

Selfconsistent molecular orbital methods. XVI. Numerically stable direct energy minimization procedures for solution of Hartree–Fock equations

Rolf Seeger and John A. Pople

Citation: [The Journal of Chemical Physics](#) **65**, 265 (1976); doi: 10.1063/1.432764

View online: <http://dx.doi.org/10.1063/1.432764>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/65/1?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Extended Numerov method for the numerical solution of the Hartree–Fock equations](#)

J. Math. Phys. **21**, 1366 (1980); 10.1063/1.524559

[Selfconsistent molecular orbital methods. XVIII. Constraints and stability in Hartree–Fock theory](#)

J. Chem. Phys. **66**, 3045 (1977); 10.1063/1.434318

[Selfconsistent timedependent HartreeFock scheme](#)

J. Chem. Phys. **61**, 710 (1974); 10.1063/1.1681950

[SelfConsistent Molecular Orbital Methods. X. Molecular Orbital Studies of Excited States with Minimal and Extended Basis Sets](#)

J. Chem. Phys. **55**, 2236 (1971); 10.1063/1.1676398

[SelfConsistent MolecularOrbital Methods. VIII. Molecular Studies with Least Energy Minimal Atomic Orbitals](#)

J. Chem. Phys. **53**, 932 (1970); 10.1063/1.1674159



Self-consistent molecular orbital methods. XVI. Numerically stable direct energy minimization procedures for solution of Hartree-Fock equations

Rolf Seeger and John A. Pople

Department of Chemistry, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213

(Received 20 June 1975)

Direct energy minimization methods are examined for solution of the Hartree-Fock equations in a basis expansion. Given a set of trial spin-orbitals, an initial direction for a univariate search of spin-orbital space is specified using general energy-weighted coordinates. The first energy minimum in such a search then provides an initial point for the next iterative step. Techniques are proposed for the maintenance of spin-orbital continuity during this process. Comparison of various energy-weighted coordinates indicates that most effective convergence is achieved if the search path is chosen to pass through the point corresponding to the classical procedure involving the construction of successive approximations to the Fock matrix. The utility of the new method is illustrated by the completion of a previous systematic study of Hartree-Fock energies for small organic species.

I. INTRODUCTION

A central problem in Hartree-Fock theory is the minimization of the energy expression

$$\mathcal{E} = \langle \Psi | \mathcal{H} | \Psi \rangle, \quad (1)$$

where \mathcal{H} is the many-electron Hamiltonian and Ψ is a single-determinant electronic wavefunction constructed from a set of orthonormal molecular spin-orbitals. If these spin-orbitals are expanded as linear combinations of a finite set of fixed basis functions, the problem is reduced to the determination of the appropriate coefficients. It is well known that the electronic energy \mathcal{E} will be stationary (not necessarily minimal) if the spin-orbitals are eigenfunctions of a one-electron Fock Hamiltonian. However, since the Fock Hamiltonian itself depends on the spin-orbitals, the problem is nonlinear and cannot be solved directly. The most common practice is the classical iterative procedure, first formulated in the basis-expansion scheme by Roothaan.¹ An initial guess is made for the coefficients which are then used to form a Fock operator from which a new set of coefficients is obtained. When self-consistency is achieved, it can be shown that the value of \mathcal{E} is stationary. This procedure has been found to converge² quite rapidly in many situations and the technique is used in most molecular orbital computer programs.

Sometimes the classical Hartree-Fock procedure fails to converge and divergent or oscillatory behavior sets in.²⁻⁴ Under these circumstances, other techniques have to be devised to achieve the desired energy minimum. One difficulty with the classical procedure is that it may involve discontinuous changes in the trial orbitals which do not necessarily lower the energy. To ensure energy descent, it is necessary to be able to follow such changes continuously. An attempt to do this was made by McWeeny⁵ who considered the consequences of small variations in the density matrix, using a steepest descent method. However, some difficulty is encountered in retaining effective orthonormality of the spin-orbitals since arbitrary variation of the density matrix is not permissible. More recently, Fletcher⁶ has proposed a transformation which allows orthonormality to be re-

tained for all variations. Some applications of these methods have been made to achieve convergence when the classical procedure has failed, but the rate of convergence has proved slow.⁷

Another method was proposed by Hillier and Saunders.⁸ They showed how each iterative step in the classical procedure can be treated by Rayleigh-Schrödinger perturbation theory. Using only first and second order terms in such an expansion, together with a harmonic approximation for energy variation, they were able to achieve better energy lowering for each iteration step than with the classical procedure, and to obtain convergence in situations where the classical method failed. They also showed that their procedure is equivalent to a steepest descent method using energy-weighted coordinates in the orbital space.

In this paper, we shall consider some new techniques which are generalizations of the work of Hillier and Saunders. We again consider continuous variations of the wavefunction, and emphasis is laid on finding an appropriate effective Fock matrix. Generalized energy-weighted coordinates will be used so that the optimum procedure can be chosen. The univariate energy search starting in steepest descent direction will be carried out in finite steps rather than by extrapolating entirely from initial derivatives.

