

Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen

Thom H. Dunning, Jr.

Theoretical & Computational Chemistry Group, Chemistry Division, Argonne National Laboratory,
Argonne, Illinois 60439

(Received 26 July 1988; accepted 28 September 1988)

In the past, basis sets for use in correlated molecular calculations have largely been taken from single configuration calculations. Recently, Almlöf, Taylor, and co-workers have found that basis sets of natural orbitals derived from correlated atomic calculations (ANOs) provide an excellent description of molecular correlation effects. We report here a careful study of correlation effects in the oxygen atom, establishing that compact sets of primitive Gaussian functions effectively and efficiently describe correlation effects *if* the exponents of the functions are optimized in atomic correlated calculations, although the primitive (*sp*) functions for describing correlation effects can be taken from atomic Hartree-Fock calculations *if* the appropriate primitive set is used. Test calculations on oxygen-containing molecules indicate that these primitive basis sets describe molecular correlation effects as well as the ANO sets of Almlöf and Taylor. Guided by the calculations on oxygen, basis sets for use in correlated atomic and molecular calculations were developed for all of the first row atoms from boron through neon and for hydrogen. As in the oxygen atom calculations, it was found that the incremental energy lowerings due to the addition of correlating functions fall into distinct groups. This leads to the concept of *correlation consistent basis sets*, i.e., sets which include all functions in a given group as well as all functions in any higher groups. Correlation consistent sets are given for all of the atoms considered. The most accurate sets determined in this way, [5s4p3d 2f1g], consistently yield 99% of the correlation energy obtained with the corresponding ANO sets, even though the latter contains 50% more primitive functions and twice as many primitive polarization functions. It is estimated that this set yields 94%–97% of the total (HF + 1 + 2) correlation energy for the atoms neon through boron.

I. INTRODUCTION

The quantitative theoretical prediction of barrier heights, bond energies, ionization potentials, electron affinities, etc., requires that electron correlation effects be included in the electronic structure calculations. For example, Hartree-Fock (HF) calculations on the $H + H_2$ reaction predict¹ a barrier to reaction of 24.5 kcal/mol. The most accurate correlated calculations to date² yield 9.65 kcal/mol which is estimated to be no more than 0.1 kcal/mol above the true barrier. Similarly, single configuration calculations predict a dissociation energy of 102.4 kcal/mol for hydrogen fluoride,³ whereas the value extrapolated from spectroscopic data⁴ is 141.2 kcal/mol. Finally, HF calculations predict an ionization potential of 10.07 eV for the oxygen atom,⁵ compared to the experimental ionization potential⁶ of 13.62 eV, and an electron affinity of -0.54 eV⁷ which is in error by 2.0 eV (the measured value is $+1.46$ eV⁸). Clearly, in any chemical process in which the number of electrons changes, correlation effects must be explicitly taken into account if accurate energy differences are desired.⁹

One of the most important correlation effects, the so-called near-degeneracy effect,¹⁰ can be readily taken into account using multiconfiguration self-consistent-field (MCSCF) wave functions.¹¹ Inclusion of these effects often dramatically improves the accuracy of the calculations especially for multiply bonded systems. For example, for N_2 a single configuration calculation predicts¹² a dissociation en-

ergy of 121.7 kcal/mol, only slightly more than half of the experimental value, 228.4 kcal/mol.¹³ A multiconfiguration wave function designed to properly describe dissociation to two (HF) nitrogen atoms predicts¹⁴ a binding energy of 167.6 kcal/mol, reducing the error in the HF estimate by 40%, while a wave function constructed from only those configurations which can be formed from the valence 2s- and 2p-atomic orbitals yields 208.9 kcal/mol,¹⁵ which is in error by less than 10%.

The accuracy of the electronic structure calculations is limited not only by the form of the wave function but also by the basis set used to expand the wave function. The choice of basis functions for molecular HF calculations is straightforward¹⁶: The (*sp*) sets are taken from optimized atomic calculations while the polarization (*d, f, ...*) sets are constructed using well established *rules of thumb*¹⁶ or by explicit optimization. Convergence of the results with increasing basis set size is reasonably rapid so that modest care in the choice of basis functions leads to accurate (within the model) results (although there are pathological cases which require more careful consideration). Experience indicates that the basis sets required to properly describe MCSCF wave functions designed to account for near-degeneracy effects are very similar to the sets required for single configuration calculations.

Although valence space MCSCF calculations on the $H + H_2$ reaction reduce the error in the calculated barrier height¹ by over 7.0 kcal/mol, to 17.1 kcal/mol, the predicted

barrier is still nearly a factor of 2 higher than that from the most accurate calculations available.² The remaining error is due to nondynamical and dynamical correlation effects.⁹ First-order¹⁷ (FO) or polarization configuration interaction¹⁸ (Pol-CI) calculations, which include configurations which have at most one electron in a nonvalence orbital, take into account nondynamical effects such as space and spin polarization. Inclusion of nondynamical correlation effects can also lead to a significant improvement in the accuracy of the calculations. For example, Pol-CI/FO calculations yield a barrier of 12.1 kcal/mol¹ for the $\text{H} + \text{H}_2$ reaction, a bond energy of 132.1 kcal/mol for HF,¹⁹ and an ionization potential of 11.57 eV and electron affinity of 1.12 eV for the oxygen atom.^{17(b)}

The description of nondynamical correlation effects requires that functions with higher angular momenta than that of the occupied atomic HF orbitals be included in the basis set even for atomic calculations. The choice of basis sets for use in FO/Pol-CI calculations has been discussed by Schaefer *et al.*^{17(b)} In practice, the basis sets used in FO/Pol-CI calculations on molecules are often taken to be the same as those used in single- and multiconfiguration calculations. Convergence appears to be adequate if reasonably large HF basis sets are used.

The remaining error in the calculated energies is due to dynamical correlation effects, i.e., to the instantaneous correlation in the motions of the electrons. It is the description of these effects which is of interest here. The theoretical description of dynamical effects has proven to be one of the most challenging problems in modern electronic structure theory. In configuration interaction theories dynamical correlation effects are described by configurations with two or more electrons in the nonvalence orbitals. As the number of such configurations increases rapidly with the size of the basis set, the goal is to minimize the number of basis functions required to achieve a given level of accuracy. Although basis sets to describe dynamical correlation effects have been considered by others,²⁰⁻²³ the only detailed studies to date are those by Ahlrichs and co-workers²² and Almlöf and Taylor.²³ The latter authors took the novel approach of taking the basis functions to be the natural orbitals obtained from correlated atomic calculations. From these studies the following conclusions were drawn:

- (i) Basis sets which include functions with high angular momenta, (d, f, g, \dots), are required to reduce the error in the correlation energy to 1 kcal/mol or less.
- (ii) The basis functions could be grouped into sets with each function in the set lowering the correlation energy by an approximately equal amount²² or falling within a given range of occupation numbers.²³

In addition, and most important, Almlöf and Taylor²³ found that basis functions optimized to describe correlation effects in atoms also describe molecular correlation effects well.

The general conclusions drawn from the present study, which includes a thorough study of basis sets and correlation effects in the oxygen atom, are in agreement with the above. In addition, we show that compact sets of primitive Gaus-

sian functions can be obtained to describe correlation effects in all of the first row atoms from boron through neon. The (sp) sets can be obtained from atomic HF calculations, while the polarization²⁴ (d, f, g, \dots) sets must be determined from correlated atomic calculations. The energies obtained with these primitive sets compare well with those yielded by the atomic natural orbital (ANO) sets of Almlöf and Taylor²³ which contain many more primitive functions.

In the following section, we report an extensive series of correlated calculations on the oxygen atom which was chosen to be the benchmark system. For this atom we determined the convergence of the correlation energy both with respect to the angular momenta of the functions and the number of functions of a given angular momentum. We then carried out calculations on the OH and O₂ molecules to determine the suitability of the atom-derived sets for use in molecular calculations. In Sec. III we report basis sets for use in correlated molecular calculations for all of the first row atoms from boron through neon. The results obtained for these atoms are consistent with those for the oxygen atom and give us confidence that these sets are also suitable for use in correlated molecular calculations. Finally, we report comparable basis sets for the hydrogen atom.

II. CORRELATED BASIS SETS FOR THE OXYGEN ATOM

An investigation of the basis sets to be used in correlated calculations which considered in detail all of the atoms of interest would be tedious, if not overindulgent. Therefore, we first carried out a thorough study of basis sets for use in configuration interaction (CI) calculations on the oxygen atom. The HF description of this atom has a singlet-coupled pair of electrons in the $2s$ orbital, another singlet-coupled pair in a $2p$ orbital and a triplet-coupled electron pair in $2p$ orbitals. It therefore has features which are representative of all of the first-row atoms. The results of this study were used to guide the calculations on the remaining first row atoms. The convergence patterns, energy lowerings, etc., observed in the latter calculations are in line with those found for oxygen, indicating that the general conclusions drawn from the oxygen calculations are applicable to all of the first row atoms.

The studies reported here are based on CI calculations which include all symmetry-allowed single and double excitations from all of the degenerate components of the Hartree-Fock configuration ($\text{HF} + 1 + 2$) of the atom, e.g., from the $2s^2 2p_x^2 2p_y^2 p_z$, $2s^2 2p_x 2p_y^2 2p_z$, and $2s^2 2p_x 2p_y 2p_z^2$ configurations of the oxygen atom. The $1s$ atomic orbital was, however, constrained to be doubly occupied in all configurations—we are focusing here only on valence correlation effects. The HF wave functions were computed using effective Hamiltonians properly averaged to produce eigenfunctions of both space and spin.²⁵

A. Polarization basis sets for the oxygen atom

We first considered the convergence of the correlation energy for the (d, f, g, \dots) polarization sets. For these calculations the (sp) set was a ($16s7p$) primitive set²⁶ contracted to $[6s4p]$ using the general contraction scheme of Raffanetti²⁸:

The first function(s) in each symmetry were the atomic HF orbitals, while the remaining functions were the most diffuse Gaussian primitives. The polarization functions were taken to be simple primitive Gaussian functions with exponents given by even-tempered expansions, i.e.,

$$\zeta_i = \alpha\beta^{i-1} \quad i = 1, \dots, N_f, \quad (1)$$

where N_f is the number of functions in the set. (α, β) were optimized for each set.

The polarization sets were added in symmetry (angular momentum) shells.²⁹ First, we added 3*d* functions to the [6*s*4*p*] set. 3*d* sets containing from one to four functions were considered. We then added 4*f* functions to the [6*s*4*p*3*d*] set obtained from these calculations. 4*f* sets containing from one to three functions were considered. Finally, we added 5*g* functions to the [6*s*4*p*3*d* 2*f*] set. Sets containing both one and two 5*g* functions were considered. The Gaussian integral program used here, ARGOS,³⁰ does not compute integrals over 6*h* and higher angular momentum functions. We were therefore not able to explicitly consider these functions; however, their effect could be estimated (see below). The results of these calculations are summarized in Table I and plotted in Fig. 1.

Two important facts are clearly evident in the results reported in Table I and Fig. 1. First, the incremental correlation energy lowerings resulting from the addition of polarization functions of a given symmetry decrease dramatically as successive functions are added to the set, e.g., the correlation energy increases by -62.2 mh^{31} (mh = millihartree) with the addition of the first 3*d* function, by an additional

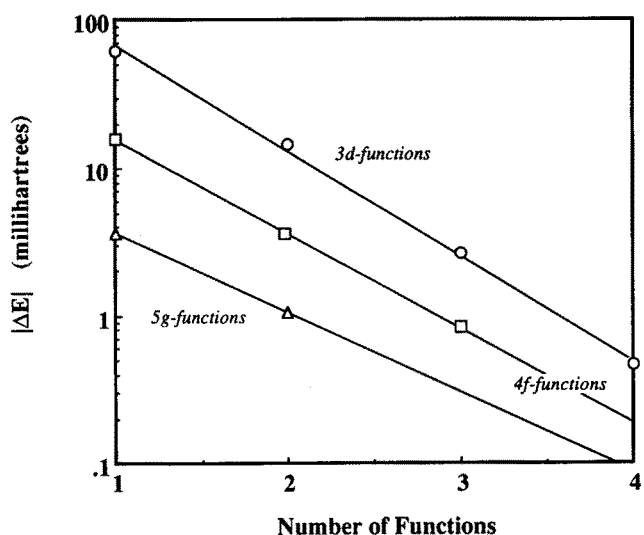


FIG. 1. Absolute values of the incremental correlation energy lowerings for the 3*d*, 4*f*, and 5*g* polarization sets for the oxygen atom. The straight lines are geometric series fits of the calculated points.

