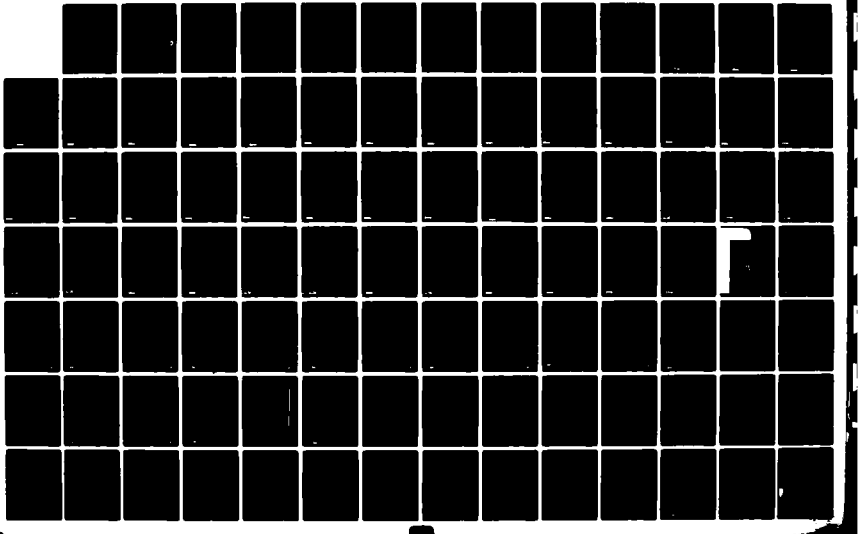
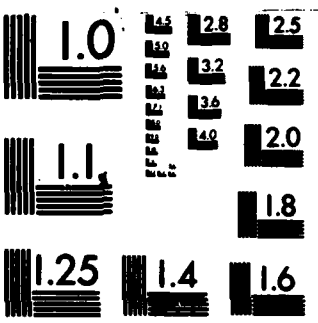
AD-A140 738 PREDICTION OF TRANSITION STATES AND THERMOCHEMISTRY FOR 1/2
COMBUSTION REACTIONS(U) BATTELLE COLUMBUS LABS OH
R J BARTLETT 07 APR 84 ARO-16929.4-PH DAAG29-80-C-0105
UNCLASSIFIED F/G 20/10 NL





Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: ARO 16929.4-PH ; 18936.4-PH	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): PREDICTION OF TRANSITION STATES AND THERMO-CHEMISTRY FOR COMBUSTION REACTIONS	5. TYPE OF REPORT & PERIOD COVERED: FINAL 8/4/80 - 12/31/83
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Rodney J. Bartlett	8. CONTRACT OR GRANT NUMBER(s): DAAG29-80-C-0105 DAAG29-82-K-0034
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Battelle, Columbus Laboratories, Columbus, OH 43201 Quantum Theory Project, Univ. of Florida, Gainesville, Florida 32611	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, Post Office Box 12211, Research Triangle Park, NC 27709	12. REPORT DATE: 4/7/84
	13. NUMBER OF PAGES: 8
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Donald Calder, Georgia Institute of Technology, Office of Naval Research Resident Representative, 206 O'Keefe Bldg. Atlanta, GA 38332	15. SECURITY CLASS. (of this report): unclassified Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE MAY 0 3 1984 S D E

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Many-body perturbation theory(MBPT), coupled-cluster(CC) theory, transient molecules, dissociation energies, transition states, activation barriers, kinetics models, ab initio quantum chemistry, molecular structure, molecular spectra, linked-diagram theorem, combustion

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Ab initio quantum chemistry methods based upon many-body perturbation theory(MBPT) and coupled-cluster(CC) theory have been developed and applied to potential energy surfaces for transient molecules. Studies included decomposition pathways for formaldehyde, methanol, the formyl radical and the inorganic radical HIF. MBPT/CC methods were extended to include effects of triple-excitations and comparisons made with full CI for HF, BH, NH_3 and H_2O. The high accuracy of MBPT/CC has been demonstrated.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

FINAL REPORT

Contract Nos. DAAG29-80-C-~~0150~~ 0105
DAAG29-82-K-0034

PREDICTION OF TRANSITION STATES AND THERMOCHEMISTRY FOR COMBUSTION REACTIONS

to

U.S. ARMY RESEARCH OFFICE

April 7, 1984

Principal Investigator:

Rodney J. Bartlett
Professor of Chemistry and Physics
Quantum Theory Project
362 Williamson Hall
University of Florida
Gainesville, Florida 32611

I. INTRODUCTION:

The Army Research Office, under a three-year contract, DAAG29-80-C-0105 began to support our research at Battelle Columbus Laboratories in the fall of 1980. I left Battelle to become professor of Chemistry and Physics at the University of Florida, Quantum Theory Project in August, 1981. Consequently, I submitted a new proposal for the second and third year of the existing contract from the University of Florida. This contract was funded under the number DAAG29-82-K-0034 beginning in January, 1982. Hence, this report covers the three-year period of the initial contract to Battelle and the subsequent two years at Florida. As a part of both of these contracts have supported work involving a strong collaboration between myself and Dr. George Adams of the Army's Ballistic Research Laboratory, Aberdeen Proving Ground, his technical reports may also be consulted for voluminous information pertaining to the ab initio many-body perturbation theory(MBPT) study of flame species. In particular, technical report ARBRL-TR-02240 by Adams and Bent, entitled "Ab Initio Prediction of Thermochemical Parameters for Flame Species," describes our initial MBPT predictions for formaldehyde, H_2CO, the formyl radical, HCO, methanol, H_3COH, and the methoxy radical, H_3CO, at length, including a discussion of our many-body methods.

In section II, I will describe the objectives of our research for ARO, and summarize our progress. A list of papers and presentations supported by ARO is presented in the third section. Our new ARO contract, DAAG29-84-K-0025, will permit us to amplify these earlier advances and use our many-body methods to develop new techniques for the treatment of large molecules. Personnel supported by ARO (9/80-12/83) are listed in section IV. Copies of papers supported by ARO are included in five copies of this report, for more detailed information about the research.

II. SUMMARY OF PROGRESS

There were three main objectives for this research:

(1) Develop new techniques built upon many-body perturbation theory and coupled cluster theory to more accurately describe potential energy surfaces for molecules.

(2) Implement highly efficient computer codes to perform accurate many-body calculations of molecular structures and reaction paths.

(3) Apply these methods to a series of examples pertaining to transient molecules as occur in flames, to identify structures, transition states for reactions, and exothermicities. This information is required in kinetic models of flame systems.

Over the past three year period we have made exceptional progress in all three areas. Some of this progress may be summarized. A more detailed account is available in the published papers.

The equations of many-body perturbation theory(MBPT) and its infinite-order generalizations, coupled-cluster(CC) theory, were derived using diagrammatic techniques[1]*. These formulae were found to be well-suited to the new generation of vector computers, since the diagrams may be evaluated as products of row and column vectors of suitably sorted molecular integrals. Codes were implemented to include single, double, and quadruple excitation terms, defining the fourth-order MBPT model(SDQ-MBPT(4)) and its infinite-order, CCD (coupled-cluster doubles) and CCSD (single and double) extensions[1].

A primary advantage of these <u>ab initio</u> MBPT/CC models[1] is that they scale properly with molecular size. We call this property size-extensivity[2] since it is the electronic equivalent of extensive properties in statistical thermodynamics. Unlike configuration interaction(CI) this makes many-body methods formally suited to large-molecule applications. It also guarantees an inherent comparative efficiency compared to CI due to the elimination of unlinked diagrams in the MBPT/CC approach.

* References refer to the publications supported by ARO listed in section III.

One other consequence of this property is the very important one that the theoretical heat of formation(ΔH_f) in some reaction A + B → C + D , may be simply obtained as $\Delta H_f = \Delta H_f(C) + \Delta H_f(D) - \Delta H_f(A) - \Delta H_f(B)$, just as when using experimental values, when the ΔH_f of each species is calculated by a many-body method. This is not true in CI. Instead one must compute the super-molecules A + B and C + D infinitely far apart to obtain this information[2]. Clearly, this property is very important in thermochemical studies of flame species.

Among our first applications of these MBPT/CC techniques to flame species were studies of the potential energy surface of formaldehyde, H_2CO. The ground state of H_2CO may decompose via at least three routes: a) $H_2CO \rightarrow H_2+CO$, b) $H_2CO \rightarrow HCO+H$, and c) $H_2CO \rightarrow HCOH \rightarrow$ products. We studied this potential energy surface using MBPT methods determining transition states and barriers for routes a) and c), with no barrier in b). The exothermicities are quite accurate. Also, we obtained excitation energies from the ground to the low-lying singlet and triplet states. In the case of the triplet state, we predict an excitation energy of 3.12eV while the experimental value is 3.14eV[3].

Our next study dealt with dissociation energies for the series of molecules $CH_3OH \rightarrow CH_3O \rightarrow CH_2O \rightarrow CO$ [4]. This made use of the size-extensive property referred to above. It also made predictions about the structure and dissociation energy of the methoxy radical, which are not known experimentally. For each species, we obtained its molecular structure using MBPT techniques, showing that bond lengths are typically accurate to 0.01Å and bond angles to less than about 3°[4].

Our study on dissociation energies in the methanol flame was followed by a more detailed study of decomposition pathways for methoxy CH_3O and its isomer, the hydroxymethylene radical, CH_2OH[5]. Transition states and barriers to hydrogen dissociation for methoxy, i.e., $CH_3O \rightarrow CH_2O + H$ and rearrangement, $CH_3O \rightarrow CH_2OH$ were determined. The barriers for the two routes are quite close,

suggesting that under some conditions the rearrangement route can compete with direct CH bond fission. These results should have a bearing on kinetics models of the methanol flame[5].

In addition to applications of our MBPT/CC techniques, we proposed to ARO our intention to augment MBPT/CC methods to properly include the effects of triple excitations. Hence, we developed formulae and wrote two separate programs to find an efficient calculational method, since adding the triple excitation terms adds an n^7 step (n= # of basis functions) to our previously n^6 MBPT/CC models[6]. We reported our results including comparisons with full CI for a series of molecules, BH, HF, NH_3 and H_2O, the latter at several geometries[6]. This work demonstrates the high accuracy of the far more efficient MBPT/CC methods compared to the best possible result (i.e. full CI), even for highly difficult cases where the reference function is a poor approximation to the correct solution[6].

Returning to applications, we made a study of methylene amidogen, CH_2N[7], which arises as a primary decomposition product of the explosive HMX. It is also expected to be an intermediate in flames oxidized by N_2O. This work reports a study of five different electronic states and ground state structural parameters[7]. Comparisons are made among different <u>ab initio</u> methods reporting complete active space multi-configuration self-consistent field (CASSCF), CI, MBPT and CC results. Except for cases where the reference function is inappropriate the MBPT/CC results are found to be quite accurate[7].

Besides the radicals that occur in hydrocarbon and nitramine flames, are radicals containing halogens. In particular, Y. Lee and co-workers observed the existence of an unexpected inorganic free-radical in the reaction of F_2 + HI + HIF + F. From their experiments, HIF was bound by ~ 30kcal/mole relative to HI + F or IF + H yet their experiment could not determine whether the structure of the molecule was HIF, HFI, or FHI? Nor could they determine whether the radical was linear? We undertook this study using MBPT to locate the geometry of

the radical[8]. As might be expected from analogy with interhalogen species, the heaviest halogen (I) is in the middle, but HIF has a bond angle of 84°, showing some interaction between H and F. The barrier to dissociation for HF + I is 12 kcal/mole. Another unusual feature is that HIF also shows a local minimum at a bond angle of 137° which should cause an unusual vibrational spectra[8]. This paper will be finalized shortly. In future work, we expect to predict the vibrational spectra from the potential energy surface we have computed.

Our initial work under ARO support has documented the applicability of high-level, correlated ab initio, many-body techniques for the detailed study of potential energy surfaces for transient molecules. Often, only by such calculations is it possible to obtain structural, spectroscopic and kinetic information for short-lived transient species like CH_2O, HNO, and HCO which have been studied in this effort. The next important development in MBPT/CC is to derive and implement equations for the analytical evaluation of the gradient on a potential energy surface. This problem, as well as the extension of ab initio techniques to much larger molecules is the goal of our new ARO contract.

III. PUBLICATIONS AND PRESENTATIONS SUPPORTED BY ARO

PUBLICATIONS

1. R.J. Bartlett, Ann. Rev. Phys. Chem. 32, 359 (1981).

2. R.J. Bartlett and G.D. Purvis, "Electron Correlation in Large Molecules with Many-Body Methods," Annals, New York Academy of Sciences, 367, 62 (1981).

3. G.F. Adams, G.D. Bent, R.J. Bartlett and G.D. Purvis, "Formaldehyde: Electronic Structure Calculations for the S_o and T_1 States," J. Chem. Phys. 75, 834 (1981).

4. G.F. Adams, G.D. Bent, G.D. Purvis and R.J. Bartlett, "Calculation of Dissociation Energies Using Many-Body Perturbation Theory," Chem. Phys. Lett. 81, 461 (1981).

5. G.F. Adams, R.J. Bartlett and G.D. Purvis, "On the Unimolecular Reactions of CH_3O and CH_2OH," Chem. Phys. Lett. 87, 311 (1982).

6. R.J. Bartlett, H. Sekino, and G.D. Purvis, "Comparisons of MBPT and Coupled-Cluster Methods with Full CI. Importance of Triple Excitations and Infinite Summations," Chem. Phys. Lett. 98, 66 (1983).

7. G.F. Adams, D.R. Yarkony, R.J. Bartlett and G.D. Purvis, "Electronic Structure and Vertical Excitation Spectrum of Methylene Amidogen, CH_2N," Int. J. Quantum Chem. 23, 437 (1983).

8. S.A. Kucharski and R.J. Bartlett, "Fourth-Order MBPT Study of the Geometry and Decomposition Path of the HIF Radical," J. Chem. Phys., to be submitted.

PRESENTATIONS

(Rodney J. Bartlett, invited presentations at conferences and universities)

1. "Theoretical Chemistry: Its Contribution to the Department of Defense," Redstone Arsenal, Huntsville, AL, June 1980.

2. "Contribution of Many-Body Methods to the Study of Energetic Materials," Workshop on Fundamental Research Directions for the Decomposition of Energetic Materials, Berkeley, CA, January 1981, organized by ONR. ARO, and AFOSR.

3. "A Perspective on Many-Body Methods in Quantum Chemistry," Southeastern Theoretical Chemistry Conference, Birmingham AL, May 1982.

4. "Progress in Many-Body Methods for Electron Correlation in Molecules," Fourth International Congress of Quantum Chemistry, Uppsala, Sweden, June 1982.

5. "The Coupled-Cluster Single and Double Excitation Model with Localized Orbitals," Meeting on Electron Correlation in Molecular Systems: Methods and Applications, Orenas, Sweden, June 1982.

6. "Many-Body Methods in Chemistry," University of Grönigen, Grönigen, The Netherlands, June 1982.

7. "Many-Body Theory and Its Applications to Molecules," Ohio State University, Columbus, OH, February 1982.

8. "Applications of Many-Body Methods to Large Molecules," Chemical Defense Research Conference, Chemical Research and Development Center, Aberdeen Proving Ground, November 1983.

9. "Many-Body Methods in Chemistry," Departmental Colloquium, Georgia Institute of Technology, Atlanta, GA, March 1983.

10. "Contributions of Ab Initio Quantum Chemistry to Defense Problems," Chemical Systems Laboratory, Edgewood, MD, May 1983.

11. "Single and Multi-Reference Many-Body Methods for Potential Energy Surfaces," Sanibel Symposium on the Quantum Theory of Matter, March 1984.

* A number of contributed presentations by graduate students and postdoctoral students have been made at the Sanibel-Palm Coast meetings and the Southeastern Theoretical Chemistry conferences, 1982-1984.

IV. PERSONNEL SUPPORTED BY ARO

No students supported by ARO have yet received their degrees, although Mr. Sohrab Zarrabian and Mr. David Magers have been working on ARO projects. Postdoctoral students who have worked on ARO projects include Dr. George Purvis, Dr. Hideo Sekino, and Dr. S.A. Kucharski. Dr. Purvis who worked on this project at Battelle and Florida, has now become an Associate in Chemistry and Manager of the Quantum Theory Project's computer systems.

Ann. Rev. Phys. Chem. 1981. 32:359–401

MANY-BODY PERTURBATION THEORY AND COUPLED CLUSTER THEORY FOR ELECTRON CORRELATION IN MOLECULES*

*2733

Rodney J. Bartlett[†]

Battelle Memorial Institute, Columbus, Ohio 43201[‡]

INTRODUCTION

Ten years ago in the *Annual Review of Physical Chemistry*, there was a review article entitled "Many-Body Theories of the Electronic Structure of Atoms and Molecules," by Karl Freed (1). In that article many-body methods were defined to be those techniques which derive their impetus from theories of the N-body problem for which $N \rightarrow \infty$. For the purposes of this review, we further specify these methods as many-body perturbation theory (MBPT) (2–5) and the closely related coupled-cluster methods (CCM) (6–9).

In the ten years since that review appeared, probably no area in theoretical chemistry has undergone more development than has the theory, methodology, and applications of such ab initio many-body methods for

†This research was sponsored by the Air Force Office of Scientific Research, the Office of Naval Research, and the Army Research Office under Contracts Nos. AFOSR 78-3661, NOOO14-79-C-0821, and DAAG29-80-C-0105, respectively. The United States Government is authorized to reproduce and distribute reprints for government purposes notwithstanding any copyright notation hereon.

‡Present address: Quantum Theory Project, University of Florida, Gainesville, FL 32611

studies of molecules and their interactions. In Freed's article the question is asked, "Do [many-body methods] provide methods or insights which are useful and/or new?" The answer is a resounding "yes!" Furthermore, these methods should have an increasingly important impact on chemistry over the next few decades, both conceptually and in providing highly accurate theoretical predictions for a constantly expanding array of problems.

In the present review, which is very much directed toward the nonspecialist, I hope to present some of the logic underlying the application of many-body theory for chemistry and to illustrate and call attention to several of the contributions that have been made over the last decade. As in any effort of this type, space limitations prohibit any attempt at completeness, but it is hoped that the essence of the developments in the area will be communicated to the reader, along with an assessment of where many important contributions remain to be made.[1]

Besides MBPT/CCM, many-body methods include techniques built upon Dyson's equation and its generalizations, which go under a variety of names such as Green's functions, propagator methods, and equations-of-motion (EOM) techniques. These methods are most often used for obtaining electronic spectra, ionization potentials, electron affinities, and the related transition moments, although they can be used to obtain the total electronic energy of a molecule as well, and thus also potential energy surfaces. In fact, one of the most attractive approaches to excited-state potential surfaces would be to combine EOM methods with ground-state

[1]Abbreviations used: ASGD, antisymmetrized Goldstone diagram; CCD, coupled cluster method limited to double excitation operators (this is also known as CPMET, coupled pair many-electron theory); CCM, coupled cluster method; CCSD, coupled cluster method limited to single and double excitation operators; CCD + ST(4) designates fourth-order single and triple excitation contributions added to the CCD result; CEPA, coupled electron pair approximation; CGTO, contracted Gaussian type orbitals; CHF, coupled Hartree-Fock; CI, configuration interaction (CI specified to include various categories of excitations are designated as D-CI, for double excitations, SD-CI for single and double excitations, and SDTQ-CI for single, double, triple, and quadruple excitations); EN, Epstein-Nesbet partitioning in perturbation theory (Ref. 65); EPV, exclusion principle violating; EOM, equations of motion; GVVT, generalized Van Vleck transformation; ICSCF, internally consistent self-consistent field orbitals; L-CCD, linearized coupled cluster double excitation model; MBPT, many-body perturbation theory [MBPT specified to include ASGD corresponding to single, (S), double, (D), triple, (T), and quadruple (Q), excitations are identified as SDTQ-MBPT(4), with the order dependence identified in the designation]; MP, Møller-Plesset partitioning in perturbation theory (Ref. 64); PCILO, perturbed CI with localized orbitals' semiempirical model; PES, potential energy surface; PPP, Pariser-Parr-Pople semiempirical model; RHF, restricted Hartree-Fock; SCF, self-consistent field model; STO, Slater type orbitals; UHF, unrestricted Hartree-Fock; VB, valence bond model.

solutions obtained by MBPT/CCM (10, 11). These EOM and propagator methods use techniques somewhat different than MBPT/CCM, and embrace a very large literature of their own; consequently, they are not discussed here. These approaches are mentioned in Freed's review article, and have been reviewed more recently by Öhrn (12), Simons (13), Jørgensen (14), and Csanak, Taylor & Yaris (15). An excellent text covering this subject is by Linderberg & Öhrn (16).

Besides the review of Freed (1), a number of other reviews of MBPT and CCM have appeared within the last decade. These include papers by Robb (17), Kutzelnigg (18), Paldus & Čížek (19), part of a review by Musher (20), and Čížek & Paldus (21). The latter paper, as well as a review of much of the effort of our group (22), appears in the proceedings of the 1979 Nobel Symposium on Many-Body Theory of Atomic Systems. These proceedings offer a quite up-to-date account of many-body theory in its most comprehensive sense—embracing atoms, molecules, nuclei, and solids—and are highly recommended.

Each of the previous reviews of MBPT/CCM deals with different aspects of the problem. Robb's review focuses on the relationship between MBPT and the electron-pair theories as originally presented by Sinanoglu (23) and Nesbet (24). Another article by Freed also addresses this aspect (25). The paper by Paldus & Čížek presents a functional, very readable account of the detailed MBPT theory. Kutzelnigg's article focuses mainly on the coupled-cluster theory and its coupled electron pair approximation (CEPA) versions (26). Also, the older review by Kelly (5) provides an excellent account of the first MBPT calculations for atoms. The well-known book (27) edited by Sinanoglu & Brueckner brings many of the principal early papers together. The second article by Čížek & Paldus (21) reviews the development of coupled-cluster theory, while the book by Hurley (28) provides a good textbook account of CCM.

Most applications of ab initio MBPT/CCM for *molecular* problems have only been made in the last five years. Hence, no prior review has covered predominantly numerical results for molecules and some of the implications of these results. This appears to be the natural subject for the present effort. To further restrict the scope of the review, with important exceptions, atomic calculations of the type pioneered by Kelly (5) and semi-empirical molecular calculations also are not considered.

Following a discussion of some of the significant concepts underlying MBPT/CCM, together with a synopsis of the theory, we proceed to some informative numerical illustrations of the techniques, demonstrating the deficiencies and the successes of the many-body methods, while emphasizing profitable future areas for research.

CRITERIA FOR A THEORETICAL MODEL CHEMISTRY

To offer a focus for the subsequent discussion, it is helpful to consider some properties that quantum mechanical methods should attempt to satisfy. In an interesting paper, Pople et al (29) proposed several criteria for approximate numerical applications for molecules, if they are to be suitable as a basis for a "theoretical model chemistry." The definition of the latter is a uniform level of calculation whose effectiveness may be assessed by comparing with experimental data. Considering their suggestions as well as some of our own (31), we think some of the conditions that such a model should satisfy are that it be

1. size-extensive (i.e. scale properly with the size of molecule)
2. generally applicable to a wide class of problems within one framework (i.e. the model should not be dependent on symmetry or specific choices of configurations)
3. invariant to classes of transformations, particularly unitary transformations among degenerate orbitals
4. efficient and cost effective
5. applicable to excited states and open shells
6. able to dissociate a molecule correctly into its fragments.

The best possible solution in a basis set is configuration interaction (CI) with all possible excitations (full CI). Except for efficiency, since full CI is impossible for anything but the smallest molecules, it satisfies all these criteria as well as being variational. However, lacking a coordinate lower bound, a variational upper bound seems to be a less important requirement than it once was in quantum chemistry. There are two primary reasons for this.

First, the quantities that are of interest, such as binding energies—or more generally, the relative energy on a potential energy surface (PES)—and excitation energies, have no variational bounds even if the separate calculations are each variational. In fact, some of the most attractive methods (e.g. EOM and perturbation techniques) for obtaining excitation energies compute the differences directly rather than via two separate calculations, an inherently nonvariational approach. Second, the many-body methods that are routinely used, although nonvariational, usually differ from rigorous variational bounds only in fourth and higher orders of perturbation theory (30). Hence, at least for nonpathological cases for which such higher-order corrections could be important, MBPT/CCM methods as usually employed are quasivariational.

At the current state of development, the first three conditions are easily accomplished by MBPT/CCM. Any approach based upon the linked-diagram theorem is size extensive, as discussed in depth in the next section. A large class of problems can be studied within the general framework of single reference MBPT/CCM calculations, provided that the reference function is an adequate starting point. As long as entire diagrams are evaluated, MBPT is invariant to unitary transformations among degenerate orbitals. CEPA models that include parts of MBPT diagrams usually lack this invariance. CCM models are typically invariant to transformations exclusively among occupied orbitals and among excited orbitals, but not necessarily when the two are mixed.

The efficiency criterion for correlated calculations is very important. To emphasize this aspect, the number of points required to obtain a potential energy surface (PES) rises astronomically with the number of degrees of freedom in the molecule. For a triatomic system, calculations at ten displacements in each degree of freedom require 10^3 points, but for four atoms, already 10^6 calculations would be required. Usually far fewer points actually need to be obtained than suggested by such a brute-force approach, but since each calculation still requires significant amounts of computer time, the importance of efficiency cannot be overestimated.

Many-body methods have traits that enable them to offer distinct computational advantages for many problems compared to some other correlated techniques:

1. A simple second-order perturbation result, which only requires a trivial addition to any SCF code, accounts for typically ~90% of the basis set correlation energy and removes most of the SCF error in other properties.
2. MBPT/CCM offer very efficient techniques for incorporating most of the effect of higher than double excitations in CI.
3. Substantial computational efficiency is gained through the fixed computational formulae (at least for a single reference function) of MBPT/CCM, which are also ideally suited to vector-based computations.

Excited states can be handled with other kinds of many-body approaches such as EOM techniques (32) or with CI; while correct separation is currently most easily achieved using CI methods. The MBPT/CCM theory for treating both these problems exists, but has not yet been implemented into a general-purpose molecular problem. Of course, excited states, as long as they are the lowest state of a given symmetry, are routinely studied with unrestricted Hartree-Fock (UHF) based MBPT/

CCM (29, 30). Also, UHF plus MBPT/CCM usually permit smooth separation on a PES, although for some cases the path toward the separated limit is not always accurate (22).

SIZE-EXTENSIVITY AND THE IMPORTANCE OF HIGHER CI-EXCITATIONS IN THE THEORY OF CORRELATION

MBPT/CCM are relatively new techniques, compared to configuration iteration (CI), for the determination of electron-correlation effects for molecular properties. MBPT/CCM were originally developed for problems in nuclear and solid state physics, where emphasis on correct size-dependence, which we refer to as "size-extensivity" (22, 30),[2] becomes mandatory. Size-extensivity is guaranteed by the evaluation of terms that the many-body development identifies as linked diagrams, hence the Brueckner-Goldstone (2–4) linked-diagram theorem of MBPT/CCM serves as the cornerstone of the theory. In a solid consisting of an infinite number of atoms, the correct size-dependence is obviously crucial, yet even in molecular problems this is a highly desirable trait for an approximate method to possess.

[2]This term is borrowed from thermodynamics, where an extensive property is one that is proportional to the size of a homogeneous system. Pople et al (29) proposed the term "size-consistency" for a closely related property. A method is considered size-consistent if the energy of a system made up of two subsystems A and B far apart is equal to the sum of the energies A and B computed separately by the same method. For closed-shell systems dissociating to closed-shell fragments, an RHF (restricted Hartree-Fock) reference function is size-consistent, and size-extensivity (or the absence of unlinked diagrams) is then a sufficient condition for size-consistency of a correlated model based upon that reference function. On the other hand, for a single-determinant reference function to be size-consistent when considering open-shell fragments, A and B, a UHF (unrestricted Hartree-Fock) solution will usually be required. Since either a RHF or a UHF function can be used to generate a linked diagram, size-extensive expansion, the condition of size-consistency imposes another requirement on the reference function, in addition to the use of a linked diagram expansion. Hence, some confusion arose over the distinction between proper scaling with size in a homogeneous system ("extensivity") and correct separation of a molecule into its fragments, which is a very different property. Adding to the confusion is the term "separability conditions," which is also sometimes used (18). Hence, we prefer the term size-extensivity to suggest only correct scaling, and discuss correct separation as a second, equally desirable property, to avoid any possible confusion. Although there are fine differences (33), the fundamental element suggested by all three terms is correct scaling with size.

The idea of size-extension is implicit in the work of Brueckner, Goldstone, and co-workers, although Primas was one of the first to emphasize the concept (34). Meyer (26) used this idea as part of the justification of the CEPA models.

Two primary reasons for this are that only approximate methods that scale properly with size are suitable for application to larger molecules, such as those encountered in quantum biochemistry (31); and size-extensivity assists in computing accurate dissociation energies (or, more generally, correct relative energies on the potential surface), which require comparison of a molecule to its smaller fragments (22). The latter aspect is also crucially dependent upon basis set effects and on whether the approximate method being used permits smooth dissociation into the different components, but size-extensivity remains an important aspect. As an added benefit, heats of formation obtained from calculations using size-extensive methods can be added together, just as experimental values are, to obtain the heats of formation of some complex molecules (22), while nonsize-extensive methods, like truncated CI, would normally require "super molecule" calculations to provide these quantities most accurately. This problem is of more than academic interest as pointed out by Ahlrichs (35), since this error is $\sim$9 kcal/mole for $2BH_3 \rightarrow B_2H_6$ (35) and is $\sim$15 kcal/mole for $CH_3F + F^- \rightarrow CH_3F_2^-$ (36).

Although full CI has the property of correct size-dependence, the failure of truncated CI models, such as CI limited to all single and double excitations (SD-CI), is related to the neglected contributions of higher excitations in the CI method. The example usually employed to illustrate this deficiency in CI is a lattice of noninteracting electron pair bonds, such as a collection of well-separated H_2 molecules or He atoms. The exact wavefunction for a single H_2 molecule may be written in terms of a complete set of natural orbitals as a reference determinant (i.e. the first natural configuration, which is close to the SCF solution) plus all paired double excitations from that determinant. Two noninteracting H_2 molecules would be exactly described by the product of two of these wavefunctions, but the product of the simultaneous double excitations on each H_2 molecule results in contributions to the product wavefunction of terms that correspond to quadruple excitations, if the two H_2 molecules are treated as a "super molecule" in CI. Similarly, a third H_2 molecule requires hextuple excitations, and so forth. However, since the number of configurations is proportional to $(n)^l$, where n is the number of basis functions and l is the level of excitation, a prohibitive $\sim 10^8$ configurations would be necessary for $n = 100$ to include just the quadruple excitations. The fact that much of the effect of such higher excitation terms is tractably included in many-body theory is one of the principal advantages of these techniques.

Since any large molecule can be viewed, in a first approximation, as a superposition of largely noninteracting electron pair bonds, the idealized

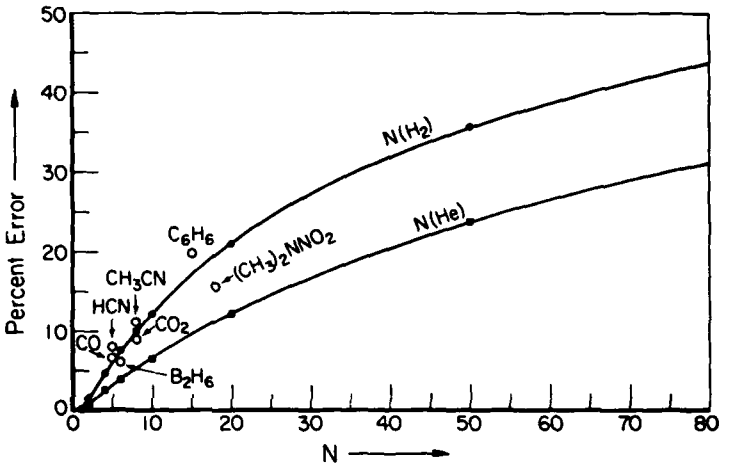


Figure 1 Percentage of error in the correlation energy as measured by D-CI for N separated H_2 molecules and He atoms. Values identified for specific molecules are obtained by actual fourth-order calculations that include quadruple excitations compared to a fourth-order approximation to D-CI. All calculations refer exclusively to the valence electrons, with the K-shell electrons frozen at the SCF level.

$N(H_2)$ model has some significance for the general description of molecular electronic structure. This model problem has been considered by several authors (37–43). Using very good full CI wavefunctions for H_2 to provide the one parameter required (31), it is possible to solve the model problem exactly for the H_2 lattice problem, and thus to determine the error in D-CI (double excitation CI) for N molecules as a function of N (31). These numbers are illustrated in Figure 1 for H_2 and He lattices.

Unlike a size-extensive correlated model, where

$$E_{corr}[N(H_2)] = N[E_{corr}(H_2)],$$

the correlation energy obtained by truncated CI for the lattice is proportional to $\sqrt{N}$ as $N \to \infty$. Even for $N = 10$ (twenty electrons), the error for $N(H_2)$ amounts to 31 kcal/mole, while for forty-electron systems, which are well within the range of current correlated studies, the error would be about 108 kcal/mole. The differences between the $N(H_2)$ curve and N(He) curve reflect the difference between inner-shell electron pairs and the pairs in covalent bonds. Although there are other errors in molecular calculations that are equally large, or even larger, it is apparent that size-extensivity has a potentially important effect in ab initio correlated calculations. A statement that size-extensivity is important is equiva-

lent to a statement that higher than single- and double-CI excitations are important.

Also shown in Figure 1 are some results for the effect of quadruple excitations on correlation energies obtained in actual calculations. These quantities are determined by comparing fourth-order MBPT results that include quadruple excitation effects with a fourth-order perturbation approximation to D-CI (22, 30). The core electrons are frozen in these examples, so that the number of electron pairs, N, for each molecule corresponds to the valence electrons only. The calculations for benzene and dimethylnitramine use a double-zeta (*DZ*) basis, while at least double-zeta-plus-polarization (*DZP*) sets are employed in the other cases. Typically, a larger quadruple-excitation effect is observed with better basis sets, so the 20% error in the correlation energy in benzene, which amounts to 64 kcal/mole, is likely to be an underestimate of the true effect for this system. The actual calculations tend more nearly to follow the $N(H_2)$ curve rather than that for N(He), as would be expected for the valence electron bonds.

Although this example pertains to the total energy, while in chemistry we are mainly concerned with energy differences, failure to maintain proper size-dependence has additional consequences. Using the H_2 lattice problem, Meunier & Levy (40) demonstrate that the density matrix and electron excitation energies will also be drastically affected if these quantities are computed by truncated CI. In each case, as $N \rightarrow \infty$, the CI density matrix and excitation energies will converge to the values corresponding to the SCF reference function, so that a great deal of effort involved in the correlation treatment for an extended system would be of little value. Even for small molecules, however, the effects of quadruple excitations can have observable consequences, as shown for the geometry and force constants of H_2O (33).

OUTLINE OF MBPT/CCM THEORY

The development of MBPT predates CCM, but the latter possesses some conceptual simplifications that recommend that CCM be treated first, with the various orders in MBPT being obtained from the CCM equations. Although, as illustrated by the H_2 lattice, the products of disjoint double-excitation terms that are neglected in CI are fundamentally simple, the standard CI framework cannot exploit this simplicity. Many-body methods accomplish this primarily through the exponential ansatz (6–8), which states that the correct wavefunction may be written in the form

$$\psi = e^T|\Phi_0\rangle, \qquad 1.$$

where T is an excitation operator which we may define as

$$T = T_1 + T_2 + \ldots + T_n. \qquad 2.$$

The subscripts refer to the number of excited electrons. In the occupation number representation,

$$T_n = 1/n! \sum_{\substack{ijk\ldots \\ abc\ldots}} t_{ijk\ldots}^{abc\ldots} X_a^\dagger X_b^\dagger X_c^\dagger \ldots X_k X_j X_i, \qquad 3.$$

where $a,b,c,\ldots$ are excited orbitals, while $i,j,k\ldots$ are orbitals occupied in Φ_0. In Eq. 1, $|\Phi_0\rangle$ is some suitable, size-extensive reference function, and the amplitudes $t_{ijk\ldots}^{abc\ldots}$, which are analogous to CI coefficients, are to be determined. A double-excitation wavefunction of this type for two separate H_2 molecules, referred to as H_2^A and H_2^B, has the property that (44)

$$\exp(T_2^A + T_2^B)|\Phi_0^A\Phi_0^B\rangle = (\exp T_2^A|\Phi_0^A\rangle)(\exp T_2^B|\Phi_0^B\rangle) \qquad 4.$$

since the cross terms vanish for the noninteracting case.

This approach should be contrasted with the truncated CI technique, for which the D-CI wavefunction for the two H_2 molecules may be written as $(1 + T_2^A + T_2^B)|\Phi_0^A\Phi_0^B\rangle$. This does not correspond to

$$[(1 + T_2^A)|\Phi_0^A\rangle] \times [(1 + T_2^B)|\Phi_0^B\rangle],$$

since the product term, $T_2^A|\Phi_0^A\rangle\, T_2^B|\Phi_0^B\rangle$ is a quadruple excitation and is thus neglected, so that the result is not size extensive.

The exponential form of the wavefunction, Eq. 1, combined with the condition that the T operator in Eq. 2 contains no disconnected parts (i.e. parts which can be resolved into products of two or more lower T_n operators), is sufficient to guarantee that the energy given by

$$E = \langle\Phi_0|He^T|\Phi_0\rangle \qquad 5.$$

is size-extensive. An alternative expression for the energy is the linked-diagram theorem (1, 2) of MBPT

$$E = \sum_{k=0}^{\infty} \langle\Phi_0|H[(E_0 - H_0)^{-1}H]^k|\Phi_0\rangle_L \qquad 6.$$

which is an expansion of orders in the perturbation $V = (H - H_0)$. When Φ_0 is a self-consistent field (SCF) function, the terms with $k > 0$ represent the correlation energy (45). H_0 is a separable n-particle Hamiltonian whose eigenfunction (at least in a matrix sense) is Φ_0. The subscript L indicates the limitation to linked diagrams, such as those shown in Figure 2. These topological figures provide a convenient mnemonic device for writing down the contribution of any order of perturbation theory solely

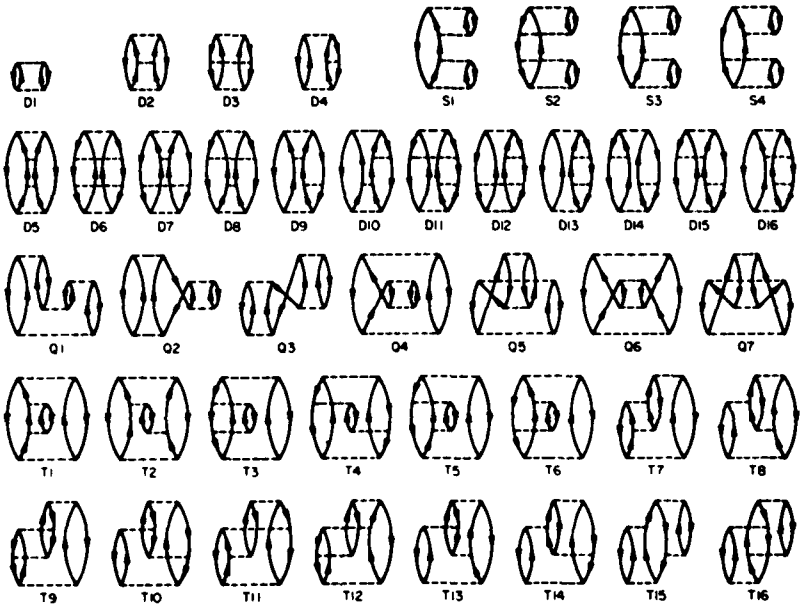


Figure 2 Antisymmetrized Goldstone diagrams (ASGDs) through fourth order MBPT. A Hartree-Fock reference state is assumed. Orders are distinguished by the number of dashed horizontal interaction lines. Particle (p) and hole (h) states are represented by upward and downward directed line segments, respectively. The excitation level of a diagram is distinguished by the number of p-h pairs intersected by an imaginary central horizontal line. In this manner the diagrams have been labeled and counted as contributions to the correlation energy arising from single (S), double (D), triple (T), and quadruple (Q) excitation types.

in terms of molecular integrals. The rules for interpreting such diagrams have been presented (46). The close relationship between Eqs. 5 and 6 will become more evident below, where the solutions of the CCM equations are considered.