II. THE CLASSICAL ITERATIVE PROCEDURE

In the independent particle approximation, the *trial* wavefunction for m electrons is⁹

$${}^t\Psi = (m!)^{-1/2} \det\{{}^t\chi_1 {}^t\chi_2 \dots {}^t\chi_m\}, \quad (2)$$

where, in the linear combination of atomic orbitals (LCAO) approximation, the m orthonormal molecular spin-orbitals ${}^t\chi_i$ are constrained to be linear combinations of N real atomic spin-orbital basis functions ω_ν ,

$${}^t\chi_i = \sum_{\nu=1}^N t_{\nu i} \omega_\nu. \quad (3)$$

The orthonormality condition is

$$\mathbf{T}^t \mathbf{S} \mathbf{T} = \mathbf{I}, \quad (4)$$

where \mathbf{T} contains the coefficients $t_{\nu i}$ of Eq. (3), \mathbf{S} is the overlap matrix with the elements

$$S_{\mu\nu} = \int \omega_{\mu} \omega_{\nu} d\tau, \quad (5)$$

and \mathbf{I} is the unit matrix. Integration here and elsewhere is over Cartesian and spin coordinates. \mathbf{T}^{\dagger} denotes the Hermitian conjugate of \mathbf{T} . The condition for \mathcal{E} to be stationary leads to the equation

$${}^a\mathbf{F}\mathbf{T} = \mathbf{S}\mathbf{T}\mathbf{E}, \quad (6)$$

where ${}^a\mathbf{F}$ is the Fock matrix in terms of the atomic basis $\{\omega\}$ and \mathbf{E} is the diagonal matrix with its eigenvalues. The elements of ${}^a\mathbf{F}$ are given by

$${}^aF_{\mu\nu} = {}^aH_{\mu\nu} + \sum_{\lambda\sigma} P_{\lambda\sigma} (\mu\lambda | \nu\sigma). \quad (7)$$

Here ${}^aH_{\mu\nu}$ is an element of the core-Hamiltonian matrix and the two-electron integrals are defined by

$$(\mu\lambda | \nu\sigma) = \int \int \omega_{\mu}(1)\omega_{\lambda}(2)(r_{12}^{-1}) \times [\omega_{\nu}(1)\omega_{\sigma}(2) - \omega_{\sigma}(1)\omega_{\nu}(2)] d\tau_1 d\tau_2. \quad (8)$$

The elements of the density matrix \mathbf{P} are given by

$$P_{\lambda\sigma} = \sum_{i=1}^m t_{\lambda i}^* t_{\sigma i}. \quad (9)$$

The expression for the electronic energy [Eq. (1)] then becomes

$$\mathcal{E} = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} ({}^aH_{\mu\nu} + {}^aF_{\mu\nu}). \quad (10)$$

These equations were first obtained by Roothaan¹ for the special case of a closed shell wavefunction in which individual molecular orbitals are doubly occupied with electrons of opposite spin. The above form is more general and encompasses spin unrestricted Hartree-Fock (UHF) theory¹⁰ (different orbitals for different spins).

Although only the spin-orbitals ${}^t\chi_1, \dots, {}^t\chi_m$ are used in the wavefunction [Eq. (2)], it is convenient to complete the set of orthonormal spin-orbitals $\{{}^t\chi\}$ with the remaining eigenfunctions of the Fock Hamiltonian ${}^t\chi_{m+1}, \dots, {}^t\chi_N$, called the *virtual* spin-orbitals. We shall denote the suffixes $m+1, \dots, N$ as a, b, \dots , and the suffixes $1, \dots, m$ as i, j, \dots . The complete set $\{{}^t\chi\}$ is then a basis of an N -dimensional Hilbert space represented by the partitioned matrix \mathbf{T} ,

$$\mathbf{T} = (\mathbf{T}_1 \vdots \mathbf{T}_2). \quad (11)$$

The suffixes 1 and 2 indicate the subsets of occupied and virtual orbitals, respectively.

Since ${}^a\mathbf{F}$ involves \mathbf{T}_1 [through Eqs. (9) and (7)], Eq. (6) is normally solved by an iterative process, starting with some guess for ${}^a\mathbf{F}$ which is constructed without knowledge of \mathbf{T}_1 . The iterative procedure may then be sketched as

$$\begin{array}{c} {}^a\mathbf{F}_{(\text{guess})} \rightarrow (\mathbf{T}_1 \vdots \mathbf{T}_2) \xrightarrow{\quad} \mathbf{P} \\ \mathcal{E}, (\mathbf{T}_1 \vdots \mathbf{T}_2) \xleftarrow{\quad} {}^a\mathbf{F}. \end{array} \quad (12)$$

The iteration is terminated if the root-mean-square

(rms) difference δ_P between the elements of two subsequent density matrices \mathbf{P} is smaller than a preset value.

One source of difficulty in this procedure is the choice of occupied orbitals in \mathbf{T} for each cycle in Eq. (12). The usual practice is to assign the m columns of \mathbf{T} with the lowest eigenvalues E_i to \mathbf{T}_1 . Normally this leads to a converged result for the electronic ground state of a molecule. However, it may sometimes happen that the occupied orbitals \mathbf{T}_1 chosen in this manner are quite different from those in the previous iteration step. This can be very evident when the symmetry of the occupied orbitals changes from one step to the next.⁴ A second source of potential trouble is that there is no guarantee that the electronic energy \mathcal{E} will decrease in the course of the iteration. Both these difficulties must be faced in a search for a more reliable Hartree-Fock procedure.

