PHYSICAL REVIEW VOLUME 127, NUMBER 5 SEPTEMBER 1, 1962

Stability of Hartree-Fock States*

WILLIAM H. ADAMS†‡
Quantum Chemistry Group, University of Uppsala, Uppsala, Sweden

(Received April 5, 1962)

The condition that must be satisfied by a Hartree-Fock wavefunction if it is to give an absolute minimum of the energy, is derived by variation of the one-electron density matrix. If the energy is not an absolute minimum, the state is unstable. Introducing spin explicitly into the equations, we find that there are two classes of variational functions which are particularly suitable in investigations of stability. The one variation is related to the *alternate orbital transformation,* while the other is connected with Hund's rule and the conditions for ferromagnetism. The first of these variations is used in two numerical examples. In the first example we investigate the stability of a restricted Hartree-Fock wavefunction for LiH relative to an unrestricted one. We find that for the chosen basis set, the restricted Hartree-Fock wavefunction is stable at the equilibrium internuclear distance (3.0 a.u.), but that at 4.0 a.u. it becomes unstable. The second example investigates the relative stability of the restricted and unrestricted Hartree-Fock approximations for the electron gas. It is shown that at a sufficiently low density, the unrestricted Hartree-Fock method gives a lower energy. The resulting state has a nonzero spin density. The importance of the stability condition in atomic, molecular, and solid-state problems is emphasized.

1. INTRODUCTION

IN evaluating Hartree-Fock wavefunctions, one usually minimizes the sum of the orbital energies, not the total energy of the system. The Hartree-Fock equation is a *necessary condition* that must be satisfied by the orbitals which minimize the total energy. It is not a sufficient condition. This is a well-known property of the equations derived by the methods of the variational calculus.[1] That the intuitive idea of putting the electrons in the orbitals of lowest energy to minimize the total energy is not sufficient, can be seen from the properties of the iron group elements.

In the case of the iron-group elements, it has been argued that the $4s$ orbitals are filled before the $3d$ orbitals because they lie lower in energy.[2] This explanation ignores the experimental set of one-electron energies given by Slater.[3] Slater's one-electron energies show that in the iron group, the $3d$ orbital always has a lower energy than the $4s$ orbital. [The smallest energy difference is 0.02 a.u.; the average, 0.066 a.u.] In addition to Slater's experimental one-electron energies, one can examine the one-electron energies found by Watson in his restricted Hartree-Fock calculations on the iron group.[4] Watson found that the orbital energy of the $4s$ electron was greater than that of the $3d$ electron and that the $3d^{n-2}4s^2$ configuration was always lower in energy than the $3d^n$ configuration. In this case, it seems safe to conclude that the intuitive idea fails; the total energy is not minimized by filling up the lowest energy orbitals.

The implication of the preceding arguments is that there exists more than one set of orbitals which satisfy the Hartree-Fock equation. In this paper, when we speak of a *Hartree-Fock state,* we shall mean the state defined by a set of orbitals which satisfy the Hartree-Fock equation. The set of orbitals which yield the lowest total energy describes the *Hartree-Fock ground state* of the system. As one hopes that the Hartree-Fock ground-state wavefunction is adequate for describing the ground state of a system, one might hope that the other Hartree-Fock state functions will describe excited states of the system.[5]

There is one problem for which two sets of self-consistent solutions of the Hartree-Fock equation exists. This is the problem of the one-dimensional Fermion gas with δ-function repulsion. In addition to the well-known plane-wave solutions, Overhauser[6] has found a set of orbitals which satisfies the appropriate Hartree-Fock equation and gives a lower total energy than the plane-wave orbitals. But until Overhauser found this Hartree-Fock state, no one knew that there was a Hartree-Fock state with lower total energy than the plane-wave state.

In this paper we shall give a criterion for deciding whether a given Hartree-Fock state corresponds to an absolute minimum of the energy or not. The criterion will be derived by requiring that the second variation of the total energy be positive definite.[1] Although this criterion is adequate for showing that a given state does or does not correspond to an absolute minimum of the total energy, we shall find that the nonlinearity of the

---

* The research reported in this paper was sponsored in part by the King Gustaf VI Adolf's 70-Years Fund for Swedish Culture, Knut and Alice Wallenberg's Foundation, the Swedish Natural Science Research Council, and in part by the Aeronautical Research Laboratory, OAR, through the European Office, Aerospace Research, United States Air Force.

† National Science Foundation postdoctoral fellow, 1960–1962.

‡ Present address: Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania.

[1] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953), Vol. I, p. 186.

[2] Compare with L. Pauling, *The Nature of the Chemical Bond* (Cornell University Press, Ithaca, New York, 1960), 3rd ed., pp. 48–49.

[3] J. C. Slater, Phys. Rev. **98**, 1039 (1955).

[4] R. E. Watson, Massachusetts Institute of Technology Solid State and Molecular Theory Group, Technical Report 12, June 15, 1959 (unpublished); Phys. Rev. **118**, 1036 (1960); **119**, 1934 (1960).

[5] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Springer-Verlag, Berlin, 1957), pp. 140–143, 354; and the references cited there.

[6] A. W. Overhauser, Phys. Rev. Letters **4**, 415, 462 (1960).

Hartree-Fock equation introduces the possibility of having more than one state which gives an absolute minimum of the energy.

The criterion for an absolute energy minimum that we shall give has been given in a more specialized form by Thouless.[7] From his equation he has obtained a connection with the theory of collective motions, showing that if a Hartree-Fock state does not correspond to an absolute minimum of the energy, it must be unstable with respect to a collective mode. Thus, he chose the name *stability condition* for the criterion for an absolute minimum of the energy. We feel that the importance of the stability condition justifies our discussing Thouless' result in a more general form.[8]

The stability condition is important for several reasons. In performing a Hartree-Fock calculation, one must either guess or know what the ground-state configuration is. One assumes that the true ground state and the Hartree-Fock ground state correspond to the same configuration. This assumption can be checked by performing Hartree-Fock calculations for other configurations, but this involves considerable extra labor. The stability condition is a less laborious way of testing the assumption. The assumption needs testing if we are to infer from Hartree-Fock calculations the properties of experimentally uninvestigated systems.

The stability condition can be used as an aid in calculations of Hartree-Fock wavefunctions. It supplements the rule of thumb that the lowest energy orbitals should be occupied. We will see that this rule of thumb is a good rule of thumb according to the stability condition, when we are solving for the Hartree-Fock orbitals which diagonalize the matrix of orbital energies. However, when one solves directly for Wannier functions[9] or some other kind of localized orbitals,[10] the rule of thumb could be less reliable. And if one solves directly for the one-electron density matrix by McWeeny's method,[11] one gives up the rule of thumb entirely. One could accidentally obtain the density matrix for a self-consistent excited state since in McWeeny's method the change in the density matrix in each step of the iteration process is determined by the lack of self-consistency.

In this paper we shall derive the stability condition in a general form and exhibit some of its properties, including its one weakness, which appears not to have been discussed previously. Then we shall consider some simple applications. A numerical application of the stability condition to the LiH molecule will be made, coupled with a brief discussion of the behavior of the Hartree-Fock wavefunction and energy when a molecule is pulled into its component parts. Finally, we shall show that the electron gas plane-wave state is unstable with respect to a state having a nonzero spin density, when the electron density is very small.