The usual choice for Φ_0 is an SCF function of the restricted or unrestricted form. Although this choice is not mandatory, it is computationally simpler, since many terms vanish for this case. For closed-shell systems near equilibrium, a restricted Hartree-Fock (RHF) function is usually a good choice for Φ_0. However, it is well-known that such a function will not separate correctly upon dissociation for any molecule that fails to separate into closed-shell fragments. An unrestricted Hartree-Fock (UHF) function will usually, but not always (47), separate correctly. However, such a function may suffer from extensive spin contamination, particularly when a UHF function is used for a multiply bonded

singlet molecule like N_2 (22). For most open-shell doublets, triplets, etc, UHF functions have only a small amount of spin-contamination, in general, and can usually be employed without excessive error (22). In the UHF or closed-shell RHF case, Brillouin's theorem eliminates any need to consider non-Hartree-Fock single-particle terms.

It is also possible to treat open-shell problems with Roothaan RHF methods (48, 48a), but these do not treat all multiplicities in an equivalent way; they require additional non-Hartree-Fock terms in the summations, and they are not useful for treating entire potential energy surfaces. Multi-reference function methods, which are more appropriate in such cases, are discussed below. For some properties, such as excitation spectra and ionization potentials, open-shell RHF MBPT should be useful.

The CCM theory, built upon Eqs. 1 and 2, originated in nuclear physics primarily through the work of Coester & Kümmel (6, 7), with Čížek (8) presenting the theory in a suitable form for quantum chemistry. Čížek, Paldus, and co-workers have made many applications to π-electron systems within the Pariser-Parr-Pople (PPP) semi-empirical framework (9). Paldus, Čížek & Shavitt made the first, and still one of the most informative, ab initio applications in a minimum basis study of BH_3 (49). This paper considers effects of T_1, T_2, T_3, and T_4 in the CCM equations.

Restricting the T operator just to bi-excitations defines the coupled-cluster doubles (CCD) model, also known as coupled-pair many electron theory (CPMET). The first results from generally applicable computer programs using the CCD model were reported by Bartlett & Purvis (30), Pople et al (50), and Taylor et al (51).

Using the CCD model, equations for the amplitudes can be obtained by back-projecting $He^{T_2}|\Phi_0\rangle$ onto the space of double excitations. These lead to a set of nonlinear coupled equations for the amplitudes of the form (8, 30)

$$\begin{aligned} 0 = \langle ab|\,|ij\rangle - D_{ijab}t_{ij}^{ab} &+ \sum_{c>d}\langle ab|\,|cd\rangle\, t_{ij}^{cd} + \sum_{k>\ell}\langle k\ell|\,|ij\rangle\, t_{k\ell}^{ab} \\ &+ \sum_{k,c}\{-\langle kb|\,|jc\rangle\, t_{ik}^{ac} + \langle ka|\,|jc\rangle\, t_{ij}^{bc} + \langle kb|\,|ic\rangle\, t_{jk}^{ac} \\ &- \langle ka|\,|ic\rangle\, t_{jk}^{bc}\} + \sum_{\substack{k>\ell\\ c>d}}\langle k\ell|\,|cd\rangle\,\{t_{ij}^{cd}t_{k\ell}^{ab} - 2(t_{ij}^{ac}t_{k\ell}^{bd} + t_{ij}^{bd}t_{k\ell}^{ac}) \\ &- 2(t_{ik}^{ab}t_{j\ell}^{cd} + t_{ik}^{cd}t_{j\ell}^{ab}) + 4(t_{ik}^{ac}t_{j\ell}^{bd} + t_{ik}^{bd}t_{j\ell}^{ac})\}. \end{aligned} \qquad 7.$$

The antisymmetrized two electron integrals

$$\langle pq|\,|rs\rangle = (pr|qs) - (ps|qr),$$

are defined relative to the molecular orbital basis set, while $D_{ijab} = (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)$ is composed of the SCF orbital energies. From the number of operators involved, the highest terms in the exponential expansion that contribute to these amplitudes are quadratic.

Notice that there are no more amplitudes to determine in the wavefunction $e^{T_2}|\Phi_0\rangle$ than in the standard D-CI, yet at the modest cost of solving a nonlinear equation, we now have a size-extensive method and the inclusion of most of the effects of CI quadruple excitations. The latter follows by comparison, since the CI quadruple excitation operator C_4 is equivalent to

$$C_4 = T_4 + (1/2)\, T_2^2 + (1/2)T_1^2T_2 + T_1T_3 + (1/4!)T_1^4. \qquad 8.$$

However, as Sinanoglu observed (23), T_4 corresponds to a true four-particle interaction and is very small, while two simultaneous two-particle interactions, which correspond to T_2^2, are far more important. This is also supported by the fact that all quadruple excitation terms in fourth-order perturbation theory come from T_2^2, with T_4 only starting to contribute in fifth-order. Since $T_1 = 0$ for Brueckner orbitals and is usually small for SCF orbitals, the last three terms are normally less important. Hence, with only little more effort than D-CI we obtain a substantial extension of the D-CI theory.

The iterative solution of Eq. 7, which defines several MBPT models, proceeds in the following fashion. Initially all amplitudes are assumed to be zero, giving

$$t_{ij}^{ab}(1) = \langle ab||ij\rangle / D_{ijab} \qquad 9.$$

with energy

$$E_2 = \sum_{\substack{a>b \\ i>j}} \langle ab||ij\rangle\, t_{ij}^{ab}(1)\,. \qquad 10.$$

This is the second-order perturbation energy, which corresponds to the antisymmetrized diagram D1 in Figure 2. The next iteration, also only of the linear terms, defines

$$t_{ij}^{ab}(2) = \sum_{c>d} \langle ab||cd\rangle t_{ij}^{cd}(1) + \sum_{k>l} \langle kl||ij\rangle\, t_{kl}^{ab}(1) + \sum_{k,c} \{-\langle kb||jc\rangle t_{ik}^{ac}(1) + \langle ka||jc\rangle t_{ik}^{bc}(1) + \langle kb||ic\rangle t_{jk}^{ac}(1) - \langle ka||ic\rangle t_{jk}^{bc}(1)\}\,, \qquad 11.$$

$$E_3 = \sum_{\substack{a>b \\ i>j}} \langle ab||ij\rangle\, t_{ij}^{ab}(2)\,. \qquad 12.$$

E_3 is given by the third-order diagrams D2, D3, and D4 (Figure 2). The next iteration of the linear term would be the same as in Eq. 11, except that $t_{ij}^{ab}(2)$ would be used in place of $t_{ij}^{ab}(1)$ to give $t_{ij}^{ab}(3;L)$. However, we show elsewhere (52) that this computation is not necessary, since

$$E_4^D = \sum_{\substack{a>b \\ i>j}} \langle ab||ij\rangle t_{ij}^{ab}(3;L) = \sum_{\substack{a>b \\ i>j}} |t_{ij}^{ab}(2)|^2/D_{ijab}. \qquad 13.$$

E_4^D corresponds to diagrams D5-D16 in Figure 2.

The first nonlinear iteration contributes

$$t_{ij}^{ab}(3;N) = \sum_{\substack{k>\ell \\ c>c}} [\langle k\ell||cd\rangle \{t_{ij}^{cd}(1)t_{k\ell}^{ab}(1)] - 2[t_{ij}^{ac}(1)t_{k\ell}^{bd}(1) + t_{ij}^{bc}(1)t_{k\ell}^{ac}(1)]$$

$$-2[t_{ik}^{ab}(1)t_{j\ell}^{cd}(1) + t_{ik}^{cd}(1)t_{j\ell}^{ab}(1)] + 4[t_{ik}^{ac}(1)t_{j\ell}^{bc}(1) + t_{ik}^{bd}(1)t_{j\ell}^{ac}(1)]\}, \qquad 14.$$

which provides the amplitudes for

$$E_4^Q = \sum_{\substack{a>b \\ i>j}} \langle ab||ij\rangle\, t_{ij}^{ab}(3;N) \qquad 15.$$

or diagrams Q1-Q7. The superscripts D and Q refer to the two components of fourth-order perturbation theory corresponding to double- and quadruple-excitation diagrams. This defines the perturbation theory model DQ-MBPT(4). A similar consideration of the coupled-cluster single and double excitation (CCSD) wavefunction, $e^{T_1+T_2}|\Phi_0\rangle$, leads in addition to the fourth-order contribution of single excitations, which we define as SDQ-MBPT (4). This adds the S1-S4 diagram (Figure 2). The triple excitation diagrams T1-T16 (Figure 2) arise from T_3 in Eq. 2.

In this manner, each of the terms in the linked-diagram expansion of MBPT, Eq. 6, can be obtained from the general CCM equations. This development emphasizes two alternative viewpoints, the infinite-order summation of selected terms, and the evaluation of all terms at some finite order. When higher-order terms are important, a model like CCSD, which neglects the triple-excitation terms in fourth-order, may well be preferable to SDTQ-MBPT(4) which includes these terms. On the other hand, when good convergence is obtained, SDQ-MBPT(4) $\approx$ CCSD and the largest remaining error will be the fourth-order triple excitation terms. The latter applies to most cases where the reference function Φ_0, as a nondegenerate RHF or UHF function, is adequate for the problem at hand. In such examples DQ-MBPT(4) usually differs from CCD by

<1 kcal/mole (22). For more difficult cases, where a single Φ_0 suffers from some near degeneracy, CCD or CCSD might be preferred (30, 52–54).

A number of formal developments in the CCM theory have been made by Paldus, Čížek, and co-workers that deal with open-shell systems (55), excitation energies (11), and the spin-symmetry adapted form of the theory (56). Nakatsuji and co-workers have developed and applied an open-shell symmetry adapted cluster theory (57–59). Harris has also considered excitation energies, grafting EOM techniques onto the CCM equations (10). Other papers by Živkovič (60) and Živkovič & Monkhorst (61) discuss excited-state solutions to the CCM equations. Monkhorst also considers CCM predictions of properties other than the energy (62). Freeman succeeded in solving the electron gas problem with CCM (63).

The correlation contributions in Eqs. 9–15 are calculated iteratively solely from a list of transformed molecular integrals and associated denominators composed of orbital energies. Very efficient computer programs for carrying out these computations can be constructed, and such programs are well-suited to vector-oriented computers.

The computational effort of second-order perturbation theory has an approximately n^4 dependence on the number of basis functions, while third-order and the fourth-order contributions of single, double, and quadruple excitations, as well as the CCD and CCSD models, are asymptotically proportional in effort to n^6. The quadruple excitation contributions in fourth order and in CCD can be factored, since these correspond to T_2^2 rather than T_4. The latter type of term, which would occur in the most general case, would require an n^8 procedure. This factorization offers a dramatic gain in efficiency over attempting to include quadruple excitations in a CI procedure. The fourth-order triple excitation terms, on the other hand, correspond to T_3 rather than T_1T_2 and hence do not factor, requiring an n^7 procedure. Perhaps of even greater significance for computational purposes is that, if higher than fourth-order contributions of T_3 are to be computed, as would be done in a CCM model that includes T_3, it would be necessary to save the t_{ijk}^{abc} amplitudes, and this would require storage space proportional to n^6.

Some other CI models use a multireference space instead of a single reference function, and include all single and double excitations relative to the several reference functions. In this case, presumably, the most important triple and quadruple excitations are introduced into the calculation. Although, just like any truncated CI model, this is not rigorously size-extensive, it should be much better than single-reference SD-CI, and may be comparable or even superior to single-reference MBPT/CCM models.

which are rigorously size extensive and incorporate higher excitations. In this review the term CI, when unspecified, means a single-reference, truncated CI model.

Before focusing on MBPT/CCM applications and the individual orders of perturbation theory, it is useful first to recognize some of the flexibility inherent in the MBPT/CCM methodology.

The Jacobi type iterative solution of the CCD equations described above is implicitly of Møller-Plesset (MP) type, since the simple denominators D_{ijab} are those used by Møller & Plesset in their classic 1934 paper on perturbation theory relative to an SCF reference function (64).[3] Numerous other choices are possible. For example, the "diagonal" terms in the linear summation of Eq. 7, $\langle ab||ab\rangle$, $\langle ij||ij\rangle$, $\langle jb||jb\rangle$, $\langle ja||ja\rangle$, $\langle ib||ib\rangle$, and $\langle ia||ia\rangle$ could be combined with D_{ijab}, with the other summation indices suitably restricted, to provide a different partitioning known as the Epstein-Nesbet (EN) approach (65). In many-body language, such denominator "shifts" are used to sum certain kinds of terms in perturbation theory to all orders. Obviously, this technique will give very different results for equivalent orders in perturbation theory, although at convergence the final results must be the same. Some study of the nonlinear terms will also lead to certain "quasidiagonal" parts that might be incorporated into the denominator of an iterative scheme. In fact, the latter is necessary to make the connection between the rigorous CCD model and the CEPA approximations to it (18, 66).

In addition to the flexibility in the partitioning used in solving the CCD equations, it is also possible to get very different order-by-order results by different choices for the molecular orbitals involved in the calculations [the equations would then include extra terms in the non-SCF case (49)]. Alternative choices could be natural orbitals, Brueckner orbitals, modified SCF orbitals proposed by Davidson (termed ICSCF for "internally consistent") (67, 68), or those obtained by incorporating modified one-electron potentials (e.g. V^{N-1}) into the calculation. In the last case, the

[3]To avoid confusion, a point should be made pertaining to the terminology in the field. In the calculations of Pople and co-workers (29, 38, 50) emphasis is placed on the SCF perturbation theory of Møller & Plesset, coupled with the specification of the reference function, rather than on the later MBPT developments employing the diagrammatic theory. This leads to the specification of their models as, for example, UMP3, for third-order Møller-Plesset perturbation theory with unrestricted Hartree-Fock reference function, and UMP4DQ, for fourth-order MB perturbation theory limited to effects of double and quadruple excitations. In our system, *UMP3 = D-MBPT*(3) and *UMP4DQ = DQ-MBPT*(4), where the reference function is understood to be UHF for an open-shell case and RHF for a closed-shell case. In most cases MP denominators are used, although Epstein-Nesbet denominators, which are often referred to as "shifted," have also been employed in some studies (52).

Arnau-Huzinaga (69) and Silverstone-Yin (70) approach provides V^{N-1} modified orbitals, which may be written as a transformation exclusively among the usual virtual SCF orbitals. For this category of transformations, when only excited orbitals are mixed among themselves (and/or occupied orbitals among themselves), the CCD equations are invariant, so again the converged results will be the same, but not the various orders in perturbation theory. The CCD equations are not invariant to more general transformations that mix occupied and excited orbitals together, like natural orbits or Brueckner orbitals, but the full CI or a full coupled-cluster result would be invariant. Insofar as a given model begins to approach the full result, invariance to even very general transformations will begin to be observed.

A third degree of flexibility is the reference function itself. An SCF function is often convenient, but for some applications just a product of bond-functions, as is used in the PCILO (perturbed CI with localized orbitals) semi-empirical perturbation method (72), or even a product of symmetrically orthogonalized atomic orbitals, may be preferable. For example, the latter two reference functions have certain localization characteristics that may be exploited in an extended system like a metal surface or a large biochemical molecule.

The enormous number of possibilities raises the question of what is the best MBPT model. It would be useful to know, for example, what choice of denominators, orbitals, or reference functions would give second-order results that are consistently closest to the basis set limit for the correlation energy or, perhaps, even to experiment.

Little work in this area has been done. Papers by Bartlett and coworkers (52, 71, 73) have considered the order-by-order perturbation convergence of the MP and EN partitionings relative to SCF orbitals, with the former found to offer much better convergence than the latter for most systems, since the pair-like terms included in the EN partitioning to all orders tend to be biased toward negative contributions. Similar results have also been found in the direct CI procedures for the solution of the CI eigenvalue problem (74). For the case of Be or LiH, and similar systems in which relatively separated electron pairs are involved, the pair-like terms included in EN do heighten the convergence, making E_2^{EN} a superior approximation compared to E_2^{MP}, but for more than four electrons this is seldom the case. [A potential exception would be for localized orbitals (41).] This is due to the pair-pair interactions, as emphasized by Micha (75) and Barr & Davidson (76), that tend to be positive and large for more complicated molecules. In related work Paldus studied the convergence of the CCD equations for Be as a continuous function of a parameter weighting the denominator from MP to EN, finding the best

convergence for a point closer to the EN denominator (77). In all studies of convergence Padé approximants are routinely used to enhance the speed of the convergence (52). The lowest $[N, N-1]$ approximant is the well-known geometric approximation.

Denominator shifts employing similar "conjoint" (33) (or less appropriately, EPV, for exclusion principle violating) terms that come from the quadratic part of the CCD equations were originally used by Kelly (5). This technique has been found to be important by Freeman & Karplus (78) in obtaining potential curves for diatomic molecules. Prime & Robb discuss a related, but more general type of denominator shift (79), building upon a development by W. J. Taylor (80). An important objection to denominator shifts is that the inclusion of "diagonal" elements of higher order diagrams can result in expressions at a given order that would not be invariant to transformations among degenerate orbitals (81) and may, in fact, give the wrong size-dependence (41).

Different types of basis functions may also be considered. Adamowicz & Sadlej (82) and Pan & King (83) have investigated the use of explicitly correlated Gaussian geminal basis functions in perturbation theory and electron pair theory.

Some work by Silver, Bartlett & Wilson (84) studied the convergence of lower-order perturbation theory with V^{N-p} orbitals, for various choices of p. Although second-order results could be quite varied for different potentials, by third-order there is already little difference in the net results in these studies. This problem has also been studied by Hiroike (85). To the contrary, Lindgren et al found important improvements when using Brueckner orbitals instead of Hartree-Fock orbitals in studies of hyperfine structures in alkali atoms (86). This might be expected due to the importance of single excitation effects for such properties.

The most drastic modification of the theory described above is the generalization to multiple reference functions. Several frameworks for the multireference MBPT/CCM problem have been proposed (87–92), although few applications have yet been made. For many molecular problems it is apparent that multireference techniques will be preferred. This area is discussed in the final section.

However, unlike traditional perturbation methods, in which the perturbation is expected to be small, it should be recognized that single reference MBPT/CCM has been developed from an inherently "infinite-order" perturbation theory viewpoint. [This terminology derives from Löwdin (93).] This means that via the infinite-order CCM models, or by using techniques like denominator shifts to employ geometric series arguments to sum classes (or components) of diagrams to all orders, conver-

gence can often be accomplished even for highly difficult cases. A prime example is found in H_2 at large separation, where the $1\sigma_g^2$ and $1\sigma_u^2$ configurations are equally important, and logic would recommend that both functions be included in the reference space. Despite this, a single reference CCD calculation based on the $1\sigma_g^2$ configuration will give a good potential curve all the way to the dissociation limit (94). Similar studies on Li_2 and N_2 using denominator modifications also illustrate this point (78).

An intermediate level between single-reference and multireference functions in MBPT has been considered by Kirtman & Cole (95). In this study they have proposed that a valence bond (VB) function should be the reference function for a perturbation approach. If this could be accomplished conveniently, one would have the advantage that, unlike an RHF function, the valence-bond solutions would frequently separate correctly, thus leading to a more accurate zeroth-order approximation to a potential energy curve. A UHF function will normally separate correctly, but as illustrated with the N_2 example (22) (described below), the spin contamination becomes too great to give a realistic potential curve in the intermediate range between equilibrium and the separated atom limit. Obviously, when the zeroth-order solution is superior, then equivalent corrections should be obtained in lower orders of perturbation theory.

There are several difficulties with the VB approach, however. The nonorthogonalities that are involved in the original VB theory make the computation time rise as $N!$ for N electrons. Hence, it is usually necessary to invoke strong orthogonality conditions in order to obtain the solution (96, 97). However even with such approximations, the different VB orbitals are eigenfunctions of different one-electron Hamiltonians, eliminating the convenient N-electron Hamiltonian of SCF theory. This feature is also true for Hartree theory and most open-shell RHF-SCF theories, although recent work suggests that a convenient, formal solution to this problem may be found (98). This is an important question, since the elimination of unlinked diagrams is the basis for the size-extensivity of the MBPT/CCM methods, and this elimination is facilitated by the separability of the N-particle Hamiltonian. Kirtman & Cole resort to a type of exchange perturbation theory to account for the different one-electron Hamiltonians. Second-order results are reported for H_2 and LiH.

Although it is important to develop and apply the multireference techniques for open-shell problems, for problems involving near degeneracies, and for the accurate description of bond breaking, the point at which efficiency versus accuracy considerations favor the multireference approach over the single-reference method, or over more general VB-based methods, has not yet been determined.

STUDIES OF MOLECULAR CORRELATION ENERGIES

The first molecular calculations using many-body methods were made by Kelly for H_2 (99) and by Miller & Kelly for H_2O (100), using one-center expansion techniques and numerical methods similar to those used for atoms. Lee, Dutta & Das made similar calculations (101), with recent work by McDowell (102). Although it is highly desirable to develop purely numerical methods for molecules to avoid the limitations inherent in basis set methods, the multicenter nature of molecular charge distributions presently requires the use of conventional finite basis sets of Slater type orbitals (STO) or of contracted Gaussian type orbitals (CGTO). In an interesting series of papers (103), McDowell has discussed the systematic elimination of the basis set error by using MBPT techniques.

Schulman & Kaufman (104) used the formal structure of MBPT and finite basis sets in calculating the second-order correlation energy and polarizability for H_2. These authors also investigated sum rules to assess the accuracy of their basis sets, an idea that should be used more frequently. This was followed by similar studies of molecular correlation energies, but with some consideration of higher-order effects, by Robb (105), Bartlett & Silver (71), and Freeman & Karplus (78). In addition, Kaldor demonstrated the accuracy of finite-basis MBPT calculations by comparing them with Kelly's numerical results for Be (106).

Additional studies without approximation at the full third-order level have been made by Urban, Kellö & Hubač (107), Bartlett et al (73), Pople et al (29), Kvasnicka & Laurinc (108), and Wilson & Silver (109). It should also be noted that when an SCF reference function is used, and until terms beyond the third order in perturbation theory are included, the first two interations of SCF-based direct CI calculations (110) are equivalent to D-MBPT(3). However, in the process of converging to the D-CI solution, the size-extensivity of the D-MBPT(3) model is lost because the truncated CI approach incorporates unlinked diagram terms in fourth and higher orders of perturbation theory (30). This results in the paradoxical situation that a second- or third-order perturbation result for the correlation effects is likely to be superior to the converged D-CI or SD-CI model, particularly for extended systems, and often even better for some properties of smaller molecules (33). Approximating these unlinked diagram terms is the basis for the widely used Davidson's (111) approximation for quadruple CI excitations (112–114).[4]

[4]Actually, two alternative viewpoints on how to approximate the quadruple and higher excitations in CI have been taken. Bartlett & Shavitt (112) proposed the viewpoint that the unlinked diagrams in fourth-order should be approximated and thereby eliminated from the

In subsequent work, Bartlett & Shavitt computed fourth- and higher-order MBPT contributions due to double excitations (52), and Krishnan & Pople (115), Bartlett & Purvis (30), and Pople et al (50) developed general purpose programs for the computation of fourth-order quadruple- and single-excitation terms. The latter two papers also report CCD results for a series of molecules. The triple-excitation diagrams have now been computed by Krishnan et al (116), Kvasnicka et al (117), and Wilson and co-workers (118, 119). At the level of quadruple and triple excitations, MBPT/CCM offers a great deal of previously unattainable information about the correlation problem for nontrivial molecules.

In a substantial achievement, Saxe, Schaefer & Handy have recently carried out an all-electron full CI calculation for H_2O in a DZ (double zeta) basis set (120). This calculation involved 256,473 configurations and required about six hours on a CDC 7600 computer (H. F. Schaefer, private communication). Since full CI is the best possible solution in the basis set, it is highly informative to compare MBPT/CCM predictions of the correlation energy with the CI results. Table 1 provides this information. Second-order perturbation theory, which is the simplest MBPT approximation, provides 94% of the full CI correlation energy, with the full fourth-order perturbation theory accounting for all but 1 mh, or 99.3% of the correlation energy. The remainder of the full CI result is associated with contributions from higher-order perturbation theory, but most of these are included in the infinite order CCD result.

Since CCD reduces to DQ-MBPT(4) in fourth-order, the difference between the results of these two models provides a measure of the higher order contributions generated by the T_2 operator, which correspond to higher order effects due to even excitations. This difference is 0.668 mh. Adding the fourth-order contribution of single and triple excitations [the latter computed by Wilson & Guest (119)] to the CCD result gives agreement with the full CI to 0.3 mh, or 99.8% of the full CI, and agreement with SDTQ-CI (single, double, triple, and quadruple excitation CI) to 99.96% or 0.06 mh.

When the very good agreement between CCD plus the fourth-order single and triple excitation terms [CCD + ST(4)] is combined with the

CI, and thereby derived Davidson's formula. This was generalized to all orders by Siegbahn (113). This approach is general for any system, although no effort was made to separate the "conjoint" components that remain in the linked diagrams from the "disjoint" terms (33). Other authors, Pople et al (38), Davidson & Silver (39), and Brändas et al (43) have obtained approximations based upon detailed considerations of a model problem like the H_2 lattice discussed in the third section. This approach would seem to be somewhat dependent on an idealized system. Luken (114) has also studied this problem from the viewpoint of Sinanoglu's electron pair theory.

Table 1 Comparison of many-body results with full CI[a] (H_2O 14 CGTO-DZ basis set, $E_{SCF} = -76.00984$)

Model	Configurations	Correlation energy (Hartree a.u.)	ΔE(full CI) (kcal/mole)	ΔE(SDTQ-CI) (kcal/mole)
SD-CI[b]	361	−0.14018	4.9	4.8
SDTQ-CI[b]	17,678	−0.14777	0.2	0.0
FULL CI[b]	256,743	−0.14803	0.0	−0.2
D-MBPT(2)		−0.13948	5.4	5.2
D-MBPT(3)		−0.14087	4.5	4.3
D-MBPT(4)		−0.14392	2.6	2.4
DQ-MBPT(4)		−0.14476	2.1	1.9
SDQ-MBPT(4)		−0.14565	1.5	1.3
SDTQ-MBPT(4)[c]		−0.14704	0.6	0.5
CCD		−0.14544	1.6	1.5
CCD + ST(4)		−0.14771	0.2	0.0

[a]MBPT/CCM calculations, R. J. Bartlett.
[b]The CI calculations are by Saxe, Schaefer & Handy (120).
[c]The triple excitation component of SDTQ-MBPT(4) is computed by Wilson & Guest (119).

fact that these MBPT/CCM calculations require only a few seconds on the CDC 7600, compared to six hours for the full CI or a few minutes for the SDTQ-CI, the potential efficiency and accuracy of the many-body methods is emphasized. Of course, H_2O at equilibrium is well-described by a closed-shell SCF reference function, which is clearly the dominant configuration, so MBPT/CCM methods based upon a single reference function are easily applicable. With open shells or near degeneracies, even a single-reference UHF-based MBPT/CCM approach may not be as reliable (30, 120*a*), and multireference function techniques are sometimes preferred, with a consequent loss in efficiency.

In another calculation on H_2O (33), a very good 39 STO basis set is used (121). For this basis the full CI or even SDTQ-CI is out of the question. However, one can still readily carry out SDQ-MBPT (4) and CCD calculations. The inclusion of fourth-order triple excitations, an n^7 problem, adds a great deal to the time for the calculation (118), but their magnitude in the examples studied is typically of the same order as the other fourth-order components. A suggested rule-of-thumb is that the triple excitation terms are about three times as large as the singles contribution (122). These results are displayed in Table 2. The CCD model with the addition of the fourth-order single and triple excitations is essen-

Table 2 Comparison of many-body results with SD-CI[a] (H_2O 39 STO basis set, $E_{SCF} = -76.06423$)

Model	Correlation energy (Hartree a.u.)	ΔE(SD-CI) (kcal/mole)
SD-CI[b] (4,120 configurations)	−0.27558	0
D-MBPT(2)	−0.28178	−3.9
D-MBPT(3)	−0.28502	−5.9
D-MBPT(4)	−0.28934	−8.6
DQ-MBPT(4)	−0.28614	−6.6
CCD	−0.28621	−6.7
SDQ-MBPT(4)	−0.28817	−7.9
SDTQ-MBPT(4)[c]	−0.29604	−12.8
CCD + ST(4)	−0.29611	−12.9
Exp.	−0.370	−59.2

[a]The MBPT/CCD calculations are reported by Bartlett, Shavitt & Purvis (33).
[b]The CI calculation is by Rosenberg & Shavitt (121).
[c]The triple excitation component is computed by Wilson (118).

tially equivalent to SDTQ-MBPT (4) for this system, and accounts for 80% of the estimated total correlation energy of H_2O. A second-order calculation recovers 95% of this, third-order recovers an additional 1.1%, and fourth-order 3.7%. The fourth-order contributions are −4.3 mh for double excitation diagrams, −2 mh for singles, −7.9 mh for triples, and +3.2 mh for quadruple excitation diagrams.

As measured by fourth-order results, the net effect of quadruple excitations in CI would be about −13.9 mh,[5] or about 4.7% of the computed correlation energy (33). Adding in the triples as well, the higher CI excitations account for 21.8 mh or about 7.4% of the correlation energy, compared to 5.1% in the DZ basis calculation.

In this study of H_2O a quartic force field was also predicted at several levels of MBPT/CCM approximation, and compared with SCF and SD-CI (33, 125). The SCF predictions of the bond length and angle are within 2% of experiment, while SD-CI is accurate to about 0.5%. The low-order many-body models D-MBPT(2) and D-MBPT(3) give some-

[5]It is important to distinguish between quadruple excitation diagrams and CI quadruple excitations, since the former contains components which derive from double excitations in the CI model (30, 33). These double-excitation terms are responsible for the quadruple excitation diagrams being positive. This also introduces the separation of the renormalization terms in fourth-order perturbation theory into the "conjoint" (or EPV) and "disjoint" parts mentioned in this review. Hubač and co-workers (123, 124) present a detailed study of these interrelationships.

what better agreement with experiment than SD-CI. The error in CCD is <0.2%, and SDQ-MBPT(4) improves this to <0.1%.

In the prediction of the force constants, the SCF results show errors of more than 30% for even some quadratic constants, while SD-CI is usually fairly close to the various MBPT/CCM results, normally <10% in error. For f_{RR}, though, there is about a 16% difference between SD-CI and SDQ-MBPT(4), with the CI result much further from experiment. There are differences between various reported experimentally derived values of the force constants, and to a lesser degree even for the bond length and bond angle, so convergence to a particular set of experimental values cannot be taken as evidence for any general superiority of MBPT/CCM to SD-CI, although one would certainly expect that the higher excitation terms like the quadruples included in SDQ-MBPT(4) and CCD should help in improving the accuracy of the SD-CI calculations. In fact, adding Davidson's approximation for quadruple excitations to SD-CI significantly improves the results of the model in this example (125), and this is a generally observed phenomenon for highly accurate CI studies (126). What is most important here, however, is that size-inextensive models show some significant differences from SD-CI even for a small molecule like H_2O. We would certainly expect this to be the case for larger systems, but even for H_2O, at the sophistication of current ab initio quantum chemistry, there are observable consequences of size-inextensive models that should be realized and corrected.

The fractions of the correlation energy attained within a given basis set with the MP denominator and ordinary SCF occupied and excited orbitals are presented in several places (22, 29, 30, 33, 50). The typical behavior is illustrated in Table 3 for some molecules we have studied. The higher than fourth-order terms are measured by means of the CCD model, which only includes the T_2 operator in Eq. 2, hence single- and triple-excitation contributions are not included.

The worst case for second-order perturbation theory is BH_3, which reflects the residual degeneracy in this system. However, even though the third-order contribution is comparatively large, the higher-order effects are modest, showing good convergence. The fourth-order single-excitation contribution to BH_3 amounts to only -0.18 mh (30). Diborane shows a somewhat similar behavior, and the single excitations give only -0.14 mh (127). The multiply-bonded molecules CO, CO_2, N_2, HCN, and CH_3CN have a large E_2, with a positive E_3, except for HCN. In a larger basis including double-polarization functions, E_3 for HCN is positive (128). The negative single-excitation contribution for each of these molecules is somewhat larger, being -7, -12, -5, -4, and -5 mh, respectively (30, 22).

Table 3 Percentage of the correlation energy obtained by different orders of perturbation theory[a]

Molecule	Second-order	Third-order	Fourth-order (DQ)	Higher order (DQ)
BH_3	80.0	16.5	3.0	0.50
H_2O	97.7	1.5	0.7	0.06
NH_3	94.3	5.0	0.6	0.12
CH_4	89.6	9.3	0.9	0.16
CO	100	−1.6	1.6	−0.09
CO_2	103.2	−4.1	0.9	0.0
HCN	98.0	0.7	1.0	0.18
N_2	101.0	−2.2	1.3	−0.11
HNO	98.9	−0.5	1.9	−0.26
HCO	99.5	−0.2	0.7	−0.05
CH_2O	97.7	0.1	2.9	−0.01
C_6H_6	95.6	0.6	3.8	—
$(CH_3)_2NNO_2$	94.1	4.4	1.5	—
CH_3CN	96.5	−2.2	3.7	—
B_2H_6	85.2	13.2	1.6	—

[a]The basis set is at least of double zeta plus polarization quality, except for C_6H_6 and $(CH_3)_2NNO_2$, where a double zeta basis is used.

The fourth-order DQ contribution is comparable in size to E_3, although it can be larger. The usual justification for this is that effects of quadruple excitations are included here for the first time, although it is too much to expect that the perturbation series will be monotonically decreasing. When the DQ terms derived from T_2 are included to all orders, as is done in the CCD model, one still generally finds little differences between CCD and DQ-MBPT(4) as shown in Table 3 (22), implying that even though the fourth-order contribution can be larger than E_3, there is really no problem with convergence through DQ-MBPT(4) for most cases. When the single-reference function MBPT/CCM method is used for problems where near degeneracies are encountered, convergence can be far worse (30, 129).

Even though the fourth-order DQ contribution can be larger than E_3, once fourth-order single and triple excitations are included, the magnitude of the total fourth-order term is even greater, since both these new contributions are negative, thus augmenting the negative DQ component. In the SCF case, due to Brillouin's theorem, this is the lowest order in which these terms can appear, accounting for their significant effect. Since there are only a few examples where higher-order T_1 contributions

are considered (49, 129), and only one example that includes T_3 to higher order (49), less justification for adequate convergence of the SDTQ-MBPT(4) model is available, although it is anticipated that this level should also be well converged.

In addition to the applications made by Battelle's group, very thorough studies of molecular correlation energies are presented by Pople and co-workers in a series of papers (29, 38, 50, 115, 116, 122). In the first major effort (29), Pople et al reported on calculations at the level of second- and third-order perturbation theory for a variety of atoms and hydride molecules, including some considerations of multiplet splittings, dissociation energies, and second-order predictions of bond lengths and angles. For geometries, these authors find that D-MBPT(2) shows a mean difference between theory and experiment of only 0.003 Å, compared to 0.01 Å at the UHF level, while bond angles are accurate to a few degrees. As seen in Table 3, E_2 is generally sufficient to account for ~90% of the basis set correlation energy and Pople et al show that it also removes at least 50% of the error remaining in the UHF predictions of geometries.

In a subsequent paper (38), Pople and co-workers compared D-MBPT (3) with variational D-CI and SD-CI predictions of correlation energies, dissociation energies, and multiplet separations for the same series of atoms and molecules. At the third-order level there is not much difference between MBPT and the CI results for small molecules, as one would expect from the fact that D-MBPT(3) corresponds to the initial iterations leading to the D-CI result. It is difficult to separate the effects of size-extensivity in D-MBPT(3) versus D-CI from the higher-order contributions included in D-CI, but this paper also considers a size-extensivity correction to D-CI that suggests that there is about a 2 kcal/mole effect in multiplet splittings and dissociation energies for the simple systems studied.

The potentially more significant differences between MBPT/CCM and SD-CI models start to appear at the fourth order of perturbation theory. As described above, the inclusion of most of the effects of CI quadruple excitations in MBPT/CCM via the factorizable T_2^2 term enables MBPT/CCM to include such higher-excitation effects comparatively easily, while only for small model problems is it possible to do CI calculations that include the full set of quadruple excitations, as in the work of Saxe et al (120). In fact, a very large number of MBPT/CCM calculations that include quadruple excitation effects have been made for rather complicated systems, and some of the results are shown in Table 3. The general size of the error due to neglect of CI quadruple excitations is illustrated in Figure 1 for a few examples, and ranges up to 20% for benzene.

A paper by Krishnan & Pople (115) reports SDQ-MBPT(4) results for

the same series of molecules studied in their previous papers on SD-CI and second- and third-order perturbation theory. These authors point out that this model, neglecting triple excitation terms, is correct through fourth-order for an assembly of isolated two-electron systems.

Bartlett & Purvis (30) discussed the fourth-order theory and its relationship to CCD, with emphasis on the cancellations in the fourth-order theory that distinguish many-body models from truncated CI. A number of results in DZP or better basis sets for BH_3, NH_3, CO, HCN, CO_2, and N_2 are presented at the SDQ-MBPT(4) and CCD levels, showing that DQ-MBPT(4) results are quite close to the CCD values. This paper also points out the convergence problems encountered when one attempts to use an RHF reference function in an MBPT calculation of the N_2 potential curve, just beyond the N_2 bifurcation into an RHF and a UHF solution. The problem is further considered in another paper (22).

Pople et al (50) reported CCD results and compared these with the linearized (L-CCD) model and DQ-MBPT(4) for their usual set of first-row atoms and hydride molecules. The L-CCD model, which neglects the nonlinear terms in Eq. 7, corresponds to the sum of just the double-excitation MBPT diagrams to all orders, or D-MBPT(∞) (30), and is also known as CEPA(0) (18). Since the nonlinear terms are generally positive, L-CCD results overestimate the CCD correlation energies by as much as 6 mh for some of the molecules studied. However, since the neglected fourth-order single- and triple-excitation diagrams are negative, the errors in L-CCD compensate to some extent for the omission of these terms. These authors also observe the close coincidence of CCD with DQ-MBPT(4) in their applications.

Nakatsuji has applied his symmetry-adapted cluster theories to Be, BH_3, and H_2O (59), including some study of the excited states.

Krishnan (130) et al defined a new 6-311G** basis set and predicted geometries and atomization energies for a series of small molecules at the D-MBPT(3), DQ-MBPT(4), and SDQ-MBPT(4) levels. Agreement with experimental bond lengths and angles is not substantially improved over the simple D-MBPT(2) predictions for most examples. In particular, SDQ-MBPT(4) tends to increase the bond lengths between first-row atoms due to the effects of the single excitations. The atomization energies at the SDQ-MBPT(4) level are within 5 to 13% of the experimental values. The effect of single substitutions are as large as 3–4 kcal/mole in multiple-bonded systems, but much smaller for hydride molecules.

A large group of molecules with single, double, and triple bonds have been studied with SDTQ-MBPT(4) to assess the effect of the triple-excitation terms (122). These calculations use a 6-31G* basis which, unlike the DZP results in Table 3, do not have polarization functions on the

H atoms. However, the general behavior is quite similar. Using the correlation energy through fourth-order as the reference point, the second-order energy accounts for a low of 79.2% for BH_3 and 82.5% for B_2H_6, to a high of 96.6% for HF and 95.9% for F_2. Of the twenty-six molecules studied, all but five exceed 90% recovery in second-order. The third-order results have a maximum of 15.9% for BH_3 and 12.8% for B_2H_6, as well as a surprisingly high 10.0% for CH_4, but are less than 10% for all the other molecules. The fourth-order contribution varies from about 2 to 7%. The maximum is given by HCN, CO, N_2, all triple-bonded molecules. As pointed out by Frisch et al (122), the heat of formation of NH_3 from N_2 and H_2 has a contribution of 5.5 kcal/mole solely from triple excitation terms.