-14.7 mh with the second 3*d* function, by another -2.6 mh with the third 3*d* function, and by only -0.5 mh with the fourth 3*d* function. The energy lowerings for all angular momenta decrease approximately geometrically, i.e.,

$$\widetilde{\Delta E}_{3d}^n = \frac{-67.81}{5.153^{n-1}} \quad \widetilde{\Delta E}_{4f}^n = \frac{-15.56}{4.351^{n-1}} \quad \widetilde{\Delta E}_{5g}^n = \frac{-3.658}{3.454^{n-1}}, \quad (2)$$

where $\widetilde{\Delta E}_l^n$ is the (estimated) incremental change in the correlation energy, i.e., the difference in correlation energy between the *l* set with *n* functions and the set with (*n* - 1) functions. An estimate of the total contribution of the 3*d*, 4*f*, and 5*g* functions to the correlation energy can be obtained by adding to the calculated energy lowerings the remainders estimated from Eq. (2), e.g.,

$$\begin{aligned} \Delta E_{3d}(\text{total}) \approx & \Delta E_{3d}^1 + \Delta E_{3d}^2 + \Delta E_{3d}^3 \\ & + \Delta E_{3d}^4 + \sum_{n=5}^{\infty} \widetilde{\Delta E}_{3d}^n. \end{aligned} \quad (3)$$

This yields

$$\Delta E_{3d}(\text{total}) \approx -80.2, \quad \Delta E_{4f}(\text{total}) \approx -20.4, \quad (4)$$

$$\Delta E_{5g}(\text{total}) \approx -5.1,$$

(in mh). Thus, addition of a single 3*d* function accounts for 78% of the estimated total energy lowering due to the addition of 3*d* functions while the second 3*d* function contributes an additional 18%. Addition of the first three 3*d* functions accounts for over 99% of the estimated total energy lowering due to 3*d* functions. Similar convergence patterns are found for the higher angular momentum functions, although the data is less complete for these sets.

The reliability of the above estimates is not known. For the 3*d* set for which the most extensive data is available, the calculated energy increments appear to bow downward slightly relative to the geometric fit. If this trend were to continue, the geometric series would overestimate the con-

TABLE I. Total energies, correlation energies, and energy lowerings from HF + 1 + 2 calculations on the oxygen atom using optimized even-tempered polarization sets. The (*sp*) set was the (16*s*7*p*) set contracted to [6*s*4*p*]. The exponents in each set are given by an even-tempered expansion with parameters (α, β). Total energies ($E_{\text{HF}+1+2}$) are in hartrees; correlation energies and energy lowerings (E_{corr} and ΔE_{corr}) are in millihartrees. For all sets the HF energy is $-74.809\,068$ hartrees.

Polarization set	(α, β)	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
...	...	-74.885 656 ^b	-76.588	...
(1 <i>d</i>)	1.185	-74.947 895	-138.827	-62.239
(2 <i>d</i>)	(0.645, 3.588)	-74.962 612	-153.544	-14.717
(3 <i>d</i>)	(0.447, 2.924)	-74.965 253	-156.185	-2.641
(4 <i>d</i>)	(0.334, 2.521)	-74.965 720	-156.652	-0.467
(1 <i>f</i>) ^c	1.428	-74.981 102	-172.034	-15.849
(2 <i>f</i>) ^c	(0.859, 3.104)	-74.984 545	-175.477	-3.443
(3 <i>f</i>) ^c	(0.611, 2.652)	-74.985 382	-176.314	-0.837
(1 <i>g</i>) ^d	1.849	-74.988 203	-179.135	-3.658
(2 <i>g</i>) ^d	(1.114, 2.870)	-74.989 262	-180.194	-1.059

^a Change in the correlation energy referenced to the preceding set with one fewer function of the given angular momentum.

^b Energy for the (16*s*7*p*)/[6*s*4*p*] set.

^c The (*sp*) basis set was augmented with the above (3*d*) primitive set.

^d The (*sp*) basis set was augmented with the above (3*d*) and (2*f*) primitive sets.

tribution of the higher order terms. However, the resulting error is estimated to be -0.03 to -0.04 mh (the geometric series predicts that the $n \geq 5$ terms will contribute only -0.12 mh). The estimates of the higher order terms will be less reliable for the $4f$ and $5g$ sets since the data here is limited and the estimated contributions of the neglected higher order terms are larger, -0.25 mh ($4f$) and -0.43 mh ($5g$). In any case, however, it appears that the above should be accurate to ± 0.1 mh.

Second, as noted earlier by Ahlrichs and co-workers,²¹ the calculated energy lowerings, and therefore the correlating functions, separate into distinct groups. The addition of a single $3d$ function increases the correlation energy by -62.2 mh; no other single function has such a dramatic effect. Addition of the second $3d$ function, on the other hand, increases the correlation energy by -14.7 mh, which is very nearly the same lowering resulting from the addition of the first $4f$ function, namely, -15.8 mh. Again, none of the remaining functions has such a large effect. Next, the third $3d$ function lowers the energy by -2.6 mh while the second $4f$ function and first $5g$ function increase the correlation energy by -3.4 and -3.7 mh, respectively. Finally, the fourth $3d$ function, third $4f$ function, second $5g$ function (as well as the first $6h$ function; see below) have nearly the same effect with the calculated (or estimated) energy lowerings ranging from -0.5 to -1.0 mh. This finding suggests that the polarization functions should be added in sets, yielding $(1d)$, $(2d\ 1f)$, $(3d\ 2f\ 1g)$, etc., sets, to provide a consistent treatment of correlation effects. This approach is also advocated by Almlöf and Taylor²³ who construct such sets by grouping together the atomic natural orbitals with occupation numbers which fell in a given range.

In Fig. 2 we plot the calculated energy lowerings resulting from addition of the first, second, and third functions of a given angular momentum l . These lowerings also decrease approximately geometrically with

$$\widetilde{\Delta E}_l^1 = \frac{-63.26}{4.125^{l-2}}, \quad \widetilde{\Delta E}_l^2 = \frac{-14.06}{3.728^{l-2}}, \quad \widetilde{\Delta E}_l^3 = \frac{-2.642}{3.155^{l-2}}. \quad (5)$$

From Eqs. (5) one can estimate that the addition of the first $6h$ function will increase the correlation energy by -0.90 mh while addition of the second $6h$ function and first $7i$ function will increase the correlation energy by an additional -0.27 and -0.22 mh, respectively. As before, we can use Eqs. (5) to estimate the contributions of the higher angular momentum functions to the total correlation energy. Doing so yields

$$\begin{aligned} \Delta E_{6h}^1 + \Delta E_{7i}^1 + \Delta E_{8k}^1 + \cdots &= -1.2, \\ \Delta E_{6h}^2 + \Delta E_{7i}^2 + \Delta E_{8k}^2 + \cdots &= -0.4, \\ \Delta E_{5g}^3 + \Delta E_{6h}^3 + \Delta E_{7i}^3 + \cdots &= -0.4, \end{aligned} \quad (6)$$

(in mh).

We can now estimate the *total* contribution of the polarization functions to the correlation energy as well as the accuracy of the *correlation consistent* groupings of the polarization functions. Combining the calculated energy differences given in Table I with the estimates of the higher order contributions given by Eqs. (2) and (5), we predict that polariza-

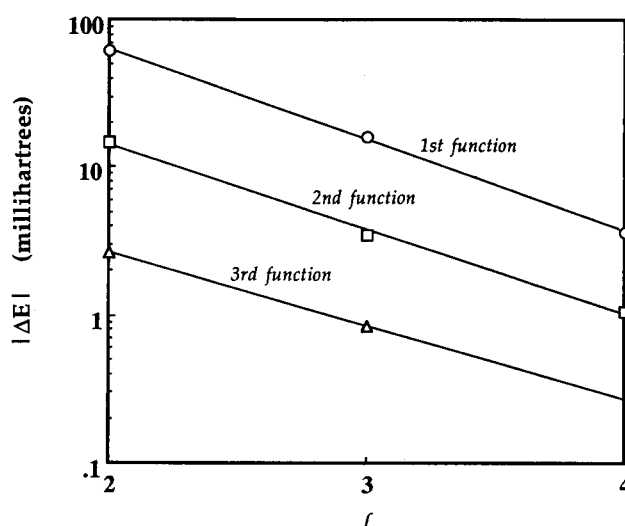


FIG. 2. Absolute values of the incremental correlation energy lowerings for the first, second, and third functions of angular momentum l for the oxygen atom. The straight lines are geometric fits of the calculated points.

tion functions contribute approximately -107.4 mh to the correlation energy of the oxygen atom (for the $HF + 1 + 2$ calculations reported here). Assuming this, we see that the $(1d)$ set accounts for $\sim 58\%$ of the total correlation energy due to polarization functions, the $(2d\ 1f)$ set $\sim 86\%$, and the $(3d\ 2f\ 1g)$ set $\sim 95\%$. The error in the correlation energy for the $(3d\ 2f\ 1g)$ set is estimated to be 4.8 mh. Addition of one more $3d$, $4f$, and $5g$ function along with a $6h$ function will reduce this error by ~ 3.2 to 1.6 mh. Thus, for the oxygen atom a $(4d\ 3f\ 2g\ 1h)$ polarization set is estimated to yield a total $(HF + 1 + 2)$ correlation energy within 1 kcal/mol of the infinite (polarization) basis set limit.

TABLE II. Comparison of the correlation energies obtained with the present optimized (even-tempered) primitive polarization sets and the atomic natural orbital (ANO) sets of Almlöf and Taylor (Ref. 23). The (sp) set for both sets of calculations was the $(16s7p)$ set contracted to $[6s4p]$. Total energies (E_{HF+1+2}) are in hartrees; correlation energy differences (ΔE_{corr}) are in millihartrees. For all basis sets the HF energy is $-74.809\ 068$ hartrees.

Polarization set	ANO E_{HF+1+2}	Present E_{HF+1+2}	ΔE_{corr}
$(1d)$	-74.956 479	-74.947 895	8.584
$(2d)$	-74.964 171	-74.962 612	1.559
$(3d)$	-74.965 450	-74.965 253	0.197
$(4d)$	-74.965 787	-74.965 720	0.067
$(1f)^a$	-74.982 379	-74.981 102	1.277
$(2f)^a$	-74.984 788	-74.984 545	0.243
$(3f)^a$	-74.985 335	-74.985 382	-0.047
$(1g)^b$	-74.988 425	-74.988 203	0.222
$(2g)^b$	-74.989 239	-74.989 262	-0.023

^a The (sp) basis set was augmented with the present $(3d)$ primitive set.

^b The (sp) basis set was augmented with the present $(3d)$ and $(2f)$ primitive sets.

TABLE III. Comparison of atom optimized polarization sets with molecule optimized polarization sets for calculations on the hydroxyl radical, OH, and oxygen molecule, O₂. The (*sp*) set for both series of calculations was the (12*sp*) set contracted to [4*s*3*p*]; for the hydrogen atom in OH we used the (5*s*1*p*) set contracted to [3*s*1*p*]; see Sec. IV. Total energies (E_{HF} and $E_{\text{HF}+1+2}$) are in hartrees; correlation energies and energy differences (E_{corr} and ΔE_{corr}) are in millihartrees.

