Gaussian Basis Sets for High-Quality ab Initio Calculations

Jan Almlof,* Trygve Helgaker,
Supercomputer Institute, University of Minnesota, Minneapolis, Minnesota 55455

and Peter R. Taylor
ELORET Institute, Sunnyvale, California 94087 (Received: August 7, 1987)

Different types of Gaussian basis sets for accurate LCAO calculations are discussed. For calculations designed to recover a substantial portion of the correlation energy, we suggest the use of basis sets comprising natural orbitals from correlated calculations on the atoms. These basis sets have proven to be very efficient in accounting for large fractions of the molecular correlation energy. For cases in which an SCF or MCSCF treatment is adequate the use of a floating basis provides a rapid convergence to the large-basis limit.

Introduction

In ab initio calculations, deviations from the exact nonrelativistic limit are caused by truncations of the one- and N-particle basis sets. It is desirable to improve these approximations in a systematic way toward a well-defined limit. Due to recent full CI benchmark calculations the truncation effects on the N-particle basis (the configuration expansion) are now rather well understood. In contrast, the construction of basis sets for LCAO calculations is an area where development has been slow. In Schaefer's words, the selection of an appropriate basis set is still "more art than science".

In correlated calculations, the size of the final contracted basis set often constitutes a bottleneck, since this determines the amount of work required to optimize the wave function. In contrast, the integral time is of minor concern. Clearly there is a need for compact basis sets even if it means a more expensive evaluation of integrals.

A configuration basis can be gradually extended, at least in principle, from one configuration to the full CI limit. Extensions can be made by adding successively less important configurations (or classes of configurations) to the basis, and truncation effects can be explored in a systematic way in a CI calculation. Similarly, a perturbation treatment of correlation energy provides a well-defined path toward the exact solution. In contrast, most studies of one-particle basis set effects have employed rather arbitrarily chosen basis sets of different size. Variations in atomic and molecular properties are erratic and difficult to interpret, since a smaller basis set does not normally span a subspace of the larger one.

The basis sets used for calculations on polyatomic molecules are almost invariably Gaussians centered on the positions of the atomic nuclei:

$$\psi(l,m,n,\alpha) = (x - x_1)^l(y - y_1)^m(z - z_1)^n e^{-\alpha(r^2)}$$

Variational optimization of Gaussian basis sets is normally restricted to orbital exponents. Recent studies have shown that although many basis sets in present use are incompletely optimized, those deficiencies have a minimal effect on computed atomic and molecular properties. Attempts to improve basis sets further must therefore focus on other parameters of the basis, i.e., contraction schemes and/or basis set centers.

Basis sets are usually optimized for the free atom (see, however, the recent work by Schlegel on exponent optimization in molecules). The conventional approach is to carry out an SCF calculation with an uncontracted basis of fixed size, optimizing the orbital exponents either with a numerical grid-based procedure or by using analytic derivative techniques. The coefficients for the expansion of the self-consistent AOs are then used as contraction coefficients. Krishnan et al. optimized the exponents and contraction coefficients of their 6-31G* sets in MP2 calculations. These basis sets were designed specifically for correlated calculations, which is an obvious advantage, even though a variational treatment of MP2 energies raises some formal objections.

Ideally, procedures such as those outlined above would lead to basis sets made up by the optimum atomic orbitals. An SCF calculation on the atom with this contracted basis would thus reproduce the results for the (usually much larger) primitive basis. However, this is seldom the case in practice. For reasons of convenience and tradition, segmented and nodeless basis sets are often used. In a segmented basis, a primitive Gaussian can contribute only to one contracted function. Clearly, for all atoms beyond He the nodal structure of the atomic orbitals is compromised in such a basis. (The popular small STO-nG basis sets are also segmented, but they do incorporate an approximate nodal structure in basis functions with n > l + 1. As a result, these basis sets sometimes produce molecular results which are competitive with those from energy-optimized, nodeless basis sets.)