2. THE STABILITY CONDITION

We discuss the Hartree-Fock approximation for a system of N electrons moving in an arbitrary external field and interacting with one another only through the Coulomb repulsion. We adopt the conventions of the transformation theory, summing over discrete indices and integrating over continuous ones. The one-electron density matrix $\boldsymbol{\rho}$ of the system in the coordinate representation is defined in terms of the Hartree-Fock spin orbitals by

$$(\boldsymbol{\rho})_{\mathbf{x}_1\mathbf{x}_2}=\rho(\mathbf{x}_1|\mathbf{x}_2)=\sum_{\mu=1}^{N}\psi_\mu(\mathbf{x}_1)\psi_\mu^*(\mathbf{x}_2). \tag{1}$$

It is an idempotent matrix. In addition to the matrix $\boldsymbol{\rho}$, we define three others in the coordinate representation. Writing $\mathcal{H}_j$ for that part of the total Hamiltonian which operates on the coordinate of the jth electron, we define

$$(\mathbf{H})_{\mathbf{x}_i\mathbf{x}_j}=\delta(\mathbf{x}_i-\mathbf{x}_j)\mathcal{H}_j. \tag{2}$$

Similarly, we define the Coulomb matrix

$$(\mathbf{B})_{\mathbf{x}_i\mathbf{x}_j}=\delta(\mathbf{x}_i-\mathbf{x}_j)\int d\mathbf{x}_k(r_{jk})^{-1}\rho(\mathbf{x}_k|\mathbf{x}_k), \tag{3}$$

and the exchange matrix

$$(\mathbf{A})_{\mathbf{x}_i\mathbf{x}_j}=(r_{ij})^{-1}\rho(\mathbf{x}_i|\mathbf{x}_j). \tag{4}$$

The important thing to remember with respect to $\mathbf{B}$ and $\mathbf{A}$ is that they are linear in $\boldsymbol{\rho}$. The total energy of the system is the trace of a sum of matrix products:

$$E=[\mathbf{H}\boldsymbol{\rho}+\tfrac{1}{2}(\mathbf{B}-\mathbf{A})\boldsymbol{\rho}]_{\text{tr}}. \tag{5}$$

In the Hartree-Fock method, one chooses the N orthonormal functions ψ_μ so that E is a minimum. It has been shown that $\boldsymbol{\rho}$ contains all possible information about the system.[12] We shall work directly with variations of the fundamental invariant $\boldsymbol{\rho}$.

We wish to replace $\boldsymbol{\rho}$ in Eq. (5) by a $\boldsymbol{\rho}'$ containing arbitrary first- and second-order variations subject to the condition that $\boldsymbol{\rho}'$ is indempotent to the second order. McWeeny has shown how to do this.[13] Let $\boldsymbol{\Delta}$ be an arbitrary matrix. Representing the unit matrix by $\mathbf{1}$ (in a continuous representation, it is the δ function), we define

$$\mathbf{v}=(\mathbf{1}-\boldsymbol{\rho})\boldsymbol{\Delta}\boldsymbol{\rho}. \tag{6}$$

Letting λ be the parameter that defines the order of the

---

[7] D. J. Thouless, Nuclear Phys. **21**, 225 (1960); *The Quantum Mechanics of Many-Body Systems* (Academic Press Inc., New York, 1961), pp. 24–29.

[8] For other discussions of the stability condition problem see I. Ia. Pomeranchuk, Soviet Phys.—JETP **8**, 361 (1959); and K. Sawada and N. Fukuda, Progr. Theoret. Phys. (Kyoto) **25**, 653 (1961).

[9] G. F. Koster, Phys. Rev. **89**, 67 (1953); G. Parzen, *ibid.* **89**, 237 (1953).

[10] W. H. Adams, J. Chem. Phys. **34**, 89 (1961).

[11] R. McWeeny, Proc. Roy. Soc. (London) **A235**, 496 (1956).

[12] P. O. Löwdin, Phys. Rev. **97**, 1490 (1955).

[13] R. McWeeny, Revs. Modern Phys. **32**, 335 (1960), Eqs. (74) and (75).

variation, and writing

$$\boldsymbol{\rho}'=\boldsymbol{\rho}+\lambda\boldsymbol{\rho}^{(1)}+\lambda^2\boldsymbol{\rho}^{(2)}+\cdots, \tag{7}$$

then an arbitrary first-order variation can be written as

$$\boldsymbol{\rho}^{(1)}=\mathbf{v}+\mathbf{v}^{\dagger}. \tag{8}$$

To the second-order variation, the first-order variation contributes $\mathbf{v}\mathbf{v}^{\dagger}-\mathbf{v}^{\dagger}\mathbf{v}$. This insures that the indempotency condition on $\boldsymbol{\rho}$ will be satisfied. Writing $\boldsymbol{\Delta}'$ for a second arbitrary matrix, and defining $\mathbf{v}'$ in terms of $\boldsymbol{\Delta}'$ as $\mathbf{v}$ was defined in terms of $\boldsymbol{\Delta}$, we can write the arbitrary second-order variation as

$$\boldsymbol{\rho}^{(2)}=\mathbf{v}\mathbf{v}^{\dagger}-\mathbf{v}^{\dagger}\mathbf{v}+\mathbf{v}'+\mathbf{v}'^{\dagger}. \tag{9}$$

The connection of these formulas with those for the variation of the orbitals has been given by McWeeny.[11]

We substitute $\boldsymbol{\rho}'$ into Eq. (5), obtaining E as a power series in λ. The requirement that E be stationary with respect to an arbitrary first-order variation of the electron density is

$$E^{(1)}=(\partial E/\partial\lambda)_{\lambda=0}=0. \tag{10}$$

The requirement that E be an absolute minimum is equivalent to

$$E^{(2)}=\tfrac{1}{2}(\partial^2E/\partial\lambda^2)_{\lambda=0}>0. \tag{11}$$

As we shall discuss later, this requirement is not as strong as one would like it to be. Defining $\mathbf{B}^{(1)}$ and $\mathbf{A}^{(1)}$, respectively, by Eqs. (3) and (4), with $\boldsymbol{\rho}$ replaced by $\boldsymbol{\rho}^{(1)}$ in each case, we obtain from (10) and (11)

$$E^{(1)}=[\mathbf{H}\boldsymbol{\rho}^{(1)}+(\mathbf{B}-\mathbf{A})\boldsymbol{\rho}^{(1)}]_{\mathrm{tr}} \tag{12}$$

$$E^{(2)}=[\mathbf{H}\boldsymbol{\rho}^{(2)}+(\mathbf{B}-\mathbf{A})\boldsymbol{\rho}^{(2)}+\tfrac{1}{2}(\mathbf{B}^{(1)}-\mathbf{A}^{(1)})\boldsymbol{\rho}^{(1)}]_{\mathrm{tr}}. \tag{13}$$

We discuss $E^{(1)}$ first.

