



Kent Academic Repository

Celeste, Ricardo, P. Maringolo, Milena, Comar Jr, Moacyr, B. Viana, Rommel, R. Guimarães, Amanda, L. A. Haiduke, Roberto and B. F. da Silva, Albérico (2015) *Accurate Gaussian basis sets for atomic and molecular calculations obtained from the generator coordinate method with polynomial discretization*. Journal of Molecular Modeling . ISSN 1610-2940.

Downloaded from

<https://kar.kent.ac.uk/115833/> The University of Kent's Academic Repository KAR

The version of record is available from

<https://doi.org/10.1007/s00894-015-2813-z>

This document version

Publisher pdf

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in **Title of Journal** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries

If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

Accurate Gaussian basis sets for atomic and molecular calculations obtained from the generator coordinate method with polynomial discretization

Ricardo Celeste¹ · Milena P. Maringolo² · Moacyr Comar Jr.³ · Rommel B. Viana² · Amanda R. Guimarães² · Roberto L. A. Haiduke² · Albérico B. F. da Silva^{2,4}

Received: 13 June 2015 / Accepted: 7 September 2015 / Published online: 29 September 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Accurate Gaussian basis sets for atoms from H to Ba were obtained by means of the generator coordinate Hartree-Fock (GCHF) method based on a polynomial expansion to discretize the Griffin-Wheeler-Hartree-Fock equations (GWHF). The discretization of the GWHF equations in this procedure is based on a mesh of points not equally distributed in contrast with the original GCHF method. The results of atomic Hartree-Fock energies demonstrate the capability of these polynomial expansions in designing compact and accurate basis sets to be used in molecular calculations and the maximum error found when compared to numerical values is only 0.788 mHartree for indium. Some test calculations with the B3LYP exchange-correlation functional for N₂, F₂, CO, NO, HF, and HCN show that total energies within 1.0 to 2.4 mHartree compared to the cc-pV5Z basis sets are attained with our contracted bases with a much smaller number of polarization functions (2p1d and 2d1f for hydrogen and heavier atoms, respectively). Other molecular calculations performed here are also in very good accordance with experimental and cc-pV5Z results. The most important point to be mentioned here is that our generator coordinate basis sets required

only a tiny fraction of the computational time when compared to B3LYP/cc-pV5Z calculations.

Keywords Gaussian basis sets · Generator coordinate method · Polynomial discretization

Introduction

Electronic structure calculations of atoms and molecules are commonly performed through expansions of orbitals in a finite set of known functions, which are called basis sets, according with the method proposed by Roothaan [1]. This development initiated the search for more compact and still accurate basis sets for application in large scale molecular calculations. Thus, Mohallem et al. [2] introduced the generator coordinate Hartree-Fock (GCHF) method in 1986 and a new perspective was devised based on a previous concept of universal basis sets developed by Silver and collaborators [3, 4]. This method furnished a new way to obtain sets of Gaussian-type functions (GTFs) and Slater-type functions (STFs) for atoms. Soon after, in 1987, Mohallem and Trsic [5] presented a universal Gaussian basis set for atoms from Li to Ne by applying the GCHF method. In that same year, da Costa et al. [6] also designed universal GCHF basis sets for He to Ar. Moreover, da Silva and collaborators [7] extended this treatment in 1989 to heavier atoms, from H to Kr, while de Castro and Jorge [8] published in 1998 an accurate universal Gaussian basis set that is intended to be used for all atoms of the periodic table. Later, de Castro et al. [9, 10] developed a new way to generate GTF sets by adjusting different groups of GCHF parameters for a given symmetry (s, p, d, etc.) through a segmentation in two or more regions and applied this proposal to first and second row atoms and ions. Finally, Jorge and Muniz [11] also designed accurate Gaussian GCFH basis

✉ Albérico B. F. da Silva
alberico@iqsc.usp.br

¹ Departamento de Química, Centro de Ciências Exatas e de Tecnologia, Universidade Estadual do Centro-Oeste, Guarapuava, PR, Brazil

² Departamento de Química e Física Molecular, Instituto de Química de São Carlos, Universidade de São Paulo, São Carlos, SP, Brazil

³ Laboratório de Bioinformática, Universidade Federal de São João del Rei, Campus Centro-Oeste Dona Lindú, Divinópolis, MG, Brazil

⁴ Av. Trabalhador São Carlense, 400 (Caixa-postal: 780). Bairro: Centro, 13566-590 São Carlos, SP, Brazil

sets for H to Xe, which are also tailored for each atom and do not follow the universal concept [9, 10]. Thus, the GCHF method has demonstrated undoubtedly, for more than two decades, a great tool for designing Gaussian basis sets of high quality [12, 13].