In general, the errors introduced by compromising the nodal structure are several orders of magnitude larger than those due to inadequacies in the primitive basis. As a remedy for the node problem, Raffenetti suggested the use of a general contraction scheme, i.e., one where each primitive function can contribute to all the contracted ones. This, of course, would make it possible to use the "true" atomic SCF orbitals as a basis, which would in turn automatically provide the correct nodal structure. There are some problems with this scheme, however. One is that only a minimal basis set is defined. Attempts have been made to circumvent this problem by augmenting the basis set with virtual orbitals, using either an N or an N - 1 electron potential, but the results were not always encouraging. Another objection is, of course, that the entire procedure is biased toward SCF calculations on atoms.

Polarization Functions

A basis set that provides a good description of the atoms is a necessary, though by no means sufficient, condition for balanced calculations on molecules. In addition to describing the atoms well, the basis must have the necessary flexibility to account for the deformation that the atoms undergo when a molecule is

---

formed. For the calculation of molecular properties, the basis must also be able to account for the change of the wave function caused by an external perturbation. The required flexibility is achieved by augmenting the basis set with diffuse functions and polarization functions.

Polarization functions are basis functions with angular momenta higher than those occupied in an atomic SCF calculation. Their purpose is to change the shape of the atomic orbitals from that adopted in the atom and to account for a shift of the center of gravity of the charge density from the atomic nucleus. Polarization functions have their radial maximum in the same region as the valence orbitals.

The center of the basis function $r_n$ in eq (1) is usually chosen to coincide with the atomic nucleus. For calculations on atoms this is a natural choice, but for molecules where the local environment of the atom has a lower symmetry, it may be expected that this position is not necessarily the optimal one. It is easily shown that the derivative of a Gaussian with respect to its center is a combination of Gaussians with the same exponent but with higher and lower $l$-values:

$$\frac{\partial \phi(1,m,n,\alpha)}{\partial r_n} = \psi(-1,m,n,\alpha) - 2\alpha \psi(1,1,m,n,\alpha)$$  (2)

Therefore, if the polarization functions have exponents similar to those of valence orbitals, they will describe the effect of displacing the basis functions from the nuclei—at least to first order in a perturbation or Taylor expansion. An alternative approach—one which will be discussed later—would be to choose the basis function to actually move off the nucleus.

Correlation Functions

In order to describe angular correlation effects, basis functions with high angular momenta are necessary. Like polarization functions they need to have their radial maxima in the region of the valence electrons, although the argument for identical orbital exponents no longer applies. Often these correlation effects demand a larger and more flexible basis set than does polarization. Correlation effects are therefore usually well accounted for in beyond-Hartree–Fock calculations with a basis set adequate to describe correlation effects.

Basis Set Superposition Errors

For most chemical applications relative energies are more important than absolute ones, and for all types of bonding it is important that any deficiencies which may exist in the description of the monomer fragments are not accidentally removed when the bond is formed. Unfortunately, this is not always the case. Deficiencies in the monomer description are often reduced by basis functions on the other fragment. This effect, which may occur on both the one- and $N$-electron level, is referred to as basis set superposition error. This undesired "improvement" is in general minimized if the description of the monomer fragments is as good as possible, while keeping the basis sets small.

ANO Basis Sets

As mentioned above, several objections can be raised against conventional basis sets: they are segmented, which compromises the nodal structure; the functions are centered on the atomic nuclei, which increases the need for polarization functions; and the contraction scheme is usually based on SCF calculations with no consideration of electron correlation (which is what ultimately places the most stringent requirement on basis sets). For highly correlated calculations, we suggest the use of a different type of basis set: the atomic natural orbital (ANO) basis, a basis made up by the natural orbitals with the highest occupation numbers from an atomic CI calculation in the uncontracted basis.