Wilson & Saunders (118), have also studied the contribution of triple-excitation diagrams to molecular correlation energies. In applications to Ne, it is found that these terms account for -1.1 mh in the largest basis studied. In H_2O, using the same basis set that had previously been used by Bartlett et al in their MBPT/CCD study for the H_2O quartic force field, the contribution of triple excitations is -7.9 mh. (It should be noted that the original papers had errors that suggested much larger contributions of -9 and -21 mh for the triple excitations in Ne and H_2O, respectively.) Other calculations by Guest & Wilson (131) emphasized that the triple-excitation terms are largest for multiple-bonded systems, including N_2, CO, SiO, and SiS. In addition, it is found that E_4^T changes from -9 mh to -34 mh in a range of internuclear separations for N_2 compared to 4.5 mh to 8.4 mh for E_4^Q. This dramatic change is probably partially due to the instability in MBPT treatments of N_2 based upon an RHF reference function at internuclear distances past equilibrium, as discussed previously (22, 30) and in the last section.

CEPA models (18, 26) are basically a modification of SD-CI, which accounts in an approximate way for higher excitations. Ahlrichs (66) has discussed in detail the relationships between CEPA with MBPT and CCD. CEPA models may be derived by making different approximations for the nonlinear term in the CCD equation, which corresponds to the fourth-order quadruple excitation diagrams of MBPT. These models have been extensively applied (18). With the recent SDQ-MBPT(4) and CCD results becoming available, it is now possible to begin to assess the accuracy of the different CEPA models for inclusion of such higher excitation terms. Ahlrichs & Zirz (132) offer a series of pertinent comparisons for the correlation energy. CEPA(1) agrees with the rigorous fourth-order effect of quadruple excitations to within 0.6% for a series of molecules. CEPA(2) usually overshoots by 1–2% and CEPA(0) by 1–3%. However,

since the fourth-order triple-excitation diagrams are negative, the latter two models may actually be closer to reality in many cases.

DISSOCIATION ENERGIES, POTENTIAL ENERGY SURFACES, AND PROPERTIES OTHER THAN THE ENERGY

In addition to papers devoted to the theory or to studies of the various components of the correlation energy, MBPT/CCM models are now being used routinely to investigate a variety of chemically interesting questions. Most of the more recent applications involve at least some fourth-order MBPT contributions, and often even CCD as in the quartic force field for H_2O (33).

Considering polyatomic systems first, in a series of studies (128, 133), the isomerization energies for the reactions $CH_3NC \rightarrow CH_3CN$, $HNC \rightarrow HCN$, $LiNC \rightarrow LiCN$, and $BNC \rightarrow BCN$ were investigated at the SDQ-MBPT(4) level as well as the activation barriers for the first two isomers. The isomerization energy and activation barrier in the methylisocyanide system are in good agreement with experiment. For the HNC→HCN isomerization an unpublished experimental value (134) is 10 ± 1 kcal/mole, but these calculations, as well as other MBPT calculations of Pople et al (50) and CI calculations of Pearson et al (135), tend to support a value of 15 ± 2 kcal/mole for this isomerization. The theoretical results seem to be vindicated by a recent, as yet unpublished ion cyclotron resonance experiment of Pau & Hehre, who report 14.8 ± 2 kcal/mole (W. J. Hehre, private communication). Thorough studies of the HCN and HNC potential surfaces near equilibrium have also been made by Taylor et al (51) using CCD and various CEPA models.

CCD and SDQ-MBPT(4) applications have been made by Adams et al (136) in a study of the stepwise decomposition of methyl alcohol, $CH_3OH \rightarrow CH_3O \rightarrow CH_2O \rightarrow CHO \rightarrow H + CO$. This paper reports a series of dissociation energies and predicted geometries for these molecules. The geometries for the known species are typically accurate to <.01 Å for bond lengths and to within a couple of degrees for angles. This work provides a prediction for the experimentally unknown structure of methoxy.

The binding energies for the borane-containing molecules, B_2H_6, H_3BCO, and H_3BNH_3 have been of substantial interest to chemists for some time. For a long period, different experiments obtained different binding energies for the first two, while the third has yet to be experimentally obtained. Redmon et al (127) studied these molecules with MBPT methods, obtaining exceptional agreement with the now accepted experi-

mental values and making a prediction of the binding energy of borazane. Good agreement with previous CEPA calculations is also observed (35).

In more complete studies of potential surfaces, a reaction path for the unimolecular decomposition of $HCO \rightarrow H + CO$ has been determined (137, 138) and used to provide a rate constant. Adams et al (138) have also provided reaction paths for the three lowest states of the HNO radical. The SDQ-MBPT(4) model was used to provide a global potential energy surface for the $O(^3P) + H_2O$ collision in order to predict vibrational excitation cross sections (139). The He + LiH surface has been studied at the D-MBPT(3) level (140).

The reaction path for formaldehyde offers a particularly interesting study. Goddard & Schaefer (141), using SD-CI techniques and Davidson's correction to estimate the effects of quadruple excitations, determined the barriers and dissociation energies for the molecular products $H_2 + CO$, the radical products $H + HCO$, and for the rearrangement to hydroxycarbene. The results of the CI calculations suggest an alternative interpretation (142) of the photodissociation experiments of Houston & Moore (143) for the $H_2CO \rightarrow H_2 + CO$ route. This system has since been studied with SDQ-MBPT(4) by Harding et al (144) as part of their H_4CO surface, and by Adams et al (145) at the SDQ-MBPT(4) and CCD level. Although the MBPT/CCM calculations show somewhat better predictions of dissociation energies, in essentials (and with the correct zero-point energy for the hydroxycarbene transition state) the predicted barriers support the CI results of Goddard & Schaefer. However, more recent SDTQ-MBPT(4) results of Frisch et al (145a) obtain a substantial ~5 kcal/mole reduction in the activation barrier for molecular product dissociation due to a change in basis set and another ~3–4 kcal/mole once triple excitation contributions are included, bringing the calculations into agreement with experiment.

Table 4 presents a summary of some of the results obtained in these efforts, compared with SCF, second-order perturbation theory, and experiment. Second-order perturbation theory removes most of the error in the SCF results for these dissociation and isomerization energies, which is an encouraging result for such a simple addition to an SCF calculation.

In addition to their thorough study of the H_4CO surface, Harding et al have investigated the unimolecular decomposition of methanol (146). Pople et al (50) have also studied the 1,2 hydrogen shifts in C_2H_2, HCN, CH_2O, and N_2H_2 at the SDQ-MBPT(4) level, finding *cis* and *trans* forms of HCOH and HNNH that differ by about 6 kcal/mole. A thorough D-MBPT(3) study of the isomers formed by 1,2 and 1,3 intramolecular hydrogen shifts in CH_3–NO and their associated transition states was reported by Adeney et al (147).

Table 4 Comparison of thermochemistry results obtained by SCF and MBPT/CCD with experiment (All basis sets are at least DZP quality.)

		$-\Delta E$(kcal/mole)				
				MBPT/CCD		
Reaction	Ref.	SCF	E_2	Results	Model	Experiment
$2\ BH_3 \rightarrow B_2H_6$	(127)	18.5	37.5	35.6	SDQ-MBPT(4)	36.6 ± 2[d]
$BH_3 + CO \rightarrow H_3BCO$	(127)	8.0	25.1	20.5	D-MBPT(4)	20.4 ± 2[d]
$BH_3 + NH_3 \rightarrow H_3BNH_3$	(127)	20.5	32.0	30.1	D-MBPT(4)	—
$HNC \rightarrow HCN$	(128)	10.7	18.0	15.0	SDQ-MBPT(4)	(10.3 ± 1)[e]
$HNC \rightarrow [H_2C^{\cdot}]$[a]	(128)	−33.4	−30.1	−29.5	SDQ-MBPT(4)	—
$BNC \rightarrow BCN$	(128)	−18.9	−9.8	−12.4	SDQ-MBPT(4)	—
$LiNC \rightarrow LiCN$	(128)	−6.4	−2.3	−3.9	SDQ-MBPT(4)	—
$CH_3NC \rightarrow CH_3CN$	(133)	19.2	26.2	22.8	SDQ-MBPT(4)	23.7 ± .14[f]
$CH_3NC \rightarrow [CH_2^{\cdot}]$[b]	(133)	−44	−40	−40	SDQ-MBPT(4)	−38.4[g]
$H + CO \rightarrow HCO$	(137)	4.8	11.8	13.6	CCD	15.7 ± 1.5[h]
$HCO \rightarrow [HCO]$[c]	(137)	−12.8	−17.4	−18.1	CCD	—
$H_2CO \rightarrow H_2 + CO$	(145)	−7.2	−3.9	−3.9	CCD	−1.9[i]
$H_2CO \rightarrow H + HCO$	(145)	−68.6	−82.8	−86.0	CCD	−86.0 ± 1.0[i]

[a]Square bracket indicates a transition state. This result includes a 4 kcal/mole zero point correction for the transition state.
[b]Square bracket indicates a transition state. This result includes a 4.8 kcal/mole zero point correction for the transition state.
[c]This result includes a 5 kcal/mole zero point correction for the transition state.
[d]Ref. (145b).
[e]Ref. (134). Ref. (128) concludes that this experimental value is in error. The result should be 15 ± 2 kcal/mole.
[f]Ref. (145c).
[g]Ref. (145d).
[h]Ref. (145e).
[i]Ref. (145f).
[j]Ref. (145g).

Kenney et al (148) used high-order D-MBPT to study the singlet-triplet separation in the series of compounds H_2C:, H_2CC:, H_2CCC:, predicting a singlet ground state for vinylidene and vinylidene carbene.

Because of its correct size-dependence, one of the natural places to apply MBPT/CCM is in the emerging area of ab initio quantum biochemistry (31). An example of this is the work of Weinstein et al (149) and Osman et al (150), who investigated the stacking of complexes of 5- and 6-hydroxytryptamine with imidazolium (represented by model compounds) to probe receptor sites for hallucinogenes. No doubt many more applications of this type will appear in the next decade.

Earlier studies from the Carnegie-Mellon group used low-order perturbation corrections to study the relative stability of the difluoroethylene isomers (151), the internal rotation of allene (152), the nature of the carbon beryllium bond in CH_2Be (153), and a series of highly unusual electron deficient carbon compounds that violate van't Hoff stereochemistry (154). De Frees et al (155) studied geometries of several complicated species like FOOF at the D-MBPT(2) level.

A number of diatomic potential energy curves have been studied with fourth-order MBPT methods and with CCD. These include Mg_2 (156, 157), Be_2 (30, 120a), and N_2 (22). These calculations offer information concerning the applicability of single-reference MBPT-CCD for entire potential curves. Later work by Chiles & Dykstra (158) studied He_2, Be_2, and Mg_2 at the CCD and CEPA levels.

In lower-order studies, Urban & Kellö used D-MBPT(3) for potential energy curves for BH, F_2, and N_2 in the vicinity of equilibrium, in order to determine spectroscopic constants and to compare with CI (159). The authors observe a substantial difference between SD-CI and D-MBPT(3) for F_2 and attribute this to the failure of SD-CI to be size-extensive. CEPA comparisons with D-MBPT(2) were also reported and found to be in very good agreement. Kellö et al have also studied the proton affinity of H_2O at the third-order level (160). Hubač & Urban have used D-MBPT(3) to obtain ionization potentials for Ne and H_2O (161). Additional third-order near-equilibrium potential curve applications to BH (162), BF, N_2, CO (163), and CH^+ (164) and full curves for He_2 (165) and Be_2 (166) have also appeared. D-MBPT(3) computations on He_2 with the EN denominator are found to have an erroneous behavior as previously observed and explained by Malrieu (41).

Except for the cases of He_2 and Be_2 and similar molecules, an RHF reference function cannot provide a potential curve of correct form all the way to separation, and the utility of single-determinant RHF as a basis for D-MBPT(3) for such problems is questionable. Either very high-order theories like CCM need to be used to attempt to overcome the erroneous behavior of RHF at large separations, or a UHF reference function should be employed, when suitable. Ultimately, multireference MBPT approaches should be developed for these categories of problems.

In studying properties other than potential curves, correlation corrections to the coupled Hartree-Fock (CHF) perturbation theory are of substantial interest. A paper by Caves & Karplus has analyzed the problem diagrammatically (167). Numerical MBPT results are now starting to become available.

Adamowicz & Sadlej (168) computed second-order correlation correc-

tions to the CHF polarizability of Be. These authors found a substantial correction of 20% due to this term, which is consistent with other work (169–171). This calculation employed electric field variant Gaussian bases, as proposed by Sadlej (172), which offer an interesting concept for such studies.

Bartlett & Purvis used finite-field methods with SDQ-MBPT(4) to obtain dipole moments, polarizabilities, and hyperpolarizabilities for the HF molecule (170). In that work particular attention was paid to the choice of basis set required to describe such properties, using numerical Hartree-Fock calculations (173) as a guide. In another effort, these authors applied this technique to H_2O (171), including an investigation of the hyperpolarizability as a function of bond stretching and bending. Correlation is found to have a very large effect on hyperpolarizabilities, implying that CHF perturbation methods cannot adequately treat this problem. Also, the hyperpolarizability tensor elements are quite sensitive to bond stretching.

Nuclear spin coupling constants in H_2 have been studied by Itagaki & Saika (174). Using a large Gaussian basis set and second-order correlation contributions, augmented by some additional terms up to fourth-order, these authors obtained a result within ~10% of the experimental value for the Fermi contact term.

In another study these authors determined the correlation energy and dipole polarizability for H_2 (175). This paper also discussed the decoupling of the energy denominators in MBPT, and employed these techniques to relate the polarizability diagrams of double-perturbation theory to the field dependent energy, as it would be employed in finite-field applications. They have also studied the electric field gradient in the HD molecule with MBPT, obtaining quite good agreement with other very extensive calculations (176).

D-MBPT(2) was used by Yoshioka & Jordan to obtain dipole moments, polarizabilities, and electron affinities for the highly polar LiF and BeO molecules (177). Using a large and flexible basis set, they found that D-MBPT(2) gave almost perfect agreement with experiment for the dipole moments of LiF. There are no experimental values for the other quantities.

The interesting work of Kelly & Carter concerning photoionization cross sections for atoms should also be mentioned (178), because of its implications for molecular studies.

Bent et al have investigated Jahn-Teller distortions in the methoxy radical, coupling MBPT methods with a clever treatment of the dynamic Jahn-Teller effect (179).

Many MBPT finite-field and related studies of electric and magnetic properties, NMR chemical shifts, and spin-spin coupling constants will doubtlessly appear soon in the literature.

MULTIREFERENCE SPACE MBPT METHODS

In order to be able to describe many processes at the level of accuracy required in current quantum chemistry applications, it frequently appears to be necessary to employ multireference determinants in MBPT/CCM. To illustrate the nature of the problem, Figure 3 shows potential energy curves for the ground state of N_2 (22). It is evident that RHF-based MBPT calculations follow the experimental curve accurately until about 2.6 a.u., where the erroneous separation of the RHF reference function can no longer be corrected by the D-MBPT(6) procedure. Using the same reference function, but treating correlation at the CCD level, which includes the effects of quadruple excitations, the applicability of the RHF + CCD model is extended to slightly longer bond lengths, where instabilities with solutions of the CCD equations begin to occur (129).

On the other hand, after the bifurcation of the SCF solution for N_2 into separate UHF and RHF results, it is also possible to use UHF + MBPT/CCM to obtain potential curves. Unlike the RHF-based models, the UHF solution separates correctly to two ^4S N atoms and, in fact, gives a dissociation energy which is only about 0.6 eV too small, but it is apparent that the path toward dissociation is in error. This is primarily due to a large amount of spin contamination for this singlet state. The N_2 UHF multiplicity along the curve is about 3.5, and correlation of the D-MBPT(6) model is unable to introduce a high enough level of correction to improve it. It should be possible to correct this problem partially by annihilating the principal (triplet) component of the spin contamination, although there are problems with this approach as discussed by Rossky & Karplus (180).

Another intriguing potential solution has its impetus in the observation that if the lowest of the various single-reference MBPT/CCD curves for N_2 could be connected together smoothly, then it would be possible to obtain a good potential curve solely from a single-reference function. The orbitals in the single reference function MBPT/CCD solution change from RHF to UHF, however. Consequently, this smooth connection could be accomplished systematically by using some variational or stationary principle to obtain the lowest correlated solution as a function of the orbitals. In other words, this would be an MBPT/CCM analogue of multiconfigurational SCF theory (181).

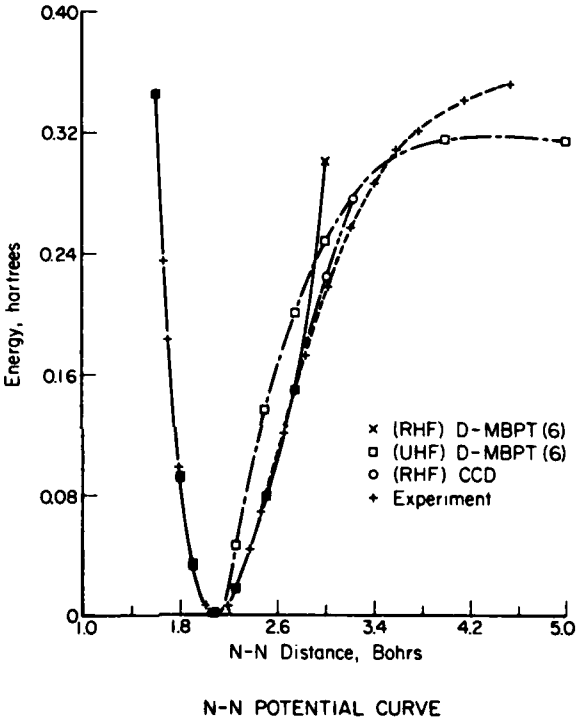


Figure 3 (UHF) and (RHF)D-MBPT(6) and (RHF)-CCD potential energy curves for N_2. The minima of the curves are superimposed. The D-MBPT(6) correlated UHF curves are higher in energy than the D-MBPT(6) RHF curves between $R = 2.0$ Bohr and $R = 2.7$ Bohr. The (RHF)-CCD result extends the reliability of the curve over the (RHF)-D-MBPT(6) approximation to somewhat larger R values, but ultimately the approximation becomes unstable.

However, the most universal solution to this type of problem is to employ multireference functions that, presumably, include within the reference space all important configurations for correct dissociation. Besides this bond breaking problem, similar difficulties, susceptible to the same approach, may be encountered with open-shell problems and with various treatments of some excited states.

The multireference analogue of the linked diagram theorem has been developed by Brandow (87), with other work by Mukherjee et al (89), Lindgren (88), Levy (91), Hose & Kaldor (90), and Kirtman (92). See also the related CCM work of Banerjee & Simons (182).

Although these developments can have important differences, the basic

structure is similar. The use of more than one reference function requires the consideration of an effective Hamiltonian matrix whose order is the number of reference functions $|\Phi\rangle = |\Phi_1^0\Phi_2^0 \ldots \Phi_m^0\rangle$. The effective Hamiltonian matrix is represented as a sum of matrices

$$\overline{\mathbf{H}} = \mathbf{H}_0 + \mathbf{H}_1 + \mathbf{H}_2 + \ldots \qquad 16.$$

that correspond to orders in the perturbation, whose diagonalization yields the energies and the "model" functions defined solely within the space of reference functions. This idea is well-known from textbook accounts of degenerate perturbation theory, but now generalized to the nondegenerate case.

One way the effective Hamiltonian is defined in many-body theory is to introduce a wave-operator, Ω, such that $\Omega\Phi_i^0 = \psi_i$, the correct eigenfunction of the Schrödinger equation. Ω is the same for each state so this is not the typical Brillouin-Wigner energy dependent wave operator (93). If one also defines a projector onto the reference space

$$P = \sum_{k=1}^{m} |\Phi_k^0\rangle \langle\Phi_k^0|,$$

it then follows that $P\psi_i = \psi_i^0$, where the $\{\psi_i^0\}$ are the "model" functions expressed solely within the reference space. Then with a slight manipulation of the Schrödinger equation, $H\psi_i = \psi_i E_i$, we obtain (88)

$$\overline{H}\psi_i^0 = \psi_i^0 E_i \qquad 17.$$

for $\overline{H} = PH\Omega P$.

From $\mathrm{PH\Omega P} = |\Phi\rangle\overline{\mathbf{H}}\langle\Phi|$ and $\psi_i^0 = |\Phi\rangle\mathbf{C}$, Eq. 16 is regained by expanding the operator Ω in powers of the perturbation. The solutions to Eq. 17 provide the exact energies and the model functions. The operator $\overline{H}$ is nonhermitian so the $\{\psi_i^0\}$ are not necessarily orthogonal, although Hermitian combinations can be constructed (183). From this point a series of equations for the Ω operator may be defined recursively from zeroth-order. The existence of a size-extensive connected diagram expansion emerges through the prescription for the different developments for the Ω operator, as presented elsewhere (87–91).

Through first order, the effective Hamiltonian matrix of Eq. 16 is Hermitian. Its eigenvalues and eigenvectors correspond to the CI solution in terms of the reference determinants $\{\Phi_\alpha^0\}$, $\alpha = 1$ to m. If this space were chosen to consist of the SCF determinant and all single excitations, the eigenvectors would correspond to the Tamm-Dancoff approximation (32) and the ground-state energy would simply be the SCF result. This usually provides a reasonable first approximation to the electronic excitation spectra. The remaining configurations start to contribute in second order. For the SCF plus all singles choice, double and triple excitations will

begin to mix in at second order, and up to pentuple excitations in third order (184).

On the other hand, if we were interested in the ground-state potential curve of H_2 we would want the two configurations $\Phi_1^0 = A(1\sigma_g^2)$ and $\Phi_2^0 = A(1\sigma_u^2)$ to be in the reference space. These configurations differ by a double excitation. Here the solution through first-order corresponds to the two-configuration CI result, while the additional configurations begin to contribute in second order. In the general (not two-electron) case, any configurations that differ from Φ_1^0 or Φ_2^0 by two or fewer excitations will contribute in second order, including double excitations from Φ_2^0 which are quadruple excitations relative to Φ_1^0. A full third-order treatment would involve hextuple excitations relative to Φ_1^0. The hope, of course, is that by using both important configurations in the reference space, the remaining effects can be treated adequately by low orders in perturbation theory.

The relatively high levels of excitation encountered in the multi-reference theory, even at low order, recommend the use of many-body, diagrammatic techniques that, in addition to facilitating size-extensivity, are more suited to handling higher categories of excitation than configuration based schemes. One also hopes that some of the partial summation techniques common to many-body and coupled-cluster theory (88) will be able to sum selectively many of the dominant, physically significant contributions to all orders.

One other comment that is pertinent to the Brandow-stimulated approaches to the multidimensional many-body theory is that the choice of reference space is not as arbitrary as one would like. In their diagrammatic development of the above equations, it turns out that only with specific choices of reference functions does a connected diagram theorem easily emerge. In particular, they require that if the configurations $A(\sigma_1^2)$, $A(\sigma_2^2)$, and $A(\sigma_3^2)$ are important in a problem, then the proper spin-coupled combinations of the configurations $A(\sigma_1\sigma_2)$, $A(\sigma_1\sigma_3)$, and $A(\sigma_2\sigma_3)$ must also be included in the reference space. Then, instead of nine matrix elements, one must compute four times as many.

The theory of Hose & Kaldor (185) permits the use of an arbitrary reference space at the cost of introducing a certain type of unlinked diagram which, however, does not destroy the size-extensivity of the model.

Another approach pursued by Kirtman (92) and discussed by Brandow (183), is the generalized Van Vleck transformation (GVVT) (186). Although it is ultimately similar to the wave operator approach above (183, 187, 188), the GVVT development generates a Hermitian effective Hamiltonian whose eigenvectors are consequently orthogonal. Full, rather than intermediate normalization is convenient. Furthermore, and potentially

most important, the GVVT method can be developed entirely within a Lie algebraic structure (34), which can ensure a properly size-extensive connected diagram description for general reference spaces.

A very limited number of applications of the multireference MBPT theory have so far been made. Kaldor applied Brandow's method to some excited states of H_2 (189), obtaining good results through third-order. Another paper by Stern & Kaldor studied states of BH (190), including their transition moments.

Lindgren's development using atomic numerical methods has been applied to open shell atoms by Morrison & Rajnak (191) and Morrison & Salomonson (192), while Salomonson, Lindgren & Mårtensson (193) have studied Be and C^{2+} as an example for cases of two important reference configurations, $1s^22s^2$ and $1s^22p^2$, for a closed-shell system. For Be, the second-order two-reference calculation recovers 93.6% of the correlation energy compared to 80.9% for second-order with only the $1s^22s^2$ configuration as reference.

The work of Hegarty & Robb based upon the Brandow approach should be mentioned (194). Also, the related CI perturbation method of Davidson & Bender for a multireference problem is pertinent (195). In the latter case, these authors used their method to study several states of Mg_2.

Hose & Kaldor applied their new general reference space approach to excited states of He_2 (185). These authors make the important point that the use of a complete model space of the type required in Brandow's theory spans a very broad energy range, while still leaving out other states within that energy range, which can cause severe convergence problems for perturbation theory [the problem of intruder states (183, 185)].

In a somewhat different vein, certain forms of the multireference function theory have also been frequently used in developing effective valence shell or pi-electron theories for molecules. Work of this type is reported by Brandow (183), Freed and co-workers (196), Westhaus and co-workers (197), Baker, Hegarty & Robb (198), and others (199, 200).

ACKNOWLEDGMENTS

The author would like to particularly thank I. Shavitt for his careful reading of the manuscript, several helpful suggestions, and assistance with locating numerous pertinent references.

Literature Cited

1. Freed, K. F. 1971. *Ann. Rev. Phys. Chem.* 22:313
2. Brueckner, K. A. 1955. *Phys. Rev.*, 97:1353; 100:36
3. Brueckner, K. A., Eden, R. J., Francis, N. C. 1955. *Phys. Rev.* 93:1445;

Brueckner, K. A., Levinson, C. A. 1955. *Phys. Rev.* 97:1344
4. Goldstone, J. 1957. *Proc. R. Soc. London Ser. A* 239:267
5. Kelly, H. P. 1969. *Adv. Chem. Phys.* 14:129
6. Coester, F. 1958. *Nucl. Phys.* 1:421
7. Coester, F., Kümmel, H. 1960. *Nucl. Phys.* 17:477; Kümmel, H. 1969. *Nucl. Phys.* 22:177
8. Čížek, J. 1966. *J. Chem. Phys.* 45: 4256; 1969. *Adv. Chem. Phys.* 14: 35
9. Paldus, J., Čížek, J. 1973. *Energy, Structure, and Reactivity*, ed. D. W. Smith, W. B. McRae, p. 389. New York: Wiley
10. Harris, F. 1977. *Int. J. Quantum Chem. Symp.* 11:403
11. Paldus, J., Čížek, J., Saute, M., Laforgue, A. 1978. *Phys. Rev. A* 17:805
12. Öhrn, Y. 1976. *The New World of Quantum Chemistry*, ed. B. Pullman, R. Parr. Dordrecht, Holland: Reidel
13. Simons, J. 1977. *Ann. Rev. Phys. Chem.* 28:15
14. Jørgensen, P. 1975. *Ann. Rev. Phys. Chem.* 26:359
15. Csanak, G., Taylor, H. S., Yaris, R. 1971. *Adv. At. Mol. Phys.* 7:287
16. Linderberg, J., Öhrn, Y. 1973. Propagators in Quantum Chemistry, London: Academic
17. Robb, M. A. 1974. *Computational Techniques in Quantum Chemistry and Molecular Physics*, ed. G. H. F. Diercksen, B. T. Sutcliffe, A. Veillard. Dordrecht, Holland: Reidel
18. Kutzelnigg, W. 1977. *Methods of Electronic Structure Theory*, ed. H. F. Schaefer, p. 129. New York: Plenum
19. Paldus, J., Čížek, J. 1975. *Adv. Quantum Chem.* 9:105
20. Musher, J. I. 1972. *Int. Rev. Sci. Phys. Chem. Ser. One* 1:1
21. Čížek, J., Paldus, J. 1980. *Phys. Scr.* 21:251
22. Bartlett, R. J., Purvis, G. D. 1980. *Phys. Scr.* 21:255
23. Sinanoglu, O. 1962. *J. Chem. Phys.* 36:706, 3198; 1964. *Adv. Chem. Phys.* 6:315
24. Nesbet, R. K. 1969. *Adv. Chem. Phys.* 14:1
25. Freed, K. F. 1968. *Phys. Rev.* 173:1
26. Meyer, W. 1971. *Int. J. Quantum Chem.* 5:341; 1973. *J. Chem. Phys.* 58:1017
27. Sinanoglu, O., Brueckner, K. A., eds. 1970. *Three Approaches to Electron Correlation in Atoms*. New Haven: Yale Univ. Press
28. Hurley, A. C. 1976. *Electron Correlation in Small Molecules*. New York: Academic
29. Pople, J. A., Binkley, J. S., Seeger, R. 1976. *Int. J. Quantum Chem. Symp.* 10:1
30. Bartlett, R. J., Purvis, G. D. 1978. *Int. J. Quantum Chem.* 14:561
31. Bartlett, R. J., Purvis, G. D. 1981. *Ann. NY Acad. Sci.* 367:62
32. Dunning, T. H., McKoy, V. 1967. *J. Chem. Phys.* 47:1735; Shibuya, T. I., McKoy, V. 1970. *Phys. Rev. A.* 2: 2208
33. Bartlett, R. J., Shavitt, I., Purvis, G. D. 1979. *J. Chem. Phys.* 71:281
34. Primas, H. 1965. *Modern Quantum Chemistry*, ed. O. Sinanoglu, 2:45. New York: Academic
35. Ahlrichs, R. 1974. *Theor. Chim. Acta* 35:59
36. Keil, F., Ahlrichs, R. 1976. *J. Am. Chem. Soc.* 98:4787
37. Meunier, A., Levy, B., Berthier, G. 1976. *Int. J. Quantum Chem.* 10: 1061
38. Pople, J. A., Seeger, R., Krishnan, R. 1977. *Int. J. Quantum Chem. Symp.* 11:149
39. Davidson, E., Silver, D. W. 1977. *Chem. Phys. Lett.* 52:403
40. Meunier, A., Levy, B. 1979. *Int. J. Quantum Chem.* 16:955
41. Malrieu, J. P. 1979. *J. Chem. Phys.* 70:4405; Malrieu, J. P., Spiegelman, F. 1979. *Theor. Chim. Acta* 52:55
42. Sasaki, F. 1977. *Int. J. Quantum Chem. Symp.* 11:125
43. Brändas, E., Dunne, L., Combs, L. L., Correia, N. S. 1981. *Int. J. Quantum Chem.* In press
44. Pople, J. A. 1978. *Post Hartree-Fock: Configuration Interaction Workshop, NRCC, LBL-8233*, Berkeley, Calif., p. 233
45. Löwdin, P. O. 1959. *Adv. Chem. Phys.* 2:207
46. Bartlett, R. J., Silver, D. M. 1975. *Int. J. Quantum Chem.* 9:183
47. Takahashi, J., Fukutome, H. 1978. *Prog. Theor. Phys.* 59:1787; Igawa, A., Fukutome, H. 1980. *Prog. Theor. Phys.* 64:491
48. Hubač, I., Čársky, P. 1980. *Phys. Rev. A* 22:2392
48a. McDowell, K. 1981. *Int. J. Quantum Chem.* 19:271
49. Paldus, J., Čížek, J., Shavitt, I. 1972. *Phys. Rev. A* 5:50
50. Pople, J. A., Krishnan, R., Schlegel, H. B., Binkley, J. S. 1978. *Int. J. Quantum Chem.* 14:545
51. Taylor, P. R., Bacskay, G. B., Hurley, A. C., Hush, N. S. 1978. *J.*

Chem. Phys. 69:1971; 1976. *Chem. Phys. Lett.* 41:444
52. Bartlett, R. J., Shavitt, I. 1977. *Chem. Phys. Lett.* 50:190; 1978. 57: 157
53. Adams, B. G., Jankowski, K., Paldus, J. 1979. *Chem. Phys. Lett.* 67:144
54. Jankowski, K., Paldus, J. 1980. *Int. J. Quantum Chem.* 18:1243
55. Saute, M., Paldus, J., Čížek, J. 1979. *Int. J. Quantum Chem.* 15:463
56. Paldus, J. 1977. *J. Chem. Phys.* 67: 303; Adams, B. G., Paldus, J. 1979. *Phys. Rev. A* 20:1
57. Nakatsuji, H., Hirao, K. 1978. *J. Chem. Phys.* 68:2053, 4279
58. Hirao, K., Nakatsuji, H. 1979. *J. Chem. Phys.* 69:4535, 4548
59. Nakatsuji, H. 1979. *Chem. Phys. Lett.* 67:329, 334
60. Živkovič, T. P. 1977. *Int. J. Quantum Chem. Symp.* 11:413
61. Živkovič, T. P., Monkhorst, H. J. 1978. *J. Math. Phys.* 19:1007
62. Monkhorst, H. 1977. *Int. J. Quantum Chem. Symp.* 11:421
63. Freeman, D. L. 1977. *Phys. Rev. B* 15:5512; 1978. *Solid State Commun.* 26:289
64. Møller, C., Plesset, M. S. 1934. *Phys. Rev.* 46:618
65. Claverie, P., Diner, S., Malrieu, J. P. 1967. *Int. J. Quantum Chem.* 1:751; Epstein, P. S. 1926. *Phys. Rev.* 28: 695; Nesbet, R. K. 1955. *Proc. R. Soc. London Ser. A* 230:312
66. Ahlrichs, R. 1979. *Comput. Phys. Commun.* 17:31
67. Davidson, E. R. 1973. *Energy, Structure, and Reactivity*, ed. D. W. Smith, W. B. McRae, p. 179. New York: Wiley
68. Davidson, E. R. 1972. *J. Chem. Phys.* 57:2005; Davidson, E. R., Bender, C. F. 1972. *J. Chem. Phys.* 56:4334
69. Huzinaga, S., Arnau, C. 1970. *Phys. Rev. A* 1:1285
70. Silverstone, H. J., Yin, M. L. 1968. *J. Chem. Phys.* 49:2020
71. Bartlett, R. J., Silver, D. M. 1974. *Int. J. Quantum Chem. Symp.* 8:271; 1974. *Phys. Rev. A* 10:1927; 1975. *J. Chem. Phys.* 62:3258; 1976. 64: 1260, 4578. Bartlett, R. J., Silver, D. M. 1976. *Quantum Science*, ed. J. Linderberg, J. L. Calais, O. Goscinski, Y. Ohrn, p. 393. New York: Plenum
72. Diner, S., Malrieu, J. P., Claverie, P. 1969. *Theor. Chim. Acta* 13:1; 1967. *Theor. Chim. Acta* 8:390
73. Bartlett, R. J., Wilson, S., Silver, D. M. 1977. *Int. J. Quantum Chem.* 12:737; Wilson, S., Silver, D. M., Bartlett, R. J. 1977. *J. Mol. Phys.* 33: 1177
74. Siegbahn, P. 1977. *Methods of Electronic Structure Theory*, ed. H. F. Schaefer III, p. 81. New York: Plenum
75. Micha, D. A. 1970. *Phys. Rev. A* 1: 755
76. Bari, T. L., Davidson, E. R. 1970. *Phys. Rev. A* 1:644
77. Paldus, J. 1978. *Post Hartree-Fock: Configuration Interaction Workshop, NRCC, LBL-8233*, Berkeley, Calif., p. 239
78. Freeman, D. L., Karplus, M. 1976. *J. Chem. Phys.* 64:2641
79. Prime, S., Robb, M. A. 1975. *Chem. Phys. Lett.* 35:86
80. Taylor, W. J. 1974. *Chem. Phys. Lett.* 26:29
81. Ostlund, N. S., Bowen, M. F. 1976. *Theor. Chim. Acta* 40:175
82. Adamowicz, L., Sadlej, A. J. 1978. *J. Chem. Phys.* 69:3992; 1977. 67:4298
83. Pan, K. C., King, H. F. 1972. *J. Chem. Phys.* 56:4667
84. Silver, D. M., Bartlett, R. J. 1976. *Phys. Rev. A* 13:1; Silver, D. M., Wilson, S., Bartlett, R. J. 1977. *Phys. Rev. A* 16:477
85. Hiroike, E. 1980. *Int. J. Quantum Chem.* 18:37
86. Lindgren, I., Lindgren, J., Mårtensson, A. M. 1976. *Z. Phys. A* 279:113; 1977. *Phys. Rev. A* 15:2123
87. Brandow, B. H. 1967. *Rev. Mod. Phys.* 39:771
88. Lindgren, I. 1978. *Int. J. Quantum Chem. Symp.* 12:33
89. Mukherjee, D., Moitra, R. K., Mukhopadhyey, A. 1975. *Mol. Phys.* 30: 1861
90. Hose, G., Kaldor, U. 1979. *J. Phys. B.* 12:3827
91. Levy, B. 1978. *Proc. 4th Seminar Comput. Meth. Quantum Chem.*, Orenas, Sweden. Rep. MPI-PAE/ Astro 171, p. 149
92. Kirtman, B. 1981. *J. Chem. Phys.* 75:798
93. Löwdin, P. O. 1968. *Int. J. Quantum Chem.* 2:867, and references therein; 1962. *J. Math. Phys.* 3:969
94. Purvis, G. D., Bartlett, R. J. 1981. *J. Chem. Phys.* 75:1284
95. Kirtman, B., Cole, S. J. 1978. *J. Chem. Phys.* 69:5055
96. Hunt, W. J., Hay, P. J., Goddard, W. A. III. 1972. *J. Chem. Phys.* 57: 738
97. Chipman, D. M., Kirtman, B., Palke, W. E. 1976. *J. Chem. Phys.* 65:2256