We can write $E^{(1)}$ in a more compact form by defining

$$\mathbf{F}=\mathbf{H}+\mathbf{B}-\mathbf{A}. \tag{14}$$

This is the familiar Hartree-Fock effective Hamiltonian. We write out $\boldsymbol{\rho}^{(1)}$ in Eq. (12) explicitly in terms of the arbitrary matrix $\boldsymbol{\Delta}$.

$$\begin{aligned}E^{(1)}&=[\mathbf{F}(\mathbf{1}-\boldsymbol{\rho})\boldsymbol{\Delta}\boldsymbol{\rho}+\mathbf{F}\boldsymbol{\rho}\boldsymbol{\Delta}^{\dagger}(\mathbf{1}-\boldsymbol{\rho})]_{\mathrm{tr}}\\&=[\boldsymbol{\rho}\mathbf{F}(\mathbf{1}-\boldsymbol{\rho})\boldsymbol{\Delta}+(\mathbf{1}-\boldsymbol{\rho})\mathbf{F}\boldsymbol{\rho}\boldsymbol{\Delta}^{\dagger}]_{\mathrm{tr}}=0.\end{aligned} \tag{15}$$

For an arbitrary $\boldsymbol{\Delta}$, this implies that

$$\boldsymbol{\rho}\mathbf{F}(\mathbf{1}-\boldsymbol{\rho})=0. \tag{16}$$

This is equivalent to

$$\mathbf{F}\boldsymbol{\rho}-\boldsymbol{\rho}\mathbf{F}=0. \tag{17}$$

This equation and the requirements that $\boldsymbol{\rho}$ be idempotent and Hermitian, and that $[\boldsymbol{\rho}]_{\mathrm{tr}}=N$, the number of occupied orbitals, are equivalent to the Hartree-Fock equation and the requirement that the Hartree-Fock orbitals be orthonormal.[13,14]

Turning to the second variation $E^{(2)}$, we can introduce one simplification due to Eq. (17). From the arbitrary part of the second-order variation the contribution to $E^{(2)}$ is

$$[\mathbf{F}(\mathbf{v}'+\mathbf{v}'^{\dagger})]_{\mathrm{tr}}=[\mathbf{F}(\mathbf{1}-\boldsymbol{\rho})\boldsymbol{\Delta}'\boldsymbol{\rho}+\mathbf{F}\boldsymbol{\rho}\boldsymbol{\Delta}'^{\dagger}(\mathbf{1}-\boldsymbol{\rho})]_{\mathrm{tr}}.$$

But this is equivalent to Eq. (15) and, hence, must vanish if Eq. (17) is satisfied. Thus, in general, the second-order variation may be written entirely in terms of $\mathbf{v}$.

$$E^{(2)}=[\mathbf{F}(\mathbf{v}\mathbf{v}^{\dagger}-\mathbf{v}^{\dagger}\mathbf{v})+\tfrac{1}{2}(\mathbf{B}^{(1)}-\mathbf{A}^{(1)})(\mathbf{v}+\mathbf{v}^{\dagger})]_{\mathrm{tr}}. \tag{18}$$

If E is an absolute minimum of the energy, then $E^{(2)}>0$ for any $\mathbf{v}$. Conversely, if we choose $\mathbf{v}$ so as to make $E^{(2)}$ as small as possible, and if this minimum value of $E^{(2)}$ is positive, then E is an absolute minimum. In other words, we can derive from Eq. (18) an equation that $\mathbf{v}$ must satisfy if $E^{(2)}$ is to be an extremal. The variation of $\mathbf{v}$ in $E^{(2)}$ is subject to the constraint that $[\mathbf{v}\mathbf{v}^{\dagger}]_{\mathrm{tr}}$ = finite number. This constraint must be applied if $E^{(2)}$ is to be finite. (This point is discussed in Appendix A.) If $E^{(2)}$ is not finite then the expansion of E in powers of λ is not valid. Applying the variational calculus to Eq. (18) with the constraint $[\mathbf{v}\mathbf{v}^{\dagger}]_{\mathrm{tr}}=1$, we obtain

$$\mathbf{F}\mathbf{v}-\mathbf{v}\mathbf{F}+(\mathbf{1}-\boldsymbol{\rho})(\mathbf{B}^{(1)}-\mathbf{A}^{(1)})\boldsymbol{\rho}=\omega\mathbf{v}. \tag{19}$$

The Hermitian conjugate of this equation is satisfied by $\mathbf{v}^{\dagger}$. Equation (19) is equivalent to Thouless' stability condition, but it is not restricted to a particular representation. (Thouless expands $\mathbf{v}$ in terms of the complete set of Hartree-Fock orbitals.) If the smallest eigenvalue of (19) is ω_0, then

$$E^{(2)}\geqslant\omega_0 \tag{20}$$

for all choices of $\mathbf{v}$. If $\omega_0>0$, then E is an absolute minimum.

A clearer understanding of the meaning of the condition $E^{(2)}>0$ can be obtained by going over to a particular representation. Letting the orbitals $\psi_\mu(\mathbf{x})$ be those that diagonalize the Hartree-Fock Hamiltonian, we write

$$(\mathbf{v})_{\mathbf{x}_1\mathbf{x}_2}=\sum_{\mu\geq N+1}\sum_{\nu=1}^{N}\psi_\mu(\mathbf{x}_1)C_{\mu\nu}\psi_\nu^*(\mathbf{x}_2). \tag{21}$$

The orbital energy eigenvalue corresponding to $\psi_\mu(\mathbf{x})$ is ϵ_μ. Using the definition of Eq. (21), we arrive at

$$[\mathbf{F}\boldsymbol{\rho}^{(2)}]_{\mathrm{tr}}=\sum_{\mu\geq N+1}\sum_{\nu=1}^{N}(\epsilon_\mu-\epsilon_\nu)|C_{\mu\nu}|^2. \tag{22}$$

If we assume that the orbital energies of the N occupied orbitals are smaller than the smallest virtual orbital energy, then the trace of Eq. (22) is positive definite. Thus, the intuitive idea of filling up the N orbitals with the lowest orbital energy first, corresponds to making one contribution to $E^{(2)}$ positive for any variation. If

---

[14] P. A. M. Dirac, Proc. Cambridge Phil. Soc. **26**, 376 (1930).

such a state is unstable, we will have

$$-\tfrac{1}{2}[(\mathbf{B}^{(1)}-\mathbf{A}^{(1)})\boldsymbol{\varrho}^{(1)}]_{\mathrm{tr}}\geqslant[\mathbf{F}\boldsymbol{\varrho}^{(2)}]_{\mathrm{tr}}>0. \tag{23}$$

In general, one finds that Coulomb and exchange integrals are relatively small, so that unless the differences between the energies of some occupied and some unoccupied orbitals are small, one should not expect to find that Eq. (23) could be satisfied. Substitution of $\mathbf{v}$, as defined in Eq. (21) into Eq. (18), and variation of the $C_{\mu\nu}$ leads to the form of the stability condition given by Thouless.[7]

At this point we can profitably examine what we mean when we say that $E^{(2)}>0$ indicates that the state under consideration corresponds to an absolute minimum of the energy. We do not mean that there exists only one absolute minimum of the energy. If the electrons did not interact we would have $E^{(2)}=[\mathbf{F}\boldsymbol{\varrho}^{(2)}]_{\mathrm{tr}}$ and there would be only one absolute minimum of the energy, namely, the state in which the orbitals of lowest energy are occupied. Because the electrons interact, a nonlinear term is included in $E^{(2)}$. This nonlinear term can be positive or negative, and can lead to the instability of the state in which the electrons fill the lowest energy orbitals. It also makes it possible for two or more absolute minima to exist. At least we cannot show, in general, that only one absolute minimum exists. If several absolute minima of the energy exist for a given system, however, these states should not differ greatly in energy from the state of lowest energy. This is a rather qualitative statement, but that is because it is based on rather qualitative reasoning. The reasoning is presented in Appendix B.

