Theor Chim Acta (1989) 75: 129–141

Theoretica Chimica Acta

© Springer-Verlag 1989

Perturbative corrections to basis incompleteness in molecular SCF calculations

Carl S. Ewig, Lidia Smentek–Mielczarek*, and B. Andes Hess, Jr.

Department of Chemistry, Vanderbilt University, Nashville, TN 37235, USA

(Received March 14, 1988/Accepted April 5, 1988)

A unified summary is presented of the mathematical approach developed by McDowell for employing perturbation theory to correct for basis-set incompleteness in *ab initio* SCF calculations. Revised expressions for the corrections to the wavefunction both in terms of orbitals and spin-orbitals are presented with explicit incorporation of the spin variables. Employing H_2O as an example, we show that this approach is considerably more powerful for computing molecular energies with standard basis sets than was indicated by previous work. In particular at the higher levels of approximation it accurately reproduces the effect of polarization functions in sets such as 6-31G* and 6-31G**. The equilibrium molecular structure of H_2O was also computed by this approach and found to give good accuracy. In each case perturbing functions coupled to both occupied and virtual orbitals are required for acceptable results.

Key words: Perturbation theory — Green's functions — Basis sets

1. Introduction

A central feature of carrying out atomic and molecular *ab initio* quantum calculations by any of the basis-set expansion methods is first gauging the degree of convergence of the basis, and then increasing the number of basis functions as needed to achieve the desired level of accuracy. This is at once both one of the greatest strengths and weaknesses of such methods since, although no basis is

---

* *Permanent address*: Institute of Physics, Nicholas Copernicus University, PL-87-100 Toruń, Poland

ever perfectly complete, comparison of computed properties in two or more differing basis sets often permits an *a priori* estimate of the reliability of the final result.

Increasing the size of a molecular basis generally consists of adding additional functions, typically diffuse or "polarization" functions of higher atomic angular momentum per atom than in the original set. These however make only a small change in the overall molecular wavefunction. In particular the addition of polarization functions, such as those of *d*, *f* and *g* symmetry, require considerable additional computational effort which in many cases makes only very small contributions to the wavefunction. This leads logically to the concept of employing perturbation theory to estimate the effect such added functions would have, without performing a full calculation in the larger basis. To our knowledge the first attempt to treat added functions by perturbative methods was made by Sadlej [1-3], who reported test calculations on the He and Be atoms. However the approach that now seems more generally applicable evolved in a series of papers by McDowell and coworkers [4-12]. The denouement of this work was a remarkable conclusion: the correction caused by including additional functions to the basis set in any SCF calculation, in terms of both the wavefunction and energy, can be found by solving a small set of inhomogeneous linear equations. Further, all quantities that appear in these equations are simply orbital energies and integrals of the same types as are formed in any conventional *ab initio* calculation. Simple test calculations on the He and Be atoms [11], the H_2 and LiH molecules [9-11] and on one polyatomic species [11], H_2O, were reported. Each of the three molecular cases employed only a small Slater basis perturbed by a single *s* gaussian function centered along each bond.

The purpose of the present study is to determine whether the perturbation approach derived by McDowell and coworkers is sufficiently powerful and general to permit accurate calculations on molecules of practical interest while using standard (e.g. split-valence) basis sets. In this paper we first present a summary of the use of Green's functions in treating perturbative corrections to the basis set in SCF calculations. In particular we give an improved equation for the coefficients of the correction-perturbation functions in spin-orbital form and a simplified expression for these functions in a paired-spin orbital representation. We then present a comparison of the results by the Sadlej and McDowell approaches for two test calculations on the He and Be atoms. The next section reports sample results for the energy of H_2O employing relatively large basis sets: 6-31G [13], 6-31G* [14], and 6-31G** [14]. Finally, since no property except the energy has thus far been computed by these perturbative procedures, we show their accuracy for determining the equilibrium molecular structure of H_2O in the same range of basis sets.

2. Summary of basis-incompleteness perturbation theory

The basic problem in the nonrelativistic theory of atoms and molecules is to find an approximate solution of the time-independent Schrödinger equation

$$H\Psi = E\Psi. \tag{1}$$

It is a common practice to start any investigation with the restricted Hartree–Fock procedure, which provides a qualitative understanding of many features of atomic or molecular systems. In order to understand the optical and chemical properties of a given system however it is usually necessary to perform an analysis beyond the Hartree–Fock model and to deal with a true many-body problem.

