronic Hamiltonian and closed-shell wavefunctions. The wavefunction then consists of spinorbitals occurring in pairs with the same spatial function, but with different s_z eigenvalues; to indicate this fact explicitly, we introduce the notations ϕ_γ and $\phi_{\overline{\gamma}}$ referring respectively to the members of a spinorbital pair with α and β spin functions. For the moment we ignore the possible effects of any spatial symmetry.

Our first concern must be to modify the derivation of the Hartree-Fock equations to take proper account of the fact that ϕ_γ and $\phi_{\overline{\gamma}}$ are not independent, but are constrained to have the same spatial dependence. As in the original derivation, we obtain an expression for $< \mathsf{H} >$,

$$
\begin{aligned}
< \mathsf{H} > = & \sum_{\gamma} (< \phi_\gamma \,|\, \mathsf{u} \,|\, \phi_\gamma > + < \phi_{\overline{\gamma}} \,|\, \mathsf{u} \,|\, \phi_{\overline{\gamma}} >) \\
& + \frac{1}{2} \sum_{\gamma\delta} (< \phi_\gamma \, \phi_\delta \,|\, \mathsf{v} \,|\, \phi_\gamma \, \phi_\delta - \phi_\delta \, \phi_\gamma > + < \phi_\gamma \, \phi_{\overline{\delta}} \,|\, \mathsf{v} \,|\, \phi_\gamma \, \phi_{\overline{\delta}} - \phi_{\overline{\delta}} \, \phi_\gamma > \\
& + < \phi_{\overline{\gamma}} \, \phi_\delta \,|\, \mathsf{v} \,|\, \phi_{\overline{\gamma}} \, \phi_\delta - \phi_\delta \, \phi_{\overline{\gamma}} > + < \phi_{\overline{\gamma}} \, \phi_{\overline{\delta}} \,|\, \mathsf{v} \,|\, \phi_{\overline{\gamma}} \, \phi_{\overline{\delta}} - \phi_{\overline{\delta}} \, \phi_{\overline{\gamma}} >)
\end{aligned}
\tag{8.109}
$$

Since we have explicitly included separate contributions for the two members of each spinorbital pair, the summations now run over spinorbitals with distinct spatial parts. Letting w_γ be the spatial of ϕ_γ and writing $\phi_\gamma = w_\gamma \alpha$, $\phi_{\overline{\gamma}} = w_\gamma \beta$, we carry out the spin integrations, causing some terms to vanish and other to become identical. The result is

$$
< \mathsf{H} > = \, 2 \sum_{\gamma} < w_\gamma \,|\, \mathsf{u} \,|\, w_\gamma > + \sum_{\gamma\delta} < w_\gamma \, w_\delta \,|\, \mathsf{v} \,|\, 2 \, w_\gamma \, w_\delta - w_\delta \, w_\gamma >
\tag{8.110}
$$

We next examine $< \mathsf{H} >$ by varying the w_γ subject to the constraint that they remain orthogonal. Introducing Lagrange multipliers, we write

$$\delta\left(< \mathsf{H} > -2 \sum_{\gamma\delta} \lambda_{\delta\gamma} \; < w_\gamma \,|\, w_\delta > \right) \; = \; 0 \tag{8.111}$$

As before, we eliminate off-diagonal Lagrange multipliers by making a suitable unitary transformation of the w_γ. Note that this step requires the invariance of the form of $< \mathsf{H} >$ under the transformation; the requirement is met here. Consideration of the effect of a general variation δw_γ^* then leads to

$$(\mathsf{u} + 2\,\mathsf{J} - \mathsf{K})\, w_\gamma(\mathbf{r}) \; = \; \epsilon_\gamma \, w_\gamma(\mathbf{r}) \qquad\qquad (\forall \;\; \gamma) \tag{8.112}$$

with

$$\mathsf{J}\, w_\gamma(\mathbf{r}) \; = \; \sum_\delta \, < w_\delta(\mathbf{r}') \,|\, \mathsf{v}(\mathbf{r}, \mathbf{r}') \,|\, w_\delta(\mathbf{r}') > \; w_\gamma(\mathbf{r}) \tag{8.113}$$

$$\mathsf{K}\, w_\gamma(\mathbf{r}) \; = \; \sum_\delta \, < w_\delta(\mathbf{r}') \,|\, \mathsf{v}(\mathbf{r}, \mathbf{r}') \,|\, w_\gamma(\mathbf{r}') > \; w_\delta(\mathbf{r}) \tag{8.114}$$

These sums are over distinct spatial orbitals w_δ. J and K are called *Coulomb* and *exchange* operators, respectively. J represents the classical electrostatic interaction with the occupied-orbital charge distribution of one spin orientation; the term $2\mathsf{J}w_\gamma(\mathbf{r})$ describes the classic interaction between an electron in orbital $w_\gamma(\mathbf{r})$ and the occupied orbitals of both spin orientations. K represents the exchange interaction with the occupied-orbital charge distribution of appropriate spin orientation; the term $-\mathsf{K}w_\gamma(\mathbf{r})$ describes the exchange interaction between an electron in orbital $w_\gamma(\mathbf{r})$ and the occupied orbitals of the same spin orientation (there is no exchange interaction between distributions of unlike spin orientation).

If we compare Eqs. (8.112) - (8.114) with Eq. (8.42) [with F given by Eq. (8.26)] we find the two sets of equations to be identical if the wavefunctions summed over in Eq. (8.26) are constrained to be a closed-shell set. This fact means that we can use all the formulas developed in our earlier discussion of the Hartree-Fock method, providing that we specialize them to reflect the restriction to a closed-shell wavefunction. In this way, we may use the formulas relating to a basis-set expansion of the Hartree-Fock equations.

Accordingly, let us consider the Hartree-Fock equations in a spinorbital basis whose members occur in pairs with the same spatial function. Let χ_μ and $\chi_{\overline{\mu}}$ be such a pair with spatial part $g_\mu(\mathbf{r})$, so $\chi_\mu = g_\mu\alpha$, $\chi_{\overline{\mu}} = g_\mu\beta$. Because the orbitals ϕ_γ and $\phi_{\overline{\gamma}}$ are similarly paired, the coefficients connecting them with the basis orbitals must satisfy

$$c_{\mu\gamma} = c_{\overline{\mu}\,\overline{\gamma}} \qquad\qquad c_{\overline{\mu}\gamma} = c_{\mu\overline{\gamma}} = 0 \tag{8.115}$$

and the elements of the first-order density matrix [cf. Eq. (1.73)] have a similar property

$$\gamma_{\mu\nu} = \gamma_{\overline{\mu}\,\overline{\nu}} \qquad\qquad \gamma_{\mu\overline{\nu}} = \gamma_{\overline{\mu}\nu} = 0 \tag{8.116}$$

An examination of Eq. (8.102) and of the expression for $\Delta_{\mu\nu}$ shows that the elements of the Fock matrix $\mathbf{F}$ and the overlap matrix $\mathbf{\Delta}$ behave correspondingly, so the matrix Hartree-Fock equations partition into two identical sets, one for each spin orientation. The restriction to a closed-shell wavefunctiion therefore halves the dimensionality of a matrix Hartree-Fock problem.

Let us now write explicitly the Fock matrix element connecting two α-spinorbitals. From Eq. (8.102), using also Eq. (8.116),

$$\begin{aligned} F_{\mu\nu} &= < g_\mu\,\alpha \,|\, \mathsf{u} \,|\, g_\nu\,\alpha > + \sum_{\lambda\sigma} (\gamma_{\sigma\lambda} \, < g_\mu\,\alpha\, g_\lambda\,\alpha \,|\, \mathsf{v} \,|\, g_\nu\,\alpha\, g_\sigma\,\alpha - g_\sigma\,\alpha\, g_\nu\,\alpha > \\ &\quad + \gamma_{\overline{\sigma}\,\overline{\lambda}} \, < g_\mu\,\alpha\, g_\lambda\,\beta \,|\, \mathsf{v} \,|\, g_\nu\,\alpha\, g_\sigma\,\beta - g_\sigma\,\beta\, g_\nu\,\alpha >) \\ &= < g_\mu \,|\, \mathsf{u} \,|\, g_\nu > + \sum_{\lambda\sigma} \gamma_{\sigma\lambda} \, < g_\mu\, g_\lambda \,|\, \mathsf{v} \,|\, 2\, g_\nu\, g_\sigma - g_\sigma\, g_\nu > \end{aligned} \tag{8.117}$$

One further notational change is needed to make Eq. (8.117) correspond with the bulk of the RHF literature, namely to replace the first-order density matrix γ by the so-called chage-bond order matrix, denoted $\mathbf{P}$. As usually defined, $\mathbf{P}$ is based on spatial orbitals rather than spinorbitals, and is related to γ by the equation

$$P_{\mu\nu} = \gamma_{\mu\nu} + \gamma_{\overline{\mu}\,\overline{\nu}} \tag{8.118}$$

For closed-shell wavefunctions, $P_{\mu\nu} = 2\,\gamma_{\mu\nu}$, so Eq. (8.117) becomes

$$F_{\mu\nu} = < g_\mu \,|\, \mathsf{u} \,|\, g_\nu > + \sum_{\lambda\sigma} P_{\sigma\lambda} \, < g_\mu\, g_\lambda \,|\, \mathsf{v} \,|\, g_\nu\, g_\sigma - 1/2\, g_\sigma\, g_\nu > \tag{8.119}$$

It should be noted that the $\{P_{\mu\nu}\}$ are the elements of an $m \times m$ matrix defined by

$$\mathbf{P} \;=\; 2\,\mathbf{R}\,\mathbf{R}^{\dagger} \qquad\qquad \mathbf{R} \;=\; (\,\mathbf{C}_1 \;\; \mathbf{C}_2 \cdots \mathbf{C}_N\,) \tag{8.120}$$

where $\mathbf{R}$ is an $m \times N$ matrix of the eigenvectors associated with the N lowest eigenvalues. The $\mathbf{R}\mathbf{R}^{\dagger}$ matrix is readily verified to be hermitian and idempotent and thus is the matrix representation of a projection operator.

We make explicit the other formulas needed for closed-shell RHF calculations

$$\Delta_{\mu\nu} \;=< g_\mu \,|\, g_\nu > \tag{8.121}$$

$$P_{\mu\nu} \;=\; 2 \sum_{\alpha} c_{\mu\alpha}\, c^{*}_{\nu\alpha} \tag{8.122}$$

$$\mathbf{F}\,\mathbf{C}_\gamma \;=\; \epsilon_\gamma\,\mathbf{\Delta}\,\mathbf{C}_\gamma \qquad\qquad (\forall\;\gamma) \tag{8.123}$$

$$\begin{aligned} < \mathsf{H} > &= \sum_{\mu\nu} P_{\nu\mu} \left[< g_\mu \,|\, \mathsf{u} \,|\, g_\nu > + \frac{1}{2} \sum_{\lambda\sigma} P_{\sigma\lambda} \; < g_\mu\, g_\lambda \,|\, \mathsf{v} \,|\, (g_\nu\, g_\sigma - 1/2\, g_\sigma\, g_\nu) > \right] \\ &= \frac{1}{2} \left(E_0 + \sum_{\mu\nu} P_{\nu\mu} \; < g_\mu \,|\, \mathsf{u} \,|\, g_\nu > \right) \;=\; \frac{1}{2} \left(E_0 + Tr\,\mathbf{P}\,\mathbf{u} \right) \end{aligned} \tag{8.124}$$

where

$$E_0 \;=\; 2 \sum_{\gamma} \epsilon_\gamma \tag{8.125}$$

Symmetries other than those associated with s_z, will lead to further factorization of the matrices $\mathbf{F}$ and $\mathbf{\Delta}$ into blocks. The maximum factorization will result if the basis function g_μ are of definite symmetry *species* and *subspecies*. These terms respectively refer to an irreducible representation of the spatial symmetry group and to individual elements of a particular basis thereof; respective examples for an atom are "p" and "p_z".

