Theor Chim Acta (1995) 90: 87–114

Theoretica
Chimica Acta

# Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wave functions

## IV. Medium size basis sets for the atoms H–Kr

**Kristine Pierloot**[1], **Birgit Dumez**[1], **Per-Olof Widmark**[2], **Björn O. Roos**[3]

[1] Department of Chemistry, University of Leuven, Celestijnenlaan 200F,
B-3001 Heverlee-Leuven, Belgium

[2] IBM Sweden, P.O.B. 4104, S-203 12 Malmö, Sweden

[3] Department of Theoretical Chemistry, Chemical Centre, P.O.B. 124, S-221 00 Lund, Sweden

Received May 4, 1994/Accepted August 11, 1994

**Summary.** Generally contracted Basis sets for the atoms H–Kr have been constructed using the atomic natural orbital (ANO) approach, with modifications for allowing symmetry breaking and state averaging. The ANO's are constructed by averaging over the most significant electronic states, the ground state of the cation, the ground state of the anion for some atoms and the homonuclear diatomic molecule at equilibrium distance for some atoms. The contracted basis sets yield excellent results for properties of molecules such as bond-strengths and -lengths, vibrational frequencies, and good results for valence spectra, ionization potentials and electron affinities of the atoms, considering the small size of these sets. The basis sets presented in this article constitute a balanced sequence of basis sets suitable for larger systems, where economy in basis set size is of importance.

**Key words:** Atomic natural orbitals – Basis sets – General contraction

## 1 Introduction

Density matrix averaged atomic natural orbital (ANO) basis sets [1] for the first and second row atoms H–Ar have recently been published [2, 3] and basis sets for atoms Sc–Zn are submitted for publication [4]. The contraction coefficients for these basis sets were obtained by computing the natural orbitals from an averaged density matrix. Singles and doubles configuration interaction (SDCI) was performed for the atom in its electronic ground state[1], the cation and anion[2] and the atom in its electronic ground state placed in a weak homogeneous electric field. The final density matrix used to construct the ANO's was obtained as an average of the density matrices obtained from these SDCI wave functions. The resulting ANO's give simultaneous accurate values for the ionization potential, electron affinity and polarizability of the atoms. The truncation errors for these properties were found to be very small. A number of calculations on small and medium sized

---

[1] For the transition metal atoms the $d^n$, $d^{n-1}s$ and $d^{n-2}s^2$ states.

[2] For the transition metal atoms only cation.

systems has been performed and shows that accurate results can be obtained with these ANO basis sets.

We present here less extensive basis sets for the atoms H–Kr using a procedure closely related to what was employed constructing the basis sets in [2–4] with a slightly different emphasis, *vide infra*. The present basis sets have not been published before, but have been used in a number of applications, see for example [5–7], and yield very satisfactory results. The primitive sets have been chosen to yield similar accuracy for all atoms, thus yielding basis sets of comparable quality throughout the periodic system. The contraction procedure omits the flexibility of the basis sets that is used to describe the electric polarizability of the atoms, thus reducing the size of the contracted basis sets needed to describe bonding and correlation in molecular systems.

When constructing these basis sets the emphasis has been put on the description of the bond formation process, thus yielding good results for bond-distances and -strengths, as well as other properties related to the shape of the potential curve close to the equilibrium. For SCF calculations, the results are expected to be close to the HF limit and for correlated wave functions the results are expected to be very close to what can be obtained with a basis set of the given size.

It must be stressed that properties such as polarizabilities and long-range forces are not well described by these basis sets unless they are augmented with extra basis functions.

All calculations have been performed with the MOLCAS-2 quantum chemistry software [27].

## 2 The primitive basis sets

The basic primitives have been taken from literature whenever basis sets of suitable size were available. In cases where it has not been possible to find primitive sets in the literature, the exponents have been optimized by varying three parameters in an expression that is an extension of the even tempered sequence [8],

$$\ln(\zeta_k) = c_{-1}/k + c_0 + c_1 k; \quad k = 1 \ldots, n; \quad \zeta_1 > \zeta_2 \ldots$$

The size of the basic primitive set has been chosen to yield approximately the same truncation errors for all atoms. For each sequence of atoms with the same number of occupied shells, the same number of primitives was used. This leads to a small bias towards a better description of the lighter elements, for example boron is slightly better described than fluorine. The selected primitive basis sets were augmented with polarization functions and diffuse functions:

- H–He: The ($6s$) set of Duijneveldt [9] has been used as the basic primitive set. This set was augmented with two $p$ polarization functions that were optimized with respect to the correlation energy of $H_2$ and He, respectively.
- Li–Be: The ($9s$) set of Duijneveldt [9] has been used as the basic primitive set. This set was augmented with three $p$ and two $d$ polarization functions that were optimized with respect to the correlation energy of $Li_2$ and Be, respectively.
- B–Ne: The ($9s5p$) set of Duijneveldt [9] has been used as the basic primitive set. This set was augmented with two $d$ polarization functions that were optimized with respect to the correlation energy of the atoms in their ground states.
- Na–Mg: The ($12s6p$) set of Huzinaga [10] has been used as the basic primitive set. This set was augmented with two $d$ polarization functions that were

optimized with respect to the correlation energy of the atoms in their ground states.

- Al–Ar: The (12*s*9*p*) set of Huzinaga [10] has been used as the basic primitive set. This set was augmented with three *d* polarization functions that were optimized with respect to the correlation energy of the atoms in their ground states.
- K–Ca: No primitive sets of suitable size were found in the literature, and sets of size (16*s*10*p*) were optimized at the SCF level. They were augmented with three *d* polarization functions, optimized with respect to the correlation energy of $K_2$ and Ca, respectively.
- Sc–Zn: The (16*s*11*p*8*d*) set of Fægri [11] has been used as the basic primitive set. This set was augmented with three *f* polarization functions that were optimized with respect to the correlation energy of the atoms in their ground states.
- Ga–Kr: No primitive sets of suitable size were found in the literature, and a set of size (16*s*14*p*8*d*) was optimized at the SCF level. These sets were not augmented with polarizing functions, except for the diffuse functions described below, due to the fact that the *d* functions are already present for the occupied *d*-shell.

All primitive sets were augmented further with one diffuse function per shell to improve the description of effects not present in SCF/CI optimized functions. The only exceptions to this rule are the alkali and alkaline earth atoms, Na, Mg, K and Ca, where a single added *p* function did not yield a satisfactory description of the 3*p*/4*p* orbitals. Two diffuse *p* functions were added for these atoms.

## 3 The contraction procedure

The basic philosophy of the contraction scheme is to produce basis sets containing the following functions:

1. Atomic Hartree–Fock orbitals with high accuracy.
2. Functions that describe the deformation of the atomic orbitals when bonds are formed.
3. Correlating functions that yield as much of the dynamic correlation as possible.
4. Functions that describe the deformation of the atomic orbitals when cations and anions are formed.
5. Functions that describe the deformation of the atomic orbitals arising from valence excitations, notably for transition metal atoms where the $d^n$, $d^{n-1}s$ and $d^{n-2}s^2$ often all contribute to the bond formation process.

Each item in the list gives rise to one or more functions describing the difference between the SCF solution of the atom and the other state/description, and the list seems to indicate that there is a need for at least 4–5 virtual basis functions per shell to do a reasonable job on any molecular system. Fortunately this is not the case, since all these functions form a nearly linearly dependent set, leading to a significant reduction in the number of degrees of freedom, and often a single virtual orbital per shell can do a reasonable job in molecular systems. The selection criteria are based on the eigenvalues of a density matrix that is the average of several density matrices from different states of the atom. Any eigenfunction with an eigenvalue that is exactly zero can be removed from the set of basis functions with no truncation error whatsoever for the states involved in the averaging. For example, averaging the SCF density matrices for the hydrogen atom and the hydrogen molecule would lead to a density matrix with two nonzero eigenvalues for the *s*-functions, since we have two slightly different *s*-functions in the two cases, and the inclusion of both in a contracted set would exactly reproduce the calculations performed with the primitive set. However, there would be a slight contraction error for the hydrogen molecule at other bond distances than the one used in the averaging.

The fact that we do get a set of near linear dependence can be made plausible by the following qualitative argument. The functions that describe bond formation at the SCF level are in many cases virtually indistinguishable from the, for each shell, first correlating natural orbitals from a CI calculation. The reason for this behaviour can be found in the physical processes involved in bond formation and dynamic correlation. When forming a bond, the occupied orbitals are deformed with the major density difference in the region where the electron density is the largest. The functions that are best suited to perform the task of moving density from the electron rich region are the following: to move the electron density in/out the function should have the *l*-quantum number as the orbital to be deformed, with a node at the distance of the deformed orbitals radial maximum. To perform an angular motion of the electron density the function should have a different *l*-quantum number with a coinciding radial maximum. The same qualitative arguments hold for correlating orbitals, thus bond forming and correlating orbitals should be similar, and in practice they turn out to be virtually indistinguishable in many cases and only the correlating orbitals need to be included.

Although the functions needed to describe the formation of cations and anions and valence excitations have the same nodal properties as the bond-forming/correlating functions, they are distinctly different and need thus to be included explicitly. This is nicely illustrated in Fig. 1 for the nitrogen atom, showing the radial shape of 3*s* orbitals defined for different purposes. The anion and cation orbitals were obtained as the third ANO of an averaged density matrix for the natural atom and the corresponding ion; the "correlating orbital" from an SDCI density matrix for the atom; the "molecule" orbital from an averaged density

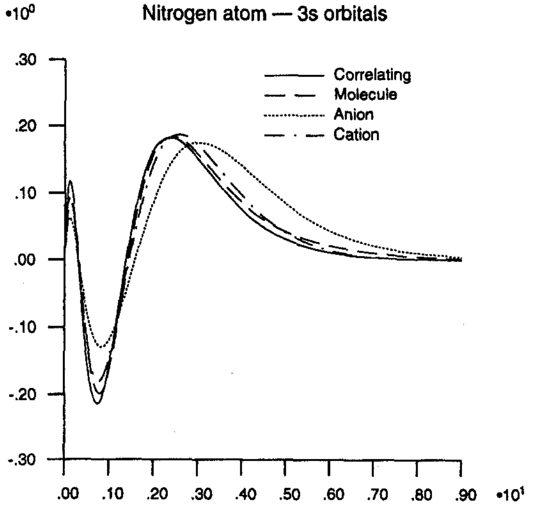

**Fig. 1.** Illustration of the shape of different types of 3*s*-functions

matrix involving the atom and the $N_2$ molecule. No polarizing orbital is shown, but it would be similar in shape to the anion orbital. In the present compilation only the cations and anions were included in the averaging, and this proved to include sufficient flexibility in the basis sets to describe the valence excitations for the transition metal atoms. See below for test results.

Functions describing polarizabilities and long-range forces are quite different in nature. The region affected the most is the outer region of the atom which is "soft" and easily deformed by small perturbations. Such degrees of freedom are not included in this compilation but can easily be included by uncontracting the outer region of the basis sets, see for example [12], or by adding functions explicitly designed for polarizabilities.

It is argued in [12] that accurate polarizabilities are necessary to describe the distortion of the atoms when bonds are formed. This is not entirely true. The bond forming functions mentioned above have the same principal shape as the functions describing the distortion of an atom in a weak electric field. The main difference is that the functions describing the polarizability are much more diffuse, and there is a smooth transition from one to the other when a bond is formed. It is possible to arrive at an accurate description of the binding situation around the equilibrium without accurate polarizabilities, but with a degradation of the long-range forces. Experience shows that the calculation of bond distances and other structure parameters are not significantly degraded by less well-described polarizabilities, provided that functions needed for a correct description of the chemical bonds are present. However, if a complete potential curve is computed for a diatomic molecule, say, the arguments in [12] are certainly valid.