In order to find such iterative procedures we shall generalize the classical one [Eq. (12)], applying the concept of pseudocanonical spin-orbitals (PCSO) introduced by Hillier and Saunders.⁸ To extend the idea of partitioning to the Fock matrix, we transform it from the atomic basis $\{\omega\}$ to the basis of the trial wavefunction $\{{}^t\chi\}$,

$${}^t\mathbf{F} = (\mathbf{T}_1 \vdots \mathbf{T}_2)^{\dagger} ({}^a\mathbf{F}) (\mathbf{T}_1 \vdots \mathbf{T}_2). \quad (13)$$

${}^t\mathbf{F}$ is then partitioned

$${}^t\mathbf{F} = \begin{pmatrix} {}^t\mathbf{F}_{11} & {}^t\mathbf{F}_{12} \\ {}^t\mathbf{F}_{21} & {}^t\mathbf{F}_{22} \end{pmatrix}. \quad (14)$$

Intraset unitary transformation of $(\mathbf{T}_1 \vdots \mathbf{T}_2)$ leaves the density matrix \mathbf{P} and hence the electronic energy \mathcal{E} unaltered. This degree of freedom may be used to achieve the most appropriate Fock matrix for application of perturbation theory. Such a Fock matrix should contain diagonal submatrices \mathbf{F}_{11} and \mathbf{F}_{22} . The appropriate basis $\{{}^p\chi\}$ is called *pseudocanonical*.⁸ The transformation of the basis $\{{}^t\chi\}$ to $\{{}^p\chi\}$ is defined by the unitary matrix \mathbf{Q} ,

$${}^p\chi = {}^t\chi \begin{pmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{pmatrix} \equiv {}^t\chi \mathbf{Q}, \quad (15)$$

where ${}^p\chi$ and ${}^t\chi$ are row vectors containing the appropriate orbitals as elements. Here \mathbf{Q}_1 and \mathbf{Q}_2 are defined by

$$\mathbf{Q}_1^{\dagger} {}^t\mathbf{F}_{11} \mathbf{Q}_1 = \boldsymbol{\epsilon}_1 \quad (16)$$

and

$$\mathbf{Q}_2^{\dagger} {}^t\mathbf{F}_{22} \mathbf{Q}_2 = \boldsymbol{\epsilon}_2, \quad (17)$$

$\boldsymbol{\epsilon}_1$ and $\boldsymbol{\epsilon}_2$ being diagonal matrices with the elements ϵ_i ($1 \leq i \leq m$) and ϵ_a ($m < a \leq N$), respectively. $\mathbf{0}$ is the zero matrix. Transforming ${}^t\mathbf{F}$ with \mathbf{Q} we obtain

$${}^p\mathbf{F} = \begin{pmatrix} \mathbf{Q}_1^{\dagger} & 0 \\ 0 & \mathbf{Q}_2^{\dagger} \end{pmatrix} \begin{pmatrix} {}^t\mathbf{F}_{11} & {}^t\mathbf{F}_{12} \\ {}^t\mathbf{F}_{21} & {}^t\mathbf{F}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_1 & 0 \\ 0 & \mathbf{Q}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\epsilon}_1 & \mathbf{A}^{\dagger} \\ \mathbf{A} & \boldsymbol{\epsilon}_2 \end{pmatrix}, \quad (18)$$

where

$$\mathbf{A} = \mathbf{Q}_2^{\dagger} ({}^t\mathbf{F}_{21}) \mathbf{Q}_1. \quad (19)$$

For convenience, we define

$$\epsilon = \begin{pmatrix} \epsilon_1 & 0 \\ 0 & \epsilon_2 \end{pmatrix} \quad (20)$$

so that the Fock matrix in terms of the pseudocanonical basis is

$${}^p\mathbf{F} = \mathbf{Q}^\dagger ({}^t\mathbf{F}) \mathbf{Q} = \epsilon + \begin{pmatrix} 0 & \mathbf{A}^\dagger \\ \mathbf{A} & 0 \end{pmatrix} \quad (21)$$

III. ALTERNATIVE SELF-CONSISTENT PROCEDURES

We now turn to the problem of carrying out a search of the complete spin-orbital space to achieve a stationary value (usually a minimum) for the electronic energy \mathcal{E} . Given some starting point with a set of trial spin-orbitals ${}^t\chi$, suppose we obtain the corresponding pseudocanonical spin-orbitals ${}^p\chi$ according to Eq. (15). We then wish to test a new set of trial spin-orbitals, obtained by some unitary transformation of this starting pseudocanonical set. In the classical procedure, this is done by diagonalizing the Fock matrix ${}^p\mathbf{F}$ given by Eq. (21) and transforming ${}^p\chi$ with the resulting eigenvector matrix. This implies a possibly large displacement in the spin-orbital space which may conceivably lead to the difficulties mentioned in the previous section.

