164
5/14/81

R4316
(1)

T.S.

14. 2648

LBL-12157
UC-4
CONF-800726

MASTER

NRCC NATIONAL RESOURCE FOR COMPUTATION IN CHEMISTRY

RECENT DEVELOPMENTS AND APPLICATIONS OF MULTI-CONFIGURATION HARTREE-FOCK METHODS

Proceedings of the Workshop held at
Texas A&M University
College Station, Texas
July 15-17, 1980

NRCC Proceedings No. 10
February 1981

LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No W-7405-ENG-48 and under a grant from the National Science Foundation Grant No. CHE-7721305.

PROCEEDINGS

of the workshop

RECENT DEVELOPMENTS AND APPLICATIONS OF MULTI-CONFIGURATION HARTREE-FOCK METHODS

Sponsored by the

NATIONAL RESOURCE FOR COMPUTATION IN CHEMISTRY

Lawrence Berkeley Laboratory,
University of Caifornia
Berkeley, California 94720

Held at
Texas A & M University
July 15-17, 1980

NRCC Proceedings No. 10

Edited by: Michel Dupuis

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No. W-7405-ENG-48 and under a grant from the National Science Foundation Grant No. CHE-7721305.

CONTENTS

Session I. Generalized Fock Operator Methods

Session II. Annihilation of Single Excitations Methods

Session III. Second-Order MCSCF Methods

Session IV. Applications of MCHF Methods

WORKSHOP PARTICIPANTS

Mr. Raymond A. Bair
Caltech
Noyes Lab 127-72
Pasadena, CA 91125

Dr. Charles W. Bauschlicher
I Case
Mail Stop 132C
NASA Langley Research Center
Hampton, VA 23665

Dr. Frank W. Bobrowicz
Monsanto Research Corporation
Mound Facility
P.O. Box 32
Maimisburg, Ohio 45342

Dr.Nickolas Camp
Department of Chemistry
State University of New York
Buffalo, NY 14214f New York

Prof. Esper Dalgaard
Department of Chemistry
Aarhus University
80000 Aarhus C
DENMARK

Dr. Guru Pada Das
Theoretical Chemistry Group
Chemistry Division
Argonne National Laboratory
Argonne, IL 60493

Prof. Ernest R. Davidson
Department of Chemistry
BG-10
University of Washington
Seattle, WA 89195ington

Dr. John Detrich
Science Applicatioons, Inc.
1211 West 22nd Street
Oak Brook, IL 60521

Dr. Michel Dupuis
NRCC, 50D
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Dr. Stephen T. Elbert
Ames Laboratory
Iowa State University
Ames, Iowa 50010

Prof. Walter England
Department of Chemistry
University of Wisconsin
Milwaukee, WI 53201

Prof. William A. Goddard, III
Department of Chemistry
California Institute of
Technology
Pasadena, CA 91125

Prof. Fritz Grein
Department of Chemistry
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
E3B 5A3, CANADA

Prof. Michael Hall
Department of Chemistry
Texas A&M University
College Station, TX 77843

Dr. Larry B. Harding
Theoretical Chemistry Group
Chemistry Division
Argonne National Laboratory
Argonne, IL 60439

Prof. Jurgen Hinze
Fakultat fur Chemie
Universitat Bielefeld
Postfact 8640
1 Universitatstrasse
4800 Bielefeld
WEST GERMANY

Prof. Harry F. King
Department of Chemistry
Acheson Hall
State University of New York
at Buffalo
Buffalo, NY 14214

Mr. Bill Laidig
Chemistry Department
University of California
Berkeley, CA 94720

Dr. Byron Lengsfield
Laser and Spectroscopy Branch
M.S.283
NASA Langley Research Center
Hampton, VA 23665

Dr. William A. Lester, Jr.
NRCC, 50D
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

Dr. Bernard Levy
Laboratoire de Chimie
Ecole Normale Superieure
de Jeunes Filles
1 rue Maurice Arnous
92120 Montrouge
FRANCE

Dr. Bowen Liu
IBM Research Laboratory
K34/281
5600 Cottle Road
San Jose, CA 95193

Prof. James W. McIver, Jr.
Department of Chemistry
Acheson Hall
State University of New York
at Buffalo
Buffalo, NY 14214

Dr. Richard Martin
Theoretical Chemistry and
Molecular Physics Group
Los Alamos Scientific Laboratory
P.O. Box 1663
Los Alamos, NM 87545

Prof. F.A. Matsen
Department of Chemistry
University of Texas
Austin, TX 78712

Dr. Connie Nellin
Department of Chemistry
The University of Texas
Austin, TX 78712

Dr. Bjorn Roos
University of Lund
Physical Chemistry 2
Chemical Center
P.O. Box 740
S-220-07 Lund
SWEDEN

Prof. Clemens C.J. Roothaan
Department of Chemistry
University of Chicago
Chicago, IL 69637

Prof. Klaus Ruedenberg
Ames Laboratory
Iowa State University
Ames, Iowa 50010

Dr. Ron Shepard
Battelle Memorial Institute
505 King Avenue
Columbus, Ohio 43201

Mr. Maurice Sheppard
Department of Chemistry
University of Chicago
Chicago, IL 60373

Prof. Jack Simons
Department of Chemistry
University of Utah
Salt Lake City, Utah 84112

Dr. Joop H. van Lenthe
16A Clyde Park
Redland, Bristol BS6-RS
GREAT BRITAIN

Dr. Arnold C. Wahl
Science Applications, Inc.
1211 West 22nd Street
Oak Brook, IL 60521

Prof. David Yarkony
Department of Chemistry
Remsen Hall
Johns Hopkins University
Baltimore, MD 21218

Dr. Danny Yaeger
Department of Chemistry
Texas A&M University
College Station, TX 77843

FOREWORD

The National Resource for Computation in Chemistry (NRCC) was established to make information on existing and developing computational methodologies available to all segments of the chemistry community, to make state-of-the-art computational facilities (hardware and software) accessible to the chemistry community, and to foster research and development of new computational methods for application to chemical problems.

Workshops form an integral part of the NRCC's program. Consultation with key workers in the field led us to the conclusion that a timely workshop for 1980 would be one on "Recent Developments and Applications of Multi-Configuration Hartree-Fock (MCHF) Methods." The NRCC is indebted to Prof. Danny L. Yeager of Texas A&M University, and to Dr. Michel Dupuis of the NRCC for organizing the scientific program. Their efforts and the contributions of the participants resulted in these Proceedings.

This work was supported by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U.S. Department of Energy under Contract No. W-7405-ENG-48) and under a grant from the National Science Foundation (Grant No. CHE-7721305).

William A. Lester, Jr.
Director, NRCC

INTRODUCTION

A workshop entitled "Recent Developments and Applications of Multi-configuration Hartree-Fock (MCFH) Methods" was held at Texas A&M University in College Station, Texas from July 15 through July 17, 1980. It was sponsored by the National Resource for Computation in Chemistry (NRCC).

For many years the independent particle Hartree-Fock (HF) model has been used with immense success to describe molecules generally in their ground state, and near the equilibrium conformation. For molecules far from equilibrium, quantum chemists have run into the limitations of the HF model, which often does not describe molecular dissociation properly. For electronic excited states of molecules, the HF model has not been as useful as for molecular ground states, because of the increasing importance of electron correlation effects.

The Configuration Interaction (CI) method is one of the techniques used to obtain a quantitative description of electronic correlation. Perturbation Theory (PT) approaches to electron correlation constitute another way of taking electron correlation into account. Advances in CI and PT methods were the subject of a workshop organized by the NRCC in 1978, entitled "Post Hartree-Fock: Configuration Interaction."

One of the difficulties of the CI method lies in the slow convergence of CI expansions, which often require several thousand configurations. This difficulty is somewhat resolved by optimizing the molecular orbitals used in the construction of a short CI expansion. This is the Multi-Configuration Hartree-Fock (MCHF) method.

A concensus from the previously mentioned CI workshop was that CI calculations can be most efficiently and reliably performed if they are based on a multiconfigurational reference space. In addition, significant advances have occurred in multiconfiguration time-dependent Hartree-Fock methods and effective hamiltonian methods. For these reasons, the organization of a workshop on the titled subject appeared opportune for an exchange of ideas to assess recent progress and to foster exploration of methods which can result in improved algorithms and computer codes.

The Fock operator approach developed in the early 1970's, the "Super-CI" approach, and the recently introduced exponential unitary transformation approach were analyzed in detail. Several application areas were discussed which can benefit from the availability of MCHF wavefunctions: i excited-state computations, ii symmetry breaking problems, and iii time-dependent Hartree-Fock studies, iv potential energy surface calculations.

The feeling emerging from these discussions was that the convergence process for ground and excited state wavefunctions is well understood. The selection of configurations to be included in the MCHF expansion in order to obtain a reliable description of chemical systems, is still a challenging task confronting theoreticians.

Michel Dupuis

Danny L. Yeager

MCSCF USING THE GENERALIZED FOCK OPERATOR

Jürgen Hinze
Department of Chemistry, University of
48 Bielefeld, Germany

A. Golebiewski
Institute of Chemistry, Jagellonian University Cracow, Poland

Introduction

It has become abundantly clear that it is necessary in many cases to use a multi-configurational reference function, if a highly accurate, correlated electronic wavefunction is desired, either via configuration interaction calculations (single and double replacements with respect to a reference) or via perturbation theory. The best method to determine the orbitals required to build the multi-configurational reference is clearly the multi-configuration selfconsistent field (MCSCF) method initially conceived by Frenkel[1] or Hartree[2] and developed into practical usefulness in the last decades by many.[3] The value and importance of the MCSCF method for the effective calculation of accurate electronic wavefunctions of atoms and molecules cannot be overemphasized and need not be belaboured here.

We will restrict ourselves here to a concise derivation of the orbital equations of the MCSCF method and outline the operator formalism to obtain solutions to these equations. To be sure, there are other procedures to solve the general MCSCF orbital problem; procedures, which make use either of the extended Brillouin conditions satisfied by the MCSCF orbitals[4] or those, which determine the elements of a Unitary matrix, which transforms approx. orbitals into MCSCF orbitals, using a many dimensional Newton Raphson technique. These approaches, which are in the general MCSCF case more efficient computationally than the operator techniques detailed here, will be described by other contributors to these conference proceedings.[5,6]

The MCSCF Equations

To derive the MCSCF equations we write the total wavefunction Ψ for state I as a superposition of N configuration state functions (CSF's) Φ as

$$\Psi_I = \Sigma_J^N \Phi_J C_{JI} \qquad (1)$$

where the expansion coefficients can be determined variationally by solving the secular equation in matrix form

$$(\mathbb{H} - \mathbb{1}E_I)\, C_I = 0 \qquad (2)$$

The matrix $\mathbb{1}$ can be chosen to be the unit matrix, since the CSF's can be constructed to be mutually orthonormal without loss of generality. The CSF's are in general specific minimal linear combinations of Slater Determinants (SD's), such as to transform as eigenfunction of the total spin and other symmetry operators commuting with the Hamiltonian. The matrix $\mathbb{H}$ is defined through its elements

$$\mathbb{H}_{IJ} = < \Phi_I | H | \Phi_J > \qquad (3)$$

where H is the Hamiltonian to be specified below of the system considered.

The SD's and thus the CSF's are constructed out of symmetry adapted spin orbitals, which may be chosen to be orthonormal without loosing generality. In general these symmetry and spin orbitals are equivalence restricted, i.e. the same spatial orbital (transforming as an irreducible representation of the symmetry group of the system) is used multiplied with either spin function α or β (and multiplied with the appropriate subspecies function in case of a degenerate symmetry species).

The spatial orbitals are in general chosen to be linear combinations of some basis functions

$$\phi_i = \Sigma_p \chi_p c_{pi} \qquad (4)$$

such that they are orthonormal, i.e.

$$\langle\phi_i|\phi_j\rangle = \delta_{ij}$$

With the equivalence restricted orbitals chosen orthonormal, the spin orbitals

$$\psi_{i\,\lambda\rho} = \phi_{i\lambda}\,\kappa(\lambda\rho) \qquad (5)$$

where $\kappa(\lambda\rho)$ is to represent the subspecies and spin component corresponding to α and β and symmetry species λ, will be orthonormal also, i.e.

$$\langle\psi_{i\lambda\rho}|\psi_{j\rho\sigma}\rangle = \delta_{ij}\delta_{\lambda\mu}\delta_{\rho\sigma} \qquad (6)$$

With this choice the SD's as well as the CSF's will turn out to be orthonormal.

It is now convenient to introduce creation and anihilation operators $a^{+}_{i\lambda\rho}$ and $a_{i\lambda\rho}$ respectively corresponding to the spin orbitals $\psi_{i\lambda\rho}$. These operators[7] are each others Hermetian conjugate and obey the anticommutation relations

$$\begin{aligned} a^{+}_i\, a_j + a_j\, a^{+}_i &= \delta_{ij} \\ a^{+}_i\, a^{+}_j + a^{+}_j\, a^{+}_i &= a_i\, a_j + a_j\, a_i = 0 \end{aligned} \qquad (7)$$

In terms of these operators we can define reduced density operators

$$\gamma^{i}_{j} = \Sigma_\rho\, a^{+}_{i\lambda\rho}\, a_{j\lambda\rho}$$

and

$$\gamma^{ik}_{jl} = \Sigma_{\rho\sigma}\, a^{+}_{i\lambda\rho}\, a^{+}_{k\mu\sigma}\, a_{l\mu\sigma}\, a_{j\lambda\rho} \qquad (8)$$

in the space of the spatial orbitals.

We will refrain from referencing symmetry from here on, since its introduction should be clear, and it would only clutter the notation.

With these definitions we can write the total (spin free) Hamiltonian as

$$H = \Sigma_{ij}\, h^{j}_{i}\gamma^{i}_{j} + 1/2\, \Sigma_{ijkl}\, g^{jl}_{ik}\gamma^{ik}_{jl}$$

with

$$h_i^j = <\phi_i|-1/2\nabla^2 + V|\phi_j>$$

and (9)

$$g_{ik}^{jl} = <\phi_i<\phi_k \;|{}^1/r_{12}|\phi_l>\phi_j>$$

Corresponding to the density operators we obtain the density matrix elements

$$\Gamma_{Ij}^{Ii} = <\psi_I|\gamma_j^i|\psi_I> \qquad (10)$$

for the MCSCF state function desired, with a similar expression for the elements of the second order reduced density matrix.

*These density matrix elements may be ex*pressed in terms of the transition density elements between the configuration state functions

$$\tilde{\Gamma}_{Jj}^{Ki} = <\Phi_J|\gamma_j^i|\Phi_K> \qquad (11)$$

which are independent of the detailed form of the orbitals. We obtain

$$\Gamma_{Ij}^{Ii} = \Sigma_{JK}\, C_{JI}\, C_{KI}\, \tilde{\Gamma}_{Jj}^{Ki} \qquad (12)$$

and a similar expression for the density matrix elements of higher order.

At times, in particular if wavefunctions for excited states are desired, it may be convenient to optimize the orbitals in a weighted averaged field for several states,[8] therefore it is useful to introduce the averaged density matrix elements

$$\Gamma_j^i = \Sigma_I\, \omega_I \Gamma_{Ij}^{Ii} \qquad (13)$$

where the ω's are the weighting coefficients which may be chosen to suite the problem at hand with the restriction

$$\Sigma_I\, \omega_I = 1 \qquad (14)$$

We will use in the following these mean density matrix elements, which permits us to remain general and suppress the state index of the final wavefunction desired.

Using these definitions we can express the expectation value of the energy

$$<E> = \Sigma_{ij}\, h_i^j\, \Gamma_j^i + 1/2\, \Sigma_{ijkl}\, g_{ik}^{jl}\, \Gamma_{jl}^{ik} \qquad (15)$$

in a form exhibiting the orbital dependence explicitely.

Variation of this expression with respect to a change of the orbitals leads to the MCSCF orbital equations

$$\Sigma_j\, G_{ij}|\phi_j> = \Sigma_j|\phi_j>\epsilon_{ji} \qquad (16)$$

where the ϵ_{ij}'s are Lagrange multipliers introduced to maintain the orthonormality constraints of the orbitals in the variation.

The operators G_{ij} are given as

$$G_{ij} = h\Gamma_j^i + \Sigma_{kl}\, U_k^l\, \Gamma_{jl}^{ik} \qquad (17)$$

with

$$h = -1/2\, \nabla^2 + V$$

and

$$U_k^l = \langle \phi_k | 1/r_{12} | \phi_l \rangle$$

Before we discuss procedures to solve eq. (16) for the determination of the orbitals it is useful to note that

$$\varepsilon_{ji} = \langle \phi_j | \Sigma_k G_{ik} | \phi_k \rangle = \langle \phi_j | F | \phi_i \rangle = F_{ji} \quad (18)$$

can be thought of as the i, j'th element of a general abstract Fock operator F in the orbital space.

The necessary and sufficient conditions for eq. (16) to be satisfied are

$$\varepsilon_{ji} = \varepsilon_{ij} \quad (19)$$

which is the same as demanding that the Matrix **F**, defined with its elements in eq. (18), be Hermitian.

In fact it is easy to show[3c] that

$$\varepsilon_{ji} - \varepsilon_{ij} = \langle \gamma_i^j \Psi | H | \Psi \rangle - \langle \gamma_j^i \Psi | H | \Psi \rangle \quad (20)$$

and thus eq. (19) is equivalent to the extended Brillouin conditions

$$\langle (\gamma_i^j - \gamma_j^i)\, \Psi | H | \Psi \rangle = 0 \quad (21)$$

satisfied by MCSCF wavefunctions.[4] We have used here the first order reduced density operators defined through eq. (8) as general single excitation operators, which they are.[5]

Solving the MCSCF equations: The non-linear second order differential equations, eq. (16), are generally solved in matrix form, a form obtained by expressing the unknown orbital ϕ_i as a linear combination of basis functions χ_p, eq. (4). Thus starting with a set of m basis functions $\{\chi_1 \ldots \chi_m\}$ we obtain the orbitals through a linear transformation

$$\{\phi_1 \ldots \phi_n \phi_{n+1} \ldots \phi_m\} = \{\chi_1 \ldots \chi_m\}\, \mathbf{C} \quad (22)$$

where **C** is the mxm matrix of the orbital expansion coefficients of eq. (4). In general only the first n orbitals, the active set, participating in the MCSCF function are of interest and can be determined variationally through solving eq. (16). The m-n orbitals $\{\phi_{n+1} \ldots \phi_m\}$, the empty or virtual set, cannot be determined uniquely without auxilliary conditions. Since these orbitals are not needed in the construction of Ψ, and transformations among them leave Ψ and therefore the energy invariant, we need not bother about them. All we need to determine are the first n columns of C and thus the orbitals ϕ_1 through ϕ_m, the active orbital set.

To obtain the active orbital set $\{\phi_1 \ldots \phi_m\}$, using operator techniques to solve eq. (16), we rewrite eq. (16) in matrix form in the mxm space of the basis functions as

$$\Sigma_j^n \mathbf{G}_{ij}\, \mathbf{C} = \Sigma_j^n \mathbf{S}\, \mathbf{C}_j\, \varepsilon_{ji} \qquad (23)$$

for i=1,...n, with

$$\mathbf{G}_{ij} = \mathbf{h}\, \Gamma_j^i + \Sigma_{kl}^n \mathbf{U}_k^l\, \Gamma_{jl}^{ik} \qquad (24)$$

The matrices in eqs. (23,24) are defined through their elements, which are molecular integrals over the original basis functions as follows.
The overlap matrix $\mathbf{S}$

$$S_{pq} = <\chi_p|\chi_q> \qquad (25)$$

The one particle operator matrix $\mathbf{h}$

$$h_{pq} = <\chi_p|-1/2\ \nabla^2 + V|\chi_q> \qquad (26)$$

and the two particle interaction matrix
$\mathbf{U}_k^l$

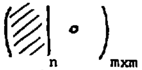
$$U_{k,pq}^l = \Sigma_{rs}^m <\chi_p<\chi_r|^1/r_{12}|\chi_s>\chi_q>C_{rk}C_{sl} \qquad (27)$$

Because of the dependence of $\mathbf{U}$ on the solution vectors, eq. (23) is non-linear. With the operator techniques, this non-linear problem is solved using a first order iterative process. Starting with some zero'th order $\mathbf{C}^o$ the $\mathbf{G}$'s are constructed, and for the fixed $\mathbf{G}$'s a new $\mathbf{C}$ is obtained by solving eq. (23). This new $\mathbf{C}$ is used in a second cycle to construct $\mathbf{G}$ and the solution process is repeated until it converges (hopefully), i.e. the input $\mathbf{C}$ is equal to the output $\mathbf{C}$.
We have to focus now on techniques to solve eq. (23) with the $\mathbf{G}_{ij}$ constructed from some $\mathbf{C}^o$ and held fixed.

a) <u>The orthogonal gradient method.</u>[9]
Given the G's we can construct the matrix of Lagrange multipliers, i.e. the matrix F of the generalized Fock operator in the orbital space defined as eq. (18)

$$\mathbf{F}_{ji} = \mathbf{C}_j^+\, \Sigma_k^n \mathbf{G}_{ik}\, \mathbf{C}_k \qquad (28)$$

Since $\mathbf{G}_{ik}$ is zero for i>n the matrix $\mathbf{F}$ has tne structure



i.e. only the first n columns are different from zero, it is essentially an nxm matrix.

For the solution orbitals we should have $\mathbf{F} = \mathbf{F}^+$ that is, only the first nxn block should be different from zero.
In order to find a transformation which brings us from the initial $\mathbf{C}^o$ orbital coefficient matrix closer to a solution we require that

$$\mathbf{U}^+\, \mathbf{F} = \mathbf{F}^+\, \mathbf{U} \qquad (29)$$

Eq. (29) is solved immediately by chosing the mxn submatrix of the orthogonal matrix $\mathbf{U}$ desired as

$$\tilde{\mathbf{U}} = \mathbf{F}\,(\mathbf{F}^+\,\mathbf{F})^{-1/2} \qquad (30)$$

The new orbitals are then obtained from the transformed orbital coefficient matrix

$$\mathbb{C} = \mathbb{C}^{\circ}\,\mathbb{U} \tag{31}$$

where $\mathbb{U}$ has as its first n columns the elements of $\tilde{\mathbb{U}}$ and the final n-m columns are chosen to complete it to a full mxm orthogonal matrix. This freedom to chose the last n-m columns of $\mathbb{U}$ reflects the invariance of the wavefunction Ψ to unitary transformations within the empty orbital set.

To ensure the approach of a minimum of <E> in this process it is important to select the negative square roots of the eigenvalues of $\mathbb{F}^{+}\,\mathbb{F}$ when forming $(\mathbb{F}^{+}\,\mathbb{F})^{-1/2}$. The iterative process can be controlled, damped or accelerated by a level shifting procedure by adding (in principle) arbitrary constants d_i to the diagonal elements of $\mathbb{F}$.

b) Direct diagonalization of $\mathbb{F}$.
Recently Polezzo[10] has suggested and tested another procedure to obtain $\mathbb{U}$ in eq. (31) such that eq. (29) is satisfied for a given $\mathbb{F}$. He suggests to find directly the singular (or principal) values Λ of the non-symmetric mxm matrix $\mathbb{F}$ together with their associated singular right and left eigenvectors $\mathbb{V}$ and $\mathbb{W}$,[11] such that

$$\mathbb{W}^{+}\,\mathbb{F}\,\mathbb{V} = \Lambda \tag{32}$$

Then $\mathbb{U}$ is given by

$$\mathbb{U} = \mathbb{W}\,\mathbb{V}^{+} \tag{33}$$

Polezzo gives also a procedure, which he believes to be satisfactory for the determination of a level shifting parameter to speed up convergence of the first order iterative sequence.

c) Generalized Jacobi Diagonalization.[12]
If no two CSF's participating in the MCSCF wavefunction differ from each other by less than a double replacement the only non zero elements of the reduced density matrices will be Γ_i^i, Γ_{ij}^{ij} and Γ_{jj}^{ii}. With this restriction the MCSCF orbital equation, eq. (16), can be written as

$$G_{ii}|\phi_i\rangle = \Sigma_j^n|\phi_j\rangle\,\varepsilon_{ji} \tag{34}$$

with

$$G_{ii} = h\,\Gamma_i^i + \Sigma_j(I_j\,\Gamma_{ij}^{ij} + K_j\,\Gamma_{jj}^{ii}) \tag{35}$$

where we have used the standard Coulomb

$$I_j\,|\,\phi_i\rangle = \langle\phi_j|{}^1/_{r_{12}}|\phi_j\rangle|\phi_i\rangle$$

and exchange operators

$$K_j|\phi_i\rangle = \langle\phi_j|{}^1/_{r_{12}}|\phi_i\rangle|\phi_j\rangle \tag{36}$$

We may now form a set of matrices, corresponding to the set of operators G_{kk} in the space of all the m orbitals which can be constructed from a given basis. The elements of these matrices are defined as

$$G_{ij}^{k} = \langle \phi_i | G_{kk} | \phi_j \rangle \qquad (37)$$

for i,j=1,2,..., n and k=1,2,...,n. It is frequently possible to reduce the number of matrices since we get for several orbitals k of the active set identical Fock type operators G_{kk}. The necessary and sufficient conditions for orbitals to satisfy the MCSCF equations, eq. (39) are now

$$\varepsilon_{ji} - \varepsilon_{ij} = G_{ji}^{i} - G_{ij}^{j} = 0 \qquad (38)$$

An angle of rotation between orbitals i and j can be calculated such as to satisfy eq. (38). Using the elementary unitary matrix thus obtained all the matrices G_{kk} in the orbital space can be transformed and the process repeated until eq. (38) is satisfied for all i and j. This process is still first order, since the Fock-like operator G_{kk} eq. (35), represented as matrices in the orbital or basis function space are kept constant even though they depend on the orbitals.

d) Expansion of the Orbital Transforming Unitary Matrix

The general MCSCF equation, eq. (16), can be solved by expanding the unitary matrix $\mathbf{U}$, which transforms initial orbitals $\{\phi^\circ\}$ represented by $\mathbf{C}^\circ$ into final orbitals $\{\phi\}$ represented by $\mathbf{C}$ via

$$(\phi_1 \ldots \phi_m) = (\phi_1^\circ \ldots \phi_m^\circ)\,\mathbf{U} \qquad (39)$$

or

$$\mathbf{C} = \mathbf{C}^\circ\,\mathbf{U} \qquad (40)$$

as

$$\mathbf{U} = \mathbf{1} + \mathbf{V} \qquad (41)$$

with $\mathbf{V}$ antihermitian (antisymmetric). To solve eq. (16) we may use the conditions, eq. (19) together with eqs. (23,24) and obtain

$$\varepsilon_{ji}-\varepsilon_{ij} = \mathbf{C}_j^{+}\Sigma_k^{n}\mathbf{G}_{ik}\mathbf{C}_k - \mathbf{C}_i^{+}\Sigma_k\mathbf{G}_{jk}\mathbf{C}_k = 0 \qquad (42)$$

where we have assumed for simplicity all the orbitals to be real; an assumption which is not necessary. Substituting now for the $\mathbf{C}$'s those obtained from eqs. (40,41)

$$\mathbf{C} = \mathbf{C}^\circ + \mathbf{C}^\circ\,\mathbf{V} \qquad (43)$$

and neglecting terms of order $\mathbf{V}^2$, we can after some algebra derive a set of linear equations for the determination of the nonredundant elements V_{ij} for $i \le m$ and $j \le n$ of the form [3c]

$$\Sigma_{kl}\, M_{ij,kl}\, V_{kl} = -\, g_{ij} \qquad (44)$$

with

$$g_{ij} = \varepsilon_{ji} - \varepsilon_{ij} \qquad (45)$$

evaluated as given in eq. (42) with $\mathbf{C}^\circ$'s.

The matrix elements of $\mathbf{M}$ can be obtained *in first order, by neglecting the dependence* of the operators G_{ij} on the $\mathbf{V}$'s as[3c]

$$M_{ij,kl} = \delta_{jl}\, \mathbf{C}_k^+ \mathbf{F}_i - \delta_{jk}\, \mathbf{C}_l^+ \mathbf{F}_i + \\ + \mathbf{C}_j^+ \mathbf{G}_{il} \mathbf{C}_k - \mathbf{C}_j^+ \mathbf{G}_{ik} \mathbf{C}_l \qquad (46) \\ - \delta_{il}\, \mathbf{C}_k \mathbf{F}_j + \delta_{ik}\, \mathbf{C}_l^+ \mathbf{F}_j \\ - \mathbf{C}_i^+ \mathbf{G}_{jl} \mathbf{C}_k + \mathbf{C}_i^+ \mathbf{G}_{jk} \mathbf{C}_l$$

evaluated again with $\mathbf{C}^{\circ}$'s. We have dropped the superscript ° for simplicity. The $\mathbf{F}_i$ vectors are defined as

$$\mathbf{F}_i = \Sigma_k^1\, \mathbf{G}_{ik}\, \mathbf{C}_k \qquad (47)$$

If one considers also the dependence of the operators G_{ij} on the V's we obtain a second order process with the following additional terms to the matrix elements of $\mathbf{M}$.

$$\bar{M}_{ij,kl} = M_{ij,kl} + \Delta_{ij,kl} \qquad (48)$$

with

$$\Delta_{ij,kl} = 2\, \Sigma_{mn}\{\mathbf{C}_j^+ \mathbf{Q}_{nm} \Gamma_{ml}^{in} \mathbf{C}_k \\ - \mathbf{C}_j^+ \mathbf{Q}_{nm} \Gamma_{mk}^{in} \mathbf{C}_l - \mathbf{C}_i^+ \mathbf{Q}_{nm} \Gamma_{ml}^{jn} \mathbf{C}_k \\ + \mathbf{C}_i^+ \mathbf{Q}_{nm} \Gamma_{mk}^{jn} \mathbf{C}_l\} \qquad (49)$$

where

$$Q_{nm,pq} = \Sigma_{rs} 1/2(\langle p\langle r|{}^1/r_{12}|q\rangle S\rangle \\ + \langle p\langle S|{}^1/r_{12}|q\rangle r\rangle)\, C_{rn}\, C_{sm} \qquad (50)$$

Discussion:

Obviously the last two procedures, in particular the last, appear computationally much more cumbersome than the processes outlined under a) through c). A large set of linear equations of dimension $n(n-1)/2 + n(m-n)$ has to be constructed and solved in order to solve for the orbitals in each SCF iterations.
However, since the most time consuming step, in a general ab-initio MCSCF calculation, is the 4-index integral transformation, a step of order nm^4, which has to be done each iteration, it is clear that the number of SCF iterations should be kept as small as possible. This is achieved when using the method described last, since it should be quadratically convergent. The first order processes remain important for approximation or semi-empirical MCSCF calculations[13] where 4-index transformations can be avoided.

Finally a few words of caution appear to be in order to prevent undue enthusiasm in favor of the quadratically convergent MCSCF procedures. For large scale calculations neither $\underset{\sim}{\mathbf{M}}$ nor the integrals needed to construct $\underset{\sim}{\mathbf{M}}$ will fit into the fast store of a computer. In this case efficient algorithmes remain to be developed for the construction of $\tilde{M}$ before these procedures can be used effectively. For the solution of the large set of linear equations, eq. (44), a conjugate gradient iterative method should be appropriate. However, as for any quadratically convergent process, the quadratic convergence is reached only near the final solution. With orbitals still far from their final MCSCF form, it may be expidient to use different approaches, as those based on the extended Brillouin theorem, at least for the first few iterations.

References

1) J. Frenkel, "Wave Mechanics" Vol. 2, "Advanced General Theory" (Clarendon, Oxford, 1934) pp. 460-462.

2) D.R. Hartree, W. Hartree and B. Swirles, Philos. Trans. Roy. Soc. London, A 238, 229 (1939).

3a) A.P. Jucys, Adv. Chem. Phys. 14, 191 (1969), review of the atomic MCSCF work done by Jucys (Yutsis, Iutsis) dating back to 1950.

b) G. Das and A.C. Wahl, J. Chem. Phys. 44, 87 (1966); 47, 2934 (1967).

c) J. Hinze, J. Chem. Phys., 59, 6424 (1973)

d) T.L. Gilbert, Phys. Rev. A6, 580 (1972)

e) A.C. Wahl and G.Das in "Modern Theoret. Chem.", H.F. Schaefer Ed., Vol. III, Chap. III (Plenum 1976) give a review.

4) B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968); (errata) 3, 247 (1969)

5) F. Grein and T.C. Chang, Chem. Phys. Lett. 12, 44 (1971); A. Banerjee and F. Grein, Int. J. Quantum Chem. 10, 123 (1976); P.A. Ruttnik and J.H. van Lenthe, Theor. Chim. Acta 44, 97 (1977); K. Ruedenberg, L.M. Cheung and S.T. Elbert, Int. J. Quant. Chem. 16, 1068 (1979).

6) J. Kendrick and I.H. Hiller, Chem. Phys. Lett. 41, 283 (1976); D.L. Yeager and P. Jørgensen, J. Chem. Phys. 71, 755 (1979); C.C.J. Roothaan, J. Detrich and D.G. Hopper, Int. J. Quantum Chem. S 13, 93 (1979); H.-J. Werner and W. Meyer, J. Chem. Phys. 73, 2342 (1980).

7) See for example D.H. Kobe, Am. J. Phys. 34, 1150 (1966).

8) K.K. Docken and J. Hinze, J. Chem. Phys. 57, 4928 (1972).

9) A. Golebiewski, J. Hinze and E. Yurtsever, J. Chem. Phys. 70, 1101 (1979).

10) S. Polezzo and P. Fantucci, Mol. Phys. 36, 1835 (1978); 39, 1527 (1980).

11) G.W. Stewart, "Introduction to Matrix Computations" (Acad. Press 1973); G.H. Golub and C. Reinsch, Num Math. 14, 403 (1970).

12) J. Hinze and E. Yurtsever, J. Chem. Phys. 70, 3188 (1979).

13) C.W. Eaker and J. Hinze, J. Am. Chem. Soc. 96, 4084 (1974).

A QUADRATICALLY CONVERGENT MCSCF THEORY FOR LARGE CONFIGURATION SETS

G. Das
Chemistry Division, Argonne National Lab., Argonne, Ill.

Summary

A multi-configuration self-consistent field method for large configuration-sets is presented. It consists in identifying a primary subset containing the predominant configurations and a secondary one containing the rest. The formalism achieves near-quadratic convergence as well as a significant simplification of the orbital-transformation problem by the introduction of Fock equations and Fock matrices.

I. INTRODUCTION

The MCSCF method[1,2] has, over the years, been demonstrated to be a powerful tool for the study of atomic and molecular interactions.[3-6] Until recently, however, the applicability of the method remained restricted to a small number of configurations and orbitals and primarily to the lowest states of a given symmetry. Customarily this shortcoming is compensated for by following up the MCSCF step with a large-scale CI based on the MCSCF orbitals. However, for problems involving a large number of valence electrons as in the case of transition metal clusters, it is very likely that the basic MCSCF problem itself contains a fairly large number of configurations(∼500). Further, for a practical calculation it is also necessary that the corresponding MCSCF process has a fast convergence and applies equally well to the ground and excited states.

The traditional method[1] of solving the MCSCF problem has been along the lines of the Hartree-Fock theory, i.e., by deriving Fock equations for the orbitals and solving them in consonance with the CI-secular equation. Since this method neglects some important coupling terms (as will be explained later), it can run into convergence problems which are then usually handled by some kind of extrapolation. This is true especially for the excited states for which the iteration-to-iteration changes in the orbitals are strongly coupled to those of the CI-coefficients. Recently a number of authors[4-6] have developed and implemented methods which take this coupling into account. However, all these methods involve huge Hessian matrices and large four-index transformation of integrals. The dimension of these matrices can be quite large depending upon the basis set and the configuration set sizes. On the other hand in the Fock equation approach the matrices required to be inverted have the same dimension as the basis set and the four-index transformation involves only the occupied orbitals.

In what follows we will present an MCSCF formulation that combines the advantages of the Fock theory approach with the generality and the efficiency of the fully-coupled orbital transformation method.

II. A NEW METHOD OF SOLUTION OF SECULAR EQUATIONS FOR LARGE MATRICES

The following method is ideally suited to the solution of secular equations of large Hamiltonian matrices in an MCSCF framework. Let us break up the configuration set into smaller groups such that the number of configurations in any one of these groups is small enough for a straightforward "in-core" diagonalization. We assume that the first group contains the important configurations of the root of the secular equation being sought as well as those of the lower-lying roots. We also assume that the mixing coefficients for the state in question are approximately known.

Let $\{A_a^0\}$ be these approximate coefficients to be used as a starting point for the following iterative process: For every group we define a secular equation:

$$\begin{vmatrix} H_{11}-\lambda & H_{12} & \cdots\cdots & H_{1n_G} & H_{10} \\ H_{21} & H_{22}-\lambda & \cdots\cdots & H_{2n_G} & H_{20} \\ \vdots & & & & \\ H_{n_G1} & H_{n_G2} & \cdots\cdots & H_{n_Gn_G}-\lambda & H_{n_G0} \\ H_{01} & H_{02} & \cdots\cdots & H_{0n_G} & H_{00}-\lambda \end{vmatrix} = 0 \qquad (1)$$

where the elements $(1,2,....n_G)$ denote the members of the configuration group G and the element "0" corresponds to the rest of the configurations. Thus

$$H_{io} = \sum_{a \not\subset G} A_a'^{(n-1)} \langle a|H|i\rangle \qquad (2)$$

where the summation goes over the configurations outside the group G and A_a' is given by

$$A_a'^{(n-1)} = A_a^{(n-1)} \Big/ \sum_{b \not\subset G} [A_b^{(n-1)}]^2 \qquad (3)$$

The superscript (n-1) denotes that the mixing coefficients are from the (n-1)th iteration. The secular equation coming from the first group of configurations is solved for the desired root and all the lower ones. The other secular equations are solved only for the lowest root. The process is iterated with each iteration yielding a new set of mixing coefficients for the next one.

III. RESUMÉ OF THE 'FOCK' APPROACH TO THE MCSCF PROBLEM

Based on the mixing coefficients obtained by solving the secular equation, Fock equations represent extremum conditions to be satisfied by the occupied orbitals and have the following general expansion form:

$$\mathbf{F}_i\mathbf{c}_i + \mathbf{f}_i = \sum \varepsilon_{ij}\mathbf{S}\mathbf{c}_j \tag{4}$$

with the Lagrangian multipliers satisfying the condition

$$\varepsilon_{ij} = \varepsilon_{ji}$$

i.e.,

$$\mathbf{c}_j^+(\mathbf{F}_i\mathbf{c}_i + \mathbf{f}_i) = \mathbf{c}_i^+(\mathbf{F}_j\mathbf{c}_j + \mathbf{f}_j) \tag{5}$$

We[1] have solved these equations in the past by the following iterative process: Given an initial set of vectors $\{\mathbf{c}_i^{0}\}$ one defines a transformation:

$$\mathbf{c}_i'^{(0)} = \sum_j u_{ij}\mathbf{c}_j^{(0)} \tag{6}$$

such that $\mathbf{c}_i'$ satisfy the Eqs (5). This is termed the occupied-space solution. Eq.(4) is now rewritten in the following approximate form:

$$\left\{\mathbf{F}_i - \sum_j [\mathbf{F}_i\mathbf{c}_j^{(0)}(\mathbf{S}\mathbf{c}_j^{(0)})^+ + (\mathbf{S}\mathbf{c}_j^{(0)})(\mathbf{F}_i\mathbf{c}_j^{(0)})^+]\right\}\delta\mathbf{c}_i^{(0)}$$

$$= -\left[\mathbf{F}_i\mathbf{c}_i^{(0)} + \mathbf{f}_i - \sum\varepsilon_{ij}^{(0)}\mathbf{S}\mathbf{c}_j^{(0)}\right] \tag{7}$$

where, obviously, $\delta\mathbf{c}_i^{(0)}$, obtained by matrix inversion, lies in the virtual space. This part of the solution is termed the virtual space solution.

The new vectors after the first iteration are, therefore, of the form:

$$\mathbf{c}_i^{(1)} = \sum_j u_{ij}\mathbf{c}_j^{(0)} + \delta\mathbf{c}_i^{(0)} \tag{8}$$

These are then orthonormalized and used to calculate new energy and mixing coefficients, whence the new Fock equations are constructed and solved yielding new vectors $\mathbf{c}_i^{(1)}$ and so on.

IV. DEFICIENCIES OF THE FOCK EQUATION APPROACH

Although in most cases the above approach would converge to a solution of Eqs.(4) and (5), very often such a solution may not correspond to an energy extremum, since the Fock operators constructed using the new vectors may be very different from the old ones making this solution unacceptable. In other words this leads to an intrinsically non-convergent process. A modification of the process, first introduced by Das[7], is to consider changes in the Fock operator in the step in which the occupied-space solution is obtained. This modification has been found very successful over the years for the ground state calculations. Except in pathological cases (involving single excitations) the method leads to convergent solutions with occasional use of 'extrapolations'.

The method, expectedly, suffers from slow convergence, since, firstly, the coupling between the occupied-space and the virtual space corrections is left out. Secondly the changes in the mixing coefficients are not coupled to orbital changes. The latter is particularly important for excited states, since such coupling may be large enough to decide the course of the solution toward the particular state under consideration. For excited states introduction[8] of a limited amount of coupling has been shown to lead to considerable improvement in the rate of convergence.

V. A REFORMULATION OF THE MCSCF SCHEME

Let us add to the occupied set $\{\mathbf{c}_i^{(0)}\}$ a set of orthonormalized vectors formed out of the normalized vectors:

$$\mathbf{v}_i' = \delta\mathbf{c}_i^{(0)} / |\delta\mathbf{c}_i^{(0)}| \tag{9}$$

obtained from the virtual-space solution of the Fock equations as described in Sec III. Let us consider a unitary transformation amongst the expanded set of vectors $\{c_i^{(0)}\}$:

$$c_i^{(1)} = u_{ii}\mathbf{c}_i^{(0)} + \sum_{j\neq i} u_{ij}\mathbf{c}_j^{(0)} \tag{10}$$

Under this transformation the change in the total energy correct to the second order in u_{ij}'s $(i\neq j)$ can be written as

$$\delta E = \delta E^c + 2\sum_{a,b}\delta A_a A_b^0 \delta H_{ab} + \sum_{a,b}\delta A_a \delta A_b (H_{ab} - \delta_{ab}E_0) \tag{11}$$

where $\{A_a^0\}$ are the mixing coefficients obtained with $\{c_i^0\}$ and $\{\delta A_a\}$, δH_{ab} are respectively the changes in $\{A_a\}$ and H_{ab} under the transformation (10), δE^C is the change due to the changes in the vectors alone and is given by

$$\delta E^C = 2\sum_{i>k} u_{ik}\Delta\varepsilon_{ik} + \sum_{i\neq k, j\neq l} u_{ik}u_{jl}(G_{ik,jl} + \delta_{ij}F_{ikl});$$

$$\Delta\varepsilon_{ik} = \underline{c}_k^{(0)^+}(\underline{F}_i^{(0)}\underline{c}_i^{(0)} + \underline{f}_i^{(0)}) - \underline{c}_i^{(0)}(\underline{F}_k^{(0)}\underline{c}_k^{(0)} + \underline{f}_k^{(0)}),$$

$$F_{ikl} = \underline{c}_k^{(0)^+}\underline{F}_i^{(0)}\underline{c}_l^{(0)} ,$$

$$G_{ik,jl} = \frac{\partial}{\partial u_{jl}}\left[\underline{c}_k^{(0)^+}\underline{F}_i\underline{c}_i^{(0)} + \underline{f}_i\right]$$

In the above it is understood that $\underline{F}_i = \underline{f}_i \equiv \underline{0}$ if $\{c_i\}$ is a virtual vector. Minimizing (11) with respect to the variation $\{\delta A_a\}$ subject to the constraint $\underline{\delta A}\cdot\underline{A}^{(0)} = 0$,

$$\delta E = \delta E^C - \underline{g}^+\underline{M}\underline{g} \tag{12}$$

where

$$\underline{g} = \delta\underline{H}\underline{A}^{(0)} - (\underline{A}^{(0)^+}\delta\underline{H}\underline{A}^{(0)})\underline{A}^{(0)}$$
$$\underline{M} = (\underline{H} - \varepsilon_0 I - \varepsilon_0|\underline{A}^{(0)}\rangle\langle\underline{A}^{(0)}|)^{-1} \tag{13}$$

Thus

$$\delta E = \sum_{i\neq k} u_{ik}\Delta\varepsilon_{ik} + \sum_{i\neq k, j\neq l} u_{ik}u_{jl}(G'_{ik,jl} + \delta_{ij}F_{ikl}) \tag{14}$$

where

$$G'_{ik,jl} = G_{ik,jl} - \underline{g}_{ik}^+\underline{M}\underline{g}_{jl} \tag{15}$$

$$g_{ik} = \frac{\partial g}{\partial u_{ik}} \tag{16}$$

The transformation coefficients are obtained by varying δE, as given by (12), with respect to u's:

$$\sum_{j\neq l} u_{jl}(G'_{ik,jl} + \delta_{ij}F_{ikl}) + \Delta\varepsilon_{ik} = 0 \tag{17}$$

VI. A SIMPLIFICATION FOR LARGE CONFIGURATION SETS

For large sets of configurations evaluation of M (Eq. 13) and g (Eq. 16) can be difficult. We present a simplification as follows: As in Sec. II we distinguish two subsets of a given set of configurations, one containing the predominant ones, the other containing the rest. We break up the corresponding wavefunction accordingly:

$$\Psi = \sum_{a=1}^{n_0} A_a\Phi_a + \sum_{a=n_0+1}^{N} B_a\Phi_a \tag{18}$$

where the first sum goes over the first subset. In the present optimization scheme during the course of an interation the mixing coefficients occurring in the second sum are kept frozen except for normalization. Thus we shall be concerned with the following energy expression:

$$E = \sum_{a,b=1}^{n_0} A_aA_bH_{ab} + 2A_0\Sigma A_aK_a + A_0^2E_0 \tag{19}$$

where

$$K_a = \langle\Phi_a|H|\Phi_0\rangle$$

$$\Phi_0 = a_0^{-1}\sum_{a=n_0+1}^{N} B_a\Phi_a \ , \ a_0 = \sqrt{\sum_a B_a^2}$$

$$E_0 = \langle\Phi_0|H|\Phi_0\rangle$$

In solving Eq.(17) we obtain M and g only for the *configurations of the first subset and for Φ_0.*

VII. COMMENTS ON THE SECOND ORDER TREATMENT OF THE MCSCF PROBLEM

In the above we have discussed an MCSCF procedure in which during every iteration we retain in the energy expression all corrections up to the second order coming from the variations of the orbitals. Barring some pathological cases (usually involving single excitations) where the third and higher order terms dominate over the first and second order terms, the process is expected to converge.

The convergence, however, is not guaranteed to correspond to the root desired. For that it is necessary to go step by step from the lowest root of the symmetry to the one under consideration.

The absence of large Hessian matrices in our theory makes each iterative step short enough such that it is easier and cheaper to monitor the course of the MCSCF iterations than what it would be for a full-blown orbital transformation method[4-6].

References

1. A. C. Wahl and G. Das in Modern Theoretical Chemistry, Vol. III, Chapter III (Edited by H. F. Schaefer III, Plenum Press, New York and London).

2. E. Dalgaard and P. Jorgensen, J. Chem. Phys. 68, 3833 (1979).

3. B. R. Brooks, W. D. Laidig, P. Saxe and H. F. Schaefer III, J. Chem. Phys. 72, 3837 (1980).

4. D. L. Yeager and P. Jørgensen, J. Chem. Phys. 71, 755 (1979).

5. E. Dalgaard, Chem. Phys. Lett. 65, 559 (1979).

6. C. C. J. Roothaan, J. Detrich and D. G. Hopper, Int. J. Quantum Chem. Symposia 13, 93 (1979).

7. G. Das, J. Chem. Phys. 46, 1568 (1967).

8. G. Das, J. Chem. Phys. 58, 5104 (1973).

Dup

GENERAL VALENCE BOND THEORY FOR CHEMICAL REACTIONS: FORMULATION OF THE SOGVB WAVEFUNCTION

Frank W. Bobrowicz
Monsanto Research Corporation
Mound Facility*
Miamisburg, Ohio 45342

SUMMARY

In this paper we present the formulation of an MCSCF wavefunction of the GVB (generalized valence - bond) type designed to deal with the changes occurring during chemical reactions. As such, this wavefunction which we have labeled SOGVB (strongly orthogonal generalized valence - bond) overcomes one of the deficiencies of the more familiar Perfect-Pairing GVB wavefunction (GVB-PP) while retaining much of the computational simplicity of the Perfect-Pairing approximation.

To understand the rationale behind the SOGVB approach[1] let us first consider the more well-known GVB-PP wavefunction.[2] In essence, this wavefunction results from a relaxation of the double-occupation restriction placed upon the restricted Hartree-Fock (RHF) wavefunction. That is, in GVB-PP, we replace an RHF doubly-occupied orbital by a pair of nonorthogonal singlet coupled orbitals. For example, the RHF wavefunction

$$\Psi = A\phi_1^2\phi_3\alpha\beta\alpha = \begin{array}{|c|}\hline \phi_1\phi_1 \\ \hline \phi_3 \\ \hline \end{array} \qquad (1)$$

becomes

$$\Psi = \frac{1}{\sqrt{2}}\, A(\phi'_{11}\phi'_{21} + \phi'_{21}\phi'_{11})\,\phi_3\,\alpha\beta\alpha = \begin{array}{|c|}\hline \phi'_{11}\phi'_{21} \\ \hline \phi_3 \\ \hline \end{array} \; . \qquad (2a)$$

However, while the orbitals in each such pair are allowed to be nonorthogonal we require that they be orthogonal to all other orbitals. Thus, in (2a) we impose the constraints

$$\langle\phi'_{11}|\phi_3\rangle = \langle\phi'_{21}|\phi_3\rangle = 0 \; . \qquad (2b)$$

If we now express each nonorthogonal pair in terms of the corresponding orthogonal natural orbitals such that

$$\phi'_{1i} = (\sqrt{\lambda_{1i}}\,\phi_{1i} + \sqrt{\lambda_{2i}}\,\phi_{2i})/(\lambda_{1i} + \lambda_{2i})^{1/2}$$

$$\phi'_{2i} = (\sqrt{\lambda_{1i}}\,\phi_{1i} - \sqrt{\lambda_{2i}}\,\phi_{2i})/(\lambda_{1i} + \lambda_{2i})^{1/2} \qquad (3)$$

we obtain an MC wavefunction involving only orthogonal orbitals. For example, (2a) then becomes (taking $\lambda_{1i}^2 + \lambda_{2i}^2 = 1$)

$$\Psi = A(\lambda_{11}\phi_{11}^2 - \lambda_{21}\phi_{21}^2)\phi_3\alpha\beta\alpha$$

$$= \lambda_{11} \begin{array}{|c|}\hline \phi_{11}\phi_{11} \\ \hline \phi_3 \\ \hline \end{array} - \lambda_{21} \begin{array}{|c|}\hline \phi_{21}\phi_{21} \\ \hline \phi_3 \\ \hline \end{array} \; . \qquad (4)$$

Since the GVB-PP wavefunction can be written as an MC wavefunction involving orthogonal orbitals in which the configurations differ by at least a doubly-occupied orbital the energy expression has a form identical to that of an open-shell HF wavefunction[3]

$$E = 2\sum_i f_i\langle\phi_i|h|\phi_i\rangle + \sum_{i,j}[a_{i,j}J_{i,j} + b_{i,j}K_{i,j}] \; . \qquad (5)$$

As a consequence, solving for a GVB-PP wavefunction is not much more difficult than solving for an RHF wavefunction. The essential difference is that the {f,a,b} energy coefficients involving paired orbitals are functions of the pair coefficients {λ} which must also be solved for self-consistentcy. In addition, there are more Fock operators to deal with since each paired orbital has a unique one of its own.[3]

Because GVB-PP uses two nonorthogonal orbitals in describing a typical chemical bond it serves as an excellent wavefunction for near-equilibrium molecular geometries. In addition, it is capable of describing rudimentary processes such as the formation of H_2 from two hydrogen atoms. However, in general it is not flexible enough to describe typical chemical reactions. For example, consider the formation of CH $^2\Pi$ from C ^3P and H ^1S. In the separated atom limit the two carbon p-orbitals are triplet couplet so that

$$\psi_\infty = \frac{1}{\sqrt{6}}\, A \cdots \phi_{PZ}\phi_{PY}\phi_H\,(2\alpha\alpha\beta-\alpha\beta\alpha-\beta\alpha\alpha) = \begin{array}{|c|c|}\hline \multicolumn{2}{|c|}{\cdots} \\ \hline \phi_{PZ} & \phi_H \\ \cline{2-2} \phi_{PY} \\ \cline{1-1} \end{array} \qquad (6a)$$

---

*Mound Facility, Miamisburg, Ohio, is operated by Monsanto Research Corporation, a subsidiary of Monsanto Company for the U. S. Department of Energy under Contract No. DE-AC04-76-DP00053.

(where . . . represents all other orbitals) whereas in the product a carbon p-orbital and the hydrogen orbital must become singlet coupled in order to describe the C-H so that

$$\psi_{eq} = \begin{array}{|c|c|}\hline \cdot\cdot\cdot & \\ \hline \phi_{PZ} & \phi_{H} \\ \hline \phi_{PY} & \\ \cline{1-1}\end{array} \quad . \tag{6b}$$

Clearly, regardless of how the orbitals change, GVB-PP wavefunction (6b) which gives a good description of CH $^2\Pi$ at equilibrium can never describe the separated atom limit. Conversely, (6a) which describes this limit cannot possibly describe a strong C-H bond (Fig. 1). The only way to properly describe this reaction is to employ a wavefunction which incorporates both of these coupling schemes, such as

$$\Psi = \xi_{eq} \begin{array}{|c|c|}\hline \cdot\cdot\cdot & \\ \hline \phi_{PZ} & \phi_{H} \\ \hline \phi_{PY} & \\ \cline{1-1}\end{array} + \xi_{\infty} \begin{array}{|c|c|}\hline \cdot\cdot\cdot & \\ \hline \phi_{PZ} & \phi_{H} \\ \hline \phi_{PY} & \\ \cline{1-1}\end{array} \tag{6c}$$

in which we must solve not only for the optimal orbitals but for the optimal coupling coefficients $\{\xi\}$ as well.

The inability of the GVB-PP wavefunction to deal with chemical reactions goes beyond the obvious example given above. It also extends to cases where the coupling between product orbitals is the same as that between reactant orbitals. An example of this is seen in the formation of CH $^2\Delta$ from C ^1D and H ^1S. Since this state arises from the formation of a bond between a hydrogen orbital and one of the carbon sp-pair orbitals (the GVB picture of carbon has two sp-orbitals in place of the RHF $2s^2$ orbital) the product wavefunction is

$$\psi_{eq} = \begin{array}{|c|c|}\hline \cdot\cdot\cdot & \\ \hline \phi_{SZ} & \phi_{H} \\ \hline \phi_{SZ} & \\ \cline{1-1}\end{array} \tag{7a}$$

whereas in the separated-atom wavefunction the hybrid sp-orbitals in C ^1D are singlet coupled so that

$$\psi_{\infty} = \begin{array}{|c|c|}\hline \cdot\cdot\cdot & \\ \hline \phi_{SZ} & \phi_{\overline{SZ}} \\ \hline \phi_{H} & \\ \cline{1-1}\end{array} \quad . \tag{7b}$$

Since in either limit the system can be described by a perfect-pairing coupling scheme it might be expected that this reaction could be studied using a GVB-PP wavefunction. However, when GVB-PP calculations are performed we find that this is not the case. Rather, if we start at the separate-atom limit and move inward we climb a repulsive curve which does not lead to the bound molecular state (Fig. 2). Likewise, if we start with the molecular state and move outward we follow a potential curve which does not tend toward the proper atomic limit. Therefore, while the GVB-PP approximation is adequate at either limit it is quite inappropriate at intermediate internuclear separations. In order to obtain a smoothly varying adiabatic description of this reaction a more flexible wavefunction is clearly required.

Relaxing the constraints imposed upon the GVB-PP wavefunction while still remaining within the framework of an independent-particle model leads to the unrestricted GVB wavefunction. In this approximation all orbitals are allowed to be nonorthogonal and, while retaining the desired overall spin multiplicity, these orbitals are coupled in a fully general manner.[4] For example, the three-electron doublet GVB wavefunction can be written as

$$\Psi = \xi_1 \begin{array}{|c|c|}\hline \phi_1 & \phi_2 \\ \hline \phi_3 & \\ \cline{1-1}\end{array} + \xi_2 \begin{array}{|c|c|}\hline \phi_1 & \phi_3 \\ \hline \phi_2 & \\ \cline{1-1}\end{array} \tag{8}$$

Unfortunately, because of the computational complexities involved, unrestricted GVB calculations are pratical only for small systems. Furthermore, extending this method through inclusion of closed-shell and/or perfect-pair orbitals is nontrivial. However, GVB calculations on small but representative systems have been performed and from the results it becomes apparent that the full generality of this wavefunction is usually not requried. To illustrate, consider the simple colinear $H_2 + D \rightarrow H + HD$ reaction. Initially the system consists of a hydrogen molecule and a deuterium atom whereas in the products we have an HD molecule and a hydrogen atom. Since the reactant and product wavefunctions are

$$\psi_R = \begin{array}{|c|c|}\hline \phi_{H} & \phi_{H'} \\ \hline \phi_{D} & \\ \cline{1-1}\end{array} \quad \text{and} \quad \psi_P = \begin{array}{|c|c|}\hline \phi_{H} & \phi_{D} \\ \hline \phi_{H'} & \\ \cline{1-1}\end{array} \tag{9a}$$

the GVB wavefunction for this reaction can be written as

$$\Psi = \xi_1 \begin{array}{|c|c|}\hline \phi_1 & \phi_2 \\ \hline \phi_3 & \\ \cline{1-1}\end{array} + \xi_2 \begin{array}{|c|c|}\hline \phi_1 & \phi_3 \\ \hline \phi_2 & \\ \cline{1-1}\end{array} \quad . \tag{9b}$$

From plots of the orbital changes involved in the GVB description of this reaction (Fig. 3)[5] we see that as the reaction proceeds, orbital ϕ_2, which is initially associated with the H' nucleus, slowly delocalizes in a symmetric manner over to the deuterium nucleus and finally relocalizes there. Simultaneously, orbital ϕ_3, which is initially centered on the deuterium, delocalizes in an antisymmetric manner over to the H' center and finally relocalizes there. The result is that orbitals ϕ_1 and ϕ_2 always remain highly overlapping while ϕ_3 remains nearly orthogonal to both of them. Thus, these orbitals remain essentially strongly orthogonal at all times. Hoever, the coupling between them does not remain fixed. In fact, the coupling changes drastically and, up through the saddle point, closely resembles what one would expect were the reaction to proceed via a purely localized orbital mechanism (Fig. 4).

As the above examples have served to illustrate,[6] in order to describe a typical chemical reaction it is necessary to allow the orbitals involved in bond breaking/forming processes to couple with one another in a completely general manner. However, it is usually not necessary to relax all orthogonality constraints since the orbitals tend to remain strongly orthogonal anyway. Therefore, in the SOGVB

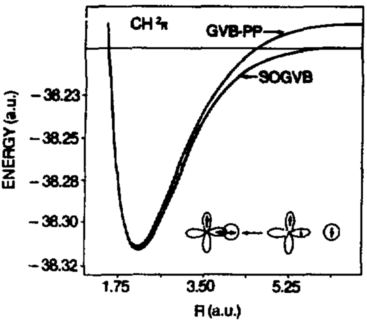


Fig. 1.

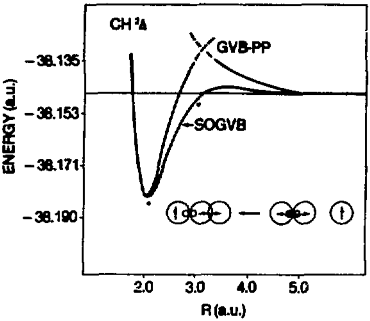


Fig. 2.

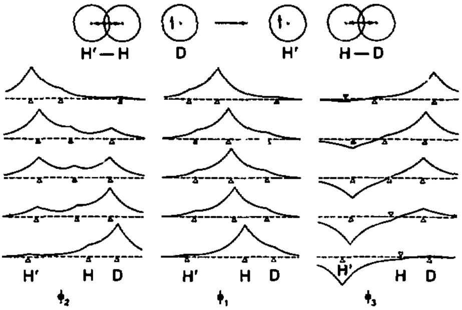


Fig. 3.

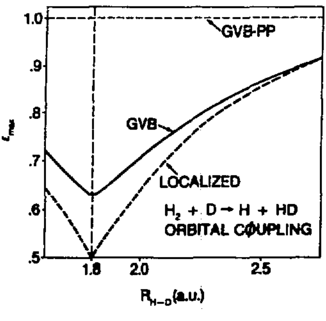


Fig. 4.

approximation we allow the orbitals to couple in a general manner, but nonetheless group them into strongly orthogonal pairs. For example, the three-electron doublet SOGVB wavefunction has the form of the unrestricted GVB wavefunction but involves orbitals constrained to be strongly orthogonal

$$\Psi = \xi_1 \begin{array}{|c|c|}\hline \phi'_{11} & \phi'_{21} \\ \hline \phi_3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + \xi_2 \begin{array}{|c|c|}\hline \phi'_{11} & \phi_3 \\ \hline \phi'_{21} & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \tag{10a}$$

$$= \begin{array}{|c|c|}\hline \phi'_{11} & \phi'_{21} \\ \hline \phi_3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array}$$

$$\langle \phi'_{11} | \phi_3 \rangle = \langle \phi'_{21} | \phi_3 \rangle = 0. \tag{10b}$$

As with the GVB-PP wavefunction, it is computationally convenient to write the SOGVB wavefunction in terms of orthogonal orbitals by expressing each nonorthogonal pair in terms of orthogonal natural orbitals. This leads to an MC wavefunction in which each orthogonal configuration either contains only one orbital of a pair with that orbital being doubly-occupied or it contains both orbitals of a pair triplet coupled. For example, substituting (3) into (10a) gives

$$\Psi = A_1 \begin{array}{|c|c|}\hline \phi_{11} & \phi_{11} \\ \hline \phi_3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + A_2 \begin{array}{|c|c|}\hline \phi_{21} & \phi_{21} \\ \hline \phi_3 & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} + A_3 \begin{array}{|c|c|}\hline \phi_{11} & \phi_3 \\ \hline \phi_{21} & \multicolumn{1}{c}{} \\ \cline{1-1}\end{array} \tag{11a}$$

where

$$A_1 = \sqrt{2}\, \xi_1 \lambda_{11};$$

$$A_2 = \sqrt{2}\, \xi_1 \lambda_{21}; \quad A_3 = 2\sqrt{\lambda_{11}\lambda_{21}}\ \xi_2 \tag{11b}$$

and the two-pair four-electron singlet wavefunction is

$$\Psi = A_1 \begin{array}{|c|}\hline \phi_{11}\phi_{11} \\ \hline \phi_{12}\phi_{21} \\ \hline\end{array} + A_2 \begin{array}{|c|}\hline \phi_{11}\phi_{11} \\ \hline \phi_{22}\phi_{22} \\ \hline\end{array} +$$

$$A_3 \begin{array}{|c|}\hline \phi_{21}\phi_{21} \\ \hline \phi_{12}\phi_{12} \\ \hline\end{array} + A_4 \begin{array}{|c|}\hline \phi_{21}\phi_{21} \\ \hline \phi_{22}\phi_{22} \\ \hline\end{array} + A_5 \begin{array}{|c|c|}\hline \phi_{11} & \phi_{21} \\ \phi_{21} & \phi_{22} \\ \hline\end{array} \tag{12}$$

subject to the constraint $A_1A_3 = A_2A_4$. As a consequence, the SOGVB energy expression for NG orbitals involving NGP pairs is

$$E = 2\sum_i^{NG} f_i h_{i,i} + \sum_{i,j}^{NG} [a_{i,j}J_{i,j} + b_{i,j}K_{i,j}]$$

$$+ 4\sum_m^{NGP} \{\sum_{i \neq 1m,2m}^{NG} c_{mi} \langle \phi_{1m} | K_i | \phi_{2m} \rangle$$

$$+ \sum_{n>m}^{NGP} [d_{nm}^{(1)} (\phi_{1m}\phi_{1n} | \phi_{2n}\phi_{2m})$$

$$+ d_{nm}^{(2)} (\phi_{1m}\phi_{2n} | \phi_{1n}\phi_{2m})]\} \tag{13a}$$

where

$$h_{i,i} = \langle \phi_i | h | \phi_i \rangle = \langle \phi_i | T+V | \phi_i \rangle$$

$$J_{i,j} = (\phi_i\phi_i | \phi_j\phi_j)$$

$$= \langle \phi_i(1)\phi_j(2) | r_{12}^{-1} | \phi_i(1)\phi_j(2) \rangle \tag{13b}$$

$$K_{i,j} = (\phi_i\phi_j | \phi_j\phi_i)$$

$$= \langle \phi_i(1)\phi_j(2) | r_{12}^{-1} | \phi_j(1)\phi_i(2) \rangle$$

and where the coefficients $\{f,a,b,c,d\}$ are functions of the configuration coefficients $\{A\}$ which in turn are functions of the pair and coupling coefficients $\{\lambda, \xi\}$.

Since (13a) contains only diagonal one-electron terms, inclusion of closed-shell, perfect-pairing, and open-shell multiplet orbitals in the wavefunction is straightforward. This is of considerable importance since in most chemical reactions only a few orbitals are actually involved in the bonding process. However, nonparticipating orbitals must also be taken into account since changes in these orbitals can be important. To treat all orbitals at a generalized coupling (GC) level is unnecessary and computationally impractical. By allowing the SOGVB wavefunction to explicitly include these less correlated orbitals we arrive at a wavefunction in which each group of orbitals is treated at an appropriate level of correlation and with an appropriate amount of computational effort. Thus, generalizing (13a) to include NS closed-shell and perfect-pair orbitals and NM multiplet-shell orbitals the energy expression for all N orbitals becomes

$$E = 2\sum_{i}^{N} f_i h_{i,i} + \sum_{i,j}^{N} [a_{i,j} J_{i,j} + b_{i,j} K_{i,j}]$$

$$+ 4\sum_{m}^{NGP} \{\sum_{i\neq 1m,2m}^{NG+NM} c_{mi} \langle\phi_{1m}|K_i|\phi_{2m}\rangle$$

$$+ \sum_{n>m}^{NGP} [d_{nm}^{(1)} (\phi_{1m}\phi_{1n}|\phi_{2n}\phi_{2m})$$

$$+ d_{nm}^{(2)} (\phi_{1m}\phi_{2n}|\phi_{1n}\phi_{2m})]\} \quad . \qquad (14)$$

Having arrived at the general energy expression for the SOGVB wavefunction we can now consider the task of developing the equations for its self-consistent solution. For the present, let us ignore the problem of optimizing the pair and coupling coefficients and concentrate on orbital optimization. From the Variational Principle we know that the orbitals will be self-consistent when the energy is stationary through first-order for any changes in the orbitals provided these changes preserve orbital orthogonality at least through first-order. Allowing the orbitals in (14) to vary according to the prescription

$$\phi_i \Rightarrow \phi_i + \delta_i \quad ;$$

$$[\langle\delta_i|\phi_j\rangle + \langle\delta_j|\phi_i\rangle + \langle\delta_i|\delta_j\rangle]^{(1)} = 0 \qquad (15a)$$

this variational condition is found to be

$$\delta E^{(1)} = \sum_{i}^{N} [\langle\delta_i|F_i|\phi_i\rangle + \langle\delta_i|R_i\rangle] = 0 \qquad (15b)$$

where

$$F_i = f_i h + \sum_{j}^{N} [a_{j,i} J_j + b_{j,i} K_j]$$

$$+ \sum_{\substack{m \\ 1m,2m\neq i}}^{NGP} c_{mi} [(\_\phi_{1m}|\phi_{2m}\_) + (\_\phi_{2m}|\phi_{1m}\_)] \qquad (15c)$$

and where we define the exchange-like operator $(\_\phi_i|\phi_j\_)$ by

$$\langle\phi_i|(\_\phi_j|\phi_k\_)|\phi_\ell\rangle = (\phi_i\phi_j|\phi_k\phi_\ell) \quad . \qquad (15d)$$

If orbital i is not a GC pair orbital then $R_i = 0$. If however it corresponds to pair orbital ϕ_{km} then

$$R_{km} = \left(\sum_{j\neq 1m,2m}^{NG+NM} c_{mj} K_j + \sum_{n\neq m}^{NGP} [d_{nm}^{(k)} (\_\phi_{1n}|\phi_{2n}\_) \right.$$

$$\left. + d_{nm}^{(3-k)} (\_\phi_{2n}|\phi_{1n}\_)] \right) |\phi_{(3-k),m}\rangle . \qquad (15e)$$

If we define the operator R_i' by

$$R_i' = R_i \langle\phi_i \;| \qquad (16a)$$

then we can write (15b) in the more usual form

$$0 = \sum_{i}^{N} \langle\delta_i|F_i'|\phi_i\rangle \qquad (16b)$$

where

$$F_i' = F_i + R_i' \qquad (16c)$$

However, we prefer to work with the quantities defined in (15) since these arise directly from (14) whereas those of (16) do not.

Now let us consider the problem of optimizing any one orbital while keeping all others fixed. If only orbital ϕ_i is varied then (15b) becomes

$$0 = \langle\delta_i|F_i|\phi_i\rangle + \langle\delta_i|R_i\rangle \qquad (17)$$

However, this is subject to the constraint that $\langle\delta_i|\phi_j\rangle = 0$ for all ϕ_j. Therefore, if only one orbital is varied it can only be changed with respect to the space orthogonal to all orbitals. Keeping this in mind, if we isolate from (14) all terms involving ϕ_i and replace this orbital with the improved orbital

$$\phi_i^1 = (\phi_i + \Delta\phi_i)/(1 + \Delta_i^2)^{1/2} \quad ;$$

$$\Delta_i^2 = \langle\Delta\phi_i|\Delta\phi_i\rangle \qquad (18a)$$

upon expansion through second-order in $\Delta\phi_i$ we obtain

$$\varepsilon_i(1 + \Delta_i^2) = <\phi_i|F_i|\phi_i> + <\phi_i|R_i>$$

$$+ 2[<\Delta\phi_i|F_i|\phi_i> + <\Delta\phi_i|R_i>]$$

$$+ <\Delta\phi_i|F_i + 2a_{i,i}K_i|\Delta\phi_i> \qquad (18b)$$

where for simplicity we have taken $b_{i,i} = 0$. Expressing $\Delta\phi_i$ in terms of the basis space $\{\chi\}$ orthogonal to all orbitals

$$\Delta\phi_i = \sum_k C_{ki}\, \chi_k \qquad (18c)$$

where

$$<\chi_k|\phi_j> = 0; \quad <\chi_k|\chi_\ell> = \delta_{k,\ell} \qquad (18d)$$

the solution through second-order for the improved orbital in the field of all other ones is obtained by diagonalizing the matrix $\underline{X}$ where

$$X_{0,0} = <\phi_i|F_i|\phi_i> + <\phi_i|R_i>$$

$$X_{0,k} = X_{k,0} = <\chi_k|F_i|\phi_i> + <\chi_k|R_i> \qquad (18e)$$

$$X_{k,\ell} = <\chi_k|F_i|\chi_\ell> + 2a_{i,i} <\chi_k|K_i|\chi_\ell> ; \quad k,\ell \neq 0$$

and choosing the desired root (usually on the basis of either lowest eigenvalue or least change).

If ϕ_i is a closed-shell orbital, $R_i = 0$. Furthermore, if we neglect the second-order self-correction term $2a_{i,i} <\chi_k|K_i|\chi_\ell>$, the $\underline{X}$ matrix simply becomes a matrix over the Fock operator F_i. Since the Fock operators for all closed-shell orbitals can be made identical, solutions for all such orbitals could then be obtained through a single diagonalization. Since this offers considerable computational advantage and since experience has shown that these terms are relatively unimportant, we normally follow this procedure. For open-shell multiplet orbitals these terms do not appear in the first place. For all other orbitals the Fock operators are different and no real advantage is gained in neglecting these terms.

Examining the terms involved in (18) we see that for closed-shell and perfect-pair orbitals these equations involve only the same operator matrices that are required for a GVB-PP calculation. For the remaining orbitals the only additional matrices required are those over the GC-pair operators $(\_\phi_{1n}|\phi_{2n}\_)$. Therefore, only a little more computational effort is required here than would be the case for the corresponding GVB-PP calculation.

Equations (18) provide us with the prescription for iteratively optimizing the orbitals with respect to unoccupied space. There now remains the problem of optimizing them with respect to one another. Since the orbitals must always remain orthogonal the only way to do this is to vary at least two of them simultaneously. If orbitals ϕ_i and ϕ_j are simultaneously varied while keeping all others fixed, (15b) becomes

$$0 = <\delta_i|F_i|\phi_i> + <\delta_i|R_i>$$

$$+ <\delta_j|F_j|\phi_j> + <\delta_j|R_j> \qquad (19a)$$

subject to the condition

$$[<\delta_i|\phi_j> + <\delta_j|\phi_i> + <\delta_i|\delta_j>]^{(1)} = 0. \qquad (19b)$$

through first-order. To obtain the equation for the optimal mixing between these two orbitals we isolate all terms involving them from (14) and substitute the improved orbitals

$$\phi_i^1 = (\phi_i + \Delta\phi_i)/(1 + \Delta_i^2)^{1/2}$$

$$\phi_j^1 = (\phi_j + \Delta\phi_j)/(1 + \Delta_j^2)^{1/2} \qquad (20a)$$

Realizing that if

$$\Delta\phi_i = \gamma_{ij}\, \phi_j \qquad (20b)$$

then we must have

$$\Delta\phi_j = - \gamma_{ij}\, \phi_i \qquad (20c)$$

we expand through second-order in γ_{ij} to obtain

$$\varepsilon_{ij}(1 + \gamma_{ij}^2)$$

$$= <\phi_i|F_i|\phi_i> + <\phi_i|R_i> + <\phi_j|F_j|\phi_j> + <\phi_j|R_j>$$

$$+ \gamma_{ij}^2\, [<\phi_i|F_j|\phi_i> + <\phi_j|F_i|\phi_j> + Q_{ij}]$$

$$+ 2\gamma_{ij}[<\phi_i|F_i-F_j|\phi_j> + <\phi_j|R_i> - <\phi_i|R_j>] \qquad (20d)$$

where the second-order correction term Q_{ij} is given by

$$\gamma_{ij}Q_{ij} = <\phi_i|\Delta F_i-\Delta F_j|\phi_j> + <\phi_j|\Delta R_i> - <\phi_i|\Delta R_j> \qquad (21a)$$

and where ΔF and ΔR are the first-order changes in the corresponding operators. For example, if either orbital is a closed-shell or perfect-pair orbital or if neither is a GC pair orbital

$$Q_{ij} = Q_{ij}^0 = 2(a_{i,i} + a_{j,j} - 2a_{i,j})K_{i,j}$$

$$+ (b_{i,i} + b_{j,j} - 2b_{i,j})\, [J_{i,j} + K_{i,j}] \qquad (21b)$$

whereas if orbital ϕ_i is in GC pair m and orbital ϕ_j is a multiplet or unpaired GC orbital

$$Q_{ij} = Q^{0}_{ij} - 2c_{mj} <\phi_{1m}|J_j + K_j|\phi_{2m}> . \qquad (21c)$$

Letting

$$C^{(1)}_{ij} = <\phi_i|F_i - F_j|\phi_j> - <\phi_j|R_i> - <\phi_i|R_j>$$

$$C^{(2)}_{ij} = <\phi_j|F_i - F_j|\phi_j> - <\phi_i|F_i - F_j|\phi_i>$$

$$- <\phi_i|R_i> - <\phi_j|R_j> + Q_{ij} \qquad (22a)$$

mixing coefficient γ_{ij} is obtained by solving

$$0 = C^{(1)}_{ij} [\gamma^2_{ij} - 1] - C^{(2)}_{ij} \gamma_{ij} \qquad (22b)$$

and choosing the appropriate root.

In principle, the orbitals can be iteratively optimized with respect to one another using (22). However, to redefine all affected matrix elements after each such orbital-pair rotation is computationally unacceptable and experience has shown that by not doing so the overall orbital changes tend to be too large. To see if there is a simple way to overcome this problem let us now consider simultaneously optimizing all orbitals with respect to one another by taking the improved orbitals to be

$$\phi^1_i = N_i [\bar{\phi}_i - \sum_{k<i} <\bar{\phi}_i|\phi^1_k> \phi^1_k] \qquad (23a)$$

with

$$\bar{\phi}_i = \phi_i + \sum_{j \neq i} \gamma_{ji} \phi_j \quad ; \quad \gamma_{ij} = - \gamma_{ji} . \qquad (23b)$$

Upon substituting these orbitals into (14) and expanding through second-order in $\{\gamma_{ij};\ i>j\}$ the result is a rather complicated expression involving many two-electron integrals not required in either (22) or (18). All such integrals, however, only appear in terms involving the product of two different mixing coefficients. If we now content ourselves with a pairwise mixing scheme by neglecting all such mixed coefficient terms, we find that the mixing coefficients can be determined by diagonalizing the matrix $\underline{B}$ where

$$B_{0,0} = 0$$

$$B_{0,ij} = B_{ij,0} = C^{(1)}_{ij}$$

$$B_{ij,kl} = 0 \text{ for } ij \neq kl \qquad (23c)$$

$$B_{ij,ij} = C^{(2)}_{ij}$$

and where $C^{(1)}_{ij}$ and $D^{(2)}_{ij}$ are given by (22a). While diagonalization of this matrix and choosing the appropriate root still constitute only a pairwise optimization scheme, it has been found to be superior to simple sequential usage of (22) since each mixing coefficient is now weighted relative to its importance in lowering the total energy.

Equations (18) and (23) provide us with the prescriptions for optimizing the orbitals to self-consistency for fixed pair and coupling coefficients. We will now turn our attention to determining these coefficients. First of all, let us consider any perfect-pair coefficients. If orbitals ϕ_{1k} and ϕ_{2k} are perfectly paired, the coefficients in (14) which depend upon pair coefficients λ_{1k} and λ_{2k} are

$$f_{ik} = a_{ik,ik} = \lambda^2_{ik}/(\lambda^2_{1k} + \lambda^2_{2k}) \quad ; \quad i = 1,2$$

$$b_{1k,2k} = b_{2k,1k} = - \lambda_{1k} \lambda_{2k}/(\lambda^2_{1k} + \lambda^2_{2k})$$

$$a_{ik,j} = a_{j,ik} = 2f_{ik}f_j \quad ;$$

$$i = 1,2 \quad ; \quad \phi_j \neq \phi_{1k},\phi_{2k}$$

$$b_{ik,j} = b_{j,ik} = -f_{ik}f_j \quad ; \qquad (24a)$$

$$i = 1,2 \quad ; \quad \phi_j \neq \phi_{1k},\phi_{2k} .$$

Isolating these terms in (14) gives

$$E_k (\lambda^2_{1k} + \lambda^2_{2k}) = \lambda^2_{1k} H_{1k,1k} + \lambda^2_{2k} H_{2k,2k}$$
$$+ 2\lambda_{1k} \lambda_{2k} H_{1k,2k} \qquad (24b)$$

where

$$H_{ik,ik} = 2 h_{ik,ik} + J_{ik,ik}$$
$$+ 2\sum^{N}_{j \neq 1k,2k} f_j (2J_{ik,j} - K_{ik,j})$$

$$H_{1k,2k} = H_{2k,1k} = - \frac{1}{2} K_{1k,2k} . \qquad (24c)$$

Therefore, we solve for λ_{1k} and λ_{2k} by diagonalizing this 2X2 $\underline{H}$ matrix and choosing the root which minimizes E_k. If there is more than one perfect pair involved, the pair coefficients for each pair are optimized iteratively through sequential application of (24) until self-consistency is achieved.

We are now left with the task of optimizing the GC pair and coupling coefficients which, in turn, define the $\{f,a,b,c,d\}$ energy coefficients involving GC orbitals. In general, the relationship between these coefficients is not straightforward and is most easily established by first expressing the wavefunction (ignoring all but GC orbitals) as

$$\Psi = \sum_i A_i \psi_i \qquad (25)$$

where the configuration coefficients $\{A\}$ are known functions of the GC pair and coupling coefficients $\{\lambda,\xi\}$. Determining the energy expression for Ψ and rearranging it to the form of (13a) then establishes the relationship between $\{A\}$ and the coefficients in (13a).

