

# Fully optimized contracted Gaussian basis sets for atoms Li to Kr

Ansgar Schäfer, Hans Horn, and Reinhart Ahlrichs

*Institut für Physikalische Chemie und Elektrochemie, Lehrstuhl für Theoretische Chemie, Universität Karlsruhe, Kaiserstr. 12, 7500 Karlsruhe, Germany*

(Received 18 March 1992; accepted 6 May 1992)

Various contracted Gaussian basis sets for atoms up to Kr are presented which have been determined by optimizing atomic self-consistent field ground state energies with respect to all basis set parameters, i.e., orbital exponents and contraction coefficients.

## INTRODUCTION

Chemical bonding usually does not affect inner shells of atoms. Atomic contributions from the valence shell, however, may contract or expand and be polarized depending on the actual bonding situation. The smallest (atom centered) basis set which can be expected to reliably account for these effects, e.g., in self-consistent field (SCF) treatments, is of SVP (split valence polarization) type, which means that inner shell atomic orbitals (AOs) are described by a single basis function, two basis functions are provided for each valence shell AO, augmented by a set of polarization functions.

CGTO (contracted Gaussian-type orbital) basis sets can, of course, be obtained by segmented or generalized contraction of atom optimized GTO basis sets. Since we will not deal with this popular and successful procedure, we refer the reader to recent reviews.<sup>1-5</sup> It should be noted, however, that the basis sets obtained by segmented contraction are clearly not fully optimized, i.e., the basis set parameters do not minimize an appropriate measure for accuracy such as the atomic ground state energy. Atom optimized basis sets in a generalized contraction are clearly fully optimized.<sup>6</sup> Since we consider it advantageous to use fully optimized segmented CGTO sets, we will present various basis sets of this kind, especially SV basis sets.

The most widely and successfully used atom optimized basis sets of SV type have been determined by Pople *et al.*, the 4-31 G (Ref. 7) and 6-31 G (Ref. 8) basis sets. The main and probably only drawback of 4-31 G bases, to give an example, concerns the relatively poor energy especially for second row elements: the 4-31 G atomic energies for Al to Cl are even higher than those of Huzinaga's (9s5p) basis.<sup>9</sup> The SV type 4-31 G basis does not appear to be an efficient contraction of a (12s8p) primitive basis.<sup>2</sup>

SV contractions can be conveniently obtained from atom optimized SZ (single zeta) basis sets by decontracting valence AOs. SZ basis sets are available for various numbers of GTOs per AO for atoms up to Rn from the work of Huzinaga and co-workers.<sup>10</sup> These SZ sets have proved invaluable to us since their study provided important hints for the design of fully atom optimized CGTO sets presented in this article.

## METHOD OF COMPUTATION, ACCURACY, AND NOMENCLATURE

The methodology of GTO basis set optimization has been developed and described in detail by Faegri and Alm-

löf.<sup>11</sup> These authors consider energy optimization of GTO basis sets by means of analytical first and second order derivatives. We refer the reader to this paper, which contains valuable remarks, and sketch technical aspects only briefly.

The basis sets presented in this article were optimized with the program system TURBOMOLE,<sup>12</sup> which allows for the calculation of analytic gradients of the SCF energy with respect to orbital exponents and contraction coefficients (for atoms and molecules). The general expression for the SCF energy gradient is as follows<sup>13</sup> (closed shell case assumed for simplicity):

$$E' = \sum_{\mu\nu} D_{\mu\nu} h'_{\mu\nu} - \sum_{\mu\nu} W_{\mu\nu} S'_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu\kappa\lambda} D_{\mu\nu} D_{\kappa\lambda} [(\mu\nu|\kappa\lambda) - \frac{1}{2}(\mu\kappa|\nu\lambda)]' \quad (1)$$

with

$$D_{\mu\nu} = 2 \sum_{a \in \text{occ}} c_{\mu a} c_{\nu a} \quad (\text{density}), \quad (2)$$

$$W_{\mu\nu} = 2 \sum_{a \in \text{occ}} \epsilon_a c_{\mu a} c_{\nu a} \quad (\text{energy weighted density}). \quad (3)$$

The indices  $\mu, \nu, \kappa, \lambda$  run over all basis functions (CGTOs in general). The density matrices  $\mathbf{D}$  and  $\mathbf{W}$  are obtained from the preceding SCF calculation and the main task of a gradient step is to calculate differentiated integrals<sup>11,14,15</sup> of the general form ( $\hat{O}$  denotes an arbitrary one-electron operator)

$$\langle \mu | \hat{O} | \nu \rangle' = \langle \mu' | \hat{O} | \nu \rangle + \langle \mu | \hat{O}' | \nu \rangle + \langle \mu | \hat{O} | \nu' \rangle, \quad (4)$$

$$(\mu\nu|\kappa\lambda)' = (\mu'\nu|\kappa\lambda) + (\mu\nu'|\kappa\lambda) + (\mu\nu|\kappa'\lambda) + (\mu\nu|\kappa\lambda'). \quad (5)$$