In searching the spin-orbital space, it is desirable to examine the energy \mathcal{E} along some path containing the point defined by the preceding PCSO set, ${}^p\chi$, as starting point. This is a *univariate search*.¹¹ A convenient way to perform such a search is to transform ${}^p\chi$ by a unitary matrix $\mathbf{V}(\lambda)$, the elements of which are functions of a dimensionless path parameter λ ,

$$\chi(\lambda) = {}^p\chi \mathbf{V}(\lambda) \quad (22)$$

An appropriate matrix $\mathbf{V}(\lambda)$ may be generated by diagonalizing a generalized Fock matrix

$$\mathcal{F}(\lambda) = \epsilon + \lambda \begin{pmatrix} 0 & \mathbf{J}^\dagger \\ \mathbf{J} & 0 \end{pmatrix}, \quad (23)$$

where \mathbf{J} is an $(N-m) \times m$ matrix and λ may be varied continuously from zero. $\mathbf{V}(\lambda)$ then satisfies

$$\mathbf{V}^\dagger(\lambda) \mathcal{F}(\lambda) \mathbf{V}(\lambda) = \mathbf{\Gamma}(\lambda), \quad (24)$$

with a diagonal matrix $\mathbf{\Gamma}(\lambda)$. The coefficient matrix $\mathbf{T}(\lambda)$ for a point of the search path may then be obtained by transforming the matrix $\mathbf{T}(0)$ of the trial spin-orbitals at the starting point of the path by the product of \mathbf{Q} [from Eq. (15)] and $\mathbf{V}(\lambda)$:

$$\mathbf{T}(\lambda) = \mathbf{T}(0) \mathbf{Q} \mathbf{V}(\lambda) \quad (25)$$

With $\mathbf{T}(\lambda)$, the electronic energy $\mathcal{E}(\lambda)$ is obtained by the sequence

$$\mathbf{T}(\lambda) \rightarrow \mathbf{P} \rightarrow {}^a\mathbf{F} \rightarrow \mathcal{E}(\lambda), \quad (26)$$

which collects Eqs. (7)–(10).

For $\lambda = 0$, $\mathbf{V}(\lambda)$ is the unit matrix, leaving ${}^p\chi$ unchanged. We consider the univariate search for small λ in order to obtain some information concerning the behavior of the search path close to the starting point. Equation (24) may then be linearized for $\mathbf{V}(\lambda)$ to give, to first order,

$$\mathbf{V}(\lambda) = \mathbf{I} + \begin{pmatrix} 0 & -\mathbf{D}^\dagger \\ \mathbf{D} & 0 \end{pmatrix}, \quad (27)$$

where \mathbf{D} is an $(N-m) \times m$ matrix of small elements

$$D_{ai} = -\lambda J_{ai} (\epsilon_a - \epsilon_i)^{-1}. \quad (28)$$

D_{ai} describes the mixing of the virtual spin-orbital ${}^p\chi_a$ with the occupied one ${}^p\chi_i$. Thus, from Eq. (22), again to first order,

$$\chi_i(\lambda) = {}^p\chi_i + \sum_a {}^p\chi_a D_{ai}. \quad (29)$$

The first order variation in the total energy \mathcal{E} is then

$$\delta\mathcal{E} = \sum_i \sum_a^{\text{occ virt}} (D_{ai} A_{ai}^* + D_{ai}^* A_{ai}) \quad (30)$$

From Eqs. (28) and (30) we obtain the initial derivative of the electronic energy \mathcal{E} with respect to the path parameter λ :

$$[d\mathcal{E}(\lambda)/d\lambda]_{\lambda=0} = - \sum_i \sum_a^{\text{occ virt}} (J_{ai} A_{ai}^* + J_{ai}^* A_{ai}) (\epsilon_a - \epsilon_i)^{-1}. \quad (31)$$

This useful expression is a major benefit of the pseudocanonical transformation.

We now turn to the choice of a specific search path as determined by \mathbf{J} and λ through the generalized Fock matrix $\mathcal{F}(\lambda)$. The most apparent procedure is to choose $\mathbf{J} = \mathbf{A}$. If $\lambda = 1$, $\mathcal{F}(\lambda)$ would then be identical with the Fock matrix ${}^p\mathbf{F}$ of the classical procedure. However, it is desirable to consider a more general choice for \mathbf{J} in which the elements contain a weighting factor depending on the energy separation $(\epsilon_a - \epsilon_i)$. This generalizes the concept of energy-weighted steepest descent.⁸ Thus we define

$$J_{ai} = A_{ai} (\epsilon_a - \epsilon_i)^q / \Delta^q. \quad (32)$$

Δ^q ensures that J_{ai} has the same dimension as A_{ai} , and q is a real parameter that may be chosen to optimize the univariate search. For Δ^q we use a weighted mean energy difference such as

$$\Delta^q = [m(N-m)]^{-1} \sum_i \sum_a^{\text{occ virt}} (\epsilon_a - \epsilon_i)^q. \quad (33)$$

Inspecting Eqs. (28) and (32), we see that for $q = 1$, the complex number D_{ai} has the same phase as A_{ai} . Hence $D_{ai} A_{ai}^*$ and $D_{ai}^* A_{ai}$ in Eq. (30) are both real quantities, and we may write Eq. (30) as

$$\delta\mathcal{E} = 2 \sum_i \sum_a^{\text{occ virt}} |D_{ai}| |A_{ai}|. \quad (30')$$

