Theor Chem Acc (2000) 104:385–391
DOI 10.1007/s002140000149

Theoretical
Chemistry Accounts

Regular article

Contracted basis sets for electrical property calculations derived from Second-order Møller–Plesset theory atomic natural orbitals

Anthony J. Russell[1], **Mark A. Spackman**[2]

[1]Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia
[2]Chemistry, School of Physical Sciences and Engineering, University of New England, Armidale, NSW 2351, Australia

Received: 20 December 1999 / Accepted: 15 February 2000 /

Abstract. Using established methods based on correlated atomic natural orbitals (ANOs), sets of contracted polarization functions are derived for use in calculations of atomic and molecular electrical properties (especially electric moments, dipole polarizabilities and related property hypersurfaces). Through test calculations on Ne, Ar, NH_3 and CO_2, these polarization functions are shown to reproduce the accuracy of larger basis sets, to incorporate dynamical electron correlation effects and are economical to use in conjunction with sophisticated electron-correlation treatments. We also show how triple-zeta polarized ANO and double-zeta polarized ANO basis sets are constructed from these contracted polarization functions for use in the calculation of reliable zero-point vibrational averages of electrical properties.

Key words: Atomic natural orbitals – Vibrational averaging – Gaussian basis sets – Electric moments – Dipole polarizability

1 Introduction

Calculations of vibrational averages of electrical properties and pure vibrational polarizabilities require basis sets that can not only reproduce the electrical properties and their derivatives with accuracy, but that can also provide a good description of the energy hypersurface (the derivatives of the energy). Reliable estimates of the geometries and vibrational frequencies of the molecule are necessary since it has been concluded from a number of studies of anharmonic force constants that the largest source of error in these constants arises from errors in the geometry and harmonic frequencies and not from errors in the cubic, quartic and higher-order force constants [1, 2]. Our recent work has been aimed at the careful determination of reliable zero-point vibrational corrections and equilibrium vibrationless values of dipole and quadrupole moments and the dipole polarizability tensor for the first-row hydrides [3], acetylene [4] and the second-row hydrides [5]. In those publications we used a variety of different basis sets, of different sizes and for different purposes. The work on the first-row hydrides employed a modified version of Sadlej's POL1 basis set [6], for acetylene we used Dunning's triple-zeta (TZ) substrate [7] augmented with correlation-consistent polarization functions [8] as well as functions chosen by a crude optimization procedure, while for the second-row hydrides the TZV substrate of Schäfer and coworkers. [9, 10] was used in conjunction with correlation-consistent polarization functions [8, 11] and modified versions of contracted *d* functions from Sadlej [6, 12]. The latter work also reported the successful use of a more general basis set constructed for the calculation of highly accurate properties at experimental r_e geometries and based on second-order Møller–Plesset theory (MP2) atomic natural orbitals (ANOs). The purpose of the present article is to provide a more detailed description of the method of construction of these MP2-ANO basis sets, especially the polarization functions, to validate their use on atoms and small molecules by comparing results with those from uncontracted basis sets, and to describe the way in which related, but more compact, TZ polarization (TZP)-ANO and double-zeta polarization (DZP)-ANO basis sets are constructed.

Much work has been reported on determining the optimum basis set requirements for computing geometries and harmonic frequencies for a variety of systems [13–20]. Basis sets for accurate prediction of electrical properties have been successfully constructed using either well-founded rules of thumb [21] or a systematic method [22–27]. Optimizing basis sets for reproducing specific properties may be a troublesome and time-consuming task with an inordinate amount of trial and error involved in the process; this is especially true if the basis set is to be restricted to a manageable size. Dynamical

---

Correspondence to: M. A. Spackman
e-mail: mspackma@metz.une.edu.au

electron-correlation effects are rarely incorporated into the one-particle basis set as most basis set optimization for electrical properties takes place at the self-consistent-field (SCF) level. This is despite the fact that the majority of electrical property calculations are subsequently performed at correlated levels of theory. Therefore, the present work was motivated by a need to obtain, via a simple method, electrical property basis sets that reproduce the accuracy of larger basis sets, incorporate dynamical electron correlation effects and that are also economical to use in combination with sophisticated electron correlation treatments.