The arguments presented lead to the following strategy for the contraction procedure employed in this compilation. All states are treated at the correlated level, either using SDCI (singles and doubles configurations interactions) or MCPF (modified coupled pair theory) wave functions, except for the one electron cases hydrogen and the alkali atoms. Only the valence electrons are correlated. Thus, functions describing core correlation effects are lacking. This implies that the present contracted sets are not suitable for correlating the semi-core $3s$,$3p$ electrons in the transition metals and the semi-core $3d$ electrons in Ga–Kr. However, the basis sets can of course always be extended by simply uncontracting the relevant primitives or by adding appropriate functions to describe these effects (see the application to Sc and ScF below). For the one electron systems hydrogen and the alkali atoms, the homo-nuclear diatomic molecules have been included in the averaging procedure. For the transition metal atoms it is essential to describe the valence excitations correctly, since very often two or more states contribute to the formation of bonds. However, the flexibility provided by the cations and anions proved to be enough, and only the $d^{n-2}s^2$ states were included in the averaging. The cations were included for all atoms except a few cases where it was deemed inappropriate, and the anions were included in all atoms that are likely to be negatively charged in any feasible system.

The result is a sequence of basis sets that will in most cases yield an accurate picture of the bonding situation provided that appropriate methods are used in the wave function calculation. For some systems these basis sets are simply too small to yield a highly accurate quantitative result, but at a qualitative and semi-quantitative level virtually all nonpathological systems should be correctly described provided that the basis sets are used within the limits of their design.

The test calculations in this paper, together with some already published studies [5–7] clearly illustrate the quality of these basis sets.

For all states with 3 or more valence electrons the MCPF method was used to generate the wave function, while for states with 2 valence electrons the SDCI method was used and finally for 1 valence electron states the SCF method was used. The elements have been grouped by their position in the periodic table, resulting in 5 different groups.

The first group of atoms, hydrogen and the alkali elements, are essentially one electron systems since the rare gas core is not correlated and relatively inert. To introduce correlation for these atoms the homo-nuclear diatomic molecules were included. In virtually all molecular system containing the alkali atoms they are in the form of cations. In a few cases these atoms are in a neutral state, but they are never negatively charged, expect for the anions in the gas-phase. Therefore, the averaging consists of the following states: X (50%), $X^+$ (20%), $X_2$ (20%) and X $^2P$ ($ns \rightarrow np$) (10%), X = Li, Na, K. Hydrogen was treated differently, since it is not uncommon to find hydrogen compounds with a negative charge on the hydrogen atom. $H^-$ is however very diffuse, and would tend to destroy the description of the atomic 1s function. Therefore, it was given only a very small weight in the averaging, which included the following states: H (49%), $H_2$ (49%) and $H^-$ (2%).

The second group of atoms, the alkaline earth elements, all lack electron affinities. Also the added electron would go into the empty *p* shell, and nothing can prevent it from escaping, except for the constraint imposed by the limited basis set size, thus yielding nonsense in any averaging. The *p* function, nearly degenerate with the highest occupied *s*-function, is better described by the combined effect of the correlating functions and the occupied orbital in the excited state $^3P$ ($ns^2 \rightarrow ns, np$), thus the averaging was performed over the states X (50%), $X^+$ (25%) and X $^3P(ns^2 \rightarrow ns, np)$ (25%), X = Be, Mg, Ca.

The third group of atoms, the transition metal elements, exhibit complex spectra for the neutral atom as well for the cations and anions, and bonds formed by these elements often involve a mixture of states. It is desirable to have the valence spectra well described by a basis set, and it is tempting to include "all" relevant states into the averaging. Fortunately it turns out that it is really not necessary to include that many different states. The reason is that what is really needed is a basis set containing the flexibility of describing all these states, i.e. the lowest virtual atomic orbitals should contain the differences in the radial extent of the orbitals between various states. Obviously there is a lot of overlap in these orbital differences, and we have found that by including the X ($d^{n-2}s^2$), $X^+$ ($d^{n-1}$) and $X^-$ ($d^{n-1}s^2$), X = Sc–Cu, states the required flexibility is obtained. For some molecular systems it is necessary to have a reasonable description of the 4*p* function that polarizes the 4*s* orbital, and for this reason the atoms in the X ($d^{n-2}s^2$) state in a weak electric field (0.05 au) was included in the averaging. All states were included with the same weight. Zinc was treated the same way except that no anion was included for obvious reasons.

The fourth group of atoms, the main group elements B–F, Al–Cl and Ga–Br, all have an *ns*, *np* valence and are treated equally. They all have positive electron affinities, except for nitrogen, and form compounds with both positive and negative charge on the atom, at least formally. It is therefore necessary to include both the cations and anions into the averaging as well as the neutral atom. It might be argued that including $N^-$ would tend to destroy the basis functions for nitrogen, but the added *p*-electron cannot escape from the nucleus at the SCF level since it is confined to the same orbital as the other *p*-electrons. The 2*p* orbital becomes more diffuse, but this is expected, and similar to what is experienced for the other atoms. Even at a correlated level the electron stays in the valence region and the

$2p_z^2 \rightarrow 2p_z3p_z$ excitations account for less than 3% of the total wave functions at the MCPF level. The averaging was performed over the states: X (50%), $X^+$ (25%) and $X^-$ (25%), X = B–F, Al–Cl, Ga–Br.

The fifth group of elements, the rare gases, are rather inert and do not participate in any bonding, except when the atom is in an ionic or excited state. The present compilation is for relatively small basis sets, and does not contain enough dynamic correlation to describe the weak complexes formed by these atoms in the ground state. There are a few cases where it is of interest to have rare gas basis sets of the present size, for example to study the effect of an argon matrix on a molecule in a matrix isolation study. Only the atom in its ground state has been included in the contraction of the atoms.

## 4 Test calculations

When assembling a collection of basis sets for general use it is imperative that a certain level of testing is performed to assess the quality of the basis sets. One reason for performing the testing is to make sure that no mistakes have been made during the process of generating the basis sets. At the SCF level the total energy of the atoms can be checked against the SCF energies cited in the literature to assess that no significant mistyping of the primitive exponents has occurred. These total energies will not be reported even though this checking has been performed.

Another, more important, reason for testing is to ensure that the desired effects are indeed included in the contracted sets. Two major atomic properties used to determine the quality of a basis set are the ionization potential and the electron affinity. Both properties have been calculated for all atoms, and are presented in Sect. 4.1. The transition metal basis sets have further been tested on the valence spectra of Ti, V and Ni (Sect. 4.4).

Apart from assessing the quality of atomic properties, there is a need to test the basis sets in actual molecular calculations. A few publications have already utilized these basis sets, and the results indicate that they are of good quality [5–7]. In this work we will present results for a few selected test molecules, namely $H_2$, CO and $P_2$ (Sect. 4.2), the halogen dimers $F_2$, $Cl_2$ and $Br_2$ (Sect. 4.3), and the ScF molecule (Sect. 4.5).

### *4.1 The ionization potential and electron affinity of the atoms*

The ionization potential of atoms is usually not too difficult to compute provided that a reasonably flexible basis set is used that includes correlating basis functions, and that major correlation effects are included in the calculation. In Tables 1–5 we present both the SCF results and the results obtained from MCPF calculations correlating only the valence electrons. As can be seen, both the basis set requirements and the demands on the correlation treatment increase with the number of valence electrons. Thus with the primitive basis sets the MCPF error, as compared to experiment, is only 0.05 eV for lithium, while it is 0.37 eV for fluorine. Similar trends are found for the higher rows: the MCPF error increases from 0.19 eV for sodium to 0.50 eV for chlorine and from 0.35 eV for potassium to 0.51 eV for bromine. For the transition metals errors ranging from 0.31 eV for titanium to 0.52 eV for zinc are found. Scandium has deliberately been left out from this series: it shows an exceptionally large error, 0.74 eV, which, as we will show in Sect. 4.5, is

**Table 1.** The ionization potential (eV) of the atoms H–He

| | SCF | | MCPF | |
|---|---|---|---|---|
| Basis | H | He | H | He |
| Primitive | 13.605 | 23.446 | | 24.484 |
| [3*s*2*p*] | 13.604 | 23.469 | | 24.502 |
| [2*s*1*p*] | 13.579 | 23.583 | | 24.551 |
| [4*s*3*p*][a] | 13.605 | 23.448 | | 24.494 |
| [4*s*3*p*2*d*][a] | | | | 24.550 |
| Exp.[b] | 13.606 | | | 24.580 |

[a] Using basis sets from Ref. [2]
[b] Experimental results from [13]. For H, the infinite mass eigenvalue is used

**Table 2.** The ionization potential (eV) of the atoms Li–Ne

| Basis | Li | Be | B | C | N | O | F | Ne |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | 5.342 | 8.046 | 7.936 | 10.795 | 13.972 | 11.961 | 15.734 | 19.856 |
| [4*s*3*p*2*d*] | 5.342 | 8.045 | 7.932 | 10.794 | 13.980 | 11.974 | 15.748 | 20.050 |
| [3*s*2*p*1*d*] | 5.340 | 8.043 | 7.913 | 10.767 | 13.946 | 11.903 | 15.782 | 20.585 |
| [5*s*4*p*3*d*][a] | 5.342 | 8.045 | 7.932 | 10.792 | 13.968 | 11.953 | 15.716 | 19.898 |
| MCPF results | | | | | | | | |
| Primitive | | 9.273 | 8.166 | 11.118 | 14.383 | 13.146 | 17.051 | 21.297 |
| [4*s*3*p*2*d*] | | 9.270 | 8.124 | 11.131 | 14.381 | 13.124 | 17.038 | 21.469 |
| [3*s*2*p*1*d*] | | 9.261 | 8.101 | 11.061 | 14.237 | 12.912 | 17.033 | 21.768 |
| [5*s*4*p*3*d*][a] | | 9.292 | 8.162 | 11.113 | 14.378 | 13.137 | 17.031 | 21.347 |
| [5*s*4*p*3*d*2*f*][a] | | 9.295 | 8.203 | 11.189 | 14.485 | 13.388 | 17.203 | 21.443 |
| Exp.[b] | 5.390 | 9.320 | 8.296 | 11.264 | 14.534 | 13.614 | 17.42 | 21.56 |

[a] Using basis sets from Ref. [2]
[b] Experimental results from [13]

mainly due to the lack of core correlation of the 3*s* and 3*p* shells. Introducing core correlation reduces the error to 0.16 eV, a most satisfactory result.

For all atoms up to Ar, MCPF calculations have also been performed using the larger ANO set of [2, 3], both with and without *f*-polarization functions. The results indicate that the main part of the error on the ionization potential for the atoms on the right-hand side of the periodic table is due to the lack of higher momentum correlating functions. Looking at the halogen atoms for example, we notice that the addition of *f*-functions reduces the error by 0.17 eV for fluorine and by 0.22 eV for chlorine. A similar effect may be expected for bromine, where it might even be necessary to include *g*-functions in order to obtain quantitatively correct results.