However, in contrast to their simplicity, universal basis sets obtained by the GCHF method show, as a general characteristic, the need for a large number of functions in order to reach the accuracy level of mHartree [8] when compared to numerical Hartree-Fock (NHF) results [14, 15]. This feature exposes a major limitation of universal sets for application in molecular systems because the computational cost for such implementation avoids their generalized usage. Moreover, energy values from universal sets deteriorate rapidly as small numbers of primitive functions are chosen. Hence, as an alternative to overcome these restrictions presented by universal basis sets, Barbosa and da Silva [16] introduced in 2003 a new proposal to design adapted basis sets of GTFs through a polynomial discretization of the GWHF equations. As a preliminary study, they applied this new technique to a very few atoms of the first, second, and third row of the periodic table. These results clearly demonstrated that this methodology is capable of generating accurate and more compact Gaussian basis sets [16].

In the work reported here we extended by a lot the number of atoms treated by means of this new polynomial approach for the discretization of the GWHF equations. The obtained GCHF basis sets of GTFs for atoms from H ($Z=1$) to Ba ($Z=56$) is discussed. Finally, a comparative study of results derived from the GCHF Gaussian basis sets obtained in this work for atoms and small molecules (N_2 , F_2 , CO, NO, HF, and HCN) is also performed.

Computational details

Polynomial expansion and the generator coordinate Hartree-Fock method

The GWHF equations are integrated numerically by using a discretization technique introduced by Mohallem [17] known as integral discretization (ID). More details about the formalism related to the GCHF method in its polynomial version can be found in the work of Barbosa and da Silva [16] and we will reproduce here only the two main equations. They proposed a discretization recipe in which a point $\Omega_k^{(w)}$ of the mesh for each w atomic symmetry ($w=s, p, d, f$, etc.) is given by:

$$\Omega_k^{(w)} = \frac{\ln \alpha_k^{(w)}}{A} = \Omega_{\min}^{(w)} + (k-1)\Delta\Omega_1^{(w)} + (k-1)^2\Delta\Omega_2^{(w)} + (k-1)^3\Delta\Omega_3^{(w)} + \cdots + (k-1)^q\Delta\Omega_q^{(w)} \quad (1)$$

in which $\alpha_k^{(w)}$ is the k th GTF exponent and A is a scaling parameter. Moreover, $\alpha_k^{(w)}$ along with a selected group of $\Delta\Omega_i^{(w)}$ parameters are optimized for each atomic system. The index k for a given symmetry varies from 1 to the number of discretization points. Each exponent is associated with a discretization point and can be determined by using the following expression:

$$\alpha_k^{(w)} = \exp A \left\{ \Omega_{\min}^{(w)} + (k-1)\Delta\Omega_1^{(w)} + (k-1)^2\Delta\Omega_2^{(w)} + (k-1)^3\Delta\Omega_3^{(w)} + \cdots + (k-1)^q\Delta\Omega_q^{(w)} \right\} \quad (2)$$

Hence, the discretization of the GWHF equations in this approach is performed by a flexible mesh in which the intervals between adjacent points in the Ω space are not restrained to be fixed, as in the original GCHF method. All higher order $\Delta\Omega_i^{(w)}$ parameters beyond $\Delta\Omega_1^{(w)}$ are neglected in the traditional GCHF formalism. Because of this limitation, the original GCHF method can only produce basis sets of satisfactory accuracy when a large number of discretization points (GTFs) is achieved. Thus, this procedure was not able to generate competitive basis sets of small and medium sizes when compared with other alternatives, being more successful when applied to heavy atoms [16].

In practice, the flexibility introduced by the implementation of Eqs. 1 and 2 allows us to use a more effective distribution of discretization points that can better represent a certain region of the atom associated with some distance range from the nucleus. The modification of the way in which the GWHF equations are discretized ensures the design of small and accurate basis sets to application in calculations of medium to large molecular systems [16]. The basis sets developed here with this polynomial expansion are labeled as p-GCHF sets.