Basis sets for molecular properties require polarization and diffuse functions to describe the electron distribution and its perturbation at a range of distances from the nuclei. If a basis set contains sufficient functions to effectively saturate the orbital space for a given symmetry type or angular momentum, then accurate molecular properties should be obtained. Even-tempered basis set expansions should satisfy this criterion provided that the parameters used to define the even-tempered series are carefully chosen. Basis sets constructed in this way have proven to be particularly useful in obtaining the convergence behaviour for many molecular properties [22, 28] and although there is no mathematical theorem to suggest that an even-tempered expansion will offer the fastest convergence to the Hartree–Fock limit or basis set limit for a given property, Woon and Dunning [22] have found it to be more reliable and stable than other construction methods. The primary difficulty that arises from the even-tempered method is the rapid increase in basis set size, thus leading to sets that either suffer from linear-dependence problems or are too large for high-level electron-correlation calculations on molecules. Correlation-consistent basis sets developed by Woon and Dunning are based on augmented correlation-consistent polarized valence X–zeta (aug-cc-pVXZ) sets where X = D, T, Q, 5 for double-, triple-, quadruple- and quintuple-zeta sets. The aug-cc-pVTZ basis set is a suitable candidate for calculation of electrical properties since it has been systematically constructed with augmenting functions optimized for anion energies. However, the size of the basis set, represented as [5s4p3d2f/4s3p2d], places it out of reach for many molecular studies, especially vibrational averaging of properties for larger polyatomic molecules. The approach taken in the present work involves construction of a slightly smaller [6s4p3d1f/4s3p1d] basis set for electrical property calculations that yields an accuracy greater than, or equal to, the accuracy of the aug-cc-pVTZ basis set for electrical properties. Atomic units are used throughout (conversion factors to SI units are provided elsewhere [29]).

2 Atomic natural orbitals

ANO contraction methods were originally proposed by Almlöf and coworkers [30–32] to deliver efficient contracted basis sets possessing small contraction errors for molecular properties with respect to large uncontracted primitive sets. In their landmark papers, large uncontracted (13s8p) substrates were augmented with both polarization and diffuse functions then generally contracted using configuration interaction singles and doubles (CISD) ANOs to yield smaller basis sets that reproduced both the SCF energy and the correlation energy with high accuracy. While their method was extremely successful for the energy and slightly less successful for multipole moments, the method appeared not to perform as well for response properties such as the polarizability [32]; extra diffuse functions were required to augment the ANO basis sets to obtain sufficient accuracy. Widmark et al. [24] have also utilized density-matrix-averaged ANOs in contracting basis sets for electrical properties with some success for computation of ionization potentials. However, the accuracy of polarizabilities using the large [6s5p4d3f] sets was less satisfactory.

The basis set construction method proposed in the present work is based on MP2 ANOs for obtaining highly accurate basis sets that reproduce as closely as possible the electrical properties of the uncontracted set. It is hoped that if the characteristics of even-tempered expansions can be reproduced within smaller basis sets by contracting the even-tempered functions through the use of ANOs, then highly accurate and cost-effective basis sets should result. Both Almlöf and Taylor [30, 32] and Widmark et al. [24] contracted the entire set (substrate and polarization functions) using ANOs to obtain improved energies and to minimize contraction errors in the substrate. However, in the present study, the substrate is left completely contracted in the construction of the basis sets so that only the polarization functions are contracted using MP2 ANOs. Although MP2 ANOs have been utilized in the present work, similar results are expected using ANOs from CISD or other correlated wavefunctions. One particular criticism of ANO sets is the heavy contraction of large numbers of primitive Gaussians into a single basis function, thereby increasing the time required to compute the two-electron integrals. This problem may be partially remedied by using general contractions; however, the majority of computational codes utilize segmented contractions and consequently these basis sets are probably unsuitable for SCF calculations where the integral calculation would be the limiting step. For correlated methods that scale approximately as n^5, iterative n^6 or n^7 (for n basis functions) the computation of integrals is no longer the limiting step of the calculation and in many instances needs only be performed once; however, as the smaller ANO contracted sets should obtain most of the accuracy of the uncontracted sets, the cost is effectively the same. The method of basis set construction is now outlined in detail.

3 Construction of ANO basis sets for first-row and second-row atoms

Restricted open-shell MP2 has been outlined by Amos et al. [33] and Knowles et al. [34]. The good quality contracted TZV substrates of Schäfer and coworkers [9, 10] were used and may be represented by [3s], [5s3p] and [5s4p] for hydrogen, first-row and second-row atoms, respectively. The substrates were left fully contracted in the construction procedure. Polarization function exponents were taken from Dunning and Woon [8, 11, 22], where the d- and f-type cc-pVTZ and s-, p-, d- and f-type aug-cc-pVTZ exponents were utilized.