In general, the coefficients $\{A\}$ in (25) are not all linearly independent. Therefore, they cannot be determined simply by diagonalizing the hamiltonian over the configurations involved. To do this we recast the wavefunction into the general form

$$\Psi = A\prod_{i}^{NGP} [\lambda_{1i}\psi_{1i}^{2} - \lambda_{2i}\phi_{2i}^{2} + \sqrt{\lambda_{1i}\lambda_{2i}}\,(\phi_{1i}\phi_{2i}-\phi_{2i}\phi_{1i})]\,\phi_m\cdots\phi_n \times (\sum_j \xi_j\, X_j) \tag{26}$$

where $\{X\}$ is a set of appropriate orthogonal spin eigenfunctions. Concentrating on pair k, (26) can be written as

$$\Psi = (\lambda_{1k}\Phi_{1k,1k} - \lambda_{2k}\Phi_{2k,2k}) \sum_u \xi\,(\underline{k})_u\, X\,(\underline{k})_u + 2\sqrt{\lambda_{1k}\lambda_{2k}}\;\Phi_{1k,2k}\sum_\nu \xi\,(k|)_\nu\, X\,(k|)_\nu \tag{27a}$$

where

$$\Phi_{ik,jk} = A\,\phi_{ik}\phi_{jk}\prod_{\ell\neq k}^{NGP} [\lambda_{1\ell}\phi_{1\ell}^{2} - \lambda_{2\ell}^{2}\phi_{2\ell} + \sqrt{\lambda_{1\ell}\lambda_{2\ell}}\;(\phi_{1\ell}\phi_{2\ell} - \phi_{2\ell}\phi_{1\ell})] \times \phi_m\cdots\phi_n \tag{27b}$$

and where $X\,(\underline{k})_u$ $[X\,(k|)_u]$ is the uth spin eigenfunction in which the electronic positions of pair k are singlet (triplet) coupled. If we now let

$$\psi_{k,\nu}^{(i)} = \Phi_{ik,ik}\, X\,(\underline{k})_\nu \quad ; \quad i = 1,2$$

$$\psi_{k,\nu}^{(3)} = \Phi_{1k,2k}\, X\,(k|)_\nu$$

$$\sigma_{k,\nu} = 2\sqrt{\lambda_{1k}\lambda_{2k}}\;\xi\,(k|)_\nu \tag{27c}$$

we can solve for $\lambda_{1k},\lambda_{2k}$ and $\{\xi(k|)\}$ by taking

$$\Psi = \lambda_{1k}\,[\sum_u \xi(\underline{k})_u\,\psi_{k,u}^{(1)}] + \lambda_{2k}\,[-\sum_u \xi(\underline{k})_u\,\psi_{k,u}^{(2)}] + \sum_\nu \sigma_{k,\nu}\,\psi_{k,\nu}^{(3)} \tag{27d}$$

and diagonalizing the hamiltonian matrix over the configurations involved. If we then rewrite (27d) as

$$\Psi = \sum_u \xi(\underline{k})_u\,[\lambda_{1k}\psi_{k,u}^{(1)} - \lambda_{2k}\psi_{k,u}^{(2)}] + \sum_\nu \sigma_{k,\nu}\,\psi_{k,\nu}^{(3)} \tag{27e}$$

we can determine $\{\xi\,(\underline{k}),\ \xi\,(k|)\}$ by diagonalizing the hamiltonian matrix over the indicated configurations. Therefore, by constructing the hamiltonian matrix over the basic configurations of (25) and then by performing a series of contractions and diagonalizations the optimum GC pair and coupling coefficients can be iteratively obtained. In the course of evaluating the required matrix elements the presence of any closed-shell or perfect-pair orbitals in the total wavefunction is taken into account by using the modified one-electron operator

$$h' = h + \sum_i^{NS} f_i\,(2J_i - K_i) \tag{28a}$$

and the presence of any multiplet-shell orbitals is taken into account by formalizing the matrix element expressions for NG + 1 electrons. This leads to terms involving the multiplet-shell exchange operator

$$K_M = \sum_i^{NM} K_i \tag{28b}$$

Early experience has shown that because of the high degree of correlation existing between the GC orbitals, it is quite advantageous to fully optimize the GC orbitals with respect to one another during each iteration of the SCF cycle. Therefore, we usually ignore terms in (23) which mix the GC orbitals and perform this mixing during the GC pair and coupling coefficient optimization stage. To do this we define the improved orbitals as in (23a, 23b) but with the summations running only over GC orbitals. Upon substituting these orbitals in (25) and expanding through second-order we obtain

$$\Psi^{1} = \Psi + \sum_{ij}\gamma_{ij}\,\psi_{ij} + \sum_{ij,k\ell}\gamma_{ij}\,\gamma_{k\ell}\,\psi_{ij,k\ell}\;. \tag{29a}$$

If we then define the matrices $\underline{H}$ and $\underline{S}$ by

$$H_{0,0} = \langle\Psi|H|\Psi\rangle$$

$$H_{ij,0} = H_{0,ij} = \langle\Psi|H|\psi_{ij}\rangle$$

$$H_{ij,k\ell} = \langle\psi_{ij}|H|\psi_{k\ell}\rangle + \langle\Psi|H|\psi_{ij,k\ell}\rangle\,(1 + \delta_{ij,k\ell})$$

$$S_{0,0} = \langle\Psi|\Psi\rangle$$

$$S_{ij,0} = S_{0,ij} = \langle\Psi|\psi_{ij}\rangle$$

$$S_{kj,k\ell} = \langle\psi_{ij}|H|\psi_{k\ell}\rangle + \langle\Psi|\psi_{ij,k\ell}\rangle\,(1 + \delta_{ij,k\ell}) \,, \tag{29b}$$

the mixing coefficients $\{\gamma_{ij};\ i>j\}$ can be obtained through second-order in the energy by solving the secular equation

$$(\underline{H} - \underline{E}\,\underline{S})\,\underline{\gamma} = 0. \tag{29c}$$

Once the improved orbitals have been found we redefine all required hamiltonian matrices for them and once again solve (27) for new pair and coupling coefficients. This process is continued until self-consistency within the current GC space is achieved.

To summarize, the SCF cycle which we employ to solve for an SOGVB wavefunction consists of the following distinct steps:

1) Optimization of all GC pair and coupling coefficients using equations (27);

2) Optimization of the GC orbitals with respect to one another using equations (29);

3) Optimization of all perfect-pair coefficients using equations (29);

4) Optimization of orbitals with respect to one another (excluding mixing between GC orbitals) using equations (23); and

5) Optimization of all orbitals with respect to unoccupied space using equations (18).

Steps 1) and 2) are performed together until self-consistent. Step 3) is continued until all perfect-pair coefficients are self-consistent. Steps 4) and 5) are performed only once per SCF iteration (i.e., they are not continued until self-consistent).

REFERENCES

1. F. W. Bobrowicz, Ph.D. Thesis, California Institute of Technology, 1974.

2. W. J. Hunt, P. J. Hay, and W. A. Goddard III, J. Chem. Phys. 57, 738 (1972).

3. F. W. Bobrowicz and W. A. Goddard III, Modern Theoretical Chemistry, vol. III, Ed. H. F. Schaefer III. New York, Plenum Press, 1977.

4. R. C. Ladner and W. A. Goddard III, J. Chem. Phys 51, 1073 (1969).

5. W. A. Goddard and R. C. Ladner, J. Amer. Chem. Soc. 93, 6750 (1971).

6. For more examples
   a) B. J. Moss, F. W. Bobrowicz, and W. A. Goddard III, J. Chem. Phys. 63, 4632 (1975).
   b) P. Baybutt, F. W. Bobrowicz, L. R. Kahn, and D. G. Truhlar, J. Chem. Phys. 68, 4809 (1978).

THE LOOP-DRIVEN GRAPHICAL UNITARY GROUP APPROACH APPLIED TO THE MCSCF PROBLEM

William D. Laidig and Henry F. Schaefer III
Department of Chemistry and Lawrence Berkeley Laboratory,
University of California
Berkeley, California 94720

and

Department of Chemistry and Institute for Theoretical Chemistry
University of Texas
Austin, Texas 78712

LDGUGA Applied to the MCSCF Problem

The unitary group approach (UGA) to the CI problem has proven over the last few years to be an efficient and effective method for estimating the correlation energy in molecules.[1-10] As an outgrowth of this we envisioned a two-step MCSCF procedure based on the LDGUGA formalism.[11] In one iteration the first step would be a CI calculstion in which the orbitals were "fixed". This is followed by a computation to determine how the orbitals should be changed to minimize the total energy. In this second step the orbitals vary but the expansion coefficients are frozen. The integrals are then transformed to the new orbital basis and we iterate until the changes in the orbitals and the expansion coefficients are suitably small. In computing the orbital changes the two-particle density matrix and the transformed integrals are needed. The ability to generate the two-particle density matrix rapidly is a key feature in our method.

Methods Used for Computing Orbital Changes

Currently we have two operational methods for computing the changes in the orbitals from one iteration to the next. The first involves the symmetrizing of the matrix of Lagrange multipliers.[11] In developing this procedure first notice that the CI energy can be expressed in a simple form as

$$E_{CI} = \sum_{ijk\ell} G_{ijk\ell}\,[ij;k\ell] + \sum_{ij} Q_{ij}\langle i|\hat{h}|j\rangle . \qquad (1)$$

If a unitary transformation was applied to the orbital basis to find a "better" set of orbitals, the CI energy in this new basis could then be approximated by

$$E \simeq \sum_{ijk\ell}\sum_{abcd} G_{ijk\ell}U_{ai}U_{bj}U_{ck}U_{d\ell}[ab;cd] +$$
$$\sum_{ij}\sum_{ab} Q_{ij}U_{ai}U_{bj}\langle a|\hat{h}|b\rangle \quad . \qquad (2)$$

This equation is an approximation because the density matrices G and Q depend on the orbital basis in a complex manner. If the unitary transformation U were close to unity, then this approximation would be fairly good. This observation suggests an iterative procedure to find the MCSCF wave function. If for any CI calculation the transformation U could be found that minimizes the energy through equation (2), then the CI calculation could be repeated in this new basis (i.e., compute the correct density matrix for this new basis). This procedure could then be repeated until the transformation matrix approaches unit, which should also yeild the MCSCF wave function.

The problem then becomes one of finding a procedure that can determine the "best" unitary transformation for each iteration within this scheme. By assuming that U is fairly close to unity, $\underset{\sim}{U}$ can be represented as

$$\underset{\sim}{U} = \underset{\sim}{1} + \underset{\sim}{U}^{(1)} \quad . \qquad (3)$$

where $U^{(1)}$ is the first-order change in the unitary transformation. When this is substituted into equation (2) and all higher terms in $U^{(1)}$ are ignored, then the first order change in the energy can be further approximated as

$$E^{(1)} \simeq \sum_{ijk\ell}\sum_{abcd} G_{ijk\ell}[U^{(1)}_{ai}\delta_{bj}\delta_{ck}\delta_{d\ell} +$$
$$\delta_{ai}U^{(1)}_{bj}\delta_{ck}\delta_{d\ell} + \delta_{ai}\delta_{bj}U^{(1)}_{ck}\delta_{d\ell} +$$
$$\delta_{ai}\delta_{bj}\delta_{ck}U^{(1)}_{d\ell}][ab;cd] + \sum_{ij}\sum_{ab} Q_{ij}[U^{(1)}_{ai}\delta_{bj}$$
$$+ \delta_{ij}U^{(1)}_{bj}]\langle a|\hat{h}|b\rangle \qquad (4)$$

which may be simplified to

$$E^{(1)} \simeq \sum_{ijk\ell r} G_{ijk\ell}\;(4\;U^{(1)}_{ir}[rj;k\ell] +$$
$$\sum_{ijr} Q_{ij}\;(2\;U^{(1)}_{ir})\langle r|\hat{h}|j\rangle \quad . \qquad (5)$$

If 4 times, the two elecrron energy and twice the one electron energy is added to both sides of equation (5) the result is

$$4E_2 + 2E_1 + E^{(1)} \simeq \sum_{ijk\ell r} G_{ijk\ell}\;(4\;U_{ir}^{(1)})$$
$$[rj;k\ell] + \sum_{ijr} Q_{ir}(2\;U^{(1)}_{ir})\langle r|\hat{h}|j\rangle +$$
$$\sum_{ijk\ell} G_{ijk\ell}\;(4\;\delta_{ir})[rj;k\ell] +$$
$$\sum_{ijr} Q_{ir}\;(2\;\delta_{ir})\langle r|\hat{h}|j\rangle \qquad (6)$$

where E_1 and E_2 are the one and two electron energies respectively. Collecting terms we have

$$K + E^{(1)} \cong 4 \sum G_{ijk\ell}\, U_{ir}\, [rj;k\ell] + 2 \sum_{ijr} Q_{ir} U_{ir} \langle r|\hat{h}|j\rangle \qquad (7)$$

where $K = 4\ E_2 + 2\ E_1$. Now it is advantageous to note that the Lagrangian matrix $\underset{\sim}{X}$ is defined as

$$X_{ir} \cong \sum_{jk\ell} 4\ G_{ijk\ell}\ [rj;k\ell] + \sum_{j} 2\ Q_{ij} \langle r|\hat{h}|j\rangle . \qquad (8)$$

Using this relation equation (7) can be reduced to

$$K + E^{(1)} \cong \sum_{ir} X_{ir} U_{ir} \quad . \qquad (9)$$

One rather direct approach is to find the unitary transformation $\underset{\sim}{U}$ that minimizes the above equation. This can be done directly using pair rotations on $\underset{\sim}{X}$. Rewriting equation (9) as

$$K + E^{(1)} \cong \sum_{ir} X_{ir} U_{ir} = \sum_{i} (\underset{\sim\sim}{XU^T})_{ii} \qquad (10)$$

and expressing the transpose of $\underset{\sim}{U}$ as a product of pair rotation matrices

$$U^T = \prod_{i>j,m} \theta_{ij,m} \qquad (11)$$

where $\theta_{ij,m}$ is the m^{th} pair rotation matrix between orbitals i and j, an iterative scheme can be devised. For each orbital pair i and j a rotation angle θ can be solved for. The largest of these is picked and the rotation is applied to the Lagrangian matrix. A new set of θ matrices are calculated and this process is repeated until the angle of rotation is below some threshold value.

To find the value of the rotation for each ij pair, multiply X by θ_{ij} and take the trace of the resulting matrix. The difference between the trace before and after rotation is

$$tr(\underset{\sim\sim}{X1}) - tr(\underset{\sim\sim}{X\theta}_{ij}) = X_{ii}+X_{jj}-X_{ii}\cos\theta+X_{ij}\sin\theta - X_{jj}\cos\theta-X_{ji}\sin\theta \quad . \qquad (12)$$

Taking the derivative with respect to θ on both sides of equation (12) and equating to a zero gives

$$(X_{jj}+X_{ii})\sin\theta + (X_{ij}-X_{ji})\cos\theta = 0 \quad . \qquad (13)$$

Solving for tanθ yields

$$\tan\theta = \frac{X_{ji}-X_{ij}}{X_{ii}+X_{jj}} \quad . \qquad (14)$$

These rotations are applied until the transformed Lagrangian is symmetric. Such an approach yields acceptable convergence for some calculations, typically when all orbitals are occupied. In systems including virtual orbitals or orbitals of forced double occupancy, we find that an approach similar to that used by Hinze[12] results in improved convergence.

The additional terms required for this method are easily obtained and can be expressed as

$$Y_{ij} = Q_{ij}\langle i|h|j\rangle + \sum_{k\ell} (2\ G_{ijk\ell}[ij;k\ell] - G_{iik\ell}[jj;k\ell] - G_{jjk\ell}[ii;k\ell]) \qquad (15)$$

and

$$Z_{ij} = Q_{ij}(\langle i|h|i\rangle - \langle j|h|j\rangle) + \langle i|h|j\rangle (Q_{jj}-Q_{ii}) + \sum_{k\ell} [2\ G_{ijk\ell}([ii;k\ell] - [jj;k\ell]) + 2[ij;k\ell]\ (G_{jjk\ell} - G_{iik\ell})] \quad . \qquad (16)$$

The pair interaction (s) can then be solved from the quadratic equation

$$[\tfrac{1}{2}(X_{ji}-X_{ij}) + 3\ Z_{ij}]s^2 + [(X_{ii}-X_{jj}) + 4\ Y_{ij}]s + (X_{ij}-X_{ji}) = 0 \qquad (17)$$

for the smallest root. The unitary orbital transformation matrix can then be given by exp $[-\underset{\sim}{s}]$. Of course, the coefficient of the terms presented here are obtained from approximations. Varying some or all of the terms could further improve convergence.

The computational effort required for each iteration of this MCSCF scheme is roughly the same as with other MCSCF procedures for large basis sets (i.e., dominated by the integral transformation step).

<u>Generation of the 2 Particle Density Matrix</u>

As stated earlier the fast and efficient generation of the two-particle density matrix is crucial for the orbital optimization stage to be competitive with the CI portion in a two-step MCSCF procedure. From the first step we have a CI vector which contains the expansion coefficients for the configurations. The formula for the two-particle density matrix is

$$G_{ijk\ell} = \sum_{IJ} C_I C_J\ b_{IJ}^{ijk\ell} \qquad (18)$$

where C_I and C_J are the CI coefficients and $b_{IJ}^{ijk\ell}$ is the coupling coefficient between these two configurations. The coupling coefficients must be generated so $\underset{\sim}{G}$ may be computed. These coefficients are exactly the same as those used in determining the CI energy and are defined by loops. The simplest method for generating the density matrix is to regenerate these same loops. The only difference here is that once a loop has been generated it is processed differently. This allows the density matrix to be generated via the loop-driven algorithm. The effort required to compute the density matrix in this manner is only slightly greater than the effort required to generate the corresponding diagonalization tape.[5,7]

In the diagonalization tape generation step, loop coefficients are combined with appropriate integrals to form an overall loop value. This loop value is then used a number of times, determined by the loop breakdown algorithm, for each diagonalization iteration. In generating the density matrices, this process is reversed. When a loop is generated, the loop breakdown algorithm is used first to determine the total loop contribution d. This total loop contribution is simply a sum of the products of C_I and C_J for each separate loop contribution and is given by:

$$d = \sum_{j}^{\bar{x}_h} \sum_{k}^{x_t} C_{Y(z_h+m_\ell+k)+j} \cdot C_{Y(z_h+m'_\ell+k)+j} \qquad (19)$$

where m_ℓ and m'_ℓ are the weights of each branch of the loop, z_h is the primary upper walk weight, $\bar{x}_h$ and x_t are the number of upper and lower walks respectively, and Y is the indexing array. Once the total loop contribution has been determined, its product with the loop coefficients is added to the proper density matrix elements.

The density matrix elements here play a similar role to that played by integrals in the diagonalization tape generation step. These elements are stored in the same manner as the MO integrals, and a particular density matrix element can be found using integral storage offset arrays.[13] The same storage method allows a block of density matrix elements to be computed simultaneously with the loop-drive algorithm.

Examples of Typical Calculations

The variety of MCSCF computations we can perform with the LDGUGA system of programs is fairly large. We essentially can choose any set of configurations that can be run by the CI programs. Our distinct row table (DRT) program, which generates the configuration set, can obtain all single or single and double replacements with respect to one reference.[5] It can also do this for most two reference cases at the present. In addition, full CI within any subspace can be done and all single and double excitations can be formed from it. Higher than double excitations can also be included if desired.

A limited amount of selection can be performed within the configuration set. One method is to doubly occupy or delete an orbital from the CI portion of the calculation. We can allow any orbital to be forced doubly occupied, to be partially occupied, or virtual. These "frozen" orbitals (virtual or doubly occupied) may be deleted from the MCSCF entirely or allowed to mix with the partially occupied space as desired. For open shell systems the inclusion of only the Hartree-Fock interacting space can be optionally chosen.[5,21]

Since our diagonalization tape generation and eigenvalue extraction programs are relatively fast and our MCSCF iterations are dominated by the integral transformation time, we have chosen to use relatively large configuration sets. Our first published calculations were on the lowest triplet of the cyclopropyne molecule.[14] The configuration set consisted of 10,115 singly and doubly-excited configurations with respect to the SCF reference.

More recently we have performed a number of calculations on water to investigate the effects of inclusion of higher than doubly excited configurations in CI and MCSCF calculations.[15] The basis set we used was a standard double-zeta basis set.[16] O(9s5p/4s2p), H(4s/2s), and the geometry was fixed at the theoretically determined minimum obtained from a previous set of CI calculations including single and double excitations. Using the SCF orbitals, CI calculations including all single (S), all single and double (SD), all single, double, and triple (SDT), and all single, double, triple, and quadruple excitations (SDTQ) were carried out. Also full MCSCF calculations were carried out for each of these four configuration sets. These results are displayed in Table 1.

There are several significant differences between the CI and MCSCF results. Perhaps the greatest differences occur for the MCSCF wave function including only single excitations. The CI energy in term of canonical SCF orbitals is of course identical to the SCF energy, due to Brillouin's theorem.[17] In striking contrast, the MCSCF wave function including all single excitations (MCS) accounts for no less than 52.3% of the correlation energy recovered by the MCSD wave function or about 50% of the full correlation energy attainable with the present basis set.

Another interesting result in Table 1 is the fact that triple excitations are roughly <u>five</u> times more important within the MCSCF function than in the straight CISDT treatment. Specifically, the energy lowering relative to CISD is 0.0011 hartrees for CISDT, but 0.0058 hartrees for MCSDT. This means that triple excitations can be made quite important by the MCSCF procedure. Another way of saying the same thing is that quadruple excitations are important, and the annihilation (in the Brillouin-Levy-Berthier (BLB) sense[18]) of certain classes of quadruples leads to a noteworthy energy lowering. This argument also explains the very small energy difference between CISD and MCSD as the annihilation of the triple excitations is expected to produce slight energy changes. Continuing in the same vein, quintuple excitations are expected to be very unimportant, and their annihilation is accordingly unimportant, as seen in the small difference between the CISDTQ and MCSDTQ energies, namely 0.0006 hartrees.

In the near future we plan to perform various calculations on the ground state of ozone, including MCSCF containing all single and double excitation from one and two reference functions. Also in the works are MCSCF calculations on methylene with a configuration set including all excitation within the valence space. We are furthermore at work incorporating MCSCF techniques that will converge at a faster rate than our present programs.[19,20]

Concluding Remarks

The LDGUGA has been shown to be readily adaptable to finding MCSCF wave functions and energies. Its main advantages are that it can be used with large configuration sets and the one and two particle density matrix generation can be accomplished rapidly and efficiently.

Acknowledgments

This research was supported by the U.S. Department of Energy, the U.S. National Science Foundation, and the Robert A. Welch Foundation. We thank Drs. Bernard R. Brooks, Paul Saxe, John D. Goddard, Nicholas C. Handy, Russell M. Pitzer, Michael Dupuis, and Yukio Yamaguchi for their contributions.

References

1. F. A. Matsen, Intern. J. Quantum Chem. Symp. 8, 379 (1974).
2. J. Paldus, J. Chem. Phys. 61, 5321 (1974); Intern. J. Quantum Chem. Symp. 9, 165 (1975); Phys. Rev. A14, 1620 (1976).
3. J. Paldus, in Theoretical Chemistry: Advances and Perspectives, Volume 2, editor H. Eyring and D. J. Henderson (Academic, New York, 1976) p. 131.
4. I. Shavitt, Intern. J. Quantum Chem. Symp. 11, 131 (1977); 12, 5 (1978).
5. B. R. Brooks and H. F. Schaefer, J. Chem. Phys. 70, 5092 (1979).
6. P. E. M. Siegbahn, J. Chem. Phys. 70, 5391 (1979).
7. I. Shavitt, Chem. Phys. Lett. 63, 421 (1979).
8. B. R. Brooks, W. D. Laidig, P. Saxe, N. C. Handy, and H. F. Schaefer, Physica Scripta 21, 312 (1980).
9. J. Paldus and M. J. Boyle, Physica Scripta, in press. 21, 295 (1980)
10. I. Shavitt, Annual Report on "New Methods in Computational Quantum Chemistry and their Application on Modern Supercomputers" to NASA Ames Research Center, June 29, 1979.
11. B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard and H. F. Schaefer, in Proceedings of the Conference on the Unitary Group for the Evaluation of Energy Matrix Elements, editor Jürgen Hinze (Springer-Verlag, Berlin, 1980).
12. J. Hinze, J. Chem. Phys. 59, 6424 (1973).
13. B. R. Brooks, Ph.D. Thesis, Department of Chemistry, University of California, Berkeley, September, 1979.
14. B. R. Brooks, W. D. Laidig, P. Saxe, and H. F. Schaefer, J. Chem. Phys. 72, 0000 (1980).
15. W. D. Laidig, P. Saxe, H. F. Schaefer, J. Chem. Phys. in press.
16. S. Huzinaga, J. Chem. Phys. 42, 1293 (1965); T. H. Dunning, J. Chem. Phys. 53, 2823 (1970).
17. L. Brillouin, Act. Sci. Ind. No. 159 (Herrmann, Paris, 1934).
18. K. Ruedenberg, L. M. Cheung and S. T. Elbert, Int. J. Quantum Chem. 16, 1069 (1979).
19. W. J. Hunt, W. A. Goddard III, T. H. Dunning, Chem. Phys. Letters 6, 147 (1970).
20. D. L. Yeager and P. Jørgensen, J. Chem. Phys. 71, 755 (1979).

Table 1. Summary of CI and MCSCF calculations on the water molecule. S,D,T, and Q refer to single, double, triple and quadruple excitations.

Wave function	Number of Configurations	Total Energy	Correlation Energy	Energy Relative to CI S+D	Coefficient C_0 of Reference Configuration
Self-Consistent-Field (SCF)	1	-76.00 98 38	0.0	--	1.0
CI All Single Excitations	20	-76.00 98 38	0.0	--	1.0
MCSCF All Single Excitations	20	-76.08 33 39	0.073 502	--	0.89326
CI All S+D	361	-76.15 00 15	0.140 177	0.000 000	0.97874
MCSCF All S+D	361	-76.15 02 69	0.140 431	0.000 254	0.97726
CI All S+D+T	3,203	-76.15 11 56	0.141 318	0.001 141	0.97819
MCSCF All S+D+T	3.203	-76.15 57 65	0.145 927	0.005 750	0.90537
CI All S+D+T+Q	17,678	-76.15 76 03	0.147 765	0.007 588	0 97543
MCSCF All S+D+T+Q	17,678	-76.15 76 64	0.147 826	0.007 649	0.96797

A NOVEL SINGLE FOCK OPERATOR APPROACH TO THE MCSCF PROBLEM

J. W. McIver Jr., M. Page and R. N. Camp
Chemistry Department
State University of New York at Buffalo
Buffalo, New York 14214

SUMMARY

An approximately quadratically convergent "Single Fock Operator" method for calculating MCSCF wavefunctions is presented. The traditional closed shell SCF equations emerge as a special case of this scheme. Computer time comparisons are made with the King and Camp method.

INTRODUCTION

The closed shell Hartree Fock problem is traditionally solved by application of the fixed point iterative scheme,(1,2b)

$$\underset{\sim}{F}(\underset{\sim}{C}^{\circ})\underset{\sim}{C} = \underset{\sim}{S}\underset{\sim}{C}\underset{\sim}{\varepsilon} \ , \tag{1}$$

with

$$\underset{\sim}{C}^{\circ\dagger}\underset{\sim}{S}\underset{\sim}{C}^{\circ} = \underset{\sim}{C}^{\dagger}\underset{\sim}{S}\underset{\sim}{C} = \underset{\sim}{I} \ . \tag{2}$$

until self consistency is obtained, i.e., $\underset{\sim}{F}(\underset{\sim}{C}) \simeq \underset{\sim}{F}(\underset{\sim}{C}^{\circ})$. Here the notation $\underset{\sim}{F}(\underset{\sim}{C}^{\circ})$ is meant to imply that the symmetric Fock matrix is constructed from the square matrix $\underset{\sim}{C}^{\circ}$ of L.C.A.O. coefficients obtained in the previous iteration. The symmetric overlap matrix $\underset{\sim}{S}$ and the diagonal matrix of Lagrange multipliers $\underset{\sim}{\varepsilon}$ have their usual meanings. The above equations are a consequence of the variational principle in the sense that, at convergence, the energy is stationary with respect to all allowed variations of $\underset{\sim}{C}$. Although self consistent solutions to these equations are variationally correct, there is no guarantee that the use of these equations recursively from any set of starting orbitals will lead to a converged solution. Although convergence difficulties have been studied and circumvented,(2) the recursive use of the above equations by and large works very well for the closed shell case.

In disappointing contrast to the closed shell case, the use of the recursion eqn. (1), with eqn. (2) for open shell and MCSCF wavefunctions has shown poor convergence.(3) In these "Single Fock Operator" methods, the prescription for constructing $\underset{\sim}{F}(\underset{\sim}{C}^{\circ})$ is obtained by using projection operators and/or coupling operators to obtain eqns. (1) and (2) which are variationally correct at self consistency. But again there are no assurances of convergence when used recursively.

In this presentation, we will derive a prescription for forming an $\underset{\sim}{F}(\underset{\sim}{C}^{\circ})$ which leads to quadratic convergence of a variationally correct solution to eqns. (1) and (2) for any wavefunction built up of orthonormal orbitals. By making some simple approximations, a computationally attractive method emerges, a method which becomes identical to the traditional Hartree-Fock method in the closed shell case. Preliminary calculations of a two-determinant MCSCF wavefunction for a distorted ethylene molecule indicate that the method is comparible to the King and Camp method in computational efficiency.(4)

THEORY

The new orbital coefficients $\underset{\sim}{C}$ are obtained from the old coefficients by

$$\underset{\sim}{C} = \underset{\sim}{C}^{\circ}\underset{\sim}{U} \ , \tag{3}$$

where $\underset{\sim}{U}$ is as yet an undetermined matrix. A necessary and sufficient condition for preserving the orthonormality of the orbitals is, from (2) and (3),

$$\underset{\sim}{C}^{\dagger}\underset{\sim}{S}\underset{\sim}{C} = \underset{\sim}{U}^{\dagger}[\underset{\sim}{C}^{\circ\dagger}\underset{\sim}{S}\underset{\sim}{C}^{\circ}]\underset{\sim}{U} = \underset{\sim}{I} \ . \tag{4}$$

Thus orthonormality is preserved if,

$$\underset{\sim}{U}^{\dagger}\underset{\sim}{U} = \underset{\sim}{I} \ . \tag{5}$$

At this point one can proceed by minimizing the expression $E = \langle\Psi|H|\Psi\rangle$ with respect to the elements of $\underset{\sim}{C}$ (or of $\underset{\sim}{U}$) where the constraint conditions are built in using the method of Lagrangian multipliers. This approach however, has not been very successful for problems other than the closed shell Hartree-Fock case.

We choose not to deal directly with the constrained elements of $\underset{\sim}{C}$ or $\underset{\sim}{U}$, but rather we seek an equivalent set of unconstrained (independent) variables. Since the eigenvectors of a real symmetric matrix are orthonormal, we view the (orthonormal) columns of the matrix $\underset{\sim}{U}$ as being the eigenvectors of a real symmetric matrix $\underset{\sim}{Q}$. Thus,

$$\underset{\sim}{Q}\underset{\sim}{U} = \underset{\sim}{U}\underset{\sim}{\varepsilon} \ . \tag{6}$$

The advantage of this approach is that the orthonormality of the orbitals is retained simply by keeping $\underset{\sim}{Q}$ symmetric. The n(n-1)/2 off-diagonal elements of $\underset{\sim}{Q}$ may then undergo <u>unconstrained</u> variations and thus form our set of independent variables. The "single Fock matrix" of eqn. (1) is then obtained from eqns. (3) and (6) as,

$$\underset{\sim}{F}(\underset{\sim}{C}^{\circ}) = \underset{\sim}{C}^{\circ\dagger-1}\underset{\sim}{Q}(\underset{\sim}{C}^{\circ})\ \underset{\sim}{C}^{\circ-1} \ . \tag{7}$$

To find the values of the elements Q_{ij} which lower the energy, we expand the energy in a Taylor series in the n(n-1)/2 off diagonal elements of $\underset{\sim}{Q}$ and choose the elements Q_{ij} via an approximate Newton-Raphson procedure.

Thus,

$$E = E(\underset{\sim}{U}(\underset{\sim}{Q})), \tag{8}$$

and in the quadratic approximation,

$$Q_{ij}^{(\text{optimum})} = -\sum_{k<l} (\underset{\sim}{M}^{-1})_{ij,kl}\, E'_{kl} \quad i<j \ , \tag{8}$$

where

$$E'_{kl} \equiv (\partial E/\partial Q_{kl})_\circ \tag{9}$$

and

$$M_{ij,kl} \equiv (\partial^2 E/\partial Q_{ij}\partial Q_{kl})_o \ . \tag{10}$$

A quadratically convergent scheme would consist of calculating the first and second derivatives with respect to Q_{ij} evaluated at zero and inverting the super-matrix $\underset{\sim}{M}$.

We instead consider only the "diagonal" second derivatives. That is, we neglect all terms of the form, $\partial^2 E/\partial Q_{ij}\partial Q_{kl}$, where $i \neq k$ and $j \neq l$. This is the "pairwise" or diagonal dominant approximation. Although this destroys the second order convergence properties, it results in enormous computational savings in that eqn. (8) can now be written as

$$Q_{ij}^{(\text{optimum})} = -\lambda(\partial E/\partial Q_{ij})_\circ/(\partial^2 E/\partial Q_{ij}^2)_\circ \ , \tag{11}$$

where the variationally determined scaling parameter λ helps overcome the effects of the quadratic and pairwise approximations. We make no assumptions about the diagonal elements of $\underset{\sim}{Q}$ except the simplifying one that they are nondegenerate.

The first derivatives, eqn. (9) are relatively straightforward to calculate. The energy expression is an explicit function of the orbital expansion coefficients which in turn depend implicitly on the Q_{ij}. Thus,

$$\begin{aligned}\frac{\partial E}{\partial Q_{ij}} &= \sum_k^n \sum_l^n \frac{\partial E}{\partial C_{kl}} \frac{\partial C_{kl}}{\partial Q_{ij}} \\ &= \mathrm{Tr}\,\Big(\underset{\sim}{\nabla E}^{\dagger} \frac{\partial \underset{\sim}{C}}{\partial Q_{ij}}\Big) \\ &= \mathrm{Tr}\,\Big(\underset{\sim}{\nabla E}^{\dagger}\underset{\sim}{C}^{\circ} \frac{\partial \underset{\sim}{U}}{\partial Q_{ij}}\Big) \ .\end{aligned} \tag{12}$$

It should be noted that the ith column of $\underset{\sim}{\nabla E}$ can be regarded as the gradient of the energy with respect to variations in the coefficients of the ith orbital:

$$\underset{\sim}{\nabla}_i E = \begin{pmatrix} \dfrac{\partial E}{\partial C_{1i}} \\ \vdots \\ \dfrac{\partial E}{\partial C_{ni}} \end{pmatrix} . \tag{13}$$

The term $\partial \underset{\sim}{U}/\partial Q_{ij}$ in eqn. (12) can be evaluated using eqns. (5) and (6) and matrix perturbation theory. Denoting differentiation with respect to a particular Q_{ij} by a prime, eqns. (6) and (5) can be differentiated to give

$$\underset{\sim}{Q}'\underset{\sim}{U} - \underset{\sim}{U}\underset{\sim}{\varepsilon}' + \underset{\sim}{Q}\underset{\sim}{U}' - \underset{\sim}{U}'\underset{\sim}{\varepsilon} = \underset{\sim}{0} \ , \tag{14}$$

and

$$\underset{\sim}{U}^{\dagger}\underset{\sim}{U}' + \underset{\sim}{U}'^{\dagger}\underset{\sim}{U} = \underset{\sim}{0} \ . \tag{15}$$

Multiplying eqn. (14) on the left by $\underset{\sim}{U}^{\dagger}$ and letting,

$$\underset{\sim}{U}' = \underset{\sim}{U}\underset{\sim}{V} \quad \text{(spectral expansion)}. \tag{16}$$

We have, with the aid of eqns. (5) and (6),

$$\underset{\sim}{U}^{\dagger}\underset{\sim}{Q}'\underset{\sim}{U} - \underset{\sim}{\varepsilon}' + \underset{\sim}{\varepsilon}\underset{\sim}{V} - \underset{\sim}{V}\underset{\sim}{\varepsilon} = \underset{\sim}{0} \ , \tag{17}$$

with

$$\underset{\sim}{V}^{\dagger} + \underset{\sim}{V} = \underset{\sim}{0} \ . \tag{18}$$

hence $\underset{\sim}{V}$ must be antisymetric. The diagonal elements of eqn. (17) give the $\underset{\sim}{\varepsilon}'$ and the diagonal elements of $\underset{\sim}{V}$ (zero here) are given by eqn. (18). The off-diagonal elements of $\underset{\sim}{V}$ are obtained from the off-diagonal elements of eqn. (17), i.e.

$$V_{kl} = (\underset{\sim}{U}^{\dagger}\underset{\sim}{Q}'\underset{\sim}{U})_{kl}/(\varepsilon_{ll} - \varepsilon_{kk}) \tag{19}$$

Since the matrix Q' contains all zero's except for $Q'_{ij} = Q'_{ji} = 1$, Eqn. (19) can be written as,

$$V_{kl} = (U_{ik}U_{jl} + U_{jk}U_{il})/(\varepsilon_{ll} - \varepsilon_{kk}) \tag{20}$$

Using eqns. (12), (16), and (20), the first derivatives can be written as,

$$\begin{aligned}\partial E/\partial Q_{ij} &= \mathrm{Tr}\,(\underset{\sim}{R}\cdot\underset{\sim}{V}) \\ &= \sum_k \sum_{l\neq k} R_{kl}\,(U_{ik}U_{jl} + U_{jk}U_{il})/(\varepsilon_{ll} - \varepsilon_{kk})\end{aligned} \tag{21}$$

where $\underset{\sim}{R} = \underset{\sim}{\nabla E}^{\dagger}\underset{\sim}{C}$. Thus, given $\underset{\sim}{R}$, calculating all of the first derivatives at any value of $\underset{\sim}{Q}$ requires $O(n^4)$ operations. A large class of quadratically convergent methods (such as conjugate gradient or variable metric methods) require only the first derivatives. When using eqn. (8), or as we do, eqn. (11), only the values at $Q_{ij}=0$ are needed. Here $U_{ij}=\delta_{ij}$, and eqn. (21) becomes

$$\Big(\frac{\partial E}{\partial Q_{ij}}\Big)_o = \frac{R_{ij} - R_{ji}}{\varepsilon^{\circ}_{jj} - \varepsilon^{\circ}_{ii}} = \frac{R_{ij} - R_{ji}}{Q_{jj} - Q_{ii}} \tag{22}$$

which requires only $O(n^2)$ operations. Note that at convergence, the proper variational conditions

$$R_{ij} - R_{ji} = \underset{\sim}{\nabla}_i \underset{\sim}{E}^{\dagger}\underset{\sim}{C}_j - \underset{\sim}{\nabla}_j \underset{\sim}{E}^{\dagger}\, \underset{\sim}{C}_i = 0 \tag{23}$$

are satisfied, since the Q_{ii} can be chosen nondegenerate and finite.

To evaluate the second derivatives, it is easiest to freeze all off-diagonal elements of $\underset{\sim}{Q}$ except Q_{ij}. The matrix $\underset{\sim}{U}$ is then an identity matrix except,

$$U_{ii} = U_{jj} = \cos\theta_{ij} ,$$

$$U_{ij} = -U_{ji} = \sin\theta_{ij} \tag{24}$$

and θ_{ij} is determined by the requirement that $\mathbf{Q}$ be diagonalized,

$$\tan 2\theta_{ij} = \frac{2Q_{ij}}{Q_{jj}-Q_{ii}} . \tag{25}$$

Under these conditions, the first derivative is

$$\frac{\partial E}{\partial Q_{ij}} = \frac{\partial E}{\partial \theta_{ij}} \frac{\partial \theta_{ij}}{\partial Q_{ij}}$$

$$= (R_{ij}-R_{ji}) \frac{\cos 2\theta_{ij}}{Q_{jj}-Q_{ii}} , \tag{26}$$

while the second derivative becomes

$$\left(\frac{\partial^2 E}{\partial Q_{ij}^2}\right)_0 = \frac{\left(\frac{\partial}{\partial Q_{ij}} [R_{ij} - R_{ji}]\right)_0}{(Q_{jj}- Q_{ij})^2}$$

$$= \frac{\left(\frac{\partial}{\partial \theta_{ij}} [\nabla_{\mathbf{i}} E^{\dagger} \mathbf{C}_j - \nabla_{\mathbf{j}} E^{\dagger} \mathbf{C}_i]\right)_0}{(Q_{jj} - Q_{ii})^2} . \tag{27}$$

At this point, matrices $\mathbf{F}_i$ are found which satisfy

$$\nabla_{\mathbf{i}} E = 4\, \mathbf{F}_i \mathbf{C}_i . \tag{28}$$

Orbitals which share the same matrix $\mathbf{F}_i$ are said to belong to the same "shell" and the matrix $\mathbf{F}_i (= \mathbf{F}_\alpha)$ is the "Fock Matrix" for shell α.

Using eqns. (3), (24) and (28), eqn. (27) for the second derivative becomes

$$\left(\frac{\partial^2 E}{\partial Q_{ij}^2}\right)_0 =$$

$$\frac{\mathbf{C}_j^{\circ} \mathbf{F}_\alpha \mathbf{C}_j - \mathbf{C}_i^{\circ\dagger} \mathbf{F}_\alpha \mathbf{C}_i^{\circ} + \mathbf{C}_i^{\circ\dagger} F_{\boldsymbol{\beta}} \mathbf{C}_i^{\circ} - C_j^{\circ\dagger} F_\beta C_j^{\circ} + \gamma_{ij}}{(Q_{jj} - Q_{ii})^2} , \tag{29}$$

where i is in shell α and j is in shell β, and,

$$\gamma_{ij} = [\mathbf{C}_i^{\circ\dagger} \frac{\partial}{\partial \theta_{ij}} (\mathbf{F}_\alpha - \mathbf{F}_\beta)\mathbf{C}_j^{\circ}]_0 . \tag{30}$$

The term γ_{ij} has been named the Fock correction term by Goddard et al.(5) If we set this term equal to zero as a further approximation, eqn. (11) for the optimum value of Q_{ij} becomes

$$Q_{ij} = \lambda \frac{(T^i_{ij} - T^j_{ij})\,(Q_{jj} - Q_{ii})}{T^i_{jj} - T^i_{ii} + T^j_{ii} - T^j_{jj}} , \tag{31}$$

where $T^\alpha_{kl} \equiv \mathbf{C}_k^{\circ\dagger} \mathbf{F}_\alpha \mathbf{C}^{\circ}_l$.

CLOSED SHELL CASE

The method outlined above reduces to the traditional closed shell Hartree Fock method by judiciously choosing the elements Q_{ij} where i and j are in the same shell. It follows by differentiating the energy expression that,

$$\nabla_{\mathbf{i}} E = 4\mathbf{F}^{Cl}\mathbf{C}_i \qquad i = \text{occupied}$$

and

$$\nabla_{\mathbf{i}} E = \mathbf{0} , \qquad i = \text{virtual}$$

where, in the usual notation,

$$\mathbf{F}^{Cl}_{\mu\nu} = \mathbf{H}^{core}_{\mu\nu} + \sum_{\lambda\sigma} P^{Cl}_{\lambda\sigma}[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] ,$$

$$\text{or } \mathbf{F}^{Cl} = \mathbf{H} + 2\,\mathbf{J}(\mathbf{P}^{Cl}) - \mathbf{K}(\mathbf{P}^{Cl}) , \tag{32}$$

where $\mathbf{P}^{Cl} = \sum_i^{occ} \mathbf{C}_i \mathbf{C}_i^{\dagger}$.

If i and j are in the same shell (both occupied or both virtual) then they share a common Fock matrix ($\mathbf{F}^{Cl}$ or $\mathbf{0}$).

$$R_{ij} - R_{ji} = C_i^{\dagger} (\mathbf{F} - \mathbf{F})\mathbf{C}_j = 0 \tag{33}$$

for all θ_{ij}. Thus $Q_{ij}^{(optimum)}$ is indeterminant if i and j are in the same shell. These Q_{ij}'s serve only to mix orbitals in the same shell and can be assigned arbitrary values. While it is usually convenient to set them all equal to zero, in the present case we choose them to be

$$Q_{ij} = \mathbf{C}_i^{\circ\dagger} \mathbf{F}^{Cl} \mathbf{C}_j^{\circ} ; \quad \text{i and j in the same shell.} \tag{34}$$

If i is occupied and j is virtual, then eqn. (31) becomes

$$Q_{ij} = \frac{C_i^{\circ} \; F^{Cl} C_j^{\circ} \, (C_j^{\circ} F^{Cl} C_j^{\circ} - C_i^{\circ} F^{Cl} C_i^{\circ})}{(\mathbf{C}_j^{\dagger} \mathbf{F}^{Cl} \mathbf{C}_j - \mathbf{C}_i^{\circ\dagger} \mathbf{F}^{Cl} \mathbf{C}_i^{\circ})} ,$$

or

$$Q_{ij} = \mathbf{C}_i^{\dagger} \mathbf{F}^{Cl} \mathbf{C}_j^{\circ} . \tag{35}$$

Exactly the same equation results when i is virtual and j is occupied. Hence in matrix form,

$$Q^{optimum} = \mathbf{C}^{\circ\dagger} \mathbf{F}^{Cl} \mathbf{C}^{\circ} , \tag{36}$$

where the diagonal elements of $\mathbf{Q}$ are also calculated by Eq. (35).

According to eqn. (7),

$$\mathbf{F}^{optimum} (C^{\circ}) = \mathbf{F}^{Cl}(C^{\circ}) .$$

and the traditional closed shell Hartree Fock method is recovered.

PRELIMINARY MCSCF CALCULATIONS

The HONDO program of King et al.(6) was modified to calculate the best wavefunction of the form

$$\Psi = \sigma_A|\phi_1\bar{\phi}_1 \cdots \phi_n\bar{\phi}_n\phi_a\bar{\phi}_a| + \sigma_B|\phi_1\bar{\phi}_1 \cdots \phi_n\bar{\phi}_n\phi_B\bar{\phi}_B| . \tag{37}$$

This is the simplest Generalized Valence Bond wavefunction$^{(5)}$ for a system of $2(n+1)$ electrons. In this case there are four shells having occupation numbers 1, σ_A^2, σ_B^2 and 0, the last being the virtuals. The four corresponding Fock matrices are

$$\underset{\sim}{F}^P = \underset{\sim}{H} + \underset{\sim}{G}(\underset{\sim}{P}) + \sigma_A^2\, \underset{\sim}{G}(\underset{\sim}{A}) + \sigma_B^2\, \underset{\sim}{G}(\underset{\sim}{B}) \;, \tag{38}$$

$$\underset{\sim}{F}^A = \sigma_A^2\, [\underset{\sim}{H} + \underset{\sim}{G}(\underset{\sim}{P}) + \underset{\sim}{G}(\underset{\sim}{A})] + \sigma_A\sigma_B\, \underset{\sim}{K}(\underset{\sim}{B}) \;, \tag{39}$$

$$\underset{\sim}{F}^B = \sigma_B^2\, [\underset{\sim}{H} + \underset{\sim}{G}(\underset{\sim}{P}) + \underset{\sim}{G}(\underset{\sim}{B})] + \sigma_A\sigma_B\, \underset{\sim}{K}(\underset{\sim}{A}) \;, \tag{40}$$

$$\underset{\sim}{F}^V = \underset{\sim}{0} \;, \tag{41}$$

where $\underset{\sim}{G} = 2\underset{\sim}{J}-\underset{\sim}{K}$, $\underset{\sim}{P}=\sum_i^n \underset{\sim}{C}_i\underset{\sim}{C}_i{}^+$, $\underset{\sim}{A}=\underset{\sim}{C}_A\underset{\sim}{C}_A{}^+$, $\underset{\sim}{B}=\underset{\sim}{C}_B\underset{\sim}{C}_B{}^+$.

Each iteration of the method consists of the following steps

1) Optimization of the coefficients σ_A and σ_B.

2) Construction of the Fock matrices for the three occupied shells (cf. equations (38)-(40)), from the coefficients, $\underset{\sim}{C}^\circ$, obtained either from the previous cycle or from the starting orbitals.

This has $O(n^4)$ operations and is the rate limiting step.

3) Calculation of the matrices,

$$\underset{\sim}{T}^P = (\underset{\sim}{C}^\circ)^+\underset{\sim}{F}^P\underset{\sim}{C}^\circ \tag{42}$$

$$\underset{\sim}{T}^P = (\underset{\sim}{C}^\circ)^+\underset{\sim}{F}^A\underset{\sim}{C}^\circ \tag{43}$$

$$\underset{\sim}{T}^B = (\underset{\sim}{C}^\circ)^+\underset{\sim}{F}^B\underset{\sim}{C}^\circ. \tag{44}$$

4) Estimation of λ^{optimum} from information obtained in the previous cycle.

5) Calculation of Q_{ij}^{optimum} from

$$Q_{ij}^{\text{optimum}} = 0, \text{ if } i \text{ and } j \text{ are in the same shell} \tag{45}$$

$$Q_{ij}^{\text{optimum}} = \lambda^{\text{optimum}} \frac{(T_{ij}^i - T_{ij}^j)(Q_{jj} - Q_{ii}),}{T_{jj}^i - T_{jj}^j - T_{ii}^i + T_{ii}^j.} \tag{46}$$

if i and j are in different shells.

6) Diagonalization of $\underset{\sim}{Q}^{\text{optimum}}$ to obtain $\underset{\sim}{U}$ (cf. equation (6)).

7) Use of equation (3) to obtain the coefficient matrix $\underset{\sim}{C}$.

8) Test for convergence and if not converged return to step 1).

The following table gives a rough indication of the convergence properties of the method. A two-determinant wavefunction for a distorted (C_1 symmetry) ethylene molecule was calculated in an STO-3G basis. For comparison, an equivalent wavefunction was calculated using the King and Camp method. Both runs were made on a CDC Cyber 174 computer (5.7 μsec/floating point multiplication). The starting orbitals were the same.

TABLE 1

	This Method (1.355 sec/cycle)	King and Camp Method (1.730 sec/cycle)
Cycle #	Energy, Au	Energy, Au
1	-110.56805242	-110.56805299
2	-110.57000382	-110.56842031
3	-110.57018313	-110.56860177
4	-110.57024124	-110.56985300
5	-110.57025596	-110.56990704
6	-110.57023561	-110.57015071
7	-110.57026603	-110.57017911
8	-110.57026909	-110.57021564
9	-110.57027267	-110.57021969
10	-110.57027454	-110.57021849
11	-110.57026760	-110.57026074
12	-110.57027499	-110.57027184
13	-110.57027507	-110.57027257
14	-110.57027512	-110.57027422
15	-110.57027516	-110.57027455
16	-110.57027488	-110.57027507
17	-110.57027523	-110.57027518
18	-110.57027524	-110.57027524
19	-110.57027524	-110.57027524
20	-110.57027524	-110.57027524

ACKNOWLEDGEMENTS

We are grateful to Harry King and Dick Stanton for many useful discussions. This work is supported by NSF grant CHE7901138.

REFERENCES

1. C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951).
2.a. R. N. Camp and H. F. King (In preparation).
 b. R. E. Stanton (In preparation).
3. See A. Golebiewski, J. Hinze and F. Yurtsever, J. Chem. Phys. 70, 1101 (1979) and references contained therein.
4. H. F. King and R. N. Camp (this proceedings).
5. F. W. Bobrowicz and W. A. Goddard III, Modern Theoretical Chemistry 3, 79, Plenum, New York (1977).
6. M. Dupuis, J. Rys and H. F. King, J. Chem. Phys. 65, 111 (1976) and Q.C.P.E. program number 338.

GENERALIZED MOLECULAR ORBITAL THEORY:
A LIMITED MULTICONFIGURATION SELF-CONSISTENT-FIELD-THEORY

Michael B. Hall
Department of Chemistry
Texas A&M University
College Station, TX 77843

The generalized molecular orbital (GMO) approach is a limited type of multiconfiguration self-consistent-field (MCSCF) calculation which divides the orbitals of a closed shell molecule into four shells: doubly occupied, strongly occupied, weakly occupied, and unoccupied. The orbitals within each shell have the same occupation number and are associated with the same Fock operator. Thus, the orbital optimization is ideally suited to solution via a coupling operator. The determination of the orbitals is followed by a configuration interaction (CI) calculation within the strongly and weakly occupied shells. Results for BH_3 show a striking similarity between the GMO's and the natural orbitals (NO's) from an all singles and doubles CI calculation. Although the GMO approach would not be accurate for an entire potential surface, results for spectroscopic constants of N_2 show that it is suitable near the equilibrium geometry. In this paper we describe the use of the GMO technique to determine the primary orbital space, but a potentially important application may be in the determination of a secondary orbital space following a more accurate MCSCF determination of the primary space.

INTRODUCTION

The need to go beyond the restricted, single-configuration, Hartree-Fock (RHF) approximation[1] and to include electron correlation is clearly evident. This need exists not only for quantitative accuracy but also for semiquantitative accuracy and, in some cases, even for qualitative results. We are interested in the electronic structure of rather large, transition-metal systems, especially those with metal-metal bonds and with metal-ligands bonds which involve strong pi acceptor ligands such as dioxygen, nitrosyl, and carbenes. In systems such as these, two factors conspire to render the RHF approximation a particularly poor one. First, transition metals, especially those of the first transition series, have rather compact valence d orbitals, a factor which leads to small metal-metal or metal-ligand overlap integrals. Second, for the bonds described above the orbitals of the two components have similar energy; thus, there are large near degeneracy correlation effects in these systems.

The most direct way to go beyond the RHF wavefunction toward a correct description of the wavefunction is the traditional configuration interaction (CI) calculation.[2] However, for very large systems the four-index transformation and the large number of configurations make this calculation prohibitively expensive. An alternative solution would be to use fewer configurations in a multiconfiguration self-consistent-field (MCSCF) calculation where the form of the molecular orbitals (MO's) is simultaneously optimized. However, each orbital active in this optimization will now require construction of a separate Fock operator or a partial four-index transformation at each iteration. If the number of active orbitals is large, this calculation may also be too costly. The procedure described in this paper was developed to avoid these two problems. Our goal is to develop a method which will rapidly determine a set of optimized primary orbitals, which are suitable for subsequent CI calculations. We will begin by describing the theory, and then some results for calculations of small molecules, which we will compare with the results of more accurate calculations.

THEORY

Shell Structure

For a 2n electron closed-shell molecule the RHF wavefunction is taken as a single determinant of doubly occupied MO's,

$$\Psi = |\phi_1\bar{\phi}_1 \text{---} \phi_t\bar{\phi}_t \text{---} \phi_n\bar{\phi}_n| . \tag{1}$$

In typical applications to molecules the MO's are expanded in a basis set

$$\phi_i = \sum_{a=1}^{m} C_{ia}\, \chi_a \tag{2}$$

and the RHF energy is minimized with respect to the C_{ia}'s by the Roothaan procedure.[1] Because of the expansion in a basis set, one also obtains a number of virtual or unoccupied MO's. Thus, the orbital space for the RHF calculation consists of two shells; the first has n doubly occupied orbitals while the second has m-n unoccupied orbitals. Schematically, we may write this partition as

$$(\phi_1 \text{---} \phi_n)^2 (\phi_{n+1} \text{---} \phi_m)^0 \tag{3}$$

One of the reasons that the solution to the RHF problem via the Roothaan procedure is computationally attractive is that all of the orbitals are eigenfunctions of the same Fock operator.

With our goal of a simple MCSCF procedure in mind, it occurred to us that to take advantage of the computationally attractive nature of standard MO theory, we should continue to treat the orbitals in groups or shells. All orbitals in each shell should have equal occupation numbers and be associated with the same Fock operator. The previously doubly occupied shell is divided into two shells, one, which remains doubly occupied, and another, which is strongly occupied with variable occupation number. Likewise, the previously unoccupied orbitals are divided into two shells, one, which is weakly occupied with variable occupation number, and another, which remains unoccupied. It is this shell structure, which can be written as

$$(\phi_1 ---)^2(---\phi_n)^x(\phi_{n+1} ---)^y(---\phi_m)^0 \qquad (4)$$

that suggested the name generalized molecular orbital (GMO) theory. As in the RHF case the notation can be extended to open shells where one would have an additional shell of singly occupied orbitals (for the high spin situation). In subsequent discussion we will refer to these shells by the following letter designations: R for the doubly occupied shell, S for the singly occupied occupied shell, T for the strongly occupied shell, U for the weakly occupied shell and V for the unoccupied shell. We will use the corresponding lower case letter, r,s,t,u, and v to refer to individual orbitals within their respective shells.

Wavefunction and Energy

One particularly simple wavefunction which satisfies this orbital partitioning is

$$\Psi = (1 - k_T k_U \lambda^2)^{1/2} \psi_{oo} + \lambda \sum_t \sum_u \psi_{tu} \qquad (5)$$

where

$$\psi_{oo} = | \phi_1 \bar{\phi}_1 --- \phi_t \bar{\phi}_t --- \phi_n \bar{\phi}_n | \qquad (6)$$

and

$$\psi_{tu} = | \phi_1 \bar{\phi}_1 --- \phi_u \bar{\phi}_u --- \phi_n \bar{\phi}_n |. \qquad (7)$$

Thus, the wavefunction consists of a dominant single determinant, ψ_{oo}, plus all determinants which can be constructed by replacing a pair of orbitals in the T shell with a pair in the U shell, ψ_{tu}. These pair-excited functions are then weighted equally by the variational parameter λ. The k_I numbers refer to the number of orbitals in the I th shell. For this wavefunction the occupation numbers, x and y, will be $2(1 - k_U\lambda^2)$ and $2k_T\lambda^2$, respectively. For the simple open shell case with k_S orbitals each containing one electron of the same spin, we would add k_S singly occupied orbitals to both ψ_{oo} and ψ_{tu}.

The total electronic energy for this wavefunction can be written as

$$E = \sum_I f_I H_I + \sum_I \sum_J (a_{IJ} J_{IJ} + b_{IJ} K_{IJ}) \tag{8}$$

where H_I, J_{IJ}, and K_{IJ} partial sums are over one-electron, coulomb, and exchange integrals, respectively.

$$H_I = 2\sum_i h_i, \; J_{IJ} = \sum_i \sum_j J_{ij}, \text{ and } K_{IJ} = \sum_i \sum_j K_{ij} \tag{9}$$

where the sums over the lower case i and j include only orbitals in the shell I and J respectively. The values for f_I, a_{IJ}, and b_{IJ} are given in Table I.

TABLE I. GMO Energy Parameters

Type	Shell	Shell / Parameter			
		R	S	T	U
f_I		1	$\frac{1}{2}$	$(1-k_U\lambda^2)$	$k_T\lambda^2$
a_{IJ}	R	2	1	$2(1-k_U\lambda^2)$	$2k_T\lambda^2$
	S		$\frac{1}{2}$	$(1-k_U\lambda^2)$	$k_T\lambda^2$
	T			$2(1-2k_U\lambda^2)$	$2(k_T-1)\lambda^2$
	U				0
b_{IJ}	R	-1	$-\frac{1}{2}$	$-(1-k_U\lambda^2)$	$-k_T\lambda^2$
	S		$-\frac{1}{2}$	$-\frac{1}{2}(1-k_U\lambda^2)$	$-\frac{1}{2}k_T\lambda^2$
	T			$-(1-3k_U\lambda^2)$	$\gamma\lambda-(k_T-1)\lambda^2$
	U				$k_T\lambda^2$

$$\gamma = (1-k_T k_U \lambda^2)^{\frac{1}{2}}$$

Coupling Operator

Because these coefficients do not depend on the individual orbitals, the wavefunction (equ. 5) and energy (equ. 8) are independent of a unitary transformation of the orbitals within a shell and depend only on the mixing between shells. Since the number of shells is small, the variational solution for the orbital optimization may be performed with a coupling operator. For each occupied shell we can define Fock operators as

$$F_I = f_I h + \sum_J (a_{IJ} J_J + b_{IJ} K_J) \tag{10}$$

and projection operators as

$$P_I = \sum_i (|\phi_i\rangle\langle\phi_i|). \tag{11}$$

The projection operator for the virtual shell can be defined as

$$P_V = (1 - \sum_I P_I) \tag{12}$$

Following Hirao[3], we define a generalized coupling operator as

$$R = \sum_I (P_V F_I P_I + P_I F_I P_V) + \sum_{I \neq J}\sum (\theta_{JI} - \theta_{IJ}) P_J (F_I - F_J) P_I + \sum_I P_I F_I P_I \tag{13}$$

where the sums are over the occupied shells. The molecular orbitals are then obtained from the eigenvalue equation

$$R \mid \phi_i \rangle = \varepsilon_i \mid \phi_i \rangle \tag{14}$$

There remains considerable freedom in the R operator. For example, one may add the operator

$$O = \sum_I P_I \Omega_I P_I \tag{15}$$

where Ω_I is any one-electron operator. This freedom allows us to make the orbitals within each shell canonical over any operator we wish. Presently, we add the following operators to our coupling operator

$$P_V F_U P_V + \omega(P_U F_T P_U + P_V F_T P_V) \tag{16}$$

The first term makes the virtual orbitals canonical over the same operator as the weakly occupied ones. The second term, which is weighted by ω, allows us to vary simultaneously the canonicalization of the weakly occupied and virtual shells from being canonical over F_U (ω=0) to being canonical over $F_U + F_T$ (ω=1). The latter case corresponds closely to making the U and V shells canonical over just F_T because its larger values will dominate those of F_U. The coefficient $(\theta_{JI} - \theta_{IJ})$ is taken as $-\theta_{IJ}$ for J > I and corresponds to a weighting factor for the first derivatives. For the case $\omega = 1$ and $\theta_{IJ} = 1$ the coupling matrix will approach the usual RHF matrix as the configuration weighting parameter, λ, approaches zero.

In the present program all the necessary Fock matrices are formed in a single pass through the atomic integral file. These matrices are then transformed into the GMO basis from the previous cycle, the R matrix is formed, the canonicalization matrices are added, and a level shifting parameter, η, is added successively to all the diagonal terms of each shell. This final coupling matrix takes the following simple form

$$\begin{array}{ccccc}
\text{R shell} & \text{T shell} & \text{S shell} & \text{U shell} & \text{V shell} \\
F_R & \theta_{RT}(F_R-F_T) & \theta_{RS}(F_R-F_S) & \theta_{RU}(F_R-F_U) & \theta_{RV}F_R \\
 & F_T+\eta & \theta_{TS}(F_T-F_S) & \theta_{TU}(F_T-F_U) & \theta_{TV}F_T \\
 & & F_S+2\eta & \theta_{SU}(F_S-F_U) & \theta_{SV}F_S \\
 & & & F_U+\omega F_T+3\eta & \theta_{UV}F_U \\
 & & & & F_U+\omega F_T+4\eta
\end{array} \qquad (17)$$

This matrix is then diagonalized to form the new GMO's for the next cycle. The value of λ (equ.5) is determined variationally just prior to each orbital cycle. The energy (equ.8) is expressed explicitly in terms of λ.

$$E = E_0 + \lambda(1 - k_T k_U \lambda^2)^{\frac{1}{2}} E_1 + E_2 \qquad (18)$$

where

$$E_0 = 2H_R + 2H_T + H_S + 2J_{RR} + 4J_{RT} + 2J_{RS} + 2J_{TT} + 2J_{TS} \quad (19)$$
$$+\tfrac{1}{2}J_{SS} - K_{RR} - 2K_{RT} - K_{RS} - K_{TT} - K_{TS} - \tfrac{1}{2}K_{SS},$$

$$E_1 = 2K_{TU}, \quad (20)$$

and

$$\begin{aligned} E_2 = {} & 2k_T(H_U + 2J_{RU} + 2J_{TU} + J_{SU} - K_{RU} - K_{TU} - \tfrac{1}{2}K_{SU} + \tfrac{1}{2}K_{UU}) \\ & -2k_U(H_T + 2J_{RT} + 2J_{TT} + J_{ST} - K_{RT} - K_{TT} - \tfrac{1}{2}K_{ST} - \tfrac{1}{2}K_{TT}) \quad (21) \\ & -2(2J_{TU} - K_{TU}) \end{aligned}$$

A one-dimenional Newton-Raphson iteration is used to find the value of λ for which $(dE/d\lambda) = 0$.

Computational Details

The GMO calculation is usually started from a set of converged or partially converged RHF molecular orbitals. The doubly occupied orbitals of this calculation can be partitioned in several ways. One might place all the valence orbitals in the strongly occupied shell while keeping only the atomic cores doubly occupied. In this situation the weakly occupied orbitals should consist of all the antibonding counterparts of the valence shell. Thus, the occupied space in this problem (R,S,T, and U shells) could consist of all those MO's one could construct from a minimal atomic basis set on each atom or of the core, singly, and strongly occupied orbitals plus a weakly occupied correlating orbital for each strongly occupied orbital. Alternatively, one might only be interested in part of the molecule such as the metal-metal bond in a metal dimer complex or a particular metal-ligand bond. In either case the strongly and weakly occupied shells would be confined to the bond of interest and the remaining valence orbitals would be in the doubly occupied shell.

Initially, ω is set to 1.0, while η is taken to be between 2.0 and 1.0, and is reduced as the calculation converges to values between 0.8 and 0.4. One may begin the cycling with all the θ_{IJ} values equal to 1.0. However, one rapidly notices that matrix elements between the doubly and strongly occupied shells and those between the weakly occupied and virtual shell are not being reduced on each cycle. The reason for this is that the matrix elements in these two blocks are first order in λ, while all other blocks are zero order in λ. Thus, these matrix elements are inherently small and the coupling matrix procedure does not make sufficient changes in the MO's at each cycle to reduce these first derivatives. If one sets $\theta_{RT} = \theta_{UV} = 1.0/\lambda$, all off-diagonal blocks of the coupling matrix will be zero order in λ and the coupling matrix will converge. When λ is very small, we may multiply the value of θ_{RT} and θ_{UV} by 2.0 or 3.0, and for some open shell problems we may set all θ_{IS} values to 1.5 or 2.0.

Comparison to Second Order Methods

Since we have effectively only one configuration variable, λ, in our wavefunction we will confine this comparison to a second order determination of the orbitals. Given a trial set of orthonormal MO's, ϕ, we may construct an improved set ϕ' via a unitary transformation.

$$\phi' = \phi\, e^{\Delta} \tag{22}$$

where Δ is a skew-symmetic matrix. The optimum values of Δ may be determined by a full matrix Newton-Raphson procedure

$$A\,\Delta = -g \tag{23}$$

where the independent matrix elements of Δ are written as a vector ($\Delta_{ij} = \Delta_l$), g is the gradient or first derivative of the energy with respect to Δ ($g_l = dE/d\Delta_l$), and A is the Hessian or second derivative matrix ($d^2E/d\Delta_l d\Delta_m$). By solving these equations exactly, one would complete one iteration of a full second order procedure (neglecting the coupling between our orbitals and our configuration parameter, λ).

We can most easily compare this procedure with our coupling matrix technique by making several unjustified assumptions. First, we will expand e^{Δ} in a power series and retain only the first order term. Thus, equation (22) becomes

$$\phi' = \phi(1 + \Delta). \tag{24}$$

Second, we assume the Hessian matrix is diagonal so that equation (23) can be written as

$$\Delta_1 = -g_1/A_{11} \tag{25}$$

or returning to matrix notation for Δ we can write equation (25) as

$$\Delta_{ij} = \frac{-(dE/d\Delta_{ij})}{(d^2E/d\Delta_{ij}^2)} \tag{26}$$

If we now return to the coupling matrix and assume that the level shift parameters are large and/or that the off-diagonal terms are small, we may diagonalize the matrix by perturbation theory. The new vectors would then be expressed as in equation (24) and Δ would be

$$\Delta_{ij} = \frac{\theta_{IJ}(F_{ij}^{I} - F_{ij}^{J})}{(F_{jj}^{J} - F_{ii}^{I} + \eta)} \tag{27}$$

where we have used superscripts to denote the shell Fock operator and subscripts to indicate the particular matrix element. The term $4(F_{iJ}^{I} - F_{iJ}^{J})$ is the negative gradient and the remainder $\theta_{IJ}/4(F_{jj}^{J} - F_{ii}^{I} + \eta)$ represents an approximation the inverse of the diagonal term of the Hessian. The level-shifting parameter is used to assure a positive second derivative, and is chosen large enough to prevent the interchange of an orbital from one shell with that of another shell (to assure small Δ_{ij} values). The θ_{IJ} values, which are always greater than 1.0, can be used to accelerate the convergence by decreasing the estimate of the second derivative.

Although there is no way in which a coupling operator technique may be called truly second order, it is legitimate to view our approach as a heavily weighted gradient technique. The number of cycles of the coupling operator necessary to converge a typical case will be much larger than the number for a complete second-order technique.[4] However, since our effort per cycle is very small, less than three times that for a standard Roothaan RHF cycle, the coupling operator technique may be competitive with other methods. This is particularly true for situations like the GMO approach which have relatively few shells.

Configuration Interaction

We do not expect the GMO wavefunction itself to recover much correlation energy. Our objective in solving for this wavefunction is to obtain a set of optimized orbitals for use in a traditional CI calculation. The lowest level of CI, which appears reasonable, is all single and double excitations from orbitals in the T shell to those in the U shell. One may include higher excitations within this occupied space, such as full CI within the active space (T and U shells) or one may introduce excitations into the virtual space by allowing only a certain number of electrons into the V shell.

RESULTS

One of our implicit assumptions is that the form of the optimum orbitals does not depend very strongly on the CI coefficients in the MCSCF wavefunction. If this assumption is adequate, the GMO wavefunction can be derived from an all paired doubles MCSCF[5] with equal CI coefficients. Of course, this assumption will not be completely true in many important cases. However, for systems near their equilibrium geometry whose "true" wavefunction is dominated by a single configuration this assumption may be quite accurate.

One way to test part of this assumption is to examine how the energy obtained from the final CI depends on the value of λ which was used in the GMO wavefunction. Previous results for H_2 indicated that the final CI energy depended only weakly on the value of λ used to obtain the orbitals[6]. Similar results are also obtained for BH_3 where the $1a_1'$ is doubly occupied, the $2a_1'$ and $1e'$ are strongly occupied while the $1a_2''$, $3a_1'$, and $2e'$ are weakly occupied. A standard double-zeta Gaussian basis[7] with polarization functions ($\underline{d}$ on B, $\underline{p}$ on H)[8] was used in these calculations. Figure 1 shows two curves; the upper one is the energy of the GMO wavefunction as a function of λ. The RHF energy corresponds to the point at $\lambda=0$. The lower curve shows the energy obtained after a CI calculation (all singles and doubles from T shell to U shell) with the orbitals obtained from the GMO wavefunction. On this curve $\lambda=0$ corresponds to a singles and doubles CI calculation with the first virtual $1a_2''$, $3a_1'$ and $2e'$ from the RHF calculation. The remaining points then correspond to the use of the weakly occupied GMO's. The upper curve, as expected, depends quite strongly on λ, but the lower curve is quite flat beyond $\lambda=0.02$. One is forced to conclude that the final CI energy obtained from the GMO's does not depend very strongly on the λ value.

A more direct assessment of the orbitals may be made by comparing the GMO's with the natural orbitals (NO's) obtained by diagonalizing the first order density matrix from an all singles and doubles CI calculation. In this calculation all of the virtual orbitals, except the core correlating one were used in the active space. Prior to the large CI calculation the virtual GMO's were made canonical over the weakly-occupied Fock operator (ω was set at zero for one iteration). The percentage of each GMO in each NO is shown in Table II. Also shown are the occupation numbers of the NO's and the GMO eigenvalues with a letter designating the shell. As

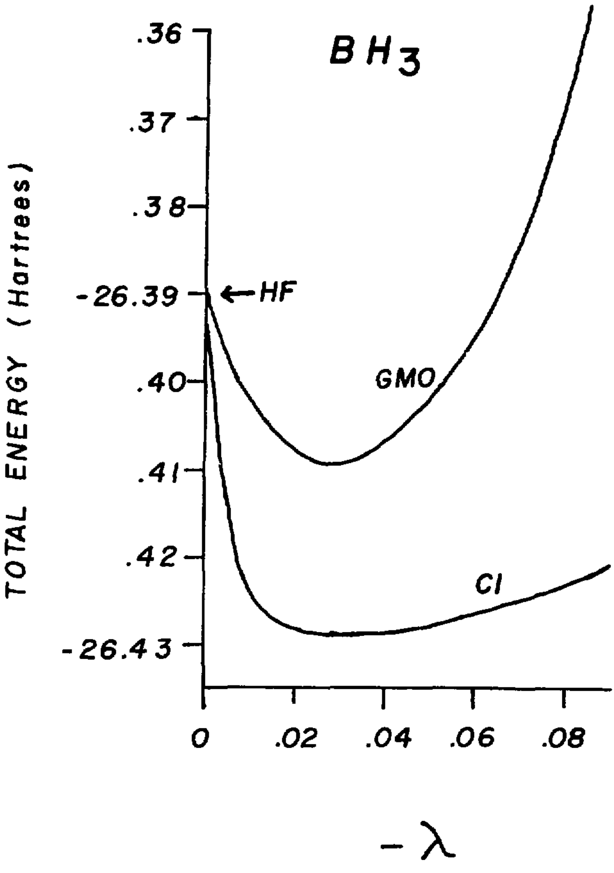

B H$_3$
.36
.37
.38
-26.39
.40
.41
.42
-26.43
TOTAL ENERGY (Hartrees)
← HF
GMO
CI
0
.02
.04
.06
.08
− λ

Fig. 1

TABLE II. Comparison of GMO's and NO's for BH_3

GMO (set) - eigenvalue	Natural orbital / occupations						
	$2a_1'$	$3a_1'$	$4a_1'$	$5a_1'$	$6a_1'$	$7a_1'$	$8a_1'$
	1.9732	0.0151	0.0052	0.0004	0.0002	0.0000	0.0000
$2a_1'$(t)-1.7101	100%	-	-	-	-	-	-
$3a_1'$(u)-0.0039	-	99%	1%	-	-	-	-
$4a_1'$(v)-0.0029	-	1%	99%	-	-	-	-
$5a_1'$(v)-0.0004	-	-	-	3%	29%	67%	1%
$6a_1'$(v)+0.0014	-	-	-	89%	9%	-	1%
$7a_1'$(v)+0.0044	-	-	-	7%	61%	32%	-
$8a_1'$(v)+0.0093	-	-	-	1%	-	1%	98%
	1e'	2e'	3e'	4e'	5e'	6e'	7e'
	1.9690	0.0141	0.0078	0.0016	0.0004	0.0003	0.0001
1e'(t)-1.4976	100%	-	-	-	-	-	-
2e'(u)-0.0043	-	100%	-	-	-	-	-
3e'(v)-0.0031	-	-	100%	-	-	-	-
4e'(v)-0.0005	-	-	-	4%	8%	87%	-
5e'(v)-0.0001	-	-	-	96%	-	4%	-
6e'(v)+0.0028	-	-	-	-	92%	8%	-
7e'(v)+0.0049	-	-	-	-	-	-	100%
	$1a_2''$	$2a_2''$	$3a_2''$				
	0.0098	0.0004	0.0000				
$1a_2''$(u)-0.0034	100%	-	-				
$2a_2''$(v)-0.0009	-	5%	95%				
$3a_2''$(v)+0.0017	-	95%	5%				
	1e"	2e"					
	0.0033	0.0003					
1e"(v)-0.0015	100%	-					
2e"(v)+0.0027	-	100%					
	$1a_2'$						
	0.0014						
$1a_2'$(v)+0.0004	100%						

expected the strongly occupied GMO's and NO's resemble each other and are not much different from the RHF orbitals. The weakly occupied GMO's, $1a''_2$, $3a'_1$, and $2e'$, have almost unit overlap with the most important natural orbitals. Even more amazing is the apparent similarity of the virtual GMO's (canonical over the operator for the weakly occupied) to the next most important NO's. In fact the GMO eigenvalues can be used to order the importance of these orbitals within each symmetry. Only as one approaches the more weakly occupied NO's does one notice that they begin to diverge from the GMO's. Similar results have been reported for H_2O and N_2.[9]

The importance of this similarity to the natural orbitals can be seen in Figure 2, where we have plotted the correlation energy obtained against the number of correlating orbitals used in the CI calculation. The three choices for the orbitals are RHF virtual (chosen by eigenvalue), GMO weakly occupied and virtual (chosen by eigenvalue but canonical over F_U) and the NO (chosen by occupation number). These are represented on Figure 2 as triangles, circles, and squares, respectively. The well known inferiority of the RHF virtual orbitals is amply illustrated (choosing the RHF orbitals in a different order makes no improvement). What is more striking is how closely the energy from the GMO's follows the energy from the NO's. The two calculations diverge only as the more weakly occupied orbitals are included.

The GMO technique has been used to calculate a number of properties such as ionization potentials of N_2 and H_2O,[9] excited state of CH_2,[10] dissociation energy of B_2H_6,[11] and the dissociation energy of a Mo-Mo triple bond.[12] Previous results on the spectroscopic constants for N_2[10] in a large Gaussian basis were only at a very low, inadequate level of CI. Table III presents the results for the equilibrium internuclear distance, dissociation energy, and stretching frequency. The basis set used in our cal-

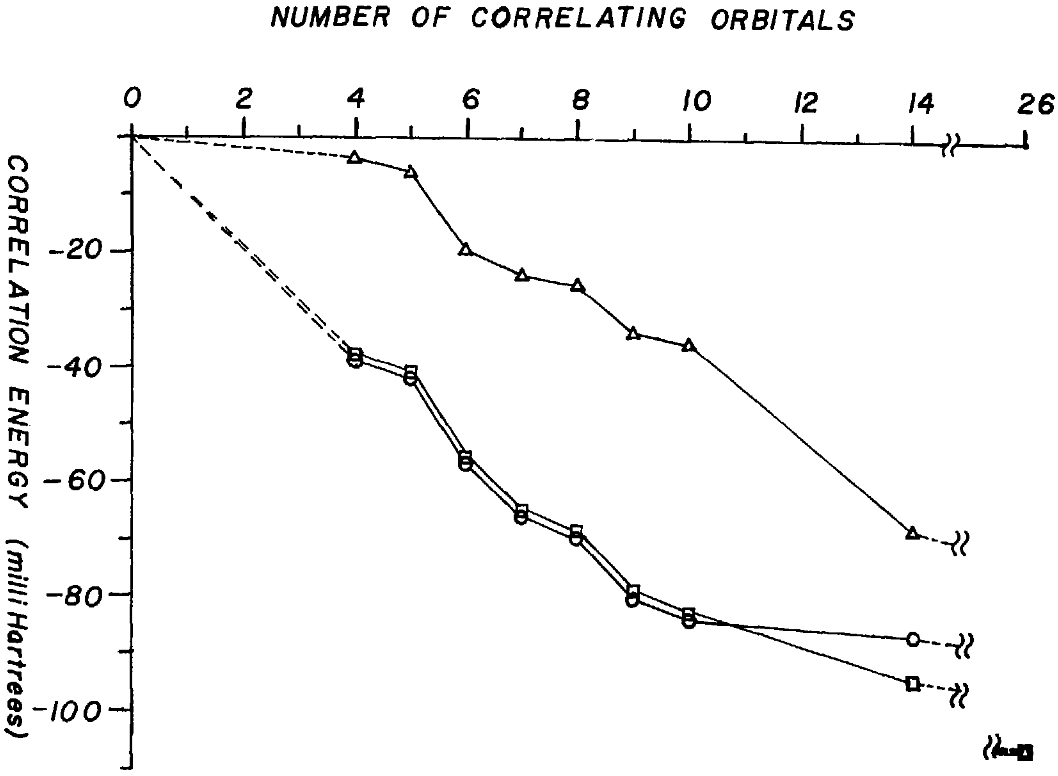

NUMBER OF CORRELATING ORBITALS
0
2
4
6
8
10
12
14
26
CORRELATION ENERGY (milli Hartrees)
-20
-40
-60
-80
-100

Fig. 2

TABLE III. Spectroscopic Constants for N_2

Method	R_e(Å)	D_e(eV)	ω_e(cm^{-1})
RHF	1.064	4.98	2769.
GMO	1.073	5.73	2645.
GMO-CI	1.103	8.82	2314.
EXP	1.098	9.91	2358.
RHF[a]	1.067	5.08	2757.
GVB(pp)[a]	1.094	7.05	2421.
GVB-CI[a]	1.106	8.93	2330.

[a] Reference 14

culation was a contracted (4s, 3p) Gaussian basis[7] augmented with a Slater 3d (exponent 2.81) expanded in two Gaussians.[13] For the GMO calculation the strongly occupied shell was $3\sigma_g$ and $1\pi_u$ (T shell), while the weakly occupied shell was $3\sigma_u$ and $1\pi_g$ (U shell). The spectroscopic constants were calculated from a five point (0.9, 1.0, 1.1, 1.2, 1.3) Dunham analysis. The improvement of the GMO over the HF is small, but when a full CI calculation is performed within the T and U shells (GMO-CI) the improvement is dramatic. The results at the GMO-CI level are nearly identical to those obtained from the GVB orbitals followed by CI in the GVB valence space.[14] The small observed differences can be attributed to a slight difference in the basis set, which is apparent in the two RHF calculation.

DISCUSSION

The examples provided above and described elsewhere[6, 9-12] show that the use of the GMO procedure to define a primary orbital space produces orbitals and subsequent energies very close to those of more sophisticated techniques. The procdeure is extremely efficient since it requires the construction of only two additional Fock matrices beyond the number required for the RHF calculation. In all cases we have tried, the coupling operator as described above has converged to the required accuracy.

However, the GMO approach does have some obvious limitations. The present formulation will not be equally accurate over the entire potential surface ; in fact, it will be very poor as one approaches dissociation. The method will work best near the equilibrium geometry, when the wavefunction is dominated by a single configuration. Although we have no direct evidence, we believe that it will be more accurate for compact molecules, where all of the orbitals of the T and U shells are in the same region of geometric space, than for extended systems. In many of our inorganic applications to coordination complexes and cluster compounds, we are dealing with just such compact molecules. For most of the larger systems we are interested only in some part of the molecule such as the metal-metal bond or a particular metal-ligand bond. In these cases the GMO procedure conveniently localizes the portion of interest in the T and U shell, while the remainder of the molecule is described by the doubly occupied R shell.

For small molecules, more accurate MCSCF techniques are available.[15] Our procedure is not intended to compete with these methods for accuracy in determining the primary orbital space. However, in some situations it might be useful, after determining the primary space by a complete MCSCF, to adapt the GMO approach to determine a secondary orbital space. This technique could be particularly important in very large basis sets since it would allow one to discard the less important virtual space.

ACKNOWLEDGEMENT

Acknowledgement is made to the National Science Foundation (Grant No. CHE 79-20993) for support of our research. This research was supported in part by the National Resource for Computation in Chemistry under a grant from the National Science Foundation and the Basic Energy Sciences Division of the United States Department of Energy under Contract No. W-7405-ENG-48.

REFERENCES

1. D.R. Hartree, *Proc. Cambridge Phil. Soc.*, 24, 89, 111 (1928); V. Fock, *Z. Physik*, 61, 126 (1930); C.C.J. Roothaan, *Rev. Mod. Phys.*, 23, 69 (1951), 32, 179 (1960).

2. I. Shavitt, "Methods of Electronic Structure Theory"; H.F. Schaefer, III, Ed; Plenum Press (New York, 1977).

3. K. Hirao, *J. Chem. Phys.* 60, 3215 (1974).

4. E. Dalgaard and P. Jørgensen, *J. Chem. Phys.* 69, 3833 (1978); D. Yeager and P. Jørgensen, *J. Chem. Phys.* 71, 755 (1979).

5. E. Clementi and A. Veillard, *J. Chem. Phys.* 44, 3050 (1966).

6. M.B. Hall, *Chem. Phys. Letters*, 61, 461 (1979).

7. T.H. Dunning, Jr. *J. Chem. Phys.* 53, 2823 (1970); S. Huzinaga, *ibid.* 42, 1293 (1965).

8. T. H. Dunning, Jr. and P.J. Hay, "Methods of Electronic Structure Theory"; H.F. Schaefer, III, Ed.; Plenum Press (New York, 1977).

9. M.B. Hall, *Intern. J. Quantum Chem. Symp.*, 13, 195 (1979).

10. M.B. Hall, *Intern. J. Quantum Chem.* 14, 613 (1978).

11. T.E. Taylor and M.B. Hall, *J. Am. Chem. Soc.* in press.

12. M.B. Hall, *J. Am. Chem. Soc.* 102, 2104 (1980).

13. R.F.Stewart, *J. Chem. Phys*. **52**, 431 (1970).

14. T.H. Dunning, Jr., D.C. Cartwright, W.J. Hunt, P.J. Hay, and F.W. Bobrowicz, *J. Chem. Phys.* **64**, 4755 (1976).

15. See other authors in this publication especially W.A. Goddard,III, B. Liu, B. Roos, C.C.J. Roothaan, K. Ruedenberg, A.C. Wahl and D.L. Yeager.