Although it is not required for this scheme, we suggest that the uncontracted basis be chosen large enough to saturate each $l$-value space for practical purposes. With modern codes the time for integral evaluation is a minor part of the total time required for a high-level correlation calculation, and with a near complete set of primitives the ANOs can adopt a more optimal form and fewer of them are needed. Such a basis has numerous advantages: (i) It accounts for electron correlation in a way that is very nearly optimal for the atom with a given number of basis functions. (ii) It has the correct nodal structure. Since the atomic fragments are so well described and the basis sets are very compact, basis set superposition errors are minimized. (iii) It can be extended in a smooth, well-defined, and physically reasonable way toward the uncontracted basis limit by lowering the occupation number cutoff. With each such extension the larger basis set continues to contain the smaller one as a subset, and the change in computed properties can be ascribed entirely to the new functions added.

This is a great advantage and is in strong contrast to most basis set investigations performed with conventional basis sets, in which smaller sets seldom survive as a subset of the larger sets used.

The development of ANO basis sets requires methods and codes that treat atoms on a level similar to the intended molecular calculation (SCD1, MRCI, etc.) while retaining full atomic symmetry in the final natural orbitals. It has recently been shown how this can be achieved with little extra computational effort even for very complicated atomic states and sophisticated treatment of electron correlation.

Finally, of course, a general contraction scheme is needed, i.e., one that allows for a primitive basis function to contribute to several contracted ones. This should not be a problem with modern codes for integral evaluation. For efficiency reasons these codes treat all integrals involving primitive functions of four different shells in parallel, in order to obtain sufficient vector lengths for modern supercomputers. This naturally produces the integrals needed for a general contraction scheme in the appropriate grouping.

More Complicated Situations

It is evident from the above that an optimal description of the atom is crucial for the ANO strategy. The choice of atomic state is one question which naturally arises. In many cases, the atomic ground state is representative for the situation in the molecule, and in other cases the difference seems to be insignificant. In the methylene molecule, for example, the singlet–triplet splitting is unaffected to within 0.1 kcal/mol whether an s$^4$ or an s$^2$ optimized basis for carbon is used. There are cases, however, where such a simple approach is likely to fail. In transition-metal chemistry, for instance, it is often necessary to have a balanced description of s$^2$d$^5$, s$^2$d$^3$d$^1$, and s$^2$d$^2$d$^1$d$^1$d$^1$ states in the same calculation. The radial d-electron distributions for these atomic configurations are known to be quite different, and it is not likely that an ANO basis optimized for one particular state would have the flexibility to describe the others on an equal footing. Similar situations arise in calculations of electron affinities, where neutral and ionic states require a balanced description. For these problems, the ANO approach offers an excellent solution. A composite density matrix containing all the relevant states can easily be computed, either from a variational calculation on an average of atomic states or by averaging density matrices for different pure states. The eigenvectors of such a composite density matrix would constitute the appropriate ANO basis molecular calculations. It is observed that when such an approach is used, the ANO spectrum is more dense than usual and more basis functions must be included before the occupation numbers reach the desired threshold. This, of course, reflects the appropriate response to the increased demands placed on the basis set in such a complex situation.

Results with ANO Basis Sets

ANO basis sets have been applied and tested for a large number of different situations. The general finding is that they perform very well for energies and closely related properties, in fact, better than STO sets of similar size for the cases where comparisons have been made. For situations where a flexible description of the diffuse regions is more important than an optimal energy, it is necessary to augment the set. This can be done by allowing the
The uncontracted basis set for Br is Dunning's \(^{11}\) 14s11p6d augmented with f and diffuse p and d to 14s13p7d4f. For H an uncontracted 8s6p4d basis was used. The uncontracted energies are \(\Delta E_{\text{SCF}} = -2572.987707\) hartree and \(E_{\text{cor}} = -0.190724\) hartree. (a) Contracted ANO \([7\,6\,4\,2/4\,3\,2]\). (b) Contracted ANO \([6+1, \,5+1, \,3+1, \,1+1, \,3+1, \,2+1, \,1+1]\). (c) Contracted ANO \([7+1, \,6+1, \,4+1, \,4+1/4+1, \,3+1, \,2+1]\). (d) Contracted segmented \([7 \, 6 \, 4 \, 2/4 \, 3 \, 2]\) \((7+2+1+1+1+1, \, 7+2+1+1+1+1, \, 4+1+1+3+1+1, \, 5+1+1+1+1, \, 4+1+1+1+1, \, 3+1+1, \, 1+1)\); contraction coefficients for \(f\) orbitals are based on natural orbitals. (e) Contracted Raffenetti \([7 \, 6 \, 4 \, 2/4 \, 3 \, 2]\). (f) Contracted ANO Br/Br' \([7 \, 6 \, 4 \, 2/4 \, 3 \, 2]\). (g) Contracted ANO Br/Br' \([6+1, \, 5+1, \, 3+1, \, 1+1/3+1, \, 2+1, \, 1+1]\).