Since the final RHF spatial orbitals will be of definite symmetry species and subspecies, the matrix $\mathbf{P}$ will have vanishing elements $P_{\mu\nu}$ unless g_μ and g_ν are of the same species and subspecies. If this symmetry is assumed for $\mathbf{P}$, and if $\mathbf{P}$ is symmetric in its description of the subspecies of each species (one of the closed-shell requirements), the symmetry properties

of the matrix elements of u and v suffice to yield a corresponding factorization of $\mathbf{F}$. It is clear that $\mathbf{\Delta}$ factorizes correspondingly.

In contrast to the spin factorization, blocks corresponding to different symmetry species need not be identical. However, because of the equivalence restriction, blocks for different subspecies of the same species *will* be identical; this fact can be used to avoid unnecessary calculation. It should be noted that the spatial-symmetry-factorized matrix blocks are coupled because elements from all blocks of $\mathbf{P}$ are needed to evaluate each block of $\mathbf{F}$.

Thus far our discussion has been based on the symmetry properties that the charge-bond order matrix $\mathbf{P}$ will have when constructed from the RHF wavefunction. It is natural, therefore, to seed the RHF solution by an iterative process in which an input approximation to $\mathbf{P}$ also has this symmetry. We note that any symmetry-blocked input $\mathbf{P}$, whether or not constructed from the RHF solution, generates a correspondingly symmetry-blocked $\mathbf{F}$, and therefore leads to output symmetry spinorbitals and therefore to an output $\mathbf{P}$ preserving the symmetry blocking. We thus see that the assumption of symmetry spinorbitals is stable under iteration, and that when an appropriate iterative process has converged, we will have exact solutions to the Hartree-Fock equations.

The fact that the RHF spinorbitals exactly satisfy the Hartree-Fock equations may lead one to ask whether the RHF procedure actually constitutes a restriction in the Hartree-Fock method. The answer to this question depends upon whether the vanishing of the variation of the Hartree-Fock functional indicates a saddle point or a local minimum, and in the latter case whether it is the absolute (i.e., lowest) minimum consistent with the state to be described. In most problems of practical interest the symmetry-adapted solution is not an absolute minimum, and therefore in fact is restricted. There is an interesting corollary to the above discussion, namely that to solve the Hartree-Fock equations without symmetry restrictions, it will be necessary to use input density matrices of broken symmetry.

H. Orthogonalization of the Basis

The basis sets that are used in molecular calculations are not orthonormal sets. The basis functions are normalized, but they are not orthogonal to each other. This gives rise to the overlap matrix in Roothaan's equations. In order to put Roothaan's equations into the form of the usual matrix eigenvalue problem, we need to consider procedures for orthogonalizing the basis functions.

If we have a set of functions $\{\chi_\mu\}$ that are not orthogonal, i.e.,

$$< \chi_\mu \,|\, \chi_\nu >= \Delta_{\mu\nu} \tag{8.126}$$

then it will always be possible to find a transformation matrix $\mathbf{X}$ (not unitary) such that a transformed set of functions $\{\chi'_\mu\}$ given by

$$\chi'_\mu \;=\; \sum_\nu X_{\mu\nu}\, \chi_\nu \qquad\qquad \mu \;=\; 1, 2, \cdots, m \tag{8.127}$$

do form an orthonormal set, i.e.,

$$< \chi'_\mu \,|\, \chi'_\nu >= \delta_{\mu\nu} \tag{8.128}$$

To derive the properties of $\mathbf{X}$, we substitute the transformation (8.127) into Eq. (8.128) to get

$$\begin{aligned} < \chi'_\mu \,|\, \chi'_\nu >&= \sum_\lambda \sum_\sigma X^*_{\lambda\mu}\, X_{\sigma\nu} \;< \chi_\lambda \,|\, \chi_\sigma > \\ &= \sum_\lambda \sum_\sigma X^*_{\lambda\mu}\, X_{\sigma\nu}\, \Delta_{\lambda\sigma} \;=\; \delta_{\mu\nu} \end{aligned} \tag{8.129}$$

The second line of this equation can be written as the matrix equation

$$\mathbf{X}^\dagger\, \mathbf{\Delta}\, \mathbf{X} \;=\; \mathbf{I}. \tag{8.130}$$

where $\mathbf{I}$ is the $m \times m$ unit matrix. Thus, Eq. (8.130) defines the relation that the matrix $\mathbf{X}$ must satisfy if the transformed orbitals are to form an orthonormal set. As we shall see later, $\mathbf{X}$ must also be nonsingular, i.e., it must possess an inverse $\mathbf{X}^{-1}$. We now proceed to

show how to obtain two different transformation matrices $\mathbf{X}$. Since $\mathbf{\Delta}$ is Hermitian it can be diagonalized by a unitary matrix $\mathbf{U}$,

$$\mathbf{U}^{\dagger}\,\mathbf{\Delta}\,\mathbf{U} = \mathbf{S} \tag{8.131}$$

where $\mathbf{S}$ is a diagonal matrix of the eigenvalues of $\mathbf{\Delta}$.

There are two ways of orthogonalizing the basis set $\{\chi_\mu\}$ in common use. The first procedure, called *symmetric orthogonalization* uses the inverse square root of $\mathbf{\Delta}$ for $\mathbf{X}$

$$\mathbf{X} \equiv \mathbf{\Delta}^{-1/2} = \mathbf{U}\mathbf{S}^{-1/2}\,\mathbf{U}^{\dagger} \tag{8.132}$$

If $\mathbf{\Delta}$ is Hermitian then $\mathbf{\Delta}^{-1/2}$ is also Hermitian. Substituting Eq. (8.132) into (8.130)

$$\mathbf{\Delta}^{-1/2}\,\mathbf{\Delta}\,\mathbf{\Delta}^{-1/2} = \mathbf{I} \tag{8.133}$$

shows that $\mathbf{X} = \mathbf{\Delta}^{-1/2}$ is indeed an orthogonalizing transformation matrix. Since the eigenvalues of $\mathbf{\Delta}$ are all positive, there is no difficulty in (8.132) of taking square roots. However, if there is linear dependence or near linear dependence in the basis set, then some of the eigenvalues will approach zero and (8.132) will involve dividing by quantities that are nearly zero. Thus symmetric orthogonalization will lead to problems in numerical precision for basis sets with near linear dependence.

A second way of obtaining an orthonormal set of basis functions is called *canonical orthogonalization*. It uses the transformation matrix

$$\mathbf{X} = \mathbf{U}\,\mathbf{S}^{-1/2} \tag{8.134}$$

that is, the columns of the unitary matrix $\mathbf{U}$ are divided by the square root of the corresponding eigenvalue

$$X_{ij} = U_{ij}/S_j^{1/2} \tag{8.135}$$

Substituting this definition of $\mathbf{X}$ into Eq. (8.130) gives

$$\mathbf{X}^{\dagger}\,\mathbf{\Delta}\,\mathbf{X} = \left(\mathbf{U}\,\mathbf{S}^{-1/2}\right)^{\dagger}\,\mathbf{\Delta}\,\mathbf{U}\,\mathbf{S}^{-1/2} = \mathbf{S}^{1/2}\,\mathbf{U}^{\dagger}\,\mathbf{\Delta}\,\mathbf{U}\,\mathbf{S}^{-1/2} = \mathbf{S}^{-1/2}\,\mathbf{S}\,\mathbf{S}^{-1/2} = \mathbf{I} \tag{8.136}$$

showing that $\mathbf{X} = \mathbf{U}\mathbf{S}^{-1/2}$ is also an orthogonalizing transformation matrix. It appears, from Eq. (8.135), that this orthogonalization procedure will also entail difficulties if there is linear dependence in the basis set, i.e., if any of the eigenvalues S_j approach zero. We can circunvent this problem with canonical orthogonalization, however. In the matrix eigenvalue problem (8.131), we can order the eigenvalues in any way in the diagonal matrix $\mathbf{S}$ provided we order the columns of $\mathbf{U}$ in the same way. Suppose we order the positive values S_j in the order $S_1 > S_2 > S_3 > \cdots$. Upon inspection we may decide that the last m of these are too small and will give numerical problems. We can then use as a transformation matrix, the truncated matrix $\tilde{\mathbf{X}}$,

$$\tilde{\mathbf{X}} = \begin{pmatrix} U_{11}/S_1^{1/2} & U_{12}/S_2^{1/2} & \cdots & \cdots & U_{1K-m}/S_{K-m}^{1/2} \\ U_{21}/S_1^{1/2} & U_{22}/S_2^{1/2} & \cdots & \cdots & U_{2K-m}/S_{K-m}^{1/2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ U_{K1}/S_1^{1/2} & U_{K2}/S_2^{1/2} & \cdots & \cdots & U_{KK-m}/S_{K-m}^{1/2} \end{pmatrix} \tag{8.137}$$

where we have eliminated the last m columns of $\mathbf{X}$ to give the $K \times (K-m)$ matrix $\tilde{\mathbf{X}}$. With this truncated transformation matrix, we get only $K - m$ transformed orthonormal basis functions

$$\chi'_\mu = \sum_{\nu=1}^{K} \tilde{\mathbf{X}}_{\nu\mu}\, \chi_\nu \qquad\qquad \mu = 1, 2, \cdots K - m \tag{8.138}$$

These would span exactly the same region of space as the original set, provided the eliminated eigenvalues were exactly zero. In practice, one often finds linear dependence problems with eigenvalues in the region $S_j \leq 10^{-4}$.

One way of dealing with the problem of a nonorthogonal basis set would thus be to orthogonalize the functions $\{\chi_\mu\}$ to obtain the transformed basis functions $\{\chi'_\mu\}$ and work with these orthonormal functions throughout. This would eliminate the overlap matrix $\mathbf{\Delta}$ from the Roothaan's equations, which could then be solved just by diagonalizing the Fock matrix. This would mean, however, that we would have to calculate all our two-electron integrals using the new orbitals or else transform all the old integrals $< \mu\nu|\lambda\sigma >$ to the set $< \mu'\nu'|\lambda'\sigma' >$. In practice this is very time consuming, and we can solve the same problem

in a more efficient way. Consider a new coefficient matrix $\mathbf{C}'$ related to the old coefficient matrix $\mathbf{C}$ by

$$\mathbf{C}' = \mathbf{X}^{-1}\,\mathbf{C} \qquad\qquad \mathbf{C} = \mathbf{X}\,\mathbf{C}' \tag{8.139}$$

where we have assumed that $\mathbf{X}$ is non-singular. This will be the case if we have eliminated linear dependences. Substituting $\mathbf{C} = \mathbf{X}\mathbf{C}'$ into the Roothaan's equations gives

$$\mathbf{F}\,\mathbf{X}\,\mathbf{C}' = \mathbf{\Delta}\,\mathbf{X}\,\mathbf{C}'\,\epsilon \tag{8.140}$$

Multiplying on the left by $\mathbf{X}^\dagger$ gives

$$\left(\mathbf{X}^\dagger\,\mathbf{F}\,\mathbf{X}\right)\mathbf{C}' = \left(\mathbf{X}^\dagger\,\mathbf{\Delta}\,\mathbf{X}\right)\mathbf{C}'\,\epsilon \tag{8.141}$$

If we define a new matrix $\mathbf{F}'$ by

$$\mathbf{F}' = \mathbf{X}^\dagger\,\mathbf{F}\,\mathbf{X} \tag{8.142}$$

and use (8.130), then

$$\mathbf{F}'\,\mathbf{C}' = \mathbf{C}'\,\epsilon \tag{8.143}$$

These are the transformed Roothaan's equations, which can be solved for $\mathbf{C}'$, by diagonalizing $\mathbf{F}'$. Given $\mathbf{C}'$, then $\mathbf{C}$ can be obtained from (8.139). Therefore, given $\mathbf{F}$ we can use (8.142), (8.143) and (8.139) to solve the Roothaan's equations $\mathbf{FC} = \mathbf{\Delta C}\epsilon$ for $\mathbf{C}$ and ϵ. The intermediated primed matrices are just the Fock matrix and expansion coefficients in the orthogonalized basis, i.e.,

$$\phi_i = \sum_{\mu=1}^{K} c'_{\mu i}\,\chi'_\mu \qquad\qquad i = 1, 2, \cdots, K \tag{8.144}$$

$$F'_{\mu\nu} = < \chi'_\mu \,|\, \mathsf{F} \,|\, \chi'_\nu > \tag{8.145}$$

The Hartree-Fock equations (8.42) produce a set $\{\phi_i\}$ of spinorbitals. The single determinant $\mathbf{\Phi}$ formed from the N spinorbitals $\{\phi_\alpha\}$ with the lowest orbital energies is the Hartree-Fock approximation to the ground state.