Consequently, the following differential form exists:

$$|A_{ai}| = \frac{1}{2} \left(\frac{\partial \mathcal{E}}{\partial |D_{ai}|} \right)_0; \quad \arg(D_{ai}) = \arg(A_{ai}). \quad (34)$$

Interpreting D_{ai} as element of a displacement vector \mathbf{D} in the spin-orbital space, and A_{ai} as an element of the energy gradient \mathbf{A} , the relationship

$$|D_{ai}| = -\lambda \left(\frac{\partial \mathcal{E}}{\partial |D_{ai}|} \right)_0 / \Delta \quad (35)$$

with any given $\arg(D_{ai})$ may be referred to as *steepest descent*. In Eq. (35), Δ guarantees a dimensionless

step parameter λ . We can generalize this expression to any real q by defining D in terms of an isotropic coordinate system X by means of a set of real anisotropy coefficients α ,

$$D_{ai} = \alpha_{ai} X_{ai}. \quad (36)$$

In the isotropic coordinate system X , the steepest descent may then be defined in a similar manner as above,

$$|X_{ai}| = -\beta \left(\frac{\partial \mathcal{E}}{\partial |X_{ai}|} \right)_0 / \Delta, \quad (37)$$

where Δ guarantees a dimensionless step parameter. The differential expression in Eq. (37) may be expanded in terms of D_{ai} ,

$$\left(\frac{\partial \mathcal{E}}{\partial |X_{ai}|} \right)_0 = \left(\frac{\partial \mathcal{E}}{\partial |D_{ai}|} \right)_0 \left(\frac{dD_{ai}}{dX_{ai}} \right), \quad (38)$$

where $\arg(X_{ai}) = \arg(D_{ai})$ by virtue of Eq. (36). Applying Eqs. (34), (28), (32), and (36)–(38) and $\lambda = -2\beta$, it is easily shown that

$$\alpha_{ai}^2 = (\epsilon_a - \epsilon_i)^{q-1} / \Delta^{q-1}. \quad (39)$$

Hence, D may be defined in terms of X and ϵ ,

$$D_{ai} = [(\epsilon_a - \epsilon_i)^{(q-1)/2} / \Delta^{(q-1)/2}] X_{ai}. \quad (40)$$

A. Initial search direction A ($q = 1$)

Equation (40) shows that this choice for q leads to a steepest descent procedure using X without weights. The elements of J are then given by Eq. (32):

$$J_{ai} = A_{ai}(\epsilon_a - \epsilon_i) / \Delta. \quad (41)$$

The initial derivative of the energy \mathcal{E} with respect to λ in direction of the search path is, according to Eq. (31),

$$[d\mathcal{E}/d\lambda]_{\lambda=0} = -2 \sum_i^{\text{occ}} \sum_a^{\text{virt}} |A_{ai}|^2 / \Delta. \quad (42)$$

This quantity is negative, provided the mean energy difference Δ [Eq. (33)] is positive.

B. Initial search direction B ($q = 0$)

This method defines a search path which will pass through the point corresponding to the next iteration step of the classical procedure at $\lambda = 1$. It is achieved by taking

$$J_{ai} = A_{ai}. \quad (43)$$

With this choice of the initial search direction, the initial derivative of the energy \mathcal{E} with respect to λ is

$$[d\mathcal{E}/d\lambda]_{\lambda=0} = -2 \sum_i^{\text{occ}} \sum_a^{\text{virt}} |A_{ai}|^2 (\epsilon_a - \epsilon_i)^{-1}. \quad (44)$$

This will certainly be negative if all energy differences $(\epsilon_a - \epsilon_i)$ are positive. In this method the mixing between frontier orbitals is emphasized relative to pairs of orbitals with larger energy separation [Eq. (40)].

C. Initial search direction with general q

In order to generalize the concept of anisotropic (energy-weighted) mixing coefficients D_{ai} and to find an optimum value for q we investigated this method involving

general q . The initial derivative of the energy \mathcal{E} with respect to λ is then

$$[d\mathcal{E}/d\lambda]_{\lambda=0} = -2 \sum_i^{\text{occ}} \sum_a^{\text{virt}} |A_{ai}|^2 (\epsilon_a - \epsilon_i)^{q-1} / \Delta^q. \quad (45)$$

Again this quantity is negative if all $(\epsilon_a - \epsilon_i)$ are positive.

Whichever of these initial directions is selected, the search path can be extended to larger absolute values of λ by calculating the energy function $\mathcal{E}(\lambda)$ applying Eq. (26). This univariate search will yield a minimum in $\mathcal{E}(\lambda)$ at some value of λ_{\min} . The corresponding point in the spin-orbital space is then used as starting point for the next univariate search. This iterative process may be terminated in the same manner as the classical one.

IV. TECHNIQUE OF THE UNIVARIATE SEARCH

In order to carry out a univariate search, it is first necessary to choose the sign of λ . To guarantee initial energy descent, this sign must be opposite to that of the initial derivative $[d\mathcal{E}/d\lambda]_{\lambda=0}$. Because this derivative is usually negative, we shall describe the search path for positive λ . However, for $[d\mathcal{E}/d\lambda]_{\lambda=0}$ positive, a similar search with $(-\lambda)$ may be carried out.