A CGTO  $G_{\mu A}$  centered at nucleus  $A$  at  $\mathbf{r}_A = (x_A, y_A, z_A)$  is defined as

$$G_{\mu A} = \sum_{i=1}^{F_{\mu}} d_i g_{iA}, \quad (6)$$

with

$$g_{iA} = N_i (x - x_A)^k (y - y_A)^l (z - z_A)^m e^{-\eta_i (r - r_A)^2}, \quad (7)$$

$$N_i = \left( \frac{2^{4L-1}}{\pi^3} \right)^{1/4} \left( \frac{1}{(2k-1)!!(2l-1)!!(2m-1)!!} \right)^{1/2} \times \eta_i^{(1/2)(L+1/2)}, \quad (8)$$

$$L = k + l + m + 1. \quad (9)$$

It is convenient for the present purpose not to require normalization of CGTOs since the SCF energy does not depend on the normalization of AOs. Differentiation of the CGTO  $G_{\mu A}$  with respect to one of its exponents  $\eta_i$  or contraction coefficients  $d_i$  yields

$$\frac{\partial G_{\mu A}}{\partial \eta_i} = d_i \frac{\partial g_{iA}}{\partial \eta_i} = \frac{1}{2} \left( L + \frac{1}{2} \right) \frac{d_i}{\eta_i} g_{iA} - d_i (\mathbf{r} - \mathbf{r}_A)^2 g_{iA}, \quad (10)$$

$$\frac{\partial G_{\mu A}}{\partial d_i} = g_{iA}. \quad (11)$$

The Eqs. (1)–(11) fully specify the gradient of  $E_{\text{SCF}}$  with respect to the individual  $\eta_i$  and  $d_i$  in CGTOs in terms of matrix elements of one- and two-electron operators (which are clearly independent of  $\eta_i$  and  $d_i$ ).

The analytic gradients computed in this way were then used to home in onto a local minimum of the energy by means of a relaxation procedure also used for geometry optimizations. We did not compute second derivatives. Since the relaxation procedure is designed to home in on minima we are confident that (at least) a local minimum of the corresponding atomic energy has been found in all cases. The relaxation procedure is not very satisfactory. There is clearly strong coupling of parameters to be optimized and the convergence occasionally requires assistance, e.g., first freezing some parameters or resetting the approximate hessian to a diagonal matrix. After some experience with a reasonable choice of starting parameters (scaling of CGTOs from a neighbor atom, using the approximate hessian from a previous optimization at a neighbor atom), the procedure converged automatically although up to 150 iterations were necessary to converge the energy to  $10^{-11}$  a.u. (in some cases only to  $10^{-10}$  a.u.), the density to  $10^{-9}$ , and all parameters to 7 decimal figures. The optimization was terminated when gradients were within numerical noise. The virial coefficient was always  $2 \pm 10^{-8}$ , although this was not enforced by global scaling.

The following nomenclature will be used to characterize basis sets:  $(k, l, m)$  denotes a GTO basis with  $k$  GTOs of  $s$  type,  $l$  GTOs of  $p$  type and  $m$  GTOs of  $d$  type;  $[k, l, m]$  similarly denotes the number of CGTOs of  $s, p, d$  type;  $\{s_1 s_2 \dots | p_1 p_2 \dots | d_1 \dots\}$  denotes the contraction pattern (e.g.,  $s_1$  GTOs of  $s$  type in the first CGTO).

## FULLY ATOM-OPTIMIZED CGTO BASIS SETS UP TO Kr

### SV basis sets

The smallest (in the valence shell) SV basis just contains two uncontracted GTOs in the valence shell. Perusing Huzinaga's SZ sets it occurred to us that such a "minimal choice" might be useful provided there is a core AO of the same  $l$  quantum number. The  $\{432/42\}$  basis of Huzinaga *et al.*<sup>10</sup> for P ( $4S$ ), to give an example, is only 16 mH higher in energy than the  $\{433/43\}$  basis. It was hoped that this loss in energy might be further reduced by fully optimizing a SV basis from the very start.

Having decided to use a 2 GTO valence shell representation, it remained to determine the SZ description of inner shells. We have been led here by considerations which will be sketched now for a typical example. Consider the  $2p$  and  $3p$  shells of second row atoms. The number  $p_1$  of  $p$  type GTOs to describe the  $2p$  AO was obtained from calculations employing  $p_1 = 4, 5, 6$ . It turned out that for  $p_1 < 5$  the valence  $(3p)$   $p$  GTOs tended to improve the  $2p$  AOs, whereas for  $p_1 = 5$  or  $p_1 = 6$  this was not the case: the 2 GTO representation of the  $3p$  shell remained stable for  $p_1 \geq 5$ .

Considerations basically identical to the one just described led to the following choice; SV basis sets: (a) 1st row elements—Li, Be:  $\{511\}$ , i.e., (7)/[3] or  $\{511/31\}$ , i.e., (7,4)/[3,2] (Be  $3P$ ), B–Ne:  $\{511/31\}$ , i.e., (7,4)/[3,2]; (b) 2nd row elements—Na, Mg:  $\{5311/5\}$ , i.e., (10,5)/[4,1] or  $\{5311/511\}$ , i.e., (10,7)/[4,3] (Mg  $3P$ ), Al–Ar:  $\{5311/511\}$ , i.e., (10,7)/[4,3]; (c) 3rd row elements—K, Ca:  $\{63311/53\}$ , i.e., (14,8)/[5,2] or  $\{63311/5311\}$ , i.e., (14,10)/[5,4] (Ca  $3P$ ), Sc–Zn:  $\{63311/53/41\}$ , i.e., (14,8,5)/[5,2,2] or  $\{63311/53/51\}$ , i.e., (14,8,6)/[5,2,2], Ga–Kr:  $\{63311/5311/41\}$ , i.e., (14,10,5)/[5,4,2].