An advantage of the Hartree-Fock method is that it provides one with a simple orbital picture of electronic structure. This picture, resulting from the single-determinantal form of the wavefunction, is a very useful one to the chemist who likes to think of electrons moving in orbitals. Nevertheless, such an orbital description is not adequate from a quantitative standpoint, however pleasing and satisfactory it may be from a qualitative viewpoint. In general, wavefunctions which go beyond the Hartree-Fock approximation tend to eliminate (or obscure) the orbital concept, so that quantitative accuracy is gained at the price of losing a pleasing physical picture. Of course, the orbital picture provided by the Hartree-Fock solution is in itself not unique, since there exist many different orbital pictures related by unitary transformations which leave the total wavefunction invariant. Thus one can choose unitary transformations to suit the particular typoe of analysis of interest.

Once one has obtained a Hartree-Fock solution to a particular problem, one can always improve the energy by using the configuratio-interaction method. In a multideterminantal representation of the exact ground state $|\mathbf{\Phi}>$, it is these determinants which might be expected, a priori, to give the leading correction to the Hartree-Fock ground state $|\mathbf{\Phi}>$. The improved wavefunctions (not normalized) can be written

$$|\mathbf{\Psi}> = c_0\,|\mathbf{\Phi}> + \sum_{r\alpha} c_\alpha^r\,|\mathbf{\Phi}_\alpha^r> + \cdots \tag{8.146}$$

where $\mathbf{\Phi}$ is the Hartree-Fock solution. The remaining determinants can be constructed from the Hartree-Fock orbitals, including those not occupied in the ground state of the system. These latter orbitals are referred to as *virtual* orbitals and can be used to represent excited states. We recall that each of the determinants in (8.146) must have the same S_z eigenvalue as $\mathbf{\Phi}$; otherwise there will be no contribution to the total wavefunction of the state represented by $\mathbf{\Phi}$.

If we consider only the singly excited determinants as corrections, then the coefficients c_α^r are determined from the variational principle by diagonalizing the Hamiltonian matrix in the basis $\{\mathbf{\Phi}, \{\mathbf{\Phi}_\alpha^r\}\}$. Consider for a moment the matrix eigenvalue problem involving one singly excited state

$$\begin{pmatrix} <\mathbf{\Phi}\,|\,\mathsf{H}\,|\,\mathbf{\Phi}> & <\mathbf{\Phi}\,|\,\mathsf{H}\,|\,\mathbf{\Phi}_\alpha^r> \\ <\mathbf{\Phi}_\alpha^r\,|\,\mathsf{H}\,|\,\mathbf{\Phi}> & <\mathbf{\Phi}_\alpha^r\,|\,\mathsf{H}\,|\,\mathbf{\Phi}_\alpha^r> \end{pmatrix} \begin{pmatrix} c_0 \\ c_\alpha^r \end{pmatrix} = E \begin{pmatrix} c_0 \\ c_\alpha^r \end{pmatrix} \tag{8.147}$$

The mixing of the two states depends on the off-diagonal element $<\mathbf{\Phi}|\mathsf{H}|\mathbf{\Phi}_\alpha^r>$. This matrix element is obtained by using the rules for evaluating matrix elements between determinants given in Sect. VII, i.e.,

$$<\mathbf{\Phi}\,|\,\mathsf{H}\,|\,\mathbf{\Phi}_\alpha^r> = <\phi_\alpha\,|\,\mathsf{u}\,|\,\phi_r> + \sum_\beta <\phi_\alpha\,\phi_\beta\,|\,\mathsf{v}\,|\,(\phi_r\,\phi_\beta - \phi_\beta\,\phi_r)> \tag{8.148}$$

The right-hand side of this equation can be simplified; the matrix elements of the Fock operator (8.60) are given by

$$<\phi_\alpha\,|\,\mathsf{F}\,|\,\phi_r> = <\phi_\alpha\,|\,\mathsf{u}\,|\,\phi_r> + \sum_\beta <\phi_\alpha\,\phi_\beta\,|\,\mathsf{v}\,|\,(\phi_r\,\phi_\beta - \phi_\beta\,\phi_r)> \tag{8.149}$$

where Eq. (8.70) has been used to express the matrix elements of u_{HF}. Therefore

$$<\mathbf{\Phi}\,|\,\mathsf{H}\,|\,\mathbf{\Phi}_\alpha^r> = <\phi_\alpha\,|\,\mathsf{F}\,|\,\phi_r> \tag{8.150}$$

The matrix element that mixes singly excited determinants with $\mathbf{\Phi}$ is thus equal to an off-diagonal element of the Fock matrix. Now, by definition, solving the Hartree-Fock eigenvalue problem requires the off-diagonal elements to satisfy $<\phi_i|\mathsf{F}|\phi_j> = 0$ $(i \neq j)$. That is

$$<\mathbf{\Phi}\,|\,\mathsf{H}\,|\,\mathbf{\Phi}_\alpha^r> = 0 \tag{8.151}$$

One can then say that solving the Hartree-Fock eigenvalue equation is equivalent to ensuiring that $\mathbf{\Phi}$ will not mix with any singly excited determinants. The Hartree-Fock ground state is in this sense "stable" since it cannot be improved by mixing it with singly excited determinants. This represents a further restriction on the determinants. The important result we have just derived is the Brillouin's theorem. This theorem states that if two determinants constructed from *exact* Hartree-Fock orbitals differ in one spinorbital, the matrix element

connecting these two determinants will vanish. The theorem depends upon the form of the Hartree-Fock orbitals. It is not necessary, however, that the orbitals used to represent the excited states be themselves eigenfunctions of the Hartree-Fock operator. The only requirement is that such orbitals be orthogonal to the occupied orbitals. The reason for this lies in the fact that the solutions of the Hartree-Fock equations form a complete set of orthogonal functions. Furthermore, these solutions can be divided into two sets (those occupied in the ground state and those not occupied in the ground state) which span mutually orthogonal subspaces. The set of unoccupied orbitals then forms a complete set for expansion of any orbital orthogonal to the occupied orbitals. Brillouin's theorem can then be applied to the expansion term by term.

A configuration differing from the ground state by one spinorbital is said to be *singly excited*; one differing by two spinorbitals is called *doubly excited*; and so on for configurations differing by more spinorbitals. Brillouin's theorem does *not* say that singly excited configurations do not contribute to the total energy, since it is possible that matrix elements between singly excited configurations and higher-excited configurations do not vanish and thus contribute to the total energy (for example, one can expect doubly excited determinants $|\mathbf{\Phi}^{rs}_{\alpha\beta} >$ to provide the leading and most important corrections to $|\mathbf{\Phi} >$). This, however, represents a higher-order effect. One can restate the Brillouin theorem as follows:
"Singly excited configurations constructed from spinorbitals orthogonal to the occupied Hartree-Fock orbitals give no first-order contribution to the total energy."
Brillouin's theorem also follows from the fact that the Hartree-Fock wavefunction is the best single-determinantal wavefunction (in the variational sense) and thus must be stationary to first-order variations. Brillouin's theorem also guarantees second-order errors for all one-electron operators *provided* the *exact* Hartree-Fock solutions are employed. Such a situation does not obtain, for example, when one solves Roothaan's equations in a limited basis set.

The calculation of the energy of an excited state constructed from the Hartree-Fock orbitals can be carried out by the use of Eqs. (7.) - (7.). We now illustrate such a calculation for the case of singly excited configurations. We assume that the Hartree-Fock problem for

the singlet ground state has been solved. The ground-state wavefunction is

$$D_0 \;=\; |\phi_1\,\overline{\phi}_1\,\phi_2\,\overline{\phi}_2\cdots\phi_N\,\overline{\phi}_N| \tag{8.152}$$

To form the singly excited configuration we promote one electron from an occupied orbital ϕ_α to an unoccupied orbital ϕ_r. For simplicity, it is assumed that ϕ_α and ϕ_r are both nondegenerate. This leads to the determinantal wavefunctions

$${}^{1,3}\mathbf{\Phi}^r_\alpha \;=\; 2^{-1/2}\,(|\phi_1\,\overline{\phi}_1\cdots\phi_\alpha\,\overline{\phi}_r\cdots\phi_N\,\overline{\phi}_N| \pm |\phi_1\,\overline{\phi}_1\cdots\overline{\phi}_\alpha\,\phi_r\cdots\phi_N\,\overline{\phi}_N|) \tag{8.153}$$

where the negative sign refers to a singlet state and the positive sign refers to the tripet state. For notational convenience we shall designate the determinantal functions appearing in Eq. (8.153) by D_1 and D_2 (in their order of appearance). The right subscript and superscript on the wavefunction symbol are a convenient notation for indicating that an electron has been promoted from ϕ_α to ϕ_r. The energy of the singlet state relative to the energy of the ground state is given by

$${}^1E - E_0 \;=\!<{}^1\,\mathbf{\Phi}^r_\alpha\,|\,\mathsf{H}\,|\,{}^1\mathbf{\Phi}^r_\alpha> - <\,D_0\,|\,\mathsf{H}\,|\,D_0\,> = \frac{1}{2}\,(\mathsf{H}_{11}+\mathsf{H}_{22}) - \mathsf{H}_{12} - \mathsf{H}_{00} \tag{8.154}$$

where

$$\mathsf{H}_{pq} \;=\!<\,D_p\,|\,\mathsf{H}\,|\,D_q\,> \tag{8.155}$$

Rather than carrying out a generalized treatment of this calculation (which involves some confusing summation notation) we shall consider the special case

$$D_0 \;=\; |\phi_1\,\overline{\phi}_1\,\phi_2\,\overline{\phi}_2| \qquad\qquad D_1 \;=\; |\phi_1\,\overline{\phi}_1\,\phi_2\,\overline{\phi}_3| \qquad\qquad D_2 \;=\; |\phi_1\,\overline{\phi}_1\overline{\phi}_2\,\phi_3| \tag{8.156}$$

Using Eqs. (7.80) and (7.84) for the one- and two-particle operators matrix elements and integrating over the spin functions, we obtain for the diagonal elements

$$\mathsf{H}_{00} \;=\; 2\,\mathsf{u}_{11} + 2\,\mathsf{u}_{22} + J_{11} + 4\,J_{12} + J_{22} - 2\,K_{12} \tag{8.157}$$

$$\mathsf{H}_{11} \;=\; \mathsf{H}_{22} \;=\; 2\,\mathsf{u}_{11} + \mathsf{u}_{22} + \mathsf{u}_{33} + J_{11} + 2\,J_{12} + 2\,J_{13} + J_{23} - K_{12} - K_{13} \tag{8.158}$$

Similarly, when we use Eq. (7.90), the off-diagonal matrix element becomes

$$\mathsf{H}_{12} = -K_{23} \tag{8.159}$$

Combining these results, we obtain

$$^1E - E_0 = \mathsf{u}_{33} - \mathsf{u}_{22} - 2\,J_{12} + 2\,J_{13} + J_{23} - J_{22} + K_{12} - K_{13} + K_{23} \tag{8.160}$$

It is convenient to rewrite this energy in terms of the Hartree-Fock eigenvalues given by

$$\epsilon_3 = \mathsf{u}_{33} + 2\,J_{13} + 2\,J_{23} - K_{13} - K_{23} \tag{8.161}$$

$$\epsilon_2 = \mathsf{u}_{22} + 2\,J_{12} + J_{22} - K_{12} \tag{8.162}$$

Substituting Eqs. (8.161) and (8.162) into Eq. (8.160) and simplifying, it is obtained

$$^1E - E_0 = \epsilon_3 - \epsilon_2 - J_{23} + 2\,K_{23} \tag{8.163}$$

Similarly, for the triplet state it is found

$$^3E - E_0 = \epsilon_3 - \epsilon_2 - J_{23} \tag{8.164}$$

The above results can be generalized to

$$^1E - E_0 = \epsilon_r - \epsilon_\alpha - J_{\alpha r} + 2\,K_{\alpha r} \tag{8.165}$$

$$^3E - E_0 = \epsilon_r - \epsilon_\alpha - J_{\alpha r} \tag{8.166}$$

The Hartree-Fock theory then predicts that the singlet-triplet separation is

$$^1E - ^3E = 2\,K_{\alpha r} > 0 \tag{8.167}$$

i.e., the lowest triplet state lies below the lowest singlet excited state.

