Theor Chem Acc (2005) 113: 69–72
DOI 10.1007/s00214-004-0617-9

Theoretical
Chemistry Accounts

Regular article

Gaussian basis sets for CO_2 molecule generated with the molecular improved generator coordinate Hartree–Fock method

M. T. Barreto, E. P. Muniz, F. E. Jorge, A. G. Cunha

Departamento de Física, Universidade Federal do Espírito Santo, 29060-900 Vitória, ES, Brazil

Received: 11 February 2004 / Accepted: 17 September 2004 / Published online: 28 January 2005

Abstract. The molecular improved generator coordinate Hartree–Fock (MIGCHF) method is used to generate increasing size atom-centered Gaussian basis sets for the CO_2 molecule. From these basis sets total HF energies and second-order correlation energies were calculated and compared with results obtained with other approaches. Considering our largest basis set, the HF energy is in error by 98 μhartree and the second-order correlation energy corresponds to ∼95.6% of an estimate of the limiting value. The relevance of the present calculations is to show the accuracy that can be achieved in studies of small polyatomic molecules with the MIGCHF method.

Keywords: Molecular improved generator coordinate Hartree–Fock method – Gaussian basis sets – CO_2 molecule – HF and second-order correlation energies

1 Introduction

Contemporary ab initio methods for calculating accurate energetics and molecular properties have two major limiting factors: incomplete incorporation of electron correlation and incomplete description of the orbitals due to the use of finite basis sets [1]. Recovering substantial fractions of the correlation energy requires sophisticated wave functions and large basis sets, and obtaining microhartree accuracy for absolute energies is prohibitively expensive for all but the smallest systems. There are procedures for approaching the complete basis set limit by systematically enlarging a basis set, for example: the even-tempered formula [2], the correlation-consistent polarized valence (cc-pVXZ, X = D, T, Q, 5) and the augmented cc-pVXZ (aug-cc-pVXZ) basis sets [3,4], the molecular generator coordinate Hartree–Fock (MGCHF) method [5], and the molecular improved GCHF (MIGCHF) method [6]. In 1979, Schmidt and Ruedenberg [7] proposed a practical scheme for systematically extending basis sets of even-tempered [2] Gaussian functions so as to approach the complete basis set limit. This scheme can be readily adapted for the development of universal even-tempered basis sets and has been used extensively in developing basis sets capable of supporting high precision. Universal basis sets for high-precision molecular applications are necessarily large and, therefore, lead to computationally demanding practical applications. Previous papers [8–10] using universal basis sets of even-tempered Gaussian functions centered not only on the atomic nuclei but also on the bond centers have shown that accuracy of a few microhartree can be achieved for various diatomic systems.

The purpose of this paper is to explore the use of atom-centered Gaussian basis sets (GBSs) in reducing the total basis set truncation error in matrix HF and second-order Møller–Plesset (MP2) calculations for the ground state of the carbon dioxide molecule at its equilibrium geometry.

Previous papers [6, 11–13] in this series have examined the accuracy of the MIGCHF method [6] for diatomic molecules at the HF level, and have demonstrated that accuracy in the total energy of 1 μhartree or less for H_2, HLi, and Li_2 molecules [6], and smaller than 77 μhartree for fourteen electron diatomic systems [11] can be achieved. In Ref. [13], the calculation of the second-order correlation energy for diatomic molecules was considered, and it was shown that 95.7, 88.4, and 96.8% of estimates of the exact second-order energies for the ground state of the N_2, BF, and CO molecules could be supported by suitably constructed atom-centered basis sets. This work, therefore, extends our previous studies to polyatomic molecules by examining, as a prototype linear triatomic system, the CO_2 molecule at both the HF and MP2 levels.

2 Method

In the MGCHF method [5] the molecular orbitals (MO) are integral transforms, i.e.,

$$\psi_i(\gamma) = \sum_n^N \sum_p^P \int f_{inp}(\alpha_{np}) \phi_p(\alpha_{np}; \vec{r}_\gamma - \vec{R}_n) d\alpha_{np}, \qquad (1)$$

Correspondence to: F. E. Jorge
e-mail: jorge@cce.ufes.br

where the index n runs for N atomic nuclei and p for the various s, p, d, … symmetries of the atomic functions ϕ_p; $\vec{r}_\gamma$ is the coordinate of the electron γ and $\vec{R}_n$ of the nucleus n. The indexes for α_{np} admit the possibility of different generator coordinates for different atomic species and symmetries. For the case of equal atoms or universal basis set one could write α_p only.