In the perturbation approach the Hamiltonian of the system is split into two parts

$$H = H_0 + V, \tag{2}$$

where H_0 is a model Hamiltonian and V is a perturbation. The zeroth-order operator H_0 is usually taken as the Hartree–Fock Hamiltonian. However at this step in the analysis we may assume only that H_0 is a hermitian operator having a complete set of eigenfunctions,

$$H_0\phi_i = E_0^i\phi_i. \tag{3}$$

If ϕ_i is a reasonable approximation to an exact solution Ψ for a given state, the purpose of the perturbation approach is to generate the corrections to the zeroth-order approximation. Thus, performing the perturbation analysis one might find how a given set of ϕ_i's is related to the exact Ψ. However it is also possible to establish the relationship between the accuracy of the results obtained for various initial basis sets used for an expansion of the zeroth-order approximation. From a practical point of view the latter aspect is of great importance and it is a main subject of the present paper.

In *ab initio* studies of perturbative corrections much computational effort is concerned with performing the summations over all excited orbitals including those in the continuum. In order to avoid this difficulty several approaches have been introduced. A very convenient and efficient way of performing perturbative calculations is based on the so-called "perturbed-function" method introduced by Sternheimer [15]. This method eliminates the need for generating a complete set of single-particle states by introducing functions obtained as solutions of the appropriate differential equations [16–18].

Another approach which avoids the summation problems in the perturbation expansion is based on the time-independent Green's function formalism. The general definitions and properties of the Green's function operators as well as their close relation to the Brillouin–Wigner perturbation theory defined in terms of the resolvent-operator technique are presented for example by Lindgren and Morrison [18].

One of the applications of the Green's function formalism is based on the possibility of solving the Dyson's equation which relates the "real" Green's function (associated with the full Hamiltonian H) and "zeroth-order" Green's function (associated with H_0) [19]. If one can solve this equation it is then possible to obtain the eigenvalues of H by finding the poles of the Green's function operator (see, for example [20]).

Following McDowell's idea [4–11] we have applied the Green's function formalism to evaluate perturbative corrections to basis incompleteness in molecular SCF calculations. (This is a different approach to this problem than the one proposed by Sadlej [1–3], which is based on variational conditions on the second-order approximation to the energy.) The details of McDowell's approach may be found in [7–10] and especially [11]. To give an overview of the approach we first summarize here the crucial points of the simplest version of the theory in which the virtual spin-orbitals are not distinguished from all the other unoccupied one-electron states.

The approach is based on a particular way of partitioning of the Hamiltonian of a given system. Namely, if the zeroth-order Hamiltonian is assumed to be the Fock operator, then the whole Hamiltonian may be written in the following form

$$H=\sum_i^N [PF(i)P+QF(i)Q] \tag{4a}$$

$$+\sum_{i<j}\frac{1}{r_{ij}}-\sum_i^N u_{HF}(i) \tag{4b}$$

$$+\sum_i^N [PF(i)Q+QF(i)P]. \tag{4c}$$

Here the partitioning technique has been applied, i.e. the entire functional space (spanned by all eigenfunctions of Eq. (3)) is split into two subspaces defined by the projection operators

$$P=\sum_i |\phi_i\rangle\langle\phi_i| \tag{5}$$

$$Q=\sum_{r\neq i} |\phi_r\rangle\langle\phi_r|. \tag{6}$$

In general the P space is associated with the eigenfunctions of H_0 for a state under consideration while Q is its orthogonal complement. Due to the completeness of the set of eigenfunctions we have

$$P+Q=1. \tag{7}$$

In the case of the Hartree–Fock approach the summation in Eq. (5) runs over all occupied spin-orbitals while the index r in the definition of Q (Eq. (6)) denotes the particle (unoccupied) states.

The first two operators in Eq. (4a) are assumed to be the zeroth-order Hamiltonian while the remaining ones form the perturbation. The operators defined by Eq. (4b) represent the non-central part of the Coulomb interaction and is generally taken as a perturbation responsible for electron-correlation effects. Furthermore, the last two operators introduced above (Eq. (4c)) are new in comparison with the standard formulation and are the source of the corrections to the Hartree–Fock energy we are concerned with.