MCSCF OPTIMIZATION THROUGH ITERATIVE CI CALCULATIONS IN THE SINGLE EXCITATION SPACE AND MCSCF WAVE FUNCTIONS IN THE FULL REACTION SPACE

Klaus Ruedenberg
Ames Laboratory* and Department of Chemistry, Iowa State University, Ames, IA 50011

1. ORBITAL SPACE AND CONFIGURATION SPACE

A multiconfiguration ("MC") N-electron wave function is a vector in a function space which we shall call the MC space. It is spanned by a basis of N-electron functions Ψ_K known as configuration state functions ("CSF's"). There exists considerable diversity as regards the construction of the CSF's from orbital products and spin functions. Ultimately every CSF can be expressed as a fixed linear combination of Slater determinants. The simplest choice is to have each CSF equal one Slater determinant. A reasonable sophisticated choice is to choose the linear combinations such that each CSF is an eigenfunction of $\hat{S}^2$, S_z and also belongs to the same irreducible representation as the state which is to be calculated. In our work, we often find an intermediate choice convenient, namely spin adapted antisymmetrized products of orbitals (SAAP's).

In any event I am assuming that the orbitals from which the CSF's are constructed do belong to irreducible representations ("irreps") of a molecular symmetry group, i.e., they are of the form $\phi_i^{\nu\mu}$, where ν denotes the irreducible representation and μ labels the various degenerate partner functions which form a multidimensional irrep. I furthermore assume that the various orbitals belonging to a given value of ν and μ are orthogonal, so that we have in fact

$$\langle\phi_i^{\nu\mu}|\phi_{i'}^{\nu'\mu'}\rangle = \delta_{ii'}\delta_{\nu\nu'}\delta_{\mu\mu'} \ . \qquad (1.1)$$

The CSF's are then functionals of the orbitals, i.e.,

$$\Psi_K = \Psi_K(\cdots\phi_i^{\nu\mu}\cdots) \qquad (1.2)$$

and I furthermore assume that they are constructed in such a manner that the orthogonalities (1.1) entail the CSF orthogonalities

$$\langle\Psi_K|\Psi_L\rangle = \delta_{KL} \ . \qquad (1.3)$$

SAAP's have this property.

In the multiconfiguration self-consistent-field (MCSCF) approach the orbitals $\phi_i^{\nu\mu}$ are not explicitly given to start with, but are to be optimally determined through calculation. Consequently, the assumption of a certain configurational basis, such as that given by the Ψ_K of Eq. (1.2) does not yet completely specify these configurations, it merely defines a certain formal structure of the configuration space in terms of MO's yet to be determined. These MOS will be called the configuration generating orbitals (CGO's). The essential characteristics of the structure of the configuration space are the number of CGO's and the type of configurations formulated in terms of them. A particular CGO may occur in one, in several, or in all configurations. However, since it does not have to occur in all configurations, it is apparent that the number of CGO's, M say, is in general larger than the minimum number of orbitals that can occur in any one configuration, namely, $\frac{1}{2}N$, where N is the number of electrons. In actual molecular problems the CGOS are nearly always divided into two groups: (i) a set of "generalized core" or "closed-shell" orbitals, all of which are doubly occupied in every configuration, and (ii) a set of "active" CGO's whose occupation numbers are less than two in at least one configuration. Suppose there are M" closed CGO's and M' active CGO's. If it is intended to free all electrons outside the closed shell completely from the "straight jacket of double occupancy," then one intuitively expects the number of active orbitals not to be smaller than the number of electrons occupying them which implies

$$M'+M''=M, \quad M'\geq N-2M'', \quad M\geq\tfrac{1}{2}(N+M') \qquad (1.4)$$

The choice $M=\frac{1}{2}(N+M')$ corresponds to having, on the average, exactly one active orbital available for each open-shell electron, a case which may be called the "extended independent particle model" for the open-shell part. On the other hand, the largest configuration space that is possible for a given number of active CGO's is obtained when the configuration space basis corresponds to all CSF's that can be constructed from the active CGO's with the closed CGO's always remaining doubly occupied. This we shall call the full configuration space generated by the active CGO's (FAS). An important feature of this full configuration space is that it is invariant against any nonsingular, in particular, orthogonal transformation among the CGO's. If the active orbitals are those which can be expected to describe a chemical reaction, we call such a space the full reaction space (FRS).

In order to optimally determine the CGO's, it is presumed that they are expressed as linear expansions in terms of a set of basis orbitals. Since these are the functions through which, ultimately, the wave function becomes a specific function in space, we shall call them the quantitative basis orbitals (QBO's). They can be molecular orbitals or atomic orbitals. In the latter case they usually are predetermined fixed superpositions ("contractions") of primitive atomic orbitals (PAO's), the ultimate building blocks of most molecular calculations. It is often convenient to generate from them a basis set of orthonormal symmetry-adapted QBO's (QBSO's) which we denote by

$$\hat{\phi}_1^{\nu\mu}, \quad \hat{\phi}_2^{\nu\mu}, \quad \hat{\phi}_3^{\nu\mu}, \ \ldots \qquad (1.5)$$

Clearly the CGO's quoted in Eq. (1.1) can be expressed in terms of the QBSO's by the expansions

$$\phi_j^{\nu\mu} = \sum_r \hat{\phi}_r^{\nu\mu} F_{rj}^{\nu\mu} \qquad (1.6)$$

In order that there be any variational freedom, there clearly must be more QBO's than CGO's. Let L be the total number of QBO's and let M be the total number of CGO's. More specifically let L_ν, M_ν be the total number of $\hat{\phi}_r^{\nu\mu}$ and $\phi_j^{\nu\mu}$ in the νth irrep, i.e., $r=1,2,\ldots,L_\nu$, $j=1,2,\ldots,M_\nu$ so that

$$L = \sum_{\nu=1}^{I} D_\nu L_\nu, \quad M = \sum_{\nu=1}^{I} D_\nu M_\nu \qquad (1.7)$$

if there are I irreps, the νth irrep having the dimension D_ν. Then we have necessarily

$$L_\nu \geq M_\nu , \quad L \geq M , \qquad (1.8)$$

with the inequality being the usual situation. The sum in Eq. (1.6) thus runs from 1 to L_ν and the transformations $F^{\nu\mu}_{rj}$ are rectangular isometric transformations. By complementing these isometric matrices in some arbitrary manner to square orthogonal matrices, we are introducing (L-M) additional MO's $\phi^{\nu\mu}_{j}$ that are orthogonal to the CGO's. These will be called virtual MO's. We shall also use the terms occupied and unoccupied MO's for the configuration generating and the virtual MO's, respectively.

The multiconfigurational approximation to the exact wave function, which is our goal, is thus of the form

$$\Psi = \sum_K C_K \Psi_K , \qquad (1.9)$$

where the Ψ_K are the configurations introduced in Eq. (1.2) and

$$\sum_K C_K^2 = 1 . \qquad (1.10)$$

In order to obtain the best possible approximation of this kind, two optimizations are required in the MCSCF approach: (i) The configurational expansion coefficients C_K, occurring in Eq. (1.9) must be optimized (MC-CI calculations) and (ii) the MO's must be optimized by optimizing the orbital expansion coefficients $F^{\nu\mu}_{rj}$ occurring in Eq. (1.6) (MC-SCF calculation). The latter optimization is the more troublesome one and requires successive iterations. Consequently, there exist M occupied MO's and (L-M) virtual MO's at each stage of this iterative process, and at each stage the occupied orbitals are improved by mixing among each other and by admixtures from the virtual space.

Clearly the MC-CI calculation yields as many eigenstates as there are CSF's and the n-th MC eigenvalue is therefore an upper bound to the n-th true state of the symmetry in question. In principle then, the orbitals can be optimized for any one of the states, and it stands to reason that the optimal orbitals will be different for different states. This circumstance entails a number of subtle consequences. For one thing, the optimal MCSCF approximations of different states are constructed from orbitals which are not only mutually (*i.e.*, between different states) non-orthogonal, but also span slightly different orbital spaces. For another, the different MCSCF states themselves will be slightly non-orthogonal to each other. Both shortcomings represent no prohibitive problems however and, these non-orthogonalities notwithstanding, the various energies nonetheless represent upper bounds.

More serious is the fact that, when the orbitals of an upper state are iteratively improved by lowering the energy of that state, then this MC eigenstate may change places energywise with a lower state (usually both states can be readily recognized by their MC mixing coefficients C_K). If this occurs for example upon lowering the energy of the second state, then *continued* lowering of the energy of the *second* state will obviously result in an oscillation between the two states. Clearly, in such a case the goal of unrestricted orbital optimization is incompatible with the requirement that the second MC root be an upper bound. Such difficulties arise from the fact that the MC space does not have sufficient degrees of freedom to accommodate both the optimal groundstate and the optimal excited state so that, when the second state is optimized the approximation to the lower state becomes so poor that its energy rises above what started out to be the second MC root. The best remedy is, of course, to give the MC space the necessary flexibility by (i) including in its basis a sufficient number of CSF's (which is straightforward) and (ii) including a sufficient number of configuration generating orbitals (which may be awkward), so that both states can be adequately represented in the MC space. An alternative option is to explicitly enforce the orthogonality to the lower states while optimizing the orbitals of an excited state. In any event, these problems must be dealt with regardless of the procedure which is being employed to achieve the MCSCF optimization.

2. VARIATIONAL CONDITIONS

A. Coefficient Variations and Orbital Variations

It is our aim to find the optimal approximation Ψ, as determined by the variation principle

$$\delta\langle\Psi|H|\Psi\rangle = 2\langle\delta\Psi|H|\Psi\rangle = 0 \qquad (2.1a)$$

while

$$\langle\Psi|\Psi\rangle \equiv 1 \qquad (2.1b)$$

where Ψ is a wave function of the type formulated in Eq. (1.9). The variation of Ψ can be written

$$\delta\Psi = \delta_{coef}\Psi + \delta_{orb}\Psi , \qquad (2.2)$$

where δ_{coef} denotes the variation of the coefficients C_K and δ_{orb} denotes the variation of the orbitals $\phi^{\nu\mu}_{j}$. For a finite number of configurations, as envisaged here, these two variations are independent and the variational Equation (2.1a) is equivalent to the two separate conditions [1]

$$\langle\delta_{coef}\Psi|H|\Psi\rangle = 0 \qquad (2.3a)$$

$$\langle\delta_{orb}\Psi|H|\Psi\rangle = 0 \qquad (2.3b)$$

In the coefficient Equation (2.3a), the orbitals are fixed and optimal and, hence, the configurations Ψ_K are fixed. Consequently, this equation yields, in familiar manner, the eigenvalue equation

$$\sum_L \langle K|H|L\rangle C_L = EC_K \qquad (2.4a)$$

where the matrix elements

$$\langle K|H|L\rangle = \langle \Psi_K|H|\Psi_L\rangle \tag{2.4b}$$

are presumed to be calculated with the configurations constructed from the optimal orbitals.

In the orbital Equation (2.3b), the coefficients C_K are presumed to be fixed and optimal. The orbital variation of Ψ is therefore given by

$$\delta_{orb}\Psi = \sum_K C_K \delta\Psi_K \tag{2.5}$$

and, since every Ψ_K is multilinear in the orbitals $\phi_i^{\nu\mu}$, the variations $\delta\Psi_K$ can be expressed in terms of the orbiral variations $\delta\phi_i^{\nu\mu}$ by means of the following sum of "single orbital variations"

$$\delta\Psi_K = \sum_\nu^{I} \sum_\mu^{D_\nu} \sum_{i=1}^{M_\nu} \Psi_K(\cdots\phi_i^{\nu\mu} + \delta\phi_i^{\nu\mu}\cdots) \tag{2.6}$$

where M_ν is the number of occupied orbitals for the irreducible representation ν. The orbital variations in turn, can be expanded in terms of the full L-dimensional orbital basis. Choosing the *optimal* occupied and corresponding virtual orbitals as that basis, we therefore write

$$\delta\phi_i^{\nu\mu} = \sum_{j=1}^{L_\nu} x_{ij}^\nu \phi_j^{\nu\mu} \tag{2.7}$$

where the x_{ij}^ν are infinitesimal, but not necessarily independent, coefficients. *Note that they are independent of μ in order that the group-theoretical-transformation properties remain intact during the orbital variations.* Substitution of the expansion (2.7) into Eqs. (2.5) and (2.6) yields

$$\delta_{orb}\Psi = \sum_\nu \sum_i^{M_\nu} \sum_j^{L_\nu} x_{ij}^\nu \Psi(\nu i \to \nu j) \tag{2.8}$$

where

$$\Psi(\nu i \to \nu j) = \sum_K C_K \sum_\mu \Psi_K(\cdots\phi_i^{\nu\mu} \to \phi_j^{\nu\mu}\cdots) \tag{2.9}$$

Since the orbitals form an orthonormal set, the *fixed* coefficients C_K satisfy Eq. (1.10), and the side condition (2.1b) will be satisfied if the orbital variations preserve the orthonormality conditions (1.1). To achieve this, the *varied* orbitals $(\phi_i^{\nu\mu} + \delta\phi_i^{\nu\mu})$ must be related to the *optimal* orbitals $\phi_i^{\nu\mu}$ by an orthogonal transformation which implies that the infinitesimal coefficients x_{ij}^ν in Eq. (2.7) are antisymmetric in their subscripts:

$$x_{ij}^\nu = -x_{ji}^\nu \tag{2.10}$$

and that the set of coefficients x_{ij}^ν given by $i<j$ represents a possible set of independent infinitesimal parameter variationa. In view of this result, the orbital varistion of Ψ of Eq. (2.8) can be written in the form

$$\delta_{orb}\Psi = \sum_\nu \sum_{i=1}^{M_\nu} \sum_{j=i+1}^{L_\nu} x_{ij}^\nu \Psi_{ij}^\nu \tag{2.11}$$

where the functions Ψ_{ij}^ν are defined by

$$\Psi_{ij}^\nu = \Psi(\nu i \to \nu j) - \Psi(\nu j \to \nu i) \tag{2.12}$$

for

$j = (i+1), (i+2), (i+3), \ldots, M_\nu =$ an occupied orbital $> i$

and

$$\Psi_{ij}^\nu = \Psi(\nu i \to \nu j) \tag{2.13}$$

for

$j = (M_\nu+1), (M_\nu+2), \ldots, L_\nu =$ a virtual orbital

We shall call these functions "singly excited MC functions" or "*MC single excitations*," and abbreviate them as *MCSX's*. *In general, they are neither normalized nor mutually orthogonal nor even necessarily linearly independent.* *They are*, however, orthogonal to the MC function Ψ itself, *i.e.*,

$$\langle \Psi_{ij}^\nu | \Psi\rangle = 0, \quad \text{for all } j>i \ . \tag{2.14}$$

B. Generalized BLB Theorem and Single Excitation Space

Insertion of $\delta_{orb}\Psi$ of Eq. (2.11) into the variational condition (2.3b) yields

$$\sum_\nu \sum_{i=1}^{M_\nu} \sum_{j=i+1}^{L_\nu} \langle \Psi_{ij}^\nu |H|\Psi\rangle x_{ij}^\nu = 0$$

and, since the x_{ij}^ν occurring in these equations are independent variational parameters, we obtain

$$\langle \Psi_{ij}^\nu |H|\Psi\rangle = 0, \quad \text{for all } i<j \tag{2.15}$$

The analogous variational conditions for the case that the Ψ_K in Eq. (1.9) are simple Slater determinants were first derived by Levy and Berthier [2]. Since they are generalizations of an earlier theorem by Brillouin [3] for Hartree-Fock wave functions, we shall refer to these equations as the BLB theorem. Our *equations differ from the original BLB equations not only in that they are for general CSF's but, moreover, in that the singly excited MC functions Ψ_{ij}^ν embody summations over the subspecies index μ shown in Eq. (2.9), as a consequence of the μ independence of the x's in* Eq. (2.7). Since the Ψ_{ij}^ν are not always mutually independent, the same holds true for Eqs. (2.15).

Equation (2.15) can be given an additional interpretation if one introduces the single-excitation (SX) space associated with a given MC function Ψ. This SX space is defined as that function space which is spanned by Ψ itself and all single excitation functions (MCSX's) Ψ_{ij}^ν derived from Ψ according to Eq. (2.12). Eq. (2.15) implies then that the hamiltonian matrix in the SX space is block diagonal, the element $\langle\Psi|H|\Psi\rangle$ being one block and all elements $\langle \Psi_{ij}^\nu |H| \Psi_{i'j'}^{\nu'}\rangle$

forming the other block. Since, moreover, Ψ is orthogonal to all Ψ^ν_{ij} [(Eq. (2.14)], it follows from this block diagonality that the MC function Ψ itself is one of the eigenfunctions of the CI problem in the SX space, when the orbitals are optimal for Ψ, regardless whether the Ψ^ν_{ij} are linearly independent of each other or not, and that Ψ cannot be improved by admixing any of the MCSX's Ψ^ν_{ij}. If Ψ is the groundstate then all Ψ^ν_{ij} will lie higher in energy than Ψ itself and it can therefore be inferred that Ψ will be the eigenfunction with the lowest energy eigenvalue in the SX/CI problem. If Ψ is the nth excited state in a given symmetry, i.e., if its energy is the n-th root of the MC-CI problem, then it is possible that the Ψ^ν_{ij}, generated from this Ψ contain sufficient admixtures from the lower MC eigenstates so that Ψ can become any one of the n lowest eigenfunctions of the SX-CI problem.

The foregoing derivations can be generalized to yield the variational conditions for those orbitals which simultaneously optimize several states [4].

3. OPTIMIZATION PROCEDURE

A. Iterative Orbital Improvement for Fixed Coefficients

Let us assume for the moment that the optimal MC coefficients C_K of Eq. (1.9) for the state in question are known, and let us consider the problem of optimizing the orbitals $\phi^{\nu\mu}_i$ without changing these fixed values C_K. If the orbitals are not optimal, but close to being so, then the solution of the CI problem in the SX space will yield one SX-CI eigenfunction

$$\Psi_{sx}=a_{00}\Psi+\sum_{\nu}\sum_{i=1}^{M_\nu}\sum_{j=i+1}^{L_\nu} a^\nu_{ij}\Psi^\nu_{ij} \qquad (3.1)$$

which is characterized by having a_{00} close to unity and all $a^\nu_{ij}<<1$. This eigenfunction is readily found by using an iterative procedure that is effective in determining a single eigenfunction which is dominated by one component [5]. If other SX-CI eigenvalues lie too close for numerical comfort, the desired root can be isolated by adding an appropriate negative constant to the $\langle\Psi|H|\Psi\rangle$ diagonal element [6]. Comparing the expansion (3.1) with the expansion of Eq. (2.11) one realizes that the orbital variation coefficients x^ν_{ij} are related to the SX-CI coefficients by the equations

$$x^\nu_{ij}=a^\nu_{ij} \quad \text{for } i<j$$
$$x^\nu_{ij}=-a^\nu_{ji} \quad \text{for } i>j \qquad (3.2)$$
$$(x^\nu_{ii})^2=1-\sum_{j(\neq i)}(x^\nu_{ij})^2 \ .$$

From this result it can be inferred that it must be possible to deduce improved approximations to the optimal MC orbitals from the CI expansion of the appropriate eigenfunction of the SX-CI problem. Furthermore, iterative execution of such improvements can be expected to lead to the optimal MC orbitals.

The essential step which has to be formulated in such an iterative scheme is the explicit method for deducing the improved orbitals from the solution of the SX-CL problem, and a variety of formalisms have been suggested to this end. We shall return to this question further below.

B. Complete MCSCF Optimization

In general the MC coefficients C_K are not known in advance, of course, and have to be variationally determined at the same time as the orbitals are determined. In general the coefficients C_K are determined as the expansion coefficients for the n-th eigenvalue of the MC-CI problem formulated in Eq. (2.4a). The matrix elements in this equation depend however on the molecular orbitals and, hence, an iterative procedure is required for the simultaneous determination of MC orbitals and MC coefficients. The most effective way to accomplish this is to solve the eigenvalue equation (2.4a) every time an improved set of molecular orbitals has been found. Hence, each iteration of the overall procedure consists of two parts: first an improvement of the MC coefficients, and then, using these improved coefficients, an improvement of the molecular orbitals. A very rough schematic outline of a total MCSCF program is then as follows:

(i) Input

Data and integrals pertaining to quantitative basis orbitals. Initial molecular orbitals in terms of quantitative basis orbitals. Spatial and spin specifications of CSF's. Specification of desired eigenvalue in MC space.

(ii) Initiating Calculations

Orthonormalization of initial MO's if needed. Generation of integrals between reference MO's. Generation of symmetry and spin information. Generation of data relevant to the SX space and integrals between the MCSX's.

(iii) Iteration

Calculation of integrals needed between current MO's from the integrals between the reference MO's. Calculation of matrix elements in the MC space and solution of the MC-CI problem. Determination of a basis in the SX space. Calculation of matrix elements in the SX space and determination of appropriate energy eigenvector for SX-CI problem. Construction of improved MO's.

(iv) Convergence Test

Convergence can be defined in terms of several criteria:

(a) Smallness of the matrix elements of Eq. (2.15). This guarantees that there is no more change in the MC expansion coefficients C_K within the stipulated accuracy.

(b) Smallness of the difference between the energy of the MC function and the (lower) energy of the SX-CI function. This difference is nearly always monotonically decreasing.

(c) Smallness of the difference between the energies of MC functions of two successive iterations.

(d) Smallness of the off-diagonal elements of the matrix which generates the new (improved) molecular orbitals from the old orbitals.

For a satisfactory MCSCF function all these criteria should be satisfied to a given accuracy and must therefore be monitored.

C. The Orbital Improvement Step

For the implementation of the outlined procedure the non-orthogonality and, possibly, linear dependence of the single-excitation functions Ψ^{ν}_{ij} is a complication which has to be taken seriously. Concretely it implies that one will first construct an orthonormal basis, in the SX space before performing the SX-CI calculation. Let us call these basis functions SX^{ν}_{τ}. The transformation from the Ψ^{ν}_{ij} to the SX^{ν}_{τ} is non-orthogonal and, in case of linear dependence among the Ψ^{ν}_{ij}, a rectangular matrix.

The most obvious choice for the orbital improvement step is to use directly the Eqs. (3.2) to calculate the x^{ν}_{ij} from the a^{ν}_{ij} and this is in fact the procedure of Grein [7]. The orbitals generated by adding the correction of Eq. (2.7) to the original orbitals are of course not strictly orthogonal and have to be reorthonormalized in some arbitrary manner. Since the a^{ν}_{ij} are the expansion coefficients in terms of the non-orthogonal Ψ^{ν}_{ij} basis, they have to be reconstructed from the CI coefficients in terms of the orthogonal SX^{ν}_{τ} basis. If a linear dependence exists between the Ψ^{ν}_{ij}, then the a^{ν}_{ij} are obviously not unique and neither are therefore the improved orbitals.

The ALIS formalism, which was developed by our group, is based on the following conjecture: The natural orbitals of the appropriate eigenfunction of the SX-CI problem, generated from an MC wavefunction with non-optimal orbitals, are closer approximations to the natural orbitals of the optimal MC wavefunction than the natural orbitals of the original MC wavefunction. This premise has led us to the following procedure [4]. First the natural orbitals $\varphi^{\nu\mu}_{n}$ of the initial MC wavefunction are determined in terms of the initial configuration generating orbitals (CGO's), i.e.,

$$\varphi^{\nu\mu}_{n} = \sum_{i}^{M_\nu} \phi^{\nu\mu}_{i} U^{\nu}_{in} , \qquad (3.3)$$

in order to establish the $M_\nu \times M_\nu$ orthogonal transformation matrices U^{ν}. Then the natural orbitals $\tilde{\varphi}^{\nu\mu}_{n}$ of the SX-CI eigenfunction are determined in terms of the quantitative basis orbitals. Finally, the inverse of Eq. (3.3) is used in order to calculate the new CGO's $\tilde{\phi}^{\nu\mu}_{i}$ from the new natural orbitals $\tilde{\varphi}^{\nu\mu}_{n}$, i.e.,

$$\tilde{\phi}^{\nu\mu}_{i} = \sum_{i=1}^{M_\nu} U^{\nu}_{in} \tilde{\varphi}^{\nu\mu}_{n} . \qquad (3.4)$$

The calculation of both types of natural orbitals is based on the totally symmetric projection of the density matrix, so that the natural orbitals are symmetry adapted. In this procedure, no additional reorthogonalization is required. Moreover, the coefficients a^{ν}_{ij} are not needed: the calculation is carried through directly in terms of the orthogonal SX^{ν}_{τ} basis and is unaffected by possible linear dependencies among the single excitation functions Ψ^{ν}_{ij}. The only point which requires careful attention to detail is the matching of the new natural orbitals to the old natural orbitals. This matching has been successfully achieved on the basis of a careful analysis described in Ref. [4].

Direct use of the natural orbitals of the SX-CI eigenfunction as improved CGO's, as used by Yarkony et al. [8], is only possible for those special MC function for which the matrix U in Eq. (3.3) is the identity or, on the other hand, in the case that the MCSCF procedure is carried out in the full active MC space (which can become very large) since, in this case, the MC function (but not the expansion coefficients) is invariant against orthogonal transformations among the CGO's. This has been pursued by B. Roos [9].

A more involved way of using the density kernel to deduce improved orbitals from the SX-CI eigenfunction has been suggested by Ruttink and van Lenthe [10].

D. Remark

The list of references given here is far from complete. As regards the past work, Ref. [4] contains a more exhaustive bibliography. The current work is essentially associated with one or the other of the participants in this workshop. I therefore feel that my illustrious colleagues will be far better able to describe and quote their contributions than I would be.

4. THE ALIS IMPLEMENTATION

The MCSCF formalism developed by this research group has been implemented within a system of computer programs called ALIS (denoting "Ames Laboratory, Iowa State University") for performing ab-initio quantum chemical calculations. It is portable, dynamic, open-ended and continually being improved and extended. It contains a flexible variety of atomic orbital basis options (currently of all atoms up to argon) and requires a minimal amount of input information, notably the reference CSF's of the MC space, expansions of the initial orbitals, specifications of the closed and the active orbitals, specifications of the orbitals which are to be optimized and of the orbitals which are to be kept frozen, specification of the MC root to be sought and specification of the convergence criteria.

The atomic orbital bases preferred by this research group are general contractions of even-tempered Gaussian primitives [11] because of the almost unlimited flexibility they offer in the choice of the number of primitives as well as in the number of quantitative basis orbitals, and because of the close relationships which have been found to exist between the principal basis orbitals of this type and the molecular MCSCF orbitals [12]. The efficient evaluation of the integrals between such general contractions is accomplished by an adaptation of Raffenetti's BIGGMOLI program [10].

The essentials of the iterative MCSCF process have been outlined in Ref. [4]. The CSF's are expressed as superpositions of spin-adapted antisymmetrized products (SAAP's) and the matrix element construction is based on the symmetric group. A formula file is being used, and particular care is given to the various large matrix transformations which the mathematical process requires.

A fair number of detailed problems have to be properly attended to in order to obtain a program which is sturdy enough to hold up under various types of mathematical and computational stresses. ALIS has proved to be stable under a considerable variety of conditions. Experience has shown, however, that it is not particularly desirable for an MCSCF program to be able to optimize any pathological MC-type ansatz. It is a useful feature that it breaks down when the investigator chooses an MC space which is physically unreasonable for the problem whose solution is sought.

The ALIS system, including a detailed documentation, is available through the NRCC library.

5. THE METHOD OF THE FULL OPTIMIZED REACTION SPACE (FORS)

As has been mentioned repeatedly, the choice of the MC space is an essential aspect for useful MCSCF calculations. For a number of years we have used in our work the Full Optimized Reaction Space and found it a very successful approach. Not only does it always yield physically reasonable results, but it also seems to satisfy the upper bound requirement. There are five essential elements to this method: (i) The choice of the reaction orbitals; (ii) The use of the full active space; (iii) The use of Raffenetti-type quantitative basis orbitals; (iv) The calculation of the initial guess from a minimal set of quantitative basis orbitals; (v) The use of systematic macroiterations to optimize in very large full reaction spaces.

A. Reaction Orbitals and Full Reaction Space

The full reaction space (FRS) is the full configuration space of the active configuration-generating orbitals, when the latter are chosen as reaction orbitals. These are those orbitals which for physical reasons are principally involved in the electronic rearrangements that occur in a chemical reaction. They are essentially a set of formal conceptual minimal basis set atomic orbitals on the participating atoms or symmetry adapted superpositions of them. In fact, since the full configuration space is being used, only the total number of reaction orbitals in each irrep is required for defining the full reaction space. When the optimal expansions of the reaction orbitals in terms of the (much larger) set of quantitative basis orbitals have been found through the MCSCF process, then we have the Full Optimized Reaction Space (FORS) [13].

B. Raffenetti-type Quantitative Basis Orbitals and Minimal Basis Set Approximation

These quantitative basis orbitals (QBO's) are defined in terms of primitive bases (of arbitrary size) of Gaussian atomic orbitals. The QBO's contributed by a specific atom are of three kinds: Principal QBO's, diffuse QBO's and polarization QBO's. The principal QBO's are the Hartree-Fock SCF AO's of the free atom; the diffuse QBO's are the most diffuse single primitive AO's occurring in the expansions of the principal QBO's; the polarization QBO's are standard polarization primitives. The advantages of such bases have been discussed elsewhere [11,12]. They do require integral programs for general contractions such as BIGCMOLI [11].

Within the present context it is pertinent that the principal QBO's invariably dominate the expansions of the reaction orbitals after optimization in the FRS. This circumstance not only has useful implications for the conceptual interpretation but, moreover, gives rise to an extremely simple and effective construction of the initial guess for the MCSCF calculation in the FRS [12]. In principle, this initial orbital set is found by making an MCSCF calculation in the FRS with the quantitative basis orbitals reduced to the minimal set of the principal QBO's. In point of fact, since the number of principal QBO's is equal to the number of reaction orbitals, no orbital optimization is necessary and a straight CI calculation in the FRS spanned by the principal QBO's suffices. The natural orbitals derived from the resulting CI wavefunction invariably turn out to be excellent approximations to the natural orbitals of the MCSCF function in the FRS when all QBO's are included. Using them as starting orbitals, one finds much faster convergence for the MCSCF procedure than one finds when using SCF MO's.

C. Optimization in Large Full Reaction Spaces through Macroiterations

When the FRS becomes very large, then one finds that only a relatively small number of dominant configurations determine the orbital optimization. This is so even though a much larger number of configurations make non-trivial contributions to the energy. In this context it is pertinent that, because of the invariance of the FRS against transformations among the configuration generating orbitals, there exists considerable freedom in the orbital choice. An effective standard choice is to use the natural orbitals of the wavefunction itself, since it leads to a rapid convergence of the wavefunction in terms of the natural-orbital-based CSF's. It is therefore well suited to identify a set of dominant configurationa which determine the optimal orbitals.

This circumstance is the basis for a relatively simple optimization method for very large full reaction spaces, which consists of a sequence of macroiterations. In each macroiteration the current orbitals are used to generate the CSF's which span the corresponding full reaction space. A straight CI calculation in this FRS then yields the current approximation to the wavefunction in the FRS. Next the natural orbitals of this approximation are determined as orthogonal transformations of the current MO's, and the wavefunction is now expressed in terms of CSF's generated from the natural orbitals. From this expansion a set of dominant configurations is aelected which are then used to obtain improved orbitals through an MCSCF calculation. The number of dominant configurations is increased from one macroiteratio [illegible] another. Convergence occurs after two or th [illegible] ations. The first macroiteration [illegible] natural orbitals from the CI calcula [illegible] terms of the principal QBO's as discussed in the preceding section.

A number of applications of the described method were discussed in the NRCC report on the Workshop on Post-Hartree-Fock Methods in Quantum Chemistry [14].

ACKNOWLEDGMENTS

In the development of the methods outlined here. I have had the able cooperation of Drs. L. M. Cheung, K. R. Sundberg, M. Gilbert. S. T. Elbert and Mr. M. Schmidt. This work was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, WPAS-KC-01-03-01-2. *Operated for the U.S. Department of Energy by Iowa State University under contract No. W-7405-Eng-82.

REFERENCES

1. J. Frenkel, Wave Mechanics, Advanced General Theory (Clarendon, Oxford, 1934), p. 460.

2. B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968); 3, 247 (1969); B. Levy, Chem. Phys. Lett. 4, 17 (1969); Int. J. Quantum Chem. 4, 297 (1970).

3. L. Brillouin, Act. Sci. Ind. No. 159 (Herrmann, Paris, 1934).

4. K. Ruedenberg, L. M. Cheung, S. T. Elbert, Int. J. Quantum Chem. 16, 1069 (1979).

5. I. Shavitt, J. Comp. Phys. 6, 124 (1970).

6. F. Grein and A. Banerjee, Int. J. Quantum Chem. Symp. 9, 147 (1975); A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054 (1977).

7. F. Grein and T. C. Chang, Chem. Phys. Lett. 12, 44 (1971).

8. D. R. Yarkony, H. F. Schaefer, and C. F. Bender, J. Chem. Phys. 64, 981 (1976); D. R. Yarkony, J. Chem. Phys. 66, 2045 (1977); R. R. Lucchese and D. R. Yarkony, J. Chem. Phys. 68, 2696 (1978).

9. B. O. Roos, P. R. Taylor, E. M. Siegbahn, Chem. Phys. 48, 157 (1980).

10. P. J. A. Ruttink and J. H. van Lenthe, Theor. Chim. Acata 44, 97 (1977).

11. R. C. Raffenetti, J. Chem. Phys. 58, 4452 (1973); see also M. W. Schmidt and K. Ruedenberg, J. Chem. Phys. 71, 3951 (1979).

12. M. W. Schmidt and K. Ruedenberg, in course of publication.

13. K. Ruedenberg and K. R. Sundberg, p. 505 of "Quantum Science" (J. L. Calais, O. Goscinski, J. Linderberg, Y. Öhrn, eds.) Plenum Press, N.Y., 1976; L. M. Cheung, K. R. Sundberg, K. Ruedenberg, Int. J. Quantum Chem. 16, 1103 (1979).

14. K. Ruedenberg, p. 46 of "Post-Hartree-Fock: Configuration Interaction", Report on 1978 Workshop, NRCC, Lawrence Berkeley Laboratory.

THE COMPLETE ACTIVE SPACE SCF PROGRAM SYSTEMS

Björn O. Roos, Chemical Centre, Lund, Sweden
Per E.M. Siegbahn, Institute of Theoretical Physics
Stockholm, Sweden

SUMMARY

A review is given of the complete active space (CAS) SCF method with special emphasis on computational aspects. The CASSCF wave function is formed from a complete distribution of a number of active electrons in a set of active orbitals, which constitute a subset of the total occupied space. No selection of configurations is made, and the wave function can subsequently comprise a large number of terms: The largest case considered to date is a calculation on the $^3A''$ state on the HNO molecule, where the wave function consisted of 10416 spin- and space-adapted configuration, obtained by distributing 12 electrons among 10 active orbitals. To be able to solve the orbital optimization problem using such large CI expansions, a density matrix formalism is used, and the CI problem is solved using the recent developments of direct CI methods.

Two approaches to the orbital optimization problem have been developed and will be discussed: the Newton-Raphson scheme and the super-CI method. Convergence is normally achieved in 5-10 iterations with the Newton-Raphson method while the super-CI method usually needs more iterations (typically 10-20). Fluctuations in the necessary number of interations is, however, large in both methods, and pathological cases can occur. They usually depend on an inappropriate choice of active orbitals.

THE CASSCF METHOD

The complete active space SCF method[1-4] represents a new way of looking at the MCSCF problem. Traditionally MCSCF calculations are based on wave functions of a very limited size, comprising a small number of pre-selected configurations. This can lead to difficulties in computations of energy surfaces (e.g. for chemical reactions), since it rarely occurs that all configurations of importance are known in advance for the entire surface. Normally, however, it is much easier to decide in advance upon an appropriate choice of occupied orbitals. This is the basic idea of the CASSCF approach. The orbital space is divided into three parts: the inactive subspace (labelled i,j,k,l,..), the active subspace (labelled t,u,v,x,...) and the secondary subspace (labelled a,b,c,d,...). The inactive orbitals are assumed to be doubly occupied in all configurations. The remaining electrons occupy the active orbitals, and the CASSCF wave functions comprises all CSF's which can be generated by distributing the active electrons among the active orbitals in all possible ways, consistent with a given overall spin and spatial symmetry of the wave function. There is thus no selection of configurations, but the wave function is complete in the active orbital subspace. Even though the number of active orbitals normally is rather limited, such a CI expansion quickly becomes very long. With recently developed CI methods, however, even expansions of the order of 10^4 terms can be handled without any particular computational problems (see below). An important aspect of the CASSCF method is that the length of the CI expansion only affects the time required to solve the secular problem, and not the orbital optimization part. By formulating the problem in terms of first- and second-order density matrices - in the small active subspace - all explicit reference to the individual CSF's is avoided.

THE SECULAR PROBLEM

With the large CI expansions often encountered in CASSCF calculations it is important to have an efficient procedure for solving the secular problem. The special structure of the CI expansion, which is complete in a small active orbital space, will on the one hand allow certain simplifications which are not possible with other general expansions. On the other hand there are high demands on the method in terms of being able to treat configurations with many open shells. For this purpose the graphical unitary group approach[5,6], which is the procedure adopted in the CASSCF CI section, is particularly well suited.

In the graphical unitary group approach Gelfand states are used as the spin adapted configuration basis. The direct CI coupling coefficients for the one electron integrals are then obtained as a segment product with one segment value for each orbital within the corresponding generator range. A similar simple formula was also recently obtained for the two electron integrals[7,8]. These type of formulas are particularly efficient in the cases of interest here, namely for generation of a large number of coefficients in a small number of orbitals.

In MCSCF procedures the same symbolic formulas are used a large number of times; in every CI iteration in each orbital iteration and for all the points on the potential energy surface. It is therefore important to simplify as much as possible the use of the formula list rather than its construction. Any diagram evaluation or index evaluation during the process of solving the secular problem should for example be avoided. For large CI-expansions it is further important to reduce the size of the formula list as much as possible. The information in the formula list is usually organized with one coupling coefficient for each index, where the index describes the two interacting configurations. By instead grouping together matrix elements which have the same coupling coefficient for the same integral the formula list can be reduced to essentially half the size and is therefore preferrable in large cases. Further reductions in the size of the formula list would still sometimes be important but do not seem to be possible without drastic changes in the method or much increased labour during the diagonalization.

In the CASCI part, iterative algorithms will be used, which as the essential step require the operation of the Hamiltonian on one or a set of trial vectors in each iteration[9]. It should be pointed out that although the CI problem has to be solved in every MCSCF iteration the total number of CI iterations is usually only about twice as many as required in one separate CI calculation. This is simply achieved by starting with the CI trial vector from the iteration before. In cases with very long CI expansions where the CI part dominates the total time is therefore only twice the time as in a CI calculation without orbital optimization.

As indicated above the only information transferred from the CI part to the orbital optimization part is in terms of the first and second order density matrices in the small active space. The number of such matrix elements is usually much smaller than the number of CI coefficients and they are also considerably more convenient to use in the orbital optimization. The density matrices are easily constructed in one pass through the coupling coefficients with the converged CI vector. In cases where there is a strong coupling between the variation of the CI coefficients and the variation of the orbitals, transition density matrices between the root of interest and closely lying roots should also be constructed at the end of the CI part. From these matrices the coupling elements are easily calculated in the orbital optimization[3].

THE NEWTON-RAPHSON SCHEME

In the orbital optimization part of the Newton-Raphson approach we look for a unitary transformation of the initial orbital set which minimizes the energy. This unitary transformation is conveniently described by an exponential parametrization where the parameter matrix is symmetric[10]. In order to obtain a quadratically convergent orbital optimization scheme both first and second derivatives of the energy with respect to the orbital rotation parameters are needed. Simple formulas for these derivatives are obtained in the form of expectation values of commutators between the Hamiltonian and the generators of the unitary group[11,12]. The disappearence at convergence of the first derivatives is through these formulas expressed as the Brillonin-Levy-Berthier (BLB) theorem[13]. An important point which is noticed in the explicit expressions for the first derivatives in the CASSCF method is that integrals with at most one index in the large secondary space are required, see equation (14) in Ref. 2. This means that a very limited integral transformation is required to obtain these gradients and this is the basis for fast, although not quadratically convergent methods. To obtain the exact expressions for the second derivatives integrals with two secondary indices are needed, which require a substantially longer transformation time. With the information from the CI-part in the form of first and second order density matrices in the active space together with the required transformed integrals all derivatives required to set up the Newton-Raphson equation system can be constructed using the explicit formulas given in Ref. 3. The dimension of the N-R equation system is set by the number of parameters which are of three types; parameters describing rotations between inactive and active orbitals, between inactive and secondary orbitals and finally between active and secondary orbitals. All first derivatives with respect to other types of rotation parameters are identically zero and these parameters are therefore discarded in the equation system. This removes all problems concerned with singularities in the Hessian matrix. A problem which can still occur in the N-R procedure far away from convergence is that the Hessian matrix is not positive definite in which case the obtained rotations would not lead to a minimum for the energy. Before solving the N-R equations all negative and small positive eigenvalues of the Hessian matrix are therefore shifted to a smallest allowed positive value following essentially the procedure described in Ref. 14.

The dominating step in a CASSCF iteration is normally the integral transformation, at least in cases where the CI expansion is not very long. As indicated above most of the transformation time goes into calculating the integrals with two external indices which are only needed for the Hessian matrix. Since it has been found that once the Hessian matrix has become positive definite it usually changes only marginally, a strategy, which has been used successfully to reduce the computation time, is to keep the same Hessian matrix in several iterations. In the cases where comparisons have been made this procedure increased the number of iterations by usually not more than one and gave an overall significant saving in computation time[3].

We should finally comment on the fact that although the N-R approach is a quadratically convergent orbital optimization procedure this process does not always lead to fast convergence. The reasons are twofold. First, the parameter surface may be far from quadratic and second there can be significant coupling between orbital and CI coefficient rotations. The latter effect has been approximately corrected for in Ref. 3 with only minor improvement in convergence rate. On the other hand drastic improvement in convergence has been obtained by changing the structure of the active orbital space which apparently sometimes may lead to a more quadratic parameter surface.

THE CASSCF SUPER-CI METHOD

The Super-CI method[15,16] can be regarded as an approximate form of the Newton-Raphson scheme. Instead of obtaining the orbital transformation matrix by solving a set of linear equations, one obtains this matrix by solving a secular equation, which corresponds to a "super-CI" expansion including the ground state |0> plus all single excitations (SX-states) obtained by operating on |0> with single generators of the unitary group. The interactions between the SX-states and the ground state is then given exactly as the BLB matrix elements (see equations (6) in ref. 4). The necessary and sufficient conditions for a converged solution - the BLB conditions - are consequently in the super-CI method equivalent to a solution of the SX-state problem, with all coefficients equal to zero.

In Grein's original work[15,16], after solution of the corresponding secular problem the super-CI expansion coefficients were used directly to set up the orbital transformation matrix. In the modified version developed by Ruedenberg et al.[17] improved orbitals are obtained instead as the natural orbitals of the super-CI wave function. This method is especially attractive in the CASSCF case, as |0> here is invariant to a unitary transformation among the active orbitals. It is thus possible to use the NO's (selecting those with the *highest occupation numbers)* as trial orbitals for the next iteration, without any further manipulation. As a result the converged solution will automatically be the natural expansion of the wave function |0>.

The BLB matrix elements can be expressed directly as matrix elements of a MCSCF Fock operator[18] (equation (9) in ref. 4). The matrix elements between the SX-states are, however, more complicated and involve the first- second- and third-order reduced density matrices for the wave function |0>[1,2]. Even if these density matrices are defined over the small active subspace only, the direct calculation of the third-order density matrix is cumbersome, and the size of this matrix easily becomes prohibitely large. Also, the SX-state matrix elements contain two-electron integrals with two indices outside the active subspace (as is the case also for the Hessian in the NR scheme). An exact solution of the super-CI problem therefore leads to a computationally more complicated procedure than the NR scheme.

The super-CI method then does not seem to be competitive with the quadratically convergent NR approach. As has been shown[1,4] it is, however, possible to approximate the density-matrix formulated super-CI method such that these difficulties dissappear to a large extent. The idea is to replace the Hamiltonian operator with an approximate one-particle Hamiltonian, when calculating the matrix elements between the SX-states. These matrix elements are then also obtained in terms of matrix elements of the MCSCF operator, and first- and second-order density matrix elements with indices in the small active subspace only[4]. The calculation of the SX Hamiltonian matrix then becomes almost trivial. The corresponding secular problem is easily solved using the direct CI method in combination with a Davidson algoritm[9]. In a direct CI method the timing is proportional to the number of cycles needed to converge. This number depends *critically* on the intitial guess of the eigenvector. Since all coefficients in the super-CI wave function approaches zero as convergence is attained, the number of necessary cycles decrease markedly as the CASSCF calculation proceeds. As a consequence the time used in the orbital optimization step decreases with the number of interations performed. This is also the case for the solution of the CASCI secular problem, as pointed out in a preceeding section. The time needed for the integral transformation of the two-electron integrals, however, remains the same, and therefore becomes a more and more dominant part of the calculation.

The super-CI method cannot be expected to have as good convergence properties as the Newton-Raphson approach. It is, however, considerably faster, especially since fewer molecular two-electron integrals are needed, and therefore often competes favourably with the NR scheme in calculations of energy surfaces were trial vectors obtained from nearby points can be used. Typically between 10 and 20 iterations are needed for convergence to 10^{-6} a.u in the energy and 10^{-4} in the BLB matrix elements, if SCF vectors are used as a starting guess of the molecular orbitals. If CASSCF orbitals of a similar calculation (nearby point on an energy surface, another electronic state, etc.) are used as trial orbitals, the number of iteration is often much less than ten.

CONVERGENCE PROBLEMS IN CASSCF

As has already been pointed out in the preceeding sections convergence is normally satisfactory both in the NR and super-CI versions of the CASSCF program system. Pathological case occur, however, now and then and it might be of interest to give some examples, which shows that these convergence problem freouently are of a physical origin instead of being a result of the numberical procedure adopted.

In a recent calculation[4] of the potential curves for the $^1\Sigma_g$, $^3\Sigma_u$ and $^3\Pi_g$ states of the N_2 molecule, convergence (to the previously mentioned thresholds) as obtained in 4-6 iterations for all points on the potential curves except in a small region between 3.5 and 4.0 a.u., where 15 iterations were needed. The super-CI approach was used and the active subspace comprised nine orbitals, including the weakly occupied $2\pi_u$ orbital. The reason for the problems was a drastic change in the structure of this orbitals. At distances shorter than 3.5 a.u. the $2\pi_u$ orbitals has mainly p-character, describing radial correlation effects, but at longer distances it becomes dominantly 3d (angular correlation of the 2p shell in the nitrogen atom). In the region 3.5-4.0 a.u. the two effects compete and convergence is slowed down considerably. An obvious solution to the problem is to include both correlation effects at all distances, that is, to add a third π_u orbital subspace (or to remove also $2\pi_u$).

Another example of convergence difficulties occurred in a calculation on the $^1\Sigma^+$ state of CN^+ at large internuclear distance. With the basis set used, and the active subspace chosen, a number of electronic states (including $C+N^+$ and C^++N) happened to be almost degenerate. The program could consequently not decide about the ground state, and the choice could vary from iteration to iteration. As a result large fluctuations of the orbital coefficients occurred in every iteration and a converged result was actually never reached. The experimental energy difference between $C+N^+$ and

and C^++N is around 0.3 eV, so there is obviously room for an improvement of the calculation.

It may also happen that a calculation converges but not to the expected (desired) wave function. If the active subspace includes weakly occupied orbitals which describe dynamical (pair) correlation effects, there is sometimes a choice between different pair correlations, and the result may depend on the starting quess for these orbitals. Such a situation has occurred in calculations on molecules containing transition metal atoms (examples are NiH and CuF_2) where weakly occupied orbitals can choose to correlate 3d electron pairs on the metal atom or electron pairs on the ligands, depending on the starting guess.

The difficulties illustrated by the examples given above are of a physical origin and cannot be handled by a improvement of the orbital optimization procedure. They are examples where the energy hypersurface has several close lying local minima, a situation which is often encountered in MCSCF calculations. One solution to these problems would be to increase the active subspace to include dynamical correlation effects in a balanced way (i.e. both radial and angular correlation in N_2), but this route is severely limited, since the number of configuations increases drastically with the number of active orbitals. Another solution is to add a fourth subspace (the "correlation space") which is allowed only two holes or two particles. This space is then used to describe pair correlation effects. In this way the primary (occupied) orbital space can be enlarged considerably and dynamical correlation effects can be included to a large extent. Work along these lines are presently in progress[19]

REFERENCES

1. B.O. Roos, P.R. Taylor and P.E.M. Siegbahn, Chem.Phys. 48, 157 (1980).
2. P.E.M. Siegbahn, A. Heiberg, B.O. Roos and B. Levy, Physica Scripta 21, 323 (1980).
3. P.E.M. Siegbahn, J. Almlöf, A. Heiberg and B.O. Roos, Submitted for publication.
4. B.O. Roos, Int.J.Quantum Chem. S14, XXX (1980).
5. J. Paldus, in Theoretical Chemistry: Advances and Perspectives (ed. H. Eyring and D. Henderson), vol. 2, p 131, Academic Press, New York, 1976.
6. I. Shavitt, Intern.J. Quantum Chem. S11, 131 (1977) and S12, 5 (1978).
7. I. Shavitt and J. Paldus, to be published.
8. J. Paldus and M.J. Boyle, Physica Scripta 21, 295 (1980).
9. E.R. Davidson, J.Comput.Phys., 17, 87 (1975).
10. B. Levy, Int.J. Quantum Chem. 4, 297 (1970).
11. B. Levy, thesis CNRS No AO 5271, Paris 1971.
12. E. Dalgaard and P. Jörgensen, J.Chem.Phys. 69, 3833 (1978).
13. B. Levy and G. Berthier, Int.J.Quantum Chem., 2, 307 (1968).
14. D.L. Yeager and P. Jörgensen, to be published.
15. F. Grein and T. Chang, Chem.Phys. Letters 12, 44 (1971).
16. F. Grein and A. Banerjee, Intern.J.Quantum Chem. S9 (1975) 147; J.Chem.Phys. 66, 1054 (1977).
17. K. Ruedenberg, L.M. Cheung and S.T. Elbert, Int.J. Quantum, Chem. 14, 1069 (1979).
18. J. Hinze, J. Chem.Phys., 591 6424 (1973).
19. J. Almlöf, A. Heiberg, P.E.M. Siegbahn and B.O. Roos, work in progress.

A MCSCF PROGRAM EMPLOYING THE GENERALIZED BRILLOUIN THEOREM

Joop H. van Lenthe*
School of Chemistry, Bristol University, Bristol BS8 1TS, U.K.
and
Paul J.A. Ruttink
Vakgroep Theoretische Chemie, Rijks Universiteit Utrecht, Utrecht, The Netherlands.

SUMMARY

A program system is described that optimises a multiconfiguration wavefunction in an iterative manner. Each iteration consists of two stages; firstly the determination of the coefficients of the Configuration State Functions (CSF's) in the reference function and secondly the performance of a CI calculation using the reference function and single excitations from it (Brillouin state Interaction (BI) calculation). Two methods of obtaining a transformation to better orbitals from the result of this BI calculation are implemented, i.e. the first order method[2] and the density matrix method[3-5] (employing natural orbitals).

The program is in practice able to handle up to 60 CSF's and 30 orbitals. It can handle an additional 30 orbitals using perturbation theory.

Convergence of both orbital transformation methods and an exponential method[6] are equally good and linear (10-15 iterations for 10^{-8} a.u. accuracy in the energy), so the first order method is to be preferred on account of its simplicity and general applicability. The four index transformation[7] in each iteration is the major time consuming step, so iteration times are proportional to n^5 where n is the number of orbitals.

Applications include calculations of the barrier for the LiH + H $\rightleftarrows$ Li + H_2 reaction,[7] potential energy curves for O_2^+[4] and CuO[8], the $^1A'$ excited state of HO_2[4,9], dipole moment of cyclopropene[10] and of the cis-trans energy difference[4] and charge distribution[11] of formic acid.

METHOD

The iterative optimisation consists of the following steps:
1) The reference function Ψ_0 is determined for the starting orbital set $\{\phi\}$ by a small CI calculation.

$$\Psi_o = \sum_k a_k \phi_k \qquad (1)$$

2) All relevant singly excited states (Brillouin states) are formed.

$$\Psi_{ij} = \Psi_o(i\rightarrow j) - \Psi_o(j\rightarrow i) \qquad (2)$$

These Brillouin states are neither normalised nor mutually orthogonal. By solving the general eigenvalue problem for the set $\{\Psi_o, \Psi_{ij}\}$ we obtain the coefficients b in the Brillouin state Interaction function

$$\Psi_{BI} = b_o\Psi_o + \sum_{i\neq j} b_{ij}\Psi_{ij} \qquad (3)$$

3) The information contained in Ψ_{BI} is used to obtain the transformation matrix $\underset{\sim}{T}$ which transforms the orbital basis to more optimal orbitals

$$\underset{\sim}{\phi'} = \underset{\sim}{\phi}\underset{\sim}{T} \qquad (4)$$

The following two methods are implemented:

A. First order method[2] $T_{ii} = b_0$ (5)
$T_{ij} = b_{ji} = -b_{ij}$

Since the orbitals ϕ' are only orthonormal to first order, an additional orthonormalisation is needed, taking care that the most important (occupied) orbitals are least spoiled, e.g. using Schmidt.

B. Density matrix method[3-5] $\underset{\sim}{T} = \underset{\sim}{T}_{BI}\underset{\sim}{T}_0^+$ (6)

where $\underset{\sim}{T}_{BI}$ and $\underset{\sim}{T}_0$ are the matrices of the coefficients in the basis $\{\phi\}$ of the natural orbitals of Ψ_{BI} and Ψ_0 respectively. This automatically yields a unitary transformation matrix.
We also tried another way to obtain a unitary T directly,

C. Exponential method[6] $\underset{\sim}{T} = e^{-\underset{\sim}{b}}$ (7)
where $\underset{\sim}{b}$ *is the matrix of the Brillouin state* coefficients b_{ij}.

4) Replace $\{\phi\}$ by $\{\phi'\}$ (and perform a four index transformation) and start again at 1).

The b_{ij}'s are used to monitor convergence. Usually a threshold value of 10^{-4} (for the maximum absolute value of a b_{ij}) is used, which corresponds to about 10^{-7}-10^{-8} a.u. in the energy.

Perturbation theory

Instead of computing all b_{ij}'s by solving the full set of secular equations, one may compute the smaller ones using second order perturbation theory[12]

$$b_{ij} = \frac{<\Psi_o|H|\Psi_{ij}>}{<\Psi_o|H|\Psi_o><\Psi_{ij}|\Psi_{ij}> - <\Psi_{ij}|H|\Psi_{ij}>} \qquad (8)$$

Treating all excitations to a certain (virtual) orbital this way has the following advantages:

1. The dimension of the set $\{\Psi_o, \Psi_{ij}\}$ needed in the BI calculation and the total number of CSF's is reduced.

2. No integrals between the "perturbation orbital" and other virtual orbitals are required.

3. All perturbation orbitals are equivalent, so only a short symbolic matrix element list is needed.

If too large a proportion of the virtual orbitals are treated using perturbation theory convergence is badly affected (cf. ref. 2).

---

*Dutch Ramsay Memorial Fellow

According to our experience a good rule of thumb is to treat about half the virtual orbitals in this way.

MCSCF-CI

The set of CSF's, from which the singly excited states (Brillouin states) are formed, may be a useful set to perform a CI calculation with, using the converged MCSCF orbitals.

Reasons for this are:

1. The H matrix for this set has already been computed in step 2 of the iteration process so the CI only involves an additional diagonalisation.

2. If Ψ_o contains the most important double excitations, the MCSCF-CI probably contains the most important single double and triple excitations.

3. The resulting CI function has the same invariance properties under orbital transformation as the reference function Ψ_o, so the result is independent of the starting orbitals[3].

The MCSCF-CI has proven to be quite effective for O_2^+ in a VB model[4] and for LiH_2[3].

PROGRAM SYSTEM[4]

The MCSCF program consists of three programs written for the Cyber 73-28 of the Utrecht University computer center.

I. FINDH, a general four index transformation program[7], which contains various options to localise[13], reorder and orthonormalise[14] the orbitals, in order to provide a proper integral input file for the iteration process. A "frozen core"[15] option, which effectively decreases the dimension of the integral file, is included.

II. MCCODE, a program that produces a formula tape, containing all information, which is independent of the actual value of the integrals, e.g. a symbolic matrix element list.

The spin functions are constructed according to the genealogical (Branching Diagram) method[16]. The order of the orbitals can be chosen freely to yield maximal interpretability. This has also been exploited to simplify the transformation from CSF's to Brillouin states, by requiring that each single excitation working on a CSF produces just one CSF.

Extensive use is made of bit manipulations (CDC-fortran) to keep the tape as short as possible; e.g. storing coefficient and position of an integral in one 60 bit word and storing the occupation scheme of a CSF in one word.

III. MCSCF, the actual iterative program that contains its own specialised four index transformation routines. By a special ordering of orbitals and CSF's, both integral file and H matrix list can be processed sequentially.

Use of symmetry can be made to eliminate unnecessary single excitations; equivalence restrictions can be applied by a contraction of the Brillouin states[17] and a fixed linear combination of CSF's in the reference function can be prescribed.

In order to accelerate convergence extrapolation may be applied. The method of Sack[18] seems to yield the best results, however, usually only if all b_{ij}'s are less than $\sim 5.10^{-3}$

Scope

In the following we present the limitations of the present MCSCF program. Limits due to fixed dimensions or available space in a computer word are underlined; other limits are approximate.

# orbitals	< 60
# variationally treated orbitals	< $\underline{30}$
# CSF's in Ψ_o	< 60 ($\underline{255}$)
# singly excited (Brillouin) states	< 300
# CSF's (total)	< $\underline{4095}$
# symmetries	< $\underline{6}$

Equivalence restrictions only for two dimensional irreducible representations. Singlets up to quartets can be treated ($0 \leq S < 3/2$)

# singly occupied orbitals in a CSF in Ψ_o:	singlets/ triplets < $\underline{4}$
	doublets/ quartets < $\underline{5}$

PERFORMANCE

Convergence

In Figure 1 we give a plot of the convergence behaviour of all three orbital transformation methods in a test calculation on H_2O using 5 CSF's[4].

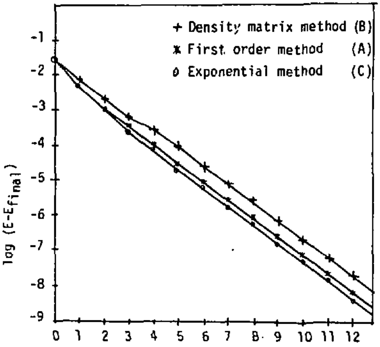


Figure 1. Convergence of various GBT methods. Logarithms of Energy-final energy (a.u.) against no. of iterations.

Convergence in all cases is linear and quite fast (~ 12 iterations for an energy constant to 10^{-8} Hartree). There is little difference between the methods. The first order method (A), being the simplest of the three, seems therefore to be the best choice.

The density matrix method (B) suffers from the defect that it cannot optimise functions with singly occupied orbitals like

$$\frac{1}{\sqrt{2}} \left(|\phi_a\bar{\phi}_b| - |\bar{\phi}_a\phi_b|\right) \qquad (9)$$

in the two orbital function space. The density matrices for Ψ_{BI} and Ψ_0 are both diagonal, so the matrices T_{BI} and T_0 (eq. 6) are unit matrices, whatever the value of b_{ij} is.

Timing.

Some timings (in seconds) for O_2^{+4} (1s shells frozen, effectively 24 orbitals) are presented below:

State	MCCODE	MCSCF (per iteration)	
		4 index	total
$^2\Pi_g$ (22 CSF's)	295	43	66
$^4\Pi_u$ (9 csf's)	78	47	53

The four index transformation is the time determining step in the MCSCF calculations (a HF calculation on $O_2\ ^3\Sigma_g^-$ in the same basis takes 16 seconds per iteration). Note the rather large time needed to produce the formula tape (MCCODE).

APPLICATIONS

Applications fall into two broad categories, viz. potential energy curves for small molecules keeping inner shells frozen and rectifying Hartree Fock deficiencies for rather larger molecules.

In the final category first calculations on O_2^{+4} and CuO[8] using typically 20-30 orbitals (effectively) and 5-30 CSF's.

The second category contains calculations on the 'A' state of HO_2^+ and on the charge distributions in cyclopropene[10] and formic acid[4]. For HO_2^+ a 2 CSF function was needed to solve a near degeneracy problem for the first excited state. The calculations on formic acid were only done in two molecular conformations in an attempt to obtain a better cis-trans energy difference. Early calculations[4] used a maximum of 38 orbitals and 2-8 CSF's, where 2 CSF's appeared to be enough for a significant improvement in the energy-difference. Since, however, dipole and quadrupole moments were far from satisfactory larger MCSCF and CI calculations are presently undertaken[11] using 10-40 CSF's and up to 50 orbitals.

REFERENCES

1. B. Levy and G. Berthier, Int.J.Quantum Chem. 2 (1968) 307.
2. F. Grein and T.C. Chang, Chem.Phys.Lett. 12 (1971) 44
   A. Banerjee and F. Grein, Int.J.Quantum.Chem. 10 (1976) 123.
3. P.J.A. Ruttink and J.H. v.Lenthe, Theoret.Chim. Acta (Berl.) 44 (1977) 97.
4. J.H. v.Lenthe, Ph.D. thesis (Utrecht, 1979): available upon request.
5. K. Ruedenberg, L.M. Cheung and S.T. Elbert, Int.J.Quantum Chem. 16 (1979) 1069.
6. B. Levy, Int.J.Quantum Chem. 4 (1970) 297.
7. M. Yoshimine, IBM Technical Report RJ 555, San José (1969)
   G.H.F. Diercksen, Theoret.Chim.Acta (Berl.) 33 (1974) 1.
8. D.H.W. den Boer and E.W. Kaleveld, Chem.Phys. Lett. 69 (1980) 389.
9. J.H. v.Lenthe and P.J.A. Ruttink, Chem.Phys. Lett. 56 (1978) 20.
10. J.H. v.Lenthe and J.L.M. Smits, Recl.Trav.Chim. Pays-Bas. 99 (1980) 130.
11. J.v.d. Linden and F.B. v. Duijneveldt, private communication (Utrecht 1980).
12. J. Hinze, J.Chem.Phys. 59 (1973) 6424.
13. V. Magnasco and A. Perico, J.Chem.Phys. 47 (1967) 971.
14. P.O. Löwdin, Adv.Quantum Chem. 5 (1970) 185.
15. R.P. Hosteny, T.H. Dunning Jr., R.R. Gilman, A. Pipano and I. Shavitt, J.Chem.Phys. 62 (1975) 4764.
16. M. Kotani, A. Amemiya, E. Ishiguro and T. Kimura, Tables of Molecular integrals, Maruzen, Tokyo (1963).
    P.J.A. Ruttink, Theoret.Chim.Acta (Berl.) 36 (1975) 289.
17. D.R. Yarkony, H.F. Schaefer and C.F. Bender, J.Chem.Phys. 64 (1976) 981.
18. R.A. Sack as quoted by C.C.J. Roothaan and P.S. Bagus in: B. Alder, S. Fernbach and M. Rotenberg (eds.), Methods in Computational Physics, Academic Press, New York (1963).

THE OPTIMISATION OF THE NON-ORTHOGONAL ORBITALS IN A GENERAL VALENCE-BOND WAVEFUNCTION - THE VB-SCF METHOD.

Joop H. v.Lenthe* and Gabriel G. Balint-Kurti
School of Chemistry, Bristol University, Bristol BS8 1TS, U.K.

SUMMARY

We present an extension of the Generalised Brillouin Theorem[1] and the corresponding MCSCF method[2,3] to allow for non-orthogonal orbitals. Essentially the same formalism as in the orthogonal case applies, but care should be taken to use the proper (singly excited) Brillouin states.

Test calculations for OH indicate a good convergence behaviour. Several small valence-bond functions have been investigated and the equilibrium distance and the dissociation energy are found to improve dramatically upon orbital optimisation. Application of valence-bond theory in this manner allows us to use an accurate wavefunction to describe chemical bonding in terms of widely used chemical concepts like hybridisation and mixing of valence-bond structures.

THEORY

Brillouin theorem

The Generalised Brillouin theorem[1] for orthogonal orbitals is well documented. We generalise it here to permit the use of non-orthogonal orbitals in order to show that the same principle still applies. Consider an orbital transformation of the form $\phi_i' = \phi_i + b_{ij}\phi_j$. For the first derivative of the energy with respect to b_{ij} we obtain[1,2]:

$$\frac{dE}{db_{ij}} = \frac{d}{db_{ij}} \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{2}{\langle\Psi|\Psi\rangle} \left[\langle\Psi|H|\Psi_{ij}\rangle - \frac{\langle\Psi|H|\Psi\rangle\langle\Psi|\Psi_{ij}\rangle}{\langle\Psi|\Psi\rangle}\right] \qquad (1)$$

Where Ψ_{ij} is the (singly excited) Brillouin state corresponding to the orbital mixing (see next section). So we get:

$$\frac{dE}{db_{ij}} = 2\,\frac{\langle\Psi|H-E|\Psi_{ij}\rangle}{\langle\Psi|\Psi\rangle} \qquad (2)$$

The requirement that the energy be stationary yields:

Generalised Brillouin Theorem: $\langle\Psi|H-E|\Psi_{ij}\rangle = 0$ (3)

This Brillouin theorem can be used in exactly the same way as that for orthogonal orbitals[2,3] to optimise the non-orthogonal orbitals in a valence-bond wavefunction.

Brillouin states

The main problem in an optimisation of non-orthogonal orbitals is to avoid dependencies, i.e. to use the correct number of degrees of freedom. To this end we distinguish two cases for mixing two orbitals ϕ_i and ϕ_j.

A. The orbitals may be varied independently. In this case two separate Brillouin states Ψ_{ij} and Ψ_{ji} exist, representing the two available degrees of freedom:

$$\begin{aligned}\Psi_{ij} &= \Psi\ (i\rightarrow j)\\ \Psi_{ji} &= \Psi\ (j\rightarrow i)\end{aligned} \qquad (4)$$

B. The orbitals may not be varied independently, i.e. it is possible to orthogonalise orbital ϕ_j to orbital ϕ_i without affecting the wavefunction. In this case we have only one degree of freedom and should use only one Brillouin state, the usual Brillouin state for orthogonal orbitals[1-3]:

$$\Psi_{ij} = -\Psi_{ji} = \Psi\ (i\rightarrow j) - \Psi\ (j\rightarrow i) \qquad (5)$$

Of course care should be taken not to introduce any orbital mixings that do not change the wavefunction at all. In the presented formalism this situation is usually easily recognised by the program.

Method

Essentially the same method as described previously[3] is incorporated in an automated and improved version of the valence-bond program Multibond[4], allowing for a maximum of 30 orbitals and 1000 Slater determinants. The orbital transformation matrix, used in the iterative optimisation, is determined by the first order method[2,3]. As many orbitals as possible are orthogonalised in order to improve convergence and facilitate various matrix-manipulations.

The Brillouin states used may be chosen freely, which makes it easy to impose restrictions on the kind of orbital mixings, that occur. An interesting case is to allow only mixings within each atom in order to determine optimal hybrid orbitals (i.e. distorted atomic orbitals on a single center).

APPLICATION

We have applied the method to the optimisation of various valence-bond wavefunctions for OH ($^2\Pi$). A double zeta + polarisation gaussian basis[5] ((9,5,1/5,1)→[4,2,1/2,1]) was used with exponents of 1.0 for the polarisation functions. No equivalence restrictions were applied so the optimised wavefunctions depart slightly from a pure Π state.

---

[*] Dutch Ramsay Memorial Fellow.

Two types of orbital optimisation were employed:

A. "intra-atomic": only orbitals on the same atom are allowed to mix.

B. "Full": the orbitals are optimised in the complete function space.

Various wavefunctions* have been used. One of them, a basic valence-bond function, is shown in Table I, with its coefficients for an internuclear distance near to the equilibrium distance (R = 1.85 a.u.; E = -75.42022 a.u.).

Table I. Basic VB functions "Intra-atomic" optimised at 1.85 bohr.

no.	Constituent atomic states	Orbital occupation[(a)]	Coefficient
1	$O(^3P) \otimes H(^2S)$	$1s^2 2s^2 y^2 x\ zh$[(b)]	-0.235
2	$O(^1D) \otimes H(^2S)$	$1s^2 2s^2 y^2 x\ zh$[(b)]	0.191
3	$O^+(^2D^o) \otimes H^-(^1S)$	$1s^2 2s^2 (z^2-y^2)xh^2$	0.053
4	$O^+(^2P^o) \otimes H^-(^1S)$	$1s^2 2s^2 (z^2+y^2)xh^2$	-0.028
5	$O^-(^2P^o) \otimes H^+$	$1s^2 2s^2 z^2 y^2 x$	0.327

(a) 1s,2s,x,y,z are orbitals on oxygen, h is the 1s orbital on hydrogen

(b) The spinfunctions are $2\alpha\alpha\beta-(\alpha\beta+\beta\alpha)\alpha$ and $(\alpha\beta-\beta\alpha)\alpha$ respectively.

Another interesting function is the covalent VB-function which consists of structures 1 and 2 of Table I. The computed equilibrium distance and dissociation energies for these functions using various optimisations are given in Table II. Also given are the number of iterations needed to converge within 10^{-4} (for $|b_{ij}|$; see ref. 3) starting from the previous level of optimisation (in brackets) and the overlap between the H orbital and the O 2s-type orbital (S_{sh}) and z-type orbital (S_{zh}), all at 1.85 bohr.

Table II. Results for various VB-functions

wavefunction	optimisation	(no of iterations)	R_e(bohr)	D_e(eV)	S_{sh}	S_{zh}
Basic	∞[(a)]	(-)	2.20	1.73	0.51	0.41
Basic	"Intra-atomic"	(13)	1.87	3.28	0.31	0.54
Basic	"Full"	(9)	1.87	3.43	0.30	0.54
Covalent	"Intra-atomic"	(7)	1.76	1.77	0.27	0.60
Covalent	"Full"	(11)	1.87	3.42	0.28	0.77
Experiment[6]			1.83	4.62		

(a) ∞ indicates that the optimised orbitals for the separate atoms are used

Convergence is quite good and the optimisation gives a dramatic improvement in equilibrium geometry and dissociation energy. The best result obtained so far is with a VB-CI (cf. ref.3) calculation using all "Intra-atomic" single excitations from a 13 structure VB-SCF function. This yielded a D_e of 4.28 eV. It is interesting to note that the covalent VB function, when fully optimised, yields the same energy as the basic VB function, indicating that allowing the orbitals on different centers to mix can eliminate the need for ionic valence-bond structures.

On inspection of the overlaps in Table II we note that, upon optimisation, the oxygen 2s-hydrogen 1s overlap is decreased, while the overlap between the oxygen z orbital and the hydrogen 1s orbital increases significantly. This is consistent with the interpretation that the oxygen 2s orbital is distorted to point away from the hydrogen while the oxygen z orbital forms an "sp"-hybrid pointing towards it. An examination of the orbital coefficients indicates, that there is a considerable s-p mixing on oxygen, but much less than would be needed to form two equivalent sp hybrids.

REFERENCES

1. B. Levy and G. Berthier, Int.J.Quantum.Chem. 2 (1968) 307
2. F. Grein and T.C. Chang, Chem.Phys.Lett. 12 (1971) 44.
   J.H. v.Lenthe, Ph.D. thesis (Utrecht, 1979).
3. J.H. v.Lenthe and P.J.A. Ruttink, A MCSCF program employing the Generalised Brillouin Theorem, these proceedings (1980).
4. G.G. Balint-Kurti and R.N. Yardley, 'MULTIBOND (Version A), Ab Initio Multistructure Valence-Bond Calculational System' QCPE 12 (1977) 335.
5. F.B. van Duijneveldt, IBM Technical Report RJ 945 (1971).
6. B. Rosen (ed.) Selected Spectroscopic data on diatomic molecules (Pergamon, London, 1970).

---

* All wavefunctions were such that the orbitals on the same atom could be mutually orthogonal.

A PARTITIONING TECHNIQUE FOR MULTICONFIGURATION SELF-CONSISTENT-FIELD THEORY

C. W. Bauschlicher, Jr.[*]
ICASE, Mail Stop 132C, NASA Langley Research Center, Hampton, Virginia 23665

P. S. Bagus
IBM Research Laboratory, 5600 Cottle Road, San Jose, California 95193

D. R. Yarkony[#+§]
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 21218

---

[*] Supported under NASA Contract No. NAS11-14101, NAS1-14472, NAS1-15810 while in residence at ICASE, NASA Langley Research Center, Hampton, Virginia 23665.

[#] Partially supported by Air Force Office of Scientific Research (AFOSR-79-0073) and National Science Foundation (CHE-7824153) grants during the course of this research.

[+] Alfred P. Sloan Research Fellow.

[§] Contributed by D. R. Yarkony.

ABSTRACT:

The limiting step in many MCSCF procedures is a requisite 4-index orbital transformation. In this work an extrapolation procedure designed to reduce the number of these transformations is presented. The algorithm is particularly well-suited to an MCSCF algorithm based on the Generalized Brillouins Theorem and Iterative Natural Orbital procedure, and has implications for the design of hybrid procedures also.

INTRODUCTION

In recent years Multiconfiguration Self-Consistent-Field (MCSCF) procedures[1] based on the generalized Brillouins theorem[2] (GBT) and the iterative natural orbital[3] (INO) procedure have been developed independently by several groups.[4-6] Although this approach has proved useful in obtaining compact MCSCF descriptions for ground and excited state wavefunctions, it suffers from at least two deficiencies:

(i) It is not a completely quadratic procedure and consequently, may show poor convergence properties toward the end of an iterative sequence.

(ii) It becomes computationally cumbersome for large MCSCF expansions.

The complete active space (CAS)[5] approach to MCSCF addresses the latter problem; however, as formulated, it is restricted to a full valence wavefunction.

In this work an extrapolation procedure based on a partitioning of the orbital space previously used to obtain convergence for large ill-conditioned SCF problems is presented. The procedure is effective in reducing the number of iterations (and, hence, the number of N^5 integral transformations) required to converge the MCSCF wavefunction. In addition the efficacy of this procedure suggests that a partitioning of the orbital space may provide the basis for a hybrid MCSCF procedure in which first- and second-order algorithms are combined.

In section II the orbital partitioning scheme and extrapolation technique are discussed in the context of the GBT-INO algorithm; section III presents the results of representative calculations and section IV discusses the implications for hybrid MCSCF procedures and concludes.

GENERAL CONSIDERATIONS

The general formulation of an MCSCF procedure based on the GBT-INO approach has been discussed previously by several authors.[4] Here, we summarize the results necessary for the present discussion.

The MCSCF conditions are given by

$$\underset{\approx}{H}\,\underset{\sim}{C}^{I} = E_{I}\,\underset{\sim}{C}^{I} \tag{1a}$$

$$<\Psi_{ij}^{R,I}|H|\Psi^{I}> = 0 \tag{1b}$$

where

$$H_{k,\ell} = <\psi_k|H|\psi_\ell>$$

$$\Psi^{I}(\underset{\sim}{C},\underset{\sim}{\phi}) = \sum_{J=1}^{M} \psi_J(\phi)\, C_{J,I} \tag{2}$$

$$\underset{\sim}{\phi} = \{\phi_\ell\}_{\ell=1}^{N}$$

and $\Psi_{ij}^{R,I}$ is the Brillouin single excitation[7] corresponding to the mixing of inequivalent orbitals ϕ_i and ϕ_j. In the usual implementations of the GBT-INO approach at the n^{th} step of the iterative determination of $\Psi^{I}(\underset{\sim}{C},\underset{\sim}{\phi})$:

(a) the appropriate eigenfunction, Ψ_{I}^{SCI}, of the super-CI (SCI)[6] matrix (H^{SCI}), i.e. the root of H in the basis $\{\Psi^{I}, \Psi_{ij}^{R,I}\}$ having the coefficient of Ψ^{I}, c_o, maximal, is determined and;

(b) the "net" first-order density matrix[4] corresponding to Ψ_{I}^{SCI} is constructed and diagonalized to define $\underset{\approx}{U}(n)$ and $\underset{\sim}{\phi}^{n+1}=\underset{\sim}{\phi}^{n}\underset{\approx}{U}(n)$.

At this stage an N^5 orbital transformation must be performed to continue the procedure. However, an extrapolatory subiteration is possible based on the following observations:

(i) Convergence of large ill-conditioned SCF problems has been achieved by examining the behavior of the iterative scheme in a reduced orbital subspace consisting of the higher doubly occupied-, partially occupied- and lower virtual-orbitals.

(ii) The natural orbital procedure tends to produce virtual orbitals with the property:

$$<\Psi_{ij}^{R,X}|H|\Psi^{I}> \xrightarrow[j\to N]{} 0 \qquad \text{(see eq. 1b)} \qquad (3)$$

(iii) The terms in 1b corresponding to the rotation $(\phi_j,\phi_i)_I$ where ϕ_j and ϕ_i are partially occupied in Ψ^I are the most difficult to eliminate and, therefore, should be treated to as high an order as possible.

These considerations suggest the following alternative procedure:

(a) Following the determination of $\underset{\sim}{\phi}(n)$, an L-orbital subset, $\underset{\sim}{\phi}'(n)$, is selected corresponding to all the occupied orbitals and (in general only) the first virtual orbital of each symmetry.

(b) The MCSCF problem is solved (to any level of approximation desired) in this reduced orbital space using the GBT-INO procedure. Considerable economies over the full N-orbital space (FOS) solution are achieved in this subiteration since the N^5 transformation can be replaced by an $L\cdot N^4$- and (by using successively transformed integrals) L^5-transformations.

(c) Following this subiteration the two orbital subspaces are combined to give $\underset{\sim}{\phi}''(n)$ and a FOS iteration performed to determine $\underset{\sim}{\phi}(n+1)$.

As a result of eq. 3 this procedure has proved effective in reducing the number of iterations required to achieve convergence to the MCSCF wavefunction using the GBT-INO procedure.

In the next section we present preliminary results of a study of the effectiveness of this procedure, which is now in routine use.[7]

APPLICATIONS

The procedure described in section II is illustrated in two tables contained in this section. The data was compiled in the course of an MCSCF/SCF study of the low-lying states of MgO which has been reported elsewhere[7] and where the basis of Slater-type orbitals and configuration state functions (CSF's) used are described.

In each table iterations are separated by double horizontal lines. An iteration begins with a 4-index transformation using the orbital set $\underset{\sim}{\phi}(n)$ and a determination of $\underset{\sim}{\phi}'(n+1)$ in the FOS using the GBT-INO procedure. This space is then partitioned into a space dimension L (denoted L) and its orthogonal compliment (denoted $L^{\perp}$) as described in the previous section and GBT-INO procedure is continued in L. The expansion coefficients of $\phi_i \epsilon L$ are well-behaved so that the use of successively transformed integrals is permitted. Consequently, while the first subiteration involves an $N^4\cdot L$ transformation subsequent subiterations involve only L^5 transformations. At the end of each subiteration sequence, L and $L^{\perp}$ are reunited to give $\underset{\sim}{\phi}(n+1)$ and the process is repeated.

Each row in the table is characterized by an ordered pair (i,j) in which i denotes the total number of N^5 transformations which precede this step and j the number of transformations in L since the last N^5 transform. One measure of the efficiency of this procedure is to compare the aggregate $\Delta^2 = 1-c_0^2$ between iterations, i.e. from (n,0) to (n+1,0) with the extrapolated result given approximately by (n,0) to (n,1). From the tables it can be seen that the subiterations are effective in reducing Δ^2.

CONCLUSIONS

In this work a partitioning method previously used to achieve convergence of large ill-condition SCF problems is applied to the GBT-INO MCSCF procedure. The resulting scheme appears to represent an efficient method for reducing the number of N^5 integral transformation required to converge the MCSCF wavefunction. The success of this procedure suggests that in general a partitioning of the orbital space may lead to a more efficient MCSCF algorithm. For example, one might consider a hybrid scheme in which a second-order procedure is used in L while a lower (first) order procedure (Fock operator[1] or CAS MCSCF approach, for example) is used in the FOS. Since first-order procedures in general replace the N^5 orbital transformation step by an N^4 construction step, considerable ecomomies could result from such a scheme.

REFERENCES

1. A. Wahl and G. Das, Modern Theoretical Chemistry, edited by H. F. Schaefer III (Plenum, New York, 1976) Vol. 3.
2. B. Levy and G. Berthier, Int. J. Quantum Chem., 2, 307 (1968).
3. C. F. Bender and E. R. Davidson, J. Chem. Phys. 70, 2675 (1966).
4. (a) D. R. Yarkony. J. Chem. Phys., 66, 2045 (1977); C. W. Bauschlicher, Jr. and D. R. Yarkony, J. Chem. Phys., 72, 1138 (1980).
(b) L. M. Cheung, S. T. Elbert, and K. Ruedenberg, Int. J. Quantum Chem., 16, 1069 (1979).
5. P. Siegbahn, A. Heiberg, B. Roos and B. Levy, Physica Scripta, 21, 323 (1980).
6. F. Grein and T. C. Chang, Chem. Phys. Letters, 12, 44 (1971).
7. For example, the results reported in C. W. Bauschlicher, D. M. Silver, and D. R. Yarkony, J. Chem. Phys., in press, use this procedure.

Table 1: MgO 1 $^1\Sigma^+$ state R=3.5b, energies in hartrees

Iteration	E(MCSCF)	E(SCI)	Δ^2
0 0	-274.4492371	-274.4595471	2.40(2)†
1 0	-274.4560762	-274.4609194	5.67(3)
1 1	-274.4604358	-274.4610953	1.43(3)
1 2	-274.4611132	-274.4611566	
2 0	-274.4611626	-274.4616116	4.10(4)
2 1	-274.4613497	-274.4616530	3.78(4)
2 2	-274.4616383	-274.4616444	
3 0	-274.4616426	-274.4616703	3.80(5)
3 1	-274.4616783	-274.4616815	6.0 (6)
4 0	-274.4616843	-274.4616843	2.0 (6)
4 1	-274.4616827	-274.4616854	
4 2	-274.4616854	-274.4616854	

†Negative of characteristic, base 10, given in parenthesis.