At present, most Hartree-Fock calculations are done by expanding the orbitals in terms of a basis set of M orthonormal functions ($M>N$). Going over to a basis set representation in Eq. (19), we see that we have M^2 simultaneous and homogeneous linear equations to solve. This gives a secular equation of degree M^2. Thus, if the basis set is very large, the size of the secular equation arising from Eq. (19) may make the exact application of the stability condition numerically impractical. Using Thouless' representation, the secular equation is reduced to one of deg $N(M-N)$. This is a considerable reduction, but unless $M-N$ is relatively small, it is not a sufficient reduction. The alternative is to work directly in terms of the condition $E^{(2)}>0$, choosing the arbitrary matrix so that it will mix states and combinations of states with respect to which one might expect the system to be unstable. We shall come back to this point in the examples.

We have assumed so far that there were no conditions on the variation other than the preservation of idempotency. In most atomic, molecular, and solid-state problems supplementary conditions are imposed on the form of $\boldsymbol{\varrho}$. Many, but not all, of these supplementary conditions can be incorporated into our equations by restricting the form of $\boldsymbol{\Delta}$. This is the way one usually introduces into the Hartree-Fock approximation the restriction of doubly occupied orbitals, and various symmetry and equivalence restrictions.[15]

Except for the remarks of the preceding paragraph, our discussion would seem to be applicable only to the unrestricted Hartree-Fock approximation. The general approach is, however, applicable to the extended Hartree-Fock schemes.[16] In an extended Hartree-Fock scheme, the energy expression is more complicated, but can be written in terms of two idempotent matrices $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$.[17] One may vary $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$ independently, again looking at the second variation.

3. INCLUSION OF THE SPIN

To introduce spin into our equations is no problem, but it does have some interesting consequences. The important consequence is that there exist two ways of choosing $\boldsymbol{\Delta}$ so that $[\mathbf{B}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\mathrm{tr}}=0$. This means that $[(\mathbf{B}^{(1)}-\mathbf{A}^{(1)})\boldsymbol{\varrho}^{(1)}]_{\mathrm{tr}}$, which can be either positive or negative, is reduced to $-[\mathbf{A}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\mathrm{tr}}$, which is always negative. It is always negative, because

$$[\mathbf{A}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\mathrm{tr}}=\int d\mathbf{x}_1d\mathbf{x}_2(r_{12})^{-1}\rho^{(1)}(\mathbf{x}_1|\mathbf{x}_2)\rho^{(1)}(\mathbf{x}_2|\mathbf{x}_1)$$

$$=\int d\mathbf{x}_1d\mathbf{x}_2(r_{12})^{-1}|\rho^{(1)}(\mathbf{x}_1|\mathbf{x}_2)|^2>0.$$

This is a well-known property of the exchange integral.[18]

With the explicit introduction of spin into the Hartree-Fock approximation, one is led to consider several variants on that approximation. The most common variants are the restricted and unrestricted Hartree-Fock approximations. Both are characterized by two fundamental invariants, $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$. Writing α and β for the spin-up and spin-down eigenfunctions, respectively, the one-electron density matrix is

$$\begin{gathered}\boldsymbol{\varrho}=\alpha\alpha^{\dagger}\boldsymbol{\varrho}_++\beta\beta^{\dagger}\boldsymbol{\varrho}_-,\\ \alpha^{\dagger}\alpha=\beta^{\dagger}\beta=1,\quad \alpha^{\dagger}\beta=\beta^{\dagger}\alpha=0.\end{gathered} \tag{24}$$

Thus, $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$ give the density of spin-up and spin-down electrons, respectively. One requires that $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$ be idempotent, thus, guaranteeing that $\boldsymbol{\varrho}$ is idempotent. With no restrictions placed on $\boldsymbol{\varrho}_+$ and $\boldsymbol{\varrho}_-$, we have the unrestricted Hartree-Fock approximation.[19]

---

[15] A thorough discussion of the various common restrictions has been given by R. K. Nesbet, Proc. Roy. Soc. (London) **A230**, 312 (1955); Revs. Modern Phys. **33**, 28 (1961).

[16] P. O. Löwdin, Phys. Rev. **97**, 1509 (1955); Revs. Modern Phys. **32**, 328 (1960). The latter article contains a large number of references.

[17] P. O. Löwdin, Colloq. intern. centre natl. recherche sci. (Paris) **82**, 23 (1958).

[18] See J. C. Slater, *Quantum Theory of Atomic Structure* (McGraw-Hill Book Company, Inc., New York, 1960), Vol. I, pp. 486–487; C. C. J. Roothaan, Revs. Modern Phys. **23**, 69 (1951); Appendix I.

[19] Professor P. O. Löwdin has emphasized to the author that it would be better to call it the Hartree-Fock approximation, since Fock's original proposal assumed there were no restrictions on the orbitals. See V. Fock, Z. Physik **61**, 126 (1930).

That the one-electron density matrix has the form given by Eq. (24), implies that the single determinant wavefunction is an eigenfunction of the z component of the total electron spin, S_z.[20] The restricted Hartree-Fock approximation is obtained when one requires that the Slater determinant be an eigenfunction of S^2, the square of the total electron spin. Let N_+ and N_- be the number of electrons of up and down spin, respectively. Then it can be shown[21] that when

$$[\boldsymbol{\varrho}_+\boldsymbol{\varrho}_-]_{\text{tr}}=N_-,\quad N_-<N_+, \tag{25}$$

the Slater determinant has the spin eigenvalues

$$S=M_S=\tfrac{1}{2}(N_+-N_-). \tag{26}$$

Finally, one might consider the case in which the orbitals are *general spin orbitals*[22] i.e.,

$$\psi_\mu(\mathbf{x})=\alpha u_\mu(\mathbf{r})+\beta v_\mu(\mathbf{r}). \tag{27}$$

In this case the Slater determinant is not an eigenfunction of S_z and $\boldsymbol{\varrho}$ will have $\alpha\beta^\dagger$ and $\beta\alpha^\dagger$ components.[20] If we require the ψ_μ's to be orthonormal, $\boldsymbol{\varrho}$ will be idempotent.

Having simultaneously generalized our notation and reviewed the usual variants of the Hartree-Fock approximation, we can now look at those variations which satisfy $[\mathbf{B}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\text{tr}}=0$. There are two such variations. We consider first a variation for which

$$\boldsymbol{\varrho}^{(1)}=\alpha\alpha^\dagger\boldsymbol{\varrho}_+^{(1)}+\beta\beta^\dagger\boldsymbol{\varrho}_-^{(1)}. \tag{28}$$

We require that

$$\boldsymbol{\varrho}_+^{(1)}=-\boldsymbol{\varrho}_-^{(1)}, \tag{29}$$

which guarantees that $[\mathbf{B}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\text{tr}}$ vanishes. If we are initially in the restricted Hartree-Fock approximation, such a variation takes us into the unrestricted approximation. The *alternate orbital transformation*,[23] which has been investigated extensively by Löwdin and coworkers in connection with the method of alternate orbitals,[24] belongs to the class of variations defined by Eq. (29).