Table II: Contraction Errors (au \(x 10^{-3}\)) in the Dipole Moment of HBr

<table>
<thead>
<tr>
<th>SCF</th>
<th>cor</th>
<th>SCF</th>
<th>cor</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>78</td>
<td>-128</td>
<td>e</td>
</tr>
<tr>
<td>b</td>
<td>151</td>
<td>-198</td>
<td>f</td>
</tr>
<tr>
<td>c</td>
<td>110</td>
<td>-23</td>
<td>g</td>
</tr>
<tr>
<td>d</td>
<td>60</td>
<td>-172</td>
<td></td>
</tr>
</tbody>
</table>

*The uncontracted values are \(\mu_{\text{SCF}} = 0.373687\) au and \(\Delta\mu_{\text{cor}} = -0.040390\) au. The notation for different contraction schemes is as in Table I.*

The basis set was then contracted, testing a number of different contraction schemes. Table I shows that all the ANO-based contractions perform well, with a total loss of less than 1 mhartree due to contraction. As expected, the segmented basis performs very poorly in this respect. The AO coefficients do not readily suggest a unique segmented contraction scheme, and indeed Dunning \(^{11}\) questioned whether a reasonable segmented contraction scheme could be found for this set (and others of the same size). The one used here has an energy loss of nearly 0.7 hartree. This is due to the incorrect nodal structure of the AOs enforced by the segmented contraction. Interestingly enough, the "Raffenetti" \(^{7}\) type contraction based on atomic SCF orbitals is not competitive with ANOs for HBr even at the SCF level. The reason is that this scheme defines only the strongly occupied space, i.e., a minimal basis set. Additional functions have to be chosen on rather arbitrary criteria—in this case the outer primitive shells were left uncontracted.

At the SCF level, the main advantage of using ANOs is the correct nodal structure of the AOs and the unique definition of the functions outside the minimal basis. For the correlation energy another interesting observation can be made. First, it should be noted that the error in correlation energy is larger than the SCF error, so that despite the fact that the contractions are optimized for correlation the higher angular momentum sets are perfectly able to describe molecular polarization effects. The contraction loss in the correlation energy for the ANO sets obtained by averaging Br and Br' ANOs is considerably larger than that for the Br ANO sets, although the SCF contraction errors are actually slightly smaller. In Table III the deviations from the uncontracted results for the HBr dipole moment are listed. As the total dipole moment is the sum of two terms which differ in sign, where the SCF and correlation contraction losses are opposite in sign the total dipole will be in better agreement with the uncontracted result than when the signs are the same. Consequently, the total dipole moment (at the CI level) is given best by the ANO contractions in which the outermost primitives have been freed. However, there is little in the way of a systematic trend in the individual results: the SCF contraction error is smallest for the SCF-derived contractions (segmented and Raffenetti), but these have the largest contraction errors at the correlated level of all except the simple ANO set. It should be noted, however, that all errors are small, none being larger than 0.002 au. Hence if it is desired to obtain good correlation energies and good dipole moments, any of the augmented ANO sets would be a suitable choice.

### Floating Basis Sets

As noted above, atom-centered basis sets require augmentation with polarization functions to describe the deformation in atomic densities attendant on formation of molecules, and this is particularly important for properties. For example, when an external electric field is applied to a molecule, the electrons and nuclei will be displaced in opposite directions, and in an atom-centered basis this effect can only be accounted for if polarization functions are included.