Table 2: MgO 2 $^1\Sigma^+$ state, R=3.0b, energies in hartrees

Iteration	E(MCSCF)	E(SCI)	Δ^2
0 0	-274.3533654	-274.354414	2.58(3)†
1 0	-274.3546429	-274.3548042	2.38(4)
1 1	-274.3548012	-274.3548345	5.9 (5)
1 2	-274.3548476	-274.3548481	
2 0	-274.3548483	-274.3548632	1.1 (5)
2 1	-274.3548599	-274.3548662	1.0 (5)
2 2	-274.3548651	-274.3548656	2.0 (6)

†Negative of characteristic, base 10, given in parenthesis.

A MULTICONFIGURATION METHOD FOR EXCITED STATES, AND SELECTION OF OPTIMAL CONFIGURATIONS

F. Grein
Department of Chemistry
University of New Brunswick
Fredericton, New Brunswick
Canada E3B 5A3

INTRODUCTION

It is assumed that multiconfiguration methods based directly on the satisfaction of the generalized Brillouin theorem (so-called super-CI methods)[1] have been reviewed earlier. It was shown[2] that such methods converge quadratically when the super CI matrix is diagonalized.

Among the early applications were all-single-excitation (ASE) wavefunctions for atoms from He to F[3,4,5,6]. Such wavefunctions are of the form

$$\Psi = C_o\Psi_o + \sum_{k=1}^{N} C_k|a_o b_o \ldots k \ldots N_o|$$

Due to the orbital optimization, the N terms in the summation constitute all possible single excitations of Ψ_o. Ψ_o is the "anti Hartree-Fock" wavefunction, or the wavefunction which has maximum interaction with its single excitations. ASE wavefunctions give from 7.6% (for B) to 32.2% (for He) of the correlation energy. Pair excitation wavefunctions of similar structure give virtually identical correlation energies.

MULTICONFIGURATION METHOD FOR EXCITED STATES

In order to use the super-CI technique for excited states of the same symmetry (higher roots of the eigenvalue problem), a level shifting method has been introduced[7]. The super-CI (SCI) wavefunction is formed for the nth root of the MC wavefunction, according to the equation

$$\Psi_{SCI,n} = \Psi^{(n)} + \sum_{i<j} x_{ij}\, \Psi^{(n)}(i\rightarrow j)$$

Without level shifting, none of the SCI eigenvectors usually contain predominantly $\Psi^{(n)}$, causing root flipping between $E^{(n)}$ and lower roots. This can be prevented by subtracting a root shifting constant q from the diagonal element $\langle\Psi^{(n)}|H|\Psi^{(n)}\rangle$ of the SCI matrix. q has to be bigger than the difference $E^{(n)}-E^{(1)}$.

SELECTION OF OPTIMAL CONFIGURATIONS

The choice of configruations is critical for MC wavefunctions with a small number of configurations. For this purpose, a method to select optimal configurations has been proposed, and tested on several states of the NH molecule[8,9,10]. It consists of a series of CI selection steps coupled with the MCSCF iterative procedure. Starting with the initial orbitals and m initial configurations, all their single and double excitations are generated. A series of diagonalizations of Hamiltanian matrices of order m+1 are performed, and those configurations which give an energy lowering in excess of a parameter δ are added to the MC wavefunction. δ may initially be 0.005 hartree, but can be lowered as the iteration-selection process is continued. If potential energy curves or surfaces are of interest, then the selection has to be performed at various points. Since the selection step is fast, it was found that this method gave all required configurations in an efficient, systematic and reliable way.

REFERENCES

1. F. Grein and T. C. Chang, Chem. Phys. Letters, 12, 44 (1971).
2. A. Banerjee and F. Grein, Intern. J. Quant. Chem., 10, 123 (1976).
3. F. Grein, Intern. J. Quant. Chem., 5, 165 (1971).
4. F. Grein and T. C. Chang, J. Phys. B:Atom. molec. Phys. 6, L237 (1973).
5. F. Grein and A. Banerjee, Chem. Phys. Letters, 25, 255 (1974).
6. F. Grein and A. Banerjee, Chem. Phys. Letters, 31, 281 (1975).
7. F. Grein and A. Banerjee, Intern. J. Quant. Chem. Symp., 9, 147 (1975).
8. A. Banerjee and F. Grein, J. Chem. Phys., 66, 1054 (1977).
9. A. Banerjee and F. Grein, J. Chem. Phys., 66, 2589 (1977).
10. A. Banerjee and F. Grein, Chem. Physics, 35, 119 (1978).

THE SUPER-CI MCHF PROGRAM IN ALCHEMY

Michel Dupuis
National Resource for Computation in Chemistry
Lawrence Berkeley Laboratory
Berkeley, CA 94720

Bowen Liu
IBM Research Laboratory
5600 Cottle Road
San Jose, CA 95193

Introduction

We will describe a computer code used for the calculation of multiconfiguration Hartree-Fock wavefunctions, available in the ALCHEMY[1] set of programs. The program represents an implementation of a method based on the generalized Brillouin theorem, also called "Super-CI" method, which has been reviewed earlier.[2] Briefly, the MCHF wavefunction Ψ_{MC} is written as a linear combination of configuration state functions (CSF) ϕ_k

$$\Psi_{MC} = \sum_k a_k \phi_k \qquad (1)$$

The singly excited Brillouin configuration $\Psi_{MC}(i \to j)$ are constructed according to

$$\Psi_{MC}(i \to j) = \sum_k a_k \; \phi_k(i \to j) - \phi_k(j \to i) \qquad (2)$$

where i j indicates that the electron in spin-orbital i has been promoted to the spin orbital j. The "Super-CI" wavefunction Ψ is written as

$$\Psi = \Psi_{MC} + \sum_{i>j} C_{ij} \Psi_{MC}(i \to j) \qquad (3)$$

The spin orbital expansion coefficients are varied until the generalized Brillouin theorem equations are satisfied: i.e.,

$$<\Psi_{MC} | H | \Psi_{MC}(i \to j)> = 0, \qquad (4)$$

for all i→j pairs. The iterative method uses the natural orbitals of the "Super-CI" wavefunction as the improved orbitals for the next iteration. Two main steps are involved in using the program: the first step creates the list of configurations and energy formulas for the corresponding CI problem.[3] The second step is the iterative process itself.

Step 1: Energy Expression

i. Generate reference CSF's (nref).

ii. Generate independent singly excited (-sx-) CSF's (nsx)

iii. Generate transformation matrix formulas of -sx- CSF's to create $\phi_k(i \to j)$

iv. Generate formula tape for energy matrix (nref + nsx)

v. Sort formula tape.

Step 2: Iterative Procedure

i. Initial guess molecular orbitals MO's.

ii. Transform electron repulsion integrals from atomic basis set to molecular basis set.

iii. Construct energy matrix (nref + nsx).

iv. Solve (nref x nref) CI to get "a_k".

v. Construct "Super-CI" matrix.

vi. Solve "Super-CI" problem.

vii. New MO's = natural orbitals of "Super-CI" wavefunction.

viii. Back to ii. if not converged.

Program Description

i. It is open ended with no restriction on the number of orbitals, the number of electrons, or the number of open shells.

ii. It uses CSF's which are linear combinations of Slater determinants built from an orthonomal set of one-particle symmetry and equivalence restricted spatial orbitals.

iii. It takes explicit advantage of $C_{\infty v}$, $D_{\infty h}$ s[illegible]y as well as D_{2h} and subgroup symmetry.

iv. The selection of CSF's can be done by specifying electronic couplings, or by partitioning the MO space.

v. There is an option for "Complete CI in Active Space" wavefunction calculations.

vi. The improved molecular orbitals can be selected by maximum overlap with the orbitals of the previous iteration.

vii. The program generates the Hartree-Fock canonical closed shells, useful for subsequent CI calculations.

viii. Calculations on averaged states can be done.

Program Characteristics

i. The iteration time is almost independent of the number of CSF's. The integral transformation time represents approximately 75% of the total iteration time.

ii. The typical number of iterations is less than 7 iterations to get

$$|E_{\text{"Super CI"}} - E_{MC}| \leq 10.0^{-7}$$

A greater number of iterations is indicative of a poor CSF selection.

iii. The practical limit of the program is about 150 CSF's due to a time consuming formula type generation.

Conclusion

The program is routinely used for problems involving up to 100 CSF's and has proven to be converging reliably. Potential energy curves and surfaces have been calculated with CSF's required to describe the proper dissociation behavior for the system, although the CSF selection is not necessarily restricted to that choice. As an example the calculation of the energy curve for Be_2 and Mg_2, using the method of interacting correlated fragments (ICF)[5] involved a 35 CSF's calculation in an extended basis set of Slater atomic orbitals.

Acknowledgment

This research was supported in part by the National Resource for Computation in Chemistry under a grant from the National Science Foundation (Grant No. CHE-7721305) and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Chemical Sciences Division of the U. S. Department of Energy (Contract No. W-7405-ENG-48).

References

1. M. Yoshimine, A. D. McLean, B. Liu, M. Dupuis, and P. S. Bagus, Nat. Resour. Comput. Chem. Software Cat., Vol. 1, Prog. No. QC03 (ALCHEMY), 1980.

2. See K. Ruedenberg in present volume.

3. E. R. Davidson, NRCC Workshop Proceedings: "Post Hartree-Fock Configuration Interaction," p. 127, 1978, available from National Technical Information Service, U.S. Department of Commerce, 5285 Park Royal Road, Springfield, VA 22161.

4. P. Siegbahn, A. Heiberg, B. Roos, and B. Levy, Phys. Script 21, 323, (1980). See also B. Roos in present volume.

5. B. Liu and A. D. McLean, J. Chem. Phys. 72, 3418 (1980)

A UNITARY EXPONENTIAL OPERATOR APPROACH TO MULTICONFIGURATIONAL HARTREE FOCK

Danny L. Yeager
Chemistry Department, Texas A&M University
College Station, Texas 77843

I. INTRODUCTION

Recently there has been renewed interest in the multiconfigurational Hartree-Fock (MCSCF) technique.[1] We have recently developed, coded, and studied[2-5] an approach to MCSCF based on a unitary operator. This unitary operator is written as an exponential of other operators. Explicit in this formulation is the coupling between the orbitals and the state expansion coefficients. By expanding the expectation value of H through second order and applying the variational principle, the Newton-Raphson equations for orbital and coefficient optimization are derived.[2,3,6,7] In this presentation we discuss the theoretical development of this technique and present some results of calculations on Be and O_2.

There have been some further recent developments.[4,5] Without constraining the Newton-Raphson equations when we are far from convergence, the technique may converge to the nearest stationary point or even diverge. By forcing all of the eigenvalues of the Hessian to be positive a downhill "walk" (through second order) on the energy hypersurface is assured. For small eigenvalues, the steplength may be reduced. Modes that are predominantly orbital or mixed or configurational are damped differently. These techniques will be discussed in detail.

Finally, it is shown how the two electron integral transformation may be replaced by an approximate orbital transformation introduced directly into the equation that defines the second order MCSCF approach. In this way, the number of two electron integral transformations required to obtain a set of MCSCF orbitals is reduced considerably.

We also note the recent contributions of other workers.[8-10] These will be discussed in more detail by other workshop participants.

II. THE MULTICONFIGURATIONAL HARTREE-FOCK PROCEDURE

A. Unitary transformations in the state and orbital space.

The multiconfigurational Hartree-Fock (MCSCF) reference state $|0\rangle$ may be regarded as a member of the set of states $\{|j\rangle\}$

$$|0\rangle = \Sigma|\phi_g\rangle C_{go} \qquad (1)$$

$$|j\rangle = \Sigma|\phi_g\rangle C_{gj} \qquad (2)$$

for which the coefficients C form a unitary matrix. The states $|\phi_g\rangle$ are given as

$$|\phi_g\rangle = \prod_{r\epsilon g} a_r^\dagger|vac\rangle \qquad (3)$$

where $\prod_{r\epsilon g} a_r^\dagger$ refer to an ordered product of creation operators. For convenience we consider real orbitals and expansion coefficients.

A unitary transformation of the states $|j\rangle$ may then be described as[6]

$$\exp(i\ \hat{S})\ |j\rangle = \Sigma|k\rangle(\exp -S)_{kj} = \Sigma|k\rangle T_{kj} \qquad (4)$$

where

$$\hat{S} = i\ \Sigma\ S_{ko}(|k\rangle\langle 0|-|0\rangle\langle k|) \qquad (5)$$

T = exp(-S) is a unitary matrix and S a real antisymmetric matrix.

A unitary transformation of the orbitals may similarly be described as[7]

$$\tilde{a}_r^\dagger = \exp(i\ \hat{\kappa})a_r^\dagger \exp(-i\ \hat{\kappa}) \qquad (6)$$

where

$$\hat{\kappa} = i\ \Sigma\ \kappa_{rs}(a_r^\dagger a_s - a_s^\dagger a_r) \qquad (7)$$

From Eqs. (6) and (7) we get

$$\tilde{a}_r^\dagger = \Sigma\ a_s^\dagger(\exp-\kappa)_{sr} = \Sigma a_s^\dagger\ X_{sr} \qquad (8)$$

where X = exp(- κ) is unitary and κ antisymmetric.

The set of excitation operators $\{a_r^\dagger a_s\}$ and $\{|k\rangle\langle 0|\}$ may be linearly dependent. The elimination of redundant operators $a_r^\dagger a_s$ for a specific reference state has been described in detail in ref. [7].

B. The second order multiconfiguration Hartree-Fock approach.

We will now determine stationary points on an energy-hypersurface where variation in the orbitals and in the expansion coefficients C is considered. A new reference state $|\tilde{0}\rangle$ may then be obtained by means of unitary transformation

$$|\tilde{0}\rangle = \exp(i\ \hat{\kappa})\ \exp\ (i\ \hat{S})|0\rangle \qquad (9)$$

The total energy then becomes

$$\begin{aligned} E(\kappa,S) &= \langle 0|\exp(-i\hat{S})\exp(-i\hat{\kappa})H\exp(i\hat{\kappa})\exp(i\hat{S})|0\rangle = \\ &\langle 0|H|0\rangle -i\langle 0|[\hat{S}+\hat{\kappa},H]|0\rangle + \tfrac{1}{2}\langle 0|[\hat{S},[H,\hat{S}]]|0\rangle \\ &+ \tfrac{1}{2}\langle 0|[\hat{\kappa},[H,\hat{\kappa}]]|0\rangle + \langle 0|[\hat{S},[H,\hat{\kappa}]]|0\rangle+\ldots \end{aligned} \qquad (10)$$

A stationary point on the energy hypersurface is obtained when δE(κ,S) = 0. We get

$$\begin{aligned} \delta E(\kappa,S) &= -i\langle 0|[\delta\hat{S}+\delta\hat{\kappa},H]|0\rangle \\ &+ \langle 0|[\delta\hat{S},H,\hat{S}]|0\rangle + \langle 0|[\delta\hat{\kappa},H,\hat{\kappa}]|0\rangle \\ &+ \langle 0|[\hat{S},[H,\delta\hat{\kappa}]]|0\rangle + \langle 0|[\delta\hat{S},[H,\hat{\kappa}]]|0\rangle+.. \end{aligned} \qquad (11)$$

where the double commutator is introduced. Neglecting third and higher order terms, Eq. (11) can be written in matrix notation as

$$-\begin{pmatrix} W \\ V \end{pmatrix} = (A - B) \begin{pmatrix} \kappa \\ S \end{pmatrix} \qquad (12)$$

where we have used Eqs. (5) and (7) and introduced the notation $Q_\nu^\dagger = a_r^\dagger a_s$, r > s and $R_n^\dagger = |n\rangle\langle 0|$.

$$W = \langle 0|[Q,H]|0\rangle \qquad (13)$$

$$V = \langle 0|[R,H]|0\rangle \qquad (14)$$

$$A = \begin{pmatrix} <0|[Q,H,Q^{\dagger}]|0> & <0|[[Q,H],R^{\dagger}]|0> \\ <0|[R,[H,Q^{\dagger}]]|0> & <0|[R,H,R^{\dagger}]|0> \end{pmatrix} \quad (15)$$

$$B = \begin{pmatrix} <0|[Q,H,Q]|0> & <0|[[Q,H],R]|0> \\ <0|[R,[H,Q]]|0> & <0|[R,H,R]|0> \end{pmatrix} \quad (16)$$

Eq. (12) may be rewritten as:

$$- \begin{pmatrix} \kappa \\ S \end{pmatrix} = (A \dot{-} B)^{-1} \begin{pmatrix} W \\ V \end{pmatrix} \quad (17)$$

An iterative procedure can now be established to get $\delta E = 0$. For a given set of orbitals and coefficients the matrices W, V, A and B are calculated from Eqs. (13), (14), (15) and (16). The vectors κ and S are then evaluated from Eq. (12) and a new coefficient matrix and a new set of orbitals are determined from Eq. (4) and (8). The procedure is repeated until the numerical value of W and V is smaller than a specific tolerance. The above procedure is a second order procedure[5] which performs the variation in the orbitals and in the coefficient in a single step.[6] *It will be denoted as* the <u>one step second order</u> approach.

A two step second order procedure may be described as follows. After an initial guess of orbitals the matrix T in Eq. (4) is determined from a configuration interaction calculation. We then have

$$<m|H|n> = E_n \delta_{mn} \quad (18)$$

and the matrix V becomes equal to zero.

The matrix κ that contains the effect of variation in the coefficient matrix through second order may then be determined from a partitioned form of Eq. (19)

$$- \kappa = (A' - B' - Y)^{-1} W \quad (19)$$

where

$$A' = <0|[Q,H,Q^{\dagger}]|0> \quad (20)$$

$$B' = <0|[Q,H,Q]|0> \quad (21)$$

$$Y = <0|[[Q,H],R^{\dagger}-R]|0><0|[R,H,R^{\dagger}]|0>^{-1} \times <0|[R^{\dagger}-R,[H,Q^{\dagger}]]|0> \quad (22)$$

From Eqs. (8) a new set of orbitals may be derived, a limited CI calculation performed etc. This procedure will be denoted the <u>two step second order</u> approach.

Calculations using these approaches and approximations to these techniques are reported in ref. [2-4].

III. CONSTRAINTS IN THE INITIAL ITERATIONS OF THE MULTICONFIGURATION HARTREE-FOCK PROCEDURE

Numerical experience has shown that in the first couple of iterations, matrix elements of the Hessian may show a great relative variation as a result of the initial guess of orbitals. As a consequence, the step length of the iterative procedure may have to be reduced until the variation in these matrix elements from one iteration to the next becomes small. We use a mode damping scheme so that we may perform a controlled walk on the energy hypersurface.

For the one step procedure we consider the spectral representation of the Hessian

$$(A-B)^{-1} = U \varepsilon^{-1} U^{\dagger} \quad (23)$$

The transformed Eq. (17) becomes

$$- \begin{pmatrix} \bar{\kappa} \\ \bar{S} \end{pmatrix} = \varepsilon^{-1} \begin{pmatrix} \bar{W} \\ \bar{V} \end{pmatrix} \quad (24)$$

where

$$\begin{pmatrix} \bar{\kappa} \\ \bar{S} \end{pmatrix} = U^{\dagger} \begin{pmatrix} \kappa \\ S \end{pmatrix} \quad (25)$$

$$\begin{pmatrix} \bar{W} \\ \bar{V} \end{pmatrix} = U^{\dagger} \begin{pmatrix} W \\ V \end{pmatrix} \quad (26)$$

The matrix elements that couple the orbital and the coefficient optimization are in most cases very small and $\bar{\kappa}$ then predominantly refers to the orbital space and $\bar{S}$ to the coefficient space. We limit the step length that may be taken by restricting the allowed size of matrix elements of $\bar{S}$ and $\bar{\kappa}$. We use different values for constraints applied for $\bar{S}$ and $\bar{\kappa}$ since the amplitudes of the dominant configurations very seldom change more than 10% during the iterative procedure, while the *orbitals may be completely changed during the* iterative procedure, e.g. if the initial guess of orbitals contains a very diffuse orbital it may become tight as convergence progresses. Numerical experience has shown that it is only for a small ε matrix elements that constraints have to be applied. We apply a constraint whenever $|\varepsilon_p|<0.1$ a.u. then the corresponding $\bar{\kappa}_p$ (or $\bar{S}_p$) is replaced by $\bar{\kappa}_p^0$ (or $\bar{S}_p^0$), if $|\bar{\kappa}_p|$ (or $|\bar{S}_p|$) is greater than $|\bar{\kappa}_p^0|$ (or $|\bar{S}_p^0|$). Note that the sign of $\bar{\kappa}_p^0$ ($\bar{S}_p^0$) is <u>not</u> changed from $\bar{\kappa}_p(\bar{S}_p)$ for positive ε_p. However, if a particular mode is predominantly orbital in nature for these calculations we automatically force the mode to have a positive ε_p (regardless of the magnitude of ε_p) and the sign of $\bar{\kappa}_p^0$ ($\bar{S}_p^0$) may change when compared to $\bar{\kappa}_p(\bar{S}_p)$.[4] For an orbital mode with negative ε_p, $\bar{\kappa}_p(\bar{S}_p)$ is replaced by $\bar{\kappa}_p^0$ ($\bar{S}_p^0$).

For some cases it may be difficult to distinguish between orbital and configuration modes due to a very degree of coupling between the orbital and coefficient space. The degree of mixing in a particular mode is defined by the number τ that is equal to the sum of the square of the coefficients in $U^{\dagger}$ that relates to the configuration excitation operators $R^{\dagger}$. In the present calculation, we arbitrarily define a mode to be mixed or configurational if τ is greater than 0.3. More complex criteria to reduce the size of the step length may be introduced in even more difficult cases. The above described criteria have, however, proven sufficient to get even the very highly correlated excited states considered so far to converge. For the two step procedure we similarly consider the Hessian matrix

$$(A' - B' - Y)^{-1} = U \varepsilon^{-1} U^{\dagger} \quad (27)$$

and the transformed second order equation becomes

$$-\bar{\kappa} = \varepsilon^{-1} \bar{W} \quad (28)$$

where

$$\bar{\kappa} = U^{\dagger} \kappa \tag{29}$$

$$\bar{W} = U^{\dagger} W \tag{30}$$

We constrain the orbital transformation if $|\epsilon_p| <$ 0.1 a.u. in which case the corresponding $\vec{\kappa}_p$ is replaced by $\vec{\kappa}_p^{0}$, if $\bar{\kappa}_p$ is greater than $\vec{\kappa}_p^{0}$. Again, the sign of $\vec{\kappa}_p^{0}$ is the same as $\bar{\kappa}_p$ unless ϵ_p is negative, in which case $\vec{\kappa}_p^{0}$ will have a different sign from $\bar{\kappa}_p$. $\bar{\kappa}_p$ is automatically replaced by $\vec{\kappa}_p^{0}$ when ϵ_p is negative. Constraints cannot as directly be applied in the configuration space, as the correlation coefficients are determined from a limited configuration interaction calculation. However, a damping in the correlation coefficient variation may indirectly be introduced by minimizing the coupling between the orbital and the coefficient optimization (minimize the effect of Y in Eq. (19)). This can e.g. be performed by adding an arbitrary number to the total energy difference $E_n - E_0$, which appears in the denominator of the matrix Y in Eq. (22). We would however, point out that constraints that are introduced in this fashion do not represent a controlled walk on the energy hypersurface, and we have for that reason not explored that possibility further.

Calculations using this technique are reported in the literature.[2-5] In, for example, a calculation on the second state of $^1\Sigma_g^+$ symmetry in C_2, a stringent damping of the mixed or configurational modes is important.

IV. A METHOD TO REDUCE THE NUMBER OF TWO ELECTRON INTEGRAL TRANSFORMATIONS IN A SECOND ORDER MULTI-CONFIGURATIONAL HARTREE-FOCK PROCEDURE

Successive unitary transformations of the orbitals may be described as

$$\tilde{a}_r^{+}(n+1) = \exp(i\hat{\kappa}(n))..\exp(i\hat{\hat{\kappa}}(2))\exp(i\hat{\kappa}(1))a_r^{+} \exp(-i\hat{\kappa}(1))\exp(-i\hat{\hat{\kappa}}(2))..\exp(-i\hat{\hat{\kappa}}(n)) \tag{31}$$

where $\{\tilde{a}_r^{+}(n+1)\}$ denotes the set of creation operators obtained after the n unitary transformations have been carried out. We can express $\hat{\tilde{a}}_r^{+}(n+1)$ in terms of the initial set of creation and annihilation operators by successive application of the relation

$$\exp(i\hat{\hat{\kappa}}(2)) = \exp(i\hat{\kappa}(1))\exp(i\dot{\kappa}(2))\exp(-i\dot{\kappa}(1)) \tag{32}$$

etc.

Operators with no tildes refer to the original MO basis where only an initial two electron integral transformations has been performed.

Similarly the multiconfigurational reference state obtained after n applications of the MCSCF Newton-Raphson equation and two electron transformations can be written as

$$|\dot{0}(n+1)\rangle = \exp(i\hat{\kappa}(1))\exp(i\hat{\kappa}(2))..\exp(i\hat{\kappa}(n)) \exp(i\hat{S}(1))\exp(i\hat{S}(2))..\exp(i\hat{S}(n))|0\rangle$$

$$= \exp(i\hat{\kappa}(1))\exp(i\hat{\kappa}(2))..\exp(i\hat{\kappa}(n))|0'(n)\rangle \tag{33}$$

The Newton-Raphson equations are set up and evaluated using the original creation and destruction operators through first order in $\{\hat{\kappa}(1),\hat{\kappa}(2),\ldots,\hat{\kappa}(n)\}$.[5] This involves evaluating the Hessian with respect to the updated reference state $|0'(n+1)\rangle$ (but zero order in κ) and W and V are evaluated using $|0'(n+1)\rangle$. In addition a term first order in $\hat{\kappa}$ must be introduced in W and V.

Preliminary results using this technique are extremely encouraging. Numerical examples[5] (see Table I) for the $^3\Sigma_g^-$, $^1\Delta_g$, and $^1\Sigma_g^+$ states of O_2 indicate that an accuracy of 10^{-6} a.u. in the total energy may be obtained by carrying out 2-3 two electron integral transformations, which is about half the number of transformations required to obtain the same accuracy in the second order MCSCF approach. An accuracy of 10^{-10} a.u. is obtained after one further iteration is carried out with a second order MCSCF scheme.

V. REFERENCES

1. A. C. Wahl and G. Das in Modern Theoretical Chemistry vol. 3 chap. 3 (Editor H. F. Schaefer III, Plenum Press, New York and London).
2. D. L. Yeager and P. Jørgensen, J. Chem. Phys. $\underline{7}$, [illegible]5 (1979).
3. D. L. Yeager and P. Jørgensen Mol. Phys. $\underline{39}$, 587 (1980).
4. P. Albertsen, P. Jørgensen, and D. L. Yeager, J. Chem. Phys., 1 Sept. 1980.
5. P. Jørgensen, P. Albertsen, and D. L. Yeager, J. Chem. Phys., 15 June 1980.
6. E. Dalgaard, Chem. Phys. Lett. $\underline{65}$, 559 (1979).
7. E. Dalgaard and P. Jørgensen, J. Chem. Phys. $\underline{68}$, 3833 (1978).
8. C. C. J. Roothaan, J. Detrich, and D. G. Hopper, Int. J. Quantum Chem. Symposia $\underline{13}$, 93 (1979).
9. B. Lengsfield III, J. Chem. Phys., to be published.
10. R. Shepard and J. Simons, Intern. J. Quantum Chem., 1980 Symposia Issue, to be published; R. Shepard and J. Simons, J. Chem. Phys., to be published.
11. C. Nellin and F. Matsen, to be published.

Table I
Multiconfigurational Hartree-Fock calculation for O_2^*

	$^3\Sigma_g^-$		$^1\Sigma_g^+$		$^1\Delta_g$	
iteration	second order appr. orb. trans. E(n)	one step second order E	second order appr. orb. trans. E(n)	one step second order E	second order appr. orb. trans. E(n)	one step second order E
1	-149.6988026494(3)	-149.6988026494	-149.6441430440(3)	-149.6441430440	-149.6682277317(2)	-149.6682277317
2	-149.7370464186(2)	-149.7274007471	-149.6749197928(2)	-149.6662568385	-149.7013552056(1)	-149.6934258436
3	-149.7379417888(2)	-149.7360408815	-149.6759382826(2)	-149.6735764049	-149.7043636264(2)	-149.7018991625
4	-149.7379421299	-149.7379190965	-149.6759392410	-149.6758765596	-149.7043799305	-149.7043337841
5		-149.7379421202		-149.6759392359		-149.7043799292
6		-149.7379421299		-149.6759392411		-149.7043799309
7						

*E is the total energy in atomic units. The number in parenthesis is the second order approximate orbital transformation procedure indicates the number of times the second order (exact or approximate) equations have been applied in between each two electron integral transformation.

GENERAL QUADRATICALLY CONVERGENT MCSCF THEORY IN TERMS OF REDUCED MATRIX ELEMENTS

Clemens C. J. Roothaan
Departments of Physics and Chemistry
University of Chicago
Chicago, Illinois 60637
and
National Resource for Computation in Chemistry
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

and

John H. Detrich
Science Applications Inc.
1211 W. 22nd Street
Oak Brook, Illinois 60521

An improved scheme is presented for calculating multi-configuration self-consistent field wave functions of electronic systems for which the symmetry group is simply reducible. Our formulation is based on expansion of the orbitals in a basis set, but is otherwise entirely general. The MCSCF equations are derived from a multi-state variation principle. In the algebraic development extensive use is made of shell and term replacement operators, permitting a transparent formal development in which symmetry properties are fully exploited. The resulting scheme is manifestly quadratically convergent in the general case. Sufficient detail is presented to provide the basis for efficient computer implementation.

DEDICATION

During the last half century or so our knowledge of electronic structure in atoms and molecules has grown from humble beginnings to an extensive and impressive body of data. While many of these data have come from ingenious and painstaking experimental work, other important information has been obtained from theoretical calculations of steadily increasing sophistication and scope. Prominent among the underpinnings of such calculations for molecules has been -- and still is -- what has been called by its inventor, Robert S. Mulliken, the MO-LCAO method: the electrons are assigned to molecular orbitals, and each molecular orbital is expanded as a linear combination of atomic orbitals. While this basic idea was already very fruitful in semi-empirical calculations, it also provided the corner stone for what is now called the a priori approach, in which an atom or molecule is represented by an all-electron wave function, and the full electronic Hamiltonian is used without further approximations. The developments presented in this paper represent in a certain sense the closing chapter of the formal development of the construction of wave functions by the LCAO principle. We consider it a privilege to dedicate this paper to Professor Mulliken, in honor of his numerous invaluable contributions to theoretical molecular structure.

I. INTRODUCTION

The description of a many-electron atom or molecule by means of approximate wave functions expressed in terms of Slater determinants (SD) built up from one-electron functions, or spin-orbitals (SO), has great appeal. It yields a computationally tractable approach while permitting a straightforward interpretation of the N-electron wave functions in terms of the constituent SO's. The conventional Hartree-Fock self-consistent field (SCF) method for atoms (Hartree, 1957; Froese

Fischer, 1977) uses a single SD, or a minimally necessary set of SD's to satisfy symmetry requirements, and numerically computed orbitals. The SCF concept can also be realized by introducing finite expansions for the SO's in terms of a basis set (Roothaan, 1951, 1960). This expansion technique currently dominates the treatment of molecules, and has also been very successful in the case of single atoms (Roothaan and Bagus, 1963). The present paper constitutes a further development of this expansion technique.

Although the SCF method often yields a good first approximation for the electronic wave function(s) of an atom or molecule, there are many instances when this approximation is inadequate. Among the many different approaches by which one can obtain wave functions of higher quality, the *multi-configuration self-consistent field* (MCSCF) method occupies a unique position as the most natural generalization of the SCF method. This approach preserves the appealing feature of compact wave functions constructed from a relatively small number of orbitals, providing a simple basis for understanding and interpreting atomic and molecular phenomena.

In the MCSCF method one constructs *term state functions* (TSF) as linear combinations of *configuration state functions* (CSF); the latter are defined as symmetry adapted linear combinations of SD's arising from a single configuration of orbital assignments (Hinze and Roothaan, 1967). In the final MCSCF wave function(s) the expansions of both the SO's in terms of basis functions, and of the TSF's in terms of the CSF's, have been optimized, so as to satisfy the variation principle for the chosen state of the system. By contrast, in the method of *configuration interaction* (CI) only the expansions of the TSF's in terms of the CSF's are optimized, while the SO's are usually taken as the solution of some type of SCF calculation.

In the usual MCSCF formulation the variation principle is applied to the energy of a single state. The idea of a *multi-state variational expression* was introduced by Docken and Hinze (1972) and refined by Ruedenberg, Cheung, and Elbert (1979); upon closer examination, it appears that this idea is entirely natural in MCSCF theory. This approach is particularly useful whenever a collection of states, rather than a single state, is relevant for a physical process; examples are radiative transitions, thermodynamical state functions at finite temperatures, target states in scattering situations, etc. The variation principle as formulated in this paper is even more general: it permits also a mixture of different states of ionization of the same species.

The determination of the optimal SO and TSF expansion coefficients, which constitute the MCSCF solution, is an algebraic but nonlinear problem, which must be solved by iteration. Many different schemes for obtaining the MCSCF solution have been proposed and used. These schemes can be roughly classified into four groups. In the first group Lagrange multipliers, which are introduced to maintain orthonormality constraints for the SO's, play a dominant role (Veillard, 1966; Clementi and Veillard, 1966; Das and Wahl, 1966, 1972; Hinze and Roothaan, 1967; Hinze, 1973; Wahl and Das, 1977; Golebiewski, Hinze and Yurtsever, 1979). In the second group the generalized Brillouin theorem (Levy and Berthier, 1968) is the dominant feature, and is invoked to formulate the so-called *super-CI scheme* (Grein and Chang, 1971; Grein and Banerjee, 1975; Banerjee and Grein, 1976, 1977; Chang and Schwartz, 1977; Ruttink and van Lenthe, 1977; Ruedenberg, Cheung and Elbert, 1979). In the third group exponential transformations of the SO's, or equivalent transformations which produce the same results, are introduced; this approach leads naturally to a Newton-Raphson type procedure for the determination of the SO's, and assigns subordinate roles to the Lagrange multipliers and the generalized Brillouin theorem (Levy, 1969, 1970, 1973; Polazzo, 1975; Kuprievich and Schramko, 1975; Kendrick and Hillier, 1976; Dalgaard and Jørgensen, 1978). In the fourth group exponential transformations are introduced for the TSF's as well as for the SO's (Dalgaard, 1979; Yeager and Jørgensen, 1979; Roothaan, Detrich and Hopper, 1979; Lengsfield, 1980; Werner and Meyer, 1980). This new TSF transformation replaces the more usual procedure of solving the TSF secular equation system, which is used in the first three groups. The present paper gives a comprehensive and detailed account of the derivation and solution of the MCSCF equations as they arise in this fourth approach.

The simultaneous exponential transformations of SO's and TSF's permit a much more powerful analytical framework for the MCSCF process than was hitherto possible. Firstly, it permits us to formulate the MCSCF variational problem straightforwardly and concisely in terms of *essential non-redundant variables*. Hence we need not introduce Lagrange multipliers to maintain constraints on the variables. In fact, orbital energies, as well as the total energy, no longer play a pivotal role in the variational process; they can of course be evaluated for monitoring purposes as the calculation proceeds, just like any other interesting quantity. Secondly, considering the variational energy as a Taylor expansion in terms of these essential non-redundant variables, it is easy to write down explicitly the linear *and* the quadratic terms. Truncating this expansion by dropping all higher order terms, the variational problem which remains is solved simply and directly by the multi-dimensional Newton-Raphson procedure. The latter is equivalent to a matrix inversion problem, which replaces the usual combination of the eigenvalue problem for the TSF's and the pseudo-eigenvalue problems for the SO's. The Newton-Raphson process yields only an approximate MCSCF solution because of the truncation, and must therefore be used iteratively. It is however quadratically convergent: generally about 3-4 iterations should suffice to obtain the converged MCSCF solution, which is in sharp contrast with the convergence behavior of the common first order methods.

The double exponential transformation effectively disposes of the two most difficult -- and related -- obstacles encountered in traditional MCSCF calculations: the coupling between the SO and TSF optimizations, and the often painfully slow rate of convergence (or worse, divergence). The coupling between SO and TSF optimization has been discussed by Das (1973); Das, Janis and Wahl (1974); Chang and Schwartz (1977); and Dalgaard and Jørgensen (1978). The idea of dealing with this coupling by use of an exponential transformation for the TSF variation, analogous to the exponential SO variation, was introduced by Dalgaard (1979) and independently by Roothaan, Detrich and Hopper (1979). The conceptual and computational advantages that are realized by expressing the orbital optimization process in terms of a unitary transformation written in exponential form were first recognized by Levy (1969, 1970, 1973), who used this approach to derive an improved multi-dimensional Newton-Raphson process for MCSCF orbital optimization. The *same process was derived, without introducing the exponential* transformation explicitly, by Kendrick and Hillier (1976). Within the framework of pseudo-eigenvalue problems, the work of Hinze and Roothaan (1967) appears to be the first successful attempt to formulate a multi-dimensional Newton-Raphson process for orbital optimization.

We regard symmetry considerations as integral to the MCSCF method. The SO's and TSF's are required to have definite symmetry, so that they belong to irreducible representations of the appropriate atomic or molecular space and spin symmetry groups. Consequently, instead of individual SO's, one actually optimizes *orbital shells*, each consisting of a set of SO's degenerate by virtue of symmetry. Likewise, instead of individual TSF's, one optimizes *spectroscopic terms* consisting of a set of TSF's degenerate due to symmetry. Full exploitation of symmetry reduces the variational problem to be solved numerically to a problem in terms of non-vanishing non-redundant quantities only. Here we accomplish this with a comprehensive and explicit formal algebra, under the assumption that the applicable symmetry group is *simply reducible* (Wigner, 1940, 1941). This covers the most important cases since it applies to atoms as well as to the vast majority of molecules. The generalizations necessary to handle the exceptions are straightforward, but are not presented in this paper.

In a practical MCSCF calculation the choice of the set of configurations, as formally defined in terms of SO's, is of great importance for the success of such a calculation. Currently available experience in this respect has been reviewed by Detrich and Wahl (1980). There are two specific methods of selection which deserve special mention. The first one is the *even-replacement MCSCF model* (Roothaan, Detrich and Hopper, 1979). It permits MCSCF calculations at a total cost hardly more than what is required for an SCF calculation, and it is consequently the most economical MCSCF model. The second one is the *fractionally saturated MCSCF model*, originated by Ruedenberg and Sundberg (1976) under the name *full orbital reaction space MCSCF method*, and recently applied by Roos, Taylor and Siegbahn (1980) under the name *complete active space SCF method*. This model is characterized by its conceptual simplicity, but it leads to very costly calculations.

The choice of configurations, combined with our insistence on symmetry requirements for the TSF's and SO's, introduces considerable complexity into the evaluation of matrix elements between TSF's of the

Coulomb repulsion. This question has been addressed for the case of atoms in the work of Racah (1942, 1943). Further work in this area, and especially the appropriate extension to molecular symmetry groups, should have a high priority in the near future. In this paper we deal with the choice of configurations only to a limited extent, primarily insofar as it has a bearing on the determination of a set of essential non-redundant variables.

Dalgaard's treatment of the double exponential transformation is carried out within the framework of second quantization. In that regime the natural constructs for one- and N-electron wave functions are the SO and SD, respectively. The explicit connections between SO's and SD's, and also the Hamiltonian operator, are written down in terms of creation and annihilation operators. The direct use of CSF's is unnatural; consequently to date all MCSCF formalisms based on second quantization deal with SD's rather than CSF's, and end up with a numerical variational problem which is not in irreducible form. If a further symmetry reduction is desired, it must be imposed a posteriori.

Our present treatment deals directly with the CSF's; wave function manipulations take place in Hilbert space (von Neumann, 1955), and instead of creation and annihilation operators we use the replacement operators introduced by Löwdin (1977). Also distinctive is our extensive use of the Trace operation, and of projection operators. Our treatment leads quite naturally to a numerical variational problem in irreducible form, thereby providing the proper basis for the construction of a computer program of maximum generality, transparency and efficiency.

II. HILBERT SPACE ALGEBRA

The manipulations required for the optimization of our multi-configuration wave functions can best be expressed as formal operations in Hilbert space (von Neumann, 1955). In this section we review the relevant concepts for this, and collect a number of general formulas which will be applied in the following sections to more specific manipulations of orbitals and configurations. In order to state these general facts with minimum complexity, we use in this and the next section a notation which is simpler, and carries less information, than what is needed in the following sections.

Let ψ_i, $i=1,2,...$ constitute a *complete orthonormal base* in the Hilbert space of our wave functions. Orthonormality is expressed by

$$\langle\psi_i|\psi_j\rangle = \delta_{ij} \quad , \tag{1}$$

while completeness permits the expansion of an arbitrary function in terms of the base according to

$$\psi = \Sigma_i \psi_i c_i \quad . \tag{2}$$

Using Eq.(1) we easily find for the expansion coefficients

$$c_i = \langle\psi_i|\psi\rangle \quad . \tag{3}$$

Taking now the scalar product of ψ with another arbitrary wave function ϕ we obtain

$$\langle\phi|\psi\rangle = \Sigma_i\langle\phi|\psi_i\rangle\langle\psi_i|\psi\rangle \quad ; \tag{4}$$

this is conveniently restated and renamed as the *closure property*

$$\Sigma_i|\psi_i\rangle\langle\psi_i| = \mathcal{I} \quad , \tag{5}$$

where $\mathcal{I}$ is the identity operator.

Let now $\mathcal{A}$ be a linear operator defined for the entire Hilbert space. Its matrix elements with respect to the base ψ_i are given by

$$A_{ij} = \langle\psi_i|\mathcal{A}|\psi_j\rangle \quad . \tag{6}$$

Conversely, using the closure property, Eq.(5), we can express the operator in terms of the matrix elements, namely

$$\mathcal{A} = \Sigma_{ij}\mathcal{P}_{ij}A_{ji} \quad , \tag{7}$$

where we have introduced the *replacement operators* (Löwdin, 1977)

$$\mathcal{P}_{ij} = |\psi_j\rangle\langle\psi_i| \quad . \tag{8}$$

According to Eq.(7) any linear operator $\mathcal{A}$ can be expanded in terms of the replacement operators $\mathcal{P}_{ij}$, with the matrix elements A_{ji} as the expansion coefficients. Hence these replacement operators play an analogous role with respect to operators as do the basis functions ψ_i with respect to wave functions.

The matrix elements of the replacement operators are easily seen to be

$$P_{ij,kl} = \langle\psi_k|\mathcal{P}_{ij}|\psi_l\rangle = \delta_{il}\delta_{kj} \quad . \tag{9}$$

For the product of two replacement operators we find

$$\mathcal{P}_{ij}\mathcal{P}_{kl} = \delta_{il}\mathcal{P}_{kj} \quad , \tag{10}$$

and consequently for the corresponding *commutator*

$$[\mathcal{P}_{ij},\mathcal{P}_{kl}] = \mathcal{P}_{ij}\mathcal{P}_{kl} - \mathcal{P}_{kl}\mathcal{P}_{ij} = \delta_{il}\mathcal{P}_{kj} - \delta_{kj}\mathcal{P}_{il} \quad . \tag{11}$$

For any operator $\mathcal{A}$ there exists a *Hermitian conjugate* operator $\mathcal{A}^*$ defined by the requirement

$$\overline{\langle\psi|\mathcal{A}^*|\phi\rangle} = \langle\phi|\mathcal{A}|\psi\rangle \quad , \tag{12}$$

where ψ and ϕ are any two functions in Hilbert space. It follows easily that

$$\mathcal{P}^*_{ij} = \mathcal{P}_{ji} \quad , \tag{13}$$

and

$$\mathcal{A}^* = \Sigma_{ij}\mathcal{P}_{ij}A^*_{ji} \quad , \tag{14}$$

where

$$A^*_{ij} = \overline{A}_{ji} \quad . \tag{15}$$

Given an algebraic expression containing operators and ordinary (complex) numbers, the Hermitian conjugate expression is obtained by taking the complex conjugate of all numbers, the Hermitian conjugate of all operators, and all operator products in reverse order. Furthermore any equation involving operators remains valid if we substitute for the operators the corresponding matrices, and vice versa.

There are certain types of operators which have special significance. They are distinguished according to properties with respect to their Hermitian conjugates and/or their eigenvalues, say λ, namely.

Hermitian: $\mathcal{H}^* = \mathcal{H}$, $H^*_{ij} = H_{ij}$, $\overline{\lambda} = \lambda$, (16)
anti-Hermitian: $\mathcal{X}^* = -\mathcal{X}$, $X^*_{ij} = -X_{ij}$, $\overline{\lambda} = -\lambda$, (17)
unitary: $\mathcal{U}^*\mathcal{U} = \mathcal{I}$, $\Sigma_k U_{ik}U_{kj} = \delta_{ij}$, $|\lambda| = 1$, (18)
idempotent: $\mathcal{M}^2 = \mathcal{M}$, $\Sigma_k M_{ik}M_{kj} = M_{ij}$, $\lambda = 1$ or $\lambda = 0$, (19)
nilpotent: $\mathcal{N}^2 = \mathcal{O}$, $\Sigma_k N_{ik}N_{kj} = 0$, $\lambda = 0$, (20)

where $\mathcal{O}$ denotes the null operator. It is easily seen that the Hermitian conjugate of the various operators are of the same type: $\mathcal{H}^*$ is Hermitian, $\mathcal{X}^*$ is anti-Hermitian, etc. Furthermore, $i\mathcal{H}$ is anti-Hermitian, and $i\mathcal{X}$ is Hermitian. Note also that $\mathcal{I}$ is Hermitian, unitary and idempotent, while $\mathcal{O}$ is Hermitian, anti-Hermitian, idempotent and nilpotent.

In a Hilbert space of finite dimension we can always choose the eigenfunctions of any Hermitian, anti-Hermitian or unitary operator so that they constitute a complete orthonormal base. In a Hilbert space of infinite dimension that statement needs to be modified somewhat; this is due to complications caused by the possibility of a continuous range of eigenvalues. On the other hand, in either a finite or an infinite Hilbert space, the eigenfunctions of an idempotent or nilpotent operator in general cannot be chosen orthogonal, nor do they always span all of Hilbert space.

The exponential of an anti-Hermitian operator yields a unitary operator, namely

$$\mathcal{U} = e^{\mathcal{X}} = \mathcal{I} + \mathcal{X} + \tfrac{1}{2}\mathcal{X}^2 + \ldots \quad . \tag{21}$$

Conversely, given any unitary operator $\mathcal{U}$ we can always find an anti-Hermitian operator $\mathcal{X}$ so that Eq.(21) is valid. Although these facts are rather well known, for the sake of completeness we present the proofs in Appendix A.

The unitary operators are particularly important for describing *transformations within Hilbert space*. When all wave functions and operators are subjected to a transformation by the same unitary operator $\mathcal{U}$, the transformed wave functions and operators are obtained by the formal definition

$$\left.\begin{aligned} \psi_i' &= \mathcal{U}\psi_i \quad , \\ \mathcal{A}' &= \mathcal{U}\mathcal{A}\mathcal{U}^* \quad ; \end{aligned}\right\} \tag{22}$$

it is well known that all relations between wave functions and operators also hold for the corresponding transformed quantities.

The role of the matrix elements U_{ij} is two-fold. On the one hand, we readily find for the transformed basis functions and replacement operators

$$\left.\begin{aligned} \psi_i' &= \Sigma_j \psi_j U_{ji} \quad , \\ \mathcal{P}_{ij}' &= \Sigma_{k\ell} U_{ik}^* \mathcal{P}_{k\ell} U_{\ell j} \quad . \end{aligned}\right\} \tag{23}$$

On the other hand, if we express arbitrary functions and operators, say ψ and $\mathcal{A}$, and also their transforms, ψ' and $\mathcal{A}'$, in terms of the *original* reference base ψ_i and operators $\mathcal{P}_{ij}$, namely

$$\left.\begin{aligned} \psi &= \Sigma_i \psi_i c_i \quad , \\ \psi' &= \Sigma_i \psi_i c_i' \quad , \\ \mathcal{A} &= \Sigma_{ij} \mathcal{P}_{ij} A_{ji} \quad , \\ \mathcal{A}' &= \Sigma_{ij} \mathcal{P}_{ij} A_{ji}' \quad , \end{aligned}\right\} \tag{24}$$

we find easily

$$\left.\begin{aligned} c_i' &= \Sigma_j U_{ij} c_j \quad , \\ A_{ij}' &= \Sigma_{k\ell} U_{ik} A_{k\ell} U_{\ell j}^* \quad . \end{aligned}\right\} \tag{25}$$

Eqs.(23) express transformations of *geometrical objects*, and are sometimes called *active transformations*. In contrast, Eqs.(25) express transformations of *coordinates*, and are sometimes called *passive transformations*.

An all-important scalar function of an operator is its *trace* defined by

$$\mathrm{Trace}(\mathcal{A}) = \Sigma_i A_{ii} \quad . \tag{26}$$

We observe that this definition is independent of the basis chosen, as of course it must be if we want to speak of a function of an operator.

The trace is a linear function with respect to both scalars and operators, namely

$$\mathrm{Trace}(\alpha\mathcal{A} + \beta\mathcal{B}) = \alpha\mathrm{Trace}(\mathcal{A}) + \beta\mathrm{Trace}(\mathcal{B}) \quad , \tag{27}$$

where $\mathcal{A}$ and $\mathcal{B}$ are any two operators, and α and β are scalars. With respect to a product of operators, the Trace is invariant under cyclical permutations of the factors in the product. It suffices to state

$$\mathrm{Trace}(\mathcal{A}\mathcal{B}\ldots\mathcal{M}) = \mathrm{Trace}(\mathcal{B}\ldots\mathcal{M}\mathcal{A}) \quad , \tag{28}$$

since repeated application of Eq.(28) can yield arbitrary cyclical permutations.

Any Hermitian, anti-Hermitian or unitary operator can always be brought into diagonal form by choosing an appropriate basis. Since the diagonal elements are then the eigenvalues, we obtain in this case

$$\mathrm{Trace}(\mathcal{A}) = \Sigma_i \lambda_i \quad , \tag{29}$$

where $\mathcal{A}=\mathcal{H}$, $\mathcal{A}=\mathcal{K}$ or $\mathcal{A}=\mathcal{U}$, and $\lambda_1, \lambda_2, \ldots$ are the eigenvalues.

We obtain another interesting use of the trace by writing the general matrix element A_{ij} of an operator $\mathcal{A}$ in the form

$$A_{ij} = \mathrm{Trace}(\mathcal{A}\mathcal{P}_{ij}) \quad . \tag{30}$$

Eq.(30) provides the basis for formulating quantum mechanics as a trace algebra (Löwdin, 1977).

If an operator $\mathcal{P}$ is both Hermitian and idempotent, it is called a projection operator, or projector:

$$\mathcal{P}^2 = \mathcal{P} = \mathcal{P}^\dagger \quad . \tag{31}$$

The eigenvalues of $\mathcal{P}$ are restricted to 1 and 0; the corresponding eigenfunctions span the image space and the shadow space of $\mathcal{P}$, respectively. These two subspaces have a null intersection, and their union constitutes the entire Hilbert space; we say that the projector $\mathcal{P}$ defines a decomposition of Hilbert space into two complementary subspaces.

It is easily seen that, given any projector $\mathcal{P}$, the operator $\mathcal{I}-\mathcal{P}$ is also a projector; it yields the same decomposition of Hilbert space as does $\mathcal{P}$, but with the roles of image space and shadow space reversed. Thus it is also true that the image spaces of $\mathcal{P}$ and $\mathcal{I}-\mathcal{P}$ are complementary; we say therefore that $\mathcal{P}$ and $\mathcal{I}-\mathcal{P}$ are complementary projectors. A special case of such a complementary pair is provided by $\mathcal{I}$ and $\mathcal{O}$.

The dimension of the image space, say d, is called the dimension of the projector. Because of Eq.(29) we have

$$d = \mathrm{Trace}(\mathcal{P}) \quad . \tag{32}$$

Obviously the null operator $\mathcal{O}$ is the (one and only) trivial projector of dimension 0; projectors of dimension 1 are called primitive; all other projectors are compound.

Two projectors, say $\mathcal{P}$ and $\mathcal{Q}$, are called orthogonal when their image spaces are orthogonal; this occurs if and only if $\mathcal{P}\mathcal{Q}=\mathcal{Q}\mathcal{P}=\mathcal{O}$. Clearly a complementary pair of projectors is also an orthogonal pair, but the converse is not necessarily true. It is useful to generalize the notion of an orthogonal pair to an orthogonal set of n projectors $\mathcal{P}_\alpha$, $\alpha=1,2,\ldots,$ which satisfy

$$\begin{aligned} \mathcal{P}_\alpha^\dagger &= \mathcal{P}_\alpha \quad , \\ \mathcal{P}_\alpha\mathcal{P}_\beta &= \delta_{\alpha\beta}\mathcal{P}_\alpha \quad . \end{aligned} \tag{33}$$

Given the orthogonal set of projectors $\mathcal{P}_\alpha$, it follows easily that

$$\mathcal{P} = \Sigma_\alpha \mathcal{P}_\alpha \tag{34}$$

is again a projector. Conversely Eq.(34) gives the decomposition of the projector $\mathcal{P}$ into the components $\mathcal{P}_\alpha$, the latter forming an orthogonal set of projectors of lesser dimensions. A fundamental theorem of Hilbert space algebra states that any n-dimensional projector can be decomposed into n orthogonal primitive projectors.

We return now to the replacement operators $\mathcal{P}_{ij}$ introduced before, Eq.(8). The off-diagonal ones, $\mathcal{P}_{ij}$ with $i \neq j$, are nilpotent; they are often called <u>shift operators</u>. The diagonal ones, $\mathcal{P}_{ii}$ for $i=1,2,\ldots$, form an orthogonal set of primitive projectors. The closure property, Eq.(5), can now be restated as the <u>decomposition of the identity into primitive projectors</u>, namely

$$\mathcal{J} = \Sigma_i \mathcal{P}_{ii} \quad . \tag{35}$$

Essentially all our wave function manipulations will be confined to a subspace of Hilbert space. Adopting a suitable basis, this subspace is spanned by a subset of the basis functions ψ_i, say for $i \in S$. The subspace is obviously the image space of the projector

$$\mathcal{P} = \Sigma_{i \in S} \mathcal{P}_{ii} \quad . \tag{36}$$

For any function ψ in the subspace the operator $\mathcal{P}$ acts like the identity, namely

$$\mathcal{P}\psi = \psi \quad , \tag{37}$$

while for any function outside of the subspace, that is, in the shadow space, $\mathcal{P}$ acts like the null operator.

Given a projector $\mathcal{P}$, we define for any operator $\mathcal{A}$ the <u>associated projected operator</u> $\hat{\mathcal{A}}$ by means of

$$\hat{\mathcal{A}} = \mathcal{P}\mathcal{A}\mathcal{P} \quad ; \tag{38}$$

it is easily seen that this is equivalent to

$$\left.\begin{aligned} \hat{A}_{ij} &= A_{ij} \quad , \quad i \in S \quad \text{and} \quad j \in S \quad . \\ \hat{A}_{ij} &= 0 \quad , \quad i \notin S \quad \text{or} \quad j \notin S \quad . \end{aligned}\right\} \tag{39}$$

Given a projector $\mathcal{P}$, if an operator $\mathcal{A}$ is identical with its associated projected operator $\hat{\mathcal{A}}$, namely

$$\hat{\mathcal{A}} = \mathcal{A} \quad , \tag{40}$$

we say that the operator is <u>internal</u> to the image space of $\mathcal{P}$. The condition (40) is equivalent to

$$\mathcal{A}\mathcal{P} = \mathcal{P}\mathcal{A} = \mathcal{A} \quad , \tag{41}$$

and also to

$$A_{ij} = 0 \quad , \quad i \notin S \quad \text{or} \quad j \notin S \quad . \tag{42}$$

From Eq.(41) follows the commutation relation for an internal operator and the projector defining the subspace, namely

$$[\mathcal{A},\mathcal{P}] = \mathcal{O} \quad ; \tag{43}$$

it should be noted however that Eq.(43) is not a sufficient condition that $\mathcal{A}$ is internal to the image space of $\mathcal{P}$.

The most important application of projected operators is the following. Let $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$,... be a set of operators internal to the image space of the projector $\mathcal{P}$, and let $\mathcal{X}$, $\mathcal{Y}$,... be any operators in general. Then in any operator product in which each general operator is flanked on both sides by an internal operator, we can replace the general operators by their associated projected operators; for instance

$$\mathcal{A}\mathcal{X}\mathcal{B}\mathcal{Y}\mathcal{C} = \mathcal{A}\hat{\mathcal{X}}\mathcal{B}\hat{\mathcal{Y}}\mathcal{C} \quad . \tag{44}$$

An interesting corollary of Eq.(44) concerns the trace of a product of operators. Since the trace involves a cyclical product, we have for instance

$$\text{Trace}(\mathcal{A}\mathcal{X}\mathcal{B}\mathcal{Y}) = \text{Trace}(\mathcal{A}\hat{\mathcal{X}}\mathcal{B}\hat{\mathcal{Y}}) \quad , \tag{45}$$

because $\mathcal{X}$ is flanked by $\mathcal{A}$ and $\mathcal{B}$, and $\mathcal{Y}$ is flanked by $\mathcal{B}$ and $\mathcal{A}$. A special case of this type of relation occurs in the evaluation of matrix elements of such a product of operators between wave functions, say ψ

and ϕ, which are confined to the image space of $\mathcal{P}$. In this case the operator

$$\mathcal{B} = |\psi\rangle\langle\phi| \tag{46}$$

is internal to the image space of $\mathcal{P}$, hence

$$\langle\phi|\mathcal{H}|\psi\rangle = \mathrm{Trace}(\mathcal{H}\mathcal{B}) = \mathrm{Trace}(\hat{\mathcal{H}}\mathcal{B}) = \langle\phi|\hat{\mathcal{H}}|\psi\rangle \quad . \tag{47}$$

An obvious generalization of Eq.(47) is

$$\langle\phi|\mathcal{A}\mathcal{H}\mathcal{B}|\psi\rangle = \langle\phi|\mathcal{A}\hat{\mathcal{H}}\mathcal{B}|\psi\rangle \quad , \tag{48}$$

which will be particularly useful later on.

III. THE MULTI-STATE VARIATION PRINCIPLE

Ordinarily one determines (a) variational wave function(s) for a bound state by demanding that the expectation value of the energy for that state, expressed in terms of (a) wave function(s) with adjustable parameters, becomes stationary to first order for any permissible infinitesimal variations of these parameters. In this section we generalize this variation principle, so that we can determine a set of wave functions which jointly describe a finite number of states of our system (Docken and Hinze, 1972; Ruedenberg, Cheung and Elbert, 1979).

Let ψ_i, i=1,2,...,n be the orthonormal set of wave functions to be subjected to the variation principle. In case any of the states under consideration is degenerate because of symmetry, the set ψ_i must of course include a full set of degenerate partners satisfying the appropriate constraints.

For the variational energy expression we adopt a weighted average of the energies of the various states given by

$$E = \Sigma_i w_i \langle\psi_i|\mathcal{H}|\psi_i\rangle \quad , \tag{49}$$

where $\mathcal{H}$ is the Hamiltonian of the system, and the weights w_i are restricted by

$$\left.\begin{aligned} \bar{w}_i = w_i &> 0 \quad , \\ \Sigma_i w_i &= 1 \quad . \end{aligned}\right\} \tag{50}$$

For degenerate states we expect of course to choose these weights identical for the degenerate partners.

The energy must be stationary with respect to all variations we permit for the wave functions ψ_i. Among these variations we must of course include unitary transformations of the set ψ_i, analogous to what we must permit for the wave functions of a degenerate state when applying a symmetry-breaking perturbation. Hence we must allow the transformation

$$\psi_i' = e^X\psi_i = \Sigma_j \psi_j (e^X)_{ji} \quad , \tag{51}$$

which leads to a variation of the energy given by

$$\delta E = E' - E = \Sigma_{ij,\, i\neq j} (w_i - w_j)\langle\psi_i|\mathcal{H}|\psi_j\rangle X_{ji} + O_2(X_{ij}) \quad . \tag{52}$$

the last term representing all terms of second and higher order in the parameters X_{ij}.

Assuming first all the weights to be different, the requirement that δE must vanish to first order for any permissible choice of the parameters X_{ij} yields

$$\langle\psi_i|\mathcal{H}|\psi_j\rangle = 0 \quad , \quad i \neq j \quad . \tag{53}$$

The proof of this statement is not quite trivial, since the X_{ij} are not completely independent; it is presented in Appendix B.

If ψ_i and ψ_j are degenerate partners due to symmetry, one must of course choose $w_i = w_j$, and Eq.(53) no longer follows from the variation principle. However, in this case Eq.(53) is a consequence of the symmetry constraints on the wave functions which must be honored by the variational process.

If ψ_i and ψ_j are not degenerate partners due to symmetry, and we still, choose $w_i = w_j$, Eq.(53) follows neither from the variation principle nor from symmetry constraints; however, we can then guarantee Eq.(53) by carrying out an additional unitary transformation of the wave functions which leaves the variational energy invariant. Clearly it is only in this last form that the wave functions can be identified with physical states.

Hence having adopted Eq.(49) as the variational energy expression, we have obtained the eminently reasonable result that the variational condition implies that the wave functions must satisfy, or may be chosen so as to satisfy the usual secular equation system. It is furthermore interesting to note that this fact does not depend on the choice for the magnitudes of the weight factors. On the other hand, the magnitudes of the weight factors will affect the variation with respect to parameters other than X_{ij}.

Finally, we note that our variation principle reduces to the standard variation principle for a single non-degenerate state for $\mathcal{N}=1$, and to the proper symmetry-weighted one for a single degenerate state, choosing of course equal weights for the different partners.

IV. CONSTRUCTION OF N-ELECTRON WAVE FUNCTIONS FROM ONE-ELECTRON SPIN-ORBITALS

The N-electron wave functions we use to describe stationary states of atoms and/or molecules are built from a finite set of M one-electron spin-orbitals (SO), where of course $M > N$. The SO's are denoted by $\phi_{\lambda i \alpha}(a)$; this notation is explicit with respect to symmetry characteristics and electron coordinates, namely

λ (or $\mu, \nu, \ldots$) designates *symmetry species*;

α (or $\beta, \gamma, \ldots$) designates *symmetry subspecies*;

i (or $j, k, \ldots$) labels spin-orbitals not distinguishable by symmetry;

a (or $b, c, \ldots$) labels electrons, representing space and spin coordinates.

For any given λ there are definite ranges for α and for i; we say that α and i are *subordinate* to λ. The range of α is the degree of degeneracy d_λ; ordinarily this will encompass spatial as well as spin degeneracy. For given λi, the set of d_λ degenerate SO's is called a *shell*: thus the compound index λi is a shell index. We establish a *canonical order* for the SO's by taking them in *dictionary order* with respect to the compound index $\lambda i \alpha$; in this arrangement the SO's are grouped in shells, and shells occur grouped by symmetry.

In general, matrices referring to shells have the index structure $\lambda i \mu j$. A special case is that of matrices which are *blocked by symmetry*. In those matrices the elements with $\lambda \neq \mu$ either vanish or are meaningless; the non-vanishing elements are conveniently indexed by $\lambda i j$ rather than $\lambda i \lambda j$.

It is convenient and customary to require that the SO's form an orthonormal set. We adhere to this practice, so that

$$\langle \phi_{\lambda i \alpha}(a) | \phi_{\mu j \beta}(a) \rangle = \delta_{\lambda i \alpha, \mu j \beta} = \delta_{\lambda \mu} \delta_{ij} \delta_{\alpha \beta} \quad . \tag{54}$$

Clearly, the SO's $\phi_{\lambda i \alpha}(a)$ span an M-dimensional subspace of the Hilbert space of one-electron functions of electron a; in general, a function in this subspace will be designated by $\phi(a)$.

Choosing now any N-membered *ordered subset* from our canonical set of SO's, we can construct the corresponding N-electron antisymmetrized product or *Slater determinant* (SD) according to

$$\Phi_{\lambda_1 i_1 \alpha_1 \ldots \lambda_N i_N \alpha_N} = (N!)^{\frac{1}{2}} \mathcal{A} \phi_{\lambda_1 i_1 \alpha_1}(1) \ldots \phi_{\lambda_N i_N \alpha_N}(N) \quad , \tag{55}$$

where $\mathcal{A}$ is the *antisymmetrizer*. Specifically, $\mathcal{A}$ is the projector defined by

$$\mathcal{A} = (N!)^{-1} \Sigma_{r=1}^{N!} \varepsilon_r \mathcal{P}_r \quad , \tag{56}$$

where $\mathcal{S}_P$ represents the distinct permutations of the electron coordinates, with $\delta_P=1$ for even, and $\delta_P=-1$ for odd permutations.

Note that we demand that the SD's *in Eq.(55) are an* _ordered_ subset of the full set. As a result the SD's defined by Eq.(55) are independent; it is easily seen that they are in fact orthonormal. The total number of SD's that can be constructed in this manner is given by the binomial coefficient $\binom{M}{N}$; they span the _SO product space_, it being understood that only antisymmetrized products of SO's are considered. In general, a function in this space will be designated by Ψ or Φ. The N-electron wave function(s) for our atom or molecule is (are) to be found in this SO product space, which is, of course, a subspace of the complete Hilbert space for our problem.

Each SD is at least partially characterized by its _shell occupation numbers_ $n_{\lambda i}$, which specify how many SO's of each shell were used to construct the SD. Obviously these occupation numbers must satisfy

$$\left.\begin{array}{l} 0 \leqslant n_{\lambda i} \leqslant d_\lambda \quad , \\ \Sigma_{\lambda i} n_{\lambda i} = N \quad . \end{array}\right\} \tag{57}$$

Each distinct solution of Eqs.(57) specifies a _configuration_, which is defined as the set of SD's associated with that solution.

In general a configuration spans a reducible representation of the symmetry group of the system. By subjecting the SD's of a configuration to a suitable unitary transformation, we can decompose this reducible representation into irreducible components. The linear combinations of SD's which result from this decomposition are called _configuration state functions_ (CSF). To do justice to the symmetry properties we denote the CSF's by $\Phi_{\Lambda P\lambda}$, analogous to $\phi_{\lambda i\mu}$ for the SO's; here Λ (or Ξ,...) refers to symmetry species, λ (or ξ,...) refers to symmetry subspecies, and P (or Q,...) labels CSF's not distinguishable by symmetry. Since the CSF's are obtained from the SD's by a unitary transformation, they form *a new orthonormal base for the SO product space*, hence

$$\langle\Phi_{\Lambda P\lambda}|\Phi_{\Xi Q\xi}\rangle = \delta_{\Lambda P\lambda,\Xi Q\xi} = \delta_{\Lambda\Xi}\delta_{PQ}\delta_{\lambda\xi} \quad . \tag{58}$$

And because the CSF's have definite symmetry properties, they, rather than the SD's, are the most useful building blocks for wave functions of actual states.

In the standard version of _self-consistent field_ (SCF) theory, the atomic or molecular wave function(s) for a particular spectroscopic term is (are) taken to be a single CSF (a degenerate set of CSF's), and the SO's are to be optimized. In the most commonly used version of _multi-configuration self-consistent field_ (MCSCF) theory one uses in general superpositions of CSF's for a spectroscopic term, and the CSF expansion coefficients as well as the SO's are to be optimized. In the present version of MCSCF theory we represent simultaneously _several spectroscopic terms_ by superpositions of CSF's. In a practical MCSCF calculation one rarely uses all possible CSF's; instead one chooses a fairly limited set of CSF's guided by physical reasoning, past experience and/or intuition.

From our chosen set of CSF's we construct _term state functions_ (TSF) according to

$$\Psi_{\Lambda I\lambda} = \Sigma_P \Phi_{\Lambda P\lambda} C_{\Lambda PI} \quad ; \tag{59}$$

for given ΛI, the set of degenerate functions $\Psi_{\Lambda I\lambda}$ represents a spectroscopic term. Demanding orthonormality for our TSF's we obtain

$$\langle\Psi_{\Lambda I\lambda}|\Psi_{\Lambda J\lambda}\rangle = \Sigma_P C^*_{\Lambda IP} C_{\Lambda PJ} = \delta_{IJ} \quad . \tag{60}$$

The particular MCSCF strategy we want to use requires that we consider at all times a set of TSF's which spans the same space as does the set of chosen CSF's. In actuality of course only one (or a few) of these functions is (are) used to represent a spectroscopic term (or terms); the larger set is needed for carrying out transformations to achieve successive improvements. Therefore the coefficients $C_{\Lambda PI}$ form a square unitary matrix which is furthermore blocked by symmetry.

Since in the construction of TSF's from CSF's the species label Λ is preserved, we may permit the number of electrons to be a function of Λ, say N_Λ. Hence our MCSCF model is used to construct in general, from a common set of SO's, many-electron wave functions representing several spectroscopic terms of different excitation and/or ionization. Thus the species label Λ now classifies wave functions as to spatial symmetry, spin symmetry, and degree of ionization. This further generalization of the multi-state variation principle of the previous section is feasible and profitable since wave functions with different numbers of electrons are now related through a common set of SO's. The shell occupation numbers for the CSF's are now properly denoted by $n_{\Lambda P,\lambda i}$; they satisfy of course

$$\left.\begin{array}{l} 0 \leq n_{\Lambda P,\lambda i} \leq d_\lambda \quad . \\ \Sigma_{\lambda i} n_{\Lambda P,\lambda i} = N_\Lambda \quad . \end{array}\right\} \tag{57'}$$

In the ideal MCSCF formulation the variation of the SO's is restricted only by orthonormality. In this paper we restrict ourselves in addition to the mathematically more tractable expansion regime for the SO's. We adopt a set of one-electron basis functions denoted by $\chi_{\lambda p\alpha}(a)$, in obvious analogy to the SO's $\phi_{\lambda i\alpha}(a)$; p (or $q, r, \ldots$) labels basis functions not distinguishable by symmetry, and is subordinate to λ.

In general, basis functions are only orthogonal if they have different symmetry; in fact their scalar products are given by

$$\langle\chi_{\lambda p\alpha}(a)|\chi_{\mu q\beta}(a)\rangle = \delta_{\lambda\alpha,\mu\beta} S_{\lambda pq} \quad . \tag{61}$$

where $S_{\lambda pq}$ is an element of the overlap matrix, which, as the notation implies, is blocked by symmetry. The overlap matrix is Hermitian, namely

$$S^{*}_{\lambda pq} = S_{\lambda pq} \quad . \tag{62}$$

Furthermore if the basis functions are linearly independent, as they of course must be, the overlap matrix is also positive definite, that is

$$\Sigma_{pq} \bar{c}_{\lambda p} S_{\lambda pq} c_{\lambda q} > 0 \tag{63}$$

for any non-trivial set of coefficients $c_{\lambda p}$ (trivial would be $c_{\lambda p}=0$ for all p).

The SO's are now put forward as expansions in terms of the basis set, namely

$$\phi_{\lambda i\alpha}(a) = \Sigma_p \chi_{\lambda p\alpha}(a) c_{\lambda pi} \quad ; \tag{64}$$

orthonormality of the SO's yields for the expansion coefficients the constraints

$$\Sigma_{pq} c^{*}_{\lambda ip} S_{\lambda pq} c_{\lambda qj} = \delta_{ij} \quad . \tag{65}$$

The MCSCF strategy with respect to the SO's is analogous to that for the TSF's. We require that the SO's $\phi_{\lambda i\alpha}$ and the basis functions $\chi_{\lambda p\alpha}$ span the same space, and carry out linear transformations of the SO's to achieve successive improvements. Hence the coefficients $c_{\lambda pi}$ are related by unitary transformations in order that Eq.(65) remains valid.

V. SHELL STRUCTURE AND MCSCF MODELS

In order to carry out an MCSCF calculation on a particular atom or molecule with a given one-electron basis set, we still have to supply a certain amount of structural information which remains constant throughout the MCSCF variational process. This structural information consists of the number and symmetry species of the spectroscopic terms, and their weights, in the variational energy formula; the number and all the relevant particulars of the CSF's to be used for the construction of the TSF's; and, in a wider sense, all the quantities derived from this which are used and remain constant during the MCSCF variational process. The specification of this structural information will be referred to as the choice of an *MCSCF model*. The choice of appropriate MCSCF models for particular physical systems or processes is a very extensive subject, much wider in scope than for instance the choice of one-electron basis functions. In this paper we touch upon the questions associated with the different models only insofar as they affect the organization of the formal algebra to be presented.

To a considerable extent, different MCSCF models are distinguished from one another by their invariance characteristics with respect to *permissible* SO and TSF transformations. These permissible transformations are unitary, and they preserve the exact symmetry -- species and subspecies -- of the SO's and TSF's; in the following "transformation" will always mean "permissible transformation", unless explicitly specified otherwise.

In order to discuss invariance with respect to SO transformations lucidly, it is useful to define, at three different levels, fully occupied or *closed shells*; partially occupied or *fractional shells*; and unoccupied, virtual or *empty shells*.

At the first level we define these categories for a single CSF (or a set of degenerate CSF's); clearly this is the definition appropriate for, and familiar from, ordinary SCF theory. Denoting the three categories by $C_{\Lambda P}$, $F_{\Lambda P}$ and $E_{\Lambda P}$, we have

$$\left.\begin{array}{llll} \lambda i \in C_{\Lambda P} & \text{if} & n_{\Lambda P,\lambda i} = d_\lambda & , \\ \lambda i \in F_{\Lambda P} & \text{if} & 0 < n_{\Lambda P,\lambda i} < d_\lambda & , \\ \lambda i \in E_{\Lambda P} & \text{if} & n_{\Lambda P,\lambda i} = 0 & . \end{array}\right\} \quad (66)$$

At the second level we define the corresponding categories for the set of all CSF's -- and therefore also all TSF's -- of the same species, say Λ. This definition is appropriate for the most common MCSCF formulation. The closed and empty shells are defined as those that are closed and empty for all CSF's of species Λ; the fractional shells are the remainder. Thus a shell which is closed in some, and empty in the other CSF's of species Λ, becomes a fractionally occupied shell at the present level. Denoting the present categories by C_Λ, F_Λ and E_Λ, they are conveniently defined in set-theoretical language by

$$\left.\begin{array}{ll} C_\Lambda = \bigcap_P C_{\Lambda P} & , \\ E_\Lambda = \bigcap_P E_{\Lambda P} & , \\ F_\Lambda = \sim C_\Lambda \cup E_\Lambda & . \end{array}\right\} \quad (67)$$

where $\sim$ represents *complementation*, $\cup$ represents the *union* operation, and $\bigcap_P$ represents the *intersection* operation to be repeated over all P.

Analogous to the second level, at the third level we define the categories for all CSF's -- and TSF's -- regardless of species. This definition is appropriate for our present most general MCSCF formulation. Denoting the present categories by C, F and E, we have

$$\left.\begin{array}{ll} C = \bigcap_\Lambda C_\Lambda = \bigcap_{\Lambda P} C_{\Lambda P} & , \\ E = \bigcap_\Lambda E_\Lambda = \bigcap_{\Lambda P} E_{\Lambda P} & , \\ F = \sim C_\Lambda \cup E_\Lambda & . \end{array}\right\} \quad (68)$$

We note that, progressing from the first through the third level, the sets of closed and empty shells generally decrease in size, while the fractional set increases, namely

$$\left.\begin{array}{ll} C_{\Lambda P} \supset C_\Lambda \supset C & , \\ E_{\Lambda P} \supset E_\Lambda \supset E & , \\ F_{\Lambda P} \subset F_\Lambda \subset F & . \end{array}\right\} \quad (69)$$

In this work we are primarily concerned with the first and third levels. The third level set of closed shells, C, is commonly called the core.

At the first level, the CSF's $\Phi_{\Lambda P A}$ are invariant under all SO transformations entirely within $C_{\Lambda P}$ or entirely within $E_{\Lambda P}$, and occasionally also under some, or perhaps all, SO transformations within $F_{\Lambda P}$. The best known case where the CSF's are invariant under all SO transformations within $F_{\Lambda P}$ occurs when $F_{\Lambda P}$ consists solely of half-filled shells with all spins parallel; we express this formally by

$$F_{\Lambda P} = H \quad . \tag{70}$$

An example of this situation is provided by the $4s3d^5$, ^7S description of the ground state of the chromium atom.

Progressing from the first through the second to the third level, we note that all CSF's, and consequently all TSF's and the MCSCF energy, will be invariant under all SO transformations entirely within $\mathbf{C}$ or entirely within $\mathbf{E}$, and occasionally also under some SO transformations within $\mathbf{F}$. An obvious example of the latter is the half-shell model. In this model all participating CSF's have the same set of half-filled shells with parallel spins, that is, Eq.(70) holds for all ΛP. Clearly $\mathbf{H}$ is then a subset of $\mathbf{F}$; we write

$$\mathbf{F} = \mathbf{H} + \mathbf{F}' \quad , \tag{71}$$

where $\mathbf{F}'$ consists of those shells which are fractional at the third level, but closed or empty at the first level in the individual TSF's.

So far we have described invariance of the MCSCF process due to invariance of the individual CSF's under SO transformations. Since the TSF's are constructed from CSF's by a linear expansion as yet to be optimized, the MCSCF process will also exhibit invariance when an SO transformation leaves the linear manifold of CSF's, rather than the individual CSF's, invariant; we say that such an SO transformation induces a linear transformation of the CSF's among themselves. In general of course CSF transformations induced by SO transformations do not have this special property.

The most extreme case of this type of invariance occurs when we let the entire SO product space, or at least its maximal subspaces of the required symmetry species, participate in the MCSCF process. In this case all orthonormal sets of SO's which satisfy the symmetry requirements are equivalent; the problem of SO optimization is trivially solved by retaining whatever SO's we have in hand, and the MCSCF solution is obtained by carrying out a straightforward configuration interaction (CI) calculation. Clearly, since SO optimization no longer plays any role whatsoever, this MCSCF model is somewhat of a perversion of the MCSCF concept. In addition to this, for all but very small systems with only a handful of electrons, the model is highly impractical due to an unmanageably large number of CSF's.

A somewhat more practical case of this type of invariance occurs in the fractionally saturated MCSCF model (Ruedenberg and Sundberg, 1976; Roos, Taylor and Siegbahn, 1980): given an assignment of SO's for the core $\mathbf{C}$ and the empty set $\mathbf{E}$, we use the SO's of the fractional set $\mathbf{F}$ to construct all possible CSF's of the required symmetry species. In this model the MCSCF process is invariant under all SO transformations within $\mathbf{F}$, as well as under those within $\mathbf{C}$ or within $\mathbf{E}$. While the model is at first sight attractive because of its conceptual simplicity, it also suffers from too many CSF's unless one uses only a handful of fractional shells. Note that for $\mathbf{C}=\mathbf{E}=\mathbf{0}$, where $\mathbf{0}$ is the null set, we obtain the complete SO product space model as a limiting case.

At the other end of the spectrum is the even-replacement MCSCF model. In this model, the CSF's of a given species and subspecies, say $\Phi_{\Lambda P A}$ for all P, are built from an even-replacement set of SO's; such a set in turn is characterized by the fact that any two SO's are related by an even replacement of SO's. Note that we do not demand that the even-replacement rule holds between SO's used for the construction of CSF's of different species or subspecies.

The even-replacement MCSCF model has the following simplifying properties: 1) there are no invariances under SO transformations other than those arising from transformations within $\mathbf{C}$ or $\mathbf{E}$, or within $\mathbf{H}$ if applicable; 2) the SO's are always natural spin-orbitals; 3) compared to more general MCSCF models, the number of non-vanishing two-electron matrix elements is drastically reduced. These inherent simplifications render this model much more economical than others: in fact calculations by this model are comparable in size and complexity to ordinary SCF calculations.

The rationale for the even-replacement MCSCF model follows from two important observations: 1) the construction of CSF's from SD's always follows the even-replacement rule; 2) for a two-electron system, an exact wave function is always expressible in terms of an (infinite) even-replacement set of SD's. The first observation is of course a necessary prerequisite for the validity of the model. The second observation, combined with our hope and/or belief that electron correlation is essentially a sum of pair interactions, suggests that the even-replacement MCSCF model is capable of producing wave functions of high quality, at least for a calculation involving only a single spectroscopic term. On the surface, this expectation is contrary to experience with CI calculations, where single replacements have proved to be important. However the findings of CI calculations are not necessarily transferable to MCSCF calculations, in which the SO's have been carefully optimized. We suggest that calculations in the near future be designed in part to throw light on this question; the relative economy of the even-replacement model is too tempting to leave that avenue unexplored.

Broadening our horizon, there is a very large variety of intermediate possibilities between the even-replacement and the fractionally saturated MCSCF models. A satisfactory theoretical understanding and appropriate classification of the possible models does not exist at this time. From a practical point of view, such an understanding would undoubtedly be very helpful in developing effective strategies for the selection of CSF's in a wide variety of physical problems. Conversely, we can expect that the experience which is now accumulating from practical calculations will help to deepen our understanding of MCSCF models.

In analogy to invariance under SO transformations, the various MCSCF models can be further distinguished according to invariance under TSF transformations. For this purpose we divide the TSF's into two sets: the <u>occupied terms</u> which are present in the energy expression, and the <u>unoccupied terms</u> which are absent. In practice it is simpler to say that the occupied and unoccupied terms are both present in the energy expression, with non-vanishing and vanishing weights, respectively. It will be shown later on, in connection with the determination of essential non-redundant variables, that the MCSCF process is invariant under all (permissible) transformations within TSF subsets of the same weight, and barring a numerical accident, under no other TSF transformations.

VI. TRANSFORMATIONS OF ONE-ELECTRON AND N-ELECTRON WAVE FUNCTIONS

As discussed before, our MCSCF strategy calls for improved wave functions in terms of <u>reference SO's</u> $\phi_{\lambda i\alpha}(a)$ and <u>reference TSF's</u> Ψ_{AIA}. We need to define transformation operators which are specifically tailored to the index structures and arguments of the SO's and TSF's. We must furthermore define additional operators by summing over subspecies and/or electrons, and discover useful properties of, and relations between those various operators.