The above reasoning does not help to fix the chosen (31) representation for the  $2p$  shell of first row atoms. Since most integral programs "like" higher contractions we have decided for a larger basis in cases of doubt. Although a (21) representation for  $2p$  could have also been considered for B to F, this appeared to be unbalanced compared to the (5) representation for  $1s$ , especially for elements at the right end of the period. Our choice has been vindicated by test calculations to be discussed below.

Some comments are to be made on alkali metals. When optimizing basis sets of the size of the present SV basis, only one of the two uncontracted  $s$  functions stays in the valence region, i.e., the total minimum is not of SV type. We therefore first optimized the corresponding SZ basis sets and then split the valence  $s$  exponent and optimized the splitting factor. It turned out that for Li and K this led to a good starting point for a full optimization, which then homed in on a local minimum of the energy with two functions in the valence region.<sup>16</sup> For Na we could not locate such a minimum, although there might be a very shallow one.

### Larger basis sets

For comparison we have also optimized CGTO basis sets of DZ (double zeta) type; DZ basis sets: (a) 1st row elements—Li, Be:  $\{5111\}$ , i.e., (8)/[4], B–Ne:  $\{5111/31\}$ , i.e., (8,4)/[4,2]; (b) 2nd row elements—Na, Mg:  $\{52111/41\}$ , i.e., (11,5)/[6,2], Al–Ar:  $\{52111/4111\}$ , i.e., (11,7)/[6,4]; (c) 3rd row elements—K, Ca:  $\{6211111/33111\}$ , i.e., (14,9)/[8,5], Sc–Zn:  $\{6211111/33111/311\}$ , i.e., (14,9,5)/[8,5,3], Ga–Kr:  $\{6211111/331211/41\}$ , i.e., (14,11,5)/[8,6,2], as well as the corresponding uncontracted basis sets. To our surprise it turned out that Huzinaga's atom optimized GTO basis sets<sup>9</sup> are in part far from being optimized. Although we achieve relatively minor lowerings of atomic ground state energies [e.g., by 3 mH for the (11,7) basis for the  $4S$  state of P], the orbital

TABLE I. RHF-SCF energies in hartrees for present basis sets for atoms Li to Ar. SV, DZ, TZ as described in the text.

Atom	State	Basis set			HF limit (Ref. 24)
		SV	DZ	TZ	
Li	$^2S$	-7.425 064	-7.430 999	-7.432 617	-7.432 727
Be	$^1S$	-14.553 727	-14.570 335	-14.572 656	-14.573 023
	$^3P$	-14.497 387	-14.508 807	-14.511 236	-14.511 50
B	$^2P$	-24.497 559	-24.522 830	-24.528 238	-24.529 061
C	$^3P$	-37.641 143	-37.676 346	-37.687 291	-37.688 619
N	$^4S$	-54.332 925	-54.379 482	-54.398 890	-54.400 934
O	$^3P$	-74.713 746	-74.773 732	-74.806 289	-74.809 398
F	$^2P$	-99.279 690	-99.354 519	-99.404 838	-99.409 349
Ne	$^1S$	-128.376 407	-128.467 528	-128.540 811	-128.547 10
Na	$^2S$	-161.791 703	-161.838 183	-161.851 998	-161.858 91
Mg	$^1S$	-199.533 593	-199.591 591	-199.607 443	-199.614 63
	$^3P$	-199.472 004	-199.527 802	-199.542 446	...
Al	$^2P$	-241.786 870	-241.855 166	-241.871 899	-241.876 71
Si	$^3P$	-288.749 785	-288.828 840	-288.848 554	-288.854 36
P	$^4S$	-340.597 368	-340.688 104	-340.711 921	-340.718 78
S	$^3P$	-397.362 235	-397.466 231	-397.496 617	-397.504 90
Cl	$^2P$	-459.315 194	-459.433 828	-459.473 030	-459.482 07
Ar	$^1S$	-526.623 385	-526.758 128	-526.802 631	-526.817 51

exponents differ by a factor of up to 1.6. [The largest exponent in Huzinaga's P( $^4S$ ) (11,7) basis is 30632, while it is 48591 in our optimized basis.] In any case it turned out that the DZ basis sets offer no appreciable advantage over SV basis sets.

For more accurate calculations and/or for extreme bonding situations, one thus needs a considerably larger basis than SV or DZ which is of genuine TZ (triple zeta) quality; TZ basis sets: (a) 1st row elements—Li, Be: {511111}, i.e., (10)/[6], B–Ne: {511111/411}, i.e., (10,6)/[6,3]; (b) 2nd row elements—Na, Mg: {512111/51}, i.e., (12,6)/[7,2], Al–Ar: {512111/51111}, i.e., (12,9)/[7,5].

We have again optimized uncontracted GTO and CGTO basis sets, and it turns out again that appreciable changes with respect to Huzinaga's basis sets are found.