As mentioned above the main aim of the approach presented here is to take into account all the energy corrections that are due to the fact that the zeroth-order eigenvalue problem has been solved for an incomplete basis set. This means that in performing the analysis of the perturbation expansion the following should be kept in mind:

1. Due to the definition of the Hartree-Fock potential the corrections to the energy appear at the second-order of the perturbation expansion.

2. From among all terms produced by the perturbation operators in Eqs. (4b) and (4c) only those associated with the single excitations should be included.

3. All terms arising from the double excitations should be excluded except those which might be described by coupled simultaneous single excitations (i.e. only those terms for which the factorization of the energy denominator is possible).

4. All terms which added together give zero, due to the definition of the Hartree-Fock potential, should be extracted from the perturbation expansion.

In order to identify the major corrections to the Hartree-Fock energy as well as to perform a summation of the most important terms along the series, the diagramatic technique of many-body perturbation theory has been applied [12, 18-20]. The Hartree-Fock scheme based on the summation of the "single-excitation" diagrams which contribute to the second-order energy has been proposed by Caves and Karplus [21] (see also [22]).

Following the rules presented above it was found that the "first correction" which contains the second-order and third-order Hartree-Fock-type graphs summed along the series has a form

$$\text{First correction} = \sum_i \langle \phi_i | F | \chi_i \rangle. \tag{8}$$

The "second-correction" representing all the new Hartree-Fock graphs appearing at the fourth-order also summed along the series has the form

$$\begin{aligned}\text{Second correction} = \sum_{i,j} \Big\{ & \langle \phi_j \chi_i | \chi_j \chi_i \rangle_A + \langle \chi_j \chi_i | \phi_j \chi_i \rangle_A - \langle \chi_i | \chi_j \rangle \\ & \times \Big[\langle \phi_j | F | \chi_i \rangle + \sum_k \left(\langle \phi_j \phi_k | \phi_i \chi_k \rangle_A + \langle \phi_j \chi_k | \phi_i \phi_k \rangle_A \right) \Big] \Big\}.\end{aligned} \tag{9}$$

Here the subscript A denotes the antisymmetrized two-electron integral.

In Eqs. (8) and (9) χ_i is the correction function to the ith spin orbital and is defined as a solution of the equation

$$\begin{aligned}|\chi_i\rangle = & \sum_r^{\text{part.}} \frac{|\phi_r\rangle\langle\phi_r|F|\phi_i\rangle}{(\varepsilon_i - \varepsilon_r)} \\ & + \sum_r^{\text{part.}} \sum_j^{\text{occ.}} \left(\frac{|\phi_r\rangle\langle\phi_r\chi_j|\phi_i\phi_j\rangle_A}{(\varepsilon_i - \varepsilon_r)} + \frac{|\phi_r\rangle\langle\phi_r\phi_j|\phi_i\chi_j\rangle_A}{(\varepsilon_i - \varepsilon_r)} \right).\end{aligned} \tag{10}$$

To avoid the difficulties of performing the explicit summation over the whole set of particle states in Eq. (10), the reduced Green's function operator G_0^r has been introduced. Thus, instead of Eq. (10) we have

$$|\chi_i\rangle = G_0^r(\varepsilon_i)F|\phi_i\rangle + \sum_j^{\text{occ.}} [\langle G_0^r(\varepsilon_i)\phi_j \,|\, \phi_i\chi_j\rangle_A + \langle G_0^r(\varepsilon_i)\chi_j \,|\, \phi_i\phi_j\rangle_A]. \tag{11}$$

The reduced Green's function operator is related to the zeroth-order Green's function operator as follows

$$G_0^r(\varepsilon_i) = QG_0(\varepsilon_i) \tag{12}$$

and has the spectral resolution

$$G_0^r(\varepsilon_i) = \sum_r^{\text{part.}} \frac{|\phi_r\rangle\langle\phi_r|}{\varepsilon_i - \varepsilon_r}. \tag{13}$$

The reduced Green's function operator may be shown to be identical to the resolvent of the Brillouin-Wigner perturbation expansion [18].