For heavy atoms, the *s*, *p* and *f* orbital spaces were extended by two functions each and the *d* orbital space by four additional functions, which should be sufficient for the present purposes. For hydrogen, the *s*, *p* and *d* orbital spaces were increased by two, four and two functions, respectively. As an example of the construction procedure, the substrate and augmenting function exponents for oxygen are illustrated in Table 1. The most diffuse s- and p-type function exponents were determined by even-tempering the aug-cc-pVTZ s- and p-type functions with respect to the most diffuse functions of the substrate. The most diffuse d-type function was generated by an even-tempered sequence of the exponents for the aug-cc-pVTZ and cc-pVTZ basis sets. A similar construction procedure was followed for hydrogen, resulting in the following exponents of the augmenting functions: two of s type (0.039469, 0.011081); four of p type (1.407, 0.388, 0.102, 0.026814); two of d type (1.057, 0.247). The most diffuse s- and p-type functions for hydrogen were derived from even-tempered expansions with the s-type exponents scaled by $(1.25)^2$ to match the scaled TZV substrate s-type exponents [9, 10]. All other exponent values for hydrogen were taken from cc-pVTZ or aug-cc-pVTZ basis sets. The size of the resulting basis set, which includes a contracted TZV substrate and uncontracted polarization and diffuse functions, is [7s5p4d2f/5s4p2d] for first-row atoms/hydrogen and [7s6p4d2f] for second-row atoms.

For each atom, closed-shell and open-shell MP2 gradient calculations were then performed for the ground electronic states of the atoms to obtain the MP2 relaxed density matrix which was subsequently diagonalized to obtain the natural orbitals and occupation numbers. The MP2 ANOs were averaged and contracted to yield a basis set for first-row and second-row atoms of [6s4p3d1f] and [6s5p3d1f], respectively. Although appropriate contraction coefficients can be obtained from the averaging formulae reported for p-, d- and f-type functions [24], the averaging for d- and f-type functions was actually carried out using the simple average of the six d-type and ten f-type component coefficients. The difference in the resulting contraction coefficients for the two methods is quite small. The contraction coefficients for the s-, p-, d- and f-type primitives were also normalized for use in molecular calculations.

Obviously, the MP2–ANO contraction scheme cannot be applied to the hydrogen atom and other one-electron species; therefore, a suitable hydrogen atom basis set was derived from MP2 calculations on the H_2 molecule, at the r_e bond length of 0.74144 Å [35], with the MP2-ANO contraction method used to give a basis set represented as [4s3p1d]. Larger uncontracted sets were also used in the present procedure by adding extra s and p functions, but linear dependence problems generally resulted.

The basis set specification for both the substrate and the resulting contracted polarization basis functions is summarized in Table 2; lists of exponents and contrac-

Table 1. Oxygen atom s-, p-, d- and f-type function exponents (in atomic units) for the second-order Møller–Plesset theory (*MP2*) – atomic natural orbital (*ANO*) basis set. *Shaded functions* belong to the triple-zeta valence (*TZV*) (11s6p)/[5s3p] substrate from Schäfer and coworkers [9, 10]

s-type exponents	
27032.382631	Substrate
4052.3871392	
922.32722710	
261.24070989	
85.354641351	
31.035035245	
12.260860728	Substrate
4.9987076005	
1.1703108158	Substrate
0.46474740994	Substrate
0.18504536357	Substrate
0.073760	aug-cc-pVTZ
0.029401	Even-tempered
p-type exponents	
63.274954801	Substrate
14.627049379	
4.4501223456	
1.5275799647	
0.52935117943	Substrate
0.17478421270	Substrate
0.059740	aug-cc-pVTZ
0.020419	Even-tempered
d-type exponents	
2.314	cc-pVTZ
0.614	cc-pVTZ
0.214	aug-cc-pVTZ
0.071002	Even-tempered
f-type exponents	
1.428	cc-pVTZ
0.5	aug-cc-pVTZ

Table 2. Summary of the basis set specification and comparison between sizes of the completely uncontracted basis sets and those with a contracted TZV substrate and both uncontracted and MP2-ANO contracted sets of polarization functions