The variation of the total energy expectation value with respect to the weight function f_{inp} leads to the molecular Griffin–Hill–Wheeler HF (GHWHF) Eq. [5]

$$\sum_n^N \sum_p^P \int F_{np,n'p'}(\alpha_{np}, \alpha_{n'p'}) - \varepsilon_i S_{np,n'p'}(\alpha_{np}, \alpha_{n'p'}) \times f_{inp}(\alpha_{np}) d\alpha_{np} = 0,\ i = 1, \ldots, I, \tag{2}$$

where the ε_i are the orbital energies and the explicit forms of $F_{np,n'p'}$ and $S_{np,n'p'}$ are given in Ref. [5].

The Eqs. (2) are solved through integral discretization (ID) technique [14], in which Eq. (1) becomes

$$\psi_i(\gamma, T) = \sum_n \sum_p \sum_t f_{inp}(\alpha_{npt}) \phi_p(\alpha_{npt}; \vec{r}_\gamma - \vec{R}_n) \Delta\alpha_{npt} \tag{3}$$

and one can interpret

$$C_{inpt} = \Delta\alpha_{npt} f_{inp}(\alpha_{npt}), \tag{4}$$

as the atomic linear combination coefficient in a MO. In order to make numerical integration through discretization efficient, a relabelling of the generator coordinate space was introduced [14] according to

$$\Omega = \ln(\alpha/A), \quad A > 1, \tag{5}$$

where A is a scaling parameter determined numerically. Thus, the coefficient that appears in Eq. (4) becomes

$$C_{inpt} = A\Delta\Omega_{npt} f_{inp}(\Omega_{npt}) \exp(A\Omega_{npt}).$$

The new generator coordinate space, Ω, is discretized for each s, p, d, … symmetry of each atom in an equally spaced mesh $\{\Omega_{np}^k\}$ so that:

$$\Omega_{np}^k = \Omega_{np}^{\min} + (k-1)\Delta\Omega_{np}, \quad k = 1, \ldots, N_{np}. \tag{6}$$

In Eq. (6) N_{np} is the number of discretization points for atom n and symmetry p, $\Omega_{np}^{\min}$ and $\Delta\Omega_{np}$ are respectively the lowest value and the constant increment for the generator coordinate. The values of $\Omega_{np}^{\min}$ and N_{np} are chosen to embrace an adequate integration range for the weight function f_{inp}. From Eq. (6) we can see that the original MGCHF method [5] uses only one arithmetic sequence of equally spaced points $\{\Omega_{np}^k\}$ to generate basis sets.

One may wonder whether the results obtained with the Eqs. (5) and (6) can be improved within the framework of the MGCHF method without adding more functions (Gaussian-type functions (GTFs) in our case). We did this by proposing a simple modification that may produce improvements in the HF wave functions. The idea is to use (when necessary) three arithmetic sequences with the same principal quantum number. This allows to form different distributions for small, intermediate, and large exponents of GTFs.

In this new approach the generator coordinate space, Ω, is discretized for each s, p, d, f,… symmetry in three independent arithmetic sequences [6]:

$$\Omega_{np}^k = \begin{cases} \Omega_{np}^{\min} + (k-1)\Delta\Omega_{np}, & k = 1, \ldots, K_{np} \\ \Omega_{np}^{'\min} + (k-1)\Delta\Omega_{np}^{'}, & k = K_{np}+1, \ldots, M_{np} \\ \Omega_{np}^{''\min} + (k-1)\Delta\Omega_{np}^{''}, & k = M_{np}+1, \ldots, N_{np}. \end{cases} \tag{7}$$

For a given value of N_{np}, the number of parameters to be optimized for each symmetry of each atom is three times that of the original MGCHF method, see Eq. (6)].

Here, we call attention to the fact that when one uses Eq. (7), one does not have equally spaced points $\{\Omega_{np}^k\}$ anymore as occur in Eq. (6), because now three independent arithmetic sequences are used to generate the basis functions exponents for each symmetry of each atom. This methodology to generate primitive GTF exponents in a molecular environment was called the MIGCHF method [6].

At each iteration of the self-consistent field procedure the integrations are implemented numerically. This approach leads formally to the HF Roothaan (HFR) equations, with the advantage of allowing the use of available HFR codes. Nonetheless, the discretization points (exponents) are chosen to preserve the integral character of the molecular GHWHF, see Eqs. (2).