Analogous to the general replacement operator defined by Eq.(8), we now define the most elementary one-electron operator, the <u>SO replacement operator</u>, by

$$\mathcal{P}_{\lambda i\alpha,\mu j\beta}(a) = |\phi_{\mu j\beta}(a)\rangle\langle\phi_{\lambda i\alpha}(a)| \quad ; \tag{72}$$

its Hermitian conjugate is given by

$$\mathcal{P}^{*}_{\lambda i\alpha,\mu j\beta}(a) = \mathcal{P}_{\mu j\beta,\lambda i\alpha}(a) \quad . \tag{73}$$

When forming the product of two such operators we must distinguish beween the cases of the <u>same</u> or <u>different arguments</u>. For the same argument we get, analogous to Eq.(1[illegible])

$$\mathcal{P}_{\lambda i\alpha,\mu j\beta}(a)\,\mathcal{P}_{\nu k\gamma,\rho l\delta}(a) = \delta_{\lambda i\alpha,\rho l\delta}\,\mathcal{P}_{\nu k\gamma,\mu j\beta}(a) \quad . \tag{74}$$

For different arguments no similar formula applies. However in general, for either the same or different arguments, we have the very useful commutation relation

$$\llbracket \mathcal{P}_{\lambda i\alpha,\mu j\beta}(a),\mathcal{P}_{\nu k\gamma,\rho\ell\delta}(b)\rrbracket$$
$$= \delta_{ab}[\delta_{\lambda i\alpha,\rho\ell\delta}\mathcal{P}_{\nu k\gamma,\mu j\beta}(a) - \delta_{\nu k\gamma,\mu j\beta}\mathcal{P}_{\lambda i\alpha,\rho\ell\delta}(a)] \quad . \tag{75}$$

A __one-electron shell replacement operator__ is defined by

$$\mathcal{P}_{\lambda ij}(a) = \Sigma_\alpha \mathcal{P}_{\lambda i\alpha,\lambda j\alpha}(a) \quad ; \tag{76}$$

its properties follow easily from Eqs.(72-75), namely

$$\mathcal{P}^\dagger_{\lambda ij}(a) = \mathcal{P}_{\lambda ji}(a) \quad , \tag{77}$$

$$\mathcal{P}_{\lambda ij}(a)\mathcal{P}_{\mu k\ell}(a) = \delta_{\lambda\mu}\delta_{i\ell}\mathcal{P}_{\lambda kj}(a) \quad , \tag{78}$$

$$\llbracket \mathcal{P}_{\lambda ij}(a),\mathcal{P}_{\mu k\ell}(b)\rrbracket = \delta_{ab}\delta_{\lambda\mu}[\delta_{i\ell}\mathcal{P}_{\lambda kj}(a) - \delta_{kj}\mathcal{P}_{\lambda i\ell}(a)] \quad . \tag{79}$$

Summing over electron arguments, we define now also the __N-electron shell replacement operator__

$$\mathcal{P}_{\lambda ij} = \Sigma_a \mathcal{P}_{\lambda ij}(a) = \Sigma_a\Sigma_\alpha \mathcal{P}_{\lambda i\alpha,\lambda j\alpha}(a) \quad ; \tag{80}$$

its properties follow from Eqs.(77,79,80), namely

$$\mathcal{P}^\dagger_{\lambda ij} = \mathcal{P}_{\lambda ji} \quad , \tag{81}$$

$$\llbracket \mathcal{P}_{\lambda ij},\mathcal{P}_{\mu k\ell}\rrbracket = \delta_{\lambda\mu}(\delta_{i\ell}\mathcal{P}_{\lambda kj} - \delta_{kj}\mathcal{P}_{\lambda i\ell}) \quad . \tag{82}$$

When applying the variation principle, the shell replacement operators permit us to perform the SO variations in the most lucid manner. In fact, the one-electron operators (76) are the basis operators needed to describe the transformations of the individual SO's, while the N-electron operators (80) play an analogous role for the transformations of the TSF's which are __induced__ by the SO transformations.

We define the one-electron anti-Hermitian operator

$$\chi(a) = \Sigma_{\lambda ij}\mathcal{P}_{\lambda ij}(a)X_{\lambda ji} \quad . \tag{83}$$

A new set of SO's is now obtained from the reference SO's by the unitary transformation

$$\phi'_{\lambda i\alpha}(a) = U(a)\phi_{\lambda i\alpha}(a) = e^{\chi(a)}\phi_{\lambda i\alpha}(a) \quad , \tag{84}$$

which can also be stated as a matrix transformation of the reference SO's, namely

$$\phi'_{\lambda i\alpha}(a) = \Sigma_j\phi_{\lambda j\alpha}(a)U_{\lambda ji} = \Sigma_j\phi_{\lambda j\alpha}(a)(e^X)_{\lambda ji} \quad . \tag{84'}$$

We note that the matrices U and X refer to shells, and that they are blocked by symmetry; this is precisely what is needed for a permissible transformation which preserves the symmetry properties of the SO's.

From the one-electron anti-Hermitian operators (83) we now construct a corresponding N-electron operator by means of

$$\mathcal{X} = \Sigma_a\chi(a) = \Sigma_{\lambda ij}\mathcal{P}_{\lambda ij}X_{\lambda ji} \quad . \tag{85}$$

Because of the commutation relations

$$\begin{aligned} \llbracket \chi(a),\chi(b)\rrbracket &= \mathcal{O} \quad , \\ \llbracket U(a),U(b)\rrbracket &= \mathcal{O} \quad , \end{aligned} \tag{86}$$

we can now define an N-electron unitary operator $\mathcal{U}$ by means of

$$\mathcal{U} = \Pi_a U(a) = \Pi_a e^{\chi(a)} = e^{\Sigma_a\chi(a)} = e^{\mathcal{X}} \quad . \tag{87}$$

Let now Ψ be any function in SO product space defined in terms of the reference SO's $\phi_{\lambda i\alpha}(a)$. If we subject all SO's simultaneously to the transformation (84), the resulting transformation in SO product space is given by

$$\Psi' = \mathcal{U}\Psi = e^{\mathcal{X}}\Psi \quad ; \tag{88}$$

we call (88) the induced TSF transformation.

It is remarkable that the anti-Hermitian operators which effect the one-electron and induced N-electron transformations, $\mathcal{X}(a)$ and $\mathcal{X}$, are so simply related: they have the same matrix elements $X_{\lambda ij}$ with respect to their natural basis operators, $\varphi_{\lambda ij}(a)$ and $\varphi_{\lambda ij}$, respectively. There is no corresponding simple relation between $\mathcal{U}(a)$ and $\mathcal{U}$. This underscores the superiority of the anti-Hermitian operators over the unitary operators in this context.

We can apply the transformation (88) to our set of CSF's $\Phi_{\Lambda P \Lambda}$, or equivalently to the TSF's $\Psi_{\Lambda I \Lambda}$, to yield

$$\left.\begin{aligned} \Phi'_{\Lambda P\Lambda} &= \mathcal{U}\Phi_{\Lambda P\Lambda} = e^{\mathcal{X}}\Phi_{\Lambda P\Lambda} \quad , \\ \Psi'_{\Lambda I\Lambda} &= \mathcal{U}\Psi_{\Lambda I\Lambda} = e^{\mathcal{X}}\Psi_{\Lambda I\Lambda} \quad . \end{aligned}\right\} \tag{89}$$

In general the set of chosen CSF's $\Phi_{\Lambda P\Lambda}$ does not span the full SO product space. It is too much to expect that under the transformation (89) the CSF's (or TSF's) would transform among themselves: the CSF (or TSF) manifold is in general not invariant under an SO-induced transformation. However it is possible that such invariance occurs for some of the operators $\mathcal{U}$; in this case the CSF and TSF transformations, and their interconnection, are given by

$$\left.\begin{aligned} \Phi'_{\Lambda P\Lambda} &= \mathcal{U}\Phi_{\Lambda P\Lambda} = \Sigma_Q \Phi_{\Lambda Q\Lambda} U_{\Lambda QP} \quad , \\ \Psi'_{\Lambda I\Lambda} &= \mathcal{U}\Psi_{\Lambda I\Lambda} = \Sigma_J \Psi_{\Lambda J\Lambda} U_{\Lambda JI} \quad , \\ U_{\Lambda IJ} &= \Sigma_{PQ} C^{*}_{\Lambda IP} U_{\Lambda PQ} C_{\Lambda QJ} \quad . \end{aligned}\right\} \tag{89'}$$

The operators $\mathcal{U}$ for which the CSF manifold is invariant will later prove to be of particular significance in connection with the elimination of non-essential variables.

In addition to the SO-induced transformation, the TSF's are also transformed by varying the expansions in terms of the CSF's; we call this the direct TSF transformation. We define, analogous to the shell replacement operators, the N-electron term replacement operator

$$\varphi_{\Lambda IJ} = \Sigma_{\Lambda} |\Psi_{\Lambda J\Lambda}\rangle\langle\Psi_{\Lambda I\Lambda}| \quad . \tag{90}$$

which has the matrix elements

$$Q_{\Lambda IJ,\Xi KL} = \langle\Psi_{\Xi K\Lambda}|\varphi_{\Lambda IJ}|\Psi_{\Xi L\Lambda}\rangle = \delta_{\Lambda\Xi}\delta_{IL}\delta_{KJ} \quad . \tag{91}$$

furthermore it is easily seen that

$$\varphi^{*}_{\Lambda IJ} = \varphi_{\Lambda JI} \quad . \tag{92}$$

$$\varphi_{\Lambda IJ}\varphi_{\Xi KL} = \delta_{\Lambda\Xi}\delta_{IL}\varphi_{\Lambda KJ} \quad . \tag{93}$$

$$[\varphi_{\Lambda IJ}, \varphi_{\Xi KL}] = \delta_{\Lambda\Xi}(\delta_{IL}\varphi_{\Lambda KJ} - \delta_{KJ}\varphi_{\Lambda IL}) \quad . \tag{94}$$

Defining now the N-electron anti-Hermitian operator

$$\mathcal{Y} = \Sigma_{\Lambda IJ}\varphi_{\Lambda IJ}Y_{\Lambda JI} \quad . \tag{95}$$

we obtain a new set of TSF's from the reference TSF's by the permissible unitary transformation

$$\Psi'_{\Lambda I\Lambda} = \mathcal{V}\Psi_{\Lambda I\Lambda} = e^{\mathcal{Y}}\Psi_{\Lambda I\Lambda} \quad . \tag{96}$$

which can also be stated as a matrix transformation of the reference TSF's, namely

$$\Psi'_{\Lambda I\Lambda} = \Sigma_J \Psi_{\Lambda J\Lambda} V_{\Lambda JI} = \Sigma_J \Psi_{\Lambda J\Lambda}(e^{Y})_{\Lambda JI} \quad . \tag{96'}$$

Finally, we must write down the combined TSF transformation which results when both the induced and the direct TSF transformation are performed. The required formula is

$$\Psi'_{AIA} = \mathcal{U}\mathcal{V}\Psi_{AIA} = e^{\mathcal{X}}e^{\mathcal{Y}}\Psi_{AIA} \quad . \tag{97}$$

We observe that in general the operators $\mathcal{X}$ and $\mathcal{Y}$ (or, equivalently, $\mathcal{U}$ and $\mathcal{V}$) do <u>not</u> commute. Hence we cannot combine the two exponents in Eq.(97), nor can we change the order of the two exponential operators. One clue to remember this order correctly is the notion that $\mathcal{Y}$ is defined with respect to the reference TSF's, and not with respect to TSF's changed by an SO-induced transformation.

VII. MATRIX ELEMENTS OF SYMMETRICAL ONE- AND TWO-ELECTRON OPERATORS

Let $\mathcal{F}(a)$ be a one-electron operator which is symmetrical with respect to the symmetry group of the system. Analogous to Eq.(61), the matrix elements of $\mathcal{F}(a)$ with respect to SO's and basis functions simplify to

$$\left.\begin{aligned} \langle\phi_{\lambda i\alpha}(a)|\mathcal{F}(a)|\phi_{\mu j\beta}(a)\rangle &= \delta_{\lambda\alpha,\mu\beta}F_{\lambda ij} \quad , \\ \langle\chi_{\lambda p\alpha}(a)|\mathcal{F}(a)|\chi_{\mu q\beta}(a)\rangle &= \delta_{\lambda\alpha,\mu\beta}f_{\lambda pq} \quad . \end{aligned}\right\} \tag{98}$$

Here the Kronecker deltas depend only on symmetry; they are a special case of what are called <u>symmetry factors</u>. The $F_{\lambda ij}$ and $f_{\lambda pq}$ do not depend on the subspecies, and are called <u>reduced matrix elements</u>.

Similarly, let $\mathcal{G}(a,b)=\mathcal{G}(b,a)$, $a\neq b$, be a two-electron operator which is symmetrical with respect to the symmetry group of the system. The most general treatment for the matrix elements of $\mathcal{G}(a,b)$ is considerably more complicated than for the one-electron operators $\mathcal{F}(a)$; the degree of complexity is a function of the structure of the symmetry group. We now make the simplifying assumption that the group is <u>simply reducible</u> (Wigner, 1940, 1941). Such a group is characterized by (1) inverse elements are in the same class; and (2) when decomposing the product of two irreducible representations into a sum of such representations, no representation in that sum occurs more than once. These two properties are sufficient to guarantee the existence of Clebsch-Gordan coefficients, which is the main reason for the resulting simplifications. As mentioned before, the assumption that the group is simply reducible covers atoms and the vast majority of molecules.

For simply reducible groups the matrix elements of $\mathcal{G}(a,b)$ with respect to SO's and basis functions simplify to

$$\left.\begin{aligned} &\langle\phi_{\lambda i\alpha}(a)\phi_{\nu k\gamma}(b)|\mathcal{G}(a,b)|\phi_{\rho l\delta}(b)\phi_{\mu j\beta}(a)\rangle \\ &\qquad = \Sigma_\kappa w_{\lambda\alpha\mu\beta,\nu\gamma\rho\delta,\kappa}\, G_{\lambda i\mu j,\nu k\rho l,\kappa} \quad , \\ &\langle\chi_{\lambda p\alpha}(a)\chi_{\nu r\gamma}(b)|\mathcal{G}(a,b)|\chi_{\rho s\delta}(b)\chi_{\mu q\beta}(a)\rangle \\ &\qquad = \Sigma_\kappa w_{\lambda\alpha\mu\beta,\nu\gamma\rho\delta,\kappa}\, g_{\lambda p\mu q,\nu r\rho s,\kappa} \quad . \end{aligned}\right\} \tag{99}$$

where now the $w_{\lambda\alpha\mu\beta,\nu\gamma\rho\delta,\kappa}$ are the symmetry factors, while $G_{\lambda i\mu j,\nu k\rho l,\kappa}$ and $g_{\lambda p\mu q,\nu r\rho s,\kappa}$ are the reduced matrix elements. The index κ, called the coupling index, identifies the resulting symmetry species arising from the decomposition of the product of species λ and μ (or ν and ρ). If the symmetry group is not simply reducible, a generalization of Eqs.(99) with a slightly more elaborate index structure applies. Since Eqs.(99) are most frequently applied to electron-electron interaction, we shall call the two-electron symmetry factors <u>electron interaction coupling coefficients</u>; they satisfy the symmetry relations

$$\bar{w}_{\mu\beta\lambda\alpha,\rho\delta\nu\gamma,\kappa} = w_{\lambda\alpha\mu\beta,\nu\gamma\rho\delta,\kappa} = w_{\nu\gamma\rho\delta,\lambda\alpha\mu\beta,\kappa} \quad . \tag{100}$$

Similarly the two-electron reduced matrix elements satisfy the symmetry relations

$$\left.\begin{aligned} G_{\lambda i\mu j,\nu k\rho l,\kappa} &= G_{\nu k\rho l,\lambda i\mu j,\kappa} \quad . \\ g_{\lambda p\mu q,\nu r\rho s,\kappa} &= g_{\nu r\rho s,\lambda p\mu q,\kappa} \quad . \end{aligned}\right\} \qquad (101)$$

Using the definition of a Hermitian conjugate operator, Eq.(12), we find for the Hermitian conjugates of the reduced matrix elements

$$\left.\begin{aligned} F^*_{\lambda ij} &= \bar{F}_{\lambda ji} \quad , \\ f^*_{\lambda pq} &= \bar{f}_{\lambda qp} \quad , \\ G^*_{\lambda i\mu j,\nu k\rho l,\kappa} &= \bar{G}_{\mu j\lambda i,\rho l\nu k,\kappa} \quad , \\ g^*_{\lambda p\mu q,\nu r\rho s,\kappa} &= \bar{g}_{\mu q\lambda p,\rho s\nu r,\kappa} \quad ; \end{aligned}\right\} \qquad (102)$$

furthermore, if the operators $\mathcal{F}(a)$ and $\mathcal{G}(a,b)$ are Hermitian, that is

$$\left.\begin{aligned} \mathcal{F}^*(a) &= \mathcal{F}(a) \quad , \\ \mathcal{G}^*(a,b) &= \mathcal{G}(a,b) \quad , \end{aligned}\right\} \qquad (103)$$

we have for the reduced matrix elements

$$\left.\begin{aligned} F^*_{\lambda ij} &= F_{\lambda ij} \quad , \\ f^*_{\lambda pq} &= f_{\lambda pq} \quad , \\ G^*_{\lambda i\mu j,\nu k\rho l,\kappa} &= G_{\lambda i\mu j,\nu k\rho l,\kappa} \quad , \\ g^*_{\lambda p\mu q,\nu r\rho s,\kappa} &= g_{\lambda p\mu q,\nu r\rho s,\kappa} \quad . \end{aligned}\right\} \qquad (104)$$

We call attention to the facts that the first two Eqs.(102) are an obvious restatement of Eq.(15) for the case of matrices blocked by symmetry, and that in order to establish the last two Eqs.(102) one needs to invoke the complex conjugate symmetry of the electron interaction coupling coefficients expressed in Eq.(100).

If $\mathcal{F}(a)$ and $\mathcal{G}(a,b)$ are the one- and two-electron operators of the Hamiltonian of the system, the reduced matrix elements $F_{\lambda ij}$, $f_{\lambda pq}$, $G_{\lambda i\mu j,\nu k\rho l,\kappa}$, $g_{\lambda p\mu q,\nu r\rho s,\kappa}$ are often referred to as the <u>one-</u> and <u>two-electron integrals over SO's</u> and <u>basis functions</u>, respectively; and since the Hamiltonian is Hermitian, they obviously satisfy Eqs.(104). These integrals over SO's and basis functions are furthermore related by

$$\left.\begin{aligned} F_{\lambda ij} &= \Sigma_{pq} c^*_{\lambda ip} f_{\lambda pq} c_{\lambda qj} \quad , \\ G_{\lambda i\mu j,\nu k\rho l,\kappa} &= \Sigma_{pqrs} c^*_{\lambda ip} c^*_{\nu kr} g_{\lambda p\mu q,\nu r\rho s,\kappa} c_{\rho sl} c_{\mu qj} \quad . \end{aligned}\right\} \qquad (105)$$

The calculation of $F_{\lambda ij}$ and $G_{\lambda i\mu j,\nu k\rho l,\kappa}$ by means of Eqs.(105) is often referred to as the <u>integrals transformation</u>.

Within our MCSCF model, we need to evaluate one- and two-electron matrix elements only between wave functions $\phi(a)$ which are confined, for each argument a, to the space spanned by the SO's $\phi_{\lambda i\alpha}(a)$, or equivalently by the basis functions $\chi_{\lambda i\alpha}(a)$. Hence we can use the machinery of projected operators developed earlier, Eqs.(36-40), for the evaluation of such matrix elements. The relevant projectors in this case are the <u>SO projectors</u> for each argument, namely

$$\mathcal{P}(a) = \Sigma_{\lambda i} \mathcal{P}_{\lambda ii}(a) = \Sigma_{\lambda i\alpha} \mathcal{P}_{\lambda i\alpha,\lambda i\alpha}(a) \quad , \qquad (106)$$

so that

$$\mathcal{P}(a)\phi(a) = \phi(a) \quad . \qquad (107)$$

The projected one- and two-electron operators are now defined by

$$\left.\begin{aligned} \hat{\mathcal{F}}(a) &= \mathcal{P}(a)\mathcal{F}(a)\mathcal{P}(a) \quad , \\ \hat{\mathcal{G}}(a,b) &= \mathcal{P}(a)\mathcal{P}(b)\mathcal{G}(a,b)\mathcal{P}(b)\mathcal{P}(a) \quad . \end{aligned}\right\} \qquad (108)$$

We now make the important observation that any function Ψ in the SO product space is, as far as the argument a is concerned, of the type $\phi(a)$; therefore

$$\mathcal{P}(a)\Psi = \Psi \quad . \tag{107'}$$

Another important observation is that $\mathcal{P}(a)$ commutes with any operator which is internal to the SO product space. In particular, $\mathcal{P}(a)$ commutes with any replacement operator defined so far, or any algebraic construct thereof; it suffices to state

$$\left.\begin{aligned} [\mathcal{P}(a), \mathcal{P}_{\lambda i\alpha,\mu j\beta}(b)] &= 0 \quad , \\ [\mathcal{P}(a), \Phi_{A\Sigma\gamma}] &= 0 \quad . \end{aligned}\right\} \tag{109}$$

For the evaluation of physical properties from our N-electron wave functions we need the <u>total</u> <u>one-</u> <u>and</u> <u>two-electron</u> <u>operators</u>

$$\left.\begin{aligned} \mathcal{F} &= \Sigma_a \mathcal{F}(a) \quad , \\ \mathcal{G} &= \tfrac{1}{2}\Sigma_{ab,a\neq b}\mathcal{G}(a,b) = \Sigma_{ab,a>b}\mathcal{G}(a,b) \quad . \end{aligned}\right\} \tag{110}$$

We define the corresponding <u>total</u> <u>projected</u> <u>one-</u> <u>and</u> <u>two-electron</u> <u>operators</u>

$$\left.\begin{aligned} \hat{\mathcal{F}} &= \Sigma_a \hat{\mathcal{F}}(a) \quad , \\ \hat{\mathcal{G}} &= \tfrac{1}{2}\Sigma_{ab,a\neq b}\hat{\mathcal{G}}(a,b) = \Sigma_{ab,a>b}\hat{\mathcal{G}}(a,b) \quad . \end{aligned}\right\} \tag{111}$$

Note that all components of $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ are internal to the SO product space, albeit by force of different projectors. Consequently, following Eq.(48), if Ψ and Φ are any wave functions within, and $\mathcal{A}$ and $\mathcal{B}$ are any operators internal to the SO product space, then

$$\left.\begin{aligned} \langle\Phi|\mathcal{A}\mathcal{F}\mathcal{B}|\Psi\rangle &= \langle\Phi|\mathcal{A}\hat{\mathcal{F}}\mathcal{B}|\Psi\rangle \quad , \\ \langle\Phi|\mathcal{A}\mathcal{G}\mathcal{B}|\Psi\rangle &= \langle\Phi|\mathcal{A}\hat{\mathcal{G}}\mathcal{B}|\Psi\rangle \quad . \end{aligned}\right\} \tag{112}$$

Eqs.(112) guarantee that $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ are projected operators associated with $\mathcal{F}$ and $\mathcal{G}$, respectively. The actual projectors which accomplish this correspondence are of no particular interest for our present purposes, and we chose to omit them.

The projected operators can now be expressed in terms of reduced matrix elements and replacement operators. After some algebra we find

$$\left.\begin{aligned} \hat{\mathcal{F}} &= \Sigma_{\lambda ij}\mathcal{P}_{\lambda ij}F_{\lambda ji} \quad , \\ \hat{\mathcal{G}} &= \Sigma_\kappa\Sigma_{\lambda i\mu j}\Sigma_{\nu k\rho l}\mathcal{P}_{\lambda i\mu j,\nu k\rho l,\kappa}G_{\mu j\lambda i,\rho l\nu k,\kappa} \quad . \end{aligned}\right\} \tag{113}$$

where we have introduced the <u>two-shell</u> <u>replacement</u> <u>operator</u>

$$\mathcal{P}_{\lambda i\mu j,\nu k\rho l,\kappa} = \Sigma_{ab,a\neq b}\Sigma_{\alpha\beta\gamma\delta}W_{\mu\beta\lambda\alpha,\rho\delta\nu\gamma,\kappa}\mathcal{P}_{\lambda i\alpha,\mu j\beta}(a)\mathcal{P}_{\nu k\gamma,\rho l\delta}(b) \quad . \tag{114}$$

This operator satisfies the symmetry conditions

$$\mathcal{P}^{\dagger}_{\mu j\lambda i,\rho l\nu k,\kappa} = \mathcal{P}_{\lambda i\mu j,\nu k\rho l,\kappa} = \mathcal{P}_{\nu k\rho l,\lambda i\mu j,\kappa} \quad . \tag{115}$$

and the commutation relation

$$\begin{aligned} &[\mathcal{P}_{\lambda i\mu j,\nu k\rho l,\kappa}, \mathcal{P}_{\sigma mn}] \\ &\quad = \delta_{\lambda i,\sigma n}\mathcal{P}_{\lambda m\mu j,\nu k\rho l,\kappa} - \delta_{\mu j,\sigma m}\mathcal{P}_{\lambda i\mu n,\nu k\rho l,\kappa} \\ &\quad + \delta_{\nu k,\sigma n}\mathcal{P}_{\lambda i\mu j,\nu m\rho l,\kappa} - \delta_{\rho l,\sigma m}\mathcal{P}_{\lambda i\mu j,\nu k\rho n,\kappa} \quad . \end{aligned} \tag{116}$$

Using Eqs.(111,113) we can now express the matrix elements of $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ between any two functions in SO product space in terms of the matrix elements of shell replacement operators and reduced matrix elements. In particular, for the matrix elements of $\hat{\mathcal{F}}$ and $\hat{\mathcal{G}}$ between CSF's we find

$$\left.\begin{aligned} \langle\Phi_{APA}|\hat{\mathcal{F}}|\Phi_{AQA}\rangle &= \Sigma_{\lambda ij}\pi_{APQ,\lambda ij}F_{\lambda ji} \quad , \\ \langle\Phi_{APA}|\hat{\mathcal{G}}|\Phi_{AQA}\rangle &= \Sigma_\kappa\Sigma_{\lambda i\mu j}\Sigma_{\nu k\rho l}\pi_{APQ,\lambda i\mu j,\nu k\rho l,\kappa}G_{\mu j\lambda i,\rho l\nu k,\kappa} \quad . \end{aligned}\right\} \tag{117}$$

where we have introduced the <u>structure constants</u>

$$\left.\begin{aligned} n_{APQ,\lambda ij} &= \langle \Phi_{APA} | E_{\lambda ij} | \Phi_{AQA} \rangle \quad . \\ n_{APQ,\lambda i\mu j,\nu k\rho l,\kappa} &= \langle \Phi_{APA} | E_{\lambda i\mu j,\nu k\rho l,\kappa} | \Phi_{AQA} \rangle \quad . \end{aligned}\right\} \qquad (118)$$

These structure constants express how the CSF's are constructed from the SO's, and are hence constant in a given calculation. They satisfy the symmetry conditions

$$\left.\begin{aligned} \bar{n}_{APQ,\lambda ij} &= n_{AQP,\lambda ji} \quad . \\ \bar{n}_{AQP,\mu j\lambda i,\rho l\nu k,\kappa} &= n_{APQ,\lambda i\mu j,\nu k\rho l,\kappa} = n_{APQ,\nu k\rho l,\lambda i\mu j,\kappa} \quad . \end{aligned}\right\} \qquad (119)$$

Putting $P=Q$ in Eqs.(118), the structure constants reduce to the usual expressions known from single configuration SCF theory, namely

$$\left.\begin{aligned} n_{APP,\lambda ij} &= \delta_{ij} n_{AP,\lambda i} \quad . \\ n_{APP,\lambda i\nu k,\mu j\rho l,\kappa} &= \delta_{\lambda i,\nu k}\delta_{\mu j,\rho l} c_{AP,\lambda i,\mu j,\kappa} \\ &\quad - \delta_{\lambda i,\rho l}\delta_{\mu j,\nu k} x_{AP,\lambda i,\mu j,\kappa} \quad . \end{aligned}\right\} \qquad (120)$$

where the $n_{AP,\lambda i}$ are the shell occupation numbers introduced before, Eqs. (57'), while $c_{AP,\lambda i,\mu j,\kappa}$ and $x_{AP,\lambda i,\mu j,\kappa}$ are the familiar <u>Coulomb interaction coefficients</u> and <u>exchange interaction coefficients</u>, respectively. They satisfy the symmetry conditions

$$\left.\begin{aligned} \bar{c}_{AP,\lambda i,\mu j,\kappa} &= c_{AP,\lambda i,\mu j,\kappa} = c_{AP,\mu j,\lambda i,\kappa} \quad . \\ \bar{x}_{AP,\lambda i,\mu j,\kappa} &= x_{AP,\lambda i,\mu j,\kappa} = x_{AP,\mu j,\lambda i,\kappa} \quad . \end{aligned}\right\} \qquad (121)$$

Most of the Coulomb and exchange interaction coefficients can be further reduced to simpler quantities by

$$\left.\begin{aligned} c_{AP,\lambda i,\mu j,\kappa} &= \delta_{\kappa 0} n_{AP,\lambda i} n_{AP,\mu j} \quad . \\ x_{AP,\lambda i,\mu j,\kappa} &= x_{\lambda\mu\kappa} n_{AP,\lambda i} n_{AP,\mu j} \quad . \end{aligned}\right\} \left\{\begin{aligned} &\lambda i \subset C_{AP} \cup E_{AP} \quad \text{or} \\ &\mu j \subset C_{AP} \cup E_{AP} \quad ; \end{aligned}\right\} \qquad (122)$$

so that the only Coulomb and exchange interaction coefficients which need to be established separately are those with $\lambda i \subset F_{AP}$ and $\mu j \subset F_{AP}$. In Eqs.(122), $\kappa=0$ is the coupling index identifying the <u>identical representation</u> of the symmetry group, and $x_{\lambda\mu\kappa}$ is a set of constants depending on the symmetry group only.

The evaluation of the structure constants for $P \neq Q$, as well as the fractional-fractional interaction constants for $P=Q$, is a complicated subject. Powerful techniques have been developed by Racah to solve this problem for atoms (Racah, 1942, 1943); the specifics of these techniques are intimately connected with the structure and properties of the applicable symmetry group. To solve this problem in general for our present MCSCF formalism, we need the analogue of Racah's algebra for arbitrary simply reducible groups. Such a general solution is not yet available, so that practical MCSCF calculations necessitate at present the piecemeal calculation of the required structure constants by specific rather than general methods.

The burden of evaluating the structure constants is considerably less for the even-replacement model than for other MCSCF models. It is easily seen that in the even-replacement model the one-electron constants between any two CSF's must vanish:

$$n_{APQ,\lambda ij} = 0 \quad , \quad P \neq Q \quad . \qquad (123)$$

For the general two-electron constants with $P \neq Q$ the simplification is relatively speaking more considerable. Namely for given APQ, that is, for a particular pair of CSF's, there are at most two non-vanishing constants for each value of the coupling index κ. As a result, calculations with the even-replacement model are significantly more economical than with other MCSCF models.

It is noteworthy that in formulations based on second quantization it is unnatural and awkward to formulate the equivalent of our two-shell replacement operator (114). This is so because in the definition (114) the two electrons must be separately identified; in second quantization it is not possible to keep track of individual electrons. Consequently a complete symmetry reduction has not been achieved to date in any SCF or MCSCF theory based on second quantization.

VIII. THE VARATIONAL PROCESS IN TERMS OF ESSENTIAL NON-REDUNDANT VARIABLES

The Hamiltonian of our system, and its projected equivalent, are given by

$$\left.\begin{array}{ll} \mathcal{H} = \mathcal{F} + \mathcal{G} & , \\ \hat{\mathcal{H}} = \hat{\mathcal{F}} + \hat{\mathcal{G}} & . \end{array}\right\} \tag{124}$$

where $\mathcal{F}$, $\mathcal{G}$, $\hat{\mathcal{F}}$, $\hat{\mathcal{G}}$ are the Hermitian total one- and two-electron operators as per Eqs.(110,111). Our multi-state variational energy is given by

$$E = \Sigma_{AI} D_A W_{AI} \langle \Psi_{AIA} | \mathcal{H} | \Psi_{AIA} \rangle = \Sigma_{AI} D_A W_{AI} \langle \Psi_{AIA} | \hat{\mathcal{H}} | \Psi_{AIA} \rangle \quad ; \tag{125}$$

where D_A is the degree of degeneracy of species A, and the weights are constrained by

$$\left.\begin{array}{ll} W^*_{AI} = W_{AI} \geq 0 & , \\ \Sigma_{AI} D_A W_{AI} = 1 & . \end{array}\right\} \tag{126}$$

Note that in general the energy sum contains occupied terms with $W_{AI} > 0$ and unoccupied terms with $W_{AI} = 0$.

It is useful to introduce the (Hermitian) weight operator $\mathcal{W}$ defined by

$$\mathcal{W} = \Sigma_{AI} \mathcal{E}_{AII} W_{AI} \quad . \tag{127}$$

Obviously $\mathcal{W}$ is internal to the SO product space; furthermore $\mathcal{W}$ is diagonal with respect to the reference TSF's Ψ_{AIA}. In the spirit of Löwdin's trace algebra (Löwdin, 1977) we can now express the MCSCF variational energy in the most compact form, namely

$$E = \mathrm{Trace}(\mathcal{W}\hat{\mathcal{H}}) \quad . \tag{128}$$

It is to be noted that during the variational process $\mathcal{H}$ and $\hat{\mathcal{H}}$ are fixed operators, while $\mathcal{W}$ moves along with the variation via the $\mathcal{Y}$-operators.

Our next objective is to define essential non-redundant variables for our variational process. A set of variables is non-redundant if there are no auxiliary conditions for, or constraints on, these variables. A single variable is non-essential if it occurs in our energy expression with vanishing coefficients. Our set of variables is essential if it, or any linear transformation thereof, contains no non-essential variables.

We subject the TSF's in the energy expression (125) to the variation (97), and transfer the transformation operators from the wave functions to the Hamiltonian, thus obtaining for the new energy

$$E' = \Sigma_{AI} D_A W_{AI} \langle \Psi_{AIA} | e^{-\mathcal{Y}} e^{-\mathcal{X}} \hat{\mathcal{H}} e^{\mathcal{X}} e^{\mathcal{Y}} | \Psi_{AIA} \rangle \quad . \tag{129}$$

This can be rewritten as a trace, in which we carry out one cyclical move, obtaining

$$E' = \mathrm{Trace}(e^{\mathcal{Y}} \mathcal{W} e^{-\mathcal{Y}} e^{-\mathcal{X}} \hat{\mathcal{H}} e^{\mathcal{X}}) \quad . \tag{130}$$

We now use for the exponential transformations of $\mathcal{W}$ and $\hat{\mathcal{H}}$ the well-known expansion in terms of commutators. The result is

$$\begin{aligned} E' = \mathrm{Trace}[(&\mathcal{W} - [\mathcal{W},\mathcal{Y}] + \tfrac{1}{2}[[\mathcal{W},\mathcal{Y}],\mathcal{Y}] - \ldots) \\ &(\hat{\mathcal{H}} + [\hat{\mathcal{H}},\mathcal{X}] + \tfrac{1}{2}[[\hat{\mathcal{H}},\mathcal{X}],\mathcal{X}] + \ldots)] \quad . \end{aligned} \tag{131}$$

Inserting in Eq.(131) the expansions for $\mathcal{X}$ and $\mathcal{Y}$, Eqs.(85,95), we obtain for E' a Taylor expansion in terms of the matrix elements of the transformations, namely

$$\begin{aligned} E' = E &+ \Sigma_{\lambda ij} e_{\lambda ij} X_{\lambda ji} + \Sigma_{AIJ} e_{AIJ} Y_{AJI} \\ &+ \tfrac{1}{2}\Sigma_{\lambda ij}\Sigma_{\mu kl} e_{\lambda ij,\mu kl} X_{\lambda ji} X_{\mu lk} \\ &+ \Sigma_{\lambda ij}\Sigma_{AIJ} e_{\lambda ij,AIJ} X_{\lambda ji} Y_{AJI} \\ &+ \tfrac{1}{2}\Sigma_{AIJ}\Sigma_{\Xi KL} e_{AIJ,\Xi KL} Y_{AJI} Y_{\Xi LK} + \ldots \quad . \end{aligned} \tag{132}$$

where E is given by Eq.(128), and

$$e_{\lambda ij} = \mathrm{Trace}(\mathcal{W}[\hat{\mathcal{H}},\rho_{\lambda ij}]) \quad , \tag{133}$$

$$e_{AIJ} = -\mathrm{Trace}([\mathcal{W},\varphi_{AIJ}]\hat{\mathcal{H}}) \quad , \tag{134}$$

$$e_{\lambda ij,\mu kl} = \tfrac{1}{2}\mathrm{Trace}[\mathcal{W}([[\hat{\mathcal{H}},\rho_{\lambda ij}],\rho_{\mu kl}] + [[\hat{\mathcal{H}},\rho_{\mu kl}],\rho_{\lambda ij}])] \quad , \tag{135}$$

$$e_{\lambda ij,AIJ} = -\mathrm{Trace}([\mathcal{W},\varphi_{AIJ}][\hat{\mathcal{H}},\rho_{\lambda ij}]) \quad , \tag{136}$$

$$e_{AIJ,\Xi KL} = \tfrac{1}{2}\mathrm{Trace}(([[\mathcal{W},\varphi_{AIJ}],\varphi_{\Xi KL}] + [[\mathcal{W},\varphi_{\Xi KL}],\varphi_{AIJ}])\hat{\mathcal{H}}) \quad . \tag{137}$$

Note that the definitions (135,137) have been chosen so as to guarantee

$$\left.\begin{aligned} e_{\lambda ij,\mu kl} &= e_{\mu kl,\lambda ij} \quad , \\ e_{AIJ,\Xi KL} &= e_{\Xi KL,AIJ} \quad . \end{aligned}\right\} \tag{138}$$

Since the symmetry group of our system is assumed to be simply reducible, we can choose our wave function representations so that all relevant matrices are real; furthermore Hermitian, anti-Hermitian and unitary matrices become symmetric, skew-symmetric and orthogonal, respectively. For more complicated structures, like the two-electron reduced matrix elements, similar statements are applicable.

Invoking the real skew-symmetry of the transformation matrices, namely

$$\left.\begin{aligned} \bar{X}_{\lambda ij} &= X_{\lambda ij} = -X_{\lambda ji} \quad , \\ \bar{Y}_{AIJ} &= Y_{AIJ} = -Y_{AJI} \quad , \end{aligned}\right\} \tag{139}$$

we note that the energy expression (132) obviously contains redundant variables. We note further that the skew-symmetry is the <u>only</u> source of redundancy, since there are no other conditions which $X_{\lambda ij}$ and Y_{AIJ} have to satisfy to be valid variational variables. The redundancy is easily removed by collecting the terms in (132) which contain equal and opposite matrix elements, and retaining as explicit variables the lower left triangular matrix elements $X_{\lambda ij}$ and Y_{AIJ}. We designate the sets of compound indices of these variables by S (for shell rotations) and T (for term rotations), so that

$$\left.\begin{aligned} \lambda ij &\in S \quad , \quad i > j \quad , \\ AIJ &\in T \quad , \quad I > J \quad . \end{aligned}\right\} \tag{140}$$

It is noteworthy that the formal and explicit removal of redundancy from the variational parameter set is a direct consequence of having used the anti-Hermitian operators $\mathcal{X}$ and $\mathcal{Y}$ rather than their unitary equivalents $\mathcal{U}$ and $\mathcal{V}$. This underscores once more the superiority of the anti-Hermitian operators over the unitary operators.

It will be shown in the next section that the sets of non-redundant variables $X_{\lambda ij}$ and Y_{AIJ} can be partitioned into subsets of essential and non-essential variables. Denoting the sets of indices of the essential variables by S_e and T_e, and of the non-essential variables by S_n and T_n, we have

$$\left.\begin{aligned} S &= S_e + S_n \quad , \\ T &= T_e + T_n \quad . \end{aligned}\right\} \tag{141}$$

In our energy formulas the sums over variational variables collapse in fact into sums over essential variables only; we shall use for such sums the simplified notation

$$\left.\begin{aligned} \Sigma'_{\lambda ij} &= \Sigma_{\lambda ij\subset S_e} \quad . \\ \Sigma'_{\Lambda IJ} &= \Sigma_{\Lambda IJ\subset T_e} \quad . \end{aligned}\right\} \tag{142}$$

In terms of the essential non-redundant variables, the new energy can now be written in the form

$$\begin{aligned} E' = E &- \Sigma'_{\lambda ij}E_{\lambda ij}X_{\lambda ij} - \Sigma'_{\Lambda IJ}E_{\Lambda IJ}Y_{\Lambda IJ} \\ &+ \tfrac{1}{2}\Sigma'_{\lambda ij}\Sigma'_{\mu kl}E_{\lambda ij,\mu kl}X_{\lambda ij}X_{\mu kl} \\ &+ \Sigma'_{\lambda ij}\Sigma'_{\Lambda IJ}E_{\lambda ij,\Lambda IJ}X_{\lambda ij}Y_{\Lambda IJ} \\ &+ \tfrac{1}{2}\Sigma'_{\Lambda IJ}\Sigma'_{\Xi KL}E_{\Lambda IJ,\Xi KL}Y_{\Lambda IJ}Y_{\Xi KL} + \dots \quad , \end{aligned} \tag{143}$$

where

$$E_{\lambda ij} = e_{\lambda ij} - e_{\lambda ji} \quad , \tag{144}$$

$$E_{\Lambda IJ} = e_{\Lambda IJ} - e_{\Lambda JI} \quad . \tag{145}$$

$$E_{\lambda ij,\mu kl} = e_{\lambda ij,\mu kl} - e_{\lambda ij,\mu lk} - e_{\lambda ji,\mu kl} + e_{\lambda ji,\mu lk} \quad , \tag{146}$$

$$E_{\lambda ij,\Lambda IJ} = e_{\lambda ij,\Lambda IJ} - e_{\lambda ij,\Lambda JI} - e_{\lambda ji,\Lambda IJ} + e_{\lambda ji,\Lambda JI} \quad , \tag{147}$$

$$E_{\Lambda IJ,\Xi KL} = e_{\Lambda IJ,\Xi KL} - e_{\Lambda IJ,\Xi LK} - e_{\Lambda JI,\Xi KL} + e_{\Lambda JI,\Xi LK} \quad , \tag{148}$$

while of course also

$$\left.\begin{aligned} E_{\lambda ij,\mu kl} &= E_{\mu kl,\lambda ij} \quad , \\ E_{\Lambda IJ,\Xi KL} &= E_{\Xi KL,\Lambda IJ} \quad . \end{aligned}\right\} \tag{149}$$

Truncating the energy expression (143) by dropping all terms of third and higher order, the variational problem is now properly solved when $X_{\lambda ij}$ and $Y_{\Lambda IJ}$ satisfy

$$\left.\begin{aligned} \Sigma'_{\mu kl}E_{\lambda ij,\mu kl}X_{\mu kl} + \Sigma'_{\Xi KL}E_{\lambda ij,\Xi KL}Y_{\Xi KL} &= E_{\lambda ij} \quad , \\ \Sigma'_{\mu kl}E_{\mu kl,\Lambda IJ}X_{\mu kl} + \Sigma'_{\Xi KL}E_{\Lambda IJ,\Xi KL}Y_{\Xi KL} &= E_{\Lambda IJ} \quad . \end{aligned}\right\} \tag{150}$$

Solving Eqs.(150) for $X_{\lambda ij}$ and $Y_{\Lambda IJ}$ is a standard problem in linear algebra. The concatenated set of matrix elements $(X_{\lambda ij},Y_{\Lambda IJ})$ is the vector of independent variables; Eq.(143) defines the energy hypersurface as a function of these variables; $(-E_{\lambda ij},-E_{\Lambda IJ})$ constitutes the vector of first derivatives of the energy, also called the <u>gradient</u>; and $(E_{\lambda ij,\mu kl},E_{\lambda ij,\Lambda IJ},E_{\Lambda IJ,\Xi KL})$ constitutes the (symmetrical) matrix of second derivatives, also called the <u>Hessian</u>. The set of equations (150) has a unique solution if and only if the Hessian is non-singular. Since the vector $(X_{\lambda ij},Y_{\Lambda IJ})$ contains only essential variables, the Hessian being singular signals a pathological case. The latter can occur for a variety of reasons, prominent among which are a poor choice of SO's, CSF's, or TSF expansion coefficients. We shall assume from here on that we are dealing with a normal case, when the Hessian is non-singular. In order to solve Eqs.(150) without unnecessary loss of accuracy, we recommend the method known as "Gaussian elimination with pivoting".

The solution $(X_{\lambda ij},Y_{\Lambda IJ})$ of Eqs.(150) now defines the appropriate unitary transformation matrices $(e^X)_{\lambda ij}$ and $(e^Y)_{\Lambda IJ}$ for the expansion coefficients. Since the entire process is only accurate through quadratic terms in $X_{\lambda ij}$ and $Y_{\Lambda IJ}$ anyway, it suffices to truncate these coefficient transformations accordingly, namely

$$\left.\begin{aligned} c'_{\lambda pi} &= \Sigma_j c_{\lambda pj}(\delta_{ji} + X_{\lambda ji} + \tfrac{1}{2}\Sigma_k X_{\lambda jk}X_{\lambda ki}) \quad , \\ C'_{\Lambda PI} &= \Sigma_J C_{\Lambda PJ}(\delta_{JI} + Y_{\Lambda JI} + \tfrac{1}{2}\Sigma_K Y_{\Lambda JK}Y_{\Lambda KI}) \quad . \end{aligned}\right\} \quad . \tag{151}$$

Note that Eqs.(151) contains full sums over all possible index values; hence the set of matrix elements $X_{\lambda ij},Y_{\Lambda IJ}$ to be used here is obtained by "padding out" the essential non-redundant set, inserting opposite values or zeros for the missing ones as appropriate.

The truncation leaves a slight discrepancy in the orthonormality of the new coefficients $c'_{\lambda p i}$ and C'_{ApI}. This will be remedied automatically, however, if these coefficients are ever used for another cycle, since a rigorous prior orthonormalization of both sets of coefficients is a prerequisite for a valid calculation; we recommend the well-known Schmidt process for this.

Clearly when

$$\left.\begin{array}{lll} X_{\lambda ij} \to 0 & , & \lambda ij \in S_e \ , \\ Y_{AIJ} \to 0 & , & AIJ \subset T_e \ , \end{array}\right\} \qquad (152)$$

both transformations (151) reduce to the identity, and the orthonormal input coefficients $c_{\lambda p i}$ and C_{ApI} then constitute the converged MCSCF solution. Accordingly we define the <u>convergence</u> <u>measure</u> δ as the root-mean square of the matrix elements, namely

$$\delta = N_{XY}^{-1}(\Sigma'_{\lambda ij} X^2_{\lambda ij} + \Sigma'_{AIJ} Y^2_{AIJ})^{\frac{1}{2}} \quad , \qquad (153)$$

where N_{XY} is the number of essential non-redundant variables $X_{\lambda ij}, Y_{AIJ}$. The MCSCF process is then considered to have converged when

$$\delta < t \quad , \qquad (154)$$

where t is the <u>convergence</u> <u>threshold</u> chosen for the particular calculation.

Since in the normal case the Hessian is non-singular, convergence occurs if and only if

$$\left.\begin{array}{lll} E_{\lambda ij} \to 0 & , & \lambda ij \subset S_e \ , \\ E_{AIJ} \to 0 & , & AIJ \subset T_e \ . \end{array}\right\} \qquad (155)$$

The first Eq.(155) is the generalized Brillouin theorem in the context of our MCSCF model, while the second Eq.(155) signals that the configuration interaction secular equation system is satisfied by the occupied terms.

IX. DETERMINATION OF THE ESSENTIAL NON-REDUNDANT VARIABLES

The transformations to which the N-electron wave functions are subjected in any variational MCSCF process constitute a Lie group, which we accordingly call the <u>variational</u> <u>group</u>. This group is separate and distinct from the <u>symmetry</u> <u>group</u> of the system, which consists of all operators which commute with the Hamiltonian. The symmetry group of course guarantees symmetry characteristics -- species and subspecies -- for the wave functions which satisfy the Schrödinger equation. In our MCSCF scheme we take advantage of this by restricting the variational wave functions accordingly, which in turn shrinks the variational group from what it would be without such an constraint. We thus obtain the permissible variational group, consisting of transformations which preserve symmetry species and subspecies.

The permissible variational group is actually the product of two groups. The SO induced transformations $\mathcal{U}$ of Eq.(88) form a group $\mathcal{U}$; similarly the direct TSF transformations $\mathcal{V}$ of Eq.(96) form a group $\mathcal{V}$. The products $\mathcal{UV}$ of Eq.(97) now form the permissible variational group, which we shall accordingly designate by $\mathcal{UV}$. Obviously $X_{\lambda ij}$, $\lambda ij \subset S$ and Y_{AIJ}, $AIJ \subset T$ are the group parameters of $\mathcal{U}$ and $\mathcal{V}$, respectively; together $X_{\lambda ij}, Y_{AIJ}$ are the group parameters of $\mathcal{UV}$. Corresponding to each of these parameters we define the elementary rotations

$$\left.\begin{array}{lll} u_{\lambda ij} = e^{X_{\lambda ij}(\Phi_{\lambda ji} - \Phi_{\lambda ij})} & , & \lambda ij \in S \ , \\ v_{AIJ} = e^{Y_{AIJ}(\Psi_{AJI} - \Psi_{AIJ})} & , & AIJ \subset T \ . \end{array}\right\} \qquad (156)$$

In view of our earlier discussion of MCSCF models, the MCSCF process is invariant if the operator $\mathcal{U}$ causes the CSF's to be transformed among themselves, namely

$$\mathcal{U}\Phi_{APA} = \Sigma_Q \Phi_{AQA} U_{AQP} \quad . \tag{157}$$

where the U_{APQ} are the matrix elements of the transformation. It is easily seen that the operators $\mathcal{U}$ which satisfy Eq.(157) again form a group, which is of course a subgroup of U; this subgroup consists of non-essential rotations, and will therefore be designated by U_n. If an elementary rotation $\mathcal{U}_{\lambda ij}$ satisfies Eq.(157), the corresponding group parameter $X_{\lambda ij}$ is non-essential; this is expressed by

$$\mathcal{U}_{\lambda ij}\Phi_{APA} = \Sigma_Q \Phi_{AQA} U_{\lambda ij,AQP} \quad , \quad \lambda ij \in S_n \quad . \tag{158}$$

where the $U_{\lambda ij,APQ}$ are the matrix elements of the transformation.

An important special case of Eq.(157) occurs when all CSF's remain unchanged, namely

$$\mathcal{U}\Phi_{APA} = \Phi_{APA} \quad . \tag{159}$$

The operators $\mathcal{U}$ which satisfy Eq.(159) again form a group, which is of course a subgroup of U_n; we shall designate this subgroup by U_i. The group parameters of U_i will be designated by S_i; hence

$$\mathcal{U}_{\lambda ij}\Phi_{APA} = \Phi_{APA} \quad , \quad \lambda ij \in S_i \quad . \tag{160}$$

The condition (160) leads to a non-essential variable because if the CSF's are invariant, so are the TSF's, and therefore also the energy. The corresponding $X_{\lambda ij}$ can simply be dropped from the algebra, which is equivalent to putting $X_{\lambda ij}=0$. For the more general condition (158) the argument is somewhat more complicated. In this case the effect of the induced transformation of the CSF's by $\mathcal{U}_{\lambda ij}$ can also be obtained by a direct TSF transformation of the type $e\Psi$. The precise matching of these two operations leads to exactly one non-essential linear combination of the applicable $X_{\lambda ij}$ and some or all of the Y_{AIJ}'s. It is not necessary, however, to nail down this linear combination explicitly; it suffices to drop the offending $X_{\lambda ij}$ from the algebra, leaving the Y_{AIJ}'s as potential essential variables (some of them may turn out to be non-essential for other reasons, see below).

We paraphrase our previous conclusions concerning shell structure and invariance of the MCSCF process as follows. Rotations $\mathcal{U}_{\lambda ij}$ which satisfy the condition (158) but not (160) are always entirely within the set of fractional shells F. Rotations $\mathcal{U}_{\lambda ij}$ which satisfy the conditions (160) are: all such rotations within the core C; all such rotations within the set of empty shells E; and occasionally some such rotations within the set of fractional shells F. On the other hand, rotations between shells belonging to any two different categories out of C,F,E cannot satisfy the condition (158), and invariance of the energy under such a rotation would be an extremely rare numerical accident. Barring such an exceptional occurence we summarize our findings by

$$\left.\begin{array}{ll}
\lambda ij \in S_n \quad \text{if} & \left\{\begin{array}{lll} \lambda i \in C & \text{and} & \lambda j \in C \; , \\ \lambda i \in E & \text{and} & \lambda j \in E \; , \end{array}\right. \\
\lambda ij \in S_e \quad \text{if} & \left\{\begin{array}{lll} \lambda i \in C & \text{and} & \lambda j \in F \; , \\ \lambda i \in C & \text{and} & \lambda j \in E \; , \\ \lambda i \in F & \text{and} & \lambda j \in C \; , \\ \lambda i \in F & \text{and} & \lambda j \in E \; , \\ \lambda i \in E & \text{and} & \lambda j \in C \; , \\ \lambda i \in E & \text{and} & \lambda j \in F \; . \end{array}\right.
\end{array}\right\} \tag{161}$$

The conditions (161) leave unresolved the ambiguous case and . The resolution of this ambiguity depends on the MCSCF model used. For the half-shell model we have

$$\left.\begin{array}{llllll} \lambda ij \subset S_m & \text{if} & & \lambda i \subset H & \text{and} & \lambda j \subset H \;, \\ & & \multirow{4}{*}{} & \lambda i \subset H & \text{and} & \lambda j \subset F' \;, \\ \lambda ij \subset S_e & \text{if} & \left\{\begin{array}{l} \\ \\ \\ \\ \end{array}\right. & \lambda i \subset F' & \text{and} & \lambda j \subset H \;, \\ & & & \lambda i \subset F' & \text{and} & \lambda j \subset F' \;, \end{array}\right\} \qquad (161')$$

while for the fractionally saturated model

$$\lambda ij \subset S_m \quad \text{if} \quad \lambda i \subset F \quad \text{and} \quad \lambda j \subset F \quad , \qquad (161'')$$

and for the even-replacement model

$$\lambda ij \subset S_e \quad \text{if} \quad \lambda i \subset F \quad \text{and} \quad \lambda j \subset F \quad . \qquad (161\ )$$

We call attention to the fact that the variational variables identified by S_i, S_m and S are the group parameters of the nested groups U_i, U_m and U. On the other hand the essential variables identified by S_e in general do not constitute a set of parameters of any group.

With respect to direct TSF transformations, it is shown in Appendix C that the necessary and sufficient condition for a non-essential rotation $\mathcal{V}$ is given by

$$[\![\mathcal{W},\mathcal{V}]\!] = \mathcal{O} \quad . \qquad (162)$$

The rotations which satisfy Eq.(162) form a group V_m, which is of course a subgroup of V. If an elementary rotation $\mathcal{V}_{AIJ}$ satisfies Eq.(162), the corresponding group parameter Y_{AIJ} is non-essential; this is expressed by

$$[\![\mathcal{W},\mathcal{V}_{AIJ}]\!] = \mathcal{O} \quad , \quad AIJ \subset T_m \quad , \qquad (163)$$

which is equivalent to

$$(W_{AI}-W_{AJ})Y_{AIJ} = 0 \quad , \quad AIJ \subset T_m \quad . \qquad (164)$$

Hence a TSF rotational variable is essential if and only if the weights of the two spectroscopic terms connected by the rotation are different, as expressed by

$$\left.\begin{array}{lll} AIJ \subset T_m & \text{if} & W_{AI} = W_{AJ} \quad , \\ AIJ \subset T_e & \text{if} & W_{AI} \neq W_{AJ} \quad . \end{array}\right\} \qquad (165)$$

It should be noted that T_m and T identify the group parameters of V_m and V, while the essential variables identified by T_e in general do not constitute a set of parameters of any group.

X. EVALUATION OF THE REFERENCE ENERGY, GRADIENT AND HESSIAN

We now turn to the actual evaluation of the reference energy, gradient and Hessian, defined by Eqs.(128,133-137,144-148), in terms of weights, structure constants and reduced matrix elements. On account of the techniques used for this evaluation, the six types of quantities fall naturally into three groups, depending on the number of SO rotation subscripts; thus the three groups consist of ($\mathcal{E}$, $\mathcal{E}_{AIJ}$, $\mathcal{E}_{AIJ,IKL}$), ($\mathcal{E}_{\lambda ij}$, $\mathcal{E}_{\lambda ij,AIJ}$) and $\mathcal{E}_{\lambda ij,\mu kl}$, respectively.

In the first group considerable economy is achieved by using as intermediate quantities the matrix elements of the Hamiltonian between TSF's defined by

$$H_{AIJ} = \langle \Psi_{AIA} | \hat{\mathcal{H}} | \Psi_{AJA} \rangle = D_A^{-1}\,\mathrm{Trace}(\mathcal{P}_{AIJ}\hat{\mathcal{H}}) \quad ; \qquad (166)$$

these matrix elements are obviously real symmetrical, namely

$$\bar{H}_{AIJ} = H_{AIJ} = H_{AJI} \quad . \qquad (167)$$

The reduction of E to the matrix elements H_{AIJ} needs no comment. For the analogous reduction of E_{AIJ} and $E_{AIJ,BKL}$ we use Eq.(94) to evaluate the relevant commutators of $\mathcal{W}$, namely

$$\llbracket \mathcal{W}, \varphi_{AIJ} \rrbracket = -(W_{AI}-W_{AJ})\varphi_{AIJ} \quad , \tag{168}$$

$$\llbracket \llbracket \mathcal{W}, \varphi_{AIJ} \rrbracket, \varphi_{BKL} \rrbracket = \delta_{AB}(W_{AI}-W_{AJ})(\delta_{KJ}\varphi_{AIL}-\delta_{IL}\varphi_{AKJ}) \quad . \tag{169}$$

Using these results we find after some algebra

$$E = \Sigma_{AI} D_A W_{AI} H_{AII} \quad , \tag{170}$$

$$E_{AIJ} = 2D_A(W_{AI}-W_{AJ})H_{AIJ} \quad , \tag{171}$$

$$\begin{aligned} E_{AIJ,BKL} = \delta_{AB} D_A [& (W_{AI}-W_{AJ}-W_{AK}+W_{AL})(\delta_{JK}H_{AIL}-\delta_{IL}H_{AJK}) \\ & - (W_{AI}-W_{AJ}+W_{AK}-W_{AL})(\delta_{JL}H_{AIK}-\delta_{IK}H_{AJL})] \quad . \end{aligned} \tag{172}$$

Expressing now the Hamiltonian matrix elements between TSF's in terms of matrix elements between CSF'S we obtain

$$H_{AIJ} = \Sigma_{PQ} h_{APQ} C_{API} C_{AQJ} \quad , \tag{173}$$

where we have used the fact that the C_{API} form a real orthogonal matrix, while the matrix elements between CSF's are given by

$$h_{APQ} = \langle \Phi_{APA} | \hat{\mathcal{H}} | \Phi_{AQA} \rangle \quad ; \tag{174}$$

they are obviously real symmetrical, namely

$$\bar{h}_{APQ} = h_{APQ} = h_{AQP} \quad . \tag{175}$$

The Hamiltonian matrix elements between CSF's are now readily evaluated in terms of structure constants and reduced matrix elements. Using Eqs.(117,119-121,124,174) we find after some algebra

$$h_{APP} = F_{AP} + \tfrac{1}{2}J_{AP} - \tfrac{1}{2}K_{AP} \quad , \tag{176}$$

$$h_{APQ} = F_{APQ} + \tfrac{1}{2}G_{APQ} \quad , \quad P \neq Q \quad , \tag{177}$$

where

$$F_{AP} = \Sigma_{\lambda i} n_{AP,\lambda i} F_{\lambda ii} \quad , \tag{178}$$

$$F_{APQ} = \Sigma_{\lambda ij} n_{APQ,\lambda ij} F_{\lambda ji} \quad , \quad P \neq Q \quad , \tag{179}$$

$$J_{AP} = \Sigma_{\kappa} \Sigma_{\lambda i} \Sigma_{\mu j} c_{AP,\lambda i,\mu j,\kappa} G_{\lambda i \lambda i,\mu j \mu j,\kappa} \quad , \tag{180}$$

$$K_{AP} = \Sigma_{\kappa} \Sigma_{\lambda i} \Sigma_{\mu j} x_{AP,\lambda i,\mu j,\kappa} G_{\lambda i \mu j,\mu j \lambda i,\kappa} \quad , \tag{181}$$

$$G_{APQ} = \Sigma_{\kappa} \Sigma_{\lambda i \mu j} \Sigma_{\nu k \rho l} n_{APQ,\lambda i \mu j,\nu k \rho l,\kappa} G_{\mu j \lambda i,\rho l \nu k,\kappa} \quad , \quad P \neq Q \quad . \tag{182}$$

For the second group we use similar procedures. We introduce the matrix elements of the commutator $\llbracket \hat{\mathcal{H}}, P_{\lambda ij} \rrbracket$ between TSF's, namely

$$\begin{aligned} H_{\lambda ij,AIJ} &= \langle \Phi_{AIA} | \llbracket \hat{\mathcal{H}}, P_{\lambda ij} \rrbracket | \Phi_{AJA} \rangle \\ &= D_A^{-1}\,\mathrm{Trace}(\varphi_{AIJ} \llbracket \hat{\mathcal{H}}, P_{\lambda ij} \rrbracket) \quad , \end{aligned} \tag{183}$$

which satisfy

$$\bar{H}_{\lambda ij,AIJ} = H_{\lambda ij,AIJ} = -H_{\lambda ji,AJI} \quad . \tag{184}$$

The reduction of $E_{\lambda ij}$ and $E_{\lambda ij,AIJ}$ in terms of these matrix elements now yields

$$E_{\lambda ij} = 2\Sigma_{AI} D_A W_{AI} H_{\lambda ij,AII} \quad , \tag{185}$$

$$E_{\lambda ij,AIJ} = 2D_A(W_{AI}-W_{AJ})(H_{\lambda ij,AIJ}+H_{\lambda ij,AJI}) \quad . \tag{186}$$

The matrix elements of $\llbracket \hat{\mathcal{H}}, P_{\lambda ij} \rrbracket$ between TSF's are now expressed in terms of the corresponding matrix elements between CSF's, namely

$$H_{\lambda ij,AIJ} = \Sigma_{PQ} h_{\lambda ij,APQ} C_{API} C_{AQJ} \quad . \tag{187}$$

where

$$h_{\lambda ij,APQ} = \langle \Phi_{APA} | [\hat{\mathcal{H}}, \mathcal{E}_{\lambda ij}] | \Phi_{AQA} \rangle \quad . \tag{188}$$

which satisfy

$$\bar{h}_{\lambda ij,APQ} = h_{\lambda ij,APQ} = -h_{\lambda ji,AQP} \quad . \tag{189}$$

To express the matrix elements (188) in terms of structure constants and reduced matrix elements we evaluate first the commutator of the Hamiltonian using Eqs.(82,113,116,124), obtaining

$$\begin{aligned}[\hat{\mathcal{H}}, \mathcal{E}_{\lambda ij}] = {} & \Sigma_k (\mathcal{E}_{\lambda ik} F_{\lambda kj} - F_{\lambda ik} \mathcal{E}_{\lambda kj}) \\ & + \Sigma_\kappa \Sigma_{\mu k} \Sigma_{\nu\ell pm} (\mathcal{E}_{\lambda i\mu k,\nu\ell pm,\kappa} G_{\mu k\lambda j,pm\nu\ell,\kappa} \\ & - G_{\lambda i\mu k,\nu\ell pm,\kappa} \mathcal{E}_{\mu k\lambda j,pm\nu\ell,\kappa}) \quad . \end{aligned} \tag{190}$$

Using now Eqs.(117,119-121,124,188,190) we find after some algebra

$$h_{\lambda ij,APP} = F_{\lambda ij,AP} + J_{\lambda ij,AP} - K_{\lambda ij,AP} \quad , \tag{191}$$

$$h_{\lambda ij,APQ} = F_{\lambda ij,APQ} + G_{\lambda ij,APQ} \quad , \quad P \neq Q \quad , \tag{192}$$

where

$$F_{\lambda ij,AP} = (n_{AP,\lambda i} - n_{AP,\lambda j}) F_{\lambda ij} \quad , \tag{193}$$

$$F_{\lambda ij,APQ} = \Sigma_k (n_{APQ,\lambda ik} F_{\lambda kj} - F_{\lambda ik} n_{APQ,\lambda kj}) \quad , \quad P \neq Q \quad , \tag{194}$$

$$J_{\lambda ij,AP} = \Sigma_\kappa \Sigma_{\mu k} (c_{AP,\lambda i,\mu k,\kappa} - c_{AP,\lambda j,\mu k,\kappa}) G_{\lambda i\lambda j,\mu k\mu k,\kappa} \quad , \tag{195}$$

$$K_{\lambda ij,AP} = \Sigma_\kappa \Sigma_{\mu k} (x_{AP,\lambda i,\mu k,\kappa} - x_{AP,\lambda j,\mu k,\kappa}) G_{\lambda i\mu k,\mu k\lambda j,\kappa} \quad , \tag{196}$$

$$\begin{aligned} G_{\lambda ij,APQ} = {} & \Sigma_\kappa \Sigma_{\mu k} \Sigma_{\nu\ell pm} (n_{APQ,\lambda i\mu k,\nu\ell pm,\kappa} G_{\mu k\lambda j,pm\nu\ell,\kappa} \\ & - G_{\lambda i\mu k,\nu\ell pm,\kappa} n_{APQ,\mu k\lambda j,pm\nu\ell,\kappa}) \quad , \quad P \neq Q \quad . \end{aligned} \tag{197}$$

The evaluation of $E_{\lambda ij,\mu k\ell}$ can be accomplished in similar fashion. Starting from Eq.(190), one derives a formula for the double commutator $[[\hat{\mathcal{H}}, \mathcal{E}_{\lambda ij}], \mathcal{E}_{\mu k\ell}]$ which is linear in the one- and two-electron $\mathcal{E}$-operators, using again the commutation relations (82,116). This formula is rather long, and since it constitutes an intermediate result, we omit its explicit presentation. Since $E_{\lambda ij,\mu k\ell}$ carries no TSF labels, we have now no particular need for the explicit matrix elements of the double commutator between TSF's and/or CSF's. Hence in this case it is profitable to first carry out the sums over the CSF's and TSF's directly on the structure constants.

We define weight factors with respect to CSF's by

$$w_{APQ} = \Sigma_I W_{AI} C_{API} C_{AQI} \quad , \tag{198}$$

which of course satisfy

$$\bar{w}_{APQ} = w_{APQ} = w_{AQP} \quad . \tag{199}$$

Using these weight factors we can now define the net structure constants

$$n_{\lambda i} = \Sigma_{AP} D_A w_{APP} n_{AP,\lambda i} \quad , \tag{200}$$

$$n_{\lambda ij} = \Sigma_{APQ,P\neq Q} D_A w_{APQ} n_{AQP,\lambda ij} \quad , \tag{201}$$

$$c_{\lambda i,\mu j,\kappa} = \Sigma_{AP} D_A w_{APP} c_{AP,\lambda i,\mu j,\kappa} \quad , \tag{202}$$

$$x_{\lambda i,\mu j,\kappa} = \Sigma_{AP} D_A w_{APP} x_{AP,\lambda i,\mu j,\kappa} \quad , \tag{203}$$

$$n_{\lambda i\mu j,\nu k p\ell,\kappa} = \Sigma_{APQ,P\neq Q} D_A w_{APQ} n_{AQP,\lambda i\mu j,\nu k p\ell,\kappa} \quad . \tag{204}$$

The final formulas for $E_{\lambda ij,\mu k\ell}$ are now given in terms of these net structure constants and reduced matrix elements. After a considerable amount of algebra we find

$$\begin{aligned} E_{\lambda ij,\mu k\ell} &= F_{\lambda ij,\mu k\ell} + F'_{\lambda ij,\mu k\ell} - 2F''_{\lambda ij,\mu k\ell} \\ &\quad + J_{\lambda ij,\mu k\ell} + 2J'_{\lambda ij,\mu k\ell} - K_{\lambda ij,\mu k\ell} - 2K'_{\lambda ij,\mu k\ell} \\ &\quad + G_{\lambda ij,\mu k\ell} - 2G'_{\lambda ij,\mu k\ell} - 2G''_{\lambda ij,\mu k\ell} \quad . \end{aligned} \tag{205}$$

where

$$\begin{aligned} F_{\lambda ij,\mu k\ell} &= \delta_{\lambda\mu}[(n_{\lambda i} - n_{\lambda j} - n_{\lambda k} + n_{\lambda\ell})(\delta_{jk}F_{\lambda i\ell} - \delta_{i\ell}F_{\lambda jk}) \\ &\qquad - (n_{\lambda i} - n_{\lambda j} + n_{\lambda k} - n_{\lambda\ell})(\delta_{j\ell}F_{\lambda ik} - \delta_{ik}F_{\lambda j\ell})] \quad , \end{aligned} \tag{206}$$

$$\begin{aligned} F'_{\lambda ij,\mu k\ell} &= \delta_{\lambda\mu}\Sigma_m[\delta_{jk}(n_{\lambda im}F_{\lambda m\ell} + F_{\lambda im}n_{\lambda m\ell}) \\ &\qquad + \delta_{i\ell}(n_{\lambda jm}F_{\lambda mk} + F_{\lambda jm}n_{\lambda mk}) \\ &\qquad - \delta_{j\ell}(n_{\lambda im}F_{\lambda mk} + F_{\lambda im}n_{\lambda mk}) \\ &\qquad - \delta_{ik}(n_{\lambda jm}F_{\lambda m\ell} + F_{\lambda jm}n_{\lambda m\ell})] \quad , \end{aligned} \tag{207}$$

$$F''_{\lambda ij,\mu k\ell} = \delta_{\lambda\mu}(n_{\lambda jk}F_{\lambda i\ell} + n_{\lambda i\ell}F_{\lambda jk} - n_{\lambda j\ell}F_{\lambda ik} - n_{\lambda ik}F_{\lambda j\ell}) \quad , \tag{208}$$

$$\begin{aligned} J_{\lambda ij,\mu k\ell} &= \delta_{\lambda\mu}\Sigma_\kappa\Sigma_{\nu m}[(c_{\lambda i,\nu m,\kappa} - c_{\lambda j,\nu m,\kappa} - c_{\lambda k,\nu m,\kappa} + c_{\lambda\ell,\nu m,\kappa}) \\ &\qquad \times (\delta_{jk}G_{\lambda i\lambda\ell,\nu m\nu m,\kappa} - \delta_{i\ell}G_{\lambda j\lambda k,\nu m\nu m,\kappa}) \\ &\qquad - (c_{\lambda i,\nu m,\kappa} - c_{\lambda j,\nu m,\kappa} + c_{\lambda k,\nu m,\kappa} - c_{\lambda\ell,\nu m,\kappa}) \\ &\qquad \times (\delta_{j\ell}G_{\lambda i\lambda k,\nu m\nu m,\kappa} - \delta_{ik}G_{\lambda j\lambda\ell,\nu m\nu m,\kappa})] \quad , \end{aligned} \tag{209}$$

$$\begin{aligned} J'_{\lambda ij,\mu k\ell} &= \Sigma_\kappa(c_{\lambda i,\mu k,\kappa} - c_{\lambda i,\mu\ell,\kappa} - c_{\lambda j,\mu k,\kappa} + c_{\lambda j,\mu\ell,\kappa}) \\ &\qquad \times (G_{\lambda i\lambda j,\mu k\mu\ell,\kappa} + G_{\lambda i\lambda j,\mu\ell\mu k,\kappa}) \quad , \end{aligned} \tag{210}$$

$$\begin{aligned} K_{\lambda ij,\mu k\ell} &= \delta_{\lambda\mu}\Sigma_\kappa\Sigma_{\nu m}[(x_{\lambda i,\nu m,\kappa} - x_{\lambda j,\nu m,\kappa} - x_{\lambda k,\nu m,\kappa} + x_{\lambda\ell,\nu m,\kappa}) \\ &\qquad \times (\delta_{jk}G_{\lambda i\nu m,\nu m\lambda\ell,\kappa} - \delta_{i\ell}G_{\lambda j\nu m,\nu m\lambda k,\kappa}) \\ &\qquad - (x_{\lambda i,\nu m,\kappa} - x_{\lambda j,\nu m,\kappa} + x_{\lambda k,\nu m,\kappa} - x_{\lambda\ell,\nu m,\kappa}) \\ &\qquad \times (\delta_{j\ell}G_{\lambda i\nu m,\nu m\lambda k,\kappa} - \delta_{ik}G_{\lambda j\nu m,\nu m\lambda\ell,\kappa})] \quad , \end{aligned} \tag{211}$$

$$\begin{aligned} K'_{\lambda ij,\mu k\ell} &= \Sigma_\kappa(x_{\lambda i,\mu k,\kappa} - x_{\lambda i,\mu\ell,\kappa} - x_{\lambda j,\mu k,\kappa} + x_{\lambda j,\mu\ell,\kappa}) \\ &\qquad \times (G_{\lambda i\mu\ell,\mu k\lambda j,\kappa} + G_{\lambda i\mu k,\mu\ell\lambda j,\kappa}) \quad , \end{aligned} \tag{212}$$

$$\begin{aligned} G_{\lambda ij,\mu k\ell} &= \delta_{\lambda\mu}\Sigma_\kappa\Sigma_{\nu m}\Sigma_{\rho n\sigma p}[\delta_{jk}(n_{\lambda i\nu m,\rho n\sigma p,\kappa}G_{\nu m\lambda\ell,\sigma p\rho n,\kappa} \\ &\qquad\qquad + G_{\lambda i\nu m,\rho n\sigma p,\kappa}n_{\nu m\lambda\ell,\sigma p\rho n,\kappa}) \\ &\qquad + \delta_{i\ell}(n_{\lambda j\nu m,\rho n\sigma p,\kappa}G_{\nu m\lambda k,\sigma p\rho n,\kappa} \\ &\qquad\qquad + G_{\lambda j\nu m,\rho n\sigma p,\kappa}n_{\nu m\lambda k,\sigma p\rho n,\kappa}) \\ &\qquad - \delta_{j\ell}(n_{\lambda i\nu m,\rho n\sigma p,\kappa}G_{\nu m\lambda k,\sigma p\rho n,\kappa} \\ &\qquad\qquad + G_{\lambda i\nu m,\rho n\sigma p,\kappa}n_{\nu m\lambda k,\sigma p\rho n,\kappa}) \\ &\qquad - \delta_{ik}(n_{\lambda j\nu m,\rho n\sigma p,\kappa}G_{\nu m\lambda\ell,\sigma p\rho n,\kappa} \\ &\qquad\qquad + G_{\lambda j\nu m,\rho n\sigma p,\kappa}n_{\nu m\lambda\ell,\sigma p\rho n,\kappa})] \quad , \end{aligned} \tag{213}$$

$$\begin{aligned} G'_{\lambda ij,\mu k\ell} &= \Sigma_\kappa\Sigma_{\nu m\rho n}(n_{\lambda j\mu k,\nu m\rho n,\kappa}G_{\lambda i\mu\ell,\nu m\rho n,\kappa} \\ &\qquad + n_{\lambda i\mu\ell,\nu m\rho n,\kappa}G_{\lambda j\mu k,\nu m\rho n,\kappa} \\ &\qquad - n_{\lambda j\mu\ell,\nu m\rho n,\kappa}G_{\lambda i\mu k,\nu m\rho n,\kappa} \\ &\qquad - n_{\lambda i\mu k,\nu m\rho n,\kappa}G_{\lambda j\mu\ell,\nu m\rho n,\kappa}) \quad , \end{aligned} \tag{214}$$

$$\begin{aligned}G''_{\lambda ij,\mu kl} = \Sigma_\kappa \Sigma_{\nu m} \Sigma_{\rho n} (& n_{\lambda j\nu m,\mu k\rho n,\kappa} G_{\lambda i\nu m,\mu l\rho n,\kappa} \\ & + n_{\lambda j\nu m,\rho n\mu k,\kappa} G_{\lambda i\nu m,\rho n\mu l,\kappa} \\ & + n_{\lambda i\nu m,\mu l\rho n,\kappa} G_{\lambda j\nu m,\mu k\rho n,\kappa} \\ & + n_{\lambda i\nu m,\rho n\mu l,\kappa} G_{\lambda j\nu m,\rho n\mu k,\kappa} \\ & - n_{\lambda j\nu m,\mu l\rho n,\kappa} G_{\lambda i\nu m,\mu k\rho n,\kappa} \\ & - n_{\lambda j\nu m,\rho n\mu l,\kappa} G_{\lambda i\nu m,\rho n\mu k,\kappa} \\ & - n_{\lambda i\nu m,\mu k\rho n,\kappa} G_{\lambda j\nu m,\mu l\rho n,\kappa} \\ & - n_{\lambda i\nu m,\rho n\mu k,\kappa} G_{\lambda j\nu m,\rho n\mu l,\kappa}) \quad . \qquad (215)\end{aligned}$$

XI. THE GLOBAL MCSCF PROCEDURE

At the beginning of an MCSCF cycle we have in hand, in logical order:

1) one- and two-electron integrals over basis functions;
2) approximate expansion coefficients which define reference SO's in terms of basis functions;
3) structure constants which describe the construction of CSF's from SO's;
4) lists of index pairs identifying the essential rotations of the SO's and TSF's;
5) approximate expansion coefficients which define reference TSF's in terms of CSF's;
6) the weights with which the TSF's enter the variational energy expression.

The integrals over basis functions, the structure constants, the index lists and the weights remain constant throughout the entire MCSCF procedure. At the end of each MCSCF cycle we will have in hand improved expansion coefficients, both for the SO's and TSF's, together with the convergence measure δ, to assess the quality of the new coefficients. When $\delta<\epsilon$, where ϵ is the convergence threshold chosen for the particular calculation, we say that the converged MCSCF solution has been obtained.

The MCSCF cycle is conveniently divided into the following steps:

1) orthonormalization of the input expansion coefficients $c_{\lambda pi}$ and C_{API}, which define the reference SO's and TSF's, respectively;
2) calculation of the integrals over SO's, $F_{\lambda ij}$ and $G_{\lambda i\mu j,\nu k\rho l,\kappa}$, from the integrals over basis functions, $f_{\lambda pq}$ and $g_{\lambda p\mu q,\nu r\rho s,\kappa}$, respectively;
3) calculation of the reference energy E and the energy derivatives $E_{\lambda ij}$, E_{AIJ}, $E_{\lambda ij,\mu kl}$, $E_{\lambda ij,AIJ}$, $E_{AIJ,BKL}$;
4) determination of the essential rotation matrices $X_{\lambda ij}$, $\lambda ij\in S_\epsilon$, Y_{AIJ}, $AIJ\in T_\epsilon$, and the MCSCF convergence measure δ;

5a) if $\delta<\epsilon$, terminate the calculation, accepting the orthonormalized input coefficients $c_{\lambda pi}$ and C_{API} as the converged MCSCF solution;

5b) if $\delta\geqslant\epsilon$, prepare for another MCSCF cycle, obtaining new coefficients $c'_{\lambda pi}$, C'_{API} by applying the unitary transformations $(e^X)_{\lambda ij}$, $(e^Y)_{AIJ}$ to the old coefficients $c_{\lambda pi}$, C_{API}.

The MCSCF process as described is the multi-dimensional generalization of the Newton-Raphson process, and it is therefore quadratically convergent. This means that if we use input coefficients which are in error by ϵ, after one MCSCF cycle the new coefficients are in error by ϵ^2. Hence starting with an error of 10^{-1}, we obtain an accuracy of 10^{-2}, 10^{-4}, 10^{-8},... after one, two, three,... iterations. And since this behavior only depends on the quadratic approximation of our energy surface, maxima, minima or saddle points are attainable with equal facility.

Occasionally MCSCF procedures involving cubic as well as quadratic terms in the variational variables have been considered. In our opinion this is not a particularly fruitful proposition for an MCSCF scheme laying claim to any sort of generality.

It should be emphasized that when choosing an MCSCF model one has to ferret out the non-essential variables $X_{\lambda ij}$ and Y_{AIJ} before proceeding with the numerical calculation. If a non-essential variable is left in inadvertently, the Hessian will be singular, and we are dealing with a pathological case. Other pathological situations, for instance when the Hessian becomes near-singular, invariably are caused by a poor choice of parameters. Whatever the reason may be for a Hessian which is singular or near-singular, the remedy for such a case is to rethink the problem and set it up properly, rather than to attempt solving an ill-behaved set of equations.

It should also be noted that it is not at all necessary that the Hessian be positive definite in order to guarantee a well-behaved and stable MCSCF process. Clearly, the MCSCF process is stable whenever the eigenvalues of the Hessian, positive or negative, do not vanish (within a reasonable threshold, of course).

Finally we like to emphasize that the procedures we have laid out permit the construction of a general MCSCF program which is equally applicable to atoms as well as molecules, while at the same time realizing the maximum possible benefits from the exploitation of symmetry.

ACKNOWLEDGEMENTS

The work described in this paper came to fruition during a period of feverish activity in this particular area of theoretical chemical physics. We had many fruitful discussions with several colleagues and students. In particular we like to thank Prof. Walter England, Dr. Arnold C. Wahl, Dr. Michel Dupuis, and graduate students Maurice Sheppard and Anthony Haymet, for their helpful comments, criticisms and encouragement. We also express our appreciation for the administrative support by Dr. Darrel G. Hopper at SAI during the initial phase, and by Dr. William A. Lester, Jr. at NRCC during the second phase when this work reached maturity.

APPENDIX A

First we prove that if $\mathcal{X}$ is an anti-Hermitian operator, then

$$\mathcal{U} = e^{\mathcal{X}} \tag{A1}$$

is unitary. Obviously $\mathcal{X}$ satisfies

$$\left.\begin{aligned} \mathcal{X}^{\dagger} + \mathcal{X} &= \mathcal{O} \quad , \\ [\![\mathcal{X}^{\dagger}, \mathcal{X}]\!] &= \mathcal{O} \quad . \end{aligned}\right\} \tag{A2}$$

Using now Eqs.(A1,A2) we find

$$\mathcal{U}^{\dagger}\mathcal{U} = e^{\mathcal{X}^{\dagger}} e^{\mathcal{X}} = e^{\mathcal{X}^{\dagger}+\mathcal{X}} = e^{\mathcal{O}} = \mathcal{1} \quad , \tag{A3}$$

that is, $\mathcal{U}$ is indeed unitary.