Basis sets

Following the work of Barbosa and da Silva [16], the numerical parameter A assumes the value 6.00 in all calculations. The letter q in Eqs. 1 and 2 represents the degree of the polynomial. All sets of GTFs were developed in this work by using polynomials with $q=3$ since this choice produced high quality results along with a reduced computational demand during optimization. The procedure to obtain the final GTF basis sets is the same for all cases reported here. The use of third degree polynomials requires four initial parameters $\Omega_{\min}^{(w)}$ and $\Delta\Omega_q^{(w)}$ for each occupied symmetry of an atom. Equation 2 shows how to obtain the exponents for a given set of parameters ($\Omega_{\min}^{(w)}$, $\Delta\Omega_1^{(w)}$, $\Delta\Omega_2^{(w)}$, and $\Delta\Omega_3^{(w)}$). Extended sets of GTFs from Partridge [18, 19] were initially used to provide the parameter set for the initial step after a polynomial fitting for each symmetry of a particular atom. Thus, the SIMPLEX

method [20] coupled with the ATOMSCF [21] program is used to optimize the parameter set. The final parameters, $\Omega_{\min}^{(w)}$ and $\Delta\Omega_q^{(w)}$, generates the definitive exponents of the basis set.

Molecular calculations

The molecular calculations were done here with the package Gaussian 03 [22]. The general contraction scheme [23] was employed in this work in order to obtain a quintuple-zeta version of our p-GCHF GTF basis sets. This means that an orbital of a certain symmetry belonging to the valence region of an atom is described by six functions, one contracted plus five single GTFs (these uncontracted GTFs are always the most diffuse ones of that particular symmetry). Our contracted p-GCHF Gaussian basis sets have also been augmented by polarization functions. Hence, in this work we have added 2p1d and 2d1f polarization functions to hydrogen and heavier atoms, respectively.

Results and discussion

Atomic calculations

In Table 1 we present the HF energy for atoms from H to Ba. The results were calculated with the Gaussian basis sets generated in this work with the polynomial discretization of the GWHF equations (p-GCHF sets), as described before. As a matter of fact, we will compare our values with those from basis sets obtained by the original GCHF method, which are labeled as adapted Gaussian basis sets (AGBSs) by the authors [11]. The AGBSs were constructed applying the original GCHF recipe atom by atom, using a different mesh of equally spaced discretization points for each system. NHF energies [14, 15] are also used to judge the quality of both families of basis sets.

First, one can observe that the number of GTFs generated with the polynomial discretization for each atom is significantly smaller than the respective number of AGBS functions. The reduction in size varies from 2s to 8s, from 1p to 6p, and from 2d to 4d functions. Heavy elements tend to show the most important size differences. Of course, a more relevant decrease in size compared to the AGBSs would also be attained for light elements if a little larger basis set error were pursued in this work for these atoms. Furthermore, it is important to highlight the fact that HF energies obtained from p-GCHF sets are in general only slightly larger than those from the respective AGBSs. The most important errors with respect to NHF data are 0.788 mHartree for indium with p-GCHF results and 0.613 mHartree for iodine with AGBSs. Hence, both basis set families respect the 1 mHartree error limit for all these atomic systems. One can also notice that the AGBSs were not developed for cesium and barium [11].

Another clear difference in p-GCHF sets and AGBSs is observed in the way new functions are included along the periodic table. The basis sets developed here only add new GTFs of that symmetry associated with the specific subshell being filled and maintains the size until a new subshell begins to be occupied. The AGBSs from the original GCHF method do not follow a so regular procedure and basis set sizes change more erratically. Thus, the results presented here reinforce the fact that the polynomial discretization is a powerful technique to generate more compact and equally accurate GTF sets when compared with those given by the original GCHF.

Molecular calculations

The contracted p-GCHF Gaussian basis sets containing polarization functions were used in some test calculations with the B3LYP exchange–correlation functional for N₂, F₂, CO, NO, HF, and HCN and the values obtained are shown in Table 2. Results from 6-31G (3df,3pd), 6-311G (3df,3pd), cc-pVTZ, cc-pVQZ, and cc-pV5Z sets are also analyzed along this section. The molecular quantities obtained were total energies, dipole moments, and harmonic vibrational frequencies.

The first point to be noticed is that the CPU time for calculations with our p-GCHF sets are similar to those with the 6-31G (3df,3pd), 6-311G (3df,3pd), and cc-pVTZ ones. On the other hand, cc-pVQZ and cc-pV5Z sets exhibit, respectively, CPU requirements 2–4 and 10–39 times larger than those from our p-GCHF Gaussian basis sets for geometry optimizations and frequency calculations. Total energies with the cc-pV5Z sets are the smallest ones for all these molecules followed closely by results from our sets with deviations up to 2.4 mHartree in CO. The largest differences with respect to total energies from cc-pV5Z are more relevant for the other basis sets studied: 5.9 mHartree for cc-pVQZ, 24.4 mHartree for cc-pVTZ, 37.7 mHartree for 6-311G (3df,3pd), and 100.7 mHartree for 6-31G (3df,3pd).