Polarization set	$\zeta(d; f)$	E_{HF}	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}
The OH radical					
(1 <i>d</i>) atom	1.185	-75.415 680	-75.589 794	-174.114	
molecule	1.206	-75.415 660	-75.589 807	-174.147	-0.033
(2 <i>d</i> 1 <i>f</i>) atom	(0.645, 2.314; 1.428)	-75.418 246	-75.626 066	-207.820	
molecule	(0.647, 2.261; 1.383)	-75.418 250	-75.626 078	-207.828	-0.008
The O ₂ molecule					
(1 <i>d</i>) atom	1.185	-149.653 616	-150.012 432	-358.816	
molecule	1.149	-149.653 906	-150.012 539	-358.633	0.183
(2 <i>d</i> 1 <i>f</i>) atom	(0.645, 2.314; 1.428)	-149.662 336	-150.081 282	-418.946	
molecule	(0.699, 2.256; 1.367)	-149.662 381	-150.081 331	-418.950	-0.004

In Table II we compare the energies obtained with the present polarization sets with the sets of Almlöf and Taylor.²³ The latter sets are derived from the natural orbitals obtained from HF + 1 + 2 calculations on the atoms

(ANOs) and are ordered by occupation number. The ANOs are expanded in sets containing six primitive functions in the 3*d* set, four primitives in the 4*f* set, and two primitives in the 5*g* set. Not unexpectedly, the differences are larger for the

TABLE IV. Comparison of the HF and correlation energies obtained with the present atomic optimized (even-tempered) polarization sets and the ANO sets of Almlöf and Taylor (Ref. 23) for the oxygen molecule, O₂. The (*sp*) set for both calculations was the (16*s*7*p*) set contracted to [6*s*4*p*]. Total energies E_{HF} and $E_{\text{HF}+1+2}$ are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{HF} and ΔE_{corr}) are in millihartrees.

Polarization set	E_{HF}	ΔE_{HF}^a	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
Present sets					
(1 <i>d</i>)	-149.657 559		-150.032 181	-374.622	
(2 <i>d</i>)	-149.660 385		-150.061 386	-401.001	
(3 <i>d</i>)	-149.661 785		-150.068 444	-406.659	
(1 <i>f</i>) ^b	-149.666 103		-150.105 117	-439.014	
(2 <i>f</i>) ^b	-149.666 107		-150.111 583	-445.476	
(1 <i>g</i>) ^c	-149.666 462		-150.121 381	-454.919	
ANO sets					
(6 <i>d</i>)/[1 <i>d</i>]	-149.660 466	-2.907	-150.047 711	-387.245	-12.623
(6 <i>d</i>)/[2 <i>d</i>]	-149.661 439	-1.054	-150.064 952	-403.513	-2.512
(6 <i>d</i>)/[3 <i>d</i>]	-149.661 803	-0.018	-150.068 722	-406.919	-0.260
(4 <i>f</i>)/[1 <i>f</i>] ^b	-149.666 083	0.020	-150.106 924	-440.841	-1.827
(4 <i>f</i>)/[2 <i>f</i>] ^b	-149.666 129	-0.022	-150.112 165	-446.036	-0.560
(2 <i>g</i>)/[1 <i>g</i>] ^c	-149.666 471	-0.009	-150.121 901	-455.430	-0.511

^a The HF and correlation energy differences are relative to the atom optimized even-tempered set with the same number of functions.

^b The (*sp*) basis set was augmented with the present (3*d*) primitive set.

^c The (*sp*) basis set was augmented with the present (3*d*) and (2*f*) primitive sets.

smaller primitive sets with lower angular momenta, i.e., (1*d*), (2*d*), and (1*f*). The errors are clearly unimportant for the (3*d*), (2*f*), and (1*g*) sets for each of which $\Delta E_{\text{corr}} \approx 0.2$ mh. For these (or larger) sets there seems to be little reason to use the ANO contracted sets. Even the small sets account for a substantial percentage of the correlation energy obtained with the corresponding ANO sets. Thus, the (1*d*) set yields 87.9% of the correlation energy obtained with the (6*d*)/[1*d*] ANO set, while the (2*d*) and (1*f*) sets yield 98.0% and 92.5%, respectively, of the correlation energy obtained with the (6*d*)/[2*d*] and (4*f*)/[1*f*] ANO sets. Thus, in any case the error resulting from the use of primitive sets is a small fraction of the error in the correlation energy.

The objective of the present studies is, of course, to determine polarization sets for use in molecular calculations. To assess the appropriateness of the above atomic sets for molecular calculations, we carried out calculations on the hydroxyl radical, OH, and oxygen molecule, O₂. Two series of calculations were carried out. In the first series we used a (12*s*6*p*) set contracted to [4*s*3*p*]²⁶ (again using the general contraction scheme of Raffinetti²⁸) augmented with the atomic (1*d*) and (2*d* 1*f*) sets determined above; in the second series, we optimized the exponents of the polarization functions in HF + 1 + 2 calculations on each molecule. The exponents and energies so obtained are listed in Table III. As can be seen, the optimum atomic and molecular exponents differ by only a few percent and the changes in the calculated correlation energies (ΔE_{corr}) are entirely negligible. In fact, for the (1*d*) set for O₂ the correlation energy obtained with the molecule optimized exponents is less than for the atom optimized exponents. This results from a lowering of the molecular HF energy which more than offsets the lowering of the HF + 1 + 2 energy; the CI energy is, of course, lower for the molecule optimized set as it has to be.

As a final check on the suitability of the present sets for use in molecular calculations, we carried out calculations on the oxygen molecule, O₂, using the ANO sets of Almlöf and Taylor²³ as well as the optimum atomic even-tempered sets. The results of the calculations are summarized in Table IV. The patterns observed in these calculations are quite similar to those found in the atomic calculations. In most cases, the differences between the optimum primitive sets and the ANO contracted sets, normalized to the number of oxygen atoms, are smaller in the molecular calculations than in the atomic calculations, e.g., for the (1*d*) set the respective differences are 6.3 mh (O₂) and 8.6 mh (O).

B. (*sp*) basis sets for the oxygen atom

Next, we considered *s* and *p* basis sets for use in correlated calculations on the oxygen atom. The calculations for the *s*-correlation set were based on the (16*s*7*p*) set contracted to [2*s*4*p*], i.e., the base *s* set consisted of only the 1*s* and 2*s* atomic orbitals. To the [2*s*4*p*] set we then added sets of primitive 1*s* functions with exponents optimized for the HF + 1 + 2 wave function. Again, the exponents of the functions in each set were taken to be even-tempered expansions. Expansions sets consisting of one to four functions were considered. The calculations for the *p* set were based on

TABLE V. Total energies, correlation energies, and energy lowerings from HF + 1 + 2 calculations on the oxygen atom using atom optimized even-tempered (*sp*) sets. For the 1*s*-correlation sets the (*sp*) set was the (16*s*7*p*) set contracted to [2*s*4*p*]; for the 2*p* sets the (16*s*7*p*) set was contracted to [6*s*1*p*]; and, for the (1*s* + 2*p*) sets the (16*s*7*p*) set was contracted to [2*s*1*p*]. The polarization set was the (2*d* 1*f*) set. Total energies ($E_{\text{HF}+1+2}$) are in hartrees; correlation energies (E_{corr}) are in millihartrees. For all sets the HF energy is -74.809 068 hartrees.

(<i>sp</i>) set	(α, β)	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
...	...	-74.946 896 ^b	-137.828	...
(1 <i>s</i>)	0.295	-74.973 127	-164.059	-26.231
(2 <i>s</i>)	(0.253, 7.461)	-74.977 438	-168.370	-4.311
(3 <i>s</i>)	(0.236, 2.769)	-74.977 947	-168.879	-0.509
(4 <i>s</i>)	(0.184, 2.465)	-74.978 288	-169.220	-0.341
...	...	-74.905 090 ^c	-96.022	...
(1 <i>p</i>)	0.289	-74.962 965	-153.897	-57.874
(2 <i>p</i>)	(0.217, 3.225)	-74.975 032	-165.964	-12.067
(3 <i>p</i>)	(0.176, 2.971)	-74.978 614	-169.546	-3.582
(4 <i>p</i>)	(0.138, 2.502)	-74.979 379	-170.311	-0.765
...	...	-74.897 438 ^d	-88.370	...
(1 <i>s</i> 1 <i>p</i>)	same as above	-74.959 732	-150.664	-62.294
(2 <i>s</i> 2 <i>p</i>)	same as above	-74.974 256	-165.188	-14.524
(3 <i>s</i> 3 <i>p</i>)	same as above	-74.978 358	-169.290	-4.102
(4 <i>s</i> 4 <i>p</i>)	same as above	-74.979 427	-170.359	-1.069

^a Change in the correlation energy referenced to the preceding set with one fewer function of the given angular momentum.

^b Energy for the [2*s*4*p*2*d* 1*f*] set.

^c Energy for the [6*s*1*p*2*d* 1*f*] set.

^d Energy for the [2*s*1*p*2*d* 1*f*] set.

the (16*s*7*p*) set contracted to [6*s*1*p*], i.e., the base *p* set consisted of just the atomic 2*p* orbital. We then added even-tempered 2*p* sets to the [6*s*1*p*] set and optimized (α, β) for each such set. Again, sets consisting of one to four functions were considered. A (2*d* 1*f*) set was added to each of the above (*sp*) sets to describe the correlation effects due to polarization functions. This approach has the advantage that the added 1*s* and 2*p* functions effectively contribute only to the correlation energy; the atomic HF orbitals are sufficiently well described by the (16*s*7*p*) set such that changes in the HF energy are negligible, <0.001 mh.

The results of the (*sp*) correlated calculations are summarized in Table V and the corresponding energy lowerings are plotted in Fig. 3. It is evident in both Table V and Fig. 3 that 2*p* correlation functions are more important than 1*s* correlation functions and that the convergence of the (*np*) series is quite similar to that observed for the polarization sets discussed above. In fact, as can be seen in the figure, the incremental energy lowerings for 2*p* series fit a geometrical progression well:

$$\Delta E_{2p}^n \approx \frac{-55.60}{4.134^{n-1}}. \quad (7)$$

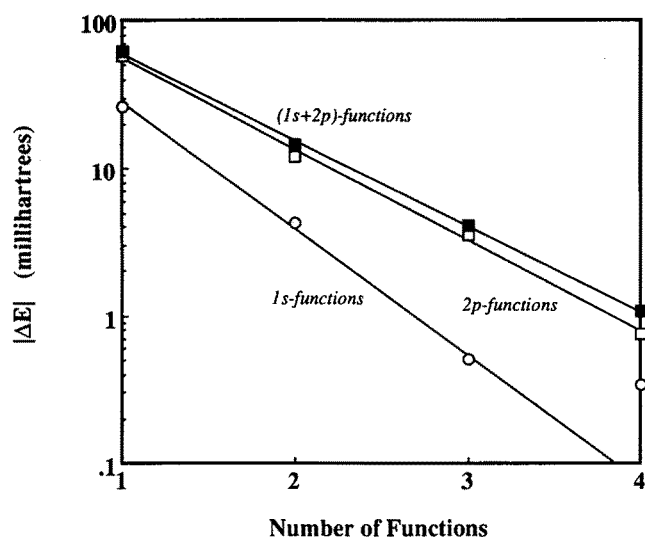


FIG. 3. Absolute values of the incremental correlation energy lowerings for the 1s, 2p, and (1s + 2p) polarization sets for the oxygen atom. The straight lines are geometrical series fits of the calculated points.

The last point in the 1s series, on the other hand, is significantly higher than would have been predicted from a geometrical progression based on the first three terms.