Atom	Contraction scheme	Completely uncontracted basis set	Contracted substrate; uncontracted polarization functions	Contracted substrate; contracted polarization functions
H	41111/1111/11	8s4p2d	5s4p2d	4s3p1d
B-Ne	6211111/41111/1111/11	13s8p4d2f	7s5p4d2f	6s4p3d1f
Al-Ar	7321111/611111/1111/11	16s11p4d2f	7s6p4d2f	6s5p3d1f

tion coefficients of these augmenting functions can be obtained from the authors. Cartesian six-membered d-type and ten-membered f-type basis functions were utilized within the ab initio program CADPAC [36] for all calculations. We note that the present method is entirely general and is only limited by the number of primitives that may be contracted into a given shell; therefore, in principal, any combination of exponents and primitive Gaussians may be used, but, in practice, integral times will be substantially increased if the number of contracted primitives is excessively large. With the exception of hydrogen, the basis sets are also derived entirely from atomic calculations and do not include any explicit optimization for the molecular environment. However, field-dependent and/or anion-optimized functions were used in constructing the even-tempered expansion, so it is likely that these functions will compensate for rearrangement of the electron density within a molecular environment.

4 Validation of ANO basis sets for electrical properties of atoms and molecules

To demonstrate the accuracy of the present basis sets, a number of example calculations, using basis sets with both uncontracted and MP2–ANO contracted polarization functions, are reported in this section. It should be noted that all the ANO sets were explicitly designed for electrical properties and may be slightly deficient in reproducing geometries, energies and frequencies since high-exponent polarization functions were not added to improve the energies calculated from these sets. Contraction errors for the various properties and the accuracy of the sets with respect to size of the basis are the primary concern. The polarization functions obtained from the contraction procedure may be transferred to different substrates such as DZ. Additionally, some of the functions may be removed from a particular atom for correlated calculations where the size of the basis set must be restricted to save computational resources. Examples of such basis sets are described in Sect. 5.

4.1 Neon and argon

Mean polarizabilities of neon and argon atoms are presented in Table 3 for various basis sets at various levels of theory and they were calculated via finite-field methods with field strengths of 0.001 au. For neon the [7s5p4d2f] basis set with uncontracted polarization functions yields SCF results that are quite close to the Hartree–Fock limit for this property, comparable in quality to results from the d-aug-cc-pVTZ [6s5p4d3f] and t-aug-cc-pVTZ [7s6p5d4f] basis sets. The Brueckner doubles with perturbatively linked triples (BD(T)) theoretical estimate for the static (i.e. zero frequency) polarizability of neon agrees well with experimental estimates. Contraction of the polarization functions leads to an SCF polarizability of 2.361 au, which is more accurate than the aug-cc-pVTZ [5s4p3d2f] result of 2.194 au and which is closer to the d-aug-cc-pVTZ value. This is particularly encouraging as one of the criteria for using MP2–ANO contractions was to obtain the accuracy of the aug-cc-pVTZ set and in favourable circumstances to acquire the accuracy of the d-aug-cc-pVTZ set for the polarizability. An alternative method of constructing basis sets for electrical properties through optimizing diffuse functions in the presence of a field is a less efficient and more tedious process than the present single MP2 gradient calculation for an atom. A strong

Table 3. Mean polarizabilities of neon and argon obtained with various basis sets and levels of theory. All polarizabilities are in atomic units. The [7s5p4d2f] and [7s6p4d2f] basis sets have uncontracted polarization functions; the [6s4p3d1f], [6s4p2d], [6s5p3d1f] and [6s5p2d] basis sets incorporate MP2-ANO contracted polarization functions. The cc-pVTZ and aug-cc-pVTZ results were taken from Woon and Dunning [22]. *n* is the number of basis functions for the basis set (the values for the Woon and Dunning sets were calculated assuming six-member d-type and ten-member f-type functions, although Woon and Dunning carried out the calculations using spherical harmonic five-member d-type and seven-member f-type basis functions)