3 Atom-centered basis sets of primitive Gaussian-type functions

As a starting point to construct accurate basis sets for CO_2, the GBSs generated by Librelon and Jorge [15] with the IGCHF method [16] were used for C (23s14p) and O (23s14p). To improve the molecular HF energy, one function of s symmetry and two of p symmetry were added to the C and O atoms and, then, all s and p exponents were reoptimized through the Eq. (7). Next, 8d (for C) and 6d (for O) polarization functions were included in the basis set and then optimized in the molecular environment using only one arithmetic sequence of Eq. (7), since the number of exponents of the d symmetry is small. The last step was repeated for higher angular momentum functions, and sequences of optimized larger atom-centered basis sets were constructed (see Table 1). It is important to say that only the additional basis functions with respect to the previous entry were optimized, and that only the spherical components of the polarization functions were used. For all calculations, the scaling parameter that appears in Eq. (5) has the same value (6.0).

4 Results and discussion

The MIGCHF method presented above has been used to generate accurate GBSs for the CO_2 molecule. The molecule is linear in its ground electronic state, belonging to the $D_{\infty h}$ point symmetry group, and the experimental C=O bond lengths are 1.160 Å[17]. The majority of the calculations reported in this work were carried out with the GAUSSIAN 94 program [18], and the correlation calculations include all electrons. It is known that the computational linear dependence can be measured by the smallest eigenvalue of the overlap matrix. All the basis sets generated in this work have overlap matrix with smallest

Table 1. Ground-state total HF energies (E_{HF}) and second-order correlation energies (E_2) for the CO_2 molecule. All energies are in hartree

Basis set	Number of GTFs	$-E_{\mathrm{HF}}$	$-E_2$	Reference
C(24s16p); O(24s16p)	216	187.600283	0.527368	This work
C(24s16p8d); O(24s16p6d)	316	187.719139	0.742880	This work
aug-cc-pV5Z	381	187.725019	0.775220	Calculated in this work
C(24s16p8d3f); O(24s16p6d3f)	379	187.725038	0.800976	This work
C(28s14p14d14f); O(28s14p14d14f)	476	187.725091	–	[19]
C(30s15p15d15f); O(30s15p15d15f)	510	187.725114	–	[19]
C(24s16p8d3f2g); O(24s16p6d3f2g)	433	187.725283	0.819551	This work
C(24s16p8d3f2g1h); O(24s16p6d3f2g1h)	466	187.725310	0.823861	This work
C(28s14p14d14f); O(28s14p14d14f); CO(24s10p11d)	694	187.725408	–	[19]
MP2-R12/A cc-pV6Z(uc)	483	–	0.861806	Calculated in this work

eigenvalues $\sim 10^{-7}$. Thus, there are no signs of significant problem associated with computational linear dependence for any of these basis sets.

Table 1 displays the ground-state total HF energies (E_{HF}) for CO_2 computed with the GBSs generated with the MIGCHF method [6], with the aug-cc-pV5Z basis sets [4], and with the even-tempered GBSs centered on the atoms [19]. The second-order correlation energies (E_2) are also calculated and compared with the result obtained with the aug-cc-pV5Z basis sets [4]. An estimate of the HF limit (−187.725408 hartree) established by Moncrieff and Wilson [19] is also included in Table 1. Besides this, the PSI3 code [20] is used for the second-order Møller–Plesset perturbation theory calculations with linear terms r_{12} (MP2-R12/A) [21,22]. For the MP2-R12/A calculations, the cc-pV6Z basis set of Wilson et al. [23] in a fully uncontracted form, denoted in the present work as cc-pV6Z(uc), is used to estimate a 'limiting' value of the second-order correlation energy of −0.861806 hartree (see the last row of Table 1).

From Table 1, one can see that the total HF energies calculated with the MIGCHF method [6] decrease monotonically with increasing size of basis set, that is, the energy is found to decrease monotonically on adding the higher harmonic basis functions with the largest decrease of ∼119 mhartree resulting from the addition of the atom-centered d functions. One obtains more stable HF energy with our C(24s16p8d3f); O(24s16p6d3f) basis set (379 primitive GTFs) than with the larger aug-cc-pV5Z basis sets (420 primitive GTFs contracted to 381 functions). Here, it is important to say that the CPU times spent to realize the HF calculations using our basis set and the widely used correlation consistent basis sets of Kendall et al. [4] are ∼28 and 165 min, respectively. The main difference between these two approaches is that the aug-cc-pV5Z basis sets are atom specific, whereas the exponents of our basis set are optimized in the molecular environment with the MIGCHF method. The C(28s14p14d14f); O(28s14p14d14f) and C(30s15p15d15f); O(30s15p15d15f) basis sets [19] and the smaller C(24s16p8d3f2g); O(24s16p6d3f2g) and C(24s16p8d3f2g1h); O(24s16p6d3f2g1h) basis sets generated with the MIGCHF method [6] have associated errors of 317, 294, 125, and 98 μhartree, respectively. From the results presented above, it is clear that employing only atom-centered basis sets of GTFs the MIGCHF method gives the best HF energies.