### Transition metals

Since only Roothaan type open shell cases are implemented in TURBOMOLE, we briefly indicate how the  $d^n$  ground states were treated. The cases  $d^1$  ( $^2D$ ;  $a=b=0$ ),  $d^4$  ( $^5D$ ;  $a=15/16$ ,  $b=15/8$ ),  $d^6$  ( $^5D$ ;  $a=35/36$ ,  $b=25/18$ ) and  $d^9$  ( $^2D$ ;  $a=b=80/81$ ) are of Roothaan type and cause no problems. The remaining  $d^n$  ground states were treated in  $O_h$  symmetry where the five  $d$  functions split into  $e_g$  ( $d_{z^2}, d_{x^2-y^2}$ ) and  $t_{2g}$  ( $d_{xy}, d_{xz}, d_{yz}$ ). One then gets the following correspondences which involve only high spin states of half filled shells (i.e.,  $a=1$ ,  $b=2$  in all cases)

$$e_g^2 \cdot {}^3A_{2g} \rightarrow {}^3F_g(d^2),$$

$$t_{2g}^3 \cdot {}^4A_{2g} \rightarrow {}^4F_g(d^3),$$

$$e_g^4 t_{2g}^3 \cdot {}^4A_{2g} \rightarrow {}^4F_g(d^7),$$

$$e_g^2 t_{2g}^6 \cdot {}^3A_{2g} \rightarrow {}^3F_g(d^8).$$

The terms so obtained in  $O_h$  have no component of other atomic terms of the same spin, e.g.,  ${}^3A_{2g}$  has no  ${}^3P_g$  component. For the  $d^7$  and  $d^8$  configurations one has to use a full contraction of the  $d$  GTOs to maintain  $L$  symmetry (same radial function for  $e_g$  and  $t_{2g}$ ), i.e., one has to optimize the energy with respect to orbital exponents and contraction coefficients by means of gradient techniques as described above, even if a (partially) uncontracted  $d$  basis is desired.

### COMPARISON OF ATOMIC SCF GROUND STATE ENERGIES

In Tables I and II we have listed atomic ground state restricted Hartree-Fock self-consistent field (RHF-SCF) energies obtained with present fully optimized basis sets. We consider N( $^4S$ ), P( $^4S$ ), and Mn( $^6S$ ) in some detail to give the reader an idea of accuracies of various basis sets (Tables III–V).

The present SV basis of type (7,4)/[3,2] for N( $^4S$ ) is 8 mH lower in energy than the 4-31 G basis, although 4-31 G is based on an (8,4) (GTO) set. The 68 mH deviation of the SV energy from the HF limit results mainly from deficiencies in the  $1s$  part. Extension from {511/31} to {611/31} recovers already 12 mH without leading to any noticeable improvement in the valence region. The DZ basis loses only 5.5 mH through contraction (with optimization of all parameters), roughly 25% of the distance to the HF limit. The present (9,5) basis is 0.4 mH lower in energy than the corresponding Huzinaga basis. Our (10,6) basis is also 0.2 mH lower than the corresponding Huzinaga basis and 1.8 mH higher than the HF limit. Contraction of this basis—and optimization of all parameters—implies a loss of only 0.2 mH.

Considering the basis sets of SV type for P( $^4S$ )—present, 4-31 G, and 6-31 G—the present SV basis is much

TABLE II. RHF-SCF energies in hartrees for present basis sets for atoms K to Kr. SV, DZ, TZ as described in the text. Energies marked with an asterisk refer to optimized basis sets with six *d* GTOs. Occupancy of transition metals is  $3d^4 4s^2$ , except for Ni( $^3D$ ): $3d^2 4s^1$ , Ni( $^1S$ ): $3d^{10}$ , Cu( $^2S$ ): $3d^{10} 4s^1$ .

Atom	State	Basis set		
		SV	DZ	HF limit (Ref. 24)
K	$^2S$	-599.076 962	-599.148 385	-599.164 79
Ca	$^1S$	-676.660 070	-676.740 256	-676.758 18
	$^3P$	-676.622 455	-676.701 640	...
Sc	$^2D$	-759.625 235	-759.713 573	-759.735 72
Ti	$^3F$	-848.282 096	-848.379 067	-848.406 00
V	$^4F$	-942.745 719	-942.851 763	-942.884 33
Cr	$^7S$	-1043.196 070	-1043.309 807	...
	$^5D$	-1043.155 118	-1043.270 664	-1043.309 8
Mn	$^6S$	-1149.694 241	-1149.819 721	-1149.866 2
Fe	$^5D$	-1262.251 794	-1262.387 575	-1262.443 7
Co	$^4F$	-1381.201 346	-1381.347 858	-1381.414 6
Ni	$^3F$	-1506.634 601	-1506.792 258	-1506.870 9
		-1506.665 423*	...	...
	$^3D$	-1506.574 107	-1506.729 374	...
		-1506.615 804*	...	...
Cu	$^1S$	-1506.397 399	-1506.550 369	...
		-1506.454 934*	...	...
	$^2S$	-1638.687 105	-1638.853 786	...
Zn		...	-1638.904 562*	...
	$^2D$	-1638.688 824	-1638.858 038	-1638.950 1
		...	-1638.896 639*	...
Zn	$^1S$	-1777.560 282	-1777.741 472	-1777.848 1
Ga	$^2P$	-1922.962 642	-1923.157 019	-1923.261 0
Ge	$^3P$	-2075.046 166	-2075.254 187	-2075.359 7
As	$^4S$	-2233.908 080	-2234.130 041	-2234.238 6
Se	$^3P$	-2399.517 577	-2399.753 936	-2399.867 6
Br	$^2P$	-2572.070 448	-2572.321 688	-2572.441 3
Kr	$^1S$	-2751.661 991	-2751.928 671	-2752.055 0

closer to 6-31 G [derived from a (16,10) primitive basis] than to 4-31 G [derived from (12,8)], although it is obtained from a (10,7) GTO basis. As further evidence for the quality of the present SV basis we have plotted in Fig. 1 the radial density distribution of the  $3p$  SCF AO. Although we "offer" only two GTOs for the valence shell, we

TABLE III. RHF-SCF energies for various basis sets for N( $^4S$ ) ground state, ordered according to the energy.