Moreover, the largest difference in vibrational frequencies between data from cc-pV5Z and p-GCHF sets is 17.6 cm^{−1} for HF. Frequencies from 6-31G (3df,3pd) and 6-311G (3df,3pd) sets showed much larger deviations with respect to those calculated with cc-pV5Z than p-GCHF (these differences can reach 85.1 and 56.4 cm^{−1} for HF, respectively). There is also some tendency for overestimation of frequencies by 6-31G (3df,3pd) and 6-311G (3df,3pd) sets in comparison with cc-pV5Z. Frequencies from cc-pVTZ and cc-pVQZ are comparable to those from p-GCHF for almost all molecules but are in closer agreement with those from the largest cc-pV5Z for HF (errors of 6.2 and 3.8 cm^{−1}, respectively). However, cc-pVTZ showed the largest error in the frequency for C-H stretching in HCN (11 cm^{−1}).

Dipole moments calculated with 6-31G (3df,3pd) and 6-311G (3df,3pd) sets are again in worse accordance with values from cc-pV5Z than those given by our p-GCHF sets and the

Table 1 Basis set size and negative values for HF energies (Hartree)

Atom	State	Size (this work)	Energy (this work)	Size ^a	Energy ^a	NHF ^b
H	(² S)	18 s	0.499999997	20s	0.499999992	0.500000000 ^c
He	(¹ S)	18 s	2.861679979	23 s	2.861679949	2.861679995
Li	(² S)	18 s	7.432726334	24 s	7.432726888	7.432726927
Be	(¹ S)	18 s	14.57302221	20s	14.57302294	14.57302316
B	(² P)	18s12p	24.52905963	24s13p	24.52906022	24.52906072
C	(³ P)	18s12p	37.68861696	23s14p	37.68861786	37.68861895
N	(⁴ S)	18s12p	54.40093022	23s14p	54.40093249	54.40093419
O	(³ P)	18s12p	74.80939542	23s14p	74.80939595	74.80939845
F	(² P)	18s12p	99.40934394	23s15p	99.40934657	99.40934933
Ne	(¹ S)	18s12p	128.5470903	23s14p	128.5470930	128.5470980
Na	(² S)	20s12p	161.8588893	25s16p	161.8589040	161.8589116
Mg	(¹ S)	20s12p	199.6146089	25s15p	199.6146261	199.6146363
Al	(² P)	20s14p	241.8766744	26s18p	241.8766995	241.8767072
Si	(³ P)	20s14p	288.8543288	26s18p	288.8543540	288.8543624
P	(⁴ S)	20s14p	340.7187436	26s18p	340.7187714	340.7187808
S	(³ P)	20s14p	397.5048553	26s18p	397.5048844	397.5048958
Cl	(² P)	20s14p	459.4820261	26s18p	459.4820589	459.4820721
Ar	(¹ S)	20s14p	526.8174608	26s18p	526.8174973	526.8175126
K	(² S)	22s14p	599.1647106	29s18p	599.1647736	599.1647865
Ca	(¹ S)	22s14p	676.7580956	28s18p	676.7581567	676.7581857
Sc	(² D)	22s14p10d	759.7356156	29s19p12d	759.7356972	759.7357178
Ti	(³ F)	22s14p10d	848.4058846	29s19p14d	848.4059730	848.4059967
V	(⁴ F)	22s14p10d	942.8842135	29s19p14d	942.8843116	942.8843374
Cr	(⁷ S)	22s14p10d	1043.356206	30s19p14d	1043.356358	1043.356376
Mn	(⁶ S)	22s14p10d	1149.866088	29s19p13d	1149.866215	1149.866251
Fe	(⁵ D)	22s14p10d	1262.443465	30s19p14d	1262.443633	1262.443665
Co	(⁴ F)	22s14p10d	1381.414337	30s19p14d	1381.414516	1381.414553
Ni	(³ F)	22s14p10d	1506.870666	29s19p14d	1506.870857	1506.870908
Cu	(² S)	22s14p10d	1638.963409	30s19p14d	1638.963691	1638.963742
Zn	(¹ S)	22s14p10d	1777.847807	30s19p14d	1777.848065	1777.848116
Ga	(² P)	22s16p10d	1923.260606	30s21p14d	1923.260940	1923.261009
Ge	(³ P)	22s16p10d	2075.359355	29s22p14d	2075.359675	2075.359733
As	(⁴ S)	22s16p10d	2234.238286	29s21p14d	2234.238586	2234.238654
Se	(³ P)	22s16p10d	2399.867225	29s21p14d	2399.867539	2399.867611
Br	(² P)	22s16p10d	2572.440936	29s21p13d	2572.441247	2572.441332
Kr	(¹ S)	22s16p10d	2752.054578	30s20p13d	2752.054867	2752.054977
Rb	(² S)	24s16p10d	2938.356998	30s19p14d	2938.357424	2938.357453
Sr	(¹ S)	24s16p10d	3131.545250	29s20p13d	3131.545463	3131.545686
Y	(² D)	24s16p13d	3331.683707	29s21p16d	3331.683932	3331.684169
Zr	(³ F)	24s16p13d	3538.994607	29s21p16d	3538.994837	3538.995064
Nb	(⁶ D)	24s16p13d	3753.597448	29s21p16d	3753.597493	3753.597727
Mo	(⁷ S)	24s16p13d	3975.549446	30s22p16d	3975.549414	3975.549499
Tc	(⁶ S)	24s16p13d	4204.788245	31s21p16d	4204.788541	4204.788736
Ru	(⁵ F)	24s16p13d	4441.538905	30s21p15d	4441.539216	4441.539487
Rh	(⁴ F)	24s16p13d	4685.881107	31s21p16d	4685.881475	4685.881703
Pd	(¹ S)	24s16p13d	4937.920458	30s21p16d	4937.920821	4937.921023
Ag	(² S)	24s16p13d	5197.697776	31s21p17d	5197.698239	5197.698472
Cd	(¹ S)	24s16p13d	5465.132488	31s20p16d	5465.132845	5465.133141
In	(² P)	24s19p13d	5740.168366	30s22p16d	5740.168703	5740.169154