Two other factors contributing to the generally deteriorating results with an increasing atomic number are the lack of core correlation (as indicated for example by the results on scandium) and the absence of relativistic corrections. The latter

**Table 3.** The ionization potential (eV) of the atoms Na–Ar

| Basis | Na | Mg | Al | Si | P | S | Cl | Ar |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | 4.949 | 6.610 | 5.500 | 7.656 | 10.044 | 9.062 | 11.801 | 14.774 |
| [5*s*4*p*3*d*] | 4.949 | 6.609 | 5.498 | 7.656 | 10.050 | 9.064 | 11.805 | 14.850 |
| [4*s*3*p*2*d*] | 4.949 | 6.607 | 5.476 | 7.641 | 10.023 | 9.050 | 11.831 | 15.016 |
| [6*s*5*p*4*d*][a] | 4.951 | 6.608 | 5.499 | 7.654 | 10.043 | 9.057 | 11.792 | 14.779 |
| MCPF results | | | | | | | | |
| Primitive | | 7.519 | 5.908 | 8.034 | 10.338 | 9.765 | 12.511 | 15.466 |
| [5*s*4*p*3*d*] | | 7.518 | 5.901 | 8.028 | 10.336 | 9.753 | 12.505 | 15.547 |
| [4*s*3*p*2*d*] | | 7.517 | 5.907 | 8.021 | 10.305 | 9.718 | 12.534 | 15.661 |
| [6*s*5*p*4*d*][a] | | 7.527 | 5.908 | 8.032 | 10.337 | 9.761 | 12.503 | 15.473 |
| [6*s*5*p*4*d*3*f*][a] | | 7.531 | 5.944 | 8.126 | 10.488 | 10.084 | 12.725 | 15.580 |
| Exp.[b] | 5.14 | 7.64 | 5.98 | 8.15 | 10.49 | 10.36 | 13.01 | 15.755 |

[a] Using basis set from Ref. [3]
[b] Experimental results from [13]

**Table 4.** The ionization potential (eV) of the atoms K–Ca, Ga–Kr

| Basis | K | Ca | Ga | Ge | As | Se | Br | Kr |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | 3.996 | 5.119 | 5.479 | 7.436 | 9.531 | 8.434 | 10.775 | 13.257 |
| [6*s*5*p*4*d*] | 3.996 | 5.119 | 5.477 | 7.435 | 9.533 | 8.433 | 10.777 | 13.314 |
| [5*s*4*p*3*d*] | 3.996 | 5.117 | 5.460 | 7.428 | 9.520 | 8.431 | 10.801 | 13.427 |
| MCPF results | | | | | | | | |
| Primitive | | 5.893 | 5.806 | 7.744 | 9.777 | 9.036 | 11.363 | 13.812 |
| [6*s*5*p*4*d*] | | 5.893 | 5.798 | 7.738 | 9.775 | 9.025 | 11.357 | 13.864 |
| [5*s*4*p*3*d*] | | 5.891 | 5.801 | 7.741 | 9.758 | 9.006 | 11.384 | 13.957 |
| Exp.[a] | 4.341 | 6.113 | 5.930 | 7.911 | 10.034 | 9.635 | 11.877 | 14.222 |

[a] Experimental results from [13]

**Table 5.** The ionization potential (eV) of the atoms Sc–Zn

| Basis | Sc | Ti | V | Cr | Mn | Fe | Co | Ni | Cu | Zn |
|---|---|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | | | |
| Primitive | 5.352 | 5.517 | 5.807 | 5.904 | 5.911 | 6.306 | 7.793 | 6.332 | 6.405 | 7.637 |
| [7*s*5*p*4*d*3*f*] | 5.352 | 5.517 | 5.806 | 5.903 | 5.912 | 6.309 | 7.790 | 6.315 | 6.398 | 7.637 |
| [6*s*4*p*3*d*2*f*] | 5.366 | 5.539 | 5.804 | 5.903 | 5.935 | 6.329 | 7.793 | 6.309 | 6.373 | 7.649 |
| MCPF results | | | | | | | | | | |
| Primitive | 5.809 | 6.532 | 6.378 | 6.414 | 7.083 | 7.528 | 7.513 | 7.194 | 7.291 | 8.869 |
| [7*s*5*p*4*d*3*f*] | 5.821 | 6.529 | 6.375 | 6.413 | 7.078 | 7.524 | 7.509 | 7.192 | 7.281 | 8.858 |
| [6*s*4*p*3*d*2*f*] | 5.858 | 6.539 | 6.357 | 6.412 | 7.085 | 7.531 | 7.486 | 7.175 | 7.203 | 8.823 |
| Exp.[a] | 6.56 | 6.84 | 6.73 | 6.76 | 7.43 | 7.90 | 7.85 | 7.62 | 7.72 | 9.39 |

[a] Experimental results from [13, 14]

are especially important for the heavier transition metals [15], and most probably also for the atoms Ga–Kr.

Overall, the results obtained with the present basis sets must be considered as satisfactory, considering the limited size of the primitive sets. With exception of the rare gases, for which the property was regarded as unimportant, the contraction errors for the ionization potential are small. With the largest contraction scheme used (cf. Tables 1–5) the errors obtained at the MCPF level are around 0.01 eV or less in almost all cases (exceptions are boron and oxygen). Deleting one more weakly occupied shell from the basis sets significantly deteriorates the results, but, except for N and O (with errors of 0.15 and 0.23 eV), the truncation errors are still less than 0.1 eV in all cases.

The electron affinity of atoms is inherently more difficult to compute. This has been demonstrated on several occasions in the literature, see for example the work of Feller et al. [17] on oxygen, where it was clearly demonstrated that the inclusion of high angular momentum functions is essential, as well as the use of extensive correlation methods. Furthermore, in order to get accurate results core correlation effects are of importance. A recent systematic study of the electron affinity for the atoms Al–Cl by Woon and Dunning [18] gives further evidence of the slow convergence of the electron affinities with respect to basis set and correlation treatment.

The results for the electron affinities obtained with the present basis sets are shown in Tables 6–10. Again, the tables include both SCF and MCPF (valence-only) results, and calculations performed with the larger sets from [2, 3] are included as a reference. These basis sets must be regarded to perform well, except for the alkali atoms for which this property was regarded as unimportant (*vide infra*). Using the MCPF method and the primitive sets, the largest errors within each $p^n$ series are obtained for $n = 4, 5, 6$, with a maximum for phosphor (0.45 eV) and arsenium (0.56 eV). For the transition metal atoms errors ranging between 0.42 eV for vanadium and 0.63 eV for iron are found. With a few exceptions, the contraction errors are of the same order of magnitude as for the ionization potential: around 0.01 eV or smaller for the largest contraction scheme, and less than 0.1 eV for the smallest contraction schemes used (Tables 6–10). Exceptionally large contraction errors are found for hydrogen (with the $2s$ $1p$ contraction scheme) and the alkali atoms. The errors are already present at the SCF level, and simply reflect the fact that the anion, characterized by a very diffuse $s$ valence orbital, was

**Table 6.** The electron affinity (eV) of the atoms H–He

| | SCF | | MCPF | |
|---|---|---|---|---|
| Basis | H | He | H | He |
| Primitive | − 0.339 | < 0 | 0.688 | < 0 |
| [3$s$2$p$] | − 0.343 | < 0 | 0.676 | < 0 |
| [2$s$1$p$] | − 0.742 | < 0 | − 0.036 | < 0 |
| [4$s$3$p$][a] | − 0.337 | < 0 | 0.699 | < 0 |
| [4$s$3$p$2$d$][a] | | | 0.720 | < 0 |
| Exp.[b] | 0.754 | < 0 | 0.754 | < 0 |

[a] Using basis set from [2]

[b] Experimental results from [16]

**Table 7.** The electron affinity (eV) of the atoms Li–Ne

| Basis | Li | Be | B | C | N | O | F | Ne |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | – 0.127 | < 0 | – 0.274 | 0.553 | < 0 | – 0.534 | 1.340 | < 0 |
| [4*s*3*p*2*d*] | – 0.354 | < 0 | – 0.271 | 0.557 | < 0 | – 0.526 | 1.345 | < 0 |
| [3*s*2*p*1*d*] | – 0.354 | < 0 | – 0.258 | 0.593 | < 0 | – 0.435 | 1.433 | < 0 |
| [5*s*4*p*3*d*][a] | – 0.126 | < 0 | – 0.274 | 0.550 | < 0 | – 0.534 | 1.335 | < 0 |
| MCPF results | | | | | | | | |
| Primitive | 0.612 | < 0 | 0.159 | 1.125 | < 0 | 1.071 | 3.086 | < 0 |
| [4*s*3*p*2*d*] | 0.483 | < 0 | 0.155 | 1.130 | < 0 | 1.064 | 3.072 | < 0 |
| [3*s*2*p*1*d*] | 0.474 | < 0 | 0.138 | 1.092 | < 0 | 0.975 | 2.925 | < 0 |
| [5*s*4*p*3*d*] | 0.616 | < 0 | 0.178 | 1.130 | < 0 | 1.097 | 3.101 | < 0 |
| [5*s*4*p*3*d*2*f*][a] | 0.616 | < 0 | 0.221 | 1.203 | < 0 | 1.207 | 3.158 | < 0 |
| Exp.[b] | 0.618 | < 0 | 0.277 | 1.263 | < 0 | 1.461 | 3.399 | < 0 |

[a] Using basis set from [2]
[b] Experimental results from [16]

**Table 8.** The electron affinity (eV) of the atoms Na–Ar

| Basis | Na | Mg | Al | Si | P | S | Cl | Ar |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | – 0.106 | < 0 | 0.041 | 0.961 | – 0.486 | 0.916 | 2.572 | < 0 |
| [5*s*4*p*3*d*] | – 0.296 | < 0 | 0.042 | 0.964 | – 0.488 | 0.919 | 2.572 | < 0 |
| [4*s*3*p*2*d*] | – 0.315 | < 0 | 0.044 | 0.964 | – 0.485 | 0.919 | 2.563 | < 0 |
| [6*s*5*p*4*d*] | – 0.115 | < 0 | 0.037 | 0.959 | – 0.492 | 0.911 | 2.567 | < 0 |
| MCPF results | | | | | | | | |
| Primitive | 0.464 | < 0 | 0.366 | 1.280 | 0.301 | 1.715 | 3.365 | < 0 |
| [5*s*4*p*3*d*] | 0.369 | < 0 | 0.366 | 1.281 | 0.293 | 1.716 | 3.364 | < 0 |
| [4*s*3*p*2*d*] | 0.358 | < 0 | 0.383 | 1.284 | 0.273 | 1.689 | 3.315 | < 0 |
| [6*s*5*p*4*d*][a] | 0.539 | < 0 | 0.359 | 1.276 | 0.292 | 1.709 | 3.358 | < 0 |
| [6*s*5*p*4*d*3*f*][a] | 0.540 | < 0 | 0.413 | 1.378 | 0.512 | 1.864 | 3.427 | < 0 |
| Exp.[b] | 0.548 | < 0 | 0.441 | 1.385 | 0.747 | 2.077 | 3.617 | < 0 |

[a] Using basis set from [3]
[b] Experimental results from [16]

**Table 9.** The electron affinity (eV) of the atoms Ki–Ca, Ga–Kr

| Basis | K | Ca | Ga | Ge | As | Se | Br | Kr |
|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | |
| Primitive | – 0.102 | < 0 | – 0.048 | 0.946 | – 0.388 | 1.005 | 2.569 | < 0 |
| [6*s*5*p*4*d*] | – 0.299 | < 0 | – 0.046 | 0.949 | – 0.389 | 1.008 | 2.567 | < 0 |
| [5*s*4*p*3*d*] | – 0.310 | < 0 | – 0.060 | 0.931 | – 0.410 | 0.980 | 2.527 | < 0 |
| MCPF results | | | | | | | | |
| Primitive | 0.480 | < 0 | 0.206 | 1.200 | 0.258 | 1.638 | 3.172 | < 0 |
| [6*s*5*p*4*d*] | 0.372 | < 0 | 0.205 | 1.200 | 0.251 | 1.637 | 3.170 | < 0 |
| [5*s*4*p*3*dp*] | 0.366 | < 0 | 0.215 | 1.198 | 0.226 | 1.609 | 3.125 | < 0 |
| Exp.[a] | 0.501 | < 0 | 0.30(2) | 1.2(2) | 0.81(3) | 2.021 | 3.365 | < 0 |

[a] Experimental results from [16]