Although 1s-correlation functions are less important than the 2p-correlation functions for describing valence correlation effects, we believe it best to symmetrically augment the HF (*sp*) set, i.e., to simultaneously add both 1s- and 2p-correlation functions. This is more straightforward than a procedure based solely on the energy lowerings: First add a

2p function, then add a 1s function, then add another 2p function, then add both a 1s and 2p function, etc. In addition, there is strong coupling between the 1s and 2p sets, i.e., the lowering due to the addition of 2p functions strongly depends on the 1s set used and vice versa (the coupling is much weaker for the polarization functions). This procedure results in the use of, at most, one extra function. In Fig. 3 we have also plotted the energy lowerings obtained by simultaneously adding 1s and 2p functions (see Table V). The calculated incremental changes also fit a geometrical progression with

$$\tilde{\Delta E}_{1s+2p}^n = \frac{-59.76}{3.842^{n-1}}. \quad (8)$$

Combining the higher order terms in Eq. (8) with the calculated correlation energy for the (4s4p) set, we estimate the total (*sp*) contribution to the correlation energy to be

$$\Delta E_{1s+2p}(\text{total}) \approx -82.4 \text{ mh} \quad (9)$$

(again, for HF + 1 + 2 calculations). Thus, the (1s1p) set recovers 76% of the estimated (*sp*) contribution to the correlation energy, the (2s2p) set recovers 93%, and the (3s3p) set recovers 98%. The absolute error for the (1s1p) set is 20.1 mh; those for the (2s2p) and (3s3p) sets are 5.5 and 1.4 mh, respectively. The (4s4p) set, which is short of the extrapolated (*sp*) limit by just 0.4 mh, accounts for 99.5% of the (*sp*) correlation energy.

In Table VI we compare the optimized even-tempered 1s and 2p primitive sets with the *s* and *p* ANO sets of Almlöf and Taylor.²³ The differences found here follow the same pattern observed earlier for the polarization sets, namely, the errors are largest for the smaller expansions, being 2.4 and

TABLE VI. Comparison of the correlation energies obtained with the present primitive 1s and 2p sets with the ANO sets of Almlöf and Taylor (Ref. 23). For the *s*-set calculations, the *p* set was the primitive (7p) set contracted to [4p]; the base *s* sets were either (i) the 1s and 2s atomic orbitals from the (16s) set (Present) or (ii) the first two natural orbitals from the (13s) set (ANO). Similarly constructed sets were used for the *p*-set calculations. The polarization set for both series of calculations was the (2d 1f) set. Total energies (E_{HF} and $E_{\text{HF}+1+2}$) are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{HF} and ΔE_{corr}) are in millihartrees.

Correlation set	E_{HF}	ANO $E_{\text{HF}+1+2}$	E_{corr}	E_{HF}	ΔE_{HF}	Present $E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}
...	-74.808 890 ^a	-74.946 656 ^a	-137.766	-74.809 068 ^b	-0.178	-74.946 896 ^b	-137.828	-0.062
(1s)	-74.808 898	-74.975 388	-166.490	-74.809 068	-0.170	-74.973 127	-164.059	2.431
(2s)	-74.808 953	-74.977 820	-168.867	-74.809 068	-0.115	-74.977 438	-168.370	0.497
(3s)	-74.808 963	-74.978 154	-169.191	-74.809 068	-0.105	-74.977 947	-168.879	0.312
(4s)	-74.808 964	-74.978 188	-169.224	-74.809 068	-0.104	-74.978 288	-169.220	0.004
...	-74.808 991 ^c	-74.904 592 ^c	-95.601	-74.809 068 ^d	-0.077	-74.905 090 ^d	-96.022	-0.421
(1p)	-74.809 030	-74.966 969	-157.939	-74.809 068	-0.038	-74.962 965	-153.897	4.042
(2p)	-74.809 222	-74.976 854	-167.632	-74.809 068	0.154	-74.975 032	-165.964	1.668
(3p)	-74.809 279	-74.979 220	-169.941	-74.809 068	0.211	-74.978 614	-169.546	0.395
(4p)	-74.809 293	-74.979 706	-170.413	-74.809 068	0.225	-74.979 379	-170.311	0.102

^aEnergy for the (13s7p2d 1f)/[2s4p2d 1f] set.

^bEnergy for the (16s7p2d 1f)/[2s4p2d 1f] set.

^cEnergy for the (16s8p2d 1f)/[6s1p2d 1f] set.

^dEnergy for the (16s7p2d 1f)/[6s1p2d 1f] set.

4.0 mh, respectively, for the single $1s$ and $2p$ sets. These sets do, however, account for more than 90% of the correlation energy obtained with the corresponding ANO sets while the two term primitive sets account for more than 97%. The errors for the $(3s)$ and $(3p)$ sets, 0.31–0.40 mh, are comparable to, although slightly larger than, the errors for the $(3d)$, $(2f)$, and $(1g)$ sets. For the four term $1s$ and $2p$ sets the differences, of 0.0 and 0.1 mh, respectively, are clearly negligible. One point which should be noted is that the differences

in the HF energies for the two series of calculations reflect the different primitive sets used in the two calculations: Almlöf and Taylor used a $(13s8p)$ set while the present calculations used a $(16s7p)$ set.

While it is certainly possible to determine sets of correlating $1s$ and $2p$ functions to be added to the HF orbitals for all of the atoms of interest, this is not necessary. In Table VII we list the results obtained with the optimized even-tempered sets along with results obtained when the augmenting

TABLE VII. Exponents and correlation energies obtained with the atom optimized even-tempered (sp) sets and with sets optimized for the HF wave function. The base (sp) set was the $(16s7p)$ set contracted to $[2s4p]$ for the s -set calculations and to $[6s1p]$ for the p -set calculations; the polarization set is the $(2d1f)$ set. Total energies ($E_{\text{HF}+1+2}$) are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{corr}) are in millihartrees. For all sets the HF energy is $-74.809\,068$ hartrees.

Set	Augmenting set Source	s/p exponents	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
<i>s</i> -set calculations					
(1s)	(8s)	0.3093	– 74.973 053	– 163.985	0.074
(1s)	(9s)	0.3023	– 74.973 109	– 164.041	0.018
(1s)	optimum	0.295	– 74.973 127	– 164.059	
(2s)	(9s)	(0.3023, 1.013)	– 74.975 037	– 165.969	2.401
(2s)	(10s)	(0.2384, 0.6882)	– 74.972 622	– 163.554	4.816
	(10s)	(0.2384, 1.752)	– 74.977 402	– 168.334	0.036
(2s)	(12s)	(0.2067, 0.5547)	– 74.973 578	– 164.510	3.860
	(12s)	(0.2067, 1.428)	– 74.977 014	– 167.946	0.424
(2s)	optimum	(0.253, 1.887)	– 74.977 438	– 168.370	
(3s)	(10s)	(0.2384, 0.6882, 1.752)	– 74.977 929	– 168.861	0.018
(3s)	(12s)	(0.2067, 0.5547, 1.428)	– 74.977 796	– 168.728	0.151
(3s)	(14s)	(0.1829, 0.4600, 1.138)	– 74.977 206	– 168.138	0.741
(3s)	optimum	(0.236, 0.653, 1.808)	– 74.977 947	– 168.879	
(4s)	(12s)	(0.2067, 0.5547, 1.428, 4.682)	– 74.978 186	– 169.118	0.102
(4s)	(14s)	(0.1829, 0.4600, 1.138, 2.776)	– 74.978 290	– 169.222	– 0.002
(4s)	(16s)	(0.1606, 0.3772, 0.8655, 1.891)	– 74.978 225	– 169.157	0.063
(4s)	optimum	(0.184, 0.454, 1.118, 2.756)	– 74.978 288	– 169.220	
<i>p</i> -set calculations					
(1p)	(4p)	0.2754	– 74.962 810	– 153.742	0.154
(1p)	(5p)	0.2146	– 74.958 219	– 149.151	4.746
(1p)	optimum	0.289	– 74.962 965	– 153.897	
(2p)	(4p)	(0.2754, 1.046)	– 74.973 238	– 164.170	1.794
(2p)	(5p)	(0.2146, 0.7172)	– 74.974 997	– 165.929	0.034
(2p)	(6p)	(0.1750, 0.5302)	– 74.974 253	– 165.185	0.779
(2p)	optimum	(0.217, 0.700)	– 74.975 032	– 165.964	
(3p)	(5p)	(0.2146, 0.7172, 2.283)	– 74.977 914	– 168.846	0.700
(3p)	(6p)	(0.1750, 0.5302, 1.531)	– 74.978 658	– 169.590	– 0.044
(3p)	(7p)	(0.1481, 0.4135, 1.104)	– 74.978 225	– 169.157	0.389
(3p)	optimum	(0.176, 0.523, 1.554)	– 74.978 614	– 169.546	
(4p)	(6p)	(0.1750, 0.5302, 1.531, 4.459)	– 74.979 055	– 169.987	0.324
(4p)	(7p)	(0.1481, 0.4135, 1.104, 2.916)	– 74.979 322	– 170.254	0.057
(4p)	(8p)	(0.1285, 0.3368, 0.8471, 2.094)	– 74.979 389	– 170.321	– 0.010
(4p)	(9p)	(0.1139, 0.2817, 0.6697, 1.565)	– 74.979 218	– 170.150	0.161
(4p)	optimum	(0.138, 0.346, 0.866, 2.167)	– 74.979 379	– 170.311	

^a Correlation energy differences are relative to the corresponding optimized set.

functions are taken from basis sets optimized for HF wave functions. As can be seen, in each case it is possible to select a HF primitive set with a distribution of exponents very nearly the same as those in the optimum even-tempered set. Thus, the most diffuse function in the (9s) set has an exponent of 0.3023 which is nearly identical to that for the optimum (1s) function, 0.295. Further, the correlation energy computed with the (1s) function from the (9s) set is just 0.018 mh above that computed with the optimum (1s) function. A

similar situation holds for the remaining s sets—the exponents of the outermost and third outermost functions in the (10s) HF set are very similar to those in the optimum (2s) set, the exponents of the outermost three functions of the (10s) HF set are very nearly the same as those in the optimum (3s) set and the exponents of the outermost four functions in the (14s) HF set differ little from those in the optimum (4s) set. In no case does the error for the *best* HF set exceed 0.1 mh.

TABLE VIII. HF and correlation energies for the O₂ molecule obtained with the atom optimized even-tempered (*sp*) sets and with sets optimized for the atomic HF wave function. The nominal (*sp*) set was the (16s7p) set contracted to [2s4p] for the s-set calculations and to [6s1p] for the p-set calculations. The polarization set is the (2d1f) set. Total energies (E_{HF} and $E_{\text{HF}+1+2}$) are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{corr}) are in millihartrees.

Augmenting set Set	Source	E_{HF}	ΔE_{HF}^a	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
<i>s</i> -set calculations						
(1s)	(8s)	− 149.659 097	1.055	− 150.083 708	− 424.611	0.116
(1s)	(9s)	− 149.659 624	0.528	− 150.084 314	− 424.690	0.037
(1s)	optimum ^b	− 149.660 152		− 150.084 879	− 424.727	
(2s)	(9s)	− 149.659 645	3.422	− 150.087 653	− 428.008	3.652
(2s)	(10s)	− 149.663 517	− 0.449	− 150.087 923	− 424.405	7.255
	(10s)	− 149.663 715	− 0.647	− 150.095 298	− 431.583	0.077
(2s)	(12s)	− 149.664 560	− 1.492	− 150.090 050	− 425.490	6.170
	(12s)	− 149.664 602	− 1.534	− 150.095 582	− 430.980	0.680
(2s)	optimum ^b	− 149.663 068		− 150.094 728	− 431.660	
(3s)	(9s)	− 149.660 189	3.828	− 150.090 758	− 430.569	2.012
(3s)	(10s)	− 149.663 921	0.096	− 150.096 472	− 432.551	0.030
(3s)	(12s)	− 149.664 730	− 0.713	− 150.097 074	− 432.344	0.237
(3s)	(14s)	− 149.665 041	− 1.024	− 150.096 460	− 431.419	1.162
(3s)	optimum ^b	− 149.664 017		− 150.096 598	− 432.581	
(4s)	(12s)	− 149.664 737	0.337	− 150.097 788	− 433.051	0.275
(4s)	(14s)	− 149.665 083	− 0.009	− 150.098 412	− 433.329	− 0.003
(4s)	(16s)	− 149.665 257	− 0.183	− 150.098 518	− 433.261	0.065
(4s)	optimum ^b	− 149.665 074		− 150.098 400	− 433.326	
<i>p</i> -set calculations						
(1p)	(4p)	− 149.662 188	− 0.632	− 150.069 717	− 407.529	0.499
(1p)	(5p)	− 149.664 069	− 2.513	− 150.062 765	− 398.696	9.332
(1p)	optimum ^b	− 149.661 556		− 150.069 584	− 408.028	
(2p)	(4p)	− 149.662 865	1.555	− 150.088 846	− 425.981	1.798
(2p)	(5p)	− 149.664 442	− 0.022	− 150.092 131	− 427.689	0.090
(2p)	(6p)	− 149.664 790	− 0.370	− 150.090 555	− 425.765	2.014
(2p)	optimum ^b	− 149.664 420		− 150.092 199	− 427.779	
(3p)	(5p)	− 149.664 567	0.457	− 150.097 578	− 433.011	1.027
(3p)	(6p)	− 149.665 026	− 0.002	− 150.099 141	− 434.115	− 0.077
(3p)	(7p)	− 149.665 257	− 0.233	− 150.098 518	− 433.261	0.777
(3p)	optimum ^b	− 149.665 024		− 150.099 062	− 434.038	
(4p)	(7p)	− 149.665 411	− 0.012	− 150.100 774	− 435.363	0.059
(4p)	(8p)	− 149.665 404	− 0.005	− 150.100 800	− 435.396	0.026
(4p)	optimum ^b	− 149.665 399		− 150.100 821	− 435.422	

^aHartree-Fock and correlation energy differences are relative to the corresponding optimized set.

