LBL-13917

Lawrence Berkeley Laboratory

UNIVERSITY OF CALIFORNIA

RECEIVED
LAWRENCE
BERKELEY LABORATORY

FEB 22 1982

LIBRARY AND
DOCUMENTS SECTION

Materials & Molecular Research Division

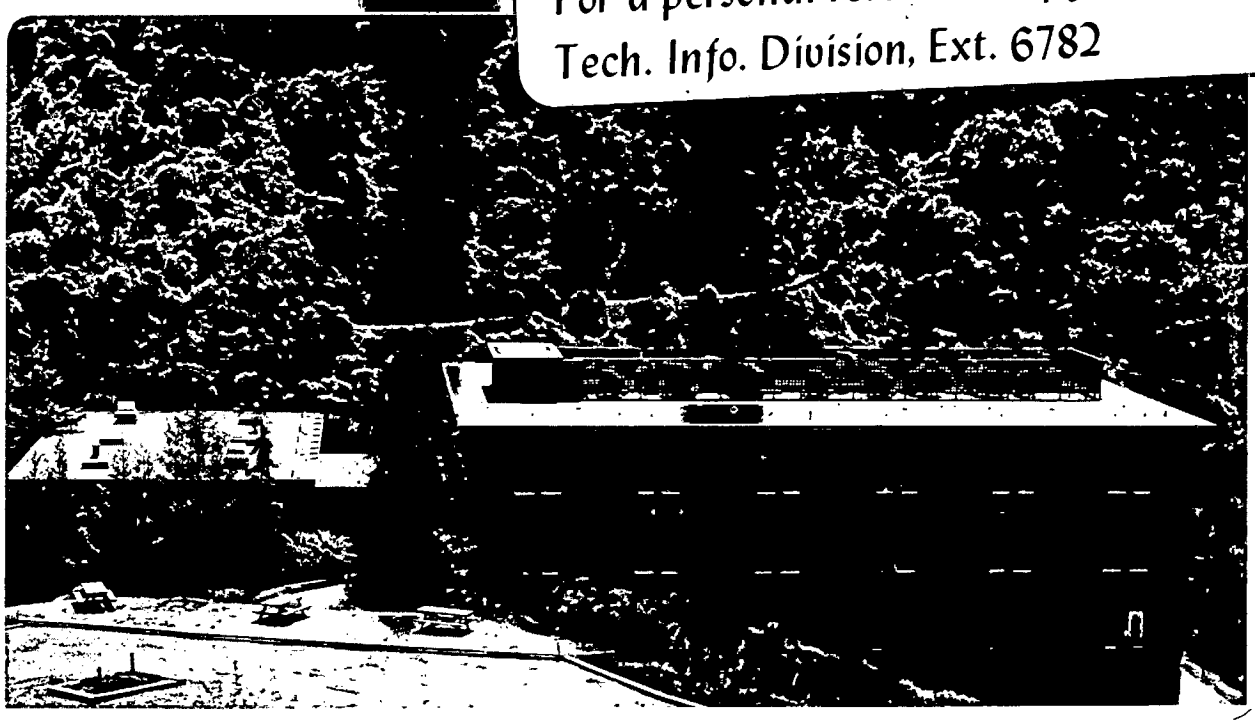
THE LOOP-DRIVEN GRAPHICAL UNITARY GROUP APPROACH MULTICONFIGURATION SELF-CONSISTENT-FIELD METHOD WITH APPLICATIONS TO WATER, OZONE AND CYCLOPROPENYL RADICAL

William David Laidig
(Ph.D. thesis)

December 1981

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

LBL-13917 c.2

LEGAL NOTICE

This book was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Lawrence Berkeley Laboratory is an equal opportunity employer.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

LBL-13917

The Loop-Driven Graphical Unitary Group Approach Multiconfiguration Self-Consistent-Field Method with Applications to Water, Ozone and Cyclopropenyl Radical

William David Laidig

Ph.D. Thesis

December 1981

Materials & Molecular Research Division
Lawrence Berkeley Laboratory
University of California
Berkeley, CA 94720

This research 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 by the U.S. National Sciences Foundation, Grant CHE-7622621, and the Robert A. Welch Foundation.

LBL-13917

The Loop-Driven Graphical Unitary Group Approach Multiconfiguration Self-Consistent-Field Method with Applications to Water, Ozone and Cyclopropenyl Radical

By

William David Laidig

Lawrence Berkeley Laboratory, University of California
Berkeley, California 94720

Abstract

A multiconfiguration self-consistent-field (MCSCF) method based on the loop-driven graphical unitary group approach (LDGUGA) has been developed that combines the best features of the LDGUGA configuration interaction (CI) procedure with the ability to optimize the molecular orbitals. In contrast to traditional MCSCF procedures which are restricted to small configuration sets, our method can handle very large numbers of configurations. Unlike other large configuration list MCSCF procedures, however, the CI expansion is not restricted to being of a particular form (such as in the CASSCF method). The formalism we adopt is an iterative two-step process, that is, during each iteration the CI coefficients are optimized for a fixed guess at the orbital expansion coefficients and then improved orbitals are computed using these CI coefficients. At the heart of the orbital optimization step is a novel algorithm for the fast and efficient

generation and storage of the one- and two-particle density matrices and the use of these density matrices in first and approximate second-order optimization procedures is discussed. To test the applicability of the LDGUGA MCSCF method a series of CI and MCSCF studies on various molecular systems were performed. These include an investigation on the effect of inclusion of higher than double excitations in MCSCF and CI calculations for the water molecule, a comparison of various large MCSCF wavefunctions for ozone and the study of non-physical symmetry breaking in cyclopropenyl radical.

Henry F. Schaefer III

Acknowledgments

First and foremost I would like to thank Fritz Schaefer for all his help and support during my four years at Berkeley. I am particularly grateful to him for seriously listening to my ideas and encouraging me to pursue the better ones. Next I would like to thank Dr. Bernard R. Brooks for introducing me to the loop-driven graphical unitary group approach and for all his help in practical computer coding. Bernie Brooks was instrumental in the development of his MCSCF procedure since the two of us worked out the formalism together and much of the subsequent program was written by him. I also owe a lot of my understanding of MCSCF theory to Professor Nick Handy who helped me with a number of difficult concepts. In addition, he informed me of the investigation on the importance of triple excitations in the neon atom carried out by Dr. Stephen Wilson which was one factor leading to our higher excitation CI and MCSCF study on the water molecule. I would also like to thank Paul Saxe who performed a number of these calculations and who along with Bernie Brooks wrote our two-particle density matrix program. I am indebted to Dr. Yoshihiro Osamura for his Hessian matrix element derivation which was very useful in writing section IIH. I also wish to thank Professor Richard Stratt, Dr. Michel Dupuis, Mr. Douglas Fox, Professor Isaiah Shavitt, Dr. Yukio Yamaguchi, Dr. John Goddard, Dr. Byron Lengsfield, Dr. Rodney Bartlett and Dr. Mark Vincent for many productive discussions. Thanks are also due to Cheryn Gliebe for all her help during the past year. Finally I would like to thank Carol Hacker who besides doing an excellent job typing this thesis

has been very helpful throughout my stay at Berkeley.

This research was supported by the U.S. National Science Foundation, Grant CHE-7622621, the Robert A. Welch Foundation, and 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.

Table of Contents

Page

I. Introduction 1

II. The LDGUGA MCSCF Method 6

A. Review of the Berkeley Loop-Driven Graphical Unitary Group Methodology 11

B. First-Order Orbital Optimization Procedure--Method 1 24

C. First-Order Orbital Optimization Procedure--Method 2 30

D. An Overview of the Berkeley MCSCF Program System . . . 39

E. Generation of the Two-Particle Density Matrix 47

F. Construction of the Lagrangian Matrix 51

G. Avoidance of Computing the Complete Set of M.O. Integrals 54

H. Construction and Use of the Hessian Matrix 55

I. A Proposed Extrapolation Procedure to Improve Convergence 60

III. An MCSCF Study of the Effect of Higher Excitations in Water 63

A. Theoretical 66

B. Water Energetic Results 69

C. Comparison with Perturbation Theory Results 75

D. Wavefunction Analysis 83

E. Conclusions for Water 88

IV. An MCSCF Study on the Ground State of Ozone 90

A. Theoretical 91

B. Ozone Energetic Results 94

C. Conclusions for Ozone 100

Page

V. An Examination of the Lowest $^2E''$ State of the Cyclopropenyl Radical 102

A. Theoretical 103

B. Results and Discussion 105

C. Conclusions for the Cyclopropenyl Radical 111

VI. Concluding Remarks 114

Appendix. A_k and B_k Theory 116

References 120

I. Introduction

The multi-configuration self-consistent field (MCSCF) method has had a long and successful history since its original inception by Frenkel[1] in 1934. The first atomic wavefunction was reported only a few years later in 1939 for oxygen by Hartree, Hartree and Swirles,[2] but the method was not of much practical use until the advent of high speed digital computers. By the mid 1960's the time was ripe for the computational development of the MCSCF procedure. In 1966 the first molecular wavefunctions were reported by Das and Wahl[3] for H_2, Li_2 and F_2. During the next few years there was much activity within the field and a number of significant advances were made.[4-9] Interest in the MCSCF approach remained high throughout the 70's and shows no signs of waning to date. Many theoretical and methodological contributions were made during this period[10-17] and the MCSCF procedure has become a standard tool among quantum chemists.

Within the last five to ten years MCSCF calculations involving small numbers of configurations (usually between 2 and 20) have become common place. Wavefunctions of this sort tend to recover much of the correlation energy, that is the difference between the SCF and the true nonrelativistic energy. They have the additional benefit of being relatively compact compared to typical configuration interaction (CI) wavefunctions. Calculations of the latter variety typically involve thousands of configurations and are thus much more difficult to qualitatively interpret. MCSCF procedures are also quite good for describing bond breaking processes and obtaining

accurate equilibrium geometries.[12]

The MCSCF method, however, does have a number of drawbacks. One of its more serious problems pertains to calculation convergence. Molecular MCSCF wavefunctions are always solved for via iterative procedures which quite often show unacceptable rates of convergence. Certain systems, typically those involving single excitation configurations, will not converge at all either displaying oscillatory or divergent behavior. To overcome this, elaborate extrapolation or damping procedures have been employed with mixed success. Another limitation of the MCSCF formalism is that it has been traditionally restricted to small configuration sets. The reasons behind this are both computational and philosophical in nature. Up until 1980 the largest reported MCSCF wavefunction contained 60 configurations.[18]

Within the last few years major advances have been made in both of these areas. With regard to convergence behavior the quadratic MCSCF procedure has been developed by a number of workers[19-28] and appears to be able to surmount many of the previous difficulties. Using this technique along with super CI[13,15-17] and/or augmented Hessian[21,29] schemes most systems converge within 5 to 10 iterations to 8 decimal places in the energy. Even pathological systems containing numerous single-excitations have been shown to converge. Another advantage of this method is that excited states of the same symmetry can be solved for with little additional effect.[23] Restriction of MCSCF calculations to small configuration sets has been virtually eliminated via the loop-driven graphical unitary group approach (LDGUGA)MCSCF method.[30-32] Our current implementation

can handle CI expansions composed of tens of thousands of configurations with the only limits being the size of core memory and the amount of computer time available. As a preliminary calculation, an MCSCF wavefunction for the lowest triplet state of cyclopropyne containing 10,115 configurations was computed.[31] This wavefunction was composed of over two orders of magnitude as many configurations as the previously largest reported MCSCF. Since then MCSCF calculations on water and ozone have employed even larger numbers of configurations.[33,34] At about the same time as our initial MCSCF work, Roos and Siegbahn developed the complete active space (CAS) MCSCF method.[35-37] Their procedure also handles large CI expansions, but these are restricted to a certain form namely a complete CI within a given subspace. CI expansions of this type have many special properties which are exploited in their CASSCF program. The major drawback of this method is that the length of the CI increases extremely rapidly as the size of the active space is enlarged. Preliminary calculations by Roos and Siegbahn were reported for N_2 employing 726 configurations and since that time a number of other systems have been studied using the CAS method. Lastly, recent work by Lengsfield on adapting the quadratic MCSCF procedure to moderately large configuration lists has been reported.[23] In that paper calculations on the X $^1\Sigma^+$ and $2^1\Sigma^+$ states of HF containing 1436 configurations are discussed.

In this dissertation the theory behind the LDGUGA MCSCF method will be discussed and the basic philosophical and computational differences inherent to our scheme compared to traditional MCSCF

procedures will be illuminated. Since our biggest departure is in the size of CI expansions that can be handled, the techniques that are employed to deal with large configuration sets will be stressed. Among these are an efficient loop-driven CI based on the graphical unitary group approach, the need for the fast generation and storage of the two-particle density matrix and its subsequent use in an orbital optimization step that has reasonable convergence properties. Other computational aspects of the system will also be presented such as restricting the integral transformation to transforming only the necessary integrals and the efficient formation and use of the Lagrangian and Hessian matrices. In a more general vein the strengths and weaknesses of the LDGUGA system will be presented. This will include what types of configuration sets can and cannot be used, the restriction of the point group symmetry to D_{2h} and its subgroups and the ability to include higher than double excitation configurations.

In addition to the theory a number of applications will be presented. The first of these involves a series of CI and MCSCF calculations on the ground state of water. The main purpose for this study was to try and elucidate the role that higher than double excitations play in CI calculations using different orbital basis sets. The most extensive computation performed on H_2O was an MCSCF containing 17,678 configurations which is the largest MCSCF result obtained for an molecule to date. Our next investigation dealt with the lowest state of ozone. In this study a number of different size MCSCF and CI wavefunctions are computed and compared. The

quality of these wavefunctions are examined in the light of their associated total energy and their ability to correctly describe the biradical nature of O_3. Lastly the $^2E''$ state of cyclopropenyl radical is probed with a number of MCSCF calculations within the valence space. Pairs of computations are performed in $^2A_2$ and $^2B_2$ symmetries and the splitting between the two resultant energies is compared.

II. The LDGUGA MCSCF Method

In developing the loop-driven graphical unitary group approach (LDGUGA) MCSCF method[30-32] a basic departure from the current philosophy was made. In a single configuration SCF calculation the orbitals are variationally determined such that the energy is minimized and for many chemical properties such as the equilibrium geometry the SCF procedure often gives a reasonably accurate result. However, for a number of systems the SCF wavefunction is a poor approximation to the true wavefunction, at least in certain regions of the potential energy surface. The multi-configuration SCF (MCSCF) method was designed to overcome these deficiencies by including the configurations that are necessary to qualitatively describe these areas of the surface. An example of this is the description of bond breaking in a molecule. Usually a number of configurations are needed to represent this process accurately including at least those configurations needed to describe both the starting molecule and its fragments.[12] Traditionally in MCSCF calculations only a small number of configurations are employed, usually the minimum required for qualitative accuracy. One problem that often arises is that since the configuration set must be predetermined, it is difficult to ascertain in advance which configurations are needed over the entire range of the potential energy surface to be examined. Another deficiency is that in many cases the wavefunction itself is not of sufficient accuracy and additional extensive CI calculations using these MCSCF orbitals must be carried out.

In implementing our MCSCF method we strove to eliminate the restriction to small configurations sets and thus minimize the resultant problems of configuration selection and wavefunction accuracy. Before going into how this was accomplished certail details of the MCSCF formalism in general need to be pointed out. In any MCSCF procedure both the orbitals and the CI expansion coefficients are variationally determined. The subsequent energy expression obtained using this wavefunction is

$$E = \sum_{IJ} C_I C_J \sum_{ij\lambda\mu} (a_{IJ}^{ij} C_{\lambda i} C_{\mu j} \langle\lambda|\hat{h}|\mu\rangle + \sum_{\substack{k\ell \\ \nu\sigma}} b_{IJ}^{ijk\ell} C_{\lambda i} C_{\mu j} C_{\nu k} C_{\sigma \ell} [\lambda\mu;\nu\sigma]) \tag{1}$$

where C_I and C_J are CI expansion coefficients, $C_{\lambda i}$, $C_{\mu j}$, $C_{\mu k}$ and $C_{\nu k}$ are orbital coefficients, a_{IJ}^{ij} and $b_{IJ}^{ijk\ell}$ are coupling constants and $\langle\lambda|h|\mu\rangle$ and $[\lambda\mu;\nu\sigma]$ are the one and two electron atomic orbital (AO) integrals. Equation (1) can be rewritten in terms of molecular orbital (MO) integrals as

$$E = \sum_{IJ} C_I C_J \sum_{ij} (a_{IJ}^{ij} \langle i|\hat{h}|j\rangle + \sum_{k\ell} b_{IJ}^{ijk\ell} [ij;k\ell]) \tag{2}$$

with

$$\langle i|\hat{h}|j\rangle = \sum_{\lambda\mu} C_{\lambda i} C_{\mu j} \langle\lambda|\hat{h}|\mu\rangle \tag{3}$$

$$[ij;k\ell] = \sum_{\lambda\mu\nu\sigma} C_{\lambda i} C_{\mu j} C_{\nu k} C_{\sigma \ell} [\lambda\mu;\nu\sigma] \quad . \tag{4}$$

Since it is not possible to directly solve for the coefficients C and c an iterative procedure must be employed. There are two general types of iterative schemes possible: one, in which corrections to both the trial CI and orbital coefficients are solved for simultaneously and two, where first the CI coefficients calculated and secondly the changes in the trial orbital coefficients are found. The former procedure is called a one-step method and the latter a two-step method. Because our MCSCF program and the large majority of others are of the two-step variety only two-step procedures will be considered.

In a typical two-step MCSCF method the calculation can generally be divided into four separate tasks. The first is the transformation of the AO integrals to the MO basis, the second is the solution of the CI secular equation to yield the expansion coefficients, the third is the construction of the one- and two-particle density matrices and the fourth is the computation of the corrections to the trial orbitals. The entire process is then repeated until the desired level of convergence has been achieved. In a traditional MCSCF calculation employing only a small number of configurations the bulk of the computer time is spent performing the integral transformation. A substantial fraction of this time may be required by the orbital optimization step, but solving the secular equation and building the density matrices takes essentially no time. This caused us to think that if an MCSCF method were built around a very fast CI (and density matrix construction) program that quite a large number of configurations could be handled before the CI

solution and density matrix computation steps began to dominate the calculation time. This is the basic idea behind the LDGUGA MCSCF approach.

At the heart of our MCSCF package is the loop-driven CI program originally developed by Brooks and Schaefer[38] with subsequent improvements by Laidig and Saxe.[39] This CI program utilizes the graphical unitary group approach (GUGA) pioneered by Paldus[40-42] and Shavitt[42-45] within the loop-driven framework. Since its introduction in 1978 the original CI system and its later modified versions have demonstrated the efficiency and applicability inherent in the LDGUGA method. During this period theoretical studies including the effects of electron correlation have been carried out on a number of molecules using the LDGUGA CI programs. Among these are the investigation of the stability of cyclopropyne,[46] a study of the triplet excited states of HCN,[47] the determination of the harmonic vibrational frequencies of NH_4^+,[48] an examination of aluminum-carbon bonding[49] and the prediction that methylnitrene is stable[50] to name just a few. Currently CI calculations including up to 40,000 configurations are routinely performed with the only size restrictions being the amount of available core memory.

Designing our MCSCF procedure around the loop-driven CI routines places certain restrictions on the type of configuration sets that can be employed. This is because only configuration lists that can be represented in a compact numerical form called a distinct row table (DRT) can be handled by the CI. At present DRT's containing all single or all single and double excitation

configurations from one reference can be generated. This can also be done for a large number of two reference cases. In addition classes of higher excitations can be included such as all single, double and triple or all single, double, triple and quadruple excitations from one or two references. Full CIs can also be performed in any subspace and all single plus double excitations can be generated from this set if desired. For a more complete discussion of the capabilities of the DRT construction program see Brooks' Ph.D. Thesis.[51] Work is currently in progress that will allow the DRT for all single and double excitations from an arbitrary set of reference functions to be generated.[52] This will greatly increase the flexibility of the method.

Within the DRT framework only limited configuration selection can be performed without destroying the advantages of the loop-driven formalism. One type of selection which does not significantly affect the performance of the CI is restriction of the configuration list to the subset which are Hartree-Fock interacting.[53] This does not alter closed-shell calculations, but can drastically reduce the number of allowed configurations from open-shell references without adversely affecting the energy.[54,55] Also a certain amount of selection can be obtained by restricting excitations into and out of specific orbitals. Within the MCSCF, orbitals can be divided into three basic types: core, virtual and active. Core orbitals are doubly occupied in all configurations in the CI and are often used to represent the inner shell orbitals in atoms and molecules. Virtuals orbitals are never occupied in the CI and usually have high

orbital energies. An orbital is active if at least some excitations are allowed into or from it and thus it has an occupancy somewhere between 0 and 2. Any DRT that can be generated in the full space can also be limited to just the active orbitals. In many cases this can dramatically reduce the size of the CI expansion.

Another aspect of the LDGUGA MCSCF method that helps in reducing the length of the configuration lists is the use of Abelian point group symmetry.[51] This is very important since the full Hamiltonian is block-diagonal with each submatrix of a particular symmetry type. For example, in water the molecule has C_{2v} symmetry and the Hamiltonian consists of four sub-matrices one for each of the four distinct symmetry types: A_1, A_2, B_1 and B_2. Because the matrix is block-diagonal the individual sub-matrices can be independently diagonalized with the resultant eigenvalues corresponding to eigenvectors of the block's symmetry type.

In our program only those configurations that are of the same symmetry as that of the reference(s) are generated. This leads to a much smaller secular program to be solved. If the molecular symmetry belongs to a degenerate point group the calculation can still be carried out in the highest symmetry non-degenerate sub-group of the degenerate point group. Since we normally do calculations on triatomic or larger systems this is seldom a problem.

A. Review of the Berkeley Loop-Driven Graphical Unitary Group Methodology

Before going on to the particulars of our MCSCF method it would be helpful to the reader if certain major concepts of the LDGUGA

approach were examined. As was mentioned earlier the key factor in designing a fast MCSCF procedure that deals with large configuration sets is in having a very efficient CI program at its heart.[51] One such program is in the Berkeley CI system of programs[38,39] and this became the basis of our MCSCF routines. The two central new ideas within this CI method are the generation and use of the distinct row table (DRT) and basing the construction of contributions to the Hamiltonian matrix around loops.[43-45,51] First the definition and subsequent use of the DRT will be discussed.