Next we must prove that, if $\mathcal{U}$ is a unitary operator, we can find an anti-Hermitian operator $\mathcal{X}$ so that Eq.(A1) is valid. We expand $\mathcal{U}$ in the operator base $\mathcal{P}_{ij}$ according to

$$\mathcal{U} = \Sigma_{ij} \mathcal{P}_{ij} U_{ji} \quad . \tag{A4}$$

Using now matrix notation, the matrix U can be diagonalized by another unitary matrix V. Since all eigenvalues of U have unit modulus, this diagonalization can be written in the form

$$V^{\dagger} U V = e^{i\Phi} \quad , \tag{A5}$$

where the matrix Φ is real diagonal, namely

$$\left.\begin{aligned} \Phi_{ij} &= \delta_{ij}\phi_i \quad , \\ \bar{\phi}_i &= \phi_i \quad ; \end{aligned}\right\} \tag{A6}$$

we can further make the angles ϕ_i unique by choosing

$$0 < \phi_i < 2\pi \quad . \qquad \text{(A7)}$$

From Eqs.(A5-A7) follows

$$U = e^{X} \quad , \qquad \text{(A8)}$$

where

$$X = i V \Phi V^{\dagger} \quad ; \qquad \text{(A9)}$$

it is easily seen that X is anti-Hermitian. The operator $\mathcal{X}$ which is related to $\mathcal{U}$ by Eq.(A1) is now given by

$$\mathcal{X} = \Sigma_{ij} \mathcal{P}_{ij} X_{ji} \quad . \qquad \text{(A10)}$$

APPENDIX B

Starting with tha variation of the energy, Eq.(32), we find after some algebra

$$\delta E = \Sigma_{i>j}(w_i - w_j)[(h_{ij} + h_{ji})\mathrm{Re}(X_{ji}) + i(h_{ij} - h_{ji})\mathrm{Im}(X_{ji})] + O_2(X_{ij}) \quad , \qquad \text{(B1)}$$

where

$$h_{ij} = \langle \psi_i | h | \psi_j \rangle \quad . \qquad \text{(B2)}$$

Since $\mathrm{Re}(X_{ij})$ and $\mathrm{Im}(X_{ij})$ are independent, and since by assumption $w_i \neq w_j$, the energy is stationary if and only if

$$\left.\begin{array}{ll} h_{ij} + h_{ji} = 0 \quad , & i > j \quad , \\ h_{ij} - h_{ji} = 0 \quad , & i > j \quad . \end{array}\right\} \qquad \text{(B3)}$$

which is equivalent to

$$h_{ij} = \langle \psi_i | h | \psi_j \rangle = 0 \quad , \quad i \neq j \quad . \qquad \text{(B4)}$$

APPENDIX C

From Eqs.(130), which expresses the energy in terms of the operators describing the variation, it is clear that the commutation of $\mathcal{W}$ and $\mathcal{Y}$, or equivalently of $\mathcal{W}$ and $\mathcal{V}$, Eq.(162), is a sufficient condition that the MCSCF process is invariant for such a variation. That it is also a necessary condition can be seen as follows.

Upon application of a general direct TSF transformation the variation of the energy is given by

$$\delta E = -2\Sigma_{\Lambda IJ \subset T}(W_{\Lambda I} - W_{\Lambda J})Y_{\Lambda JJ}H_{\Lambda JI} + \ldots \quad . \qquad \text{(C1)}$$

If the MCSCF process is invariant under such a direct TSF transformation, the corresponding non-essential parameters are designated by T_m, so that

$$\Sigma_{\Lambda IJ \subset T_m}(W_{\Lambda I} - W_{\Lambda J})Y_{\Lambda IJ}H_{\Lambda JI} = 0 \quad . \qquad \text{(C2)}$$

We now observe that invariance under a direct TSF transformation implies that Eq.(C2) must hold regardless of what we have chosen our SO's to be, provided of course that they form a proper orthonormal set.

Thus, except for a rare numerical accident, Eq.(C2) must hold for arbitrary choices of the matrix elements H_{AIJ}, $AIJ \in T_m$. Hence the variational parameters must satisfy

$$(W_{AI} - W_{AJ})Y_{AIJ} = 0 \quad , \quad AIJ \in T_m \quad , \tag{C3}$$

which is easily seen to be equivalent to

$$\left.\begin{aligned} [W, y] &= 0 \quad , \\ y &= -\Sigma_{AIJ \in T_m} Y_{AIJ}(\varphi_{AIJ} - \varphi_{AJI}) \quad . \end{aligned}\right\} \tag{C4}$$

which, in turn, is equivalent to Eq.(162).

REFERENCES

Banerjee, A., and F. Grein, 1976, Int. J. Quantum Chem. 10, 123.
Banerjee, A., and F. Grein, 1977, J. Chem. Phys. 66, 1054.
Chang, T. C., and W. H. E. Schwartz, 1977, Theor. Chim. Acta 44, 45.
Clementi, E., and A. Veillard, 1966, J. Chem. Phys. 44, 3050.
Dalgaard, E., 1979, Chem. Phys. Lett. 65, 559.
Dalgaard, E., and P. Jørgensen, 1978, J. Chem. Phys. 69, 3833.
Das, G., 1973, J. Chem. Phys. 58, 5104.
Das, G., T. Janis, and A. C. Wahl, 1974, J. Chem. Phys. 61, 1274.
Das, G., and A. C. Wahl, 1966, J. Chem. Phys. 44, 76.
Das, G., And A. C. Wahl, 1972, J. Chem. Phys. 56, 1769.
Detrich, J., and A. C. Wahl, 1980, MCHF Workshop Report, National Resource for Computation in Chemistry, Lawrence Berkeley Laboratory.
Docken, K. K., and J. Hinze, 1972, J. Chem. Phys. 57, 4928.
Froese Fischer, C., 1977, The Hartree-Fock Method for Atoms -- A Numerical Approach (Wiley, New York).
Grein, F., and A. Banerjee, 1975, Int. J. Quantum Chem. S9, 147.
Grein, F., and T. C. Chang, 1971, Chem. Phys. Lett. 12, 44.
Golebiewski, A., J. Hinze, and E. Yurtsever, 1979, J. Chem. Phys. 70, 1101.
Hartree, D. R., 1957, The Calculation of Atomic Structures (Wiley, New York).
Hinze, J., 1973, J. Chem. Phys. 59, 6424.
Hinze, J., and C. C. J. Roothaan, 1967, Progr. Theor. Phys. Suppl. 40, 37.
Kendrick, J., and I. H. Hillier, 1976, Chem. Phys. Lett. 41, 283.
Kuprievich, V. A., and O. V. Schramko, 1975, Int. J. Quantum Chem. 9, 1009.
Lengsfield, B. H., III, 1980, J. Chem. Phys. 73, 382.
Levy, B., 1969, Chem Phys. Lett. 4, 17.
Levy, B., 1970, Int. J. Quantum Chem. 6, 297.
Levy, B., 1973, Chem. Phys. Lett. 18, 59.
Levy, B., and G. Berthier, 1968, Int. J. Quantum Chem. 2, 307; errata ibid. 3, 247 (1969).
Löwdin, P.-O., 1977, Int. J. Quantum Chem. 12, Suppl. 1, 197.
Polezzo, S., 1975, Theor. Chim. Acta 38, 211.
Racah, G., 1942, Phys. Rev. 62, 438.
Racah, G., 1943, Phys. Rev. 63, 367.
Roos, B. O., P. R. Taylor, and P. E. M. Siegbahn, 1980, Chem. Phys. 48, 157.
Roothaan, C. C. J., 1951, Revs. Mod. Phys. 23, 69.
Roothaan, C. C. J., 1960, Revs. Mod. Phys. 32, 179.
Roothaan, C. C. J., and P. S. Bagus, 1963, in Methods in Computational Physics, Vol. II, B. Alder, S. Fernbach and M. Rotenberg, Eds. (Academic, New York).
Roothaan, C. C. J., J. Detrich, and D. G. Hopper, 1979, Int. J. Quantum Chem. S13, 93.
Ruedenberg, K., L. M. Cheung, S.T. Elbert, 1979, Int. J. Quantum Chem., 16, p1069.
Ruedenberg, K., and K. R. Sundberg, 1976, in Quantum Science -- Methods and Structure, J.-L. Calais, O. Goscinski, J. Linderberg, and Y. Ohrn, Eds. (Plenum, New York).
Ruttink, P. J. A., and J. H. van Lenthe, 1977, Theor. Chim. Acta 44, 97.

Veillard, A., 1966, Theor. Chim. Acta 4, 22.

von Neumann, J., 1955, Mathematical Foundations of Quantum Mechanics, R. T. Beyer, trans. (Princeton University, Princeton).

Wahl A. C., and G. Das, 1977, in Methods of Electronic Stucture Theory, H. F. Schaefer, III, Ed., Vol. 3 of Modern Theoretical Chemistry (Plenum, New York).

Werner, H.-J., and W. Meyer, 1980, J. Chem. Phys. 73, 2342.

Wigner, E. P., 1940, unpublished manuscript, later included in L. C. Biedenharn and H. van Dam, Quantum Theory of Angular Momentum (Academic, New York, 1965), pp. 87-133.

Wigner, E. P., 1941, Am. J. Math. 63, 57.

Yeager, D. L., and P. Jørgensen, 1979, J. Chem. Phys. 71, 755.

ON THE CONVERGENCE PROPERTIES OF THE DENSITY MATRIX DIRECTED GENERAL SECOND-ORDER MCSCF ALGORITHM

Byron H. Lengsfield, III*,†
Laser and Spectroscopy Branch
NASA Langley Research Center, Hampton, VA 23665

Charles W. Bauschlicher, Jr.†
Institute for Computer Applications in Science and Engineering
NASA Langley Research Center, Hampton, VA 23665

The density matrix directed (DMD) second order MCSCF algorithm[1] (see also ref. 2-6) is based on the fact that one can construct the Hessian and the gradient of the energy expression from the unique elements of one and two particle density matrices. With the recent development of the Unitary Group CI method,[7,8,9] density matrix elements can be obtained even for very large MCSCF problems without excessive computational effort.[10,11] Alternatively, one can obtain these density matrix elements by sorting a conventional CI formula tape,[1] and this technique has proven to be quite practical for traditional MCSCF problems.

The energy of a general MCSCF wavefunction

$$\Psi^I = \sum_{K}^{N} C_k^I \phi_k \tag{1}$$

can be expressed as follows

$$E^I = \sum_{i\geq j} D_{ij}^I (\phi_i \,|h|\, \phi_j) + \sum_{\substack{i\geq j \\ (ij)\geq(k\ell)}} \sum_{k\geq \ell} D_{ijk\ell}^I (\phi_i \phi_j \,|\, g \,|\, \phi_k \phi_\ell) \tag{2}$$

where

$$D_{ij}^I = \sum_{o\geq p}^{N} C_o^I C_p^I SF_{op}^{ij} \tag{3a}$$

$$D_{ijk\ell}^I = \sum_{o\geq p}^{N} C_o^I C_p^I SF_{op}^{ijk\ell} \tag{3b}$$

and $SF_{k\ell}^{ij}$ are structure factors. Variations in eqn. (2) are introduced by means of exponential unitary transformations[12] of the molecular orbitals (ϕ_i)

$$\tilde{U}_{mo} = e^{\tilde{\Delta}} \cong \tilde{1} + \tilde{\Delta} + \tfrac{1}{2}\tilde{\Delta}^2$$

$$\tilde{\Delta} = \begin{pmatrix} 0 & 0 & \Delta_{13} & \Delta_{14} \\ 0 & 0 & \Delta_{23} & \Delta_{24} \\ -\Delta_{13} & -\Delta_{23} & 0 & \Delta_{34} \\ -\Delta_{14} & -\Delta_{24} & -\Delta_{34} & 0 \end{pmatrix}$$

$$\tilde{\Delta} = \begin{pmatrix} \tilde{P}_1 \\ \tilde{P}_2 \\ \tilde{P}_3 \\ \tilde{P}_4 \end{pmatrix}$$

Full = 2

Part = 2

$$\Psi = c_1(f_1^2\ f_2^2\ p_1^2) + c_2(f_1^2\ f_2^2\ (p_1\ p_2))$$

$$\tilde{\Delta}^2 = \begin{pmatrix} (-\Delta_{13}^2 - \Delta_{14}^2) & (-\Delta_{13}\Delta_{23} - \Delta_{14}\Delta_{24}) & (-\Delta_{14}\Delta_{34}) & (\Delta_{13}\Delta_{34}) \\ (-\Delta_{13}\Delta_{23} - \Delta_{14}\Delta_{24}) & (-\Delta_{23}^2 - \Delta_{24}^2) & (-\Delta_{24}\Delta_{34}) & (\Delta_{23}\Delta_{34}) \\ \vdots & \cdot & \cdot & \cdot \end{pmatrix}$$

$$\tilde{\Delta}^2 = \begin{pmatrix} \tilde{Q}_1 \\ \tilde{Q}_2 \\ \tilde{Q}_3 \\ \tilde{Q}_4 \end{pmatrix} \tag{4}$$

and the CI vectors (C^I).

$$\tilde{U}_{CI} = e^{\tilde{\gamma}} \cong \tilde{1} + \tilde{\gamma} + \tfrac{1}{2}\tilde{\gamma}^2$$

$$\tilde{\gamma} = \begin{pmatrix} 0 & \gamma_{12} & 0 & 0 \\ -\gamma_{12} & 0 & \gamma_{23} & \gamma_{24} \\ 0 & -\gamma_{23} & 0 & 0 \\ 0 & -\gamma_{24} & 0 & 0 \end{pmatrix} \text{ — } P_2$$

$$\tilde{\gamma}^2 = \begin{pmatrix} -\gamma_{12}^2 & 0 & (\gamma_{12}\gamma_{23}) & (\gamma_{12}\gamma_{24}) \\ 0 & (-\gamma_{12}^2-\gamma_{23}^2-\gamma_{24}^2) & 0 & 0 \\ \gamma_{12}\gamma_{23} & 0 & -\gamma_{23}^2 & -\gamma_{23}\gamma_{24} \\ \gamma_{12}\gamma_{24} & 0 & -\gamma_{23}\gamma_{24} & -\gamma_{24}^2 \end{pmatrix} \text{ — } Q_2 \tag{5}$$

where $\tilde{P}_i$ and $\tilde{Q}_i$ are row generators which define the non-redundant orbital and CI mixings at first and second order respectively, (i.e., $\tilde{P}_i$

---

[*] a contribution by B. H. Lengsfield, III. NRC-NASA Research Associate, 1978-1980. Present address: IBM, 5600 Cottle Road, San Jose, CA 95193.

[†] Work was supported under NASA Contracts No. NAS1-15810 and NAS1-14472 while in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665.

generates the non-zero elements of the i^{th} row of $\tilde{\Delta}$ and $\tilde{Q}_I$ generates the non-zero elements of the i^{th} row of Δ^2). The energy expression is then expanded to second order in terms of the non-redundant variables of the generators of the two unitary transformations.

The contributions to gradient and Hessian can be obtained quite simply in terms of the $\tilde{P}$ and $\tilde{Q}$ row generators.

GRADIENT TERMS

ORBITAL MIXINGS

$$\tilde{P}_i \overset{I}{D}_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l) = \overset{I}{D}_{ijkl} \sum_{z}^{i} (\phi_z \phi_j |g| \phi_k \phi_l)\ \Delta_{iz}\ ph(i,z)$$

$$ph(i,z) = 1 \text{ if } z > i$$
$$= -1 \text{ if } z < i \qquad (6)$$

C.I. MIXINGS

$$\tilde{P}_I \overset{I}{D}_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l) = (\phi_i \phi_j |g| \phi_k \phi_l) \sum_{J}^{i} \gamma_{IJ}\ ph(I,J)$$

$$\cdot \sum_{o \geq p}^{N} (C_o^I C_p^J + C_o^J C_p^I)\ SF_{op}^{ijkl} \qquad (7)$$

HESSIAN TERMS

ORBITAL-ORBITAL INTERACTIONS

(1) $\tilde{P}_i \tilde{P}_j D_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l)$

$$= D_{ijkl} \sum_{y}^{i} \sum_{z}^{i} (\phi_z \phi_y |g| \phi_k \phi_l)\ \Delta_{iz}\ \Delta_{jy}\ ph(i,z)\ ph(j,y)$$

(2) $\tilde{Q}_i\ D_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l)$

$$= D_{ijkl} \sum_{q}^{n} (\phi_q \phi_j |g| \phi_k \phi_l) \sum_{m}^{m} \Delta_{im}\ \Delta_{qm}\ ph(i,m)\ ph(m,q)$$

$$\Delta_{im}\ \Delta_{jm} = -1$$

$$\Delta_{im}\ \Delta_{mq} = +1 \qquad (8)$$

C. I.-ORBITAL INTERACTIONS

$$\tilde{P}_I\ \tilde{P}_i\ \overset{I}{D}_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l)$$

$$= \left\{\tilde{P}_I\ \overset{I}{D}_{ijkl}\right\} \left\{\tilde{P}_i\ (\phi_i \phi_j |g| \phi_k \phi_l)\right\} \qquad (9)$$

$$\therefore\ \tilde{P}_{CI} \cdot \tilde{P}_{MO} = \left(\partial^2 E' / \partial\gamma\, \partial\Delta\right)$$

C. I. - C. I. INTERACTIONS

$$\tilde{P}_I \tilde{P}_I\ D_{ijkl} (\phi_i \phi_j |g| \phi_k \phi_l)$$

$$= (\phi_i \phi_j |g| \phi_k \phi_l) \sum_{o \geq p} SF_{o,p}^{ijkl} \sum_{J}^{N} \sum_{K}^{N} C_o^J C_p^K\ \gamma_{IJ} \gamma_{IK}\ ph(I,J)\ ph(I,K)$$

$$= \sum_{J}^{N} \sum_{K}^{N} \langle \tilde{C}^J | \tilde{H} | \tilde{C}^K \rangle\ \gamma_{IJ} \gamma_{IK}\ ph(I,J)\ ph(I,K) \qquad (10)$$

RECALL THERE IS ONLY ONE Q TYPE C. I. TERM

DIAGONAL ELEMENT OF THE C. I. - C. I. HESSIAN

$$\left\{\langle C^J | \tilde{H} | C^J \rangle - \langle C^I | \tilde{H} | C^I \rangle\right\} \gamma_{IJ}^2$$

$$\longrightarrow \quad \{E_J - E_I\} \qquad \text{if} \quad \tilde{H}\ \tilde{C}^K = E_K\ \tilde{C}^K \qquad (11)$$

The DMD-MCSCF algorithm is outlined in Figure 1.

MCSCF ALGORITHM

1. CONFIGURATION GENERATION
2. C. I. FORMULA GENERATION
3. D. M. SORT OF THE C. I. FORMULAS
4. HESSIAN FORMULA GENERATION
5. MCSCF ITERATIONS
   a. ORTHOGONALIZATION OF THE ORBITALS
   b. INTEGRAL TRANSFORMATION
   c. CONSTRUCTION AND DIAGONALIZATION OF THE HAMILTONIAN
   d. CONSTRUCTION OF THE HESSIAN AND THE GRADIENT, FOLLOWED BY THE SOLUTION OF THE LINEAR EQUATIONS
   e. APPROXIMATE UNITARY TRANSFORMATION OF THE ORBITALS

Figure 1.

As noted by Siegbahn, Heilberg, Roos, & Levy[5] (see ref. 2), the density matrix elements which involve only core orbitals need not be treated separately and various two electron density matrix elements in which two of the indices refer to core orbitals may also be grouped together. The Fock operator expressions employed by Siegbahn, et al. in the construction of the gradient may also be used to advantage in the construction of the CI-orbital portion of the Hessian. Furthermore, the density matrix elements whose indices refer only to core orbitals do not contribute to the CI-orbital portion of the Hessian.

Finally we note several advantages obtained by employing the eigenvectors of the Hamiltonian in this technique. First one need not include all of the CI vectors in the expansion of the energy to obtain a variational algorithm. This, of course, allows one to address much larger MCSCF problems. Second, the CI-CI portion of the Hessian and the CI terms in the gradient are diagonal and zero respectively. This allows the equations for the elements of the unitary generators to be simplified. Finally, redundant variables are not always easily identified in large MCSCF calculations. However, when the root of the Hamiltonian which is being optimized is an eigenvector of the Hamiltonian, redundant variables generally give rise to zeros in the gradient and can be easily detected.

There are perhaps two points of primary interest in evaluating the convergence properties of this alogrithm. The first point being the behavior of this procedure when poor starting orbitals are employed. The second point is concerned with the number of CI vectors which must be included in large problems to obtain a reasonable rate of convergence.

In a large number of MCSCF problems one does not possess a very good choice of starting orbitals. This is especially true if the MCSCF wavefunction contains several configurations which differ from one another by (spin orbital) single excitations. In this case the Hessian may posses very small and even negative eigenvalues. Various means have been proposed to shift the eigenvalues (or alternatively the diagonal elements)[13] of the Hessian or to take a step in the direction indicated by the eigenvector which corresponds to the negative eigenvalue of the Hessian.[2] These techniques often yield disappointing results especially when applied to problems where the MCSCF reference contains CSF's which differ by a single excitation.[1] Instead, a super CI technique has been found to possess a much larger radius of convergence[1] as demonstrated in Table I. In this method one constructs an augmented Hessian matrix [1,14,15] in analogy with the Singles Hamiltonian constructed in Generalized Brillouin Theorem algorithms. It is important to note that a density matrix oriented approach allows this matrix to be constructed in a much more efficient manner than the traditional contraction type procedures.[16-22] This method is particularly attractive as quadratic convergence[1] is very often obtained (when all CI coupling terms are included) in this procedure when a reasonable set of orbitals has been obtained. Moreover, a simple scheme can be devised to correct the eigenvector of this matrix, when it is not dominated by the MCSCF reference and further increase the radius of convergence attained by this technique.

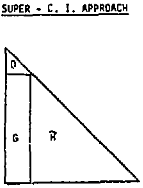


Figure 2.

In the following tables the results of several MCSCF calculations are presented in which the number CI vectors included in the Hessian has been varied.

TABLE I. BeO[a,b] 3 CSF SUPER CI ALGORITHM

Iteration	Energy (a.u.)	ΔE	Δ[c]
1[d]	-89.424317	-	2.56 E-1
2	-89.466285	4.20 E-2	2.29 E-1
3	-89.495331	2.90 E-2	1.13 E-1
4	-89.503767	8.44 E-3	6.62 E-2
5	-89.505765	2.00 E-3	1.53 E-2
6	-89.506026	2.60 E-4	5.07 E-4
7	-89.506033	7.99 E-6	6.57 E-7
8	-89.506034	1.13 E-8	1.42 E-12

a) core $4\sigma^2\pi^4$, core $4\sigma5\sigma\pi^4$ and core $4\sigma^2\pi^3\pi$

b) Bauschlicher-Yarkony Basis, J. Chem. Phys. 72, 1138 (1980)

c) $\Delta \equiv \Sigma \Delta_{ij}^2$, see eqn. (4)

d) Damping performed, the Hessian possessed two negative eigenvalues. SCF starting guess employed.

TABLE II. APPROXIMATE SUPER CI[a]

Iteration	Energy (a.u)	ΔE	Δ
1	-89.424317	-	.162
2	-89.462099	3.78 E-2	.137
3	-89.487145	2.50 E-2	8.36 E-2
4	-89.497525	1.04 E-2	2.93 E-2
5	-89.501648	4.12 E-3	1.40 E-2
6	-89.503637	1.99 E-3	8.40 E-3
7	-89.504713	1.08 E-3	5.00 E-3
8	-89.505311	5.98 E-4	2.82 E-3
9	-89.505642	3.31 E-4	1.54 E-3
10	-89.505824	1.82 E-4	8.19 E-4

a) The second order CI terms were not included in this calculation.

The importance of including or excluding a particular CI vector can be placed on a more quantitative basis by means of the simple perturbation arguments presented below.

Consider the Newton-Raphson linear equations induced by a two CSF, two orbital problem,

$$\begin{pmatrix} B & C \\ C & A \end{pmatrix} \begin{pmatrix} \Delta \\ Y \end{pmatrix} = \begin{pmatrix} g \\ 0 \end{pmatrix} \tag{12}$$

where B, C and A represent the orb.-orb., CI-orb., and CI-CI portions of the Hessian respectively. Δ and γ are the unique elements of the generators of the unitary transformations and g is the orbital gradient (the CI gradient is zero by virtue of the fact that the secular is assumed to have been solved on the preceding iteration). We then have,[2]

$$\left(B - \frac{C^2}{A}\right)\Delta = g \tag{13}$$

$$\begin{aligned}\Delta &= \left(\frac{g}{B - \frac{C^2}{A}}\right) \\ &= \frac{g}{B}\left(\frac{1}{1-\frac{C^2}{AB}}\right) \\ &\approx \frac{g}{B}\left(1+\frac{C^2}{AB}+\cdots\right) .\end{aligned} \tag{14}$$

(*Recall* $A \alpha E^2 - E^1$ *where* 1 *is the root being* optimized.) This perturbation may be generalized to account for the interaction between a particular CI vector and all of the orbital mixings. For the purposes of this study it suffices to consider these interactions a sum of decoupled two by two problems and monitor the largest perturbation term (C^2/AB) associated with the last CI vector to be included in the problem. These terms are also presented in Tables III and IV. It is interesting to note that this perturbation term differs from the second-order perturbation expression obtained by Das[21] in that the energy difference appearing in the denominator is weighted a *diagonal element from the orbital section of the* Hessian.

The results of a number of MCSCF calculations are summarized in the following tables.

TABLE IIIa. 3 CSF BeO CALCULATION WITH ONE CI ROOT EXCLUDED

	E	ΔE	Δ
1	-89.424317	-	2.61 E-1[a]
2	-89.465898	4.16 E-3	2.28 E-1
3	-89.494027	2.81 E-2	1.03 E-1
4	-89.502397	8.37 E-3	2.82 E-2
5	-89.504267	1.87 E-3	8.63 E-3
6	-89.505109	8.41 E-4	3.71 E-3
7	-89.505550	4.41 E-4	1.86 E-3
8	-89.505782	2.32 E-4	9.58 E-4
9	-89.505903	1.21 E-4	4.92 E-4
10	-89.505967	6.3 E-5	2.51 E-4

[a] Damping employed this iteration

TABLE IIIb. BeO 3 CSF MCSCF

Root	Largest Perturbation Term	Sum of Perturbation Contributions
1	.18	.72
2	.31	1.48

TABLE IV. FULL VALENCE $^1\Sigma^+$ STATE OF MgO[a,b]

Number of CI Roots	0		30		90	
Iteration	ΔE	Δ	ΔE	Δ	ΔE	Δ
1	-	9.E-5	-	1.E-4	-	2.E-4
2	-3.E-6	6.E-6	-4.E-6	1.E-5	-5.E-6	3.E-6
3	-9.E-7	2.E-6	-5.E-7	2.E-6	-8.E-7	1.E-7
Final Energy	-274.514267		-274.514268		-274.514268	

[a] Yoshimine-McLean molecular DZP Slater basis with a diffuse 3s-function (.855) on Mg. 142 CSF's in C_{2v}

[b] The maximum perturbation term (.43) was obtained from the 22nd eigenvector of the Hamiltonian.

The perturbation contributions of the higher eigenvectors of the Hamiltonian are often larger than the perturbation estimate of the contribution of many of the lower roots of the Hamiltonian. While the inclusion of a few of the CI vectors in the varitional problem can dramatically affect the convergence characteristic of this MCSCF algorithm far from convergence. One can not expect any substantial advantage from this procedure near convergence for a general (containing single excitations) MCSCF wavefunction.

The DMD-MCSCF algorithm provides a simple and efficient means for constructing the Hessian and the gradient of a general MCSCF energy expression. The studies reported in the paper and recent work on MgO and BeO[22] indicates that this method is capable of providing the convergence behavior needed to perform practical quantum chemical calculations.

References

[1] B. H. Lengsfield, III, J. Chem. Phys. xx, xxx (1980).

[2] L. G. Yaffe and W. A. Goddard, III, Phys. Rev. A 13, 1682 (1976).

[3] E. Daalgard and P. Jorgensen, J. Chem. Phys. 69, 3833 (1979).

[4] C. C. J. Roothaan, J. Detrich and D. G. Hopper, Int. J. Quantum. Chem. Symp. 13, 93 (1979).

[5] P. Siegbahn, A. Heilberg, B. Roos and B. Levy, Physica Scripta 21, 323 (1980).

[6] P. Jorgensen and D. Yeager, J. Chem. Phys. 71, 757 (1979).

[7] J. Paldus and M. J. Boyle, Physica Scripta, 21 295 (1980), and references therein.

[8] I. Shavitt, Int. J. Quant. Chem. Symp. 11, 131 (1977), 12, 5 (1978).

[9] B. R. Brooks and H. F. Schaefer, III, J. Chem. Phys. 70, 5092 (1979)

[10] B. R. Brooks, W. D. Laidig, P. Saxe, N. C. Handy, and H. F. Schaefer, III, Physica Scripta 21, 312 (1980).

[11] B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi and H. F. Schaefer, III, J. Chem. Phys. 72, 4652 (1980).

[12] K. M. Sando and J. E. Harriman, J. Chem. Phys. 47, 180 (1967).

[13] V. A. Kuprevich and V. E. Klymeno, Int. J. Quantum Chem. 12, 327 (1977).

[14] T. C. Chang and W. H. E. Schwartz, Theor. Chem. Acta 44, 45 (1977).

[15] F. W. Bobrowicz and W. A. Goddard, III, pp. 79-127, Vol. 3, *Modern Theoretical Chemistry*, ed. H. F. Schaefer, III (Plenum, New York, 1979).

[16] B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968).

[17] F. Grein and A. Banarjee, Int. J. Quantum Chem. Symp. 9, 147 (1975).

[18] F. Grein and T. C. Chang, Chem. Phys. Lett. 12, 44 (1971).

[19] P. J. A. Ruttink and J. H. van Lenthe, Theor. Chem. Acta 44, 97 (1977).

[20] L. M. Cheung, T. S. Elbert, and K. Ruedenberg, Int. J. Quantum Chem. 16, 1069 (1979).

[21] G. Das, J. Chem. Phys. 58, 5104 (1973).

[22] C. W. Bauschlicher, Jr., B. H. Lengsfield, III, D. M. Silver, and D. R. Yarkony, to be published.

RECENT DEVELOPMENTS IN MULTICONFIGURATION WAVEFUNCTION OPTIMIZATION

Ron Shepard
Battelle Columbus Labs
Columbus, Ohio

INTRODUCTION

A formal and numerical comparison of two of the most promising types of multiconfiguration wavefunction optimization methods, the super-CI methods and the exponential operator methods, is performed. The super-CI methods display superior convergence properties when the initial wavefunction is far from correct while the exponential operator methods possess superior convergence characteristics when the initial wavefunction is close to correct. The formal comparison of these approaches suggests hybrid methods which have some of the advantages of both methods. Unitary group methods have been successfully employed in the implementation of a general and efficient multiconfiguration wavefunction optimization program using these exponential operator and hybrid methods and a discussion of this implementation is included.

SUPER-CI METHODS

In the super-CI wavefunction optimization method developed by Grein[1], the matrix representation of the hamiltonian operator is constructed within the single excitation basis $\{|mc\rangle, |ij\rangle\}$ where $|ij\rangle \equiv (i^+j - j^+i)|mc\rangle$.
The super-CI wavefunction is determined from the solution of the secular equation:
$\underline{H}(sci)\underline{Y}=\underline{S}(sci)\underline{Y}E(sci)$ where $|SCI\rangle = Y(mc)|mc\rangle + \sum_{i>j} Y(ij)|ij\rangle$

If the orbital change $\underline{\phi}' = \underline{\phi}(\underline{1}+\underline{X}) \approx \underline{\phi}(e^{\underline{X}})$ where $X_{ij}=-X_{ji}=Y_{(ij)}$ is performed, then the MC energy of the next iteration may be compared with the current SCI energy:

$$E(sci) = E(mc) + 2\sum_{i>j} X_{ij}\langle mc|H|ij\rangle + \sum_{\substack{i>j\\k>l}} X_{ij}X_{kl}\langle ij|H|kl\rangle$$

$$E'(mc) = E(mc) + 2\sum_{i>j} X_{ij}\langle mc|H|ij\rangle + \sum_{\substack{i>j\\k>l}} X_{ij}X_{kl}\langle ij|H|kl\rangle$$
$$+ \sum_{ijkl} X_{ij}X_{kl}\langle mc|H(ij)(kl)|mc\rangle + O(X^3)$$

It is apparent from the comparison of these two expressions that minimization of E(sci) via the SCI secular problem is equivalent to minimization of the MC energy of the next iteration except for the last second order term and the higher order terms in the E(mc)' expression. This expression for E(mc)' also assumes that any coupling of the mixing coefficient change and the orbital expansion coefficient change may be neglected in the calculation of the optimimum X parameters. A serious limitation of this method is that the individual matrix elements of the $\underline{H}$(sci) matrix are difficult to calculate since they require either summations over all configurations within the MC space or the calculation of some 3- and 4-particle density matrix elements. The advantage of the SCI method is that the X parameters are calculated to lower the energy each iteration and not simply to extremize it. Thus at convergence, not only is the generlized Brillouin theorem (GBT) satisfied indicating an extremum solution, but a change of the orbitals will increase the energy. The first attempt to determine the importance of the neglected second order terms was performed by Grein and Banerjee[1]. The inclusion of coupling between the mixing coefficients and orbital expansion coefficients in the SCI equations was first performed by Chang and Schwarz[2]. They included this coupling by appending the excited states $\{|n\rangle\}$ where $\langle n|H|n'\rangle = E_n\delta_{nn'}$ to the single excitation basis and solving the resulting SCI secular problem.

EXPONENTIAL OPERATOR METHODS

The exponential operator method of Dalgaard and Jørgensen[3] is based on expressing a trial MC wavefunction in terms of a unitary operator
$|mc'\rangle = e^{i\hat{\Lambda}}|mc\rangle$ where $\hat{\Lambda} = \sum_{r>s} K_{rs}\, i(r^+s - s^+r) = \sum_{rs} K_{rs}\hat{T}_{rs}$
The orbitals used to construct the MC' are obtained as
$$\underline{\Phi}' = \underline{\Phi}(e^{-\underline{K}}) \quad \text{where} \quad K_{rs} = -K_{sr} = K_{(rs)}$$
and where the K parameters are to be determined. The resulting commutator expansion of the energy expression is truncated to include second order terms. Stabilization of this expression leads to the linear equation $\underline{B}\underline{K}+\underline{W}=\underline{0}$ where
$$B_{pq,rs} = \tfrac{1}{2}\langle [[H, i\hat{T}_{pq}], i\hat{T}_{rs}] + [[H, i\hat{T}_{rs}], i\hat{T}_{pq}]\rangle$$
is an element of the orbital Hessian matrix and
$W_{rs} = \langle [H, i\hat{T}_{rs}]\rangle$ is a GBT term and a component of the gradient. Explicit coupling of the orbital change and mixing coefficient change may be included using the exponential operator notation of Yeager and Jørgensen[4] with the trial MC wavefunction expressed as
$$|mc'\rangle = e^{i\hat{\Lambda}}e^{i\hat{S}}|mc\rangle \quad \text{where} \quad \hat{S} = \sum_{n\neq mc} s_n i(|n\rangle\langle mc| - |mc\rangle\langle n|) = \sum_n{}' s_n\hat{P}_n$$
Stabilization of the resulting second order energy expression leads to the linear equations involving the full wavefunction Hessian matrix and gradient vector $\begin{pmatrix}\underline{B} & \underline{C}\\ \underline{C}^t & \underline{M}\end{pmatrix}\begin{pmatrix}\underline{K}\\ \underline{s}\end{pmatrix} + \begin{pmatrix}\underline{W}\\ \underline{W}\end{pmatrix} = \begin{pmatrix}\underline{0}\\ \underline{0}\end{pmatrix}$ which may be expressed as the partitioned Hessian expression
$$(\underline{B} - \underline{C}\,\underline{M}^{-1}\underline{C}^t)\underline{K} + \underline{W} = \underline{0} \quad \text{when} \quad \langle n|H|n'\rangle = \delta_{nn'}E_n$$
This partitioned form of the orbital Hessian emphasizes the relationship between the Rayleigh Schroedinger procedure of Das[5] and of the corrected Hessian of Dalgaard and Jørgensen[3] to the exponential operator methods.

The advantage of these operator based methods are that only 1- and 2-particle density matrix contributions are required. This is due to the fact that all the terms of the Hessian and gradient are expressed directly as commutators instead of simple products. All second order terms are included so that second order convergence in the GBT terms and the energy is observed. However these Newton Raphson methods converge not only to energy minima but also to saddle points and other types of solutions for which the GBT is satisfied but where the Hessian matrix is not positive definite. These solutions do not necessarily correspond to excited states of the molecular systems although excited state solutions do necessarily possess a negative eigenvalue of the wavefunction Hessian for each lower state. The desired excited state solutions correspond to those for which their variational energy is minimized. These solutions are characterized by positive definite partitioned orbital Hessian matrices.

DISCUSSION OF COMPARISON

Numerical comparisons[10] of the convergence properties of the SCI method and the exponential operator methods, both with and without the perturbation corrections, reveal that the SCI method possesses superior convergence properties when the initial wavefunction guess is far from correct. This is a result of the energy minimization approach of the SCI iterative procedure compared to the energy stabilization approach of the exponential operator methods. The local convergence properties of the SCI methods are similar to those of the exponential operator methods when the perturbation corrections are neglected. The local convergence properties of the exponential methods are improved when the perturbation corrections are included and second order convergence is observed. The matrix construction step of the exponential methods is more efficient than the matrix construction step of the SCI methods because the former requires only the non-zero contributions of the 1- and 2-particle reduced density and transition density matrices while the latter requires effort equivalent to the construction of a very large Hamiltonian matrix and subsequent contraction to form the $\underline{H}$(sci) matrix.

HYBRID METHODS

A more detailed investigation of the exponential operator methods and the SCI methods reveals the relationship between these solutions and suggests methods which have the desirable properties of both methods. The energy expression of the $e^{i\hat{\lambda}}|mc\rangle$ trial function may be written as

$$\Delta E = \tfrac{1}{2}(\underline{\kappa}^{\dagger}, 1)\begin{pmatrix}\underline{B} & \underline{W}\\ \underline{W}^{\dagger} & 0\end{pmatrix}\begin{pmatrix}\underline{\kappa}\\ 1\end{pmatrix}$$

and requiring $\partial\Delta E/\partial\kappa = 0$ results in the usual linear equation $\underline{B}\underline{\kappa} + \underline{W} = \underline{0}$. If an error of order $|\underline{W}|^2$ is introduced into this equation however

$$\Delta E = \tfrac{1}{2}(\tilde{\underline{\kappa}}^{\dagger}, 1)\begin{pmatrix}\underline{B} & \underline{W}\\ \underline{W}^{\dagger} & 0\end{pmatrix}\begin{pmatrix}\tilde{\underline{\kappa}}\\ 1\end{pmatrix}\left(1 + |\tilde{\underline{\kappa}}|^2\right)^{-1}$$

then requiring $\partial\Delta E/\partial\kappa = 0$ leads to the eigenvalue problem

$$\begin{pmatrix}\underline{B} & \underline{W}\\ \underline{W}^{\dagger} & 0\end{pmatrix}\begin{pmatrix}\tilde{\underline{\kappa}}\\ 1\end{pmatrix} = 2\Delta E\begin{pmatrix}\tilde{\underline{\kappa}}\\ 1\end{pmatrix}$$

which is an approximation to the SCI eigenvalue problem. The $\tilde{\underline{\kappa}}$ parameters may be expressed as the solution of the linear equation

$$(\underline{B} - 2\Delta E)\tilde{\underline{\kappa}} + \underline{W} = \underline{0}$$

where $2\Delta E$ is the lowest eigenvalue of the approximate SCI problem. An identical analysis of the $e^{i\hat{\lambda}}e^{i\hat{S}}|mc\rangle$ exponential method results in the partitioned equation

$$\left(\underline{B} - 2\Delta E - \underline{C}(\underline{M} - 2\Delta E)^{-1}\underline{C}^{\dagger}\right)\tilde{\underline{\kappa}} + \underline{W} = \underline{0}$$

where $2\Delta E$ is the eigenvalue of the Chang and Schwarz SCI problem. Since the addition of a constant to the diagonal terms of the Hessian matrix shifts all the eigenvalues by that same constant, it is clear that the SCI solution is a special case of the mode damping or level shifting techniques of Yeager and Jørgensen and of Dalgaard and Jørgensen[3]. The expansion of $\tilde{\underline{\kappa}}$ in terms of the full second order $\underline{\kappa}$

$$\tilde{\underline{\kappa}} = \underline{\kappa} + 2\Delta E\,\underline{B}^{-1}\underline{\kappa} + O(\Delta E^2)$$

shows the ΔE relationship between these two solutions. The improved convergence that is made available in the approximate SCI solution results from the fact that there are many solutions to the secular equation. The one that is chosen is the one that lowers the energy. This flexibility is obtained however only with the sacrifice of the local second order convergence properties. These hybrid methods attempt to exploit the energy minimization characteristics of the SCI methods and the local convergence properties of the Newton Raphson methods by switching to the full second order procedures only when convergence can be guaranteed. One method consists of using

$$\begin{pmatrix}\underline{B} & \underline{W}\\ \underline{W}^{\dagger} & 0\end{pmatrix}\begin{pmatrix}\tilde{\underline{\kappa}}\\ 1\end{pmatrix} = 2\Delta E\begin{pmatrix}\tilde{\underline{\kappa}}\\ 1\end{pmatrix}$$

until $|\tilde{\underline{\kappa}}|$ is small, then using $(\underline{B} - \underline{C}\underline{M}^{-1}\underline{C}^{\dagger})\underline{\kappa} + \underline{W} = \underline{0}$ until convergence is reached. Another approach is to always solve

$$(\underline{B} - \lambda)\underline{\kappa} + \underline{W} = 0 \quad \text{or} \quad \left(\underline{B} - \lambda - \underline{C}(\underline{M} - \lambda)^{-1}\underline{C}^{\dagger}\right)\underline{\kappa} + \underline{W} = \underline{0}$$

treating λ as an emperical parameter to switch from the SCI solution ($\lambda \approx 2\Delta E$) which converges to an energy minima to the full second order solution ($\lambda = 0$) which converges to the closest extremum solution. Our numerical experience[10] shows that the energy may usually be obtained from the approximate SCI solution without switching to the second order solution. *However the full second order convergence* is sometimes useful for energy convergence and is almost always necessary for wavefunction convergence (or equivalently for problems where $|\underline{W}|$ is required to be small). These observations are consistant with other researchers (Banerjee and Grein, Chang and Schwarz, Lengsfield) who have reported on the numerical convergence properties of the SCI methods.

MATRIX ELEMENT EVALUATION

If the $\hat{\Lambda} = \sum X_{rs}\hat{T}_{rs}$ operator is expressed in terms of the generators of the unitary group, $\hat{T}_{rs} = i(e_{rs} - e_{sr})$, then the matrix elements of the $\underline{B}$ and $\underline{C}$ matrices and the $\underline{W}$ vector may be expressed directly in terms of the reduced spatial 1- and 2-particle density matrix elements, $\langle e_{ij}\rangle$ and $\langle e_{ij,kl}\rangle$ and the transition density matrices $\langle mc|e_{ij}|n\rangle$ and $\langle mc|e_{ij,kl}|n\rangle$ along with the molecular 1- and 2-electron integrals h_{ij} and $(ij|kl)$.

$$\begin{aligned}
B_{pq,rs} = {} & 2h_{qs}\langle e_{pr}\rangle + 2h_{pr}\langle e_{qs}\rangle - 2h_{qr}\langle e_{ps}\rangle - 2h_{ps}\langle e_{qr}\rangle \\
& + (H_{sp} + H_{ps})\delta_{qr} + (H_{rq} + H_{qr})\delta_{ps} \\
& - (H_{rp} + H_{pr})\delta_{qs} - (H_{qs} + H_{sq})\delta_{pr} \\
& - 2\sum_{ij}(ij|qr)\langle e_{ps,ij}\rangle - 2\sum_{ij}(qj|ir)\langle e_{pj,is}\rangle \\
& + 2\sum_{ij}(qj|is)\langle e_{pj,ir}\rangle + 2\sum_{ij}(ij|qs)\langle e_{pr,ij}\rangle \\
& + 2\sum_{ij}(pj|ir)\langle e_{qj,is}\rangle + 2\sum_{ij}(ij|pr)\langle e_{qs,ij}\rangle
\end{aligned}$$

$$-2\sum_{ij}(ij|ps)\langle e_{pr,ij}\rangle \; -2\sum_{ij}(pj|is)\langle e_{qj,ir}\rangle$$

$$+2\sum_{ij}(ip|jr)\langle e_{iq,js}\rangle \; -2\sum_{ij}(is|jp)\langle e_{ir,jq}\rangle$$

$$-2\sum_{ij}(ir|jq)\langle e_{is,jp}\rangle \; +2\sum_{ij}(iq|js)\langle e_{ip,jr}\rangle$$

$$-(Ysq+Yqs)\delta pr \quad -(Yrp+Ypr)\delta sq$$

$$+(Ysp+Yps)\delta rq \quad +(Yrq+Yqr)\delta sp$$

$$W_{rs} = 2Hrs - 2Hsr - 2Ysr + 2Yrs$$

In these expressions H and Y are the partial sums:

$$Hrs = \sum_{i} h_{is}\langle e_{ri}\rangle$$

$$Yrs = \sum_{ijk}(si|jk)\langle e_{ri,jk}\rangle$$

The $\underline{C}$ matrix involves transition density matrix elements:

$$\begin{aligned} C_{(rs),n} &= -2\langle mc|[H, iT_{rs}]|n\rangle \\ &= 2\sum_{i} h_{ir}\langle mc|e_{is}+e_{si}|n\rangle \; -2\sum_{i} h_{is}\langle mc|e_{ri}+e_{ir}|n\rangle \\ &\quad +2\sum_{ijk}(ir|jk)\langle mc|e_{is,jk} + e_{kj,si}|n\rangle \\ &\quad -2\sum_{ijk}(is|jk)\langle mc|e_{ri,jk} + e_{kj,ir}|n\rangle \end{aligned}$$

And finally, if a CI calculation is performed at the beginning of each iteration, the $\underline{M}$ matrix takes the simple form:

$$M_{nn'} = 2(E(n) - E(mc))\,\delta n,n'$$

In the implementation of a general wavefunction optimization procedure, no restrictions should be placed on the density matrix elements appearing in these expressions. In particular, no assumptions such as "full valence CI", "full active space CI", or "doubly occupied CI" should be made. In order to use the simplifying features of certain reference spaces when they are employed however, it is convenient to use the non-zero density matrix contributions to control the construction of the required matrices. It should be mentioned that only T operators which may result in non-zero gradient contributions need to be included in the optimization procedure. *For example, indices p and q which* correspond to molecular orbitals of different symmetries produce zero gradient contributions $\langle[H, iT_{pq}]\rangle$ because the electronic Hamiltonian operator is completely symmetric. The formalism used by Rudenburg in the SCF method[7] and by Roothaan in the exponential method[8] emphasizes this aspect. For a given MC space, other orbital pairs sometimes result in trivial zero contributions to the gradient and to zero eigenvalues of the wavefunction Hessian matrix. These redundant variables may be removed to reduce the dimension of the matrices without affecting either the convergence properties or the final results.

Although the 1- and 2-particle density matrices and transition density matrices are usually sparse, we have chosen not to attempt to explicitly construct them each iteration. Instead, we calculate a <u>contribution</u> to a density matrix element or transition density matrix element, use that contribution with any available integrals, and update the appropriate Hessian or gradient matrix elements. When all non-zero density contributions have been exhausted, new integrals are made available, the density contributions are again retrieved, and the process is repeated. This procedure allows the construction of the Hessian and gradient matrix elements to be performed with only one reading of the integral list. Because the unitary group procedures we have used result in very efficient evaluation and very compact storage of the density contributions, the multiple readings of the resulting short formula tape are only a small part of the wavefunction optimization procedure. Since the total number of formula tape readings required each time the matrices are constructed is equal to the number of blocks of integrals, an attempt is made to minimize the number of blocks of integrals with dynamic memory allocation using subroutine argument lists and array offsets. In the case where all the integrals may be made available at one time, only one reading of the formula tape is required.

In the unitary group approach, a density contribution is determined by a loop within a Shavitt graph[9] Each loop is associated with many configuration pairs. The number of configuration pairs is given as the product of the number of upper walks from the loop head and the number of lower walks from the loop tail. The total contribution of a loop to a density matrix element is the sum of all the products of pairs of mixing coefficients which share the loop. In the most favorable cases, there is only one loop which contributes to a density matrix element. Although this loop may be shared by many configuration pairs, the complete contribution may easily be precomputed before using this density contribution in the $\underline{B}$ and $\underline{W}$ construction. Account must also be taken of the operator relations $e_{ij}=e^{+}_{ji}$ and $e_{ij,kl}=e_{kl,ij}=e^{+}_{ji,lk}=e^{+}_{lk,ji}$. Since only a unique set of loops are constructed and stored on the formula tape, the appropriate permutations of the indices must be performed to include all possible non-zero density matrix contributions. The $\underline{C}$ matrix construction may be broken into two parts

$$\begin{aligned} C_{pq,n} &= 2\langle mc|[H, i\hat{T}_{pq}]|n\rangle \\ &= -2\sum_{a}\Big(\sum_{b} U_{b,mc}\langle b|[H, i\hat{T}_{pq}]|a\rangle\Big)U_{a,n} \\ &= -2\sum_{a} C'_{pq,a}U_{a,n} \end{aligned}$$

where the construction of the $\underline{C}'$ matrix requires only the current mixing coefficients for the MC state. Construction of the intermediate $\underline{C}'$ matrix also avoids multiple references to excited state mixing coefficients for each transition density matrix contribution and replaces it by the indicated matrix product which is independent of the formula tape and integral list lengths. Since the $\underline{C}$ matrix construction requires the transition 1- and 2-particle density matrices, or equivalently the density contributions from the primitive configurations, its construction becomes the dominating matrix construction step as the number of configurations in the MC space is increased.

In our implementation of the unitary group approach, the configuration list is specified by a distinct row table (DRT). The DRT is constructed automatically by a small interactive program which requires only a limited amount of spatial symmetry and orbital occupation restriction information. The formula tape which corresponds to this DRT is then constructed and used by either the CI program or the wavefunction optimization program.

SUMMARY

We have found that the unitary group approach used in our wavefunction optimization programs results in not only efficient matrix element construction but also general and flexible configuration specification for many molecular systems. Wavefunction optimization procedures using second order methods and hybrid methods consisting of approximate SCI methods and second order methods have been implemented. Wavefunctions corresponding to both ground and excited states of the same symmetry have been obtained and reported using these methods.

REFERENCES

1. A. Banerjee and F. Grein, Int. J. Quantum Chem., 10 123(1976)
   A. Banerjee and F. Grein, J. Chem. Phys. 66, 1050 (1977)
   A. Banerjee and F. Grein, Int. J. Quantum Chem. Symp. 9, 147 (1975)

2. T. Chang and W Schwarz, Theoret. Chim. Acta 44, 45 (1977)

3. E. Dalgaard and P. Jørgensen, J. Chem. Phys. 69 (8), 3833 (1978)

4. D. Yeager and P. Jørgensen, J. Chem. Phys. 71 (2), 755 (1979)

5. G. Das, J. Chem. Phys. 58, 5104 (1973)

6. B. Lengsfield, J. Chem. Phys. 0, 0 (1980)

7. L. Cheung, S. Elbert, and K. Ruendenberg, Int. J. Quantum Chem., 16, 1069 (1979)

8. C. Roothan, J. Detrich, and D. Hopper, Int. J. Quantum Chem. Symp. 13, 93 (1979)

9. I. Shavitt, Int. J. Quantum Chem. Symp. 12, 5 (1978)
   I. Shavitt, Annual Report to NASA Ames Research Center, June 1979

10. R. Shepard and J. Simons, J. Chem. Phys. 0, 0 (0)
    R. Shepard and J. Simons, Int. J. Quantum Chem. 0,0 (1980)

FORMAL AND NUMERICAL CONVERGENCE COMPARISON OF SCI AND EXPONENTIAL METHODS

Dr. Ronnie Shepard
Battelle Memorial Institute
Columbus, Ohio 43201

and

Prof. Jack Simons
University of Utah
Salt Lake City, Utah 84112

FORMAL COMPARISON

By examining the forms used to describe modifications in the orbital and configuration expansion coefficients, the energy functions which are made stationary, the resultant equations to be solved, and the iterative procedures which are commonly employed in solving such equations, we see the similarities and differences among the super CI(SCI) and exponential MCSCF methods.

The SCI method can be viewed as using a trial variational wavefunction written as a linear combination of the current (in the iterative scheme) MC wavefunction $X_o|0\rangle$ plus single excitations $X_{ij}(i^+j - j^+i)|0\rangle \equiv |ij\rangle$. The SCI secular equation resulting from making the energy of this trial wavefunction stationary yields an eigenvector which is then used to generate modified MCSCF orbitals. The modification of the CI expansion coefficients $|0\rangle = \sum_g C_{go} |\phi_g\rangle$ is then achieved by performing a CI calculation using configurations $|\phi_g\rangle$ involving the "new" MCSCF orbitals. The fact that both orbital and CI expansion coefficient optimization are achieved through solutions of secular problems (which have the well known eigenvalue separation properties) makes the SCI method look promising for use on excited states. Critics would remark that the SCI does not treat coupling between orbital and CI coefficient variation in a balanced or fully coupled manner and that the SCI energy expression ignores certain terms $X_{ij} X_{k\ell}\langle 0|H(i^+j-j^+i)(k^+\ell-\ell^+k)|0\rangle$ which are second order in the orbital variation parameters X_{ij}. This latter fact makes the SCI method not quadratically convergent.

The exponential MCSCF method expresses the modification of the orbitals ($\exp(i\lambda)$) and CI coefficients ($\exp(iS)$) in terms of exponential unitary transformations. By expanding the MC energy through second order in λ and S (thereby neglecting third and higher order factors) and making the resultant expression stationary with respect to variations in λ and S, one arrives at a set of linear algebraic equations for the parameters in λ and S. This method is quadratically convergent since its energy expression contains all terms in λ and S through second order. However, it is not variational since the "energy function" which was made stationary is not an upper bound to the true ground-state energy. The coupling between orbital and configuration mixing coefficient optimization is treated in a balanced or coupled manner in this exponential approach. However, the fundamental assumption that second order in λ is equivalent (in an energy sense) to second order in S remains to be tested.

Given the basic working equations of the SCI and exponential MCSCF methods, one next must attempt to find stable iterative procedures for solving these equations. In solving the SCI equations, one is faced with choosing the eigenvectors of the SCI and CI secular equations which are proper for the state of interest and with damping the modifications of the orbitals resulting from the SCI secular problem. One is also faced with orthonormalizing the MCSCF orbitals on each iteration because the transformation matrix resulting from the SCI secular equation is not unitary.

Several procedures have been suggested for solving the linear equations obtained in the exponential method. The so called one-step procedure involves solving the coupled linear equations for λ and S and then using the λ, S values to transform the MC wavefunction. Damping is often employed to accelerate the convergence of this process. If the orthogonal complement MC wavefunctions $\{|n\rangle\}$ are chosen to diagonalize H, $\langle n|H|m\rangle = \delta_{nm} E_n$, and the block of the Hessian matrix arising from $\langle 0|[S,[S,H]]|0\rangle$ becomes diagonal. This then permits the linear equations for λ and S to be decoupled, thereby reducing the dimension of the matrices to be handled. This procedure is still viewed as a one-step approach. In the most common two-step approach, one obtains the CI expansion coefficients by diagonalizing H within the full MC space $\{|0\rangle, |n\rangle\}$. Then, given $|0\rangle$, one solves the partitioned linear equations (which have been partitioned to eliminate S in favor of λ) for the orbital modification parameters λ. The S parameters are not obtained from the linear equations. Such a procedure is called two-step because it involves computing the CI coefficients (from the CI secular problem) and the λ parameters (from the linear equations) in two sequential steps.

Because the SCI achieves both CI coefficient and orbital optimization via secular problems, one can monitor the state to which one is converging by following the eigenvalues of the SCI and CI secular equations from one iteration to the next. In the exponential methods, whose working equations are not secular equations, one can also monitor convergence to a specific state by examining the number of negative eigenvalues of the Hessian matrix.

NUMERICAL STUDIES

We have carried out a large number of exponential and SCI MCSCF calculations on ground and excited states of Be, HCN, HCF$^-$, MgH$_2$, and C$_5$H$_4$ using MC wavefunctions containing single and double excitations. We have examined the convergence rate of these two methods under several sets of conditions for ground and excited ^1S Be atoms and HCN.

1. By freezing the CI expansion coefficients at their optimal converged values, we explored their rates of convergence and ranges of convergence for orbital optimization. The latter was achieved by using starting orbitals which were successively further and further from the proper converged orbitals.
2. The SCI method was compared to a two-step method in which the coupling matrix elements between orbital and CI coefficient optimization $\langle 0 [S,[\lambda,H]]|0\rangle$ were neglected and the CI coefficients were determined via a secular problem.
3. The full (including $\langle 0|[S,[\lambda,H]]|0\rangle$) exponential two-step and one-step procedures were also compared to the SCI method.

The results of these exploratory calculations together with our recommendations and observations concerning optimal implementation of MCSCF wavefunction optimization techniques will be put forth in the second part of this presentation.

MCHF CONVERGENCE USING INCOMPLETE SECOND-DERIVATIVE INFORMATION

Harry F. King and R. Nicholas Camp
Department of Chemistry
Acheson Hall
SUNY at Buffalo
Buffalo, New York 14214

Orbital optimization is achieved using a generalized conjugate gradient algorithm due to Hestenes, Eq. (5). The $\underline{R}$ matrix consists of exact first derivatives with respect to orbital rotations. An arbitrary amount of second derivative information can be incorporated into the $\underline{F}$ matrix. Efficient computation of $\underline{R}$ is described. An application to the nitric oxide dimer is discussed.

INTRODUCTION

We are certainly impressed by the second-order MCHF methods of Jorgensen, Dalgaard and Yeager[1] and of Roothaan, Detrich and Hopper,[2] but somewhat dismayed at the prospect of working with the full second derivative matrix for large basis sets. Thus Dr. Camp and I decided to try our hand at writing our own MCHF program in which the Newton-Raphson equation is replaced by one invented by Hestenes. The Hestenes algorithm contains an arbitrary positive definite matrix, but for rapid convergence it should be at least a rough approximation to the inverse second-derivative matrix. We will have more to say about this presently.

We call our MCHF program HONDO since it is an extension of the single-configuration SCF program with that name. It is a working program but still far from being a finished product. HONDO presently employs a variational wavefunction corresponding to what Bjorn Roos calls "complete CI in the active space." The orbital space is partitioned into three subspaces: N_c inactive (alias "core" or "spectator") orbitals with fixed occupation numbers equal to two, N_a active orbitals with variable occupation numbers, and N_v virtual orbitals with occupation numbers equal to zero. Let there be N spatial orbitals and N_e^{total} electrons so that

$$N = N_c + N_a + N_v, \tag{1}$$

$$N_e^{total} = 2N_c + N_e. \tag{2}$$

The CI expansion includes all configurations with a given spin that can be constructed for N_e electrons in N_a orbitals. In writing the program we have focussed on applications in which N is an order of magnitude greater than N_a, e.g. $N \sim 100$ and $N_a = 2$ to 6. Although we have yet to carry out calculations for such systems, we are thinking specifically of transition metal complexes. Please keep in mind that the active orbitals constitute a small fraction of the total number of M.O.'s. Thus it is significant that HONDO performs the four-index transformation and computes various density matrices only for the small set of active orbitals.

In the absence of point group symmetry the number of independent first derivatives with respect to orbital rotations is

$$N_f = N_cN_a + N_aN_v + N_vN_c \tag{3}$$

An upper bound on the value of N_f is $N^2/3$, and a typical value is approximately $N^2/5$. Thus the number of unique second-derivative matrix elements is about $N^4/50$. The size of the second-derivative matrix is approaching that of the two-electron integral file. We want to work with much less second derivative information than that. An obvious suggestion is to use only the diagonal elements, but there is probably a better tactic than that. We are certainly open to suggestions.

We have further comments to make on three aspects of this work: (1) method of orbital optimization, (2) computation of derivatives with respect to orbital rotations, and (3) some applications to the computation of ground and excited state potential energy surfaces.

ORBITAL OPTIMIZATION

HONDO employs the exponential operator notation of Bernard Levy.[3] Let $\underline{\Phi}$ be an N-dimensional row vector whose elements are real, orthonormal, spatial orbitals. Let λ be a real parameter and $\underline{A}$ a real N by N, antisymmetric matrix. A unitary transformation on $\underline{\Phi}$ generates a new set of orbitals $\underline{\Phi}(\lambda)$.

$$\Phi(\lambda) = \underline{\Phi}\,\underline{U}(\lambda) = \underline{\Phi}\exp(\lambda\underline{A}) \tag{4}$$

A "major cycle" in our method consists of a one-dimensional search with fixed $\underline{A}$ that minimizes the CI energy with respect to the variable lambda. First let us discuss the computation of the optimum lambda value, then we will return to the more interesting matter of determining the search direction $\underline{A}$. The derivative $(dE/d\lambda)$ at $\lambda=0$ is easily computed from first derivative information computed in the previous major cycle. The program then evaluates $E(\lambda)$ and $(dE/d\lambda)$ at $\lambda=1$. This constitutes a "minor cycle" and involves a four-index transformation, construction and diagonalization of the CI matrix, and computation of first derivatives at $\lambda=1$. One minor cycle requires about as much CPU time as three to four ordinary closed-shell SCF cycles. Knowing the energy and its first derivative at two points determines a cubic approximation to $E(\lambda)$ from which values of λ_{min} and $E(\lambda_{min})$ are estimated. If $\underline{A}$ has been well chosen then λ_{min} is sufficiently close to unity and the search is terminated after one minor cycle. Otherwise, the CI energy and its derivative are recomputed at the predicted lambda, the results are fitted by a cubic spline, and the process is repeated until convergence of the one-dimensional search has been achieved. It is rare, in practice, that a major cycle requires more than two minor cycles. This one-dimensional search plays an important role in the theory described below, it is also responsible for the remarkable stability of the convergence algorithm even when the starting orbitals correspond to a second derivative matrix with one or more negative eigenvalues. (This situation has been encountered starting from an apparently reasonable initial guess.)

The $\underline{A}$ matrix for the nth major cycle is computed by the following algorithm based on a generalized conjugate gradient method invented by Hestenes[4] for the solution of large linear systems.

$$\underset{\sim}{A}^{(n)} = \underset{\sim}{F}\,\underset{\sim}{R}^{(n)} + b^{(n-1)}\,\underset{\sim}{A}^{(n-1)} \qquad (5)$$

Like $\underline{A}$ itself, the $\underline{R}$ matrix is antisymmetric. Element R_{ij} is the first derivative of the CI energy with respect to a rotation of the ij-th pair of orbitals. The theory assumes only that $\underline{F}$ is a positive definite Hermitian matrix.[5] The first term in (5) by itself corresponds to Newton-Raphson when $\underline{F}$ is exactly the inverse of the matrix of second derivatives. The second term effectively accumulates information from previous major cycles and so corrects for errors in second-order terms and for third and higher-order effects. Hestenes derives the following expression for the b coefficient:

$$b^{(n-1)} = \rho^{(n)}/\rho^{(n-1)} \qquad (6)$$

where

$$\rho^{(n)} = \mathrm{tr}\,(\underline{R}^{(n)}\ \underline{F}\,\underline{R}^{(n)}). \qquad (7)$$

In effect the Newton Raphson equation has been combined with the iterative conjugate gradient method for solving large linear systems in such a way as to eliminate the explicit appearance of the second derivative matrix from the formalism. To show this, let us set up the Newton Raphson equation in N_f dimensional space. Let $\underline{x}$ be a vector in this space that denotes the orbitals. For example, we can pick a set of reference orbitals and define any others in terms of these by an exponential unitary transformation $\underline{U} = \exp(\underline{A}')$. Then $\underline{x}$ consists of the independent elements of A', e.g. the elements in the lower triangle. Similarly let $\underline{r}_0$ consist of the independent elements of $\underline{R}$ evaluated for the reference orbitals. (We are aware that there exist Unitary matrices that cannot be expressed by the exponential formula with A' restricted to being real. That problem is easily overcome by adopting a suitable phase convention.) Let $\underline{M}$ be the N_f by N_f matrix of second derivatives with respect to elements of $\underline{x}$. Expand the CI energy through quadratic terms.

$$E(\underline{x}) = E_0 - \underline{r}_0^{\dagger}\,\underline{x} + 1/2\ \underline{x}^{\dagger}\underline{Mx} + \ldots \qquad (8)$$

In this notation the solution of the Newton Raphson equation is $\underline{x}=\underline{h}$ where

$$\underline{Mh} = \underline{r}_0. \qquad (9)$$

Equation (5) is based on the generalized conjugate gradient method of Hestenes which yields $\underline{h}$ as the limit of a finite, converging sequence of vectors $\underline{x}^{(1)}, \underline{x}^{(2)}, \ldots, \underline{h}$. Each major cycle in our method generates a new member of the sequence. In effect we redefine $\underline{M}$ at each step to be the second derivative matrix at the new $\underline{x}^{(n)}$ so as to correct for cubic and higher order terms in (5) as well as for errors associated with non commutivity of rotation operations. The Hestenes algorithm is one of a class of methods for solving large linear systems such as (9). Each of these iterative methods generates the vectors $\{\underline{M}^n\underline{v},\ n = 1,2,3,\ldots\}$ for some generator $\underline{v}$. This is known generally as a "Krylov sequence" or Krylov space. At each step $\underline{h}$ is approximated by the best vector in the Krylov space according to some criterion. Following Hestenes our criterion is that $\underline{x}$ minimize the generalized norm of the residual. The residual vector corresponding to (9) is $\underline{r}_0 - \underline{Mx}$.

$$<\underline{r}_0-\underline{Mx}|\underline{H}|\underline{r}_0-\underline{Mx}> \; = \; <\underline{h}-\underline{x}|\underline{M}^{\dagger}\underline{H}\ \underline{M}|\underline{h}-\underline{x}> \qquad (10)$$

Here $\underline{H}$ can be any positive definite matrix. (Do not confuse $\underline{H}$ with the arbitrary matrix $\underline{F}$ introduced above.) A particular choice of $\underline{H}$ discussed by Hestenes and employed by us is $\underline{H} = \underline{M}^{-1}$. This tends to avoid overemphasis of small components of $\underline{h}-\underline{x}$ corresponding to large eigenvalues of $\underline{M}$. It follows immediately that this is equivalent to minimizing the quadratic approximation to (8) in the Krylov space. Our method replaces this step with minimization of $E(\underline{x})$ itself using the one-dimensional search technique already discussed. In the absence of higher order terms this reduces to the Hestenes algorithm. The $\underline{A}$ matrix in (5) corresponds to the change in $\underline{x}$

$$\underline{x}^{(n+1)} = \underline{x}^{(n)} + \lambda A^{(n)} \qquad (11)$$

and $\underline{R}^{(n)}$ corresponds to the residual vector evaluated for the current orbitals $\underline{x}^{(n)}$.

It is, of course, <u>unnecessary to generate the M matrix</u> explicitly if there exists an equivalent, alternative technique for extending the Krylov sequence. This is, in fact, what we do by explicitly computing $\underline{r}^{(n)}$. Note that if one drops higher order terms from (8) then the negative gradient is just the residual vector.

$$\underline{r}(\underline{x}) = \underline{r}_0 - \underline{Mx} \qquad (12)$$

A Newton-Raphson method using the Hestenes algorithm evaluates the right hand side of (12), we evaluate the left hand side. The explicit matrix-vector multiplication is several times faster computationally than evaluation of $\underline{r}$. On the other hand we completely avoid constructing $\underline{M}$ and at the same time pick up a small advantage from the fact that we effectively redefine $\underline{M}$ at each major cycle and so correct for non quadratic effects more frequently.

Hestenes shows that (5) together with (6) minimizes $E(\underline{x})$ in the Krylov space, in particular, it minimizes (10) in the subspace spanned by the two terms in (5). Our computational experience shows, however, that the CI energy is not always well minimized in this two-dimensional subspace when b is computed by (6). This is almost certainly due to the neglect of cubic and other higher order effects not considered by Hestenes. Thus we have recently replaced (6) with a two-dimensional search procedure that minimizes the

CI energy with respect to the coefficients of both terms in (5). This two dimensional search has sometimes resulted in a marked improvement during early MCHF cycles, but is a waste of time in later cycles. A simple test could be coded into the program to improve efficiency in this respect. This is just one of several illustrations of the need for a general "tune up" of the program.

Improvement in the selection of the $\underline{F}$ matrix is of central interest. To date we have employed only diagonal $\underline{F}$ matrices whose elements are computed by a simple finite difference formula applied to $\underline{R}$ matrix elements for consecutive major cycles. This has actually worked reasonably well, but we suspect that certain small second derivative elements should be computed accurately. We want to maintain an exceedingly sparse $\underline{F}$ matrix, but we would certainly be willing to include a small number of off-diagonal elements if we could only identify the really important ones. We hoped to generate some discussion at this meeting concerning what second derivative information is of greatest importance.

COMPUTATION OF FIRST-ORDER TERMS

The CI energy is given in terms of integrals over MO's as follows:

$$E = E_{nuc} + \sum_{ij} (i|h_n|j)\, P_{ij} + \sum_{ijk\ell} (ij||k\ell)\, P^{(2)}_{ijk\ell}, \tag{13}$$

where $\underline{P}$ and $\underline{P}^{(2)}$ are one and two-particle density matrices respectively. They are related by:

$$P^{(2)}_{ijk\ell} = \frac{1}{2} P_{ij}\, P_{k\ell} - \frac{1}{8} (P_{ik}\, P_{j\ell} + P_{i\ell}\, P_{jk}) + P'_{ijk\ell}. \tag{14}$$

$$P^{(2)}_{ijk\ell} = P^{(2)}_{jik\ell} = P^{(2)}_{k\ell ij} = \ldots \tag{15}$$

It is important to appreciate that $P'_{ijk\ell}$ is non zero only if all four indices are in the active space. For the problems under consideration it requires a trivial amount of core storage to hold the non zero elements of P' and the corresponding two electron integrals. Only those elements in the active-active block of the symmetric, one-particle density matrix are needed since the others are given by:

$$P_{ij} = 2\delta_{ij} \quad \text{if i and/or j are spectators,} \tag{16}$$

$$P_{ij} = 0 \quad \text{if i and/or j are virtuals.} \tag{17}$$

Let the orbitals vary according to (4) for fixed $\underline{A}$ and variable λ. Consider the CI calculation to be repeated for each λ. We want the values of E and $dE/d\lambda$ at each step during the variation.

$$\underline{\phi}(\lambda + \varepsilon) = \underline{\phi}(\lambda)\, [\underline{I} + \varepsilon \underline{A} + O(\varepsilon^2)] \tag{18}$$

$$E(\lambda + \varepsilon) = E(\lambda) + 2\, \varepsilon \sum_{i'ij} (i'|h_N|j)\, A_{i'i}\, P_{ij} +$$

$$4\varepsilon \sum_{i'ijk\ell} (i'j||k\ell)\, A_{i'i}\, P^{(2)}_{ijk\ell} + O(\varepsilon^2). \tag{19}$$

Note that first-order variations in the density matrices make no first-order contribution to the energy by virtue of having diagonalized the CI matrix. It follows immediately that

$$dE/d\lambda = \mathrm{tr}(\underline{AR}') = \frac{1}{2}\, \mathrm{tr}(\underline{AR}). \tag{20}$$

where

$$R'_{ii'} = 2\sum_{j} (i'|h_N|j)\, P_{ij} + 4 \sum_{jk\ell} (i'j||k\ell)\, P^{(2)}_{ijk\ell}, \tag{21}$$

$$\underline{R} = \underline{R}' - (\underline{R}')^{\dagger}. \tag{22}$$

To evaluate the $\underline{R}$ matrix substitute (14) into (21) and separate the resulting expression into two terms. The first term contains everything not involving $\underline{P}'$ contributions. Computationally, evaluation of this first term is in every respect like the evaluation of a closed-shell Fock matrix. It can be evaluated in an AO basis using point group symmetry[6] and a PK file,[7] and then be transformed to an MO basis. For large basis sets this incurs the same I/O charges and just slightly less CPU time than for an ordinary SCF cycle.

The second term, i.e. the $\underline{P}'$ contribution to $\underline{R}'$, is evaluated in a mixed MO, AO basis. It reduces to evaluation of the inner product of a pair of N-dimensional vectors.

$$4 \sum_{jk\ell} (i'j||k\ell)\, P'_{ijk\ell} = \sum_{\sigma=1,N} B_{\sigma i}\, C_{\sigma i'} \tag{23}$$

Here $\underline{C}$ is the MO coefficient matrix which expresses the MO's, ϕ_i, in terms of AO's, χ_σ.

$$\underline{\chi} = (\chi_1, \chi_2, \ldots, \chi_n) \tag{24}$$

$$\underline{\Phi} = \underline{\chi C} \tag{25}$$

Evaluation of (23) is performed for each element of $\underline{R}'$ in the active-virtual and active-spectator blocks. This is a fast step. The $\underline{B}$ matrix has N rows and N_a columns.

$$B_{\sigma i} = 4\sum_{j} \sum_{k \geq \ell} (\sigma j||k\ell)\, P'_{ijk\ell}\, (2 - \delta_{k\ell}) \tag{26}$$

Evaluation of (26) is not a heavy computation. There are N_t terms in the sum. (See Table 1.) The two-electron integral, $(\sigma j||k\ell)$, is preserved from the penultimate step in an earlier four-index transformation. The number of such 3/4 transformed integrals is N x N_t. For small problems these can be held in core.

Note that the strategy is to divide the computation into two types of terms. The first involves only one-particle density matrices, and these are evaluated in an A.O. basis. The second involves only active orbitals, and these are evaluated in an MO basis. This organization has also proved useful for the computation of energy derivatives with respect to nuclear displacement, i.e. analytic gradients.

TABLE 1. DIMENSIONS OF ARRAYS

N_a	N_p	N_t	N_q
2	3	6	6
3	6	18	21
4	10	40	55
5	15	75	120
6	21	126	231
7	28	196	406
8	36	288	666

N_a = number of active orbitals

N_p = number of unique pairs

$N_t = N_a \times N_p$ = number of unique triple indices

N_q = number of unique P' elements

POTENTIAL ENERGY SURFACES FOR A WEAKLY INTERACTING BIRADICAL SYSTEM

A variety of experimental evidence indicates that nitric oxide forms a dimer in the gas phase with an association energy of about 2 kcal per mole of dimer. Its structure has not yet been determined, but the nitrogen-nitrogen separation is almost certainly unusually long. Various features of the electronic structure can be expected to more nearly resemble that of a transition state complex rather than that of a normal covalently bonded molecule. Unlike a transition state complex, however, the dimer is far more accessible to spectroscopic observation.[8] These weak complexes would seem, therefore, to provide interesting test cases for quantum chemical methods.

The ground electronic state of NO is a doublet pi state. Thus there are four nearly degenerate π^* orbitals in the dimer for two electrons. Not surprisingly, earlier calculations using single configuration SCF theory yield a qualitatively incorrect description of the dimer structure. We have recently begun a rather extensive exploration of the ground state potential energy surface using $N_e = 2$ and $N_a = 4$. Less extensive calculations are being carried out for the low-lying excited state surfaces. We wish to discuss some rather puzzling aspects of these results. In particular, this level of theory predicts a completely repulsive ground state surface.

REFERENCES

(1) E. Dalgaard and P. Jorgensen, J. Chem. Phys. 69, 3833 (1978). D. L. Yeager and P. Jorgensen, J. Chem. Phys. 71, 755 (1979).

(2) C. C. J. Roothaan, J. Detrich and D. G. Hopper, Int. J. Quant. Chem. Symp. 13, 93 (1979).

(3) B. Levy, Chem. Phys. Letters 4, 17 (1969), 18, 59 (1973); Int. J. Quant. Chem. 4, 297 (1970). J. Douady, Y. Ellinger, R. Subra and B. Levy, J. Chem. Phys. 72, 1452 (1980).

(4) M. R. Hestenes, in Proceedings of Symposia in Applied Math., vol. 6 (Numerical Analysis), pp. 83-102 (McGraw-Hill, N. Y., 1956).

(5) Our Eq. (5) is (6:2f) in Ref. 4, where our a, F, r correspond to p, K, g. The original method of conjugate gradients, contained nothing analogous to the F matrix, but as discussed in Section 8 of Ref. 4, the generalized method is just the original algorithm applied in a transformed space. Our F is the square of the matrix for the linear transformation.

(6) M. Dupuis and H. F. King, Int. J. Quant. Chem. 11, 613 (1977).

(7) R. C. Raffenetti, Chem. Phys. Letters 20, 335 (1973).

(8) C. E. Dinerman and G. E. Ewing, J. Chem. Phys. 35, 626 (1970).

THE UNITARY GROUP FORMULATION OF MCSCF THEORY AND EXTENSIONS*

C. J. Nelin and F. A. Matsen
Departments of Chemistry and Physics
University of Texas at Austin
Austin, TX 78712

*Supported by the Robert A. Welch Foundation of Houston, Texas.

I. INTRODUCTION

The unitary group formulation of the many-body problem[1] is a viable alternative to the Slater determinant and second quantized formulation when the Hamiltonian is spin-free. In the unitary group formulation the Hamiltonian is expressed as a Hermitian second-degree polynomial in the infinitesimal generators of $U(\rho)$, where ρ is the number of spin-free orbitals. Each irreducibly invariant representation space of $U(\rho)$ is uniquely characterized by a partition of a non-negative integer and the associated Young diagram. Each space is invariant under the Hamiltonian, so the partition is a quantum number. The Pauli-allowed spaces are characterized by Young diagrams containing no more than two columns such that the particle number, N, is given by the number of squares and the spin, S, is given by one-half the difference in the lengths of the two columns. Thus the Pauli-allowed irreducible invariant spaces of $U(\rho)$ are pure spin states and spin projection is not required. These spaces are spanned by either Gel'fand states or generator states,[2] and the matrix elements of the Hamiltonian over either set of basis vectors are evaluated straightforwardly by standard Lie algebraic techniques.[3] The unitary group formulation has been shown to be directly applicable to large CI calculations.[4] Further, the unitary group formulation provides a basis for various many-body theories.

In previous papers we have given the unitary group formulation to single-configuration SCF,[5] multiconfiguration SCF,[6] and coupled-cluster[7] theories. The present paper is a review and an extension of these ideas. Section II contains a review of the relevant unitary group theory. Section III contains a brief outline of the unitary group formulation of MCSCF theory. In Section IV we treat the effective Hamiltonian for a multiconfigurational reference (primary) space, a device for enfolding states from the external secondary space into the multiconfigurational primary space. In Section V these two concepts are combined into an MCSCF effective Hamiltonian. We outline in the appendix a test calculation for H_2O with a double zeta basis.

II. THE UNITARY GROUP FORMULATION

The group

$$U(\rho) = \{x,y,z,\ldots\},\ x^{+} = x^{-1} \tag{2.1}$$

is the set of all unitary transformations on a set of ρ orthonormal spin-free (freeon) orbitals:

$$V(\rho)\colon \{|r\rangle,\ r = 1 \text{ to } \rho\} \tag{2.2}$$

The group elements have an exponential form

$$x = e^{X} \text{ where } X^{+} = -X \text{ (skew Hermitian)} \tag{2.3}$$

where X is an element of the Lie algebra,

$$LAU(\rho) = \{X,Y,Z,\ldots\} \tag{2.4}$$

This nonassociative algebra is closed under the Lie product (commutator),

$$[X,Y] \in LAU(\rho) \tag{2.5}$$

The Lie algebra has a basis given by

$$LAU(\rho)\colon \{E_{rs};\ r,s = 1 \text{ to } \rho\} \text{ with } E^{\dagger}_{rs} = E_{sr} \tag{2.6}$$

whose elements satisfy the Lie product

$$[E_{rs},E_{tu}] = \delta_{st}E_{ru} - \delta_{ru}E_{ts} \tag{2.7}$$

The covering algebra is

$$CAU(\rho)\colon \{I,E_{rs},E_{rs}E_{tu},\ldots\} \equiv E(i,\tau) \tag{2.8}$$

where i is the degree of the element in $CAU(\rho)$ and τ labels the indices.

The unitary group many-electron Hamiltonian, which lies in $CAU(\rho)$, is

$$H = H^{0} + V \tag{2.9}$$

with

$$H^{0} = \overset{\rho}{\overline{\sum_{r}\sum_{s}}}\, h_{rs}E_{rs} \tag{2.10}$$

and

$$V = \overset{\rho}{\overline{\sum_{r}\sum_{s}\sum_{t}\sum_{u}}}\, v_{rstu}(E_{rs}E_{tu} - \delta_{st}E_{ru}) \tag{2.11}$$

where

$$h_{rs} = \langle r_i|h_i|s_i\rangle \tag{2.12}$$

and

$$v_{rstu} = \langle t_j|\langle r_i|h_{ij}|s_i\rangle|u_j\rangle = v_{turs} \tag{2.13}$$

The Hilbert space of this Hamiltonian is the representation space of $U(\rho)$:

$$V_{\rho} = \sum_{[\lambda]} \oplus V_{\rho}^{[\lambda]} \tag{2.14}$$

where each $V_{\rho}^{[\lambda]}$ is invariant under $CAU(\rho)$ and where $[\lambda]$ is a partition of the positive integer N,

$$[\lambda] = [\lambda_1,\lambda_2,\ldots,\lambda_{\rho}] \tag{2.15}$$

with

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{\rho} \tag{2.16}$$

and

$$\sum_{i=1}^{\rho} \lambda_i = N \tag{2.17}$$

$[\lambda]$ is graphically represented by a Young diagram $YD[\lambda]$, with N boxes where the ith row has λ_i boxes. Each invariant space is spanned by $f_\rho[\lambda]$ Gel'fand states labeled by Gel'fand tableaux constructed by adding integers to $YD[\lambda]$ in nondescending order along rows and in descending order down columns and there exist explicit formulae for the evaluation of matrix elements over Gel'fand states, where each Gel'fand state can be represented by a Paldus array and a walk on a Shavitt graph (see Appendix). For the Gel'fand basis represented by

$$V_\rho^{[\lambda]}: \{|F\rangle,\ \dim = f_\rho[\lambda]\} \tag{2.18}$$

a matrix element, $\langle F|H|F'\rangle$, has a closed form given as a function of the walks corresponding to $|F\rangle$ and $|F'\rangle$.[4]

For large systems it is often impractical to deal with the full space $V_\rho^{[\lambda]}$ even after limiting the number of allowed excitations. Thus we partition the full space into a primary (or reference) space and a second (or virtual) space:

$$\begin{aligned} V_F = V_P: \{|P\rangle;\ \dim = d_P\} \oplus V_Q\{|Q\rangle;\\ \dim = d_Q\} \end{aligned} \tag{2.19}$$

For the primary space there is selected a small set of states which interact strongly with the zero order ground state. Frequently the primary space is selected by first dividing the orbitals into three classes -- (a) core, inactive frozen (internal); (b) active, valence, mixed (internal); and (c) external, empty -- and then retaining for the primary space only those states with configurations that employ (a) and (b). This primary space has been called the complete active subspace (CAS).[8] We will represent this orbital partitioning by

$$\begin{aligned} V(\rho) = V_{\text{internal}} &: \{|i\rangle, |j\rangle, |k\rangle, |\ell\rangle, \ldots;\\ &\quad \dim = \rho_i\}\\ \oplus V_{\text{external}} &: \{|e\rangle, |e'\rangle, |e''\rangle, \ldots;\\ &\quad \dim = \rho_e\} \end{aligned} \tag{2.20}$$

Furthermore,

$$\begin{aligned} V_{\text{internal}} = V_{\text{core}} &: \{|c\rangle, |c'\rangle, |c''\rangle, \ldots;\\ &\quad \dim = \rho_c\}\\ \oplus V_{\text{mixed}} &: \{|m\rangle, |m'\rangle, |m''\rangle, \ldots;\\ &\quad \dim = \rho_m\} \end{aligned} \tag{2.21}$$

where

$$\rho = \rho_i + \rho_e = \rho_c + \rho_m + \rho_e \tag{2.22}$$

An example of this partitioning is given for the water molecule in Appendix A.