So far we have confined our discussion to a particular form of the Hartree-Fock method, a form applicable to atoms whose wavefunctions can be developed on the basis of closed-shell configuration, so that m_l and m_s quantum numbers do not have to be specified. In all cases this leads to a wavefunction which is a ^1S state and in which the orbitals turn

out to satisfy so-called ıequivalence restrictions. The equivalence restrictions are satisfied whenever the radial portions of the orbitals are independent of the m_s and m_l quantum numbers and the orbitals are associated with definite l quantum numbers. This means that the resulting wavefunction is an eigenfunction of S^2 and L^2. For wavefunctions obtained in other ways than by reference to traditional electronic configurations, the equivalence restrictions are not automatically satisfied, even in the case of systems for which a closed-shell configuration can be written. Such a situation has been well known in the case of atoms whose wavefunctions are based on open-shell configurations, but it has not been generally appreciated that the same situation occurs for closed-shell atoms or molecules if one removes the closed-shell configurations basis for the wavefunctions. For example, there exist Hartree-Fock solutions for the ions H^-, Li^-, O^{--}, and Na^- which lead to energies lower than those obtained from wavefunctions derived from the closed-shell configurations $1s^2$, $1s^22s^2$, $1s^22s^22p^6$, and $1s^22s^22p^63s^2$, respectively. Each of these results is obtained by exhibiting a Slater determinant not satisfying the conventional equivalence restrictions. This means that the popular notion of ^1S states representing the ground states of such species is not necessarily valid.

The Hartree-Fock method in which the equivalence restrictions hold or are imposed is known as the *traditional, conventional, or restricted* Hartree-Fock method. Most Hartree-Fock calculations reported in the literature have been carried out by this method. In the case of open-shell configurations in which the equivalence restrictions are *imposed*, it is not always possible to write the total wavefunction as a single determinant. It is possible, however, to apply the SCF theory of closed-shell systems to *certain* open-shell systems if the equivalence restrictions are imposed; e.g., half-filled open shell may be treated in this way.

In the case of configurations involving an odd number of electrons, one can drop the restriction that the radial functions be independent of m_s. This means that orbitals of opposite spin need not have the same spatial functions in configurations with a singly occupied orbital. For example, one could write the lithium wavefunction as $|1s\overline{1s'}2s|$ instead of $|1s\overline{1s}2s|$. This arises from consideration of the fact that in a system of unbalanced spins

(where $< \mathsf{S}_z > \neq 0$) the electrons of one spin will be affected by an exchange potential other than that from electrons of opposite spin, since exchange potentials occur only between electrons of parallel spins. This suggests that electrons with different spins should be placed into different spatial orbitals in order to account for this exchange polarization. This variant is usually referred to as the the *spin-polarized* Hartree-Fock method. If the trial wavefunction is a single Slater determinant with no a priori constraints on the form of the occupied orbitals, one has the so-called *unrestricted* Hartree-Fock method (see Sect.VIII.I). This name is also sometimes applied to any Hartree-Fock method in which at least one of the restrictions of the conventional Hartree-Fock method is removed. It should be noted that the conventional Hartree-Fock method for systems described by closed-shell configurations is a special case of the unrestricted Hartree-Fock method, since the equivalence relations are not imposed but occur automatically. It is to the unrestricted Hartree-Fock method that Brillouin's theorem applies; for any other Hartree-Fock methods, e.g., open-shell systems with imposed equivalence restrictions, Brillouin's theorem does not apply exactly.

Reference is also made to an *extended* Hartree-Fock method, in which the trial function is obtained from the Slater determinant of an open-shell configuration by use of L^2 and S^2 projection operators. The extended Hartree-Fock wavefunction can be written

$$\mathbf{\Phi}_p = \mathcal{O}\,\mathbf{\Phi} \tag{8.168}$$

where $\mathcal{O}$ is a desired projection operator and $\mathbf{\Phi}$ is a single-determinantal wavefunction. The energy is then given by

$$< \mathsf{H} > = \frac{< \mathbf{\Phi} \mid \mathcal{O}^\dagger \mathsf{H}\, \mathcal{O} \mid \mathbf{\Phi} >}{< \mathbf{\Phi} \mid \mathcal{O} \mid \mathbf{\Phi} >} \tag{8.169}$$

The minimization process is applied in the usual way, but nos the relevant operator is $\mathcal{O}^\dagger \mathsf{H} \mathcal{O}$ and not H. So far it has not proved feasible to apply this method to large systems.

There are certain difficulties in restricted Hartree-Fock methods. For example, consider an arbitrary hermitian operator G which commutes with the total Hamiltonian. It follows at once from Theorem 3, Sect. I that the *exact* eigenfunctions of H are also eigenfunctions of G

or, in the case of energy degeneracies, can always be chosen as eigenfunctions of G. However, in the case of *approximate* eigenfunctions of H, there exists no theorem which also requires these to be eigenfunctions of G. Any imposed requirement that such H eigenfunctions also be G eigenfunctions then *must* represent a *constraint*, which necessarily raises the energy above the absolute minimum. In the same way, restricting the wavefunction to one that can be constructed from a closed-shell electronic configuration constitutes a constraint. It is therefore apparent why the H^-, Li^-, O^{--}, and Na^- have Hartree-Fock solutions with energies below that of the usual ^1S states.

The off-diagonal Lagrangian multipliers occurring in open-shell Hartree-Fock calculations disappear if one drops the restriction of doubly occupied orbitals. In the case of systems with an odd number of electrons this is just the spin-polarized Hartree-Fock method. Now, however, the single-determinantal wavefunctions, for example, $|1s\overline{1s'}|$ and $|1s\overline{1s'}2s|$, are not eigenfunctions of S^2. In general, the determinantal wavefunction of this form is a mixture of different spin states and can be written

$$\mathbf{\Phi} = \sum_{j=0}^{M+2j} {}^{M+2j}\Phi_j \, c_j \tag{8.170}$$

where $M = 2S + 1$ and S is the spin quantum number of the state of lowest possible multiplicity. The number of terms in Eq. (8.170) depends upon the total number of electrons and can be found from the branching diagram of Fig. VI.2 . The functions ${}^{M+2j}\Phi_j$ are eigenfunctions of S^2, with eigenvalues $(S + j)(S + j + 1)$. An example of this type of determinant is $|\phi_1\overline{\phi}_2|$, which can be written

$$|\phi_1\overline{\phi}_2| = c_0 \, {}^1\Phi_0 + c_1 \, {}^3\Phi_1 \tag{8.171}$$

We recall from Sect. VI.B that

$${}^1\Phi_0 = |\phi_1 \, \overline{\phi}_2| - |\overline{\phi}_1 \, \phi_2| \qquad\qquad {}^3\Phi_1 = |\phi_1 \, \overline{\phi}_2| + |\overline{\phi}_1 \, \phi_2| \tag{8.172}$$

For a many-electron determinant the number of multiplets in (8.170) will be rather large. One can use the determinant (8.170) as a trial wavefunction, but the resulting energy represents some average of the energies of the component multiplets. However, one can project

the spin state of the desired multiplicity and evaluate the energy of this multiplet. This energy will be lower than that of the mixture and, furthermore, will be lower than that obtained by the conventional (restricted) Hartree-Fock method. This circumstance arises because the projected determinant will, in general, be a linear combination of determinants and thus leads to a sort of configuration-interaction effect, in which the coefficients are fixed by the projection instead of by a variational procedure.

J. "Unrestricted" and Generalized Hartree-Fock Methods

The closed-shell RHF method outlined in Sect. VIII.G is found to be highly satisfactory for the systems for which it is most appropriate. The restrictions are too extreme, however, for systems possessing a partially occupied set of orbitals of a given symmetry species, or for those where the spin state requires the existence of singly-occupied orbitals. Some of these systems can be well handled by relaxing the requirement of double-occupancy, while retaining the symmetry and equivalence restrictions. Calculations of this type are referred to as *open-shell RHF*.

An alternate possibility is the so-called *Unrestricted Hartree-Fock* (UHF) method, "so-called" because not all of the RHF restrictions are removed. In the usual UHF approach, the double-occupancy restriction is dropped, but the spinorbitals are still required to have definite spatial symmetry and to be s_z eigenfunctions. An energy improvement relative to RHF is usually gained, but its cost is that $\mathbf{\Phi}$ is no longer and S^2 eigenfunction and that spatial orbitals of different spin orientation are no longer identical or orthogonal.

Single-particle states and improved energies can alternatively be obtained using generalizations of the Hartree-Fock method. One important class of generalizations is reached by allowing the wavefunction to be a nondeterminantal function from a single determinant. Wavefunctions of this type include those where a pure-symmetry component is *projected* out of a determinantal function. Normally, the projected function will be a linear combination of determinants with coefficients whose values are fixed by the projection process. A simple

example of a projected function is the singlet state obtained from the mixed-spin determinantal function $\mathbf{\Phi} = w_1\alpha w_2\beta - w_2\beta w_1\alpha$, namely $\mathbf{\Phi}_s = 1/2(w_1w_2 + w_2w_1)(\alpha\beta - \beta\alpha)$. We see that $\mathbf{\Phi}_s$ is a component of the original $\mathbf{\Phi}$ by noting that $\mathbf{\Phi}$ could have been written $\mathbf{\Phi} = \mathbf{\Phi}_s + \mathbf{\Phi}_t$, where $\mathbf{\Phi}_t$, the other component of $\mathbf{\Phi}$, is the triplet function $\mathbf{\Phi}_t = 1/2(w_1w_2 - w_2w_1)(\alpha\beta + \beta\alpha)$. We also note that $\mathbf{\Phi}_s$ is not a determinantal function but is the linear combination $\mathbf{\Phi}_s = 1/2Det\,(w_1\alpha, w_2\beta) - 1/2Det\,(w_1\beta, w_2\alpha)$. Staes of pure spatial symmetry may also be projected from mixed-symmetry determinantal functions.

Symmetry projection can improve a wavefunction, as it eliminates components that cannot be present in an exact solution to the Schrödinger equations. Conceptually, it may also be regarded as falling within the independent-particle model, as it does not impair the introduction and use of single-particle concepts such as orbitals and orbital energies. In fact, the antisymmetrizer $\mathcal{A}$ is a permutational-symmetry projection operator, and is fundamentally very similar to the projectors for other types of symmetry.