Basis set	*n*	SCF	MP2	BD	BD(T)	Experiment
Ne						
[7s5p4d2f]	56	2.376	2.699	2.619	2.673	2.67(2)[a], 2.669[b]
[6s4p3d1f]	36	2.361	2.667	2.593	2.641	
[6s4p2d]	30	2.325	2.632	2.573	2.618	
cc-pVTZ	35	1.026	1.028			
aug-cc-pVTZ	55	2.194	2.437			
d-aug-cc-pVTZ	75	2.375	2.711			
t-aug-cc-pVTZ	95	2.375	2.712			
Ar						
[7s6p4d2f]	69	10.697	11.132	11.063	11.153	11.092(9)[a], 11.08[b]
[6s5p3d1f]	49	10.626	10.937	10.882	10.950	
[6s5p2d]	33	10.204	10.468	10.481	10.554	
aug-cc-pVTZ	59	10.438	10.814			
d-aug-cc-pVTZ	79	10.701	11.169			
t-aug-cc-pVTZ	99	10.702	11.170			

[a] Static mean polarizability from a quadratic extrapolation of refractivity data [42]

[b] Static dipole oscillator strength distributions estimate [43, 44]

disadvantage of any exponent-optimization method is that removal of the optimized diffuse functions often leads to poor predictions of the properties if other diffuse functions are not present in the basis set. This is not the case for the MP2–ANO contracted set, as demonstrated by the results in Table 3 for neon, where the removal of the third d-type polarization function and f-type function, to give the [6s4p2d] set, does not markedly affect the value of the mean polarizability.

The MP2–ANO [6s5p3d1f] contracted basis set for argon appears to be less successful in reproducing the accuracy of the uncontracted set since, at the MP2 level, the differences between polarizabilities are larger than found for neon. Nevertheless, the MP2–ANO set still betters the accuracy of the aug-cc-pVTZ basis, yielding an MP2 mean polarizability of 10.937 au, which is in reasonable agreement with (2.1% below) the d-aug-cc-pVTZ result of 11.169 au. Core correlation may be responsible for some of the discrepancy at the MP2 level between the present [7s6p4d2f] set with uncontracted polarization functions and the results of Woon and Dunning for the largest augmented basis sets, since the frozen-core approximation was used by Woon and Dunning for their finite-field polarizability calculations. However, it should also be noted that atoms perhaps represent a worst-case scenario for polarizability calculations as there is no possibility of interatomic polarization, such as that found for molecules [37].

4.2 Ammonia

Although the MP2–ANO basis sets were developed from atomic calculations, they perform equally well for molecular property calculations. For the calculations reported in Table 4, the TZPOL1 basis set (reported in Ref. [5] as TZP1) was utilized for the hydrogen atom, with an additional *d*-type function (exponent of 1.057 au). This basis set was used for hydrogen in these calculations in order to focus on the contraction error for the nitrogen atom rather than utilizing contracted sets for both nitrogen and hydrogen. The experimental r_e geometry of Benedict and Plyler [38] ($r_{NH} = 1.0124$ Å, $\theta_{HNH} = 106.67°$) was used for these calculations. Energies and electrical properties of NH_3 are presented in Table 4 for basis sets with uncontracted and MP2–ANO contracted polarization functions. The contraction error for the SCF energy is only 5.8 m Hartrees and this is only slightly increased to 6.1 m Hartrees at the MP2 level. For dipole and quadrupole moments, differences between the uncontracted and contracted basis set values for both the SCF and MP2 levels of theory are relatively insignificant; less than 0.0014 au and less than 0.0025 au for the dipole and quadrupole moments, respectively. Mean polarizability differences are also quite small (0.003 au) for the SCF level and are only a little larger at the MP2 level (0.005 au). Generally for NH_3, the contraction errors for the properties from this basis set on nitrogen are exceptionally small.

Table 4. Self-consistent field (SCF) and MP2 energies and electrical properties of ammonia using basis sets with sets of uncontracted and MP2-ANO contracted polarization functions on the nitrogen atom. All properties are in atomic units. For all calculations the basis set for hydrogen is the TZPOL1 [4s3p1d] basis set [5] with an additional d-type function (exponent of 1.057 au). The uncontracted basis set has uncontracted polarization functions on nitrogen, represented as [7s5p4d2f]; 123 basis functions for the molecule. The contracted basis set has polarization functions on nitrogen constructed from MP2 ANOs, represented as [6s4p3d1f]; 103 basis functions for the molecule. Quadrupole moments refer to an origin at the centre of mass

Property	SCF		MP2	
	Uncontracted	Contracted	Uncontracted	Contracted
E	−56.222803	−56.222745	−56.484156	−56.478072
μ_z	−0.6374	−0.6372	−0.6027	−0.6013
Θ_{zz}	−2.1515	−2.1490	−2.2405	−2.2395
α	12.948	12.945	14.262	14.257
$\Delta\alpha$	0.530	0.520	1.844	1.840