An alternative to polarization functions is to allow the orbitals to move off the nuclei and occupy their optimal positions in the molecule. This adds considerable flexibility to the basis set without increasing its size. Due to the high power dependence of electronic structure calculations on the basis set size, such "floating orbital" basis sets have considerable advantages. On the other hand, the optimization of the orbital positions is a nontrivial task, being at least as expensive as a conventional geometry optimization. A final judgment on the merits of floating orbitals is therefore difficult, especially as new technical developments change computational efficiencies.

Interest in floating functions is, of course, not new. Hurley \(^{12}\) investigated floating functions in the 1950's, Frost introduced floating spherical Gaussian orbitals 20 years ago, \(^{13}\) and more recently Huber \(^{14}\) has carried out extensive studies using his floating orbital geometry optimization (FOGO) model. However, with recent developments in second-order techniques for optimization of geometries and calculation of molecular properties as energy derivatives, the use of floating orbitals has become more attractive.

There are two main steps in floating orbital calculations: optimization of the wave function and the calculation of properties. The optimization may be carried out as a conventional geometry optimization by treating the orbitals as charge-free atoms. With a second-order convergence scheme (the trust-region method) the orbital positions usually converge (to within \(10^{-7}\) in the gradient) in less than five iterations, even though the initial Hessian often has several negative eigenvalues.

The analytical calculation of properties can also be carried out in a straightforward manner by treating the floating orbitals as charge-free atoms. However, when second-order properties are calculated, an additional (inexpensive) computational step is required at the end of the calculation to account for the relaxation of the orbital positions after the perturbation. For example, if we calculate the perpendicular component of the polarizability of H\(_2\)O using a fixed double-\(\tilde{\text{p}}\) basis set, we obtain 1.29 au, considerably less than the Hartree–Fock value, which is close to 8.

---


If we now optimize the orbital positions and recalculate the polarizability using the same procedure as for the fixed orbitals, we obtain 1.13—almost the same as before. The reason is that we have only taken into account the relaxation of the orbital centers in the absence of a field. An additional relaxation term is required to describe the response of the centers to a field. This "dynamic contribution" is easily calculated from the second derivatives with respect to displacements of the orbital coordinates and the dipole derivatives with respect to the orbital coordinates. We obtain 5.03 au for this term, which gives a total perpendicular polarizability of 6.16, considerably closer to the Hartree–Fock limit.

The results of our calculations are summarized in Tables III and IV and are based on Hartree–Fock calculations on HF, H2O, NH3, CH4, CO, and H2CO. For each molecule we have calculated geometries, dipole moments, static polarizabilities, harmonic frequencies, and double-harmonic infrared intensities using four different floating basis sets. All calculations were carried out with the SiRUS15 and ABACUS16 programs. The calculated numbers have been compared with those obtained with the corresponding fixed orbitals and also with results from high-quality Hartree–Fock calculations at the 6-311++G(3df,3pd) level. The complete results will be published elsewhere,17 and we concentrate here on some general features.

All properties have been calculated at the theoretical minima rather than the experimental geometries. In this way we compare results that have been obtained without any a priori knowledge about the system. Although calculations at the experimental geometry in many cases give values closer to observed properties, our purpose is to examine the convergence of the one-electron basis, and experimental values are therefore not of immediate interest.

The double-\(\alpha\) basis (DZ) is van Duijneveldt's (8s4p/4s) set18 contracted to (4s2p/2s). DZP is the same basic set augmented with one set of polarization functions on all atoms, using the exponents of Krishnan et al.19 Finally the DZ+ and DZP+ basis sets were obtained by adding one set of diffuse orbitals to each atom (s and p on heavy atoms and s on hydrogen). The four floating basis sets are identical with the fixed basis sets except that the orbital centers are variationally optimized. For reasons explained elsewhere20 we did not float the orbitals independently; the two innermost s functions on the heavy atoms were fixed at the nuclei, while the two outermost s functions were constrained to have the same (floating) coordinates. The diffuse functions in DZ+ and DZP+ are attached to these same floating centers.

