INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

73-21,075

HARIHARAN, Puthugramam Chidambaram, 1940-
A STUDY OF THE USE OF POLARIZATION FUNCTIONS IN MOLECULAR ORBITAL CALCULATIONS ON ORGANIC COMPOUNDS.

Carnegie-Mellon University, Ph.D., 1973
Chemistry, physical

University Microfilms, A XEROX Company, Ann Arbor, Michigan

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Carnegie-Mellon University

CARNEGIE INSTITUTE OF TECHNOLOGY

AND

MELLON INSTITUTE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

TITLE A STUDY OF THE USE OF POLARIZATION FUNCTIONS IN MOLECULAR ORBITAL CALCULATIONS ON ORGANIC COMPOUNDS

PRESENTED BY P. C. HARIHARAN

ACCEPTED BY THE DEPARTMENT OF CHEMISTRY

John A Pope — MAJOR PROFESSOR — March 16, 1973 — DATE

Albert A. Caretto, Jr. — DEPARTMENT HEAD — March 16, 1973 — DATE

APPROVED BY THE COLLEGE COUNCIL

C Omomar — DEAN — April 8, 1973 — DATE

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ABSTRACT

This thesis presents a systematic study of the influence of polarization functions on (i) the energies of reactions of polyatomic molecules containing the atoms C, F, H, N and O; (ii) the conformations of simple carbocations; and (iii) the geometries of the hydrides of C, F, N and O (containing only one heavy atom). The split-valence 6-31G basis developed by Hehre, Ditchfield and Pople was found to furnish an adequate *sp* basis to which the polarization functions were added. Comparison with similar work reported in the literature reveals that the exponents of the polarization functions do not vary drastically in the atomic series considered. The same standard value could be assigned to the *d* exponent for C, N, O and F and, similarly, a single *p* exponent for hydrogen in all hydrides. Two polarization basis sets, with the standard polarization exponents, were thus developed: 6-31G* containing *d* functions on the first row atom, and 6-31G** which additionally has *p* functions on hydrogen. Consideration of the energies of reactions as the basis sets were systematically improved has enabled the tentative prediction of (i) hydrogenation energy errors at the Hartree-Fock limit, and (ii) the magnitudes of correlation energies of double and triple bonds compared to that of a single bond. The study of the conformations of simple carbocations has shown that polarization functions are important in non-classical structures involving three-center bonds. Finally, investigation of the geometries of hydrides has revealed that polarization functions are essential for proper calculation of bond angles.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE OF CONTENTS

		Page
	List of Tables	ii
	Acknowledgments	iv
I.	Introduction	1
II.	Systematic Use of Polarization Functions	11
III.	Development of the Polarization Basis Set	19
IV.	Systematic Approach to Calculating One- and Two-Electron Integrals in Gaussian Basis Sets	28
V.	The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies	41
VI.	The Effect of Polarization Functions on Conformations of Simple Carbonium Ions	59
VII.	The Effect of *d* Functions on the Geometries of the First Row Hydrides	69
VIII.	Summary and Conclusions	89
	Appendix. Explanation of Some Terms Used in LCAO-MO Theory	91
	References	94

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

LIST OF TABLES

Table No.	Title	Page
I.	Total Energies (hartrees) and Optimum d-Type Exponents α_d	24
II.	Energies of Some Representative Molecules with Five and Six d Functions	25
III.	Some Values of d Exponents for First-Row Atoms	26
IV.	Some Values of α_p for Hydrogen	27
V.	Total Energies (hartrees) with Standard Polarization Functions and Estimated Hartree-Fock Limits	52
VI.	Complete Hydrogenation Energies ΔE (kcal/mole)	54
VII.	Energy Lowerings per Atom (kcal/mole) Due to Addition of d-Functions	55
VIII.	Energy Lowerings per Hydrogen Atom (kcal/mole) Due to Addition of p-Functions	56
IX.	Energies of Multiple Bond Separation Reactions (kcal/mole)	57
X.	Hydrogenation Energy Errors $\Delta E_{th} - \Delta E_{exp}$ with the 6-31G** Basis (kcal/mole)	58
XI.	Estimated Hydrogenation Energy Errors $\Delta E_{th} - \Delta E_{exp}$ at the Hartree-Fock Limit (kcal/mole)	58
XII.	Systematic Approach to the Hartree-Fock Limit	66
XIII.	Total Energies of Some Simple Carbocations	67
XIV.	Energy Differences	68
XV.	Optimized Energies of First-Row Hydrides at Various Levels	84
XVI.	Optimized Geometries of First-Row Hydrides at Various Levels	86
XVII.	Geometry of H_2O as the Hartree-Fock Limit is Approached	88

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ACKNOWLEDGMENTS

A major part of the work reported in this thesis rests upon the firm foundation laid by my colleagues, the 'Pople People', and others. I am grateful to all of them for their help, both active and passive. In particular, I wish to thank: Warren Hehre and Robert Ditchfield, who labored hard to develop the constrained basis sets; Robert Ditchfield for generous advice, useful discussions and test output from POLYATOM; Larry Curtiss for help in optimizing the geometries of CH, OH and NH; John Bentley for taking the time to listen to and react to my arguments; David Rodgers, whose expertise both as engineer and programmer, coupled with his ready willingness to listen to and resolve my problems with the UNIVAC 1108, tremendously lightened my programming burden; Professor R.F. Stewart for many illuminating and instructive discussions.

I owe a lot to Professor J.A. Pople, who provided the inspiration for this work. I am grateful to him for his patient instruction and skillful guidance and for his constant encouragement.

Mrs Kathryn Severn expertly deciphered my illegible scrawl into the finished typescript without the benefit of my interpretation. I am thankful to her for this service and for numerous other acts of kindness through the years of my stay in the group.

Larry Curtiss and John Bentley have also helped with their criticism and proofreading. I must also acknowledge my gratefulness

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to: Gil Arnold and Mrs Laura Beach for the figures; Leland Bard for miscellaneous assistance.

P.C. Hariharan

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

I. INTRODUCTION

Atomic functions of higher angular quantum number than the orbitals occupied in the atomic ground state are important in molecular orbital (MO) expansions. This fact has been recognized since the very early days of quantum chemistry. Wang[1] was able to improve the Heitler-London wave function for the hydrogen molecule, of the form

$$\Psi(\vec{r}_1,\vec{r}_2) = c\,[a(\vec{r}_1)b(\vec{r}_2) + b(\vec{r}_1)\,a(\vec{r}_2)] \tag{1.1}$$

where a and b are hydrogenic $1s$ wave functions, by introducing a scale factor and adjusting it to minimize the energy. Rosen[2] obtained a better wave function by setting

$$\begin{aligned} a(\vec{r}) &= c_1\, e^{-\delta r_a} + c_2 Z_a\, e^{-\delta r_a} \\ b(\vec{r}) &= c_1\, e^{-\delta r_b} - c_2 Z_b\, e^{-\delta r_b} \end{aligned} \tag{1.2}$$

In (1.2) the second term is a function of $2p$ type symmetry; however, it shares the same exponent as the $1s$ type function. The $2p$ exponent in the free hydrogen atom is only half as large as the $1s$ exponent and so the $2p$ type function in (1.2) is an artificially contracted one. The optimum value of δ found by Wang was 1.166, and this increased slightly to 1.19 when Rosen added the $2p$ type function. With the addition of the p-type function, the energy of the wave function (1.1) fell by 0.0095 hartrees. The H-H bond length found by Rosen disappointingly increased to $1.416a_o$ from $1.406a_o$ for Wang's wave function; the experimental value is $1.4016a_o$.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Equations (1.2), in which s and p type functions are mixed at an atomic center, represent mathematically the simplest form of hybridization, a concept familiar to all chemists. The term hybridization is usually reserved, however, for cases where the functions of higher angular quantum number either have the same principal quantum number, or a smaller principal quantum number, than the function of lowest angular quantum number with which they are mixed. Thus the tetrahedral hybrids postulated in CH_4 involve mixing of the carbon $2s$ and $2p$ orbitals. $[Cr(H_2O)_6]^{2+}$ and $[Cr(CN)_6]^{4-}$ exemplify outer- and inner-orbital complexes respectively; the former, high-spin complex is considered to involve hybridization of $4d$ orbitals with $4s$ and $4p$ and the latter, low-spin complex typifies hybridization of $3d$, $4s$ and $4p$ orbitals of Cr. In Rosen's wave function, on the other hand, a $2p$ orbital on hydrogen is mixed with a $1s$ orbital. The term *polarization* is used to describe this mixing. The distorted wave function that describes a hydrogen atom in its ground state when subjected to a uniform electric field is similar to equation (1.2) and it is this similarity which has led to the use of the term polarization. The atomic orbitals (AO's) of a free atom, being the eigenfunctions of a spherically symmetric Hamiltonian, can be classified as s, p, d, ... etc. When a molecule is formed which destroys this symmetry the distortion of the AO's may be described by addition of p, d, f, ... to the s, p, d, ... orbitals.

Wang's results for the hydrogen molecule show an important effect associated with the formation of the molecule. The exponent of the $1s$ orbital has increased from the free-atom value of 1.0 to 1.166,

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

indicating that the orbital has contracted in the molecule. It is obvious that this change in *size* of the atomic orbital is accompanied by a change in *shape* also, as indicated by the improvement obtained by Rosen.

Similar results have been obtained for the simplest molecular system, H_2^+. Following Rosen's work Dickinson[3] added p-type polarization functions to the wave function of the form determined by Finkelstein and Horowitz:[4]

$$\psi = 1S_A(\alpha) + 1S_B(\alpha) + c\{2P_{Z_A}(\beta) + 2P_{Z_B}(\beta)\} \qquad (1.3)$$

here $1S_A(\alpha)$, $2P_{Z_A}(\beta)$ represent hydrogen-like $1s$ and $2p_Z$ functions on center A with exponents α and β; the Z-axis is taken to be the internuclear axis. The optimum values of α and β were found to be 1.247 and 1.434 at $r_{AB} = 2.003\ a_o$. Finkelstein and Horowitz had obtained $\alpha = 1.228$ and $r_{AB} = 2.003a_o$. Their energy of -15.78 eV lay 0.48 eV above that found by Dickinson who had accounted for all but 0.05 eV of the total energy of this system.

In the case of H_2^+ it is not difficult to form a physical picture of the situation as improvements are systematically made in the wave function. The increase in the value of the $1s$ orbital exponent from that for the free hydrogen atom shows that the orbitals have contracted around the atoms in the molecule. The region of largest electronic density is that between the two nuclei since this is where both AO's simultaneously have appreciable value. The p-type polarization function further enhances this buildup of electronic density in the internuclear region by providing greater directional character to the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

linear combination of atomic orbitals (LCAO), equation (1.3), describing the wave function. At the same time, because of the reduced size of the 1s orbital, less of the electronic density is spread in regions off the bond axis.

Another simple way to increase the electronic density in the internuclear region, while continuing to use 1s type functions for the hydrogen molecule, is to displace the centers of these functions towards each other as shown in Figure 2. This suggestion was put forward

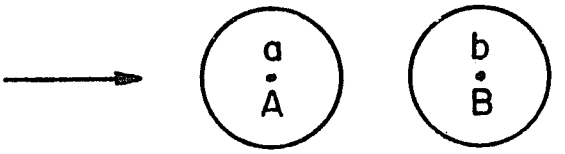


Fig. 1

A and B represent the two hydrogen nuclei; the 1s functions are centered at a and b

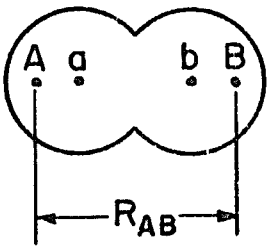


Fig. 2

The 1s functions have drifted or floated towards each other, withdrawing density from non-bonding regions and increasing the density in the bonding regions. Nuclear centers A and B no longer coincide with the centers of the 1s functions a and b.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

by Gurnee and Magee[5] in 1950 and was applied to H_2^+ and to H_2 by Hurley.[6] He displaced 1s orbitals (with $\alpha = 1.23$) 0.061Å along the molecular axis and obtained a marginally better energy than that of Dickinson at $r_{AB} = 1.05$Å. It must be remembered that Hurley's wave function had just two adjustable parameters (the 1s orbital exponent and the displacement of the 1s orbital from the nucleus) as compared with Dickinson's three-parameter wave function. For H_2 itself Gurnee and Magee had found it necessary to vary both the exponent of the 1s orbital and its displacement in order to obtain a wave function with a lower energy than that of Rosen. This is because increasing the value of the orbital exponent for a fixed internuclear separation also has the effect of decreasing the ratio of the average potential energy to the average kinetic energy of the electrons.[7]

Hurley's aim was to set up a simple scheme for calculating binding energies of molecules from electrostatic considerations. When a completely general method for all molecules could not be set up he began to investigate the conditions under which an approximate wave function could simultaneously satisfy both the virial theorem and the Hellmann-Feynman theorem. It was then that he discovered that even the Gurnee-Magee wave function had a hidden variable which had not been optimized: the internuclear distance. According to Hurley 'most calculations violate the generalized electrostatic theorem by the inclusion in the wave function of variable parameters whose values are not chosen by the variational principle...In nearly all of the wave function calculations, the approximate wave functions have been built up from AO's rigidly attached to one of the nuclei of the molecule. In order

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

to construct approximate wave functions which do satisfy the conditions, it is necessary to detach the atomic functions from the nuclei and allow them to take up optimum positions determined by the variational principle. This feature of the approximate wave functions which we shall use has been emphasized by referring to them as floating functions. A "floating function" is defined as a wave function in which *all* variational parameters are determined by the variational principle.'

The method of floating orbitals has some disadvantages, most notably that even in a diatomic molecule they give rise to three-center integrals. For this reason gaussian type orbitals (GTO's) have been preferred for such calculations.

The formation of a molecule induces a two-fold polarization of the valence AO's of the constituent atoms: *radial polarization* which can be taken into account by varying the exponent of the orbital, and *angular polarization* which can be allowed for either by inclusion of functions of higher angular quantum number, or by letting the functions float free from the nuclear centers, or both.

The use of polarization functions in MO calculations has caused considerable confusion in the minds of many. Even to those used to the concept of LCAO the difference between the terms polarization and hybridization is not always clear. The LCAO-MO method is after all just one convenient way of obtaining an approximate molecular wave function. The addition of p functions centered on hydrogen to describe an MO in a hydride is a mathematical artifice to improve the quality of the MO and thus that of the total wave function. As pointed out earlier, the electronic density due to a bonding MO in a diatomic molecule is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

greatest in the internuclear region. One way of accomplishing this was illustrated in Figure 2. However, the addition of a small amount of the p function of appropriate symmetry (e.g., p_Z in the hydrogen molecule if the molecular axis were the Z-axis) achieves the same, as shown in Figures 3 and 4:

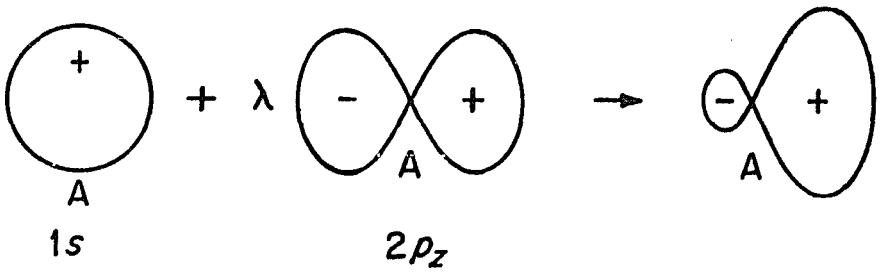


Fig. 3

Addition of a small amount of $2p_z$ to $1s$ makes the resulting function more directional.

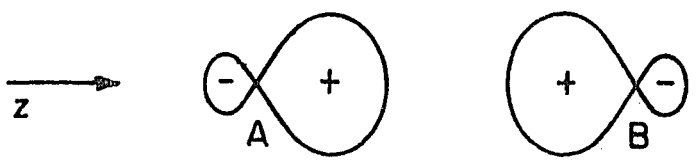


Fig. 4

The MO obtained by superposing such directed functions also enhances the electronic density in the internuclear region.

One need not refer to this situation as 'p orbital participation' in bonding, any more than referring to the situation depicted in Figure 2 as 's orbital drifting'.

In a similar way, a $p\pi$ orbital can be bent into bonding regions, or away from them, by addition of small amounts of $d\pi$ orbitals of the right symmetry.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

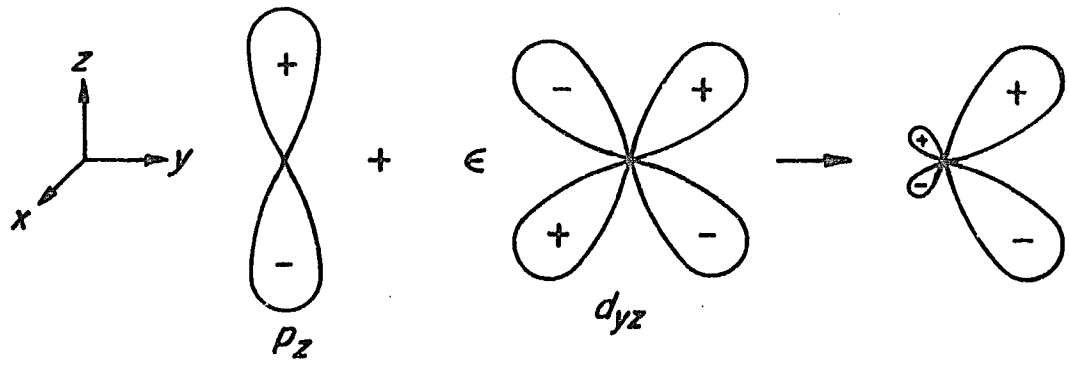


Fig. 5
Addition of a small amount of d_{yz} to $2p_z$ helps to 'bend' the $2p_z$ function towards or away from a neighboring center.

Thus, in ethylene, a better description of the π bond can be achieved by addition of a small amount of a polarizing d orbital to the p functions forming the π bond. Again, this does not mean that carbon d functions are involved in this bond. The functions which make up an atomic basis are rather rigidly anchored to their atomic centers. The two degrees of freedom normally available to them in a self-consistent field (SCF) LCAO-MO calculation, namely, exponent optimization and variation of the LCAO coefficient, do not permit sufficient relaxation of the kind shown in Figures 2, 4 and 5. The 'drifting' and 'bending' tendencies can be satisfied by the two methods mentioned above. In ring systems having small bond angles, e.g. cyclopropane, it is known[8] that basis sets which do not contain p functions on hydrogen, or d functions on heavy atoms, overestimate the strain. Bending is thus important not only in describing multiple bonding, but also in giving an adequate account of bonding in ring systems.