Basis	$E_{SCF}/a.u. + 54$
4-31 G (Ref. 7)	-0.325 289 <sup>a</sup>
SV {511/31}	-0.332 925
{611/31}	-0.344 963
DZ {5111/31}	-0.379 482
6-31 G (Ref. 8)	-0.382 575 <sup>a</sup>
(8,4)	-0.385 045
(9,5) Huzinaga (Ref. 9)	-0.395 473
{51111/311}	-0.395 684
(9,5)	-0.395 885
TZ {511111/411}	-0.398 890
(10,6) Huzinaga (Ref. 9)	-0.398 925
(10,6)	-0.399 097
HF limit (Ref. 24)	-0.400 934

<sup>a</sup>RHF-SCF energies calculated with TURBOMOLE. Pople *et al.* give the following UHF energies: -54.327 925 a.u. for 4-31 G, -54.385 385 a.u. for 6-31 G.

TABLE IV. RHF-SCF energies for various basis sets for P( $^4S$ ) ground state, ordered according to the energy.

Basis	$E_{SCF}/a.u. + 340$
4-31 G (Ref. 7)	-0.320 156 <sup>a</sup>
SV {5311/511}	-0.597 368
{6311/511}	-0.656 000
DZ {521111/4111}	-0.688 104
6-31 G (Ref. 8)	-0.688 937 <sup>a</sup>
(11,7) Huzinaga (Ref. 9)	-0.689 484
(11,7)	-0.692 914
TZ {5121111/511111}	-0.711 921
(12,9)	-0.712 856
HF limit (Ref. 24)	-0.718 78

<sup>a</sup>RHF-SCF energies calculated with TURBOMOLE. Pople *et al.* give the following UHF energies: -340.320 188 a.u. for 4-31 G, -340.690 20 a.u. for 6-31 G.

have the same quality for the  $3p$  AO for  $0 < r < 2 a_0$  (roughly the maximum of the radial density distribution) as the 4-31 G basis, whereas the latter leads to markedly larger deviations (more than a factor of two) from the near HF limit [e.g., a (17,12) GTO basis set] result at  $r > 2 a_0$ . For the  $3s$  AO the relative accuracy of the SV basis is even better.

The energy error of the present SV basis {5311/511} with respect to the HF solution (of 121 mH) is mainly due to deficiencies in the description of the  $1s$  AO. An extension from {5311/511} to {6311/511} recovers already 48% of the distance to the HF limit. Again we have not found any effect of this improved  $1s$  AO representation on the valence shell.

The present DZ contraction of an (11,7) GTO basis implies a loss of only 4.8 mH in energy as compared to a deviation of 31 mH to the HF limit. Our test calculations, in part discussed below, do not point to a sufficient advantage of this DZ basis over the present SV basis. For more accurate calculations we therefore recommend to go directly to the TZ (12,9)/[7,5] basis with contraction pattern {5121111/511111}. The corresponding energy is only 6.9 mH above the HF limit and the contraction of the (12,9) basis (with optimization) incurs a loss of only 1 mH.

A series of tests was performed for the [Ar]( $4s^2 3d^2$ )  $^6S$  state of Mn to assess accuracies and potential shortcomings

TABLE V. RHF-SCF energies for various basis sets for Mn ( $^6S$ ) ground state, ordered according to the energy.

Basis	$E_{SCF}/a.u. + 1149$
Huzinaga (Ref. 10)	{5333/53/5} -0.535
Present	{53311/53/41} -0.536
SV	{63311/53/41} -0.694
Wachters (Ref. 17)	{62111111/33111/311} -0.787
Present	{83311/73/5} -0.794
DZ	{62111111/33111/311} -0.820
Wachters (Ref. 17)	(14,9,5) -0.832
Present	(14,9,5) -0.835
Clementi (Ref. 25)	STO <sup>a</sup> -0.866
HF limit (Ref. 24)	-0.866

<sup>a</sup>Slater-type orbitals (STO).