Within this partitioning scheme we obtain three distinct classes of single excitations from our primary space, c → e, m → e, and c → m. Furthermore, we obtain the relations

$$E_{ce}|P\rangle = E_{me}|P\rangle = E_{cm}|P\rangle = 0 \tag{2.23}$$

and

$$E_{cc'}|P\rangle = 2\delta_{cc'}|P\rangle \tag{2.24}$$

(Recall that our primary space contained all configurations involving core and mixed (valence) orbitals only and that our core orbitals were all doubly occupied.) Using (2.23) and (2.24), we can calculate nonzero matrix elements for the primary space.[6] For another choice for the primary space, see Appendix B.

III. MCSCF THEORY

In the present section we give the unitary group formulation of *MCSCF theory*,[6] *which will* also serve to introduce section V.

The variational state is constructed as follows. We apply the state and orbital variation simultaneously. However, since $[S,T] \neq 0$, the order of application is significant. We follow previous work[9] and apply the state variation first. The variational state $|\tilde{0}\rangle$ is then

$$|\tilde{0}\rangle \equiv e^{T}e^{S}|0\rangle \tag{3.1}$$

where T and S are orbital and state variation operators, respectively. The energy of this state is

$$E(\tilde{0}) = \langle\tilde{0}|H|\tilde{0}\rangle = \langle 0|e^{-S}e^{-T}He^{T}e^{S}|0\rangle \tag{3.2}$$

where H is defined by (2.9). On applying the Baker-Campbell-Hausdorff expansion, and including only terms to second order, (3.2) becomes

$$\begin{aligned} E(\tilde{0}) = {} & \langle 0|H|0\rangle + \langle 0|[H,T]|0\rangle + \langle 0|[H,S]|0\rangle\\ & + \tfrac{1}{2}\langle 0|[[H,T],T]|0\rangle + \tfrac{1}{2}\langle 0|[[H,S],S]|0\rangle\\ & + \langle 0|[[H,T],S]|0\rangle \end{aligned} \tag{3.3}$$

The energy is an extremum when

$$\begin{aligned} \delta E(\tilde{0}) = {} & \langle 0|[H,\delta T]|0\rangle + \langle 0|[H,\delta S]|0\rangle\\ & - \langle 0|[\delta S,H,S]|0\rangle - \langle 0|[\delta T,H,T]|0\rangle\\ & - \langle 0|[[\delta T,H],S]|0\rangle - \langle 0|[[T,H],\delta S]|0\rangle\\ = {} & 0 \end{aligned} \tag{3.4}$$

where we have used the symmetric double commutator defined by

$$[A,B,C] = \frac{1}{2}\,[[A,B],C] + [A,[B,C]] \tag{3.5}$$

Thus to second order the extremum of $E(\tilde{0})$ (3.4) is defined by the matrix equation

$$-\begin{Bmatrix} W_{rs} \\ V_{K0} \end{Bmatrix} = \begin{Bmatrix} C_{rs,r's'} & C_{rs,K'} \\ C_{K,r's'} & C_{K,K'} \end{Bmatrix} \begin{Bmatrix} T_{r's'} \\ S_{K'0} \end{Bmatrix} \tag{3.6}$$

where, using the orbital classes of Roos et al.,[8]

$$W_{rs} = \begin{Bmatrix} W_{ce} \\ W_{cm} \\ W_{me} \end{Bmatrix} \tag{3.7}$$

$$C_{rs,r's'} = \begin{pmatrix} C_{ce,c'e'} & C_{ce,c'm'} & C_{ce,m'e'} \\ C_{cm,c'e'} & C_{cm,c'm'} & C_{cm,m'e'} \\ C_{me,c'e'} & C_{me,c'm'} & C_{me,m'e'} \end{pmatrix} \tag{3.8}$$

$$C_{rs,K} = \begin{Bmatrix} C_{ce,K} \\ C_{cm,K} \\ C_{me,K} \end{Bmatrix} \tag{3.9}$$

and

$$C_{K,rs} = \{C_{K,ce} \quad C_{K,cm} \quad C_{K,me}\} \tag{3.10}$$

The elements in (3.6) are defined by the equations

$$W_{rs} \equiv \langle 0|[E_{sr},H]|0\rangle \tag{3.11}$$

$$V_{K0} \equiv \langle 0|[|0\rangle\langle K|,H]|0\rangle \tag{3.12}$$

$$\begin{aligned} C_{rs,r's'} \equiv\ & \langle 0|[E_{sr},H,E_{r's'}]|0\rangle \\ & - \langle 0|[E_{sr},H,E_{s'r'}]|0\rangle \end{aligned} \tag{3.13}$$

$$\begin{aligned} C_{rs,K'} \equiv\ & \langle 0|[[E_{sr},H],|K'\rangle\langle 0|]|0\rangle \\ & - \langle 0|[[E_{sr},H],|0\rangle\langle K'|]|0\rangle \end{aligned} \tag{3.14}$$

$$\begin{aligned} C_{K,r's'} \equiv\ & \langle 0|[[E_{r's'},H],|0\rangle\langle K|]|0\rangle \\ & - \langle 0|[[E_{s'r'},H],|0\rangle\langle K'|]|0\rangle \end{aligned} \tag{3.15}$$

and

$$\begin{aligned} C_{K,K'} \equiv\ & \langle 0|[|0\rangle\langle K|,H,|K'\rangle\langle 0|]|0\rangle \\ & - \langle 0|[|0\rangle\langle K|,H,|0\rangle\langle K'|]|0\rangle \end{aligned} \tag{3.16}$$

The explicit expressions for these elements are given elsewhere.[6]

We can rewrite equation (3.6) in condensed form:

$$-\begin{Bmatrix} W \\ V \end{Bmatrix} = \{C\} \begin{Bmatrix} T \\ S \end{Bmatrix} \tag{3.17}$$

or

$$-\begin{Bmatrix} T \\ S \end{Bmatrix} = \{C\}^{-1} \begin{Bmatrix} W \\ V \end{Bmatrix} \tag{3.18}$$

where the values of T and S will be used to update the orbital coefficients and configuration coefficients, respectively. Calculations are carried out iteratively until W and V vanish (or become less than the convergence criterion). That W vanish is equivalent to the Brillouin-Levy Berthier condition (BLB).[10] When V vanishes, orthogonality of the multiconfigurational states is insured.

The calculations begin with an SCSCF determination of the orbitals and a diagonalization of the resulting representation of H in the primary space to obtain initial multiconfigurational states. The iteration can be carried out in one or two steps.

One-Step Procedure

a. Compute $W^{(1)}$, $V^{(1)}$, and $C^{(1)}$.

b. Compute $T^{(1)}$ and $S^{(1)}$ from (3.18).

c. Compute new orbital and configuration coefficients.

d. Repeat steps (a) through (c) until self-consistency is obtained.

Two-Step Procedure

In a two-step procedure the orbital and state coefficients are obtained separately. This is done by setting V = 0. The new state coefficients will be obtained via a Hamiltonian diagonalization. When V = 0 in (3.6), we obtain

$$-\begin{Bmatrix} W_{rs} \\ 0 \end{Bmatrix} = \begin{Bmatrix} C_{rs,r's'} & C_{rs,K'} \\ C_{K,r's'} & C_{K,K'} \end{Bmatrix} \begin{Bmatrix} T_{r's'} \\ S_{K'0} \end{Bmatrix} \tag{3.19}$$

or

$$-\{W_{rs}\} = \{C_{rs,r's'}\}\{T_{r's'}\} + \{C_{rs,K'}\}\{S_{K'0}\} \tag{3.20}$$

and

$$0 = \{C_{K,r's'}\}\{T_{r's'}\} + \{C_{K,K'}\}\{S_{K'0}\} \tag{3.21}$$

Then from (3.20) and (3.21) we obtain

$$\begin{aligned} -\{W_{rs}\} = \Big(&\{C_{rs,r's'}\} + \{C_{rs,K'}\}\{C_{K,K'}\}^{-1} \\ &\{C_{K,r's'}\}\Big)\{T_{r's'}\} \end{aligned} \tag{3.22}$$

or

$$\begin{aligned} \{T_{rs}\} = -\Big(&\{C_{rs,r's'}\} + \{C_{rs,K'}\}\{C_{KK'}\}^{-1} \\ &\{C_{K,r's'}\}\Big)^{-1}\{W_{rs}\} \end{aligned} \tag{3.23}$$

Furthermore, now

$$C_{KK'} = \delta_{K,K'}\big(E(K) - E(0)\big) \tag{3.24}$$

as the Hamiltonian is diagonalized between iterations. The steps in the two-step procedure are

a. Diagonalize H to obtain configuration coefficients.

b. Compute C_s (3.12)-(3.15) and W (3.11).

c. Obtain T from (3.17).

d. Compute new orbital coefficients.

e. Repeat steps (a) through (d) until convergence is obtained.

The optimum procedure for carrying out an MCSCF calculation is a subject of active debate.[11]

Questions in need of further investigation are orbital space selections and hence partitions, as well as techniques for dealing with convergence difficulties. When the orbital space becomes large enough, their quality becomes unimportant; however, the problem becomes too large to handle. What we propose in the next section is to combine MCSCF and effective Hamiltonian theories to bring in the contributions of a larger orbital space without drastically increasing the size of the MCSCS calculation.

IV. THE EFFECTIVE HAMILTONIAN[12]

One technique for the construction of an effective Hamiltonian is by means of an exponential and the Baker-Campbell-Hausdorff (BCH) expansion

$$\mathcal{H} = e^{-G}\, He^{G} \tag{4.1}$$

$$= H + [H,G] + \frac{1}{2!}\Big[[H,G],G\Big] + \ldots \tag{4.2}$$

where

$$G = \begin{cases} \in |P><Q|,\ |Q><P| \\ \quad \text{or} \\ \in \ \mathrm{CAU}(p) \end{cases} \tag{4.3}$$

and require that

$$[\mathcal{H},\mathcal{P}] = [\mathcal{H},\mathcal{Q}] = 0 \tag{4.4}$$

where $\mathcal{P}$ and $\mathcal{Q}$ are projection operators for V_P and V_Q, respectively. We can employ two methods for the construction of $\mathcal{H}$, where G is determined either by perturbation or by coupled-cluster (variational) theory.

Perturbation Theory

1. We perform a perturbation expansion of $\mathcal{H}$ and G :

$$\mathcal{H} = \sum_{n=0} \mathcal{H}^{(n)} \tag{4.5}$$

and

$$G = \sum_{n=0} G^{(n)} \tag{4.6}$$

Substituting (4.5) and (4.6) into (4.2) we obtain

$$\mathcal{H}^{(0)} = H^0 + [H^0, G^{(0)}] \tag{4.7}$$

$$\mathcal{H}^{(1)} = V + [H^0,G^{(1)}] \tag{4.8}$$

$$\mathcal{H}^{(2)} = [H^0,G^{(2)}] + [V,G^{(1)}] + \frac{1}{2}\Big[[H^0,G^{(1)}],G^{(1)}\Big],\ \text{etc.} \tag{4.9}$$

2. The explicit form of G (4.6) is defined by the conditions

$$G^{(0)} = 0 \tag{4.10}$$

and

$$\mathcal{P}G^{(n)}\mathcal{P} = \mathcal{Q}G^{(n)}\mathcal{Q} = 0 \tag{4.11}$$

We take

$$G^{(n)} = \sum_{P\in V_P} \sum_{Q\in V_Q} G^{(n)}_{PQ}\ \{|P><Q| - |Q><P|\} \tag{4.12}$$

3. The representation of $\mathcal{H}$ on V_P is then given by

$$\mathcal{H}^{(0)}_{PP'} = \delta_{P,P'}E^0_P \tag{4.13}$$

$$\mathcal{H}^{(1)}_{PP'} = V_{PP'} \tag{4.14}$$

$$\mathcal{H}^{(2)}_{PP'} = \sum_Q \Big\{-V_{QP'}G^{(1)}_{PQ} - V_{PQ}\,G^{(1)}_{P'Q} - \frac{1}{2}\Big(E^0_P - E^0_Q\Big)\,G^{(1)}_{PQ}G^{(1)}_{P'Q} - \frac{1}{2}\Big(E^0_{P'} - E^0_Q\Big)\,G^{(1)}_{PQ}G^{(1)}_{P'Q}\Big\} \tag{4.15}$$

etc.

Using (4.7) and (4.8) we obtain

$$G^{(1)}_{PQ} = -\frac{V_{QP}}{(E^0_P - E^0_Q)} \tag{4.16}$$

Thus (4.16) becomes

$$\mathcal{H}^{(2)}_{PP'} = \sum_Q \frac{1}{2}\left\{\frac{V_{PQ}V_{QP'}}{(E^0_P - E^0_Q)} + \frac{V_{PQ}V_{QP'}}{(E^0_{P'} - E^0_Q)}\right\} \tag{4.17}$$

Higher order terms are obtained in a similar manner.

Coupled Cluster Theory

For the complete active space decomposition

$$V_P:\ \{\ |\ >,\ F_P{}^\dagger|\ >\}\ ,\ F_P{}^\dagger \equiv E_{mh}E_{m'h'}\ldots \tag{4.18}$$

where h, h',... are active orbitals which are occupied in | >, the highest wright (single reference) state and

$$V_Q:\ \{\ F_Q{}^\dagger|\ >\ \}\ ,\ F_Q{}^\dagger \equiv E_{ec}\ldots E_{mc}\ldots E_{hc}\ldots E_{em}\ldots E_{eh} \tag{4.19}$$

and where care is taken to generate a set of linearly independent states. The effective Hamiltonian is defined

$$\mathcal{H}|K> = E_K\ |K> \tag{4.20}$$

where

$$|K> = \sum_P K_P\ F_P{}^\dagger|\ >\ \in V_P \tag{4.21}$$

is a multiconfigurational reference state. Then

$$\mathcal{P}\mathcal{H}|K> = E_K\ |K> \text{ and } \mathcal{Q}\mathcal{H}|K> = 0 \tag{4.22}$$

In (4.1) we take

$$G = \sum_Q G_Q F_Q^{+} \qquad (4.23)$$

where G includes one-, two-,...particle clusters. Then by (4.2) and (4.23)

$$< |F_Q(H + [H,G] +\ldots) \sum_P K_P F_P^{+}| >= 0 \qquad (4.24)$$

which is solved iteratively with $\mathcal{H}^{(0)} = H \rightarrow K_P^{(0)} \rightarrow G^{(1)} \rightarrow \mathcal{H}^{(1)} \rightarrow K_P^{(1)} \rightarrow G^{(2)}$ etc. The matrix elements of the resulting products of generators over the highest weight state are evaluated algebraically[2].

V. MCSCF-EH THEORY

We propose to combine the MCSCF and effective Hamiltonian approaches. In principle this procedure will provide the best definition of the primary space, thereby reducing the size of the H_{PQ} terms and reducing the number to be included in the effective Hamiltonian.

We take the same variational state as before,

$$|\tilde{0}> = e^T e^S|0> \qquad (5.1)$$

where now we compute its energy using an effective Hamiltonian

$$E(\tilde{0}) = <0|e^{-S}e^{-T}\mathcal{H}e^{T}e^{S}|0> \qquad (5.2)$$

$$= <0|e^{-S}e^{-T}e^{-G}He^{G}e^{T}e^{S}|0> \qquad (5.3)$$

$$= <0|H|0> + <0|[H,G]|0> + <0|[H,S]|0>$$

$$+ <0|[H,T]|0> + \frac{1}{2} <0|[[H,G],G]|0>$$

$$+ \frac{1}{2} <0|[[H,S],S]|0> + \frac{1}{2} <0|[[H,T],T]|0>$$

$$+ <0|[[H,G],S]|0> + <0|[[H,G],T]|0>$$

$$+ <0|[[H,T],S]|0> \qquad (5.4)$$

The most general treatment would be to solve for G as a third variational parameter (along with T and S). This would increase the size and complexity of equation (3.17). Although this completely variational procedure would be bounded, it appears to be too cumbersome, and so we choose to use a predetermined G (from either perturbation or coupled cluster theory). An extremum of the energy then occurs when

$$\delta E(\tilde{0}) = 0 = <0|[(H+[H,G]),\delta S]|0>$$

$$+ <0|[(H+[H,G]),\delta T]|0> - <0|[\delta S,H,S]|0>$$

$$- <0|[\delta T,H,T]|0> + <0|[[H,T],\delta S]|0>$$

$$+ <0|[[H,\delta T]S]|0> \qquad (5.5)$$

or, in matrix form,

$$- \begin{Bmatrix} W + W' \\ V + V' \end{Bmatrix} = \{C\} \begin{Bmatrix} T \\ S \end{Bmatrix} \qquad (5.6)$$

where C, W, and V are defined by (3.21)-(3.25) and the new terms are

$$W' = <0|[E_{sr},[H,G]]|0> \qquad (5.7)$$

$$V' = <0|[|0><K|,[H,G]]|0> \qquad (5.8)$$

We now reach convergence when

$$<0|[E_{sr},(H+[H,G])]|0> \rightarrow 0 \qquad (5.9)$$

and

$$<0|[|0><K|,(H+[H,G])]|0> = <K|(H + [H,G])|0> \rightarrow 0 \qquad (5.10)$$

Equation (5.9) is simply a BLB condition for our truncated effective Hamiltonian (i.e., $\mathcal{H} = H+[H,G]$), and (5.10) insures that the states determined diagonalize the effective Hamiltonian. This process is to be carried out self-consistently. Since the procedure yields a new set of MCSCF orbitals and primary space basis vectors, we must compute an updated effective Hamiltonian (i.e., update the G's) each iteration. As with conventional MCSCF there are both "one" step (5.6) and "two" step forms. Within the two-step procedures there are two classes:

a. Uncoupled. We (1) perform a conventional MCSCF calculation to obtain orbitals (section 3) and then with those orbitals (2) calculate an effective Hamiltonian (section 4) to obtain configuration coefficients. We repeat steps (1) and (2) until we reach self-consistency.

b. Coupled. We now set V + V′ = 0; then equation (5.6) becomes

$$\{T_{rs}\} = - \Big(\{C_{rs,r's'}\} + \{C_{rs,K'}\}\{C_{KK'}\}^{-1} \{C_{K,r's'}\}\Big)^{-1}\{(W+W')_{rs}\} \qquad (5.11)$$

Then we (1) obtain orbital coefficients from (5.11) and (2) obtain configuration coefficients from an effective Hamiltonian calculation. We repeat steps (1) and (2) until we reach self-consistency.

A self-consistent MCSCF effective Hamiltonian is a unique object and provides the optimum condensation of the quantum mechanical information content of the system which can be encapsulated within a given primary space and should provide a valid basis for chemical interpretation.

APPENDIX A - H_2O

We propose to test our procedures on H_2O (C_{2v} symmetry) using a double zeta basis and a 361 dimensional V_P composed of all single and double excitations (of $^1A_1$ symmetry) from the closed shell singlet ground state. We propose to partition the orbitals on the basis of their SCF energies and obtain an 11 dimensional V_P:

Basis: double zeta O(9s 5p/4s 2p) H(4s/2s)

SCF Ground State: $(1a_1)^2 (2a_1)^2 (3a_1)^2 (1b_2)^2 (1b_1)^2$; symmetry--singlet A_1.

Full Space: All single and double excitations.
Dimension = 361 configurations.
Symmetry = singlet A_1.

SCF Orbital Energies:

$1=1a_1$	-20.559	$3=1b_2$	-0.717	$8=2b_1$	0.866
$2=2a_1$	- 1.361	$4=3a_1$	-0.567	$9=6a_1$	0.891
		$5=1b_1$	-0.506	$10=2b_2$	0.915
		$6=8a_1$	0.218	$11=7a_1$	1.223
		$7=4b_2$	0.310	$12=3b_2$	1.233
				$13=5a_1$	1.675
				$14=4a_1$	43.335

Orbital Partitions:

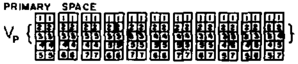
V_{core}: $\{|1\rangle,|2\rangle\}$

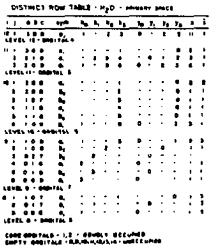
V_{mixed}: $\{|3\rangle,|4\rangle,|5\rangle,|6\rangle,|7\rangle\}$

V_{empty}: $\{|8\rangle,|9\rangle,|10\rangle,|11\rangle,|12\rangle,|13\rangle,|14\rangle\}$

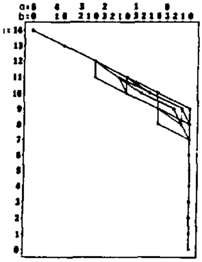
Configuration Space Partition:

V_P dimension = 11	V_Q dimension = 350
(Primary Space)	(Secondary Space)





SHAVITT GRAPH



APPENDIX B - THE PAIR-PRIMARY SPACE

The pair-primary space extends the number of active orbitals but is of smaller dimension than the corresponding complete active primary space. This space is based on the concept of geminals (see for example Linderberg or Goscinski[13]), and it may provide a better surface for a dissociating molecule or a transition state. The difference between the two formulations can be illustrated by a four-electron, four-orbital nonsymmetric space:

Orbital Partition

V_{mixed}: $\{|1\rangle,|2\rangle,|3\rangle\}$

V_{empty}: $\{|4\rangle\}$

Configuration Space Partition

Complete Active Space (CAS) - V_P dimension = 6

V_P(CAS):

Pair Space (PS) - V_P dimension = 3

V_P(PS):

The form of the PS-MCSCF equations is simpler than those for CAS-MCSCF. For PS calculations GUGA is no longer necessary, as the resulting MCSCF equations depend only on the weights (occupation numbers) of the Gel'fand states. (For details see Matsen and Nelin.[14])

The use of the pair-primary space is applicable to certain (symmetrical) symmetry states. The extension to non-singlet states is under consideration.

REFERENCES

1. See for example F.A. Matsen: Int. J. Quantum Chem. 58, 379 (1974), and 10, 525 (1976); Adv. Quantum Chem. 2, 223 (1978).

2. F.A. Matsen, Proceedings of 1979 Conference on the Unitary Group and Applications to Ab Initio Quantum Mechanics (in press).

3. See for example M. Moshinsky, J. Math. Phys. 4, 436 (1963); J. Paldus, J. Chem. Phys. 61, 5321 (1974); L.C. Biedenharn, J. Math. Phys. 4, 436 (1963).

4. See for example J. Paldus, in "Theoretical Chemistry: Advances and Perspectives," vol. 2, ed. H. Eyring and D.J. Henderson (Academic Press, New York, 1976), pp. 131-290; I. Shavitt Int. J. Quantum Chem. S11, 131 (1977) and S12, 5 (1978); B.R. Brooks and H.F. Schaefer, J. Chem. Phys. 70, 5092 (1979); B.O. Roos and P.E.M. Siegbahn, Int. J. Quantum Chem. 17, 485 (1980).

5. F.A. Matsen and C.J. Nelin, Int. J. Quantum Chem. 15, 751 (1979).

6. F.A. Matsen and C.J. Nelin, Int. J. Quantum Chem. (submitted for publication).

7. F.A. Matsen, Int. J. Quantum Chem. S13, 305 (1980).

8. B.O. Roos, P.R. Taylor, and P.E.M. Siegbahn, Chem. Phys. 48, 157 (1980); B.O. Roos, Int. J. Quantum Chem. (in press).

9. See for example E. Dalgaard and P. Jørgensen, J. Chem. Phys. 69, 3833 (1978); D.L. Yeager and P. Jørgensen, Mol. Phys. (submitted for publication); C.C.J. Roothaan, J. Detrich, and I.G. Hopper, Int. J. Quantum Chem. S13, 93 (1979); K. Ruedenberg, L.M. Cheung, and S.T. Elbert, Int. J. Quantum Chem. 16, 1069 (1979); R. Shepard and J. Simons (submitted for publication); D.L. Yeager and P. Jørgensen, J. Chem. Phys. 71, 755 (1979).

10. B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968), and 3, 247 (1969).

11. See for example P. Jørgensen, P. Albertson, and D.L. Yeager, J. Chem. Phys. (submitted for publication); R. Shepard and J. Simons (submitted for publication).

12. See for example M. Karplus and P. Rossky, J. Chem. Phys. (submitted for publication); J. Cizek, J. Chem. Phys. 45, 4256 (1966); K.R. Roby, Int. J. Quantum Chem. 6, 101 (1972) F. Harris, Int. J. Quantum Chem. S11, 403 (1977); H. Monkhurst, Int. J. Quantum Chem. S11, 421 (1977); J. Paldus and J. Cizek, Adv. Quantum Chem. 9, 105 (1975); I. Lindgren, Int. J. Quantum Chem. S12, 33 (1978); B.H. Brandow, Adv. Quantum Chem. 10, 187 (1977); B.H. Brandow, Int. J. Quantum Chem. 15, 207 (1979); R.J. Bartlett and D.M. Silver, Int. J. Quantum Chem. S9, 183 (1975); R.J. Bartlett and G.D. Purvis, Int. J. Quantum Chem. 14, 561 (1978); M.G. Sheppard, K.F. Freed, M. Herman, and D.L. Yeager, Chem. Phys. Letters, 61, 577 (1979).

13. J. Linderberg, Israeli J. of Chem. (in press); O. Goscinski, "Horizons of Quantum Chemistry," ed. K. Fukui and B. Pullman (D. Reidel, Holland, 1980), pp.

14. F.A. Matsen and C.J. Nelin (to be submitted).

STABILITY CONDITION OF THE MCHF ENERGY AND CHOICE OF THE STARTING VECTORS

B. LEVY
Laboratoire de chimie, ENSJF, 1, rue Maurice Arnoux ,
92 120 Montrouge, France.

An approximate stability condition of the MCHF energy is derived. It leads to a criterium for the choice of the starting vectors.

1. Second derivatives

If the matrix of the MO components in a fixed basis is written in an exponential form

$$C = \exp(X) ,$$

then the useful second derivatives of the total energy with respect to the X_{ij} 's matrix elements are[1-4]

$$\frac{1}{2}\left[\left(\frac{\partial}{\partial X_{\alpha\beta}} - \frac{\partial}{\partial X_{\beta\alpha}}\right)\left(\frac{\partial}{\partial X_{\gamma\delta}} - \frac{\partial}{\partial X_{\delta\gamma}}\right) E\right]_{X=0} = E''^{(dir)}_{\alpha\beta,\gamma\delta} + E''^{(ind)}_{\alpha\beta,\gamma\delta} \tag{1}$$

$$E''^{(dir)}_{\alpha\beta,\gamma\delta} = \langle \psi_0 | (\overset{*}{\beta}\alpha - \overset{*}{\alpha}\beta) H (\overset{*}{\gamma}\delta - \overset{*}{\delta}\gamma) | \psi_0 \rangle + \frac{1}{2} \langle \psi_0 | \left[(\overset{*}{\beta}\alpha - \overset{*}{\alpha}\beta), (\overset{*}{\gamma}\delta - \overset{*}{\delta}\gamma) \right]_+ H | \psi_0 \rangle$$

$$E''^{(ind)}_{\alpha\beta,\gamma\delta} = \sum_{K \neq 0} \frac{\langle \psi_0 | \left[(\overset{*}{\beta}\alpha - \overset{*}{\alpha}\beta), H \right] | \psi_K \rangle \langle \psi_K | \left[(\overset{*}{\delta}\gamma - \overset{*}{\gamma}\delta), H \right] | \psi_0 \rangle}{E_0 - E_K}$$

where $\alpha^*, \beta^*, \gamma^*, \delta^*$ are creation operators, $\alpha, \beta, \gamma, \delta$ are annihilation operators, $[\ ]_+$ means the anti-commutation, ψ_0 is the eigenfunction of the CI corresponding to the state under consideration and the ψ_K's $(K \neq 0)$ are the remaining eigenfunctions.

The first term, $E''^{(dir)}_{\alpha\beta,\gamma\delta}$, is obtained by differentiating the MO's alone and the second one comes from the CI coefficients.

2. Stability condition

It is seen on these expressions that for a diagonal second derivative ($\alpha = \gamma$, $\beta = \delta$) the second term is always negative. Therefore, a necessary condition for the stability of the energy is

$$E''^{(dir)}_{\alpha\beta,\alpha\beta} > |E''^{(ind)}_{\alpha\beta,\alpha\beta}|$$

We introduce now the following hypotheses

i) ψ_0 has a leading component Φ_0. This allows one to introduce the following notation
- i,j,k ... are the indices used for the core orbitals (never substituted)
- t,u,v ... for the active orbitals occupied in Φ_0
- t^*, u^*, v^*... for the active orbitals not occupied in Φ_0
- a,b,c ... for virtual orbitals

ii) The most important components of ψ_0 are obtained by closed shell double substitutions of type $t\bar{t} \to t^*\bar{t}^*$. Thus

$$\psi_0 \simeq a_0 \Phi_0 + \sum_t a_{tt^*} \Phi_{tt^*} \tag{2}$$

$$|a_0| \gg |a_{tt^*}|$$

iii) All the wavefunctions ψ_i have a leading component like ψ_0 (no near degeneracy).

It appears from these hypotheses that the contributions of the diagonal matrix elements of H to second derivatives of type $(\overset{*}{t}a , \overset{*}{t}a)$ are smaller than the the contributions of the off diagonal matrix elements of H that have a cofactor of type $a_0\, a_{tt^*}$.

Thus

$$E''^{(dir)}_{\overset{*}{t}a,\overset{*}{t}a} \simeq 2\, a_0\, a_{tt^*} (K_{at} - K_{tt^*})$$

$$E''^{(ind)}_{\overset{*}{t}a,\overset{*}{t}a} \simeq 4\, |(K_t)_{\overset{*}{t}a}|^2 / (E_0 - E_{tt^*})$$

where K is an exchange integral and K_t is the exchange operator for the orbital φ_t.

Then, using

$$a_{tt^*} \simeq K_{tt^*} / E_0 - E_{tt^*}$$

$$a_0 \simeq 1$$

we find for the stability condition

$$K_{tt^*}(K_{tt^*}-K_{at}) > 2\,|(K_t)_{ta}^*|^2 \qquad (3)$$

3. Discussion

The last above condition, Eq.(3) is satisfied if the orbital φ_a is sufficiently different from φ_{t^*}. Indeed, if φ_a looks like φ_{t^*} one has

$$K_{tt^*} \neq K_{at} = |(K_t)_{ta}^*|$$

and the condition is not satisfied.

Starting vectors satisfying Eq.(3) can be easily constructed : φ_{t^*} is obtained by using the same combinations of basis AO's as in φ_t (with appropriate changes of sign) and then projected in the space complementary to φ_t. Then φ_a is obtained by Schmidt orthogonalization with respect to φ_t and φ_{t^*}. An example of such a procedure is given in Table I for an MC HF calculation of CH_4 in a double ζ basis were ψ_0 is exactly given by Eq. (2). The first two lines give the SCF vectors of F_2 symmetry which were leading to a near unstability ($E''\sim 10^{-2}$) due to the contribution of $E''(ind)$. The last two lines give the starting vectors of F_2 symmetry obtained by projection as described just above : these vectors lead to convergence in 5 iterations with a single calculation of E'' (at the first iteration) and the final vectors are differing from these starting vectors in decimal places that are below the ones given in the Table.

		Carbon		Hydrogen	
		2p	2p'	1s	1s'
SCF MO's	φ_t	.25	.45	.23	.04
	φ_{t^*}	-.19	-1.01	-.14	1.02
MCHF MO's	φ_t	.25	.45	.23	.05
	φ_{t^*}	.56	.80	-.41	-.38

Table I. SCF and MCHF MO's on CH_4 in a double ζ basis.

It is concluded that

i) it is easy to obtain starting vectors satisfying Eq.(3) for which neglecting $E''(ind)$ will result in no hidden unstability and consequently will allow quadratic convergence ;

ii) The Eq.(3) also provides a theoretical explnation of the character of the MO's that are generally out coming from MC HF calculations.

References

1. B. Lévy, Int. J. Quant.Chem., 4 , 297 (1970)

2. B. Lévy, Chem. Phys. Lett., 18 , 59 (1973)

3. D.L. Yeager and P. Jörgensen, to be published

4. E. Dalgaard and P. Jörgensen, J. Chem. Phys., 69 , 3833 (1978)

RESPONSE FUNCTIONS ASSOCIATED WITH MULTICONFIGURATIONAL REFERENCE STATES

Esper Dalgaard
Department of Chemistry
Aarhus University
DK-8000 Aarhus C, Denmark

SUMMARY

The linear and quadratic response of an atomic or molecular system to an external perturbation is studied within the framework of time dependent multiconfigurational Hartree-Fock theory. A time dependent reference state is set up as $|\psi\rangle = \exp(i\Lambda)\exp(i\kappa)|0\rangle\exp(-i\varepsilon t)$, where Λ and κ are Hermitian operators, which generate unitary transformations of orbitals and expansion coefficients, respectively, in the MCHF ground state, $|0\rangle = \Sigma|\phi_g\rangle C_{go}$. When the variation in time of Λ and κ is derived from Frenkel's variation principle, the response functions are consistent with Ehrenfest's theorem $\frac{d}{dt}\langle\psi|O|\psi\rangle = \langle\psi|\partial O/\partial t|\psi\rangle - \langle\psi|[O,H]|\psi\rangle$ for arbitrary one electron operators. Due to this feature the dipole length and velocity expressions for the oscillator strengths of electronic transitions will give the same result apart from basis set truncation errors. The emphasis is on linear response, but quadratic response functions are defined and a method for their calculation within MCHF theory is described.

I. INTRODUCTION

Descriptions of the dynamic response properties of many electron systems in terms of the time dependent Hartree- and Hartree-Fock theories occurred in the literature almost as early in the history of quantum mechanics as the corresponding ground state approximations.[1] Over the years, the random phase approximation, which is equivalent to the time dependent Hartree-Fock approach, has remained a popular tool for estimates of the frequency dependent polarizabilities.[2] Formally, this method has the attractive features of gauge invariance[3] and variational derivation[4] from a bounded functional, and for closed shell systems for which the Hartree-Fock state is stable, the results are generally quite satisfactory for the lowest lying transitions and the corresponding oscillator strengths.

Deficiencies in the time dependent Hartree-Fock approximation are numerous, however, particularly with regard to the continuous part of photoabsorption spectra[5] and to the lacking self-consistency of the polarization propagator.[6,7] The method fails completely if the Hartree-Fock state is unstable[8], and this is probably the most serious problem in the molecular applications. Naturally, many improvements have been developed to include correlation effects. Most of these treat the electron interaction in a perturbative manner.[9]

The purpose of this contribution is to report on a recently formulated generalization of the time dependent Hartree-Fock approximation, which is based on a multiconfigurational representation of ground state[10,11]:

$$|0\rangle = \sum_{g=1}^{N} |\phi_g\rangle C_{go} \tag{1}$$

where $|\phi_g\rangle$ denotes a single determinant or a symmetry projected single determinant. Improved optimization procedures[12,13] as well as encouraging numerical results from relatively large scale MCHF calculations[14] of static molecular properties indicate that dynamic response functions might be calculated effectively in terms of the variables of MCHF theory. Initial numerical results seem to justify such expectations.[10]

Our aim is to analyze the response of the system to an adiabatically switched on perturbation of the form ($\eta > 0$):

$$W(t) = V\exp(-i[\omega+i\eta]t) + V^{+}\exp(-i[-\omega+i\eta]t) \tag{2}$$

As an ansatz for a time dependent reference state we introduce the expression

$$|\psi\rangle = U|0\rangle\exp(-i\varepsilon t) \tag{3}$$

The state vector $|\psi\rangle$ will also be of the multiconfigurational form if the unitary operator U is constructed as

$$U = \exp(i\Lambda)\exp(i\kappa) \tag{4}$$

where Λ and κ are Hermitian operators, which generate transformations of orbitals and expansion coefficients C_{go} respectively:

$$\exp(i\kappa)|0\rangle = \sum|\phi_g\rangle\langle\phi_g|\exp(i\kappa)|0\rangle \tag{5}$$

$$\exp(i\Lambda)|\phi_g\rangle = |\tilde{\phi}_g(t)\rangle \tag{6}$$

Explicit expressions for Λ and κ will be given later. Presently, we take $\Lambda = \Lambda(t)$, $\kappa = \kappa(t)$, and $\varepsilon = \varepsilon(t)$ to be the basic variables of the theory. Frenkel's variation principle as modified by Langhoff, Epstein, and Karplus,[15]

$$\mathrm{Re}\,\langle\delta\psi|i\frac{d}{dt} - H - W|\psi\rangle = 0 \tag{7}$$

yields the following requirement on the operator U,

$$2\,\mathrm{Im}\langle\delta U^{+}\dot{U}\rangle + \delta\langle U^{+}(H+W)U\rangle = 0 \tag{8}$$

The phase function $\varepsilon(t)$ shall not concern us here since it does not enter the calculation of response functions. Equation (8) does not in general correspond to a minimization of a positive semi-definite measure of the error introduced by the ansatz for $|\psi\rangle$,[4] but this variation method is convenient to employ in practice. It reduces to the ordinary Rayleigh-Ritz variation principle in the static limit, an aspect which ensures that the dynamic polarizabilities will equal the static polarizabilities at zero-frequency as obtained from coupled MCHF calculations.[13,16]

In the next section we discuss the formal spectral representations of linear and quadratic response functions, while the remaining parts of the paper is devoted to their calculations in the MCHF framework.

II. ELEMENTS OF RESPONSE THEORY

Response theory is a standard topic in quantum mechanics and statistical mechanics. An account on the subject, which is particularly useful for our purpose, has been written by Zubarev.[17] Assume for the moment that $|0\rangle$ is the exact unperturbed reference state and $|\psi\rangle$ the corresponding solution to the time dependent Schrödinger equation in the presence of the perturbation given by equation (2). The result, which we shall need is then that the expectation value at time t of an arbitrary Hermitian operator R takes the form

$$\langle\psi|R|\psi\rangle - \langle 0|R|0\rangle = \tag{9}$$

$$2\,\mathrm{Re}\int ds \langle\langle R(t);\, V(s)\rangle\rangle \exp(-iEs)$$

$$+2\,\mathrm{Re}\iint ds\,ds' \langle\langle R(t);\, V(s);\, V(s')\rangle\rangle \exp(-iE(s+s'))$$

$$+2\,\mathrm{Re}\iint ds\,ds' \langle\langle R(t);\, V(s);\, V^{+}(s')\rangle\rangle \exp(-iEs-iE's')$$

$$+\ldots$$

where the integrations over the time variables s, s' are from $-\infty$ to $+\infty$. Several definitions have been introduced here. First, R(t) and V(s) are operators in the interaction representation, i.e. $R(t) = \exp(iHt)\, R \exp(-iHt)$, and $E = \omega+i\eta$ while $E' = -\omega+i\eta$. The quantities in the double brackets are the linear and quadratic response functions or propagators in terms of time variables.[17,18] Zubarev defines the retarded two-time propagators or Green's functions as

$$\langle\langle R(t);\, V(s)\rangle\rangle = -i\langle 0|[R(t),\, V(s)]|0\rangle\Theta(t-s) \tag{10}$$

and it seems natural to extend this concept to include quadratic response functions. Thus, we define the three time propagators as

$$\langle\langle R(t);\, V(s);\, V^{+}(s')\rangle\rangle = \tag{11}$$

$$-\langle 0|[[R(t),V(s)],V^{+}(s')]|0\rangle\Theta(t-s)\Theta(s-s')$$

In equations (10) and (11), Θ denotes the Heaviside step function, $\Theta(\tau) = 1$ for $\tau > 0$ and zero for $\tau < 0$. It is generally most convenient to employ the Fourier transforms of the propagators[18], and since these are functions of time differences only, the transforms are obtained as

$$\langle\langle R;V\rangle\rangle_{E} = \tag{12}$$

$$\int d\tau \exp(iE\tau)\langle\langle R(0);V(-\tau)\rangle\rangle$$

and

$$\langle\langle R;V;V^{+}\rangle\rangle_{EE'} = \tag{13}$$

$$\iint d\tau\, d\tau' \exp(iE\tau+iE'\tau')\langle\langle R(0);V(-\tau);V^{+}(-\tau')\rangle\rangle$$

The functions $\langle\langle ;\rangle\rangle_{E}$ and $\langle\langle ;\ ;\rangle\rangle_{EE'}$ are analytic in the upper half of the complex plane. Introducing formally a complete set of states $\{|m\rangle\}$, which are eigenstates of H, we find that the spectral representations may be written

$$\langle\langle R;V\rangle\rangle_{E} = \tag{14}$$

$$\sum \left\{\frac{\langle 0|R|m\rangle\langle m|V|0\rangle}{E-E_{m}+E_{0}} - \frac{\langle 0|V|m\rangle\langle m|R|0\rangle}{E+E_{m}-E_{0}}\right\}$$

and

$$\langle\langle R;V;V^{+}\rangle\rangle_{EE'} = \tag{15}$$

$$\sum \frac{\langle 0|R|m\rangle\langle m|V|m'\rangle\langle m'|V^{+}|0\rangle}{(E+E'-E_{m}+E_{0})(E'-E_{m'}+E_{0})}$$

$$+\sum \frac{\langle 0|V^{+}|m\rangle\langle m|V|m'\rangle\langle m'|R|0\rangle}{(E+E'+E_{m'}-E_{0})(E'+E_{m}-E_{0})}$$

$$-\sum \frac{\langle 0|V|m\rangle\langle m|R|m'\rangle\langle m'|V^{+}|0\rangle}{(E+E'+E_{m}-E_{m'})(E'-E_{m'}+E_{0})}$$

$$-\sum \frac{\langle 0|V^{+}|m\rangle\langle m|R|m'\rangle\langle m'|V|0\rangle}{(E+E'+E_{m}-E_{m'})(E'+E_{m}-E_{0})}$$

Somewhat simpler expressions can be established for the special cases $\langle\langle R,V,V\rangle\rangle$ and $\langle\langle V,V,V\rangle\rangle$; these will not be delt with here.

It is apparent that the poles and residues of the linear response functions provide transition energies and transition moments for transitions involving the ground state directly. In addition, the quadratic response functions furnish us with transition amplitudes between two excited states. Such amplitudes are needed for the description of two-photon processes[19], which are of much interest within the field of laser spectroscopy, and also for calculations of radiative lifetimes, except for the lowest excited state[20]. There is therefore ample motivation for extending current efforts in the theory of propagators to include the quadratic ones.

Most approximate calculations of propagators have been based on the equation of motion[18,20]

$$E\langle\langle R;V\rangle\rangle_{E} = \tag{16}$$

$$\langle 0|[R,V]|0\rangle + \langle\langle [R,H];V\rangle\rangle_{E}$$

A similar equation connects the linear and quadratic response functions

$$(E+E')\langle\langle R;V;V^{+}\rangle\rangle_{EE'} = \tag{17}$$

$$\langle\langle [R,V];V^{+}\rangle\rangle_{E'} + \langle\langle [R,H];V;V^{+}\rangle\rangle_{EE'}$$

Formally, we may express the connection between the two types of propagators as a moment expansion

$$<<R;V;V^{+}>>_{EE'} = \sum_{n=0}^{\infty} \frac{<<A_n;V^{+}>>_{E'}}{(E+E')^{n+1}} \tag{18}$$

where $A_n = (-)^n[(\hat{H})^n R,V]$. Here $\hat{H}$ denotes the superoperator which on an arbitrary operator has the effect $\hat{H}X = [H,X]$. Equation (18) suggests approximation schemes for the quadratic response functions in terms of Padé approximants[21], but this aspect will not be pursued here.

Introducing the Fourier transforms of the propagators into equation (9), we find that

$$<\psi|R|\psi> = <0|R|0> \tag{19}$$

$$+ 2\,\mathrm{Re}[<<R;V>>_E \exp(-iEt)$$

$$+ <<R;V;V>>_{EE} \exp(-i2Et)$$

$$+ <<R;V;V^{+}>>_{EE'} \exp(2\eta t)] + \ldots$$

where $E = \omega+i\eta$ and $E' = -\omega+i\eta$.

In the time dependent multiconfigurational Hartree-Fock formalism, the response functions will be expressed in terms of the operators Λ and κ in equation (4) as calculated from equation (8). As shown previously[11], a perturbation expansion of the type

$$\Lambda = \Lambda(1) + \Lambda(2) + \ldots \tag{20}$$

$$\kappa = \kappa(1) + \kappa(2) + \ldots$$

derives from equation (8) only if the unperturbed reference state $|0>$ is fully optimized so that the generalized Brillouin theorem

$$<0|[\delta\Lambda+\delta\kappa,H]|0> = 0 \tag{21}$$

is valid. Then we find from equations (3) and (4) that

$$<\psi|R|\psi> = <0|R|0> \tag{22}$$

$$+ i<0|[R,\Lambda(1)+\kappa(1)]|0>$$

$$+ i<0|[R,\Lambda(2)+\kappa(2)]|0>$$

$$- <0|[[R,\Lambda(1)],\kappa(1)]|0>$$

$$- \tfrac{1}{2}<0|[[R,\Lambda(1)],\Lambda(1)]|0>$$

$$- \tfrac{1}{2}<0|[[R,\kappa(1)],\kappa(1)]|0>$$

$$+ \ldots$$

The next section is devoted to the equations from which Λ and κ are calculated. The resulting MCHF propagators are in subsequent sections identified by a comparison of equations (19) and (22).

III. TIME DEPENDENT MCHF EQUATIONS

Consider now again equations (3), (4), and (8). The objective of this section is to deduce the perturbation expansion indicated by equation (20) from equation (8). We assume that our unperturbed reference state has been fully optimized as discussed by professor Yeager at this symposium.

Following a suggestion by Linderberg and Öhrn[7], we first express the operators Λ and κ as linear combinations of Hermitian operators Λ_j and κ_j,

$$\Lambda = \sum a_j\Lambda_j \quad ; \kappa = \sum b_j\kappa_j \tag{23}$$

with real coefficients a_j and b_j. For orbital transformations we need the set

$$\{\Lambda_j\} = \{(a_r^{+}a_s+a_s^{+}a_r);\ i(a_r^{+}a_s-a_s^{+}a_r)\} \tag{24}$$

where a_r^{+} denote electron creation operators referring to an orthonormal spin orbital basis set $\{\phi_r(\xi)|r=1,\ldots,M\}$. Transformations of expansion coefficients $\{C_{go}\}$ in equation (1) are generated by the elements of the set

$$\{\kappa_j\} = \{|\phi_g><\phi_{g'}|+|\phi_{g'}><\phi_g|;i|\phi_g><\phi_{g'}|-i|\phi_{g'}><\phi_g|\} \tag{25}$$

Detailed descriptions of how to carry out these transformations in practice have been provided previously.[12,13]

A general infinitesimal change in the unitary operator U is given by[7]

$$\delta U = i(\chi+\tilde{K})U \tag{26}$$

where χ and $\tilde{K}$ are given by the equations

$$\chi = \int_0^1 ds\ \exp(is\Lambda)\delta\Lambda\exp(-is\Lambda) \tag{27}$$

$$K = \int_0^1 ds\ \exp(is\kappa)\delta\kappa\exp(-is\kappa) \tag{28}$$

and

$$\tilde{K} = \exp(i\Lambda)K\ \exp(-is\Lambda) \tag{29}$$

The operator $\tilde{K}$ has the same form as K but is expressed in terms of the transformed orbitals. It will prove convenient shortly to employ variations $\delta\Lambda$ such that $\chi = \Lambda_j$. In order to find such variations, we first define the superoperator $\hat{\Lambda}X = [\Lambda,X]$ so that equation (27) reads

$$\chi = \int_0^1 ds\ \exp(is\hat{\Lambda})\delta\Lambda \tag{30}$$

$$= \sum_{n=0}^{\infty} \frac{(i)^n}{(n+1)!} (\hat{\Lambda})^n\ \delta\Lambda$$

The numbers $(i)^n/(n+1)!$ are the coefficients for the Maclaurin expansion of the function $f(x) = (\exp(ix)-1)/ix$. Let d_n denote the Maclaurin coefficients for the reciprocal function $1/f(x)$ and take $\delta\Lambda$ to be

$$\delta\Lambda = \sum_{n=0}^{\infty} d_n(\hat{\Lambda})^n\Lambda_j \tag{31}$$

This sum is convergent if all eigenvalues, Θ_j, of $\hat{\Lambda}$ lie in the interval $0 \leq \Theta_j < 2\pi$, a requirement which can always be fulfilled since those eigenvalues in connection with orbital transformations always occur as $\exp(i\Theta_j)$. For this choice of $\delta\Lambda$ we obtain $\chi = \Lambda_j$ as desired. Similarly, we may choose $\delta\kappa$ such that $\tilde{K} = \tilde{\kappa}_j = \exp(i\Lambda)\kappa_j\ \exp(-i\Lambda)$.

Employing equations (26), ..., (31) in equation (8) we conclude that Frenkel's variation principle in the present context becomes equivalent to the equations

$$\langle\psi|\Lambda_j|\dot{\psi}\rangle + \langle\dot{\psi}|\Lambda_j|\psi\rangle = \tag{32}$$

$$-i\langle\psi|[\Lambda_j,H+W]|\psi\rangle \qquad \text{(all j)}$$

$$\langle\psi|\tilde{\kappa}_j|\psi\rangle + \langle\psi|\tilde{\kappa}_j|\psi\rangle = \tag{33}$$

$$-i\langle\psi|[\tilde{\kappa}_j,H+W]|\psi\rangle \qquad \text{(all j)}$$

An interesting observation can be made here, namely that when equations (32) and (33) have been solved for Λ and κ then the generalized Ehrenfest's theorem[22]

$$\frac{d}{dt}\langle\psi|O|\psi\rangle = \langle\psi|\frac{\partial O}{\partial t}|\psi\rangle - i\langle\psi|[O,H+W]|\psi\rangle \tag{34}$$

will hold for any operator, which is a linear combination of the generators Λ_j and $\tilde{\kappa}_j$. This result will be used in the next section to demonstrate the equivalence of the dipole length and velocity expressions for oscillator strengths within time dependent MCHF theory.

So far, we have assumed that the set $\{\Lambda_j\}$ contains all the M^2 generators, $a_r^+a_s$, of the unitary group U(M) of all possible unitary transformations of the spin orbital basis set, but this set is often highly redundant. If there is a subset of generators $\{\Lambda_j'\}\subset\{\Lambda_j\}$ which are generators of a subgroup of U(M) and if

$$\Lambda_j'|\phi_g\rangle = \sum_{g'=1}^{N} |\phi_{g'}\rangle\langle\phi_{g'}|\Lambda_j'|\phi_g\rangle \tag{35}$$

then this subset, $\{\Lambda_j'\}$, may without loss of generality be omitted in the representation of Λ as given by equation (23). Such generators provide no additional flexibility in the ansatz for $|\psi\rangle$. The set $\{\kappa_j\}$ given by equation (25) is also redundant. A nonredundant set of generators for unitary transformations of expansion coefficients, C_{go}, in equation (1) is conveniently expressed in terms of a set of states $\{|n\rangle\}$ which spans the orthogonal complement of the reference state $|0\rangle$ within the subspace spanned by the manifold $\{|\phi_g\rangle$, g = 1,..., N}. This set may be chosen to be

$$\{\kappa_j\} = \{|n\rangle\langle 0| + |0\rangle\langle n|; i|n\rangle\langle 0| - i|0\rangle\langle n|\} \tag{36}$$

Elimination of redundant variables is essential for the following, since otherwise equations (32) and (33) would not provide a unique solution for the variables $\{a_i,b_j\}$ which are included in our ansatz for $|\psi\rangle$. For a more detailed description of the identification of the redundant variables the reader is referred to recent publications.[11,12,13]

Finally, we are prepared to set up equation systems for the calculation of linear and quadratic response properties. Using equation (20) in (32) and (33), and collecting terms of the same order in the external perturbation, we obtain a set of first order equations as

$$i\langle[\Lambda_j,\dot{\Lambda}(1)+\dot{\kappa}(1)]\rangle - \langle[[\Lambda_j,H],\Lambda(1)+\kappa(1)]\rangle =$$

$$-i\langle[\Lambda_j,W]\rangle \tag{37}$$

$$i\langle[\kappa_j,\dot{\Lambda}(1)+\dot{\kappa}(1)]\rangle - \langle[\kappa_j,[H,\Lambda(1)+\kappa(1)]]\rangle$$

$$= -i\langle[\kappa_j,W]\rangle$$

while the second order equations become

$$i\langle[\Lambda_j,\dot{\Lambda}(2)+\dot{\kappa}(2)]\rangle - \langle[[\Lambda_j,H],\Lambda(2)+\kappa(2)]\rangle = \tag{38}$$

$$\langle[[\Lambda_j,W],\ \Lambda(1)+\kappa(1)]\rangle$$

$$+ \langle(\Lambda_j,\Lambda(1),\dot{\Lambda}(1)) + (\Lambda_j,\kappa(1),\dot{\kappa}(1))\rangle$$

$$+ \langle[[\Lambda_j,\dot{\Lambda}(1)],\kappa(1)] + [[\Lambda_j,\Lambda(1)],\ \dot{\kappa}(1)]\rangle$$

$$+ \frac{i}{2}\langle[[[\Lambda_j,H],\Lambda(1)],\Lambda(1) + 2\kappa(1)]\rangle$$

$$+ \frac{i}{2}\langle[[[\Lambda_j,H],\kappa(1)],\kappa(1)]\rangle$$

and

$$i\langle[\kappa_j,\dot{\Lambda}(2)+\dot{\kappa}(2)]\rangle - \langle[\kappa_j,[H,\Lambda(2)+\kappa(2)]]\rangle = \tag{39}$$

$$\langle[\kappa_j,[W,\Lambda(1)+\kappa(1)]]\rangle + \langle[[\kappa_j,\dot{\Lambda}(1)],\kappa(1)]\rangle$$

$$+ \tfrac{1}{2}\langle[\kappa_j,[\dot{\Lambda}(1),\Lambda(1)]]\rangle$$

$$+ \frac{i}{2}\langle[\kappa_j,[[H,\Lambda(1)],\Lambda(1)+2\kappa(1)]\rangle$$

In these equations a "." is used to designate the derivative with respect to time and we have introduced the abbreviation

$$(\Lambda_j,\Lambda,\dot{\Lambda}) = \tfrac{1}{2}[[\Lambda_j,\Lambda],\dot{\Lambda}] + \tfrac{1}{2}[[\Lambda_j,\dot{\Lambda}],\Lambda]$$

All expectation values are with respect to the unperturbed reference state $|0\rangle$.

The non-linear orbital relaxation effects which are included in the time dependent MCHF formalism appear explicitly in the equations for the quadratic or second order part of the operators Λ and κ. These terms, i.e. the last four parts of equation (38) and the last three parts of equations (39) are spurious in the sense that they would vanish identically, if Λ and κ were linear combinations of true excitation operators. A detailed analysis of those terms will be published elsewhere.

Two special cases of equations (38) and (39) may be of interest. If the operator Λ everywhere is replaced by the unit operator, the present method becomes equivalent to the configuration interaction approach and the non-linear response terms are obtained from

$$i\langle[\kappa_j,\dot{\kappa}(2)]\rangle - \langle[\kappa_j,[H,\kappa(2)]]\rangle = \tag{40}$$

$$\langle[\kappa_j,[W,\kappa(1)]]\rangle$$

On the other hand, if κ is replaced by the unit operator and the reference state $|0\rangle$ is chosen to be the Hartree-Fock state then the random phase approximation is recovered. In this case equation (38) becomes

$$i\langle[\Lambda_j,\dot{\Lambda}(2)]\rangle - \langle[[\Lambda_j,H],\Lambda(2)]\rangle = \tag{41}$$

$$\langle[[\Lambda_j,W],\Lambda(1)] \ + \frac{i}{2}\langle[[[\Lambda_j,H],\Lambda(1)],\Lambda(1)]\rangle$$

since the expectation value of the product of any three particle-hole excitations or deexcitations will vanish. Equation (41) is equivalent to the

the non-linear time dependent Hartree-Fock equations recently described by Tillieu and Groenendael[23], using effective one particle operators and Lagrangian multipliers.

IV. MCHF RESPONSE FUNCTIONS

Turning now to the problem of solving equations (37), (38), and (39) for Λ and κ we first observe that a unique solution requires a boundary condition. Since the perturbation operator vanishes when t tends to minus infinity the appropriate boundary condition is

$$\lim_{t\to-\infty} |\psi\rangle\langle\psi| = |0\rangle\langle 0| \qquad (42)$$

or equivalently

$$\lim_{t\to-\infty} \Lambda = \lim_{t\to-\infty} \kappa = 0 \qquad (43)$$

The form of the perturbation operator W(t) and the linear nature of equations (37) allow us to conclude that the time dependence of the first order operators $\Lambda(1)$ and $\kappa(1)$ will be

$$\Lambda(1) = \alpha(1;E)\ \exp(-iEt) + \alpha^{+}(1;E)\exp(-iE't) \qquad (44)$$

$$\kappa(1) = \beta(1;E)\ \exp(-iEt) + \beta^{+}(1;E)\exp(-iE't)$$

where $E = \omega+i\eta$ and $E' = -\omega+i\eta$. The following expansions apply for the frequency dependent operators α and β

$$\alpha(1;E) = \sum_j a_j^{1}(E)\Lambda_j \qquad (45)$$

$$\beta(1;E) = \sum_j b_j^{1}(E)\kappa_j$$

We find then that equations (37) become replaced by the requirement

$$E\langle[\Lambda_j,\alpha(1)+\beta(1)]\rangle-\langle[[\Lambda_j,H],\alpha(1)+\beta(1)]\rangle = \qquad (46)$$

$$- i\langle[\Lambda_j,V]\rangle$$

$$E\langle[\kappa_j,\alpha(1)+\beta(1)]\rangle-\langle[\kappa_j,[H,\alpha(1)+\beta(1)]]\rangle = \qquad (47)$$

$$- i\langle[\kappa_j,V]\rangle$$

When the expansions for $\alpha(1)$ and $\beta(1)$ in terms of the generators Λ_j and κ_j are introduced in (46) and (47), a linear system of equations is obtained for the variables $a_j^{1}(E)$ and $b_j^{1}(E)$. Standard methods may be employed to invert the coefficient matrix for this system for each energy value, but if the spectral representation of the response functions is desired, it is most convenient to employ a set of matrix transformations suggested by Linderberg and Öhrn for a generalized random phase approximation[7].

We assume in the following that the operators $\alpha(1;E)$ and $\beta(1;E)$ have been determined. A compact formula for the MCHF linear response functions can then be inferred from equations (19) and (22):

$$\langle\langle R;V\rangle\rangle_E = i\langle[R,\alpha(1;E)+\beta(1;E)]\rangle \qquad (48)$$

Equation (46) may be viewed as a direct consequence of the generalized Ehrenfest's theorem, equation (34), which in the present scheme is valid for all one-electron operators (even if some of the generators $a_r^{+}a_s$, are redundant and therefore omitted in the ansatz for $|\psi\rangle$). Thus, it holds that[11]

$$E\langle\langle R;V\rangle\rangle_E - \langle\langle[R,H];V\rangle\rangle_E = \langle[R,V]\rangle \qquad (49)$$

if R is a one-electron operator. Suppose now that an operator P is related to R through the Heisenberg equation

$$iP = [R,H] \qquad (50)$$

as, e.g., the dipole moment and linear momentum operators. Since

$$\langle\langle[R,H];V\rangle\rangle_E = \langle\langle R;[H,V]\rangle\rangle_E \qquad (51)$$

for the exact propagators, and also for the approximate ones obtained in the MCHF scheme, we have a sequence of identities

$$E^2\langle\langle R;R\rangle\rangle_E = iE\langle\langle P;R\rangle\rangle \qquad (52)$$

$$= \langle\langle P;P\rangle\rangle + i\langle[P,R]\rangle$$

According to equation (14) the linear response have poles of first order at transition energies $\omega_J = E_J-E_0$. The oscillator strength for the transition $|0\rangle\to|J\rangle$ is in the dipole length formulation

$$f(L;0\to J) = \frac{2}{3}(E_J-E_0)|\langle 0|\underline{R}|J\rangle|^2 \qquad (53)$$

while the dipole velocity formulation reads

$$f(V;0\to J) = \frac{2}{3}(E_J-E_0)^{-1}|\langle 0|\underline{P}|J\rangle|^2 \qquad (54)$$

Atomic units are used here, and $\underline{R}$ and $\underline{P}$ denote the dipole moment and linear momentum vector operators, respectively. The residue at ω_J of the propagator

$$Sp\langle\langle\underline{R};\underline{R}\rangle\rangle_E = \qquad (55)$$

$$\langle\langle R_x;R_x\rangle\rangle_E + \langle\langle R_y;R_y\rangle\rangle_E + \langle\langle R_z;R_z\rangle\rangle_E$$

equals the square of the transition moment $\langle 0|\underline{R}|J\rangle$. Therefore we see that

$$f(L;0\to J) = \frac{2}{3}\omega_J Res[Sp\langle\langle\underline{R};\underline{R}\rangle\rangle_E] \qquad (56)$$

$$= \frac{2}{3}\,\omega_J^{-1}Res[E^2 Sp\langle\langle\underline{R};\underline{R}\rangle\rangle_E]$$

$$= \frac{2}{3}\,\omega_J^{-1}Res[Sp\langle\langle\underline{P};\underline{P}\rangle\rangle_E] \equiv f(V;0\to J)$$

All residues above are evaluated at the pole ω_J. It follows that apart from a possible violation of equation (50) due to the finite orbital basis the two expressions for the oscillator strengths are equivalent in MCHF theory. In fact, all the various formulations are equivalent in the present scheme[24] and the Thomas-Reiche-Kuhn sum rule $\Sigma f(0\to J) = \langle N_{op}\rangle$ holds as well[11]. These features of time dependent MCHF are attractive, but it should be pointed out that Kobe[25] recently has demonstrated, on the basis of gauge invariance, that it is the length form of the interaction between charged particles and radiation, which must be used, when the unperturbed Hamiltonian is chosen in the usual manner.

Changing now the topic to the quadratic response functions, we consider again equations (38) and (39). The right hand sides of these equations contain the oscillatory factors $\exp(-i2Et)$, $\exp(-i(E+E')t)$, and $\exp(-i2E't)$, where as in equation (44) $E = \omega+i\eta$ and $E' = -\omega+i\eta$. The operators $\Lambda(2)$ and $\kappa(2)$ must therefore have the form

$$\Lambda(2) = \alpha(2;E)\exp(-i2Et) + \alpha^{+}(2;E)\exp(-i2E't)$$

$$+ \gamma(E,E')\exp(-i(E+E')t) \qquad (57)$$

and

$$\kappa(2) = \beta(2;E)\exp(-i2Et) + \beta^{+}(2;E)\exp(-i2E't) \qquad (58)$$

$$+ \mu(E,E')\exp(-i(E+E')t)$$

The operators $\alpha(2)$ and $\beta(2)$ should satisfy the equations

$$2E\langle[\Lambda_j,\alpha(2)+\beta(2)]\rangle - \langle[[\Lambda_j,H],\alpha(2)+\beta(2)]\rangle = F_j(E) \qquad (59)$$

$$2E\langle[\kappa_j,\alpha(2)+\beta(2)]\rangle - \langle[\kappa_j,[H,\alpha(2)+\beta(2)]]\rangle = G_j(E)$$

while the operators γ and μ should be a solution to the system

$$(E+E')\langle[\Lambda_j,\gamma+\mu]\rangle - \langle[[\Lambda_j,H], \gamma+\mu]\rangle = F_j(E,E') \qquad (60)$$

$$(E+E')\langle[\kappa_j,\gamma+\mu]\rangle - \langle[\kappa_j,[H,\gamma+\mu]]\rangle = G_j(E,E')$$

The four functionals $F_j(E)$, $G_j(E)$, $F_j(E,E')$ and $G_j(E,E')$ are identified by introducing equations (2) and (44) in the right hand sides of (38) and (39). We shall not need the explicit expressions for these functionals in the present discussion but we note that they can relatively easily be calculated from the two electron transition density matrix elements $\langle\phi_g|a_r^{+}a_{r'}^{+}a_{s'}a_s|\phi_{g'}\rangle$ and the first order variables $a_j^{1}(E)$ and $b_j^{1}(E)$. When expansions analogous to those applied for $\alpha(1)$ and $\beta(1)$ are introduced in (59) and (60), we get equation systems of the same form as the first order equations and the same computer code may be used for their solution.

Finally, we compare equations (22) and (19) in order to obtain expressions for the quadratic response functions

$$\langle\langle R;V;V\rangle\rangle_{\omega+i\eta,\omega+i\eta} = i\langle[R,\alpha(2)+\beta(2)]\rangle \qquad (61)$$

$$- \tfrac{1}{2}\langle[[R,\alpha(1)],\alpha(1)+2\beta(1)]\rangle - \tfrac{1}{2}\langle[[R,\beta(1)],\beta(1)]\rangle$$

and

$$2\mathrm{Re}\langle\langle R;V;V^{+}\rangle\rangle_{\omega+i\eta,-\omega+i\eta} = \qquad (62)$$

$$\langle\langle R;V;V^{+}\rangle\rangle_{\omega+i\eta,-\omega+i\eta} + \langle\langle R;V^{+};V\rangle\rangle_{-\omega+i\eta,\omega+i\eta}$$

$$= i\langle[R,\gamma+\mu]\rangle - \langle[[R,\alpha(1)], \beta^{+}(1)]\rangle$$

$$-\langle[[R,\alpha^{+}(1)], \beta(1)]\rangle$$

$$- \langle(R,\alpha(1),\alpha^{+}(1))\rangle - \langle(R,\beta(1), \beta^{+}(1))\rangle$$

Equation (62) indicates that it may at times be more convenient to use an alternative to the forms (11) and (15), namely

$$G(RVV^{+};EE') = \tfrac{1}{2}\langle\langle R;V;V^{+}\rangle\rangle_{EE'} + \tfrac{1}{2}\langle\langle R;V^{+};V\rangle\rangle_{E'E} \qquad (63)$$

The definition (11) appears to be a natural extension of Zubarev's retarded Green's functions and it also seems to give the simplest spectral representation for a non-linear response function. On the other hand, the function $\langle\langle R;V;V^{+}\rangle\rangle_{EE'}$, has a singularity for $E+E' = E_{m'}-E_m$ where both m and m' denote excited states, while all such singularities are absent in the function $G(RVV^{+}; EE')$. A more complete description of the calculation of quadratic response functions within the time dependent MCHF theory will be published elsewhere.

V. DISCUSSION

Multiconfigurational Hartree-Fock methods have enjoyed an increasing popularity during the last decade as a natural means of including correlation effects, when describing the electronic structure of atoms and molecules[26]. The present symposium is a manifestation of this fact.

Several authors have found the MCHF method useful for the generation of an orbital basis set, which improves the convergence of a configuration interaction expansion. Such calculations are generally directed toward highly accurate descriptions of a few low lying states and to a lesser extent toward the dynamic response functions, albeit mainly for economic reasons. This study has been incited by the propagator methods as developed for quantum chemistry[18] and at the same time by the difficulties that those methods encounter, when an explicit ground state representative is abandoned.[6,7] I would like to advocate the view that well defined propagator approximations emerge in a natural way within the time dependent MCHF framework. Recent numerical results for oscillator strengths and excitation energies for the O_2 molecule obtained by Albertsen, Yeager and Jørgensen[27] support this view. These authors employ an MCHF propagator formalism, which is equivalent to the present formulation for the linear response functions[10].

So far, no computations have been made for the quadratic response functions as defined here. However, the rapid development of laser spectroscopy in the non-linear domain calls for theoretical descriptions of second order processes. It seems that the functions $\langle\langle R;V;V^{+}\rangle\rangle_{EE'}$, or $G(RVV^{+};EE')$ given by equations (11) and (63), respectively, will be useful tools for such descriptions.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to the organizers of this workshop on MCHF methods, and in particular to professor D. Yeager, for making my participation possible and enjoyable.

REFERENCES

1. D.R. Hartree, Proc. Camb. Phil. Soc. 24, 89 (1927); P.A.M. Dirac, Proc. Camb. Phil. Soc. 26, 376 (1930).

2. P. Jørgensen, Ann. Rev. Phys. Chem. 26, 359 (1975), and references therein.

3. G. Baym and L. Kadanoff, Phys. Rev. 124, 287 (1961).

4. A.D. McLachlan, Mol. Phys. 8, 39 (1964);

E. Dalgaard, Int. J. Quantum Chem. 17, 449 (1980).

5. G. Wendin, J. Phys. B: Atom. Molec. Phys. 6, 42 (1973).

6. J. Linderberg, P. Jørgensen, J. Oddershede and M. Ratner, J. Chem. Phys. 56, 6213 (1972).

7. J. Linderberg and Y. Öhrn, Int. J. Quantum Chem. 12, 161 (1977), 15, 343 (1979).

8. D.J. Thoules, The Quantum Mechanics of Many-Body Systems (Academic, New York, 1961); J. Cizek and J. Paldus, J. Chem. Phys. 47, 3976 (1967); P. Jørgensen and J. Linderberg, Int. J. Quantum Chem. 4, 587 (1970).

9. J. Oddershede and P. Jørgensen, J. Chem. Phys 66, 1541 (1977).

10. D.L. Yeager and P. Jørgensen, Chem. Phys. Letters, 65, 77 (1979).

11. E. Dalgaard, J. Chem. Phys. 72, 816 (1980).

12. D.J. Rowe and G. Rosensteel, Int. J. Theoret. Phys. 15, 501 (1976); J. Kendrick and I. H. Hillier, Chem. Phys. Letters 41, 283 (1976); E. Dalgaard, Chem. Phys. Letters 65, 559 (1979). C.C.J. Roothaan, J. Detrich and D.G. Hopper, Int. J. Quantum Chem. Symp. 13 , 93 (1979).

13. E. Dalgaard and P. Jørgensen, J. Chem. Phys. 69, 3833 (1978).

14. B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Physics 48, 157 (1980).

15. J. Frenkel, Wave Mechanics, Advanced General Theory (Clarendon, Oxford, 1935); P.W. Langhoff, S.T. Epstein, and M. Karplus, Rev. Mod. Phys. 44, 602 (1972).

16. M. Jaszunski and A.J. Sadlej, Int. J. Quantum Chem. 11, 233 (1977); W.J. Stevens and F.P. Billingsley II, Phys. Rev. 8A, 2235 (1973).

17. D.N. Zubarev, Nonequilibrium Statistical Thermodynamics (Consultants Bureau, Plenum, New York 1974) p. 161.

18. J. Linderberg and Y. Öhrn, Propagators in Quantum Chemistry (Academic, London, 1973).

19. See e.g. J.I. Steinfeld, Molecules and Radiation, An Introduction to Modern Molecular Spectroscopy (Harper and Row, New York, 1974) ch. 13.

20. J. Oddershede, Adv. Quantum Chem. 11, 275 (1978).

21. O. Goscinski and B. Lukman, Chem. Phys. Lett. 7, 573 (1970).

22. P. Ehrenfest, Z. Phys. 45, 455 (1927).

23. F. Smet, J. Tillieu, and A. van Groenendael, Int. J. Quantum Chem. 17, 531 (1980).

24. A.E. Hansen, Int. J. Quantum Chem. Symp. 4, 473 (1971).

25. D.H. Kobe, Int. J. Quantum Chem. Symp. 12, 73 (1978).

26. C.F. Fischer, The Hartree-Fock method for atoms (Wiley, New York, 1977); A. C. Wahl and G. Das, in Modern theoretical chemistry Vol. 3, ed. H. F. Schaefer III (Plenum, New York, 1977) ch. 3.

27. P. Albertsen, P. Jørgensen and D.L. Yeager, to be published in Int. J. Quantum Chem. Symp. (1980).

MCSCF REFERENCE STATES FOR RESPONSE FUNCTIONS

Prof. Jack Simons
Department of Chemistry
University of Utah
Salt Lake City, Utah 84112

GREEN'S FUNCTIONS

Until quite recently Green's function (GF) and other response function developments have been restricted to single configuration SCF or first-order Rayleigh-Schrödinger perturbation theory reference wavefunctions. Following the lead of Yeager and Jørgensen, David Chuljian and the author have been able to formulate one-particle GF theory for an MCSCF reference function $|0\rangle$. The choice of ionization operator manifold p^+, $p^+|n\rangle\langle 0|$, where $\{|n\rangle\}$ is the MC orthogonal complement space, gives rise to quite tractable expressions for the elements of the GF matrix. By choosing the MC space to include all single and double excitations out of one of the dominant configurations of $|0\rangle$, we are able to significantly simplify the transition density matrix elements which appear in our working equations. In fact, we have been able to link this MC based GF program to our unitary group MCSCF programs quite easily.

ENERGY GRADIENTS

The fact that a converged MCSCF wavefunction $|0\rangle$ obeys the Generalized Brillouin Theorem makes such wavefunctions especially attractive for use in calculating how the energy E varies when the molecule's nuclei are moved from $\vec{R}_a$ to $\vec{R}_a + \vec{\alpha}_a$. By collecting the 3N-6(5) nuclear displacement vectors $\vec{\alpha}_a$ (whose lengths are to be determined) into a single vector $\vec{\alpha}$, one can express the Hamiltonian at $\vec{R} + \vec{\alpha}$ in terms of orbitals centered at $\vec{R}$: $H_{\vec{\alpha}} = H + \vec{\alpha}\cdot\vec{H}^{(1)} + \vec{\alpha}\cdot\overleftrightarrow{H}^{(2)}\cdot\vec{\alpha}$. Here $\vec{H}^{(1)}$ contains both the Hellmann-Feynman factors $\sum_{i,j} \langle\phi_i|\frac{\partial}{\partial R_a}|\vec{r}-\vec{R}_a|^{-1}|\phi_j\rangle i^+j$ and the <u>atomic</u> basis function derivative terms $\frac{\partial}{\partial R_a}\langle\chi_a||\vec{r}-\vec{R}_b|^{-1}|\chi_c\rangle$. Likewise, $\overleftrightarrow{H}^{(2)}$ contains both kinds of second derivative terms, but neither $\vec{H}^{(1)}$ nor $\overleftrightarrow{H}^{(2)}$ contain any effects of MCSCF orbital variation or CI coefficient variation. These latter effects are treated via the exp(iλ) and exp(iS) operators. The resultant energy expression, when made stationary with respect to variations in λ and S, gives rise to the following expression (through second order in $\vec{\alpha}$): $E = E^0 + \vec{\alpha}\cdot\langle 0|\vec{H}^{(1)}|0\rangle + \vec{\alpha}\cdot\langle 0|\overleftrightarrow{H}^{(2)}|0\rangle\cdot\vec{\alpha} - \frac{1}{2}(\vec{\alpha}\cdot\underline{\vec{F}}, \vec{\alpha}\cdot\underline{\vec{G}})\,(\underline{A}-\underline{B})^{-1}\,(\underline{\vec{F}}\cdot\vec{\alpha}, \underline{\vec{G}}\cdot\vec{\alpha})$, where $\underline{A}-\underline{B}$ is the MCSCF Hessian matrix (at $\vec{R}$) and $\vec{F}$ and $\vec{G}$ are Generalized Brillouin-like matrix elements involving the Hamiltonian derivative $\vec{H}^{(1)}$: $\vec{F}_{ij} = \langle 0|[i^+j, \vec{H}^{(1)}]|0\rangle$, $\vec{G}_n = \langle 0|[|n\rangle\langle 0|, \vec{H}^{(1)}]|0\rangle$. This quadriatic form in $\vec{\alpha}$ can then be used to find stationary points where $\partial E/\partial\vec{\alpha} = 0$. It should be noticed that the orbital and CI expansion coefficients' responses to the nuclear displacement $\vec{\alpha}$ are computed in a coupled MCSCF method; we do not resort to using single configuration coupled SCF to estimate the orbital changes λ.

AB INITIO CALCULATION OF THE EFFECTIVE VALENCE SHELL HAMILTONIAN

Maurice G. Sheppard, Hosung Sun, and Karl F. Freed
The James Franck Institute and The Department of Chemistry
The University of Chicago, Chicago, Illinois 60637

ABSTRACT

The study of molecular electronic structure has been divided into two separate and often acrimonious branches, ab initio and semiempirical theories. We have shown that the clue to providing a rigorous theoretical basis for understanding and systematically improving semiempirical methods of electronic structure lies in the study of the structure and properties of the effective valence shell Hamiltonian, $\mathcal{H}^v$.[1] $\mathcal{H}^v$ can be calculated by purely ab initio methods, and its eigenvalues, associated with the valence states of interest, are exactly equivalent to those generated by the solution of the full molecular electronic Schrödinger equation. The results of our recent calculations of $\mathcal{H}^v$ are of interest both as a novel ab initio formalism and for the insights they provide into the hidden underlying assumptions of semiempirical theories. The pursuit of this work naturally leads into the use of multiconfigurational self-consistent-field (MCSCF) methods to provide input for $\mathcal{H}^v$ calculations.

INTRODUCTION

Given a set of orbitals which is partitioned into core, valence, and excited orbitals, it is possible to define an effective valence shell Hamiltonian, $\mathcal{H}^v$, which is an exact transformation of the original Schrödinger equation for the states of interest (the valence states).[2] The orbital space is assumed to be large enough to accurately describe the valence states which in zeroth order are represented in terms of linear superpositions of configuration functions containing a full core and a partially occupied set of valence orbitals. The resulting $\mathcal{H}^v$ has the following properties: (1) $\mathcal{H}^v$ contains explicit reference only to the designated valence shell orbitals. Nevertheless, the eigenvalues of $\mathcal{H}^v$ are exactly identical to the corresponding valence state energies (i.e., potential energy surfaces) which result from the solution of the full molecular electronic Schrödinger equation within the given orbital basis. (2) The exact eigenvalues of $\mathcal{H}^v$ are obtained from a full valence space configuration interaction (CI) calculation, so $\mathcal{H}^v$ is the quantity which is mimicked through parametrizations of the model Hamiltonians, $\mathcal{H}^v_M$, of semiempirical valence methods.[2] Many semiempirical $\mathcal{H}^v_M$'s are fit directly to a form which ignores the valence CI because of the added expense. (3) In addition to the usual one- and two- electron effective operators ($\mathcal{H}^v_i$ and $\mathcal{H}^v_{ij}$), $\mathcal{H}^v$ also contains three-, four-, ..., N_v-electron effective operators ($\mathcal{H}^v_{ijk}$, $\mathcal{H}^v_{ijkl}$, ...) which have no counterparts in semiempirical theories. Here N_v is the number of valence electrons. (4) $\mathcal{H}^v$ uses the same set of orbitals to describe <u>all</u> valence states of a system including <u>all</u> charge states.[3-5] The frozen nature of the valence orbitals in the calculation of $\mathcal{H}^v$ significantly differs from conventional ab initio methods where the orbitals vary with the state and charge of the system. However, the same integrals of the effective many-electron operators of $\mathcal{H}^v$ may be used for all of these valence states since $\mathcal{H}^v$ is formally exact. We have numerically verified this fact, as described below, by using one set of orbitals to calculate all valence excitation energies and ionization potentials of the fluorine atom from F^{+7} through F^- with an average deviation from experiment of 0.27 eV.

In this paper we present a survey of our $\mathcal{H}^v$ calculations for the fluorine atom and for the CH molecule preceded by a brief discussion of the theory of $\mathcal{H}^v$. A discussion is also given of several calculations which are planned or in progress that require the use of multiconfigurational self-consistent-field (MCSCF) methods to generate a reasonable set of valence orbitals. Since the calculations for fluorine and CH involve third order quasidegenerate perturbation theory we are able to address questions of convergence and to analyze the effects of three- and (the never-before-calculated) four-electron effective integrals. Our calculations for CH provide the first ab initio test of the fundamental transferability hypothesis of semiempirical methods.[6]

THEORY

Definitions

In order to define $\mathcal{H}^v$ for an N electron system, it is necessary to first describe the subspace of the full N-electron Hilbert space within which $\mathcal{H}^v$ exists. This subspace is called the valence space and is spanned by the set of all N-electron symmetry adapted Slater determinants which have all core orbitals fully occupied with $N_c = N - N_v$ core electrons and the remaining N_v valence electrons distributed amongst the valence shell orbitals in all unique ways. The basis of orthogonal complement space, or excited space, incorporates all other possible configurations. These are characterized by having at least one vacancy in a core orbital and/or at least one occupied excited orbital.

Emphasizing this partitioning of the full N-electron Hilbert space, the Schrödinger equation in matrix form is represented by

$$\left(\begin{array}{c|c} H_{PP} & H_{PQ} \\ \hline H_{QP} & H_{QQ} \end{array}\right) \left(\begin{array}{c} C_P \\ \hline C_Q \end{array}\right) = E \left(\begin{array}{c} C_P \\ \hline C_Q \end{array}\right) \tag{1}$$

where C_P and C_Q are coefficient vectors in the valence (P) and excited (Q) spaces, respectively. Using the well known transformation described by Löwdin [7] produces a representation of the exact $\mathcal{H}^v$ which is defined on the valence space and which has the matrix representation,

$$\mathcal{H}^v C_P = [H_{PP} + H_{PQ}(E\mathbf{1}_{QQ} - H_{QQ})^{-1} H_{QP}] C_P = E C_P . \tag{2}$$

Since all configurations in the valence space have a full core and no reference to any excited orbital, they are totally specified by their valence shell configuration. Therefore, $\mathcal{H}^v$ makes explicit reference only to valence orbitals. It can be shown that $\mathcal{H}^v$ of (2) has off-diagonal matrix elements between configurations which differ by more than two valence spin orbitals. [1-4] In operator language, these two observations mean that $\mathcal{H}^v$ must have the form

$$\begin{aligned}\mathcal{H}^v = E_c &+ \sum_{i=1}^{N_v} \mathcal{H}_i^v + \frac{1}{2!}\sum_{i=1}^{N_v}\sum_{j=1}^{N_v}{}' \mathcal{H}_{ij}^v \\ &+ \frac{1}{3!}\sum_{i=1}^{N_v}\sum_{j=1}^{N_v}{}'\sum_{k=1}^{N_v}{}'' \mathcal{H}_{ijk}^v \\ &+ \frac{1}{4!}\sum_{i=1}^{N_v}\sum_{j=1}^{N_v}{}'\sum_{k=1}^{N_v}{}''\sum_{l=1}^{N_v}{}''' \mathcal{H}_{ijkl}^v + \cdots\end{aligned} \tag{3}$$

where E_c is the correlated energy of the bare core, and $\mathcal{H}_i^v$, $\mathcal{H}_{ij}^v$, $\mathcal{H}_{ijk}^v$, ..., are effective one-, two-, three-, ... up to N_v-electron operators.