In practice it is necessary to use finite intervals τ for the variable λ . With a given τ , a sequence of trial spin-orbitals $\chi(k\tau)$ and the corresponding electronic energies $\mathcal{E}(k\tau)$ for $k = 0, 1, \dots$ is computed. Since we need only a local minimum for $\mathcal{E}(k\tau)$ this series may be terminated at $k = k_e$ where first $\mathcal{E}[(k_e - 1)\tau] < \mathcal{E}(k_e\tau)$. The sequence $\mathcal{E}(k\tau)$ with $k = 0, 1, \dots, k_e$, in conjunction with the initial derivative $[d\mathcal{E}/d\lambda]_{\lambda=0}$ of Eq. (31), defines a polynomial of degree $k_e + 1$ which approximates the search path. This polynomial is then used to yield an approximate λ_{\min} . With Eqs. (25) and (26) we evaluate $T(\lambda_{\min})$ and $\mathcal{E}(\lambda_{\min})$. $T(\lambda_{\min})$ in turn leads to the starting point and the energy gradient [Eq. (34)] for the next univariate search.

In order to maintain continuity of the trial wavefunction along the search path, we have to avoid interchanges of occupied and virtual orbitals, which may occur if the eigenvectors in $V(\lambda)$ are ordered according to ascending eigenvalues Γ_i [Eq. (24)]. At any given point of the search path with $\lambda = k\tau$, the problem may be reduced to the decision whether or not a given eigenvector in the matrix $V(k\tau)$ represents an occupied or a virtual spin-orbital, the corresponding information for the previous matrix $V[(k-1)\tau]$ being available. If v_i and v_a are eigenvectors in $V[(k-1)\tau]$ corresponding to occupied and virtual spin-orbitals respectively, then any normalized vector t of appropriate dimension satisfies the Bessel equation

$$1 = |(t, t)|^2 = \sum_i^{\text{occ}} |(t, v_i)|^2 + \sum_a^{\text{virt}} |(t, v_a)|^2. \quad (46)$$

Now let t be an eigenvector in $V(k\tau)$. t will then be assigned to the occupied set if the first sum in Eq. (46) is close to unity, and to the virtual set if this sum is close to zero. In practice, we use the following criterion for assigning a given eigenvector t to the occupied set:

$$a = \sum_i^{\text{occ}} |(t, v_i)|^2 > 0.75. \quad (47)$$

However, for t representing an occupied spin-orbital, a can be made to approach unity as closely as desired by reducing τ appropriately. We consider τ to be small enough if the criterion [Eq. (47)] leads to exactly m occupied eigenvectors. Usually values of τ between 0.2 and 0.4 are found to be suitable.

Rather than using fixed τ throughout, it may be desirable to use information from one univariate search to obtain an improved τ for the next. Our investigations have shown that the following technique usually leads to a reduced number of search points without compromising numerical stability. Let us define a minimum number of points evaluated for each univariate search, k_{\min} . If λ_{\min} is the path parameter corresponding to the minimum of a given search, then τ for the next search is chosen as

$$\tau = 2\lambda_{\min}/(k_{\min} - 1). \quad (48)$$

With this τ we perform the next search testing the criterion [Eq. (47)] at each point. If it should be violated, τ is reduced, the current point discarded, and an appropriate one with smaller τ is evaluated. The search then proceeds with this reduced τ until $k \geq k_{\min} - 1$ and $\delta[(k-1)\tau] < \delta(k\tau)$.

V. EVALUATION OF SEARCH PROCEDURES

In order to optimize the univariate search with respect to q we have investigated the iterative procedure for the range $-2 \leq q \leq 2$ for several molecules. Since the numerical behavior was quite similar in all those molecules we shall restrict ourselves to discussing one example. This is the lowest triplet state of the methylene radical, CH_2 . We performed calculations of the spin unrestricted Hartree-Fock (UHF) type¹⁰ with the minimal STO-3G basis,¹² using the geometry of lowest

energy ($R_{\text{CH}} = 1.082 \text{ \AA}$, $\angle \text{HCH} = 125.5193^\circ$).¹³ The steepest descent procedures described above have been programmed and incorporated into the Gaussian-70 system.¹⁴

The results of the first univariate search, starting with a set of trial functions obtained from the core-Hamiltonian as initial guess [Eq. (12)], are shown in Fig. 1. Apparently the search path for $q=0$ contains the lowest minimum (-44.578 hartree at $\lambda_{\min} = 1.1133$) of all search paths explored. This minimum is 0.003 hartree lower than the appropriate value obtained by the classical procedure (-44.575 hartree with $\lambda=1$, $q=0$). This particular choice for q corresponds to our method B with an anisotropy factor

$$\alpha_{ai} = (\epsilon_a - \epsilon_i)^{-1/2} \Delta^{1/2}. \quad (49)$$

Compared with the path for $q=1$, corresponding to method A which implies a uniformly weighted coordinate system, method B provides a much larger energy lowering. For infinitesimal τ a search with method B would generally lead to a lower energy than the single calculation ($\lambda=1$) of the classical procedure. In practice, however, we use a finite mesh size τ . If the function $\delta(\lambda)$ is highly anharmonic, its approximation by a polynomial of low degree may lead to an inaccurate minimum. This will reduce the energy lowering for a given univariate search, possibly so much that it might be less than for the classical procedure. However, as the iteration proceeds, $\delta(\lambda)$ is more harmonic and such effects become less important.