It is seen from Eq. (11) that the correction function obeys an integral equation. However, taking advantage of the relation [5]

$$G_0^r(\varepsilon_i)^{-1}G_0^r(\varepsilon_i) = Q \tag{14}$$

it is possible to obtain the differential equation. From the point of view of applications to molecular systems this step is particularly useful since the difficulties with the multicenter integrals over the Green's functions are thus avoided. Multiplying Eq. (11) from the left by $G_0^r(\varepsilon_i)^{-1}$ we have the differential equation

$$Q(\varepsilon_i - F)Q|\chi_i\rangle + \sum_j^{\text{occ.}} [\langle Q\phi_j \,|\, \chi_j\phi_i\rangle_A + \langle Q\chi_j \,|\, \phi_j\phi_i\rangle_A] = QF|\phi_i\rangle. \tag{15}$$

The problem of solving these equations for χ's is reduced to algebraic form if each correction function to a particular spin-orbital is expressed in terms of known functions. Thus, in practice each function χ_k has the form

$$\chi_k = Q\sum_a^s c_{ka}f_a\gamma_k \equiv Q\sum_a^s c_{ka}f_{ak}, \tag{16}$$

where s is the dimension of the set of added basis functions f_a, γ_k denotes the spin function associated with kth spin-orbital and the Q operator in the expansion assures the orthogonality of the correction function to all occupied orbitals. Substituting expansion (16) into Eq. (15) and multiplying the whole expression from the left by f_{bi} we get the linear inhomogeneous equations for the unknown coefficients c_{ia},

$$\begin{aligned} &\sum_a^s c_{ia}[\langle Qf_{bi}|\varepsilon_i - F|Qf_{ai}\rangle + \langle Qf_{bi}\phi_i \,|\, Qf_{ai}\phi_i\rangle_A] \\ &\quad + \sum_j^{\text{occ.}} (1 - \delta(\phi_i, \phi_j)) \sum_a^s c_{ja}[\langle Qf_{bi}\phi_j \,|\, Qf_{aj}\phi_i\rangle_A \\ &\quad + \langle Qf_{bi}Qf_{aj} \,|\, \phi_j\phi_i\rangle_A] = \langle Qf_{bi}|F|\phi_i\rangle. \end{aligned} \tag{17}$$

This expression is a more generally useful form of the equation for the spin-orbital coefficients than that given by McDowell and Lewis, Eq. (8) of [11], in which the spin variables were omitted.

Since we are mainly interested in the applications of this approach to closed-shell systems, the transformation from spin-orbitals in Eq. (17) to orbitals is required. For a closed-shell system each orbital is doubly occupied, i.e.

$$\phi_k \equiv \psi_k \gamma_k \quad \text{for } k=1,\ldots,N \Rightarrow \begin{cases} \psi_k \alpha \\ \psi_k \beta \end{cases} \quad \text{for } k=1,\ldots,N/2.$$

After performing the integration over the spin variables, the equation analogous to Eq. (17) has the form

$$\begin{aligned} &\sum_a^s c_{ia}[\langle Qf_b | \varepsilon_i - F | Qf_a\rangle + \langle Qf_b\psi_i \,|\, Qf_a\psi_i\rangle - 3\langle Qf_b\psi_i \,|\, \psi_i Qf_a\rangle] \\ &\quad + \sum_{j\neq i}^{N/2} \sum_a^s c_{ja}[\langle Qf_b\psi_j \,|\, Qf_a\psi_i\rangle - 4\langle Qf_b\psi_j \,|\, \psi_i Qf_a\rangle \\ &\quad + \langle Qf_b Qf_a \,|\, \psi_j\psi_i\rangle] = \langle Qf_b | F | \psi_i\rangle. \end{aligned} \tag{18}$$

The same procedure may be applied to obtain the expressions for the corrections to the energy. Thus, for a closed-shell system and with the correction functions orthogonal to the occupied orbitals, the corrections to the energy in terms of the known functions f_a are determined by the expressions

$$\text{First correction} = 2\sum_i^{N/2}\left[h_{ii} + B_{ii,ii} + \sum_{j\neq i}^{N/2}(2B_{ij,ij} - B_{ij,ji})\right] \tag{19}$$