The results in Table 5 demonstrate the contraction error for the properties when basis sets with MP2-ANO contracted polarization functions are used on both nitrogen and hydrogen atoms. Although the contraction error for the SCF dipole moment is exceedingly small (less than 0.0001 au), contraction errors for other properties are somewhat larger for this basis set. A different basis set optimization procedure should probably be used for hydrogen as the polarization functions from the H_2 molecule MP2 ANOs may not be well-suited towards general molecular calculations. Nevertheless, the contracted basis set yields relatively good quality results which compare quite well with the original [7s5p4d2f/5s4p2d] basis set values reported in the table.

4.3 Carbon dioxide

Highly accurate electrical property calculations using electron-correlation treatments such as fourth-order Møller–Plesset theory (MP4), BD(T) and coupled-clusters single-double and perturbative triple excitation [CCSD(T)] require considerable computational time and disk storage. Approximate costs for BD(T) and

Table 5. SCF and MP2 calculations on ammonia using basis sets with the uncontracted and MP2-ANO contracted polarization functions for both nitrogen and hydrogen atoms. All properties are in atomic units. The uncontracted basis set has uncontracted polarization functions on nitrogen and hydrogen, [7s5p4d2f/5s4p2d]; 153 basis functions for the molecule. The contracted basis set has contracted polarization functions on both nitrogen and hydrogen constructed from MP2 ANOs, [6s4p3d1f/4s3p1d]; 103 basis functions for the molecule. The basis set for hydrogen was constructed using MP2 ANOs for the H_2 molecule

Property	SCF		MP2	
	Uncontracted	Contracted	Uncontracted	Contracted
E	−56.222863	−56.222053	−56.485223	−56.472205
μ_z	−0.6375	−0.6376	−0.6022	−0.6033
Θ_{zz}	−2.1459	−2.1554	−2.2359	−2.2527
α	12.954	12.935	14.293	14.302
$\Delta\alpha$	0.516	0.543	1.847	1.909

CCSD(T) theory are iterative n^6 + a single n^7 step for n basis functions, while MP4 requires a single n^7 step. The basis sets incorporating MP–ANO contracted polarization functions offer one way to circumvent the computational cost problems since they reduce the number of basis functions to an acceptable size while still retaining nearly all the accuracy of the uncontracted set. Calculations of the mean polarizability and polarizability anisotropy for CO_2 using a variety of basis sets, at the experimental r_e geometry of Graner et al. [39, 40] (r_{co} = 1.15995884 Å), are presented in Table 6. For these calculations, the finite-field method was utilized with field strengths of 0.002 au. As calculations with f-type functions on all atoms with the [6s4p3d1f] basis set exceeded available computational resources for BD(T) calculations, f-type functions on oxygen were removed and the d-type functions were uncontracted, giving a [6s4p4d1f/6s4p4d] set for carbon/oxygen, thereby saving 20 basis functions. The contracted set used in Table 6 is built from MP2–ANO contracted d-type functions.

From Table 6 we see that the contraction of the d-type functions only leads to small differences in the mean polarizability with a maximum contraction error of 0.033 au found for MP2 theory (amounting to less than 0.2% of α from the uncontracted basis set). Differences between results for the polarizability anisotropy for the two basis sets are also quite small with respect to the uncontracted property. However, the uncontracted basis set calculation required 1.8 times the disk storage and 1.7 times the computation time of the contracted basis set calculation. Obviously, the required computational cost must be weighed against the accuracy desired from the calculation. From the results in Tables 3–6, the MP2–ANO contraction scheme offers another way of retaining accuracy in a truncated basis set while avoiding the time-consuming optimization of exponents for a given molecule or basis set.

5 Basis sets used in vibrational averaging calculations

As described in the Introduction, we used the contracted polarization functions obtained in this work in a variety of ways, sometimes coupled with TZ substrates [5] and for larger molecules, such as the fluoromethanes, ethylene and ethane, coupled with the smaller DZ substrates. In addition, for some purposes additional polarization functions were added, while other calculations dictated the judicious removal of some polarization functions. Here we summarize the derivation of the basis sets used in those studies; many other combinations are also possible.