98. Ishikawa, Y. 1980. *Z. Naturforsch. Teil A* 35:408
99. Kelly, H. P. 1969. *Phys. Rev. Lett.* 23:455
100. Miller, J. H., Kelly, H. P. 1971. *Phys. Rev. Lett.* 26:679; 1971. *Phys. Rev. A* 4:480
101. Lee, T., Dutta, N. C., Das, T. P. 1970. *Phys. Rev. Lett.* 25:204; Lee, T., Das, T. P. 1972. *Phys. Rev. A* 6: 968
102. McDowell, K. 1979. *Chem. Phys. Lett.* 66:423
103. McDowell, K. 1978. *J. Chem. Phys.* 68:4151; 69:5050
104. Schulman, J. M., Kaufman, D. N. 1970. *J. Chem. Phys.* 53:477
105. Robb, M. A. 1973. *Chem. Phys. Lett.* 20:274
106. Kaldor, U. 1973. *Phys. Rev. A* 7:427
107. Urban, M., Kellö, V., Hubač, I. 1977. *Chem. Phys. Lett.* 51:170
108. Kvasnicka, V., Laurinc, V. 1977. *Theor. Chim. Acta* 45:197
109. Wilson, S., Silver, D. M. 1976. *Phys. Rev. A* 14:1949
110. Roos, B. 1972. *Chem. Phys. Lett.* 15: 153
111. Davidson, E. R. 1974. *The World of Quantum Chemistry*, ed. R. Daudel, B. Pullman, p. 17. Dordrecht, Holland: Reidel; Langhoff, S. R., Davidson, E. R. 1974. *Int. J. Quantum Chem.* 8:61
112. Bartlett, R. J., Shavitt, I. 1977. *Int. J. Quantum Chem. Symp.* 11:165; 1978. 12:543
113. Siegbahn, P. E. M. 1978. *Chem. Phys. Lett.* 55:386
114. Luken, W. L. 1978. *Chem. Phys. Lett.* 58:421
115. Krishnan, R., Pople, J. A. 1978. *Int. J. Quantum Chem.* 14:91
116. Krishnan, R., Frisch, M. J., Pople, J. A. 1980. *J. Chem. Phys.* 72:4244
117. Kvasnicka, V., Laurinc, V., Biskupic, S. 1979. *Chem. Phys. Lett.* 67:81; 1980. *Mol. Phys.* 39:143
118. Wilson, S. 1979. *J. Phys. B* 12:L657; Wilson, S., Saunders, V. R. 1979. *J. Phys. B* 12:403; Errata 1980 *J. Phys. B* 13:1505
119. Wilson, S., Guest, M. F. 1980. *Chem. Phys. Lett.* 73:607
120. Saxe, P., Schaefer, H. F. III, Handy, N. 1981. *Chem. Phys. Lett.* 79:202
120a. Blomberg, M. R. A., Siegbahn, P. E. M. 1978. *Int. J. Quantum Chem.* 14:583; Blomberg, M. R. A., Siegbahn, P. E. M., Roos, B. O. 1980. *Int. J. Quantum Chem. Symp.* 14:229
121. Rosenberg, B. J., Shavitt, I. 1975. *J. Chem. Phys.* 63:2162
122. Frisch, M. J., Krishnan, R., Pople, J. A. 1980. *Chem. Phys. Lett.* 75:66
123. Hubač, I. 1980. *Int. J. Quantum Chem.* 17:195
124. Urban, M., Hubač, I., Kellö, V., Noga, J. 1980. *J. Chem. Phys.* 72: 3378
125. Rosenberg, B. J., Ermler, W. C., Shavitt, I. 1976. *J. Chem. Phys.* 65: 4072
126. Kraemer, W. P., Diercksen, G. H. F. 1976. *Astrophys. J.* 205:97
127. Redmon, L. T., Purvis, G. D., Bartlett, R. J. 1979. *J. Am. Chem. Soc.* 101:2856
128. Redmon, L. T., Purvis, G. D., Bartlett, R. J. 1980. *J. Chem. Phys.* 72: 986
129. Purvis, G. D., Bartlett, R. J. 1981. *J. Chem. Phys.* 75:1284
130. Krishnan, R., Binkley, J. S., Seeger, R., Pople, J. A. 1980. *J. Chem. Phys.* 72:650
131. Guest, M. F., Wilson, S. 1980. *Chem. Phys. Lett.* 72:49
132. Ahlrichs, R., Zirz, C. 1980. *Proc. Mol. Phys. Quantum Chem. Workshop*, ed. P. G. Burton, Univ. Wollongang, Australia
133. Redmon, L. T., Purvis, G. D., Bartlett, R. J. 1978. *J. Chem. Phys.* 69: 5386
134. Maki, A. G., Sams, R. L. Unpublished, private communication
135. Pearson, P. K., Schaefer, H. F. III, Walgren, U. 1975. *J. Chem. Phys.* 62:350
136. Adams, G. F., Bent, G. D., Purvis, G. D., Bartlett, R. J. 1981. *Chem. Phys. Lett.* 81:461
137. Adams, G. F., Bent, G. D., Purvis, G. D., Bartlett, R. J. 1979. *J. Chem. Phys.* 71:3697
138. Adams, G. F., Bent, G. D., Bartlett, R. J., Purvis, G. D. 1981. *Potential Energy Surfaces and Dynamics Calculations*, ed. D. G. Truhlar, p. 133. New York: Plenum
139. Redmon, M., Bartlett, R. J., Garrett, B. C., Purvis, G. D., Saatzer, P. M., Schatz, G. C., Shavitt, I. 1981. *Potential Energy Surfaces and Dynamics Calculations*, ed. D. G. Truhlar, p. 771. New York: Plenum
140. Silver, D. M. 1980. *J. Chem. Phys.* 72:6445
141. Goddard, J. D. Schaefer, H. F. III. 1979. *J. Chem. Phys.* 70:5117
142. Miller, W. H. 1979. *J. Am. Chem. Soc.* 101:6810
143. Houston, P. L., Moore, C. B. 1976. *J. Chem. Phys.* 65:757

144. Harding, L. B., Schlegel, H. B., Krishnan, R., Pople, J. A. 1980. *J. Phys. Chem.* 83:3394
145. Adams, G. F., Bent, G. D., Bartlett, R. J., Purvis, G. D. 1981. *J. Chem. Phys.* 75:834
145a. Frisch, M. J., Krishnan, R., Pople, J. A. 1981. *J. Phys. Chem.* In press
145b. Fehlner, T. P., Mappes, G. W. 1969. *J. Phys. Chem.* 73:873
145c. Baghal-Vayjooee, M. H., Collister, J. L., Pritchard, H. O. 1977. *Can. J. Chem.* 55:2634
145d. Schneider, F. W., Rabinovitch, B. S. 1962. *J. Am. Chem. Soc.* 82:4215
145e. Warneck, P. 1971. *Z. Naturforsch.* A26:2047
145f. Yamada, K., Nagakuru, T., Kuchitsu, K., Morino, Y. 1971. *J. Mol. Spectrosc.* 38:70
145g. Walsh, R., Benson, S. W. 1966. *J. Am. Chem. Soc.* 88:4570
146. Harding, L. B., Schlegel, H. B., Krishnan, R., Pople, J. A. 1981. *Potential Energy Surfaces and Dynamics Calculations*, p. 169. New York: Plenum
147. Adeney, P. D., Bouma, W. J., Raddom, L., Rodwell, W. R. 1980. *J. Am. Chem. Soc.* 102:4069
148. Kenney, J. W., Simons, J., Purvis, G. D., Bartlett, R. J. 1978. *J. Am. Chem. Soc.* 100:6930
149. Weinstein, H., Osman, R., Edwards, W. D., Green, J. P. 1978. *Int. J. Quantum Chem.* QBS5:449
150. Osman, R., Topiol, S., Weinstein, H., Eilers, J. E. 1980. *Chem. Phys. Lett.* 73:399
151. Binkley, J. S., Pople, J. A. 1977. *Chem. Phys. Lett.* 45:197
152. Seeger, R., Krishnan, R., Pople, J. A., Schleyer, P. v. R. 1977. *J. Am. Chem. Soc.* 99:7103
153. Binkley, J. S., Seeger, R., Pople, J. A., Dill, J. D., Schleyer, P. v. R. 1977. *Theor. Chim. Acta* 45:69
154. Krogh-Jespersen, K., Cremer, D., Poppinger, D., Pople, J. A., Schleyer, P. v. R., Chandrasekhar, J. 1979. *J. Am. Chem. Soc.* 101:4843
155. DeFrees, D. J., Levi, B. A., Pollack, S. K. Hehre, W. J., Binkley, J. S., Pople, J. A. 1979. *J. Am. Chem. Soc.* 101:4085
156. Purvis, G. D., Bartlett, R. J. 1978. *J. Chem. Phys.* 68:2114
157. Purvis, G. D., Bartlett, R. J. 1979. *J. Chem. Phys.* 71:548
158. Chiles, R. A., Dykstra, C. E. 1981. *J. Chem. Phys.* 74:4544
159. Urban, M., Kellö, V. 1979. *Mol. Phys.* 38:1621
160. Kellö, V., Urban, M., Hubač, I., Čársky, P. 1978. *Chem. Phys. Lett.* 58:83
161. Hubač, I., Urban, M. 1977. *Theor. Chim. Acta* 45:185
162. Wilson, S., Silver, D. M. 1977. *J. Chem. Phys.* 66:5400
163. Wilson, S., Silver, D. M. 1980. *J. Chem. Phys.* 72:2159
164. Wilson, S. 1979. *J. Phys. B* 12:1623
165. Silver, D. M. 1980. *Phys. Rev. A* 21: 1106
166. Robb, M. A., Wilson, S. 1980. *Mol. Phys.* 40:1333
167. Caves, T. C., Karplus, M. 1969. *J. Chem. Phys.* 50:3649
168. Adamowicz, L., Sadlej, A. J. 1978. *Chem. Phys. Lett.* 53:377
169. Werner, H. J., Meyer, W. 1976. *Mol. Phys.* 31:855
170. Bartlett, R. J., Purvis, G. D. 1979. *Phys. Rev.* 20:1313
171. Purvis, G. D., Bartlett, R. J. 1981. *Phys. Rev.* 23:1594
172. Sadlej, A. J. 1977. *Chem. Phys. Lett.* 47:50
173. McCullough, E. A. 1975. *Chem. Phys. Lett.* 24:55; 1975. *J. Chem. Phys.* 62:3991; 63:5050
174. Itagaki, T., Saika, A. 1979. *J. Chem. Phys.* 71:4620
175. Itagaki, T., Saika, A. 1979. *J. Chem. Phys.* 70:2378
176. Itagaki, T., Saika, A. 1980. *Int. J. Quantum Chem.* 18:1505
177. Yoshioka, Y., Jordan, K. D. 1980. *J. Chem. Phys.* 73:5899
178. Kelly, H. P., Carter, S. L. 1980. *Phys. Scr.* 21:448
179. Bent, G. D., Adams, G. F., Bartram, R. H., Bartlett, R. J., Purvis, G. D. 1981. *J. Chem. Phys.* In press
180. Rossky, P. J., Karplus, M. 1980. *J. Chem. Phys.* 73:6196
181. Purvis, G. D., Bartlett, R. J. To be published
182. Banerjee, A., Simons, J. 1981. *Int. J. Quantum Chem.* 19:207
183. Brandow, B. H. 1979. *Int. J. Quantum Chem.* 15:207
184. Paldus, J., Čížek, J. 1974. *J. Chem. Phys.* 60:149
185. Hose, G., Kaldor, U. 1980. *Phys. Scr.* 21:357
186. Van Vleck, J. H. 1929. *Phys. Rev.* 33: 467; Jordahl, O. M. 1934. *Phys. Rev.* 45:87; Kemble, E. C. 1937. *Fundamental Principles of Quantum Mechanics*. New York: McGraw-Hill
187. Klein, D. J. 1974. *J. Chem. Phys.* 61: 786
188. Shavitt, I., Redmon, L. T. 1981. *J. Chem. Phys.* 73:5711

189. Kaldor, U. 1973. *Phys. Rev. Lett.* 31: 1338; 1975. *J. Chem. Phys.* 63:2199
190. Stern, P. S., Kaldor, U. 1976. *J. Chem. Phys.* 64:2002
191. Morrison, J., Rajnak, K. 1971. *Phys. Rev. A* 4:536
192. Morrison, J., Salomonson, S. 1980. *Phys. Scr.* 21:343
193. Salomonson, S., Lindgren, I., Mårtensson, A. M. 1980. *Phys. Scr.* 21: 351
194. Hegarty, D., Robb, M. A. 1979. *Mol. Phys.* 37:1455
195. Davidson, E. R., Bender, C. 1978. *Chem. Phys. Lett.* 59:369
196. Freed, K. F. 1977. *Semiempirical Methods of Electronic Structure Calculations*, Vol. 1, ed. G. A. Segal. New York: Plenum; Iwata, S., Freed, K. F. 1976. *J. Chem. Phys.* 65:1071; Sheppard, M. G., Freed, K. F., Herman, M. F., Yeager, D. C. 1979. *Chem. Phys. Lett.* 61:577
197. Westhaus, P. 1973. *Int. J. Quantum Chem. Symp.* 7:463; Westhaus, P., Moghtaderi, M. 1980. *J. Chem. Phys.* 72:4174
198. Baker, H., Hegarty, D., Robb, M. A. 1980. *Mol. Phys.* 41:653
199. Harris, R. A. 1967. *J. Chem. Phys.* 47:3967; 1968. 48:3600
200. Kvasnicka, V. 1975. *Phys. Rev. A* 12: 1159; Fischer, S. 1970. *Int. J. Quantum Chem. Symp.* 3:651

ELECTRON CORRELATION IN LARGE MOLECULES WITH MANY-BODY METHODS

Rodney J. Bartlett and George D. Purvis III

Reprinted from
ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Volume 367 Pages 62-82
May 22,1981 28113

ELECTRON CORRELATION IN LARGE MOLECULES WITH MANY-BODY METHODS

Rodney J. Bartlett and George D. Purvis III

Battelle
Columbus Laboratories
Columbus, Ohio 43201

INTRODUCTION

In the application of quantum chemical methods to problems involving biochemical molecules and their interactions, there is a natural progression from empirical or semiempirical models and methods to ab initio self-consistent field (SCF) approaches, and, eventually, to ab initio approaches that properly include the effects of electron correlation. The purpose of this contribution is to discuss the many-body methods[1-15] (i.e., many-body perturbation theory, MBPT,[1-3] and coupled cluster methods, CCM[4-7]) for including electron correlation. The emphasis is on the advantages that these methods offer over the more traditional configuration interaction (CI) approaches in large molecule applications.

Semiempirical models and methods, which should be used synergistically with experimental results, are most properly employed to investigate trends among a series of similar molecules. Such methods can be used for rather large molecules relatively inexpensively, and are thus finding wide use in biochemistry and particularly in drug design.[16-20] On the other hand, ab initio methods can, in principle, provide hard, quantitative results for molecular systems, which can, potentially, be used to complement various experimental methods by providing answers to classes of problems that are not as amenable to experiment. An example would be identifying the transition state and activation barrier in a reaction.

In practice, however, ab initio quantum chemistry suffers from severe limitations, which have permitted highly accurate results to be obtained for only comparatively small molecules. These limitations are basically of three types:

1. The number of degrees of freedom in molecular systems
2. The limited size of the basis set that can be used
3. The required degree of accuracy of the ab initio approach.

The problem in the first category essentially revolves around the Born-Oppenheimer (or fixed nuclei) approximation, since the calculation of the electronic structure and energy must be repeated for each choice of coordinates for the nuclei. Limitations 2 and 3 pertain to each of these calculations. For example, mapping out a potential energy surface for a four-atom system with $3N - 6 = 6$ degrees of freedom and computing 10 points for each degree of freedom would amount to a million calculations. In quantum biochemistry, fortunately, one is not often interested in a complete energy surface, but usually only in a few crucial bond lengths and angles that need to be optimized, but this is still a formidable problem. The development of self-consistent field (SCF)[21-23]

0077-8923/81/0367-0062 $1.75/0

and correlated gradient methods[24-26] is a welcome addition to the quantum chemist's repertoire, but even these techniques are only applicable to a few degrees of freedom.

To take an example in quantum biochemistry, consider a solvated molecule in which it is recognized that the solvation characteristics are partially responsible for the conformation of the molecule, which can, in its turn, directly affect a highly specific interaction. The only feasible approach to such a problem at present is to determine analytic model potentials of the Lennard-Jones, generalized Morse, and other types, with parameters determined either empirically[27] or, perhaps, from highly accurate quantum chemical calculations of the component of the dynamical movement of the molecule and solvent, allowing the more accurate quantum chemical methods, augmented by gradient techniques, to focus on the most crucial active site interactions. The results of this procedure, however, are no more accurate than the individual calculations, which are subject to limitations 2 and 3.

In FIGURE 1 is shown a schematic drawing that illustrates the dependence of an ab initio quantum chemical calculation on basis set and caliber of method. Improving only one of the two is not enough, but rather a systematic improvement in both is required.

First, consider the basis set problem. Depending upon the property of interest, it is a matter of opinion just how many basis functions are required to obtain good SCF results for molecules, but one would certainly want at least a minimum basis set of one Slater orbital (or contracted Gaussian orbital, i.e., SZ—single zeta) for each electron and probably two (DZ, double zeta) or more (DZP, double zeta + polarization). The number of molecular integrals needed to do an SCF calculation rises formally as n^4, where n is the number of basis functions, although this dependence can be reduced to n^2 for sufficiently large molecules. The largest SCF calculations that have been done employ no more than ~300 functions. This imposes a limit of, at most, 300 electrons or, more realistically, ~100 to 150 electrons explicitly considered.

The problem is further compounded when electron correlation is included. Except for second-order perturbation theory, which will be considered in more detail below, correlated methods have a dependence on the number of basis functions of $\sim n^6$. Again, it is possible to reduce this by perhaps two orders of magnitude for sufficiently large molecules, but it is evident that even fewer problems can be studied at the correlated level than at the SCF level.[30,31]

No really good idea for eliminating the basis set problem in quantum chemistry has yet appeared. Completely numerical SCF calculations have only been accomplished for a few diatomic molecules[32] and nothing of general utility has yet emerged. At the cost of using unrealistic potentials, the numerical procedures of the type used in MS-X_α have had some success.[33] The development of effective potentials for the chemically inert electrons in heavy atom molecules is very useful from the viewpoint of basis set quantum chemical computations.[34-36] Also, using Gaussian lobe functions[37] to represent the bonds in a molecule rather than locating them on the various atomic centers has reduced the number of basis functions while simplifying the calculations of the integrals.[40] Various integral approximations and other clever schemes can also aid in making

AB INITIO CALCULATIONS

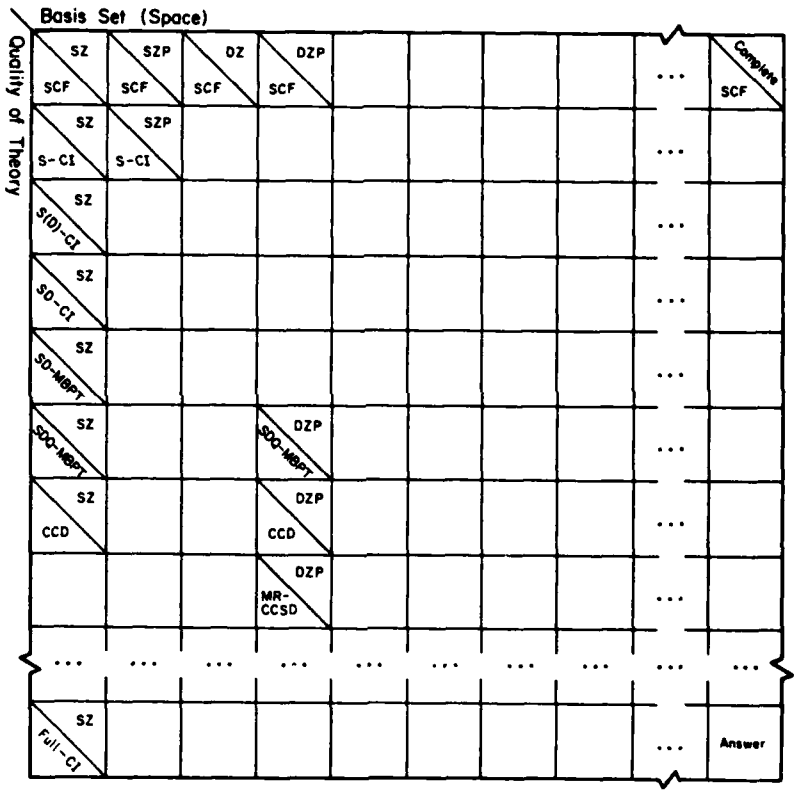


FIGURE 1. An illustration of the dependence of an ab initio calculation on the basis set and on the quality of the theory. SZ, SZP, DZ, and DZP are, respectively, single zeta, single zeta plus polarization, double zeta, etc. Configuration interaction (CI) is usually accomplished by adding single and double excitations. MBPT and CCM, in general, exceed SD-CI in accuracy, since the effects of higher excitations are included to some degree. MR-CCSD indicates a coupled-cluster theory that is limited to $e^{T_2+T_2}$ but relative to more than one reference function. The best possible solution in a basis set is full CI.

the calculations more efficient,[32] but the basis set problem is still a fundamental limitation.

The third limitation above, illustrated in FIGURE 1, pertains to the degree of accuracy of method that is required for the property of interest; this is the main concern of this article. Generally, SCF theory is considered to be adequate ($\pm 10\%$) for molecular conformations, equilibrium molecular structure, and first-order properties; that is, properties obtained as an expectation value over the

SCF density, such as the electrostatic potential or dipole moments. On the other hand, correlated methods are considered absolutely necessary in predicting electronic and photoelectronic spectra, in studying binding energies and other thermochemical quantities in reactions where bond breaking is occurring, and, in investigating most second and higher order properties, such as polarizabilities, shielding constants, magnetic succeptibilities. Since many questions in quantum biochemistry revolve around one or another of the properties that need an accurate treatment of electron correlation, it is important to consider the characteristics that a correlated method should have if it is to be applied to the large molecules that occur in quantum biochemistry.

A few desirable characteristics for such a correlated approach are that the method should be[11]

1. Size-extensive (i.e., should scale properly with the size of molecule)
2. Generally applicable to a wide class of problems (i.e., avoid specific formulations or choices of configurations)
3. Efficient and cost-effective (i.e., provide large correlation corrections inexpensively)
4. Applicable to both open-shells and excited states
5. Able to correctly separate a molecule into its fragments.

Another condition that one might expect is that the method be variational, giving an upper bound for the total energy. Lacking a coordinate lower bound, we believe this to be an unnecessary restriction because the quantities of interest in quantum chemistry are, invariably, energy differences like binding energies, which possess no rigorous variational properties even if the individual calculations are variational. Furthermore, except for SCF, full CI, and a few other isolated cases (e.g., generalized valence bond, GVB[41]), a variational requirement is not consistent with the size-extensive condition above, which is felt to be much more important to satisfy for large molecules.

In the present state of the art of correlated theory, the first three conditions are easily accomplished with MBPT. Any approach based upon the linked-diagram theorem is size-extensive. A large class of problems can be studied with single reference MBPT/CCM calculations, provided that RHF (Restricted Hartree-Fock) (or a UHF (Unrestricted) open-shell solution) is an adequate starting point.

As to the cost-effective property—it will be shown that second order perturbation theory, which is the simplest MBPT approximation, typically accounts for ~90% of the correlation energy in a basis set and significantly improves the SCF predictions of dissociation energies and molecular geometries. Since this requires only marginally more effort than an SCF calculation, is size-extensive and has rather general utility, it is a very attractive lowest-order approximation.

The fourth requirement can be handled with many-body approaches, such as equation-of-motion techniques,[42-44] or with CI; the fifth is currently most easily achieved using CI methods. The MBPT/CCM theory does exist for this latter problem, but it has not yet been implemented in a general purpose program.[45-47] In many cases, a UHF reference function will permit correct separation, but the path toward the separated limit is not always accurate.[9]

In the next section, the size extensive property of MBPT/CCM will be discussed in some detail, since this is an extremely important condition for potential applications of correlated methods to large molecules. We will then present a brief discussion of some other concepts in many-body theory that are important for large molecules, while the final section will focus on some applications to benzene in order to demonstrate the nature of correlation effects due to higher excitations in this prototype system. Emphasis will also be placed on the accuracy of the simplest approximation, second order perturbation theory, which typically provides a very large part of the electron correlation effect as an inexpensive by-product of the SCF calculation.

Size-Extensivity in Molecular Calculations

Probably the best way to illustrate the importance of quantum mechanical methods that scale properly with the size of a molecule is to consider the model problem of a lattice of separated electron pair bonds, such as H_2 molecules, since this serves as a first approximation to any large molecule. This problem has been worked out by several investigators,[11,48,49] but it is pertinent enough to the discussion that it is worth presenting a version here.

Assume that the H_2 molecules are either sufficiently far apart or separated by barriers so that they can be considered noninteracting. For simplicity's sake, we will further assume that the component set of molecular orbitals for each H_2 molecule are natural orbitals so that single excitations in the H_2 wavefunction need not be considered. Then the intermediately normalized wavefunction for each molecule, i, may be written,

$$\Phi^M(i) = \Phi_0^M(i) + \chi^M(i). \tag{1}$$

where $\Phi_0^M(i)$ is the first natural determinant (close to the Hartree-Fock solution) and $\chi^M(i)$ is the sum of doubly-excited determinants, including their appropriate coefficients. The norm of the function in (1) is $1 + S$, where

$$S(i) = \langle \chi^M(i) | \chi^M(i) \rangle, \tag{2}$$

$$\langle \Phi_0^M(i) | \chi^M(i) \rangle = 0. \tag{3}$$

The wave function for the lattice is

$$\Phi^L = \prod_{i=1}^{N} (\Phi^M(i) + \chi^M(i)). \tag{4}$$

Antisymmetry is disregarded because the molecules are noninteracting.

With $H^L = \Sigma H(i)$, the energy of the lattice

$$E^L = \frac{\langle \Phi^L | H^L | \Phi^L \rangle}{\langle \Phi^L | \Phi^L \rangle} = N E_{H_2}^M, \tag{5}$$

where $E_{H_2}^M$ is the energy of the H_2 molecule. With $\beta = \langle \Phi_0^M(i) | H(i) | \chi^M(i) \rangle$, which is essentially the correlation energy of the molecule,

$$E_{H_2}{}^M - E_0{}^M + \beta. \tag{6}$$

A method is said to be "size-extensive" if the total energy calculated by the method is appropriately linear in N, as in (5).

Notice that the product wavefunction in (4) includes quadratic and higher product terms like $\chi^M(i)\chi^M(j)$, which correspond to simultaneous double-excitations on different centers, but are quadruple and higher excitations in a supermolecule CI description. Since these terms arise from disjoint double-excitations, they are fundamentally simple, but the standard CI framework is not able to exploit this inherent simplicity. This causes an innate error in truncated CI that becomes most important for large molecules.

To investigate this, we can consider a reference wavefunction for the lattice of the form

$$\Psi_0{}^L - \prod_{i=1}^{N} \Phi_0{}^M(i), \tag{7}$$

with energy

$$E_0{}^L - \sum_{i=1}^{N} E_0{}^M(i) - NE_0{}^M. \tag{8}$$

Using this reference function, the double-excitation CI (DCI) wavefunction for the lattice is constructed as

$$\begin{aligned}\Psi_{\mathrm{DCI}}^L &- \Psi_0{}^L + \sum_{k=1}^{N} c_k \Psi^L(k) \\ &- \Phi_0(1)\Phi_0(2)\cdots\Phi_0(N) + \sum_{k=1}^{N} c_k \Psi_0{}^L \chi^M(k)/\Phi_0{}^M(k) \\ &- \Psi_0{}^L + c\sum_{k=1}^{N} \Psi^L(k).\end{aligned} \tag{9}$$

The weighting coefficient, c, is the same for each H_2 molecule in a noninteracting lattice. Using the expressions

$$\langle \chi^M(i) | \chi^M(j) \rangle - \delta_{ij} S, \tag{10}$$

$$\langle \chi^M(i) | H^L | \chi^M(j) \rangle - \delta_{ij}[E_{H_2}{}^M S - \beta],$$

$$\Delta E - E^L - NE_0{}^M,$$

it follows that

$$\langle \Psi_0{}^L | H^L | \Psi^L(k) \rangle - \beta, \tag{11}$$

$$\langle \Psi^L(k) | H^L | \Psi^L(l) \rangle - \delta_{kl}[NSE_0{}^M + (S-1)\beta]. \tag{12}$$

From these matrix elements, the DCI secular equation becomes

$$\Delta E - Nc\beta - 0, \tag{13a}$$

$$|\beta - [S\Delta E + (S-1)\beta]|c - 0. \tag{13b}$$

Solving (13) simultaneously for ΔE and c,

$$\Delta E_{\mathrm{DCI}} = \beta\left[\frac{-(1-S) \pm \sqrt{(1-S)^2 + 4SN}}{2S}\right] \underset{N\to\infty}{\longrightarrow} N^{\mathrm{eff}}\beta \underset{N\to\infty}{\longrightarrow} \beta\sqrt{N/S}. \qquad (14)$$

The positive sign is required, since $\Delta E < 0$ and $\beta < 0$. Since the correct $\Delta E = N\beta$, DCI is not size-extensive.

With the aid of a value for S in (2) and (14), it is possible to get some feel for the size of these effects. From a natural orbital study by Davidson and Jones[50] of the 50 term Kolos-Roothaan wavefunction for H_2 at $R = 1.40$,[51] S for H_2 is 0.0181. Some representative values are shown in TABLE 1, along with values for a lattice of He atoms for comparison ($S_{He} = 0.0083$). It is apparent that the error in the correlation energy as determined by DCI can be significant even for modest numbers of electrons. It is also apparent that the errors are greater for typical

TABLE 1
ERRORS IN DCI AS A FUNCTION OF N
FOR A LATTICE OF NH_2 MOLECULES AND NHe ATOMS

	H_2		He	
N	N^{eff}	% Error in Correlation Energy	N^{eff}	% Error in Correlation Energy
2	1.97	1.5	1.98	0.8
4	3.81	4.8	3.91	2.4
6	5.54	7.7	5.77	3.8
8	7.19	10.1	7.58	5.2
10	8.77	12.3	9.35	6.5
20	15.8	21.1	17.6	12.1
50	32.0	35.6	38.2	23.6
100	51.0	48.0	65.2	34.8
1,000	209.5	79.1	292.2	70.8

covalent bonds than for inner-shell electron pairs, as in He atoms. In fact, we shall find that TABLE 1 can provide a rather accurate estimate of the effects of higher excitations if one simply counts the number of electrons in covalent bonds and inner-shell electron pairs.

Since the product terms $\chi^M(i)\chi^M(j)$ correspond to quadruple excitations in a supermolecule CI, triple products correspond to hextuple excitations, and so on, the size-extensive property of MBPT/CCM, which is a consequence of the linked-diagram theorem,[1,2] is essentially a result of a more proper treatment of quadruple and higher excitations than in CI. Hence, the statement that size extensivity is important in correlated calculations is equivalent to the statement that quadruple and higher excitations are important. Since the number of configurations in CI is proportional to the number of basis functions raised to the level of excitation included, the number of quadruple excitations generated from 100 basis functions would require $\sim(100)^4$, or 10^8, configurations. Hence, better computational methods for including effects of higher excitations in correlated

calculations are extremely important. Many-body methods tend to take the intelligent viewpoint that removing the erroneous terms (i.e., unlinked diagrams) in DCI is preferable to a more explicit inclusion of higher excitations. In practice, this viewpoint leads to computationally more tractable equations that are closer to those in DCI.[7]

Since the correct density is in error because of the neglect of the product terms in the wavefunctions, further analysis discussed elsewhere demonstrates that the density matrix obtained from a truncated CI reduces to just the density computed from the reference function.[49] If the latter is an SCF function, then we have

$$\lim_{N\to\infty} \rho^{\mathrm{TCI}} = \rho^{\mathrm{SCF}}. \tag{15}$$

Similarly, for an excitation energy,[49]

$$\lim_{N\to\infty} (E_f^{\mathrm{TCI}} - E_i^{\mathrm{TCI}}) = E_f^{\mathrm{SCF}} - E_i^{\mathrm{SCF}}. \tag{16}$$

Hence, size-extensivity affects more than the total energy.

One additional consequence worth mentioning is that, in a typical reaction,

$$A + B \rightarrow C + D, \tag{17}$$

the heat of the reaction, $\Delta H_{rxn} = \Delta H_f(C) + \Delta H_f(D) - \Delta H_f(A) - \Delta H_f(B)$. However, if these individual quantities are determined by a truncated CI, this simple addition is not entirely justified, since the truncated CI ignores the simultaneous excitations that prohibit $\Delta H_f(C + D)$ at $R_{CD} \rightarrow \infty$ from being $\Delta H_f(C) + \Delta H_f(D)$. In practice, this frequently requires that one compute the supermolecules $C + D$ and $A + B$ in CI to make the energy difference most accurate. This should be contrasted to predictions made with a size extensive method in which a table of results for species obtained at a given level of approximation may be added and subtracted just like the experimental values.

Synopsis of Many-Body Theory

MBPT/CCM has been discussed in detail in several places.[1-7] In particular, Reference 7 provides a fairly detailed mathematical description from the viewpoint taken in this article. The theory, as originally developed, uses second-quantization and diagram techniques, which, being unfamiliar to many quantum chemists, tend to camouflage the important concepts that emerge from the many-body approach. Instead of presenting any detailed mathematical development here, we will sketch the basis for the two significant concepts that emerge from MBPT/CCM, namely the linked-diagram theorem, which guarantees size extensivity, and the cluster decomposition of CI excitations into separate, more physically satisfying pieces that lead to tractable equations that include the effects of higher excitations. Consult Reference 7 for detailed equations and the original references[1-6] for the complete formal development. For simplicity's sake, we will limit ourselves in the following to a single reference function, such

as an unrestricted Hartree-Fock [SCF] solution. Various versions of the multi-reference function theory are available.[46-47]

It is well known that one way to solve a CI eigenvalue equation is with perturbation theory.[52] Using the Rayleigh-Schrödinger form, we can separate the Hamiltonian $H = H_0 + V$, where H_0 is the sum of the SCF one-electron Hamiltonians and V is the two-electron part minus the SCF effective one-particle Hamiltonian; we then have

$$H_0\phi_0 = E_0\phi_0,$$

$$H_0 D^{abc\ldots}_{ijk\ldots} = E^{abc\ldots}_{ijk\ldots} D^{abc\ldots}_{ijk\ldots}, \tag{18}$$

for ϕ_0, the SCF solution, and $D^{abc\ldots}_{ijk\ldots}$, the various determinants that can be formed by replacing occupied SCF orbitals with excited SCF orbitals. The CI eigenvalue through the fourth order then becomes

$$E = E_0 + \langle\phi_0|V|\phi_0\rangle + \langle\phi_0|VR_0V|\phi_0\rangle + \langle\phi_0|VR_0(V - \langle V\rangle)R_0V|\phi_0\rangle$$
$$+ \langle\phi_0|VR_0(V - \langle V\rangle R_0(V - \langle V\rangle)R_0V|\phi_0\rangle$$
$$- \langle\phi_0|VR_0V|\phi_0\rangle\langle\phi_0|VR_0^2V|\phi_0\rangle. \tag{19}$$

The resolvent, R_0, has the expression

$$R_0 = |\underline{h}\rangle(\underline{h}|E_0 - H_0|\underline{h}\rangle^{-1}\langle\underline{h}|, \tag{20}$$

where $|\underline{h}\rangle$ is composed of all the CI excitations, $D^{abc\ldots}_{ijk\ldots}$. Even though $|\underline{h}\rangle$ is formally complete, Slater's rules for matrix elements choose from all possibilities only the few that have nonvanishing contributions.

Subject to an SCF reference function, only double excitations can mix across V with ϕ_0, so the second and third order terms in (19) involve only double excitations. The first of the two terms in the fourth order, however, can mix single, double, triple, and quadruple excitations at the middle R_0 although the second fourth order term (i.e., the renormalization term) still has no contributions, except from double excitations.

From the model problem of separated H_2 molecules presented in the previous section, it is easy to check whether each of the terms in (19) is size extensive. Considering the second order energy of the H_2 lattice

$$E_2^L = \langle\phi_0|VR_0V|\phi_0\rangle, \tag{21}$$

as an example, we have, from Slater's rules,

$$E_2^L = \frac{1}{4}\sum_{i,j}^{\text{occ}}\sum_{a,b}^{\text{vac}} |\langle ij||ab\rangle|^2/(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b). \tag{22}$$

The notation $\langle ij\|ab\rangle = \langle ij|ab\rangle - \langle ij|ba\rangle = (ia|bj) - (ib|aj)$, and $\{\epsilon_i\}$ and $\{\epsilon_a\}$ are the SCF orbital energies for the occupied and excited orbitals, respectively. Using a little algebra,

$$E_2^L = 2\sum_{i<j}\sum_{a<b} [|(ia|jb)|^2 - (ia|jb)(ib|ja)]/(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b). \tag{23}$$

Since this expression is invariant to any unitary transformation among degenerate orbitals, we may choose the orbitals to be localized on the H_2 molecules in the lattice to make the argument most transparent. In this case, the only nonvanishing integrals have the charge distribution (ia), (jb), (ib), or (jc) on the same H_2 molecule. Hence, it follows that

$$E_2^L = NE_2^M \tag{24}$$

and second order perturbation theory is size extenstive. It can be similarly shown that this is also true for E_3.

We shall now consider the fourth order. The renormalization term is composed of an E_2 term and a similar term, $\Delta = \langle\phi_0|VR_0^2V|\phi_0\rangle$, which differs from E_2 only by the squared denominator. Since

$$E_2^L = NE_2^M, \qquad \Delta^L = N\Delta^M, \tag{25}$$

the product of the two has an N^2 dependence, which is erroneous. If E_4 is to be size extensive, then the first term in E_4 must also have an equal and opposite N^2 dependence to cancel out these uncharacteristic terms. The single, double, and triple excitation contributions to the first part of E_4 can be shown to be size extensive. Hence, to resolve the problem, it is necessary to consider the contributions of the quadruple excitations. After a great deal of algebra,[7] the quadruple excitation part, E_4^Q, may be written in the form

$$E_4^Q = E_2\Delta + Q, \tag{26}$$

where Q is properly size extensive. Hence, $E_2\Delta$ cancels the renormalization term and, with it, the erroneous N^2 dependence. This is the substance of the linked-diagram theorem. The algebraic analysis that leads to (26) represents Q as linked diagrams, while $E_2\Delta$ corresponds to unlinked diagrams. A similar analysis applies in all higher orders; this is the linked-diagram theorem.[2] This provides the expression for the energy

$$E = E_0 + \sum_{p=0}^{\infty} \langle\phi_0|V[(E_0 - H_0)^{-1}V]^p|\phi_0\rangle_L. \tag{27}$$

where L limits the terms to linked diagrams.

It should be evident that if quadruple excitations had not been included in E_4, then the $E_2\Delta$ term with the erroneous N^2 dependence would remain. This is exactly what happens when a truncated CI calculation is made. Limiting the configurations to single and double excitations, for example, will necessarily retain these erroneous terms, thus destroying the size extensivity of the method. If quadruple excitations were included in the CI, the result would be size extensive through the fifth order, but would fail in the sixth and higher orders due to hextuple excitations. On the other hand, any approximation to the linked-diagram theorem (equation 27) is size extensive. This means that even second order perturbation theory can be much better than very good CI calculations for sufficiently large molecules.

For small molecules, multireference CI techniques that incorporate at least the most important quadruple excitations as double excitations from a double

excitation reference space will be size extensive for most practical purposes. GVB calculations are size extensive, but GVBCI will be only approximately size extensive unless all excitations into the GVB orbitals are included. Since GVB provides a better choice of orbitals than SCF, and since one includes higher level excitations than is normal for SCF-based CI approaches, GVBCI will usually be more nearly size extensive than other CI methods. An added advantage is that, within the GVB method, it is often possible to ensure correct separation.

The other important idea developed in many-body theory is that of the cluster expansion of the wavefunction. The basic concept is that the correct wavefunction may be written as $e^T|\phi_0\rangle$, where T is an operator. This form of the wavefunction ensures the linked-diagram, size extensive basis of the theory.

T has the form

$$T = T_1 + T_2 + T_3 + \cdots, \tag{28}$$

where T_1, T_2, ... are one-body, two-body, ... cluster operators. The T_2 operator generates double excitations with amplitudes to be determined by the coupled-cluster equations,[5-7] but the exponential form

$$e^T = T + \tfrac{1}{2}T^2 + \tfrac{1}{3}!T^3 + \cdots \tag{29}$$

causes some very different things to happen than in the CI approach. To see this, consider the CI operator for quadruple excitations, C_4. By equating the CI and coupled-cluster expressions for the quadruple excitations, we have

$$C_4 = T_4 + \tfrac{1}{2}T_2^2 + \tfrac{1}{4}T_1^4 + \tfrac{1}{2}T_1^2T_2 + T_1T_3. \tag{30}$$

Physically, what does this mean? Roughly, T_4 represents an interaction among four electrons while T_2^2 represents two simultaneous interactions of two electrons. A transformation to Brueckner orbitals makes T_1 vanish, while T_1 is usually small, even for SCF orbitals, so the final three terms are negligible most of the time. Since the normal electrostatic Hamiltonian involves only one- and two-electron operators, simultaneous two-electron interactions would seem to be much more frequent in molecules than "true" four-electron interactions. From another viewpoint, the NH_3 lattice problem emphasizes the neglect of simultaneous double excitations on different H_2 molecules, which is exactly what T_2^2 offers. Thirdly, from perturbation theory, it may be shown that all the fourth order quadruple excitation terms arise from T_2^2,[7] with T_4 only contributing in fifth and higher orders. Consequently, it was suggested by Sinanoglu that $C_4 = \tfrac{1}{2}T_2^2$ is a very good approximation.[14] Using this ansatz,[14] we have the coupled-cluster doubles (CCD) approximation for the wavefunction, $e^{T_2}|\phi_0\rangle$.[5] This leads to a set of nonlinear equations for the T_2 amplitudes, but there are only as many of these amplitudes as there are in a DCI.[5,7] This provides the benefit that we have a size extensive method: it is of infinite order, although restricted to T_2, and we have no more amplitudes than in a DCI calculation, even though the effect of quadruple excitations is included. Since, in fourth order perturbation theory, CCD reduces to double and quadruple excitation diagrams, it is straightforward to solve the CCD equations as successive iterations of a fourth order MBPT calculation.[7] Hence, coupled-cluster methods may be viewed as complementary to MBPT when higher order corrections are needed, as in pathological cases.[9]

Illustration of MBPT/CCM Results for Some Small Molecules

The simplest approximation to the correlation energy in MBPT (assuming an SCF reference function for simplicity) is given by the second order perturbation theory expression of (22).