Table 1 (continued)

Atom	State	Size (this work)	Energy (this work)	Size ^a	Energy ^a	NHF ^b
Sn	(³ P)	24s19p13d	6022.931027	30s21p16d	6022.931115	6022.931694
Sb	(⁴ S)	24s19p13d	6313.484727	30s22p16d	6313.484911	6313.485319
Te	(³ P)	24s19p13d	6611.783487	30s23p16d	6611.783735	6611.784058
I	(² P)	24s19p13d	6917.980336	30s21p16d	6917.980282	6917.980895
Xe	(¹ S)	24s19p13d	7232.137809	30s22p16d	7232.137986	7232.138363
Cs	(² S)	27s19p13d	7553.933052	32s22p15d ^d	7553.932730 ^d	7553.933658
Ba	(¹ S)	27s19p13d	7883.543109	32s22p16d ^d	7883.542961 ^d	7883.543827

^a Values for a basis set developed by the traditional GCHF method from reference [11]^b Numerical Hartree-Fock energies from reference [14]^c Analytical result^d Reference [8]**Table 2** Total energies (Hartree), dipole moments (Debye), harmonic vibrational frequencies (cm⁻¹), and CPU time for B3LYP calculations with several basis sets along with some experimental data

	Property	6-31G(3df,3pd)	6-311G(3df,3pd)	cc-pVTZ	cc-pVQZ	cc-pV5Z	p-GCHF ^a	Exp[20]
N₂	Energy	-109.5372259	-109.5645535	-109.5684291	-109.5769839	-109.5798174	-109.5781259	
	CPU time (opt)	1 min 46 s	1 min 37 s	1 min 12 s	4 min 27 s	37 min 18 s	1 min 51 s	
	Frequency	2443.2	2451.6	2449.4	2447.6	2448.6	2450.0	2358.6
	CPU time (freq)	1 min 17 s	1 min 18 s	1 min 6 s	5 min 43 s	37 min 57 s	1 min 36 s	
F₂	Energy	-199.510131	-199.5730564	-199.5863638	-199.6048666	-199.6107937	-199.6096775	
	CPU time (opt)	3 min 59 s	2 min 24 s	2 min 4 s	6 min 32 s	1 h 1 min 46 s	2 min 47 s	
	Frequency	1074.6	1047.8	1050.7	1045.9	1045.9	1042.4	916.6
	CPU time (freq)	2 min 46 s	1 min 34 s	1 min 15 s	5 min 39 s	38 min 0 s	1 min 43 s	
CO	Energy	-113.3231666	-113.3541964	-113.3572536	-113.3667645	-113.3695062	-113.3670758	
	CPU time (opt)	1 min 43 s	1 min 50s	1 min 40s	7 min 30s	1 h 24 min 47 s	2 min 59 s	
	Dipole moment	0.03	0.14	0.13	0.11	0.10	0.08	0.11
	Frequency	2216.6	2223.4	2212.4	2213.9	2213.1	2207.0	2169.8
	CPU time (freq)	1 min 20s	1 min 34 s	1 min 24 s	6 min 58 s	49 min 58 s	2 min 0 s	
NO	Energy	-129.9029357	-129.9355227	-129.9405218	-129.9514118	-129.954936	-129.953102	
	CPU time (opt)	2 min 34 s	2 min 37 s	2 min 34 s	9 min 29 s	1 h 51 min 34 s	3 min 12 s	
	Dipole moment	0.08	0.13	0.11	0.12	0.13	0.12	0.16
	Frequency	1987.3	1988.3	1976.5	1975.1	1973.9	1973.6	1904.2
	CPU time (freq)	1 min 36 s	1 min 44 s	1 min 23 s	8 min 41 s	1 h 4 min 30s	2 min 18 s	
HF	Energy	-100.4438313	-100.4804524	-100.4835704	-100.4957801	-100.500068	-100.4990551	