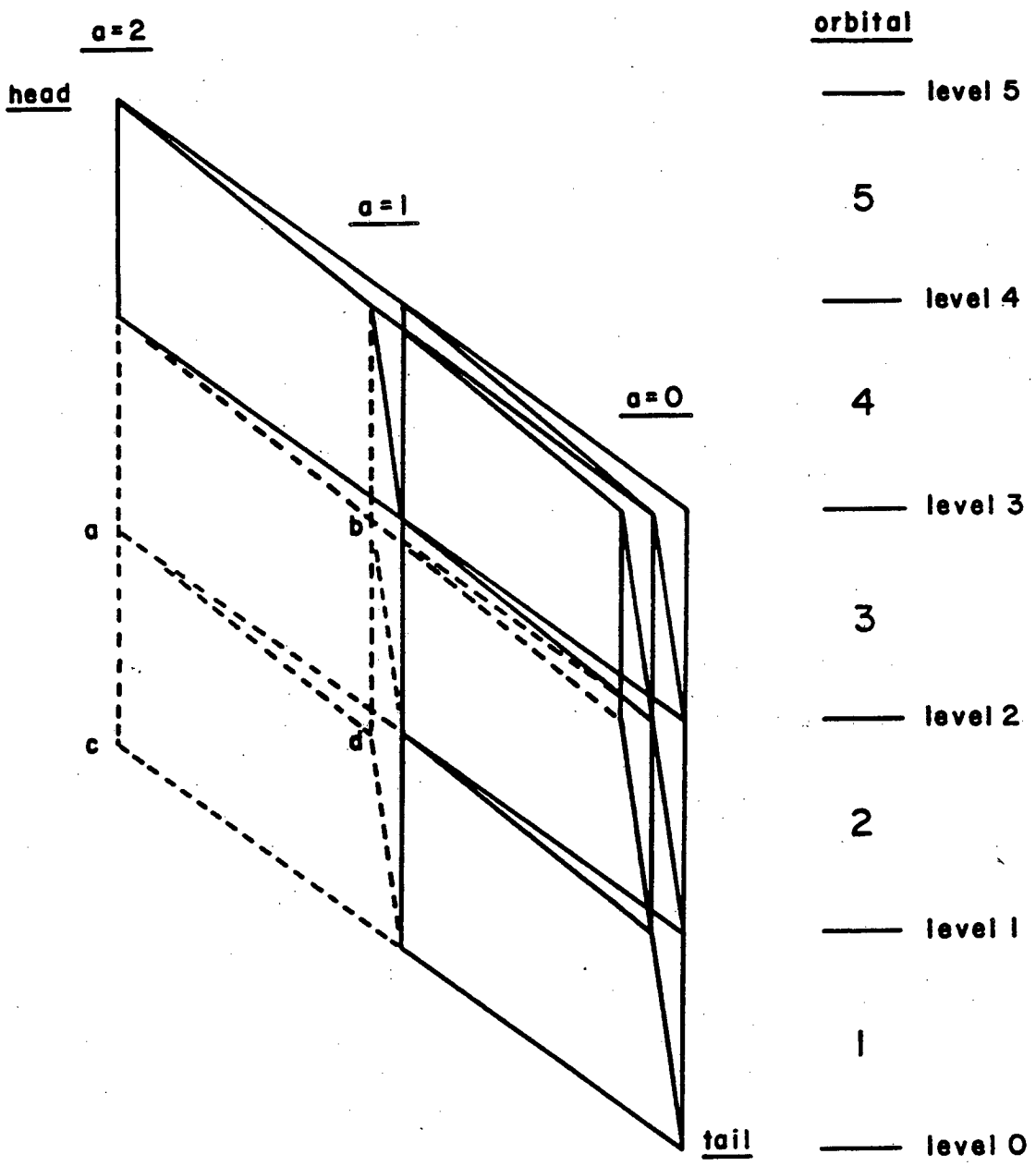
Basically the DRT is the numerical analog of a Shavitt graph[43,44] which is a pictoral representation of a given configuration set. A sample graph is shown in Figure I. As one can see from the figure a Shavitt graph is composed of a framework of dots called distinct rows and a series of line segments connecting them labeled arcs. The top most row is the graph head while the bottom vertex is the graph tail. Any path from the head to the tail which passes through one row per level defines a particular Gelfand[56] state or configuration and is called a walk. The totality of all walks define the configuration set. Each row can be defined by three parameters a_{ij}, b_{ij} and c_{ij} such that

$$a_{ij} + b_{ij} + c_{ij} = i \tag{5}$$

$$a_{ij} = 1/2\ N_{ij} - S_{ij} \tag{6}$$

$$b_{ij} = 2\ S_{ij} \tag{7}$$

Figure I. Sample Shavitt Graph.

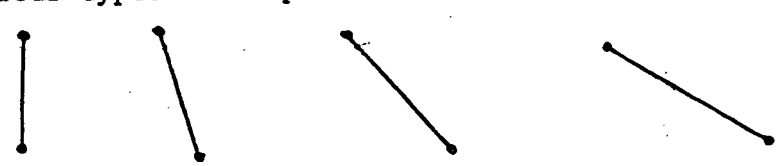


All solid lines - DRT of all single and double excitations from the reference $4a_1^2 5a_1^2$.

All solid plus dashed lines - DRT for the full CI.

with i being the level index and j being the vertex index on that particular level. N_{ij} and S_{ij} are the number of electrons and the total spin of the ij^{th} row.[40] Looking again at the figure one notices that the rows are separated into three columns by the value of a. A group of vertices on the same level with identical a values differ in b with the right-most point being b=0. Moving to the left each row within the group has $\Delta b=+1$.

To work out a Shavitt graph, one first calculates the values of a, b and c needed for the graph head using N equal to the number of electrons in the particular system and S equal to the total spin. The graph tail is always assigned a = b = c = 0. For the ground state of water, for example, N=10 and S=0. This yields a=5 and b=0. c is determined by how many orbitas are used in the calculations. Once the graph head is found segments connecting adjacent levels are drawn. Four distinct shapes are possible: one, a vertical line segment corresponding to an orbital occupancy of zero in which $\Delta a=0$ and $\Delta b=0$; two, a slightly slanted line segment with $\Delta a=0$ and $\Delta b=-1$ corresponding to an alpha occupied orbital; three, a more slanted line segment corresponding to a beta occupied orbital where $\Delta a=-1$ and $\Delta b=+1$; and four, a nearly 45° line segment with $\Delta a=-1$ and $\Delta b=0$ corresponding to a doubly occupied orbital. Pictorially the four types are represented as

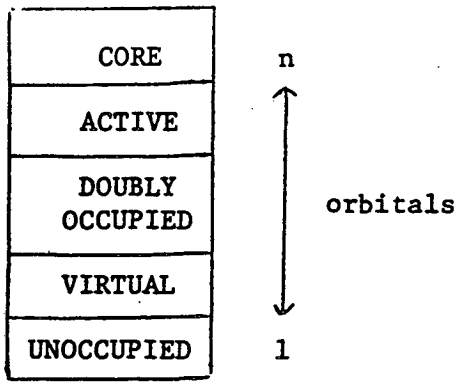


1 2 3 4 . (8)

In a given Shavitt graph all allowed walks begin at the graph head and end at the graph tail. In Figure I the graph for the case of a system with 5 orbitals, 4 electrons and a total spin of zero is shown composed of both the solid and dashed line segments. If the dashed arcs are eliminated the CI is restricted to all single and double replacement configurations from the rightmost walk which is the reference.

As discussed above the Shavitt graph contains all configurations allowed in the CI. To actually perform the computation, however, some method must be found to identify each particular configuration. The location scheme[43,44] that was adopted assigns a unique number m to each configuration with the leftmost walk defined as m=1. If two configurations coincide from the graph head down until some level i then the state with the lower m value is the leftmost below this level. An ordering of this sort is labeled lexical. Each configuration is composed of n arcs where n is the number of orbitals in the CI step. A property of this lexical ordering is that a value can be assigned to each arc such that the index m can be computed for each walk by summing over the n parameters associated with its components. These parameters are called the arc weights.

Besides defining the configuration ordering the orbital sequence must also be determined. In the full CI case the ordering of the orbitals is essentially arbitrary. If a restricted form of the graph is used, however, it is advantageous to order the orbitals in a certain fashion. In our MCSCF program the ordering is as follows.



In this diagram core are the MCSCF core orbitals, active are the open-shell or partially filled orbitals in the reference(s), doubly occupied are the doubly occupied orbitals in the reference(s), virtual are the orbitals not occupied in the MCSCF and occupied are the unoccupied orbitals in the reference(s). Having the active block above the doubly-occupied orbitals and the doubly-occupied above the unoccupied orbitals is useful in non-closed shell and multi-reference cases. The placing of the core and virtual orbital blocks is not essential, but was chosen to accommodate restrictions within the CI gradient programs. In a future implementation the core and virtual sections will both be above the actives. The main reason for positioning the unoccupied orbitals at the bottom is that this allows the simplicity of the Shavitt graph in this region to be exploited.

If spatial symmetry is treated correctly or if the configuration list is restricted to only the Hartree-Fock interacting set[53] additional complications arise. In the case of spatial symmetry (only valid for nondegenerate point groups) all rows except the graph head and tail are expanded into a set of vertices, one for each symmetry type. Next all walks are generated with a total

symmetry equal to that of the reference(s). If no walks pass through a given point it is then eliminated from the graph. So, therefore, it now takes four quantities to define a row: a, b, c and s where s is the a symmetry label. A fifth parameter t is necessary if the Hartree-Fock interacting space restrictions are to be implemented. Using these five parameters then any Shavitt graph or corresponding DRT can be constructed subject to the limitations presented in the previous section.[51]

In the computer the Shavitt graph is stored as its numerical analog, the distinct row table. Instead of vertices and arcs a series of arrays are formed which describe the rows and how they interconnect. Example DRT's and a detailed description of the necessary arrays are presented in reference 51.

The second major concept which needs to be examined is that of the loop.[43,44] When two configurations m and m' are compared in the DRT they will coincide above a certain level j and also below a specific level i-1. (In the worst case the jth level will be the graph head and the i-1 level will be the graph tail). The section between the levels j and i-1 which need not coincide is called a loop with the loop head being the distinct row at level j and the loop tail associated with the row at level i-1. The segments joining the graph and loop heads is labeled the upper walk and the segment between the graph and loop tails the lower walk. An example of a typical loop is shown in Figure II.

The reason why the loop concept is important in GUGA implementations can best be understood if the Hamiltonian operator is cast

into its second quantized form[45]

$$H = \sum_{ij} h_{ij}E_{ij} + \sum_{ijk\ell} [ij;k\ell]e_{ij,k\ell} \tag{9}$$

where the summation is over the orbitals allowed in the CI and E_{ij} and $e_{ij,k\ell}$ are the one- and two-body unitary group operators. The two-body operator can be written in terms of the one-body operators as

$$e_{ij,k\ell} = E_{ij}E_{k\ell} - \delta_{jk}E_{i\ell} \tag{10}$$

with the one-body operator defined by

$$E_{ij} = \sum_{\sigma} X^{\dagger}_{i\sigma} X_{j\sigma} \tag{11}$$

where $X^{\dagger}_{i\sigma}$ and $X_{j\sigma}$ are the spin-orbital creation and annihilation operators respectively. The summation index σ in (7) is over α and β spin. The Hamiltonian matrix element between configurations m and m' can now be expressed as

$$H_{mm'} = \sum_{ij} h_{ij}\langle m|E_{ij}|m'\rangle + \frac{1}{2}\sum_{ijk\ell} [ij;k\ell]\langle m|e_{ij,k\ell}|m'\rangle \quad . \tag{12}$$

Now if equation (2) is examined it becomes obvious that the matrix elements $\langle m|E_{ij}|m'\rangle$ and $\frac{1}{2}\langle m|e_{ij,k\ell}|m'\rangle$ are just the coupling coefficients $a^{ij}_{mm'}$ and $b^{ijk\ell}_{mm'}$ respectively defined in equation (1). Since for the majority of configurations m and m' in a large CI there are only two non-zero coupling coefficients, it is found that $\langle m|E_{ij}|m'\rangle$ and $\langle m|e_{ij,k\ell}|m'\rangle$ are zero for most choices of i, j, k

Figure II. Example of a loop.

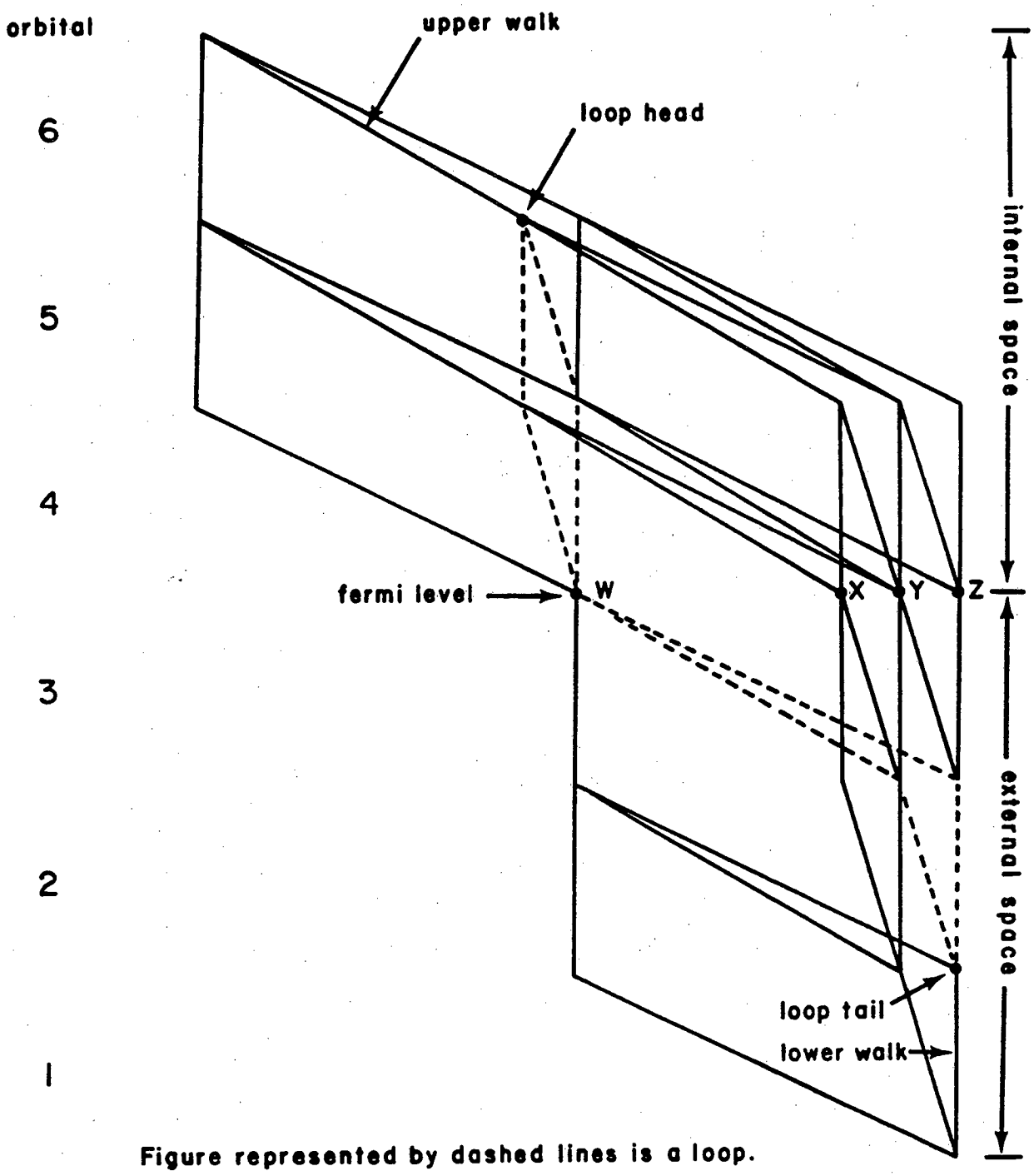


Figure represented by dashed lines is a loop.

and ℓ. It can be shown that if $\langle m|E_{ij}|m'\rangle$ is to ever by non-zero the configurations m and m' must coincide above level j and below level i-1. Furthermore, using equation (10)

$$\langle m|e_{ij,k\ell}|m'\rangle = \sum_{m''} \langle m|E_{ij}|m''\rangle\langle m''|E_{k\ell}|m'\rangle + \delta_{k\ell} \langle m|E_{i\ell}|m'\rangle \tag{13}$$

and, therefore, if $\langle m|e_{ij,k\ell}|m'\rangle$ is to be non-zero the arcs above the largest and below the smallest value of i, j, k and ℓ must be coincident also. If the element $\langle m|E_{ij}|m'\rangle$ or $\langle m|e_{ij,k\ell}|m'\rangle$ is non-zero it is also known that their values only depend on the shape and position of the loop in the DRT and not in the upper and lower walk portions of the two configurations involved. What this suggests is then a loop-driven CI procedure.[38] In this scheme instead of constructing each element of the Hamiltonian matrix separately as in a conventional configuration-driven program or sequentially reading through the integral list and forming all contributions to the H matrix that are possible as in an integral-driven CI, the set of pairs of distinct rows on different levels are looped over and all non-zero loops between a given pair are generated. Once a loop is formed its value is computed. What is necessary now is a procedure for computing this loop value and also how to determine which configuration pairs involve a given loop.

In the LDGUGA system of programs the loop value is found through the use of segments.[45] A segment is defined as the pair of arcs, one from each configuration m and m', between any two adjacent levels.

For example, the two-body loop that is pictured in Figure II has four segments. These segments are classified according to their shapes and to each distinct shape a segment value W is associated. These segment values are either constants or depend on the b value at the top of the m' arc. The unitary group matrix elements can be expressed as[41,45]

$$\langle m|E_{ij}|m'\rangle = \prod_{k=i}^{j} W(Q_k,b_k) \tag{14}$$

$$\langle m|e_{ij,k\ell}|m'\rangle = \prod_{p\varepsilon S_1} W(T_p,b_p)\left(\sum_{x=Q,1} \prod_{p\varepsilon S_2} W_x(T_p,b_p)\right) \tag{15}$$

where S_2 and S_1 are the interacting and non-interacting regions respectively and W_x are segment values that also depend on a spin parameter x. In our program then the loop values are found through first multiplying and adding the appropriate segment values together to form the coupling coefficients and then multiplying these by the appropriate MO integrals. Since many different loops share common segments it is not necessary to entirely calculate each coupling coefficient from scratch. In the loop-driven program loops are generated in an order that takes advantage of this fact.

Once a loop value has been computed how do we determine which Hamiltonian matrix element(s) it belongs to? This is done using the loop breakdown algorithm of Brooks and Schaefer.[38,51] A loop will contribute to the H matrix elements between any pair of configurations formed by the loop plus one of its upper and one of its lower walks. In Figure II the example loop has only a single upper and lower walk

so it contributes to only one Hamiltonian element (the matrix H is stored in triangular form), however, in general loops can have n upper and m lower walks and will contribute to n x m matrix elements. In this algorithm the lexical numbers of the leftmost configurations containing the left and right sides of the given loop are computed by summing up the individual arc weights. These configurations are called the primary upper walks of the left and right sides respectively and are symbolized as IUWK and JUWK. Due to the nature of lexical ordering all subsequent configurations formed with the same upper walk but different lower walks are sequential. This, unfortunately, is not the case for a fixed lower walk with variable upper walks. If, however the Shavitt graph is inverted so the graph tail is now the head and vice-versa, a different lexical ordering can be defined. Using this scheme all configurations with the same lower walk but different upper walks are now sequential. This ordering is labeled reverse lexical. In our program an array INDX is computed and stored which converts from lexical to reverse lexical ordering. If NUWK and NLWK are defined as the number of upper and lower walks respectively and VAL is the loop value then the Hamiltonian matrix element contributes for a given loop can be constructed using the following code.

```
DO   100  I = 1, NLWK
II = INDX (IUWK)
JJ = INDX (JUWK)
```

```
      DO    90   J = 1, NUWK
      H(II,JJ) = VAL
      II = II + 1
      JJ = JJ + 1

90    Continue
      IUWK = IUWK + 1
      JUWK = JUWK + 1

100   Continue
```

The advantage of this algorithm is that no multiplications or other complicated arithmetic are needed to compute the subsequent contributions to H. The calculation of the coupling coefficients through the use of segment values and the use of the loop breakdown algorithm to construct the Hamiltonian contributions are two of the features that make the LDGUGA formalism extremely fast and efficient.

Before closing this section one improvement to the original loop-driven method should be mentioned. In the original implementation[38] the loops were constructed by a tree-search algorithm which means that from a given loop head all possible segment shape combinations were investigated subject to a list of formation rules. The newest versions of the program[51,57] take advantage of the simplicity of the DRT among the unoccupied orbitals.[58-60] This section of the graph is called the external space while the remainder which includes the doubly-occupied and active orbitals is called the internal space. The interface level between the two sub-spaces is the Fermi level. In the external space there are at most four distinct rows per level

in the no symmetry single plus double excitation case. In Figure II these are labeled W, X, Y and Z. Because the structure of this space is so simple all possible types of loops (or partial loops if the loop begins in the internal space but ends in the external space) can be programmed up separately. When this was implemented substantial time savings were gained in the loop value construction stage of the CI step.

For a more complete review of the concepts presented here see either Brooks' thesis[51] or Shavitt's excellent summary article.[45]

B. First-Order Orbital Optimization Procedure--Method I.

Currently two methods for computing the changes in the orbitals from one iteration to the next have been implemented in the LDGUGA MCSCF program. The first involves the symmetrizing of the matrix of Lagrange multipliers.[30,31] In developing this procedure first note that the CI energy expressed in equation (1) can be written,

$$E = \sum_{ijk\ell} G_{ijk\ell}[ij;k\ell] + \sum_{ij} Q_{ij}\langle i|\hat{h}|j\rangle \tag{16}$$

where Q and G are the one and two particle density matrices defined below.

$$Q_{ij} = \sum_{IJ} C_I C_J \, a_{IJ}^{ij} \tag{17}$$

$$G_{ijk\ell} = \sum_{IJ} C_I C_J \, b_{IJ}^{ijk\ell} \quad . \tag{18}$$

The simple form E takes above is a convenient starting point for deriving the iterative equations needed to obtain the MCSCF solution.

In equation (16) both the density matrices and the MO integrals are independent of the choice of orbitals. However, for the present we will assume that the density matrices are independent of orbital rotation. Applying a unitary transformation, U, to the set of MO integrals in equation (16) then yields,

$$E \cong \sum_{abcd} G_{ijk\ell}\, U_{ai} U_{bj} U_{ck} U_{d\ell} [ab;cd] + \sum_{ab} Q_{ij} U_{qi} U_{bj} \langle a|\hat{h}|b\rangle \quad .(19)$$

If the orbitals used in the original CI (equation (16)) were a reasonable starting guess for the MCSCF orbitals and the transformation matrix $\underset{\sim}{U}$ was picked in such a way as to yield a "better" set of orbitals in an energetic sense then $\underset{\sim}{U}$ could be assumed to be close to the unit matrix $\underset{\sim}{1}$. If this was the case a two-step MCSCF procedure could be envisioned. First for a given set of MO integrals the CI step could be performed and the configuration coefficients could be used to construct $\underset{\sim}{G}$ and $\underset{\sim}{Q}$. Next using these fixed values for $\underset{\sim}{G}$ and $\underset{\sim}{Q}$ the MO to MO transformation matrix $\underset{\sim}{U}$ could be computed. Now the integrals are retransformed to this new MO basis and the steps are repeated until convergence. That is until $\underset{\sim}{U}$ is the unit matrix.