The majority of the calculations in the literature, like the ones in this thesis, take polarization into account by including higher angular momentum functions in the atomic basis. Since the alternative way is not so well-known to practicing chemists, attempts have been made to assess the extent of 'd participation'

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

in compounds of the first row and second row elements. This problem received considerable attention at an international symposium on valence and reactivity held in Oxford under the auspices of the Chemical Society on January 9-11, 1968.[9] Coulson[10] has given a readable discussion of this confusing issue. His main conclusion is that if the *d*-orbital population on a first- or second-row atom is small, then the situation should be described as polarization. In such a case, the population will change perceptibly when the basis set is changed. There are, however, cases where an atom, in its ground state, has occupied *d* orbitals, as in the transition metals. These can undeniably participate in bonding. Coulson believes that only in this latter case 'is there any really significant chemical meaning in talking about the role of *d* electrons in bonding'. Mulliken[11] is against this attempt to distinguish polarization and chemical participation. He finds it 'to be scarcely possible to draw a line anywhere between small and large *d* participation. Basically, the question can be dismissed as trivial semantics; what is really involved is just the mathematical description of the forms of the MO's, which in themselves are merely convenient expressions for use in one way of building up wave functions of molecules. Still, MO's and AO's for atoms, are convenient concepts. Granting this, and noting that there is no *d* participation (except through configuration interaction) in AO's in *atoms* such as N, O, Cl, P and S, it can reasonably be argued that all *d* participation of these atoms *when they enter molecules* can properly be thought of as a chemical effect'.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Mulliken[12] distinguishes two types of polarization: *valence polarization* occurs in the LCAO-MO method when the AO's of atoms are polarized when a molecule is formed. The hydrogen molecule is an example. If the molecule is heteropolar, *Coulomb polarization* is superposed on this. Valence polarization is largely due to quantum mechanical valence forces whereas Coulomb polarization is produced by an electric field. He finds that the concept of polarization is in general broader, more appropriate and more helpful than that of hybridization in thinking about what happens to AO's of atoms when they form molecules.

In concluding this section we point to a simple distinction between polarization and hybridization if one were needed. The concept of hybridization was put forward by Pauling to explain the characteristic valence angles and bond lengths in compounds of C, N, and O. In CH_4, where sp^3 hybrids are postulated, there are enough electrons to fill all four hybrid orbitals. If we look at equation (1.2) as representing *sp* hybridization in the hydrogen molecule there are enough electrons to fill just half of the hybrid orbitals. This illustrates the usefulness of looking on hybridization as a special case of the more general phenomenon of polarization; both are 'chemical' in the sense that they occur when atoms come together to form molecules.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

II. SYSTEMATIC USE OF POLARIZATION FUNCTIONS

Nesbet,[13] in his study of the HF molecule, first drew attention to the need for including *d* functions in the basis sets of first row atoms to obtain improved wave functions. Polarization functions must be present in basis sets in order to obtain wave functions approaching the Hartree-Fock (HF) limit.

The first systematic study of the use of polarization functions, however, came from the Laboratory of Molecular Structure and Spectra of the University of Chicago, in a series of papers authored by Wahl,[14] Huo,[15] and by Cade, Sales and Wahl.[16] These authors used large basis sets of Slater type orbitals (STO's) in their computations. Two types of basis set synthesis were tried out by this group. The first kind, termed 'arduous' by Wahl[14] started with a minimal basis which was then gradually and arduously built up by addition of more functions of the same and different symmetries. The second type used a large starting atomic basis which was then augmented by polarization functions. Both from the point of the computation involved and the molecular energy the latter constituted the easier and better approach. One of the features of the Chicago basis which is absent in most other basis sets is that the atomic functions entering MO's of different symmetry had different exponents. This anisotropy came about partly because of extensive exponent optimization. Coupled with the large size of the basis this extra flexibility enabled these workers to calculate wave functions very close to the HF limit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In building up the basis, functions of higher angular momentum were added only when further addition of functions of existing angular momentum types produced negligible improvement in energy. Since the basis set was saturated with *s* and *p* functions, the energy lowerings obtained by these authors when *d* and *f* functions were included in the basis represent almost completely the contribution of these added functions as polarization functions. It is important to bear this in mind since in a basis that is not near-complete with respect to *s* and *p* functions the *d* functions try partly to take the place of the *s* and *p* functions. For the two homonuclear diatomic molecules N_2 and F_2 for which the complete study has been reported it is interesting that *d* functions play the major role of polarization functions. In the arduous basis, the best *sp* basis energy for N_2 is -108.9022 hartrees (1 hartree = 627.5 kcal/mole = 27.210 eV). With five sets of *d* functions (σ_g 3*d*, σ_g 3*d*', π_u 3*d*, π_u 3*d*', σ_u 3*d*) and two *f* functions (σ_g 4*f*, π_u 4*f*) the energy fell to -108.9875 hartrees. Of this drop of 85.3 millihartrees (mh), σ_g 3*d* and π_u 3*d* together account for 79.7 mh. The remaining three *d* functions are responsible for 2 mh. Thus, in N_2, about 94% of the energy lowering due to polarization comes from *d* functions. The first set of *d* functions added also shows that while the energy lowering due to σ_g 3*d* is 27.6 mh, that due to π_u 3*d* is 52.1 mh. Wahl[14] has not reported as extensively on F_2, but the *d* functions again account for the major drop in energy, 22.1 mh, while the *f* functions contribute 4.4 mh. Similarly, in CO, when $3d\sigma_c$ and $3d\sigma_o$ are added the energy drops by 27.9 mh; addition of $3d\pi_o$ and $3d\pi_c$ results in a further lowering of 44.52 mh. The only *f* type function added, $4f\pi_c$, lowers the energy by 1.57 mh. With a smaller

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

basis of double-zeta quality Nesbet[17] found that d functions decreased the energy of N_2, CO and BF by 109.7 mh, 85.52 mh and 39.09 mh respectively; inclusion of an f type function in the basis produced a further lowering of less than 4 mh for N_2.

The consensus of these studies is that for the first row atoms the principal polarization in molecular environments is provided by d functions. While the energy lowering due to d functions in the first row diatomics spans the range from $\sim$100 mh in N_2 to $\sim$25 mh in F_2, the limited number of available calculations indicate near-constancy of the lowering due to f-functions; 4.4 mh in F_2 and $\sim$4 mh in N_2.

It may be conjectured from the above that the major polarization for hydrogen in molecular environments is provided by p functions. This surmise is supported by the computations of Cade and Huo[18] on hydrides of the first-row elements.

All of the calculations discussed thus far in this section have used Slater type (ST) polarization functions in extended STO basis sets. There have been some calculations in STO minimal bases which have employed ST polarization functions. Nesbet[17] obtained initial values for the polarization exponents by optimizing them in a minimal basis. Lipscomb[19,20] et. al. have added polarization functions to minimal STO basis sets in molecules like SiH_4, PH_3, and H_2S. Mulliken and Liu[11] have cautioned that 'linear combinations of d functions on different atoms are not in general independent of similar linear combinations of s or p functions, and so can in part take the place of the latter'. The d exponents obtained in such cases do not correspond to those for pure polarization effects. For example, the 'polarization' function may try to make up the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

deficiency in the basis of a neighboring atom; in doing so, it expands its size and takes on a smaller exponent. Such diffuse functions act more as filler or substitute functions for poorly represented, or absent functions in the underlying *sp* basis. Populations derived for these diffuse orbitals may be unrealistically high and probably constitute one reason for the confusion about the role of *d* functions in bonding of first- and second-row atoms. Nesbet[17] has also cautioned against trying to add on polarization functions to a minimal basis set.

Gaussian basis sets for atomic and molecular calculations have been used in two forms. In the uncontracted form the variational principle is used to determine the contribution of each primitive to the various MO's. The number of basis functions in the calculation in such cases is equal to the total number of primitive gaussians in the basis. To reduce the time needed to set up the Fock matrix and to decrease the size of the Fock matrix the primitives have been grouped to form contracted basis functions. The 6-31G basis[24] developed in this laboratory is an example. The basis consists of 10 1*s* primitive gaussians and 4 2*p* type primitive gaussians. Six 1*s* gaussians have been contracted to represent the 1*s* atomic orbital, three 1*s* gaussians have been contracted to represent the inner part of the 2*s* atomic function and the remaining 1*s* gaussian is left uncontracted. Similarly, the 2*p* primitives have been grouped into a contracted set of three and an uncontracted one. In the process of contraction the number of basis functions is reduced from 22 (the number of primitives) to 9 (the number of contracted functions).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Dunning[21] has carried out a systematic study of the influence of polarization functions in a gaussian basis on H_2O and N_2. Roos and Siegbahn[22] found that for CH_4, NH_3 and H_2O hydrogen polarization functions were more important than heavy atom polarization functions; it must be noted that this observation has not been made by any other workers in the field. Additionally their studies showed that

(i) the curve of heavy atom polarization exponent plotted versus energy lowering due to the polarization function was fairly flat for the first row atoms. These exponents could therefore be optimized in a coarse grid

(ii) the curve of hydrogen *p* polarization exponent plotted against the energy lowering was narrower.

It is instructive to compare these findings with those of Dunning[21] who states that in H_2O, based on considerations of the energy alone, it would appear that addition of a set of 2*p* functions on the hydrogens is *nearly as satisfactory* as addition of a set of oxygen 3*d* functions. Consideration of the charge distribution, as reflected in a number of one-electron properties, shows that this is not true. From this it is clear that it is more important to add polarization functions to the oxygen basis than to the hydrogen basis, although 2*p* functions on the hydrogens are certainly required in order to obtain a wave function of near Hartree-Fock accuracy. His calculations also show that the energy surface for H_2O is quite flat with respect to changes in the hydrogen 2*p* exponents.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Having obtained the exponents for the heavy atom polarization functions in H_2O and H_2S, Roos and Siegbahn assumed that the ratio of the radius of the charge density maximum of the *d* function to the corresponding maximum of the valence *p* function was a constant for each row of the Periodic Table. Using this assumption they determined the ratio for the first and second rows from their optimizations on H_2O and H_2S. Since they had *sp* basis sets for B, C, N and F also, Roos and Siegbahn were able to assign values for the *d* exponent for these atoms without further effort.

Dunning[21] has carried out a characteristically thorough study of the influence of polarization functions on H_2O and N_2. His investigation included the following points:

(i) the efficacy of a gaussian type (GT) polarization function as compared with that of an ST polarization function

(ii) the use of a single set of GT or ST polarization functions compared with the use of more than one set

(iii) the effect of asymmetric exponent optimization compared with isotropic exponent optimization.

His results may be summarized as follows: a single set of 3*d* Slater functions is more effective for representing molecular polarization effects than is a primitive set of 3*d* gaussian functions. For H_2O the gaussian-fitted Slater polarization function lowered the energy by an additional 6% over the primitive gaussian polarization function. A lowering of the same order was obtained by using two sets of uncontracted primitive gaussians. The improvement obtained by employing three primitive sets was marginal. In the absence of polarization functions

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

on hydrogen, therefore, two sets of primitive *d* gaussians on oxygen sufficed. When polarization functions were added to the hydrogen basis the energy difference between the one and contracted two term expansions was nearly the same as before, but that between the contracted and uncontracted two-term expansions fell markedly from 2.3 mh to 0.9 mh. Dunning has pointed out the reasons for this. 'In the absence of hydrogen polarization functions, the oxygen polarization functions will in a limited way be used to describe the polarization of the electronic density in the vicinity of the hydrogens. Since two 3*d* functions on oxygen are substantially more effective in this respect an excessive lowering occurs with the uncontracted set. Increasing the length of the (contracted) expansion of a polarization function is an internal change in the form of the function and is only slightly affected by the presence or absence of functions on other centers. On the other hand, uncontracting the expansion is an external change and will be influenced by the basis on neighboring centers.' The *d* exponent on oxygen increased slightly when it was optimized after polarization *p* functions were placed on hydrogen; this reflects the decreased need for the *d* functions to describe polarization effects near the hydrogen atoms.

These results were corroborated by similar computations on the N_2 molecule. The high symmetry of the N_2 molecule made possible a second test: anisotropy of the polarization functions entering the σ and π molecular orbitals. This asymmetric optimization showed that the energy surface for σ polarization had a much smaller curvature than that for the π polarization part. The additional flexibility of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

the anisotropic basis resulted in an energy lowering approximately half that obtained by uncontracting and symmetrically optimizing. Asymmetric optimization of the hydrogen polarization exponent resulted in only marginal improvement in H_2O. Since Dunning experimented with both gaussian and gaussian-fitted ST polarization functions he obtained additional information not available from other calculations reported in the literature. Thus, the effective Slater exponent for a single primitive gaussian is larger than that for the two-term expansion of the Slater orbital. This provides for better match of the Slater and gaussian function for small values of r. The effective Slater exponent for a contracted two-term expansion is larger than that for the uncontracted expansion.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

III. DEVELOPMENT OF THE POLARIZATION BASIS SET

In the SCF approximation the polarization functions do not occur in the ground state configuration of the first row atoms. Thus, while these functions contribute directly to the binding in the molecule, it also necessitates that they be selected in a molecular environment. It is essential that a representative set of molecules be chosen which span the set of different bonding types and, if possible, geometries. Since our main interest is in the group of organic compounds containing O, N and F in addition to C and H, we chose to optimize the *d* exponent in the following compounds: CH_4, NH_3, H_2O and HF. As a test of the influence of the environment the exponent was also optimized in HC≡CH and N≡N. Standard model geometries, as specified by Pople and Gordon[23] were used in these tests. The 6-31G basis developed by Hehre, Ditchfield and Pople[24] provided the underlying *sp* basis. In this basis all inner-shell orbitals (1*s* orbitals for the first row atoms) are represented by a sum of six gaussian 1*s* type functions. The valence shell orbitals have inner and outer parts; the inner parts are expressed as a linear combination of three gaussians while the outer part consists of just one gaussian of the appropriate symmetry. This basis is of double-zeta type for the valence shell, but despite the representation of the inner shell by just one contracted gaussian, has been shown comparable to the double-zeta basis of Basch, Robin and Kuebler.[25] This basis may therefore be considered to be an extended basis. A single set of (uncontracted) *d* type gaussians

$$(x^2, y^2, z^2, xy, xz, yz) \times exp\ (-\alpha_d r^2) \qquad (3.1)$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

was added to the 6-31G basis and the energies of the six molecules listed in Table I were minimized with respect to α_d. Table I lists the optimum d exponents, the corresponding 6-31G* energies and, for comparison, the 6-31G energies. The results of this table suggest that α_d has some dependence on chemical environment (as with pairs NH_3, N_2 and CH_4, C_2H_2) but, on the other hand, there is little evidence of significant change along the atomic series C, N, O, F. On the basis of the optimum values given here, $\alpha_d = 0.8$ was selected as the standard value for all the heavy atoms C, N, O, F. This value has been used in all subsequent calculations. The augmented basis containing a single set of uncontracted second-degree functions with $\alpha_d = 0.8$ has been called the 6-31G* basis.[28]

This value of α_d is close to the value found by others for gaussian basis sets. Dunning gives $\alpha_d = 0.75$ for oxygen in H_2O and $\alpha_d = 0.98$ for N in N_2. Rothenberg and Schaefer[29] found $\alpha_d = 0.74$ for C in CH_4. Kari and Csizmadia[30] obtained $\alpha_d = 0.75$ for N in NH_3 and $\alpha_d = 0.7969$ for C in CH_3. Table III summarizes some of the α_d values determined by other workers.

At the time the optimizations were started by us it had been hoped that the exponent would match one of those in the inner or outer parts of the valence shell in the 6-31G basis. Had this occurred, evaluation of two-electron integrals would have been speeded up appreciably. Unfortunately, the polarization function is too diffuse to interact with the inner part and too small to participate in the outer part.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Of the six second-degree functions (3.1) only the following five combinations are pure d-type

$$(xy, xz, yz, x^2-y^2, 3z^2-r^2) \times \exp(-\alpha_d r^2) \tag{3.2}$$

while the sixth combination obtainable from (3.1)

$$(x^2 + y^2 + z^2) \times \exp(-\alpha_d r^2) \tag{3.3}$$

is a $3s$-type gaussian. Both Dunning[21] and Neumann and Moskowitz[37] used only the five d-type gaussians in their calculations while all six have been employed in the computations reported here. If the sp basis is saturated with respect to s functions then the deletion of (3.3) from the basis should have a negligible effect on the energy. For comparison the energies of CH_4, NH_3, H_2O, HF and C_2H_2 are listed with both six and five d functions in Table II. Inspection of the final column reveals that the sp basis for C is most satisfactory and that for F the least satisfactory.

The computations on diatomic molecules carried out by the Chicago group showed that with STO basis sets sp saturation required a multiple-zeta basis. The largest energy lowering in the arduously built basis occurs in going from the best minimal MO to the double-zeta basis. The fall in energy was less steep thereafter, but not insubstantial. Clementi and Popkie[26] have published calculations on H_2O in various extended basis sets with and without polarization functions. [Their selection of the basis sets does not seem to have been as systematic as that of the Chicago group. The energy of H_2O in a $(13s8p/7s)$ basis is 0.21 mh above the energy in a $(13s8p/6s)$ basis. Since the definitive paper on these basis sets has not been published yet we refrain from further comment.] Their best sp energy, in a $(13s8p/6s)$ basis is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

-76.02573 hartrees, while a single set of polarization functions, in a $(13s^8 8p^5 1d/6s^4 1p)$ basis, yields an energy of -76.06027 hartrees, representing a gain of 34.64 mh. Table 13 of their paper lists the energy gain due to addition of polarization functions to *sp* basis sets of various sizes.