Since the molecular orbitals $i, j, \ldots$ and $a, b, \ldots$ and their orbital energies $\epsilon_i, \epsilon_j, \ldots \epsilon_a, \epsilon_b, \ldots$ are obtained from an SCF calculation, all the necessary information is available for a correlated calculation. The SCF calculation generates a set of two-electron integrals relative to atomic (i.e., primitive) basis functions and, in the general case, either an integral transformation, which depends on the number of basis functions as n^5, is required to obtain the integrals relative to the molecular orbitals, i.e., $\langle ab \parallel cd \rangle$, or, alternatively, a direct calculation of E_2 and E_3, in terms of the integrals relative to atomic orbitals (probably orthogonalized) is required. In the case of E_2, however, only a very small number of integrals are

Table 2
Percentage Contribution of Different Orders of Perturbation Theory

Molecule*	Second-Order	Third-Order	Fourth-Order (DQ)	Higher Order (DQ)
BH_3	80.0	16.5	3.0	0.50
H_2O	97.7	1.5	0.7	0.06
NH_3	94.3	5.0	0.6	0.12
CH_4	89.6	9.3	0.9	0.16
CO	100.0	−1.6	1.6	−0.09
CO_2	103.2	−4.1	0.9	0.0
HCN	98.0	0.7	1.0	0.18
N_2	101.0	−2.2	1.3	−0.11
C_6H_6	95.6	0.6	3.8	—
CH_3CN	96.5	−2.2	3.7	—
B_2H_6	85.2	13.2	1.6	—

*In every example except C_6H_6, the basis set is at least of double zeta plus polarization quality. In C_6H_6, a double zeta basis is used.

required, since each integral involves only two occupied and two unoccupied orbitals. Consequently, E_2 requires no more than $n_{occ}^2 n_{exc}^2 < n^4$ operations; fewer operations than are required in the SCF calculation itself. In a sufficiently large molecule, one in which the primitive integrals $(\alpha\beta \mid \gamma\delta)$ vanish unless α and β are in the same neighborhood as γ and δ, and unless the charge distributions $(\alpha\beta)$ and $(\gamma\delta)$ are not too far apart, the SCF calculations goes as $\sim n^2$. In this case, evaluating E_2 directly in terms of $(\alpha\beta \mid \gamma\delta)$ will permit a similar simplification; hence, E_2 can always be evaluated as a by-product of large SCF calculations at negligible additional expense.

This approximation is certainly recommended by convenience, but how reliable an approximation is it for the correlation energy? In Table 2 are shown the fractions of the correlation energy within a basis set given by E_2, E_3 and the fourth order contributions just from double and quadruple excitation diagrams for a variety of molecules.[7,8] Using an SCF starting point, E_2 and E_3 are solely

determined by double excitations, with single, double, triple, and quadruple excitations appearing in the fourth order, but, in the interest of also comparing the higher order corrections obtained by the CCD (coupled-cluster doubles) approximation, which includes only double excitations and the disjoint (i.e., T_2^2) quadruple excitations to all orders, the single and triple excitation contributions are omitted from TABLE 2.

It is apparent from the table that the simple second order approximation accounts for the vast majority of the correlation energy obtainable within the basis set. A few generalizations about the results may be made. In multiply bonded systems such as N_2, CO, and CO_2, E_2 tends to slightly overestimate the net correlation energy in the basis set, while, for saturated systems like H_2O, CH_4, etc., it is more likely to underestimate the effect. HCN and benzene are intermediate. In a case where near-degeneracy plays a role, such as BH_3, convergence of the perturbation theory is much slower, making E_2 a poorer approximation. No particular differences are observed for open-shell molecules when a UHFSCF instead of an RHFSCF solution is used as the unperturbed solution. On the average, it is clear that E_2 accounts for ~90% of the correlation energy obtainable within the basis set. Since these basis sets are good enough that they account for ~70% of the "experimental" valence shell correlation energy,[9] E_2 gives ~60% of the experimental valence correlation energy. It is also clear that DQ-MBPT(4) is usually very close to the infinite order CCD model.[9] This is a common occurrence, except in cases where near-degeneracy is a problem.

In TABLE 3, we show some thermochemical results obtained from E_2 compared to higher order correlation approximations. Although E_2 predictions are not as good as the better approximations, they are clearly superior to the SCF predictions, as they provide most of the observed correlation corrections.

A similar result can be obtained for second order predictions of molecular structure, where, on the average, 50% of the error in the SCF geometries is removed.[11]

To obtain the exceptional accuracy reflected in TABLE 3 and elsewhere for various properties of small molecules,[53,54] it is necessary to go beyond the second order, but, for large molecules, the simplicity and comparatively high accuracy of this approximation demands that it be used to augment any large-scale SCF calculation of biochemical interactions.

Benzene serves as a prototype of many of the large conjugated molecules that occur in biochemistry. As such, it is appropriate to analyze the higher order MBPT/CCM description of electron correlation in benzene to develop some feeling for the effect of quadruple excitations.

The basis set is a standard Dunning double zeta contraction of Huzinaga's 9s5*p* primitive Gaussian basis for carbon and the two H 1*s* orbitals corresponding to a Slater exponent of 1.2, giving 72 CGTO (Contracted Gaussian Type Orbitals). The SCF energy of −230.6369 differs by 0.113 atomic units (a.u.) from the SCF result for a DZP basis and 0.18 a.u. from the estimated SCF limit.[55] The C($1s^2$) electrons are kept frozen at the SCF level so the correlated calculations only pertain to the valence correlation energy. Polarization functions are usually found to be more important for correlation effects than in the SCF calculation itself, so the current DZ predictions should underestimate the magnitude of the

valence correlation energy. Even so, it is apparent from TABLE 4 that quadruple excitations amount to ~20% of the predicted correlation energy.

In an attempt to study the origin of the quadruple excitation effects, the occupied and excited pi-orbitals were removed and the calculation repeated to give a value for just the sigma framework excited solely into unoccupied sigma orbitals. The same procedure was carried out for the pi-electrons. These results are reported in the second and third columns of TABLE 4.

TABLE 3

COMPARISON OF THERMOCHEMISTRY RESULTS OBTAINED BY SCF AND MBPT WITH EXPERIMENT*

Reaction	Method	$-\Delta E$ (kcal/mole) SCF	E_2	MBPT/CCD	Experiment
$2\,BH_3 \rightarrow B_2H_6$†	SDQ-MBPT(4)	18.5	37.5	35.6	36.6 ± 2**
$BH_3 + CO \rightarrow H_3BCO$†	D-MBPT(4)	8.0	25.1	20.5	20.4 ± 2**
$BH_3 + NH_3 \rightarrow H_3BNH_3$†	D-MBPT(4)	*20.5*	*32.0*	30.1	—
$HNC \rightarrow HCN$‡	SDQ-MBPT(4)	10.7	18.0	15.0	(10.3 ± 1)††
$HNC \rightarrow [H_N^C]$‡	SDQ-MBPT(4)	−33.4	−30.1	−29.5	—
$BNC \rightarrow BCN$‡	SDQ-MBPT(4)	−18.9	−9.8	−9.4	—
$LiNC \rightarrow LiCN$‡	SDQ-MBPT(4)	−6.4	−2.3	−3.8	—
$CH_3NC \rightarrow CH_3CN$§	SDQ-MBPT(4)	19.2	26.2	22.8	23.7 ± 0.14‡‡
$CH_3NC \rightarrow [CH_2^N]$§	SDQ-MBPT(4)	−44	−40	−40	−38.4§§
$H + CO \rightarrow HCO$‖	CCD	4.8	11.8	13.6	15.7 ± 1.5‖‖
$HCO \rightarrow [HCO]$‖	CCD	−12.8	−17.4	−18.1	—
$H_2CO \rightarrow H_2 + CO$¶	CCD	−7.2	−3.9	−3.9	−1.9¶¶
$H_2CO \rightarrow H + HCO$¶	CCD	−68.6	−82.8	−86.0	−86.0 ± 1.0***

*All basis sets are at least DZP quality.

†Reference 53.

‡Reference 55. The square bracket indicates a transition state. This result includes a 4 kcal/mol zero point correction for the transition state.

§Reference 54. The square bracket indicates a transition state. This result includes a 4.8 kcal/mol zero point correction for the transition state.

‖Reference 56.

¶Reference 57.

**Reference 58.

††Reference 59. Reference 55 concludes that this experimental value is in error. The result should be 15 ± 2 kcal/mol.

‡‡Reference 60.

§§Reference 61.

‖‖Reference 62.

¶¶Reference 63.

***Reference 64.

The sigma framework accounts for over half the net quadruple excitation effect, while the correlation effects of the delocalized pi-electrons are relatively independent of the quadruple excitations. Since the former involves 12 roughly independent covalent bonds, one would estimate, from TABLE 1 and **(14)**, an effect of ~14%, which is in reasonable agreement with the calculated 12%. The effect of the quadruple excitations on the pi-electron bonds is much smaller, but this is

primarily due to the fact that only three bonds are possible. If the appropriate S for the pi-structure were as small as for He, the estimated effect of quadruple excitations would be 1.6%. The remaining correlation effects come from the sigma-pi interactions. It is interesting that configurations corresponding to excitations of sigma electrons into pi-excited orbitals and vice versa contribute ~13% of the correction energy.

The DZ basis used here is capable of providing only about 49% of the experimental valence correlation energy.[65,66] Polarization functions would improve this result by about 20%.[9] Since the quadruple excitations are responsible for more than 20% of the correlation energy, the size of error encountered in SDCI is as severe as excluding polarization functions from the basis set. Since the effects of quadruple and higher CI excitations will inevitably increase as

TABLE 4

EFFECT OF QUADRUPLE EXCITATIONS ON THE CORRELATION ENERGY OF BENZENE*

Method	All valence electrons	Sigma electrons (only)	Pi electrons (only)
E_{SCF}	−230.6369	—	—
E_2	−0.4922	−0.3622	−0.0539
E_3	−0.0030	−0.0005	−0.0177
S	−0.0028	−0.0021	−0.0002
D	−0.0272	−0.0123	−0.0079
Q	+0.0079	+0.0074	+0.0012
E_4(SDQ)	0.0229	−0.0070	−0.0069
SDQ-MBPT(4)	−0.5172	−0.3697	−0.0785
Experimental Valence Correlation Energy†	(−1.065)		
CCD	—	−0.3671	−0.0850
SD-RSPT(4) ~ SD-CI	−0.4160	−0.3259	−0.0774
Net effect of quadruple excitations	−0.1012 (20%)	−0.0438 (12%)	−0.0011 (1%)

*All energies are in Hartree atomic units.
†Reference 65.

molecules larger than benzene are studied, a correlation method that does not account for such effects has little chance of contributing to the investigation of biochemical interactions.

I would like to thank Dr. H. Weinstein and Dr. J. P. Green for the opportunity to speak at this meeting.

SUMMARY

One of the goals of quantum chemistry in the biomedical sciences is to provide accurate calculations of molecular interactions among biochemical molecules, drugs, carcinogens, etc. In this effort, there is a natural progression

from semiempirical quantum chemistry to ab initio self-consistent field theory to methods that properly include electron-correlation. As ab initio theories continue to develop, many more problems of biomedical interest can be addressed by accurate correlated methods. In this article were discussed many-body approaches to the correlation problem, i.e., many-body perturbation theory and coupled-cluster methods. Unlike most configuration interaction methods, MBPT/CCM offers a number of important features in its extension to larger molecules. These include the proper dependence of the correlated calculation on the size of the molecule (i.e., size extensivity). This has significant consequences for predictions of ground and excited-state properties. These features were illustrated by applications to selected molecules. It was demonstrated that MBPT/CCM offers a natural generalization of SCF theory that is formally suitable for applications to some of the molecules that occur in biomedical studies.

References

1. Brueckner, K. A. 1955. Phys. Rev. **97:** 1353; **100:** 36; Brueckner, K. A., R. J. Eden & N. C. Frances. 1955. Ibid. **98:** 1445; Brueckner, K. A. & C. A. Levinson. 1955. Ibid. **97:** 1344.
2. Goldstone, J. 1957. Proc. R. Soc. London Ser. A **239:** 267.
3. Kelly, H. P. 1969. Adv. Chem. Phys. **14:** 129; 1963. Phys. Rev. **131:** 684; 1964. Phys. Rev. B **136:** 896.
4. Coester, F. 1958. Nucl. Phys. **1:** 421; Coester, F. & H. Kümmel. 1960. Ibid. **17:** 477; Kümmel, H. 1969. Ibid. **22:** 177.
5. Čížek, J. 1966. J. Chem. Phys. **45:** 4256; 1969. Ibid. **14:** 35; Čížek, J., J. Paldus & L. Sroubkova. 1969. Int. J. Quantum Chem. **3:** 149.
6. Paldus, J., J. Čížek & I. Shavitt. 1972. Phy. Rev. A. **5:** 50; Paldus, J. & J. Čížek. 1973. *In* Energy, Structure, and Reactivity. D. W. Smith and W. B. McRae, Eds.: 389. Wiley. New York.
7. Bartlett, R. J. & G. D. Purvis. 1978. Int. J. Quantum Chem. **14:** 561.
8. Bartlett, R. J. & I. Shavitt. 1978. Chem. Phys. Lett. **50:** 190; 1978. Ibid. **58:** 157.
9. Bartlett, R J. & G. D. Purvis. 1980. Phys. Scr. **21:** 255.
10. Bartlett, R. J. & D. M. Silver. 1974. Int. J. Quantum Chem. Symp. **8:** 271; 1974. Phys. Rev. A **10:** 1927; 1975. J. Chem. Phys. **62:** 3258; 1976. Ibid. **64:** 1260; 1976. Ibid. **64:** 4578; 1975. Int. J. Quantum Chem. Symp. **9:** 183.
11. Pople, J. A., J. S. Binkley & R. Seeger. 1976. Int. J. Quantum Chem. Symp. **10:** 1; Pople, J. A., R. Seeger & R. Krishnan. 1976. Int. J. Quantum Chem. Symp. **11:** 149.
12. Pople, J. A., R. Krishnan, H. B. Schlegel & J. S. Binkley. 1978. Int. J Quantum Chem. **14:** 545.
13. Robb, M. A. 1973. Chem. Phys. Lett. **20:** 274.
14. Sinanoğlu, O. 1962. J. Chem. Phys. **36:** 706, 3198; 1964. Adv. Chem. Phys. **6:** 315.
15. Freeman, D. L. & M. Karplus. 1976. J. Chem. Phys. **64:** 2641.
16. Pullman, B. & A. Pullman. 1963. Quantum Biochemistry. Wiley. New York.
17. Green, J. P. & S. Kang. 1970. *In* Molecular Orbital Studies in Chemical Pharmacology. L. B. Kier, Ed: 105. Springer-Verlag. New York.
18. Weinstein, H., S. Maayani, S. Srebrenik, S. Cohen & M. Sokolovsky. 1973. Mol. Pharmacol. **9:** 820.
19. Weinstein, H., S. Srebrenik, R. Paungz, S. Maayani, S. Cohen & M. Sokolovsky. 1974. *In* Chemical and Biochemical Reactivity, Proc. 6th Jerusalem Symp. E. D. Bergman and B. Pullman, Eds.: 493.
20. Kaufman, J. J. & W. S. Koski. 1975. *In* Drug Design, Vol. 5. E. J. Ariens, Ed.: 251. Academic Press. New York.

21. PULAY, P. 1977. In Modern Theoretical Chemistry, Vol. 4. H. F. Schaefer, III, Ed.: 153. Plenum Press, New York; 1969. Mol. Phys. **17:** 197; 1970. Ibid. **18:** 473.
22. MCIVER, J. W. & A. KOMORNICKI. 1979. Chem. Phys. Lett. **10:** 303; 1972. J. Am. Chem. Soc. **94:** 2625.
23. DUPUIS, M. & H. F. KING. 1978. J. Chem. Phys. **68:** 3998.
24. POPLE, J. A., R. KRISHNAN, H. B. SCHLEGEL & J. S. BINKLEY. 1979. Int. J. Quantum Chem. Symp. **13:** 225.
25. GODDARD, J. D., N. C. HANDY & H. F. SCHAEFER, III. 1979. J. Chem. Phys. **71:** 1525.
26. KOTA, S. & K. MOROKUMA. 1979. Chem. Phys. Lett. **65:** 19.
27. SCHERAGA, H. A. 1971. Chem. Rev. **71:** 195.
28. CLEMENTI, E., F. CAVALLONE & R. SCORDAMAGLIA. 1977. J. Am. Chem. Soc. **99:** 5531; SCORDAMAGLIA, R., F. CAVALLONE & E. CLEMENTI. 1977. Ibid. **99:** 5545; BOLIS, G. & E. CLEMENTI. 1977. Ibid. **99:** 5550.
29. SWAMINATHAN, S. & D. L. BEVERIDGE. 1977. J. Am. Chem. Soc. **99:** 8392.
30. WEINSTEIN, H., D. CHON, S. KANG, C. L. JOHNSON & J. P. GREEN. 1976. Int. J. Quantum Chem. Quantum Biol. Symp. **3:** 135.
31. KAUFMAN, J. J. 1979. Int. J. Quantum Chem. **16:** 221.
32. MCCULLOUGH, E. A. 1974. Chem. Phys. Lett. **24:** 55; 1975. J. Chem. Phys. **62:** 3991; 1975. Ibid **63:** 5050.
33. JOHNSON, K. H. 1967. Int. J. Quantum Chem. Symp. **1:** 361; SMITH. F. C., Jr. & K. H. JOHNSON. Phys. Rev. Lett. **22:** 1168; JOHNSON, K. H. & F. C. SMITH, Jr. 1971. Int. J. Quantum. Chem. Symp. **5:** 429.
34. KAHN, L. R., P. BAYBUTT & D. TRUHLAR. 1976. J. Chem. Phys. **65:** 3826. KAHN, L. R. & W. A. GODDARD, III. 1968. Chem. Phys. Lett. **2:** 667; 1972. J. Chem. Phys. **56:** 2686.
35. MELIUS, C. F. & W. A. GODDARD, III. 1974. Phys. Rev. A **10:** 1528.
36. TOPIOL, S., M. A. RATNER & J. W. MOSKOWITZ. 1977. Chem. Phys. **70:** 1; 1977. Chem. Phys. Lett. **46:** 256.
37. EWIG, C. S., P. COFFEY & J. R. VAN WAZER. 1976. Inorg. Chem. **14:** 1848; OSMAN, R. C., S. EWIG & J. R. VAN WAZER. 1976. Chem. Phys. Lett. **39:** 27.
38. OSMAN, R. & H. WEINSTEIN 1977. Chem. Phys. Lett. **49:** 69.
39. NELSON, J. L. & A. A. FROST. 1972. J. Am. Chem. Soc. **94:** 3727.
40. CHRISTOFFERSEN, R. E., D. SPANGLER, G. G. HALL & G. M. MAGGIORA. 1973. J. Am. Chem. Soc. **95:** 8526; DAVIS, T. D., R. E. CHRISTOFFERSEN & G. M. MAGGIORA. 1975. Ibid. **97:** 1347.
41. GODDARD, W. A., III, T. H. DUNNING, Jr., W. J. HUNT & P. J. HAY. 1973. Acc. Chem. Res. **6:** 368.
42. DUNNING, T. H. & V. MCKOY. 1967. J. Chem. Phys. **47:** 1735; SHIBUYA, T. & V. MCKOY. 1970. Phys. Rev. A **2:** 2208.
43. SIMONS, J. & W. D. SMITH. 1973. J. Chem. Phys. **58:** 4899; SIMONS, J. & P. JØRGENSEN. 1976. Ibid **64:** 1413.
44. HARRIS, F. E. 1977. Int. J. Quantum Chem. Symp. **11:** 403.
45. BRANDOW, B. H. 1967. Rev. Mod. Phys. **39:** 771.
46. LINDGREN, I. 1978. Int. J. Quantum Chem. Symp. **12:** 33.
47. HOSE, G. & U. KALDOR. 1979. J. Phys. B **12:** 3827; 1980. Phys. Scripta **21:** 357.
48. MALRIEU, J. P. 1979. J. Chem. Phys. **70:** 4405.
49. MEUNIER, A. & B. LEVY. 1979. Int. J. Quantum Chem. **16:** 953.
50. DAVIDSON, E. R. & L. L. JONES. 1962. J. Chem. Phys. **37:** 2966.
51. KOLOS, W. C. & C. J. ROOTHAAN. 1960. Rev. Mod. Phys. **32:** 219.
52. LÖWDIN, P. O. 1965. J. Math. Phys. **6:** 1341; 1964. Phys. Rev. A **139:** 357.
53. REDMON, L. T., G. D. PURVIS & R. J. BARTLETT. 1979. J. Am. Chem. Soc. **101:** 2856.
54. REDMON, L. T., G. D. PURVIS & R. J. BARTLETT. 1978. J. Chem. Phys. **69:** 5386.
55. REDMON, L. T., G. D. PURVIS & R. J. BARTLETT. 1980. J. Chem. Phys. **72:** 986.
56. ADAMS, G. F., G. D. BENT, G. D. PURVIS & R. J. BARTLETT. 1979. J. Chem. Phys. **71:** 3697.
57. ADAMS, G. F., G. D. BENT, R. J. BARTLETT & G. D. PURVIS III. 1981. J. Chem. Phys. In press.
58. FEHLNER, T. P. & G. W. MAPPES. 1969. J. Phys. Chem. **73:** 873.
59. MAKI, L. Unpublished results.

60. BAGHAL-VAYJOVEC, M. H., J. L. COLLISTER & H. O. PRITCHARD. 1977. Can. J. Chem. **55**: 2634.
61. SCHNEIDER, F. W. & B. S. RABINOVITCH. 1969. J. Am. Chem. Soc. **85**: 1794.
62. WARNECK, P. 1971. Z. NATURFORSCH. Teil A **27**: 2047.
63. YAMADA, K., T. NAGAKURU, K. KUCHISTU & Y. MORINO. 1971. J. Mol. Spectrosc. **38**: 70.
64. WALSH, R. & S. W. BENSON, 1966. J. Am. Chem. Soc. **88**: 4570.
65. ERMLER, W. C. & C. W. KERN. 1973. J. Chem. Phys. **58**: 3458.
66. NESBET, R. K. 1968. Phys. Rev. **175**: 2.

DISCUSSION

W. GODDARD (*California Institute of Technology, Pasadena*): That was really great work. I might point out that there is one approach, the generalized valence bond plus CI, which is size extensive and which does go to the right limits.

M. KARPLUS (*Harvard University, Cambridge, Mass.*): Is the percentage of the correlation energy that you presented calculated with reference to the best that one obtains with the given basis set or does it really refer to the correlation energy of the system?

BARTLETT: The basis set. These are fractions of the total correlation energy determined in a given basis set, usually by coupled cluster calculations. This is the best model we have right now. It's a question of fractions within a basis set.

J. KAUFMAN (*Johns Hopkins University, Baltimore, Md.*): We're interested in reactive surfaces for large molecules, for which we either have to do configuration interaction or use many-body pertubation theory on multi-reference determinants. Can you give us an estimate of when these methods are going to be usable?

BARTLETT: Within a year, I suspect.

KAUFMAN: How do you think they'll compare to potential energy surfaces obtained from CI calculations?

BARTLETT: We are essentially comparing single reference many-body approaches with single reference CI approaches. But these approaches have advantages of size, extensiveness, etc. The same thing will apply to the multi-reference approaches. There will be a linked diagram theorem, an analogue of the ground state theorem that will apply to excited states. It will eliminate the problem of getting transition energies and excitation energies that go to zero as the number of atoms goes to infinity. You have to use these types of methods because they're formally suitable for larger molecules. It's quite another thing to say that they will be applicable to chlorophyll or something that large because of the basis set problem.

E. CLEMENTI (*IBM, Poughkeepsie, N.Y.*): Could you compare second order perturbation theory with the density functional approach? In the latter, you could

also get the correlation energy plus or minus 20%. And you'll also have to have more than one reference state. The density functional approach, however, is very cheap: it costs roughly one iteration, and you can go to large molecules.

GODDARD: There are all sorts of problems with the density functional approach. If you want to look at potential surfaces, you don't necessarily go to the right limits at infinity; you have a bias towards high spin states when you start to break bonds. All sorts of things blow up, although impressive results were obtained for geometries. You don't know what the correlation is in the density functional approach because you don't get correct excitation energies. How can you talk about percentage of correlation there? The real test will come when they actually try open shell systems. Ozone would be a good case. The approach is well known for atoms, and they make errors of one or two eV and also errors in excitation energies.

CLEMENTI: I think it's on the order of 0.2 eV from helium to an atomic number of 54.

GODDARD: At a symposium in Paris a couple of weeks ago, W. L. Jones presented results for some transition metal atoms showing excitation energies and how you get the wrong order for states within an eV.

J. SCHULMAN (*Queens College—CUNY, New York*): Since these errors can be plus or minus relative to the complete many-body result, wouldn't it be rather risky to use second order energies for thermochemistry? It could accentuate, rather than cancel, the error.

BARTLETT: Our experience is that it does give most of the answer. It's true that you cannot expect, say, the second order energy for B_2H_6 compared with twice the second order energy for BH_3 to be as accurate for the dissociation energy as our higher order calculations.

SCHULMAN: Suppose that product C and D have greater than 100% error (i.e., 120% correlation), whereas reactants A and B have the lower estimate (i.e., 80%). The total error of reaction is going to be greater, perhaps, than if you did an SCF, isn't it?

BARTLETT: It could be, but our experience is that the SCF would be much further off. See TABLE 3.

SCHULMAN: Do you ever experience any convergence difficulties?

BARTLETT: This can happen if you use a single reference function. Then, anytime you have some near-degeneracies, as with beryllium, you would like to use a $1s^22s^2$ configuration and a $1s^22p^2$ configuration in your reference space. If you try to do the problem without the $2p^2$, it will be comparatively slowly convergent; it will converge, however, and you will get a good answer. But you will have to do more work and go to higher order terms and pay a great deal of attention to the coupled cluster type of contributions.

SCHULMAN: Do you ever converge to the wrong energy?

BARTLETT: That could happen, but I don't remember an example. Coupled

cluster theory itself, because it is done in an iterative fashion, can converge to the wrong energy purely for numerical reasons. If you use superior numerical methods (we use something that we call reduced partitioning procedure) then you get much better convergence of the coupled cluster equation.

M. ZERNER (*University of Guelph, Ontario*): You would have trouble with this when there are near-degeneracies between the reference determinant and other states.

SCHULMAN: That's the point I'm making.

ZERNER: This happens quite often when you get to things like transition state complexes.

BARTLETT: We are now studying insertion of Be into H_2. In this case, we have two reference functions in Be that should be important, and we're breaking the H_2 bonds so that the $(1\sigma_u)^2$ configuration will be degenerate with the $(1\sigma_g)^2$. If we can handle that problem, we think we're in pretty good shape with single reference approaches. But we're also doing that problem with a multi-reference function.

ZERNER: In most cases the reference space will not require more than two determinants.

GODDARD: There might be some good examples for transition metal atoms. Take the case of nickel with the s^1d^9 and the s^2d^8 states, which are experimentally degenerate. You get the triplet D states for both configurations. That would be a terrific case to test.

BARTLETT: It would be a very difficult problem.

A. ROSSI (*University of Connecticut, Storrs*): How would you calculate a potential energy surface using your method? Would you include correlation at each point on the surface?

BARTLETT: Yes. We would do an SCF calculation at each point, UHF if it happened to be an open shell problem, and add on SDQ MBPT or coupled cluster theory at each point.

GODDARD: Then, for something like N_2, do you use UHF only for part of the way? What happens the rest of the way on the potential surface?

BARTLETT: This is a notorious problem. N_2 is very well described with an RHF reference function near equilibrium, but, a little beyond equilibrium, there's a bifurcation into a UHF and an RHF solution. You can converge both solutions. If you want it to separate correctly, you'd use a UHF function, but it's very spin-contaminated. If you take the RHF function, the coupled cluster theory will push it even a little bit farther toward dissociation, but it won't take you all the way to the separated nitrogens. This is the worst possible case for something like UHF because you're dealing with a singlet, so it has as much spin contamination as possible from higher states. As long as you have an open shell problem, you do a UHF calculation and the multiplicity will come out correct to a couple of decimal places, so we're willing to use UHF plus correlation to do thermochemi-

cal calculations in which we'd have to compare open shell species with closed shell species. But we cannot do something like N_2 well without a superior method.

GODDARD: What kind of error does it actually make in the N_2 when you take the coupled cluster?

BARTLETT: In the dissociation energy we are off by about 0.5 eV.

SCHULMAN: At what order would it cease to be worthwhile to do Hartree-Fock at all, and be better to use some linear combination and then do the many-body correction? People have suggested earlier that it might be a relatively low order, that there's a crossing point between the two. Have you looked at that?

BARTLETT: I haven't really looked at that. I would be speculating, but many of the things we can do, such as change our basis sets or sum certain terms to all orders with denominator shifts, become less important after third order in perturbation theory. So you might expect that, if you wanted to do something like *what is done in the PCILO method, where they take a bond orbital and perturb* from it, you might actually be able to avoid Hartree-Fock if you went to fourth order.

SCHULMAN: I notice that, for acetonitrile, the fourth order correction is larger than the third. Is that correct?

BARTLETT: Yes. Since I was using the coupled cluster method as my reference point for these first several calculations, I limited myself to DQ. This means double and quadruple excitation diagrams. The MBPT quadruples are positive because they involve part of the renormalization term. The magnitude of this number will actually become larger when the singles and triples are added to it, and, in fact, it will generally be bigger than third order. That's to be expected because, if we're starting from the SCF calculation, this is the first time that singles, triples, and quadruples have an opportunity to contribute. However, even if we sum some category of terms to high orders, as we do in CCD, we get no significant change. So we believe the results that we get through fourth order, even if the fourth is bigger than the third. It's asking too much to expect the series to be a monotonically decreasing set of numbers.

Formaldehyde: Electronic structure calculations for the S_0 and T_1 states[a)]

George F. Adams and Gary D. Bent[b)c)]

ARRADCOM, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005

Rodney J. Bartlett and George D. Purvis

Battelle Columbus Laboratory, Columbus, Ohio 43201

(Received 4 June 1980; accepted 9 April 1981)

Many-body perturbation theory (MBPT) and coupled-cluster method calculations are reported for the $S_0(X\,^1A_1)$ and $T_1(a\,^3A'')$ electronic states of formaldehyde. The structural parameters for the S_0 minimum ($R_{CH} = 1.102$ Å, $R_{CO} = 1.211$ Å, HCH = 116.2°) and the T_1 minimum ($R_{CH} = 1.085$ Å, $R_{CO} = 1.327$ Å, HCH = 118°, "out-of-plane" angle = 37° 12') agree well with experimentally deduced values. Calculated heats of reaction for dissociation to radical products and molecular products agree well with literature values. The energy barriers for dissociation to molecular products and rearrangement to hydroxycarbene are presented. Vertical and adiabatic transition energies are reported for $S_0 \rightarrow T_1$, while a vertical transition energy for $S_0 \rightarrow S_1$ is reported.

I. INTRODUCTION

Formaldehyde is an important and ubiquitous chemical species. In recent years the photochemistry resulting from excitation to the first excited singlet state S_1 has been the subject of a number of experimental[1-4] and theoretical[5-7] reports which raise a number of important questions about interpretation of experimental data. Formaldehyde also occurs as a crucial species in models that describe the flame-zone chemistry pertinent to the combustion of carbon-containing molecules.[8] In fact, Fifer concludes that the vapor-phase combustion of nitrate esters is dominated by the reaction sequence that describes the oxidation of formaldehyde by nitrogen dioxide.[8]

Numerous reports of theoretical studies on various electronic states of formaldehyde occur in the literature. Recent papers by Goddard and Schaefer (GS),[5] Goddard, Yamaguchi, and Schaefer (GYS),[7(a)] and Bell[6] provide excellent results as well as thorough discussions of previous molecular obital calculations.[9,10] In particular, the results in the GS paper would support a tunneling mechanism[11] for the decomposition of H_2CO to $H_2 + CO$. However, the question of whether this is required to interpret the experimental results is based upon energy differences of only 4 or 5 kcal mol^{-1}. Since the CI calculations of GS[5] and GYS[7(a)] estimate the effects of unlinked diagrams associated with the CI quadruple excitations by the formula of Davidson,[12] and since the estimate amounts to approximately 3 kcal mol^{-1} for formaldehyde, it seems pertinent to study the formaldehyde decomposition with many-body methods that properly treat the effects of unlinked diagrams and quadruple excitations.[13-16] Consequently, dissociation energies and energy barriers are predicted by MBPT/CCD for the dissociation to radical products

$$CH_2O \rightarrow HCO + H\ , \qquad \text{(I)}$$

molecular products

$$CH_2O \rightarrow H_2 + CO\ , \qquad \text{(II)}$$

and the rearrangement to *trans*-hydroxycarbene

$$CH_2O \rightarrow HCOH\ . \qquad \text{(III)}$$

Other important questions pertain to the electronic excitations in formaldehyde. Equations-of-motion calculations for a number of excitation energies have been reported.[17,18] We report here MBPT/CCD calculations for the ground ($X\,^1A_1$, S_0) and first excited ($a\,^3A''$, T_1) electronic states of formaldehyde. The results include equilibrium structural parameters for each state, the adiabatic excitation energy for $S_0 \rightarrow T_1$, and vertical excitation energies for $S_0 \rightarrow T_1$ and $S_0 \rightarrow S_1$.

II. OUTLINE OF THE CALCULATIONS

For the closed-shell states included in this study, the Hartree–Fock representation of a state is given by a restricted Hartree–Fock (RHF) wave function, whereas open-shell states are described by unrestricted Hartree–Fock (UHF) wave functions. The integrals over atomic functions were computed using Dunning's $4s3p$ contraction of Huzinaga's $9s5p$ Gaussian primitive set for first-row atoms, and Dunning's scaled ($\zeta = 1.2$) $3s$ contraction of Huzinaga's $4s$ primitive set for hydrogen.[19,20] A single set of d-type polarization functions augments the atomic basis sets for carbon and oxygen, with orbital exponents 0.75 and 0.85, respectively.[21] A set of p-type polarization functions, with $\alpha = 1.0$, augments the hydrogen basis set.[21] Integral calculations for the MBPT/CCD calculations were performed using the MOLECULE integral program[22]; electronic structure calculations were performed using the GRNFNC and UMBPT programs.[23]

Structural parameters for the formaldehyde conformers were obtained in several ways. Optimal parameters for the equilibrium positions on both the S_0 and T_1 surfaces were determined by fitting the results

a) This research was sponsored in part by the U. S. Army Research Office under grant No. DAAG29-80-C-0105.

b) NAS–NRC Resident Research Associate.

c) Present address: Department of Physics, University of Connecticut, Storrs, Conn. 06268.

 0021-9606/81/140834-09$01.00

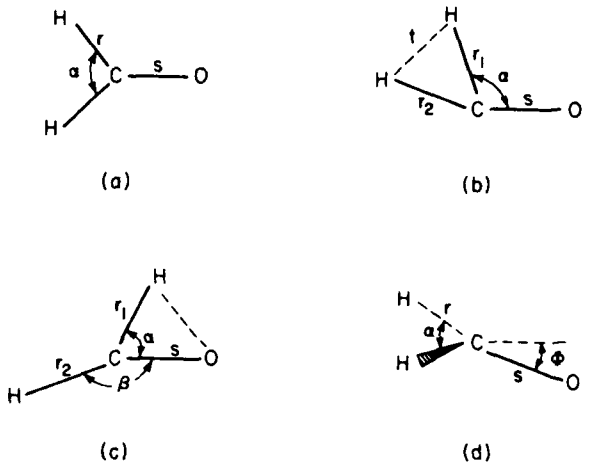


FIG. 1. Definition of structural parameters for formaldehyde calculations: (a) equilibrium, (b) molecular products transition state, (c) rearrangement transition state, (d) equilibrium for T_1 state.

of frozen-core, double-excitation MBPT(4) calculations. Such a procedure is impractical when seeking transition state structures. Structural parameters for saddlepoints on the S_0 surface resulted from analytic gradient based calculations. In particular, we used the GRADSCF computer codes.[24] In order to be consistent in computing energy differences on the S_0 surface, equilibrium structural parameters for formaldehyde were also determined using the GRADSCF codes. Finally, vibrational frequencies were estimated using the gradient codes. The same basis set was used for all calculations. The structural parameters for each configuration are depicted in Fig. 1. Theoretical and experimental structural parameters for each configuration are listed in Table I. We report dissociation energy and energy barrier results for several levels of computation: SCF; fourth-order MBPT including all double-excitation diagrams D-MBPT(4); fourth-order MBPT including all single-, double-, and quadruple-excitation diagrams SDQ-MBPT(4); and coupled cluster double calculations, CCD based upon the wave function $\exp(T_2)\Phi_0$. Excitation energies are reported at both SCF and MBPT/CCD levels of computation. All reported energies include the core electrons.

TABLE I. Structural parameters for formaldehyde (S_0).[a]

Equilibrium configuration	r	s	α
SCF			
This work	1.092	1.182	115.9
DZ[b]	1.084	1.217	116.8
DZ + P[b]	1.096	1.189	116.3
DZ + P CI[c]	1.100	1.212	116.2
D-MBPT(4)	1.102	1.211	116.2
Experiment[d]	1.099	1.203	116.5

Molecular products transition state	r_1	r_2	s	α	t
SCF					
This work	1.606	1.097	1.141	113.3	1.219
DZ + P[b]	1.586	1.104	1.151	112.5	1.203
DZ + P CI	1.572	1.096	1.179	110.8	1.213

Rearrangement transition state	r_1	r_2	s	α	β
SCF					
This work	1.213	1.095	1.266	56.4	116.6
DZ + P[b]	1.262	1.087	1.316	55.0	117.3
DZ + P CI	1.232	1.106	1.303	55.2	115.0

[a]Units: angstroms, degrees.
[b]Reference 5.
[c]Reference 7(a).
[d]Reference 24.

TABLE II. Electronic structure calculations for $CH_2O(X\,^1A_1)$.[a]

Level of theory	Structure SCF	Structure D-MBPT(4)
SCF	−113.8990	−113.8974
D-MBPT(2)	−114.2672	−114.2692
D-MBPT(3)	−114.2730	−114.2738
D-MBPT(4)	−114.2818	−114.2832
SDQ-MBPT(4)	−114.2814	−114.2830
CCD	−114.2760	−114.2772

[a]Units: hartree.

To compute dissociation energies and energy barriers requires accurate calculation of energy differences. In addition to the electronic energy, vibrational, rotational, and translational energy differences must be estimated. We assume that the rotational and translational degrees of freedom are classical. Vibrational frequencies for the various formaldehyde conformers were predicted using SCF-gradient calculations.[24]

III. THE $S_0(X\,^1A_1)$ STATE OF FORMALDEHYDE

At its ground state equilibrium configuration, formaldehyde possesses C_{2v} symmetry. The structural parameters obtained by several theoretical calculations are compared with experimentally[25] derived values in Table I. The agreement between theory and experiment is excellent, especially for the CI calculations[7(a)] and the MBPT results.