What is missing in the procedure outlined above is a method for determining the matrix $\underset{\sim}{U}$. If $\underset{\sim}{U}$ is close to $\underset{\sim}{1}$ then it is reasonable to express $\underset{\sim}{U}$ as,

$$\underset{\sim}{U} = \underset{\sim}{1} + \underset{\sim}{U}^{(1)} \tag{20}$$

where $\underset{\sim}{U}^{(1)}$ is just the difference between $\underset{\sim}{1}$ and $\underset{\sim}{U}$. The matrix $\underset{\sim}{U}^{(1)}$ is expected to be nearly antisymmetric since any unitary matrix can

be expressed as

$$\underset{\sim}{U} = e^{\underset{\sim}{T}} = \underset{\sim}{1} + \underset{\sim}{T} + \underset{\sim}{T}^2/2 + \ldots \tag{21}$$

where $\underset{\sim}{T}$ is antisymmetric. The higher powers of $\underset{\sim}{T}$ in this case will be negligible making

$$\underset{\sim}{U}^{(1)} \cong \underset{\sim}{T} \quad . \tag{22}$$

Expanding equation (19) using equation (20) and discarding terms higher than first order in $\underset{\sim}{U}^{(1)}$ yields for the first order change in the energy,

$$E^{(1)} \cong \sum_{ijk\ell} \sum_{abcd} G_{ijk\ell}[U^{(1)}_{ai}\ \delta_{bj}\delta_{ck}\delta_{d\ell} + \delta_{ai}U^{(1)}_{bj}\delta_{ck}\delta_{d\ell} + \delta_{ai}\delta_{bj}U^{(1)}_{ck}\delta_{d\ell} + \delta_{ai}\delta_{bj}\delta_{ck}U^{(1)}_{d\ell}][ab;cd] + \sum_{ij} \sum_{ab} Q_{ij}[U^{(1)}_{ai}\delta_{bj} + \delta_{ij}U^{(1)}_{bj}]\langle a|\hat{h}|b\rangle. \tag{23}$$

This equation can be further simplified to

$$E^{(1)} \cong \sum_{ijk\ell r} G_{ijk\ell}\ (4\ U^{(1)}_{ir}[rj;k\ell] + \sum_{ijr} Q_{ij}(2\ U^{(1)}_{ir})\langle r|\hat{h}|j\rangle\ . \tag{24}$$

If 4 times the two electron plus twice the one electron energies are added to both sides of equation (24) the result is

$$4E_2 + 2E_1 + E^{(1)} \cong \sum_{ijk\ell r} G_{ijk\ell}\ (4\ U_{ir}^{(1)})$$

$$[rj;k\ell] + \sum_{ijr} Q_{ir}(2\ U_{ir}^{(1)})\langle r|\hat{h}|j\rangle +$$

$$\sum_{ijk\ell r} G_{ijk\ell}\ (4\ \delta_{ir})[rj;k\ell] +$$

$$\sum_{ijr} Q_{ir}\ (2\ \delta_{ir})\langle r|\hat{h}|j\rangle \tag{25}$$

where E_1 and E_2 are the one and two electron energies. Collecting terms leaves

$$K + E^{(1)} \cong 4 \sum_{ijk\ell r} G_{ijk\ell}\ U_{ir}[rj;k\ell] +$$

$$2 \sum_{ijr} Q_{ir}U_{ir}\langle r|\hat{h}|j\rangle \tag{26}$$

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

$$X_{ir} \cong \sum_{jk\ell} 4\ G_{ijk\ell}\ [rj;k\ell] + \sum_{j} 2\ Q_{ij}\langle r|\hat{h}|j\rangle \quad . \tag{27}$$

Using this relation equation (26) can be reduced to

$$K + E^{(1)} \cong \sum_{ir} X_{ir}U_{ir} \tag{28}$$

What one would like now is a way to minimize the above equation to calculate the optimal matrix U. One straight forward way of doing this is to perform pair rotations on $\underset{\sim}{X}$. Rewriting equation (28) as

$$K + E^{(1)} \cong \sum_{ir} X_{ir} U_{ir} = \sum_{i} (\underset{\sim}{X}\underset{\sim}{U}^{T})_{ii} \tag{29}$$

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

$$U^{T} = \prod_{i>j,m} \theta_{ij,m} \tag{30}$$

where $\theta_{ij,m}$ is the mth 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 a given ij pair, multiply $\underset{\sim}{X}$ by the general pair rotation matrix $\underset{\sim}{\theta}_{ij}$ and take the trace of the resultant matrix. $\underset{\sim}{\theta}_{ij}$ is defined below as

$$\underset{\sim}{\theta}_{ij} = \begin{array}{c} \\ \\ \\ i \\ j \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccc} i & & & & & j \end{array} \\ \begin{bmatrix} 1 & & & & & 0 \\ & 1 & & & & \\ & & \ddots & & & \\ & & \cos\theta & & \sin\theta & \\ & & & \ddots & & \\ -\sin\theta & & & & \cos\theta & \\ & & & & & \ddots \\ 0 & & & & & & 1 \\ & & & & & & & 1 \end{bmatrix} \end{array} \tag{31}$$

The difference between the trace before and after rotation is

$$\mathrm{tr}(\underset{\sim}{X}\underset{\sim}{1}) - \mathrm{tr}(\underset{\sim}{X}\underset{\sim}{\theta}_{ij}) = X_{ii} + X_{jj} - X_{ii}\cos\theta + X_{ij}\sin\theta - X_{jj}\cos\theta - X_{ji}\sin\theta \quad . \tag{32}$$

Taking the derivative with respect to θ on both sides of equation (32) and equating to zero gives

$$(X_{jj} + X_{ii}) \sin\theta + (X_{ij} - X_{ji}) \cos\theta = 0 \quad . \tag{33}$$

Solving for tanθ yields

$$\tan\theta = \frac{X_{ji} - X_{ij}}{X_{ii} + X_{jj}} \quad . \tag{34}$$

Since θ is expected to be small and tanθ = θ as θ approaches zero equation (34) can be approximated by

$$\theta = \frac{X_{ji} - X_{ij}}{X_{ii} + X_{jj}} \quad . \tag{35}$$

This approximation saves considerable computation time since $\tan^{-1}$ operations are very slow on most computers. In practice using this approximation did not significantly effect the rate of convergence.

This method has been used with mixed success in a number of MCSCF calculations. Cases that converge adequately are typically those in which all orbitals are active, that is for all orbitals at least one configuration exists that excites one or two electrons into or out of each orbital. Cases where convergence is poor or non-existent usually involve at least a few core or virtual type orbitals. In Table I the results of several iterations of an MCSCF calculation on the ground state of water are presented. The MCSCF included all 361 single and double excitations from the SCF reference in the calculation set. More details on this calculation are given in

Section IIIB. In column 2 of the table the sum of the squares of the difference of the Lagrangian matrix is tabulated. This number is an indicator of convergence since the Lagrange multiplier matrix is symmetric upon convergence. The next column labeled overr. is what we call the overrotation or damping factor. Each predicted rotation in U is multiplied by this prefactor. If the calculation is diverging this factor is made less than 1.0 and if it is converging slowly a value greater than 1.0 generally improves the rate of convergence. Lastly the total energy and the coefficient of the reference configuration are tabulated. From perusing the table one notices reasonable convergence for the first 5 iterations but divergence for the next three. If however one returns to iteration 4 and changes the overrotation factor ot 0.5 the calculation converges. If appropriate overrotation factors are chosen convergence to 8 figures in the energy can be achieved in about 30-50 iterations.

C. <u>First-Order Orbital Optimization Procedure--Method 2</u>.

In developing a second approach it is necessary to examine the MCSCF procedure in a different light. This method was originally devised by Hinze[11] and we adopted it with only minor modifications.[30] For an MCSCF wavefunction

$$\langle\Psi|H|\Psi\rangle = E = \sum_{ij}^{m} Q_{ij}\langle i|\hat{h}|j\rangle + \sum_{ijk\ell}^{m} G_{ijk\ell}[ij;k\ell]$$

$$\sum_{ij}^{m} Q_{ij}\langle\psi_i|\hat{h}|\psi_j\rangle + \sum_{ijk\ell}^{m} G_{ijk\ell}\langle\psi_i\psi_k|\hat{g}|\psi_j\psi_\ell\rangle \qquad (36)$$

Table I. Convergence behavior of double-zeta water using method 1 employing 361 configurations.

Method 1

iter	sq. dif.	overr	E	C_1
1	1.1×10^{-4}	1.0	-76.15001468	.978735
2	4.4×10^{-5}	1.0	-76.15013779	.978465
3	2.8×10^{-5}	1.0	-76.15018605	.978258
4	2.6×10^{-5}	1.0	-76.15020970	.978100
5	3.4×10^{-5}	1.0	-76.15022125	.977981
6	5.7×10^{-5}	1.0	-76.15022293	.977879
7	1.1×10^{-4}	1.0	-76.15021200	.977801
8	2.1×10^{-4}	1.0	-76.15017986	.977239
4'	2.6×10^{-5}	0.5	-76.15020970	.978100
5'	1.3×10^{-5}	0.5	-76.15022211	.978042
6'	1.0×10^{-5}	0.5	-76.15022922	.977988
7'	8.9×10^{-6}	0.5	-76.15023498	.977938
8'	7.6×10^{-6}	0.5	-76.15023979	.977892
9'	6.6×10^{-6}	0.5	-76.15024384	.977851
10'	5.8×10^{-6}	0.5	-76.15024727	.977811

where ψ_i represents orbital i and $\hat{g}$ is the two electron operator and m is the number of total orbitals. (The two-electron integral written using the explicit orbital notation ψ_i is now $<(1)(2)|g(1,2)|(1)(2)>$ where 1 and w refer to electrons 1 and 2). The restrictions to equation (36) are that

$$<\psi_i|\psi_j> = \delta_{ij} \quad .$$

That is that the orbitals form an orthonormal set. Since Ψ is completely optimized with respect to the orbitals

$$\delta(\sum_{ij}^{m} Q_{ij}<\psi_i|\hat{h}|\psi_j> + \sum_{ijk\ell}^{m} G_{ijk\ell} <\psi_i\psi_k|\hat{g}|\psi_j\psi_\ell> - <\psi_i|\psi_j>\varepsilon_{ji}) = 0 \tag{37}$$

where the parameters ε_{ji} are the Lagrange multipliers. Expanding (37) to first order in δ and neglecting variations in the density matrices yields

$$\begin{aligned}\sum_{ij}^{m} Q_{ij}<\delta\psi_i|\hat{h}|\psi_j> + \sum_{ijk\ell}^{m} G_{ijk\ell}<\delta\psi_i\psi_k|\hat{g}|\psi_j\psi_\ell> - <\delta\psi_i|\psi_j>\varepsilon_{ji} \\ + \sum_{ijk\ell}^{m} G_{ijk\ell}<\psi_i\delta\psi_k|\hat{g}|\psi_j\psi_\ell> \\ + \sum_{ij}^{m} Q_{ij}<\psi_i|\hat{h}|\delta\psi_j> + \sum_{ijk\ell}^{m} G_{ijk\ell}<\psi_i\psi_k|\hat{g}|\delta\psi_i\psi_\ell> - <\psi_i|\delta\psi_j>\varepsilon_{ji} \\ + \sum_{ijk\ell}^{m} G_{ijk\ell}<\psi_i\psi_k|\hat{g}|\psi_j\delta\psi_\ell> = 0 \quad .\end{aligned} \tag{38}$$

Since the labels of the summation indices are arbitrary equation (38)

can be rewritten as

$$\sum_{ij}^{m} Q_{ij}\langle\delta\psi_i|\hat{h}|\psi_j\rangle + 2\sum_{ijk\ell}^{m} G_{ijk\ell}\langle\delta\psi_i\psi_k|\hat{g}|\psi_j\psi_\ell\rangle$$

$$+ \langle\delta\psi_i|\psi_j\rangle\varepsilon_{ji} + \text{c.c.} = 0 \quad . \tag{39}$$

where c.c. stands for the complex conjugate of the first two terms. The term that is explicitly written out may be now set equal to zero independently of the c.c. term. This is because a pure complex perturbation will change the sign of the complex conjugate part of equation (39) and therefore, both the explicit and c.c. terms can be separately equated to zero. The first term in (39) now becomes

$$\sum_{j}^{m}[\hat{h}\nu_{ij} + \sum_{k\ell}^{m} V_{k\ell}\, G_{ijk\ell}]\psi_j = \sum_{j}^{m}\psi_j\varepsilon_j \; ; \tag{40}$$

with

$$V_{k\ell}(1) = \int d(2)\; \psi_k^*(2)g(1,2)\psi_\ell(2) \quad . \tag{41}$$

Examining equation (36) one realizes that this system of equations is analogous to the Fock equations in SCF theory. The quantity in brackets in (40) is, therefore, defined as the generalized Fock operator F and the corresponding Fock-like equations for MCSCF wave-functions can be written as

$$\sum_{j}^{m} F_{ij}\psi_j = \sum_{j}^{m}\psi_j\varepsilon_{ji} \quad . \tag{42}$$

It is also known that $\varepsilon_{ji} = \varepsilon_{ij}^*$ (see Hinze[11] page 6430) or $\varepsilon_{ji} = \varepsilon_{ij}$ in our case since we only use real orbitals. Hinze also goes on to show that these are the necessary and sufficient conditions of convergence. If j is relabeled k in equation (42) and ψ_j is multiplied on both sides and integrated, the result is

$$\varepsilon_{ji} = \sum_k^m \langle\psi_j|F_{ik}|\psi_k\rangle \quad . \tag{43}$$

On convergence then

$$\sum_k^m (\langle\psi_j|F_{ik}|\psi_k\rangle - \langle\psi_i|F_{jk}|\psi_k\rangle) = 0 \quad . \tag{44}$$

It should now be noted that

$$X_{ij} = 2 \sum_k^m \langle\psi_j|F_{ik}|\psi_k\rangle \tag{45}$$

where $\underset{\sim}{X}$ is the matrix defined in equation (27) of the previous section as the Lagrangian matrix. On convergence, then, $\underset{\sim}{X}$ really is twice the transpose of the Lagrange multiplier matrix and, therefore, as stated in the preceeding section $\underset{\sim}{X}$ becomes symmetric upon convergence.

The MCSCF problem now reduces to finding an iterative procedure that upon convergence will satisfy equation (44). If a "better" set of orbitals is defined as

$$\psi' = \underset{\sim}{U}\psi \tag{46}$$

where $\underset{\sim}{U}$ is a unitary matrix and the Fock operator $\underset{\sim}{F}$ is assumed to be independent of the orbital basis, then equation (44) can be written in this new basis as

$$\sum_{k}^{m} (\langle\psi_j'|F_{ik}|\psi_k'\rangle - \langle\psi_i'|F_{jk}|\psi_k'\rangle) = 0 \quad . \tag{47}$$

Expanding the new orbitals in terms of the old yields

$$\sum_{k}^{m} \sum_{\ell n} (U_{\ell j}\langle\psi_\ell|F_{ik}|\psi_n\rangle - U_{\ell i} \langle\psi_\ell|F_{jk}|\psi_n\rangle) \; U_{nk} = 0 \quad . \tag{48}$$

The equation above can be solved directly for the set of all u's, however, this can be very time consuming. An alternative to this is to notice that U can be decomposed as follows

$$U = \prod_{i>j} U_{ij} \tag{49}$$

where U_{ij} is the pair rotation matrix between orbitals i and j (see pictorial representation in preceeding section). The diagonal elements of U_{ij} are unit and the off-diagonal elements are zero except for U_{ii}, U_{ij}, U_{ji} and U_{jj}. They have the following values

$$U_{ii} = U_{jj} = \cos\theta \tag{50}$$

$$U_{ij} = -U_{ji} = \sin\theta \tag{51}$$

where θ is the angle of rotation between the two orbitals. If the assumption is made that the set of rotation matrices commute then equation (48) can be solved independently for each pair of orbitals. For the pair i,j this equation becomes

$$\varepsilon'_{ji} - \varepsilon'_{ij} \cong \sum_{k\neq i,j}^{m} (s\langle\psi_i|F_{ik}|\psi_k\rangle + c\langle\psi_j|F_{ik}|\psi_k\rangle$$

$$- c\langle\psi_i|F_{jk}|\psi_k\rangle + s\langle\psi_j|F_{jk}|\psi_k\rangle)$$

$$+ c(s\langle\psi_i|F_{ii}|\psi_i\rangle + c\langle\psi_j|F_{ii}|\psi_i\rangle$$

$$- c\langle\psi_i|F_{ji}|\psi_i\rangle + s\langle\psi_j|F_{ji}|\psi_i\rangle)$$

$$- s(s\langle\psi_i|F_{ii}|\psi_j\rangle + c\langle\psi_j|F_{ii}|\psi_j\rangle$$

$$- c\langle\psi_i|F_{ji}|\psi_j\rangle + s\langle\psi_j|F_{ji}|\psi_j\rangle)$$

$$+ s(s\langle\psi_i|F_{ij}|\psi_i\rangle + c\langle\psi_j|F_{ij}|\psi_i\rangle$$

$$- c\langle\psi_i|F_{jj}|\psi_i\rangle + s\langle\psi_j|F_{jj}|\psi_i\rangle)$$

$$+ c(s\langle\psi_i|F_{ij}|\psi_j\rangle + c\langle\psi_j|F_{ij}|\psi_j\rangle$$

$$- c\langle\psi_i|F_{jj}|\psi_j\rangle + s\langle\psi_j|F_{jj}|\psi_j\rangle) \tag{52}$$

where $s = \sin\theta$ and $c = \cos\theta$. If the additional approximation is made that

$$\cos\phi = \sqrt{1-\sin^2\phi} \cong 1 - 1/2 \sin^2\phi \tag{53}$$

equation (48) can be expanded in powers of $\sin\theta$ alone. If one notices that since only real orbitals are used $F_{ij} = F_{ji}$ and

expands (52) through second order in $\sin\theta$ the resultant quadratic equation in s is

$$As^2 + Bs + C = 0 \tag{54}$$

with

$$A = 1/2\ (X_{ji} - X_{ij}) + 3\ Z_{ij} \tag{55}$$

$$B = X_{ii} + X_{jj} + 2\ Y_{ij} \tag{56}$$

$$C = X_{ij} - X_{ji} \quad . \tag{57}$$

The expressions for Y and Z are

$$Y_{ij} = 2\ Q_{ij}\ \langle i|\hat{h}|j\rangle - Q_{ii}\langle j|\hat{h}|j\rangle - Q_{jj}\langle i|\hat{h}|i\rangle$$

$$+ \sum_{k\ell} (4\ G_{ijk\ell}[ij;k\ell] - 2\ G_{iik\ell}[jj;k\ell] - 2\ G_{jjk\ell}[ii;k\ell]) \tag{58}$$

$$Z_{ij} = Q_{ij}(\langle i|\hat{h}|i\rangle - \langle j|\hat{h}|j\rangle) + (Q_{jj} - Q_{ii})\langle i|\hat{h}|j\rangle$$

$$+ \sum_{k\ell} [2G_{ijk\ell}([ii;k\ell - [jj;k\ell]) + 2[ij;k\ell](G_{jj}k\ell - G_{ii}k\ell)] \quad . \tag{59}$$

Equation (54) can now be solved for $\sin\theta$. If the values of s for a given A, B and C turn out to be complex, however, the second order terms are dropped and

$$s = -C/B \quad . \tag{60}$$

Once the set of s is found the unitary matrix U is formed via

$$\underset{\sim}{U} = e^{-\underset{\approx}{S}} \tag{61}$$

where $\underset{\sim}{S}$ is the antisymmetric matrix whose i,j element is $\sin\theta_{ij}$ for j > i. This last approximation is much less severe than the assumption that the set of pair rotation matrices commute.

Next let us compare the convergence of this method with method 1 described in the preceeding section. In Table II ten iterations of the MCSCF calculation involving all single and double excitations from the SCF reference on double-zeta H_2O are displayed. This calculation is identical to the water calculation presented in Table I with the only difference being their relative rates of convergence. Comparing the first few iterations in Table I and the upper set of iterations from Table II it is seen that method 1 converges at a faster rate. After iteration 4, however, the calculation begins to diverge. Though method II appears to converge more slowly in the initial iterations, it continues to converge in all subsequent iterations. This tended to be our general impression of method 2, namely that it is much more stable than method 1.

As in the case of method 1 an overrotation factor was also included. In the current procedure this factor is post-multiplied by all rotations in the antisymmetric matrix $\underset{\sim}{S}$. It has been our experience that if all orbitals are active in the MCSCF than a factor of around 3.0 leads to enhanced convergence. Comparing the two sets of iterations in Table II one notices that the lower table (overr. = 3.0) converges much more rapidly. There is some danger, however, in

selecting an overrotation factor that is too large. If this happens the calculations will tend to overshoot and the iterations will oscillate or diverge. Usually we increase this factor slowly from one iteration to the next to avoid this problem.

In Table III a more typical MCSCF calculation is presented. In this table all 36 iterations in a two configuration MCSCF calculation done on the ground state of ozone are listed. More details about this computation can be found in section IVB. This is a more representative MCSCF calculation because not all of the orbitals are active; in fact, just two are active while 11 are of the core and 17 are of the virtual varieties. Furthermore, only two configurations are used, putting this within the calculation range of all MCSCF practioners. Also due to these facts the computation should be inherently more difficult to converge. In practice, however, we found that this particular MCSCF converged rather well considering that only a first-order method was employed. In examining the table one notices that the calculation converges roughly linearly with 4-6 iterations required per order of magnitude obtained in the sum of the squares of the difference of the Lagrangian matrix. In 36 iterations the energy was good to at least eight digits with the principle coefficient converged to around the 5th decimal place. From reviewing Tables I-III it should become obvious why method 2 is normally favored.

D. An Overview of the Berkeley MCSCF Program System.

In the previous sections the theoretical details of our MCSCF method have been presented. Starting with this section the

Table II. Convergence behavior of double-zeta water using method 2 employing 361 configurations.

Method 2

iter	sq. dif.	overr	E	C_1
1	1.1×10^{-4}	1.0	-76.15001468	.978735
2	7.4×10^{-5}	1.0	-76.15008792	.978594
3	5.3×10^{-5}	1.0	-76.15013139	.978469
4	4.1×10^{-5}	1.0	-76.15015918	.978362
5	3.3×10^{-5}	1.0	-76.15017837	.978273
6	2.8×10^{-5}	1.0	-76.15019256	.978198
7	2.5×10^{-5}	1.0	-76.15020366	.978133
8	2.2×10^{-5}	1.0	-76.15021267	.978076
9	2.0×10^{-5}	1.0	-76.15022018	.978025
10	1.7×10^{-5}	1.0	-76.15022654	.977979

Method 2

iter.	sq. dif.	overr	E	C_1
1	1.1×10^{-4}	3.0	-76.15001468	.978735
2	3.4×10^{-5}	3.0	-76.15017392	.978221
3	2.3×10^{-5}	3.0	-76.15020942	.978052
4	1.6×10^{-5}	3.0	-76.15023072	.977931
5	1.2×10^{-5}	3.0	-76.15024408	.977824
6	8.4×10^{-6}	3.0	-76.15025256	.977731
7	6.1×10^{-6}	3.0	-76.15025802	.977650
8	4.5×10^{-6}	3.0	-76.15026156	.977582
9	3.3×10^{-6}	3.0	-76.15026389	.977525

Table II continued.

iter.	sq. dif.	overr	E	C_1
10	2.5×10^{-6}	3.0	-76.15026542	.977477
⋮	⋮	⋮	⋮	⋮
∞	0.0		-76.15026865	.977256

Table III. Convergence behavior of a 2 configuration MCSCF calculation on double-zeta ozone.

iter	sq. dif.	overr.	C_1	C_2	E
1	1.83×10^{-2}	1.0	.938595	-.345021	-224.26895991
2	8.88×10^{-3}	1.0	.917032	-.398813	-224.29521303
3	4.67×10^{-3}	1.0	.908605	-.417657	-224.30227497
4	2.48×10^{-3}	1.0	.900898	-.434030	-224.30569014
5	1.65×10^{-3}	1.0	.895989	-.444077	-224.30763307
6	8.99×10^{-4}	1.0	.891033	-.453938	-224.30877964
7	6.12×10^{-4}	1.0	.888099	-.459652	-224.30945695
8	3.34×10^{-4}	1.0	.885058	-.465481	-224.30985798
9	2.30×10^{-4}	1.0	.883359	-.468697	-224.31009471
10	1.26×10^{-4}	1.0	.881512	-.472162	-224.31023464
11	8.82×10^{-5}	1.0	.880543	-.473967	-224.31031727
12	4.84×10^{-5}	1.0	.879422	-.476043	-224.31036604
13	3.45×10^{-5}	1.0	.878876	-.477051	-224.31039488
14	1.91×10^{-5}	1.0	.878195	-.478303	-224.31041194
15	1.38×10^{-5}	1.0	.877890	-.478862	-224.31042204
16	7.73×10^{-6}	1.0	.877475	-.479622	-224.31042803
17	5.68×10^{-6}	1.0	.877308	-.479928	-224.31043104
18	3.23×10^{-6}	1.0	.877054	-.480392	-224.31043370
19	2.93×10^{-6}	1.0	.876964	-.480556	-224.31043496
20	1.39×10^{-6}	1.0	.876808	-.480841	-224.31043571
21	1.03×10^{-6}	1.0	.876761	-.480926	-224.31043616
22	6.11×10^{-7}	1.0	.876664	-.481103	-224.31043644
23	4.56×10^{-7}	1.0	.876641	-.481145	-224.31043660

Table III continued.