Once one accepts the notion of using a symmetry-projected state, the most logical approach would be to find the $\mathbf{\Phi}$ whose projection $\mathcal{P}\mathbf{\Phi}$ minimizes the expression

$$< \mathcal{H} > = < \mathcal{P}\,\mathbf{\Phi} \,|\, \mathsf{H} \,|\, \mathcal{P}\,\mathbf{\Phi} > / < \mathcal{P}\,\mathbf{\Phi} \,|\, \mathcal{P}\,\mathbf{\Phi} > \tag{8.173}$$

where $\mathcal{P}$ is a projector for all symmetries of the state under study. This is the procedure referred to as the *Projected Hartree-Fock* (PHF) method. Note that the PHF wavefunction is not what would be obtained by projection of even an exact ordinary Hartree-Fock function. Since the latter function has been determined by optimal choice of a necessarily mixed-symmetry form, its pure-symmetry component will not itself be optimum, and will be a poorer wavefunction than that obtained from the PHF process, where it is the projected function which is optimized.

Considerable study has been directed toward spin-projected states. The situation is complicated by the fact that in most cases the coupling of individual spins to a given total-spin resultant is not unique, with different couplings defining the wavefunctions of different energies. Many different types of spin-projected calculations are possible, depending upon

whether all, or only certain spin couplings are considered. The three most used formulations are (1) those in which the spins associated with specific pairs of orbitals are coupled to zero resultants (*valence-bond or perfect-pairing* structures); (2) those in which all spinorbitals of the same spin orientation are coupled symmetrically; and (3) those in which an optimum spin coupling is found (spin-optimized Hartree-Fock).

As independent-particle wavefunctions, the unrestricted functions are underniably superior to the RHF function, and PHF results have in fact been used as a basis for discussing a number of chemically interesting systems. However, the situation becomes less clear when an independent-particle calculation is used as a starting point for correlation-energy studies. The deficiencies of the RHF model are then easily corrected by subsequent calculation steps, and increasing importance must be attached to the convenience of the starting point. Most correlation studies use a basis set of orbitals derived from independent-particle calculations, and are easier to carry out when the basis contains one orthogonal set of spatial orbitals.

K. Properties

Although much can be learned from a direct examination of the Hartree-Fock energies and wavefunctions, a full understanding of the physical content of the independent-particle model can only be obtained by the calculation of appropriate properties based on Hartree-Fock wavefunctions. Some properties, such as the electric dipole moment, can be thought of as describing features of the wavefunction. But other, as for example the electric polarizability, can only be understood in terms of the response of a many-particle system to an appropriate external perturbation. In particular, the energy change caused by the perturbation may be related to the observed value of the corresponding property. In the case of the polarizability, the perturbation is the interaction of all charges of the system with an electric field $\mathcal{E}$, and (in the absence of a preexisting dipole moment) the perturbation changes the energy by an amount $-1/2\mathcal{E}.\alpha\mathcal{E}$ where α is the polarizability tensor. If $\mathcal{E}$ is directed along a principal axis of the polarizability tensor (say the z axis), then this energy reduces to $-1/2\alpha_{zz}\mathcal{E}^2$.

Properties that describe features of the wavefunction can also be identified in formulations involving the energetic response of a system to an appropriate perturbation. For example, the application of an electric field $\mathcal{E}$ to a system with electric dipole moment μ will result in an energy change $-\mu.\mathcal{E} - 1/2\mathcal{E}.\alpha\mathcal{E}$. If the field is made large enough, the dipole moment it induces will no longer be linear in the field strength, and a complete representation of the change in energy with $\mathcal{E}$ will contain further terms. The coefficient of each power of $\mathcal{E}$ in the energy change is therefore in a certain sense a property of the many-particle system.

It may appear that we have identified two divergent methods for defining those properties that are directly identifiable with features of the wavefunction of a many-particle system. However, it is clear that from an operational point of view, the experimental basis for the measurement of a property must be in terms of the response of a system to a suitable perturbation. As we have already pointed out, in the case of the electric dipole moment the perturbation will be that caused by an electric field (presumably weak enough that its presence will not produce a significant induced dipole moment). If, as for the electric dipole moment, a linear relation exists (in the limit of small perturbations) between the magnitude of the perturbation and the energy change thereby produced, the property is termed *first-order*. Properties that are identifiable with features of a system wavefunction are necessarily first-order. It is possible to show that for exact wavefunctions, a first-order property defined by response to a perturbation can also be equivalently defined as an expectation value based on the system wavefunction. However, it is possible to find approximate wavefunctions for a many-particle system that do not preserve this equivalence, and in such cases the discrepancy will be an indication of inaccuracies in the calculations by one, or possibly both approaches.

Some properties, such as the electric polarizability, inherently involve energy changes associated with rearrangements of a many-particle wavefunction. If, as for the polarizability, the energy change is proportional to the square of the magnitude of the relevant perturbation, the property is termed *second- order*. As we shall see, second-order properties have values that depend not only upon the unperturbed wavefunction but also upon the entire energy spectrum of the many-particle system. Although we shall not discuss them, it is possible

to analyze the energy changes proportional to third or higher powers of the magnitude of a perturbation; when the perturbation is that due to an electromagnetic field, these third and higher-order properties are called hyperpolarizabilities.

Properties may also be classified according to the number of particles whose coordinates are coupled in the relevant perturbation. Thus, dipole moments and polarizabilities are *one-particle* properties, while spin-spin interactions are *two-particle* properties. Properties may also be time-independent or time-dependent.

L. Calculation of Properties

We proceed now to a discussion of those properties associated with a general perturbation to a many-particle Hamiltonian H. We write the perturbation in the form $g\mathsf{W}$, where W is an operator characteristic of the perturbation and g is a parameter controlling its magnitude. For example, if the properties under consideration are the electric- dipole component μ_z and the polarizability component α_{zz}, g would be the magnitude $\mathcal{E}$ of an electric field in the $+z$ direction, and (suppressing constant factors) $\mathcal{W}$ would be a sum of z coordinates: $\mathsf{W} = -\sum_i z(i)$. We write the perturbed Hamiltonian as

$$\mathsf{H}(g) \;=\; \mathsf{H} + g\,\mathsf{W} \tag{8.174}$$

The perturbed problem defined by $\mathsf{H}(g)$ will have eigenfunctions and eigenvalues that may be expanded about the corresponding unperturbed quantities:

$$\Psi(g) \;=\; \Psi + g\,\Psi^{(1)} + g^2\,\psi^{(2)} + \cdots \tag{8.175}$$

$$E(g) \;=\; E + g\,E^{(1)} + g^2\,E^{(2)} + \cdots \tag{8.176}$$

We can without loss of generality require $\Psi^{(1)}$, $\Psi^{(2)}$, to be orthogonal to Ψ. The energies $E^{(1)}$, $E^{(2)}, \cdots$ are proportional to the first, second, $\cdots$ order properties associated with the perturbation $g\mathsf{W}$. For the electric-field perturbation, we have already observed $\Delta E(\mathcal{E}) = -\mu_z\mathcal{E} - 1/2\alpha_{zz}\mathcal{E}^2 + \cdots$, so that (with $g = \mathcal{E}$) we may identify $\mu_z = -E^{(1)}$ and $1/2\alpha_{zz} = -E^{(2)}$.

Another way of representing $E^{(1)}$ and $E^{(2)}$ is in terms of derivatives of $E(g)$ evaluated at $g = 0$:

$$E^{(1)} = \left.\frac{dE(g)}{dg}\right|_{g=0} \tag{8.177}$$

$$E^{(2)} = \frac{1}{2}\left.\frac{d^2E(g)}{dg^2}\right|_{g=0} \tag{8.178}$$

If we assume that Ψ and E are known exactly, we can proceed to a formal derivation of $E^{(1)}$, $E^{(2)}, \cdots$ We start by writing the Schrödinger equation for $\mathsf{H}(g)$:

$$[\mathsf{H}(g) - E(g)]\,\Psi(g) = 0 \tag{8.179}$$

Expanding all quantities in Eq. (8.179) according to Eqs. (8.174 - 8.176), we have

$$(\mathsf{H} + g\,\mathsf{W} - E - g\,E^{(1)}\cdots)(\Psi + g\,\Psi^{(1)} + \cdots) = 0 \tag{8.180}$$

Since Eq. (8.180) must be satisfied for arbitrary g, the coefficient of each power of g on its left-hand side must vanish. From the coefficients of g^0, g^1, g^2, $\cdots$,

$$(\mathsf{H} - E)\,\Psi = 0 \tag{8.181}$$

$$(\mathsf{H} - E)\,\Psi^{(1)} + (\mathsf{W} - E^{(1)})\,\Psi = 0 \tag{8.182}$$

$$(\mathsf{H} - E)\,\Psi^{(2)} + (\mathsf{W} - E^{(1)})\,\Psi^{(1)} - E^{(2)}\,\Psi = 0 \tag{8.183}$$

The first of the preceding equations is satisfied by virtue of the fact that Ψ and E satisfy the unperturbed Schrödinger equation. The second equation permits the determination of $\Psi^{(1)}$ and $E^{(1)}$, and the third (and subsequent) equations lead to expression for $\Psi^{(2)}$, $E^{(2)}$, and subsequent contributions. Taking the scalar product of Eq. (8.182) with $< \Psi|$, and using the fact that $< \Psi|(\mathsf{H} - E) = 0$, we find (choosing Ψ to be normalized)

$$E^{(1)} = < \Psi \,|\, \mathsf{W} \,|\, \Psi > \tag{8.184}$$

This result confirms our earlier observation that in an exact calculation, a first-order property may be specified equivalently by the energy response $E^{(1)}$ or by an expectation value based on the unperturbed wavefunction Ψ.

In order to obtain the energy response $E^{(2)}$, we will need the first-order wavefunction change $\Psi^{(1)}$. We proceed by a method that is valid only if the unperturbed energy E is nondegenerate. Degeneracy of E leads to complications beyond the scope of the current presentation. We solve Eq. (8.182) for $\Psi^{(1)}$ through left-multiplication by $\mathcal{Q}/(\mathsf{H}-E)$, where $\mathcal{Q}$ is the projector for the orthogonal complement to Ψ and the notation is that introduced in Eq. (1.100). Because E is nondegenerate $\mathcal{Q}/(\mathsf{H}-E)$ can be inverted in the subspace defined by $\mathcal{Q}$ so $\mathcal{Q}/(\mathsf{H}-E)$ exists and the procedure we are using will be well defined. The left-multiplication yields

$$\frac{\mathcal{Q}}{\mathsf{H}-E}\,(\mathsf{H}-E)\,\Psi^{(1)}+\frac{\mathcal{Q}}{\mathsf{H}-E}\,(\mathsf{W}-E^{(1)})\,\Psi\;=\;0 \tag{8.185}$$

which, according to Eq. (1.102), reduces to

$$\mathcal{Q}\,\Psi^{(1)}\;=\;-\frac{\mathcal{Q}}{\mathsf{H}-E}\,(\mathsf{W}-E^{(1)})\,\Psi \tag{8.186}$$

Because $\mathcal{Q}\Psi=0$, we may drop the $E^{(1)}$. Remembering that $\Psi^{(1)}$ has been assumed orthogonal to Ψ, we replace $\mathcal{Q}\Psi^{(1)}$ by $\Psi^{(1)}$, reaching

$$\Psi^{(1)}\;=\;-\frac{\mathcal{Q}}{\mathsf{H}-E}\,\mathsf{W}\,\Psi \tag{8.187}$$