To determine chemical energy differences requires a high-quality *ab initio* calculation of the electronic energy of those species involved in the chemical reaction. The results of several levels of electronic structure calculation for the $S_0(X\,^1A_1)$ state of formaldehyde are listed in Table II. The Hartree–Fock energy for the structure determined using the gradient-based SCF calculations is lower than that corresponding to the structure found by fitting the results of D-MBPT(4) calculations. However, through fourth order, the MBPT calculations yield lower energies for the structure obtained using the D-MBPT(4) structure. In addition, the results of CCD calculations predict a lower energy for the structure determined using the D-MBPT(4) structure. Each of the SCF energies reported here is a few millihartrees lower than the single-configuration DZ + P result reported by Goddard and Schaefer.[5] We expect, therefore, that our basis set is comparable

in quality to that used by GS in their initial study of the S_0 surface of formaldehyde.[5] The two fourth-order MBPT calculations, D-MBPT(4) and SDQ-MBPT(4), agree almost exactly with one another. Clearly the contribution of the single- and quadruple-excitation diagrams almost cancel at fourth order. The fourth-order triple excitation contributions would lower the electronic energies. The CCD calculation, which includes higher order effects of double and quadruple excitations, but neglects single- and triple-excitation effects, predicts a smaller contribution of electron correlation than do the fourth-order MBPT calculations. This is due to the neglect of single- and triple-excitation effects, both of which are negative. Although CCD is not rigorously variational, since the principal terms neglected in the calculation are negative, it is quasivariational and yields the lowest energy to date for the S_0 state of formaldehyde.

IV. DISSOCIATION REACTIONS OF FORMALDEHYDE

Thermal dissociation of formaldehyde is assumed to produce radical products, hydrogen atom and formyl radical.[26] Interestingly, the dissociation to molecular products H_2 and CO is approximately thermoneutral, while the observed thermal dissociation energy is approximately 88 kcal mol^{-1}.[27] Both sets of products correlate by symmetry with the ground electronic state.

The first singlet excited state of formaldehyde $[S_1(A\,^1A'')]$, has an excitation energy threshold lower than the threshold for thermal dissociation.[1] Photochemical studies of reactions subsequent to excitation of formaldehyde to the S_1 state are interpreted in terms of the following three processes[1]:

$$H_2CO(S_1, \nu') \rightarrow H_2 + CO\ , \qquad \text{(IV)}$$

$$\rightarrow H + HCO\ , \qquad \text{(V)}$$

$$\rightarrow H_2CO(S_0, \nu'') + h\nu\ . \qquad \text{(VI)}$$

Most investigators agree that the formation of molecular products (IV) dominates for excitation energies near the S_1 threshold (3.495 eV, 80.5 kcal mol^{-1}). The importance of the radical products channel (V) increases with increasing excitation energy.[2] For an excitation energy equal to 91.3 kcal mol^{-1}, the sum of quantum yields for processes (IV) and (V) is almost one, with the molecular quantum yield 0.32 ± 0.03 and the radical quantum yield 0.68 ± 0.03.[2]

The crucial datum in understanding the photochemical dissociation of formaldehyde is the energy difference between the S_1 threshold (80.5 kcal mol^{-1}) and the potential energy barrier to the formation of $H_2 + CO$ on the S_0 surface. The appearance of carbon monoxide is much slower than the decay rate for S_1, and the formation of carbon monoxide requires a collision. This behavior suggests that the high-pressure (>0.1 Torr) photochemistry of formaldehyde excited to S_1 involves a transition to an intermediate species or to a different electronic state of formaldehyde. The identity of the intermediate(s) remains a point of contention, but the results of the CI calculations by GS[5] and GYS[7(a)] imply that the intermediate occurs on the S_0 surface.

Goddard and Schaefer reported extensive configuration interaction calculations, including all single- and double excitations relative to their reference functions, that elucidated many features of the potential energy hypersurface that pertain to the photodissociation of formaldehyde.[5] Subsequently, GYS[7(a)] reinvestigated the S_0 surface using analytic gradient methods that included correlation effects. The calculations addressed reactions (IV) and (V) and the isomerization of formaldehyde to form *trans*-hydroxycarbene, all on the S_0 potential energy surface. The SD-CI study[5] provided improved values for energy barriers to photochemical reactions of formaldehyde. A significant result of the study was the prediction that the energy required for isomerization to hydroxycarbene is only slightly greater than the energies needed for radical or molecular photodissociation.[5] The SD-CI calculations also predicted a small energy barrier above the S_1 origin to any photochemical reaction of formaldehyde.

In subsequent research Goddard, Yamaguchi, and Schaefer[7(a)] reported features of the formaldehyde potential energy surface using analytic configuration interaction techniques.[28,29] The results of that research differ little from the results obtained by GS.[5] A noteworthy conclusion of the GYS work is that the results of CI calculations using the SCF optimized geometry differs little from those obtained using the CI determined transition state. For a double zeta plus polarization basis set, the two predictions of the molecular products dissociation barrier differ by only 0.8 kcal mol^{-1}.

Miller[11] employed the barrier height and vibrational frequencies predicted by GS[5] to calculate the collisionless decay rate for $S_0 - H_2$ to CO as a function of energy. The results of Miller's analysis suggest that S_0 can decay rapidly to molecular products even at energies as much as 10 kcal mol^{-1} lower than the barrier height. Since that analysis, however, additional studies of the S_0 surface by Harding *et al.*,[30] Goddard *et al.*,[7(a)] and this work have modified the original estimate of the vibrational frequencies for the rearrangement transition state.[5] The effect of these studies is to reduce the barrier height for the rearrangement of formaldehyde to *trans*-hydroxycarbene relative to the barrier for molecular products dissociation.

A. $CH_2O \rightarrow CHO + H$

The theoretical investigation of the formyl radical using MBPT and CCD has been published.[31] We report in Table III the results of those calculations that pertain to the study of the dissociation of formaldehyde. The experimental zero-point vibrational energies of formaldehyde and formyl radical are 16.2[32] and 7.8

TABLE III. Electronic structure results for $HCO(X\,^2A')$.[a]

	UHF	D-MBPT(4)	SDQ-MBPT(4)	CCD
Energy	−113.2769	−113.6343	−113.6364	−113.6288

[a]Units: hartree.

TABLE IV. Heat of reaction for the dissociation of formaldehyde to radical products.[a]

Level of calculation	ΔE_{el}	ΔE_v	ΔE_R
SCF	77.0	−8.4	68.6
D-MBPT(4)	94.8	−8.4	86.4
SDQ-MBPT(4)	93.4	−8.4	85.0
CCD	94.5	−8.4	86.1
SD-CI+QC[b]	87.4	−8.3	79.1
Experiment: Ref. 3			86.0±1.0
Experiment: Ref. 27			86.7±1.5

[a]Units: kcal mol^{-1}.
[b]Reference 4. QC implies that Davidson's formula was used to estimate the effect of quadruple excitations.

kcal mol^{-1},[33] respectively. The electronic energy of the hydrogen atom, using our basis set, is −0.49778 hartree. Goddard and Schaefer computed the zero-point energies using vibrational frequencies derived from force constants predicted by *ab initio* SCF calculations.[5] Although the magnitude of individual frequencies differed from the experimental values, the zero-point energy difference determined from SCF theory differs little from the experimental value.

A series of computed dissociation energies is presented in Table IV. Our self-consistent field calculations predict a dissociation energy (68.6 kcal mol^{-1}) that is much lower than the experimental values[3,34] listed in the table. The two MBPT results (D-MBPT(4), 86.4 kcal mol^{-1}); SDQ-MBPT(4), 85.0 kcal mol^{-1}) and the CCD result (86.1 kcal mol^{-1}) agree well with the experimental values,[3,34] while the GS-CI calculations predict a dissociation energy of approximately 79 kcal × mol^{-1}.[5] The present MBPT/CCD calculations employ UHF reference functions for the radical products, while the SD-CI calculations[5] use a two-configuration RHF-SCF root function. In the latter calculation, an estimate of higher excitation effects is obtained by using a two-configuration modification of Davidson's formula,[12] although the justification for using such a formula with a two-configuration reference is not clear.[35] For formaldehyde the net effect is less than 1 kcal mol^{-1} in any event. The linked diagram and SD-CI results together demonstrate that high quality theoretical calculations can predict energy differences with good accuracy.

B. $CH_2O \rightarrow H_2 + CO$

Since MBPT/CCD results are size-extensive, dissociation energies may be obtained without requiring supermolecule calculations. Table V contains a summary of electronic energies for the supermolecule $H_2 \cdots CO$ obtained by addition. The zero-point energy change, computed with experimental frequencies is −7.1 kcal mol^{-1}. Predicted heats of reaction, for $T = 0$ K, are gathered in Table VI. The experimental value (−1.9 kcal mol^{-1}) corresponds to the value derived from heat-of-formation data suggested by Benson,[36] corrected to zero degrees. The SCF results yield a heat of reaction that is more negative than that obtained from either of the MBPT or the CCD calculations. The MBPT and CCD calculations predict a heat of reaction that is less than that estimated from the thermochemical data.

TABLE V. Summary of electronic structure results for $H_2 + CO$.[a]

Level of calculation	Total energy
SCF	−113.8977
D-MBPT(4)	−114.2762
SDQ-MBPT(4)	−114.2780
CCD	−114.2709

[a]Units: hartree.

TABLE VI. Predicted and experimental values for the dissociation of formaldehyde to molecular products.[a]

Level of calculation	ΔE_{el}	ΔE_v	ΔE_R
SCF	0.1	−7.1	−7.0
D-MBPT(4)	4.4	−7.1	−2.7
SDQ-MBPT(4)	3.1	−7.1	−4.0
CCD	3.9	−7.1	−3.2
Experiment: Ref. 36			−1.9

[a]Units: kcal mol^{-1}.

In both dissociation energy estimates described here, the high-quality (DZP) *ab initio* electronic structure calculations predict dissociation energies that are below the experimentally obtained values. Since the experimental studies for these two reactions have good precision, it seems relevant to inquire whether theory will generally underestimate dissociation energies. One aspect of the computations suggests that theory will generally underestimate the change in electronic energy in going from reactant to dissociation products. The reason for this is that the finite basis sets used in the molecular calculations should be better suited to computations on the smaller product fragments than on the reactant molecule.[37] In a future article describing a series of hydrogen-dissociation reactions, we will report results that suggest that MBPT calculations of the quality used here predict dissociation energies that are generally low by about 1–3 kcal mol^{-1}.[38] In each case reported, however, the zeroth-order reference function was well-represented by a single determinant.

V. MOLECULAR PRODUCTS TRANSITION STATES

The dissociation of formaldehyde to form molecular products may occur by one of several proposed competing pathways. Two possible reaction channels were studied by Goddard and Schaefer, leading to the identification of two reaction intermediates.[5] One of these intermediates, represented schematically in Fig. 1(b), occurs on the S_0 potential surface in the channel leading to the direct formation of molecular products. The other intermediate, represented schematically in Fig. 1(c), occurs on the S_0 potential surface in the channel leading to formation of the metastable species, hydroxycarbene. Because the dissociation energies pre-

dicted by the fourth-order MBPT and the CCD calculations differed by several kcal mol^{-1} from the dissociation energy predicted by the SD–CI calculations,[5] it is important to determine whether the same difference occurs for energy barrier predictions.

The calculations reported in Tables VII and VIII refer to structural parameters determined at the SCF level using the GRADSCF computer codes.[24] As noted by GYS,[7(a)] the results of the barrier predictions should not be greatly affected by use of structural parameters determined from SCF results, since the transition states are well described by a single determinant wave function. The vibrational frequencies for the three formaldehyde structures, determined using our (431/31) basis set, are presented in Table IX, along with experimental values[39] and several other theoretical estimates.[7(b),40] All calculations were done using basis sets that included polarization functions. Although the vibrational frequencies predicted using correlated wave function calculations agree better with experimental harmonic frequencies than do the SCF-level estimates, none of the theoretical methods succeeds in obtaining good agreement with experimental values. Yamaguchi *et al.*[7(b)] suggest that the effects of higher excitations, perhaps unlinked cluster contributions, have a significant effect on the theoretical computation of frequencies. A study by Pople *et al.*[41] supports that conclusion. Calculations for 12 vibrational frequencies using second-order Moller–Plesset theory, with a molecular basis that includes polarization functions on heavy atoms, yield results that differ, on the average, from experimental harmonic frequencies by 47.7 cm^{-1}, or 2.5%.[41]

It is more germane to the study of the barriers on the S_0 surface, however, to note the variation in zero-point energy differences occasioned by the different levels of theory. The results reported by GYS[7(a)] clearly demonstrate that the zero-point energy differences shift little when different levels of theory are used. For the two barriers reported in that study, the SCF and CI calculations based upon the DZ + P basis lead to vibrational energy correction differences of 0.4 and 0.2 kcal mol^{-1}. Larger differences occur between the SCF and CI for the calculations that did not include polarization functions. Since our basis set is of the same quality as the DZ + P basis used by GYS[7(a)] the zero-point energy difference predicted by our SCF calculations should be as accurate as those obtained by those authors. The zero-point energy differences were computed using the frequencies presented in Table IX. The vibrational correction for the molecular products dissociation is −5.7 kcal mol^{-1}, and that for the rearrangement reaction is −4.3 kcal mol^{-1}. These values agree well with those determined by GYS[7(a)] and Harding *et al.*,[30] irrespective of the theoretical method used to estimate the frequencies.

TABLE VII. Electronic structure calculations for the molecular products transition state.[a]

Theoretical method	Electronic energy
SCF	
This work	−113.7319
DZ + P[b]	−113.7265
DZ + P CI[b]	−114.0391
D-MBPT(4)[c]	−114.1321
D-MBPT(4)[d]	−114.1310
SDQ–MBPT(4)[d]	−114.1314
CCD[c]	−114.1241
CCD[d]	−114.1230

[a]Units: hartree.
[b]Reference 7(a).
[c]Structural parameters, Ref. 5.
[d]Structural parameters, this work.

TABLE VIII. Electronic structure calculations for the rearrangement transition state.[a]

Theoretical method	Electronic energy
SCF	
This work	−113.7397
DZ + P[b]	−113.7366
DZ + P CI[b]	−114.0481
D-MBPT(4)[c]	−114.1365
D-MBPT(4)[d]	−114.1365
SDQ–MBPT(4)[d]	−114.1380
CCD[c]	−114.1293
CCD[d]	−114.1295

[a]Units: hartree.
[b]Reference 7(a).
[c]Structural parameters, Ref. 5.
[d]Structural parameters, this work.

Table X contains a summary of the predicted energy barriers for the formaldehyde reactions. We report the results of SCF, fourth-order MBPT calculations including double-excitation diagrams, and including all single-, double- and quadruple-excitation diagrams that contribute at that order, and CCD calculations. These results may be compared to the CI results,[7(a)] the Moller–Plesset results of Harding *et al.*,[30] and to the predicted dissociation energy for the formation of radical products. In each case, the SCF calculations predict barriers that exceed those predicted by the correlation methods. Further, the correlated calculations agree well with one another in the prediction of the barrier heights.

Isomerization of formaldehyde to form hydroxycarbene exemplifies the 1,2-hydrogen shift, an important class of chemical reactions. A recent review by Schaefer elucidates the features of this reaction and provides a review of the experimental literature.[42] The energy barriers predicted by the various correlated methods range from 85.7 (SDQ–MBPT) to 88.2 kcal mol^{-1} (SD–CI).[7(a)] Each theoretical technique predicts that the barrier to formation of molecular products exceeds the rearrangement barrier. However, the correlated predictions for this barrier range from 88.0 (RMP-4)[30] to 92.8 kcal mol^{-1} (SD–CI),[7(a)] with the difference in the barrier heights equal to approximately 3.0 kcal mol^{-1}.

TABLE IX. Vibrational frequencies: $CH_2O(S_0)$.

Equilibrium structure mode	ν_1 (a_1)	ν_2 (a_1)	ν_3 (a_1)	ν_4 (b_1)	ν_5 (b_2)	ν_6 (b_2)
Experiment[a]	2944	1764	1563	1191	3009	1287
Experiment[b]	2782	1746	1500	1167	2843	1247
SCF						
This work	3191	2004	1659	1343	3113	1368
Ref. 7(b)	3149	2006	1656	1335	3226	1367
Ref. 40	2971.2	1882.6	1600.1	1301.4	3070.6	1349.7
DZ + P CI [Ref. 7(b)]	3074	1869	1596	1243	3155	1306
CEPA	2925.6	1797.2	1580.9	1220.9	3032.3	1308.9

Rearrangement transition state	ν_1 (a')	ν_2 (a')	ν_3 (a')	ν_4 (a')	ν_5 (a')	ν_6 (a'')
SCF						
This work	3155	2806	1637	1408	2645i	637
Ref. 7(b)	3180	2832	1639	1408	2650i	625
DZ + P CI						
Ref. 7(b)	3057	2748	1512	1356	2356i	~600[d]

Molecular products transition state	ν_1 (a')	ν_2 (a')	ν_3 (a')	ν_4 (a')	ν_5 (a')[c]	ν_6 (a'')
SCF						
This work	3250	2101	1484	819	2295i	1043
Ref. 7(b)	3243	2092	1526	829	2305i	1024
DZ + P CI						
Ref. 7(b)	3263	1939	1555	876	2124i	~950[d]

[a]Reference 39(a).
[b]Reference 39(b).
[c]Reaction coordinate.
[d]Estimated.

The MBPT/CCD results suggest that the energy requirements for all three unimolecular processes are approximately equal. Since it is often assumed that even good correlated wave function calculations overestimate barrier heights and underestimate dissociation energies, the MBPT/CCD results imply that the rearrangement to hydroxycarbene is the energetically favored reaction pathway. Such a result agrees with the conclusions of Goddard *et al.*[5,7(a)] However, basis set changes and the inclusion of triple excitation diagrams may change this result.

It is interesting to note the substantial difference between the energy difference obtained from correlated calculations and the results obtained using just SCF calculations. As stated, the results obtained with MBPT, CCD, and SD–CI imply that the energy requirements for all three reaction pathways are almost equal. The SCF calculations predict that the barriers for dissociation to molecular products (95.7 kcal mol^{-1}) and the barrier for rearrangement to *trans*-hydroxycarbene (99.2 kcal mol^{-1}) differ by a relatively small amount of energy. However, the SCF results imply that each of these barriers is much larger than the barrier for dissociation to radical products (68.6 kcal mol^{-1}). These energy differences demonstrate the importance of including

TABLE X. Predicted energy barriers for formaldehyde reactions.[a]

Theoretical method	Reaction: 1,2-shift	Reaction: Molecular products	Reaction: Radical products
SCF			
This work	95.7	99.2	68.6
Ref. 7(a)	95.2	100.2	...[b]
D-MBPT(4)	86.9	89.0	86.4
SDQ–MBPT(4)	85.7	88.4	85.0
CCD	87.6	90.3	86.1
CI [Ref. 7(a)]			
SD–CI	88.2	92.8	80.0
SD–CI + QC[c]	85.9	88.9	79.1
RMP4			
Ref. 30	86.2	88.0	...[b]

[a]Units: kcal mol^{-1}.
[b]No comparable SCF results reported.
[c]Includes correction for quadruple excitations based upon Davidson's formula.

correlation energy effects to obtain chemically accurate results.

An additional result of the configuration interaction study was that each decomposition reaction possesses an energy barrier relative to the S_1 band origin. To determine whether the linked-diagram related methods will also predict the existence of this barrier, we must consider the first excited state of formaldehyde.

VI. THE $a\,^3A''$ STATE OF FORMALDEHYDE

The ultraviolet absorption spectrum of formaldehyde vapor consists of two band systems: a strong band with apparent origin at 3530 A (3.454 eV) and a weak band with apparent origin at 3960 A (3.124 eV). Walsh predicted that these states would have nonplanar structure,[43] and this was verified for both the singlet[44,45] and the triplet[46,47] states. Jones and Coon derived experimental structures for the excited states, including a specification of the out-of-plane bending angle by a distribution function for each vibrational level of the bending mode.[48] Recent theoretical studies of the excited states include *ab initio* calculations by Garrison, Schaefer, and Lester,[10] Bell,[6] and Goddard and Schaefer,[5] as well as equations-of-motion calculations of the vertical excitation energies.[17,18] This report complements and extends Bell's work.

Recent calculations[5,6] used excited state structural parameters determined by optimizing the electronic energy at the SCF level. This study reports structural parameters for the $a\,^3A''$ state determined by optimizing the energy at the correlated level. Fourth order, double-excitation MBPT calculations, with core orbitals frozen, were performed for structures obtained by varying the CH and CO bond lengths, the HCH angle, and the out-of-plane bending angle. The CH bond length and the HCH angle were reoptimized after determination of the optimal out-of-plane bending angle. We report the optimal structural parameters in Table XI.

The carbon–hydrogen bond length (1.085 Å) reported here is slightly less than the experimental value (1.0962 A) while the carbon–oxygen bond length (1.327 Å) exceeds the experimental value (1.307 A). The optimum HCH and out-of-plane bending angles (118° and 37°12') agree well with the experimental values (118° and 37°54'). Comparison of these theoretical results with Bell's structure[6] shows that the inclusion of correlation effects improves the agreement between theory and experiment.

TABLE XI. Formaldehyde structural parameters, $a\,^3A''$ state.

	r_{CH}[a]	r_{CO}[a]	HCH	ϕ
Experiment[b]	1.0962	1.3069	118°	37°54'
Theory				
DZP–SCF[c]	1.0779	1.3404	119.67°	38°30'
D-MBPT(4)–FC	1.0848	1.3271	118°	37°12'

[a]Units: angstroms.
[b]Reference 48.
[c]Reference 6.

TABLE XII. Adiabatic excitation energies, $S_0 \rightarrow T_1$.

	ΔE (eV)	ΔE (kcal mol^{-1})
Experiment[a]	3.124	71.9
D-MBPT(4)	3.1455	72.5
SDQ–MBPT(4)	3.1436	72.5
CCD	3.093	71.3
SD–CI[b]	3.01	69.5
SD–CI[c]	2.664	61.4
SCF[d]	2.183	50.3
SCF, this work	1.912	44.1

[a]Reference 32.
[b]Reference 5.
[c]Reference 6.
[d]Reference 10.

We report, in Table XII, three theoretical results for the adiabatic excitation energy corresponding to $S_0 \rightarrow T_1$. Both the D-MBPT(4) and SDQ–MBPT(4) calculations yield 72.5 kcal mol^{-1} for the excitation energy, while the CCD calculations yielded a value of 71.3 kcal mol^{-1}. These results may be compared to the near Hartree–Fock value, 50.3 kcal mol^{-1},[10] two SD–CI values 67.4 kcal mol^{-1}[5] and 68.9 kcal mol^{-1},[49] and a value obtained by including the unlinked cluster correction of quadruple-excitation effects 69.5 kcal mol^{-1}.[5] All the theoretical values compare well with the experimental value 71.9 kcal mol^{-1}.[32] None of the theoretical values, however, include any correction for zero-point energy differences that may contribute to the observed value.

Although no effort has been made to determine the structure of formaldehyde in the $A\,^1A''(S_1)$ state, vertical excitation energies are computed for both the $S_0 \rightarrow S_1$ and $S_0 \rightarrow T_1$ transitions. Since the UHF wave function for the S_1 state was contaminated by higher-order spin states, particularly T_1, the computed excitation energy for this state is expected to be less accurate than that for the T_1 state. Nonetheless, vertical excitation energies for both transitions are presented in Table XIII, along with other theoretical predictions and experimental estimates. The excitation energy for $S_0 \rightarrow T_1$ is less than that for $S_0 \rightarrow S_1$, although the latter is probably lower than would be obtained were the ref-

TABLE XIII. Vertical excitation energies for formaldehyde.[a]

	$S_0 \rightarrow T_1$	$S_0 \rightarrow S_1$
D-MBPT(4)	84.2	89.3
MRD–CI[b]	78.6	87.9
EOM[c]	79.8	93.2
Experiment[d]	83.0	98.2

[a]Units: kcal mol^{-1}.
[b]Reference 50.
[c]Reference 18.
[d]Estimated excitation energies, Ref. 51.

TABLE XIV. Energy prediction comparisons for photochemical processes.[a]

	Reference 5	Reference 7(a)	SDQ–MBPT(4)	CCD	Experiment
$S_0 - T_1$	69.5		72.5	71.3	71.9[a]
D_0	~79.		85.0	86.1	86.0 ± 1[b]
E_0 (H_2CO)	87.5	88.9	88.4	90.3	
E_0	83.9	85.9	85.7	87.6	
$S_0 - S_1$	78(est.)		81.0(est.)	79.8(est.)	80.5[c]

[a]Reference 32. [b]Reference 3. [c]Reference 1.

erence function not spin-contaminated. The difference in the excitation energy predictions is much less than the difference obtained from the equation of motion calculations[17] or estimated from experimental data.[51] However, it should be noted that MRD–CI calculations predict a significantly smaller excitation energy difference (9.3 kcal mol^{-1})[50] than is estimated by Chutjian (15.2 kcal mol^{-1}).[51] Since the difference between the two adiabatic excitations is 8.1 kcal mol^{-1}, it seems best to conclude that the $S_0 - S_1$ vertical excitation energy remains uncertain.

VII. DISCUSSION

The present calculations demonstrate that linked-diagram related computational methods predict relative energies of small molecular systems to high accuracy. For the results that can be compared with experiment, the dissociation energy for formation of radical products and the $S_0 - T_1$ adiabatic excitation energy, the SDQ–MBPT(4) and the CCD results agree well with the experimental values.

It is instructive to review results pertinent to photochemical dissociation and rearrangement of formaldehyde. We collect, in Table XIV, the excitation energy, dissociation energy (D_0), and reaction barrier (E_0) results obtained in this study and in work by Goddard and co-workers.[5,7(a)] Note that the values for $S_0 - S_1$ are estimates based on the theoretical results for the $S_0 - T_1$ excitation energy, and the observed $T_1 - S_1$ separation. In analyzing the CI results, the authors noted that high level theoretical calculations underestimate dissociation energies and slightly overestimate rearrangement barriers.[52,53] If it is assumed that the CI calculations underestimate the dissociation energy by 3 to 5 kcal $\times mol^{-1}$, and overestimate the energy barriers by a similar amount, then the CI results imply that the energy requirements for the radical products and molecular products dissociation pathways are similar.[5] Furthermore, the energy requirements suggest that the hydroxycarbene intermediate cannot be ruled out. The subsequent analytic gradient CI calculations support these results, but each of the energy barriers increases slightly.[7(a)] Finally, those authors noted that the barriers exceed the estimated adiabatic excitation energy for $S_0 - S_1$. Thus, there all small barriers above the S_1 origin to all three photochemical processes.

Examinations of the SDQ–MBPT CCD results reveals no significant difference from the results of the CI calculations, even though our basis set is somewhat larger than the GS DZ + P basis. However, our SCF result for the molecular products transition state shows an energy lowering of 3.4 kcal mol compared to the DZ + P results, suggesting that basis set changes may still be significant in this problem. We cannot make realistic estimates of the errors associated with the barrier calculations, since we have no data for comparison. We assume that these calculations overestimate the energy barriers by 2 to 4 kcal mol^{-1}. With these assumptions, the results of these calculations lead to the same conclusions obtained by GS[5] and Goddard *et al.*[7(a)] However, recent work by Frisch *et al.*[54] suggests that basis set effects and triple excitations lower the computed barriers by ~8 kcal mol, which is below the origin of the H_2CO excited singlet state.

[1]R. D. McQuigg and J. G. Calvert, J. Am. Chem. Soc. 91, 1590 (1979).
[2]A. Horowitz and J. G. Calvert, Int. J. Chem. Kinet. 10, 713 (1978).
[3]J. H. Clark, C. B. Moore, and N. S. Nogar, J. Chem. Phys. 68, 1264 (1978).
[4]J. C. Weisshaar and C. B. Moore, J. Chem. Phys. 70, 5135 (1979).
[5]J. D. Goddard and H. F. Schaefer, J. Chem. Phys. 70, 5117 (1979).
[6]S. Bell, Mol. Phys. 37, 255 (1979).
[7](a) J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer (to be published); (b) Y. Yamaguchi and H. F. Schaefer, J. Chem. Phys. 73, 2310 (1980).
[8]R. A. Fifer, *Eighteenth Symposium (International) on Combustion* (Combustion Institute, Pittsburgh, 1979).
[9]R. L. Jaffe and K. Morokuma, J. Chem. Phys. 64, 4881 (1976).
[10]B. J. Garrison, H. F. Schaefer, and W. A. Lester, J. Chem. Phys. 61, 3039 (1974).
[11]W. Miller, J. Am. Chem. Soc. 101, 6810 (1979).
[12]S. R. Langhoff and E. R. Davidson, Int. J. Quantum Chem. 8, 61 (1974); E. R. Davidson and D. W. Silver, Chem. Phys. Lett. 52, 403 (1978).
[13]H. P. Kelly, Adv. Chem. Phys. 14, 129 (1969).
[14]J. Čížek, J. Chem. Phys. 45, 4256 (1969).
[15]R. J. Bartlett and G. D. Purvis, Int. J. Quantum Chem. 14, 561 (1978).
[16]R. J. Bartlett and G. D. Purvis, Phys. Scr. 21, 255 (1980), R. J. Bartlett, I. Shavitt, and G. D. Purvis, J. Chem. Phys. 71, 281 (1979).
[17]D. L. Yeager and V. McKoy, J. Chem. Phys. 60, 2714 (1974).
[18]C. W. McCurdy, Jr., T. N. Rescigno, D. L. Yeager, and V. McKoy, in *Modern Theoretical Chemistry*, Vol. 3, *Methods of Electronic Structure Theory*, edited by H. F. Schaefer III (Plenum, New York, 1977).
[19]T. H. Dunning, Jr., J. Chem. Phys. 53, 2823 (1970).
[20]S. Huzinaga, J. Chem. Phys. 42, 1293 (1965).
[21]T. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry*, Vol. 3, *Methods of Electronic Structure Theory*,

edited by H. F. Schaefer III (Plenum, New York, 1977).
[22]J. Almlöf, *Proceedings of the Second Seminar on Computational Problems in Quantum Chemistry*, Strasbourg, France, 1972 (1973); USIP Report 74-29 (1974).
[23]The program GRNFNC, written by G. D. Purvis, does SCF iterations and integral transformations. The program UMBPT, written by R. J. Bartlett and G. D. Purvis, does MBPT, CCD, and CI.
[24]A. Komornicki, *National Resources Computer Chemistry Software Catalog*, Vol. 1, Prog. No. QH04 (GRADSCF), 1980.
[25]K. Yamada, T. Nagakura, K. Kuchitsu, and Y. Morino, J. Mol. Spectrosc. 38, 70 (1971).
[26]R. Walsh and S. W. Benson, J. Am. Chem. Soc. 88, 4570 (1966).
[27]P. Warneck, Z. Naturforsch Teil A 26, 2047 (1971).
[28]B. R. Brooks, E. D. Laidig, P. Saxe, N. C. Handy, and H. F. Schaefer, Phys. Scr. 21, 312 (1980).
[29]B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. 72, 4652 (1980).
[30]L. B. Harding, H. B. Schlegel, R. Krishnan, and J. A. Pople, J. Phys. Chem. 84, 3394 (1980).
[31]G. F. Adams, G. D. Bent, G. D. Purvis, and R. J. Bartlett, J. Chem. Phys. 71, 3697 (1979).
[32]G. Herzberg, *Electronic Spectra of Polyatomic Molecules* (Van Nostrand, New York, 1966).
[33]J. M. Brown and D. A. Ramsey, Can. J. Phys. 53, 2232 (1975).
[34]J. P. Reilly, J. H. Clark, C. B. Moore, and G. C. Pimentel, J. Chem. Phys. 69, 4381 (1978).
[35]R. J. Bartlett and I. Shavitt, Int. J. Quantum Chem. Symp. 11, 165 (1977); 12, 543 (1978).
[36]S. W. Benson, *Thermochemical Kinetics* (Wiley, New York, 1976), 2nd ed.
[37]On the other hand, basis set superposition effects tend to act in the opposite direction. However, these effects are usually important only for van der Waals molecules.
[38]G. F. Adams, G. D. Bent, R. J. Bartlett, and G. D. Purvis, Chem. Phys. Lett. (in press).
[39]J. L. Duncan and P. S. Mallinson, Chem. Phys. Lett. 23, 597 (1979); J. L. Duncan, Mol. Phys. 28, 1177 (1974).
[40]R. Joquet, W. Kutzelnigg, and U. Staemmler, Theor. Chim. Acta 54, 205 (1980).
[41]J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. 13, 225 (1979).
[42]H. F. Schaefer, Acc. Chem. Res. 12, 288 (1979).
[43]A. D. Walsh, J. Chem. Soc. 1953, 2260.
[44]J. C. D. Brand, J. Chem. Soc. 1956, 858.
[45]G. W. Robinson, Can. J. Phys. 34, 699 (1956).
[46]G. W. Robinson and V. E. DeGiorgio, Can. J. Chem. 36, 31 (1958).
[47]S. E. Hodges, J. R. Henderson, and J. B. Coon, J. Mol. Spectrosc. 2, 99 (1958).
[48]V. T. Jones and J. B. Coon, J. Mol. Spectrosc. 31, 137 (1969).
[49]R. R. Lucchese and H. F. Schaefer, J. Am. Chem. Soc. 100, 298 (1978).
[50]S. D. Peyerimhoff, R. J. Buenker, W. E. Kammer, and H. Hsu, Chem. Phys. Lett. 8, 129 (1971).
[51]A. Chutjian, J. Chem. Phys. 61, 4279 (1974).
[52]R. L. Jaffe, D. M. Hayes, and K. Morokuma, J. Chem. Phys. 60, 5108 (1974).
[53]R. L. Jaffe and K. Morokuma, J. Chem. Phys. 64, 4881 (1976).
[54]M. J. Frisch, R. Krishnan, and J. A. Pople, J. Phys. Chem. (in press); we appreciate receiving a copy of this article prior to publication.

CALCULATION OF DISSOCIATION ENERGIES USING MANY-BODY PERTURBATION THEORY *

George F. ADAMS, Gary D. BENT ‡
ARRADCOM, US Army Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland 21005, USA

and

George D. PURVIS and Rodney J. BARTLETT
Battelle Columbus Laboratories, Columbus, Ohio 43201, USA

Received 10 April 1981; in final form 28 April 1971

Dissociation energies for the step-wise removal of hydrogen from methanol, $CH_3OH \rightarrow CH_3O \rightarrow CH_2O \rightarrow CHO \rightarrow CO$, are obtained by many-body perturbation theory. The heat of formation of CH_3O is predicted as 2 ± 3 kcal mole^{-1}. Computed geometries are in excellent agreement with experiment, where available, and provide a prediction for CH_3O.

1. Introduction

A major task for theoretical chemists is the development of methods to predict energy differences with chemical accuracy. Most quantum chemists agree that accurate prediction of relative energies requires application of theories that include electron correlation effects, effects not treated in self-consistent-field (SCF) calculations [1–4]. Estimates of molecular correlation energy have been obtained by using the method of configuration interaction to improve the wavefunction obtained from an ab initio SCF calculation. Recently, however, theoretical techniques based upon the linked-diagram theorem have been extended to molecular applications [1,2, 4], and these methods provide an alternative approach to the correlation problem. Several comparisons of the various theoretical methods have appeared [1,2,4].

This report describes the application of many-body perturbation theory (MBPT) to predict the dissociation energies for a series of reactions that occur in flame zone combustion models. The reaction set comprises the unimolecular reactions that oxidize methanol to carbon monoxide:

$$CH_3OH \rightarrow CH_3O + H, \quad (1)$$

$$CH_3O \rightarrow CH_2O + H, \quad (2)$$

$$CH_2O \rightarrow CHO + H, \quad (3)$$

$$CHO \rightarrow CO + H. \quad (4)$$

Reactions (1) and (2) involve the methoxy radical. Since heat of formation data are available for the other moieties in those two reactions, accurate predictions of the dissociation energies for these reactions will enable us to predict the heat of formation for the methoxy radical.

The theoretical techniques used are outlined in section 2, and the results for the series of dissociation reactions follow in section 3.

2. Outline of the computations

For the closed-shell molecules included in this study, a restricted Hartree–Fock (RHF) wavefunction represents the ground state, whereas unrestricted Hartree–Fock (UHF) wavefunctions represent the ground state for open-shell molecules. The integrals

* This research was sponsored in part by the US Army Research Office under Grant No. DAAG29-80-C-0105.
‡ NAS-NRC Resident Research Associate. Present address: Department of Physics, University of Connecticut, Storrs, Connecticut 06268, USA.

over atomic functions were computed using Dunning's 4s3p contraction of Huzinaga's 9s5p gaussian primitive set for first-row atoms, and Dunnings scaled (δ = 1.2) 3s contraction of Huzinaga's 4s primitive set for hydrogen [5,6]. A single set of d-type polarization functions (gaussian) augments the atomic basis sets for carbon and oxygen, with exponents 0.75 and 0.85 respectively [7]. A set of p-type polarization functions (gaussian), α = 1.0, augments the hydrogen basis set. All integral calculations were performed using the MOLECULE integral program [8]. Electronic structure calculations were performed using the GRNFNC and UMBPT programs ‡.

The total energy reported for each molecule pertains to structural parameters that have been optimized relative to double-excitation, frozen-core, fourth-order MBPT calculations. The structural parameters optimized for each molecule are defined in fig. 1. Computed structural parameters are compared with available experimental values in table 1. In the absence of experiment, we offer a prediction of the methoxy structure which we estimate to be accurate to 0.01 Å in bond length and 2–3° in bond angles.

We report dissociation energy predictions for three levels of calculation: SCF, fourth-order MBPT including all double-excitation diagrams, D MBPT(4), and fourth-order MBPT including all single-, double-, and quadruple-excitation diagrams that contribute at that order, SDQ MBPT(4).

The prediction of dissociation energies requires

‡ The program GRNFNC, written by G.D. Purvis, does SCF iterations and integral transformations. The program UMBPT, written by R.J. Bartlett and G.D. Purvis, does MBPT, CCD and VP CI.

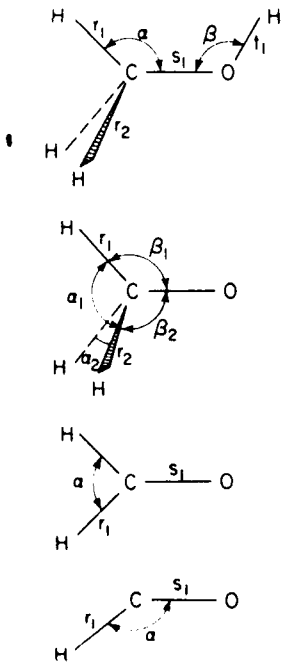


Fig. 1. Definition of bond lengths and of bond angles for CH_3OH, CH_3O, CH_2O, and CHO.

Table 1
Optimized structural parameters for molecules included in this study [a]

	Parameter	This work	Experiment
CH_3OH [b]	r_1	1.093	1.0937
	r_2	1.093	1.0937
	α	107° 9′	108° 32′
	methyl tilt	2° 12′	3° 12′
	s_1	1.432	1.4214
	t_1	0.963	0.963
	β	107° 2′	108° 2′
CH_3O	r_1	1.085	
	r_2	1.081	
	α_1	108° 42′	
	α_2	112° 12′	
	β_1	103° 48′	
	β_2	111° 30′	
	s_1	1.405	
CH_2O [c]	r_1	1.102	1.099
	s_1	1.211	1.203
	α	116° 11′	116° 30′
CHO [d]	r_1	1.111	1.125
	s_1	1.188	1.175
	α	124°	124° 57′
CO [e]	s_1	2.125	2.127

[a] Bond lengths in A.
[b] Experimental values: ref. [12].
[c] Experimental values: ref. [13].
[d] Experimental values: ref. [14].
[e] Experimental values: ref. [15].

accurate computation of energy differences. In addition to the change in the electronic energy, vibrational, rotational, and translational energy differences must also be obtained. We assume that rotational and translational energy effects are well approximated by the equipartition of energy among those degrees of freedom. Since vibrational frequencies are established for all the molecular species addressed in this study except the methoxy radical, empirical zero-point energy changes are used for these cases.