$$\begin{aligned} \text{Second correction} = 2\sum_i^{N/2} \bigg\{ &2D_{ii,ii} - A_{ii}(h_{ii} + 3B_{ii,ii}) \\ &+ \sum_{j\neq i}^{N/2} \bigg[4D_{ij,ij} - 2D_{ii,jj} - A_{ii}(2B_{ij,ij} - B_{ij,ji} + 4B_{ji,ji} - 2B_{ji,ij}) \\ &- A_{ij}(h_{ji} + 2B_{ii,ij} + B_{ii,ji} + 3B_{ji,jj} - B_{ij,ij} + 2B_{ij,jj} - B_{jj,ij} \\ &+ \sum_{\substack{k\neq i\\ k\neq j}}^{N/2} (2\langle \psi_k\psi_j \,|\, \psi_k\psi_i\rangle - B_{ik,ij} \\ &- B_{ki,jk} - B_{kj,ik} + 4B_{ki,kj})) \bigg] \bigg\}. \end{aligned} \tag{20}$$

In order to simplify the notation, the following matrices have been introduced in Eqs. (19) and (20):

$$\begin{aligned} \boldsymbol{h}\colon h_{ji} &\equiv \langle \psi_j | h | \chi_i\rangle \\ &= \sum_a^s c_{ia}\left[\langle \psi_j | h | f_a\rangle - \sum_l^{N/2} \langle f_a \,|\, \psi_l\rangle \langle \psi_j | h | \psi_l\rangle\right], \end{aligned} \tag{21}$$

where h is an one-particle part of the Fock operator;

$$\begin{aligned} \boldsymbol{A}\colon A_{ij} &\equiv \langle \chi_i \,|\, \chi_j \rangle \\ &= \sum_a^s c_{ia} \sum_b^s c_{jb} \left[\langle f_a \,|\, f_b \rangle - \sum_l^{N/2} \langle f_a \,|\, \psi_l \rangle \langle f_b \,|\, \psi_l \rangle \right] \end{aligned} \tag{22}$$

$$\begin{aligned} \boldsymbol{B}\colon B_{ij,mn} &\equiv \langle \chi_i \psi_j \,|\, \psi_m \psi_n \rangle \\ &= \sum_a^s c_{ia} \left[\langle f_a \psi_j \,|\, \psi_m \psi_n \rangle - \sum_l^{N/2} \langle f_a \,|\, \psi_l \rangle \langle \psi_l \psi_j \,|\, \psi_m \psi_n \rangle \right] \end{aligned} \tag{23}$$

$$\begin{aligned} \boldsymbol{D}\colon D_{ij,nw} &\equiv \langle \chi_i \chi_j \,|\, \chi_n \psi_w \rangle \\ &= \sum_a^s c_{ia} \sum_b^s c_{jb} \sum_d^s c_{nd} \bigg\{ \langle f_a f_b \,|\, f_d \psi_w \rangle \\ &\quad - \sum_m^{N/2} \bigg[\langle f_a \,|\, \psi_m \rangle \langle f_d f_b \,|\, \psi_m \psi_w \rangle + \langle f_b \,|\, \psi_m \rangle \langle f_a \psi_m \,|\, f_d \psi_w \rangle \\ &\quad + \langle f_d \,|\, \psi_m \rangle \langle f_a f_b \,|\, \psi_m \psi_w \rangle \\ &\quad - \sum_l^{N/2} \bigg(\langle f_a \,|\, \psi_l \rangle \langle f_d \,|\, \psi_m \rangle \langle f_b \psi_l \,|\, \psi_w \psi_m \rangle \\ &\quad + \langle f_a \,|\, \psi_l \rangle \langle f_b \,|\, \psi_m \rangle \langle f_d \psi_m \,|\, \psi_l \psi_w \rangle \\ &\quad + \langle f_b \,|\, \psi_l \rangle \langle f_d \,|\, \psi_m \rangle \langle f_a \psi_l \,|\, \psi_m \psi_w \rangle \\ &\quad - \langle f_a \,|\, \psi_l \rangle \langle f_b \,|\, \psi_m \rangle \sum_k^{N/2} \langle f_d \,|\, \psi_k \rangle \langle \psi_l \psi_m \,|\, \psi_k \psi_w \rangle \bigg) \bigg] \bigg\}. \end{aligned} \tag{24}$$

Note that in Eqs. (17)-(24) all integrals of the form $\langle f_a f_b \,|\, f_c f_d \rangle$ which would be required in the analogous conventional calculation do not appear. Similarly in only one term are there any integrals over three of the added basis functions, i.e. integrals of the form $\langle f_a f_b \,|\, f_c \psi_w \rangle$. Of course the contributions of the added functions to the starting orbitals need not be computed in any case. These factors combined lead to the potential computational efficiency of this procedure.