^bOptimum 1s/2p functions for the oxygen atom.

For the $2p$ sets it is the outermost function from the $(4p)$ set which best matches the optimum $(1p)$ set—the exponents differ by less than 5% and the calculated correlation energies differ by only 0.15 mh. Using the outermost function from the $(5p)$ set, on the other hand, increases the correlation energy by 4.6 mh (out of a total lowering of 57.9 mh; see Table V). The optimum $(2p)$ set is best described by the outermost two functions from the $(5p)$ set, while the optimum $(3p)$ and $(4p)$ sets are best described by the $(6p)$ and $(8p)$ HF sets, respectively. In fact, in the latter two cases the correlation energy obtained with the best HF sets are slightly larger than those obtained with the optimum *even-tempered* sets. The above results clearly show that (sp) sets optimized for the HF wave function can also provide an excellent description of correlation effects.

Table VII contains another interesting result. By and large, the changes observed in the optimum even-tempered exponents with increasing basis set size are as expected, i.e., as the set becomes larger, the smallest exponent decreases while the largest exponent increases with the additional functions covering the intermediate region ever more finely (decreasing β). This pattern was also observed for the exponents of the polarization functions. The first members of the s set, on the other hand, do not follow this pattern. Upon expanding the s set from one to two functions, the exponent of the tight function becomes substantially larger than expected, i.e., $\beta = 7.461$ for the $(2s)$ set whereas $\beta = 3.225$ and 3.588 for the $(2p)$ and $(2d)$ sets. Then, upon adding a third s function, we find that the exponent of the extra function lies inbetween the exponents of the $(2s)$ set, i.e., $(0.253, 1.887) \rightarrow (0.236, 0.653, 1.808)$, and that the exponent of the tightest function actually decreases slightly. The $(3s)$ to $(4s)$ expansion again follows the normal pattern. The reason for this behavior is not obvious, although it may be connected with the nodal structure of the $2s$ atomic orbital.

In Table VIII we summarize the results of HF and HF + 1 + 2 calculations on the oxygen molecule, O_2 , with the atom optimized even-tempered (sp) sets and the (sp) sets derived from atomic HF calculations. The calculated correlation energies for the molecular calculations follow the same pattern as for the atomic calculations, i.e., the outermost function from the $(9s)$ HF set yields nearly the same correlation energy as that from the optimum $(1s)$ set, the outermost function from the $(4p)$ set yields nearly the same correlation energy as that from the optimum $(1p)$ set and so on. The calculated HF energies, on the other hand, almost always favor the larger HF sets, i.e., the $(14s)$ set always yields a lower HF energy than the $(12s)$ set which in turn always yields a lower HF energy than the $(10s)$ set, etc., independent of which correlation set is being considered. Only the $(8p)$ set does not follow this trend and here the difference is negligible, 0.007 mh.

Although the results given above clearly show that it is possible to obtain (sp) sets which consistently describe both the HF and HF + 1 + 2 wave functions, as the results in Table VIII caution, in doing so we must balance the errors in the HF energy against the errors in the correlation energy. This is even more true when using the HF (sp) sets themselves to describe both the HF and correlated wave func-

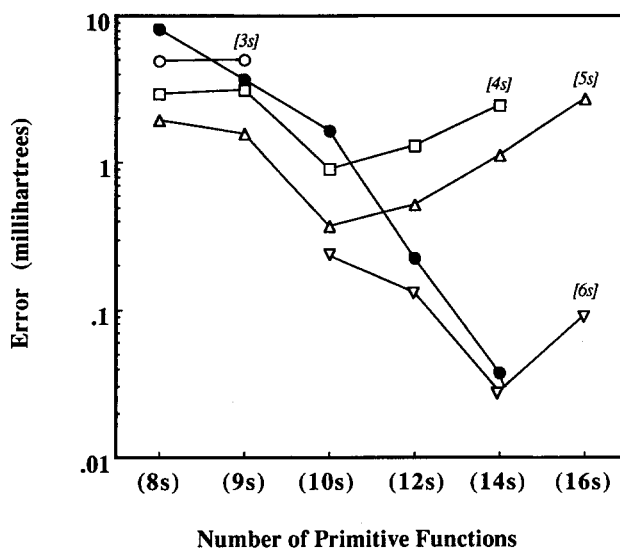


FIG. 4. Errors in the atomic HF and HF + 1 + 2 energies for an (ns) primitive set contracted to $[ms]$. The solid circles are the errors in the HF energies; the open polygons are the errors in the HF + 1 + 2 energies.

tions, i.e., not just as correlation functions to be added to the HF $1s$ and $2s$ orbitals. In Figs. 4 and 5 we plot the errors in the atomic HF and HF + 1 + 2 energies resulting from the use of a given HF s set (Fig. 4) and p set (Fig. 5). For the s -set calculations an (ns) HF set was contracted to $[ms]$ and augmented with the $(7p)$ set contracted to $[4p]$. For the p -set calculations an (np) HF set was contracted to $[mp]$ and augmented with the $(16s)$ set contracted to $[6s]$. Both sets of calculations included the $(2d\ 1f)$ polarization set. The errors in the HF energies are relative to that obtained with a $(16s9p)$ set; the errors in the correlation energy are relative to that obtained with the $(16s9p2d\ 1f)/(6s6p2d\ 1f)$ set. Choosing the optimum s or p set then requires a balance

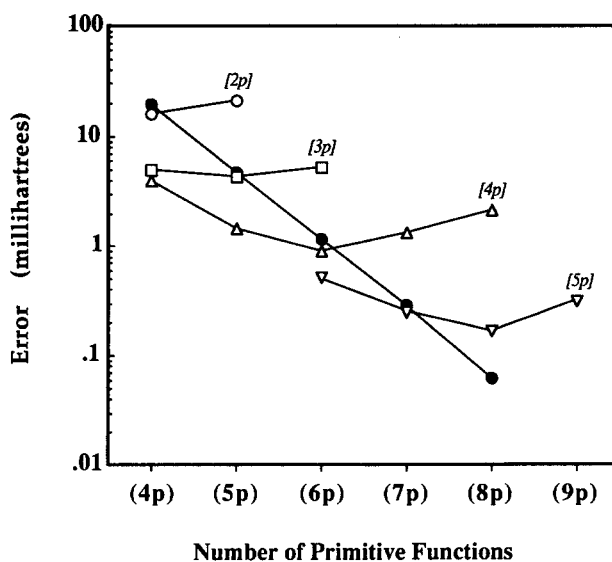


FIG. 5. Errors in the HF and HF + 1 + 2 energies for an (np) primitive set contracted to $[mp]$. The solid circles are the errors in the HF energies; the open polygons are the errors in the HF + 1 + 2 energies.

between (i) minimizing the error in the HF energy and (ii) minimizing the error in the correlation energy.

As can be seen in the figures, the error in the HF energy decreases monotonically with increasing primitive set size. This is expected since the contracted basis set always contains the HF atomic orbitals. The errors in the correlation energy for a given contracted set, on the other hand, exhibit a minimum for a specific primitive set. This is in line with the results presented in Table VII. Based on the data plotted in Fig. 4, the recommended *s* sets are: The (9*s*) set contracted to [3*s*], the (10*s*) set contracted to [4*s*], the (12*s*) set contracted to [5*s*], and the (14*s*) set contracted to [6*s*]. We recommend the use of the (12*s*) set contracted to [5*s*], rather than the (10*s*) set. Although the (10*s*) set yields a correlation energy which is 0.15 mh lower than that obtained with the (12*s*) set, the HF energy for the (12*s*) set is lower than that for the (10*s*) set by 1.4 mh. In this case we believe that use of the larger *s* set is a better compromise.

The recommended *p* sets are: the (4*p*) set contracted to [2*p*], the (5*p*) set contracted to [3*p*], the (6*p*) set contracted to [4*p*], and the (8*p*) set contracted to [5*p*]. The only difficult choice was for the appropriate primitive set for the [2*p*] set—the correlation energy obtained with the (4*p*) set is 5.6 mh lower than that obtained with the (5*p*) set, but the HF energy for the (5*p*) set is 14.8 mh lower than that for the (4*p*) set. The above choice minimizes the difference between the errors in the HF and correlation energies.

III. CORRELATED BASIS SETS FOR THE FIRST ROW ATOMS: BORON THROUGH NEON

The above calculations on the oxygen atom were used to guide the development of correlated basis sets for the first row atoms boron through neon. Again, we first considered polarization basis sets. We determined (1*d*)–(3*d*) sets, (1*f*)–(2*f*) sets, and a (1*g*) set for each atom, first optimizing the *d* exponents in the (1*d*)–(3*d*) sets, then optimizing the *f* exponents in the (2*d* 1*f*) and (3*d* 2*f*) sets and finally optimizing the *g* exponent in the (3*d* 2*f* 1*g*) set. In this case the exponents of the (3*d*) set were fully optimized, i.e., they were not constrained to be given by an even-tempered expansion.³² The optimum exponents, correlation energies, and incremental energy lowerings are summarized in Table IX where, for comparative purposes, the quoted results for the (1*f*) set are based on the (3*d*) set rather than the (2*d*) set; the correlation energy lowerings are plotted in Fig. 6. The incremental lowerings have been normalized to the total correlation energy of the *L* shell, $E_{\text{corr}}(L \text{ shell})$, determined from calculations with the (12*s*6*p*3*d* 2*f* 1*g*)/[5*s*4*p*3*d* 2*f* 1*g*] set (see below).

The same general trends are observed in the results in Table IX as were observed for the oxygen atom: addition of the first 3*d* function results in the largest energy lowering by far, followed by smaller but comparable energy lowerings upon addition of the second 3*d* and first 4*f* functions, followed by even smaller, but again comparable, energy lowerings for the third 3*d*, second 4*f*, and first 5*g* functions. There are differences, however. For example, the higher angular momentum functions are relatively more important for the

latter half of the row than for the first half of the row—this is undoubtedly due to the increased number of *p* electrons.

The correlation energy lowerings resulting from the addition of functions from selected HF (*sp*) sets follow the same trend observed in the oxygen atom calculations: The most diffuse functions in the (9*s*) and (4*p*) sets provide the best single correlating functions, the first and third most diffuse functions in the (10*s*) set and the two most diffuse functions in the (5*p*) set provide the best two correlating functions, etc. The incremental energy lowerings for the best (*nsnp*) sets are also plotted in Fig. 6. The first (1*s* + 2*p*) set is less important than the first 3*d* function for the atoms B–N whereas the 3*d* function is less important for F–Ne. The third (1*s* + 2*p*) set, on the other hand, is far more important at the beginning of the row than the corresponding polarization functions.