We recall that the lowest matrix HF energy (−187.725408 hartree) so far reported in the literature was obtained with a basis set containing functions centered on both the atomic and bond centers, namely: C(28s14p14d14f); O(28s14p14d14f); CO(24s10p11d) (694 GTFs) [19]. The idea of using bond functions in molecular structure calculations is an old one [24]. Indeed, a number of authors have advocated the use of basis functions located at the bond center in both SCF calculations and in calculations taking account of correlation effects [25, 26]. Such functions were introduced to provide the same effects as the addition of a set of polarization functions, but at a lower cost.

The first two E_2 results presented in Table 1 show clearly that the addition of d-type polarization functions yields the largest energy decrease in the correlation energy. The next two studies summarized in this table indicate that the second-order correlation energy evaluated with the MIGCHF method [6] is ∼26 mhartree lower than the result obtained with the larger aug-cc-pV5Z basis set [4]. We recall that all the GTF exponents of our C(24s16p8d3f); O(24s16p6d3f) basis set were optimized at the HF level, and that the aug-cc-pV5Z basis set (which contains up to *h* functions) is designed for valence correlation calculations. Even so, it is notable that the basis set generated with the MIGCHF method reproduces all electron correlation energy better than aug-cc-pV5Z. Helgaker et al. [27] investigating the basis set convergence of the correlation energy in the water molecule verified that the correlation-consistent polarized core-valence (cc-pCVXZ) basis sets [28] give all electron correlation energies closer to the MP2-R12 results than the corresponding values calculated with the cc-pVXZ basis sets. One knows that for calculations of electron correlation effects it is necessary to include functions of higher angular symmetry in the basis set. The best E_2 value (−0.823861 hartree) presented in Table 1 is obtained with our C(24s16p8d3f2g1h); O(24s16p6d3f2g1h) atom-centered basis set, which corresponds to ∼95.6% of the limiting value. The results of the second-order correlation energies confirm the good accuracy that can be achieved with the MIGCHF method [6]

Table 2. Comparison of the occupied molecular orbital (MO) energies (in hartree) calculated with Gaussian basis sets (GBSs)

Orbital	$-\varepsilon$ (GBS)[a]	$-\varepsilon$ (GBS)[b]
$1\sigma_g^2$	20.64939	20.64939
$1\sigma_u^2$	20.64934	20.64935
$2\sigma_g^2$	11.45889	11.45889
$3\sigma_g^2$	1.53106	1.53104
$4\sigma_g^2$	1.47790	1.47788
$2\sigma_u^2$	0.80078	0.80081
$3\sigma_u^2$	0.74385	0.74388
$1\pi_u^4$	0.71610	0.71610
$1\pi_g^4$	0.54504	0.54502

[a] MO energies obtained with our largest GBS C(24s16p8d3f2g1h);O(24s16p6d3f2g1h)

[b] MO energies obtained with the largest GBS of the Ref. [19] C(28s14p14d14f);O(28s14p14d14f);CO(24s10p11d)

in the generation of total wave functions for polyatomic molecules.

A comparison of the occupied MO energies calculated with the largest basis sets reported in this work and in Ref. [19] is made in Table 2. From this table, it is clear that the largest difference between corresponding MO energies computed with these two basis sets is only three units in the last figure.

An extensive discussion about the main differences between the MIGCHF method [6] and the algebraic approximation used by Moncrieff and Wilson [19] is included in Ref. [11].

5 Conclusions

Using the MIGCHF method [6], which has previously been shown to be capable of providing accurate total wave functions for the ground states of some diatomic systems [6, 11–13], we have carried out HF and MP2 calculations for the ground state of the carbon dioxide molecule, which we consider as a prototype for polyatomic systems.