	CPU time (opt)	1 min 30s	1 min 31 s	1 min 28 s	3 min 41 s	16 min 9 s	1 min 35 s	
	Dipole moment	1.74	1.65	1.83	1.82	1.83	1.91	1.83
	Frequency	4166.7	4138.0	4087.8	4085.4	4081.6	4064.0	4138.3
	CPU time (freq)	1 min 11 s	1 min 13 s	1 min 1 s	4 min 13 s	14 min 23 s	1 min 23 s	
HCN	Energy	-93.4347956	-93.4580869	-93.4613719	-93.4686271	-93.4708114	-93.4692365	
	CPU time (opt)	2 min 16 s	1 min 29 s	1 min 16 s	5 min 57 s	1 h 16 min 23 s	1 min 58 s	
	Dipole moment	2.98	2.93	2.98	3.02	3.04	3.04	2.98
	Frequency	3509.0	3452.6	3453.1	3441.7	3442.1	3442.7	3311
	Frequency	2215	2205	2201	2201	2200	2199	2097
	Frequency	779.3	746.5	760.5	760.1	757.1	760.9	712
	CPU time (freq)	2 min 1 s	2 min 23 s	1 min 44 s	12 min 47 s	1 h 34 min 6 s	3 min 7 s	

^a Contracted and augmented by polarization functions (2p1d and 2d1f for hydrogen and heavier atoms, respectively)

Table 3 The average error for some properties calculated with the B3LYP method and several Gaussian basis sets when compared with B3LYP/cc-pV5Z results

Property	6-31G(3df,3pd)	6-311G(3df,3pd)	cc-pVTZ	cc-pVQZ	p-GCHF ^a
Energy (mHartree)	55.6	20.0	14.7	3.6	1.6
Dipole moment (Debye)	0.07	0.08	0.03	0.01	0.03
Frequencies (cm ⁻¹)	30.0	14.0	3.8	1.4	4.3

^a Contracted and augmented by polarization functions (2p1d and 2d1f for hydrogen and heavier atoms, respectively)

largest discrepancies are observed in HF (0.09 and 0.18 Debye, respectively). The 6-31G (3df,3pd), and 6-311G (3df, 3pd) values for HF and HCN leads to some underestimation of values for this electric property. The largest deviation between p-GCHF and cc-pV5Z dipole moments, 0.08 Debye, also occurs for HF. Further, cc-pVTZ is better for this property in HF than p-GCHF but the error with respect to cc-pV5Z for HCN is now 0.06 Debye. Finally, cc-pVQZ is in excellent agreement with cc-pV5Z for this property with deviations up to 0.02 Debye.

Experimental values are also shown in Table 2 [24]. The overall concordance of calculated quantities with the measured data is quite good though error cancelations may occur between the approximated B3LYP treatment and small basis sets. This fortuity can lead to closer agreement with experiment for a particular combination of method and basis set as noticed for example in the vibrational frequency calculated at the B3LYP/6-311G (3df,3pd) level for HF, which is almost exactly the same as the experimental determination. Hence, these error cancelations impose some limitation in the use of experimental data to judge the performance of new basis sets.