We now proceed to obtain $E^{(2)}$ by taking the scalar product of Eq. (8.183) with Ψ^*. We see that the term $<\Psi|(\mathsf{H}-E)|\Psi^{(2)}$ vanishes by letting $\mathsf{H}-E$ act to the left, and the term $<\psi|E^{(1)}|\Psi^{(1)}$ vanishes through the orthogonality of Ψ and $\Psi^{(1)}$. The surviving terms yield

$$E^{(2)}\;=<\Psi\,|\,\mathsf{W}\,|\,\Psi^{(1)}> \tag{8.188}$$

Substituting from Eq. (8.187),

$$E^{(2)}\;=\;-\left\langle\Psi\,\middle|\,\mathsf{W}\,\frac{\mathcal{Q}}{\mathsf{H}-E}\,\mathsf{W}\,\middle|\,\Psi\right\rangle \tag{8.189}$$

As pointed out before, $E^{(2)}$ depends upon the entire spectrum of H as well as the wavefunction Ψ. To evaluate Eq. (8.189), $\mathcal{Q}$ can be written in terms of the eigenfunctions Ψ_i of $\mathcal{H}$ with respective eigenvalues E_i. Then

$$\frac{\mathcal{Q}}{\mathsf{H}-E}\;=\;\sum_i{}'\,|\Psi_i>\,\frac{1}{E_i-E}\,<\Psi_i| \tag{8.190}$$

where the prime means Ψ is omitted from the sum, and

$$E^{(2)} = -\sum_i{}' \frac{<\Psi\,|\,\mathsf{W}\,|\,\Psi_i><\Psi_i\,|\,\mathsf{W}\,|\,\Psi>}{E_i - E} \tag{8.191}$$

The foregoing analysis is based on the traditional assumption that $\Psi^{(1)}$ is orthogonal to Ψ. As already pointed out, this is not a restrictive assumption because any $\Psi(g)$ can be written as Ψ plus a correction orthogonal thereto. However, if we look more closely at the equation defining $\Psi^{(1)}$, we see that that equation is, in fact, indeterminate with respect to component of $\Psi^{(1)}$ proportional to Ψ. In other words, addition of a term $c\Psi$ to $\Psi^{(1)}$ would have no effect in Eq. (8.182), as such a term would be annihilated by $\mathsf{H} - E$. Moreover, the derivation of Eq. (8.184) does not depend upon the form of $\Psi^{(1)}$, so that the value of $E^{(1)}$ would also remain unaffected by addition of $c\Psi$ to $\Psi^{(1)}$.

Looking now at the derivation of $E^{(2)}$, we find that if Ψ and $\Psi^{(1)}$ are nonorthogonal, Eq. (8.188) must be replaced (for normalized Ψ) by

$$E^{(2)} = <\Psi\,|\,\mathsf{W} - E^{(1)}\,|\,\Psi^{(1)}> \tag{8.192}$$

If we now substitute

$$\Psi^{(1)} = c\,\Psi + \mathcal{Q}\,\Psi^{(1)} \tag{8.193}$$

where $\mathcal{Q}\Psi^{(1)}$ is what is actually given by Eq. (8.187), we get

$$\begin{aligned} E^{(2)} &= c\,<\Psi\,|\,\mathsf{W}\,|\,\Psi> + <\Psi\,|\,\mathsf{W}\,|\,\mathcal{Q}\,\Psi^{(1)}> - c\,E^{(1)} \\ &= <\Psi\,|\,\mathsf{W}\,|\,\mathcal{Q}\,\Psi^{(1)}> \end{aligned} \tag{8.194}$$

which leads to Eqs. (8.189) - (8.191). We conclude that both $E^{(1)}$ and $E^{(2)}$ remain invariant even though the Ψ component of $\Psi^{(1)}$ is arbitrary. This fact is particularly relevant because perturbative developments of $\Psi^{(1)}$ are not ordinarily orthogonal to Ψ. The true significance of the assumption $<\Psi|\Psi^{(1)}>= 0$ is that it yields the $\Psi^{(1)}$ of minimum norm, in which case $\Psi(g)$ is as close as possible to Ψ

The derivations of the preceding subsection are not sufficient for treatment of approximate wavefunctions, because they made liberal use of the assumptions that Ψ was an exact wavefunction of the unperturbed Hamiltonian. For approximate wavefunctions it will be necessary to reesamine the analysis and to proceed in a manner consistent with the approximations already in use. Here we conduct such a further study for wavefunctions obtained by the Hartree-Fock procedure. For simplicity, we restrict discussion to one-particle properties, for which the operator W is of the form $\sum_i \mathsf{w}(i)$.

The quantity ultimately needed, $E(g)$, is obtained by constructing a Fock operator $\mathsf{F}(g)$, finding its eigenfunctions $\phi_\mu(g)$ and eigenvalues $\epsilon_\mu(g)$, and forming

$$E(g) \;=\; \sum_\alpha \epsilon_\alpha(g) - \frac{1}{2}\sum_{\alpha\beta} < \phi_\alpha(g)\,\phi_\beta(g)\,|\,\mathsf{v}\,|\,\phi_\alpha(g)\,\phi_\beta(g) - \phi_\beta(g)\,\phi_\alpha(g) > \tag{8.195}$$

We see that $E(g)$ is affected not only by changes in the Fock eigenvalues, but also through the two-particle term by which the eigenalue sum is modified. The Fock operator $\mathsf{F}(g)$ also contains terms involving the perturbed single-particle states:

$$\mathsf{F}(g) \;=\; \mathsf{u} + g\,\mathsf{w} + \sum_{\mu\nu\alpha} |\phi_\mu >< \phi_\mu\,\phi_\alpha(g)\,|\,\mathsf{v}\,|\,\phi_\nu\,\phi_\alpha(g)\,\phi_\nu >< \phi_\nu| \tag{8.196}$$

Notice that in Eq. (8.196) the μ and ν summations simply provide resolutions of the identity, and the ϕ_μ and ϕ_ν could have been any complete sets of states. For later convenience we have chosen to take them as the unperturbed states. We are continuing with our usual convention that μ and ν summations are unrestricted, while α summations are over occupied, single-particle states only.

As before, we expand all the g-dependent quantities in series:

$$E(g) \;=\; E + g\,E^{(1)} + g^2\,E^{(2)} + \cdots \tag{8.197}$$

$$\mathsf{F}(g) \;=\; \mathsf{F} + g\,\mathsf{F}^{(1)} + g^2\,\mathsf{F}^{(2)} + \cdots \tag{8.198}$$

$$\phi_\mu(g) \;=\; \phi_\mu + g\,\phi_\mu^{(1)} + g^2\,\phi^{(2)} + \cdots \tag{8.199}$$

$$\epsilon_\mu(g) \;=\; \epsilon_\mu + g\,\epsilon_\mu^{(1)} + g^2\,\epsilon_\mu^{(2)} + \cdots \tag{8.200}$$

By equating the coefficients of successive powers of g to zero in the Fock equation $[\mathsf{F}(g) - \epsilon_\mu(g)]\phi_\mu(g) = 0$, we reach

$$(\mathsf{F} - \epsilon_\mu)\,\phi_\mu \;=\; 0 \tag{8.201}$$

$$(\mathsf{F} - \epsilon_\mu)\,\phi_\mu^{(1)} + (\mathsf{F}^{(1)} - \epsilon_\mu^{(1)})\,\phi_\mu \;=\; 0 \tag{8.202}$$

$$(\mathsf{F} - \epsilon_\mu)\,\phi_\mu^{(2)} + (\mathsf{F}^{(1)} - \epsilon_\mu^{(1)})\,\phi_\mu^{(1)} + (\mathsf{F}^{(2)} - \epsilon_\mu^{(2)})\,\phi_\mu \;=\; 0 \tag{8.203}$$

Equations (8.201) - (8.203) are similar to Eqs. (8.181) - (8.183), with the exception that Eq. (8.138) does not contain a term analogous to the $\mathsf{F}^{(2)}$ that appears in Eq. (8.203). Proceeding as for the earliner set of equations, we may deduce

$$\epsilon_\mu^{(1)} \;=< \phi_\mu \,|\, \mathsf{F}^{(1)} \,|\, \phi_\mu > \tag{8.204}$$

$$\phi_\mu^{(1)} \;=\; -\left(\frac{\mathcal{Q}_\mu}{\mathsf{F} - \epsilon_\mu}\right) \mathsf{F}^{(1)}\,\phi_\mu \tag{8.205}$$

$$\epsilon_\mu^{(2)} \;=< \phi_\mu \,|\, \mathsf{F} \,|\, \phi_\mu > + < \phi_\mu \,|\, \mathsf{F}^{(1)} \,|\, \phi_\mu^{(1)} > \tag{8.206}$$

The notation $\mathcal{Q}_\mu$ refers to the projector for the orthogonal complement to ϕ_μ; the formulation is valid only if the (unperturbed) ϵ_μ are nondegenerate.

In order to proceed, we will need to make more explicit the formulas for $E^{(1)}$, $E^{(2)}$, $\mathsf{F}^{(1)}$, and $\mathsf{F}^{(2)}$. Introducing Eqs. (8.199) and (8.200) into Eqs. (8.195) and (8.196), and using the space-saving notation

$$< \phi_\mu\,\phi_\nu \,|\, \tilde{\mathsf{v}} \,|\, \phi_\lambda\,\phi_\sigma > =< \phi_\mu\,\phi_\nu \,|\mathsf{v}\,|\, \phi_\lambda\,\phi_\sigma - \phi_\sigma\,\phi_\lambda > \tag{8.207}$$

we obtain

$$E^{(1)} \;=\; \sum_\alpha \epsilon_\alpha^{(1)} - \sum_{\alpha\beta} \left[< \phi_\alpha^{(1)}\,\phi_\beta \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\,\phi_\beta > + < \phi_\alpha\,\phi_\beta \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha^{(1)}\,\phi_\beta > \right] \tag{8.208}$$

$$E^{(2)} = \sum_{\alpha} \epsilon_{\alpha}^{(2)} - \sum_{\alpha\beta} \Big[< \phi_{\alpha}^{(2)}\, \phi_{\beta} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}\, \phi_{\beta} > + < \phi_{\alpha}\, \phi_{\beta} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}^{(2)}\, \phi_{\beta} >$$
$$+ \frac{1}{2} < \phi_{\alpha}^{(1)}\, \phi_{\beta}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}\, \phi_{\beta} > + \frac{1}{2} < \phi_{\alpha}\, \phi_{\beta} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}^{(1)}\, \phi_{\beta}^{(1)} >$$
$$+ < \phi_{\alpha}^{(1)}\, \phi_{\beta} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}^{(1)}\, \phi_{\beta} > + < \phi_{\alpha}^{(1)}\, \phi_{\beta} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\alpha}\, \phi_{\beta}^{(1)} > \Big] \tag{8.209}$$

$$\mathsf{F}^{(1)} = \mathsf{w} + \sum_{\mu\nu\alpha} |\phi_{\mu} > \Big[< \phi_{\mu}\, \phi_{\alpha}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\nu}\, \phi_{\alpha} > + < \phi_{\mu}\, \phi_{\alpha} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\nu}\, \phi_{\alpha}^{(1)} > \Big] < \phi_{\nu}| \tag{8.210}$$

$$\mathsf{F}^{(2)} = \sum_{\mu\nu\alpha} |\phi_{\mu} > \Big[< \phi_{\mu}\, \phi_{\alpha}^{(2)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\nu}\, \phi_{\alpha} > + < \phi_{\mu}\, \phi_{\alpha} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\nu}\, \phi_{\alpha}^{(2)} >$$
$$+ < \phi_{\mu}\, \phi_{\alpha}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\nu}\, \phi_{\alpha}^{(1)} > \Big] < \phi_{\nu}| \tag{8.211}$$