iter.	sq. dif.	overr.	C_1	C_2	E
24	2.75×10^{-7}	1.0	.876581	-.481255	-224.31043670
25	2.05×10^{-7}	1.0	.876570	-.481275	-224.31043676
26	1.26×10^{-7}	1.0	.876532	-.481344	-224.31043679
27	9.37×10^{-8}	1.0	.876528	-.481352	-224.31043682
28	5.86×10^{-8}	1.0	.876504	-.481395	-224.31043683
29	4.33×10^{-8}	1.0	.876502	-.481397	-224.31043684
30	2.75×10^{-8}	1.0	.876487	-.481425	-224.31043684
31	2.03×10^{-8}	1.0	.876488	-.481424	-224.31043685
32	1.31×10^{-8}	1.0	.876478	-.481442	-224.31043685
33	9.5×10^{-9}	1.0	.876479	-.481441	-224.31043685
34	6.2×10^{-9}	1.0	.876472	-.481452	-224.31043685
35	4.5×10^{-9}	1.0	.876473	-.481450	-224.31043685
36	3.0×10^{-9}	1.0	.876469	-.481458	-224.31043685

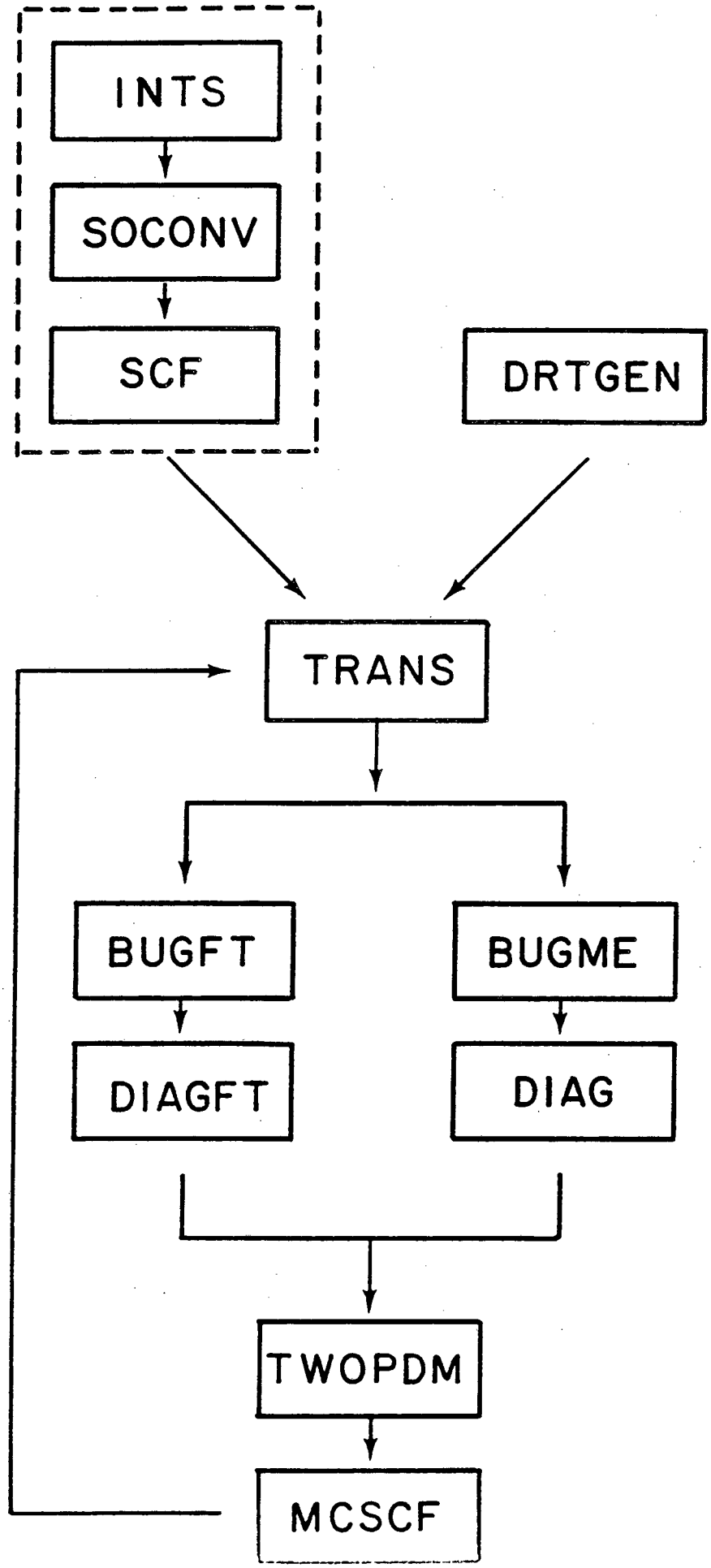
computational aspects will be considered. Before going on and describing the individual programs in detail an overview of the entire Berkeley MCSCF program system will be described. A flowchart of the system is laid out in Figure III.

At the beginning of the MCSCF calculation the symmetry orbital (SO) integrals are computed and stored. This is done using a modified version of a program (written by Pitzer) entitled INTS.[61] This program calculates its integrals using the method of Rys polynomials.[62] The SO integrals are a simple, fixed linear combination of atomic orbitals and have the correct molecular symmetry. These integrals are sorted in SOCONV and only those integrals needed in the SCF are stored. An SCF calculation is performed to get a set of starting orbitals. If many points are to be calculated on an energy hypersurface, the orbitals obtained at an earlier point can be used instead and the SCF and sorting steps can be omitted.

After the SCF, the DRT is calculated using DRTGEN. The DRT needs to be computed only once for a given potential energy surface. However since it is very fast to construct and the output file needs to be saved, it is usually calculated at each energy point. The limitations on the type of configuration sets that can be handled in the CI are presented in section II. After the DRT is obtained the SO to MO integral transformation is carried out using TRANS. The orbitals are sorted from the SCF ordering to a more convenient ordering specified by the program DRT. More details about our transformation program will be presented in section IIG.

Next the CI is solved for the desired eigenvalue(s) and eigenvector(s). One will notice from the figure that there are two paths following TRANS. Either the programs BUGME and DIAG or the programs BUGFT and DIAGFT are executed. The first two programs, BUGME and DIAG, are the diagonalization tape CI programs presented by Brooks in his Ph.D. thesis.[51] In BUGME the Hamiltonian matrix is computed and stored on an output tape in the form of loops. DIAG reads this tape every iteration of the Davidson's diagonalization and computes the CI energy along with the corresponding eigenvector. The only problem with this is that the BUGME output tape grows very rapidly with the size of the configuration set. On the Harris 800 minicomputer we have a limit of roughly 12,000 configurations. This leads to an output tape of around 70 M bytes. Clearly for cases of more than 10,000 configurations it would be advantageous to eliminate this tape if possible. This was the impetus for developing two new CI programs BUGFT and DIAGFT. In the case of BUGFT this long tape has been shortened by a factor of approximately 100. This is accomplished by splitting up the work differently between the two programs. BUGFT now outputs only partial loops within the internal space (see section IIB). These are read in by DIAGFT and the lower halves of the loops are computed each iteration by specialized subroutines taking advantage of the particular characteristics of the external space as first suggested by Siegbahn.[58-60] Part of these special features were exploited in BUGME-DIAG, however, full advantage is now obtained in BUGFT-DIAGFT. The additional external space subroutines were written by Douglas Fox. Computationally BUGFT-DIAGFT is nearly identical in CPU time to BUGME-DIAG. Since

Figure III. The Berkeley LDGUGA MCSCF Programs



partial loops are constructed during each iteration in the former case but not the latter it is obvious that performing the multiplications necessary in the Davidson diagonalization is the major time step.

After the CI is finished the eigenvector is read in by TWOPDM, the two-particle density matrix construction program. TWOPDM constructs both the one and two-particle density matrices and writes them to an output tape. More on TWOPDM is presented in the next section.

The program MCSCF next reads in the TWOPDM output tape and calculates the "improved" unitary transformation matrix. MCSCF first reads and sorts both the MO integrals and the density matrix elements into a suitable ordering and outputs each list to a scratch tape. Now the scratch tapes are sequentially read and the Fock matrix (Lagrangian) is constructed. Using either method 1 or 2 a new unitary transformation matrix is constructed. This is multiplied onto the old matrix and written on an output file, completing one iteration. For subsequent iterations the matrix U is read by TRANS along with the SCF vector. The two are multiplied to yield a new SO to MO transformation matrix and the integrals are retransformed. The iterations are continued until both the change in the energy and in the sum of the squares of the differences of the Lagrangian elements are below some threshold.

E. Generation of the Two-Particle Density Matrix.

In any MCSCF method that handles large configuration sets the ability to generate the two-particle density matrix fast and efficiently is crucial for the orbital optimization stage to be competitive with the CI portion in a two-step MCSCF procedure. For methods designed

to only handle a few select configurations this is not as important a consideration. In a small MCSCF the vast majority of computer time is spent in the orbital transformation step. Constructing the Hessian matrix is also time consuming if a second-order MCSCF procedure is used, but all other steps are negligible if only a few configurations are employed. We designed the LDGUGA MCSCF program to be able to handle very large numbers of configurations and in cases like this the relative timings of the individual procedures can change drastically. In our largest MCSCF calculations of 17,678 configurations the CI Davidson's diagonalization took the majority of the CPU time followed by the generation of the two-particle density matrix and the diagonalization tape construction steps. In our more typical MCSCF calculations I estimate the time required to generate the two-particle density matrix at between 10-30% of the total.

The traditional way to generate the two-particle density matrix is to read in the CI coupling coefficients, $b_{IJ}^{ijk\ell}$, from an external tape and combine these elements with the corresponding configuration coefficients, C_I, to form $G_{ijk\ell}$. The general formula for $G_{ijk\ell}$ is,

$$G_{ijk\ell} = \sum_{IJ} C_I C_J b_{IJ}^{ijk\ell} \quad . \tag{62}$$

In a conventional CI the coupling coefficients are stored on what is called the formula tape along with identifiers for its six indices. Since there are on the order of N^2m^4 coupling coefficients where N is the number of configurations and m is the number of orbitals this tape can get very large and is usually stored on an external device like a disc drive. If the number of density matrix elements is too

large to hold in central memory as is often the case either the tape must be read several times or the tape must be sorted so that all coupling coefficients that contribute to a given subset of $\underset{\sim}{G}$ are sequential on the formula tape. Either of these solutions was deemed too expensive in our case. Furthermore, the fact that using a formula tape would severely limit the size CIs that could be handled.

The approach we originally decided to implement did not use a formula tape. It was actually a modified version of the diagonalization tape construction program, BUGME. In BUGME a block of integrals are read into core. This integral block consists of at least all integrals $k\ell$ for a fixed i,j or a minimum of n^2 where n is the number of orbitals. Usually we try to hold at least n^3 integrals or all $jk\ell$ for a fixed index i, since the program is more efficient if this is possible. Then all loops involving the integrals in core are generated. For each loop a value is calculated which is a sum of products of coupling coefficients and integrals, and the loop information is written onto the diagonalization tape. These loops are read in and broken down using the loop-breakdown algorithm[51] (see section IIA) and in the next program DIAG the contributions to the individual Hamiltonian matrix elements are generated.

In TWOPDM, our two-particle density matrix generation program, instead of holding a block of integrals in core the same space is assigned for the storage of the density matrix elements. In other words, the density matrix is stored in the same structure as that of the integrals.[51] In this scheme unique density matrix elements with the same indices are stored sequentially (an example of this is that

the elements G_{4231}, G_{4321} and G_{4132} are grouped together) and the one-particle elements are stored together with the two-particle elements. Initially the arrays holding the density matrix are zeroed. Then using the loop-driven algorithm all loops contributing to the density matrix elements in core are generated. For each loop one or two coupling coefficients are computed. A given loop is next broken down using the loop-breakdown algorithm and the coupling coefficients are multiplied by the appropriate CI coefficients. These contributions are then summed into their respective density matrix elements. The address for a particular element is computed using the integral offset arrays defined by Brooks.[51] These arrays take advantage of full Abelian point group symmetry and no element which is zero by symmetry is stored. After all loops contributing to the block of density matrix elements currently in core are processed the finished block is written to an external file and the in-core block is zeroed. This entire process is repeated until all blocks have been constructed. Below is an outline of this procedure.

1. Read in DRT arrays.
2. Read in CI vector and store in array C.
3. Compute block structure for $\underset{\sim}{G}$, the two-particle density matrix (D.M.)
   A. Loop over blocks of $\underset{\sim}{G}$ and zero out current block.
      I. Loop over groups of D.M. elements where a group contains all j, k and ℓ elements from a fixed i.
         a. Calculate loops involving D.M. elements from group i and break down loops to obtain loop value d.

1. loop from 1 to the number of lower walks (NLWK). m = 1, NLWK
2. loop from 1 to the number of upper walks (NUWK). n = 1, NUWK
3. calculate d

$$d = \sum_{m}^{NLWK} \sum_{n}^{NUWK} C(INDX(IUWK+m)+n) \times C(INDX(IUWK+m)+n)$$

where C is the CI coefficient array, INDX is the lexical to reverse lexical ordering array and IUWK and JUWK are the primary upper walks[51] for the right and left sides of the loop.

b. d is multiplied by the appropriate coupling coefficients and summed into the correct $\underset{\sim}{G}$ element.

II. Write out finished block of D.M. elements to external tape and zero in core elements.

4. Write an end-of file marker to the D.M. output tape.

F. <u>Construction of the Lagrangian Matrix</u>.

Since the Lagrangian matrix is used both in determining the new unitary transformation matrix and as an indicator of convergence (it is symmetric upon convergence) it is necessary to generate it every iteration. The formula for this matrix is

$$X_{ir} = 4 \sum_{jk\ell}^{m} G_{ijk\ell}[rj;k\ell] + 2 \sum_{j}^{m} Q_{ij} \langle r|\hat{h}|j\rangle \tag{63}$$

where m is the number of non-virtual orbitals. These sums do not involve virtual type indices since density matrix elements involving these orbitals are zero. Therefore,

$$X_{ij} = 0 \quad , \quad j \in \text{virtual space} \quad . \tag{64}$$

One also notices in equation (63) that the two-electron integrals involve at most one index in the virtual space. This means that integrals with two through four virtual indices are not needed in the MCSCF calculation and, therefore, need not be computed in the integral transformation step. This can result in a considerable savings and will be discussed in the next section.

Constructing X then is a nm^4 process where n is the total number of orbitals. In the extreme case of all orbitals being active, m=n, the calculation becomes n^5 in complexity. This is equivalent in effort to roughly 25% of the integral transformation time. So if our program is to be efficient in the general case then an efficient algorithm must be used to find X. The algorithm we adopted begins by sorting both the MO integrals and the density matrix elements and writing them to external storage. The integrals and density matrix elements are sorted into triangles, one for each i and j index, and each triangle holds all k and ℓ indices. Elements that are zero by symmetry are stored explicitly. After the sort is complete the $k\ell$ triangle of each is brought in and the following contributions to X are found

$$X_{ir}(k\ell) = \sum_j G_{ijk\ell}[rj;k\ell] \quad , \tag{65}$$

After the sort it is only necessary to read through the external storage tapes once to construct X.

We also take advantage of many of the symmetry properties of $\underset{\sim}{X}$, $\underset{\sim}{G}$, and the integrals. Since our CI is restricted to non-degenerate point groups only symmetry of this sort will be considered. In such cases if i and r belong to different symmetry species $X_{ir} = 0$. This is tested for by the program and if i and r are of different symmetries the contribution to X_{ir} is not computed. Also since both the MO integrals and the density matrices must be totally symmetric then the symmetry of i or r must equal the cross-product symmetry of j, k and ℓ. Elements that do not satisfy this condition are equal to zero on the storage tapes, so in the program zero elements are checked for. These two procedures result in a considerable time savings in cases where molecular symmetry is present and do not appreciably slow down the program when none is present.

In determining the matrix $\underset{\sim}{U}$ using method 1 or 2 only the differences $X_{ij}-X_{ji}$ are needed never either X_{ij} or X_{ji} alone. Therefore, if one knew in advance that a particular difference $X_{ij}-X_{ji}$ were zero it would not be necessary to compute either X_{ij} or X_{ji}. This turns out to be the case whenever the energy is invariant to rotations between orbitals i and j. This happens when any two orbitals are indistinguishable in the calculation such as when both are doubly occupied or empty in all configurations within the MCSCF. Therefore, core-core and virtual-virtual rotations are invariant and the corresponding off-diagonal Lagrangian elements $X_{core,core}$ and $X_{virtual,virtual}$ need not be computed. Also for specialized

configuration sets many of the active-active rotations are energy invariant. In the case of the complete active space (CAS)[35-37] no active-active $\underset{\sim}{X}$ elements need be computed. For all single and double excitations from a closed-shell reference all active-active rotations between any pair of orbitals that are either both doubly occupied or both unoccupied in the reference are invariant. In the future minor modifications will be made to the optimization program so that these Lagrangian elements will not be computed.

G. Avoidance of Computing the Complete Set of MO Integrals.

In the preceeding section it was seen that integrals involving 2, 3 or 4 indices in the virtual space are not needed in a first-order MCSCF method. If the total number of orbitals is n then the number of integrals in the complete set are proportional to n^4. If m is the number of non-virtual orbitals the number of necessary integrals is nm^3. In many MCSCF calculations m is less than half the total n. If $m = \frac{1}{2}n$ then one finds that only 1/2 of the total integrals need be calculated. This results in a substantial savings in space.

Computing only the necessary integrals also saves a considerable amount of time. The transformation from the A.O. to the MO integrals is usually broken up into 4 steps as follows

$$[i\nu,\mu\sigma] = \sum_{\lambda}^{n} C_{\lambda i}[\lambda\nu;\mu\sigma] \tag{66}$$

$$[ij;\mu\sigma] = \sum_{\nu}^{n} C_{\nu j}[i\nu;\mu\sigma] \tag{67}$$

$$[ij;k\sigma] = \sum_{\mu}^{n} C_{\mu k}[ij;\mu\sigma] \tag{68}$$

$$[ij;k\ell] = \sum_{\sigma} C_{\sigma\ell}[ij;k\sigma] \quad . \tag{69}$$

In the full integral transformation each step is proportional to n^5 in effort. If, however, i, j and k are restricted to the set of non-virtual indices then the amount of work necessary for each of the four steps becomes proportional to mn^4, m^2n^3, m^3n^2, and m^3n^2 respectively. If m is much smaller than n then the time to do the first quarter transformation dominates the calculation.

In second-order MCSCF methods[35-37] and in CI gradient programs[30,51] additional integrals are needed. In the former case 2 virtual index integrals are used and in the latter case 3 virtual index integrals are required. These integrals can still be computed in a time proportional to mn^4 so this is no major problem. In the near future we hope to have an integral transformation program which generates all integrals except the 4 virtual index case. This, however, requires major rewriting of the program. Recently a version of our standard transformation program has been adapted so that all 1 and a subset of the 2 and 3 index integrals are computed. This program is proportional to mn^4 in CPU time and but can't be used for second-order MCSCF or CI gradient calculations since not all integrals have been generated for these cases.

H. Construction and Use of the Hessian Matrix.

Up to this point all of the orbital optimization schemes presented have been of the first-order type. That is they involved some expansion of the unitary transformation matrix that was truncated to first order. For systems in which the CI coefficients are loosely coupled to the

orbital expansion coefficients one finds roughly linear convergence in the MCSCF. One could expand the unitary matrix U to second order in the antisymmetric matrix T as follows

$$\underset{\sim}{U} = \underset{\sim}{1} + \underset{\sim}{T} + \underset{\sim}{T}^2/2 \tag{70}$$

and solve the resulting equations for T. However, for most MCSCF calculations this does little to improve convergence. This is because usually the coupling of the CI and orbital coefficients is of the same order or more important than the second order terms in the orbital expansion equations. The second order changes in the CI coefficients and the second order coupling terms can also be computed to yield what is called a quadratic MCSCF. However, these additional terms are extremely expensive both in storage space and in time necessary to compute. There are, though, certain classes of MCSCF wavefunctions in which the CI and orbital coefficients are only slightly coupled.

One class of these are complete active space (CAS) wavefunctions.[35-37] The configuration sets for these wavefunctions include a full CI amongst the active orbitals. For CAS systems then there are no active-active rotations as these are all invariant to the energy. This is quite advantageous since rotations of this type usually couple strongly with the orbital expansion coefficients.

Since CASSCF's are well behaved and also predict reasonable properties for chemical systems we would like to be able to routinely compute them. Though they can currently be calculated using the methods 1 and 2 described earlier they tend to show nearly quadratic convergence if the second-order equations for T are solved to yield U.

Expanding equation (16) using (70) and discarding terms higher than order 2 gives

$$
\begin{aligned}
E = & \sum_{ij} Q_{ij}h_{ij} + \sum_{ijk\ell} G_{ijk\ell}[ij;k\ell] \\
& + 2\sum_{ir} T_{ir}\,[\sum_{j} Q_{ij}h_{rj} + 2\sum_{jk\ell} G_{ijk\ell}[rj;k\ell]] \\
& + 2\sum_{ir}[\sum_{s} T_{is}T_{sr}\,(\sum_{j} Q_{ij}h_{rj} + 2\sum_{jk\ell} G_{ijk\ell}[rj;k\ell])] \qquad (71) \\
& + 2\sum_{ij}[\sum_{mn} T_{im}T_{jn}\,(\sum_{k\ell} G_{ijk\ell}[mn;k\ell] + 2\sum_{k\ell} G_{ikj\ell}[mk;n\ell])] \\
= & \; E_0 + 2\sum_{ij}[T_{ij}\varepsilon_{ij} + (\sum_{s} T_{is}T_{sj})\varepsilon_{ij}] \\
& + 2\sum_{ij}[\sum_{mn} T_{im}T_{jn}\,(\sum_{k\ell} G_{ijk\ell}[mn;k\ell] + 2\sum_{k\ell} G_{ikj\ell}[mk;n\ell])] \quad . \qquad (72)
\end{aligned}
$$

The second equality makes use of the definition of the Lagrange multipliers defined in equation (43). Differentiating (72) with respec t to T_{ab} and remembering that $T_{ij}=-T_{ji}$ one finds that

$$
\begin{aligned}
\frac{\partial E}{\partial T_{ab}} = & \; 2(\varepsilon_{ab}-\varepsilon_{ba}) + \sum_{i>a} T_{ia}(\varepsilon_{ib}+\varepsilon_{bi}) - \sum_{i>b} T_{ib}(\varepsilon_{ia}+\varepsilon_{ai}) \\
& - \sum_{a>j} T_{aj}(\varepsilon_{bj}+\varepsilon_{jb}) + \sum_{b>j} T_{bj}(\varepsilon_{aj}+\varepsilon_{ja}) \\
& + 2\sum_{i>j} T_{ij}(Y_{abij}-Y_{baij}-Y_{abji}+Y_{baji}) \qquad (73)
\end{aligned}
$$

with

$$Y_{abij} = Q_{bj}h_{ai} + 2\sum_{k\ell} (G_{bjk\ell}[ai;k\ell] + 2.G_{bkj\ell}\ [ak;i\ell]) \quad . \tag{74}$$

If the energy as a function of $\underset{\sim}{T}$ is expanded in a Taylor's series about the unperturbed energy to second order the result is

$$E(\underset{\sim}{T}) = E_0 + \sum_{ab} T_{ab}(\frac{\partial E}{\partial T_{ab}})_0 + \frac{1}{2}\sum_{\substack{ab\\cd}} T_{ab}T_{cd}(\frac{\partial^2 E}{\partial T_{ab}\partial T_{cd}})_0 \tag{75}$$

or

$$E = E_0 + \underset{\sim}{g}\cdot\underset{\sim}{T} + \frac{1}{2}\underset{\sim}{T}^{\dagger}\underset{\sim}{\underset{\sim}{H}}\underset{\sim}{T} \tag{76}$$

where $\underset{\sim}{g}$ is the gradient and $\underset{\sim}{H}$ is the Hessian matrix. If the derivative of (76) is taken

$$\partial E/\partial\underset{\sim}{T} = 0 = \underset{\sim}{g} + \underset{\sim}{\underset{\sim}{H}}\underset{\sim}{T} \quad . \tag{77}$$

Comparing (73) with the equation (77) above one notices that

$$g_{ij} = 2(\varepsilon_{ij}-\varepsilon_{ji}) \tag{78}$$

$$\begin{aligned} H_{ijk\ell} = \frac{\partial^2 E}{\partial T_{ij}\partial T_{k\ell}} &= 2(Y_{ijk\ell}-Y_{jik\ell}-Y_{ijk\ell}+Y_{ji\ell k}) \\ &+ \delta_{i\ell}(\varepsilon_{jk}+\varepsilon_{kj}) - \delta_{j\ell}(\varepsilon_{ik}+\varepsilon_{ki}) \\ &- \delta_{ik}(\varepsilon_{j\ell}+\varepsilon_{\ell j}) + \delta_{jk}(\varepsilon_{i\ell}+\varepsilon_{\ell i}) \quad . \end{aligned} \tag{79}$$

Finally the system of equations

$$\underset{\sim}{\underset{\sim}{H}}\underset{\sim}{T} = -\underset{\sim}{g} \tag{80}$$

is solved for $\underset{\sim}{T}$ and the unitary matrix $\underset{\sim}{U}$ is formed via $\underset{\sim}{U} = e^{\underset{\sim}{T}}$.