The other variation which we wish to point out is characterized by

$$\boldsymbol{\varrho}^{(1)}=\alpha\beta^\dagger\mathbf{v}+\beta\alpha^\dagger\mathbf{v}^\dagger. \tag{30}$$

This choice guarantees the vanishing of $[\mathbf{B}^{(1)}\boldsymbol{\varrho}^{(1)}]_{\text{tr}}$. Furthermore, if we are considering the restricted or unrestricted Hartree-Fock approximation, the form of $\boldsymbol{\varrho}^{(1)}$ indicates that states of different M_S are being mixed. Equation (30) is written explicitly for the case in which one mixes a state of given M_S with those of higher M_S. The variation given in (30) can be restricted so that one stays within the restricted Hartree-Fock approximation. If the energies of some occupied and unoccupied orbitals do not differ significantly, one would expect that $E^{(2)}$ could be made negative by considering variations of the type defined by Eq. (30). This is, of course, related to Hund's rule that in a given term, the state of highest multiplicity will lie lowest. It is also related to the well-known fact that at low densities [$r_s>6$ a.u. $(4\pi/3)r_s^3$ is the volume per electron] the electron gas is ferromagnetic.[25] This property of the electron gas implies that a solid which has a partially filled energy band with a sufficiently high density of states, will be ferromagnetic. This is, of course, just the assumption made in the band theory of ferromagnetism.[26]

In addition to a variation of the kind exhibited in Eq. (30), one can adopt a more general spin mixing variation. For example one might choose

$$\boldsymbol{\varrho}^{(1)}=\alpha\beta^\dagger(\mathbf{v}_-+\mathbf{v}_+^\dagger)+\beta\alpha^\dagger(\mathbf{v}_-^\dagger+\mathbf{v}_+).$$

This form gives a test of stability relative to a state described by the general spin orbitals [see Eq. (27)] even if the state we are investigating is described by a restricted or unrestricted Hartree-Fock wavefunction. Thus, if we wish to test the stability of a restricted or unrestricted Hartree-Fock wavefunction relative to other restricted or unrestricted wavefunctions, the above variation is the wrong one to choose. At present it is not clear that the use of general spin orbitals will lead to lower energies than those found in the unrestricted Hartree-Fock approximation.

We see that the inclusion of spin leads in a natural way to variations which take maximum advantage in $E^{(2)}$ of the exchange term involving $\boldsymbol{\varrho}^{(1)}$. In view of this one may well ask if the orbital angular momentum can also be incorporated easily. Unfortunately, it cannot. If one allows only those variations which have the same symmetry as the wavefunction, one can build up an approach similar to the one treated here. Variations which alter the symmetry of the wavefunction are more easily treated by working with the total wavefunction rather than with the orbitals. We have not required the Hartree-Fock states we have considered to be eigenfunctions of the square of the orbital angular momentum. For this reason, we have not investigated the interesting problem of the stable states of the iron group transition elements discussed in the introduction. (Two members of this group are known to have only one electron in a 4s level but it has not been shown that this is the case in the Hartree-Fock approximation.)

---

[20] R. McWeeny, Proc. Roy. Soc. (London) **A253**, 242 (1959).

[21] P. O. Löwdin, Advances in Chem. Phys. **2**, 207 (1959), Sec. II D. (2).

[22] R. McWeeny, Revs. Modern Phys. **32**, 335 (1960). See Sec. 4.2. This case has been investigated numerically in Uppsala by R. McWeeny and K. F. Berggren.

[23] The alternate orbital transformation is used to obtain different orbitals for electrons of different spin. In the method of alternate orbitals one defines the wavefunction for a system to be the spin projection of a Slater determinant of alternate orbitals.

[24] P. O. Löwdin, Proceedings of the Japanese Conference on Theoretical Physics, Nikko Symposium, 1953 (unpublished); Phys. Rev. **97**, 1509 (1955); Revs. Modern Phys. **32**, 328 (1960); R. Pauncz, J. de Heer and P. O. Löwdin, J. Chem. Phys. **36**, 2247, 2257 (1962).

[25] F. Bloch, Z. Physik **57**, 545 (1929).

[26] N. F. Mott, Proc. Phys. Soc. (London) **47**, 571 (1935). J. C. Slater, Phys. Rev. **49**, 537, 931 (1936).

4. MOLECULAR ORBITALS

The restricted Hartree-Fock approximation works well for a molecule when the atoms are at their equilibrium distances. When one separates the atoms, however, one finds that the binding energy does not approach zero, but some relatively large positive number. (The only exception is the case of a system of closed shell atoms.) This has been explained by the observation that the doubly occupied molecular orbitals describe at infinite internuclear distances a state which is a mixture of ionic and nonionic states, or a pure ionic state.[27] It was shown by Coulson and Fischer[28] that when the requirement that the orbitals be doubly occupied is dropped, the Hartree-Fock approximation leads to the correct state at infinite nuclear separation for the hydrogen molecule, i.e., the neutral atoms. Furthermore, they found that at the equilibrium distance, the doubly occupied molecular orbitals gave the lowest energy.

The calculation of Coulson and Fischer for H_2 implies that the Hartree-Fock approximation will correctly describe the behavior of a molecule as it is pulled apart, if we let it. That this should be true for other systems may be inferred from a comparison of the total energies of systems of atoms and ions at infinite internuclear distance. For example, for the alkali halides one finds that the stable state at infinite separation is always the neutral alkali and halogen atoms. (Data are available for Li, Na, F, Cl, Br and their singly charged ions.)[29] This is in agreement with experience. Application of the restricted Hartree-Fock approximation to a molecule results in the ionic state being the stable species. But this is only because we have required the orbitals to be doubly occupied. This requirement is introduced to force the single determinant wave function to be in a definite spin multiplet state. At infinite separation it may be important to have each atom in a particular spin state, but does not matter what spin state the system as a whole is in.

The question that we wish to investigate is the one raised by the observation that at the equilibrium internuclear distance, the doubly occupied molecular orbitals minimized the energy of H_2. We shall investigate the stability of the restricted Hartree-Fock approximation for LiH with respect to an unrestricted Hartree-Fock approximation. We have chosen to investigate LiH because it is a conveniently small molecule, and because Karo and Olson have published values for all of their integrals.[30,31] We have found that at the equilibrium distance the doubly occupied molecular orbitals are stable.

The calculations have been performed in two ways. In neither case have we used the full power of the stability condition, but we have instead used our physical intuition to decide how $\mathbf{\Delta}$ should be chosen. Our first observation was an obvious one, namely, that the 1σ orbital would remain doubly occupied. (This is in part a consequence of the limited basis set.) Then we asked how the orbitals of the unrestricted Hartree-Fock approximation should differ from those of the restricted. Obviously in the former, one 2σ orbital will be increased about the Li nucleus, the other about the H nucleus. Following Karo and Olson, we use the symbols h for the $1s$ orbital of H; s, S, and p for the $1s$, $2s$, and $2p$ orbitals, respectively, of Li. The $n\sigma$ orbital is denoted by φ_n. Using this notation the variation we choose is

$$\Delta_{\mathrm{I}}(1|2)=\Delta_+(1|2)=-\Delta_-(1|2)=S(1)\varphi_2(2). \qquad \text{(I)}$$

This variation is of the class defined by Eq. (29). From the definition of $\mathbf{v}$ and the restriction that $[\mathbf{v}\mathbf{v}^\dagger]_{\mathrm{tr}}=1$, we obtain from (I)

$$\begin{aligned} v_{\mathrm{I}}=v_{\mathrm{I}^+}=-v_{\mathrm{I}^-}=&[(\varphi_3|S)^2+(\varphi_4|S)^2]^{-1/2}\\ &\times[\varphi_3(\varphi_3|S)+\varphi_4(\varphi_4|S)]\varphi_2. \qquad (31)\end{aligned}$$

The second variation we choose is

$$v_{\mathrm{II}}=v_{\mathrm{II}^+}=-v_{\mathrm{II}^-}=(1+x^2)^{-1/2}(\varphi_3+x\varphi_4)\varphi_2. \qquad \text{(II)}$$

This variation also belongs to the class defined by Eq. (29) and uses the Hartree-Fock orbitals as Thouless does.[7] We minimize $E^{(2)}$ with respect to x.