As a benchmark for the smaller floating and fixed calculations, we have carried out calculations using the 6-311++G(3df,3pd) basis set of Frisch et al.19 We do not claim that this basis gives the Hartree–Fock limit for all properties considered, but it should be large enough to represent a limit on what can be obtained in routine calculations. Also, although more accurate calculations have been carried out for selected properties using larger basis sets, most calculations of near-Hartree–Fock quality have been performed at the experimental geometry. Since many properties are strongly geometry dependent, any comparison with such results would be misleading. For example, using the floating DZP+ and the large 6-311++G(3df,3pd) basis, we obtained 0.060 and 0.058 au for the dipole moment of CO (with carbon positively charged). From the calculated dipole derivatives we may further estimate that at the experimental geometries we would obtain 0.104 and 0.107 au, respectively. This in good agreement with numerical Hartree–Fock, which gives 0.104 at the experimental geometry.20

One way to measure the quality of a basis set is to compare its energy with that obtained at the Hartree–Fock limit. In Table III we have listed how much of the difference between the Hartree–Fock energy and the DZ energy (the "residual energy") is recovered by each basis set. It is clear from Table III that floating functions cannot compete with the addition of polarization functions if the only criterion is energy lowering.

The errors in the calculated properties are listed in Table IV. The properties fall into two classes: geometrical (geometries and

---

**Table IV: Errors in Calculated Properties (%) Relative to the Hartree–Fock Limit**

<table>
<thead>
<tr>
<th>Property</th>
<th>DZ</th>
<th>DZ+</th>
<th>DZP</th>
<th>DZP+</th>
</tr>
</thead>
<tbody>
<tr>
<td>bond distances</td>
<td>1.5</td>
<td>1.6</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>bond angles</td>
<td>0.8</td>
<td>0.9</td>
<td>0.2</td>
<td>0.3</td>
</tr>
<tr>
<td>frequencies</td>
<td>4.4</td>
<td>5.0</td>
<td>0.6</td>
<td>0.2</td>
</tr>
<tr>
<td>dipole moments</td>
<td>21.3</td>
<td>28.5</td>
<td>7.3</td>
<td>9.1</td>
</tr>
<tr>
<td>polarizabilities</td>
<td>35.8</td>
<td>21.2</td>
<td>28.4</td>
<td>15.1</td>
</tr>
<tr>
<td>geometrical</td>
<td>69.7</td>
<td>73.3</td>
<td>12.4</td>
<td>19.0</td>
</tr>
<tr>
<td>electrical</td>
<td>42.2</td>
<td>41.0</td>
<td>16.0</td>
<td>14.4</td>
</tr>
</tbody>
</table>

*Calculated using 6-311++G(3df,3pd) basis. The results for floating orbitals are given in italics. 6-311++G(3df,3pd).*

---

**Table V: Properties Calculated for H2O**

<table>
<thead>
<tr>
<th>Property</th>
<th>DZ</th>
<th>DZ+</th>
<th>DZP</th>
<th>DZP+</th>
</tr>
</thead>
<tbody>
<tr>
<td>rOH, Å</td>
<td>0.949</td>
<td>0.950</td>
<td>0.941</td>
<td>0.941</td>
</tr>
<tr>
<td>dOH, deg</td>
<td>111.9</td>
<td>112.9</td>
<td>105.6</td>
<td>106.2</td>
</tr>
<tr>
<td>dipole moment, au</td>
<td>105.1</td>
<td>106.6</td>
<td>105.3</td>
<td>106.2</td>
</tr>
<tr>
<td>polarizabilities</td>
<td>0.977</td>
<td>1.016</td>
<td>0.835</td>
<td>0.874</td>
</tr>
<tr>
<td>polarizabilities</td>
<td>0.753</td>
<td>0.785</td>
<td>0.731</td>
<td>0.777</td>
</tr>
<tr>
<td>frequencies, cm⁻¹</td>
<td>6.56</td>
<td>7.34</td>
<td>6.74</td>
<td>7.26</td>
</tr>
<tr>
<td>in-plane σ</td>
<td>8.27</td>
<td>8.57</td>
<td>8.12</td>
<td>8.41</td>
</tr>
<tr>
<td>in-plane π</td>
<td>3.70</td>
<td>4.87</td>
<td>4.90</td>
<td>5.78</td>
</tr>
<tr>
<td>out-of-plane</td>
<td>1.29</td>
<td>3.86</td>
<td>2.58</td>
<td>5.22</td>
</tr>
<tr>
<td>intensities, km/mol</td>
<td>6.16</td>
<td>7.06</td>
<td>6.10</td>
<td>7.26</td>
</tr>
</tbody>
</table>