Table I compares the energy convergence of the methods A, B, and the classical procedure. For method B, two values of the initial mesh size τ_0 were used. The larger value ($\tau_0 = 0.4$) leads to a first-step energy above the value obtained with the classical procedure, but this does not occur with $\tau_0 = 0.2$. The superiority of method B with either value of τ_0 is reflected in the rapidity of convergence of the total energy.

If the iterative procedure is terminated by reaching a present rms accuracy of the density matrix, the number of univariate searches required is lowest with q close to zero (Fig. 2). Method A ($q=1$) is shown to be rather poor in convergence. In this and other molecules studied, the number of univariate searches required by

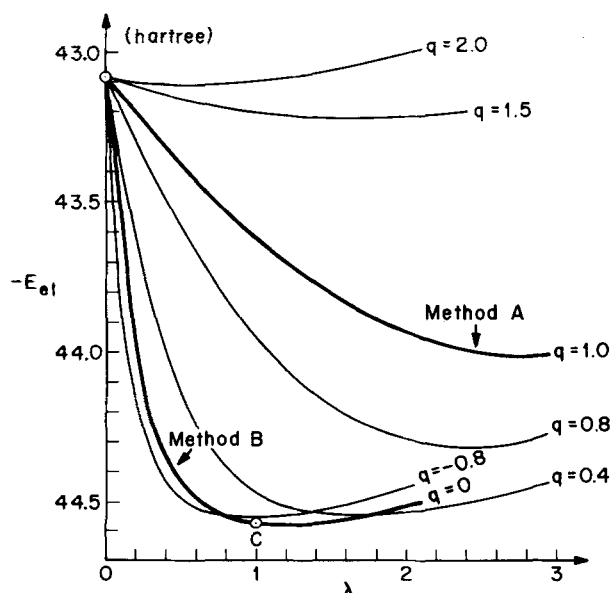


FIG. 1. The electronic energy as a function of the path variable λ of the univariate search with the parameter q . — denotes the paths for the methods A and B. The point C corresponds to the value obtained with the classical procedure.

TABLE I. The electronic energy (hartree) for the ground triplet state of CH_2 in the process of iteration for method A, the classical method, and method B. τ_0 is the initial mesh size of the univariate search.

No. of iteration step	Method A	Classical method	Method B	
	$\tau_0 = 0.4$		$\tau_0 = 0.4$	$\tau_0 = 0.2$
0	-43.0853853	-43.0853853	-43.0853853	-43.0853853
1	-44.0264858	-44.5754896	-44.5678053	-44.5774499
2	3395800	795504	796233	798555
3	4642383	799208	799978	800099
4	5224182	800039	800029	319
5	502269	260	351	56
6	639830	327	62	63
7	711061	51	64	64
8	748514	59		
9	769314	62		
10	781044	64		

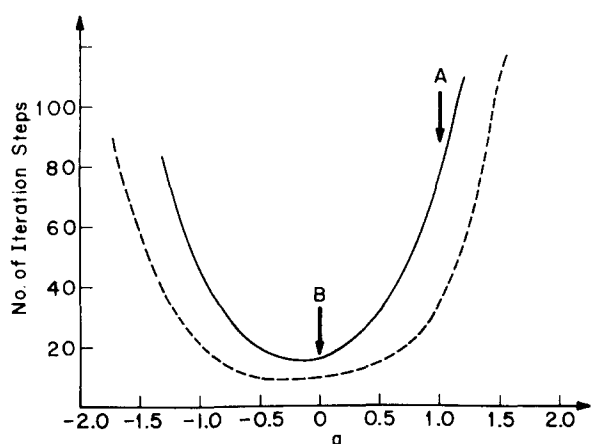


FIG. 2. The number of iteration steps necessary to achieve a preset accuracy δp of the density matrix as a function of the parameter q . $\delta p < 10^{-7}$, $\delta p < 5 \times 10^{-5}$. The arrows A and B indicate the special cases of method A and B, respectively.

method B is about $\frac{2}{3}$ of the number of iterations used by the classical procedure (provided the latter converges at all). However it should be noted that each univariate search involves the evaluation of at least two Fock matrices, so that the required computation time may be greater than for the classical procedure. Nevertheless, method B does provide superior stability at moderate cost.

VI. APPLICATIONS

In order to demonstrate the utility of method B in circumstances where the classical procedure fails, we have completed a previous systematic study of small neutral molecules and cations.³ In this work, equilibrium geometries with the minimal STO-3G basis were found for all systems H_mABH_n (where A, B are carbon, nitrogen,

oxygen, or fluorine) and single Hartree-Fock calculations were then performed using the larger, split-valence 4-31G basis.¹⁵ Some of the 4-31G calculations failed to converge with the classical procedure.⁴ By applying method B, we were able to obtain converged energies in all cases. These are listed in Table II. The number of iteration steps was considerable for some of the molecules (varying from 10 for N_2^+ to 161 for NO with a density cutoff $\delta p < 10^{-6}$), but the iterations all converged monotonically. Once they started with a suitable set of occupied orbitals, the orbitals of the converged results had the same symmetries.