Table 6. Carbon dioxide mean polarizability, α, and polarizability anisotropy, $\Delta\alpha$, using basis sets with uncontracted and MP2-ANO contracted polarization functions. All properties are in atomic units. The uncontracted basis set used uncontracted d-type polarization functions, [6s4p4d1f/6s4p4d] (carbon/oxygen); 136 basis functions for the molecule. The contracted basis set has d-type polarization functions contracted using MP2 ANOs, [6s4p3d1f/6s4p3d]; 118 basis functions for the molecule

Basis set	SCF	MP2	BD	BD(T)
α				
Uncontracted	15.849	17.835	17.299	17.521
Contracted	15.836	17.802	17.277	17.497
Difference	0.013	0.033	0.022	0.024
$\Delta\alpha$				
Uncontracted	11.711	14.545	13.880	13.943
Contracted	11.812	14.659	13.997	14.061
Difference	−0.101	−0.114	−0.117	−0.118

5.1 TZP-ANO1 basis set

The TZP-ANO1 [6s6p4d1f/4s3p1d] basis set used for the second-row hydrides [5] was constructed using the MP2–ANO contracted polarization functions for second-row atoms and hydrogen and the following additional functions for second-row atoms:

1. A sharper d-type function (exponents for Si: 1.455101 au; P: 1.968074 au; S: 2.493535 au; Cl:3.180570 au) was added to the MP2–ANO [6s5p3d1f] basis set.
2. The diffuse p-type functions were uncontracted.

5.2 DZP-ANO*x* basis sets

For studies on the fluoromethanes, ethylene and ethane, the DZP-ANO basis set was constructed from the DZ substrate of Thakkar et al. [41] and the complete set of MP2–ANO contracted polarization functions. Recipes for obtaining the relevant DZP-ANO*x* sets (where *x* denotes the number of d-type polarization functions) are

1. DZP-ANO3: the f-type function was deleted from DZP-ANO.
   DZP-ANO2: the third d-type function was deleted from DZP-ANO3.
   DZP-ANO2(*f*): the third d-type function was deleted from DZP-ANO.
   DZP-ANO1: the first and third d-type functions were deleted from DZP-ANO3.
2. The basis set for hydrogen was constructed from the (DZ) substrate of Thakkar et al., with one s and two p augmenting functions as follows:
   s type 0.025260 au
   p type (1.407 au, 0.388 au), contraction coefficients of (0.5, 0.5)
   p type (0.150568 au, 0.077927 au), contraction coefficients of (0.5, 0.5)

6 Concluding remarks

It has been shown that basis sets possessing MP2–ANO contracted polarization and diffuse functions are capable of reproducing electrical properties of atoms and small molecules with an accuracy close to that of their uncontracted counterparts. Polarization and diffuse functions obtained from the present contraction scheme should be transferable to other substrates, such as DZ, with minimal loss of accuracy. Full contraction of the substrates and polarization functions would prove to be computationally very costly and likely to lead to inefficient basis sets; this procedure has been avoided in the present work. The transferability approach has been used for the construction of the DZP-ANOx type basis sets ($x = 1, 2, 3$), which have been shown to be successful in our more recent studies on the fluoromethanes, ethylene and ethane, where the MP2-ANO polarization and diffuse functions were combined with the DZ substrate of Thakkar et al. [41].

Although electrical properties, such as the polarizability, have been targeted in this study, the MP2–ANO contraction is open-ended and may be used with any combination of Gaussian primitives or choice of molecular properties. It is not recommended that these sets be used for SCF calculations since the time to compute two-electron integrals is expected to dominate such computations. However, the MP2–ANO sets are suitable for accurate electrical property investigations involving high-level electron-correlation methods, where the cost of the calculations is dominated by the approximate iterative n^6 + single step n^7 cost for n basis functions and not by the evaluation of two-electron integrals.

Acknowledgements. This work has been supported by the Australian Research Council through research funding as well as the award of an Australian Postgraduate Research Award to A.J.R. We are grateful to D. Jayatilaka, University of Western Australia, for many enlightening discussions about ANOs. M.A.S. is grateful to J.A.K. Howard, University of Durham, for generous access to facilities in Durham, where this work was completed.