The next step was to expand the 6-31G* basis further by including *p* functions on hydrogen and to determine the optimum values for α_p. This search was restricted to CH_4 and HF with α_d fixed at 0.8. For CH_4 the optimum value of α_p was found to be 1.2, leading to a total energy of -40.20162 hartrees. In a comparable study Rothenberg and Schaefer[29] found $\alpha_p = 1.08$. For HF the optimum value was 1.0 giving an energy of -100.01124 hartrees. The choice of an average value of $\alpha_p = 1.1$ seemed appropriate and this value was used as standard in all subsequent calculations. The 6-31G* basis augmented by *p*-type polarization (with $\alpha_p = 1.1$) on hydrogen has been called the 6-31G** basis.[28] No attempt was made to optimize α_p and α_d simultaneously for two reasons. First, Dunning[21] has already noted that the exponent tends to increase when polarization functions are present on hydrogen. The relative flatness of the energy surface with respect to α_p and α_d that has been noted both by Dunning and us suggests that further reoptimization would lead to small corrections and only marginal improvements in the energy. The choice of the standard exponents would be nullified if one wished to resort to this procedure.

Table IV displays the polarization exponent α_p determined for H by other workers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Finally, attention must be drawn to the work of Ratner and Sabin[31] who attempt to distinguish between the numerical and stereochemical need to include *d* functions in basis sets. Though they dismiss polarization as 'not of fundamental importance in the qualitative bonding picture' they do make the point that in certain point groups *s* and *p* functions may not suffice as bases for some irreducible representations; if *d* functions can provide the basis then there is a stereochemical need for their presence in a basis. This point is again considered in Section VII.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE I. Total Energies (hartrees) and Optimum *d*-Type Exponents α_d

Molecule	Optimum α_d	Energy	
		6-31G	6-31G + *d* Functions
CH_4	0.73	-40.18038	-40.19518
NH_3	0.75	-56.16320	-56.18381
H_2O	0.74	-75.98508	-76.00994
HF	0.84	-99.98348	-100.00287
N_2	0.92	-108.86762	-108.94336
C_2H_2	0.83	-76.79261	-76.81737

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE II. Energies of Some Representative Molecules with Five and Six *d* Functions (*d* exponent = 0.8)

Molecule[a]	- Energy (hartrees) 6-31G + 5*d*	6-31G + 6*d*	Energy Difference −(6*d*-5*d*) mh
CH_4	40.19472	40.19506	0.33
	0	*0*	
NH_3	56.18311	56.18374	0.63
	1.768	*1.776*	
H_2O	76.00850	76.00987	1.37
	2.131	*2.138*	
HF	100.00070	100.00281	2.11
	1.990	*1.988*	
C_2H_2	76.81684	76.81732	0.49
	0	*0*	

Values in italics are dipole moments in units of Debye.

[a]Standard model geometry; see Reference 23

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE III. Some Values of d Exponents for First-Row Atoms

Authors	Molecule	Atom	α_d	Basis[a]	Reference
Diercksen	H_2O	O	1.0	$(11s^5 7p^4 1d/6s^3 1p)$	32
Dunning	H_2O	O	0.75	$(9s^4 5p^3 1d/4s^2)$	21
	H_2O	O	0.88	$(9s^4 5p^3 1d/4s^2 1p)$	21
	N_2	N	0.98	$(9s^4 5p^3 1d)$	21
Hankins, Moskowitz & Stillinger	H_2O	O	0.897	$(10s^5 5p^3 1d/4s^2 1p)$	33
Kari & Csizmadia	NH_3	N	0.75	See reference	30
	CH_3^- (planar)	C	0.7969	$(10s6p1d/8s)$	
Roos & Siegbahn	H_2O	O	1.325	$(7s^3 3p^2 1d/4s^2 1p)$	22
		B	0.37	These were not obtained by energy optimization but through an assumption	
		C	0.63		
		N	0.95		
		F	1.62		
Rothenberg & Schaefer	CH_4	C	0.74	$(10s^4 5p^2 1d/5s^2)$	29

[a] $(as^x bp^y cd^z/a's^{x'} b'p^{y'})$ stands for a primitive $1s$ gaussians on the heavy atom contracted to x $1s$ gaussians, b primitive $2p$ gaussians contracted to y $2p$ gaussians, c primitive $3d$ gaussians contracted to z $3d$ gaussians, a' primitive $1s$ gaussians on hydrogen contracted to x' $1s$ gaussians, b' primitive $2p$ gaussians contracted to y' $2p$ gaussians. A missing superscript implies no contraction.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE IV. Some Values of α_p for Hydrogen

Authors	Molecule	α_p(H)	Basis[a]	Ref.
Diercksen	H_2O	0.75	$(11s^5 7p^4 1d/6s^3 1p)$	32
Dunning	H_2O	1.00	$(9s^4 5p^3/4s^2 1p)$	21
		1.16	$(9s^4 5p^3 2d/4s^2 1p)$	
Hankins, Moskowitz & Stillinger	H_2O	0.9650	$(10s^5 5p^3 1d/4s^2 1p)$	33
Roos & Siegbahn	H_2O	0.8	$(7s^3 3p^2 1d/4s^2 1p)$	22
	H_2S	0.8	$(9s^6 5p^4/4s^2 1p)$	
Rothenberg & Schaefer	CH_4	1.08	$(10s^4 5p^2/5s^2 1p)$	29

[a]See footnote a, Table III for explanation of notation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

IV. SYSTEMATIC APPROACH TO CALCULATING ONE- AND TWO-ELECTRON INTEGRALS IN GAUSSIAN BASIS SETS

The two preceding sections have emphasized the need to include polarization functions in the basis sets of hydrogen and the first row atoms and the growing use of gaussian atomic basis sets for MO calculations. The principal reason for preferring gaussians is the ease with which integrals needed for energy and many properties can be computed over them. Many authors[34] have pointed out that, in molecular calculations, between two and three times as many gaussians as ST functions are needed to achieve the same accuracy in energy. Unless efficient procedures are devised for calculations involving gaussians these would prove too expensive despite the exact analytic forms of the integrals.

Boys[35] was the first to point out the advantage of using gaussians as expansion functions in molecular calculations. For unnormalized 1s gaussians he derived the expressions for the energy integrals in closed form. In addition to introducing the gaussian function, Boys also pioneered the use of the electronic computer for wave function calculations and performed a number of computations on simple molecules with these functions.[36]

The most significant fact about gaussians that simplifies the evaluation of the energy integrals is that *the product of two gaussians is a third gaussian located on the line joining the first two.* This means, for example, that a two-center charge distribution can be expressed as a one-center charge distribution. Boys[35] also showed how expressions for higher functions could be derived from those for the 1s function.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Huzinaga[37] has summarized the various efforts to solve the problem of molecular integrals. For ST basis functions three- and four-center integrals involve numerical integration at some stage or other. A disadvantage of such a procedure is that the accuracy of a given kind of integral will depend not only on the angular momentum types of the functions involved, but also on the number of centers.[38] Since exact expressions for molecular energy integrals are available for gaussian functions it should be possible to calculate all integrals over such functions to an accuracy limited only by the word-size of the computer.

One of the lessons to be learned from the exact solution of the hydrogen atom is that for a given principal quantum number all the orbitals share the same common exponent; for principal quantum level n this is $1/n$. When atomic basis sets are constructed, better energies are obtained by using variationally flexible forms for the basis functions. Such flexibility can be achieved by increasing the number of variational parameters; in this process exponent-sharing between s and p functions of the second principal quantum level, for example, is lost. When large basis set calculations on polyatomic molecules began to be carried out the advantages of exponent sharing were apparently not realized. Most efforts at writing efficient integral programs have tended to use features of the point-group symmetry of the molecule to avoid redundant operations. If exponent sharing is present in the basis then the actual calculation of the integrals, even without using special symmetry relations, can be speeded up substantially. Of course, the decreased flexibility

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of the basis, because of the fewer variational parameters, will yield poorer atomic energies, but a careful choice will usually not worsen the value of the constrained basis for molecular calculations. The 4-31G[39] and 6-31G[24] basis sets developed in this laboratory have been shown to be almost as good as double-zeta basis sets for absolute energies, and even better for relative energies; yet in these basis sets s and p functions in the valence shells share common exponents. The exact advantage of the constrained basis sets in accelerating integral evaluation is outlined later on.

The unnormalized 1s gaussian centered at point A will be denoted

$$|1s_A) = \exp\{-a|\vec{r}-\vec{R}_A|^2\} \tag{4.1}$$

where $\vec{R}_A$ is the vector from the space-fixed origin to the point A. Unnormalized $2p$ and $3d$ gaussians may be obtained from (4.1) by differentiation:

$$|2px_A) = \frac{1}{2a}\frac{\partial}{\partial A_x}|1s_A) = (x-A_x)\exp\{-a|\vec{r}-\vec{R}_A|^2\}$$

$$= x_A \exp\{-a|\vec{r}-\vec{R}_A|^2\} \tag{4.2}$$

$$|3d_{xy_A}) = \frac{1}{2a}\frac{\partial}{\partial A_y}|2p_{x_A}) = x_A y_A \exp\{-a|\vec{r}-\vec{R}_A|^2\} \tag{4.3}$$

In (4.2) and (4.3) A_x, A_y, A_z represent the coordinates of A, and x_A, y_A denote the appropriate coordinates with reference to the axis system located at A. The axes at A are taken to be parallel to the space-fixed axes. For a one-electron operator Ω the corresponding one-electron integral is defined as

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\ell_1\ell_2\ell_3|\Omega|\ell_1'\ell_2'\ell_3') = \int d\vec{r}\; x_A^{\ell_1} y_A^{\ell_2} z_A^{\ell_3} \exp\{-a|\vec{r}-\vec{R}_A|^2\}\;\Omega$$

$$\times\; x_B^{\ell_1'} y_B^{\ell_2'} z_B^{\ell_3'} \exp\{-b|\vec{r}-\vec{R}_B|^2\} \qquad (4.4)$$

The basic recurrence relations may be established from (4.4):

$$(\ell_1+1\;\ell_2\ell_3|\Omega|\ell_1'\ell_2'\ell_3') = \frac{1}{2a}\frac{\partial}{\partial A_x}(\ell_1\ell_2\ell_3|\Omega|\ell_1'\ell_2'\ell_3') +$$

$$\frac{\ell_1}{2a}(\ell_1-1\ell_2\ell_3|\Omega|\ell_1'\ell_2'\ell_3') \qquad (4.5)$$

with similar relations for ℓ_2, ℓ_3, ℓ_1', ℓ_2' and ℓ_3'. To obtain the specific recurrence relation for a given operator the form of the expression for $(\ell_1\ell_2\ell_3|\Omega|\ell_1'\ell_2'\ell_3')$ must be known as well as the exact expression for $(000|\Omega|000)$. Thus, if Ω = 1 (overlap), Boys has provided the expression for the 1s integrals:

$$(000|000) \equiv (000|1|000) = \left(\frac{\pi}{a+b}\right)^{3/2} \exp\left(-\frac{ab}{a+b}\,\overline{AB}^2\right) \qquad (4.6)$$

Application of the operators $\frac{\partial}{\partial A_x}$, $\frac{\partial}{\partial B_y}$, etc. to (4.6) reveals the form of the general integral:

$$(\ell_1\ell_2\ell_3|\ell_1'\ell_2'\ell_3') = \sum_{\substack{n_1n_2n_3 \\ ijk}} S^{\ell_1\ell_2\ell_3\ell_1'\ell_2'\ell_3'}_{n_1n_2n_3\,ijk}\; AB_x^{n_1} AB_y^{n_2} AB_z^{n_3}\; a^{-i}b^{-j}f^{k}$$

$$\times \exp\left(-\frac{f}{2}\,\overline{AB}^2\right) \qquad (4.7)$$

In (4.7) $AB_x = B_x - A_x$, etc. (4.5) and (4.7) together lead to (4.8):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$\Sigma S^{\ell_1+1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1n_2n_3\ ijk} AB_x^{n_1} AB_y^{n_2} AB_z^{n_3}\ a^{-i}b^{-j}f^{k} =$$

$$\frac{1}{2}\Sigma\, S^{\ell_1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1-1\ n_2n_3\ i-1j\ k-1} \cdots -\left(\frac{n_1+1}{2}\right) S^{\ell_1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1+1n_2n_3\ i-1jk} \cdots$$

$$+\ \frac{\ell_1}{2}\ \Sigma\, S^{\ell_1-1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1n_2n_3\ i-1jk} \cdots \qquad (4.8)$$

which establishes the recurrence relation for the coefficients S:

$$S^{\ell_1+1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1n_2n_3\ ijk} = \frac{1}{2}\ S^{\ell_1\ell_2\ell_3\ell_1\ell_2\ell_3}_{n_1-1n_2n_3\ i-1jk-1} - \left(\frac{n_1+1}{2}\right)\ S^{\ell_1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1+1n_2n_3\ i-1jk}$$

$$+\ \frac{\ell_1}{2}\ S^{\ell_1-1\ell_2\ell_3\ell'_1\ell'_2\ell'_3}_{n_1n_2n_3\ i-1jk} \qquad (4.9)$$

with

$$S^{000\ 000}_{000\ 000} = 1.0$$

(4.9) yields, for example,

$$S^{100\ 000}_{100\ 101} = 0.5 \qquad S^{000\ 100}_{100\ 011} = -0.5$$

The coefficients S in the expression for the overlap have only six subscripts. The analysis corresponding to the steps (4.7) to (4.9) has been performed for the following operators: ∇^2 (kinetic energy), V_C (potential energy of attraction due to the nucleus at center C), x, y and z (the three components of the electric dipole moment),

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

x^2, y^2, z^2, xy, xz, and yz (second moments of the charge distribution), third moments of the charge distribution, electric fields at the nuclei, and electric field gradients at the nuclei. The most complicated expressions had coefficients with eleven subscripts. The derivations were carried out with a view to programming the process corresponding to (4.9); it is therefore possible that (4.9) is not at its most elegant. Once (4.9) had been programmed, a series of routines was written to optimize the totality of coefficients for the case where both centers A and B had d functions. The output of the optimizer, which detected the occurrence of common subexpressions, was then converted by another routine to the form of FORTRAN statements. The resulting program works most efficiently when the largest number of functions share a common exponent (a, b). This is not a critical restriction for the one-electron integrals since their number is proportional only to the square of the number of functions in the basis.

If Ω_{12} is a two-electron operator the corresponding two-electron integral is defined as

$$(\ell_1\ell_2\ell_3|\Omega_{12}|\ell_1'\ell_2'\ell_3') = \int d\vec{r} \int d\vec{r}'\; x_P^{\ell_1} y_P^{\ell_2} z_P^{\ell_3} \exp\{-e|\vec{r}-\vec{R}_P|^2\}$$

$$\Omega_{12}\; x_Q'^{\ell_1'} y_Q'^{\ell_2'} z_Q'^{\ell_3'} \exp\{-f|\vec{r}'-\vec{R}_Q|^2\} \tag{4.10}$$

where, following Boys

$$\begin{aligned} e &= a + b \\ f &= c + d \\ P_x &= \frac{aA_x + bB_x}{a + b} \\ Q_x &= \frac{cC_x + dD_x}{c + d} \end{aligned} \tag{4.11}$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The electronic repulsion integral, with $\Omega_{12} = |\vec{r}-\vec{r}'|^{-1}$, was the only two-electron integral for which programs were written. Since the number of two-electron integrals in a given basis is proportional to the fourth power of the number of functions in the basis special care was exercised in setting up the programs. To appreciate the problem fully a more detailed discussion follows, with the time-saving devices clearly pointed out.

Figure 6 illustrates the situation one has to deal with. A, B, C, D are the locations of the four functions in a repulsion integral. Points P and Q have been defined as in (4.11). The expression given by Boys for the repulsion integral over unnormalized $1s$ gaussians is

$$(1s_A 1s_B \left|\frac{1}{r_{12}}\right| 1s_C 1s_D) = \frac{2\pi^{5/2}}{ef\sqrt{e+f}} \exp\left(-\frac{ab}{a+b}\overline{AB}^2 - \frac{cd}{c+d}\overline{CD}^2\right) F_o\left[\frac{ef}{e+f}\overline{PQ}^2\right] \tag{4.12}$$

This may equivalently be written as a 'two-center' integral with centers P and Q for the charge distributions representing electrons 1 and 2:

$$(1s_P \left| \frac{1}{r_{12}} \right| 1s_Q) = (1s_A 1s_B \left|\frac{1}{r_{12}}\right| 1s_C 1s_D) \tag{4.13}$$

the point P being obtained from A and B by (4.11), and, similarly Q from C and D. The repulsion integral (4.12) may be a one-, two-, three-, or four-center integral whereas the corresponding 'two-center' form (4.13) is either one- or two-center.

Proceeding as for the one-electron integrals it is easy to establish that the form of (4.13) for the general two-electron integral is

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

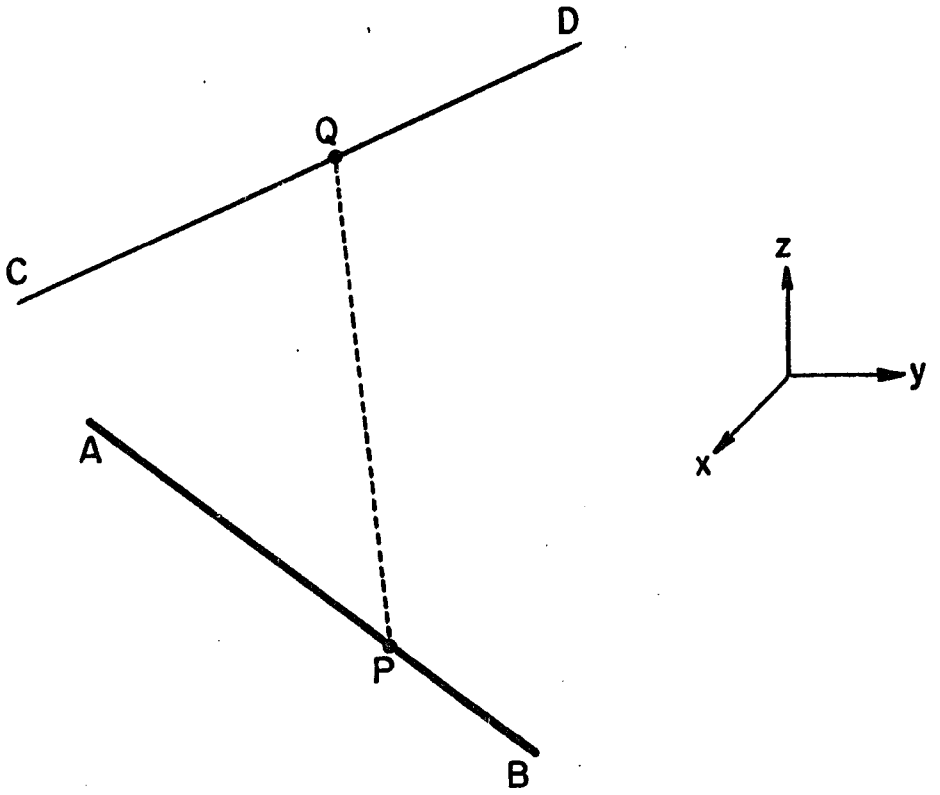


Fig. 6

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\ell_1\ell_2\ell_3 \left|\frac{1}{r_{12}}\right| \ell_1'\ell_2'\ell_3') = \frac{2\pi^{5/2}}{ef\sqrt{e+f}} \exp\left(-\frac{ab}{a+b}\overline{AB}^2 - \frac{cd}{c+d}\overline{CD}^2\right)$$

$$\times \sum_{\substack{n_1 n_2 n_3 \\ ijm}} G^{\ell_1\ell_2\ell_3\ell_1'\ell_2'\ell_3'}_{n_1 n_2 n_3 ijm} PQ_x^{n_1} PQ_y^{n_2} PQ_z^{n_3} e^{-i} f^{-j} F_m \qquad (4.14)$$

In (4.14) e^{-i} does *not* mean exp $(-i)$, but rather $(a+b)^{-i}$. The general practice is to express each function in (4.12) as a linear combination of a number of primitive gaussians. Thus there is a four-fold summation on the right-hand side of (4.14) if (4.14) is looked on as an integral over contracted gaussians. To accelerate the calculation of these integrals as many of the terms as possible should be moved to the outer summation signs or eliminated altogether. The method developed in this laboratory is a combination of both.