Accordingly, the average error for these three properties with respect to results from cc-pV5Z in Table 3 reinforces that p-GCHF sets are quite accurate for total energies. On the other hand, the comparison of dipole moments and vibrational frequencies shows that p-GCHF results are approximately at the same level as cc-pVTZ ones, while cc-pVQZ is in slightly closer agreement with results from the largest cc-pV5Z set. Finally, p-GCHF sets also are better than 6-31G (3df,3pd) and 6-311G (3df,3pd) for these two molecular properties. Considering that the overall performance of 6-31G (3df,3pd) and 6-311G (3df,3dp) sets is worse than that observed with our basis sets or with cc-pVTZ and taking into account the fact that the first two sets own more polarization functions than the last two, it is possible to argue that a more careful treatment of fundamental (primitive) functions in a basis set can result in higher accuracy than that attained by the simple augmentation of polarization functions to a not so precise (poor primitive) basis set.

Conclusions

Based on the atomic and molecular results shown here, the use of a polynomial expansion to discretize the Griffin-Wheeler-

Hartree-Fock equations is capable of generating flexible and accurate Gaussian basis sets with promising perspectives of being largely used in molecular calculations. Another important aspect to be emphasized is the reduction in the size of sets of GTFs achieved due to the polynomial discretization.

The molecular calculations done with the B3LYP method also show that total energies given by the basis sets developed here are quite accurate with respect to cc-pV5Z results, with average deviations of only 1.6 mHartree. The results for this property indicate that our basis set is better than the other alternatives also investigated: 6-31G (3df,3pd), 6-311G (3df, 3pd), cc-pVTZ, and cc-pVQZ. In addition, the computational cost from calculations with our basis sets is only a tiny fraction of that from the cc-pV5Z set and is also significantly smaller than the one assigned to cc-pVQZ.

Moreover, dipole moments and vibrational frequencies also indicate that our basis sets are superior when compared to the 6-31G (3df,3pd) and 6-311G (3df,3pd) basis sets and are extremely competitive with the very large cc-pVQZ and cc-pV5Z basis sets.

This work also points to the conclusion that a more advanced treatment of sets of fundamental (primitive) functions for an atom may result in better molecular results of electronic structure calculations than those achieved only by increasing the number of polarization functions of a less accurate (poor primitive) basis set. However, further studies are needed in order to investigate the perspectives of this alternative to design more competitive basis sets than the options commonly in use nowadays.

Acknowledgments The authors thank CNPq (Brazilian agency) for financial support. R.B. Viana thanks CAPES for the research fellowship.

References

1. Roothaan CCJ (1951) New Developments in Molecular Orbital Theory. *Rev Mod Phys* 23:69–89. doi:10.1103/RevModPhys.23.69
2. Mohallem JR, Dreizler MR, Trsic M (1986) A Griffin-Hill-Wheeler version of the Hartree-Fock equations. *Int J Quantum Chem Symp* 20:45–55. doi:10.1002/qua.560300707
3. Silver DM, Nieuwpoort WC (1978) Universal atomic basis sets. *Chem Phys Lett* 57:421–422. doi:10.1016/0009-2614(78)85539-0
4. Silver DM, Wilson S, Nieuwpoort WC (1978) Universal basis sets and transferability of integrals. *Int J Quantum Chem* 14:635–639. doi:10.1002/qua.560140510