VII. CONCLUSIONS

Of the various techniques considered in this paper for the solution of Hartree-Fock equations, method B is most effective and reliable. Its main features may be enumerated as follows.

1. The iteration proceeds by a series of univariate searches, each leading to an energy lowering. This energy descent feature must ultimately lead to a solution of the Hartree-Fock equations representing a stationary point on the corresponding energy surface.
2. The energy-weighting used in selecting the initial direction for each univariate search is close to optimal for the molecules studied. The rate of convergence of method B is clearly superior to that of method A which uses nonweighted mixing coefficients.
3. The relationship of method B to the classical procedure is close. Each univariate search leads to an energy lowering at least as great as the corresponding iteration step in the classical procedure.
4. By proceeding in sufficiently small steps and selecting the occupied trial spin-orbitals by a criterion of overlap with the previous set, the technique ensures

TABLE II. Calculated total energy (hartree) for some neutral molecules and cations containing the heavy atoms. The geometrical parameters are listed in Ref. 3. The numbers in brackets indicate the number of iteration steps necessary to achieve convergence of $\delta p < 10^{-6}$ with method B. The energy values without this number are obtained by the classical procedure.

Stoichiometric formula	Molecule		State	Energy	
	Structure	Point group symmetry		STO-3G	4-31G
CN	CN	$C_{\infty v}$	$^2\Sigma^+$	-91.026 39	-92.062 64 (50)
CO ⁺	CO ⁺	$C_{\infty v}$	$^2\Pi$	-110.774 60	-112.006 11 (49)
	CO ⁺	$C_{\infty v}$	$^2\Sigma^+$	-110.795 04	-112.075 96 (60)
N_2^+	NN ⁺	$C_{\infty h}$	$^2\Sigma_g^+$	-106.986 75	-108.161 37
	NN ⁺	$C_{\infty h}$	$^2\Pi_u$	-106.988 05	-108.158 27 (10)
NO	NO	$C_{\infty v}$	$^2\Pi$	-127.552 76	-129.040 29 (161)
NOH ⁺	HNO ⁺	C_s	$^2A'$	-127.793 06	-129.188 34 (37)
	HON ⁺	C_s	$^2A'$	-127.814 16	-129.202 64 (42)
NOH ₂	H ₂ ON nonplanar	C_s	$^2A'$	-128.556 16	-130.104 74 (27)
NOH ₄	H ₂ NOH ₂ twisted	C_s	$^2A''$	-129.811 72	-131.381 26 (15)

stability and avoids switching of orbital occupation.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under Grant GP-25617. R. S. wishes to thank the Swiss National Foundation for a generous fellowship.

¹C. C. J. Roothaan, *Rev. Mod. Phys.* **23**, 69 (1951).

²T. Claxton and W. Smith, *Theor. Chim. Acta* **22**, 399 (1971).

³W. A. Lathan, L. A. Curtiss, W. J. Hehre, J. B. Lisle, and J. A. Pople, *Progress in Physical Organic Chemistry* (Wiley, New York, 1974), Vol. II.

⁴For example, in an SCF calculation of nonplanar H₂ON using the geometry described in Ref. 3, and an extended 4-31G basis set,¹⁵ starting with the core Hamiltonian and applying Eq. (12), the wavefunction oscillates between a ²A'' state (with $\epsilon \approx -150.57$ hartree) and two ²A' states (one with $\epsilon \approx -131$ hartree, the other with $\epsilon \approx -143$ hartree). Convergence is not achieved after 128 iteration steps.

⁵R. McWeeny, *Proc. R. Soc. London Ser. A* **235**, 496 (1956).
R. McWeeny, *Rev. Phys.* **32**, 335 (1960).

⁶R. Fletcher, *Mol. Phys.* **19**, 55 (1970).

⁷D. H. Sleeman, *Theor. Chim. Acta* **11**, 135 (1968); R. Kari and B. T. Sutcliffe, *Chem. Phys. Lett.* **7**, 149 (1970); B. T. Sutcliffe, *Theor. Chim. Acta* **33**, 201 (1974).

⁸I. H. Hillier and V. R. Saunders, *Int. J. Quantum Chem.* **4**, 503 (1970); I. H. Hillier and V. R. Saunders, *Proc. R. Soc. London Ser. A* **320**, 161 (1970).

⁹J. C. Slater, *Phys. Rev.* **34**, 1293 (1929).

¹⁰J. A. Pople and R. K. Nesbet, *J. Chem. Phys.* **22**, 571 (1954).

¹¹D. Garton and B. T. Sutcliffe, *Theoretical Chemistry* (The Chemical Society, London, 1974), Vol. 1. D. A. Pierre, *Optimization Theory with Applications* (Wiley, New York, 1969).

¹²W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969).

¹³W. A. Lathan, W. J. Hehre, and J. A. Pople, *J. Am. Chem. Soc.* **93**, 808 (1971).

¹⁴W. J. Hehre, W. A. Lathan, R. Ditchfield, M. D. Newton, and J. A. Pople, Program No. 236, Quantum Chemistry Program Exchange, University of Indiana, Bloomington, Indiana.

¹⁵R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971).