Since the question of stability is rather a qualitative one, we present our results for the variations (I) and (II) as a graph of $E^{(2)}$ vs R. Our calculations cover the range $3.0<R<6.0$ a.u. which is only part of the range covered by the calculations of Karo. We see that $E^{(2)}$ is positive at the equilibrium internuclear distance ($R=3.0$ a.u.), but becomes negative for R a little larger than 4.0 a.u. It is for $R>4.0$ a.u. that the restricted Hartree-Fock approximation fails. That the restricted Hartree-Fock state should be unstable before $R=5.0$

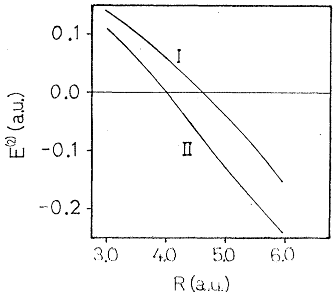

FIG. 1. The second variation of the electronic energy as a function of the internuclear distance in LiH for two different choices of $\mathbf{\Delta}$.

---

[27] J .C. Slater, Phys. Rev. **35**, 509 (1930).

[28] C. A. Coulson and I. Fischer, Phil. Mag. **40**, 386 (1949). See also J. C. Slater, Phys. Rev. **82**, 538 (1951).

[29] Li, Na, and F: L. M. Sachs, Technical Report, Argonne National Laboratory, May, 1961 (unpublished). Li: Roothaan, Sachs and Weiss, Revs. Modern Phys. **32**, 186 (1960). F: L. C. Allen, J. Chem. Phys. **34**, 1156 (1961). Cl: R. E. Watson and A. J. Freeman, Phys. Rev. **123**, 521 (1961). Br: R. E. Watson and A. J. Freeman, *ibid.* **124**, 1117 (1961). There is a misprint in their energy for Br. It should be −2582.443 a.u.

[30] A. M. Karo and A. R. Olson, J. Chem. Phys. **30**, 1232 (1959).

[31] A. M. Karo, J. Chem. Phys. **30**, 1241 (1959).

a.u. can be inferred from the potential energy curve given by Karo (his Fig. 1).[31] At $R=5.0$ a.u., the total energy found by Karo is larger than the energy of an isolated Li atom and an isolated H atom. (This energy is, in turn, larger than the energy obtained from a Slater determinant wave function with a doubly occupied s, and an up-spin in S and a down-spin in h.)[32] One cannot infer in this way, however, that the restricted Hartree-Fock state for LiH is stable out to $R=4.0$ a.u. One can only infer that with the small basis set used in this calculation, the restricted Hartree-Fock approximation cannot be improved upon by going over to the unrestricted approximation for $R<4.0$ a.u.

Finally, we remark on the helpfulness of our physical intuition. Quite plainly the variation (II) yielded a significantly smaller $E^{(2)}$ than variation (I). However, using variation (I) is simpler and gives an indication of where the stability breaks down. [In the example we have considered (I) is only a little simpler than (II), a situation which should not be true in general.] Thus, we really needed to use (II) only at one point, $R=4.0$ a.u. The other calculations with (II) were done merely to exhibit the behavior of $E^{(2)}$ as a function of R.

5. THE ELECTRON GAS

It is well known that plane waves provide a self-consistent solution to the Hartree-Fock equation for an electron gas. Recently Overhauser[6,33] has raised a question as to the existence of other self-consistent solutions which have a lower total energy than the plane-wave solutions. For a one-dimensional Fermion gas with δ-function repulsions, Overhauser found a self-consistent solution of lower energy than the plane-wave state. (Other examples have been discussed since.)[34] This state exhibits a nonzero spin density, or, as Overhauser says, "giant spin-density waves". No one, however, has published an Overhauser-type solution to the electron gas problem.[35] Indeed arguments have been presented to demonstrate that in an electron gas an Overhauser state can exist only at very low densities.[36] In this section we shall use the stability condition to investigate this problem.

In our investigation of the electron gas we have not looked for a self-consistent solution of the Hartree-Fock equation, but have merely tried to find a simple variation function $\mathbf{v}$ which gives rise to a spin density. (We cannot imagine how E can be improved without introducing a spin density, if we stay within the Hartree-Fock approximation.) By our definition $\mathbf{v}$ was simple if it led to integrals that could be easily evaluated. Here we present only one of these $\mathbf{v}$'s, and reserve a discussion of the physics of such states for a later publication.

The variation we consider is of the type defined by Eq. (29). Writing k_F for the Fermi momentum, and considering the case in which there are the same number of up- and down-spin electrons, the spinless Fock-Dirac density matrix is

$$s=|\mathbf{r}_1-\mathbf{r}_2|,$$

$$\rho_0(\mathbf{r}_1|\mathbf{r}_2)=(2\pi^2s^3)^{-1}[\sin(2\pi k_F s)-2\pi k_F s\cos(2\pi k_F s)].$$

Defining $K=(2/\sqrt{3})k_F$, the variation we have chosen is

$$\boldsymbol{\rho}_+{}^{(1)}=-\boldsymbol{\rho}_-{}^{(1)}=\mathbf{v}+\mathbf{v}^\dagger,$$

$$v(\mathbf{r}_1|\mathbf{r}_2)=2\sqrt{2}\rho_0(\mathbf{r}_1|\mathbf{r}_2)\cos(2\pi Kx_1)\times\cos(2\pi Ky_1)\cos(2\pi Kz_1). \tag{32}$$

This variation gives rise to a nonzero spin density in the first order in λ, and a nonzero density periodicity in the second order. The integrals arising in $E^{(2)}$ from the $\mathbf{v}$ defined in (32) may be evaluated rather easily. (Only one integral had to be evaluated by numerical integration, and then only over one variable.) We find that the plane-wave state is unstable with respect to the above variation, for $r_s>32$ a.u., a density well below those found in metals. [One defines r_s so that $(4\pi/3)r_s^3$ is the volume per electron.]

Although the plane-wave state is stable with respect to the variation defined by Eq. (32) for $r_s<32$ a.u., this does not mean that there does not exist another choice of $\mathbf{v}$ with respect to which the plane-wave state is unstable at higher densities. Indeed $\mathbf{v}$ can be chosen so that $[\mathbf{F}\boldsymbol{\rho}^{(2)}]_{\text{tr}}$ is smaller than for the above variation. Unfortunately these other choices lead to integrals which are much less easy to evaluate. Furthermore, even though the $\mathbf{v}$ we have chosen gives rise to a periodicity in the electron density, this does not imply that the Hartree-Fock ground state of the electron gas has this property. That the variation leads to a nonzero spin density does imply that the Hartree-Fock ground state in the unrestricted approximation has a spin density.