From Eqs. (8.210) and (8.211) we may obtain the following matrix element formulas

$$< \phi_{\mu} \,|\, \mathsf{F}^{(1)} \,|\, \phi_{\mu} > = < \phi_{\mu} \,|\, \mathsf{w} \,|\, \phi_{\mu} > + \sum_{\alpha} \Big[< \phi_{\mu}\, \phi_{\alpha}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}\, \phi_{\alpha} >$$
$$+ < \phi_{\mu}\, \phi_{\alpha} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}\, \phi_{\alpha}^{(1)} > \Big] \tag{8.212}$$

$$< \phi_{\mu} \,|\, \mathsf{F}^{(1)} \,|\, \phi_{\mu}^{(1)} > = < \phi_{\mu} \,|\, \mathsf{w} \,|\, \phi_{\mu}^{(1)} > + \sum_{\alpha} \Big[< \phi_{\mu}\, \phi_{\alpha}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}^{(1)}\, \phi_{\alpha} >$$
$$+ < \phi_{\mu}\, \phi_{\alpha} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}^{(1)}\, \phi_{\alpha}^{(1)} > \Big] \tag{8.213}$$

$$< \phi_{\mu} \,|\, \mathsf{F}^{(2)} \,|\, \phi_{\mu} > = \sum_{\alpha} \Big[< \phi_{\mu}\, \phi_{\alpha}^{(2)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}\, \phi_{\alpha} > + < \phi_{\mu}\, \phi_{\alpha} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}\, \phi_{\alpha}^{(2)} >$$
$$+ < \phi_{\mu}\, \phi_{\alpha}^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_{\mu}\, \phi_{\alpha}^{(1)} > \Big] \tag{8.214}$$

We also have, because $\epsilon_{\mu}^{(2)}$ must be real,

$$< \phi_{\mu} \,|\, \mathsf{F}^{(1)} \,|\, \phi_{\mu}^{(1)} > = < \phi_{\mu}^{(1)} \,|\, \mathsf{F}^{(1)} \,|\, \phi_{\mu} > \tag{8.215}$$

When these formulas are inserted in the expressions for $\epsilon_{\alpha}^{(1)}$ and $\epsilon_{\alpha}^{(2)}$, Eqs. (8.208) and (8.209) can be simplified considerably, to

$$E^{(1)} = \sum_{\alpha} < \phi_{\alpha} \,|\, \mathsf{w} \,|\, \phi_{\alpha} > \tag{8.216}$$

$$E^{(2)} = \frac{1}{2} \sum_{\alpha} \left[< \phi_\alpha \,|\, \mathsf{w} \,|\, \phi_\alpha^{(1)} > + < \phi_\alpha^{(1)} \,|\, \mathsf{w} \,|\, \phi_\alpha > \right] \tag{8.217}$$

Equation (8.216) shows that the intuitive result has been recovered: that a firs-order property is, within the Hartree-Fock model, represented by an appropriate expectation value. The expression for $E^{(2)}$, however, reveals a more complex situation. Equation (8.217) gives $E^{(2)}$ in terms of the wavefunction changes $\phi_\alpha^{(1)}$, which in turn are given by Eq. (8.205). But when an explicit expression for $\mathsf{F}^{(1)}$ is inserted on the right side of Eq. (8.205), the $\phi_\alpha^{(1)}$ occur on both sides of that equation, so that the set of such equations is not in a solved format. More explicitly, Eq. (8.205) may be written

$$\phi_\mu^{(1)} = -\sum_{\nu;\nu\neq\mu} \left(\frac{1}{\epsilon_\nu - \epsilon_\mu} \right) \left[< \phi_\nu \,|\, \mathsf{w} \,|\, \phi_\mu > + \sum_\alpha < \phi_\nu \, \phi_\alpha^{(1)} \,|\, \tilde{\mathsf{v}} \,|\, \phi_\mu \, \phi_\alpha > \right.$$
$$\left. + \sum_\alpha < \phi_\nu \, \phi_\alpha \,|\, \tilde{\mathsf{v}} \,|\, \phi_\mu \, \phi_\alpha^{(1)} > \right] \phi_\nu \qquad\qquad (each\ \mu) \tag{8.218}$$

If we introduce expansions of the $\phi_\mu^{(1)}$ in terms of the unperturbed single-particle states,

$$\phi_\mu^{(1)} = \sum_\lambda c_{\lambda\mu} \, \phi_\lambda \tag{8.219}$$

Eq. (8.218) leads to

$$c_{\nu\mu} = \left(\frac{-1}{\epsilon_\nu - \epsilon_\mu} \right) \left[< \phi_\nu \,|\, \mathsf{w} \,|\, \phi_\mu > + \sum_{\alpha\lambda} c_{\lambda\alpha}^* \, < \phi_\nu \, \phi_\lambda \,|\, \tilde{\mathsf{v}} \,|\, \phi_\mu \, \phi_\alpha > \right.$$
$$\left. + \sum_{\alpha\lambda} c_{\lambda\alpha} \, < \phi_\nu \, \phi_\alpha \,|\, \tilde{\mathsf{v}} \,|\, \phi_\mu \, \phi_\lambda > \right] \qquad (each\ \mu, \nu; \mu \neq \nu) \tag{8.220}$$

Now, the $c_{\nu\mu}$ are not completely independent, as they must be consistent with the orthogonality of the perturbed single-particle states $\phi_\mu(g)$. The coefficient of g in the orthogonality integral $< \phi_\mu(g) | \phi_\nu(g) >$ is $c_{\nu\mu}^* + c_{\mu\nu}$, and upon equating this coefficient to zero we obtain

$$c_{\nu\mu}^* = -c_{\mu\nu} \qquad\qquad (\mu \neq \nu) \tag{8.221}$$

Inserting this result into Eq. (8.220), and separating the λ summations into occupied-state and unoccupied-state parts, we find that the occupied-state portions of the two summations cancel against each other, leaving λ restricted to unoccupied states. This result is to be expected, as the $\phi_\mu^{(1)}$ should not be influenced by rearrangements completely within the

occupied states. A similar argument applies if the expansion of Eq. (8.219) is inserted into Eq. (8.217) for $E^{(2)}$. Therefore, even though the coefficients $c_{\nu\alpha}$ will be nonzero for both occupied and unoccupied ν, the relevant equations are

$$E^{(2)} = \frac{1}{2}\sum_{\alpha r}(c_{r\alpha} < \phi_\alpha \,|\, \mathsf{w} \,|\, \phi_r > + c^*_{r\alpha} < \phi_r \,|\, \mathsf{w} \,|\, \phi_\alpha >) \tag{8.222}$$

$$c_{r\alpha} = \frac{-1}{\epsilon_r - \epsilon_\alpha}\left[< \phi_r \,|\, \mathsf{w} \,|\, \phi_\alpha > + \sum_{\beta s}(c^*_{s\beta} < \phi_r\, \phi_s \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_\beta > + c_{s\beta} < \phi_r \phi_\beta \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_s >)\right] \tag{8.223}$$

We now see that the characterization of the $\phi^{(1)}_\alpha$ entails solving the simultaneous linear inhomogeneous equation set given in Eq. (8.223). For a basis of M single-particle states, of which N are occupied, the set will contain $N(M-N)$ complex equations. If, as is normally the case, the unperturbed basis has been chosen in such a way that all nonvanishing two-particle integrals are real, the equations can easily be separated into noninteracting real and imaginary components. We have

$$Re(c_{r\alpha}) = \frac{-1}{\epsilon_r - \epsilon_\alpha}\left[Re < \phi_r \,|\, \mathsf{w} \,|\, \phi_\alpha > + \sum_{s\beta} Re(c_{s\beta})\,(< \phi_r\, \phi_s \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_\beta > + < \phi_r\, \phi_\beta \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_s >)\right] \tag{8.224}$$

$$Im(c_{r\alpha}) = \frac{-1}{\epsilon_r - \epsilon_\alpha}\left[Im < \phi_r \,|\, \mathsf{w} \,|\, \phi_\alpha > + \sum_{s\beta} Im(c_{s\beta})\,(< \phi_r\, \phi_\beta \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_s > - < \phi_r\, \phi_s \,|\, \tilde{\mathsf{v}} \,|\, \phi_\alpha\, \phi_\beta >)\right] \tag{8.225}$$

If, in addition, the operator w has only real matrix elements, we can see that the $c_{r\alpha}$ will also be real.

For many practical problems the dimensionality of the equation systems impied by Eqs. (8.224) and (8.225) is small enough that the necessary manipulations can be carried out in a straightforward manner. However, for very large systems it may prove advisable to proceed in an approximate fashion. A number of schemes can be considered for eliminating

or estimating the effect of the s, β summations in Eqs. (8.224) and (8.225); because these summations have the effect of coupling different $c_{r\alpha}$ the procedures for handling them have been identified by the name *coupled Hartree-Fock*.

IX. THE CONFIGURATION-INTERACTION METHOD

A. Preliminary Remarks

In the previous section we introduced a variety of independent-particle models for the wavefunction of a many-fermion system. As we have observed, independent- particle concepts provide an intuitively appealing framework for the discussion of many-fermion problems, and in addition can often yield qualitatively correct results for quantities such as the total energy. Unfortunately, however, there are many physically important problems whose solutions depend upon the deviations of actual wavefunctions from the independent-particle form, and such problems require the introduction and use of more sophisticated forms of many-fermion wavefunctions. Deviations from the independent-particle results are most extreme for properties that depend on the details of the interparticle interaction or on the variation in its magnitude between different states. Thus, quantities such as excitation energies, molecular dissociation energies, polarizabilities, and relativistic effects really must be treated by methods beyond the independent-particle model. This section is concerned with correcting the independent-particle model in such a way as to include those physical effects needed for accurate calculation of important propeties.

In this section we will be concerned with the most widely used and mathematically simplest technique for correcting an independent-particle wavefunction, namely the *configurration-interaction* (CI) expansion. The CI method, sometimes (and perhaps more appropriately) referred to as the *method of superposition of configurations*, involves the expansion of the many-fermion wavefunction as a linear combination of determinantal functions (configurations), with the coefficients of the various determinants found by application of the variational principle.

B. The CI Wavefunction

Given an orthonormal set of one-particle functions $\{\phi_\mu\}$, the most general, fully-antisymmetric N-fermion wavefunction can be written as a linear combination of determinantal functions, each of which is of the form illustrated in Eq. (7.1), i.e.,

$$[\phi_\mu\,\phi_\nu\cdots\cdots] \equiv (N!)^{1/2}\,\mathcal{A}\,(\phi_\mu\,\phi_\nu\cdots\cdots) \tag{9.1}$$

A CI wavefunction consists of a linear combination of those determinantal functions that upon practical considerations it is decided to include. It is convenient to use the notation of Section 7.C and introduce a reference determinant (assumed not orthogonal to the final CI wavefunction)

$$\Phi \;=\; [\phi_\alpha\,\phi_\beta\cdots\cdots] \;=\; \mathbf{a}^\dagger\,\mathbf{a}^\dagger\cdots\cdots|0> \tag{9.2}$$

We will also require "excited determinants" in which one or more of $\phi_\alpha, \phi_\beta, \cdots\cdots$ are replaced by a corresponding number of states $\phi_r, \phi_s, \cdots\cdots$ not present in Φ. The excited determinants are written

$$\begin{aligned}
\phi^r_\alpha \;&=\; \mathbf{a}^\dagger_r\,\mathbf{a}_\alpha\,\Phi \;=\; \mathsf{A}^r_\alpha\,\Phi \\
\Phi^{rs}_{\alpha\beta} \;&=\; \mathbf{a}^\dagger_r\,\mathbf{a}^\dagger_s\,\mathbf{a}_\beta\,\mathbf{a}_\alpha\,\Phi \;=\; \mathsf{A}^{rs}_{\alpha\beta}\,\Phi \\
&\cdots\cdots \\
\Phi^R_\Delta \;&=\; \mathsf{A}^R_\Delta\,\Phi
\end{aligned} \tag{9.3}$$

This notation makes Φ^R_Δ the determinantal function resulting if the one-particle functions in Φ and listed in Δ are replaced, in order, by those listed in R. The CI wavefunction Ψ is therefore of the form

$$\Psi \;=\; c_0\,\Phi + \sum_{\alpha r} c^r_\alpha\,\Phi^r_\alpha + \sum_{(\alpha\beta)(rs)} c^{rs}_{\alpha\beta}\,\Phi^{rs}_{\alpha\beta} + \cdots \tag{9.4}$$

with the coefficients $c_0(\neq 0)$, $c^r_\alpha, \cdots$ to be determined variationally. Observe that the summation index notation limits the expansion to terms having distinct index sets; terms differing only by an index permutation would be linearly dependent. To avoid awkwardness in later

equations, we adopt the convention that Eq. (9.4) contains only index sets arrayed in lexical order; if a c_Δ^R with other orderings occurs in any equation, it is to be reduced to the corresponding lexically ordered coefficient with the introduction of a factor $(-)^P$, where P is the parity of the permutation to the lexical order.