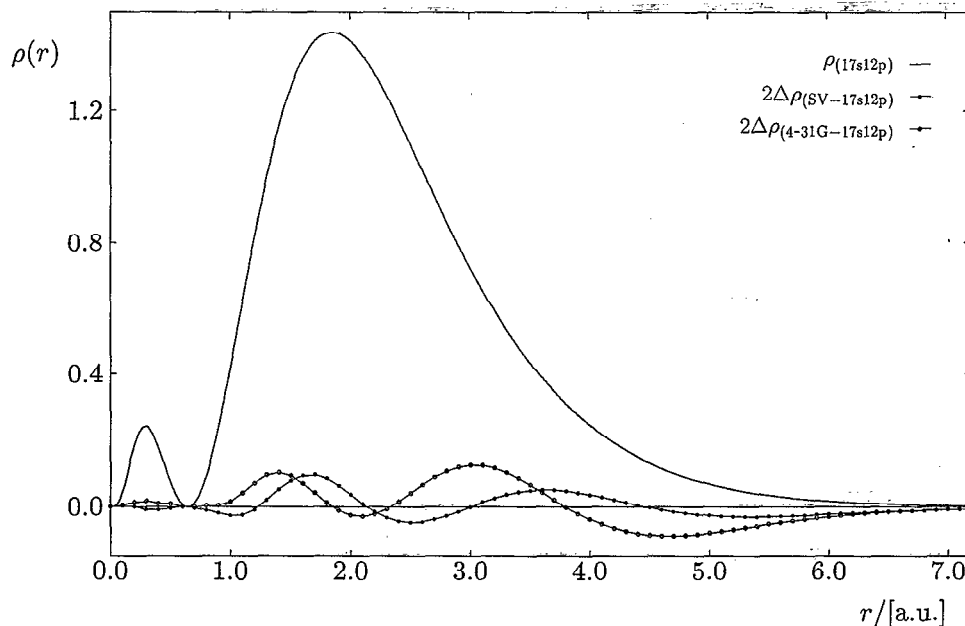


FIG. 1. Radial density distribution  $\rho(r) = 4\pi r^2 |\varphi(r)|^2$  of the  $3p$  AO of  $P(^4S)$  with a near HF basis (17,12) and deviations for the 4-31 G and present SV basis sets.

of various basis sets. Table V contains the SCF energies we have obtained, ordered according to the energy, where "present" always implies full optimization of all parameters. On the basis of these results we selected {63311/53/41} as optimal basis of SV type. The SCF energy differs by 172 mH from the HF limit, of which 100 mH (58%) can be recovered by just extending to {83311/73/41}, i.e., by a better description of the  $1s$  and  $2p$  AOs. This extension has no visible effect on the outer shell description. The remaining discrepancy, i.e., {83311/73/41} vs the HF limit (of 72 mH) reflects errors in the approximation of outer shells, mainly the  $2s$  AO. If a better description than the present SV basis is desired, we recommend the present fully optimized Wachters type contraction {62111111/33111/311} which is only 46 mH higher in energy than the HF limit and only 12 mH higher than the uncontracted Wachters basis.<sup>17</sup> We finally note that careful optimization lowers the energy of the Wachters basis by 3.7 mH (11% of the distance to the HF limit) and that the present {53311/53/41} energy is even lower than the best SZ {5333/53/5}<sup>10</sup> result.

As an additional indication of the appropriateness of the present SV basis we compare the present orbital exponents with those obtained by Wachters for the most diffuse functions: 0.703 (present)/0.652 (Wachters), 0.106/0.084, 0.040/0.031 for the three outermost  $s$  functions; 3.337/3.477, 1.379/1.341, 0.539/0.505 for the three outermost  $p$  functions; 1.212/1.103, 0.388/0.337 for the two most diffuse  $d$  functions. The relatively minor deviations strongly indicate similar quality in the description of  $3s$ ,  $3p$ ,  $3d$ , and  $4s$ , at least if the most diffuse GTO of  $3s$  and  $3p$  is taken out of the contraction (and corresponding contractions are used for  $3d$ ).

## SIMPLE MOLECULAR TEST CALCULATIONS

Equilibrium geometry determinations of ethane, ethene, ethine, benzene, and azulene with SV, SVP, DZP and the fullerene  $C_{60}$  with SV and SVP basis sets (with an optimized (4)/[2] basis for H) show the following trends.<sup>18</sup>

The SV structure constants differ by at most 1.0 pm and  $0.2^\circ$  from SVP and by at most 1.4 pm and  $0.3^\circ$  from DZP results. For  $C_{60}$ , e.g., we get the C-C distances 145.7 and 138.1 pm (SV) and 145.1/137.6 pm (SVP) which have to be compared to the TZP MP2 (second order Møller-Plesset) equilibrium distances of 144.6/140.6 pm.<sup>19</sup>

As typical examples of molecules with second row atoms we report results for  $P_4$  as a typical covalent molecule and for  $AlCl_3$  and  $SF_6$  as more ionic cases. The following results were obtained for various basis sets in comparison to experiment,<sup>20</sup> listed in the order SVP/DZP/TZP/exp ( $d$  exponents: Al 0.3, P 0.4, S 0.6, Cl 0.65, F 1.0)

$$P_4(T_d) \quad R_e/\text{pm} = 219.8/219.8/219.5/221,$$

$$\Delta E_{\text{SCF}}/\text{kJ mol}^{-1} = -414/-405/-408,$$

$$AlCl_3(D_{3h}) \quad R_e/\text{pm} = 208.0/207.9/208.0/206,$$

$$\Delta E_{\text{SCF}}/\text{kJ mol}^{-1} = -1057/-1036/-1028,$$

$$SF_6(O_h) \quad R_e/\text{pm} = 154.6/155.2/154.4/156,$$

$$\Delta E_{\text{SCF}}/\text{kJ mol}^{-1} = -970/-923/-876,$$

$\Delta E_{\text{SCF}}$  is the SCF bonding energy, i.e.,

$$\Delta E_{\text{SCF}} = E_{\text{SCF}}(\text{molecule}) - \sum E_{\text{SCF}}(\text{atoms}).$$

The results demonstrate the quality of SVP basis sets in SCF treatments, e.g., in comparison to DZP or TZP, and confirms the quantitative accuracy of the SCF approximation as far as structure constants are concerned. The SVP basis also shows no shortcomings with respect to SCF binding energies for  $\text{P}_4$  or  $\text{AlCl}_3$ . Only for  $\text{SF}_6$  which shows a rather delicate balance of ionic ( $\text{S}^{6+}\text{F}_6^-$ ) and hypervalent ( $3d$  on S) contributions to bonding, a larger basis is required for a quantitative account of energetics of bonding.