Characterization of the methoxy radical was the most difficult problem encountered in this study. Bent et al. [9] obtained optimal structural parameters for the Jahn - Teller distorted ground-state conformation after extensive D MBPT(4) calculations. An analysis of the potential energy function obtained in that study provided guidelines to estimating the vibrational frequencies for the e modes. Engelking et al. [10] and Inoue et al. [11] have published experimental values for several frequencies. By combining theoretical predictions and experimental data with observations on similar molecules, such as CH_3F^+, we assign the following values to the vibrational eigenvalues for the methoxy radical (directly observed values are underlined: 3006 (e), 2930 (a_1), 1380 (e), 1325 (a_1), 1060 (a_1), 690 cm^{-1} (e).

These frequencies correspond to a methoxy radical zero-point vibrational energy equal to 22.1 kcal mole^{-1}.

3. Dissociation energy predictions

Because linked-diagram derived methods are size extensive [2], the electronic energy change for a dissociation reaction, A → B + C, may be written

$$\Delta E_{el} = E_{el}(B) + E_{el}(C) - E_{el}(A).$$

Since each of the reactions considered here contains the hydrogen atom as one of the products, ΔE_{el} requires a correlated calculation for only the reactant and the molecular product. Table 2 contains predicted electronic energies for methanol, CH_3OH, methoxy radical, CH_3O, formaldehyde, CH_2O, formyl radical, HCO, carbon monoxide, CO, and the SCF energy for the hydrogen atom. Table 2 also contains the zero-point vibrational energy for each molecule in the study [10–15].

The information summarized in table 2 is sufficient to determine the zero-degree heat of reaction for each of the four reactions included in this study. We summarize the values for each of the molecules in the third row of table 3. For most molecules, there is no experimental data available to compare with the predicted $\Delta E_R(0\ K)$ values. An exception is formaldehyde. Reilly et al. [16] have confirmed that the threshold energy for the dissociation of formaldehyde to radical products is 86.0 ± 1.0 kcal mole^{-1}. The SDQ MBPT(4) calculations predict that the threshold energy equals 85.0 kcal mole^{-1}. In this case, the agreement between theory and experiment is very good.

In addition to the experimental data of Reilly et al. [16], experimental values for room-temperature heats of reaction are available for methanol [17], formaldehyde [18] and the formyl radical [19]. To compare the theoretical predictions with these heat of reaction data, we must estimate the effect of the rotational and translational degrees of freedom, and the effect of the pV work term must be included.

Table 2
Electronic energy predictions and zero-point energies for CH_nO molecules

Molecule	Electronic energy (hartree)			Zero-point energy (kcal mole^{-1})
	SCF (UHF)	D MBPT(4)	SDQ MBPT(4)	
CH_3OH	−115.0744	−115.4942	−115.4920	31.1 [12]
CH_3O	−114.4532	−114.8248	−114.8232	22.1 [9]
CH_2O	−113.8974	−114.2832	−114.2829	16.2 [13]
CHO	−113.2769	−113.6343	−113.6363	7.8 [14]
CO	−112.7665	−113.1105	−113.1124	3.1 [15]
H	−0.49778			

Table 3
Dissociation energies predicted by SDQ MBPT(4) results (in kcal mole^{-1})

	CH_3OH	CH_3O	CH_2O	CHO
ΔE^{el}	107.3	26.6	93.4	16.4
ΔE^{vib}	−9.0	−5.9	−8.4	−4.7
ΔE_R(0 K)	98.3	20.7	85.0	11.7
ΔE^{tr}(300 K)	0.9	0.9	0.9	0.9
ΔE^{rot}(300 K)	0	0	0	−0.3
ΔE_R(300 K)	99.2	21.6	85.9	12.3
$\Delta H^0_{R,300}$	99.8	22.2	86.5	12.9

Translational and rotational energy contributions are estimated by assuming that each degree of freedom contributes $\frac{1}{2}kT$. Thus, for T = 300 K, each degree of freedom contributes ≈ 0.3 kcal mole^{-1}. The translational and rotational energy changes for each reaction are listed in table 3. Combining these contributions with the zero-degree dissociation energy, we obtain the energy of the reaction for T = 300 K. Finally, we assume that $pV = \Delta n(RT)$, and add this contribution to ΔE_R(300 K) to obtain predicted values for the heat of reaction, ΔH^0_R(300 K), which are listed in the last line of table 3.

Table 4 contains a comparison of the heats of reaction predicted by our SDQ MBPT(4) calculations with experimental values. Several conclusions may be gleaned from the theoretical results summarized in table 4. First, the inclusion of correlation effects is necessary to predict dissociation energies with chemical accuracy. Ab initio self-consistent-field calculations predict dissociation energies substantially different from experimental values. Each of the correlated treatments, D MBPT(4) and SDQ MBPT(4), predicts dissociation energies that agree well with experiment. There is little reason to prefer the more expensive SDQ MBPT(4) calculations to the D MBPT(4) calculations for these examples. In each case reported here, the contributions of the single- and quadruple-excitation diagrams at fourth order almost cancel. This is not always the case, so there are instances where SDQ MBPT(4) is preferred. The fourth-order triple excitations can also be important for some examples [20]. Finally, in each case where theoretical predictions can be compared with experimental results, the predicted dissociation energy is less than the experimental value. This is not surprising. The finite basis sets used in molecular calculations describe the product molecules better than they describe the reactant. This characteristic implies that the electronic energy difference between products and reactant usually will be too low. Reliable estimates of the magnitude of this error are not available, but our experience [21–24] suggests that dissociation energies predicted by MBPT, for breaking a single bond in a polyatomic molecule, and for the quality of basis set used in these calculations, typically underestimate experiment by 1–3 kcal mole^{-1}. On the other hand, atomization energies are much poorer due to the limited basis sets [21].

4. Heat of formation of the methoxy radical

Two experimental values for the heat of formation

Table 4
Comparison of predicted and experimental dissociation energies (in kcal mole^{-1})

Molecule	Predicted $H^0_{R,300}$			Experiment	
	SCF (UHF)	D MBPT(4)	SDQ MBPT(4)	$H^0_{R,300}$	ref.
CH_3OH	69.9	100.2	99.8	100.9	[10]
				104.0	[17]
CH_3O	32.9	23.1	22.2		
CH_2O	70.1	86.4	86.5	88.2 ± 1.6	[18]
				87.5 ± 1.0 [a]	[16]
CHO	4.4	12.9	12.9	15.5 ± 1.5	[18]
				17.0 ± 2.0	[19]

[a] We estimate this value based upon the $\Delta E^0_{R,300}$ value of 86.0 ± 1 [16].

Table 5
Heat of reaction for H_2 elimination from CH_3O

	CH_3O →	CHO +	H_2
electronic energy [a]	−114.8232	−113.6363	−1.1656
ΔE^{el}	(0.0213 [a])	13.4 [b]	
zero-point energy [b]	22.1	7.8	6.2 [15]
ΔE^{zp}		−8.1 [b]	
ΔE_R(0 K)		5.3 [b]	
rotational energy (300 K) [b]	0.9	0.9	0.6
ΔR^{rot}(300 K)		0.6 [b]	
translational energy (300 K)	0.9	0.9	0.9
ΔE^{T}		0.9 [b]	
ΔE_c(300 K)		6.8 [b]	
$\Delta H^0_{R,300}$		7.4 [b]	

[a] In hartree. [b] In kcal mole^{-1}.

of the methoxy radical have been reported recently [10,17]. Batt and McCulloch [17] derived a heat of formation equal to 3.9 kcal mole^{-1} using parameters obtained from chemical kinetic data. Engelking et al. [10] measured the electron affinity of the methoxy radical, and estimated the heat of formation for methoxy to be 0.7 kcal mole^{-1}. Since our calculations predict heats of reaction for two processes that include the methoxy radical, by combining these results with published heat of formation data we can obtain a heat of formation for CH_3O. We assume the following heats of formation: CH_3OH, −48.0 kcal mole^{-1}; CH_2O, −26.0 kcal mole^{-1}, H, 52.1 kcal mole^{-1} [25]. Combining the dissociation energy of methanol with the heats of formation of methanol and hydrogen, we predict the methoxy heat of formation to be −0.3 kcal mole^{-1}. On the other hand, combining the methoxy radical dissociation energy with heats of formation for formaldehyde and hydrogen, we predict the methoxy heat of formation to be 3.9 kcal mole^{-1}.

In addition to the two hydrogen elimination reactions already considered, the heat of formation for CH_3O can be obtained by considering the elimination of the hydrogen molecule from the methoxy radical, giving the formyl radical and H_2 as products. The pertinent information needed to estimate the heat of reaction for this process is summarized in table 5. The results of SDQ MBPT(4) calculations predict a classical dissociation energy D_e equal to 13.4 kcal mole^{-1}. Combining the estimated frequencies for the methoxy radical with empirical values for the formyl radical and H_2, we estimate a zero-point energy change of −8.1 kcal mole^{-1} for the reaction. When rotational and translational energy effects and our estimate of the pV terms are included, we obtain a value for the heat of reaction equal to 7.4 kcal mole^{-1}. Assuming that $\Delta H^0_{f,300}$ for the formyl radical is 9.0 kcal mole^{-1} [25], we predict that the heat of formation for the methoxy radical is 1.6 kcal mole^{-1}.

The lack of agreement among the predicted values for the heat of formation is not surprising. It is apparent from the results of this work and the results of theoretical studies using alternative methods [1,3] that theoretical calculations underestimate dissociation energies. For the H-dissociation reactions reported here, and in other work done using these methods [21,26], our dissociation energy errors range from 1 to 3 kcal mole^{-1}. Considering the sign of this error, it is reasonable to suggest that the heat of formation of the methoxy radical should be greater than −0.3 kcal mole^{-1} and less than 3.9 kcal mole^{-1}. The heat of formation obtained by studying the hydrogen molecule elimination reaction of CH_3O suggests a heat of formation for the radical of 1.6 kcal mole^{-1} that lies between the two extreme values. Taking the likely errors into account we suggest a value equal to 2.0 ± 3.0 kcal mole^{-1}.

References

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

[2] R.J. Bartlett, I. Shavitt and G.D. Purvis, J. Chem. Phys. 71 (1979) 281.
[3] J.D. Goddard and H.F. Schaefer III, J. Chem. Phys. 70 (1979) 5117;
J.D. Goddard, Y. Yamaguchi and H.F. Schaeffer III, to be published.
[4] R.J. Bartlett and G.D. Purvis, Intern. J. Quantum Chem. 14 (1978) 561.
[5] T.H. Dunning, J. Chem. Phys. 53 (1970) 2823.
[6] S. Huzinaga, J. Chem. Phys. 42 (1965) 1293.
[7] T.H. Dunning and P.J. Hay, in: Modern theoretical chemistry, Vol. 3. Methods of electronic structure theory, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
[8] J. Almlöf, Proceedings of the 2nd Seminar on Computational Problems in Chemistry, Strasbourg, France, 1972 (1973).
[9] G.D. Bent, G.F. Adams, R.H. Bartram, R.J. Bartlett and G.D. Purvis, submitted for publication.
[10] P.C. Engelking, G.B. Ellison and W.C. Lineberger, J. Chem. Phys. 69 (1978) 1826.
[11] G. Inoue, H. Akimoto and M. Okuda, Chem. Phys. Letters 63 (1979) 213.
[12] M.L.C. Gerry, R.M. Lees and G. Winnewisser, J. Mol. Spectry. 31 (1967) 231.
[13] K. Yamada, T. Nagakura, K. Kuchistu and Y. Morino, J. Mol. Spectry. 38 (1971) 70.
[14] J.M. Brown and D.A. Ramsay, Can. J. Phys. 53 (1975) 2232.
[15] K. Huber and G. Herzberg, Molecular spectra and molecular structure, Vol. 4. Constants of diatomic molecules (Van Nostrand, Princeton, 1979).
[16] J.P. Reilly, J.H. Clark, C.B. Moore and G.C. Pimentel, J. Chem. Phys. 69 (1978) 4381.
[17] L. Batt and R.D. McCulloch, Intern. J. Chem. Kinetics 8 (1976) 491.
[18] P. Warneck, Z. Naturforsch. A26 (1971) 2047.
[19] M.A. Haney and J.C. Franklin, Trans. Faraday Soc. 65 (1969) 1794.
[20] M.J. Frisch, R. Krishnan and J.A. Pople, Chem. Phys. Letters 75 (1980) 66.
[21] R.J. Bartlett and G.D. Purvis, Physica Scripta 21 (1980) 255.
[22] L.T. Redmon, G.D. Purvis and R.J. Bartlett, J. Am. Chem. Soc. 101 (1979) 2856.
[23] G.F. Adams, G.D. Bent, R.J. Bartlett and G.D. Purvis, submitted for publication.
[24] G.F. Adams, G.D. Bent, G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 71 (1979) 3697.
[25] S.W. Benson, Thermochemical kinetics, 2nd Ed. (Wiley, New York, 1976).
[26] G.F. Adams, G.D. Bent, R.J. Bartlett and G.D. Purvis, in: Potential energy surfaces and dynamics calculations, ed. D.G. Truhlar (Plenum Press, New York), to be published.

ON THE UNIMOLECULAR REACTIONS OF CH_3O AND CH_2OH *

George F. ADAMS
USA ARRADCOM/Ballistics Research Laboratory, Aberdeen Proving Ground, Maryland 21005, USA

and

Rodney J. BARTLETT and George D. PURVIS
Quantum Theory Project, University of Florida, Gainesville, Florida 32611, USA

Received 21 September 1981; in final form 4 January 1982

Recent experiments suggest that the isomerization reaction $CH_3O \rightleftharpoons CH_2OH$ may compete with the hydrogen dissociation reaction $CH_3O \rightarrow CH_2O + H$. We report correlated many-body calculations of barrier heights for these reactions and other decompositions of CH_3O and CH_2OH. The barriers for isomerization and dissociation are similar, while the reaction $CH_3O \rightarrow HCO + H_2$ has a higher energy barrier.

1. Introduction

The methoxy radical CH_3O is thought to play an important role in hydrocarbon oxidation processes relevant to combustion and photochemical air pollution [1]. Recently, Radford [2] has suggested that the unimolecular isomerization of methoxy to form the hydroxymethylene radical,

$$CH_3O \rightarrow CH_2OH, \quad (1)$$

followed by the fast reaction

$$CH_2OH + O_2 \rightarrow CH_2O + HO_2, \quad (2)$$

may provide an alternate route for the oxidation of methyl radicals. Subsequently, Batt et al. [3] pointed out that although the isomerization reaction should be competitive with the hydrogen-dissociation reaction of the methoxy radical,

$$CH_3O \rightarrow CH_2O + H, \quad (3)$$

neither reaction should occur quickly enough to compete with other faster reactions of the radical that are important in atmospheric chemistry. However, at the higher temperatures encountered in combustion processes, the unimolecular reactions may be more significant.

Here we report computed classical barriers for several unimolecular reactions of CH_3O and CH_2OH. In addition to (1) and (3), we also discuss the hydrogen-dissociation reactions of hydroxymethylene,

$$CH_2OH \rightarrow CH_2O + H, \quad (4)$$

and the hydrogen molecule elimination reaction of methoxy,

$$CH_3O \rightarrow CHO + H_2. \quad (5)$$

2. Outline of the calculations

For each of the open-shell species in this study an unrestricted Hartree–Fock (UHF) wavefunction is chosen for the ground electronic reference state. The structural parameters for the reactants and the various transition states are determined using the analytic gradient based GRADSCF codes [4] and a 6-31G** basis

* This work was supported, in part, by the Army Research Office under contract No. DAAG29-80-C-0105 to the University of Florida.

[5]. The many-body perturbation theory (MBPT) calculations use double zeta-quality basis sets [6] augmented with polarization functions on all atoms [6]. Since self-consistent-field calculations provide poor estimates of energy differences, the electronic energy is determined using MBPT [7]. In particular, fourth-order MBPT calculations that include all single-, double-, and quadruple excitation diagrams [SDQ MBPT(4)] are reported.

3. Equilibrium configurations for methoxy and hydroxymethylene

The structural parameters determined for the methoxy and hydroxymethylene radicals are shown in fig. 1. Our SCF calculations predict structural parameters for CH_3O that agree closely with those of Yarkony et al. [8]. Each of the SCF calculations, however, disagrees somewhat with the results of our previous correlated D MBPT(4) calculation for this radical [9]. In that study, Bent et al. showed that the C_{3v} symmetry axis is broken by a reduction in the in-plane OCH bond angle, while, to the contrary, the SCF calculations predict that the symmetry changes via an increase of the in-plane OCH bond angle. A correlated calculation of the CH_3O, CH_3OH, CH_2O, and HCO structures has also been published [10]. Since the structures of the radicals and transition states described in this work are determined using UHF calculations, for consistency, all energy differences for the methoxy reactions are computed at the structural parameters determined in the UHF calculations.

The electronic energies for the two radicals, including both the UHF and SDQ MBPT(4) results, are reported in table 1. Interestingly, the UHF calculations suggest that the methoxy radical CH_3O is more stable

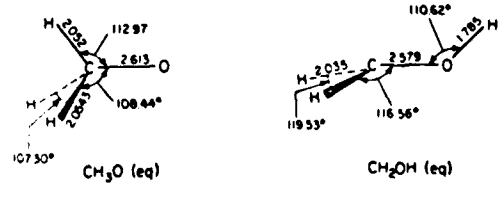


Fig. 1. UHF predicted equilibrium structures for CH_3O and CH_2OH (bond lengths in bohr).

Table 1
Electronic energy predictions (hartree)

Structure	UHF	SDQ MBPT(4)
CH_3O, equilibrium	−114.4484	−114.7621
CH_2OH, equilibrium	−114.4419	−114.7684
1,2-shift SP [a]	−114.3668	−114.7054
$CH_3O \rightarrow CH_2O + H$ SP	−114.3845	−114.7060
$CH_2OH \rightarrow CH_2O + H$ SP	−114.3679	−114.6923

[a] SP = saddle point.

than the hydroxymethylene radical CH_2OH. The energy difference equals 4.07 kcal/mole. On the other hand, the correlated SDQ MBPT(4) calculations predict that the hydroxymethylene radical is more stable, by 3.90 kcal/mole. Experimental studies [11,12] indicate that hydroxymethylene is 5–10 kcal/mole more stable than methoxy, although a more recent experimental [13] determination of the heat of formation of the methoxy radical would reduce this range by 3 kcal/mole. Our result of 3.90 is in good agreement with this lower (2–7 kcal/mole) estimate.

4. Isomerization and dissociation reactions

The isomerization reaction that converts methoxy to hydroxymethylene belongs to the reaction class identified as 1,2-hydrogen shifts. A recent review by Schaefer [14] discusses this class of reactions. The transition state identified for the isomerization reaction is shown in fig. 2. Such isomerizations are often ignored in chemical kinetic models, since it is assumed that the potential energy barrier for the reaction is large. As the results in table 1 show, however, the electronic energy for this transition state does not differ significantly from those values calculated for the two hydrogen-dissociation reactions.

In particular, as summarized in table 2, the classical barrier for the isomerization, reaction (1), is 35.6 kcal/mole (39.5 kcal/mole for the reverse reaction) with the competing H-dissociation, reaction (3), having a 35.2 kcal/mole barrier. The UHF transition state structures are shown in fig. 2. Since the 1,2 shift transition state has a more compact structure, one would expect that the zero-point vibrational energy differences would re-

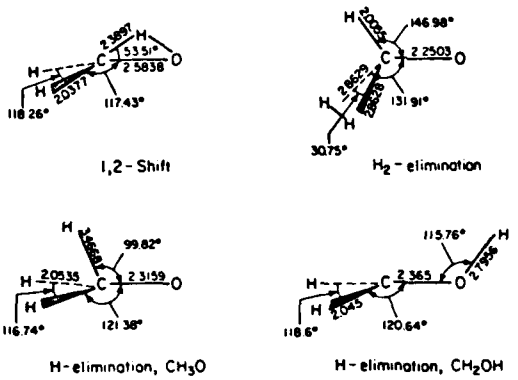


Fig. 2. UHF predictions of transition states for CH_3O undergoing 1,2-H shift, H elimination, H_2 elimination, and for H elimination from CH_2OH (bond lengths in bohr).

sult in a lower 0 K activation barrier for (3) than for (1). Nevertheless, the classical barriers are so nearly equal that it is expected that both reactions will occur under most conditions.

In addition to the above, the H-dissociation reaction, reaction (4), is also of interest. The competing reaction

$$CH_2OH \rightarrow HCOH + H \qquad (6)$$

is ruled out since its energy barrier is greater than 85 kcal/mole [15–17].

The saddle-point structure for (4) is shown in fig. 2. SDQ MBPT(4) calculations at the UHF determined geometries suggests a barrier of 47.8 kcal/mole for (4). Consequently, it would appear that the preferred unimolecular reaction of CH_2OH is the isomerization to CH_3O.

Table 2
Reaction barriers (kcal/mole)

Reaction	Barrier height [a]
$CH_3O \rightarrow CH_2O + H$	35.2
$CH_3O \rightarrow CH_2OH$	35.6
$CH_2OH \rightarrow CH_3O$	39.5
$CH_2OH \rightarrow CH_2O + H$	47.8

[a] All barrier heights are classical. No attempt has been made to correct for differences in the zero-point vibrational energies.

5. The H_2-elimination reaction of methoxy

The reaction of methoxy to produce the formyl radical and hydrogen molecule is similar to the molecular products reaction channel of formaldehyde [16–22]. Like formaldehyde, the methoxy radical has a hydrogen-dissociation channel, an isomerization channel, and an H_2-elimination channel. Recent MINDO/3 results suggest that the dominant dissociation of the radical is the H_2-elimination reaction [23]. Furthermore, those results suggest that the isomerization reaction and the H_2-elimination reaction have nearly the same activation energy.

We have made numerous efforts to locate the H_2-elimination reaction saddle point using the GRADSCF computer codes. The structure shown in fig. 2 corresponds to a structure for which the norm of the gradient satisfies a stringent convergence criterion. Proper identification of a saddle point, however, requires the calculation of the vibrational frequencies for the structure. If one of the vibrational frequencies is imaginary, then the structure corresponds to a saddle point. The structure shown in fig. 2 fails this test, because the vibrational analysis yields two imaginary frequencies. Additional study of the hypersurface in this region, however, suggests that the second imaginary frequency is due to an artifact in the SCF results for this structure. In particular, our UHF wavefunction breaks the C_s symmetry assumed for the problem. The symmetry breaking may be viewed as due to the existence of a large-amplitude wagging motion characterized by a double-minimum potential energy fun on. This vibration may be viewed as a rocking of the H_2 fragment perpendicular to the plane defined by the formyl radical fragment.

To test this view, we extended the "transition state" along the apparent reaction coordinate, and searched for a minimum on the hypersurface. This calculation leads to a minimum that corresponds to the products H_2 and HCO. Distortion of the structure along the normal mode of the second imaginary frequency followed by steepest-descent minimization appears to lead to *no* products. Instead, the search routine wanders in hyperspace near the transition state structure. After twelve evaluations, the structure remains almost equivalent to that located by the transition state search routine.

Finally, to test whether the symmetry breaking was an artifact of the UHF calculation, a transition-state search was carried out using an analytic gradient based

restricted Hartree–Fock calculation. This calculation also converged, in forty-eight function evaluations, to a structure with no plane of symmetry. The bond lengths and bond angles predicted in that calculation are close to those predicted by the UHF calculations. Furthermore, when force constants for the structure are calculated at the RHF level, we still obtain two imaginary frequencies. Thus there is no substantial difference between the UHF and RHF predictions.

In sum, our calculations to date demonstrate that the transition state does lead to the hydrogen molecule product. The calculations do not conclusively demonstrate the existence of a double-minimum potential well about the transition state. To determine whether this potential is, as we believe, an artifact of the SCF level calculations requires a re-examination of this transition-state region using a theoretical method that includes correlation effects. We anticipate reporting such calculations soon.

The electronic energy including the correlation energy for this transition state is −114.6732 au, and the reaction barrier, 55.8 kcal/mole, can be compared with the other reaction barriers in table 2. The barrier for the H_2-elimination reaction exceeds the barriers determined for the other methoxy radical reactions by more than 20 kcal/mole.

6. Discussion and conclusion

The energy barrier results determined by these calculations and summarized in table 2 support Radford's contention that the hydrogen-dissociation and isomerization reactions of methoxy radical could compete favorably under appropriate conditions. In addition, these results suggest that for hydroxymethylene, the isomerization reaction should be faster than the hydrogen-dissociation reaction. Finally, our preliminary results indicate that the H_2-elimination reaction of the methoxy radical has a much larger energy barrier than either of the other methoxy reactions considered here. In all cases, the calculations do not consider the sometimes important effect of the fourth-order triplet-excitation diagrams [22,24,25] and are limited to UHF determined structures. For example, the triple-excitation contribution in $H_2CO \rightarrow H_2 + CO$ reduced the barrier by ≈3 kcal/mole [22] and it is possible to obtain similar changes by using better, correlated structures. The spin contamination present in UHF calculations can sometimes cause additional problems. For the H_2-elimination pathway, the UHF multiplicity is 2.013 and SDQ MBPT(4) improves this. We expect that a correlated determination of the (5) transition state should eliminate any possible artifacts with the SCF transition state.

[1] K.L. Demerjian, J.A. Kerr and J.A. Calvert, Advan. Environ. Sci. Technol. 4 (1974) 1.
[2] H.E. Radford, Chem. Phys. Letters 71 (1980) 195.
[3] L. Batt, J.P. Burrows and G.N. Robinson, Chem. Phys. Letters 78 (1981) 467.
[4] A. Komornicki, National Resource for Computations in Chemistry, Software Catalog, Vol. 1, Program No. QHO4 (GRADSCF) (1980).
[5] P.C. Hariharan and J.A. Pople, Mol. Phys. 27 (1974) 209.
[6] T.H. Dunning Jr. and P.J. Hay, in: Modern theoretical chemistry, Vol. 3. Methods of electronic structure theory, ed. H.F. Schaefer III (Plenum Press, New York, 1977); L.T. Redmon, G.D. Purvis and R.J. Bartlett, J. Am. Chem. Soc. 101 (1979) 2856.
[7] G.F. Adams, G.D. Bent, R.J. Bartlett and G.D. Purvis, in: Potential energy surfaces and dynamics calculations, ed. D.G. Truhlar (Plenum Press, New York, 1981) p. 133.
[8] D.R. Yarkony, H.F. Schaefer III and S. Rothenberg, J. Am. Chem. Soc. 96 (1974) 656.
[9] G.D. Bent, G.F. Adams, R.J. Bartram, R.J. Bartlett and G.D. Purvis, J. Chem. Phys., to be published.
[10] G.F. Adams, G.D. Bent, G.D. Purvis and R.J. Bartlett, Chem. Phys. Letters 81 (1981) 461.
[11] J.A. Kerr, Chem. Rev. 66 (1966) 465.
[12] M.A. Haney and J.L. Franklin, Trans. Faraday Soc. 65 (1969) 1794.
[13] P.C. Engelking, G.B. Ellison and W.C. Lineberger, J. Chem. Phys. 69 (1978) 1826.
[14] H.F. Schaefer III, Accounts Chem. Res. 12 (1979) 288.
[15] S.W. Benson, Thermochemical kinetics, 2nd Ed. (Wiley, New York, 1976).
[16] J.D. Goddard and H.F. Schaefer III, J. Chem. Phys. 70 (1979) 5117.
[17] J.D. Goddard, Y. Yamaguchi and H.F. Schaefer III, to be published.
[18] L.B. Harding, H.B. Schlegel, R. Krishnan and J.A. Pople, J. Phys. Chem. 84 (1980) 3394.
[19] G.F. Adams, G.D. Bent, R.J. Bartlett and G.D. Purvis, J. Chem. Phys. 75 (1981) 834.
[20] A. Horowitz and J.G. Calvert, Intern. J. Chem. Kinetics 10 (1978) 713.
[21] P.C. Houston and C.B. Moore, J. Chem. Phys. 65 (1976) 757.
[22] M.J. Frisch, R. Krishnan and J.A. Pople, J. Phys. Chem. 85 (1981) 1467.
[23] J. Ritchie and M.J.R. Dewar, private communication.
[24] R.J. Bartlett and G.D. Purvis, Intern. J. Quantum Chem. 14 (1978) 561.
[25] R. Krishnan, M.J. Frisch and J.A. Pople, J. Chem. Phys. 72 (1980) 4244.

INTERNATIONAL JOURNAL OF QUANTUM CHEMISTRY, VOL. XXIII, 437-446 (1983)

Electronic Structure and Vertical Excitation Spectrum of Methylene Amidogen CH_2N

GEORGE F. ADAMS
USA ARRADCOM/BRL, Aberdeen Proving Ground, Maryland 21005, U.S.A.

DAVID R. YARKONY*
Department of Chemistry, The Johns Hopkins University, Baltimore, Maryland 20218, U.S.A.

RODNEY J. BARTLETT AND GEORGE D. PURVIS†
Quantum Theory Project, University of Florida, Gainesville, Florida 32611, U.S.A.

Abstract

Ab initio electronic structure calculations are reported for five electronic states of the methylene amidogen radical. Structure parameters for the ground electronic state are predicted by RHF and D-MBPT (4) calculations. Vertical excitation energies were determined using four different theoretical chemical models: complete active space (CAS) MCSCF, CAS/MCSCF plus singles and doubles CI, fourth-order many-body perturbation theory SDQ-MBPT(4), and coupled-cluster theory.

1. Introduction

Experimental study of the spectra and properties of molecules allows the accurate determination of the structural parameters, electronic state separations, and the distribution of electrons in the molecule. However, not all molecular species are susceptible to detailed experimental study. Some free radicals possess such short lifetimes that even the demonstration of existence may challenge the most sophisticated experiment. Other radicals, such as methylene amidogen, have been experimentally detected [1–6], but enjoy such a transient existence that the structure and properties have not been determined. As noted by Schaefer [7], theory may play an important role in the study of evanescent species, since the tools used by the electronic structure theorist are equally applicable to stable well-characterized molecules and transient radical species. Results obtained via theoretical methods should be as accurate for a short-lived radical, such as CH_2N, as those obtained for a molecule such as formaldehyde. This works, provided that the method used accounts properly for open-skill character. We present here the results of *ab initio* electronic structure calculations on the methylene amidogen radical CH_2N.

The first experimental evidence for the CH_2N radical was provided by the ESR spectrum of Cochran, Adrian, and Bowers [1]. Subsequent work by Banks and Gordy

* Supported, in part, by the U.S. Army Armaments Research and Development Command.
† Supported by the Army Research Office under contract No. DAAG29-80-C-0105 to the University of Florida.

 CCC 0020-7608/83/020437-10$02.00

[2] confirmed the identification and provided a qualitative description of the molecular orbital occupied by the unpaired electron. The ESR experiments, and the issue of hyperconjugation, spawned a series of self-consistent field molecular orbital calculations, culminating in Hinchliffe's recent report [8–10].

The electronic spectrum of methylene amidogen was first described by Ogilvie and Horne [3]. Subsequently, Horne and Norrish published a detailed electronic spectrum of this radical [4]. Their spectrum consisted of a diffuse absorption with a maximum intensity at 280.8 nm and a second, weaker absorption in four diffuse bands with intensity maxima at 284.5, 284.75, 285.0, and 285.4 nm. Ogilvie interpreted the spectrum of CH_2N to be due to two different electronic transitions originating in the ground state, and he reiterated the validity of his maximum absorption coefficients [5].

Methylene amidogen was identified by Morgan and Beyer as a primary decomposition product of the 1,3,5,7-tetranitro 1,3,5,7-tetrazocyclooctane (HMX) molecule [6]. The radical has been suggested to play a role in the reaction of H atom with HCN [11], and has been suggested as an intermediate in flames oxidized by N_2O. The interest of flame diagnosticians spurred this study of the electronic structure and the excitation energy spectrum of this radical. A recent report by So [12] reports structural parameters for two electronic states of CH_2N. Here we report structural parameters for the ground electronic state of this radical, as well as a series of calculations designed to elucidate the details of the observed absorption spectrum of this radical. Furthermore, the calculations provide an interesting comparison of results obtained by application of MCSCF-CI methods with those obtained using two linked cluster methods MBPT and CCM.

2. Description of Theoretical Methods

All the calculations reported here were performed using basis sets of double-zeta plus polarization quality. Those many-body perturbation theory (MBPT) calculations performed to determine the structural parameters for the ground electronic state of methylene amidogen were performed using Dunning's [13] double-zeta contractions of Huzinaga's [14] $9s5p$ primitive set for carbon and nitrogen and the double-zeta contraction of the $4s$ primitive set for hydrogen. A single set of d-type polarization functions [15] augments the basis for the first row atoms, while the hydrogen basis set includes a single set of p-type functions [15]. All other calculations reported here use the 6-31G^{**} basis set proposed by Hariharan et al. [16]. These include the RHF predictions of structural parameters for the ground electronic state of CH_2N, and all the vertical excitation energy calculations.

In its ground electronic state, methylene amidogen has C_{2v} structural symmetry. The unpaired electron occupies an in-plane nitrogen orbital; thus the ground state has B_2 symmetry. The orbital occupation is $(1a_1)^2 \cdots (5a_1)^2(1b_2)^2(2b_2)(1b_1)^2$. This state is represented by the orbital diagram [Fig. 1(a)]. Figures 1(b) to 1(d) represent electronic excited states formed by excitations involving the $5a_1$, $2b_2$, and $1b_1$ orbitals; the excitations lead to the $1\,^2B_1$, $1\,^2A_1$, and $2\,^2B_1$ electronic states. The calculations described here refer to vertical excitations from $X\,^2B_2$ to each of these excited states, plus the excitation to the lowest quartet excited state ($^4B_2$).

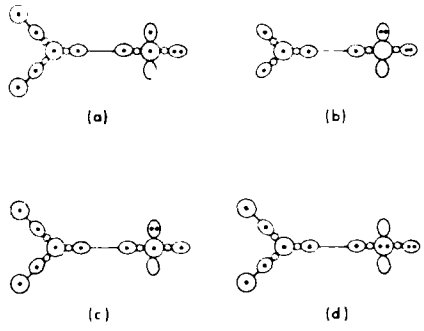


Figure 1. Valence bond representations of the first four electronic states of methylene amidogen. The figure plane is the XZ plane. (a) $1\,^2B_2$, (b) $1\,^2B_1$, (c) $1\,^2A_1$, (d) $2\,^2B_2$.

The application of MBPT has been described in a series of papers by the present authors and others [17–19]. A recent review paper by Bartlett [20] offers a description of the methods, summarizes recent results, and provides many references. The MBPT calculations for methylene amidogen include results through fourth order; the fourth-order calculations include all single-, double-, and quadruple-excitation diagrams which contribute.

The coupled-cluster method calculations reported here use the computational procedures described by Purvis and Bartlett [21]. In particular, the cluster operator includes the single- and double-excitation operators, viz.,

$$T = T_1 + T_2.$$

The particulate *Ansatz* described by Purvis and Bartlett includes both single- and double-excitation cluster effects, as well as disconnected triple excitations (T_1T_2 and $\frac{1}{2}T_1^3$) and quadruple excitations (T_2^2, $\frac{1}{2}T_1^2T_2$ and $\frac{1}{4}T_1^4$).

Finally, the states in question were characterized at the MCSCF/CI level using a procedure originally applied to describe the low-lying states of MgO [22]. Using an active space comprising five molecular orbitals ($5a_1$, $1b_1$, $2b_1$, $1b_2$, and $2b_2$), complete active space comprising five molecular orbitals ($5a_1$, $1b_1$, $2b_1$, $1b_2$, and $2b_2$), complete active space (CAS) MCSCF calculations were performed for the five electronic states discussed here. In each case, the CAS space was reduced to a compact configuration cients greater than 0.075 in the CAS MCSCF. The resulting compact CSF descriptions (see Table I), provide a qualitatively accurate description of these states and the reference space for a single- and double-excitation CI (SDCI) described below.

3. Structure of the $X\,^2B_2$ Ground State

While spectroscopic data demonstrate that methylene amidogen exists as a stable species with two equivalent hydrogen atoms, no empirical structural parameters have been derived for this radical. As a part of this study, we obtained optimal structural parameters for the ground electronic state of this radical using the HHL BRL analytical gradient driven SCF–MCSCF programs [23]. These results may be compared with

AD-A140 738 PREDICTION OF TRANSITION STATES AND THERMOCHEMISTRY FOR COMBUSTION REACTIONS(U) BATTELLE COLUMBUS LABS OH R J BARTLETT 07 APR 84 ARO-16929.4-PH DAAG29-80-C-0105 2/2

UNCLASSIFIED F/G 20/10 NL

END

DATE FILMED

6-84

DTIC

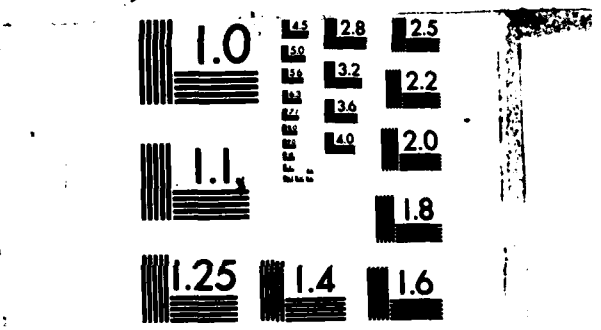


TABLE I. MCSCF expansions.

State	Label	Electron Occupations[a]	Dimension of SD-CI	
			Unselected	Selected
1^2B_2	A	$4a_1^2\ 5a_1^2\ 1b_2^2\ 1b_1^2\ 2b_2$	16879	9209
		$4a_1^2\ 5a_1^2\ 1b_2^2\ 2b_1^2\ 2b_2$		
1^2B_1	B	$4a_1^2\ 5a_1^2\ 1b_2^2\ 2b_2^2\ 1b_1$	16965	8686
		$4a_1^2\ 1b_2^2\ 2b_2^2\ 1b_1^2\ 2b_1$		
1^2A_1	C	$4a_1^2\ 1b_2^2\ 2b_2^2\ 1b_1^2\ 5a_1$	16760	9220
		$4a_1^2\ 1b_2^2\ 2b_2^2\ 2b_1^2\ 5a_1$		
1^4B_2	D	$4a_1^2\ 5a_1^2\ 1b_2^2\ 2b_2\ 1b_1\ 2b_1$	18219	7962
$1,\ 2^2B_1$	E	$4a_1^2\ 5a_1^2\ 1b_2^2\ 2b_2^2\ 1b_1$		
		$4a_1^2\ 1b_2^2\ 2b_2^2\ 1b_1^2\ 2b_1$	33852	16194
		$4a_1^2\ 5a_1^2\ 1b_2^2\ 1b_1^2\ 2b_1$		
		$4a_1^2\ 5a_1^2\ 1b_2^2\ 2b_2^2\ 2b_1$		

[a] Fully occupied $1a_1$, $2a_1$, $3a_1$ orbitals are omitted for clarity.

results obtained in an earlier study using fourth-order many-body perturbation theory calculations [24]. Both sets of structural data are presented in Table II, along with the results of previously published SCF calculations [10,12]. The notable result is that each of the methods predicts the ground state to have C_{2v} structural symmetry. The modestly larger basis sets used in Ref. 10 and these MBPT calculations do not lead to results startling in their diversity, although the C—H bond length predicted in Ref. 10 is longer than normally observed for such a bond, while the bond angle predicted

TABLE II. Structural parameters for CH_2N.[a]

	R_{C-H}	R_{C-N}	< HCH
Reference 12	1.078	1.263	118.23
Reference 10	1.14	1.24	119
This work			
SCF	1.08	1.24	114.2
MBPT	1.095	1.23	117

[a] Units: bond lengths, angstroms, angles, degrees.

in our own SCF calculations is significantly outside the range of the other three predictions.