**Table 10.** The electron affinity (eV) of the atoms Sc–Zn

| Basis | Sc | Ti | V | Cr | Mn | Fe | Co | Ni | Cu | Zn |
|---|---|---|---|---|---|---|---|---|---|---|
| SCF results | | | | | | | | | | |
| Primitive | $< 0$ | $-0.931$ | $-0.591$ | $-0.549$ | $< 0$ | $-2.127$ | $-1.754$ | $-0.096$ | $-0.018$ | $< 0$ |
| [7*s*5*p*4*d*3*f*] | $< 0$ | $-0.931$ | $-0.591$ | $-0.542$ | $< 0$ | $-2.125$ | $-1.741$ | $-0.078$ | $-0.008$ | $< 0$ |
| [6*s*4*p*3*d*2*f*] | $< 0$ | $-0.927$ | $-0.583$ | $-0.548$ | $< 0$ | $-2.110$ | $-1.729$ | $-0.056$ | 0.027 | $< 0$ |
| MCPF results | | | | | | | | | | |
| Primitive | $< 0$ | $-0.420$ | 0.108 | 0.219 | $< 0$ | $-0.463$ | 0.115 | 0.649 | 0.701 | $< 0$ |
| [7*s*5*p*4*d*3*f*] | $< 0$ | $-0.422$ | 0.107 | 0.216 | $< 0$ | $-0.474$ | 0.108 | 0.639 | 0.647 | $< 0$ |
| [6*s*4*p*3*d*2*f*] | $< 0$ | $-0.413$ | 0.116 | 0.197 | $< 0$ | $-0.511$ | 0.080 | 0.588 | 0.575 | $< 0$ |
| Exp.[a] | $< 0$ | 0.079 | 0.525 | 0.666 | $< 0$ | 0.163 | 0.661 | 1.156 | 1.228 | $< 0$ |

[a] Experimental results from [16]

not included (or with a very small weight for H) in the density matrix averaging used to generate the basis set for these atoms. A proper description of the alkali anions was however considered as less important, since these atoms are actually never found with a negative charge in any bonding situation.

### *4.2 The molecules $H_2$, CO and $P_2$*

For $H_2$ calculations were performed at the CASSCF level with $\sigma$ and $\sigma^*$ in the active space, and at the CI level. The results are presented in Table 11. At the CASSCF level the results are compared to the results obtained with a [4*s* 3*p* 2*d*] basis set taken from [2]. Compared to this large basis set, the present set gives the following deviation:

- The primitive set: $\Delta r_e = 0.0001$ Å, $\Delta D_0 = 0.0027$ eV, and $\Delta\omega_e = 2.84\ \mathrm{cm}^{-1}$.
- The [4*s*3*p*] contraction: $\Delta r_e = 0.0000$ Å, $\Delta D_0 = 0.0031$ eV, and $\Delta\omega_e = 6.26\ \mathrm{cm}^{-1}$.
- The [3*s*2*p*] contraction: $\Delta r_e = 0.0005$ Å, $\Delta D_0 = 0.0111$ eV, and $\Delta\omega_e = 3.19\ \mathrm{cm}^{-1}$.
- The [2*s* 1*p*] contraction: $\Delta r_e = 0.0091$ Å, $\Delta D_0 = 0.0091$ eV, and $\Delta\omega_e = 53.13\ \mathrm{cm}^{-1}$.

$r_e$ is virtually identical to the reference basis set except possibly for the [2*s* 1*p*] contraction where the error is close to 0.01 Å. $D_0$ is also close to the reference basis set with a small degradation showing up at the [3*s*2*p*] contraction, and increasing at the [2*s* 1*p*] contraction, which yields an error of opposite sign. The errors for $\omega_e$ are within 7 $\mathrm{cm}^{-1}$ except for the [2*s* 1*p*] contraction where it is 53 $\mathrm{cm}^{-1}$. These results are most satisfactory. At the CI level the results are compared to experimental data. The following deviations are obtained for the primitive set: $\Delta r_e = -0.0001$ Å, $\Delta D_0 = 0.0730$ eV, and $\Delta\omega_e = 8.87\ \mathrm{cm}^{-1}$. The corresponding errors obtained with the reference basis set are: $\Delta r_e = 0.0007$ Å, $\Delta D_0 = 0.0253$ eV, and $\Delta\omega_e = 0.77\ \mathrm{cm}^{-1}$. The accuracy of $r_e$ must be considered fortuitous, while the errors of $D_0$ and $\omega_e$ are what might be expected for a set without *d*-functions.

**Table 11.** Effect of basis set contraction on some ground state properties of the $H_2$ molecule

| Basis set | $r_e$(Å) | $D_e$(eV) | $D_0$(eV) | $\omega_e(\mathrm{cm}^{-1})$ |
|---|---|---|---|---|
| CASSCF results | | | | |
| Primitive | 0.7548 | 4.1399 | 3.8809 | 4244.59 |
| [4*s*3*p*] | 0.7547 | 4.1397 | 3.8805 | 4248.01 |
| [3*s*2*p*] | 0.7552 | 4.1314 | 3.8725 | 4244.94 |
| [2*s*1*p*] | 0.7638 | 4.1543 | 3.8927 | 4294.88 |
| [4*s*3*p*2*d*][a] | 0.7547 | 4.1425 | 3.8836 | 4241.75 |
| CI results | | | | |
| Primitive | 0.7413 | 4.6754 | 4.4051 | 4410.08 |
| [4*s*3*p*] | 0.7416 | 4.6705 | 4.3994 | 4421.31 |
| [3*s*2*p*] | 0.7423 | 4.6439 | 4.3726 | 4427.73 |
| [2*s*1*p*] | 0.7578 | 4.5707 | 4.3012 | 4405.94 |
| [4*s*3*p*2*d*][a] | 0.7421 | 4.7225 | 4.4528 | 4401.98 |
| Exp. results | 0.7414 | | 4.4781 | 4401.21 |

[a] With the basis sets from [2]

For the contraction [4*s*3*p*] the errors are: $\Delta r_e = 0.0002$ Å, $\Delta D_0 = 0.0734$ eV, and $\Delta\omega_e = 20.10\ \mathrm{cm}^{-1}$, while for the [3*s*2*p*] contraction: $\Delta r_e = 0.0009$ Å, $\Delta D_0 = 0.1055$ eV, and $\Delta\omega_e = 26.52\ \mathrm{cm}^{-1}$, and for the [2*s*1*p*] contraction: $\Delta r_e = 0.0164$ Å, $\Delta D_0 = 0.1769$ eV, and $\Delta\omega_e = 4.73\ \mathrm{cm}^{-1}$. The contraction errors for $r_e$ are small for the [4*s*3*p*] and [3*s*2*p*] contractions with, again, a substantial effect for the [2*s*1*p*] contraction, quite expected for such a small basis set. $\omega_e$ is more sensitive to the contraction, but the error is still only about 20 $\mathrm{cm}^{-1}$ for the [4*s*3*p*] and [3*s*2*p*] contractions, while it drops to 5 $\mathrm{cm}^{-1}$ for the [2*s*1*p*] contraction, which must be regarded as a fortuitous result.

Table 12 shows the results of the test calculations on some ground state properties of the CO molecule. Calculations were performed at the SCF level, and at the correlated level using the MCPF approach and correlating the ten valence electrons. The results obtained with the present basis sets are compared to analogous results obtained with the larger primitive sets from [2].

The SCF results are further confronted with results from numerical Hartree–Fock calculations [19], while for the MCPF calculations the results obtained by Barnes et al. [20], using a large basis set ((16*s*8*p*6*d*4*f*2*g*1*h*) contracted to [6*s*5*p*4*d*3*f*2*g*1*h*]) serve as a reference.

The present basis sets yield accurate results for $r_e$ and $\omega_e$. The SCF results obtained with the primitive set are close to the HF-limit (with errors of 0.002 Å and 8 $\mathrm{cm}^{-1}$ respectively), while the MCPF results are close to the results obtained with a much larger primitive set [20] (with deviations of 0.004 Å and 28 $\mathrm{cm}^{-1}$). Contraction to [4*s*3*p*2*d*] set still yields satisfactory results, with an almost negligible

**Table 12.** Effect of basis set contraction on some ground state properties of the CO molecule

| | $R_e$(Å) | $D_e$(eV) | $\omega_e(\mathrm{cm}^{-1})$ | $\mu$(au) |
|---|---|---|---|---|
| SCF results | | | | |
| Uncontracted | 1.103 | 7.847 | 2423 | − 0.055 |
| [4*s*3*p*2*d*] | 1.105 | 7.778 | 2427 | − 0.052 |
| [3*s*2*p*1*d*] | 1.113 | 7.607 | 2437 | − 0.071 |
| [4*s*3*p*2*d*][a] | 1.107 | 7.713 | 2422 | − 0.054 |
| [5*s*4*p*3*d*2*f*][a] | 1.103 | 7.880 | 2424 | − 0.054 |
| HF-limit[b] | 1.101 | | 2431 | − 0.110 |
| MCPF results | | | | |
| Uncontracted | 1.130 | 10.452 | 2175 | 0.114 |
| [4*s*3*p*2*d*] | 1.134 | 10.342 | 2168 | 0.116 |
| [3*s*2*p*1*d*] | 1.148 | 10.171 | 2146 | 0.107 |
| [4*s*3*p*2*d*][a] | 1.138 | 10.277 | 2159 | 0.110 |
| [5*s*4*p*3*d*2*f*][a] | 1.129 | 10.663 | 2185 | 0.109 |
| Barnes[c] | 1.126 | | 2203 | 0.107<br>(0.033)[d] |
| Exp.[e] | 1.128 | 11.09 | 2170 | 0.044 |

[a] With the basis sets from [2]

[b] From [19]

[c] From [20]. Basis set (16*s*8*p*6*d*4*f*2*g*1*h*)/[6*s*5*p*4*d*3*f*2*g*1*h*]

[d] Computed using a finite field (energy derivative) approach

[e] Experimental results from [21]

truncation error both for $r_e$ and $\omega_e$. The truncation errors for [$3s2p1d$] are significantly larger, especially for $r_e$, with errors of 0.010 and 0.018 Å at the SCF and MCPF level, respectively. It is worth noting that the [$4s3p2d$] contracted set is actually performing slightly better than the same contracted set obtained from a larger primitive set [2].

The binding energy is 0.64 eV too low at the MCPF level using the primitive set. This may seem discouraging at first but, as indicated by the results obtained with the [$5s4p3d2f$] sets, higher angular momentum functions are needed in order to obtain significantly improved results. Basis set contraction further increases the error, to 0.75 eV for [$4s3p2d$], and to 0.92 eV for [$3s2p1d$]. The major contraction error is already present at the SCF level: The decrease in the SCF binding energy is 0.07 and 0.24 eV for the [$4s3p2d$] and [$3s2p1d$] contractions, respectively, while the differential dynamic correlation contribution is only little affected, being reduced by 0.04 eV for both contractions. But also in this case, the results obtained with the [$4s3p2d$] set are superior to the ones obtained with the [$4s3p2d$] set originating from a large primitive set [2]. Obviously there is no need to start with a large primitive set for the calculation of the properties under consideration, when only a limited number of contracted functions is actually used in the calculation.

Not surprisingly, the results obtained for the dipole moment are less accurate than those for $r_e$, $D_e$ and $\omega_e$. This could have been expected *a priori*, since the basis sets were not designated to yield good values for such properties. It may be noted however that the present basis sets do not appear to be significantly less accurate than the larger sets of [2]. Overall, large discrepancies are obtained between the MCPF results and the experimental value of the dipole moment. However, the results cited in Table 12 are expectation values. It has been shown in [20] that the MCPF density matrix is unreliable for the computation of properties such as the dipole moment via an expectation value. Much more reliable results can be obtained with an energy derivative approach. The MCPF dipole moment from Barnes et al. [20], computed as an energy derivative, has been included in Table 12 for comparison.