Relationship to Semiempirical Model Hamiltonians

The similarity between equation (2) and the general form of the complete valence shell CI interpretation of semiempirical methods is striking. The general idea of these methods is to postulate the existence of a model Hamiltonian, $\mathcal{H}_M^v$, which is defined on the valence space and has eigenvalues which are the valence state energies. From this description it is obvious that $\mathcal{H}^v$ is the object which is being mimicked by semiempirical $\mathcal{H}_M^v$'s. In semiempirical theories it is further assumed that $\mathcal{H}_M^v$ is composed of effective one- and sometimes two-electron operators; however, three-and higher-electron operators are always ignored. Semiempirical theories then consider the individual integrals of the effective one- and two-electron operators between valence orbitals to be parameters that may be "adjusted" to include correlation effects and to reproduce the valence state energies. The parametrization scheme for these effective integrals incorporates empirical data based on the particular brand of chemical intuition used by the method's inventor at the time of its conception. With each scheme is associated a different acronym, and consequently the entire collection of semiempirical methods appears as alphabet soup to the uninitiated.

In an effort to understand the theoretical basis of semiempirical theories, we have calculated, from first principles, both the valence state energies and the integrals of the effective one-, two-, three-, and four-electron operators of $\mathcal{H}^v$ between the valence orbitals. By analogy with semiempirical theories, we designate these integrals as true parameters. Our calculations are designed to address the following questions: (1) How well can we calculate the valence state energies; and, does $\mathcal{H}^v$ have the potential to be a viable ab initio method? (2) How do the true parameters compare with semiempirical parameters? Especially, how large are the nonclassical many-electron effective integrals which are ignored in semiempirical methods? (3)Are the one-center true parameters really transferable between molecular systems with similar environments as is assumed in semiempirical theories? (4) What is the effect of many-electron effective operators on the calculation of excitation energies and ionization potentials?

Perturbation Expansion for $\mathcal{H}^v$

The energy-dependent representation of $\mathcal{H}^v$ in equation (2) has several nice properties which are discussed elsewhere[1-3]; However, it suffers from the drawback that individual configuration-independent true parameters cannot be directly calculated. This is particularly important for making comparisons with semiempirical methods. Hence, Iwata and Freed introduced an energy-independent generalized perturbation expansion couched in the language of second quantization.[3] With subsequent modifications,[4] which significantly improve the convergence properties of the expansion, it is possible to derive formulas for the individual true parameters. These true parameters have the following physically appealing properties[3,4]: (1) The true parameters are independent of the valence electron configuration. (2) The true parameters are independent of the number of valence electrons. This parallels the use of semiempirical parameters which are independent of the molecular environment and the net (or local) charge of the system. (3) The true parameters can be evaluated within a theory which utilizes matrix energy denominators, thus eliminating the necessity of tedious resummation of dominant terms through all orders in perturbation theory. Properties (1) and (2) mean that after the true parameters are calculated once for a given set of valence orbitals, they can then be used to evaluate the energies of <u>all</u> valence states of <u>all</u> charge states of the system.

Using the simplest form of our generalized quasi-degenerate many-body perturbation theory (GQDMBPT) in which the energy denominators are taken to be strictly diagonal, our expansion reduces to a symmetrized version of Brandow's diagrammatic perturbation theory. [8] All of our calculations to date use this form. In second order our method is equivalent to a B_k-type calculation of Gershgorn and Shavitt,[9] and in thrid order the off-diagonal elements of H_{QQ} in equation (1) are kept in lowest order of perturbation theory.

The explicit details of third order GQDMBPT and the calculation of $\mathcal{H}^v$ will be presented elsewhere with emphasis on the formal theory of GQDMBPT,

TABLE I. Selected effective integrals (true parameters) for fluorine in eV. Atomic basis orbitals are obtained from ground state RHF calculations for the indicated ion. Both second and third order results are presented to indicate convergence. MINDO/3 values are from Reference 13 and the empirical values are in Reference 12. s, x. y, z are shorthand notation for 2s, $2p_x$, $2p_y$, $2p_z$.

	F		F^+		F^{++}		F^{+3}		F^{+7}		Empirical	
True Parameter	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd		MINDO/3
E_c	-2054.97	-2055.05	-2054.97	-2055.05	-2054.98	-2055.05	-2054.98	-2055.05	-2055.00	-2055.05	-2057.11	
$\langle s\|\mathcal{H}^{v}_{i}\|s\rangle$	- 185.57	- 185.00	- 185.21	- 185.40	- 184.98	- 185.12	- 184.91	- 184.98	- 184.92	- 184.92	- 185.19	-129.86
$\langle x\|\mathcal{H}^{v}_{i}\|x\rangle$	- 169.28	- 173.54	- 169.33	- 171.64	- 169.91	- 170.96	- 170.37	- 170.82	- 170.79	- 170.79	- 171.19	-105.83
$\langle sx\|\mathcal{H}^{v}_{ij}\|sx\rangle$	31.85	33.76	31.50	32.62	31.45	32.02	31.54	31.80	31.72	31.68	31.63	17.25
$\langle sx\|\mathcal{H}^{v}_{ij}\|xs\rangle$	5.82	6.00	5.73	5.80	5.70	5.70	5.71	5.68	5.69	5.65	5.56	4.83
$\langle xy\|\mathcal{H}^{v}_{ij}\|xy\rangle$	31.86	35.46	31.61	33.79	31.82	33.01	32.19	32.77	32.69	32.67	32.66	14.91
$\langle xy\|\mathcal{H}^{v}_{ij}\|yx\rangle$	1.40	1.49	1.42	1.48	1.46	1.48	1.50	1.49	1.51	1.49	1.41	0.90
$\langle sxy\|\mathcal{H}^{v}_{ijk}\|sxy\rangle$ [a]	-2.03	-3.19	-1.86	-2.63	-1.76	-2.27	-1.77	-2.08	-1.90	-2.00	-1.90	
$\langle sxy\|\mathcal{H}^{v}_{ijk}\|syx\rangle$	0.06	0.09	0.07	0.08	0.08	0.07	0.07	0.07	0.06	0.06	0.06	
$\langle sxy\|\mathcal{H}^{v}_{ijk}\|xsy\rangle$	-0.30	-0.38	-0.26	-0.30	-0.24	-0.25	-0.23	-0.23	-0.17	-0.24	-0.23	
$\langle xyz\|\mathcal{H}^{v}_{ijk}\|xyz\rangle$	-2.63	-4.50	-2.44	-3.75	-2.36	-3.29	-2.45	-3.07	-3.00	-3.08	-2.79	
$\langle xyz\|\mathcal{H}^{v}_{ijk}\|xzy\rangle$	-0.10	-0.15	-0.10	-0.14	-0.09	-0.14	-0.10	-0.14	-0.14	-0.18	-0.13	
$\langle zyxs\|\mathcal{H}^{v}_{ijk\ell}\|zyxs\rangle$ [b]		0.37		0.30		0.25		0.21		-0.02	-0.07	
$\langle zyxs\|\mathcal{H}^{v}_{ijk\ell}\|yzsx\rangle$		-0.02		-0.02		-0.02		-0.02		-0.03	-0.04	
$\langle zyxs\|\mathcal{H}^{v}_{ijk\ell}\|zsyx\rangle$		0.01		0.00		-0.01		-0.01		-0.01	0.01	
$\langle zyxs\|\mathcal{H}^{v}_{ijk\ell}\|szxy\rangle$		-0.02		-0.01		0.01		0.01		0.04	0.04	

a. $\langle uvw|\mathcal{H}^{v}_{ijk}|xyz\rangle = \iiint d\underline{r}_i d\underline{r}_j d\underline{r}_k\, u^*(i)v^*(j)w^*(k)\, \mathcal{H}^{v}_{ijk}\, [x(i)y(j)z(k)-x(k)y(j)z(i)]$

b. $\langle abcd|\mathcal{H}^{v}_{ijk\ell}|wxyz\rangle = \iiiint d\underline{r}_i d\underline{r}_j d\underline{r}_k d\underline{r}_\ell\; a^*(i)b^*(j)c^*(k)d^*(\ell)\; \mathcal{H}^{v}_{ijk\ell}\; [w(i)x(j)y(k)z(\ell)-w(i)x(\ell)y(k)z(j)-w(k)x(j)y(i)z(\ell)+w(k)x(\ell)y(i)z(j)]$

convergence properties of the expansion, calculational difficulties in third order, and special techniques used in computer algorithms.[10]

TABLE 2. Average deviation from experiment of all calculated valence excitation energies and ionization potentials for all charge states of fluorine.

Atomic Orbital Base	2nd order	3rd order	3rd order*
F	0.86	0.41	6.90
F^{+}	0.72	0.39	5.75
F^{++}	0.69	0.56	5.16
F^{+3}	0.74	0.68	4.99
F^{+7}	0.59	0.27	4.96

*without three-and four-electron true parameters.

CALCULATION FOR THE FLUORINE ATOM

In addition to the questions raised in the last section, our calculations for fluorine are also designed to be a stringent test of the convergence properties of the perturbation expansion. The calculation begins with the choice of a primitive basis which for fluorine consists of 5s4p2d Slater type functions (STF's) and is unchanged throughout all calculations. Quantum numbers and exponents for the basis set are reported in reference (11). We determine five different sets of atomic orbitals from restricted Hartree-Fock (RHF) calculations for the ground states of F, F^{+}, F^{++}, F^{+3}, and F^{+7} all using the same primitive basis. Each of these is used in a different third order approximate calculation of $\mathcal{H}^{v}$. The 2s and 2p atomic orbitals are designated as the valence shell orbitals. The zeroth order Hamiltonian, H_o, for the perturbation expansion ($H = H_o + V$) is chosen to be the N-electron operator which is the diagonal portion of the Fock operator constructed from the $(1s)^2$ configuration for each atomic orbital basis. Note that only for the orbitals obtained from the F^{+7} RHF calculation is the $(1s)^2$ Fock operator truly diagonal. The diagonal elements of the $(1s)^2$ Fock operator are the "bare core" orbital energies.

Values for selected ab initio true parameters through third order are presented in Table 1 along with a uniquely determined empirical set of true parameters[12] and one example of a set of semiempirical parameters.[13] Because there are 46 valence state energies to be reported for each calculation we list, in Table 2, just the atomic orbital basis and the average deviation from experiment for all known excitation energies and ionization potentials.[14] A more thorough report will be made elsewhere.[15]

From Table 1, there are several important observations. (1) Some of the three-electron true parameters are as large as 3 eV! (2) There is approximately an order of magnitude reduction in the size of the four-electron true parameters relative to the three-electron ones. (3) Notice the conspicuous absence of three- and four- electron semiempirical parameters. (4) There is a large discrepancy between the ab initio true parameters and the semiempirical MINDO/3 parameters. (5) The empirical true parameters represent the theoretical limit of infinite order complete basis function calculations.[12] A comparison of the ab initio and empirical true parameters shows that convergence is better for atomic orbitals that are determined from RHF calculations in which the subshell is not more than half full. This is also demonstrated in Table 2; and can be explained on the basis of the structure of the theoretical expressions for the true parameters.[10] This fact is more evident with different choices of H_o.[15] Finally from Table 2, the effect of ignoring three-electron true parameters are important for describing intrashell and intrasubshell correlations (in this case 2s↔2p and 2p↔2p).[5,11]

CALCULATION OF $\mathcal{H}^{v}$ for CH

We have performed two distinct types of calculations for CH; the first is used to test the applicability of ab initio effective Hamiltonian methods for calculating the valence properties of molecules, while the second is designed to investigate the degree of transferability for the true parameters of $\mathcal{H}^{v}$. The primitive basis set for both calculations is a 4s3p1d STF basis on carbon with a 2s1p STF basis on hydrogen. Both calculations are performed at the ground state equilibrium internuclear separation of 2.124 au. The valence shell orbitals for each $\mathcal{H}^{v}$ calculation are taken to be the 2σ, 3σ, $1\pi_x$, $1\pi_y$, and 4σ molecular orbitals obtained from a RHF calculation for the $^2\Pi$ ground state of the neutral molecule.

Direct Calculation

The first CH calculation is straightforward and follows very similar procedures to those used in the fluorine calculation, except that we report the results from only one molecular orbital basis and we use an average bare core orbital energy in H_o for the valence shell orbitals. The second and third order $\mathcal{H}^{v}$ results for valence shell ionization potentials and excitation energies are presented in Table 3. Since we did not have access to an MCSCF wavefunction, special pains are taken to determine the 4σ virtual orbital from a V^{N-1} potential. This is necessary to insure that the 4σ orbital has the character of a valence shell orbital. Otherwise, nonconvergence of the $\mathcal{H}^{v}$ expansion becomes a problem.[10] Averaging of the valence shell orbital energies is also important for convergence of the $\mathcal{H}^{v}$ expansion for CH.

Constrained Calculation

As a preliminary investigation of transferability, we compare the one-center two-electron effective integrals, generated for the valence shell of a pure carbon atom, with the corresponding effective integrals extracted from a calculation for CH. In order to make a valid comparison, it is important that the valence space for the pure carbon atom and the carbon in CH be as similar as possible.

TABLE 3. Excitation energies, ionization potential, and electron affinity of CH in eV. Calculations are at internuclear separation of 2.124 au. The $^2\Pi$ ground state energy of CH is in au.

State	Direct		Constrained			
	2nd	3rd	2nd	3rd	CI[a]	Experimental[b]
CH($^2\Pi$)[c]	-38.4642	-38.4014	-38.4515	-38.3983	-38.4103	-38.490
$^4\Sigma^-$	0.673	0.479	0.778	0.491	0.67	0.742
$^2\Delta$	2.946	3.084	3.119	3.105	2.93	2.88
$^2\Sigma^-$	3.445	3.211	3.530	3.224	3.28	3.19
$^2\Sigma^+$	4.433	4.080	4.644	4.055	4.02	3.94
$^2\Pi$	7.668	6.379	7.988	6.408		7.31
CH$^+$ ($^1\Sigma^+$)						
$^3\Pi$	1.535	0.845	1.302	1.147	1.14	
$^1\Pi$	3.529	2.821	3.336	3.105	3.18	
$^3\Sigma^+$	13.482	12.383	13.417	12.461	11.65 (?)	
CH$^-$($^3\Sigma^-$)						
$^1\Delta$	1.054	1.578	1.141	1.542		0.845
I. P.[d]	10.488	10.411	9.972	10.586		10.64
E. A.[e]	0.937	0.820	0.241	- 0.376		1.238
D. E.[f]	3.884	2.690	3.539	2.621		3.65

a. CI data for CH is from ref. 16 and CH$^+$ is from ref. 17. b. Ref. 18. c. Ground state energy for CH. d. Ionization potential. e. Electron affinity. f. Dissociation energy (using pure carbon calculation).

TABLE 4. One-center two-electron effective integrals from pure carbon calculation and corresponding effective integrals from constrained CH calculation in eV.

Integral	Pure Carbon Calculated	Pure Carbon Empirical[a]	Constrained CH[b]	
			σ	π
$\langle ss\|\mathcal{H}^v_{ij}\|ss\rangle$	18.42	17.47	18.34	
$\langle sx\|\mathcal{H}^v_{ij}\|sx\rangle$	18.96	18.20	18.94	19.33
$\langle sx\|\mathcal{H}^v_{ij}\|xs\rangle$	3.20	3.10	3.68	3.15
$\langle ss\|\mathcal{H}^v_{ij}\|xx\rangle$	3.34	2.48	3.89	3.08
$\langle xx\|\mathcal{H}^v_{ij}\|xx\rangle$	20.98	19.93	21.22	22.25
$\langle xy\|\mathcal{H}^v_{ij}\|xy\rangle$	19.32	18.17	19.64	20.32
$\langle xy\|\mathcal{H}^v_{ij}\|yx\rangle$	0.58	0.52	0.63	0.65
$\langle xx\|\mathcal{H}^v_{ij}\|yy\rangle$	1.08	1.23	0.99	1.27

a. From ref. 12.

b. There are two columns for the constrained CH calculation because the atom 2p orbitals are split by linear symmetry into σ and π orbitals.

This arduous task is accomplished as described below. Atomic orbitals are determined from RHF calculations separately for the pure carbon ^3P ground state and for hydrogen. The carbon atomic orbitals are used for an $\mathcal{H}^v$ carbon atom calculation. Then an unorthodox RHF calculation for the $^2\Pi$ ground state of CH is carried out such that the valence shell orbitals of CH (2σ, 3σ, $1\Pi_x$, $1\Pi_y$, and 4σ) are constrained to be representable by the basis which is the union of the C and H atomic valence bases. This set of molecular orbitals is fed into the approximate $\mathcal{H}^v$ calculation for CH. This restriction on the valence space produces orbitals which are less than optimal, therefore the rate of convergence for this constrained calculation is slower. Valence excitation energies and ionization potentials are reported in Table 3 with those from the direct calculation.

The $\mathcal{H}^v$ for CH is then back transformed to the nonorthogonal atomic valence orbitals to enable a comparison between the carbon atom two-electron true parameters and the corresponding one-center two-electron effective integrals of CH. (See Table 4) Included in the table are the carbon atom true parameters. There are two columns associated with the constrained CH calculations in Table 4 because the atomic p orbitals are split by symmetry into σ and π orbitals.

Results

Unlike the case for fluorine atom, there is not sufficient experimental data to determine a set of empirical parameters for CH as a check for the $\mathcal{H}^v$ true parameters individually.[12] The same trends for the true parameters, noticed in fluorine, are present in CH also. Calculated values for the vertical excitation energies and ionization potentials from both calculations are compared with the CI results of Lie, Hinze, and Liu for CH[16] and Green et al for CH^+[17] in Table 3. Also in Table 3 are the experimental adiabatic quantities of which we are aware.[18]

Table 3 indicates that our $\mathcal{H}^v$ calculations are in excellent agreement with conventional CI results, and hence we conclude that $\mathcal{H}^v$ can be applied with accuracy to molecular systems. Table 3 also demonstrates, as expected, that the constrained calculation is not as well converged in second order; and consequently, at least a third order treatment is required to make accurate predictions.

From Table 4 we see that the idea of transferability is not inconsistent with our calculations. However, there remains much work to be done before this aspect of $\mathcal{H}^v$ is more fully understood and can be utilized.

THE FUTURE

In contrast to the relative insensitivity of the $\mathcal{H}^v$ expansion to the choice of atomic orbital basis in first row atoms like fluorine, molecules and larger atoms (i.e., transition metals) require a more accurate zeroth order description of the orbitals. To insure convergence of the $\mathcal{H}^v$ expansion for these systems, especially polyatomic molecules and molecules at large internuclear separations, it will be necessary to begin with an MCSCF procedure to determine the valence (and excited) molecular orbitals.

Some of the work, which is planned but remains to be started and/or finished in the continuing saga of $\mathcal{H}^v$, involves a series of third order $\mathcal{H}^v$ calculations for NH and OH as well as CH.[19] A series of third order $\mathcal{H}^v$ calculations will be presented for selected first and second row atoms.[20] Thought has been given to an $\mathcal{H}^v$ calculation for a heavier diatomic like CN, CO, or F_2.

One planned project is to calculate a set of potential curves for the valence states of CH.[21] This constitutes a stringent test of the usefulness of our ab initio $\mathcal{H}^v$ method for molecules as both a direct first principles calculation and as a source of insight into the fundamental basis of semiempirical theories. The calculation will provide the first ab initio investigation of the bond length dependence of two-center semiempirical-like effective integrals. Obviously a calculation of this type must begin with MCSCF wavefunctions.

Another large project is the calculation of $\mathcal{H}^v$ for the π shell of butadiene. This calculation has been delayed since Iwata and Freed finished their benchmark investigation for ethylene.[22] Westhaus has since published $\mathcal{H}^v$ ethylene calculations using a different formalism.[23] With the knowledge obtained from calculations on first and second row atomic and atomic hydride systems, and the experience gained through implementation of the third order expansion, we now have the technology to undertake the calculation of $\mathcal{H}^v$ for butadiene. This will be the first ab initio test of semiempirical π-Hamiltonian theories for a system with an extended π-valence shell using $\mathcal{H}^v$. The orbitals are by necessity determined from a three configuration valence MCSCF calculation. Motivated by analogy with the semiemperical multiconfigurational work of Eaker and Hinze,[24] a set of well localized π orbitals can be obtained. These localized orbitals will facilitate the investigation of semiempirical assumptions concerning model π-Hamiltonians.

Preliminary calculations for titanium have recently been reported.[25] Transition metals are particularly interesting for three reasons. (1) Semiempirical theories have been unsuccessful for these systems, (2) little experimental data is available, and (3) not much ab initio work has been done on their correlation energies, especially for excited states. We plan to continue our work on titanium and also to investigate iron since it resides at the other end of the first transition metal series.

SUMMARY

To reiterate, our calculations show that $\mathcal{H}^v$ is an accurate and novel ab initio method for small systems. It is particularly interesting and exciting since one calculation of the effective integrals yields all valence state energies of all charge states accurately.

The calculation for atoms already provides a first principles check of the one-center integrals of semiempirical theories. The large discrepancy noted in Table I is explained by an averaging of effective three- and four-electron integrals into the one- and two-electron semiempricial parameters in analogy with the way that two-electron integrals are averaged into the one-electron Fock operator.[26]

Finally, our calculation for CH shows that $\mathcal{H}^v$ can be applied accurately to molecules. It provides a test of transferability and will generate an ab initio bond length dependence for two-center effective integrals which can be compared with semiempirical theories.

REFERENCES

1. K. F. Freed, in *Modern Theoretical Chemistry*, Vol. 7, ed. G. A. Segal (Plenum, New York, 1977).

2. K. F. Freed, Chem. Phys. Lett. 13, 181 (1972); 15, 331 (1972); 24, 275 (1974); J. Chem. Phys. 60, 1765 (1974); Chem. Phys. 3, 463 (1974); 4, 80 (1974).

3. S. Iwata and K. F. Freed, J. Chem. Phys. 65, 1071 (1976). Erratum, 66, 1765 (1977).

4. M. G. Sheppard, K. F. Freed, M. F. Herman, and D. L. Yeager, Chem. Phys. Lett. 61, 577 (1979).

5. D. L. Yeager, H. Sun, K. F. Freed, and M. F. Herman, Chem. Phys. Lett. 57, 490 (1978).

6. R. G. Parr, *The Quantum Theory of Molecular Electronic Structure* (Benjamin, New York, 1963); M. J. S. Dewar, *The Molecular Orbital Theory of Organic Chemistry* (McGraw-Hill, New York, 1969); J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory* (McGraw Hill, New YOrk, 1970).

7. P. O. Lowdin, J. Mol. Spectrosc. 14, 112 (1964); J. Math. Phys. 3, 969 (1962).

8. B. H. Brandow, Rev. Mod. Phys. 39, 771 (1967); Adv. Quantum Chem. 10, 187 (1977); Int. J. Quantum Chem. 15, 207 (1979).

9. Z. Gershgorn and I. Shavitt, Int. J. Quantum Chem. 2, 751 (1968).

10. M. G. Sheppard and K. F. Freed, "Third Order $\mathcal{H}^v$: Theory and Calculation" to be published.

11. H. Sun, K. F. Freed, M. F. Herman and D. L. Yeager, J. Chem. Phys. 72, 4158 (1980).

12. D. L. Yeager, M. G. Sheppard, and K. F. Freed, J. A. C. S. 102, 1270 (1980).

13. R. C. Bingham, M. J. S. Dewar, and D. H. Lo, J. A. C. S. 97, 1285 (1975).

14. C. Moore, Ed., *Atomic Energy Levels*, Natl. Stand. Ref. Data Ser., Natl. Bur. Stand. No. 35 (1971); S. Bashkin and J. O. Stoner, *Atomic Energy Levels and Grotrian Diagrams*, Vol. I (North-Holland, Amsterdam, 1975).

15. M. G. Sheppard and K. F. Freed "Third Order $\mathcal{H}^v$: Application to Fluorine" to be published.

16. G. C. Lie, J. Hinze, and B. Liu, J. Chem. Phys. 57, 625 (1972); 59, 1872 (1973).

17. S. Green, P. S. Bagus, B. Liu, A. D. McLean, and M. Yoshimine, Phys. Rev. A, 5, 1614 (1972).

18. G. Herzberg and J. W. C. Johns, Astrophys. J. 158, 399 (1969); A. Kasdan, E. Herbst, and W. C. Lineberger, Chem. Phys. Lett. 31, 78 (1975); ref. cited in P. E. Cade and W. M. Huo, J. Chem. Phys. 47, 614 (1967); ref. cited in ref. 16 and 17.

19. H. Sun, M. F. Herman, M. G. Sheppard, and K. F. Freed, to be published.

20. M. G. Sheppard and K. F. Freed, to be published.

21. H. Sun and K. F. Freed, to be published.

22. S. Iwata and K. F. Freed, J. Chem. Phys. 61, 1500 (1974); Chem. Phys. Lett. 28, 176 (1974).

23. P. Westhaus, E. G. Bradford, and D. Hall, J. Chem. Phys. 62, 1607 (1975); P. Westhaus and E. G. Bradford, J. Chem. Phys. 63, 5416 (1975); 64, 4276 (1976); P. Westhaus and Moghtaderi, Quantum Theory Research Group, Notes 80, 81, Oklahoma State University, Nov. 1978;J.Chem.Phys. 72, 4174 (1980).

24. C. W. Eaker and J. Hinze, J. A. C. S. 96, 4084 (1974).

25. Y. S. Lee, H. Sun, M. G. Sheppard, and K. F. Freed, J. Chem. Phys. in press.

26. K. F. Freed and H. Sun, Israel J. Chem. 19, 99 (1980).

ACKNOWLEDGEMENT

This work was supported in part by the National Science Foundation Grant NSF CHE 77-24652.

MCSCF FOR EXCITED STATES

Charles W. Bauschlicher, Jr.[*]
Institute for Computer Applications in Science and Engineering
NASA Langley Research Center, Hampton, VA 23665

Byron H. Lengsfield III[†]
Laser and Spectroscopy Branch
NASA Langley Research Center, Hampton, VA 23665

David R. Yarkony[‡]
Department of Chemistry, Johns Hopkins University
Baltimore, MD 21218

Two basic types of configuration interaction (CI) calculations are routinely used in accurate ab initio calculations. They are the First Order CI[1] (FOCI) and the all single and double excitation CI, CI(SD). In both cases the orbitals are initially optimized using a limited MCSCF. Recent advances in MCSCF Theory[2] now permit MCSCF calculation containing ∿200 CSF's to be performed with little effort in excess of the previous small MCSCF calculations. Thus, it is now possible to optimize the orbitals for a FOCI, by performing a full valence MCSCF (FVMCSCF). If the FOCI is too large and a POLCI[3] (a POLCI is best viewed as a FOCI with CSF selection) is to be used or a CI(SD) is to be run, the FVMCSCF can be employed to identify the dominant configurations in the wavefunction. FVMCSCF calculations are performed at various points on the surface, the natural orbitals (NO's) obtained and FVCI repeated in the NO basis. The union of all important configurations is then used in a small MCSCF. This MCSCF is followed by either the POLCI or CI(SD), using the orbitals obtained in the small MCSCF. By performing the FVMCSCF, one simplifies the procedure for identifying the important CSF's which should be included in the MCSCF.

Limited MCSCF calculations followed by CI calculations are a standard treatment for the lowest state in each symmetry. One exception would be when transition moments are desired. In this case a common set of orbitals would be used, but even in this case the separate MCSCF orbital optimization followed by CI is usually performed to calibrate the common orbital set. In principle one could use the same techniques on the higher states of a given symmetry, but in practice new problems arise. As one optimizes an upper root, the description of the lower root is degraded. This can lead to a flipping of these two states and the loss of the upper bound to the desire root.

Hinze[4] has proposed optimizing the orbitals based on the averaged field of the states of interest. He also noted that one must compensate for the poorer description of the orbitals by adding additional CSF's. The addition of CSF's to the MCSCF greatly increased the size of the CI(SD) and can quickly make the problem intractable. If the configuration list is not increased, it seems unlikely the states treated by averaged field method will be as accurate as those state for which an independent optimization of the orbitals is performed. Averaging all states of interest would encounter problems if all states but one were well described by one set of orbitals. The average field technique would treat one state more poorly than the rest. While the averaged field technique avoids the problem of variational collapse, it may not treat all states equivalently for a configuration list small enough to allow a CI(SD). Ideally, one would like to perform an independent variational calculation on each state. Grein[5] hss noted if variational collapse begins to occur, a small CI would indicate which CSF's are needed to keep the lower roots in place. While in principle the MCSCF could become prohibitively large, in practice this does not appear to be a problem. We find for ionic systems such as BeO, MgO, and CaO that FVMCSCF - FOCI calculations yield poor separations between the low-lying states and instead a limited MCSCF followed by CI(SD) was employed. The addition of the CSF's needed to prevent variational collapse can make the CI(SD) prohibitively large. Also for a FVMCSCF - FOCI if an upper root collapsed, the addition of the CSF's needed to prevent the variational collapse would represent a redefinition of the valence space. Liu[6] has found that in some cases the addition of extra orbitals to the valence space in a MCSCF - FOCI calculation can actually yield poorer results and therefore should be avoided.

Our general procedure is to perform a FVMCSCF at several representative points on the surface. The important and marginally important configurations are included in a small MCSCF. Those CSF's with a coefficient smaller than 0.1 at all points are dropped and the final CSF list is chosen. This final list is used in an MCSCF and followed by a CI(SD). If one is only considering the lowest root, repeating the FVCI in terms of NO's can simplify the indentification of the important CSF, but for several roots this is not advisable. In BeO, for example, in the NO basis the $1^1\Sigma^+$ state would be described by

---

[*] a contribution by C. W. Bauschlicher, Jr. Work was supported under NASA Contracts No. NAS1-14101 and NAS1-14472 while in residence at ICASE, NASA Langley Research Center, Hampton, VA 23665

[†] NRC-NASA Research Associate, 1978-1980. Present address: IBM, 5600 Cottle Road, San Jose, CA 95193.

[‡] Alfred P. Sloan Research Fellow. Work was supported by the Air Force Office of Scientific Research (AFOSR-79-0073) and the National Science Foundation (CHE-7824153) during the course of this research.

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 4\sigma^2\ \ 1\pi^4 \tag{1}$$

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 5\sigma^2\ \ 1\pi^4 \tag{2}$$

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 4\sigma^2\ \ 1\pi^2\ 2\pi^2 \tag{3}$$

but the second and third roots would be nominally described as

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 4\sigma\ 5\sigma\ \ 1\pi^4 \tag{4}$$

$$1\sigma^2\ 2\sigma^2\ 3\sigma^2\ 4\sigma^2\ \ 1\pi^3\ 2\pi \tag{5}$$

Starting from (1), (4) and (5) provides a good description of the first and second roots, but starting from (1), (2) and (3) for the first root would necessitate adding other CSF's to describe the second root.

The cases where it is impossible to perform separate optimizations on higher roots fall into two categories: 1) the FVMCSCF calculation collapses and in 2) the limited MCSCF undergoes variational collapse. In the former we do not add any additional configurations. In the case of CaO $^1\Sigma^+$ states, we overcome this collapse by performing a FVMCSCF (where 8σ, 9σ, 10σ, 3π and 4π are valence orbitals) on the lowest $^1\Sigma^+$ state, obtained the natural orbitals, then used these NO's in FVMCSCF for the $2^1\Sigma^+$ with the 8σ orbital frozen to be the 8σ NO. The important CSF's from both roots included in an MCSCF calculation, and the orbitals optimized for the 1st root. The second root orbitals are then optimized with the 8σ orbital to be the 8σ orbital of the ground state. The second case where the limited MCSCF collapses, arises when the roots are close and the marginal CSF's of the first root are important in preventing root flipping. If the number of CSF's to be added is too large, we freeze an orbital (or orbitals) to be an orbital obtained in limited MCSCF calculation of the ground state. Since we have performed the FVMCSCF calculations we have some measure of the severity of these constraints.

We have noted that if the marginally important CSF's are added to the MCSCF, the energy of the limited MCSCF is within a few milli-hartrees of the FVMCSCF. For MgO the number of CSF's need to approach the FVMCSCF is less than 15 CSF's for all the low lying states. In order to minimize the uncertainty in our calculations we avoid CSF selection whenever possible and if forced to select we keep the cumulative[7] selection threshold an order of magnitude smaller than the accuracy we seek (1 milli-hartree accuracy). For this reason we choose to include only the important CSF's in the MCSCF - CI. Calculations for MgO using an STO basis set show that this procedure yields Re's and Te's in excellent agreement with experiment.

We should note a few technical aspects of our procedure. The FVMCSCF is made possible by using a second-order MCSCF procedure. Instead of solving the simultaneous equations, we employ a Super-CI technique with a procedure for damping the eigenvector far from convergence. In our implimation a variable number of CI roots can be included in the Hessian. We find that far from convergence including all lower roots improves convergence and only near convergence are all roots coupled in and quadratic convergence observed. The inclusion of all lower CI vectors is an important aspect of the second order MCSCF treatment of excited states. This is to be contrasted to the traditional generalized Brillouin theorem MCSCF where one often needs to shift the diagonal elements of the Super - CI Hamiltonian in order to prevent root flipping. In the case of higher roots of ionic molecules the lower state orbitals or the orbitals of a non-variational SCF do not always provide a good set of starting orbitals. In these situations the use of damping, can be very important. In some cases, the roots flipped as the higher root orbitals were optimized. However, convergence was obtained in these cases by freezing orbitals in the lower root for a few iteration and the eliminating the constraint. In the cases where the apparent variational collapse was a result of a poor choice of starting orbitals, this procedure works well. However one must be careful that this procedure does not lead to a local minima. We should also note that we have found corresponding orbitals very useful in comparing two sets of orbitals. For example, by computing the corresponding orbitals[9] between the lowest root and a root just as it undergoes variational collapse, it becomes trivial to observe which orbital or orbitals differ in the two states.

The procedure discussed in this paper are based on the assumption that MCSCF calculations containing more than a few hundred CSF's are not routinely possible, however Schaefer[10] has recently reported an MCSCF including more than 10,000 CSF's Since Schaefer is using a first order method to optimize the orbitals, it is not clear that reasonable convergence will be obtained for a CI(SD) using a general MCSCF reference. Second order MCSCF techniques[2] have been shown to provide excellent convergence for a general list of CSF's and this method in principle could be used to treat problems including large numbers of CSF's. However, the amount of work needed to construct the Hessian and solve the simultaneous equations could become prohibitively large. If these techniques lead to the ability to routinely perform very large MCSCF calculations, the need to perform a separate CI calculation will be eliminated for most calculations. However, the procedures described will still be useful in determining the list of reference configurations and an initial set of orbitals for the larger MCSCF.

REFERENCES

[1] H. F. Schaefer, III, R. H. Klemm, and F. E. Harris, Phys. Rev. 181, 137(1969).

[2] B. H. Lengsfield, J. Chem. Phys. in press. While much work has been performed on second order MCSCF theory, this appears to be one of the few completely general second-order MCSCF procedure.

[3] P. J. Hay and T. H. Dunning, Jr., J. Chem. Phys. 64, 5077(1976).

[4] K. K. Docken and J. Hinze, J. Chem. Phys. 59, 6424(1973).

[5] A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054(1977).

[6] B. Liu Private Communication

[7] R. C. Raffenetti, K. Hsu, and I. Shavitt, Theor. Chim. Acta 45, 33(1977).

[8] C. W. Bauschlicher, B. H. Lengsfield, D. M. Silver, D. R. Yarkony, unpublished work.

[9] A. T. Amos and G. G. Hall, Proc. Roy. Soc. A263, 483(1961), R. L. Martin and E. R. Davidson, Phys. Rev. A16, 1341(1977).

[10] B. R. Brooks, W. D. Laidig, P. Saxe, and H. F. Schaefer, J. Chem. Phys. 72 3837(1980).

THE PROJECTION OF BROKEN SYMMETRY WAVEFUNCTIONS

Richard L. Martin
Theoretical Division, MS-569
Los Alamos Scientific Laboratory
Los Alamos, NM 87545

I. INTRODUCTION

There are a number of problems for which the simplest qualitatively acceptable wavefunction is written as a linear combination of nonorthogonal Slater determinants; i.e. in valence bond (VB) form. For example, Jackels and Davidson[1] showed that certain portions of the NO_2 potential energy surface are preferably described as "resonance" combinations of the two valence bond structures,

which arise as broken (spatial) symmetry solutions to the Hartree-Fock equations. The principal advantage of a nonorthogonal expansion of this form is that all of the polarization, or left-right correlation effects, implicit in the broken symmetry determinant are included in the wavefunction; a CI expansion based on symmetry restricted orbitals requires high order excitations in order to span the same space. On the other hand, if the VB structures need to be "dressed" with additional correlations, a nonorthogonal expansion quickly becomes impractical. A compromise MCSCF/CI expansion for NO_2 has been suggested by Engelbrecht and Liu,[2] and is discussed elsewhere in these proceedings. In the present contribution,[3] the encouraging results of an application of the valence bond model to the low-lying excited states of pyrazine and para-benzoquinone are presented, and the features of the corresponding orbital transformation of Amos and Hall[4] which make these "projected broken symmetry," or "projected spatially-unrestricted Hartree-Fock," calculations possible for large systems are outlined.[5]

II. CORRESPONDING ORBITALS

Let the orthonormal set of n occupied spin orbitals which define the left-hand determinant, Ψ_L, be denoted by $\{\psi^\ell\}$, and those defining the right-hand determinant Ψ_R, by $\{\psi^\ell\}$. These two sets need not be related, but for the applications considered here they are connected by a symmetry operation. In NO_2, e.g.,

$$\psi_i^r = \hat{\sigma}_{yz}\psi_i^\ell \quad ,$$

where $\hat{\sigma}_{xz}$ is the reflection operator which sends one oxygen into the other. The problems associated with evaluation of the Hamiltonian between the two determinants arise because the two sets are not mutually orthogonal. Amos and Hall,[4] however, showed that it is possible to generate two new sets of orbitals, defined by the unitary transformations

$$(\phi_1^\ell,\phi_2^\ell,\ldots\phi_n^\ell) = (\psi_1^\ell,\psi_2^\ell,\ldots\psi_n^\ell)\underline{U} \quad ,$$

and

$$(\phi_1^r,\phi_2^r,\ldots\phi_n^r) = (\psi_1^r,\psi_2^r,\ldots\psi_n^r)\underline{V} \quad ,$$

such that the transformed overlap matrix

$$S'_{ij} = (\phi_i^\ell,\phi_j^r) = \delta_{ij}d_i$$

is diagonal. The matrices $\underline{U}$ and $\underline{V}$ are determined by solving the eigenvalue equations

$$\underline{S}\underline{S}^+u_i = d_i^2u_i$$

$$\underline{S}^+\underline{S}v_i = d_i^2v_i \quad ,$$

with

$$\underline{U} = (u_1,u_2,\ldots u_n)$$

$$\underline{V} = (v_1,v_2,\ldots v_n) \quad .$$

In this basis, the determinantal overlap is simply

$$S_{LR} = (\Psi_L,\Psi_R) = \prod_i^{p+q} d_i \quad ,$$

and the Hamiltonian matrix element is

$$\begin{aligned} H_{LR} &= (\Psi_L,H\Psi_R) \\ &= \sum_i^{p+q} D^{-i}h_{ii} \\ &+ \frac{1}{2}\left\{\sum_{ij}^{p+q} D^{-ij}J'_{ij} - \sum_{ij}^{p}D^{-ij}K_{ij} - \sum_{ij}^{q} D^{-ij}K'_{ij}\right\} \end{aligned} \quad (1)$$

where p and q refer to the number of α and β spin-orbitals, respectively, and

$$D^{-i} = \prod_{k\neq i}^{p+q} d_k$$

$$D^{-ij} = \prod_{k\neq i,j}^{p+q} d_k$$

$$h'_{ii} = \int\phi_i^\ell(1)h\phi_i^r(1)d\tau_1 .$$

$$J'_{ij} = \iint\phi_i^\ell(1)\phi_i^r(1)\frac{1}{r_{12}}\phi_j^\ell(2)\phi_j^r(2)d\tau_1d\tau_2$$

$$D'_{ij} = \iint\phi_i^\ell(1)\phi_j^\ell(1)\frac{1}{r_{12}}\phi_j^r(2)\phi_i^r(2)d\tau_1d\tau_2$$

In practice, Eq. (1) is evaluated from appropriately defined transition density matrices and the original *one- and two-electron integral list. The time* required for the entire procedure is roughly equivalent to that needed for an SCF iteration.

Before discussing the results, it is important to point out that the corresponding orbitals also aid in the interpretation of the wavefunction.[6] Because they maximize the overlap between pairs, they *bring the two orbital sets into maximum* justaposition without changing either wavefunction, thus permitting a straightforward analysis of what really differs between Ψ_L and Ψ_R. This is illustrated in Table I, where the eigenvalues d_1^2 are presented for the corresponding orbitals associated *with the mirror image* $^2A'$ *broken symmetry solutions* in NO_2. Notice that only two orbital pairs have $d_1^2 < 0.99$. The 9a′β set describes the "unpaired" σ electron. One partner is a combination of p_y and p_z functions strongly localized on the left oxygen, the other member is its' mirror image. *In terms of the original broken symmetry* canonical molecular orbitals, it is predominantly the open shell 9a′ orbital (88%), with some admixture of 8a′ (7%), and 7a′ (5%). The only other orbital to markedly differ between Ψ_L and Ψ_R is the 2a″ orbital, an anti-bonding combination of O_{px} orbitals. *It is certainly not surprising* that a π orbital is the most responsive (polarizable) to the localization of the σ electron.

Table I also contains an analysis of the $^3A''$ broken symmetry determinant in CO_2, a mixture of $^3A_2$ $(n_-\to\pi^*)$ and $^3B_1$ $(n_+\to\pi^*)$. If differs from NO_2 in two major respects. First of all, the localization is not so strong. The unpaired σ electron has substantial amplitude on the carbon center. Secondly, the polarization effects are limited to the σ space. The open shell π* orbital is nearly a symmetry function.

Table I. Corresponding Orbital Overlap Integrals[a] in $NO_2(^2A')$ and $CO_2(^3A'')$

Orbital	$NO_2(^2A')$ d_1^2 α	β	$CO_2(^3A'')$ d_1^2 α	β
(1a′,1a′)	1.0000	1.0000	1.0000	1.0000
(2a′,2a′)	1.0000	1.0000	1.0000	1.0000
(3a′,3a′)	1.0000	1.0000	1.0000	1.0000
(4a′,4a′)	1.0000	1.0000	1.0000	0.9999
(5a′,5a′)	0.9999	0.9999	1.0000	0.9998
(6a′,6a′)	0.9999	0.9999	0.9999	0.9998
(7a′,7a′)	0.9994	0.9955	0.9985	0.9920
(8a′,8a′)	0.9977	0.9941	0.9972	0.5086
(9a′,9a′)	0.9903	0.0470	0.7764	
(10a′,10a′)	0.9886			
(1a″,1a″)	0.9991	0.9991	0.9999	0.9999
(2a″,2a″)	0.8597	0.8597	0.9999	0.9973
(3a″,3a″)			0.9992	

[a] The calculations were performed with the bases of ref. 2 at C_{2v} geometries: NO_2, $\bar{r} = 2.45a_o$, $\theta = 100°$; CO_2, $\bar{r} = 2.39a_o$, $\theta = 128°$.

These examples should demonstrate that an analysis of this sort has advantages for defining a suitable one-electron basis for further CI.[7] In much the same spirit, Martin, Davidson, and Eggers[8] have compared SCF wavefunctions in the absence and presence of a perturbing electric field in order to define an appropriate CI basis Similarly, a corresponding orbital comparison of the symmetry restricted Hartree-Fock (SRHF) determinant with the broken symmetry wavefunction should be of assistance in determining an active space for an MCSCF calculation.

III. LOCALIZED EXCITATIONS IN PYRAZINE AND PARA-BENZOQUINONE

Table II summarizes a recent theoretical study[3] of the $^{1,3}(n\to\pi^*)$ excited states and the ^2n ionic states of para-benzoquinone (PBQ), O =  = O. The SCF calculations were performed at the experimental geometry, and, aside from slight differences attributable to the larger (double zeta) Gaussian basis set used in the present work, reinforce the earlier conclusions of Jonkman.[9] The "anomalous" correlation energies associated with the symmetry restricted ΔSCF results (SRHF, column 1) are removed by the broken symmetry wavefunctions (column 2). The n→π* excitation energies are now roughly correct, and the ionization potential is underestimated by ∿1 eV, an amount typical of an oxygen lone pair correlation energy. The g-u splittings in the VB approximation are of the order of a few hundred wavenumbers, nearly an order of magnitude smaller than the SRHF results, and in general qualitative agreement with the experimental observations.[10] Note that the gerade coupling of the excitations lies lower in each case.

Table II Excitation Energies in PBQ (eV)[a]

Transition		$SCF(D_{2h})$	$SCF(C_{2v})$	S_{LR}	VB	Expt.
$^3(n\to\pi^*)$	$^3B_{1g}$	4.24	1.91	$+4.42\times10^{-3}$	1.90	2.32[b]
	^3A_u	$+1600cm^{-1}$			$+230cm^{-1}$	$11cm^{-1}$
$^1(n\to\pi^*)$	$^1B_{1g}$	4.63	2.14	$+3.25\times10^{-3}$	2.14	2.49[b]
	^1A_u	$+1300cm^{-1}$			$+100cm^{-1}$	$-54cm^{-1}$
$^2(n)$	$^2B_{3g}$	11.38	8.91	$+7.77\times10^{-3}$	8.88	9.99[c]
	$^2B_{2u}$	$+2100cm^{-1}$			$+450cm^{-1}$	$+2400cm^{-1}$

[a] All calculations were performed at the experimental geometry with Dunning's [3s,2p/2s] contraction of Huzinaga's (9s,5p/4s) set of primitive Gaussian basis functions (T. H. Dunning, Jr., and P. J. Hay, Modern Theoretical Chemistry, II. Electronic Structure Ab Initio Methods, edited by H. F. Schaefer, III (Plenum, New York, 1975). The first entry for a given pair of transitions is the excitation energy relative to a total energy of −379.11435 a.u. for the ^1A_g ground state and the second is the pair splitting in cm^{-1}.

[b] Ref. 10.

[c] D. Dougherty and S. P. McGlynn, J. Am. Chem. Soc. 99, 3234 (1977).

The interaction between the localized excitations is larger in pyrazine, , primarily because the distance between the nitrogen lone pairs is smaller, and the overlap between them larger. Although the system is thus closer to a "delocalized" (MO) limit, Wadt and Goddard[11] discovered that the SCF equations still prefer broken symmetry solutions for the $^{1,3}(n\to\pi^*)$ and ^2n states. Table III shows that the VB model accurately predicts the excitation energy to the lowest singlet and triplet but it consistently underestimates the experimental splittings by about 0.5 eV. The CI excitation energies of Wadt and Goddard,[11] estimated from independent σ and π space calculations in the SRHF one-electron basis, fall roughly *half-way between the VB model* and experiment. It is interesting to note that the SRHF splittings in this case are in good agreement with experiment; the net effect of the broken symmetry is simply to depress the center of gravity of the excitation energy by 1-2 eV. Jackels and Davidson[1] observed similar behavior for the two lowest electronic states of NO_2.

In summary, the VB approximation appears to provide reasonably accurate excitation energies and splittings for only a modest increase in computational effort. If more accurate studies are required, the natural orbitals of the VB wavefunction should provide a good one-electron basis for more sophisticated CI calculations. The coupling of the localized excitations with nuclear distortions of b_{1u} symmetry have been examined for the $^3(n\to\pi^*)$ states of PBQ and the $^1(n\to\pi^*)$ states of pyrazine and will be reported elsewhere.[12]

The author is indebted to E. R. Davidson, D. A. Kleier, and W. R. Wadt for many helpful discussions. This work was performed under the auspices of the Department of Energy.

Table III. Excitation Energies in pyrazine (eV)[a]

Transition		$SCF(D_{2h})$	$SCF(C_{2v})$	S_{LR}	VB	Expt.
$^3(n\to\pi^*)$	$^3B_{3u}$	4.60	3.60	-.197	3.24	3.2[b]
	$^3B_{2g}$	6.28			4.21	4.7[c]
$^1(n\to\pi^*)$	$^1B_{3u}$	5.36	4.25	-.159	3.82	3.8[b]
	$^1B_{2g}$	7.18			4.84	~5.5[d]
$^2(n)$	^2A_g	11.32	9.27	.227	8.81	9.63[e]
	$^2B_{1u}$	13.55			10.00	11.35[e]

[a] All calculations were performed at the experimental geometry with a double-zeta quality Gaussian basis set (see ref. 11). The excitation energies are relative to a ^1A_g ground state energy of -262.58320 a.u.

[b] K. K. Innes, J. P. Byrne, and I. G. Ross, J. Mol. Spectrosc. 22, 125 (1967); C. Fridh, L. Asbrink, B. O. Jonsson, and E. Lindholm, Int. J. Mass. Spectrom. Ion Physics 8, 101 (1972); H. N. Pisanias, L. S. Christophorou, J. G. Carter, and D. L. McCorkle, J. Chem. Phys. 58, 2110 (1973).

[c] A. Inoue, D. Webster, and E. C. Lim, J. Chem. Phys. 72 1419 (1980).

[d] J. H. Cordes and W. R. Moomaw, private communication to W. R. Wadt, quoted in ref. 11.

[e] R. Gleiter, E. Heilbronner, and V. Hornung, Angew. Chem. 82, 678 (1970); Helv. Chim. Acta. 55, 255 (1972).

References

1. C. F. Jackels and E. R. Davidson, J. Chem. Phys. 69, 2908 (1976).
2. L. Engelbrecht and B. Liu, to be published.
3. R. L. Martin, submitted to J. Chem. Phys.
4. A. T. Amos and G. G. Hall, Proc. R. Soc. A263, 483 (1961).
5. This transformation has also been used recently to simplify the calculation of transition moments between nonorthogonal determinants: R. Bair and W. A. Goddard III, Phys. Rev. A, xxx (1980), B. Lengsfeld, private communication.
6. R. L. Martin and E. R. Davidson, Phys. Rev. A 16, 1341 (1977), and references therein.
7. L. E. Nitzche and E. R. Davidson, Chem. Phys. Letters 58, 171 (1979).
8. R. L. Martin, E. R. Davidson, and D. F. Eggers, Chem. Phys. 38 341 (1979).
9. H. F. Jonkman, G. Van der Velde, and W. C. Nieuwpoort, SRC Atlas Symp. No. 4, Oxford, 1974; H. T. Jonkman, Broken Orbital Symmetry in Excited and Ionized States of Molecules, Ph.D. Thesis, University of Gronigen, 1975.
10. J. Goodman and L. E. Brus, J. Chem. Phys. 69, 1604 (1979); G. Ter Horst and J. Kommandeur, Chem. Phys. 44, 287 (1979); and references therein.
11. W. R. Wadt and W. A. Goddard, J. Am. Chem. Soc. 97, 2034 (1975); W. A. Wadt, W. A. Goddard, and T. H. Dunning, J. Chem. Phys. 65, 438 (1976).
12. D. A. Kleier, R. L. Martin, and W. R. Wadt, manuscript in prepartion.

A PRACTICAL GUIDE TO APPLICATIONS OF THE MCSCF METHOD

John H. Detrich and Arnold C. Wahl
Science Applications, Inc.
Oak Brook, Illinois 60521

I. INTRODUCTION

The considerable virtues of the MCSCF method are now well recognized and need not be exhorted here. Also the formal analysis and its recent development underlying MCSCF technology is well covered in the major papers of this workshop and in the literature[1-21] and will not be repeated. Rather, this paper will concentrate on the practical application of the MCSCF method with specific attention to three important and somewhat subjective questions:

1) the choice of configurations
2) the choice of initial orbitals, and
3) the design of the basis set.

We will begin with an inventory of MCSCF calculations intended to illustrate various choices made on specific systems and to assess how extensively the MCSCF method has been applied (Table I). We shall summarize the choices made in regard to the above three questions in MCSCF calculations made since 1966. These questions have often been downplayed or only vaguely discussed in published work even though they must be faced and are essential to the success and in fact mere execution of any MCSCF calculation. We hope that a guide to this practical experience will prove useful in implementing the new generation of MCSCF technology now under development in various laboratories.

Table I is not intended to be complete however we believe that it is representative of practical MCSCF applications performed during the past one and one half decades. We welcome comments and criticism drawing our attention to calculations and novel ideas we may have missed.

II. APPLICATIONS INVENTORY

It is valuable as a guide and background for current workers to inventory the MCSCF calculations performed to date on various systems using the several MCSCF strategies. In Table I we have arranged these calculations in chronological order to track progress and indicated some characteristics of the computations performed. These are:

1) how were the MCSCF configurations chosen.
2) how many configurations were included in the MCSCF process.
3) how many configurations were included in the CI step.
4) was a potential curve or surface generated.
5) were properties generated.
6) were excited states handled.

We see from this table that since the first molecular MCSCF calculations there has been a steady evolution and improvement and the method has been applied to excited states , electron affinities , extended to heavy diatomics , and polyatomics , orbital and mixing coefficients have been coupled , more and more configurations have been included in the MCSCF process , and the method has been joined with a larger CI step using the generated MCSCF orbitals (MCSCF/CI) .

An evaluation of each calculation reported will not be made in this paper however it should be pointed out that many of the calculations cited achieved chemical accuracy, some did not, and others were model test calculations. The reader is encouraged to assess for himself from the literature the merits of various approaches.

All of the above represents progress however it is not clear what the most effective thrust of MCSCF technology should be. With this in mind, we will proceed to our discussion.

III. CHOICE OF CONFIGURATIONS

Among the major advantages of the MCSCF method are the possibility of relatively short configurations list and the resulting chemical interpretability. In Table I the criteria used for the selection of configurations for inclusion in the MCSCF process is indicated. This selection process has ranged from one of trial and error to schemes which incorporate the physics of molecular formation and excitation. Below we give a brief description with key references for each configuration selection scheme:

T&E: Trial and Error: configurations are tried in the MCSCF list and retained if they produce a lowering of the energy or significant mixing.

INTU: Intuition: configurations are selected on some chemical or physical basis specific to the particular system under study.

CI-SORT: A CI is performed and configurations with mixing coefficients larger than some threshold are included in the MCSCF list. New MCSCF orbitals are generated and sort can be repeated. Usually only a few CI iterations are required.

LR-IO-A: Left-right in-out, and angular correlation of an electron pair: 3 MCSCF orbitals are chosen to provide these three types of correlation for each electron pairs treated. Often used in selection schemes which follow.

PD: Proper Dissociation: the minimum number of configurations required to dissociate the molecule to Hartree-Fock fragments are included. This concept was introduced in the OVC and GVB methods. It would seem foolish not to include these configurations in any MCSCF list.

ODC: Optimized Double Configurations: Proper Dissociation when only 2 configurations are required.[1]

OVC: Optimized Valence Configurations: This scheme involves a "base" function which accounts for proper dissociation. In addition there are MECE (Molecular Extra Correlation Energy) configurations which vanish in the dissociated atoms. These include: Interatomic doubles, split-shell charge transfer, and valence charge redistribution terms. The OVC selection scheme has been the most widely applied.[16,22]

ICF: Interacting correlated fragments: "designed for weakly interacting systems" - so far Be_2 and Mg_2 and He_2. A series of wave functions with increasing levels of correlation. "balanced" inclusion of inter and intrafragment correlation: in the Be_2 calculation this is $2s^2 \rightarrow 2p^2$ for Be atom and in the Mg_2 calculation this is $3s^2 \rightarrow 3p^2$; for He atom in He_2 it is $1s^2 \rightarrow 2s^2$ or $2p^2$.[23]

FRS: Full Reaction Space. Preceded the CAS model, and is very nearly equivalent to it. The FRS model seems to be more flexible in that open-shell orbitals in the dominant configuration need not be included in the reaction (or active) orbital space.[24]

CAS: or CASSCF: Complete Active Space SCF method: Basically, specification of the CSF list involves only symmetry specification and orbital lists for <u>inactive</u> (or core) orbitals and <u>active</u> orbitals. The active orbitals generate <u>all possible</u> occupancies and spin-couplings in construction of the CSF list.[25]

SRFV: Symmetry Restricted Full Valence: Not completely clear. Apparently includes only molecular orbitals (symmetry restricted) which can be manufactured from the valence shells of the constituent atoms.[21]

PE: Pair excitations. Consists of including configurations generated by replacing each doubly and each singly occupied orbital in the reference configuration, one at a time, by another orbital.[5]

INTER: Interatomic excitations only: involves double excitation with each electron coming from separate fragments or atoms. Also referred to as dispersion excitations.[26]

INTRA: Interatomic excitations only: involves simultaneous excitation from same fragment or atom. Accounts for changes in atomic correlation which is important for high accuracy and long range potentials.[26]

INTER-INTRA: Inter-Intra coupling: involves excitation which couple the above two classes of configurations. Important effect in long range potentials.[26]

DCE: Differential Correlation Energy Method: includes only those configurations which contribute directly to the correlation energy difference. These include configurations involving the "active" orbital and an "excluded space" effect. The idea is similar in spirit to the OVC method for potential curves; it may provide a reasonable starting point for MCSCF calculations of spectra, electron affinities and ionization potentials.[27]

GVB: Generalized Valence Bond. In its original form, this approach deals with the atomic orbital distortions describing bond formation between constituent atoms in the molecule, hence it has more in common with unrestricted Hartree-Fock type approaches than MCSCF. However, the GVB wave function can be re-expressed as a superposition of configurations and handled by MCSCF methods. This approach has relatively recently been adopted for computational convenience and brings GVB into the MCSCF mainstream.[28]

CI-SEL: Starting with m basic MC configuration functions (CF) (Proper dissociation, curve crossings, etc.) all singly and double excited CFs are generated, and for each one the (m + 1) dimensional Hamiltonian matrix is diagonalized. CFs giving an energy lowering in excess of some threshold are added to the MC list. New MCSCF mixing coefficients and orbitals are generated, and new selections are performed, until the MC list stabilizes. For excited states of a given symmetry, occasional selections for the next lower root are performed and important CFs are also added to the MC list.[13]

ASE: All Single Excitations. Applied to atoms, and consists of including configurations generated by replacing single orbitals in the reference configuration by a (correlation) orbital.[29]

MO-ORDER: Molecular Orbital Order. Assigns rough molecular orbital energies to orbitals on the basis of intuition and experience, then selects configurations with particularly low energy sums. Appears well-adapted to handle near-degeneracy and valence effects.[30]

CMC: Complete Multi-Configuration. all double excitations from the reference configuration with singlet spin paring. There appear to be two distinct varients, depending on whether localized or symmetry adopted orbitals are used. 31

BMC: Bond Multi-Configuration. Double excitations to correlate orbital pairs for a single bond, i.e., intrabond correlation. 31

BDMC: Bond and Dispersion Multi-Configuration. Adds to BMC double excitations one each from differing bond pairs, i.e., cross-bond correlation. 31

There is much to be learned from the work listed in Table I and the concepts discussed above. Among the most obvious lessons are:

1) Changes in the correlation energy of a properly chosen core can be neglected.

2) It is reasonable as a starting point for the MCSCF/CI procedure to select configurations to account for the Molecular Extra Correlation Energy (include proper dissociation).

3) Often additional configurations must be included to describe significant changes in the intra-atomic correlation energy.

4) High accuracy (better than .1 ev) requires many more configurations than those prescribed by simple rules. The MCSCF/CI procedure appears to be an effective method for achieving such higher accuracy.

5) Molecular properties can often be affected by configurations which do little to the potential curve or surface.

6) Sometimes even qualitatively correct results cannot be obtained with simple configuration selection rules.

7) What is learned about configuration choice from the first member of an isovalent series can be used efficiently in subsequent members.

8) It is difficult to apply simple configuration selection rules to polyatomic systems.

IV. BASIS SETS FOR MCSCF CALCULATIONS

Most of the art of basis set design developed for single configuration SCF calculations is directly applicable to MCSCF calculations. The reason for this is that effective correlating orbitals span the same space as the orbital they are meant to correlate. Thus basis sets of double zeta quality[32] with polarization functions are usually adequate. Obvious attention however, must be given to proper design and testing of basis sets for excited states, ions, and weak interactions. For weak interactions the basis set expansion errors in the dissociated atoms must be significantly smaller than the magnitude of the interactions. For heavy atoms this requires basis sets better than double zeta quality. With the expansion method there are often special cases and surprises.

Finally, since it seems to be perennially forgotten, good basis sets are essential for achieving reliable and accurate results. CI or MCSCF procedures, no matter how extensive, cannot compensate for expansion errors in any predictable way.

V. CHOICE OF INITIAL ORBITALS

No single decision is as crucial to the successful convergence of both the SCF and MCSCF processes as the choice of starting orbitals. Many schemes have been tried which include:

1) the use of virtual SCF orbitals,

2) orbitals obtained by maximizing coupling integral between dominant and excited configurations,

3) orbitals obtained by maximizing overlap of correlating orbitals with correlated orbitals.

4) orbitals based on intuition,

5) the concept of in-out, left-right, and angular correlation, and

6) other ideas.

Some of these have been done manually and others automatically programmed. Experience has shown that it is well worth the effort to start the MCSCF process with the converged SCF orbitals.[33] Also starting guesses can be aided by converged calculations on other states of the same system or other members of an isovalent series. None of the above schemes is fool proof and the construction of initial orbitals in the MCSCF process is clearly an area needing significant attention.

VI. CONVERGENCE AND TIMING

One of the major advantages of the new MCSCF technology which has been emerging over the past several years is anticipated increases in the efficiency of convergence of the MCSCF process. In order to evaluate and document these performances, some sort of standard test case is required to monitor convergence behavior. Not only is the number of configurations important but also the time required for each interaction and the starting orbitals used. In Table II we suggest a list of parameters which should be included in convergence assessment and would be useful to have in the literature. Sample values are given for CN^- and CO_2^- using the CDC 6600 version[34] of the BISON-MC[35] code.

VII. FUTURE GOALS

Considerable attention is being given by several groups to the efficient formalism, implementation, and convergence properties of the MCSCF procedure. We can realistically anticipate continuing improvements and success in these areas. However, it is our belief that more attention is needed in the three areas which have formed the focus of this paper. Specifically:

1) Systematic configuration selection rules should be developed and explored. Ideally successful rules should be amenable to automation and be incorporated as a user option into the computer program.

2) Systematic initial orbital selection rules should be developed and explored. These also should be incorporated into the computer program.

3) Despite our successes, basis set uncertainties for new situations or molecular environments remain a limitation on accuracy, reliability, and level of confidence. Much greater effort should be given to the development of numerical or semi-numerical procedures for molecules thus eliminating basis set anxiety.

We hope that in this brief contribution to the NRCC-MCSCF workshop that we have drawn your attention to not only past work, but also to several important challenges facing computational chemistry in this new decade.

ACKNOWLEDGEMENTS

We are grateful to our colleagues for their assistance in assembling Table I.

REFERENCES

1. G. Das and A. C. Wahl, J. Chem. Phys. 44, 87 (1966).
2. J. Hinze and C. C. J. Roothaan, Progr. Theor. Phys. Suppl. 40, 37 (1967).
3. B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968).
4. B. Levy, Chem. Phys. Lett. 4, 17 (1969).
5. F. Grein and T. C. Chang, Chem. Phys. Lett.

12. 12, 44 (1971).

6. G. Das and A. C. Wahl, J. Chem. Phys. 56, 1769 (1972).
7. K. K. Docken and J. Hinze, J. Chem. Phys. 57, 4928 (1972).
8. G. Das, J. Chem. Phys. 58, 5104 (1973).
9. J. Hinze, J. Chem. Phys. 59, 6424 (1973).
10. G. Das, T. Janis and A. C. Wahl, J. Chem. Phys. 61, 1274 (1974).
11. F. Grein and A. Banerjee, Int. J. Quantum Chem. 59, 147 (1975).
12. A. Banerjee and F. Grein, Int. J. Quantum Chem. 10, 123 (1976).
13. A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054 (1977).
14. T. C. Chang and W. H. E. Schwartz, Theor. Chim. Acta 44, 45 (1977).
15. P. J. A. Ruttink and J. H. van Lenthe, Theor. Chim. Acta 44, 97 (1977).
16. E. Dalgaard and P. Jørgensen, J. Chem. Phys. 69, 3833 (1978).
17. E. Dalgaard, Chem. Phys. Lett. 65, 559 (1979)
18. D. L. Yeager and P. Jørgensen, J. Chem. Phys. 71, 755 (1979).
19. C. C. J. Roothaan, J. Detrich and D. G. Hopper, Int. J. Quantum Chem. S13, 93 (1979).
20. K. Ruedenberg, L. M. Cheung and S. T. Elbert, Int. J. Quantum Chem. 16, 1069 (1979)
21. B. H. Lengsfield III, J. Chem. Phys. 73, 382 (1980).
22. A. C. Wahl and G. Das, Advances in Quantum Chem., 5, 261 (1970): also see calculations on OH and CO listed in table I and Methods of Electronics Structure Theory page 51, (Plenum 1977.)
23. B. Liu and A. D. McLean, J. Chem. Phys. 72, 3418 (1980).
24. K. Ruedenberg and K. R. Sundberg, Quantum Science-Methods and Structure, J.-L. Calais, O. Goscinski, J. Linderberg and Y. Öhrn, Eds. (Plenum, New York, 1976.)
25. B. O. Roos, P. R. Taylor and P. E. M. Siegbahn, Chem. Phys. 48, 157 (1980).
26. See Calculations on He_2, ArH and Be_2 listed in table I.
27. A. C. Wahl, E. S. Sachs and J. Detrich, Int. J. Quantum Chem. (in press).

28. W. J. Hunt, P. J. Hay and W. A. Goddard III., J. Chem. Phys. 57, 738 (1972).

29. F. Grein, Int. J. Quantum Chem. 5, 165 (1971).

30. W. B. England, Chem. Phys. (in press).

31. (1) A. Veillard, Theoret.Chim. Acta 7, 133 (1967).
(2) B. Levy, Int. J. Quantum Chem. 4, 297 (1970).
B. Levy, Chem. Phys. Lett. 18, 59 (1973)

32. By double zeta quality we mean 2 or more functions for each loop of the orbital. For Gaussians this means two or more contracted basis functions.

33. There are notable cases where this is not true, for example O_3. However, usually the single Hartree-Fock configuration is the dominant one in the CI expansion and it is well worth the effort of converging it first.

34. The BISON-MC MCSCF code was linked to POLYATOM by Allan Hinds and installed on the CDC6600 by Richard Hosteny.

35. G. Das and A. C. Wahl, ANL Report (1972)

TABLE I. A DIRECTORY OF ATOMIC AND MOLECULAR MCSCF AND MCSCF/CI CALCULATIONS

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
N		2				No	D.R. Hartree, W. Hartree and B. Swirles, Phil. Trans. Roy. Soc. (London) A238, 229 (1939).
He		6				No	J. Vizbaraite, J. Sirancs, V. Kaveckis and A. Jucys, Opt.i Spectroscopiya, 1,277 (1956).
Li		5				No	A. Jucys, J. Vizbaraite, J. Batarunas and V. Keveckis, Lietuvos TSR Mokslu, Akad. Darbai, Ser.B 2,3 (1958).
Be		10				No	
H_2	PD	2		Yes	No	No	E.R. Davidson and L. L. Jones, J. Chem. Phys. 37 1918 (1962).
H_2	ODC OVC	2 4		Yes Yes	Yes Yes	No	G. Das and A. C. Wahl. J. Chem. Phys. 44, 87 (1966).
Li_2	ODC OVC	2 4		Yes	Yes	No	G. Das and A. C. Wahl, J. Chem. Phys. 44,87 (1966).
F_2	ODC	2		Yes	Yes	No	G. Das and A. C. Wahl, J. Chem. Phys. 44, 87 (1966).
Be-O isoelectronic series	$2s^2 \to 2p^2$	2			No	No	E. Clementi and A. Veillard J. Chem. Phys. 44, 3050 (1966).
Li_2	OVC	7		Yes	Yes	No	G. Das, J. Chem. Phys. 46, 1568 (1967).
$He, H^-, Li^+, Be^{2+}, B^{3+}, C^{4+}, N^{5+}, O^{6+}$	INTU	10,10,9,9,9,9, 9,9			Various	No	N. Sabelli and J. Hinze J. Chem. Phys. 50, 684 (1969).
H_2	ODC	2		Yes	Yes	Yes	W. T. Zemke, P. G. Lykos, and A. C. Wahl, J. Chem. Phys. 51, 5635 (1969).
F_2	OVC	6		Yes	Yes	No	G. Das and A. C. Wahl, J. Chem. Phys. 56, 3532 (1972): Phys. Rev. Letts. 24, 440 (1970).

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
He_2	INTER INTRA	9 20		Van der Waals	No	No	P. Bertoncini and A. C. Wahl, Phys. Rev. Letts 25, 991 (1970).
H_2CO_2	CMC BDMC			No	No	No	B. Levy, Int. J. Quant. Chem. 4, 297 (1970).
NaLi	BMC OVC	4		Yes	Yes	No	P. J. Bertoncini, G. Das, and A. C. Wahl, J. Chem. Phys. 52. 5112 (1971).
LiHe	INTER INTRA	5 7		Van der Waals	No	No	G. Das and A. C. Wahl, Phys. Rev. A4, 825 (1971).
He,Li,Be	ASE & PE	3,5			No	No	F. Grein and T. C. Chang, Chem. Phys. Letters, 12, 44 (1971).
LiH	INTER	15,5,4,5		Yes	Various	Yes (Weighted avg. ref)	K. K. Docken and J. Hinze, J. Chem. Phys. 57, 4928, 4936 (1972).
He,Be	ASE & PE	2-7			No	Yes	T. C. Chang and F. Grein, J. Chem. Phys., 57, 5270 (1972).
B	CAS	28,17,26,12			Osc. Strengths	Yes	Z. Sibincic, Phys. Rev.A5, 1150 (1972).
N	INTU, TE	9,5			Fine structure	No	J. Detrich, Phys. Rev. A5, 2014 (1972).
O_2^-	OVC-INTRA			No	Electron affinity	No	W. T. Zemke, G. Das, and A. C. Wahl, Chem. Phys. Lett. 14, 310 (1972).
O_2^-	PD	8		Yes	No	Yes	M. Krauss, D. Neumann, A. C. Wahl, G. Das, & W. Zemke, Phys. Rev A7, 69 (1973).
NH	OVC	8		Yes	Yes	No	W. J. Stevens. J. Chem. Phys. 58, 1264 (1973).
CH	OVC	8		Yes	Yes	No	P. Julienne and M. Krauss, Molecules in the Environment, J. Wiley Page 354 (1973).
CH_4 C_2H_4 C_2H_6	CMC BMC BDMC			No	No	No	B. Levy, Chem. Phys. Lett. 18, 59 (1973).
H_2O	OVC	8		Yes	No	Yes	R. P. Hosteny, A. R. Hinds, A. C. Wahl & M. Krauss, Chem. Phys. Lettr. 23, 9 (1973)
Li,Be,B,C	INTRA	1-8		No	Polarizabilities	No	A. J. Stevens & F. P. Billingsly, II Phys. Rev. A82,236 (1973).
Be-F	ASE	3-4			No	No	F. Grein and T. C. Chang, J. Phys. B. Atom, Molec.Phys. 6,L237 (1973).

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
CH	INTU	8	4147	Yes	Various	No	G. C. Lie, J. Hinze and B. Liu, J. Chem. Phys. 59, 1872, 1887 (1973).
CN	OVC	17		Yes	Yes	Yes	G. Das & A. C. Wahl, J. Chem. Phys. 61 1274 (1974).
AlO	OVC	10		Yes	Yes	Yes	
ArH	INTER OT	10 2		Van der Waals	Cross Section	No	A. F. Wagner, G. Das and A. C. Wahl, J. Chem. Phys. 60, 1885 (1974).
Ne_2	INTER	9		Van der Waals	Yes	No	W. J. Stevens, A. C. Wahl, M.A. Gardner and A. M. Karo, J. Chem. Phys. 60, 2195 (1974)
OH	OVC	14		Yes	Yes	No	W. J. Stevens, G. Das, and A. C. Wahl, D. Neumann and M. Krauss, J. Chem. Phys. 61, 3686 (1974).
FH	OVC	8		Yes	Yes	No	D. Neumann and M. Krauss, Mol. Phys. 27 917 (1974).
CO	OVC	7		Yes	Yes	No	F. P. Billingsley and M. Krauss J. Chem. Phys. 60, 4130 (1974).
CO	OVC	9		Yes	Yes	No	F. P. Billingsley and M. Krauss, J. Chem. Phys. 60, 4130 (1974).
NO^+	OVC	9		Yes	Yes	No	F. P. Billingsley and M. Krauss, J. Chem. Phys. 60, 2767 (1974).
He - F	ASE & PE	2-4			No	No	F. Grein and A. Banerjee, Chem. Phys. Lett. 25, 255 (1974).
Be	INTU	2-5			No	Yes	F. Grein and A. Banerjee, Chem. Phys. Letters, 31, 281 (1975).
Li,B	INTU	2-8			No	Yes	F. Grein and A. Banerjee, Int. J. Quantum. Chem. Symp. 9, 147 (1975).
O	INTU, TE	10			Fine structure	No	J. Detrich, Phys. Rev. A11, 1498 (1975).
NO_2	OVC	18	99	Yes	Yes	Yes	G. D. Gillespie, A. V. Khan, A. C. Wahl, R. P. Hosteny, and M. Krauss, J. Chem. Phys. 63, 3425 (1975).
VH	OVC	4		Yes	Yes	Yes	G. A. Henderson, G. Das, and A. C. Wahl, J. Chem. Phys. 63, 2805, (1975).
Li_2H	INTU	8	30	Yes	No	Yes	W. B. England, N. H. Sabelli, and A. C. Wahl, J. Chem. Phys. 63, 4596 (1975).

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
CO_2	CI-SORT	5	200	No	Yes	Yes	W. B. England, B. J. Rosenberg, P. J. Fortune, and A. C. Wahl, J. Chem. Phys. 65, 694 (1976).
HgH	INTER INTRA	7		Yes	Yes	Yes	G. Das, and A. C. Wahl, J. Chem. Phys. 64, 4672 (1976).
H_2O^+	CI-SORT	17	99	Yes		Yes	P. J. Fortune, B. J. Rosenberg, and A. C. Wahl, J. Chem. Phys. 65, 2201 (1976).
OH	OVC	17	61	Yes	Yes	No	J. O. Arnold, E. E. Whiting, and L. F. Sharbough J. Chem. Phys. 64, 3251 (1976).
NH	CI-SEL	14		Yes	Yes	Yes	A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054, 2509 (1977).
Na_2	OVC	8		Yes			W. J. Stevens, M. M. Hessel, P. J. Bertoncini and A. C. Wahl, J. Chem. Phys. 66, 1477 (1977).
NO_2, NO_2^+, NO_2^-	CI-SORT	15	99	No			G. Das, A. F. Wagner, and A. C. Wahl, J. Chem. Phys. 68, 4917 (1978).
ClO	OVC	61	825	Yes	Yes	No	J. O. Arnold, E. E. Whiting and S. R. Langoff, J. Chem. Phys. 66, 4459 (1977).
LiH_2	INTU, TE	3,4	168	Yes	No	No	P. J. A. Ruttink and J. H. van Lengthe, Theoret. Chem. Acta 44, 97 (1977).
Li_2	OVC			Yes			M. L. Olson and O. D. Konowalow, Chem. Phys, 21, 393 (1977) ibid., 22, 29 (1977); D. D. Konowalow and M. L. Olson, J. Chem. Phys. 67, 590 (1977); 71, 450 (1979).
I_2	OVC-INTRA			Yes		Yes	G. Das and A. C. Wahl, J. Chem. Phys. 69, 53 (1978).
LiH_2	OVC	10		Yes	Cross Section	No	A. F. Wagner, A. C. Wahl A. M. Karo, and R. Kreijci J. Chem. Phys. 69, 3756 (1978).
HD_2^+	INTU,TE	2		Yes	No	No	J.H. van Lenthe and P.J.A Ruttink Chem.Phys. Lett. 56, 20 (1978).
O_2^-	OVC + INTRA	76		Yes		Yes	G. Das, A. C. Wahl, W. T. Zemke and W. C. Stwalley J. Chem. Phys. 68, 4252 (1978).
HeH, NeH, ArH KrH, XeH	INTER INTRA INTER-INTRA			Van der Waals	Cross Section	No	G. Das, A. F. Wagner, and A. C. Wahl, J. Chem. Phys. 68, 4917 (1978).
NH	CI-SEL	9,13		Yes	Yes	Yes	A. Banerjee and F.Grein,Chem. Phys. 35, 119 (1978).

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
CaO	INTU	2		Yes	No	No	C. W. Bauschlicher and D. R. Yarkony, J. Chem. Phys. 68, 3990 (1978).
KOH	MO Order	7	1465	Yes	Yes	No	W. England, J. Chem. Phys. 68, 4898 (1978).
C_2H_4	FRS			Yes	No	No	L. J. Cheung, K. R. Sundberg, and K. Ruedenberg J. Am. Chem. Soc. 100, 8024 (1978), Int. J. Quantum Chem. 16, 1103 (1979).
C_3H_6							S. Kato and K. Morokuma, Chem.Phys. Lett. 65,19 (1979).
N_2	OVC	20		Yes			D. Neumann and M. Krauss, Mol. Phys. 37, 1661 (1979).
O_2^+	PD	12	3854	No	Transition Moments	No	K. Tanka and M. Yoshimine, J. Chem. Phys. 70, 1626 (1979).
SiO	INTU	45	2500	Yes	No	No	S. R. Langhoff and J. O. Arnold J. Chem. Phys. 70, 852 (1979).
CF	GVB	2-8		Yes	Yes	No	T. H. Dunning, Jr., W. P. White R. M. Pitzer, and C. W. Mathews, J. Mol. Spect. 75, 297 (1979).
CO_2	CI-Sort	15	3573	No	Yes	Yes	W. B. England and W. C. Ermler, J. Chem. Phys. 70, 1711 (1979).
Na_2	OVC	12		Yes	Yes	Yes	D. D. Konowalow, M. E. Rosenkrantz, and M. L. Olson, J. Chem. Phys. 72, 2612 (1980).
C_3H_4	INTU, TE	10	2047	No	Yes	No	J. H. van Lenthe and J. L. M. Smits, Recueil, J. Roy, Netherlands Chem. Soc. 99, 130 (1980).
CuO	PD	2		Yes	Yes	No	D. H. W. den Boer and E. W. Keleveld, Chem. Phys. Lett. 69, 389 (1980).
O_2	CAS			No	No	Yes	D. L. Yeager and P. Jørgensen, Mol. Phys. 39, 587 (1980).
Be		2		No	No	Yes	D. L. Yeager and P. Jørgensen, Mol. Phys. 39, 487 (1980).
BeO		4,3		No	No	Yes	C. W. Bauschlicher, Jr. and D. R. Yarkony, J. Chem. Phys. 72, 1138 (1980).
Li_2	GVB	2		No	No	Yes	B. H. Lengsfield III, J. Chem. Phys. 73, 382 (1980).

Atom or Molecule	Configuration Choice Criteria	Number of MCSCF Configurations	Number of MCSCF/CI Configurations	Potential Curve or Surface	Properties	Excited States	References
Li_4		4		No	No	Yes	B. H. Lengsfield III, J. Chem. Phys. 73, 382 (1980).
BeO	SRFV (CAS)	81		No	No	Yes	B. H. Lengsfield III, J. Chem. Phys. 73, 382 (1980)/
CaO	MO Order	8	6102	Yes	Yes	Yes	W. B. England, Chem. Phys. in press).
Be_2	CAS	1832	11688	Yes	No	No	Blomberg, Siegbahn and Roos (pre-print).
Be		106	1148		No	Yes	R. Shepard & J. Simons (pre-print).
C_5H_4	INTU	2		Yes	No	No	R. Shepard and J. Simons (pre-print).
C_3H_2	CI-SORT	6,4	100	Yes	No	No	R. Shepard, A. Banerjee and J. Simons (pre-print).
N_2	CAS	726,948,950		Yes	No	Yes	B. Roos (pre-print).
CCO	GVB	16	2523	Yes	No	Yes	S. P. Walch, J. Chem. Phys. 72, 5679 (1980).
$OH + H_2$	GVB	6	4560	Yes	No	No	S. P. Walch and T. H. Dunning Jr., J. Chem. 72, 1303 (1980).
Zn,Cd,Hg	LR-IO-A				Polarizabilities	Yes	M. E. Rosenkrantz, W. J. Stevens M. Krauss and D. Konowalow, J. Chem. Phys. 72 2525 (1980).
Be_2	ICF			Yes	No	No	B. Liu and A. D. McLean, J. Chem. Phys. 72, 3418 (1980).
H_2O	CAS	1380			Various	Yes	B. Roos (pre-print).
PbO	OVC	9		Yes			H. Basch, M. Krauss, W. J. Stevens (to be published).
CN	DCE	17	41	Yes		No	A. C. Wahl, E. S. Sachs and J. Detrich, Int. J. Quantum, Chem. (1980).
CN^-	OCE	15	26	Yes	Electron Affinity	No	A. C. Wahl, E. S. Sachs, and J. Detrich, Int. J. Quantum Chem. (1980).

TABLE II

SYSTEM	BASIS SET	#ORBS	#CONFGS	ITERATION TIMES		#ITERATIONS	THRESHOLD VECTORS	TOTAL MCSCF		MACHINE	CODE
				CPU	IO			CPU	IO		
CN^-	4s2p1d +diffuse	9	15	20	25	23	10^{-5}	530	600	CDC 6600	BISON-MC
CO_2^-	3s2p	15	19	34	16	12	10^{-5}	434	207	CDC 6600	BISON-MC

USGPO 1981 - 789-009 57:11