The final integral transmitted to the SCF section must appear referred to the space-fixed axes of the molecule. One is not restricted to the space-fixed axes when performing any or all of the summations in (4.14). As many intermediate axes as needed may be chosen to reduce the amount of calculation within the various summations. However, when a new axis system is selected it is necessary to rotate *all* the intermediate stage integrals from the old to the new system. (4.14) is rewritten to simplify further explanation.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(\ell_1\ell_2\ell_3|\ \frac{1}{r_{12}}\ |\ell_1'\ell_2'\ell_3') = 2\pi^{5/2} \sum_{n_D} \sum_{n_C} \exp\left(-\frac{cd}{c+d}\ \overline{CD}^{\,2}\right)$$

$$\times\ G^{\ell_1\ell_2\ell_3\ell_1'\ell_2'\ell_3'}_{n_1\ n_2\ n_3\ ijm}\ f^{-j-1} \sum_{n_B} \sum_{n_A} \exp\left(-\frac{ab}{a+b}\ \overline{AB}^{\,2}\right)\ \frac{e^{-i}-1}{\sqrt{e+f}} \qquad (4.15)$$

$$\times\ PQ_x^{n_1}\ PQ_y^{n_2}\ PQ_z^{n_3}\ F_m$$

n_A, n_B, n_C, n_D represent the degree of contraction at centers A, B, C and D. To reduce the amount of calculation in (4.15) one must either eliminate the terms PQ_x, PQ_y and PQ_z or move them into the outer summations. Our technique, when performing the inner summations (at A and B) is to orient the local z-axis along AB, the local y-axis as the perpendicular to the plane ABQ so that the local x-axis lies in the plane ABQ. With this choice of axes PQ_x can be calculated in the outer loop (over C and D), since PQ_x is the perpendicular distance between Q and the line AB. PQ_y is altogether eliminated by the particular choice of the local y-axis. Once the summation over A and B has been completed a rotation is performed about the local z-axis so that the local y-axis is brought into coincidence with the y-axis in a second axis system. This rotation, it must be noted, is performed outside the A and B summations. The functions at Q are then 'split' into those centered on C and D and, when the summation over C and D is completed, the second set of axes is rotated into coincidence with the space-fixed set. The distances PQ_x, PQ_y, PQ_z are determined by the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

spatial distribution of the four centers and the orbital exponents of the primitive gaussians. If a primitive s at each center shared its exponent with the three primitive p functions at that center then, for this set of four functions a common set of the distance components can be used. If s, p and d functions shared exponents then the sequence outlined above would enable one to calculate up to 10 000 integrals from one set of e's, f's, F_m's and the distance components. Within the AB loop this would lead to just 1822 integrals. The rotation in the CD loop would increase this to 3500 and, finally, the 'splitting' at Q transforms the 3500 numbers into 10 000 all of which must be rotated to the space-fixed axes.

It has been explained that in basis sets developed in this laboratory, exponent sharing by functions of different angular quantum numbers is resorted to when possible. In what follows a *shell* is a set of contracted gaussians whose primitives share exponents. An s shell consists of just one function. There are four functions in an sp shell, one s function and three p functions with which it shares exponents. A d shell consists of *either* five *or* six functions, depending on whether the set (2.4) or the set (2.5) is used. An spd shell consists of either nine or ten functions.

There are more computer-oriented ways of speeding up the calculation of integrals. When possible, operations are carried out in single precision. If an integral has a small value, then conceivably *all* operations leading up to the final summation can be performed in single precision. Parallel efforts by Clementi[45] have come to our attention.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In a crude way an integral may be considered to be small if the shells at A and B are not identical, or if the shells at C and D are not identical. More sophisticated methods are being tried out. Estimates of the degree of smallness of the largest integral in a set of four shells may be obtained on the basis of differential overlap. Depending on the smallness index the two-electron integrals could be:

(i) calculated in double precision, or

(ii) calculated in single precision, or

(iii) calculated in a related basis; a basis with a lower degree of contraction can be substituted for the original basis, and the integrals can be calculated in single precision over the shorter basis.

Even with the crude test of smallness, we have obtained an improvement of 15%-20% in the time taken to calculate the two-electron integrals. Clementi,[45] who has been investigating similar techniques, has been able to bring down to 80 min. a calculation that took more than 12 000 min., a substantial reduction by a factor of about 150. Though these procedures do give total energies to the specified accuracy for moderately large molecules it remains to be established that these approximations do not degrade the wave function itself.

Our integral programs have been checked out in several ways. For a test case involving 250 000 two-electron integrals, all of the integrals from our program agreed to within a microhartree of the output from the Bell Laboratory's version of POLYATOM. As a check of the transformation of the six second-degree gaussians (3.1) to the five pure

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

d type gaussians (3.2), the water calculation of Neumann and Moskowitz[27] was reproduced. The program was tested for internal consistency by repeating various calculations with rotated axes.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

V. THE INFLUENCE OF POLARIZATION FUNCTIONS ON MOLECULAR ORBITAL HYDROGENATION ENERGIES

Although it has long been recognized that single-determinant molecular orbital theory does not give good bond dissociation energies because of neglect of electron correlation, there remains the possibility that it is more effective in predicting the energies of reactions involving only closed shell species, where correlation corrections may largely cancel. This hypothesis was put forward by Snyder.[40] One important class of reactions of this type is hydrogenations of organic molecules to products containing only one non-hydrogen atom (methane, ammonia, water, etc.). The energies of such reactions relate the strength of bonds between heavy atoms to bonds involving hydrogen. The validity of such a hypothesis is important since an adequate theory of hydrogenation energies would permit evaluation of the total energy of larger molecules, the energies of the hydrogenation products being well known.

Ideally, molecular orbital calculations should be carried out with a large enough basis for the Hartree-Fock limit to be reached. In practice, however, this is only possible for a number of rather small systems. Consequently, attention has been devoted to reaction energies using a more limited basis set, requiring that the same basis be used for reactants and products. In such investigations it is hoped that some of the errors due to the limitation of basis set will also cancel to some extent. An extensive study of this sort was first carried out by Snyder and Basch,[41] who examined a number of reactions using a

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

contracted gaussian basis of the 'split-shell' or 'double zeta' type--that is with two basis functions per atomic orbital (ten functions for carbon, nitrogen, etc. and two for hydrogen). They found moderate agreement with experiment for heats of hydrogenation, the theoretical values being mostly too negative. Further work with a rather simpler 'split valence shell' basis (4-31G) gave similar results.[42,43]

All the systematic molecular orbital studies of hydrogenation energies to date have used only *s*- and *p*-type atomic functions in the basis. For a limited number of the molecules involved, more extensive basis sets including polarization functions (*d*-type on heavy atoms and *p*-type on hydrogen) have been used in individual calculations, so that Snyder and Basch[41] were able to make some observations about the difference between errors with their 'double-zeta' basis and errors at the Hartree-Fock limit. However, nobody has yet attempted a systematic study of hydrogenation energies at a uniform level with a basis including polarization functions. Such a study is presented here for a set of molecules involving atoms H, C, O, N and F with two non-hydrogen atoms. The aims are (1) to find whether addition of polarization functions gives overall improved agreement with experimental hydrogenation energies, and (2) to make a comparative study of the consequent total energy lowering as a function of atom and type of chemical environment.

The molecules considered in this study are those containing H, C, N, O and F which can be represented by a classical valence structure without formal charges and which contain up to two heavy atoms. Standard model geometry[23,43] was used throughout (see footnote a to Table V).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Carbon monoxide was added to the list with a bondlength of 1.13 Å. The standard model was used partly because experimental geometries are incompletely known and partly so that the results could be fitted in with a systematic study of larger molecules. Various studies not reported in detail here indicate that hydrogenation energies calculated at experimental geometries are generally within 1 kcal/mole of the standard geometry values.

Molecular orbital methods are standard using the Roothaan equations (with the appropriate unrestricted generalization for the triplet ground state of O_2). We begin with already published results using the 4-31G basis set.[42,43] This has an *s* inner shell function which is a linear combination of four *s*-type gaussians and a valence shell (*s* and *p*) represented by inner (three-gaussian) and outer (one-gaussian) parts. For hydrogen, similar (31) *s*-type functions are used. This basis has the special feature that gaussian *s* and *p* exponents are shared in the valence shell for computational efficiency.

The 4-31G basis is improved in three successive steps. To begin with, it is replaced by 6-31G which differs by improvement of the inner shell function.[24] This is found to lower total energies substantially, but should not lead to large changes in chemical properties such as reaction energies. Nevertheless, it is desirable to test this before adding polarization functions.

The next step is the addition of a set of single (uncontracted) *d*-gaussian functions to the 6-31G set for heavy atoms, the hydrogen functions being unmodified. The additional six (unnormalized) functions are

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

$$(x^2,\ y^2,\ z^2,\ xy,\ yz,\ zx)\ \exp\ (-\alpha_d\ r^2)$$

These are equivalent to the five pure d-type functions

$$(3z^2-r^2,\ xz,\ yz,\ xy,\ x^2-y^2)\ \exp\ (-\alpha_d\ r^2)$$

together with the further s-type function $r^2\ \exp\ (-\alpha_d\ r^2)$. We have retained all six functions rather than just the five pure-d functions, since the appropriate integrals can be evaluated more directly.

The addition of the d-functions would be most economically carried out if the exponent α_d were chosen to be identical to that in the outer s and p functions of 6-31G. This is because the integral evaluation program actually treats all ten functions s, p and d together. However, this turns out to be somewhat ineffective since the outer s and p functions are very diffuse and the d-type polarization functions do not interact strongly with the inner valence shell functions. α_d were chosen as reported in Section III.

The next step is to extend the basis further by adding p-polarization functions

$$(x,\ y,\ z)\ \exp\ (-\alpha_p\ r^2)$$

on the hydrogen atoms and find optimum values for α_p as described in Section III.

We have now introduced two extensions of the 6-31G basis. In the first (which we refer to as 6-31G*), there are six extra uncontracted gaussian functions on heavy atoms. The second, more elaborate, set (6-31G**) also has three p-type gaussian functions on each hydrogen atom.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

All three of these bases have been applied to the complete set of molecules under investigation. The total energies are listed in Table V.

In the following discussion, we shall attempt to make some comparisons with results at the Hartree-Fock limit even though these are not well established for many molecules. The final column of Table V gives a set of estimates of the energy at this limit for molecules which have been studied with more extensive bases than any used here. These figures are slightly below the best calculated energies in most cases and necessarily rather tentative.

Energies of hydrogenation reactions (ΔE) are easily derived from the results in Table V. The theoretical values of ΔE for the 4-31G, 6-31G and the full polarization set 6-31G** are listed in Table VI. The corresponding estimates at the Hartree-Fock limit are also given for C_2H_2, HCN, CO, N_2 and F_2. Again, these values are subject to considerable uncertainty and may have an error of ± 5 kcal/mole. The final column of Table VI lists experimental values corrected for temperature effects and zero-point motion. For some of the molecules these may also be subject to an uncertainty of several kilocalories.

The theoretical results are clearly poorest for O_2 which is reasonable, since this molecule has a triplet ground state and hydrogenation involves a change in the number of unpaired electrons. If O_2 is omitted from the list, we may test the relative success of the basis sets by evaluating the mean absolute deviation between theory and experiment when the latter is available. This gives 6.8, 7.1 and 4.0 kcal/mole for 4-31G, 6-31G and 6-31G** respectively. We deduce that

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

improvement of inner shell orbitals causes little change in theoretical hydrogenation energies but that addition of polarization functions leads to a marked overall improvement. If we consider the available results for single hydrogenations (addition of only one molecule of hydrogen), the mean deviation is reduced to 3.1 kcal/mole for the basis with full polarization functions.

Before analyzing these results in more detail, it is useful to present the energy lowerings due to *d*-function inclusion (6-31G→6-31G*) for various valence environments of the heavy atoms. These are given for symmetrical molecules in Table VII and show a number of interesting features. The energy lowerings are generally smallest for carbon. Comparing various types of carbon, it is found that the lowering is greatest for saturated tetrahedral carbon. This is a slightly surprising result in view of the high symmetry involved. It may perhaps be rationalized by noting that tetrahedral symmetry does not include an inversion operation and does, therefore, permit mixing of *d*-functions into those molecular orbitals which already include the valence *p*-functions. The carbon atoms in acetylene, on the other hand, have an approximate local center of inversion, in the sense that there are bonds in opposite directions. Mixing between *d* and *p* functions may well be reduced under these circumstances. *s*-*d* mixing may occur, but this may be less effective because the *s*-type orbitals are more tightly bound. The trigonal carbon in ethylene is apparently intermediate between acetylene and ethane.

For nitrogen and oxygen atoms, the energy lowerings are larger. This can perhaps be attributed to greater anisotropy and non-uniformity of the charge distribution within the atom, leading to stronger

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

polarization effects. Unlike the carbon series, the effect is greatest for the atom with unsaturated valency. Triply-bonded nitrogen and doubly-bonded oxygen have no pseudo-center of symmetry and *d*-polarization is large. The doubly-bonded oxygen value obtained from O_2 is not, of course, a suitable general model, but the strong *d*-lowering in formaldehyde (35.0 kcal/mole) suggests that the carbonyl oxygen contribution is of the order of 26 kcal/mole.

Table VIII gives a corresponding set of energy lowerings due to the addition of *p*-functions on hydrogen subsequent to *d*-function inclusion (6-31G*→6-31G**). These are smallest for H-H and C-H bonds but increase sharply in going to more polar bonds. This can be attributed partly to additional polarization effects in the σ-bond, but another contributing factor in a molecule such as HF is delocalization of π-electron onto hydrogen if *p*π functions are available there.

We now return to the discussion of the hydrogenation energies shown in Table VI. The results for the hydrocarbons show little change due to the addition of polarization functions, all hydrogenation energies being too negative. Using the 6-31G** basis, the results for acetylene give $\Delta E_{th} - \Delta E_{exp} = -12.5$ kcal/mole which is the poorest result for all the singlet-state molecules. The fact that no improvement is achieved by addition of polarization functions suggests that an error of this order may remain at the Hartree-Fock limit. The estimate based on more accurate calculations lends support to this. All of the one-step hydrogenation energies for the C_2 hydrocarbons are also too negative, 6-31G** values for $\Delta E_{th} - \Delta E_{exp}$ being -5.1, -3.8 and -3.6 kcal/mole for acetylene, ethylene and ethane respectively.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The behavior of the three N_2 compounds is quite different. With the *sp* bases (4-31G and 6-31G), the theoretical hydrogenation energy of N_2 is too negative. Snyder and Basch[41] also found this. However, addition of polarization functions changes this sharply and the final (6-31G**) result gives $\Delta E_{th} - \Delta E_{exp} = +4$ kcal/mole. The estimated Hartree-Fock limit, noted by Rauk, Allen and Clementi[44] is close to this. Inspection of Table VII shows that the large change in ΔE_{th} due to inclusion of polarization functions is primarily due to the large extra *d*-lowering for the N_2 molecule, partly offset by extra hydrogen *p*-lowering for hydrogens in NH bonds. The other N_2 compounds diimide and hydrazine show similar changes of rather smaller magnitude.

We have already noted that the theoretical hydrogenation energy of the oxygen molecule is too positive, which is consistent with a smaller correlation correction for a triplet-state. For the remaining symmetrical single bonds in hydrogen peroxide and the fluorine molecule, the theoretical hydrogenation energies are too negative. In the case of F_2, an even more negative value of -149 kcal/mole is obtained if the estimated Hartree-Fock energies are used. This gives $\Delta E_{th} - \Delta E_{exp} = -15$ kcal/mole, a discrepancy already noted by Snyder and Basch.[41] It is evident that for this molecule, a significant change in ΔE_{th} occurs as the best of our basis sets is extended to the Hartree-Fock limit. This may well be due to the limitation of using only four primitive *p*-gaussians in the valence shell of fluorine since Snyder and Basch (using a basis with five primitive *p*-gaussians but without polarization functions) obtain a better result than 6-31G. It should also be noted that the fluorine

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

molecule is somewhat exceptional in that no binding relative to separate fluorine atoms is obtained at the Hartree-Fock limit.[14]

For the unsymmetrical molecules with two heavy atoms, the results of Table VI are generally intermediate between those of the corresponding symmetrical systems. Thus, we have already noted that C_2 molecules give $\Delta E_{th} - \Delta E_{exp}$ too negative and N_2 molecules slightly too positive. For CN compounds, the corresponding values are slightly too negative. In the case of hydrogen cyanide, the value of $\Delta E_{th} - \Delta E_{exp}$ is about -4 kcal/mole with 6-31G** and with the estimated Hartree-Fock values. Among the other molecules, we may note that the value of ΔE_{th} for carbon monoxide changes sharply on addition of polarization functions. This is similar to the result for the isoelectronic molecule N_2.