Note also that in these expressions the value of N, the number of spin-orbitals, is not uniquely specified. In nearly all prior work N has been set equal to the number of occupied spin-orbitals. However it has been suggested [11] that an improved level of accuracy may be achieved when the virtual orbitals are included as well. This means that from among all particle states in definition (6) the virtual states are extracted and the correction functions are required to be orthogonal to these as well as to the occupied states. Thus, in the more general version of this approach the operator Q in (16) is replaced by

$$Q \to Q + Q' = 1 - \sum_i^{\text{occ.}} |\phi_i \rangle \langle \phi_i |, \tag{25}$$

where

$$Q = \sum_{j}^{\text{virt.}} |\phi_j^\nu\rangle\langle\phi_j^\nu|$$

$$Q' = \sum_{r \neq \text{virt.}}^{\text{part.}} |\phi_r\rangle\langle\phi_r|$$

and the correction function is then defined as

$$\chi_k = \sum_{j}^{\text{virt.}} c_{kj}^\nu |\phi_j^\nu\rangle + Q' \sum_{a}^{s} c'_{ka} f_{ak}. \tag{26}$$

The algebraic equations for the two sets of coefficients c_{kj}^ν and c'_{ka} may be obtained in the same way as above and are given explicitly by Eqs. (21) and (22) in [11]. In order to obtain the expressions for corrections to the energy, Eq. (26) is simply inserted into (8) and (9).

Finally, five ways of approximating these procedures have been proposed by McDowell and Lewis [11], which may be summarized as follows:

Approximation I. Only occupied orbitals are included. In addition the coupling terms between correction function f_b and all orbitals other than i (the double summation in Eq. (18)) are neglected.

Approximation II. Only occupied orbitals are included. Correction function/other-orbital terms are included.

Approximation III. Both occupied and virtual orbitals are included. Correction function/other-orbital terms are neglected.

Approximation IV. Both occupied and virtual orbitals are included. Only correction function/other-orbital terms connected to occupied orbitals are included.

Approximation V. Both occupied and virtual orbitals are included throughout. All terms are included.

Therefore in the following we shall be concerned not only with the accuracy of the overall perturbation method, but also with these individual levels of approximation.

3. Sample calculations

3.1. *Comparison of methods for* He *and* Be

As noted above, the first perturbative scheme of this type was proposed by Sadlej [1-3]. He also reported test calculations on the He and Be atoms in Slater basis sets. However of the six cases he reported, only two represent an appreciable correction in the energy (greater than 10^{-5} hartree). The first consisted of He with only one Slater function perturbed by the addition of a second Slater function. We shall refer to this as a $1\zeta \rightarrow 2\zeta$ perturbation. The second case was a Be atom with four Slater functions perturbed by addition of a fifth, or $4\zeta \rightarrow 5\zeta$. We studied these same two cases employing the equations described above, however approximating each Slater function by an expansion in six gaussian functions. The results are shown in Table 1. In the case of He the Sadlej result is qualitatively correct, the error being primarily a measure of the importance of the neglected third- and higher-order terms. This relative accuracy is to be expected from the absence of virtual orbitals in this case, which are not treated in the Sadlej approach. For the same reason approximations I and III must be identical, as must II, IV and V. However the Be test case is much different and represents a rigorous test for both methods. Here the correct treatment of the virtual orbitals is seen to be critical. Not only is the Sadlej result in error by a factor of roughly 2.5 but approximations I–IV, which lack full virtual-orbital orthogonalization and coupling, are even worse. However, as often seems to be the case for atoms approximation V, which includes a full treatment of the virtual orbitals, gives almost exactly the same result as would be obtained in the larger basis.

3.2. Split-valence test calculations on H_2O

3.2.1. Energy corrections. As noted above, the test calculations using basis-incompleteness perturbation theory that have been reported thus far for molecules have been quite different from cases of practical interest in that small Slater basis sets were employed with the perturbing functions consisting of either one or two gaussian primitive functions centered along bonds. Therefore we carried out a series of calculations to determine the applicability to gaussian basis sets and to the addition of several perturbing functions, particularly atom-centered polarization functions.