Finally we report results for  $\text{TiCl}_4$  as a transition metal compound which can be treated on the SCF level. The equilibrium geometry was determined with SVP and DZP basis sets [ $\eta_p=0.065$  (Ti),  $\eta_d=0.34$  (Cl)] with following results listed in the order SVP/DZP/exp.<sup>20</sup>

$$\text{TiCl}_4(T_d) \quad R_e/\text{pm}=217.3/219.5/219,$$

$$\Delta E_{\text{SCF}}/\text{kJ mol}^{-1} = -1081/-1106.$$

Again the SVP basis shows relatively good accuracy with respect to molecular structure and bonding energy as compared to DZP.

## RECOMMENDATIONS FOR MOLECULAR CALCULATIONS

The use of atom optimized basis sets in molecular calculations requires some modifications and extensions which have been extensively discussed in the literature.<sup>1-5</sup> We only make a few comments in connection with the basis sets presented in this article.

*Polarization functions* should be added to SV or larger basis sets to obtain a reliable account of bonding effects in SCF calculations. An SVP basis appears to yield equilibrium geometries sufficiently close to the HF limit. Polarization functions for H are here of lesser importance,<sup>5</sup> and hydrocarbons, benzene or related compounds, and fullerenes can even be treated on the SV level.<sup>18</sup>

*Alkaline earth metals.* Since it is absolutely necessary to have valence  $p$  functions added to the basis sets of alkaline earth elements in molecular calculations, we have optimized the corresponding SV and DZ basis sets for the excited state  $^3P$  (i.e.,  $ns^1np^1$ ). Alternatively one can add two  $p$  sets to the SV basis obtained for the ground state with exponents similar to those of the valence  $s$  functions.

For *boron* we recommend to scale the  $p$  exponents by a factor of 1.3, since the  $p$  contribution to molecular orbitals (MOs) is always markedly contracted as compared to the (relatively weakly bound)  $2p$  AO of the free atom.

*SVP versus larger basis sets.* If more extended basis sets than SVP are to be used, we recommend to go directly to TZP [or TZDP] for atoms up to Ar and the DZP basis sets for K to Kr. The present SV basis may be decontracted for the shell below the valence shell (by taking out the most diffuse function), but we have not encountered cases so far where this led to marked improvements. We recommend in any case that an SVP treatment for a structure determina-

tion is followed by a single point calculation with a larger basis to obtain a more reliable energy—unless sufficient experience about the performance of SVP bases is already available.

*Transition metal compounds* show marked variations of  $d$  occupation. To be on the safe side it is therefore recommended to contract the ( $5d$ ) GTO set {311} in a TZ type fashion. For the electron rich atoms Ni or Cu a {411} contraction of the ( $6d$ ) basis may be necessary since this gives a better balanced description of  $3d^n4s^2$  vs  $3d^{n+1}4s^1$  (see Table II). Transition metal complexes with formal oxidation of +2 or larger for the metal atom show only weak  $4s$  and  $4p$  occupations. In these cases one can actually drop the more diffuse of the two GTOs describing the  $4s$  AO and add a  $4p$  set with the same exponent as the steeper  $s$  GTO. In complexes with formal charges  $M^0$  or  $M^{+1}$ , we recommend to add two  $p$  GTO sets with exponents as for the  $4s$  subshell to account for polarization effects or to use the  $4p$  functions determined by Wachters for excited states.<sup>17</sup>

*Relativistic effects* are already noticeable for post  $3d$  elements, so that it might be better in some cases to use effective core potentials (ECPs) with relativistic corrections instead of all electron basis sets.<sup>21</sup>

*Small basis sets for methyl,  $-\text{CH}_3$ , and tertiary butyl,  $-\text{C}(\text{CH}_3)_3$ .* Methyl (Me) and tertiary butyl ( $^t\text{Bu}$ ) groups are often used in preparative chemistry for steric protection. A proper description of steric requirements can be achieved by relatively small basis sets of basically SZ quality. A simultaneous optimization of basis sets and structure constants has been performed for methane and neopentane as parent compounds for Me and  $^t\text{Bu}$ . The basis sets chosen for this purpose are H: {3}, C: {42/3} or {511/31}, where the larger SV basis is used only for the central carbon atom which then connects these groups to the rest of the molecule. The following geometries were obtained [values in parentheses: SV basis at all atoms ({31} is used as SV and DZ basis at H)/DZP basis [ $\eta_p=0.46$  (H);  $\eta_d=0.46$  (C)/experiment]

$$\text{CH}_4:\text{C-H}=108.3(109.3/108.8/108.5^{22}) \text{ pm},$$

$$\text{C}(\text{CH}_3)_4:\text{C-C}=154.3(154.2/153.3/153.4 \pm 0.3^{23}) \text{ pm},$$

$$\text{C-H}=108.7(109.5/109.1/111.4 \pm 0.8) \text{ pm},$$

$$\angle\text{CCH}=111.2^\circ(111.0^\circ/111.2^\circ/112^\circ \pm 3^\circ).$$

The present fully optimized minimal basis sets clearly give a reliable description of the steric demand of these bulky groups.