5. Mohallem JR, Trsic M (1987) A universal Gaussian basis set for atoms Li through Ne based on a generator coordinate version of the Hartree-Fock equations. *J Chem Phys* 86:5043–5044. doi:[10.1063/1.452680](https://doi.org/10.1063/1.452680)
6. da Costa HFM, Trsic M, Mohallem JR (1987) Universal Gaussian and Slater type basis sets for atoms He To Ar based on an integral version of the Hartree-Fock equations. *Mol Phys* 62:91–95. doi:[10.1080/00268978700102071](https://doi.org/10.1080/00268978700102071)
7. da Silva ABF, da Costa HFM, Trsic M (1989) Universal Gaussian and Slater-type bases for atoms H To Xe based on the generator coordinate Hartree-Fock method. I Ground And Certain Low-Lying Excited States Of The Neutral Atoms. *Mol Phys* 68:433–445. doi:[10.1080/00268978900102271](https://doi.org/10.1080/00268978900102271)
8. de Castro EVR, Jorge FE (1998) Accurate universal Gaussian basis set for all atoms of the periodic table. *J Chem Phys* 108:5225–5229. doi:[10.1063/1.475959](https://doi.org/10.1063/1.475959)
9. Jorge FE, de Castro EVR (1999) Improved generator coordinate Hartree Fock method: application to first-row atoms. *Chem Phys Lett* 302:454–460. doi:[10.1016/S0009-2614\(99\)00144-X](https://doi.org/10.1016/S0009-2614(99)00144-X)
10. de Castro EVR, Jorge FE, Pinheiro JC (1999) Accurate gaussian sets for second-row atoms and ions generated with the improved generator coordinate Hartree-Fock method. *Chem Phys* 243:1–7. doi:[10.1016/S0301-0104\(99\)00034-8](https://doi.org/10.1016/S0301-0104(99)00034-8)
11. Jorge FE, Muniz EP (1999) Accurate adapted Gaussian basis sets for the atoms from H through Xe. *Int J Quantum Chem* 71:307–312. doi:[10.1002/\(SICI\)1097-461X\(1999\)71:4<307::AID-QUA2>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-461X(1999)71:4<307::AID-QUA2>3.0.CO;2-4)
12. Jorge FE, de Castro EVR, da Silva ABF (1997) A universal Gaussian basis set for atoms cerium through lawrencium generated with the generator coordinate Hartree-Fock method. *J Comput Chem* 18:1565–1569. doi:[10.1002/\(SICI\)1096-987X\(199710\)18:13<1565::AID-JCC1>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1096-987X(199710)18:13<1565::AID-JCC1>3.0.CO;2-P)
13. Pinheiro JC, da Silva ABF, Trsic M (1997) Generator coordinate Hartree-Fock method applied to the choice of a contracted Gaussian basis for the second-row atoms. *Int J Quantum Chem* 63:927–934. doi:[10.1002/\(SICI\)1097-461X\(1997\)63:5<927::AID-QUA3>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-461X(1997)63:5<927::AID-QUA3>3.0.CO;2-Y)
14. Bunge CF, Barrientos JA, Bunge AV, Cogordan C (1992) Hartree-Fock and Roothaan-Hartree-Fock energies for the ground states of He through Xe. *Phys Rev A* 46:3691–3696. doi:[10.1103/PhysRevA.46.3691](https://doi.org/10.1103/PhysRevA.46.3691)
15. Froese-Fischer C (1977) The Hartree-Fock method for atoms. Wiley, New York
16. Barbosa RC, da Silva ABF (2003) A new proposal for the discretization of the Griffin-Wheeler-Hartree-Fock equations. *Mol Phys* 101:1073–1077. doi:[10.1080/0026897021000044007](https://doi.org/10.1080/0026897021000044007)
17. Mohallem JR (1986) A further study on the discretisation of the Griffin-Hill-Wheeler equation. *Z Phys D* 3:339–344. doi:[10.1007/BF01437189](https://doi.org/10.1007/BF01437189)
18. Partridge H (1987) Near Hartree-Fock quality GTO basis sets for the second-row atoms. *J Chem Phys* 87:6643–6647. doi:[10.1063/1.453450](https://doi.org/10.1063/1.453450)
19. Partridge H (1989) Near Hartree-Fock quality GTO basis sets for the first- and third-row atoms. *J Chem Phys* 90:1043–1047. doi:[10.1063/1.456157](https://doi.org/10.1063/1.456157)
20. Nelder JA, Mead R (1965) A Simplex Method for Function Minimization. *Comput J* 7:308–313. doi:[10.1093/comjnl/7.4.308](https://doi.org/10.1093/comjnl/7.4.308)
21. Clementi E, Chakravorty SJ, Coriנגiu C, Flores JR, Sonnad V (1991) Independent electron models: Hartree-Fock for many electron atoms. In: Clementi E (ed) *Modern techniques in computational chemistry: MOTECC-91*. ESCON, Leiden, pp 23–110
22. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas O, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2004) *Gaussian 03*, revision C.02. Gaussian Inc, Wallingford
23. Davidson ER (1996) Comment on “Comment on Dunning’s correlation-consistent basis sets”. *Chem Phys Lett* 260:514–518. doi:[10.1016/0009-2614\(96\)00917-7](https://doi.org/10.1016/0009-2614(96)00917-7)
24. Lide DR (2000) *CRC Handbook of Chemistry and Physics – a ready-reference book of chemical and physical data*, 81st edn. CRC, Boca Raton