Krishnan *et al.*²¹ determined a (2*s*2*p*1*d*) primitive set of correlating functions for all of the first row atoms and for the hydrogen atom. For the atoms boron through neon the exponents of the functions were determined by minimizing the second order Moller–Plesset perturbation energy (UMP2). The 3*d* exponents reported by these authors are only slightly larger (0.06–0.10) than the exponents of the (1*d*) sets reported here. The exponents for each of the two pairs of 1*s* and 2*p* functions in the (2*s*2*p*) sets were constrained to be identical. Although the calculations reported here show that $\zeta_2(2s) > 2\zeta_2(2p)$, the net effect of the constraint is probably not large since (i) the exponents of the first 1*s*- and 2*p*-correlating functions are similar [compare the outermost functions of the (10*s*) and (5*p*) sets] and (ii) the second 2*p* correlating function is far more important than the second 1*s* correlating function, e.g., for oxygen $\Delta E_{2p}^2 \approx 3\Delta E_{1s}^2$.

We define *correlation consistent basis sets* to contain all of the correlating functions which lower the correlation energy by similar amounts as well as all correlating functions which lower the energy by larger amounts. Thus, the (9*s*5*p*1*d*) primitive set contracted to [3*s*2*p*1*d*] is the simplest correlation consistent basis set, since the (1*s*1*p*) correlation set lowers the correlation energy by approximately the same amount as the (1*d*) set (see Fig. 6); all other functions lower the correlation energy by substantially smaller amounts. We refer to this set as the correlation consistent polarized valence double-zeta (cc-pVDZ) set. Other correlation consistent sets are the polarized valence triple-zeta (cc-pVTZ) set, the (10*s*5*p*2*d* 1*f*) set contracted to [4*s*3*p*2*d* 1*f*], and the polarized valence quadruple-zeta (cc-pVQZ) set, the (12*s*6*p*3*d* 2*f* 1*g*) set contracted to [5*s*4*p*3*d* 2*f* 1*g*]. The HF and correlation energies obtained with the correlation consistent sets are listed in Table X along with the energies obtained with the ANO sets of Almlöf and Taylor.²³ Also included in this table are accurate numerical HF energies³³ as well as estimates of the total (HF + 1 + 2) correlation energies. To obtain the estimated correlation energies we fit the incremental energy lowerings from Table X to a geometric series (see Fig. 7) and then added the higher order contributions predicted by the series to the correlation energy obtained with the cc-pVQZ set. Although it is difficult to estimate the accuracy of the extrapolation, we expect the errors in the estimated total

TABLE IX. Exponents, correlation energies, and correlation energy lowerings obtained with the atom optimized polarization sets for the boron through neon atoms. The (*sp*) set was the (16*s7p*) set contracted to [6*s4p*]. Total energies (E_{HF} and $E_{\text{HF}+1+2}$) are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{corr}) are in millihartrees.

Polarization set	Exponents	$E_{\text{HF}+1+2}$	E_{corr}	ΔE_{corr}^a
The boron atom ($E_{\text{HF}} = -24.529\,035$)				
...	...	-24.574 918 ^b	-45.883	
(1 <i>d</i>)	0.343	-24.595 567	-66.532	-20.649
(2 <i>d</i>)	(0.199, 0.661)	-24.598 299	-69.264	-2.732
(3 <i>d</i>)	(0.145, 0.402, 1.110)	-24.598 640	-69.605	-0.341
(3 <i>d</i> 1 <i>f</i>)	0.490	-24.600 353	-71.318	-1.713
(3 <i>d</i> 2 <i>f</i>)	(0.311, 0.882)	-24.600 653	-71.618	-0.300
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	0.673	-24.600 988	-71.953	-0.335
The carbon atom ($E_{\text{HF}} = -37.688\,544$)				
...	...	-37.733 875 ^b	-45.331	
(1 <i>d</i>)	0.550	-37.772 437	-83.893	-38.562
(2 <i>d</i>)	(0.318, 1.097)	-37.778 497	-89.953	-6.060
(3 <i>d</i>)	(0.228, 0.649, 1.848)	-37.779 308	-90.764	-0.811
(3 <i>d</i> 1 <i>f</i>)	0.761	-37.783 827	-95.283	-4.519
(3 <i>d</i> 2 <i>f</i>)	(0.485, 1.419)	-37.784 623	-96.079	-0.796
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	1.011	-37.785 483	-96.939	-0.860
The nitrogen atom ($E_{\text{HF}} = -54.400\,771$)				
...	...	-54.447 265 ^b	-46.494	
(1 <i>d</i>)	0.817	-54.500 019	-99.248	-52.754
(2 <i>d</i>)	(0.469, 1.654)	-54.509 599	-108.828	-9.580
(3 <i>d</i>)	(0.335, 0.968, 2.837)	-54.510 976	-110.205	-1.377
(3 <i>d</i> 1 <i>f</i>)	1.093	-54.518 987	-118.216	-8.011
(3 <i>d</i> 2 <i>f</i>)	(0.685, 2.027)	-54.520 460	-119.689	-1.473
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	1.427	-54.522 023	-121.252	-1.563
The oxygen atom ($E_{\text{HF}} = -74.809\,068$)				
...	...	-74.885 656 ^b	-76.588	
(1 <i>d</i>)	1.185	-74.947 895	-138.827	-62.239
(2 <i>d</i>)	(0.645, 2.314)	-74.962 612	-153.544	-14.717
(3 <i>d</i>)	(0.444, 1.300, 3.775)	-74.965 253	-156.185	-2.641
(3 <i>d</i> 1 <i>f</i>)	1.428	-74.981 101	-172.033	-15.848
(3 <i>d</i> 2 <i>f</i>)	(0.859, 2.666)	-74.984 547	-175.479	-3.446
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	1.846	-74.988 205	-179.137	-3.658
The fluorine atom ($E_{\text{HF}} = -99.408\,778$)				
...	...	-99.515 053 ^b	-106.275	
(1 <i>d</i>)	1.640	-99.585 707	-176.929	-70.654
(2 <i>d</i>)	(0.855, 3.107)	-99.605 698	-196.920	-19.991
(3 <i>d</i>)	(0.586, 1.725, 5.014)	-99.609 658	-200.880	-3.960
(3 <i>d</i> 1 <i>f</i>)	1.917	-99.631 072	-222.294	-21.414
(3 <i>d</i> 2 <i>f</i>)	(1.148, 3.562)	-99.636 226	-227.448	-5.154
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	2.376	-99.642 303	-233.525	-6.077
The neon atom ($E_{\text{HF}} = -128.546\,201$)				
...	...	-128.685 520 ^b	-139.319	
(1 <i>d</i>)	2.202	-128.762 876	-216.675	-77.356
(2 <i>d</i>)	(1.096, 4.014)	-128.788 201	-242.000	-25.325
(3 <i>d</i>)	(0.747, 2.213, 6.471)	-128.793 564	-247.363	-5.363
(3 <i>d</i> 1 <i>f</i>)	2.544	-128.818 377	-272.176	-24.813
(3 <i>d</i> 2 <i>f</i>)	(1.524, 4.657)	-128.824 921	-278.720	-6.544
(3 <i>d</i> 2 <i>f</i> 1 <i>g</i>)	2.983	-128.833 665	-287.464	-8.744

^a Change in the correlation energy referenced to the preceding set with one fewer function of the given angular momentum.

^b Energy for the (16*s7p*)/[6*s4p*] set.

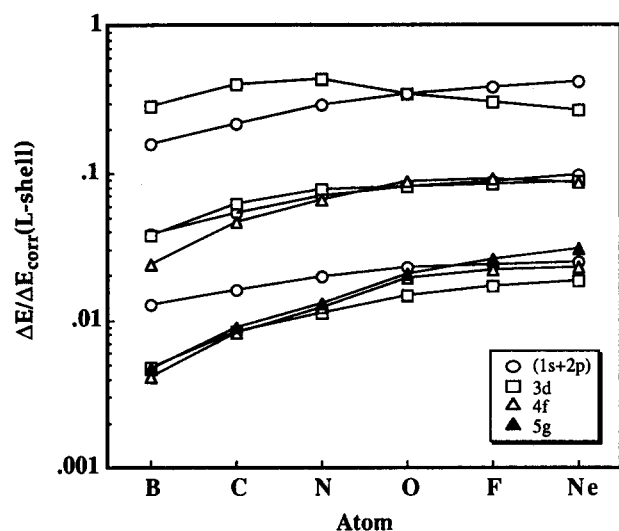


FIG. 6. The normalized incremental correlation energy lowerings for the first row atoms boron through neon. The $(1s + 2p)$ sets were taken from the appropriate HF basis sets. The total correlation energy for the L shell was taken from calculations with the $(12s6p3d\ 2f\ 1g)/[5s4p3d\ 2f\ 1g]$ set.

(HF + 1 + 2) correlation energies to range from ± 0.1 mh for boron to ± 4 mh for neon.

The smallest correlation consistent set, the cc-pVDZ set, yields between 61% (Ne) and 87% (B) of the estimated total correlation energy; the error ranges from 9.3 (B) to 120.4 mh (Ne). The cc-pVTZ set yields 85% (Ne) to 97% (B) of the correlation energy while the cc-pVQZ set yields 94.2% (Ne) to 99.6% (B) of the estimated correlation energy. In comparison, the $[5s4p3d\ 2f\ 1g]$ ANO set²² yields 95.4% of the estimated correlation energy of neon. Thus, the differences in the correlation energies obtained with the cc-pVQZ sets and the ANO sets of Almlöf and Taylor²³ are, in general, a small fraction of the remaining errors in the correlation energies. The geometric series predict that use of a $(6s5p4d\ 3f\ 2g\ 1h)$ set will reduce the errors in the calculated correlation energies to: <0.1 mh (B), 0.2–0.3 mh (C), 0.7 mh (N), 2.4 mh (O), ~ 5 mh (F), and ~ 7 mh (Ne); the estimate for oxygen is in line with the estimates given earlier.

In general, the errors in the HF energies are only a fraction of the errors in the correlation energies. It should also be

TABLE X. The HF and correlation energies obtained with the *correlation consistent basis sets* and the ANO sets of Almlöf and Taylor (Ref. 23). The energy differences (ΔE_{HF} and ΔE_{corr}) are relative to the accurate HF calculations of Froese Fischer (Ref. 33) and to the estimated total (HF + 1 + 2) correlation energies (see the text). Total energies (E_{HF} and $E_{\text{HF} + 1 + 2}$) are in hartrees; correlation energies (E_{corr}) and energy differences (ΔE_{HF} and ΔE_{corr}) are in millihartrees.

	Boron	Carbon	Nitrogen	Oxygen	Fluorine	Neon
cc-pVDZ: $(9s4p1d)/[3s2p1d]$						
E_{HF}	−24.526 564	−37.682 391	−54.388 414	−74.786 188	−99.371 080	−128.488 776
ΔE_{HF}	2.497	6.228	12.520	23.210	38.269	58.32
$E_{\text{HF} + 1 + 2}$	−24.589 568	−37.759 783	−54.477 112	−74.907 257	−99.523 619	−128.673 617
E_{corr}	−63.004	−77.392	−88.698	−121.069	−152.539	−184.842
ΔE_{corr}	9.3	20.7	35.2	65.2	93.3	120.4
cc-pVTZ: $(10s5p2d\ 1f)/[4s3p2d\ 1f]$						
E_{HF}	−24.528 098	−37.686 662	−54.397 358	−74.803 078	−99.399 194	−128.531 862
ΔE_{HF}	0.963	1.957	3.576	6.320	10.155	15.24
$E_{\text{HF} + 1 + 2}$	−24.598 193	−37.779 611	−54.511 779	−74.968 451	−99.612 199	−128.791 916
E_{corr}	−70.095	−92.949	−114.421	−165.373	−213.005	−260.054
ΔE_{corr}	2.2	5.2	9.5	20.9	32.8	45.1
cc-pVQZ: $(12s6p3d\ 2f\ 1g)/[5s4p3d\ 2f\ 1g]$						
E_{HF}	−24.528 898	−37.688 234	−54.400 176	−74.807 975	−99.406 980	−128.543 470
ΔE_{HF}	0.163	0.385	0.758	1.423	2.369	3.630
$E_{\text{HF} + 1 + 2}$	−24.600 877	−37.785 188	−54.521 402	−74.987 069	−99.640 537	−128.831 057
E_{corr}	−71.979	−96.954	−121.226	−179.094	−233.557	−287.587
ΔE_{corr}	0.3	1.1	2.7	7.2	12.2	17.6
ANO: $(13s8p6d\ 4f\ 2g)/[5s4p3d\ 2f\ 1g]$						
E_{HF}		−37.688 518	−54.400 774	−74.809 174		−128.546 511
ΔE_{HF}		0.101	0.160	0.224		0.589
$E_{\text{HF} + 1 + 2}$		−37.786 002	−54.522 889	−74.989 805		−128.837 748
E_{corr}		−97.485	−122.115	−180.631		−291.237
		0.6	1.8	5.7		14.0
Accurate HF^a						
E_{HF}	−24.529 061	−37.688 619	−54.400 934	−74.809 398	−99.409 349	−128.547 10
Estimated E_{corr}	−72.3 \pm 0.3	−98.1 \pm 0.4	−123.9 \pm 0.5	−186.3 \pm 1.0	−245.8 \pm 2.0	−305.2 \pm 3.0

^a Reference 32.