---

frequencies) and electrical (dipole moments, polarizabilities, and infrared intensities). The geometrical properties are well described at all levels except fixed DZ and DZ+. The improvements obtained by floating are small, particularly for bond distances. Floating DZ gives results slightly inferior to fixed DZ but is still an improvement on the fixed DZ results.

The situation is different for electrical properties, which are poorly described at the fixed-orbital level: the best agreement is obtained for the DZP basis dipole moment values, which differ from the Hartree–Fock limit values by some 7%. Floating produces drastic improvements in all cases. Indeed, for all the properties considered even the smallest floating basis set gives properties which are similar to or better than the best fixed basis set. Dipole moments at the fixed DZ level are in error by about 20%, compared to 3–4% when orbital floating is allowed. With addition of diffuse functions the computed values are somewhat larger than the Hartree–Fock limit, while the inclusion of both diffuse orbitals and polarization functions (floating DZP+) gives agreement with the Hartree–Fock limit to within 1%. Since experimental dipole moments are usually a few percent smaller than the Hartree–Fock limit, a floating DZ basis often gives dipole moments which are in fortuitously good agreement with experiment, as noted by Huber.21

The requirements for a fixed-orbital basis for calculating polarizabilities are quite demanding, and it is not surprising that even the largest fixed basis set (DZP+) produces polarizabilities which are 15% too low. When floating is allowed, the errors are 7% or smaller in all cases. In particular, when diffuse functions are added, the agreement with the Hartree–Fock limit is remarkably good.

---


An Open-Shell Spin-Restricted Coupled Cluster Method: Application to Ionization Potentials in \( \text{N}_2 \)

Magnus Rittby and Rodney J. Bartlett*

Quantum Theory Project, University of Florida, Gainesville, Florida 32611 (Received: July 15, 1987; In Final Form: October 1, 1987)

In order to circumvent the problem of spin contamination in unrestricted Hartree–Fock based coupled cluster (CC) calculations, we present a new method of calculation for certain classes of open-shell systems. The approach ensures that the proper spin component of the resulting correlated wave function is projected out in the energy evaluation by the use of a reference function constructed from suitably chosen restricted open-shell Hartree–Fock or other orbitals. This single-reference open-shell spin-restricted CC method is applied to the calculation of ionization potentials in the \( \text{N}_2 \) molecule, and it is shown that highly accurate results can be obtained in a 5s4p1d basis. The mean error for all the principal ionization potentials of \( \text{N}_2 \) compared to experiment is 0.45%.

---

Introduction

In single-determinant reference coupled cluster (CC)\(^2\)–\(^5\) and many-body perturbation theory (MBPT) methods,\(^6\)\(^,\)\(^7\) the use of an unrestricted Hartree–Fock (UHF) wave function for open-shell systems is frequently employed.\(^2\)\(^,\)\(^4\) A UHF reference generally leads to a lowering of the total energy at the Hartree–Fock (HF) level since the \( \alpha \)-spin electrons are chosen to occupy different spatial orbitals than the \( \beta \)-spin electrons. This procedure introduces important spin-polarization effects on the spin density, for example, while the corresponding open-shell RHF results would erroneously give zero.\(^8\) Also, UHF functions facilitate the correct separation of a single-determinant wave function in bond breaking.\(^2\)\(^,\)\(^3\) However, there are two prices paid for the additional flexibility introduced by a UHF function. First, the wave function is not an eigenfunction of \( \hat{S}^2 \) and second, having to treat different

---