Considering only atom-centered basis sets of GTFs, the accuracies achieved in this work at the HF level are better than the corresponding ones obtained with the basis sets generated by Kendall et al. [4] and by Moncrieff and Wilson [19]. The energy computed with our C(24s16p8d3f2g1h); O(24s16p6d3f2g1h) basis set (466 GTFs) is $98\,\mu$hartree above an estimate of the HF limit [19] obtained with the larger C(28s14p14d14f); O(28s14p14d14f); CO(24s10p11d) basis set (694 GTFs).

The second-order correlation energy evaluated in the present paper with the C(24s16p8d3f); O(24s16p6d3f) basis set is better than the result obtained with the widely used larger aug-cc-pV5Z basis sets [4]. The second-order correlation energy evaluated with our largest basis set (−0.823861 hartree) accounts for ~95.6% of the limiting value.

Thus, considering only atom-centered basis sets of GTFs, we believe that the degree of accuracy achieved in this work with the MIGCHF method at the HF and MP2 levels is among the best so far presented in the literature, and that this method is appropriate to realize calculations of electronic structures of small polyatomic molecules.

Acknowledgements. We acknowledge the financial support of CNPq (Brazilian Agency). We employed computational facilities at Universidade Federal do Espírito Santo and Universidade Estadual Paulista (IQ Araraquara).

References

1. Jensen F (1998) Introduction to computational chemistry. Wiley, New York
2. Reeves CM (1963) J Chem Phys 39:1
3. Dunning TH Jr (1989) J Chem Phys 90:1007
4. Kendall RA, Dunning TH Jr, Harrison RJ (1992) J Chem Phys 96:6796
5. Da Costa HFM, da Silva ABF, Mohallem JR, Simas AM, Trsic M (1991) Chem Phys 154:379
6. Jorge FE, Centoducatte R, de Castro EVR (2000) Theor Chem Acc 103:477
7. Schmidt MW, Ruedenberg K (1979) J Chem Phys 71:3951
8. Kobus J, Moncrieff D, Wilson S (1994) J Phys B 27:2867
9. Kobus J, Moncrieff D, Wilson S (1994) J Phys B 27:5139
10. Moncrieff D, Wilson S (1998) J Phys B 31:3819
11. Canal Neto A, Jorge FE, Centoducatte R (2002) Int J Quantum Chem 90:244
12. Barreto MT, Muniz EP, Jorge FE, Centoducatte R (submitted)
13. Jorge FE, Canal Neto A (2002) J Mol Struct (Theochem) 589–590:359
14. Mohallem JR (1986) Z Phys D 3:339
15. Librelon PR, Jorge FE (2003) Int J Quantum Chem 95:190
16. Jorge FE, de Castro EVR (1999) Chem Phys Lett 302:454
17. Mills IM (1974) Theor Chem (Specialist Periodical Reports 1) 1:110
18. Frisch MJ, Trucks GW, Schlegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Keith TA, Peterson GA, Montgomery JA, Raghavachari K, Al-Laham MA, Zakrzewski VG, Ortiz JV, Foresman JB, Cioslowski J, Stefanov BB, Nanayakkara A, Challacombe M, Peng CY, Ayala PY, Chen W, Wong MW, Andres JL, Replogle ES, Gomperts R, Martin RL, Fox DJ, Binkley JS, Defrees DJ, Baker J, Stewart JP, Head-Gordon M, Gonzalez C, Pople JA (1995) Gaussian 94, Revision A.I, Gaussian Inc., Pittsburgh
19. Moncrieff D, Wilson S (1995) J Phys B 28:4007
20. Crawford TD, Sherrill CD, Valeev EF, Fermann JT, King RA, Leininger ML, Brown S, Janssen CL, Seidl ET, Kenny JP, Allen WD (2003) psi 3.2, PSITECH Inc., Watkinsville
21. Klopper W, Kutzelnigg W (1987) Chem Phys Lett 134:17
22. Kutzelnigg W, Klopper W (1991) J Chem Phys 94:1985
23. Wilson AK, van Mourik T, Dunning TH Jr (1996) J Mol Struct (Theochem) 388:339
24. Preuss H (1964) Z Naturf a 19:1335
25. Rothenberg S, Schaefer HF III (1971) J Chem Phys 54:2765
26. Mach P, Kysel O (1985) J Comput Chem 6:312
27. Helgaker T, Klopper W, Koch H, Noga J (1997) J Chem Phys 106:9639
28. Woon DE, Dunning TH Jr (1995) J Chem Phys 103:4572