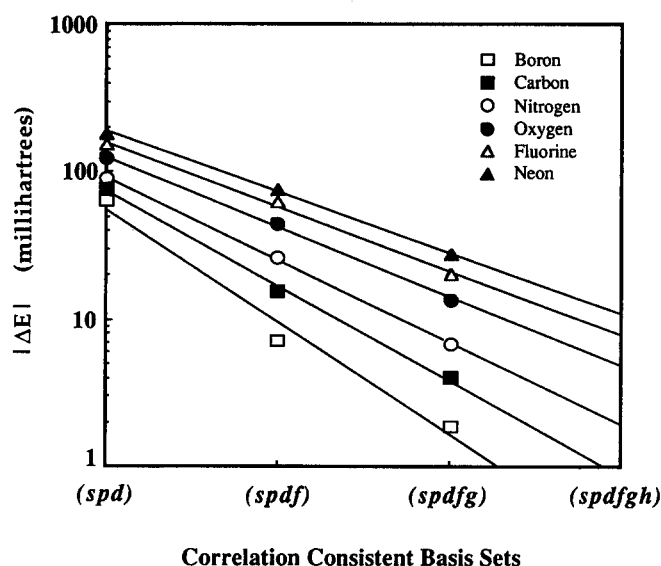


FIG. 7. Absolute values of the incremental energy lowerings for the first row atoms boron through neon obtained with the correlation consistent basis sets. The straight lines are geometric series fits of the calculated points. The series so obtained are: $-54.60/5.783^{(i-1)}$ (B), $-74.26/4.396^{(i-1)}$ (C), $-90.06/3.610^{(i-1)}$ (N), $-124.49/2.970^{(i-1)}$ (O), $-156.51/2.724^{(i-1)}$ (F), and $-188.13/2.591^{(i-1)}$ (Ne).

noted that an important source of error in the HF calculations is the description of the $1s$ orbital. This is in contrast to the correlation energy calculations which explicitly consider only the valence electrons. The error in the HF description of the valence orbitals will thus be substantially less than the total errors quoted in Table X.

IV. CORRELATED BASIS SETS FOR THE HYDROGEN ATOM

To determine correlated basis sets for the hydrogen atom poses a special problem; atomic calculations are clearly not useful. Following the suggestion of Almlöf and Taylor,²³

we determined correlation sets for the hydrogen atom from calculations on the hydrogen molecule. As before, we first added $(1p)-(3p)$ polarization sets to the $(7s)$ set of Duijneveldt²⁷ contracted to $[4s]$ ³⁴ and individually optimized the exponents in $HF + 1 + 2$ calculations. To the $[4s3p]$ set so obtained we then added $(1d)$ and $(2d)$ sets, again optimizing the exponents. Finally, we added a $(1f)$ set to the $[4s3p2d]$ set and optimized its exponent. The exponents, energies, and incremental HF and correlation energy lowerings so obtained are collected in Table XI. The results listed in this table show the by-now standard convergence pattern: (i) the first $2p$ function lowers the correlation energy substantially, by -10.1 mh in this case; (ii) the second $2p$ function and the first $3d$ function lower the energy by comparable amounts, -1.3 to -2.0 mh; (iii) finally, the third $2p$ function, second $3d$ function, and first $4f$ function lower the energy by -0.25 to -0.29 mh.

The energies reported in Table XI differ little from those reported by Almlöf and Taylor.²³ With an $(8s6p4d)$ set contracted to $[4s3p2d]$ these authors reported a total energy of $-1.173\,756$ hartrees; the comparable set here yields $-1.173\,617$ hartrees. The energy obtained here with the largest set considered, $-1.173\,868$ hartrees for the $[4s3p2d\,1f]$ set, also agrees well with more accurate energies reported by others, e.g., Liu³⁵ reported a CI energy for H_2 of $-1.174\,142$ hartrees obtained with a $[5s3p3d\,2f]$ set of Slater functions, while Kolos and Wolniewicz³⁶ calculated an "exact" energy of $-1.174\,475$ hartrees.

To ascertain whether the hydrogen polarization sets obtained above are appropriate for calculations on other molecules, we reoptimized the exponents of the $(1p)$ and $(2p1d)$ sets for both CH and OH. The results of these calculations, along with those for the H_2 -optimized sets, are summarized in Table XII. Although the changes observed in the optimum hydrogen exponents are significantly larger than those found for oxygen (compare with Table III), the resulting energetic changes are small, <0.05 mh. It remains to be seen whether the noted differences in the hydrogen exponents are

TABLE XI. Polarization function exponents for the hydrogen atom obtained from calculations on molecular hydrogen along with the corresponding energies and incremental energy lowerings. The (s) set used in the calculations was the $(7s)$ set of Duijneveldt (Ref. 27) contracted to $[4s]$. Total energies (E_{HF} and E_{HF+1+2}) are in hartrees; correlation energies and energy differences (E_{corr} and ΔE_{corr}) are in millihartrees.

Polarization set	Exponents	E_{HF}	ΔE_{HF}^a	E_{HF+1+2}	E_{corr}	ΔE_{corr}^a
...	...	$-1.128\,424^b$...	$-1.154\,535^b$	-26.111	...
$(1p)$	0.727	$-1.132\,846$	-4.422	$-1.169\,060$	-36.214	-10.103
$(2p)$	(0.388, 1.407)	$-1.133\,423$	-0.577	$-1.171\,612$	-38.189	-1.975
$(3p)$	(0.292, 0.838, 2.292)	$-1.133\,488$	-0.065	$-1.171\,934$	-38.446	-0.257
$(1d)^c$	1.057	$-1.133\,527$	-0.039	$-1.173\,310$	-39.783	-1.337
$(2d)^c$	(0.662, 2.062)	$-1.133\,543$	-0.016	$-1.173\,617$	-40.074	-0.291
$(1f)^d$	1.397	$-1.133\,546$	-0.003	$-1.173\,868$	-40.322	-0.248

^a Changes in the HF and correlation energy referenced to the preceding set with one fewer function of the given angular momentum.

^b Energies for the $(7s)/[4s]$ set.

^c The (s) basis set was augmented with the above $(3p)$ primitive set.

^d The (s) basis set was augmented with the above $(3p)$ and $(2d)$ primitive sets.

TABLE XII. Comparison of H_2 -optimized hydrogen polarization sets with molecule optimized hydrogen polarization sets for CH and OH. The basis sets used for carbon and oxygen atoms were the (12s6p) sets contracted to [4s3p] augmented with the (2d 1f) polarization sets. For the hydrogen atom we used the (5s) set contracted to [3s]. Total energies (E_{HF} and E_{HF+1+2}) are in hartrees; correlation energies and energy differences (E_{corr} and ΔE_{corr}) are in millihartrees.

Polarization set		$\zeta(p;d)$	E_{HF}	E_{HF+1+2}	E_{corr}	ΔE_{corr}^a
The CH radical						
(1p)	H ₂	0.727	-38.278 172	-38.401 264	-123.092	
	CH	0.935	-38.278 336	-38.401 452	-123.116	-0.024
(2p1d)	H ₂	(0.388, 1.407; 1.057)	-38.278 581	-38.404 384	-125.803	
	CH	(0.355, 1.257; 0.916)	-38.278 577	-38.404 422	-125.845	-0.042
The OH radical						
(1p)	H ₂	0.727	-75.416 954	-75.623 349	-206.395	0.0
	OH	1.006	-75.417 179	-75.623 571	-206.392	0.003
(2p1d)	H ₂	(0.388, 1.407; 1.057)	-75.418 671	-75.628 106	-209.435	0.0
	OH	(0.344, 1.649; 1.262)	-75.418 714	-75.628 145	-209.431	0.004

^a Change in the correlation energy referenced to the calculation involving exponents optimized for H₂.

important for properties other than the energy.

To determine the optimum (*s*) sets for use in correlated calculations involving hydrogen, we contracted the (*ns*) HF sets of Duijneveldt²⁷ to [*ms*], augmented the resulting sets

with the (2p1d) polarization set determined above, and carried out HF + 1 + 2 calculations on both molecular hydrogen and OH. The results of these calculations are summarized in Table XIII. From the data in this table the optimum

TABLE XIII. Calculations on molecular hydrogen and OH using (*ns*) HF sets for the hydrogen atom contracted to [*ms*] and augmented with the (2p1d) polarization set. Total energies (E_{HF} and E_{HF+1+2}) are in hartrees; correlation energies and energy differences (E_{corr} and ΔE_{corr}) are in millihartrees.

(s) set						
Contracted	Primitive	E_{HF}	ΔE_{HF}	E_{HF+1+2}	E_{corr}	ΔE_{corr}^a
Molecular hydrogen						
[2s]	(4s)	-1.129 427	...	-1.168 017	-38.590	...
	(5s)	-1.128 905	0.522	-1.166 731	-37.826	0.764
[3s]	(4s)	-1.131 772	...	-1.171 066	-39.294	...
	(5s)	-1.132 953	-1.181	-1.172 336	-39.383	-0.089
	(6s)	-1.133 024	-0.071	-1.172 498	-39.474	-0.091
[4s]	(5s)	-1.133 069	...	-1.172 666	-39.597	...
	(6s)	-1.133 374	-0.305	-1.172 985	-39.611	-0.014
	(7s)	-1.133 472	-0.098	-1.173 041	-39.569	0.042
The OH radical						
[2s]	(4s)	-75.417 475	...	-75.626 414	-208.939	...
	(5s)	-75.417 847	-0.372	-75.626 685	-208.838	0.101
[3s]	(4s)	-75.417 941	...	-75.627 390	-209.449	...
	(5s)	-75.418 671	-0.730	-75.628 106	-209.435	0.014
	(6s)	-75.418 833	-0.162	-75.628 196	-209.363	0.072
[4s]	(5s)	-75.418 680	...	-75.628 413	-209.733	...
	(6s)	-75.418 867	-0.187	-75.628 624	-209.757	-0.024
	(7s)	-75.418 927	-0.060	-75.628 634	-209.707	0.050

^a Change in the correlation energy referenced to the preceding set with one fewer primitive *s* function.

[2s] and [4s] sets are obvious: the (4s) set should be contracted to [2s] and the (6s) set should be contracted to [4s]. Although the (6s) set contracted to [3s] yields a lower correlation energy for H₂, this is not the case for OH. In fact, for OH the (4s) set yields the lowest correlation energy. We consider the (5s) set contracted to [3s] to be a reasonable compromise, especially in light of the relatively large difference between the HF energies obtained with the (4s) and (5s) sets, 0.7–1.2 mh.

V. CONCLUSIONS

In the past basis sets for use in correlated molecular calculations have largely been taken from Hartree–Fock calculations. The recent work of Almlöf, Taylor, Bauschlicher, Langhoff, Walch, and co-workers^{23,37} indicates that basis sets taken from correlated atomic calculations provide an excellent description of molecular correlation effects. It is therefore important to understand the basis set requirements for describing correlation effects in atoms.