6. CONCLUSION

The stability condition when used in its complete form [Eq. (19) and the requirement $\omega_0>0$], tells whether a given Hartree-Fock state is stable or not. If the state is stable then it will correspond to an absolute minimum of the energy. Unfortunately it has not been possible to show that there is only one absolute minimum of the energy in the Hartree-Fock approximation. It may be that the number of absolute minima will depend upon the nature of the system considered. Whatever the case may be, one may well ask if the

---

[32] William H. Adams (unpublished calculation).

[33] A. W. Overhauser, Bull. Am. Phys. Soc. **5**, 433 (1960).

[34] E. M. Henley and Th. W. Ruijgrok, Ann. Phys. (New York) **12**, 409 (1961); E. M. Henley and L. Wilets, *ibid.* **14**, 120 (1961).

[35] Overhauser has described one. See reference 32. After this manuscript had been completed an abstract of a report on a stability calculation on the electron gas was found. Iwamoto and Sawada report that the plane wave state is unstable for $r_s<4.5$. See F. Iwamoto and K. Sawada, Bull. Am. Phys. Soc. **7**, 66 (1962). *Note added in proof.* Their paper has now been published: F. Iwamoto and K. Sawada, Phys. Rev. **126**, 887 (1962).

[36] W. Kohn and S. J. Nettel, Phys. Rev. Letters **5**, 8 (1960).

stability of the wavefunction matters. If we are interested only in having approximate charge densities magnetic form factors, etc., for experimentally known states of a system, it should not matter. If, however, we wish to use the Hartree-Fock approximation as a model of the physical world, then stability does matter. In chemistry in particular, one is interested in the stable state. There is no guarantee that the Hartree-Fock approximation will predict the stable states found in nature, but we may hope that it will. The stability condition gives us a tool for investigating this equation.

As a tool the stability condition has its weaknesses. One of these weaknesses is the question that exists regarding the number of absolute minima. Another weakness is that in its complete form it is numerically as complicated as solving the Hartree-Fock equations. We have given two examples in which we have used intuition to pick out variations for which one might expect to find instability. But in doing so we did not answer the question, is this state stable? Instead we answered the question, is this state unstable with respect to this variation? Whenever we get a "no" answer to this question, we still cannot be certain that the state is stable. Whenever we get a "yes" answer, then we know the state in question is unstable. By investigating the stability of a Hartree-Fock state with respect to particular choices of $\boldsymbol{\Delta}$, we can make it seem improbable that the state is unstable, even though we cannot prove that it is stable.

ACKNOWLEDGMENTS

The author thanks Professor J. E. Mayer for asking him how one knew that the energy was minimized by a self-consistent solution of the Hartree-Fock equation, and Professor P. O. Löwdin and Dr. R. McWeeny for a number of helpful discussions, suggestions, and criticisms. To Professor Löwdin the author expresses his gratitude for the hospitality shown him in Uppsala and for the opportunity of being in the stimulating atmosphere of the Quantum Chemistry Group.

To F. K. Klaus Appel the author is indebted for programming for the ALWAC III-E computer and carrying through the calculation of the two electron integrals between the molecular orbitals of LiH.

APPENDIX A

We have stated that one should have

$$[\mathbf{v}\mathbf{v}^\dagger]_{\mathrm{tr}}=\text{finite number,} \tag{A1}$$

if $E^{(2)}$ is to be finite. This can be rather easily seen from Eq. (22). Let us assume that the smallest nonzero difference between the orbital energy of an occupied orbital and an unoccupied orbital is $\Delta\epsilon$, and that $\Delta\epsilon>0$. Then we have the inequality

$$[\mathbf{F}\boldsymbol{\rho}^{(2)}]_{\mathrm{tr}}>\Delta\epsilon[\mathbf{v}\mathbf{v}^\dagger]_{\mathrm{tr}}. \tag{A2}$$

Thus, if $[\mathbf{v}\mathbf{v}^\dagger]_{\mathrm{tr}}$ is not finite, $[\mathbf{F}\boldsymbol{\rho}^{(2)}]_{\mathrm{tr}}$ is not finite, and only for very special variations will $E^{(2)}$ be finite. This argument can be generalized by observing that only a finite number of the differences $\epsilon_\mu-\epsilon_\nu$ in Eq. (22) can be negative.

APPENDIX B

The purpose of this Appendix is to present the quantitative and qualitative reasoning behind the statements made in Sec. 2 regarding the number of absolute minima of the energy. We asserted there that if several absolute minima of the energy exist, these minima do not differ greatly from the lowest. To see this we shall test the stability of a state of energy E_2 (we shall call it *state 2*) relative to a state of energy E_1. We assume that $E_2>E_1$. We shall see that we can suppose that states 1 and 2 are both stable without meeting a contradiction. We define the orbital matrix, which is composed of the N occupied orbitals, by writing down its elements.

$$(\boldsymbol{\Psi})_{\mathrm{x},\mu}=\psi_\mu(\mathbf{x}). \tag{B1}$$

Let $\boldsymbol{\Psi}_1$ and $\boldsymbol{\Psi}_2$ be the orbital matrices of states 1 and 2, respectively. The Fock-Dirac density matrices of these states can be written in terms of their orbital matrices

$$\boldsymbol{\rho}_n=\boldsymbol{\Psi}_n\boldsymbol{\Psi}_n^\dagger. \tag{B2}$$

Defining $\mathbf{1}_N$ to be the $N\times N$ unit matrix, then

$$\boldsymbol{\Psi}_n^\dagger\boldsymbol{\Psi}_n=\mathbf{1}_N. \tag{B3}$$

The orbitals of states 1 and 2 will overlap, in general.

$$\mathbf{S}=\boldsymbol{\Psi}_2^\dagger\boldsymbol{\Psi}_1. \tag{B4}$$

It has been necessary to introduce these orbital matrices, so that we can easily write down a suitable expression for $\boldsymbol{\Delta}$.

We choose our variation to be

$$\boldsymbol{\Delta}_2=\boldsymbol{\Psi}_1\boldsymbol{\Psi}_2^\dagger. \tag{B5}$$

The subscript 2 on $\boldsymbol{\Delta}$, indicates that $\boldsymbol{\Delta}_2$ is used to test the stability of the state 2. We shall attach similar subscripts to $E^{(2)}$, $\mathbf{B}^{(1)}$, and $\mathbf{A}^{(1)}$. Although E_2 and $\boldsymbol{\rho}_2$ are unchanged by unitary transformations of $\boldsymbol{\Psi}_2$ or by linear transformations that do not decrease the number of linearly independent orbitals, our expression for $\boldsymbol{\Delta}_2$ would appear to depend on our definition of the orbitals that make up $\boldsymbol{\Psi}_2$. In reality $\boldsymbol{\Delta}_2$ depends only on how $\boldsymbol{\Psi}_1$ is chosen, for any linear transformation of $\boldsymbol{\Psi}_2$ can be regarded as a linear transformation of $\boldsymbol{\Psi}_1$ in Eq. (B5). From Eq. (6) and the definitions given in (B3) and (B5) we obtain

$$\mathbf{v}_2=\boldsymbol{\Psi}_1\boldsymbol{\Psi}_2^\dagger-\boldsymbol{\Psi}_2\mathbf{S}\boldsymbol{\Psi}_2^\dagger. \tag{B6}$$