Another related class of formal reactions that may be investigated in a similar way is the separation of multiple bonds into single bonds. For example, the reaction

$$H_2C = CH_2 + 2CH_4 \longrightarrow 2\ H_3C - CH_3$$

is one in which the two 'parts' of the CC double bond of ethylene are separated into CC single bonds in separate ethane molecules, the number of CH bonds remaining constant. It is the bond separation reaction[42] for ethylene, treated as a member of the cycloalkane series $(CH_2)_n$. The energy of the reaction is also the amount by which the double bond is stronger than two single bonds. The theoretical values for the energies of these reactions using the 6-31G, 6-31G* and 6-31G** bases are compared with experimental data in Table IX. It is evident that addition of

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

polarization functions improves the results leading to excellent agreement with experiment for both 6-31G* and 6-31G**. (Close correspondence between 6-31G* and 6-31G** results is expected since all hydrogens remain bonded to the same type of other atom in these reactions.)

The success of the comparison with experiment displayed in Table IX suggests that the energies of these reactions would continue to be given well at the Hartree-Fock limit. If this is so, it would follow that the correlation contribution to the energy of a double (or triple) bond would be close to two (or three) times the correlation contribution to the corresponding single bond.

Next, it is useful to summarize the differences between theoretical and experimental hydrogenation energies in chart form so that trends are apparent. Table X does this for the 6-31G** basis. It may be noted that the values for double and triple bonds are roughly two and three times those for single bonds. This reflects the successful comparison in Table IX. Finally, we may make an attempt to estimate the residual error in all the theoretical hydrogenation energies at the Hartree-Fock limit. Table XI displays such an estimate based on this discussion. The numbers are arrived at in the following steps.

(1) Values for double and triple bonds are taken to be respectively twice and thrice those for single bonds in accordance with the hypothesis put forward above.

(2) Available evidence suggests that the 6-31G** values for CC, CN and NN bonds are close to the Hartree-Fock limit.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

(3) For the FF in the fluorine molecule, the Hartree-Fock value of about -15 kcal/mole is 12 kcal/mole more negative than the 6-31G** result. It is therefore suggested that the value for the C-F bond in methyl fluoride should be reduced by 6 kcal/mole, giving -3 kcal/mole as the estimated limiting error.

(4) Values for the remaining bonds are interpolated roughly to give the complete table.

It should be emphasized that the numbers given in Table XI are only tentative and may have to be modified in the light of future studies with larger basis sets. However, they do sum up our present assessment of the performance of Hartree-Fock theory in predicting hydrogenation energies.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE V. Total Energies (hartrees) with Standard Polarization Functions and Estimated Hartree-Fock Limits

Molecule[a]	6-31G	6-31G*	6-31G**	Limit[c]
H_2	-1.12676	-1.12676	-1.13129	-1.1336
CH_4	-40.18038	-40.19506	-40.20159	-40.225
NH_3	-56.16320	-56.18374	-56.19499	-56.225
OH_2	-75.98508	-76.00987	-76.02255	-76.065
FH	-99.98343	-100.00281	-100.01122	-100.071
HC≡CH	-76.79261	-76.81732	-76.82138	-76.860
$H_2C{=}CH_2$	-78.00317	-78.03037	-78.03754	
$H_3C{-}CH_3$	-79.19651	-79.22774	-79.23724	
HC≡N	-92.82763	-92.87317	-92.87515	-92.920
$H_2C{=}NH$	-93.97527	-94.02105	-94.02824	
$H_3C{-}NH_2$	-95.16717	-95.20819	-95.22220	
$C^{-}{\equiv}O^{+}$	-112.66722	-112.73718	-112.73718	-112.791
$H_2C{=}O$	-113.80789	-113.86370	-113.86711	
$H_3C{-}OH$	-114.98682	-115.03387	-115.04501	
$H_3C{-}F$	-138.99200	-139.03445	-139.03961	
N≡N	-108.86762	-108.94234	-108.94234	-108.997
HN=NH	-109.92792	-109.99123	-109.99768	
$H_2N{-}NH_2$	-111.11852	-111.16733	-111.18147	
HN=O	-129.71179	-129.78123	-129.78426	
$H_2N{-}OH$	-130.92160	-130.97505	-130.98818	
$H_2N{-}F$	-154.90449	-154.95130	-154.95897	
$O{=}O$[b]	-149.54546	-149.61440	-149.61440	
HO-OH	-150.70287	-150.75299	-150.76540	
HO-F	-174.68169	-174.72327	-174.72980	
F-F	-198.64605	-198.67290	-198.67290	-198.770

continued--

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE V (continued)

[a]Bond lengths and angles as specified in reference 23. $r(C^{-}\equiv O^{+}) = 1.13\text{Å}$. Other conformational information: Staggered bonds for single bonds to carbon, HN=NH trans, H_2N-NH_2 with 90° dihedral angle between nitrogen lone pair directions, H_2N-OH with nitrogen lone pair cis to OH, HO-OH with 90° dihedral angle.

[b]Lowest triplet state energies.

[c]Estimated Hartree-Fock limits based on the following references or calculations reported therein: H_2, Kolos, W., Roothaan, C. C. J.: Rev. Mod. Physics 32, 219 (1960); CH_4, reference 29; NH_3, reference 44; OH_2, reference 27; FH, reference 18; HC≡CH and HC≡N, McLean, A. D., Yoshimine, M.: 'Tables of Linear Molecule Wave Functions' (IBM Corp., 1967); $C^{-}\equiv O^{+}$, reference 15; N≡N, reference 16; F-F, reference 14.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE VI. Complete Hydrogenation Energies ΔE (kcal/mole)

Reaction	Calculated 4-31G[a]	6-31G	6-31G**	Limit[b]	Exp.[c]
$HC\equiv CH + 3H_2 \rightarrow 2CH_4$	-117.8	-117.9	-117.9	-119	-105.4
$H_2C=CH_2 + 2H_2 \rightarrow 2CH_4$	-65.9	-65.3	-64.7		-57.2
$H_3C-CH_3 + H_2 \rightarrow 2CH_4$	-23.5	-23.5	-21.7		-18.1
$HC\equiv N + 3H_2 \rightarrow CH_4 + NH_3$	-83.4	-85.1	-80.0	-81	-76.8
$H_2C=NH + 2H_2 \rightarrow CH_4 + NH_3$	-72.1	-72.0	-66.4		
$H_3C-NH_2 + H_2 \rightarrow CH_4 + NH_3$	-30.9	-31.2	-27.2		-25.7
$C^{-}\equiv O^{+} + 3H_2 \rightarrow CH_4 + H_2O$	-72.4	-74.0	-58.4	-64	-63.9
$H_2C=O + 2H_2 \rightarrow CH_4 + H_2O$	-64.3	-65.3	-59.3		-57.3
$H_3C-OH + H_2 \rightarrow CH_4 + H_2O$	-32.0	-32.6	-30.0		-30.3
$H_3C-F + H_2 \rightarrow CH_4 + HF$	-27.4	-28.3	-26.3		-29.5
$N\equiv N + 3H_2 \rightarrow 2NH_3$	-47.3	-49.3	-33.7	-33	-37.7
$HN=NH + 2H_2 \rightarrow 2NH_3$	-90.7	-91.0	-81.4		-83.5
$H_2N-NH_2 + H_2 \rightarrow 2NH_3$	-50.4	-50.9	-48.5		-50.0
$HN=O + 2H_2 \rightarrow NH_3 + H_2O$	-114.2	-114.8	-107.1		-102.9
$H_2N-OH + H_2 \rightarrow NH_3 + H_2O$	-62.0	-62.7	-61.5		
$H_2N-F + H_2 \rightarrow NH_3 + HF$	-71.2	-72.4	-72.8		
$O=O + 2H_2 \rightarrow 2H_2O$	-107.1	-107.4	-105.5		-125.1
$HO-OH + H_2 \rightarrow 2H_2O$	-86.3	-88.2	-93.1		-86.8
$HO-F + H_2 \rightarrow H_2O + HF$	-98.2	-100.4	-108.4		
$F-F + H_2 \rightarrow 2HF$	-118.9	-121.8	-137.0	-149	-133.8

[a] From reference 43 except for carbon monoxide which is based on E(4-31G) = -112.55234 hartrees for a bondlength of 1.13Å.

[b] Based on estimated Hartree-Fock limiting energies listed in Table V.

[c] Heats of hydrogenation at 0°K corrected for zero-point vibrations. See reference 42 for details

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE VII. Energy Lowerings per Atom (kcal/mole) Due to Addition of *d*-Functions

Atom	Molecule	Energy Lowering
	CH_4	9.2
	C_2H_6	9.8
	C_2H_4	8.5
	C_2H_2	7.8
	NH_3	12.9
	N_2H_4	15.3
	N_2H_2	19.9
	N_2	23.4
	H_2O	15.6
	H_2O_2	15.7
	O_2	22.1
	HF	12.2
	F_2	8.4

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE VIII. Energy Lowerings per Hydrogen Atom (kcal/mole) Due to Addition of *p*-Functions

Bond	Molecule	Energy Lowering
H-H	H_2	1.4
C-H	CH_4	1.0
	C_2H_6	1.0
	C_2H_4	1.1
	C_2H_2	1.3
N-H	NH_3	2.4
	N_2H_4	2.2
	N_2H_2	2.0
O-H	H_2O	4.0
	H_2O_2	3.9
F-H	FH	5.3

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE IX. Energies of Multiple Bond Separation Reactions (kcal/mole)

Reaction	Calculated 6-31G	6-31G*	6-31G**	Experimental
$HC{\equiv}CH + 4CH_4 \rightarrow 3H_3C{-}CH_3$	-47.3	-53.8	-52.7	-51.1
$H_2C{=}CH_2 + 2CH_4 \rightarrow 2H_3C{-}CH_3$	-18.3	-22.0	-21.2	-21.0
$HC{\equiv}N + 2CH_4 + 2NH_3 \rightarrow 3H_3C{-}NH_2$	+8.3	+3.9	+1.4	+0.3
$H_2C{=}O + CH_4 + OH_2 \rightarrow 2H_3C{-}OH$	-0.2	+0.6	+0.8	+3.3
$N{\equiv}N + 4NH_3 \rightarrow 3H_2N{-}NH_2$	+103.4	+110.0	+111.6	+112.3
$HN\ NH + 2NH_3 \rightarrow 2H_2{-}NH_2$	+10.8	+15.1	+15.5	+16.5

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE X. Hydrogenation Energy Errors $\Delta E_{th} - \Delta E_{exp}$ with the 6-31G** Basis (kcal/mole)

	Single Bonds				Double Bonds			Triple Bonds	
	C	N	O	F	C	N	O	C	N
C	-3.6	-1.5	+0.3	+3.2	-7.5		-2.0	-12.5	-3.2
N		+1.5				+2.1	+4.1		+4.0
O			-6.3				+19.6[a]		
F				-3.2					

[a]Triplet state

TABLE XI. Estimated Hydrogenation Energy Errors $\Delta E_{th} - \Delta E_{exp}$ at the Hartree-Fock Limit (kcal/mole)

	Single Bonds				Double Bonds			Triple Bonds	
	C	N	O	F	C	N	O	C	N
C	-4	-1	-2	-3	-8	-2	-4	-12	-3
N		+1	-2	-5		+2	-4		+3
O			-6	-10					
F				-15					

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

VI. THE EFFECT OF POLARIZATION FUNCTIONS ON CONFORMATIONS OF SIMPLE CARBOCATIONS

The determinantal wave functions obtained by the LCAO-SCF procedure have been called the Hartree-Fock-Roothaan (HFR) wave functions[16]. If the size of the AO basis used for the expansion is gradually increased, the HFR wave functions approach the Hartree-Fock wave function in a limiting manner. One may increase the size of the basis by adding more functions of types already present and, additionally, including functions of different types. When the basis is allowed to grow in this fashion, the time taken to calculate the two-electron integrals increases rapidly. It has been pointed out in section IV that this time may be substantially reduced if one uses constrained basis sets in which the primitive gaussians of different symmetry types share exponents. Both minimal[46,47] and extended[24,39] basis sets of this kind have been developed by Hehre, Ditchfield, Stewart and Pople. To obtain near-Hartree-Fock wave functions, it is necessary to determine all parameters variationally: orbital exponents and geometric parameters. In principle, exponent optimization is not necessary if the basis is complete. Cade and Huo[18] have noted that if the basis is large energy improvements due to exponent optimization are marginal. However, even if the basis is complete, the best energy can be obtained only by variationally determining the geometric parameters of a system. A systematic way of approaching the HF limit is to use progressively better basis sets and determine the best geometry of the system for each basis. The efforts in this direction from this laboratory may be summarized in two dimensions as shown in Table XII. Advance along the diagonal [1-1]→[2-2]→[3-3]→... constitutes the fastest path to the HF limit. In this description the

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

first numeral denotes the basis set used for the energy calculation and the second numeral specifies the basis for which the geometry has been optimized. Thus, [3-2] stands for use of basis set 3 (6-31G) to calculate the energy at the geometry optimized for basis set 2 (4-31G). The [1-1], [2-1] and [2-2] geometries of the hydrides of C, N, O and F have been published.[48] For larger systems and the more extended basis sets complete geometry optimizations become more expensive. A more modest effort then is to advance directly down, e.g., from [1-1] to [2-1], rather than attempt the expensive route from [1-1] to [2-2]. The geometries of the two-heavy atom hydrides have been obtained[49] for STO-3G, but not for 4-31G. In this section we report some [1-4], [1-5], [2-4] and [2-5] structures for the simple carbocations,[50] CH_5^+, $C_2H_3^+$ and $C_2H_5^+$.

The fullest *ab initio* studies published to date suggest the following:

1. CH_5^+ is more stable in the C_s form (Figure 7) than in the trigonal bipyramid form (Figure 8).

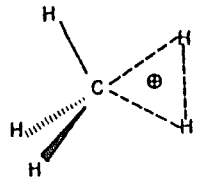


Fig. 7

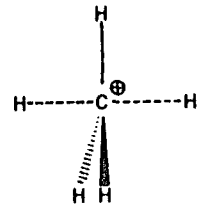


Fig. 8

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

2. $C_2H_3^+$ is most stable in the classical (planar) vinyl form (Figure 9), the symmetrically bridged form (Figure 10) being 18.5 kcal/mole less stable.(51)

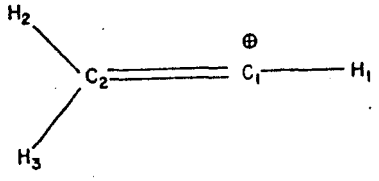


Fig. 9

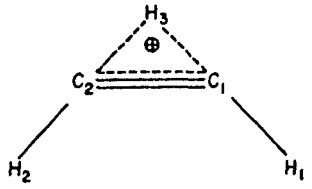


Fig. 10

3. $C_2H_5^+$ in the classical ethyl form (Figure 11) is slightly more stable than the alternative classical form (Figure 12). The symmetrically bridged form (Figure 13) is predicted to be 6-9 kcal/mole less stable than the classical forms.(52,53)

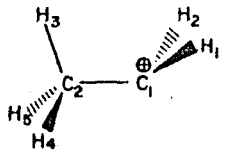


Fig. 11

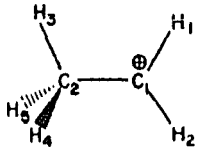


Fig. 12

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

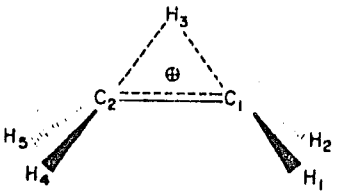


Fig. 13

These results are based on a number of different basis sets and varying degrees of geometrical searching for energy minima. Since the energy differences involved are quite small, further work is desirable to test these conclusions qualitatively as the HF limit is more closely approached. In order to give predictions comparable significance for the various species involved we have used a systematic series of theoretical procedures involving successive refinements of basis sets and geometrical searching. The calculations were carried out in the following steps:

Step 1. All geometrical parameters are varied (subject to the symmetry constraints implicit in Figures 7-13) to minimize the energy with the minimal STO-3G basis. The [1-1] structures thus obtained have been reported[53] (except for Figure 10).

Step 2. The extended 4-31G basis is then used to carry out single calculations with the geometry obtained in Step 1. These [2-1] points have also been published.[53]

Step 3. The geometries are now revaried with the 4-31G basis. The [2-2] geometries have been obtained for the forms in Figures 7, 8, 9, 10 and 13. For Figure 11 and Figure 12 only a partial refinement was possible because of the larger number of geometrical parameters. The new geometries are (H_{12} and H_{45} denote midpoints of H_1H_2 and H_4H_5 respectively; r in Å, angles in degrees):

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Figure 9: $r(CC) = 1.263$, $r(C_1H_1) = 1.068$,
$r(C_2H_2) = 1.083$, $C_1\hat{C}_2H_2 = 121.1$.

Figure 10: $r(CC) = 1.214$, $r(C_1H_1) = 1.065$,
$r(C_1H_3) = 1.301$, $C_1\hat{C}_2H_2 = 177.0$,

Figure 11: $r(CC) = 1.440$, $H_1\hat{C}_1H_2 = 116.4$,
$\hat{C}_2H_1H_{12} = 177.5$,
$r(C_1H_1) = 1.077$, $C_1\hat{C}_2H_3 = 105.3$,
$r(C_2H_3) = 1.110$, $C_1\hat{C}_2H_{45} = 136.3$,
$r(C_2H_4) = 1.091$, $H_4\hat{C}_2H_5 = 113.6$.

Figure 12: $r(CC) = 1.443$, $C_2\hat{C}_1H_1 = 123.0$,
$r(C_1H_1) = 1.077$, $C_2\hat{C}_1H_2 = 120.8$,
$r(C_1H_2) = 1.077$, $C_1\hat{C}_2H_3 = 114.9$,
$r(C_2H_3) = 1.088$, $C_1\hat{C}_2H_{45} = 121.0$,
$r(C_2H_4) = 1.101$, $H_4\hat{C}_2H_5 = 105.9$.

Figure 13: $r(CC) = 1.375$, $C_2\hat{C}_1H_{12} = 176.8$,
$r(C_1H_1) = 1.074$, $H_1\hat{C}_1H_2 = 117.8$,
$r(C_1H_3) = 1.329$.

Step 4. The geometries obtained in Step 3 are used for single calculations with the extended 6-31G basis which differs from 4-31G by having a more accurate carbon inner shell atomic orbital.

Step 5. The same [2-2] geometry is used, with the 6-31G* basis which contains polarization functions on carbon. Addition of similar functions has previously been shown to lower the energy of CH_3^+ significantly,[54]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

but no comparable study on the other ions has yet been published. Dyczmons *et.al*,[55] however, used basis functions centered in the C-H bonds which, as has been shown in Section I, may have comparable effects.