From a thorough study of correlation effects in the oxygen atom we have established that:

(i) Primitive Gaussian functions effectively and efficiently describe correlation effects if the exponents of the functions are optimized in atomic correlated calculations. The (3s), (3p), (3d), (2f), and (1g) primitive correlation sets reported here yield correlation energies within a few tenths of a millihartree of those obtained with the ANO sets of Almlöf and Taylor.²³ Even the small primitive sets yield a substantial fraction of the correlation energy obtained with the corresponding ANO sets: the (1s), (2s), (1p), and (2p) sets yield 91.3%, 98.2%, 92.8%, and 97.1%, respectively, of the correlation energy obtained with the comparable ANO sets, while the (1d), (2d), and (1f) sets yield 87.9%, 98.0%, and 92.5%.

(ii) The (*sp*) functions for describing correlation effects can be taken from atomic Hartree–Fock calculations if the appropriate primitive set is used to obtain a given contracted set. The valence functions in (9s) and (4p) sets provide the best single correlating function, the valence functions in the (10s) and (5p) sets the best two correlating functions, etc. In general, the (*sp*) correlation sets are obtained by augmenting the atomic orbitals with the most diffuse primitive functions in the set. The only exception to this rule is for the two function (*s*) correlating set; in this case the first and third most diffuse functions are best.

(iii) For the oxygen atom the energy lowerings resulting from the successive addition of functions of a given angular momentum decrease approximately geometrically with ratios ranging from 3.5⁻¹ (for 5g functions) to 5.1⁻¹ (for 3d functions). The energy lowering resulting from the successive addition of functions with higher angular momenta also decrease approximately geometrically with ratios ranging from 3.2⁻¹ (for the third function of a given symmetry) to 4.1⁻¹ (for the first function of a given symmetry). Using these geometric series to extrapolate to the infinite basis set limit, we estimate that a (4s4p4d3f2g1h) correlation set will be required to converge the (HF + 1 + 2) correlation energy to ~1 kcal/mol.

Use of the above primitive basis sets in molecular calcu-

lations on OH and O₂ indicate that they also provide an excellent description of molecular correlation effects.

The calculations on the oxygen atom served as a guide for calculations on the atoms in the first row from boron through neon. (*spdfg*) sets for use in correlated molecular calculations were determined for all these atoms. Again, it was found that the incremental energy lowerings due to the successive addition of correlating functions fall into distinct groups. Thus, the (1s1p) set and the (1d) function both decrease the correlation energy by comparable amounts, the incremental lowerings for the (2s2p), (2d), and (1f) sets are similar, etc. This leads to the concept of *correlation consistent basis sets*. For the first row atoms these sets are:

(sp) set			
Primitive	Contracted	Polarization set	
cc-pVDZ	(9s4p)	[3s2p]	(1d)
cc-pVTZ	(10s5p)	[4s3p]	(2d 1f)
cc-pVQZ	(12s6p)	[5s4p]	(3d 2f 1g)

Correlation consistent sets are proposed for all of the first row atoms from boron through neon. The largest set, the cc-pVQZ set, yields $99.2 \pm 0.4\%$ of the correlation energy obtained with the ANO set of Almlöf and Taylor²³ with a comparable number of basis functions even though it contains only $\frac{2}{3}$ of the number of primitive functions and $\frac{1}{2}$ the number of primitive polarization functions. Estimates of the higher order terms in the basis set expansions suggest that the correlation energies obtained with the cc-pVQZ sets are in error by as little as 0.3 ± 0.3 mh (boron) to as much as 18 ± 3 mh (neon). Comparable correlated basis sets were determined for the hydrogen atom.

ACKNOWLEDGMENTS

I wish to thank Dr. R. L. Shepard, Dr. R. J. Harrison, Dr. P. R. Taylor, Dr. J. Almlöf, and Dr. I. Shavitt for their comments on the manuscript. This work was performed under the auspices of the Office of Basic Energy Sciences, Division of Chemical Sciences, U.S. Department of Energy, under Contract No. W-31-109-Eng-38.

¹T. H. Dunning, Jr. and L. B. Harding, in *Theory of Chemical Reaction Dynamics*, edited by M. Baer (Chemical Rubber, Boca Raton, FL, 1985), Vol. I, Chap. 1.

²B. Liu, *J. Chem. Phys.* **80**, 581 (1984). See also, B. Liu, *J. Chem. Phys.* **58**, 1925 (1973) and P. Siegbahn and B. Liu, *ibid.* **68**, 1794 (1978).

³P. E. Cade and W. M. Huo, *J. Chem. Phys.* **47**, 614 (1967); A. D. McLean and M. Yoshimine, *ibid.* **47**, 3256 (1967).

⁴J. W. C. Johns and R. F. Barrow, *Proc. R. Soc. London Ser. A* **251**, 504 (1959); G. DiLorenzo and A. E. Douglas, *Can. J. Phys.* **51**, 434 (1973).

⁵C. C. J. Roothaan and P. S. Kelley, *Phys. Rev.* **131**, 1177 (1963).

⁶C. E. Moore, *Atomic Energy Levels*, Natl. Bur. Stand. No. 467 (U.S. GPO, Washington, D.C., 1949), Vol. 1.

⁷F. Sasaki and M. Yoshimine, *Phys. Rev. A* **9**, 26 (1974).

⁸F. A. Elder, D. Villarejo, and M. G. Inghram, *J. Chem. Phys.* **43**, 758 (1965).

⁹Pople and co-workers have proposed a scheme for computing quantities such as reaction energies which attempts to balance the errors in the correlation energy; see, for example, J. A. Pople, M. J. Frisch, B. T. Luke, and J. S. Binkley, *Int. J. Quantum Chem. Quantum Chem. Symp.* **17**, 307 (1983).

¹⁰H. J. Silverstone and O. Sinanoglu, *J. Chem. Phys.* **44**, 1899 (1966).

- ¹¹See, for example, W. A. Goddard III, *Phys. Rev.* **157**, 81 (1967); B. J. Moss, F. W. Bobrowicz, and W. A. Goddard III, *J. Chem. Phys.* **63**, 4632 (1975); T. H. Dunning, Jr., D. C. Cartwright, W. J. Hunt, P. J. Hay, and F. W. Bobrowicz, *ibid.* **64**, 4755 (1976); K. Ruedenberg and K. R. Sundberg, in *Quantum Science*, edited by J.-L. Calais, O. Goscinski, J. Lindenberg, and Y. Ohm (Plenum, New York, 1976); L. M. Cheung, K. R. Sundberg, and K. Ruedenberg, *Int. J. Quantum Chem.* **16**, 1103 (1979); K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, *Chem. Phys.* **71**, 41 (1982); B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, *ibid.* **48**, 157 (1980); B. O. Roos, *Int. J. Quantum Chem. S* **14**, 175 (1980). A recent review of multiconfiguration techniques is given in R. Shepard, in *Ab Initio Methods in Quantum Chemistry-II*, *Adv. Chem. Phys.*, edited by K. P. Lawley (Wiley, New York, 1987).
- ¹²Calculated from the molecular energy given in P. E. Cade, K. D. Sales, and A. C. Wahl, *J. Chem. Phys.* **44**, 1973 (1966), and the atomic energy given in Ref. 5.
- ¹³B. Rosen, *Spectroscopic Data Relative to Diatomic Molecules* (Pergamon, New York, 1970), p. 268.
- ¹⁴G. C. Lie and E. Clementi, *J. Chem. Phys.* **60**, 1288 (1974); see also, T. H. Dunning, Jr., D. C. Cartwright, W. J. Hunt, P. J. Hay, and F. W. Bobrowicz, *J. Chem. Phys.* **64**, 4755 (1976).
- ¹⁵K. Ruedenberg, M. W. Schmidt, M. M. Gilbert, and S. T. Elbert, *Chem. Phys.* **71**, 65 (1982).
- ¹⁶(a) T. H. Dunning, Jr. and P. J. Hay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 2, Chap. 1; (b) E. R. Davidson and D. Feller, *Chem. Rev.* **86**, 681 (1986).
- ¹⁷(a) H. F. Schaefer III, R. A. Klemm, and F. E. Harris, *Phys. Rev.* **181**, 137 (1969); (b) H. F. Schaefer III, R. A. Klemm, and F. E. Harris, *J. Chem. Phys.* **51**, 4643 (1969).
- ¹⁸P. J. Hay and T. H. Dunning, Jr., *J. Chem. Phys.* **64**, 5077 (1976), and references therein.
- ¹⁹T. H. Dunning, Jr., *J. Chem. Phys.* **65**, 3854 (1976).
- ²⁰D. Gervy and G. Verhaegen, *Int. J. Quantum Chem.* **12**, 115 (1977); L. G. M. Pettersson and P. E. M. Siegbahn, *J. Chem. Phys.* **83**, 3538 (1985).
- ²¹R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, *J. Chem. Phys.* **72**, 650 (1980).
- ²²K. Jankowski, R. Becherer, P. Scharf, H. Schiffer, and R. Ahlrichs, *J. Chem. Phys.* **82**, 1413 (1985); R. Ahlrichs, P. Scharf, and K. Jankowski, *Chem. Phys. Lett.* **98**, 381 (1985); R. Becherer and R. Ahlrichs, *Chem. Phys.* **99**, 389 (1985).
- ²³J. Almlöf and P. R. Taylor, *J. Chem. Phys.* **86**, 4070 (1987). See also, J. Almlöf, T. Helgaker, and P. R. Taylor, *ibid.* **92**, 3029 (1988).
- ²⁴Although the (*d, f, g, ...*) sets to be determined here describe dynamical correlation effects in addition to polarization effects, we shall refer to the collection of functions of higher angular momentum than the functions occupied in the HF atomic wave function as polarization functions.
- ²⁵F. W. Bobrowicz and W. A. Goddard III, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 2, Chap. 4.
- ²⁶T. H. Dunning, Jr. (unpublished). The initial set of exponents for the (12*s*), (6*p*), and (7*p*) sets were taken from Duijneveldt (Ref. 26); the reoptimization had little effect on either the exponents or the energy.
- ²⁷F. B. van Duijneveldt, IBM Res. Rep. RJ 945 (1971).
- ²⁸R. C. Raffanetti, *J. Chem. Phys.* **58**, 4452 (1973).
- ²⁹In all calculations reported here the Cartesian Gaussian functions were transformed to pure (real) angular momentum functions. In addition, the *s* component was removed from the (*l + m + n*) = 2 Cartesian set, the *p* components from the (*l + m + n*) = 3 set and the *s* and *d* components from the (*l + m + n*) = 4 set.
- ³⁰R. M. Pitzer (unpublished).
- ³¹1 mh (millihartree) = 0.001 h (hartree) = 0.027 21 eV = 0.6275 kcal/mol.
- ³²This has only a minor effect; compare, e.g., the results for the (3*d* 2*f* 1*g*) sets quoted in Tables I and IX.
- ³³C. Froese Fischer, *The Hartree-Fock Method for Atoms. A Numerical Approach* (Wiley, New York, 1977), Chap. 2.
- ³⁴The (7*s*) set was contracted using the prescription established in the preceding two sections, namely, the three most diffuse Gaussian primitives were added to the 1*s* orbital of the hydrogen atom (see also below).
- ³⁵B. Liu, *J. Chem. Phys.* **80**, 581 (1984).
- ³⁶W. Kolos and L. Wolniewicz, *J. Chem. Phys.* **43**, 2429 (1965).
- ³⁷See, for example, C. W. Bauschlicher, Jr. and P. Taylor, *J. Chem. Phys.* **86**, 858, 1420, 2844 (1987); C. W. Bauschlicher, Jr. and S. R. Langhoff, *ibid.* **86**, 5595 (1987); C. W. Bauschlicher, Jr. and P. Taylor, *ibid.* **86**, 5600 (1987); C. W. Bauschlicher, Jr., S. P. Walch, S. R. Langhoff, P. R. Taylor, and R. L. Jaffe, *ibid.* **88**, 1743 (1988).