## CONCLUSION

Full optimization of all basis set parameters (exponents and contraction coefficients) by means of gradient techniques leads to basis sets of relatively small size with least possible loss in accuracy. These basis sets are expected to be superior to those created by just contracting atom

optimized GTO sets. The present fully optimized SV basis sets are considered to be the best compromise of expense and accuracy.

## AVAILABILITY

The basis sets presented in this article are available via FTP (file transfer protocol) at internet address "tchibm3.chemie.uni-karlsruhe.de" (internet number 129.13.108.8) with login-ID "anonymous" in the directory "/pub/basis." Besides those mentioned in Tables I and II, the files include also basis sets of type (7,4), (8,4), (9,5), (9,5)/[5,3]/{51111/311}, (10,6), (11,7), (11,7)/[6,4]/{611111/4111}, (13,9) for first row, (10,7), (11,7), (12,8), (12,9), (13,10), (13,10)/[8,6]/{61111111/511111}, (15,12), (17,12) for second row, (14,9,5) for third row transition metals, and (14,11,5), (14,11,5)/[8,6,2]/{62111111/611111/41} for third row main group elements.

## ACKNOWLEDGMENTS

This work was supported by the Volkswagen-Stiftung and by the Fonds der Chemischen Industrie.

- <sup>1</sup>T. H. Dunning and P. J. Hay, in *Modern Theoretical Chemistry: Methods of Electronic Structure Theory*, edited by H. F. Schaefer III (Plenum, New York, 1977), Vol. 3.  
<sup>2</sup>R. Ahlrichs and P. R. Taylor, *J. Chim. Phys.* **78**, 315 (1981).  
<sup>3</sup>S. Huzinaga, *Comp. Phys. Rep.* **2**, 279 (1985).  
<sup>4</sup>E. R. Davidson and D. Feller, *Chem. Rev.* **86**, 681 (1986).

- <sup>5</sup>W. J. Hehre, L. Random, P. v. R. Schleyer, and J. A. Pople, *Ab initio Molecular Orbital Theory* (Wiley, New York, 1986).  
<sup>6</sup>R. C. Raffanetti, *J. Chem. Phys.* **58**, 4452 (1973).  
<sup>7</sup>R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971); W. J. Hehre and W. A. Lathan, *ibid.* **56**, 5255 (1972).  
<sup>8</sup>W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem. Phys.* **56**, 2257 (1972); M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees, and J. A. Pople, *ibid.* **77**, 3654 (1982).  
<sup>9</sup>S. Huzinaga, *Approximate Atomic Functions I, II* (Division of Theoretical Chemistry, Department of Chemistry, University of Alberta, 1971).  
<sup>10</sup>*Gaussian Basis Sets for Molecular Calculations*, edited by S. Huzinaga (Physical Sciences Data, Elsevier, Amsterdam, 1984), Vol. 16.  
<sup>11</sup>K. Faegri, Jr. and J. Almlöf, *J. Comput. Chem.* **7**, 396 (1986).  
<sup>12</sup>R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, *Chem. Phys. Lett.* **162**, 165 (1989); Recent extensions required for the optimization of CGTOs are described in A. Schäfer, diploma thesis, Karlsruhe, 1991.  
<sup>13</sup>P. Pulay, *Mol. Phys.* **17**, 197 (1969).  
<sup>14</sup>J. Almlöf and P. R. Taylor, *Int. J. Quantum Chem.* **27**, 743 (1985).  
<sup>15</sup>G. Tonachini and H. B. Schlegel, *J. Chem. Phys.* **87**, 514 (1987).  
<sup>16</sup>F. B. van Duijneveldt, IBM Technical Research Report No. RJ-945, (1971).  
<sup>17</sup>A. J. H. Wachters, *J. Chem. Phys.* **52**, 1033 (1970).  
<sup>18</sup>S. Richard (private communication).  
<sup>19</sup>M. Häser, J. Almlöf, and G. E. Scuseria, *Chem. Phys. Lett.* **181**, 497 (1991).  
<sup>20</sup>A. F. Wells, *Structural Inorganic Chemistry*, 5th ed. (Oxford University, Oxford, 1984).  
<sup>21</sup>P. A. Christiansen, W. C. Ermler, and K. S. Pitzer, *Annu. Rev. Phys. Chem.* **36**, 407 (1985).  
<sup>22</sup>L. S. Bartell, K. Kuchitsu, and R. J. deNeui, *J. Chem. Phys.* **33**, 1254 (1960).  
<sup>23</sup>L. S. Bartell and W. F. Bradford, *J. Mol. Struct.* **37**, 113 (1977).  
<sup>24</sup>C. Froese Fischer, *The Hartree-Fock Method for Atoms. A Numerical Approach* (Wiley, New York, 1977).  
<sup>25</sup>E. Clementi and C. Roetti, *At. Data Nucl. Data Tables* **14**, 177 (1974).