To accurately determine the bonding properties of the $P_2$ molecule is not trivial. A recent study on the phosphorus dimer [22] has indicated that an extensive correlation treatment including core polarization effects is necessary in order to obtain quantitatively correct results. In this work, test calculations on $P_2$ have been performed using the CASSCF/MRCI approach. CASSCF calculations were performed with an active space consisting of the $3s$ and $3p$ orbitals, while the Ne cores were kept inactive. The full CASSCF active space was used as the reference space for subsequent MRCI calculations. No excitations from the core orbitals were included in the MRCI treatment, so that we cannot expect to obtain quantitative accuracy. Rather than comparing our results directly to experiment, we will therefore use the [$5s4p3d2f$] basis set of [2] as a reference. All results are shown in Table 13.

Overall, satisfactory results are obtained with the present basis sets, both for the bond distance and stretching frequency. The distances obtained with the primitive set are 0.001, 0.004 and 0.004 Å longer than our reference basis set, at the CASSCF, MRCI and MRCI + Q level, respectively. There is a definite degradation of the quality by contraction, but the truncation errors are within tolerable limits: not larger than 0.005 Å for [$5s4p3d$] and than 0.015 Å for [$4s3p2d$], at all considered levels of approximation. The results for $\omega_e$ are all within a 8 cm$^{-1}$ discrepancy range at the CASSCF level and a 23 cm$^{-1}$ for the MRCI and MRCI + Q treatments.

**Table 13.** Effect of basis set contraction on some ground state properties of the $P_2$ molecule

| | $R_e$(Å) | $D_e$(eV) | $\omega_e$(cm$^{-1}$) |
|---|---|---|---|
| CASSCF results | | | |
| Uncontracted | 1.926 | 4.084 | 747 |
| [5*s*4*p*3*d*] | 1.928 | 4.071 | 748 |
| [4*s*3*p*2*d*] | 1.932 | 4.074 | 751 |
| [5*s*4*p*3*d*2*f*][a] | 1.925 | 4.14 | 755 |
| MRCI results | | | |
| Uncontracted | 1.922 | 4.234 | 744 |
| [5*s*4*p*3*d*] | 1.927 | 4.150 | 740 |
| [4*s*3*p*2*d*] | 1.936 | 4.107 | 740 |
| [5*s*4*p*3*d*2*f*][a] | 1.918 | 4.66 | 762 |
| MRCI + Q results | | | |
| Uncontracted | 1.925 | 4.179 | 737 |
| [5*s*4*p*3*d*] | 1.930 | 4.094 | 733 |
| [4*s*3*p*2*d*] | 1.940 | 4.029 | 732 |
| [5*s*4*p*3*d*2*f*][a] | 1.921 | 4.71 | 755 |
| Exp.[b] | 1.893 | 5.08 | 781 |

[a] With the basis set from [3]

[b] Experimental results from [21]

The results for the binding energies are also acceptable, considering the small size of the present basis sets. At the CASSCF level, no substantial deterioration of the result with respect to the larger basis set [3] is obtained. With the primitive basis set, the binding energy is only 0.06 eV smaller, and the truncation errors are almost negligibly small. The lack of higher angular momentum functions is however felt much more strongly at the MRCI and MRCI + Q level, resulting in discrepancies of 0.43 eV (MRCI) and 0.53 eV (MRCI + Q) with respect to the basis set containing *f*-functions, if the primitive sets are considered. The truncation errors are also larger, up to 0.15 eV at the MRCI + Q level using a [4*s*3*p*2*d*] set. It is clear that errors of this size in the binding energy cannot be avoided when using limited basis sets, simply due to the lack of correlating functions. The present test has indicated however that reliable bond lengths and vibrational frequencies can still be obtained.

### *4.3 The halogen dimers $F_2$, $Cl_2$ and $Br_2$*

The behaviour of the present basis sets was considered in some detail for the series of halogen dimer molecules $F_2$, $Cl_2$, $Br_2$. Table 14 shows the results of two different types of correlation treatment. A first set of calculations consists of an MCPF treatment based on a RHF wave function, while in the second set the orbitals were optimized at the CASSCF level and dynamical correlation was added using multiconfigurational second-order perturbation theory (CASPT2) [23]. Only the *s*, *p* valence shell electrons were included in the dynamical correlation treatment, MCPF or CASPT2. The CASSCF calculations were performed with an active space consisting of the 2*p*, 3*p*, 4*p* orbitals for $F_2$, $Cl_2$ and $Br_2$, respectively, while the

**Table 14.** Effect of basis set contraction on some ground state properties of the molecules $F_2$, $Cl_2$ and $Br_2$

| | $F_2$ | | | | $Cl_2$ | | | | $Br_2$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Basis | $R_e$(Å) | $D_0$(eV)[a] | $\omega_e$(cm$^{-1}$) | Basis | $R_e$(Å) | $D_0$(eV)[a] | $\omega_e$(cm$^{-1}$) | Basis | $R_e$(Å) | $D_0$(eV)[a] | $\omega_e$(cm$^{-1}$) |
| SCF results | | | | | | | | | | | |
| Uncontracted | 1.330 | − 1.232 | 1272 | Uncontracted | 1.996 | 0.935 | 605 | Uncontracted | 2.295 | 0.804 | 347 |
| [*4s3p2d*] | 1.334 | − 1.288 | 1265 | [*5s4p3d*] | 1.997 | 0.904 | 602 | [*6s5p4d*] | 2.296 | 0.787 | 348 |
| [*3s2p1d*] | 1.341 | − 1.570 | 1211 | [*4s3p2d*] | 2.001 | 0.812 | 591 | [*5s4p3d*] | 2.299 | 0.708 | 344 |
| [*4s3p2d*][b] | 1.337 | − 1.311 | 1252 | [*5s4p3d*][b] | 1.992 | 0.944 | 608 | | | | |
| [*5s4p3d2f*][b] | 1.328 | − 1.142 | 1265 | [*6s5p4d3f*][b] | 1.977 | 1.160 | 616 | [*6s5p4d3f*] | 2.275 | 0.999 | 355 |
| CASSCF results | | | | | | | | | | | |
| Uncontracted | 1.467 | 0.716 | 704 | Uncontracted | 2.062 | 1.564 | 491 | Uncontracted | 2.364 | 1.385 | 285 |
| [*4s3p2d*] | 1.475 | 0.699 | 692 | [*5s4p3d*] | 2.063 | 1.536 | 487 | [*6s5p4d*] | 2.365 | 1.369 | 285 |
| [*3s2p1d*] | 1.505 | 0.494 | 602 | [*4s3p2d*] | 2.071 | 1.462 | 474 | [*5s4p3d*] | 2.371 | 1.306 | 280 |
| [*4s3p2d*][b] | 1.481 | 0.646 | 685 | [*5s4p3d*][b] | 2.057 | 1.568 | 493 | | | | |
| [*5s4p3d2f*][b] | 1.465 | 0.716 | 709 | [*6s5p4d3f*][b] | 2.038 | 1.698 | 506 | [*6s5p4d3f*] | 2.340 | 1.493 | 294 |
| CASPT2 results | | | | | | | | | | | |
| Uncontracted | 1.434 | 1.414 | 856 | Uncontracted | 2.045 | 1.962 | 507 | Uncontracted | 2.343 | 1.761 | 297 |
| [*4s3p2d*] | 1.441 | 1.396 | 845 | [*5s4p3d*] | 2.047 | 1.927 | 505 | [*6s5p4d*] | 2.347 | 1.728 | 297 |
| [*3s2p1d*] | 1.470 | 1.193 | 771 | [*4s3p2d*] | 2.061 | 1.836 | 487 | [*5s4p3d*] | 2.356 | 1.660 | 292 |
| [*4s3p2d*][b] | 1.448 | 1.355 | 830 | [*5s4p3d*][b] | 2.042 | 1.964 | 510 | | | | |
| [*5s4p3d2f*][b] | 1.426 | 1.500 | 872 | [*6s5p4d3f*][b] | 2.006 | 2.240 | 540 | [*6s5p4d3f*] | 2.309 | 1.990 | 311 |
| MCPF results | | | | | | | | | | | |
| Uncontracted | 1.423 | 1.266 | 894 | Uncontracted | 2.042 | 1.966 | 520 | Uncontracted | 2.340 | 1.756 | 299 |
| [*4s3p2d*] | 1.431 | 1.258 | 880 | [*5s4p3d*] | 2.044 | 1.930 | 517 | [*6s5p4d*] | 2.345 | 1.721 | 303 |
| [*3s2p1d*] | 1.458 | 1.093 | 799 | [*4s3p2d*] | 2.058 | 1.845 | 500 | [*5s4p3d*] | 2.355 | 1.657 | 298 |
| [*4s3p2d*][b] | 1.438 | 1.221 | 865 | [*5s4p3d*][b] | 2.040 | 1.969 | 523 | | | | |
| [*5s4p3d2f*][b] | 1.415 | 1.356 | 916 | [*6s5p4d3f*][b] | 2.005 | 2.217 | 551 | [*6s5p4d3f*] | 2.309 | 1.952 | 316 |
| Exp.[c] | 1.417 | 1.600 | 891 | | 1.988 | 2.475 | 565 | | 2.283 | 1.971 | 323.2 |

[a] The SCF and MCPF values of the atomization energy were calculated as the molecular energy minus twice the atomic energy. A zero-point energy correction was added in all cases except SCF for $F_2$

[b] Using the basis sets from [2, 3]

[c] Experimental results from [21]

corresponding He, Ne and Ar cores were kept frozen using RHF orbitals. For $F_2$ and $Cl_2$, the quality of the present basis sets can again be tested by confronting the results to analogous results obtained with the larger primitive sets from [2, 3]. For $Br_2$ no such sets are available. The effect of adding higher angular momentum functions to the basis set was examined in this case by adding three *f*-functions to the [6*s*5*p*4*d*] contracted set. The exponents for the *f*-functions (0.658, 0.219, 0.073) were chosen such as to obtain maximum overlap with the 4*p*-orbitals.

The results in Table 14 illustrate the inability of the RHF method in describing the dissociation of the closed-shell dimers into open-shell halogen atoms. The situation is most severe for $F_2$, for which dissociation along the asymptotic potential curve results in a total energy which is 9.17 eV too high (with the primitive basis set). The erratic behaviour is even clearly reflected at the $F_2$ equilibrium geometry, for which a much too short bond distance and a much too large stretching frequency are obtained. The atomization energies in Table 14 were calculated as the molecular energy minus twice the atomic energy. The situation is now, of course, completely reversed for $F_2$, being unbound by 1.23 eV at the SCF level. For $Cl_2$ and $Br_2$ the errors are less dramatic, although also in this case the SCF atomization energy is too low by more than 1 eV.

A considerable improvement of the results is obtained at the CASSCF level. Dissociation is now described correctly, and all three molecules are bound, although much too weakly, with remaining errors of 0.88, 0.91 and 0.59 eV for $F_2$, $Cl_2$ and $Br_2$, respectively (with the primitive basis sets). The weakness of the bonding is also reflected in a too large bond distance and a too low stretching frequency. The errors are obviously connected to the method used, and cannot be reduced substantially by increasing the size of the basis set. More accurate results are only obtained at the fully correlated level, either MCPF or CASPT2. Actually, both levels of theory give similar results for $Cl_2$ and $Br_2$. For $F_2$ there are some differences, related to the different nature of the RHF versus CASSCF reference wave function for the different approaches. With the primitive basis set, the bond distance is 0.011 Å shorter and the stretching frequency 38 $cm^{-1}$ larger at the MCPF than at the CASPT2 level. CASPT2 performs better for the binding energy, with a result that is 0.15 eV closer to experiment than the MCPF result (calculated as the molecular minus twice the atomic energy).