References

1. Allen WD, Yamaguchi Y, Csaszar AG, Clabo DA, Remington RB, Schaefer HF (1990) Chem Phys 145: 427
2. Allen WD, Császár AG (1993) J Chem Phys 98: 2983
3. Russell AJ, Spackman MA (1995) Mol Phys 84: 1239
4. Russell AJ, Spackman MA (1996) Mol Phys 88: 1109
5. Russell AJ, Spackman MA (1997) Mol Phys 90: 251
6. Sadlej AJ (1988) Collect Czech Chem Commun 53: 1995
7. Dunning TH Jr (1971) J Chem Phys 55: 716
8. Dunning TH (1989) J Chem Phys 90: 1007
9. Schäfer A, Horn H, Ahlrichs R (1992) J Chem Phys 97: 2571
10. Schäfer A, Huber C, Ahlrichs R (1994) J Chem Phys 100: 5829
11. Woon DE, Dunning TH (1994) J Chem Phys 98: 1358
12. Sadlej AJ (1991) Theor Chim Acta 79: 123
13. Simandiras ED, Rice JE, Lee TJ, Amos RD, Handy NC (1988) J Chem Phys 88: 3187
14. Lee TJ, Rendell AP (1991) Chem Phys Lett 177: 491
15. Thomas JR, DeLeeuw BJ, Vacek G, Crawford TD, Yamaguchi Y, Schaefer HF (1993) J Chem Phys 99: 403
16. Green WH, Jayatilaka D, Willetts A, Amos RD, Handy NC (1990) J Chem Phys 93: 4965
17. Martin JML (1994) J Chem Phys 100: 8186
18. Martin JML, Taylor PR (1996) Chem Phys Lett 248: 336
19. Martin JML, Taylor PR, Lee TJ (1993) Chem Phys Lett 205: 535
20. Martin JML, Lee TJ, Taylor PR (1992) J Chem Phys 97: 8361
21. Dunning TH, Hay PJ (1977) In: Schaefer HF (ed) Modern theoretical chemistry. Plenum, New York, p 1
22. Woon DE, Dunning TH (1994) J Chem Phys 100: 2975
23. Pierloot K, Dumez B, Widmark P-O, Roos BO (1995) Theor Chim Acta 90: 87
24. Widmark P-O, Malmqvist P-Å, Roos BO (1990) Theor Chim Acta 77: 291
25. Spackman MA (1989) J Phys Chem 93: 7594
26. Dougherty J, Spackman MA (1994) Mol Phys 82: 193
27. Papadopoulos MG, Waite J, Buckingham AD (1995) J Chem Phys 102: 371
28. Feller D (1993) J Chem Phys 98: 7059
29. Maroulis G, Thakkar AJ (1990) J Chem Phys 93: 4164
30. Almlöf J, Taylor PR (1987) J Chem Phys 86: 4070
31. Almlöf J, Helgaker T, Taylor PR (1988) J Phys Chem 92: 3029
32. Almlöf J, Taylor PR (1990) J Chem Phys 92: 551
33. Amos RD, Andrews JS, Handy NC, Knowles PJ (1991) Chem Phys Lett 185: 256
34. Knowles PJ, Andrews JS, Amos RD, Handy NC, Pople JA (1991) Chem Phys Lett 186: 130
35. Huber KP, Herzberg G (1979) Constants of diatomic molecules. Van Nostrand-Reinhold, New York
36. Amos RD, Alberts IL, Andrews JS, Colwell SM, Handy NC, Jayatilaka D, Knowles PJ, Kobayashi R, Noga N, Laidig KE, Maslen PE, Murray CW, Rice JE, Sanz JE, Simandiras ED, Stone AJ, Su M-D (1994) CADPAC: The Cambridge Analytic Derivatives Package, issue 5.2
37. Dykstra CE (1988) Ab initio calculation of the structures and properties of molecules. Elsevier, Amsterdam
38. Benedict WS, Plyler EK (1957) Can J Phys 35: 1235
39. Graner G, Rossetti C, Bailly D (1986) Mol Phys 58: 627
40. Graner G (1992) In: Domenicano A, Hargittai I (eds) Accurate molecular structures. Their determination and importance. Oxford University Press, New York, p 65
41. Thakkar AJ, Koga T, Saito M, Hoffmeyer RE (1993) Int J Quantum Chem Quantum Chem Symp 27: 343
42. Cuthbertson C, Cuthbertson M (1909) Proc R Soc Lond Ser A 83: 149
43. Kumar A, Meath WJ (1985) Mol Phys 54: 823
44. Kumar A, Meath WJ (1985) Can J Chem 63: 1616