As an initial example we computed the energy of H_2O in the 6-31G basis [12] and employing the molecular structure optimized in that basis. We then added as the perturbing functions a six-fold set of *d* gaussian primitives to the oxygen

Table 1. Comparison of Sadlej [3] and McDowell methods for basis-incompleteness corrections to atomic energies[a]

Approximation	He $1\zeta \rightarrow 2\zeta$	Be $4\zeta \rightarrow 5\zeta$
Sadlej	−0.006879	−0.000854
I	−0.007331	−0.000737
II	−0.007341	−0.000747
III	−0.007331	−0.000593
IV	−0.007341	−0.000594
V	−0.007341	−0.002269
Exact difference between basis sets	−0.007341	−0.002269

[a] Energy in the larger basis set minus the energy in the smaller basis, in hartrees. The total energies to which these values are to be added are He: −2.846299 and Be: −14.564580

atom, forming the 6-31G* basis [13]. This we will refer to as the 6-31G→6-31G* perturbation. The results for the five types of approximations are shown in the first column of data in Table 2. Here we list the energy computed in the larger set minus that in the smaller set, in hartrees. The "exact" value is the difference between the energies obtained from two full SCF calculations at the same molecular structure. Approximations I and II are seen to be only qualitatively correct. Approximations III through V, which employ orthogonalization to the virtual orbitals, are much more accurate, with V being of considerably higher accuracy than the others.

A second test of computed energies consisted of the addition of both d functions on oxygen and p functions on hydrogen, to form the 6-31G** basis [14]. This will be termed the 6-31G→6-31G** perturbation. This is a stringent test of the method, since now the perturbing functions nearly double the size of the basis. The results are shown in the second column of data in Table 2. The same trends are observed as for the 6-31G→6-31G* case. At each level the differences between approximated and exact energy differences are larger, although approximation V is still quite accurate, being in error by only 0.000034 hartrees or less than 0.1% of the difference.

3.2.2. Molecular structure corrections. Thus far no property has been computed by basis-incompleteness perturbation methods except the energy. Therefore as an additional test we carried out an optimization of the molecular structure of H_2O by minimizing the exact and approximated energies with respect to the OH bond lengths and the HOH bond angle. Since analytic expressions for the derivatives of the approximated energy with respect to internal coordinates have not yet been developed, the optimization was performed by computing the energies over a grid of discrete points in the internal coordinate space. The results are listed in Table 3 which shows the differences in structural parameters, i.e. values optimized in the larger basis minus those in the smaller basis. The general

Table 2. Corrections to molecular energy[a] of H_2O from the basis-incompleteness perturbations 6-31G→6-31G* and 6-31G→6-31G**

Approximation	6-31G→6-31G*	6-31G→6-31G**
I	−0.020938	−0.034474
II	−0.021071	−0.034892
III	−0.023927	−0.036540
IV	−0.024017	−0.036778
V	−0.024327	−0.037238
Exact difference between basis sets	−0.024322	−0.037272

[a] Energy in the larger basis set minus the energy in the smaller basis in hartrees at the 6-31G optimized molecular structure. The total energy in the 6-31G basis, to which these values are to be added, is −75.985359 hartrees

Table 3. Effect on computed equilibrium structure[a] of H_2O from the basis-incompleteness perturbations 6-31G→6-31G* and 6-31G→6-31G**

Approximation	6-31G→6-31G*		6-31G→6-31G**	
	Δ OH length	Δ HOH angle	Δ OH length	Δ HOH angle
I	−0.00084	−5.12	−0.00453	−4.69
II	−0.00080	−5.13	−0.00479	−4.76
III	−0.00242	−5.95	−0.00630	−5.38
IV	−0.00236	−5.94	−0.00626	−5.36
V	−0.00230	−6.06	−0.00653	−5.57
Exact difference between basis sets	−0.00232	−6.04	−0.00658	−5.58

[a] Bond lengths and angles computed in the larger basis minus those in the smaller basis. Lengths in Å, angles in degrees. The values in the 6-31G basis, to which these values are to be added, are OH=0.94963 Å and HOH=111.55°

trends here are the same as seen for the energies in Table 2 except that the 6-31G→6-31G** is not consistently less accurate than 6-31G→6-31G*. Approximations I and II are surprisingly poor in predicting corrections to the bond lengths. Approximation V is quite accurate in all cases, the perturbed 6-31G basis predicting the bond length obtained the larger basis sets with an accuracy of better than 5×10^{-5} Å and angles within 0.02°.