On the whole, the present basis sets perform well at the correlated level for the three calculated properties. As was the case for CO and $P_2$, the results obtained around equilibrium geometry are satisfactory. With the primitive sets, the binding distances are accurate to within 0.06 Å and the stretching frequencies to within 60 $cm^{-1}$ in all cases. The largest error on the atomization energy is obtained for $Cl_2$, with an error of 0.51 eV, both at the MCPF and CASPT2 level. Truncation errors are only minimal for the largest contraction schemes ([4*s*3*p*2*d*] for $F_2$, [5*s*4*p*3*d*] for $Cl_2$ and [6*s*5*p*4*d*] for $Br_2$). Deleting an additional weakly occupied shell from the ANO basis however significantly deteriorates the results. The most severe truncation errors are obtained for the [3*s*2*p*1*d*] set in $F_2$: 0.04 Å on $r_e$, up to 0.22 eV on $D_e$ (CASPT2 result) and 95 $cm^{-1}$ on $\omega_e$ (MCPF result). For $F_2$, the [4*s*3*p*2*d*] set obtained from the larger primitive set from [2] leads to results that are slightly worse than the results obtained with the present [4*s*3*p*2*d*] set. For $Cl_2$ on the other hand, the larger primitive set from [3] is clearly superior: contracted to [5*s*4*p*3*d*], it produces results that are even slightly more accurate than the results obtained with the present primitive set.

However, as could be expected, a substantial improvement of the results for all three properties can only be obtained at the expense of adding higher angular momentum functions to the primitive sets. The final MCPF and CASPT2 results, obtained with the largest basis sets, are satisfactory, with a remaining error of less than 0.025 Å on $r_e$, less than 0.25 eV on $D_e$ and less than 25 $\mathrm{cm}^{-1}$ on $\omega_e$ in all cases. It is especially gratifying to see the results for the [6*s*5*p*4*d*3*f*] set in $Br_2$. The accuracy obtained with this set is similar as for the [6*s*5*p*4*d*3*f*] set in $Cl_2$, contracted from a much larger primitive set [3]. This indicates that the present sets can also be used in calculations requiring a high accuracy, providing that an appropriate number of higher angular momentum polarization functions is added. The CASPT2 result for the binding energy in $Br_2$ is even slightly too high. This is somewhat surprising: in a series of test calculations of the CASPT2 method on molecules built from first-row atoms [24] it has been shown that the method systematically tends to underestimate the value of atomization energies by 0.1–0.25 eV per number of electron pairs formed. This conclusion is also confirmed by our results for $F_2$ and $Cl_2$. However, we do not yet have enough experience to tell whether the same systematic error will prevail for systems with heavier atoms. Some recent experience on transition metal systems seems to indicate that the errors may be smaller, or even that CASPT2 may lead to a slight overestimate of the binding energy in these cases [26]. Another uncertainty lies in the relativistic correction to the binding energy, which has not been added for $Br_2$.

Finally, we have also considered the basis set superposition error (BSSE) on the bonding in the halogen dimers, obtained with the present basis sets. Normally, one would expect ANO basis sets to give only small errors due to BSSE [1]. Due to the relatively small size of the present sets, both primitive and contracted, one may expect the errors to become slightly larger. In order to investigate the effect of BSSE, we have performed calculations on one halogen atom with a set of a halogen ghost orbitals situated at the experimental bond distance in the dimer. The errors obtained at the MCPF and CASPT2 level are similar (with differences of less than 10 meV in all cases), and we report here only the CASPT2 results. For F, the computed errors for the [4*s*3*p*2*d*] and [3*s*2*p*1*d*] sets were 55 and 80 meV, respectively. For Cl, errors of 66 and 70 meV were obtained with the [5*s*4*p*3*d*] and [4*s*3*p*2*d*] sets, while for Br the errors were 54 and 58 meV with the [6*s*5*p*4*d*] and [5*s*4*p*3*d*] sets, respectively. These results must be regarded as satisfactory: the errors are small and of the same magnitude for the three molecules, indicating that the present sets constitute a well-balanced series of basis functions for the different rows. For F and Cl the errors are comparable to the errors obtained with the larger primitive sets [2, 3]: 63 meV for a [4*s*3*p*2*d*] set on F, 66 meV for the [5*s*4*p*3*d*] set on Cl. They can of course be further reduced by adding *f*-functions to the basis sets: errors of 25 and 24 meV were obtained with a [5*s*4*p*3*d*2*f*] and a [6*s*5*p*4*d*3*f*] set on F and Cl, respectively. For Br the error is reduced to 40 meV by adding three *f*-functions to the [6*s*5*p*4*d*] contracted set.

### *4.4 Valence excitation energies in Ti, V and Ni*

A main requirement for an ANO type transition metal basis set to work well is that it is flexible enough to produce a well-balanced description of electronic states belonging to the different configurations $d^{n-2}s^2$, $d^{n-1}s^1$ and $d^ns^0$. This is especially important for the bonding in transition metal compounds, which quite often is built from a mixture of several atomic states, belonging to a different 4*s*-occupation. One way to obtain the required flexibility would be to include all relevant atomic states in the density matrix averaging used to generate the basis set. A study along these lines has been published recently [12]. It was shown there that in some cases it may be necessary to uncontract the outermost primitive functions in order to obtain an equivalent treatment of valence states belonging to a different number of $3d$ occupied orbitals.

As a first test for the flexibility of the present transition metal basis sets, we consider the energy separation between the lowest-lying states belonging to $3d^{n-2}4s^2$, $3d^{n-1}4s^1$ and $3d^n$ in Ti ($^3$F, $^5$F, $^5$D), V ($^4$F, $^6$D, $^6$S) and Ni ($^3$F, $^3$D, $^1$S). Each state is individually optimized in a CASSCF calculation, and dynamical correlation is treated by MRCI, using the full CASSCF space as reference, and including only the valence $3d$, $4s$ electrons. For V and Ni, the CASSCF active space consists of the $3d$, $4s$ shells only. For Ti the $4p$ shell was also included in the active space, thus allowing us to compare our results with the results obtained by Bauschlicher and Taylor [12], using an alternative contraction scheme. The results are shown in Tables 15–17. The present treatment cannot be expected to yield results that are accurate enough to be verifiable directly against the experimental spectrum. This would require a more elaborate treatment, including, at least for Ni, the radial $3d$–$3d'$ correlation in the CASSCF space, and adding also $3s$,$3p$ core correlation. Also, no attempt was made to include relativistic corrections in the calculations. It was shown in Ref. [12] that much larger basis sets are needed to do so successfully. Instead, the relativistic contribution to the energy separation between the different states was taken from Ref. [15].

For Ti, we have included in Table 15 both the total energies and the energy separations. From the total energies one can see that the contraction procedure used introduces only minimal errors for the three states considered: less than 1 mH at the CASSCF level for the contraction schemes, and less than 1 mH for a $[7s5p4d3f]$ contraction scheme, around 2 mH for $[6s4p3d2f]$ and around 2–5 mH for $[5s4p3d2f]$ at the MRCI and MRCI + Q level. More important, with the exception of the $[5s4p3d2f]$ contracted set, the errors are remarkably constant for the different states, resulting in negligible truncation errors on the relative energies. A significant deterioration of the balance between the different occupations is only observed when less than six $s$-functions are included in the contracted set. A comparison of the total MRCI energies obtained with the $[6s4p3d2f]$ and the $[5s4p3d2f]$ set reveals that deleting the sixth $s$-function from the basis set significantly increases the contraction error for the $^3$F($3d^24s^2$) ground state, while leaving the $^5$D($3d^4$) energy almost invariant. This result is by no means surprising. Contracting the basis set to $[5s4p3d2f]$ only leaves one weakly occupied $s$ orbital to account for $4s$–$4s'$ radial correlation, which is absent for the $^5$D state, but becomes more important with each $3d \rightarrow 4s$ transition. Consequently, the results obtained with the $[5s4p3d2f]$ set are biased against the $^3$F ground state, and the $4s$–$3d$ excitation energies are calculated too low.

Very similar results are obtained for the spectrum of V and Ni. We have included in Tables 16 and 17 only the relative energies of the different states, but it is clear that the observations made for Ti remain valid in both cases. Again, truncation errors are almost negligible for a $[7s5p4d3f]$ and $[6s4p3d2f]$ contraction scheme, while the same bias against the $4s^2$ state is observed when the sixth $s$-function is deleted from the basis set.

Our results are in a marked contrast to the recent results obtained for Ti by Bauschlicher and Taylor, included in Table 15. The starting primitive set, $(21s16p10d6f)$, was much larger, and consequently, the total energies significantly lower than with the primitive set used here. Contraction was performed by averaging the one-electron density matrices of the relevant $^3$F, $^5$F and $^5$D states,

Table 15. Atomic calculations on titanium

| State | CASSCF energy (+ 848)(Ha) | MRCI energy (+ 848)(Ha) | MRCI + Q energy (+ 848)(Ha) | CASSCF separation (eV) | MRCI separation (eV) | MRCI + Q separation (eV) | + Rel.[a] correction (eV) | Exp.[b] (eV) |
|---|---|---|---|---|---|---|---|---|
| Uncontracted basis | | | | | | | | |
| $^3F(3d^24s^2)$ | − 0.434076 | − 0.457120 | − 0.457521 | | | | | |
| $^5F(3d^34s^1)$ | − 0.385006 | − 0.424786 | − 0.426068 | 1.34 | 0.88 | 0.86 | 1.00 | 0.81 |
| $^5D(3d^4)$ | − 0.253422 | − 0.320142 | − 0.324826 | 4.92 | 3.73 | 3.61 | 3.83 | 3.35 |
| Contraction [7*s*5*p*4*d*3*f*] | | | | | | | | |
| $^3F(3d^24s^2)$ | − 0.434047 | − 0.456714 | − 0.457110 | | | | | |
| $^5F(3d^34s^1)$ | − 0.384996 | − 0.424379 | − 0.425656 | 1.33 | 0.88 | 0.86 | 1.00 | 0.81 |
| $^5D(3d^4)$ | − 0.253400 | − 0.319707 | − 0.324340 | 4.92 | 3.73 | 3.61 | 3.83 | 3.35 |
| Contraction [6*s*4*p*3*d*2*f*] | | | | | | | | |
| $^3F(3d^24s^2)$ | − 0.433705 | − 0.454722 | − 0.455086 | | | | | |
| $^5F(3d^34s^1)$ | − 0.384741 | − 0.422722 | − 0.424058 | 1.33 | 0.87 | 0.84 | 0.98 | 0.81 |
| $^5D(3d^4)$ | − 0.253276 | − 0.318255 | − 0.322569 | 4.91 | 3.71 | 3.61 | 3.83 | 3.35 |
| Contraction [5*s*4*p*3*d*2*f*] | | | | | | | | |
| $^3F(3d^24s^2)$ | − 0.433523 | − 0.452050 | − 0.452330 | | | | | |
| $^5F(3d^34s^1)$ | − 0.384240 | − 0.421169 | − 0.422392 | 1.34 | 0.84 | 0.82 | 0.96 | 0.81 |
| $^5D(3d^4)$ | − 0.253070 | − 0.318073 | − 0.322395 | 4.91 | 3.62 | 3.54 | 3.76 | 3.35 |
| Uncontracted basis (21*s*16*p*10*d*6*f*][c] | | | | | | | | |
| $^3F(3d^24s^2)$ | | − 0.462123 | | | | | | |
| $^5F(3d^34s^1)$ | | − 0.427318 | | | 0.95 | | | 0.81 |
| $^5D(3d^4)$ | | − 0.323807 | | | 3.76 | | | 3.35 |
| (21*s*16*p*10*d*6*f*)/[7*s*6*p*4*d*2*f*] from the average of $^3F$, $^5F$ and $^5D$.[c] | | | | | | | | |
| $^3F(3d^24s^2)$ | | − 0.461550 | | | | | | |
| $^5F(3d^34s^1)$ | | − 0.426135 | | | 0.96 | | | 0.81 |
| $^5D(3d^4)$ | | − 0.316285 | | | 3.95 | | | 3.35 |