Step 6. In the final step single calculations with the [2-2] geometries were carried out with the full polarization basis, 6-31G**. The energies obtained in Steps 3-6 are listed in Table XIII together with the lowest values in the current literature. Table XIV gives the corresponding differences tabulated in a similar manner. The results show that improvement of the carbon inner shell function (Step 4) has only a small influence on the relative energies, but that the addition of polarization functions (Steps 5 and 6) leads to substantial changes.

For CH_5^+ the addition of *d* functions on carbon lowers the energy of the C_s form (Fig. 7) by 12.5 kcal/mole; but in the D_{3h} form (Fig. 8) the lowering is only 7.0 kcal/mole. This can probably be attributed to the higher local symmetry at the carbon atom in Figure 8 which implies less mixing with *d* functions. These energy changes open up the energy gap between the forms in Figure 7 and 8. The subsequent addition of *p* functions on hydrogen further increases the separation to 16.0 kcal/mole. The final results are in good agreement with those of Dyczmons, *et al.*[55]

In $C_2H_3^+$ the carbon *d* functions stabilize the bridged form (Figure 10) much more than the classical vinyl cation (Figure 9), leading to a reduction of the barrier for 1,2-hydride transfer. Again, hydrogen *p* functions lead to a further change in the same direction and the final barrier is reduced from 19.2 kcal/mole (Step 3) to 5.7 kcal/mole (Step 6).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

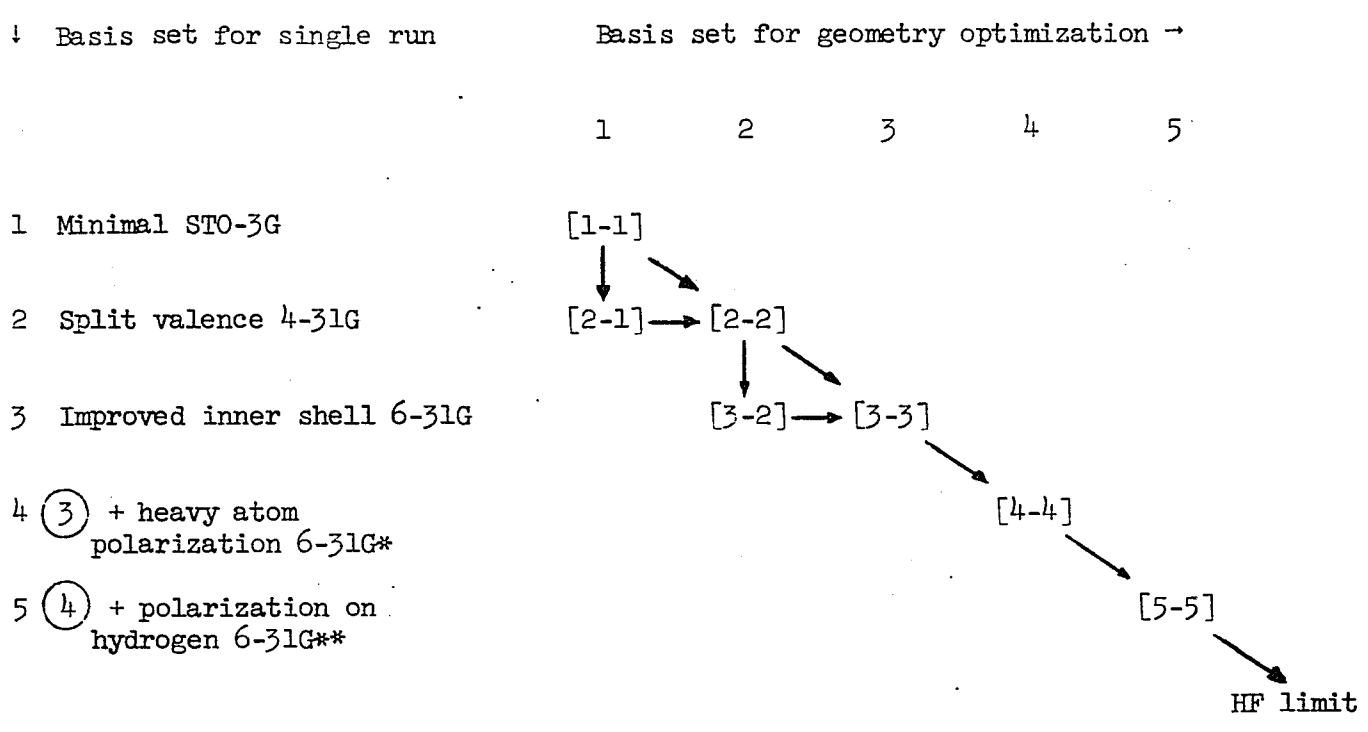
The relative energies for $C_2H_5^+$ are affected in a similar manner. The small energy difference between the classical forms (Figures 11 and 12) is unaltered by the addition of *d* functions, but the barrier to hydride transfer (Figure 11-Figure 13) is greatly reduced. With the full set of basis functions on carbon and hydrogen (Step 6), the symmetrically bridged form Figure 13 turns out to have the lowest energy. This suggests that the classical forms (Figures 11 and 12) may not be close to local minima in the potential surface and that the bridged form may be the equilibrium structure at the HF limit.

The actual equilibrium structure of $C_2H_5^+$ will, of course, be influenced by the relative values of correlation energies for the various structures. Little is yet known about these, but we may note that studies[56] of H_3^+ indicate that the correlation energy for this three-center bond system is slightly greater than the value for the isoelectronic two-center bond in H_2. If this comparison can be carried over to $C_2H_5^+$, the bridged form (Figure 13) is favored further.

The main conclusion to be drawn from these results is that the energies of all structures involving three-center bonds (Figures 7, 10, and 13) are strongly lowered by inclusion of higher-order angular functions. This result is not surprising since, as noted in Section I, basis sets lacking polarization functions are too rigid and the addition of polarization functions permits 'bending' and 'drifting'. These points should be taken into consideration in theoretical discussions of larger carbonium ions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XII. Systematic Approach to the Hartree-Fock Limit



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XIII. Total Energies[a] of Some Simple Carbocations

	Literature	Step 3 4-31	Step 4 6-31	Step 5 6-31 + *d*	Step 6 6-31 + *d* + *p*
$CH_5^+(C_s)$	40.40393[c]	40.32715	40.36824	40.38822	40.40571
$CH_5^+(D_{3h})$	40.38065[c]	40.31571	40.35650	40.36763	40.38021
$C_2H_3^+$ (classical)	77.0563[d]	76.97973	77.06058	77.08661	77.09293
$C_2H_3^+$ (bridged)	77.0268[d]	76.94913	77.03006	77.07483	77.08382
$C_2H_5^+$(classical)	78.2422[e]	78.19852	78.27720	78.30766	78.31688
$C_2H_5^+$(classical)		78.19788	78.27675	78.30692	
$C_2H_5^+$ (bridged)	78.2279[e]	78.18680	78.26649	78.30703	78.31831

[a]Hartree units (negative signs omitted).

[c]Reference 55

[d]Reference 51

[e]Reference 52

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XIV.

Energy Differences (kcal/mole)

		Step 3	Step 4	Step 5	Step 6
CH_5^+	Fig 8 - Fig 7	7.2	7.4	12.9	16.0
$C_2H_3^+$	Fig 10- Fig 9	19.2	19.2	7.4	5.7
$C_2H_5^+$	Fig 12- Fig 11	0.4	0.3	0.5	
$C_2H_5^+$	Fig 13- Fig 11	7.3	6.7	0.4	-0.9

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

VII. THE EFFECT OF *d* FUNCTIONS ON THE GEOMETRIES OF THE FIRST ROW HYDRIDES.

One of the aims of non-empirical MO theory is the prediction of the equilibrium geometries of various chemical species. Reference has already been made to the systematic optimization of the geometries of hydrides containing one or two of the first-row atoms C, N, O and F. The species treated in the foregoing investigations were either neutral or cationic. There is ample evidence that, at the HF level, anions are unstable with respect to the disproportionation

$$A^- \longrightarrow A + e^- \qquad (7.1)$$

Kaplan and Kleiner[57] have shown that for most of the negative atomic ions for which HF calculations are available, the calculated wave functions differ from the ground HF wave functions in symmetry and by chemically significant energies. Anions derive their binding mostly from correlation and are therefore best treated by techniques which transcend the HF method. For example, Millie and Berthier,[54] in their study of the CH_3 system report the following energies (in hartrees): -39.246 (CH_3^+), -39.571 (CH_3), -39.512 (CH_3^-). It can be seen that CH_3^- is unstable relative to CH_3. However, these authors have estimated the correlation energies in the three cases to be -0.227, -0.263 and -0.373 so that, if their estimate is accepted, the total energies of the three species would be -39.473 (CH_3^+), -39.834 (CH_3) and -39.885 (CH_3^-). There is experimental evidence for the existence of anions, but the HF method is incapable of accounting for the binding of the extra electron.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

In this connection it must be remembered[14] that even the F_2 molecule is unbound relative to the atoms near the HF limit.

The present study forms part of the continuing effort in this laboratory to obtain near-HF wave functions for small organic molecules. This investigation is the first systematic attempt to assess the influence of basis set size and composition on the geometries of small polyatomic systems. Cade and Huo have carried out a thorough study of the diatomic hydrides using a large STO basis. Some of the polyatomic hydrides have been the subject of investigation by the one-center expansion method by Moccia,[58] Bishop[59] and Joshi.[60] More recently both gaussian and Slater basis sets have been applied to H_2O,[61,62] NH_3[44,63] and CH_4.[64,65]

The geometries of most of the hydrides of the first row elements are known and, with the advent of high-resolution photoelectron spectroscopy, structural information about more low-lying ionic states is becoming available. For these molecules we thus have both experimental information and other theoretical calculations for comparison. The present study is mainly directed at assessing the influence of polarization functions in heavy atom basis sets on calculated geometries. The 6-31G* basis introduced earlier has been used for all the calculations reported below. In the notation introduced in Section VI we have obtained the [4-4] geometries of the hydrides.

Table XV lists the energies at the [4-4] level for the systems reported here and, for comparison, includes the energies at the [1-1] and [2-2] levels also from reference 48. The geometries at the [4-4]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

level are compared with those at the [2-2] and [1-1] levels in Table XVI.

We now proceed to a discussion of each system in turn.

$\underline{CH^+, CH}$. The work of Cade and Huo[18] has established the ordering of the MO's to be

$$1\sigma<2\sigma<3\sigma<1\pi$$

CH^+ has just enough electrons to fill the first three MO's and therefore has a $^1\Sigma^+$ ground state. The additional electron in CH leads to a $^2\pi$ ground state. The ionization of CH to CH^+ causes the C-H bond to lengthen by .003 Å at the [4-4] level, the corresponding values at the [2-2] and [1-1] levels being .01 Å and .04 Å. The experimental result is closer to the [2-2] value. Cade and Huo[18] obtained a bond length of 1.104 Å for CH, the corresponding energy being -38.27958 hartrees.

$\underline{CH_2^+, CH_2}$. The CH_2 radical has been known for quite some time now. Herzberg[66,67] assigned a linear structure to the $^3B_1$ state of CH_2 but later[68] revised it to a bent form. The ESR results of Wasserman, *et al*[69,70] support a bent conformation for the $^3B_1$ state.

CH_2^+ is isoelectronic with BH_2 whose spectrum has been observed by Herzberg and Johns;[71] for its $^2A_1$ ground state they give the following structural parameters: r(B-H) = 1.18 Å, $\widehat{HBH}$ = 131°. According to Walsh's rules[72] isoelectronic systems should have similar geometries. Since there are no direct observations of ground state CH_2^+, the effect of the positive charge on the HCH angle is not experimentally known. In the isoelectronic system NH_2-OH_2^+ ($^2B_1$) the angle opening in the cation is known (see below) to be ∿7°. Bender and Schaefer[73] have

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

optimized the geometry of CH_2^+ at both the SCF and configuration interaction (CI) level. Their SCF bond length of 1.083 Å increased to 1.107 Å when CI was used and there was a small opening of the HCH angle from 139.7° (SCF) to 140.3°. The energies they report are: -38.55686 (SCF) and -38.61459 (CI). Their SCF energy is higher than the [4-4] value because the basis they used lacked polarization functions. Bender and Schaefer also carried out parallel calculations on BH_2, obtaining r (B-H) = 1.21 Å and $H\hat{B}H$ = 129°. It would have been interesting to have included calculations on $^2A_1$ BH_2 in this work; unfortunately we do not have a 6-31G* basis for Boron yet.

The $^2A_1$ state of CH_2^+ is obtained by removing an electron from the $3a_1$ orbital of $^1A_1$ CH_2. If yz is the molecular plane and z is the two-fold axis, then the $3a_1$ orbital is mainly a $2p_z$ orbital centered on the carbon atom, but includes the bonding combination of the hydrogen 1s orbitals. Depleting its population would therefore tend to open the angle compared to the parent $^1A_1$ CH_2.

Three states, $^3B_1$, $^1B_1$ and $^1A_1$ of the CH_2 radical have been observed spectroscopically. The $^3B_1$-$^1A_1$ separation is not known experimentally, but is thought to be ∿1 eV. Carr, Eder and Toper[74] propose a much smaller value of 12 kcal/mole based on the 3500 Å photolysis of ketene. Since the 'bending' $3a_1$ orbital is occupied by a pair of electrons in the $^1A_1$ state we expect the bond angle in this state to be fairly small. The B_1 states are obtained by promoting an electron from the $3a_1$ orbital to the $1b_1$ orbital (following Walsh[72] we define the b_1 orbital to be the p

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

orbital on the central atom which is perpendicular to the molecular plane). Since the $1b_1$ orbital is nonbonding it is not expected to have much influence on the geometry, but because an electron has been removed from the 'bending' $3a_1$ orbital, the B_1 states will have larger bond angles; experimentally, the opening is $\sim 30°$. As expected, the bond angle in the B_1 states of CH_2 are fairly close to the angle in $^2A_1$ CH_2^+. At the [4-4] level the extra opening due to the positive charge is 8°.

The three states of CH_2 mentioned above have been the subject of several theoretical studies. For the $^3B_1$ state, O'Neil, Schaefer, and Bender,[75] using a $(9s^4 5p^2/4s^2)$ basis, obtained an energy of -38.9136 hartrees at the optimum geometry characterized by r(C-H) = 1.075 Å, $\hat{HCH}$ = 130.4°. The same basis, with CI, yielded an energy of -38.9826 hartrees, the new geometrical parameters being r(C-H) = 1.095 Å and $\hat{HCH}$ = 133.3°. Using the Generalized Valence Bond (GVB) method, Hay, Hunt and Goddard[76] found the lowest energy form of $^3B_1$ CH_2 had $\hat{HCH}$ = 135°, r(C-H) = 1.11 Å. The energy of this form, -38.9483, fell slightly to -38.9598 with CI. The basis used by these workers had polarization functions on carbon. The experimental value of $\hat{HCH}$ = 136° is an ESR value. The ultra-violet (U-V) spectral results would support assignments between 128° and 148° if the standard deviation of the rotation constant B were taken into account.[68] For the $^1A_1$ state O'Neil, *et al* obtained an energy of -38.8620 hartrees for r(C-H) = 1.103 Å, $\hat{HCH}$ = 106.5°. When CI was used the bond length increased to 1.133 Å and there was a slight closing of the angle to 104.4°. At this geometry they obtained an energy of -38.9472 hartrees. The singlet-triplet separation at the [4-4] level, 30.83 kcal/mole, is clearly larger than the estimate of Carr, *et al*.[74]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

The ionization potential at the [4-4] level (calculated as the $^3B_1$ CH_2 - $^2A_1$ CH_2^+ difference) is 9.67 eV, the experimental value being 10.396 eV.[68]

CH_3^+, CH_3. The electronic configuration of CH_3 in the planar and pyramidal forms can be written as

$$D_{3h} \quad (1a_1')^2(2a_1')^2(1e')^4(1a_2'') \qquad {}^2A_2''$$

$$C_{3v} \quad (1a_1)^2(2a_1)^2(1e)^4(3a_1) \qquad {}^2A_1$$

The orbitals $2a_1'$ and $1e'$ describe the three C-H bonds. $1a_2''$ is a p orbital on the central atom with its node in the molecular plane. In C_{3v} symmetry the $1a_2''$ orbital becomes $3a_1$ and acquires more and more s character as the bending angle increases. Populating the $1a_2''$ orbital should therefore favor a non-planar structure. However, experimental evidence favors the planar form overwhelmingly. Andrews and Pimentel[77] obtained the methyl radical in solid argon using the matrix reaction of Li atoms and methyl halides at 15°K. They deduce that if the radical is non-planar, the out-of-plane angle is less than 5°. Herzberg,[66] from an analysis of the u-v spectrum, has concluded that the deviation from planarity cannot exceed 10°. ESR spectra too point to a planar structure. NH_3 is well known to be pyramidal in its ground state, but in its lowest excited state ($\tilde{A}$ of A_2'' symmetry) in which the $1a_2''$ orbital contains only one electron (thus resembling the CH_3 radical) Douglas[78] has found it to be planar. Since the pyramidalizing $1a_2''$ orbital is vacant in CH_3^+ the cation is certainly planar. We find both the cation and the radical to be planar. The best values reported in the literature for these two species are by Millie and Berthier.[54] For CH_3 and CH_3^+ they obtained planar forms

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

while holding the bond length fixed at the experimental value of 1.079 Å. The energies were -39.57148 hartrees and -39.24592 hartrees respectively for CH_3 and CH_3^+.

CH_4. This molecule was examined only under tetrahedral symmetry. For r(C-H) = 1.094 Å Clementi and Popkie[64] have obtained an energy of -40.2136 hartrees.

NH^+, NH. Experimentally the ground state of NH^+ has been identified[79] as $^2\pi$ corresponding to the configuration $(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi)$. From a study of the isotopic species $^{14}NH^+$, $^{15}NH^+$ and $^{14}ND^+$ Colin and Douglas have established that a $^4\Sigma^-$ state lies 354 cm^{-1} above the $^2\pi$ state. At the [4-4] level, as at the [1-1] and [2-2] levels, the $^4\Sigma^-$ state lies below the $^2\pi$ state. Since the experimental separation between the two states is small, correlation energy differences could account for the reversal. Of the three states $^3\Sigma^-$, $^1\Delta$ and $^1\Sigma^+$ arising from the configuration $(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi)^2$ of NH, we have examined only the first. Cade and Huo[18] obtain r(N-H) = 1.018 Å and E = -54.97838 hartrees.