Using the definition of Eq. (9), we obtain with the above

$$\boldsymbol{\rho}_2^{(2)}=\boldsymbol{\rho}_1-\boldsymbol{\rho}_2-\boldsymbol{\rho}_1\boldsymbol{\rho}_2-\boldsymbol{\rho}_2\boldsymbol{\rho}_1+\boldsymbol{\rho}_2\boldsymbol{\rho}_1\boldsymbol{\rho}_2+\boldsymbol{\Psi}_2\mathbf{S}^\dagger\mathbf{S}\boldsymbol{\Psi}_2^\dagger. \tag{B7}$$

Note that only the last term of $\boldsymbol{\varrho}_2^{(2)}$ depends upon how the orbitals of $\boldsymbol{\Psi}_1$ are defined. We have not written down $\boldsymbol{\varrho}_2^{(1)}$ because we have not been able to introduce any simplifications by considering it explicitly.

Introducing $\boldsymbol{\varrho}_2^{(2)}$ as given in Eq. (B7) into the expression for $E_2^{(2)}$ we can show that the difference E_1-E_2 is introduced. We have first to extend our notation a little. To do this, we define $\mathbf{B}_n$ and $\mathbf{A}_n$ in terms of $\boldsymbol{\varrho}_n$ in the same way that $\mathbf{B}$ and $\mathbf{A}$ are defined in terms of $\boldsymbol{\varrho}$. [See Eqs. (3) and (4).] With these two matrices we can define

$$\mathbf{F}_n=\mathbf{H}+\mathbf{B}_n-\mathbf{A}_n. \tag{B8}$$

The usual form of the Hartree-Fock equation is

$$\mathbf{F}_n\boldsymbol{\Psi}_n=\boldsymbol{\Psi}_n\mathbf{E}_n. \tag{B9}$$

In terms of these matrices we have

$$\begin{aligned}[\mathbf{F}_2\boldsymbol{\varrho}_2^{(2)}]_{\mathrm{tr}}&=[\mathbf{F}_2\boldsymbol{\varrho}_1]_{\mathrm{tr}}+[\mathbf{E}_2(-\mathbf{1}_N+\mathbf{S}^\dagger\mathbf{S}-\mathbf{S}\mathbf{S}^\dagger)]_{\mathrm{tr}}\\ &=E_1-E_2+[\mathbf{E}_2(\mathbf{S}^\dagger\mathbf{S}-\mathbf{S}'\mathbf{S}^\dagger)]_{\mathrm{tr}}\\ &\quad-\tfrac{1}{2}[(\mathbf{B}_2-\mathbf{B}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)\\ &\quad\quad-(\mathbf{A}_2-\mathbf{A}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)]_{\mathrm{tr}}.\end{aligned} \tag{B10}$$

Note that linear transformations of $\boldsymbol{\Psi}_1$ affect only one term of $[\mathbf{F}_2\boldsymbol{\varrho}_2^{(2)}]_{\mathrm{tr}}$. Thus, by linear transformations of $\boldsymbol{\Psi}_1$ we can minimize

$$[\mathbf{E}_2\mathbf{S}^\dagger\mathbf{S}]_{\mathrm{tr}}+\tfrac{1}{2}[(\mathbf{B}_2^{(1)}-\mathbf{A}_2^{(1)})\boldsymbol{\varrho}_2^{(1)}]_{\mathrm{tr}}.$$

Although Eq. (B10) is exact, we have not been able to proceed further without introducing approximations. The most revealing approximation is to assume that $\mathbf{S}$ is diagonal, the diagonal elements being either zero or one. Then the orbitals of $\boldsymbol{\Psi}_2$ are either identical with those of $\boldsymbol{\Psi}_1$, or linear combinations of the virtual orbitals of state 1. Thus, $\mathbf{v}_2$ is a bilinear combination of orbitals which are orthogonal. If we were to construct in analogy to $\mathbf{v}_2$, a variation $\mathbf{v}_1$ to test the stability of state 1 relative to state 2, then in view of our assumptions regarding $\mathbf{S}$,

$$\mathbf{v}_1=\mathbf{v}_2^\dagger. \tag{B11}$$

An immediate consequence of this, is

$$\boldsymbol{\varrho}_1^{(1)}=\boldsymbol{\varrho}_2^{(1)}, \tag{B12a}$$

$$\boldsymbol{\varrho}_1^{(2)}=-\boldsymbol{\varrho}_2^{(2)}. \tag{B12b}$$

If the orbitals of states 1 and 2 can be chosen so that $\mathbf{S}$ has large elements (≈ 1) only along the diagonal, then the Eqs. (B12) can be regarded as approximately valid. In what follows we shall assume that these corrections are too small to alter the sign of the second variation of E.

Introducing the assumption that $\mathbf{S}$ is diagonal, having the elements zero or one, we obtain, using Eq. (B10),

$$\begin{aligned}E_2^{(2)}&=E_1-E_2+\tfrac{1}{2}[(\mathbf{B}_2^{(1)}-\mathbf{A}_2^{(1)})\boldsymbol{\varrho}_2^{(1)}]_{\mathrm{tr}}\\ &\quad-\tfrac{1}{2}[(\mathbf{B}_2-\mathbf{B}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)-(\mathbf{A}_2-\mathbf{A}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)]_{\mathrm{tr}}.\end{aligned} \tag{B13}$$

To construct $E_1^{(2)}$, we first note that in view of assumptions concerning $\mathbf{S}$

$$\boldsymbol{\varrho}_2^{(2)}=\boldsymbol{\varrho}_1-\boldsymbol{\varrho}_2.$$

With the aid of Eqs. (B12), we find that

$$\begin{aligned}E_1^{(2)}&=E_2-E_1+\tfrac{1}{2}[(\mathbf{B}_2^{(1)}-\mathbf{A}_2^{(1)})\boldsymbol{\varrho}_2^{(1)}]_{\mathrm{tr}}\\ &\quad-\tfrac{1}{2}[(\mathbf{B}_2-\mathbf{B}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)-(\mathbf{A}_2-\mathbf{A}_1)(\boldsymbol{\varrho}_2-\boldsymbol{\varrho}_1)]_{\mathrm{tr}}.\end{aligned} \tag{B14}$$

Then we see that $E_1^{(2)}$ and $E_2^{(2)}$ differ only in the sign of E_2-E_1. If $E_2>E_1$, as we assumed at the beginning of this appendix, then

$$E_1^{(2)}>E_2^{(2)}. \tag{B15}$$

This inequality is valid only under the rather restrictive assumption we have introduced concerning $\mathbf{S}$. The important point is that we cannot say definitely that $E_2^{(2)}<0$. The Coulomb and exchange integrals which are common to $E_1^{(2)}$ and $E_2^{(2)}$ can be either positive or negative. But on the other hand, we know that Coulomb and exchange integrals are rather small, so that we would expect that $E_2^{(2)}$ can be positive only if E_2-E_1 is not too large. This forms the basis of our statement, that if there exists more than one absolute minimum of the energy, the energy differences between the various stable states should not differ greatly from the lowest energy.