[a] Relativistic corrections taken from [15]

[b] Experimental results from [14]

[c] Results from [12]

**Table 16.** Atomic calculations on vanadium

| State | CASSCF separation (eV) | MRCI separation (eV) | MRCI + Q separation (eV) | + Rel.[a] correction (eV) | Exp.[b] (eV) |
|---|---|---|---|---|---|
| Uncontracted basis | | | | | |
| $^4F(3d^34s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^6D(3d^44s^1)$ | 0.21 | 0.18 | 0.24 | 0.41 | 0.25 |
| $^6S(3d^5)$ | 3.37 | 2.48 | 2.45 | 2.73 | 2.47 |
| Contraction [7*s*5*p*4*d*3*f*] | | | | | |
| $^4F(3d^34s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^6D(3d^44s^1)$ | 0.21 | 0.18 | 0.24 | 0.41 | 0.25 |
| $^6S(3d^5)$ | 3.37 | 2.48 | 2.44 | 2.72 | 2.47 |
| Contraction [6*s*4*p*3*d*2*f*] | | | | | |
| $^4F(3d^34s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^6D(3d^44s^1)$ | 0.21 | 0.17 | 0.22 | 0.39 | 0.25 |
| $^6S(3d^5)$ | 3.36 | 2.46 | 2.42 | 2.70 | 2.47 |
| Contraction [5*s*4*p*3*d*2*f*] | | | | | |
| $^4F(3d^34s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^6D(3d^44s^1)$ | 0.22 | 0.15 | 0.20 | 0.37 | 0.25 |
| $^6S(3d^5)$ | 3.36 | 2.39 | 2.34 | 2.62 | 2.47 |

[a] Relativistic correction taken from [15]

[b] Experimental results from [14]

**Table 17.** Atomic calculations on nickel

| State | CASSCF separation (eV) | MRCI separation (eV) | MRCI + Q separation (eV) | + Rel.[a] correction (eV) | Exp.[b] (eV) |
|---|---|---|---|---|---|
| Uncontracted basis | | | | | |
| $^3F(3d^84s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^3D(3d^94s^1)$ | 1.28 | 0.01 | – 0.04 | 0.32 | – 0.03 |
| $^1S(3d^{10})$ | 5.43 | 2.25 | 1.94 | 2.51 | 1.71 |
| Contraction [7*s*5*p*4*d*3*f*] | | | | | |
| $^3F(3d^84s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^3D(3d^94s^1)$ | 1.28 | 0.01 | – 0.04 | 0.32 | – 0.03 |
| $^1S(3d^{10})$ | 5.43 | 2.24 | 1.93 | 2.50 | 1.71 |
| Contraction [6*s*4*p*3*d*2*f*] | | | | | |
| $^3F(3d^84s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^3D(3d^94s^1)$ | 1.29 | – 0.01 | – 0.07 | 0.29 | – 0.03 |
| $^1S(3d^{10})$ | 5.44 | 2.21 | 1.91 | 2.48 | 1.71 |
| Contraction [5*s*4*p*3*d*2*f*] | | | | | |
| $^3F(3d^84s^2)$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| $^3D(3d^94s^1)$ | 1.36 | – 0.05 | – 0.11 | 0.25 | – 0.03 |
| $^1S(3d^{10})$ | 5.46 | 2.12 | 1.79 | 2.36 | 1.71 |

[a] Relativistic correction taken from [15]

[b] Experimental results from [14]

using an equal weight for the three states. This contraction procedure however resulted in a strong bias against the $^5$D state which could only be removed after uncontracting the outermost primitive functions. We notice indeed that the MRCI energy obtained for the $^5$D state with a [$7s6p4d2f$] contracted set is higher than any of our results, while, more comprehensible, the reverse situation holds for the other two states. The resulting contraction error for the $^5$D–$^3$F transition energy is almost 0.2 eV. The reason for the large discrepancy between the results obtained by Bauschlicher and our results is unclear at the moment. From the results obtained for Ti, V and Ni, we may safely conclude that the contraction scheme used in the present work is flexible enough to give a balanced description of the atomic spectra of the transition metal atoms, even if only the $4s^2$ state of the neutral atom is actually included in the density matrix averaging. Yet it is highly unlikely that our averaging procedure could be so obviously superior to a scheme in which the relevant states themselves are included in the averaging.

### *4.5 Sc$^+$ and the ScF molecule*

The ScF molecule represents a severe test to the flexibility of the present basis sets. The bonding in ScF is primarily ionic, and its charge distribution much more closely resembles the constituent ions Sc$^+$ and F$^-$ than the neutral atoms. In order to get an accurate value for the dissociation energy to the neutral atoms, one therefore needs basis sets that are flexible enough to give a well-balanced description both of the atoms and their ions. The problem can be partly overcome by calculating the dissociation energy in a step-wise manner, using the theoretical results for the dissociation to the ionic limits, and correcting to the neutral ground state atomic limits by using experimental values for the ionization energy of Sc and the electron affinity of F. Both approaches, dissociation to ground state atoms and to the ionic limits, will be compared here.

The formation of ScF out of Sc$^+$ and F$^-$ is attended by a change in ground state configuration: the ground state in ScF, $X^1\Sigma^+$, corresponds to the $^1$S($4s^2$) excited state in Sc$^+$, while the $^3$D($3d^14s^1$) ground state of Sc$^+$ ends up as a low-lying $a^3\Delta$ state in the molecule. We notice further that neither of the states $^3$D($3d^14s^1$) or $^1$S($4s^2$) have been included in the density matrix averaging used to generate the Sc basis set, where instead we used the $^3$F($d^2$) state. We will therefore start by testing the performance of the Sc basis set on the relative energies of the Sc $^2$D($d^1s^2$) ground state and Sc$^+$ $^3$D and $^1$S states.

A detailed *ab initio* study of the $X^1\Sigma^+$ and $a^3\Delta$ states in ScF has already been presented by Langhoff et al. [25], both at the SDCI and CPF level. Large flexible basis sets were used: for scandium, the ($14s9p5d$) set of Wachters was contracted to [$8s4p3d$], and further augmented with diffuse and polarization functions yielding a [$8s6p4d3f$] contracted set. For fluorine, the ($9s5p$) primitive set of Huzinaga was contracted to [$4s2p$], and augmented further to [$4s3p2d1f$]. The present calculations are performed at the MCPF level, except for the valence only calculations on Sc$^+$, which include only two electrons, and where SDCI was used instead. We will compare the results for ScF to the CPF results from Langhoff et al. [25]. Since both the $X^1\Sigma^+$ and $a^3\Delta$ states in ScF are relatively well described by a single configuration, we expect CPF and MCPF to give similar results.

Two sets of calculations were performed. In a first set, 8 electrons, originating from scandium $3d$,$4s$ and fluorine $2p$, were correlated. In a second set, the $3s$ and $3p$ electrons on scandium and the $2s$ electrons on fluorine were added to the correlation treatment, thus including 18 electrons. Since the present contracted basis sets are not designed to treat 3*s*,3*p* core correlation on Sc accurately, the second set of calculations was performed with an enlarged set, obtained by uncontracting four *s*, three *p* and three *d* primitives in the core-valence region.

First we consider the calculations on Sc and $Sc^+$. The results are shown in Table 18. The table includes the SCF results, the results obtained by correlating only 3*d*,4*s* electrons (denoted as MCPF(v), but notice that for $Sc^+$ this corresponds to SDCI), and the results obtained by correlating also 3*s*,3*p* (denoted as MCPF (c-v)). The first thing to note is the extremely large effect of 3*s*,3*p* correlation, both on the Sc ionization potential and on the $^1S$–$^3D$ splitting in $Sc^+$. Including core correlation raises the ionization potential by as much as 0.59 eV (with the primitive set). The final MCPF(c-v) result, 6.40 eV, is in excellent agreement with the experimental value of 6.56 eV. Adding the relativitistic correction suggested by Martin and Hay [15] slightly further improves the result of 6.42 eV. The $^1S$–$^3D$ splitting in $Sc^+$ is lowered by 0.30 eV by including 3*s*,3*p* in the correlation treatment, to a final value of 1.54 eV, only 0.10 eV above the experimental splitting. Introducing relativistic effects would slightly improve the result here too, since these effects always tend to stabilize the configuration with the largest number of

**Table 18.** MCPF calculations on Sc and $Sc^+$

| Basis | SCF | MCPF(v) | MCPF(c-v) |
|---|---|---|---|
| | $Sc^+$ $^3D$–Sc $^2D$ (eV) | | |
| Uncontracted | 5.35 | 5.81 | 6.40 |
| [7*s*5*p*4*d*3*f*][b] | 5.35 | 5.82 | 6.40 |
| [6*s*5*p*4*d*3*f*][b] | 5.37 | 5.85 | 6.40 |
| [6*s*4*p*3*d*2*f*][b] | 5.37 | 5.86 | 6.41 |
| [5*s*4*p*3*d*2*f*][b] | 5.40 | 5.90 | 6.45 |
| Exp.[a] | | 6.56 | |
| | $Sc^+$ $^1S$–$Sc^+$ $^3D$ (eV) | | |
| Uncontracted | 2.17 | 1.84 | 1.54 |
| [7*s*5*p*4*d*3*f*][b] | 2.17 | 1.83 | 1.51 |
| [6*s*5*p*4*d*3*f*][b] | 2.21 | 1.85 | 1.51 |
| [6*s*4*p*3*d*2*f*][b] | 2.25 | 1.88 | 1.37 |
| [5*s*4*p*3*d*2*f*][b] | 2.64 | 2.24 | 1.61 |
| Exp.[a] | | 1.44 | |
| | $Sc^+$ $^1S$–Sc $^2D$ (eV) | | |
| Uncontracted | 7.52 | 7,65 | 7.93 |
| [7*s*5*p*4*d*3*f*][b] | 7.52 | 7.65 | 7.91 |
| [6*s*5*p*4*d*3*f*][b] | 7.58 | 7.70 | 7.91 |
| [6*s*4*p*3*d*2*f*][b] | 7.62 | 7.74 | 7.78 |
| [5*s*4*p*3*d*2*f*][b] | 8.04 | 8.14 | 8.06 |
| Exp.[a] | | 8.00 | |

[a] Experimental results from [14]

[b] For MCPF(c-v) primitive functions in the core-valence region were uncontracted, as described in the text

valence *s* electrons ($^1$S in this case; the paper of Martin and Hay, however, does not include the relativistic correction for this state).

The results obtained with the contracted set for the ionization potential are still satisfactory, with truncation errors comparable to the errors obtained for the atomic valence excitation energies in the previous section. The same remains true when 3*s*,3*p* correlation is included, indicating that the present ANO basis sets are capable of capturing also this type of correlation, as long as the necessary functions are uncontracted or, alternatively, additional primitives are added. However, much more severe truncation errors occur for the $^1$S–$^3$D splitting in Sc$^+$. First it is clear that the [5*s*4*p*3*d*2*f*] contracted set gives an extremely bad description of the $^1$S(4$s^2$) state in Sc$^+$, even at the SCF level. The truncation error on the total energy is much larger for this state than for the two other states considered, leading for example to a $^1$S–$^3$D SCF excitation energy that is 0.47 eV higher than for the uncontracted set. The error is related to the fact that the $^1$S(4$s^2$) state of Sc$^+$ was not included in the density matrix averaging used for creating the ANO basis set. The present contraction scheme obviously is too limited to provide an accurate description of the difference in radial extent of the 4*s*-orbital between different states in all cases, when only a limited number of contracted *s*-functions (essentially two) is used. It should be noted however that Sc may be somewhat exceptional in this respect, in that the 4*s*-orbital is characterized by a large radius both in Sc and Sc$^+$. Adding one more *s*-function to the contracted set takes care of the major part of the error: with a [6*s*4*p*3*d*2*f*] contracted set, the truncation error on the $^1$S–$^3$D transition energy is reduced to less than 0.1 eV, both at the SCF and MCPF(v) level. A significantly larger truncation error does however appear at the MCPF(c-v) level. The error is now related to the dynamical correlation of the 3*s*,3*p* electrons. Apparently, the [6*s*4*p*3*d*2*f*] set simply lacks the necessary number of correlating functions to describe this correlation properly, and [6*s*5*p*4*d*3*f*] should be considered as a minimum (note that additional primitives have been added in the core-valence region, as described above).