NH_2^+, NH_2. There is no direct experimental information on NH_2^+ which is isoelectronic with CH_2. But Gassman[80] has pointed out that nitrenium ion chemistry is similar to carbonium ion chemistry. The structure should be similar to that of $^3B_1$ CH_2 with a slightly larger bond angle. Chu, Siu and Hayes[81] find that with an *sp* basis the angle in NH_2^+ is greater by ~10° over that in CH_2.

NH_2 was the first triatomic radical to be observed experimentally. Both the ground $^2B_1$ and first excited $^2A_1$ states have been characterized by optical or u-v spectroscopic techniques.[82,83] The ground state

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is definitely known to be bent, but the upper state ($^2A_1$) was thought to be nearly linear. More recently Dixon[84] has deduced the bond angle in the $^2A_1$ state to be 144° ± 5°. Accurate calculations by Bender and Schaefer,[85] both at the SCF and CI levels, with a basis containing polarization functions, confirms these findings. The energies and geometry parameters obtained by these workers are: E ($^2B_1$, SCF) = -55.5757 hartrees, r(N-H) = 1.019 Å, $\hat{HNH}$ = 105.4°; E ($^2A_1$, SCF) = -55.5237 hartrees, r(N-H) = 0.997 Å, $\hat{HNH}$ = 141.9°; E ($^2B_1$, CI) = -55.6799 hartrees, r(N-H) = 1.055 Å, $\hat{HNH}$ = 102.7°; E ($^2A_1$, CI) = -55.6185 hartrees, r(N-H) = 1.01 Å, $\hat{HNH}$ = 144.7°. The ground state of NH_2 has the configuration $^2B_1$ $(1a_1)^2$ $(2a_1)^2$ $(2b_2)^2$ $(3a_1)^2$ $(1b_1)$. $3a_1$ is bonding between the two hydrogen atoms whereas $1b_1$, a p orbital on N perpendicular to the molecular plane, is non-bonding. In the first excited state, the configuration is $^2A_1$ $(1a_1)^2$ $(2a_1)^2$ $(2b_2)^2$ $(3a_1)$ $(1b_1)^2$. Since an electron has been removed from $3a_1$ the bond angle increases in the $^2A_1$ state. The experimental separation between the $^2B_1$ and $^2A_1$ states is ∿29 kcal/mole; the [4-4] value is 35.21 kcal/mole.

NH_3^+, NH_3. Being isoelectronic with CH_3, NH_3^+ should be planar. Its ESR spectrum[86,87] is consistent with a planar structure. NH_3 was examined in both pyramidal and planar forms to obtain the inversion barrier. Rauk, Allen and Clementi[44] showed that with a large enough basis the barrier could be calculated in the HF approximation. The structures and energies obtained by them are: E(C_{3v}) = -56.2219 hartrees, r(N-H) = 1.000 Å, $\hat{HNH}$ = 107.2°; E(D_{3h}) = -56.2138 hartrees, r(N-H) = 0.984 Å. It may be noted that the bond shortening in the planar form

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is fortuitously the same at level [4-4] as in the above-mentioned calculation.[88] The experimental value of the inversion barrier in NH_3 is 5.77 kcal/mole. Rauk *et al* obtain 5.08 kcal/mole while at level [4-4] it is 6.5 kcal/mole. Pipano, Gilman, Bender and Shavitt[88] found the barrier to be 3.8 mh at SCF, but this increased to 8.3 mh with CI, just 0.9 mh less than the experimental value. They therefore conclude that correlation corrections to the barrier cannot be ignored. Kari and Csizmadia[30] also have carried out both SCF and CI calculations on the two forms of NH_3. Their SCF barrier of 16.1 mh was reduced to 11.1 mh by CI. To this extent it bears out the statement of Pipano, *et al*, but it must be noted that the correlation correction has a different sign. Stevens[63] has obtained an SCF inversion barrier of 9.4 mh (the closest to date to the experimental value) with a large ST basis and concludes that the inversion barrier can be accurately predicted in the HF approximation provided the basis is close to being complete.

One reason for the discrepancy between the calculations of Pipano, Gilman, Bender and Shavitt and Kari and Csizmadia is the absence of *d* functions in the basis used by the former group. Ratner and Sabin[31] point out that *s* and *p* functions do not provide a basis for the irreducible representation E" of D_{3h}. If configurations in which MO's of this symmetry are populated are important in CI then *d* functions are *essential* in the basis.

OH, OH^+. OH^+ is isoelectronic with NH and its ground state, $^3\Sigma^-$, has the configuration $(1\sigma)^2\ (2\sigma)^2\ (3\sigma)^2\ (1\pi)^2$. The neutral hydroxyl radical, whose ground state is $^2\pi$, has a shorter bond. The lengthening on

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

ionization at level [4-4], 0.055 Å, is close to the experimental value, 0.058 Å. The near-HF results of Cade and Huo for OH, r(O-H) = 0.95 Å and E = -75.42127 hartrees, again indicate that the SCF bond length is smaller than the experimental value.

OH_2^+, OH_2. The electronic configuration of the ground state water molecule is $(1a_1)^2\ (2a_1)^2\ (1b_2)^2\ (3a_1)^2\ (1b_1)^2$. The b_1 orbital is antisymmetric with respect to reflection in the molecular plane. The three lowest ionized states of water that can be formed from this configuration are:

$$^2B_1 \quad (1a_1)^2\ (2a_1)^2\ (1b_2)^2\ (3a_1)^2\ (1b_1)$$
$$^2A_1 \quad (1a_1)^2\ (2a_1)^2\ (1b_2)^2\ (3a_1)\ (1b_1)^2$$
$$^2B_2 \quad (1a_1)^2\ (2a_1)^2\ (1b_2)\ (3a_1)^2\ (1b_1)^2$$

$2a_1$ and $1b_2$ are the main bonding orbitals. When the H-O-H angle is increased to 180°, both $3a_1$ and $1b_1$ go to the degenerate $1\pi_u$ orbitals of the linear conformation. If, as in H_2O^+, there are not enough electrons to fill the π orbitals, then the system is susceptible to the *Renner-Teller* effect, an example of the breakdown of the Born-Oppenheimer approximation: a bending vibration of the linear form induces a bent geometry which may be stabler than the linear form with orbital degeneracy. Pople and Longuet-Higgins[90] have pointed out the three different kinds of potential curves which can result and have thoroughly discussed the case in which the lower curve has a minimum for a non-linear geometry. This discussion was then applied by them to the $^2B_1$ state of NH_2. Dixon[84]

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

has interpreted the $^2A_1$ state of NH_2 as belonging to the third type of potential in which the upper curve also has a minimum for a non-linear geometry, but for smaller departure from linearity than for the $^2B_1$ state. In general, the Renner-Teller effect, in breaking up the degeneracy of the linear form, produces a highly bent state and a nearly linear one. Since H_2O^+ is isoelectronic with NH_2, the Renner-Teller effect may be expected to be operative for its states too. The photoelectron spectrum of water has been obtained by Brundle and Turner;[91] Asbrink and Rabalais[92] and by Potts and Price;[93] three bands, showing 'extensive, deceivingly discernible vibrational fine structure', have been noted. On the basis of the electronic configuration given for H_2O the three bands have been assigned to the $^2B_1$, $^2A_1$ and $^2B_2$ states. Brundle and Turner[91] have obtained ionic parameters for the $^2B_1$ state by fitting calculated Franck-Condon factors to bond angle and distance. The $1b_1$ orbital is non-bonding; ionization of an electron from this orbital should not result in extensive geometry changes. In the photoelectron spectrum the most intense transition is between the 0, 0, 0 levels of the two states ($^2B_1 \leftarrow {}^1A_1$). The experimental opening of the angle from $^2B_1$ NH_2 to $^2B_1$ OH_2^+ is 6.6°; at level [4-4] it is calculated to be 7.8°. High resolution photoelectron spectra of the $^2A_1 \leftarrow {}^1A_1$ and $^2B_2 \leftarrow {}^1A_1$ transitions have been obtained by Rabalais. The $^2B_2$ state predissociates by curve crossing with the $^4A''$ state (C_s symmetry) to OH and the lines are therefore broadened. The departure from linearity of the $^2A_1$ state is far smaller in OH_2^+ than it is in NH_2 at the [4-4] level.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

OH_3^+. The hydronium ion is well known to flame chemists, being the most abundant ion in hydrocarbon flames.[94,95] Broad-line NMR spectra of crystalline monohydrates of HNO_3, $HClO_4$ and H_2SO_4 have been interpreted by Richards and Smith[96] to point to the probable presence of OH_3^+ in these solids. In the nitric acid monohydrate the interatomic distance was found to be 1.72±.02Å which has been taken by them as being consistent with a pyramidal shape for OH_3^+ with OH bond lengths of 1.02Å. In a later study of the crystal structure of the monohydrate of HCl at -35°C Yoon and Carpenter[97] find that OH_3^+ is pyramidal with a bond angle of ∿117° and a bond length of 0.96 ± 0.08 Å. This ion is isoelectronic with NH_3 and should therefore be pyramidal with $\hat{HOH}$ ∿ 114° (opening of ∿8° from NH_3 extrapolated from the sequence NH_2-OH_2^+). Extended basis sets which do not contain polarization functions predict the ground state of this ion to be planar. Both Moskowitz and Harrison[98] and Diercksen and Kraemer,[99] who have carried out calculations with basis sets comparable to 6-31G*, appear to have optimized the geometry *without* the polarization functions and obtain a planar geometry. One-center studies with basis functions of high angular quantum number, on the other hand, support our finding of a pyramidal conformation.[58,59,60] The experimental situation has been admirably summarized by Bishop.[59]

FH, FH_2^+: In line with the foregoing systems, the FH bond length is found to be shorter than the experimental one. FH_2^+ is isoelectronic with H_2O and has a similar geometry. The bond angle (at level [4-4]) has opened by 8.4° from H_2O, slightly larger than the opening in the sequence NH_2-OH_2^+.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Inspection of Table XVI shows that, in the majority of cases, bond lengths at level [4-4] are smaller than those at level [2-2]. It has been remarked in the literature that bond lengths obtained using minimal basis sets are often close to the experimental values. Extending the basis to multiple-zeta levels, and adding polarization functions, results in shorter bond lengths. Cade and Huo[18] found that the calculated bond lengths of six of the seven first row hydrides considered by them were below the experimental values, the mean deviation (for all seven cases) being 0.016 Å. At the minimal STO-3G level, bonds were mostly predicted to be too long with a mean absolute deviation of 0.023 Å. For the extended 4-31G basis, the bonds were too short with a mean absolute deviation of 0.01 Å. At level [4-4] bond lengths are, as mentioned above, mostly shorter than those at level [2-2], but better agreement with experiment is obtained for the following cases: $^3B_1$, CH_2, CH_3, CH_4, $^2A_1$, NH_2, and NH_3. The mean absolute deviation is reduced to 0.008 Å. The work of Cade and Huo on the diatomic hydrides, the study of NH_3 by Rauk, *et al*, and the study of H_2O by Dunning, *et al*, definitely shows that as the HF limit is more closely approached the bond lengths become shorter. In the case of HF, Nesbet[13] has pointed out that configuration interaction due to the HF bond antibonding orbital will increase as the molecule dissociates; the correlation energy therefore increases with internuclear separation near equilibrium, implying that the Hartree-Fock estimate of the equilibrium bond length should be somewhat smaller than the experimental value.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Deviations in bond angles are not so systematic. At level [1-1] bond angles are too small (mean deviation of 5.3°) whereas the 4-31G values are too large (mean deviation of 4.1°). The mean bond angle deviation for the seven species $^1A_1$ CH_2, $^3B_1$ CH_2, $^2A_1$ NH_2, $^2B_1$ NH_2, NH_3, $^2B_1$ OH_2^+ and OH_2 is 1.8° at level [4-4]. Table XVII lists the geometry of H_2O optimized using various basis sets by Dunning, Pitzer and Aung[61] and in this laboratory. The most striking feature of this table is the rapid movement of the calculated bond angle towards the experimental value when *d*-type polarization functions are added to the heavy atom basis set. Even though the [4-4] energy is 2 mh above the energy in the $(9s^4 5p^3/4s^2)$ basis, the bond angle is much closer to the experimental value. *Polarization functions are therefore essential for proper calculation of bond angles.* NH_3 furnishes a less dramatic demonstration of the same effect. This is not surprising in view of the fact that polarization functions make it possible for AO's to bend.

Dunning, Pitzer and Aung used two sets of *d* functions on oxygen in their optimization of the geometry of H_2O. Our studies reveal, however, that one set is adequate. Lehn and Munsch[100] have concluded from their calculations on PH_3 that two sets of *d* functions on the heavy atom are needed for obtaining good geometries: a compact set to reproduce bond lengths and an expanded set to reproduce bond angles. It must be remembered that more *p* functions are needed for second-row atoms than for the corresponding first row atoms. For first row atoms, however, our studies have not revealed the need for a second set of *d* functions.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Walsh's rules[72] are still found useful for understanding the shapes of simple polyatomic molecules despite the fact that, in his discussions, polarization was not taken into account.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV. Optimized Energies of First-Row Hydrides at Various Levels

Molecule	Point Group Symmetry	State	Energy [1-1]	[2-2]	[4-4]
CH^+	$C_{\infty v}$	$^1\Sigma^+$	-37.45638	-37.83988	-37.89554
CH	$C_{\infty v}$	$^2\Pi$	-37.77026	-38.20801	-38.26493
CH_2^+	C_{2v}	$^2A_1$	-38.11894	-38.51257	-38.56619
CH_2	C_{2v}	$^1A_1$	-38.37230	-38.81035	-38.87237
		$^3B_1$	-38.43623	-38.86963	-38.92150
CH_3^+	D_{3h}	$^1A_1'$	-38.77948	-39.17512	-39.23064
CH_3	D_{3h}	$^2A_2''$	-39.07671	-39.50497	-39.55899
CH_4	T_d	$^1A_1$	-39.72686	-40.13977	-40.19517
NH^+	$C_{\infty v}$	$^2\Pi$	-53.83593	-54.41055	-54.48706
	$C_{\infty v}$	$^4\Sigma^-$	-53.88021	-54.45593	-54.51962
NH	$C_{\infty v}$	$^3\Sigma^-$	-54.26378	-54.88495	-54.95942
NH_2^+	$C_{\infty v}$	$^3B_1$	-54.55042	-55.13968	-55.20852
NH_2	C_{2v}	$^2A_1$	-54.76094	-55.43244	-55.50159
		$^2B_1$	-54.83930	-55.47473	-55.55770
NH_3^+	D_{3h}	$^2A_2''$	-55.20701	-55.80167	-55.87323
NH_3	C_{3v}	$^1A_1$	-55.45542	-56.10669	-56.18434
	D_{3h}	$^1A_1'$	-55.43767	-56.10600	-56.17398
OH^+	$C_{\infty v}$	$^3\Sigma^-$	-76.04893	-76.87323	-74.96875
OH	$C_{\infty v}$	$^2\Pi$	-76.36489	-75.28716	-75.38228

continued--

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XV (continued)

OH_2^+	C_{2v}	$^2A_1$	-74.62473	-75.49268	-75.57906
	C_{2v}	$^2B_1$	-74.66974	-75.51285	-75.61530
OH_2	C_{2v}	$^1A_1$	-74.96590	-75.90864	-76.01075
OH_3^+	C_{3v}	$^1A_1$	-75.33044	-76.20060	-76.28656
FH	$C_{\infty v}$	$^1\Sigma^+$	-98.57285	-99.88729	-100.00291
FH_2^+	C_{2v}	$^1A_1$	-98.86411	-100.07787	-100.19782

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI. Optimized Geometries of First-Row Hydrides at Various Levels

Molecule	Point Group Symmetry	State	Geometric[a] Parameter	[1-1]	[2-2]	[4-4]	Exp.[b]
CH^+	$C_{\infty v}$	$^1\Sigma^+$	r(C-H)	1.185	1.108	1.105	1.131
CH	$C_{\infty v}$	$^2\Pi$	r(C-H)	1.143	1.118	1.108	1.120
CH_2^+	C_{2v}	$^2A_1$	r(C-H)	1.132	1.079	1.081	
			$\hat{HCH}$	136.0	141.8	138.8	
CH_2	C_{2v}	$^1A_1$	r(C-H)	1.123	1.100	1.096	1.11
			$\hat{HCH}$	100.5	105.4	103.0	102.4
	C_{2v}	$^3B_1$	r(C-H)	1.082	1.069	1.071	1.078
			$\hat{HCH}$	125.5	132.0	130.8	136.0
CH_3^+	D_{3h}	$^1A_1'$	r(C-H)	1.120	1.076	1.078	
CH_3	D_{2h}	$^2A_2''$	r(C-H)	1.078	1.070	1.073	1.079
CH_4	T_d	$^1A_1$	r(C-H)	1.083	1.081	1.084	1.085
NH^+	$C_{\infty v}$	$^2\Pi$	r(N-H)	1.140	1.049	1.046	
	$C_{\infty v}$	$^4\Sigma^-$	r(N-H)	1.136	1.060	1.060	
NH	$C_{\infty v}$	$^3\Sigma^-$	r(N-H)	1.082	1.033	1.024	1.0481
NH_2^+	C_{2v}	$^3B_1$	r(N-H)	1.079	1.020	1.019	
			$\hat{HNH}$	147.4	157.5	151.5	
NH_2	C_{2v}	$^2A_1$	r(N-H)	1.015	0.985	0.988	1.004
			$\hat{HNH}$	131.3	144.9	141.8	144.0
	C_{2v}	$^2B_1$	r(N-H)	1.058	1.015	1.013	1.024
			$\hat{HNH}$	100.2	108.3	104.3	103.4
NH_3^+	D_{3h}	$^2A_2''$	r(N-H)	1.056	1.010	1.012	

continued--

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVI (continued)

NH_3	C_{3v}	$^1A_1$	r(N-H)	1.033	0.991	1.004	1.012
			HÑH	104.2	115.8	107.5	106.7
	D_{3h}	$^1A_1'$	r(N-H)	1.006	0.986	0.988	
OH^+	$C_{\infty v}$	$^3\Sigma^-$	r(O-H)	1.084	1.023	1.013	1.0289
OH	$C_{\infty v}$	$^2\Pi$	r(O-H)	1.014	0.968	0.958	0.9706
OH_2^+	C_{2v}	$^2A_1$	r(O-H)	1.014	0.987[c]	0.985	
			HÔH	162.2	180.0	176.0	
	C_{2v}	$^2B_1$	r(O-H)	1.033	0.991	0.987	1.00
			HÔH	109.8	119.9	112.1	110.0
OH_2	C_{2v}	$^1A_1$	r(O-H)	0.990	0.951	0.948	0.957
			HÔH	100.0	111.2	105.5	104.5
OH_3^+	C_{3v}	$^1A_1$	r(O-H)	0.990	0.964	0.969	
			HÔH	113.9	120.0	113.2	
FH	$C_{\infty v}$	$^1\Sigma$	r(F-H)	0.956	0.922	0.911	0.917
FH_2^+	C_{2v}	$^1A_1$	r(F-H)	0.974	0.969	0.962	
			HF̂H	112.0	125.5	113.9	

[a] Distances in Angstroms, angles in degrees

[b] Unless otherwise noted we use the same references as Ref. 48.