The effectiveness of the CI expansion depends, of course, upon the rapidity of its convergence. Best results are normally obtained if the reference function Φ provides a reasonable independent-particle model for the system under study. In that case the succeeding terms of the CI expansion describes corrections to the reference function depending upon the coordinates of one, two, ... particles, and under certain circumstances the higher-order terms may become small. For that reason, and for further reasons of computational convenience to be identified later, Φ is frequently taken to be a Hartree-Fock wavefunction. However, the CI formalism does not require use of any particular reference state, and there may be applications where other choices could be more suitable. We also note that in principle it is possible to make expansions of the form of Eq. (9.4) with nonorthogonal one-particle functions. Except for systems with very few particles, the nonorthogonality procduces such severe computational complications that it becomes impractical to complete the CI calculations.

Another way of expressing the relationship between Φ and Ψ given in Eq. (9.4) is to write

$$\Psi = (c_0 + \mathsf{C})\,\Phi \tag{9.5}$$

where C is an operator with definition

$$\mathsf{C} = \sum_{\alpha r} c_\alpha^r\, \mathsf{A}_\alpha^r + \sum_{(\alpha\beta)(rs)} c_{\alpha\beta}^{rs}\, \mathsf{A}_{\alpha\beta}^{rs} + \cdots \tag{9.6}$$

Since the normalization of Ψ is at our disposal and since c_0 is by hypothesis nonzero, we may arbitrarily set $c_0 = 1$, with corresponding proportional changes in all the other coefficients. We then have a formula that proves useful in the comparison of many-fermion wavefunctions:

$$\Psi = (1 + \mathsf{C})\,\Phi \tag{9.7}$$

The Ψ given by Eq. (9.7) possesses the so-called intermediate normalization

$$< \Phi \mid \Psi > = < \Phi \mid \Phi > = 1 \tag{9.8}$$

Although we realize that it is not the most convenient choice for actual CI calculations, for formal reasons, we will use intermediate normalization. It is easier to solve simulataneous homogeneous equations for arbitrary scale than to handle a less symmetrical but equivalent inhomogeneous problem.

C. The CI Equations

With a normalized (to unity) wavefunction Ψ of the form given in Eq. (9.4), the expectation value of the energy E for a Hamiltonian H is

$$E = < \Psi \mid \mathsf{H} \mid \Psi > = \sum_{R\Delta; R'\Delta'} c_{\Delta}^{R^*} \; < \Phi_{\Delta}^{R} \mid \mathsf{H} \mid \Phi_{\Delta'}^{R'} > \; c_{\Delta'}^{R'} \tag{9.9}$$

where the summations are over all distinct index sets consistent with N-fermion wavefunctions, including the empty set (correspondient to c_0 and Φ). The unit normalization requires that

$$< \Psi \mid \Psi > = \sum_{R\Delta} |c_{\Delta R}|^2 \; = \; 1 \tag{9.10}$$

We now invoke the variational principle and seek the values of the c_{Δ}^{R} that make the E of Eq. (9.9) a minimum, subject to the constraint of Eq. (9.10). It is most convenient to proceed by the method of Lagrange multipliers, which dictates that we seek an unconstrained minimum of $< \Psi|\mathsf{H}|\Psi > - \lambda < \Psi|\Psi >$, with λ ultimately chosen so that Eq. (9.10) is satisfied. The variation of this quantity with respect to c_{Δ}^{R} is

$$\begin{aligned} \delta[< \Psi \mid \mathsf{H} \mid \Psi > - \lambda < \Psi \mid \Psi >] \; = \; & \sum_{R'\Delta'} \left[\left(\delta c_{\delta}^{r^*} \right) \; < \Phi_{\Delta}^{R} \mid \mathsf{H} \mid \Phi_{\Delta'}^{R'} > \; c_{\Delta'}^{R'} \right. \\ & \left. + c_{\Delta'}^{R'^*} \; < \Phi_{\Delta'}^{R'} \mid \mathsf{H} \mid \Phi_{\Delta}^{R} > \; \left(\delta c_{\Delta}^{R} \right) \right] - \lambda \left[\left(\delta c_{\Delta}^{R^*} \right) c_{\Delta}^{R} + c_{\Delta}^{R^*} \left(\delta c_{\Delta}^{R} \right) \right] \; = \; 0 \end{aligned} \tag{9.11}$$

In deriving Eq. (9.11) we have used the fact that the Φ_{Δ}^{R} are orthonormal. Since the terms involving $\delta c_{\Delta}^{R^*}$ are the complex conjugates of those involving δc_{Δ}^{R}, it is necessary and

sufficient to require the quantities multiplying $\delta c_\Delta^{R^*}$ to vanish. This leads to the following CI equations:

$$\sum_{R'\Delta'} < \Phi_\Delta^R \,|\, \mathsf{H} \,|\, \Phi_{\Delta'}^{R'} > c_{\Delta'}^{R'} - \lambda\, c_\Delta^R \;=\; 0 \qquad\qquad (\forall\; R, \Delta) \tag{9.12}$$

As is well known, Eqs. (9.12) have nonzero values of the c_Δ^R as solutions only for certain values of λ. For these λ values the nonzero sets of c_Δ^R are arbitrarily scaled and can be chosen to satisfy Eq. (9.10) or, for that matter, any desired normalization including the intermediate normalization described by Eq. (9.8).

To find the physical significance of λ, we multiply each of the Eqs. (9.12) by the corresponding $c_\Delta^{R^*}$ and add all of the resulting equations together. We then arrive at

$$\sum_{R\Delta;R'\Delta'} c_\Delta^{R^*} \; < \Phi_\Delta^R \,|\, \mathsf{H} \,|\, \Phi_{\Delta'}^{R'} > c_{\Delta'}^{R'} - \lambda \sum_{\Delta R} |c_\Delta^R|^2 \;=\; 0 \tag{9.13}$$

Comparing with Eqs. (9.9) and (9.10), we find that $\lambda = E$ (irrespective of the normalization chosen). Noting that $< \Phi_\Delta^R | \Phi_{\Delta'}^{R'} >$ vanishes unless $R\Delta = R'\Delta'$ (in which case it is unity), we rewrite Eqs. (9.12) as

$$\sum_{R'\Delta'} < \Phi_\Delta^R \,|\, \mathsf{H} - E \,|\, \Phi_{\Delta'}^{R'} > c_{\Delta'}^{R'} \;=\; 0 \qquad\qquad (\forall\; R, \Delta) \tag{9.14}$$

The derivation of Eq. (9.14) remains valid if the expansion for Ψ is truncated to a finite set of $R\Delta$. In that case, there will be an Eq. (9.14) for each retained $R\Delta$, and the solution of these equations will give the minimum-energy Ψ of the truncated form.

Another way of obtaining Eqs. (9.14) is to start from the Schrödinger equation as applied to an (in principle) untruncated and therefore completely general wavefunction expanded as in Eq. (9.4):

$$(\mathsf{H} - E)\, \Psi \;=\; (\mathsf{H} - E) \sum_{\Delta' R'} c_{\Delta'}^{R'}\, \Phi_{\Delta'}^{R'} \;=\; 0 \tag{9.15}$$

If Eq. (9.15) is successively projected against all the Φ_Δ^R retained in the expansion of Ψ (including Φ itself), Eqs. (9.14) result. We may thus interpret Eqs. (9.14) as defining an approximate solution Ψ to the Schrödinger equation that satisfies the equation in projection upon the space in which the approximate Ψ is found. We show this explicitly by writing

$$\mathcal{P}\,(\mathsf{H} - E)\,\mathcal{P}\,\Psi \;=\; 0 \tag{9.16}$$

where $\mathcal{P}$ projects onto the space spanned by the configuration to be used. As we have already pointed out, the derivation of Eq. (9.14) shows hat the CI procedure has the additional feature of yielding he minimum energy expectation value within the space searched for Ψ.

D. Further Study of the CI Equations

Most actual calculations using the CI method have been accomplished by identifying Eq. (9.14) as a matrix eigenvalue problem, then solving it by standard techniques. However, some successive approximation procedures for large problems, and much additional insight, can be obtained if we consider the CI equations to be a system that couples terms describing increasing degrees of excitation from the reference state Φ. This viewpoint becomes particularly useful when we introduce approximations as to the nature of the terms that describe high degrees of excitation.

We begin the analysis of this subsection by examing the member of the equation set (9.14) corresponding to projection on Φ (i.e., the equation with $\Phi_\Delta^R = \Phi$). In intermediate normalization, that equation is

$$< \Phi \mid \mathsf{H} - E \mid \Phi > + \sum_{\alpha r} c_\alpha^r \; < \Phi \mid \mathsf{H} - E \mid \Phi_\alpha^r > + \sum_{(\alpha\beta)(rs)} c_{\alpha\beta}^{rs} < \Phi \mid \mathsf{H} - E \mid \Phi_{\alpha\beta}^{rs} >= 0 \quad (9.17)$$

Recall that $(\alpha\beta)$ and (rs) denote summations over distinct index sets. Writing $\mathsf{H} = \mathsf{U} + \mathsf{V}$, where U and V are its one- and two-electron parts, respectively, and using Eqs. to evaluate the matrix elements, we obtain

$$\begin{aligned} E - E_\Phi \; = \; & \sum_{\alpha r} c_\alpha^r \left[< \phi_\alpha \mid \mathsf{u} \mid \phi_r > + \sum_\beta < \phi_\alpha \, \phi_\beta \mid \tilde{\mathsf{v}} \mid \phi_r \, \phi_\beta > \right] \\ & + \frac{1}{4} \sum_{\alpha\beta;rs} c_{\alpha\beta}^{rs} \; < \phi_\alpha \, \phi_\beta \mid \tilde{\mathsf{v}} \mid \phi_r \, \phi_s > \end{aligned} \quad (9.18)$$

with E_Φ the energy expectation value associated with Φ:

$$E_\Phi \; = < \Phi \mid \mathsf{H} \mid \Phi > = \sum_\alpha < \phi_\alpha \mid \mathsf{u} \mid \phi_\alpha > + \frac{1}{2} \sum_{\alpha\beta} < \phi_\alpha \, \phi_\beta \mid \tilde{\mathsf{v}} \mid \phi_\alpha \, \phi_\beta > \quad (9.19)$$

In Eqs. (9.18) and (9.19) we have removed the restrictions on the index sums and inserted appropriate factors $\frac{1}{4}$ and $\frac{1}{2}$. This change requires the assumption that the $c_{\alpha\beta}^{rs}$ are

antisymmetric under index permutation. Remember that the $\tilde{\mathsf{v}}$ matrix elements are also antisymmetrized. Because H contains no terms that depend on the coordinates of more than two particles, triply and higher excited terms in Eq. (9.17) make no contribution to E.