We now turn to ScF and the MCPF results for this molecule, presented in Table 19, and consider first the $a^3\Delta$–$X^1\Sigma^+$ separation. With the uncontracted set, accurate $T_e$ values are obtained. The result from the 8-electron treatment exactly equals the CPF results from Langhoff et al. Somewhat surprisingly, in view of the large core correlation effect on the $^1$S–$^3$D separation in Sc$^+$, the $a^3\Delta$–$X^1\Sigma^+$ separation does not change to any significant extent with the inclusion of Sc 3*s* and 3*p* correlation. A slightly decreasing effect was obtained from the CPF calculations: our 18-electron MCPF result deviates by 0.04 eV from the CPF result from Langhoff. The basis set truncation errors on $T_e$ are acceptable, and merely reflect the errors already observed for Sc$^+$: [7*s*5*p*4*d*3*f*] leads to negligible errors, while [6*s*4*p*3*d*2*f*] should be used with care, especially when including core correlation.

As was the case for all other molecules considered in the series of test calculations, the present basis sets provide satisfactory results for the bond distance and stretching frequency of ScF. The bond distances obtained with the primitive basis sets are slightly (at most 0.006 Å) larger than the CPF results reported by Langhoff et al. The ground state distance is however still in agreement with experiment, with an error of 0.024 Å for the 8-electron correlation treatment, reduced to 0.012 Å when 18 electrons are correlated. Truncation errors are small, at most 0.006 Å. The stretching frequencies for both states compare well both to the CPF results and to the experimental ground state value, and also for this property the truncation errors are almost negligibly small.

**Table 19.** Spectroscopic constants for the $X^1\Sigma^+$ and $a^3\Delta$ states of ScF

| Basis Set | $r_e$(Å) | $\omega_e$(cm$^{-1}$) | $D_0$(eV) | $D'_0$(eV) | $T_e$(eV) | $\mu$(au) |
|---|---|---|---|---|---|---|
| $X^1\Sigma^+$ 8 electrons correlated | | | | | | |
| Uncontracted | 1.812 | 724 | 6.04 | 5.91 | | 0.661 |
| [7543]/[432] | 1.812 | 723 | 5.97 | 5.86 | | 0.752 |
| [6432]/[432] | 1.815 | 718 | 5.90 | 5.87 | | 0.791 |
| [6432]/[321] | 1.818 | 712 | 5.97 | 5.98 | | 0.906 |
| [7543]/[4321] | 1.812 | 725 | 6.03 | 5.90 | | 0.729 |
| CPF[a] | 1.807 | 725 | 5.91 | | | 0.555 |
| Expt.[b] | 1.788 | 735.6 | 6.1(7) | 6.1(7) | | |
| $X^1\Sigma^+$ 18 electrons correlated | | | | | | |
| Uncontracted | 1.800 | 729 | 5.92 | 6.16 | | 0.730 |
| [7543][c]/[432] | 1.798 | 729 | 5.92 | 6.16 | | 0.793 |
| [6432][c]/[432] | 1.801 | 734 | 5.99 | 6.09 | | 0.792 |
| [6432][c]/[321] | 1.798 | 734 | 6.15 | 6.40 | | 0.837 |
| [7543][c]/[4321] | 1.797 | 732 | 6.02 | 6.21 | | 0.783 |
| CPF[a] | 1.794 | 713 | 5.85 | | | 0.677 |
| Expt.[b] | 1.788 | 735.6 | 6.1(7) | 6.1(7) | | |
| $a^3\Delta$ 8 electrons correlated | | | | | | |
| Uncontracted | 1.890 | 628 | | | 0.26 | 1.244 |
| [7543]/[432] | 1.891 | 625 | | | 0.25 | 1.395 |
| [6432]/[432] | 1.894 | 625 | | | 0.21 | 1.419 |
| [6432]/[321] | 1.896 | 622 | | | 0.18 | 1.526 |
| [7543]/[4321] | 1.891 | 628 | | | 0.26 | 1.369 |
| CPF[a] | 1.886 | 612 | | | 0.26 | 1.076 |
| $a^3\Delta$ 18 electrons correlated | | | | | | |
| Uncontracted | 1.872 | 638 | | | 0.26 | 1.282 |
| [7543][c]/[432] | 1.871 | 637 | | | 0.26 | 1.384 |
| [6432][c]/[432] | 1.871 | 642 | | | 0.33 | 1.406 |
| [6432]/[321] | 1.868 | 644 | | | 0.34 | 1.457 |
| [7543][c]/[4321] | 1.869 | 641 | | | 0.26 | 1.369 |
| CPF[a] | 1.868 | 640 | | | 0.22 | 1.157 |

[a] CPF results from [25]

[b] Experimental results from [21]

[c] For MCPF(c-v) primitive functions in the core-valence region were uncontracted, as described in the text

In Table 19 $D_0$ denotes the dissociation energy with respect to the ground state atoms, while $D'_0$ stands for the dissociation to the ionic limits. The difference between both quantities is of course exclusively determined by the (compensating) errors on the computed values of the $Sc^+(^1S)$–$Sc(^2D)$ ionization energy (given in Table 19) and the fluorine electron affinity (shown in Table 7). In the valence-only (8 electrons) treatment, this difference is dominated by the large error on the scandium ionization energy, and $D_0$ is larger than $D'_0$. Including 3*s* and 3*p* in the correlation treatment takes care of the major part of the error for scandium, and the difference between $D'_0$ and $D_0$ is now determined by the error on the fluorine electron affinity. In view of the large $3s$, $3p$ correlation effect on the Sc ionization energy, we believe that the results obtained from the 18-electron correlation treatment should be regarded as superior to the valence-only treatment. The value of $D'_0$ obtained when correlating 18 electrons is in excellent agreement with the experimental value of the dissociation energy. A somewhat smaller, but still acceptable, result is obtained for $D_0$, due to the inherent lack of correlating functions in the fluorine basis, indispensable for an accurate description of its electron affinity. The value of $D_0$, obtained with the primitive sets, is slightly higher than the CPF result from Langhoff. The effect of the truncation is small, except for the calculation performed with a $[3s2p1d]$ set on F.

As was the case for CO (Sect. 4.2), the calculated dipole moment is much more sensitive to basis set contraction than the other properties. The results obtained with the primitive sets are larger than the CPF results from Langhoff, both for the $X^1\Sigma^+$ ground state and for the $a^3\Delta$ excited state, and basis set contraction further increases the values. Yet, our calculations do agree with the CPF results on the relative value of the dipole moments, being almost twice as large for the $a^3\Delta$ state than it is for the $X^1\Sigma^+$ ground state.

## 5 Conclusions

The basis sets of the present work have been obtained by an averaging procedure, which includes electronic states of the atoms and ions that are important in chemical bond formation. Thus good results are obtained for the atomic ionization potentials and electron affinities. Further it has been demonstrated that these basis sets yield excellent results for various molecular properties, notably properties related to the potential curve close to equilibrium such as bond-lengths and vibrational frequencies, even if they lack the functions necessary to produce accurate polarizabilities. Also valence spectra for transition metal compounds have been demonstrated to be well described by these basis sets. Like for any ANO type basis sets small superposition errors are found. In spite of the limited size of the present basis sets no significantly larger errors are obtained than for the larger sets such as [2, 3].

*Acknowledgements.* The research reported in this communication has been supported by a grant from the Swedish Natural Science Research Council (NFR), by IBM Sweden under a joint study contract. K. P. and B. D. thank the Belgian National Science Foundation (NFWO) and the Belgian Government (DPWB) for a research grant. We also thank Dr. Per-Åke Malmqvist and Dr. Markus Fülscher for helpful discussions.

## Appendix

The contracted ANO basis sets for H–Kr can be obtained via e-mail from one of the authors,

Kristin.Pierloot@chem.kuleuven.ac.be (KP),
ibmpow@garm.teokem.lu.se (POW) or
teobor@garm.teokem.lu.se (BOR).

The basis sets form part of the basis set library of the MOLCAS-3 quantum chemistry software [28].

## References

1. Almlöf J, Taylor PR (1987) J Chem Phys 86:4070
2. Widmark P-O, Malmqvist P-Å, Roos BO (1990) Theor Chim Acta 77:291
3. Widmark P-O, Persson BJ, Roos BO (1991) Theoret Chim Acta 79:419
4. Pou R, Merchán M, Widmark P-O, Roos BO, to be published.
5. Pierloot K, Roos BO (1992) Inorg Chem 31:5353
6. Pierloot K, Van Praet E, Vanquickenborne LG, Roos BO (1993) J Phys Chem 97:12220
7. Pierloot K, Tsokos E, Roos BO (1993) Chem Phys Lett 214:583
8. Schmidt MW, Ruedenberg K (1979) J Chem Phys 71:3951
9. Van Duijneveldt, IBM technical research report RJ945
10. Huzinaga S (1971) Approximate atomic functions II. Department of Chemistry report, University of Alberta, Edmonton, Alberta, Canada
11. Fægri K, Speis HJ (1986) J Chem Phys 86:7035
12. Bauschlicher CW Jr, Taylor PR (1993) Theoret Chim Acta 86:13
13. Moore CE (1949) Circ Nat Bureau Stand 467
14. Sugar J, Corliss C (1977) J Phys Chem Ref Data 6:317; (1977) 6:1256; (1978) 7:1194; (1979) 8:1; (1981) 10:1097; (1982) 11:135
15. Martin RL, Hay PJ (1981) J Chem Phys 75:4539
16. Hotop H, Lineberger WC (1985) J Phys Chem Ref Data 14:731
17. Feller D, Davidson ER (1989) J Chem Phys 90:1042
18. Woon DE, Dunning TH (1993) J Chem Phys 99:3730
19. Huo WM (1965) J Chem Phys 43:625
20. Barnes L, Lindh R (1993) J Chem Phys 98:3972
21. Huber KP, Herzberg G (1979) Constants of Diatomic Molecules. Van Nostrand-Reinhold, New York
22. Petterson LGM, Persson BJ (1993) Chem Phys 170:149
23. Andersson K, Malmquist PÅ, Roos BO, Sadlej AJ, Wolinski K (1990) J Phys Chem 94:5483; Andersson K, Malmqvist PÅ, Roos BO (1992) J Chem Phys 96:1218
24. Andersson K, Roos BO (1993) Int J Quantum Chem 45:591
25. Langhoff SR, Bauschlicher CW Jr, Partridge H (1988) J Chem Phys 89:396
26. Pou-Amérigo R, Merchán M, Nebot-Gil I, Roos BO (1994) J Chem Phys 101:4893; Persson BJ, Roos BO, Pierloot K (to be published).
27. Andersson K, Blomberg MRA, Fülscher MP, Kellö V, Lindh R, Malmqvist PÅ, Noga J, Olsen J, Roos BO, Sadlej AJ, Siegbahn PEM, Urban M, Widmark P-O (1992) MOLCAS Version 2 User's Guide. Dept of Theor Chem., Chem Centre, Univ of Lund, Lund
28. Andersson K, Blomberg MRA, Fülscher MP, Kellö V, Lindh R, Malmqvist PÅ, Noga J, Olsen J, Roos BO, Sadlej AJ, Siegbahn PEM, Urban M, Widmark P-O (1994) MOLCAS Version 3 User's Guide. Dept of Theor Chem., Chem Centre, Univ of Lund, Lund