[c] Point group symmetry, $C_{\infty v}$, state $^2\Pi$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

TABLE XVII. Geometry of H_2O as the Hartree-Fock Limit is Approached

	$(2s1p/1s)$ STO-3G	$(2s1p/1s)^{a,b}$	$(8s^3 4p^2/4s^2)$ 4-31G	$(10s^3 4p^2 1d/4s^2)$ 6-31G*
r(O-H) Å	0.990	0.990	0.951	0.948
HOH degrees	100.0	100.3	111.2°	105.5°
-E (hartrees)	74.96590	75.70545	75.90864	76.01075

TABLE XVII (continued from above)

	$(9s^4 5p^3/4s^2)^b$	$(9s^4 5p^3 2d^2/4s^2 1p)^b$	$(11s^6 7p^5 2d^2/5s^3 1p)^b$
r(O-H) Å	0.951	0.941	0.941
HOH degrees	112.6°	106.1°	106.6°
-E(hartrees)	76.01271	76.05103	76.06283

[a] Slater type orbitals

[b] Reference 61

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

VIII. SUMMARY AND CONCLUSIONS

It has been shown in the preceding sections that, in order to reach the Hartree-Fock limit even for molecules containing only first row atoms, provision must be made for taking into account the polarization of the AO's when a molecule is formed. The polarization can be allowed for in two ways: either by letting the AO's 'float' from the nuclei on which they are centered, or by including in the AO basis, functions of higher angular quantum number than are occupied in the ground states of the constituent atoms. The latter method has been the more popular and studies by various groups of workers has revealed that p functions account for the major part of the polarization around hydrogen and the addition of d functions to the basis of first row atoms explains the principal polarization around these atoms in their molecules.

Computationally efficient basis sets, useful for calculations on a large number of organic molecules, have been developed and applied uniformly and systematically to a group of molecules containing either one or two of the atoms C, N, O and F in addition to hydrogen. This has revealed that the energy lowering due to the addition of d functions depends on the *local* symmetry around the heavy atom. Similarly, the energy lowering due to the addition of p functions in hydrogen is sensitive to the nature and number of atoms the hydrogen is bonded to.

Consideration of energies of reactions involving closed shell species has shown that overall improvement of the predicted energies occurs when polarization functions are present in the basis. Based on hydrogenation energy errors at the Hartree-Fock limit it has been

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

conjectured that the correlation energies of double and triple bonds would be twice and thrice respectively that of the corresponding single bond.

Polarization functions have been found to be important in geometries involving three-center bonds. They are also shown to be indispensable for predicting proper values of bond angles in the Hartree-Fock scheme.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

APPENDIX: EXPLANATION OF SOME TERMS USED IN LCAO-MO THEORY

There are two principal types of basis sets in use in LCAO-MO theory. Slater [101] introduced the exponential form

$$S(\zeta) \propto \exp(-\zeta r) \qquad (A.1)$$

patterned after the solutions of the hydrogen-like atoms. Till about 1965, these were the most widely used functions in MO calculations. The principal difficulty with using ST functions is that there is no exact expression for electronic repulsion integrals over more than two centers and one has to go through tedious expansions, or transforms, to approximate the integral. Boys [35] circumvented this difficulty by introducing the gaussian function

$$G(\alpha) \propto \exp(-\alpha r^2) \qquad (A.2)$$

for which he showed that all the integrals needed for calculating the energy could be obtained in closed form. Gaussians, however, did not become popular till the mid-sixties when programs like POLYATOM, IBMOL, MOSES, etc. began to become available.

Basis sets can classed as minimal or extended, whatever be the nature of the constituent functions [102]. A basis is called minimal if to each orbital of an atom (in its ground state) there is just one function of the appropriate symmetry in the basis. Thus, the minimal basis for hydrogen contains only one 1s function. The minimal basis for carbon has one 1s function, one 2s (or corresponding) function and three 2p functions. In general, when orbitals for which $\ell \neq 0$ are occupied in the ground atomic state, the full set of $(2\ell + 1)$ functions

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

is present in the minimal basis. In an extended basis, on the other hand, more than one function in the basis will correspond to a given atomic orbital. Thus, an extended basis for hydrogen will have two 1s functions, generally denoted 1s and 1s', in it. The 4-31G and 6-31G basis sets are therefore extended basis sets of the 'split-valence' kind.

More adjectives have been coined to describe specific kinds of extended basis sets. Thus, there are two functions in the 'double zeta' basis for every occupied orbital of an atom. The basis sets used in the diatomic calculations of the Chicago group are of the 'multiple zeta' type. The terminology thus far introduced was originally devised for Slater basis sets, but apply equally to gaussian basis sets.

Two more terms have come into vogue with the development of gaussian basis sets. The gaussian (A.2) is called a 'primitive' gaussian

The number of GT functions required to obtain a given energy is about twice the number of ST functions. Since the number of two-electron integrals in a given LCAO-MO calculation is proportional to n^4 , where n is the number of functions in the basis, doubling the size of a basis increases the number of two-electron integrals by a factor of sixteen. When the Fock matrix is formed, each two-electron integral has to be processed and, in an SCF procedure, the cycle is repeated to convergence. The time-consuming portion of the SCF cycle is the formation of the Fock matrix since, in general, this involves an input operation and a number

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

of multiplications and additions. If the number of two-electron integrals can be reduced, there can be a substantial saving of time in the calculation. One way of achieving this reduction is to group a set of primitive gaussians together and manipulate them as a single function. If $\{\varphi_i\}$ is a set of primitive gaussians, then a contracted set may be obtained from them

$$\chi_j = \sum_i c_{ji} \varphi_i \qquad (A.3)$$

c_{ji} are called the contraction coefficients. If the primitive set has n functions and the contracted set has m functions ($m < n$), then contraction reduces the number of two-electron integrals to be manipulated by $(m/n)^4$. Shavitt [103] was the first to note the advantage of contraction, but practical realization came much later. Early efforts concentrated on obtaining the contraction coefficients solely by calculations on atoms, but it is known that these are inadequate for molecules. Dunning [104] has given a comprehensive discussion and has formulated rules for optimum contraction schemes for molecules. In principle, contraction is applicable to ST basis sets also. It should be noted that even though there is a reduction in the total number of two-electron integrals (expressed over contracted functions) these must still be obtained over the primitives whose number is the same as before. There is therefore no special saving in time when the integrals are being calculated.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

REFERENCES

1. S. Wang, Phys.Rev. 31, 579 (1928).
2. N. Rosen, Phys.Rev. 38, 2099 (1931).
3. B. N. Dickinson, J.Chem.Phys. 1, 317 (1933).
4. B. N. Finkelstein and G. E. Horowitz, Z.Physik. 48, 118 (1928).
5. E. F. Gurnee and J. L. Magee, J.Chem.Phys. 18, 142 (1950).
6. A. C. Hurley, Proc.Roy.Soc. (London) A226, 179 (1954); see also pages 170 and 193 of same reference.
7. J. O. Hirschfelder and J. W. Linnett, J.Chem.Phys. 18, 130 (1950).
8. P. C. Hariharan and J. A. Pople, Chem.Phys.Lett. 16, 217 (1972).
9. See report from a correspondent in Nature (London) 217, 220 (1968).
10. C. A. Coulson, Nature (London) 221, 1106 (1969).
11. R. S. Mulliken and B. Liu, J.Amer.Chem.Soc. 93, 6738 (1971).
12. R. S. Mulliken, J.Chem.Phys. 36, 3428 (1962).
13. R. K. Nesbet, J.Chem.Phys. 36, 1518 (1962).
14. A. C. Wahl, J.Chem.Phys. 41, 2600 (1964).
15. W. M. Huo, J.Chem.Phys. 43, 624 (1965).
16. P. E. Cade, K. D. Sales, and A. C. Wahl, J.Chem.Phys. 44, 1973 (1966).
17. R. K. Nesbet, J.Chem.Phys. 40, 3619 (1964).
18. P. E. Cade and W. M. Huo, J.Chem.Phys. 47, 614 (1967).
19. D. B. Boyd and W. N. Lipscomb, J.Chem.Phys. 46, 910 (1967).
20. F. P. Boer and W. N. Lipscomb, J.Chem.Phys. 50, 989 (1969).
21. T. H. Dunning, Jr., J.Chem.Phys. 55, 3958 (1971).
22. B. Roos and P. Siegbahn, Theor.Chim.Acta 17, 199 (1970).
23. J. A. Pople and M. S. Gordon, J.Amer.Chem.Soc. 89, 4253 (1967).
24. W. J. Hehre, R. Ditchfield and J. A. Pople, J.Chem.Phys. 56, 2257 (1972).
25. H. Basch, M. B. Robin and N. A. Kuebler, J.Chem.Phys. 47, 1201 (1967).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

26. E. Clementi and H. Popkie, J.Chem.Phys. 57, 1077 (1972).

27. D. Neumann and J. W. Moskowitz, J.Chem.Phys. 49, 2056 (1968).

28. P. C. Hariharan and J. A. Pople, 'The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies' (accepted for publication in Theor.Chim.Acta).

29. S. Rothenberg and H. F. Schaefer III, J.Chem.Phys. 54, 2764 (1971).

30. R. E. Kari and I. G. Csizmadia, J.Chem.Phys. 56, 4337 (1972).

31. M. A. Ratner and J. R. Sabin, J.Amer.Chem.Soc. 93, 3342 (1971).

32. G.H.F. Diercksen, Theor.Chim.Acta 21, 335 (1971).

33. E. D. Hankins, J. W. Moskowitz and F. H. Stillinger, J.Chem.Phys. 53, 4544 (1970).

34. See Reference 21, especially Reference 2 therein.

35. S. F. Boys, Proc.Roy.Soc. (London) A200, 542 (1950).

36. See, for example, J. M. Foster and S. F. Boys, Rev.Mod.Phys. 32, 303 (1960) and references therein.

37. S. Huzinaga, Supp.Prog.Theor.Phys., Extra No. 40, 52 (1967).

38. See, for example, R. M. Stevens, J.Chem.Phys. 52, 1397 (1970).

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

40. L. C. Snyder, J.Chem.Phys. 46, 3602 (1967); 'Thermochemistry and the Electronic Structure of Molecules', the Robert A. Welch Foundation Research Bulletin No. 29 (Houston, Texas, August, 1971).

41. L. C. Snyder and H. Basch, J.Amer.Chem.Soc. 91, 2189 (1969).

42. W. J. Hehre, R. Ditchfield, L. Radom and J. A. Pople, J.Amer.Chem.Soc. 92, 4796 (1970).

43. L. Radom, W. J. Hehre and J. A. Pople, J.Amer.Chem.Soc. 93, 289 (1971).

44. A. Rauk, L. C. Allen and E. Clementi, J.Chem.Phys. 52, 4133 (1970).

45. E. Clementi, Proc.Nat.Acad.Sci. USA 69, 2942 (1972).

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

47. R. Ditchfield, W. J. Hehre and J. A. Pople, J.Chem.Phys. 52, 5001 (1970).

48. W. A. Lathan, W. J. Hehre, L. A. Curtiss and J. A. Pople, J.Amer.Chem. Soc. 93, 6377 (1971).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

49. W. A. Lathan, L. A. Curtiss, W. J. Hehre, J. B. Lisle and J. A. Pople, 'Molecular Orbital Structures for Small Organic Molecules and Cations', to be published in *Progress in Physical Organic Chemistry*.

50. P. C. Hariharan, W. A. Lathan and J. A. Pople, Chem.Phys.Lett. 14, 385 (1972).

51. A. C. Hopkinson, K. Yates and I. G. Csizmadia, J.Chem.Phys. 55, 3835 (1971).

52. R. Sustmann, J. E. Williams, M.J.S. Dewar, L. C. Allen and P. v. R. Schleyer, J.Amer.Chem.Soc. 91, 5350 (1969).

53. W. A. Lathan, W. J. Hehre and J. A. Pople, J.Amer.Chem.Soc. 93, 808 (1971).

54. P. Millie and G. Berthier, Inter.J.Quantum Chem. IIS, 67 (1968).

55. V. Dyczmons, V. Staemmler and W. Kutzelnigg, Chem.Phys.Lett. 5, 361 (1970).

56. M. E. Schwartz and L. J. Schaad, J.Chem.Phys. 47, 5325 (1967).

57. T. A. Kaplan and W. H. Kleiner, Phys.Rev. 156, 1 (1967).

58. R. Moccia, J.Chem.Phys. 40, 2164, 2176, 2186 (1964).

59. D. M. Bishop, J.Chem.Phys. 43, 4453 (1965).

60. B. D. Joshi, J.Chem.Phys. 47, 2793 (1967).

61. T. H. Dunning, Jr., R. M. Pitzer and S. Aung, J.Chem.Phys. 57, 5044 (1972).

62. G. P. Arrighini, C. Guidotti and O. Salvetti, J.Chem.Phys. 52, 1037 (1970).

63. R. M. Stevens, J.Chem.Phys. 55, 1725 (1971).

64. E. Clementi and H. Popkie, J.Amer.Chem.Soc. 94, 4057 (1972).

65. G. P. Arrighini, C. Guidotti, M. Maestro, R. Moccia and O. Salvetti, J.Chem.Phys. 49, 2224 (1968).

66. G. Herzberg, Proc.Roy.Soc. (London) A262, 291 (1961).

67. G. Herzberg and J. W. C. Johns, Proc.Roy.Soc. (London) A295, 107 (1966).

68. G. Herzberg and J. W. C. Johns, J.Chem.Phys. 54, 2276 (1971).

69. E. Wasserman, W. A. Yager and V. J. Kuck, Chem.Phys.Lett. 7, 409 (1970).

70. E. Wasserman, V. J. Kuck, R. S. Hutton and W. A. Yager, J.Amer.Chem.Soc. 92, 7491 (1970).

71. G. Herzberg and J. W. C. Johns, Proc.Roy.Soc. A298, 142 (1967).

72. A. D. Walsh, J.Chem.Soc. 2260, 2296 (1953).

73. C. F. Bender and H. F. Schaefer III, J.Mol.Spec. 37, 423 (1971).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

74. R. W. Carr, Jr., T. W. Eder and M. G. Topor, J.Chem.Phys. 53, 4716 (1970).

75. S. V. O'Neil, H. F. Schaefer III, and C. F. Bender, J.Chem.Phys. 55, 162 (1971).

76. P. J. Hay, W. J. Hunt and W. A. Goddard III, Chem.Phys.Lett. 13, 30 (1970).

77. L. Andrews and G. C. Pimentel, J.Chem.Phys. 47, 3637 (1965).

78. A. E. Douglas, Discuss. Faraday Soc. 35, 158 (1963).

79. R. Colin and A. E. Douglas, Can.J.Phys. 46, 61 (1968).

80. P. G. Gassman, Accounts Chem.Res. 3, 27 (1970).

81. S. Y. Chu, A. K. Q. Siu, and E. F. Hayes, J.Amer.Chem.Soc. 94, 2969 (1972).

82. D. A. Ramsay, J.Chem.Phys. 25, 188 (1956).

83. K. Dressler and D. A. Ramsay, Proc.Roy.Soc. (London) A251, 553 (1959).

84. R. N. Dixon, Mol.Phys. 9, 357 (1965).

85. C. F. Bender and H. F. Schaefer III, J.Chem.Phys. 55, 4798 (1972).

86. J. S. Hyde and E. S. Freeman, J.Phys.Chem. 65, 1636 (1961).

87. T. Cole, J.Chem.Phys. 35, 1169 (1961).

88. A. Pipano, R. R. Gilman, C. F. Bender and I. Shavitt, Chem.Phys.Lett. 4, 583 (1970).

89. A. Pipano, R. R. Gilman, and I. Shavitt, Chem.Phys.Lett. 5, 285 (1970).

90. J. A. Pople and H. C. Longuet-Higgins, Mol.Phys. 1, 372 (1958).

91. C. R. Brundle and D. W. Turner, Proc.Roy.Soc. (London) A307, 27 (1968).

92. L. Asbrink and J. W. Rabalais, Chem.Phys.Lett. 12, 182 (1971).

93. A. W. Potts and W. C. Price, Proc.Roy.Soc. (London) A326, 181 (1972).

94. S. De Jaegere, J. Deckers and A. Van Tiggelen, *Proceedings of the Eighth Symposium (International) on Flame and Combustion*, 155 (Baltimore: Williams and Wilkins, 1962).

95. J. A. Green and T. M. Sugden, *Proceedings of the Ninth Symposium (International) on Flame and Combustion*, 607 (New York and London: Academic Press, 163).

96. R. E. Richards and J. A. S. Smith, Trans. Faraday Soc. 47, 1261 (1951).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

97. Y. K. Yoon and G. B. Carpenter, Acta Cryst. 12, 17 (1959).

98. J. W. Moskowitz and M. C. Harrison, J.Chem.Phys. 43, 3550 (1965).

99. G. H. F. Diercksen and W. P. Kraemer, Theor.Chim.Acta 23, 387 (1972).

100. J. M. Lehn and B. Munsch, Mol.Phys. 23, 91 (1972).

101. J.C. Slater, Phys. Rev. 36, 57 (1930).

102. R.S. Mulliken, Rev. Mod. Phys. 32, 232 (1960).

103. I. Shavitt, Methods of Computational Physics (Academic Press, New York, 1963), Vol. II, p 30.

104. T. H. Dunning, Jr, J. Chem. Phys. 53,2823 (1970); 55, 716 (1971).

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.