One other detail of the calculations should be noted. The MBPT and CCSD calculations employ single-determinental unrestricted Hartree-Fock wave function as a reference function. This does not imply a lack of generality of these methods, but identifies a restriction on the application of these techniques. Since UHF wave functions are not constructed to be eigenfunctions of the total spin, there is the possibility that the reference function may contain contributions from wave functions of higher spin, than the state under consideration. In this study of CH_2N, both the ground electronic state, $X\ ^2B_2$, and the first excited state of A_1 symmetry have spin multiplicities much greater than two (see Table III). Thus, some caution should be exercised in interpreting energy differences that include these two states. As noted by Adams et al. [25], spin contamination in reference functions degrades the accuracy of MBPT/CCM calculations, thus affecting the computed electronic excitation energy results.

4. Vertical Excitation Energies of Methylene Amidogen

As noted in Section 1, the observed electronic spectrum of CH_2N comprises two diffuse absorptions; a strong, diffuse band with a maximum intensity at 280.8 nm (35612.5 cm^{-1}) and a weaker, absorption in four diffuse bands, with maximum intensities at 284.5 nm (35149.3 cm^{-1}), 284.75 nm (35118.5 cm^{-1}), 285.0 nm (35087.0 cm^{-1}), and 285.4 nm (35038.5 cm^{-1}) [4]. Ogilvie assigns the two absorptions to separate electronic excitations originating in the $X\ ^2B_2$ ground state, but he makes no assignment of the upper states [5]. Here we present two distinct sets of quantum chemical calculations that provide a reasonable explanation of the electronic spectrum of this radical. All of these calculations were performed using the structural parameters predicted by the present RHF calculations. Thus, the carbon-hydrogen bond length is 1.08 Å, the carbon-nitrogen distance is 1.24 Å, and the HCH bond angle is 114.2°.

All MBPT and CCSD calculations were performed relative to single determinant reference functions. The UHF wave functions for the $X\ ^2B_2$ and $B\ ^2A_1$ electronic states

TABLE III. UHF spin multiplities for methylene amidogen.

State	Spin-Multiplicity
1^2B_2	2.178
1^2B_1	2.004
1^2A_1	2.216
2^2B_2	2.011
1^4B_1	4.008

had spin multiplicities 2.178 and 2.216, respectively. These results suggest the occurrence of at least one low-lying quartet state. The UHF wave functions for the four states studied had the orbital occupations

$$1\,^2B_2 \quad (1a_1)^2\cdots(5a_1)^2(1b_1)^2(1b_2)^2(2b_2),$$
$$1\,^2B_1 \quad \cdots(5a_1)^2(1b_1)(1b_2)^2(2b_2)^2,$$
$$1\,^2A_1 \quad \cdots(4a_1)^2(5a_1)(1b_1)^2(1b_2)^2(2b_2)^2,$$
$$2\,^2B_1 \quad \cdots(5a_1)^2(1b_1)^2(2b_1)(1b_2)^2,$$
$$1\,^4B_2 \quad \cdots(5a_1)^2(1b_1)(2b_1)(1b_2)^2(2b_2).$$

The excitation energies predicted by the UHF results are collected in Table IV. At this level of theory, the excitation to the quartet state exceeds the lowest excitation energy by less than 1000 cm^{-1}. Thus, a low-lying quartet state has, apparently, been located. In addition, the UHF results predict that the excitations to the $1\,^2A_1$ and the $2\,^2B_1$ states differ by only 299 cm^{-1}, suggesting the relatively small observed separation in the absorption spectrum (ca. 520 cm^{-1}). However, the results of the UHF calculations for the $2\,^2B_1$ state do not yield an upper bound to the electronic energy, so we assume a substantial error in the calculation of the corresponding excitation energy, even improved by the MBPT and CCSD calculations. The excitation energy must be too low.

The results obtained from the MBPT and CCSD calculations are presented in Tables IV and V. Both of these methods predict a large increase in the excitation energy for the quartet state relative to UHF results. Furthermore, the calculations agree well with one another on the excitation energy for the $X\,^2B_2 \rightarrow 1\,^2A_1$ process: MBPT, 36,064 cm^{-1} and CCSD, 36,024 cm^{-1}. However, the MBPT calculations predict a much larger difference between the first and second excitation energies than do the CCSD calculations, while predicting a smaller difference between the second and third excitation energies. Were one to attempt to assign the absorption spectrum based upon the results of the MBPT calculations, it would be tempting to assign the processes as $1\,^2B_2 \rightarrow 1\,^2A_1$ and $1\,^2B_2 \rightarrow 2\,^2B_1$. Such an assignment, however, seems unlikely if one recalls

TABLE IV. Electronic energies for CH_2N: UHF, MBPT, and CCSD results.[a]

State	UHF	SDQ-MBPT(4)	CCSD
$1\,^2B_2$	-93.43539	-93.70776	-93.71195
$1\,^2B_1$	-93.29176	-93.55510	-93.55641
$1\,^2A_1$	-93.27033	-93.54368	-93.54502
$2\,^2B_1$	-93.26897	-93.54003	-93.54253
$1\,^4B_2$	-93.29669	-93.51969	-93.52101

[a] Units: hartrees.

TABLE V. Excitation energies for CH_2N: UHF, MBPT, and CCSD results.[a]

Excitation	UHF	MBPT	CCSD
$1^2B_2 \rightarrow 1^2B_1$	31534	33513	34146
$1^2B_2 \rightarrow 1^2A_1$	36239	36064	36024
$1^2B_2 \rightarrow 2^2B_1$	36538	36882	37192
$1^2B_2 \rightarrow 1^4B_2$	32465	41449	41917

[a] Units: cm^{-1}.

that the higher energy absorption is significantly stronger than the diffuse, longer wavelength transition. Since a $B_2 \rightarrow B_1$ transition is symmetry forbidden, intensity arguments lead us to reject this proposed assignment.

The only significant difference between the MBPT and CCSD calculations is the spacing of the $^2A_1$ state relative to the two $^2B_1$ states. The CCSD results reduce the energy gap between the $1\,^2B_1$ state and the $1\,^2A_1$ state while increasing the gap between the $1\,^2A_1$ and $2\,^2B_1$ states. The predicted $1\,^2B_2 \rightarrow 1\,^2A_1$ separation is almost equal to that predicted by the MBPT calculations. These results do not lend credence to the MBPT based interpretation of the experimental spectrum, nor do they provide an alternative suggestion. An analysis of the results of the MBPT and CCSD calculations reveals the importance of single- and double-excitation contributions involving the $1b_1$ and $2b_1$ molecular orbitals. The reasons for this dominance will become apparent as we discuss the results of the MCSCF calculations for this radical.

The results of the MCSCF calculations are summarized in Tables VI and VII. Note that of the two CSF expansions B and E reported for the $^2B_1$ states, expansion B is appropriate for a description of the $1\,^2B_1$ state relative to the $1\,^2B_2$, $1\,^2A_1$, and $1\,^4B_2$ states, while expansion E is required to treat the $1\,^2B_1$ and $2\,^2B_1$ states in an even-handed manner. Those results presented in Table VI show that the $1\,^4B_2$ state is well

TABLE VI. Energies of lowest root CH_2N states at $1\,^2B_2$ equilibrium geometry.[a]

	Method		MCSCF Coefficients[b]	
State	MCSCF	SDCI	C_1	C_2
1^2B_2	-93.47720	-93.70540	0.979	-0.203
1^2B_1	-93.29265	-93.54353	0.998	-0.056
1^2A_1	-93.29114	-93.54126	0.978	-0.206
1^4B_2	-93.27667	-93.50992	1.0	

[a] Units: hartrees.
[b] See Table I for definition of corresponding CSF's.

TABLE VII. MCSCF wave functions for the 1,2 $^2B_1$ states of CH_2N.

State	Energy[b]	MCSCF Coefficients[a]			
		C_1	C_2	C_3	C_4
1^2B_1 [c]	-93.30438	0.598	-0.032	-0.120	0.791
2^2B_1 [c]	-92.84998	0.084	0.984	-0.152	-0.047
1^2B_1 [d]	-93.29309	0.781	-0.507	0.022	0.362
2^2B_1 [d]	-93.26699	0.468	-0.134	0.855	0.179

[a] See Table I for definition of corresponding CSF's.
[b] Units: hartrees.
[c] From 1 $^2B_1$ optimization.
[d] From 2 $^2B_1$ optimization.

described by a single reference CSF, the 1 $^2B_2$ and 1 $^2A_1$ states have appreciable multiconfigurational character, while the 1 $^2B_1$ state has slight multiconfigurational character. The results gathered in Table VII show that the 2 $^2B_1$ state has substantial multiconfigurational character; these results also indicate that the attribute "multiconfigurational" depends upon the method used to derive the corresponding MCSCF expansion. This is exemplified by the results for the 1 $^2B_1$ state presented in Table VII, wherein the state has more multiconfigurational character than indicated by the data presented in Table VI.

The MCSCF wave functions comprise the basis for the construction of extended CI wave functions for these electronic states. Initially all CSF's which are single- or double-excitations away from any of the reference states were included in the CSF expansion. The *K* shells of carbon and nitrogen were kept fully occupied in this procedure. This expansion was reduced by an a priori selection procedure which approximates the Hartree-Fock interacting space for geneologically coupled CSF's [26]. The dimension for each SDCI space is presented in Table I. Since the CSF space describing the 1,2 $^2B_1$ states is almost twice as large as any of the other expansions in the table, it is clear that these roots must be treated in a manner distinct from that appropriate to the treatment of the lowest root of each symmetry.

The results of the SDCI calculations are presented in Tables VI and VIII. Table IX summarizes the excitation energy results. These results provide an obvious interpretation of the absorption spectrum. The low-energy weak absorption with appreciable fine structure corresponds to the 1 $^2B_2 \rightarrow$ 1 $^2B_1$ transition. Its weak intensity and structure are attributed to the transition being vibronically allowed by electronically forbidden in C_{2v} symmetry. The strong absorption about 500 cm^{-1} to the blue is assigned to the symmetry allowed 1 $^2B_2 \rightarrow$ 1 $^2A_1$ transition. The position of the 2 $^2B_1$ state, 6100 cm^{-1} to the blue of the 1 $^2B_1$ state, is reassuring, since this eliminates the state as a possible participant in the absorption spectrum. This result is significant,

TABLE VIII. SDCI Energies of the 1,2 $^2B_1$ states of CH_2N.[a]

Orbitals	1^2B_1	2^2B_1
Orbitals optimized for 1^2B_1 state	-93.54724	-93.451216
Orbitals optimized for 2^2B_1 state	-93.54964	-93.52181

[a] Units: hartrees.

since the agreement of the SDCI excitation energies with the experimental values must be regarded as fortuitous.

The lack of agreement between the SDCI results and the MBPT/CCSD results could have been anticipated on two grounds. There is no adequate single CSF description for the $2\,^2B_1$ state, so one expects that a substantial error should occur in treating both of the $^2B_1$ states. Furthermore, spin contamination of the $^2A_1$ and $^2B_2$ wave functions suggests a contribution to the UHF wave function that is not ignorable. That both the MBPT and CCSD methods predict the excitation energy for the $1\,^2B_2 \rightarrow 1\,^2A_1$ transition as does the SDCI technique is both intriguing and mystifying.

5. Summary

We have reported an extensive set of electronic structure theory calculations for several excited states of the methylene amidogen radical. The results of MCSCF/SDCI calculations permit the assignment of the observed electronic absorption spectrum for the radical. A comparison of the results of MCSCF/SDCI calculations with those obtained using the MBPT and CCSD techniques show reasonable agreement for several excitation energies, but poor agreement when considering excitations to two $^2B_1$ excited states. It seems clear that the single reference CSF's used as reference functions in the MBPT/CCSD calculations are not adequate descriptions of the $^2B_1$ electronic states. Several other sources of error should be noted. None of the theoretical methods considered the major triple excitation contributions to the correlation energy. Further, the atomic orbital basis set should have included diffuse functions to provide a better description of the $2\,^2B_1$ state.

TABLE IX. Excitation energies for CH_2N.

Transition	METHOD		
	MBPT	CCSD	SDCI
$1^2B_2 \leftarrow 1^2B_1$	33513	34146	35525
$1^2B_2 \leftarrow 1^2A_1$	36064	36024	36024
$1^2B_2 \leftarrow 2^2B_1$	36882	37192	41633
$1^2B_2 \leftarrow 1^4B_2$	41449	41917	42903

Bibliography

[1] E. L. Cochran, F. J. Adrian, and V. A. Bowers, J. Chem. Phys. **35**, 1938 (1962).
[2] D. Banks and W. Gordy, Mol. Phys. **26**, 1555 (1973).
[3] J. F. Ogilvie and D. G. Horne, J. Chem. Phys. **48**, 2248 (1968).
[4] D. G. Horne and R. G. W. Norrish, Proc. R. Soc. London, Ser. A **315**, 301 (1970).
[5] J. F. Ogilvie, Can. J. Spectrosc. **19**, 89 (1974).
[6] C. U. Morgan and R. A. Beyer, Combust. Flame (to appear).
[7] H. F. Schaefer III, Acc. Chem. Res. **79**, 288 (1979).
[8] T. A. Claxton and N. A. Smith, Trans. Faraday Soc. **67**, 1859 (1971); T. A. Claxton, *ibid.* **66**, 897 (1970).
[9] N. C. Baird and R. F. Barr, Theor. Chim. Acta **36**, 125 (1974).
[10] A. Hinchliffe, J. Mol. Structure **67**, 101 (1980).
[11] W. R. Anderson (private communication).
[12] S. P. So, Chem. Phys. Lett. **82**, 370 (1981).
[13] T. H. Dunning, J. Chem. Phys. **53**, 2823 (1970).
[14] S. Huzinaga, J. Chem. Phys. **42**, 1293 (1965).
[15] T. Dunning and P. J. Hay, *Modern Theoretical Chemistry. Vol. III. Methods of Electronic Structure Theory*, H. F. Schaefer III, Ed. (Plenum, New York, 1977).
[16] P. C. Hariharan and J. A. Pople, Mol. Phys. **27**, 209 (1974).
[17] G. F. Adams, G. D. Bent, R. J. Bartlett, and G. D. Purvis, J. Chem. Phys. **75**, 834 (1981).
[18] R. J. Bartlett and G. D. Purvis, Phys. Scrip. **21**, 255 (1980).
[19] J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. **14**, 545 (1978).
[20] R. J. Bartlett, Ann. Rev. Phys. Chem. **32**, 359 (1981).
[21] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).
[22] C. W. Bauschlicher, B. L. Lengsfield, D. M. Silver, and D. R. Yarkony, J. Chem. Phys. **74**, 2397 (1981).
[23] D. R. Yarkony, J. Chem. Phys. (to appear).
[24] G. F. Adams, ARRADCOM Tech. Rept. (to appear).
[25] G. F. Adams, G. D. Bent, R. J. Bartlett, and G. D. Purvis, in *Potential Energy Surfaces and Dynamics Calculations*, D. G. Truhlar, Ed. (Plenum, New York, 1977).
[26] W. C. Swope, H. F. Schaefer III, and D. R. Yarkony, J. Chem. Phys. **73**, 407 (1980).

Volume 98, number 1 CHEMICAL PHYSICS LETTERS 10 June 1983

COMPARISON OF MBPT AND COUPLED-CLUSTER METHODS WITH FULL CI. IMPORTANCE OF TRIPLET EXCITATIONS AND INFINITE SUMMATIONS *

Rodney J. BARTLETT, Hideo SEKINO and George D. PURVIS III
Quantum Theory Project, Departments of Chemistry and Physics, University of Florida, Gainesville, Florida 32611, USA

Received 21 March 1983

Results from full fourth-order perturbation theory [SDTQ MBPT(4)], and the coupled-cluster single- and double-excitation model (CCSD), are compared with recent full CI results for BH, HF, NH_3, and H_2O. For H_2O, studies include large symmetric displacements of the OH bonds, which offer a severe test for any single-reference MBPT/CC method. In every case, CCSD plus fourth-order triple-excitation terms provide agreement with the full CI to < 2 kcal/mole. SDTQ MBPT(4) has an error 10 kcal/mole for displaced H_2O.

1. Introduction

Recently, a series of full CI calculations on molecules with small basis sets have become available [1,2]. Such calculations provide interesting benchmark results that may be compared with results from different methods for electron correlation. The first calculation of this type was for H_2O in a double-zeta (DZ) basis set reported by Saxe et al. [1]. Comparison many-body perturbation theory (MBPT) and coupled-cluster doubles (CCD) results have been reported by Bartlett [3], and coupled-cluster singles and doubles (CCSD) results by Purvis and Bartlett [4]. Comparison MBPT and CCSD results in a crude localized-orbital basis set have now been added to this list [5]. The MBPT/CC results are in excellent agreement with the full CI for these examples. However, H_2O at its equilibrium geometry is well described by a single SCF reference function. This determinant has a coefficient of 0.975 in the full CI wavefunction, built upon the usual SCF orbitals. Consequently, this is not a difficult case for single-reference-function-correlated methods like MBPT/CC.

Harrison and Handy [2] have now extended this study of DZ H_2O to large symmetric displacements of the OH bonds, extending to twice the equilibrium distance. For such displacements other single and double excitations in the CI wavefunction are of comparable importance to the SCF determinant. Consequently, these displaced geometries offer a far more stringent test of the applicability of single-reference MBPT/CC methods. In particular, one would expect the infinite-order summations of correlation corrections that are built into coupled-cluster methods to be more important than is the case at the equilibrium geometry [6].

In addition to DZ H_2O at displaced geometries, Harrison and Handy have also reported full CI results for BH, HF, and NH_3. In the following, we report full fourth-order MBPT, and CCSD results for all of these examples. In the case of H_2O displaced from equilibrium, we have also considered the differences in the results of MBPT/CC calculations based upon restricted (RHF) and unrestricted (UHF) reference functions [6,7].

2. Theory

There are three currently popular approaches to the accurate inclusion of electron correlation effects in molecules; configuration interaction (CI) [8,9],

* This research has been supported by the Army Research Office under contract No. DAAG29-82-K-0034.

many-body perturbation theory (MBPT) [10,11], and coupled-cluster (CC) methods [12,13]. These differ in how the approximations are introduced, which, in turn, result in different properties for the calculations. For example, truncated CI results are variational but not size-extensive while MBPT and CC methods are not variational but scale correctly with size [3,14]. The full CI, where all possible N-tuple excitations are included for N electrons, is both variational and size-extensive and also separates correctly into fragments when a bond is broken.

The CI wavefunction relative to some single-determinant reference function Φ_0 (often the SCF determinant) may be written (intermediately normalized) as

$$\psi_{CI} = \Phi_0 + \sum_{n}^{N} \hat{C}_n |\Phi_0\rangle , \tag{1}$$

$$\hat{C}_n = \sum_{\substack{i<j<k... \\ a<b<c...}}^{(n)} C_{ijk...}^{abc...} a^\dagger i b^\dagger j c^\dagger k... \quad (n = 1, 2, 3, ..N) , \tag{2}$$

where (n) indicates the level of excitation and N is the number of electrons. The operators and indices $i, j, k, ...$ represent orbitals occupied in Φ_0 while $a, b, c...$ are excited orbitals, and the coefficients $C_{ijk...}^{abc...}$ are to be determined for each n variationally. If all possible excitations are included, i.e. $n = N$, we obtain the full CI result. Except for small problems, the CI wavefunction must be truncated. For many years, the level of all single and double excitations from Φ_0, i.e. CISD, has been routinely applicable to realistic problems [3,15]. More recently, modified CI methods that first determine an MC SCF or CI reference space, and then form all single and double excitations from that entire space, are becoming applicable [16,17]. Such methods presumably introduce the dominant higher-excitations relative to a single reference, Φ_0.

The fundamental formula of MBPT is the linked-diagram theorem [10], which gives the wavefunction and energy as

$$\Psi_{MBPT} = (1 + V)\Phi_0 + \sum_{k=1}^{\infty} [(E_0 - H_0)^{-1} V]^k |\Phi_0\rangle_L , \tag{3}$$

$$E = E_0 + \sum_{k=0}^{\infty} \langle\Phi_0| V [(E_0 - H_0)^{-1} V]^k |\Phi_0\rangle_L . \tag{4}$$

The subscript L indicates the limitation to linked diagrams. Although various categories of infinite-order summations are obtainable from eq. (4), the method is usually limited by terminating at some order of perturbation theory. The most attractive current MBPT model is the complete fourth-order result, which consists of all single-, double-, triple- and quadruple-excitation terms [18]. Because of the contribution of triple and quadruple excitations, this model should exceed CISD in accuracy. Of course, any linked-diagram approximation is size extensive [3].

The coupled-cluster theory uses the ansatz [12,13] for the wavefunction

$$\Psi_{CC} = \exp(T)|\Phi_0\rangle , \tag{5}$$

$$T = T_1 + T_2 + T_3 + ... , \tag{6}$$

$$T_n = \sum_{\substack{i<j<k... \\ a<b<c...}}^{(n)} t_{ijk...}^{abc...} a^\dagger i b^\dagger j c^\dagger k... \quad (n = 1, 2, ..N) .$$

Essential approximations are introduced by restricting T to some categories of excitation. The current state of the art is CCSD [4], where

$$\Psi_{CCSD} = \exp(T_1 + T_2)|\Phi_0\rangle \tag{8}$$

and

$$E = \langle\Phi_0|H|\Psi_{CCSD}\rangle .$$

Since the complete solution to Ψ_{CCSD} includes *all* powers of T_1 and T_2 and their coupling, a comparison with CI demonstrates the inclusion of large parts of higher excitation terms. For example,

$$C_4 = \tfrac{1}{2}T_2^2 + T_1 T_3 + \tfrac{1}{2}T_1^2 T_2 + \tfrac{1}{4}T_1^4 + T_4 , \tag{9}$$

hence by including non-linear T_2 and T_1 terms most of the quadruple contribution is obtained without requiring explicit evaluation of T_4, e.g refs. [6,7]. The MBPT linked-diagram expansion can be obtained by iteratively solving the CC equations. All single, double and quadruple terms that arise in fourth-order MBPT (the SDQ MBPT(4) model) come from low-order iterations of CCSD. However, the fourth-order triples arise from the first contribution of T_3 to the CCSDT equations. Coupled-cluster theory is clearly size-extensive as guaranteed by the exponential ansatz, and unlike MBPT, is an infinite-order method. MBPT/CC methods using a multi-reference space are also possible but are not discussed here [19].

3. Boron hydride

In table 1 we present results for the BH molecule, at the experimental equilibrium in a double-zeta plus polarization (DZP) basis set. The basis set and geometry are the same as Harrison and Handy [2]. Results are compared including the core-correlation effects, and with the 1s electrons unexcited.

The BH molecule exhibits slower convergence than average in MBPT calculations [6,20] due to the near degeneracy of the 2p and 2s orbitals. This is reflected in table 1 by the better convergence of the infinite-order CCSD model compared to its fourth-order MBPT approximation, SDQ MBPT(4). The best agreement of the many-body methods with the full CI is given by CCSD + T(4), where the fourth-order triple excitation terms that are neglected in CCSD are added (see ref. [21] for the method for including the triple-excitation contributions). For this example, these terms amount to 0.4 kcal/mole.

The best agreement of the MBPT/CC methods compared to full CI for the dissociation energy is also obtained by the CCSD + T(4) model, which is within 0.2 kcal/mole of the full CI result. Since B is an open-shell atom, the MBPT/CC calculations for the atom are based upon an unrestricted Hartree–Fock (UHF) reference function. In this example, as is usually the case for open-shell problems, the computed CCSD multiplicity is quite accurate (see ref. [22] for the procedure for evaluating S^2), giving 2.000016, compared to 2.00505 for the UHF reference function.

4. Hydrogen fluoride

The HF molecule serves as a second example of MBPT/CC comparisons with full CI within a DZP basis set. These results are in table 2. Unlike BH, HF exhibits good convergence. In BH there is a 2.4 kcal/mole energy correction for higher than fourth-order terms obtained by comparing CCSD to its fourth-order approximation, SDQ MBPT(4), while in HF, this difference is only 0.3 kcal/mole, but with an opposite sign. The triple-excitation contribution is 1.8 kcal/mole in HF compared to 0.4 for BH, causing SDTQ MBPT(4) to be lower than the full CI in this example.

Table 1
Correlation corrections for boron hydride (DZP basis set; R = 2.329 bohr) [a]

	BH (core + valence) correlation)	Δ (result − full CI) (kcal/mole)	BH (valence correlation, au) [c]	Δ (result − full CI) (kcal/mole)	B(au) (core + valence)	D_e (kcal/mole) [e]
SCF	25.125260	64.2	25.125360	55.1	24.529577 (UHF)[d] 24.526781 (RHF)[d]	61.5
D MBPT(2)	25.198988	18.0	25.185557	17.2	24.579623	76.4
D MBPT(3)	25.216566	6.94	25.202040	6.89	24.593577	78.7
D MBPT(4)	25.222817	3.02	25.208222	3.01	24.599104	79.1
SD MBPT(4)	25.223125	2.83	25.208520	2.83	24.599309	79.2
SDQ MBPT(4)	25.221940	3.57	25.207381	3.54	24.598446	79.0
SDTQ MBPT(4)	25.222567	3.18	25.207964	3.18	24.598781	79.1
CCSD	25.225834	1.13	25.211367	1.04	24.601545	79.5
CCSD + T(4)	25.226461	0.73	25.211949	0.67	24.601870	79.6
CISD [b]	25.222417	3.27	25.208933	2.57	24.600618 [d]	77.9 [d]
CISDTQ [b]	25.227600	0.02	25.213024	0.00	–	–
full CI [b]	25.227627	0.00	25.213024	0.00	24.602641	79.9

a) All results for total energies are in hartree and are negative.
b) CI results are taken from ref. [2].
c) Valence correlation energy is defined by freezing the B(1s) electrons and dropping the corresponding virtual orbital.
d) CI results are based upon a RHF solution for B(^{2}P).
e) MBPT/CC results for D_e are based upon an UHF solution for B(^{2}P). E(H) = −0.497639 au.

Table 2
Correlation corrections for hydrogen fluoride (DZP basis; R = 1.732 bohr)

	−Energy (au)	Δ (result − full CI) (kcal/mole)	% valence correlation energy [a]
SCF	100.048009	125.3	0.0
D MBPT(2)	100.240133	4.7	96.2
D MBPT(3)	100.242166	3.5	97.2
D MBPT(4)	100.244665	1.9	98.5
SD MBPT(4)	100.246181	0.9	99.3
SDQ MBPT(4)	100.245247	1.5	98.8
SDTQ MBPT(4)	100.248129	−0.3	100.2
CCSD	100.244833	1.8	98.6
CCSD + T(4)	100.247715	−0.02	100.02
CISD [b]	100.238526	5.7	95.4
CISDT [b]	100.240685	4.4	96.5
CISDTQ [b]	100.247410	0.2	99.9
full CI [b]	100.247684	0.0	100.0

[a] Valence correlation energy is defined by freezing the F(1s) electrons at the SCF level and dropping the corresponding *virtual orbital*.
[b] CI results are from ref. [2].

Table 3
Correlation corrections for ammonia (DZ basis; R = 1.91165 bohr, ϕ = 106.7°) [a]

	−Energy (au)	Δ (result − full CI) (kcal/mole)	% valence correlation energy [a]
SCF	56.165931	79.5	00.0
D MBPT(2)	56.277352	9.6	88.0
D MBPT(3)	56.285281	4.6	94.2
D MBPT(4)	56.288908	2.3	97.1
SD MBPT(4)	56.289359	2.0	97.4
SDQ MBPT(4)	56.289584	1.9	97.6
SDTQ MBPT(4)	56.290692	1.2	98.5
CCSD	56.290904	1.1	98.7
CCSD + T(4)	56.292011	0.4	99.5
CISD [b]	56.285574	4.4	94.4
CISDT [b]	56.286825	3.6	95.4
CISDTQ [b]	56.292405	0.1	99.8
full CI [b]	56.292612	0.0	100.0

[a] Valence correlation energy is defined be freezing the N(1s) electrons at the SCF level and dropping the corresponding *virtual orbital*.
[b] CI results are from ref. [2].

CCSD is slightly higher than its fourth-order, SDQ MBPT(4) approximation, reflecting the residual positive character of the non-linear contributions of T_1 to the CCSD result [4].

By comparing CISD and CISDT with CISDTQ, the contributions of CI triple excitations are 1.4 kcal/mole, and quadruple excitations, 4.2 kcal/mole. The triples number compares well with the fourth-order MBPT values of 1.8 kcal/mole. To obtain the fourth-order MBPT value for CI quadruple excitations, it is necessary to add the quadruple diagram values E_{4Q} and the value due to the unlinked quadrupole diagrams, which is $E_2\Delta$, where $\Delta = \langle\phi_1|\phi_1\rangle$ is the overlap of the first-order perturbed wavefunction. This value is 4.6, again in good agreement with the CI quadruple estimate. As usual, the unlinked portion of the quadruples, which is automatically eliminated from the linked-diagram expansion of MBPT/CC, is the dominant portion, in this case being 113% of the value. This should also apply to higher categories of excitations, and this part (the unlinked part) of the hextuple, octuple, and higher excitations is automatically eliminated from the linked-diagram expansion. Harrison and Handy's statement that hextuples would be difficult to include in MBPT/CC should be qualified to read that the hextuples from T_2^3 or any hextuple combination of T_1 and T_2 are included in CCSD. Furthermore, all unlinked hextuples that would contribute to CI hextuples, which are probably the dominant part, are appropriately eliminated, via the linked-diagram theorem. The remaining hextuples from T_3^2, $T_1T_2T_3$, T_2T_4, T_6, etc. would, indeed, be very difficult to include.

5. Ammonia

The NH_3 results of Harrison and Handy employ an unscaled, double-zeta basis set. In MBPT/CC there is usually better convergence for larger basis sets, since once polarization functions are included, there is a better representation of low-lying virtual orbitals. Previous work has compared the convergence of MBPT/CC in different basis sets [7]. In comparing NH_3(DZ) with HF(DZP) (see table 3) it is clear there is a larger residual error at second-order and between

the fourth-order SDQ MBPT(4) and CCSD results. Again CI triples amount to 0.8 kcal/mole and compare well with 0.7 from E_{4T}, while CI quadruples at 2.5 kcal/mole are well approximated by $E_{4Q} + E_2\Delta$ = 3.0. Only in very small basis sets is E_{4Q} computed to be negative [7], as in this example. Again the unlinked diagram part, $E_2\Delta$ amounts to 95% of the total CI quadruple excitation contribution.

6. Water

The most interesting example in this study is the H_2O molecule in a DZ basis set, which has been subjected to a symmetric OH stretch to 1.5R and 2.0R, where R is the equilibrium bond distance (see table 4). Most previous comparisons of full CI with MBPT/CC have been for geometries near equilibrium, where a single RHF reference function is a good approximation to the correct wavefunction. For example, at equilibrium, the RHF function has a coefficient of C_0 = 0.975 in the full CI, while at 1.5R, C_0 = 0.923, and at 2.0R, C_0 = 0.764. Consequently, any single-reference-function procedure suffers from a much poorer starting point. However, with some reservations, even very difficult cases can often be described by CCSD [23].

Comparing DZ NH_3 with DZ H_2O at equilibrium, by third order the convergence is very similar. Comparison with CI triples gives 0.85 kcal/mole from E_{4T} and 0.72 from CI. Similarly, quadruples from CI are 4.0 kcal/mole and $E_{4Q} + E_2\Delta$ = 4.1, of which the linked part is only 0.5, with 88% from $E_2\Delta$. The higher than fourth-order corrections are only 0.3 kcal/mole.

Once we consider displaced geometries, the convergence is much worse. The infinite-order corrections at 1.5R are 2.8 kcal/mole, closer to that observed in BH, increasing to 12 kcal/mole at 2.0R. The CI triples vary from 0.7 at R to 2.3 at 1.5R and 6.7 at 2.0R. These may be compared to the corresponding E_{4T} numbers of 0.9, 2.5 and 7.6 kcal/mole which show some overestimate in fourth order as in HF. E_{4T} is sufficiently large at 2.0R to cause the straightforward addition of E_{4T} to CCSD to exceed the full CI result. Presumably, higher-order contributions of triples, either introduced directly or via allowing T_3 to affect the T_1 and T_2 amplitudes, would reduce the net effect of triples to be more in line with the CI value. The quadruple excitations have an even more dramatic dependence on geometry, varying from 4.0 to 11.0 to 28.5 kcal/mole in CI compared to 4.1, 10.6, and 25.8 as measured in fourth order.

At large displacements, unrestricted Hartree–Fock

Table 4
Correlation corrections for H_2O as a function of symmetric bond displacement (DZ basis set; R = 1.84345 bohr, θ = 110.5°) [a]

	R		1.5R		2.0R (RHF)		2.0R (UHF)	
	−energy (au)	Δ (result − full CI) (kcal/mole)	−energy (au)	Δ (result − full CI) (kcal/mole)	−energy	Δ (result − full CI) (kcal/mole)	−energy	Δ (result − full CI) (kcal/mole)
SCF	76.009836	92.9	75.803529	132.0	75.595180	194.6	75.801918	64.8
D MBPT(2)	76.149315	5.4	75.994577	12.5	75.852461	33.1	75.866403	24.4
D MBPT(3)	76.150707	4.5	75.989393	15.8	75.834803	44.2	75.872822	20.3
D MBPT(4)	76.153790	2.6	76.000665	8.7	75.871252	21.3	75.874345	19.4
SD MBPT(4)	76.154698	2.0	76.004625	6.2	75.885014	12.7	75.875826	18.5
SDQ MBPT(4)	76.155513	1.5	76.004468	6.3	75.876785	17.9	75.875853	18.4
SDTQ MBPT(4)	76.156876	0.6	76.008395	3.8	75.888867	10.3	75.876654	17.9
CCSD	76.156076	1.1	76.008931	3.5	75.895913	5.9	75.892927	7.7
CCSD + T(4)	76.157439	0.3	76.012858	1.0	75.907997	−1.7	75.893727	7.3
CISD [b]	76.150015	4.9	75.992140	14.0	75.844817	37.9		
CISDT [b]	76.151156	4.2	75.995843	11.7	75.855527	31.2		
CISDTQ [b]	76.157626	0.2	76.013418	0.7	75.900896	2.7		
full CI [b]	76.157866	0.0	76.014521	0.0	75.905247	0.0		

[a] All total energies are in hartree and are negative. [b] CI results are from ref. [2].

(UHF) results might be preferred, in order to enable the molecule to separate correctly into $O(^3P)$ plus $2H(^2S)$ atoms. However, in any UHF calculation, one has to be careful to obtain the correct UHF solution of the plethora of possibilities. Although the UHF results reverses the trend that the RHF energy is progressively farther from the correct answer due to its inability to separate correctly, the convergence toward the correct answer is not very good. This is largely due to the fact that only one of the several almost orthogonal and equivalent UHF configurations is used as a reference function for H_2O, introducing the effects of the remaining equivalent configurations from the complementary space painfully slowly. After an excessive number of CCSD iterations the CCSD + T(4) result gains 11.8 kcal/mole over the fourth-order value, to give 89% of the full CI correlation energy result. Unlike the RHF case, the fourth-order triplet excitations are only 0.5 kcal/mole and quadruples are 0.7.

Using UHF to correctly separate a singlet molecule into open-shell atoms is probably the worst possible application of the method. In true open-shell systems, the UHF + CCSD multiplicity is nearly always a pure doublet, triplet, etc., to ≈4 decimal places, as in the $B(^2P)$ atom discussed above. However, for a singlet case it is very difficult to eliminate higher contaminating multiplicities. Straightforward spin annihilation of the reference function can be introduced, but this introduces some formal questions about how spin annihilation is preserved throughout the iterations of the CCSD equations. Alternatively, spin-adapted cluster methods as developed by Nakatsuji [24] could be appropriate. However, the ultimate resolution of bond breaking problems would seem to require a multi-reference function MBPT/CC approach [25].

To summarize, CCSD + T(4) provides 99.7, 99.2 and 101.2% of the full CI correlation energy as a function of displacement in H_2O. In each case, the variational CISDTQ gives, respectively, 99.8, 99.5 and 99.0%. In both cases, the error is non-uniform, oscillating to within 2 kcal/mole for CCSD + T(4) and 3 kcal/mole for CISDTQ. Multi-reference-function CI, MBPT or CC techniques potentially offer smoother behavior as a function of displacement. In most examples, the fourth-order triple- and quadruple-excitation effects overestimate those measured by CI. This suggests that higher-order coupling involving triples and quadruples would tend to slightly reduce the net effect of such terms.

Acknowledgement

We would like to thank Professor Handy and Dr. Harrison for providing us with their full CI results prior to publication.

References

[1] P. Saxe, H.F. Schaefer III and N.C. Handy, Chem. Phys. Letters 79 (1981) 202.
[2] R.J. Harrison and N.C. Handy, Chem. Phys. Letters 95 (1983) 386.
[3] R.J. Bartlett, Ann. Rev. Phys. Chem. 32 (1981) 359.
[4] G.D. Purvis and R.J. Bartlett, J. Chem. Phys. 76 (1982) 1910.
[5] W.D. Laidig, G.D. Purvis and R.J. Bartlett, Intern. J. Quantum Chem. 16S (1982) 561.
[6] R.J. Bartlett and G.D. Purvis, Intern. J. Quantum. Chem 14 (1978) 561.
[7] R.J. Bartlett and G.D. Purvis, Physica Scripta 21 (1980) 255.
[8] I. Shavitt, in: Modern theoretical chemistry, Vol. 3, ed. H.F. Schaefer III (Plenum Press, New York, 1977) p. 189.
[9] H.F. Schaefer III, Electronic structure of atoms and molecules: A survey of rigorous quantum mechanical results (Addison-Wesley, Reading, 1972).
[10] K.A. Brueckner, Phys. Rev. 97 (1955) 1353; 100 (1955) 36;
J. Goldstone, Proc. Roy. Soc. A239 (1957) 267.
[11] H.P. Kelly, Advan. Chem. Phys. 14 (1969) 129.
[12] F. Coester, Nucl. Phys. 1 (1958) 421;
F. Coester and H. Kummel, Nucl. Phys. 17 (1960) 477.
[13] J. Čížek, J. Chem. Phys. 45 (1969) 4256; Advan. Chem. Phys. 14 (1969) 35.
[14] R.J. Bartlett, I. Shavitt and G.D. Purvis, J. Chem. Phys. 71 (1979) 281.
[15] B.O. Roos and P.E.M. Siegbahn, in: Modern theoretical chemistry, Vol. 3, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
[16] P.E.M. Siegbahn, J. Almloft, A. Heiberg and B.O. Roos, J. Chem. Phys. 74 (1981) 2384.
[17] H. Lischka, R. Shepard, F.B. Brown and I. Shavitt, Intern. J. Quantum Chem. 15S (1981) 91.
[18] M.J. Frisch, R. Krishnan and J.A. Pople, Chem. Phys. Letters 75 (1980) 66.
[19] I. Lindgren, Intern. J. Quantum.Chem. 12S (1978) 33.
[20] R.J. Bartlett and D.M. Silver, J. Chem. Phys. 62 (1975) 3258.
[21] R.J. Bartlett and G.D. Purvis, to be published.
[22] G.D. Purvis, H. Sekino and R.J. Bartlett, to be published.
[23] G.D. Purvis, R. Shepard, F.B. Brown and R.J. Bartlett, Intern. J. Quantum Chem. 23 (1983) 835.
[24] H. Nakatsuji, Chem Phys. Letters 67 (1979) 329.
[25] W. Laidig and R.J. Bartlett, to be published.


D
FI
6