In order to solve the system of equations (76) the vector $\underset{\sim}{g}$ and matrix $\underset{\sim}{H}$ must be constructed. The gradient $\underset{\sim}{g}$ is easily formed once the Lagrange multipliers have been computed. The real problem then is making the Hessian $\underset{\sim}{H}$. The ordering of the integral and density matrix elements used to construct the Lagrangian matrix $\underset{\sim}{X}$ (see section IIF) is the same as needed for building $\underset{\sim}{H}$ so an additional sort is not necessary. Currently our existing MCSCF program is nearly completed and we hope to have it operational by October 1981. This Hessian matrix construction code is to be integral driven, meaning that a block containing all integrals with fixed indices i and j and variable indices k and ℓ are brought into core and every H contribution using these integrals is computed. Once an integral block is in core the density matrix $\underset{\sim}{G}$ is read sequentially block by block until all elements of $\underset{\sim}{G}$ have been processed. Then the next integral block is read in and so forth. If the integral list is to be read only once enough core space must be allocated so that the Hessian matrix will fit entirely in core. This is not as severe a requirement as it would seem since only elements of H between pairs of non-redundant rotations need to be computed.[37] In other words, this means that only terms involving two rotational pairs both having non-zero gradients need to be calculated. In a typical MCSCF this results in a substantial reduction of the number of Hessians element to be formed and also reduces the amount of core storage necessary. In the special case of a CAS configuration set the necessary Hessian elements are very few in number since active-active, core-core, and

virtual-virtual terms in the gradient are zero.

Once $\underset{\sim}{H}$ is found the system (80) must be solved. This can be done directly by finding $\underset{\sim}{H}^{-1}$ and solving for $\underset{\sim}{T} = -\underset{\sim}{H}^{-1}\,\underset{\sim}{g}$. However, to find $\underset{\sim}{H}^{-1}$ requires all elements of the Hessian matrix, not just the non-redundant terms. Also matrix inversion is an order N^3 process where N is the number of orbitals pairs. We instead use the iterative procedure of Pople et al.[64] Their scheme requires only the non-redundant elements of $\underset{\sim}{H}$ and is order M^2 where M is the number of non-redundant orbital pairs. Another advantage of the iterative procedure is that the elements of H can be processed in any order and the entire Hessian matrix need not be in core at any given time. In our CI gradient system of programs we have previously employed this procedure to solve quite large systems of equations.[30,51,65] We have found the procedure to be very successful and to have adequate convergence behavior.

I. A Proposed Extrapolation Procedure to Improve Convergence.

Before closing section II I want to discuss an extrapolation procedure that is currently being developed for the LDGUGA system. The purpose of this is to speed up the rate of convergence by using information from the previous iteration in the formation of the new unitary transformation matrix $\underset{\sim}{U}$. Up to this point the only variable parameter in our MCSCF method has been the overrotation factor α. This arbitrary constant was chosen by the operator to increase convergence with the resultant matrix $\underset{\sim}{U}$ being defined as

$$\underset{\sim}{U} = e^{\alpha \underset{\sim}{T}} \tag{81}$$

in method 2 where $\underset{\sim}{T}$ is the usual antisymmetrix matrix of orbital rotations. What we would like to do is find a different factor for each i,j pair of orbitals such that i>j. If the antisymmetric matrix of these parameters is defined as $\underset{\sim}{r}$ then $\underset{\sim}{U}$ can be written as

$$\underset{\sim}{U} = e^{\underset{\sim}{T}'} \text{ where } T'_{ij} = r_{ij}T_{ij} \quad . \tag{82}$$

What is now needed is a procedure to compute $\underset{\sim}{r}$.

The method we decided to implement is the following. After the kth MCSCF iteration the matrices $\underset{\sim}{U}(k)$, $\underset{\sim}{r}(k)$ and $\underset{\sim}{g}(k)$ are written to an external tape (g(k)is the current gradient matrix whose ij element is defined as $X_{ij}(k)-X_{ji}(k)$). At the start of each orbital optimization step the new gradient, $\underset{\sim}{g}(k+1)$, is generated. The storage tape from the previous iteration is read in and g(k+1) and g(k) are compared. How the ij elements of the two matrices differ determines the new multiplicative factor $c_{ij}(k+1)$ such that

$$r_{ij}(k+1) = c_{ij}(k+1)r_{ij}(k) \quad . \tag{83}$$

This is done for all ij pairs (i j), r(k+1) is computed and the matrix U(k+1) is constructed. This procedure is then repeated every iteration until the MCSCF has converged. The problem now reduces to finding $\underset{\sim}{r}(1)$ and $\underset{\sim}{c}(k)$. Initially we set $r_{ij}(1) = \alpha$ for all ij pair (i>j). Usually α is set to 1.0 unless prior information is known indicating that another value should be used. At the end of the second iteration the parameters $\underset{\sim}{c}$ can be calculated. A relatively simple formula for computing $\underset{\sim}{c}$ is as follows

$$c_{ij}(k+1) = \frac{g_{ij}(k)}{g_{ij}(k)-g_{ij}(k+1)} \quad . \tag{84}$$

We also have adopted a few additional restrictions to keep the orbital changes down to a reasonable size each iteration. First, no element of r can ever be greater than 10.0 or less than 0.10 and second

$$0.333 < \left|\frac{r_{ij}(k+1)}{r_{ij}(k)}\right| < 3.0 \quad . \tag{85}$$

If any element exceeds these bounds it is set equal to the maximum or minimum value allowed.

While this procedure is not very complex it is, nonetheless, expected to enhance convergence by a considerable amount. Extrapolation, however, should only be used once one is reasonably sure of being within the vicinity of the desired minimum. That is, there are usually several distinct minima in the MCSCF equations and if one is not careful a method like this will tend to force convergence to the wrong solution.

III. An MCSCF Study of the Effect of Higher Excitations in the Water Molecule

The second application of the LDGUGA MCSCF method was in performing a series of computations on the water molecule.[33] These results were part of a test study of the effects of inclusion of higher than double excitations in CI and MCSCF calculations. We chose to do this study for a number of reasons, the first having to do with the relative importance of triple and quadruple excitations.

Traditionally, triple excitations were thought to contribute little to the total correlation energy. In a CI using canonical SCF orbitals double excitations alone account for at least 90% of the correlation energy in closed-shell ground state systems of 10 electrons or less[66-68] and it has been numerically shown that single excitations contribute very little to the correlation energy (at most a few percent).[69] This implies that triple, quadruple and higher excitations are responsible for less than 10% of the total correlation energy and it is widely believed that the quadruple excitations account for most of this remaining fraction.[70-73] From many-body perturbation theory arguments it is generally assumed "that the dominant contributions of quadruple and higher even excitations in an extended CI expansion are to cancel the unlinked diagram errors of the double-excitations CI".[73]

Unlinked diagram contributions are terms that are summed in at a given level of perturbation theory, but are cancelled in higher orders. According to the Linked Cluster Theorem,[70,74] if the perturbation expansion is summed to infinite order only "linked" terms are

included, the unlinked terms summing to zero. Since the relative size of the unlinked cluster terms increases as the square of the number of electrons N^2 while the linked terms are linear in N it is apparent for large N that the unlinked cluster terms will dominate in a truncated expansion if they are included.[73,75-83]

Recent calculations by Wilson, et al.[84-87] however, have challenged these long held views. Their work suggests that the fourth order linked triples contribution to the correlation energy is much larger than had been previously believed. Originally they reported that this contribution for the neon atom employing a sizeable basis set of Slater type orbitals (STOs) was -0.009073 hartrees compared to the fourth order linked quadruple result of 0.000405 hartrees.[84] This would make the linked triples more important by over a factor of 20. This triples term was later found to be in error with a corrected value of -0.001134.[85] Though this contribution is much smaller, it is still much larger than anticipated. Around the same time period Wilson also computed the fourth order linked triples energy for water using a 39 STO basis set.[86-87] Earlier Bartlett et al.[75] calculated the fourth order linked single, double and quadruple diagram contributions for this system and if all four terms are summed together the total fourth order energy involving only linked terms is -0.0110 hartrees. Bartlett also determined that the unlinked terms in the same order add up to -0.0100 hartrees. Since single and double excitation CI's are often corrected for unlinked cluster effects is it not then also necessary to correct for the linked terms as well?

One reason for choosing H_2O as our test molecule is that we needed a system with a reasonably large number of single and double excitations in which the number of triple and quadruple excitations would be small enough to be computationally feasible. As it turns out if a double-zeta basis set is employed both of these conditions are met. In our CI studies of this 14 orbital system 17,678 configurations are generated if all fourfold or less excitations are included.

Another goal of this research is to determine if the relative importance of classes of excitations changes when different sets of molecular orbitals are used. To the best of our knowledge little work has been done in this area previously. We were primarily interested in the effects found in going from the SCF to the MCSCF orbitals within a fixed configuration set. For example, as stated earlier it was expected that for a straight CI employing SCF orbitals triple excitations would have little importance, however, if MCSCF orbitals were used instead would this still be the case?

Lastly the water molecule was selected because it appeared to be well behaved in the MCSCF procedure. For our largest MCSCF calculation of 17,678 configurations it was essential that the computation converge fairly quickly since each iteration was quite expensive. This particular MCSCF wavefunction at the time it was calculated included over 100 times as many configurations as the next largest published calculation (60 configurations).[18] Though unpublished calculations of several hundred configurations had been performed this is still a very large increase in the size of the

MCSCF problem.

Very few variational studies on the effects of triple and quadruple excitations in molecules had been performed prior to this work. Of the existing previous investigations a number were done by Shavitt and co-workers.[88-92] These studies were carried out of a number of small molecules: BH_3, H_2O, $H_2C{=}CH_2$ and BeH_2. The BH_3 investigation was performed using a minimum basis set and found that the triple and quadruple excitations contribute 0.8% and 1.9% of the full CI correlation energy respectively.[88] Next water was examined employing a STO double-zeta basis set and the triples and quadruples were found to recover 1.5% and 3.0% of the CI including all excitations through quadruple correlation energy.[90] In this study, however, only the valence electrons were correlated so these percentages are only approximate. What is really interesting is their results for water are very basis set dependent. When they switched to a contracted Gaussian type orbital (GTO) basis the corresponding values for the triple and quadruple excitations became 0.8% and 4.3%. In a later study, minimum basis ethylene was found to have a triple excitation contribution of 0.5% compared to a 6.7% figure for the quadruples.[91] Lastly the triples and quadruples correlation energy contributions have been estimated for double-zeta plus polarization BeH_2 at 0.4% and 2.4% respectively.[91] These earlier studies then seem to suggest that for small systems the triple excitations are any where from 10 to 50% of the importance of the quadruple excitations.

A. Theoretical.

For our study of the water molecule we adopted a procedure

similar to one used previously by Hosteny, Gilman, Dunning, Pipano and Shavitt (HGDPS).[90] In these calculations a contracted Gaussian double-zeta basis set was employed as in the previous HGDPS study. The precise designation of this basis set is O(9s5p/4s2p), H(4s/2p).[93-94] We felt that this basis set was adequate for our purposes as we wished to observe the trends in a series of CI and MCSCF calculations and not predict quantitatively chemical properties such as the equilibrium geometry. For the hydrogen atoms a scale factor of 1.2 was multiplied to their Gaussian s function coefficients. The exact cartesian geometry used was O(0.0,0.0,-0.009000), H(±1.515263,0.0,-1.058898). This geometry is very close to the single and double excitation CI theoretical equilibrium geometry determined for double-zeta H_2O. In internal coordinates this geometry corresponds to an O-H bond distance of 0.976 Å and an H-O-H bond angle of 110.6°.

The major difference between our study and that of HGDPS is that in the present investigation all ten electrons were correlated whereas in their earlier study only the eight valence electrons were correlated. The freezing of the core electrons, that is restricting the core orbitals to be doubly occupied in all configurations within a CI calculation, generally has little effect on the predicted chemical behavior of the system involved.[95] However, the amount of correlation energy obtained from these core excitations is generally a large fraction of the total. The reason these excitations are usually neglected is that the energy lowering due to the core-core and core-virtual type excitations is roughly constant with respect to conformational changes in the molecule or between low-lying energetic states. Since unlinked cluster contributions to the energy increase as the

number of correlated electrons goes up,[70-72,74] the relative importance of higher excitations should increase if excitations are allowed out of the core orbitals.

Initially a canonical SCF wavefunction was determined for the $^1A_1$ ground state of water. Following this a series of CI computations were performed to investigate the effects of inclusion of various classes of excitations. The SCF and subsequent CI and MCSCF calculations were carried out in C_{2v} symmetry. Four different CI calculations were performed using the SCF orbitals: the first containing all single excitations (CIS), the second all single and double excitations (CISD), the third all single, double and triple excitations (CISDT) and the fourth all single, double, triple and quadruple excitations (CISDTQ). The number of spin and spatially adapted $^1A_1$ configurations of each of these sets is 20, 361, 3203 and 17678 respectively.

In addition to carrying out CIs employing the SCF orbitals for these four configuration sets MCSCF calculations were also performed in each of these spaces. The four computations will be labeled MCS, MCSD, MCSDT and MCSDTQ respectively. Since excitations from all occupied into all virtual type orbitals were included, all orbitals have a partial occupancy in the MCSCF sense. Therefore, instead of the typical four rotation classes: core-active, virtual-active, core-virtual and active-active, only the last class, active-active, involves non-zero rotations. This fact was anticipated to aid in MCSCF convergence since each of these classes has a different convergence behavior. For these water calculations convergence method

2 was used for the majority of iterations with method 1 being employed for the remainder. Additional discussion in convergence can be found in sections IIB and IIC.

B. Water Energetic Results.

Our energetic results for water are summarized in Table IV. Also included in this table is the full CI result of Saxe, Handy and Schaefer (SHS)[96] for H_2O. Their computation was done with the identical basis set and at the same geometry employed here. The full CI calculation was actually carried out about six months after the completion of the present study. Therefore, at the time of our calculation the "true" correlation energy was not known.

For the $^1A_1$ ground state of water an SCF energy of -76.009838 hartrees was found. Comparing this result to the 256,473 configuration complete CI of SHS one notes that the actual correlation energy for a double-zeta basis set at this geometry is -0.148028 hartrees which corresponds to a total energy of -76.157866 hartrees. The largest straight CI we performed was CISDTQ. For this 17,678 configuration CI a total energy of -76.157626 hartrees was obtained. In the SDTQ case 0.147765 hartrees of correlation energy was recovered or 99.8% of the true correlation energy. This result supports the view long held by quantum chemists that excitations higher than the quadruple level are unimportant for the size systems we are dealing with.[66-68] Indeed in the present example higher than quadruple excitations account for less than 0.2% of the correlation energy.

What percentage of the correlation energy is recovered in the standard CISD calculation? It turns out that 94.7% of this energy

Table IV. 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	-76.009 838	0.0	...	1.0
CI All single excitations	20	-76.009 838	0.0	...	1.0
MCSCF All single excitations	20	-76.083 339	0.073 502	...	0.893 26
CI All S + D	361	-76.150 015	0.140 177	0.000 000	0.978 74
MCSCF All S + D	361	-76.150 269	0.140 431	0.000 254	0.977 26
CI All S + D + T	3,203	-76.151 156	0.141 318	0.001 141	0.978 19
MCSCF All S + D + T	3,203	-76.155 765	0.145 927	0.005 750	0.905 37
CI All S + D + T + Q	17,678	-76.157 626	0.147 788	0.007 611	0.975 42
MCSCF All S + D + T + Q	17,678	-76.157 690	0.147 852	0.007 675	0.962 96
Full CI[96]	256,473	-76.157 866	0.148 028	0.007 851	0.975 28

is obtained. If this result is subtracted from the 99.8% value for CISDTQ one finds that together the triple and quadruple excitations account for only 5.13% of the true correlation energy. The total energy of CISDT is only slightly lower than the CISD result (-76.151156 compared to -76.150015 hartrees). This implies that in the CISDT calculation the addition of the triple excitations only recovers 0.77% of the full CI correlation energy. This finding agrees with earlier work done by Shavitt and co-workers,[88-92] who found the addition of triple excitations to have only a small effect on the total energy. If the CISDTQ and CISDT energies are subtracted it was found that the inclusion of the quadruple excitations lowers the energy by 0.006470 hartrees which is 4.4% of the real correlation energy.

Since the CIS energy is equal to the SCF energy due to Brillouin's theorem[97] the CI containing only all double excitations of SHS is compared to our CISD calculation to estimate the effect of including single excitations in CIs containing at least all single and double excitations. Subtracting the two energies yields a difference of 0.000837 hartrees or 0.57% of the correlation energy. However, the total energy is not additive in the sense that, $E(\text{full CI}) \neq E_{SCF} + \Delta E(\text{CIS}) + \Delta E(\text{CID}) + \ldots$. The effects of singles in CISD was found to be 0.57%, but in CIS it is 0.00% of the correlation energy. Since it had been suggested that triple excitations were more important than previously believed[84-87] (see Section III) might not then the triple excitations be correspondingly more important in CISDTQ than in CISDT. In the last paragraph the energy difference between the two CIs had been entirely attributed to the addition of the quadruple excitations, however, some of this lowering could be due to an

increased importance of the triple excitations through their interaction with the quadruples. To see if this is so the single excitation lowering energy of 0.000837 hartrees is added to the CI containing all double plus quadruple excitations (CIDQ) result of SHS and this sum is compared to the total energy of CISDTQ. This procedure furnishes a value for the triples energy lowering of 0.001069 hartrees or 0.72% of the correlation energy. This value is nearly the same as the previous estimate of 0.77%. Therefore, unless the effect of single excitations in "CISDQ" is much different from the value found for CISD it is safe to conclude that the triple excitation contribution is indeed small, on the order of 1% of the true correlation energy.

Next the effects of changing from the SCF to the MCSCF orbitals for our four CI expansions are examined. The greatest change in the amount of correlation energy recovered is found in going from CIS to MCS. As stated before the CIS energy is equal to the SCF energy and has a value of -76.009838 hartrees. However, the MCS energy is much lower than this at -76.083339 hartrees. This wavefunction contains only 20 configurations but accounts for approximately half the full CI correlation energy (49.7%). It was shown earlier that single excitations in CISD and most likely in CIs containing additional higher excitation classes as well are associated with roughly 0.5% of the correlation energy. This implies that <u>the relative importance of a particular class of excitations can be drastically changed by modifying the orbital basis</u>. One way of gauging the extent of the differences between two orbital basis sets is to compare their reference configuration energies.

If the reference configuration is designated (1), then the reference energy of (1) is given by $H_{11} + E_{nucrep}$ where H represents the Hamiltonian and E_{nucrep} is the nuclear repulsion energy. In the MCS case the reference energy is -75.60935 hartrees, up 0.40049 hartrees from the SCF energy. Since this is a very large energy change the CIS and MCS orbitals must be substantially different.

What accounts for this dramatic energy lowering in the MCS case? This is easiestly explained if one invokes the "super CI" or Brillouin-Levy-Berthier (BLB) outlook on the MCSCF problem.[5,13,15,16] The BLB approach takes advantage of what is called the extended Brillouin's theorem.[5] This theorem proves that for a converged MCSCF wave function the Hamiltonian matrix elements between the MCSCF wave-function and the set of singly excited configurations formed from ψ_{MC} are identically zero. Using this theorem an iterative MCSCF procedure can be envisioned. First the Hamiltonian matrix within the space of ψ_{MC} and all singly excited configurations is formed. Then the secular equation is solved and the lowest eigenvector is then used to calculate the changes in the orbitals. Next, the integrals are retransformed and the process is repeated. At convergence the top row and column of the super CI Hamiltonian is zero except for the diagonal element which is the MCSCF energy. What this implies is that for a converged MCSCF solution all relative single excitations have been annihilated. Now lets return to the MCS case. The single excitations from MCS are just double excitations with respect to the SCF reference. As was shown earlier double excitations from the SCF account for roughly 94% of the correlation energy in CISD and,

therefore, annihilation of these configurations should lead to a large energy lowering for MCS over CIS. As we discovered this was indeed the case with MCS accounting for over half of the CISD correlation energy.

Moving on to the MCSD results a total energy of -76.150269 hartrees is obtained. This result is only 0.0003 hartrees below the corresponding CISD energy. This can be understood by employing the same reasoning as used to explain the massive energy lowering for MCS. In the BLB[5,13,15,16] picture the MCSD wavefunction annihilates certain classes of triple excitations, these being the corresponding single excitations for this MCSCF. Since the triple excitations are relatively unimportant in an energetic sense, we would expect little difference between the energies of CISD and MCSD. Also interesting to note is that while the difference between CISD and CISDT is .001141 hartrees the splitting of CISD and MCSD is only 0.00254 hartrees. Therefore, in this instance annihilation of the triple excitations in MCSD recovered less than a quarter of the triples' contribution to the energy of CISDT.

The triple excitations in MCSDT are, however, very important. For this MCSCF a total energy of -76.155765 hartrees was found or an energy lowering relative to CISD of 0.005750 hartrees. This difference is over 5 times as large as the corresponding lowering of CISDT relative to the single and doubles CI. The extended Brillouin's theorem[5] can again be invoked to explain this result. In the CISDT case the triple excitations were found to be rather unimportant, however, in the MCSDT the annihilation of the corresponding quadruple

excitations increased the importance of the triple excitations drastically. This is due to the fact that in the CIs performed using SCF orbitals the quadruples were second only to the double excitations in energetic importance. 98.6% of the full CI correlation energy was obtained for MCSDT (compared to 95.5% for CISDT) or over 70% of the remaining correlation energy beyond the single and double excitation CI.

For MCSDTQ we would expect very little energy lowering for two reasons. One that quadruple excitations are expected to be unimportant and their annihilation would recover little additional correlation energy and two, that CISDTQ has already accounted for over 99.8% of the correlation energy leaving very little left to include. As anticipated MCSDTQ was only marginly lower in energy than CISDTQ, recovering only 27% of the remaining correlation energy. Therefore in the straight CI and in the MCSCF calculations the relative importance of triple and quadruple excitations was switched. In the case of the SCF orbital CIs, quadruple excitations accounted for 4.4% of the correlation energy while the triple excitations only netted 0.8%. The order is inverted in the MCSCF calculation with the triple excitations worth 3.7% versus the quadruple excitations garnering 1.3% of the correlation energy. This result shows that the level of importance of a given class of excitations is not an absolute, but is dependent of the choice of orbital basis.

C. Comparison with Perturbation Theory Results.

In this section the current results are comapred with the perturbation theory predictions of Bartlett[98] and of Wilson and

Guest.[99] Their calculations were all performed using the Möller-Plesset variant of many-body perturbation theory (MBPT)[100] and were carried out using the same double-zeta basis set and molecular geometry specified in section IIIA. Their results are summaried in Table V.