4. Conclusions

We have shown, in Sect. 2 above, that the very powerful perturbation approach developed by McDowell and coworkers can be written in a simple and useful form, both in terms of spin-orbitals, when the spin variable is explicitly included, and in terms of orbitals as in the case of restricted Hartree-Fock calculations. In the test calculations on the He and Be atoms and H_2O our findings concur with those of McDowell and Lewis [11] that approximations I and II are useful only in the special case where there are no virtual orbitals, as the He energies shown in Table 1. The calculation on Be in comparison with Sadlej's value and with the exact result shows that inclusion of virtual orbitals is critical to obtaining a useful level of accuracy, (although it requires typically ten times more computational effort than methods I and II in which virtual orbitals are neglected). Further, based on these and the test calculations on H_2O, we can make a stronger statement: since nearly the same computational effort is required by approximations III, IV and V with V nonetheless invariably being appreciably more accurate than III and IV for both the energy and structural parameters, there is in general no advantage in using any of the more approximate methods.

For perturbations connecting the widely employed split-valence basis sets 6-31G, 6-31G* and 6-31G** approximation V consistently gives results of useful accuracy. For H_2O the calculation of total molecular energy in the 6-31G basis corrected to 6-31G** (6-31G→6-31G**) is in error relative to a conventional SCF calculation in the larger basis by only 0.000034 hartrees (0.089 kJ/mol), suggesting that the method will be accurate enough for *ab initio* studies of thermodynamic quantities. Similarly the same approximation gives, in this case, the bond length accurate within 5×10^{-5} Å and the bond angle to 0.02°. We believe much of the remaining error in the structural parameters computed employing approximation V is due to roundoff and similar numerical errors. Although test calculations employing approximation V on much larger systems are clearly needed, as well as efficient computer programs to determine its ultimate computational efficiency, it now appears that basis-incompleteness perturbation methods may prove to be valuable extensions of all *ab initio* methods based on basis-set expansions of molecular orbitals.

Acknowledgments. It is a pleasure to acknowledge the assistance of Dr. Keith McDowell. This research was sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-86-0146.

References

1. Sadlej AJ (1972) Acta Phys Pol A42:341
2. Sadlej AJ (1972) Mol Phys 23:1031
3. Sadlej AJ (1973) Chem Phys Lett 19:604
4. McDowell HK, Porter RN (1976) J Chem Phys 65:658
5. McDowell HK, Porter RN (1976) J Chem Phys 65:4210
6. McDowell HK, Porter RN (1977) J Chem Phys 66:4725
7. McDowell K (1978) J Chem Phys 68:4151
8. McDowell K (1979) J Chem Phys 70:3148
9. McDowell K (1980) J Chem Phys 73:4138
10. McDowell K (1980) Int J Quantum Chem XVII:895
11. McDowell K, Lewis L (1982) J Chem Phys 77:3068
12. Lewis L (1981) Thesis. Clemson University
13. Hehre WJ, Ditchfield R, Pople JA (1972) J Chem Phys 56:2257
14. Hariharan PC, Pople JA (1973) Theor Chim Acta 28:213
15. Sternheimer R (1950) Phys Rev 80:102
16. Morrison JC (1972) Phys Rev A 6:643
17. Jankowski K, Smentek-Mielczarek L, Sokolowski A (1986) Mol Phys 59:1165
18. Lindgren I, Morrison J (1981) Atomic many-body theory. Springer, Berlin Heidelberg New York
19. Fetter AL, Walecka JD (1971) Quantum theory of many-particle systems. McGraw-Hill, San Francisco
20. Szabo A, Ostlund NS (1982) Modern Quantum theory: introduction to advanced electronic structure theory. Macmillan, New York
21. Caves TC, Karplus M (1969) J Chem Phys 50:3649
22. Rossky PJ, Karplus M (1980) J Chem Phys 72:6085