Summing the second and third order energies together with the SCF energy yields -76.150707 hartrees which is very close to the CISD result of -76.150015 hartrees. A calculation of this type is abbreviated MP3. In general one would expect to find the single and double excitation CI energy to be only slightly more reliable than the sum of the perturbation energies through third order. In fact there exist CI schemes in which the sum to third order using Möller-Plesset partitioning is obtained on the first iteration of the CI procedure.[101] With reference to CISD, the missing ingredients in MP3 are the effects of single excitations (which do not enter until the fourth order) and relaxation of the reference configuration with the remainder of the configuration set.

Adding the fourth order single, double and quadruple excitation contributions to MP3 yields what is called SDQ-MBPT(4) by Bartlett. In the case of water the energy obtained for SDQ-MBPT(4) is -76.155513 hartrees and this calculation recovers -0.145675 hartrees of correlation energy. Computing this energy is proportional in time to finding the CISD solution, that is they are both 0^6 processes with respect to the number of basis functions.[100,102] However, the SDQ-MBPT(4) result comes much closer to the exact energy recovering 98.4% of the correlation energy and 70.0% of the energy attainable beyond CISD.

Table V. Comparison between exact results for double zeta H_2O and approximate results obtained via variational and perturbation techniques.

Variational Results

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	-76.009 838	0.0	---	1.0
CI All Double Excitations (D)	342	-76.149 178	-0.139 340	+0.000 837	0.979 38
CI All Single and Double Excitations (S+D)	361	-76.150 015	-0.140 177	0.0	0.978 74
CI All S + D + T	3,203	-76.151 156	-0.141 318	-0.001 141	0.978 19
CI All D + Q	14,817	-76.155 697	-0.145 859	-0.005 682	0.976 68
CI All S + D + T + Q	17,678	-76.157 626	-0.147 788	-0.007 611	0.975 42
Full CI	256,473	-76.157 866	-0.148 028	-0.007 851	0.975 28

Many-Body Perturbation Theory (Bartlett)[a]

	Correlation Contribution	Cummulative Correlation Energy	Energy Relative to CI S + D
Second-Order Energy	-0.139478	-0.139478	+0.000699
Third-Order Energy	-0.001391	-0.140869	-0.000692
Fourth-Order			
Linked Diagram Double Excitations	-0.003083	-0.143952	-0.003775
Linked Diagram Single Excitations	-0.000908	-0.144860	-0.004683
Linked Diagram Quadruple Excitations	-0.000815	-0.145675	-0.005498
Linked Diagram Triple Excitations[b]	-0.001364	-0.147039	-0.006862

	Correlation Contribution	Cummulative Correlation Energy	Energy Relative to CI S + D
Coupled Clusted Doubles (CCD)		-0.145435	-0.005258
CCD + Fourth-Order Singles and Triples		-0.147707	-0.007530

[a] R. J. Bartlett, private communication.

[b] Reference 14.

What is missing now from the fourth order energy is the triples diagram contribution. Calculating this term, however, is inherently more difficult than performing a singles and doubles CI and requires an 0^7 process with respect to the number of basis functions.[100] Wilson and Guest found this term to be equal to -0.001364 hartrees and when summed into the SDQ-MBPT(4) result yields -0.147039 hartrees of correlation energy. A calculation of this sort is labeled SDTQ-MBPT(4) and should include the majority of the triple and quadruple excitation effects found in CISDTQ. This is indeed the case with SDTQ-MBPT(4) obtaining 99.5% of the CISDTQ and 99.3% of the full CI correlation energy.

Also performed by Bartlett was a coupled cluster calculation[100] which is very similar to an MBPT computation except that now certain diagrams are summed to infinite order. The calculation that was carried out is called coupled cluster doubles (CCD) and in it double and disconnected quadruple excitation diagrams are summed to all orders. The straight CCD result is -76.155273 hartrees and recovers 67.0% of the additional correlation energy past CISD. What is neglected from this calculation is the effect of single and triple excitations. Adding the fourth order singles and triples contribution to CCD yields -76.157545 hartrees or a correlation energy of -0.147707 hartrees. This is the most extensive perturbation theory calculation done on this particular system and obtains 99.8% of the correlation energy.

All three calculations SDQ-MBPT(4), SDTQ-MBPT(4) and CCD plus the fourth order single and triple contributions are much less expensive to compute than CISDTQ, but recover almost all of the

additional correlation energy beyond CISD. This would seem to suggest that these pertubation theory approaches are the way to go if additional correlation energy is required past the CISD level. Though this appears to be the case for water there are many problems associated with MBPT which would tend to diminish the usefulness of these procedures in general. First, Möller-Plesset partitioning is not well-defined for open-shell systems unless Φ_0 is found using the unrestricted Hartree-Fock (UHF) formalism.[103] The problem with UHF is that it's solutions are not generally eigenfunctions of $\hat{S}^2$. Secondly, if the molecule is not well described by a single configurration, MBPT expansions about the dominate reference often converge very poorly.

In closing this section a series of A_k and B_k[104-107] calculations performed on H_2O using the same basis set and geometry are presented and compared to our CI and MCSCF results. These computations are unpublished work done by Laidig and Schaefer and are summarized in Table VI. For readers who are unfamiliar with the A_k and B_k procedures, their theory is presented in Appendix I.

The basic idea behind this series of calculations was to correctly include the effects of single and double excitation configurations while using the A_k or B_k perturbation procedure to estimate the triples and quadruples contribution to the correlation energy. If the perturbation procedures are carried out to infinite order the CISDTQ energy of -76.157626 hartrees will be obtained, so all energy comparisons are relative to CISDTQ and not the full CI result. In both the A_k and B_k calculations the primary configuration space P

Table VI. Results for A_k and B_k calculations in water.

Wavefunction	Total Energy	Correlation Contribution	Cummulative Correlation Energy
SCF	-76.009838	0.0	0.0
CISD	-76.150015	-0.140177	-0.140177
CISD + Davidson corr.	-76.155913	-0.005898	-0.146075
CISDTQ	-76.157626	-0.007611	-0.147788
Full CI[a]	-76.157866	-0.007851	-0.148028
A_{k_2}	-76.158226	-0.008211	-0.148388
A_{k_3}	-76.156789	+0.001436	-0.146951
B_{k_2}	-76.158871	-0.008856	-0.149033
B_{k_3}	-76.157125	+0.001746	-0.147287

[a]Obtained from reference 96.

was chosen as the set of the reference plus all its single and double excitations and will be designated SD. The secondary space Q was composed of the set of all triple and quadruple excitations henceforth called TQ. In the A_k procedure the P space Hamiltonian is diagonalized to yield ψ_0 and Epstein-Nesbet perturbation theory[108-109] is applied to this wavefunction. If this expansion is truncated after the second or third orders[110] the subsequent calculations will be labeled A_{k_2} or A_{k_3} respectively. Analogously $E_{A_{k_2}}$ and $E_{A_{k_3}}$ are defined as the second and third order energies. In the second order B_k method (B_{k_2}) the off-diagonal Q-Q block of the Hamiltonian is set to zero and the matrix is diagonalized. $E_{B_{k_2}}$ is found by subtracting the P space CI energy from the energy associated with the modified P + Q Hamiltonian. B_{k_3} calculations can also be carried out if desired and $E_{B_{k_3}}$ is found using equation (95) found in Appendix I. The time required to perform an A_{k_2} or B_{k_2} calculation is much less than in the corresponding CISDTQ case as the vast majority of the Hamiltonian elements are never computed. If the calculations are extended to the third order, however, all elements of the H matrix must be determined. It is nevertheless a simpler procedure than the CI since the Hamiltonian elements are only used once as compared to once per diagonalization iteration in the CI.

For water the A_{k_2} procedure yielded a total energy of -76.158226 hartrees which corresponds to a correlation energy of 0.148388 hartrees. This is 100.4% of the CISDTQ correlation energy and 107.9% of the triples and quadruples contribution. Using the B_{k_2} method a somewhat lower energy of -76.158871 hartrees was found and the resulting

correlation energy is -0.149033 hartrees. The B_{k_2} energy is lower than the A_{k_2} value by around 600 microhartrees. This is expected because relaxation of the P by the Q space is allowed in the B_{k_2}, but not the A_{k_2} procedures. The fact that both the A_{k_2} and B_{k_2} energies were below the full CI result was not anticipated, however. This is most likely due to the oscillatory nature of Epstein-Nesbet perturbation theory.[73] In most cases E_3 is positive and a substantial fraction of the size of E_2 in magnitude. Since ψ_{CISD} is a very complete zeroth order wave function it is reasonable that E_2 overshoots and E_3 corrects back toward the CISDTQ energy.

Moving on to the third order results, $E_{A_{k_3}}$ and $E_{B_{k_3}}$ are indeed positive with values of 0.001436 and 0.001746 hartrees respectively. Since $E_{A_{k_2}}$ and $E_{B_{k_2}}$ are -0.008211 and -0.008856 hartrees, this makes the third order energies between a factor of 5 and 6 smaller than the corresponding second order results. In terms of total energies A_{k_3} yields -76.156789 hartrees while B_{k_3} is slightly less at -76.157125. This means the A_{k_3} procedure recovers 99.4% of the CISDTQ correlation energy and that B_{k_3} net 99.7%. These results are impressive since they obtain almost all the correlation energy of CISDTQ and are substantially simpler to obtain. To get comparable accuracy in MBPT all fourth order perturbation terms must be computed.

D. Wavefunction Analysis.

In this section the various wavefunctions for water will be considered. The principle CI coefficient C_1 for our eight CI and MCSCF calculations and for the full CI of SHS are listed in the last column of Table IV. In all cases this configuration is the

reference configuration

$$\Phi_1 = 1a_1^{\ 2}\ 2a_1^{\ 2}\ 1b_2^{\ 2}\ 3a_1^{\ 2}\ 1b_1^{\ 2} \quad .$$

In addition the ten largest CI coefficients and their corresponding configurations are displayed in Table VII. There is no table for the ordinary CI results, as wavefunctions of this sort have been examined previously in the literature[111] and we found no peculiar features warranting additional attention.

For the three calculations using the SCF orbitals, CISD, CISDT, and CISDTQ, the corresponding values of c_1 are 0.97874, 0.97819 and 0.97542. All three of these coefficients are typical of previously reported values for the $^1A_1$ ground state of water.[111] Also the fact that c_1 drops in magnitude with the addition of each new class of excitations is to be expected since if the off-diagonal Hamiltonian elements connecting the new class to the previous ones are small as expected, the coefficients for this new class will also be small. This will tend to decrease the importance of all the principle configurations by a constant factor. Continuing in this trend is the primary coefficient for the full CI of SHS with a value of 0.97528 which is very close to the CISDTQ aresult. Moving on to the MCSCFs the values for MCSD and MCSDTQ are 0.9773 and 0.9630 respectively. Each of these values is slightly less than the analogous straight CI result. This is reasonable since one usually expects the value of the principle coefficient to decrease in MCSCF versus ordinary CI calculations with the same configuration list. The reason behind this is that the SCF orbitals used in the standard CI particularly favor

the reference configuration due to the nature of the SCF procedure. Also the MCSDTQ c_1 is smaller than the MCSD value as expected.

In the MCS and MCSDT cases one notices for the first time a qualitative change in the form of the wavefunction. In both of these calculations the value of the most important coefficient is much smaller than in the previous examples, namely 0.89326 for MCS and 0.90537 for MCSDT. Since using the SCF orbitals in a CI involving only single excitations leads to zero correlation energy one would expect the orbitals to change dramatically in the MCS. The low value of c_1 for MCS suggests that the orbitals did indeed change and the large amount of correlation energy obtained for this wavefunction implies that the reference-single excitation Hamiltonian matrix elements were relatively big. The situation for MCSDT is very similar to MCS. This can be seen both in comparing their principle coefficients and also in examining their orbitals. The final MCSDT orbitals were much more like the MCS than the MCSD orbitals and would have made a much better set of starting orbitals in the MCSCF calculation. The fact that c_1 is slightly larger in MCSDT than in MCS is due to the presence of double excitations which tend to favor a greater value for the reference coefficient.

Next lets examine the ten most important configurations for the four MCSCF calculations. Since the wavefunction for MCSD most closely resembles its corresponding CI its coefficients will be examined first. As expected the majority of the most important configurations are double excitations namely 8 of the first 10 largest coefficients. Besides the reference the list also contain one single excitation,

Table VII. Ten most important configurations in the four multiconfiguration self-consistent-field wavefunctions for H_2O. The reference configuration Φ_0 is $1a_1^2\ 2a_1^2\ 1b_2^2\ 3a_1^2\ 1b_1^2$.

MCS		MCSD		MCSDT		MCSDTQ	
Φ_0	0.8933	Φ_0	0.9773	Φ_0	0.9054	Φ_0	0.9630
$1b_1 \to 2b_1$	0.2429	$1b_1^2 \to 2b_1^2$	0.0526	$1b_2 \to 2b_2$	0.1715	$1b_1 \to 2b_1$	0.1079
$3a_1 \to 4a_1$	0.2102	$3a_1^2 \to 4a_1^2$	0.0439	$1b_2 \to 3b_2$	0.1645	$1b_2 \to 2b_2$	0.0819
$1b_2 \to 2b_2$	0.1751	$3a_1 1b_1 \to 4a_1 2b_1$	0.0431[a]	$3a_1 \to 4a_1$	0.1643	$1b_1^2 \to 2b_1^2$	0.0508
$1b_2 \to 3b_2$	0.1410	$1b_2^2 \to 2b_2 3b_2$	0.0411	$1b_1 \to 2b_1$	0.1477	$3a_1^2 \to 4a_1^2$	0.0459
$2a_1 \to 5a_1$	0.1136	$1b_2^2 \to 2b_2^2$	0.0400	$1b_2 \to 4b_2$	0.1223	$3a_1 \to 4a_1$	0.0444
$2a_1 \to 6a_1$	0.1015	$3a_1 1b_1 \to 4a_1 2b_1$	0.0395[b]	$3a_1 \to 5a_1$	0.0880	$3a_1 1b_1 \to 4a_1 2b_1$	0.0439[a]
$3a_1 \to 5a_1$	0.0938	$1b_1 \to 2b_1$	0.0395	$3a_1 \to 6a_1$	0.0837	$1b_2^2 \to 2b_2 3b_2$	0.0435
$1b_2 \to 4b_2$	0.0851	$3a_1 1b_2 \to 5a_1 2b_2$	0.0349[a]	$2a_1 \to 7a_1$	0.0832	$1b_2^2 \to 3b_2^2$	0.0420
$2a_1 \to 7a_1$	0.0632	$1b_2^2 \to 3b_2^2$	0.0329	$2a_1 \to 5a_1$	0.0810	$3a_1 1b_1 \to 4a_1 2b_1$	0.0390[b]

[a]There are two linearly independent singlet (S=0) spin eigenfunctions for four unpaired electrons. This is the first.

[b]Second spin coupling.

$1b_1 \rightarrow 2b_1$, as the eighth most important configuration. However, the double excitation between these same orbitals, $1b_1^2 \rightarrow 2b_2^2$ is second only to the reference in importance and, therefore, one would expect the corresponding single excitation to be relatively important also.

Comparing the MCSDTQ wavefunction to that of MCSD we note that 9 out of the 10 most important configurations are on both lists. This implies that qualitatively the two wavefunctions are very similar. Furthermore, the major configurations for these two MCSCF wavefunctions are nearly identical with their corresponding straight CI results. The most noticeable difference between MCSD and MCSDTQ is in the relative importance of the $1b_1 \rightarrow 2b_1$ excitation. In MCSDTQ the configuration is now the second in importance, while the double excitation, $1b_1^2 \rightarrow 2b_1^2$ has moved down to the fourth position. This is the opposite of the MCSD ordering. We have no explanation for this result except to note that the $1b_1 \rightarrow 2b_1$ excitation is the most important excitation in MCS.

Turning now to the MCS and MCSDT wavefunctions one notices that the nine excitations, No. 2-No. 10, are all single excitations. Since we only included single excitations in MCS this is trivially obvious, but why is this the case in MCSDT which contains double and triple excitations as well. One possible explanation for this can best be examined by reverting to the BLB picture.[5,13,15,16] In this framework the triple excitations at convergence have annihilated the quadruple excitations. If the final orbitals are similar to the MCS orbitals then the reference and its single excitations are strongly coupled. Since triple excitations are double excitations from the

class of singly excited configurations than the corresponding Hamiltonian matrix element between a given single and triple excitation is expected to be relatively large. In diagonalizing the H matrix this allows the triples to indirectly couple rather strongly to the reference configuration.

Up to now we have been stressing the differences between the various orbitals used and between the values of the configuration coefficients. However, it should be noted that the overlaps between the total wavefunctions for any of these CI or MCSCF calculations is very high, at least 99% and their AO density matrices are very similar. In the cases where the principle coefficient was much smaller than the typical value of 0.97 (MCS and MCSDT) the orbitals also varied a lot. Since all of these wavefunctions are very complete from a total energy viewpoint (our worst correlated calculation MCS recovered 99.9% of the total energy) the orbital and configuration coefficient changes must be coupled in a cancelling manner.

E. Conclusions for Water.

The most interesting result obtained from the series of MCSCF and CI calculations performed on water was that the relative importance of classes of excitations depend largely on the orbital choice. The most dramatic example we found was for the CI involving all single, double and triple excitations. In the CI employing SCF orbitals the double excitations were the most important class. However, using the MCSCF orbitals the single excitations became the most important. In CISDT the triple excitations accounted for only 0.8% of the correlation energy, but in MCSDT 3.7% of this energy was recovered. Therefore, in

the latter orbital basis the triple excitations were more important by nearly a factor of five.

It was also found that the small MCS calculation recovered almost 50% of the true correlation energy. This calculation with a total of 20 configurations obtained 49.7% of the correlation energy of the 256,473 configuration full CI. It would seem then for molecules MCS calculations would be a relatively inexpensive way to get a significant fraction of the correlation energy.

Lastly it was discovered, at least in these calculations, that triple excitations in CIs using canonical SCF orbitals were much less important than Wilson had initially claimed. As expected the quadruple excitations turned out to be the dominant excitations beyond doubles and as a class recover over 6.5 times as much energy as the triple excitations. This conclusion agrees well with past higher excitation CI studies of Shavitt and co-workers.

IV. An MCSCF Study on the Ground State of Ozone

In our second MCSCF study we chose to investigate the $^1A_1$ ground state of the ozone molecule.[34] In contrast to water, ozone is not correctly described using simple single-configuration SCF theory. In fact it was found in an earlier study that a triplet state of O_3 was predicted to be below the lowest $^1A_1$ state by 2 eV at the SCF level.[112] Therefore, to attain qualitative accuracy additional configurations must be included in the theoretical framework. The simplest such treatment is the two-configuration SCF (TCSCF) procedure in which both the orbitals and the CI coefficients are variationally optimized. In this paper a number of different results are presented from calculations performed beyond the SCF level. These include the TCSCF and a series of larger MCSCF and CI results.

A feature we wish to explore is the extent of biradical character present in the ground state of O_3. In the mid 70's Goddard and co-workers[112,113] discovered that this state of ozone could be well represented as a biradical, that is, it could be described as a system containing two singly occupied π orbitals that do not interact. They found that if the highest π orbital is represented by a GVB pair the resultant non-orthogonal orbitals are centered on each of the terminal oxygen atoms. Furthermore, these two singly occupied orbitals have a very small overlap ($S = 0.28$) and are, therefore, only weakly interacting. The corresponding wavefunction composed of these two non-orthogonal orbitals can be transformed to an equivalent TCSCF wavefunction involving only orthogonal orbitals. Using these orthogonal orbitals the two configurations in the TCSCF are

$$\Phi_1 = 5a_1^2\ 3b_2^2\ 1b_1^2\ 6a_1^2\ 4b_2^2\ 1a_2^2 \qquad \text{reference 1}$$

$$\Phi_2 = 5a_1^2\ 3b_2^2\ 1b_1^2\ 6a_1^2\ 4b_2^2\ 2b_1^2 \qquad \text{reference 2} \quad .$$

Due to the small overlap between the two localized orbitals it is expected that Φ_1 alone is a poor description of the exact ozone wavefunction and that the coefficient of Φ_2 in a TCSCF procedure will be relatively large. This is indeed the case in our present TCSCF calculation in which $c_1 = 0.876$ and $c_2 = -0.481$.

Another reason for chosing the ozone molecule is that it should present more of a challenge computationally than water. Since all orbitals were active in the water MCSCFs the calculations tended to display well behaved convergence. In at least one of the ozone MCSCF calculations, however, the orbitals belong to all three categories: core, active and virtual. In general calculations involving core and virtual orbitals are more difficult to converge. See sections IIB and IIC for more details and examples of the rates of convergence.

A. Theoretical.

For ozone as in the case of water we employed the standard contracted Gaussian double-zeta basis set of Dunning and Huzinaga.[93,94] This basis set was considered adequate since our study was mainly qualitative and we were primarily interested in the varying amounts of correlation energy obtained using different CI expansions and orbital sets. The geometry was fixed at $\theta(000) = 116.8°$ and $r(O-O) = 1.271$ Å. This corresponds to a precise cartesian geometry of (0.0, 0.0,0.0) for the central oxygen and (0,±2.0457412,1.2585484) for the

two terminal atoms. This particular choice for the geometry is very close to the experimental geometry[114] of $r_0(0-0) = 1.278$ Å and $\theta_0(000) = 116.8°$. The SCF and subsequent CI and MCSCF calculations were carried out in C_{2v} symmetry.

As stated before from previous studies[113-114] it was found that the $^1A_1$ ground state of ozone has two dominant configuration (1) and (2). The standard procedure for a system like this would be to first carry out a two-configuration SCF(TCSCF) calculation. This wavefunction should be of at least qualitative accuracy in the region of the O_3 potential energy surface near the equilibrium geometry. If a more quantitative result were desired the TCSCF orbitals could be used in a CI calculation containing all single and double excitations from both of the two important reference configurations. In this study we intend to see if the procedure outlined above is in fact the most economical method for obtaining the most correlation energy. That is, is there another CI expansion and/or orbital basis set choice of at least the same quantitative accurach which is cheaper from a computational viewpoint.

One can envision an alternate procedure to the standard one presented above in which an MCSCF calculation involving all single and double excitations from just the most important reference is performed. First notice that reference (2) is included in this configuration set as it is the $1a_1^2 \rightarrow 2b_1^2$ double excitation from reference (1). Also a large percentage of single and a small percentage of double excitations from (2) are included. Since this is an MCSCF procedure both the orbital and configuration coefficients

are variationally optimized to minimize the total energy. Thus it is possible that reference (2) could end up with a large CI coefficient analogous to the TCSCF case if this leads to the lowest energy.

The major drawback to the one reference single and double excitation MCSCF is that most of the single and double excitations from (2) are missing from the configuration set. We reasoned that if the number of configurations were large enough and the orbitals and CI coefficients were given complete freedom to adjust these additional configurations may not be necessary. The question we set out to answer then was, "can a deficiency in the wavefunction, i.e., the omission of certain single and double excitations, be overcome by allowing the wavefunction additional flexibility?" In this case the additional flexibility would be in the complete optimization of the molecular orbitals.

Three additional calculations were also performed. The first of these was a standard single plus double excitation CI from (1) using the straight SCF orbitals. This calculation was needed to give us some idea of how much extra correlation energy is obtained if an MCSCF procedure is carried out in this same configuration space. The second calculation was a CI containing all single and double excitations from both references (1) and (2) using the one-configuration SCF (OCSCF) orbitals. This was performed to determine if the TCSCF orbitals were necessary or if the OCSCF orbitals are adequate. Lastly, an MCSCF in the full two reference single plus double excitation configuration space was computed. The energy associated with this wavefunction is a lower bond with respect to the other calculations

done in our study on ozone and will be treated as a benchmark with which to compare the other wavefunctions.

B. Ozone Energetic Results.

We begin this section by first comparing the OCSCF results to the analogous TCSCF calculation (all energetic results are collected in Table VIII). At the chosen geometry using a double-zeta basis set and an energy of -224.3104 hartrees is found for the TCSCF case. This very large energy lowering along with the TCSCF coefficients for the references (1) and (2) (c_1 = 0.876, c_2 = -0.481) confirm that the OCSCF description is quite poor and that the TCSCF gives a much more qualitatively accurate picture of the $^1A_1$ ground state of ozone. The fact that the TCSCF wavefunction is far superior to the OCSCF wavefunction came as no surprise since it is a consequence of the biradical nature of O_3 established earlier by Dunning, Hay and Goddard.[112,113]

A single and double excitation CI from reference (1) was next carried out using the standard SCF orbitals. This CI contained a total of 6825 configurations and for brevity's sake will be labeled CI6825. Examining the list of important configurations for this calculation one discovers that c_1 = 0.933 and c_2 = -0.149. The next largest coefficient is for the configuration $\Phi_3 = \ldots 5a_1^2\ 3b_2^2\ 6a_1^2\ 4b_2^2\ 1a_2^2\ 2b_1^2$ with a coefficient of -0.065. Since this value is much smaller than c_2 we felt justified in only carrying out a TCSCF as opposed to a three- or four-configuration SCF calculation to obtain a qualitatively accurate wavefunction. In the CI6825 case the c_2 coefficient while still large is much smaller than in the corresponding

TCSCF case. This suggests that the CI6825 wavefunction contains little of the biradical character inherent in the ozone molecule and may be qualitatively lacking in this regard. Examining the energy of CI6825, however, one notices a nearly four-fold increase in the amount of correlation energy recovered compared to the earlier TCSCF result. In an absolute energetic sense then the 6825 configuration CI is far superior to the two-configuration SCF.

Optimizing the orbital coefficients in the CI6825 leads to an MCSCF in the same single and double excitation space. This computation will be abbreviated as MC6825. Performing this calculation leads to an energy lowering of 0.0016 hartrees compared to the CI using OCSCF orbitals, or a total energy of -224.6155 hartrees. The energy difference for the analogous water calculations (the difference between the CISD and MCSD energies) is 0.0003 hartrees or roughly 6.5 times less than in ozone. If this difference is converted into the fraction of the total energy in both cases, one finds the effect of using MCSCF instead of the OCSCF orbitals is still more important in O_3 than in H_2O, namely by a factor of around ?.[2] This energy lowering is much less than we expected for O_3, however, in light of the arguments presented in the preceeding section. The coefficient of reference (2) is only -0.162 in the MC6825 case, only slightly larger than the CI6825 value of -0.149. As in the case of the CI, the MC wavefunction badly underestimates the biradical character of ozone, and thus is not acceptable for studying the $^1A_1$ potential energy surface near the minimum.

For the single and double excitation CI from both important

Table VIII. Summary of large MCSCF results for the $^1A_1$ ground state of ozone near its equilibrium geometry.

Number of Configurations	Description of Configurations	Type of Wavefunction	Energy (hartrees)	Coefficients of Two Most Important Configurations C_1	C_2
1	(1) in Text	SCF	-224.207 656	1.000	---
2	(1) and (2) in Text	TCSCF	-224.310 437	0.8765	-0.4815
6,825	All S + D Excitations Relative to (1)	CI[a]	-224.613 929	0.9331	-0.1493
6,825	" "	MCSCF	-224.615 493	0.9343	-0.1620
13,413	All S + D Excitations Relative to (1) and (2)	CI[a]	-224.640 183	0.8935	-0.2957
13,413	" "	CI[b]	-224.647 645	0.8762	-0.3553
13,413	" "	MCSCF	-224.649 326	0.8870	-0.3371

[a]CI used one-configuration SCF orbitals.

[b]CI used two-configuration SCF orbitals.

reference configurations (1) and (2) using the TCSCF orbitals a total energy of -224.64765 hartrees was determined. This configuration set is the largest used in this study and is composed of 13,413 individual configurations. When the OCSCF orbitals are employed instead, the energy is found to be higher by 0.0075 hartrees leading to a total energy of -224.640183 hartrees. To avoid confusion the CI using TCSCF orbitals will be called TC13413 and the equivalent calculation employing OCSCF orbitals, OC13413. In both of these calculations the magnitude of the coefficient of (2) was substantially larger than in the one reference CIs, CI6825 and MC6825. In the case of OC13413 C_2 is -0.2957, over twice as large as the CI6825 value of -0.1493, and in TC13413 the coefficient is still greater at -0.3553. It is obvious then that the single and double excitations that are missing from the 6825 configuration results are necessary for a qualitatively accurate description of ozone, and that just optimizing the orbitals for this smaller CI expansion will not overcome this deficiency. As it turns out the MC6825 calculation only recovered 6.0% of the difference between the OC13413 and CI6825 energies and 4.6% of the TC13413-CI6825 energy splitting.

An MCSCF calculation within the full 13413 configuration set was also performed yielding a total energy of -224.649326 hartrees giving us our lowest variational result. This result is 0.0091 hartrees below the OC13413 energy, but is only 0.0017 hartrees below the TC13413 value. This is very interesting in that it illustrates that while either choice of OCSCF or TCSCF orbitals gives a qualitatively accurate description of ozone, the TCSCF orbitals are much superior

from an energetic standpoint. Also very little is gained in fully optimizing the complete 13413 configuration list as just optimizing the important references (1) and (2) yields 81.6% of the recoverable energy. It appears that from a cost effectiveness standpoint the TCSCF followed by a CI containing all single and double excitations from the two optimized references is the most inexpensive way to get a quantitatively accurate wavefunction with the correct biradical description for O_3.

A slightly puzzling aspect of this research concerns the amount of energy lowering in going from CI6825 to MC6825 compared to the analogous lowering between TC13413 and MC13413. Since the larger CI is much more complete than the one reference single and double excitation CI, i.e., has nearly twice as many configurations, it was expected that the energy recovered in going from the OCSCF to the MCSCF orbitals would be less than for performing the same procedure in the smaller CI. However, this was not the case. In the 6825 configuration CI 0.0016 hartrees were gained, while for the bigger 13,413 configuration set a value of 0.0017 hartrees was obtained. At this time we have found no straight-forward explanation for this.

Returning to our best calculation, MC13413, the values for the coefficients of (1) and (2) are found to be 0.887 and -0.337 respectively. Using this wavefunction we would like to be able to predict the extent of pure biradical character in the $^1A_1$ ground state of ozone. In the case of an idealized pure biradical the coefficients of the two important configurations would be equal except for a phase factor of ± 1 and all the other configuration coefficients would

be zero. Therefore, if ozone were a perfect biradical one would expect that $c_1 = -c_2 = 1/\sqrt{2}$. Hayes and Siu[115] about a decade ago devised a simple formula for estimating the fraction of biradical character B employing just the coefficient of the second most important reference configuration,

$$B = \frac{c_2^{\,2}}{(1/\sqrt{2})^2} = 2\ c_2^{\,2} \quad . \tag{86}$$

Notice that if $c_2 = 1/\sqrt{2} \approx 0.707$ then B = 1. (The last column of Table VIII lists the B values for our series of calculations). Applying this procedure to MC13413 yields our best estimate for the fraction of biradical character in O_3 of 0.227 or 22.7%. Comparing this best result to the MC6825 and CI6825 values of B = 0.052 and B = 0.044 respectively one sees that these calculations seriously underestimate the amount of biradical character. On the opposite extreme the TCSCF over estimates the biradical nature of ozone by about 100% with a corresponding B value of 0.464. The standard 13413 configuration CIs, however, agree quite well with our estimate of 23%. Since any wavefunction for the ground state of ozone in which the values of the coefficients of (1) and (2) are known can be used to determine B we can compare our best guess with estimates predicted from the published coefficients of others. As an example, the calculated B value for the 1557 configuration CI wavefunction of Hay and Dunning[116] has a value of .204 or 20.4%. One reason behind choosing their calculation is that it was done using a larger double-zeta plus polarization basis set. The largest deficiency

of our study is that we omitted polarization functions from our basis set as the CI expansions became too large to handle. In many cases, however, these higher angular momentum functions are necessary if quantitative chemical properties such as the equilibrium geometry are desired. In this particular system though we believe that these polarization effects are small with respect to the determination of the biradical character of O_3 since the Hay and Dunning[116] result of 20.4% agrees well with our best estimate of 22.7%.

C. Conclusions for Ozone.

In section IV a number of large MCSCF and CI calculations have been presented. Our largest MCSCF wavefunction for ozone contained 13,413 configurations which is the second largest MCSCF calculation reported to date (excluding full CIs). We discovered that the energy of the 13,413 configuration MCSCF is only marginally lower than the 13,413 configuration CI using TCSCF orbitals. Since the former calculation is much more expensive than the latter the inclusion of the 13000 plus additional configurations in the orbital optimization procedure was found to be unnecessary. In contrast, though, the use of the TCSCF instead of the OCSCF orbitals in this CI computation leads to an energy lowering of 0.0075 hartrees which is over 4 times as much energy as was gained in going from TCSCF to MCSCF orbitals. Our recommendation would thus be to first carry out a TCSCF and follow it with a CI including all single and double excitations from both references.

It was also discovered that an MCSCF in the one reference single plus double excitation space is very close in energy to the straight CI result and only recovered between 4-6% of the energy splitting between

the one and any of the two reference single and double excitation CIs. In addition, the magnitude of c_2 was much closer to the CI6825 result than any of the 13,413 configuration calculations. We therefore must answer the question posed earlier, "can a deficiency in the wavefunction be overcome by allowing the wavefunction additional flexibility", with a qualified no. That is to say that, at least in the case of ground state O_3, omission of certain classes of configurations from the CI expansion cannot be compensated for by allowing the orbitals to change.

Lastly we predict a biradical character for ozone of approximately 23%. This compares favorably with estimates made employing previously published high quality wavefunctions.[116] It was also found that both the TCSCF and the one reference CI wavefunction predict substantially different amounts of biradical character. It was expected that the TCSCF overestimates by around 100%. It will be interesting to see if this same trend is found in other molecules with a sizeable biradical nature such as the cyclopentadiyl system.[117]

V. An Examination of the Lowest $^2E''$ State of the Cyclopropenyl Radical

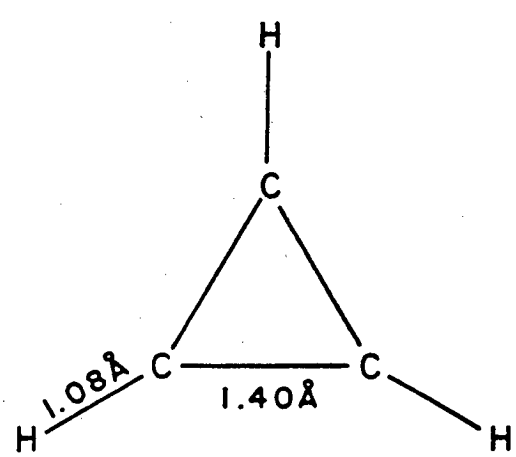
The cyclopropenyl radical C_3H_3 has received much attention because it is a simple example of a molecule that is expected to undergo a Jahn-Teller distortion.[118-125] If an electron is added to the lowest unoccupied molecular orbital (LUMO) of the D_{3h} ground state of the cyclopropenyl cation it will enter a degenerate e" orbital to form the lowest $^2E''$ state of the radical. According to the Jahn-Teller Theorem[126] a non-linear molecule of E symmetry will spontaneously distort in such a way as to lift the degeneracy. In the case of the C_3H_3 radical this distortion will lower the symmetry to C_{2v} or possibly C_s or C_2.[125] If the true wavefunction within the Born-Oppenheimer approximation were known at the D_{3h} and surrounding geometries it would be easy to find out what the "real" equilibrium structure looks like. However, when this complete wavefunction is not known certain complications arise. If a full CI within a given basis set could be computed at any D_{3h} geometry it would lead to a wavefunction having D_{3h} symmetry whether the calculation was forced to be of this symmetry or not. However, for calculations done below the full CI level the wavefunction symmetry may break (i.e., belong to a point group of lower symmetry) unless the correct symmetry is forced by the program.[124,125] Just such a situation arises for the C_3H_3 radical.

In the case of the cyclopropenyl radical if SCF or truncated CI or MCSCF calculations are performed at a D_{3h} geometry the resulting wavefunctions will be of only C_{2v} symmetry unless otherwise constrained.[127] The $^2E''$ state corresponds to a pair of states in the lower C_{2v} point group: $^2A_2$ and $^2B_1$, which should physically be degenerate in energy

at the D_{3h} geometry. However, if the calculations are carried out in C_{2v} symmetry the energies of the two states will be different. The purpose of this research is to study the effects of using different CI expansions and orbital basis sets on the $^2B_1$-$^2A_2$ energy splitting. What we would like to find is a reasonably inexpensive calculation procedure that predicts a very small splitting.

A. Theoretical.

In our study of the cyclopropenyl radical we employed the standard Dunning and Huzinaga contracted Gaussian double-zeta (DZ) basis set[93,94] whose precise designation is C(9s6p/4s2p) and H(4s/2s). All calculations were performed at a single D_{3h} geometry with all C-C bond distances fixed at 1.40 Å and all C-H bond lengths set at 1.08 Å. The exact cartesian geometry used was (0.0,0.0,0.0), (0.0,±1.322826, -2.291202) for the three carbon atoms and (0.0,0.0,2.040932), (0.0, ±3.090325, -3.311668) for the three hydrogen atoms. Since the geometry was not optimized in this study and our primary purpose was to determine the $^2B_1$-$^2A_2$ energy splitting at this D_{3h} geometry we deemed a DZ basis set to be adequate. Also the geometry selected does not correspond to the minimum on the D_{3h} surface for this basis set, but was picked because it is in the range of previous theoretical work.



To determine how the $^2B_1$-$^2A_2$ splitting varies as the level of theory is changed a series of calculations were carried out on both the lowest $^2B_1$ and $^2A_2$ states of C_3H_3. All of these calculations were done in C_{2v} symmetry at the D_{3h} geometry. Initially one reference SCF calculations were performed on these two states with the $^2B_1$ and $^2A_2$ reference configurations being

$$\Phi_1(^2B_1) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1^2\ 6a_1^2\ 2b_1$$

$$\Phi_1(^2A_2) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1^2\ 6a_1^2\ 1a_2 \quad .$$

Beyond the SCF level a number of CI and MCSCF calculations were obtained. First a set of three single and double excitation CI wavefunctions were computed for each state. In the largest, involving 11829 and 11800 configurations for the $^2B_1$ and $^2A_2$ states respectively, all orbitals were allowed in the CI. In the second the 18 orbitals with the highest orbital energies were forced to have zero occupancy in the CI (in other words, one half of the orbitals of each symmetry type were frozen) reducing the number of configurations to 1069 for either state. In the third CI the three lowest orbitals (in an energetic sense) were also frozen resulting in 549 configurations in both the $^2B_1$ and $^2A_2$ cases. These three calculations will be abbreviated as CIALL, CIMIN and CI549. In this last CI the orbitals are partitioned such that there are 18 of the virtual, 3 of the core and 15 of the active varieties. This limits excitations to being within the valence space. The concept of the valence space is easy

to visualize if one envisions the C-C and C-H bond lengths being stretched to infinity. When this happens the set of 15 active orbitals reduces to the 2s and 2p shells on the carbon atoms and the 1s orbitals of the hydrogens.

Besides these three CIs a number of additional calcuations were performed within the valence space. The first of these is an MCSCF computation within the same 549 configuration space defined above and is called MC549. To investigate the effect of higher then double excitations in this subspace on the $^2B_1$-$^2A_2$ splitting, three further calculations were carried out. The first was a straight CI using SCF orbitals whose configuration set is composed of the reference plus all single, double and triple excitations within the valence space. Following the standard CI an MCSCF calculation employing the same configuration list was carried out to ascertain the effects of optimizing the orbitals. These two calculations will be labeled CISDT and MCSDT. Lastly a single, double, triple and quadruple excitation CI using the SCF orbitals was performed and will be identified as CISDTQ. Originally MCSCFs containing the reference plus valence space single excitations were to be computed for the $^2B_1$ and $^2A_2$ states. However, extreme convergence difficulties were encountered and these two calculations were abandoned.

B. Results and Discussion.

In Table IX the total energies for the calculations specified in the preceeding section are presented. First, the SCF results are examined. For the $^2A_2$ state a total energy of -115.104667 hartrees was obtained and for the corresponding $^2B_1$ state a lower energy of

-115.109144 hartrees was found. This leads to a splitting of 0.004477 hartrees or 2.81 kcal/mole. This difference is of the same order as the minimum basis SCF results of Davidson and Borden[124] in which a value of around 5 kcal/mole was reported. Poppinger, Radom and Vincent[125] (PRV) also published SCF $^2B_1$-$^2A_2$ splittings for both the minimum and the split valence 4-31G basis sets for which they obtained 4.79 and 3.23 kcal/mole respectively.

Moving on to the straight CI results the energies of the $^2B_1$ and $^2A_2$ states are found to be -115.389963 and -115.387954 hartrees respectively in the all orbital single and double excitation CI approximation. These calculations which involve nearly 12,000 configurations reduce the splitting to 0.002009 hartrees or 1.26 kcal/mole. This energy difference is less than half the SCF value. If the 18 highest reference unoccupied orbitals are now frozen in the single and doubles CI the corresponding $^2B_1$ and $^2A_2$ energies are -115.165587 and -115.167098 hartrees. This leads to a splitting of -0.001512 hartrees or -0.95 kcal/mole, where the minus sign indicates that the ordering of the two states has changed. This state reversal is most likely due to the choice of the frozen orbitals in the CI. If all the reference unoccupied orbitals are kept in the CI the energy is invariant to unitary transformations among this orbital subset. However, if any of these orbitals are frozen the energy is no longer invariant. From the SCF point of view, however, the unoccupied orbitals are only required to be orthogonal to the occupied space and amongst themselves and are thus rather arbitrary. What probably happened was that by chance the set of $^2A_2$ unoccupied orbitals that were retained

in the CI were energetically more favorable than the corresponding $^2B_2$ orbitals. If in addition to the 18 virtual orbitals 3 core orbitals are frozen the single and doubles CI energies are found to be -115.165484 and -115.166984 hartrees for the $^2B_1$ and $^2A_2$ states. Here one notices that the ordering of the states is again reversed with the splitting being -0.001510 hartrees or -0.94 kcal/mole.

Proceeding onto the first of our MCSCF results the $^2B_1$-$^2A_2$ energy gap is found to be +0.001336 hartrees for the valence space singles and doubles MCSCF (MC549). This 0.84 kcal/mole difference is roughly 2/3 of the analogous CIALL splitting (1.26 kcal/mole), but the CI contains only about 1/20 of the number of configurations. This result is encouraging since MCSCF calculations of this variety can be performed on nearly all systems that can be handled using SCF theory alone. In terms of absolute energies, -115.290106 and -115.288770 hartrees were obtained for the $^2B_1$ and $^2A_2$ states respectively. These results are surprising in that these small 549 configurations MCSCF calculations recovered around 65% of the full CI singles plus doubles correlation energy while the CIs using SCF orbitals in the same valence space recovered about 20%. Optimizing the orbitals then leads to roughly 3 times as much correlation energy and the correct prediction of the state ordering.

Three additional sets of calculations were carried out within the valence space to test the effect of inclusion of higher than double excitations. The first two of these employed a configuration set consisting of all valence single, double and triple excitations. The total energies for the straight CISDT computations are -115.166563

Table IX. Summary of Energetic Results for the Cyclopenyl Radical

Calculation Type	A_2 Symmetry		B_1 Symmetry		Energy Difference $E_{B_2}-E_{A_2}$	
	No. of Configurations	Total Energy (hartrees)	No. of Configurations	Total Energy (hartrees)	hartrees	kcal/mole
SCF	1	-115.104667	1	-115.109144	+0.004477	2.81
CI-all orbitals-SD[a]	11800	-115.387954	11829	-115.389963	+0.002009	1.26
CI-all orbitals-SD plus Davidson's Correction	11800	-115.417994	11829	-115.418878	+0.000884	0.55
CI-18 virtual-SD	1069	-115.167098	1069	-115.165587	-0.001512	-0.95
CI-18 virtual-SD 3 core	549	-115.166984	549	-115.165484	-0.001510	-0.94
MCSCF-18 virtual-SD 3 core	549	-115.288770	549	-115.290106	+0.001336	0.84
CI-18 virtual-SDT 3 core	7470	-115.168311	7474	-115.166563	-0.001748	-1.10
MCSCF-18 virtual-SDT 3 core	7470	-115.2914	7474	-115.29252	+0.0011	∿0.70
CI-18 virtual-SDTQ 3 core	60645	-115.170483	60685	-115.168276	-0.002207	-1.38

[a]Symbols S, D, T and Q represent all single, double, triple and quadruple configurations respectively.

and -115.168311 hartrees for the $^2B_1$ and $^2A_1$ states respectively which leads to a splitting of -0.001748 hartrees or -1.10 kcal/mole. The ordering of the two states is again inverted as in the CIVIR and CI549 cases with the gap now larger by roughly 0.15 hartrees. This result is at first glance surprising since the CISDT calculation is much more extensive than either of the CIVIR or CI549 calculations. On closer examination, however, since the $^2A_2$ state is above the $^2B_1$ state in the SCF approximation one would expect that as the size of the CI increases so will the ratio of the $^2A_2$ and $^2B_1$ correlation energies.

The second type of calculation performed using the all valence single, double and triple configuration set was MCSDT. Certain difficulties arose in converging these two wavefunctions expecially for the $^2A_2$ state. Because of this the energetic results in Table IX are only reported to 4 or 5 decimal places for the $^2A_2$ and $^2B_1$ states respectively. These results are believed to be accurate to the number of digits listed, however, additional work is being done to make absolutely certain. As in the case of MC549 the $^2B_1$ is below the $^2A_2$ state and for MCSDT a splitting of 0.0011 hartrees or around 0.70 kcal/mole is found. This difference is the smallest obtained employing variational energies and at least 0.14 kcal/mole less than the MC549 result.

We lastly carried straight CI calculations employing all valence single, double, triple and quadruple excitations for both the two states of cyclopropenyl radical. As in the other three CIs with 18 frozen SCF virtual orbitals, CISDTQ predicts the $^2A_2$ to be below

the $^2B_1$ state. This time an energy splitting of -0.002207 hartrees or -1.38 kcal/mole is found. What the four calculations: CIVIR, CI549, CISDT and CISDTQ show us is that if the CI expansion is restricted to the minimum basis or valence space it is necessary to optimize the orbitals if the $^2A_2$-$^2B_1$ energy ordering is to be correctly predicted even if a very large number of configurations are included.

In addition to the variational results if CIALL is corrected for the effects of unlinked clusters using the Davidson's correction a splitting of only 0.55 kcal/mole is obtained. This result is less than 1/2 of the straight CIALL difference and maintains the "correct" ordering of the two states. This seems to suggest the validity of employing this correction to all singles and doubles CI energies.

Before closing this section certain features of the CI and MCSCF wave functions should be discussed. It was in the course of performing these calculations that we discovered that even though cyclopropenyl radical is adequately described by one reference, two additional configuration in both the $^2A_2$ and $^2B_1$ are moderately important also. For the $^2A_2$ state these configurations are

$$\Phi_2(^2A_2) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1\alpha\ 6a_1^2\ 2b_1\alpha\ 1a_2\beta$$

$$\Phi_3(^2A_2) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1^0\ 6a_1^2\ 2b_1^2\ 1a_2$$

and for the $^2B_1$ state they have

$$\Phi_2(^2B_1) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1^0\ 6a_1^2\ 2b_1\ 1a_2^2$$

$$\Phi_3(^2B_1) = \ldots\ 5a_1^2\ 3b_2^2\ 1b_1\ 6a_1^2\ 1a_2^2 \quad .$$

The values of the coefficients of these configurations plus the SCF references are compiled in Table X.

In all the CI and MCSCF calculations the principle coefficients were roughly 0.95 with the values of the other two coefficients being between about 0.06 and 0.15. In the four calculations that contained frozen SCF virtual orbitals the values of C_2 are much higher in both states (around 0.15) than in the other calculations. Also C_3 for the $^2B_1$ is much larger roughly 0.10 compared to the MCSCF values of around 0.08. The MCSDT coefficients appears to be quite similar to the MC549 values, though the importance of single excitations in the former is far greater. Since there is some question at this time as to the validity of our MCSDT results due to certain convergence difficulties the wave functions may in fact be somewhat different. In particular the last digit reported for the coefficients of MCSDT in Table X may be suspect. Currently more work is being done in an attempt to converge these MCSCF wave functions to a greater degree.

C. Conclusions for the Cyclopropenyl Radical

For the cyclopropenyl radical it was found that at a D_{3h} geometry an artificial energy splitting of nearly 3 kcal/mole exists for the truly degenerate $^2A_2$ and $^2B_1$ states at the SCF level of theory if the calculations are carried out in C_{2v} symmetry. This energy gap can be substantially lessened if subsequent MCSCF or CI calculations are performed. In the case of all single and double

Table X. Tabulated Coefficients for the Three Most Important $^2B_1$ and $^2A_2$ References in MC549.

		SCF	CIALL	CIVIR	CI549	MC549	CISDT	MCSDT	CISDTQ
	C_1	1.0000	0.9454	0.9669	0.9669	0.9495	0.9651	0.947	0.962
A_2	C_2		0.0869	-0.1523	-0.1524	-0.1241	-0.1557	-0.128	-0.164
	C_3		-0.0703	-0.1046	-0.1047	-0.0994	-0.1055	- .100	-0.109
	C_1	1.0000	0.9471	0.9729	0.9729	0.9530	0.9716	0.951	0.969
B_1	C_2		0.0796	- .1200	-0.1201	-0.1114	-0.1207	-0.112	-0.125
	C_3		0.0581	- .1007	-0.1007	0.0778	0.1041	0.083	0.110

excitation CIs the splitting is reduced by over 50%. If these two energies are corrected for unlinked cluster effects the $^2A_2$-$^2B_1$ difference is only 0.55 kcal/mole.

Good results are also obtained for valence space MCSCF calculations. In the MCSCF employing all valence single and double excitations the splitting is roughly 2/3 of the full CISD result and only about 1/20 of the number of configurations are needed. If triple excitations are added the gap is reduced still more to around 0.70 kcal/mole. However, if the orbitals are not optimized in the valence space CIs poor agreement is obtained and the incorrect $^2A_2$-$^2B_1$ ordering is found.

As out outgrowth of this work a number of additional calculations will be carried out on these two states of the cyclopropenyl radical. The first of these consists of a small MCSCF followed by a CI containing all single and double excitations from each reference. This small MCSCF will be made up of a full CI among the lowest a_2 and the two lowest b_1 orbitals for both states. Also an MCSCF including all valence single and double excitations from the 3 important references will be obtained. In addition MCSCFs in both the full single and double and the all valence single, double, triple and quadruple excitation spaces will be computed. These results plus the present calculations will be presented in a future publication.

VI. Concluding Remarks

In this dissertation it was shown that the loop-driven graphical unitary group approach can be extended to include the ability to determine large MCSCF wavefunctions. The largest wavefunction computed using this method contained 17,678 configurations and is the largest MCSCF calculation performed anywhere[33] to date. In contrast to other MCSCF procedures which can handle large configuration sets the orbital optimization step does not assume any special form for the configuration list (such as in the CASSCF method where the configuration set is required to be a full CI within an orbital subspace.)[35-37] In fact the LDGUGA MCSCF method can employ very flexible configuration lists including all single and double excitations from one or two references, full CI in any subspace and inclusion of classes of higher than double excitations.

At the heart of this MCSCF formalism is the efficient computation and storage of the two-particle density matrix. This feature coupled with the loop-driven CI method[38,39] developed by Bernie Brooks leads to a very fast and versatile MCSCF procedure. Though at present our program is restricted to only first-order convergence various extensions such as the construction and implementation of the orbital Hessian matrix and a proposed extrapolation procedure were discussed herein.

To demonstrate the power of the LDGUGA method a number of molecular calculations were presented. These included a series of CI and MCSCF results for water,[33] ozone[34] and the cyclopropenyl radical. A number of these computations explored the effect of higher

excitations in MCSCF calculations, an area that previously had not been extensively investigated.

Appendix I. A_k and B_k Theory

The A_k and B_k methods are examples of truncated CI schemes and were initially presented by Gershgorn and Shavitt.[104] The methods are based on partitioning perturbation theory[128] and have been implemented and extended by a number of authors.[105-107] In practice A_k and B_k techniques allow the use of much larger configuration sets than can be handled in a non-truncated CI and the hope is that relatively little energy will be sacrificed.

In developing the theory it is useful to envision that the CI Hamiltonian H is partitioned as follows:

$$\left[\begin{array}{c|c} H^0 & h^T \\ \hline h & H' \end{array}\right] \begin{array}{c} C_0 \\ \hline C' \end{array} = E \begin{array}{c} C_0 \\ \hline C' \end{array} \quad . \tag{87}$$

In constructing this Hamiltonian the configuration space (composed of N configurations) has been divided into two subspaces P and Q. P denotes the primary space and usually includes all the important configurations as determined through some selection process. Q, the secondary space, is composed of the remainder of the total configuration space. $\underset{\sim}{H}^0$, $\underset{\sim}{h}$ and $\underset{\sim}{H}'$ are then PxP, PxQ and QxQ sub-matrices respectively. If (87) is multiplied out the following two equations are found

$$\underset{\sim}{H}^0 \underset{\sim}{c}_0 + \underset{\sim}{h}^T \underset{\sim}{c}' = E\underset{\sim}{c}_0 \tag{88}$$

$$\underset{\sim}{h}\underset{\sim}{c}_0 + \underset{\sim}{H}'\underset{\sim}{c}' = E\underset{\sim}{c}' \quad . \tag{89}$$

Solving (89) for c' and substituting into (88) yields

$$[\underset{\sim}{H}^0 + \underset{\sim}{h}^T(E\underset{\sim}{1}-\underset{\sim}{H}')^{-1}\underset{\sim}{h}]\underset{\sim}{c}_0 = \underset{\sim}{H}_{eff}\,\underset{\sim}{c}_0 = E\,\underset{\sim}{c}_0 \quad . \tag{90}$$

This reduces the full NxN CI equations to the solution of a PxP problem using $\underset{\sim}{H}_{eff}$. However, calculating $\underset{\sim}{H}_{eff}$ is at least as much work as solving (87) directly so nothing is gained. The time consuming aspect in constructing H_{eff} involves finding the inverse of $(E\underset{\sim}{1}-\underset{\sim}{H}')$. What one wants to use then is an approximation to this inverse that is relatively simple to obtain and that still gives the majority of the NxN CI energy.

Since $\underset{\sim}{H}'$, the QxQ Hamiltonian, is diagonally dominate a reasonable approximation would be that

$$(E\underset{\sim}{1}-\underset{\sim}{H}')^{-1} \approx (E\underset{\sim}{1}-\underset{\sim}{D})^{-1} \tag{91}$$

with

$$D_{ij} = 0, \text{ if } i \neq j \qquad i,j \in \Gamma$$

$$D_{ii} = H_{ii} \quad .$$

Solving equation (90) using the approximation (91) is called the B_k method. The most straight-forward way to solve for E is to diagonalize $\underset{\sim}{H}$ in equation (87) with $\underset{\sim}{H}'$ set zero to D or in other words to set all off-diagonal elements of $\underset{\sim}{H}'$ to zero. The neglect of these elements makes this a second-order perturbative scheme since only the matrix elements that directly interact with $\underset{\sim}{H}_0$ are kept.

The A_k method is very similar to the B_k except that H_0 is diagonalized first to yield

$$\psi_0 = \sum_{n=1}^{k} c_n \Phi_n \tag{92}$$

where Φ_n is configuration n and the sum is over all elements of P. The subscript k in both the A_k and B_k methods refers to the k configurations in the primary space. Next straight first-order Rayleigh-Schrödinger perturbation theory is applied to (92) yielding

$$\psi_1 = \Psi_0 + \sum_{n=k+1}^{N} \frac{\langle \Phi_n | H | \Psi_0 \rangle}{E_0 - H_{nn}} \Phi_n = \Psi_0 + \sum_{n=n+1}^{N} c_n \Phi_n \tag{93}$$

The second-order energy associated with (93) is

$$E_2 = \sum_{n=k+1}^{N} \frac{|\langle \Phi_k | H | \Psi_0 \rangle|^2}{E_0 - H_{nn}} = \sum_{i,j=1}^{k} c_i c_j \left(\sum_{n=k+1}^{N} \frac{H_{in} H_{nj}}{E_0 - H_{nn}} \right) . \tag{94}$$

The major difference between the second-order A_k and B_k results is that B_k is usually lower in energy since the primary space coefficients are allowed in relax while in the A_k case they are fixed. Since constructing the off-diagonal elements of $\underset{\sim}{H}'$ is the major time step in the full NxN CI the A_k and B_k calculations are much faster if the P space is much less than the Q space as is normally the situation.

In going to the third-order in energy the off-diagonal elements of H' must now be calculated. In both cases A_k and B_k, the correction formula is

$$E_3 = \sum_{m,n=k+1}^{N} C_m C_n H_{mn} \quad .^{110} \tag{95}$$

The B_k coefficients in the Q space (C_m, $m \in Q$) are obtained from the diagonalization of the approximate B_k Hamiltonian and in the A_k scheme are calculated while forming E_2 using equation (94). E_3 for the two methods will be different since their two sets of coefficients generally differ.

In both methods as presented the Epstein-Nesbet (EN) partitioning of the Hamiltonian was used. This has advantages in that it is well defined for open-shell and multi-reference states as well as for closed shells. Also it is easier to adapt to existing CI codes. The disadvantage with respect to Möller-Plesset (MP) partitioning is that the magnitude of E_3 is usually much larger in EN than in the MP scheme. This makes calculating the third order correction more important than if we had used the other partitioning.

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. R. Soc. London Ser. A 238, 229 (1939).
3. G. Das and A. C. Wahl, J. Chem. Phys. 44, 87 (1966).
4. E. Clementi and A. Veillard, J. Chem. Phys. 44, 3050 (1966).
5. B. Levy and G. Berthier, Int. J. Quantum Chem. 2, 307 (1968).
6. N. Sabelli and J. Hinze, J. Chem. Phys. 50, 684 (1969).
7. G. Das and A. C. Wahl, J. Chem. Phys. 47, 2934 (1967).
8. K. J. Miller and K. Ruedenberg, J. Chem. Phys. 48, 3414 (1968).
9. W. A. Goddard, Phys. Rev. 157, 81 (1967).
10. W. J. Hunt, P. J. Hay, and W. A. Goddard, J. Chem. Phys. 57, 738 (1972); F. W. Brobrowicz and W. A. Goddard, Modern Theoretical Chemistry, edited by H. F. Schaefer (Plenum, New York, 1977), Vol. 3, pp. 79-127.
11. J. Hinze, J. Chem. Phys. 59, 6424 (1973).
12. A. C. Wahl and G. Das, Modern Theoretical Chemistry, edited by H. F. Schaefer (Plenum, New York, 1977), Vol. 3, pp. 51-78; A. C. Wahl and G. Das, Adv. Quantum Chem. 5, 261 (1970).
13. A. Banerjee and F. Grein, J. Chem. Phys. 66, 1054 (1977).
14. J. O. Arnold, E. E. Whiting, and S. R. Langhoff, J. Chem. Phys. 66, 4459 (1977).
15. L. M. Chueng, K. R. Sundberg, and K. Ruedenberg, J. Amer. Chem. Soc. 100, 8024 (1978).

16. R. R. Lucchese and D. R. Yarkony, J. Chem. Phys. 68, 2696 (1978).

17. M. Dupuis and B. Liu (unpublished research, 1977).

18. J. O. Arnold, E. E. Whiting, and S. R. Langhoff, J. Chem. Phys. 66, 4459 (1977).

19. E. Dalgaard and P. Jorgensen, J. Chem. Phys. 69, 3833 (1978).

20. D. L. Yeager and P. Jorgensen, J. Chem. Phys. 71, 755 (1979).

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

22. B. H. Lengsfield III, J. Chem. Phys. 73, 382 (1980).

23. B. H. Lengsfield III, (to be published).

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

25. C. C. J. Roothaan, Recent Developments and Applications of Multi-Configuration Hartree-Fock Methods, NRCC proceeds No. 10, Edited by M. Dupuis (1981).

26. R. Shepard and J. Simons, J. Chem. Phys. (in press).

27. R. Shepard and J. Simons, Int. J. Quantum Chem. Symp. 14, 211 (1980).

28. H. J. Werner and W. Meyer, J. Chem. Phys. 73, 2342 (1980).

29. D. R. Yarkony, Recent Developments and Applications of Multi-Configuration Hartree-Fock Methods, NRCC proceedings No. 10, Edited by M. Dupuis (1980).

30. 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 Electronic Energy Matrix Elements, edited by Jürgen Hinze (Springer, Berlin 1980).

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

32. W. D. Laidig and H. F. Schaefer, Recent Developments and Applications of Multi-Configuration Hartree-Fock Methods, NRCC proceedings No. 10, Edited by M. Dupuis (1981).

33. W. D. Laidig, P. Saxe, and H. F. Schaefer, J. Chem. Phys. 73, 1765 (1980).

34. W. D. Laidig and H. F. Schaefer, J. Chem. Phys. 74, 3411 (1981).

35. P. Siegbahn, A. Heiberg, B. Roos and B. Levy, Phys. Scr. 21, 323 (1980).

36. B. O. Roos, P. R. Taylor and P. E. M. Siegbahn, Chem. Phys. 48, 157 (1980).

37. P. E. M. Siegbahn, J. Almlof, A. Heiberg and B. O. Roos, J. Chem. Phys. 74, 2384 (1981).

38. B. R. Brooks and H. F. Schaefer, J. Chem. Phys. 70, 5092 (1979).

39. B. R. Brooks, W. D. Laidig, P. Saxe, N. C. Handy and H. F. Schaefer, Phys. Scr. 21, 312 (1980).

40. J. Paldus, in Theoretical Chemistry: Advances and Perspectives, Vol. 2, p. 131, edited by H. Eyring and D. Henderson (Academic Press, New York, 1976).

41. J. Paldus and M. J. Boyle, Phys. Scr. 21, 295 (1980).

42. I. Shavitt and J. Paldus (to be published).

43. I. Shavitt, Int. J. Quantum Chem. Symp. 11, 131 (1977).

44. I. Shavitt, Int. J. Quantum Chem. Symp. 12, 5 (1978).

45. I. Shavitt, in Proceedings of the Conference on the Unitary Group for the Evaluation of Electronic Energy Matrix Elements,

edited by Jürgen Hinze (Springer, Berlin, 1980).

46. P. Saxe and H. F. Schaefer, J. Amer. Chem. Soc. 102, 3239 (1980).

47. W. D. Laidig and H. F. Schaefer, J. Chem. Phys. 73, 1470 (1980).

48. Y. Yamaguchi and H. F. Schaefer, J. Chem. Phys. 73, 2310 (1980).

49. D. J. Fox, D. Ray, P. C. Rubesin and H. F. Schaefer, J. Chem. Phys. 73, 3246 (1980).

50. J. Demuynck, D. J. Fox, Y. Yamaguchi and H. F. Schaefer, J. Amer. Chem. Soc. 102, 6204 (1980).

51. B. R. Brooks, Ph.D. Thesis, Department of Chemistry, University of California, Berkeley, September 1979.

52. P. Saxe, D. J. Fox, N. C. Handy, H. F. Schaefer (to be published).

53. A. Bunge, J. Chem. Phys. 53, 20 (1970).

54. C. F. Bender and H. F. Schaefer, J. Chem. Phys. 55, 4798 (1971).

55. R. R. Lucchese and H. F. Schaefer, J. Amer. Chem. Soc. 99, 6766 (1977).

56. I. M. Gelfand and M. L. Zetlin, Dokl. Adad. Nauk. SSSR 71, 825, 1017 (1950).

57. Since Brooks Ph.D. thesis was published the implicit one and two external cases have been coded by Douglas Fox.

58. P. E. M. Siegbahn, J. Chem. Phys. 70, 5391 (1979).

59. P. E. M. Siegbahn, J. Chem. Phys. 72, 1647 (1980).

60. P. E. M. Siegbahn, in Proceedings of the Conference on the Unitary Group for the Evaluation of Electronic Energy Matrix Elements, edited by Jürgen Hinze (Springer, Berlin 1980).

61. R. M. Pitzer, J. Chem. Phys. 58, 3111 (1973).

62. M. Dupuis, J. Rys, and H. F. King, J. Chem. Phys. 65, 111 (1976).

63. E. R. Davidson, J. Computational Phys. 17, 87 (1975).

64. J. A. Pople, R. Krishnan, H. B. Schlegel and J. S. Binkley, Int. J. Quantum Chem 14, 545 (1978).

65. B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. 72, 4652 (1980).

66. H. F. Schaefer, The Electronic Structure of Atoms and Molecules: A Survey of Rigorous Quantum Mechanical Results (Addison-Wesley, Reading, Massachusetts, 1972).

67. P. S. Bagus, B. Liu, A. D. McLean and M. Yoshimine, Wave Mechanics: The First Fifty Years, edited by W. C. Price, S. S. Chissick, and T. Ravensdale (Butterworth, London, 1973), pp. 99-118.

68. I. Shavitt, Modern Theoretical Chemistry, edited by H. F. Schaefer, (Plenum, New York, 1977), Vol. 3, pp. 189-275.

69. F. Grimaldi, A. Lecourt, and C. Moser, Intern. J. Quantum Chem. Symp. 1, 153 (1976).

70. K. A. Brueckner, Phys. Rev. 100, 36 (1955).

71. O. Sinanoglu, J. Chem. Phys. 36, 706 (1962).

72. J. Cizek, Adv. Chem. Phys. 14, 35 (1969).

73. R. J. Bartlett and I. Shavitt, Chem. Phys. Lett. 50, 190 (1977).

74. J. Goldstone, Proc. R. Soc. London Ser. A 239, 267 (1957).

75. R. J. Bartlett, I. Shavitt, and G. D. Purvis, J. Chem. Phys. 71, 281 (1979).

76. A. C. Hurley, Electron Correlation in Small Molecules, (Academic, London, 1976).

77. E. R. Davidson, of The World of Quantum Chemistry, edited by R.

Daudel and B. Pullman (Reidel, Dordecht, Holland, 1974, pp. 17ff.

78. W. Meyer, Modern Theoretical Chemistry, edited by H. F. Schaefer (Plenum, New York, 1977), Vol. 3, pp. 413-446.

79. R. Aldrichs, H. Lischka, V. Staemmler, and W. Kutzelnigg, J. Chem. Phys. 62, 1225 (1975).

80. E. R. Davidson and D. W. Silver, Chem. Phys. Lett. 52, 403 (1977).

81. J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Intern. J. Quantum Chem. 14, 545 (1978).

82. P. E. M. Siegbahn, Chem. Phys. Lett. 55, 386 (1978).

83. C. E. Dykstra and H. F. Schaefer, The Chemistry of Ketenes and Allenes, edited by S. Patai (Wiley/Interscience, London, 1979), pp. 1-44.

84. S. Wilson and V. R. Saunders, J. Phys. B 12, L403 (1979).

85. S. Wilson and V. R. Saunders, J. Phys. B 13, 1505 (1980).

86. S. Wilson, J. Phys. B 12, L657 (1979).

87. S. Wilson, J. Phys. B 13, 505 (1980).

88. A. Pipano and I. Shavitt, Intern. J. Quantum Chem. 2, 741 (1968).

89. J. Paldus, J. Cizek, and I. Shavitt, Phys. Rev. A. 5, 50 (1972).

90. R. P. Hostenry, R. R. Gilman, T. H. Dunning, A. Pipano, and I. Shavitt, Chem. Phys. Lett. 7, 325 (1970).

91. I. Shavitt, unpublished work of BeH_2 and C_2H_4, 1971.

92. C. W. Bauschlicher and I. Shavitt, J. Amer. Chem. Soc. 100, 739 (1978).

93. S. Huzinaga, J. Chem. Phys. 42, 1293 (1965).

94. T. H. Dunning, J. Chem. Phys. 53, 2823 (1970).

95. A. Pipano, R. R. Bilman and I. Shavitt, Chem. Phys. Lett. 5, 285

(1970).

96. P. Saxe, H. F. Schaefer and N. C. Handy, Chem. Phys. Lett. 79, 202 (1981).

97. L. Brillouin, Act. Sci. Ind. No. 159 (Herrmann, Paris, 1934).

98. R. J. Bartlett, private communication.

99. S. Wilson and M. F. Guest, Chem. Phys. Lett. 73, 607 (1980).

100. See, for example, R. J. Bartlett, Annual Review of Physical Chemistry 32, XXXX (1981).

101. J. A. Pople, R. Seeger and R. Krishnan, IJQC S11 (1977).

102. This is assuming that the number of electrons is roughly the same as the number of basis functions.

103. H. P. Kelly, Phys. Rev. 136B, 896 (1964).

104. Z. Gershgorn and I. Shavitt, Intern. J. Quantum Chem. 2, 751 (1968).

105. G. A. Segal, R. W. Wetmore, and K. Wolf, Chem. Phys. 30, 269 (1978).

106. L. E. Nitzsche and E. R. Davidson, J. Chem. Phys. 68, 3103 (1978).

107. T. H. Dunning, Chem. Phys. 42, 249 (1979).

108. P. S. Epstein, Phys. Rev. 28, 695 (1926).

109. R. K. Nesbet, Proc. Roy. Soc. A230, 312 (1955).

110. G. B. Backsay, Chem. Phys. 26, 47 (1977).

111. See, for example, C. F. Bender and H. F. Schaefer, J. Chem. Phys. 55, 1720 (1971).

112. W. A. Goddard, T. H. Dunning, W. J. Hunt, P. J. Hay, Accts. Chem. Res. 6, 368 (1973).

113. P. J. Hay, T. H. Dunning and W. A. Goddard, J. Chem. Phys. 62, 3912

(1975).

114. G. Herzberg, Electronic Spectra of Polyatomic Molecules, (Van Nostrand, Princeton, New York, 1967).

115. E. F. Hayes and A. K. Q. Siu, J. Amer. Chem. Soc. 93, 2090 (1971).

116. P. J. Hay and T. H. Dunning, J. Chem. Phys. 67, 2290 (1977).

117. M. P. Conrad, R. M. Pitzer, and H. F. Schaefer, J. Amer. Chem. Soc. 101, 2246 (1979).

118. L. C. Snyder, J. Phys. Chem. 66, 2299 (1962).

119. M. Shansal, Z. Naturforsch. 26A, 1336 (1971).

120. T. K. Ha, F. Graf and H. H. Günthard, J. Mol. Struct. 15, 335 (1973).

121. G. Cirelli, F. Graf and H. H. Günthard, Chem. Phys. Lett. 28, 494 (1974).

122. N. C. Baird, J. Org. Chem. 40, 624 (1975).

123. J. Pancir and R. Zahradnik, Tetrahedron 32, 2257 (1976).

124. E. R. Davidson and W. T. Borden, J. Chem. Phys. 67, 2191 (1977).

125. D. Poppinger, L. Radom and M. A. Vincent, Chem. Phys. 23, 437 (1977).

126. H. A. Jahn and E. Teller, Proc. R. Soc. London Ser. A 161, 220 (1937).

127. R. Manne, Mol. Phys. 24, 935 (1972).

128. P.-O. Löwdin, Perturbation Theory and its Applications in Quantum Mechanics, edited by C. H. Wilcox (Wiley, New York, 1966).

This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.

TECHNICAL INFORMATION DEPARTMENT
LAWRENCE BERKELEY LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720