

Regular article

# On the influence of the basis set superposition error on calculated vibrational frequencies

Jose Manuel Hermida-Ramón<sup>1</sup>, Gunnar Karlström<sup>1</sup>, Bengt Nelander<sup>2</sup>

<sup>1</sup>Theoretical Chemistry, Chemical Center, P.O. Box 124, 22100 Lund, Sweden

<sup>2</sup>Chemical Physics, Chemical Center, P.O. Box 124, 22100 Lund, Sweden

Received: 5 March 2002 / Accepted: 13 August 2002 / Published online: 6 October 2003  
© Springer-Verlag 2003

**Abstract.** The relation between the so called basis set superposition error and intramolecular vibrational frequencies calculated at the Hartree Fock SCF level of approximation was investigated. A linear conformation of HF dimer was chosen as test system for the investigation. It was found that the direct basis set superposition error for the studied system is rather small. It was further found that the shifts are mainly determined by the geometry parameters of the system.

**Keywords:** Infrared frequencies – Intermolecular complex – Geometry influence

## Introduction

The study of intermolecular interactions between small and medium sized molecules is a research field where both theoretical and experimental techniques can provide valuable information about the properties of a system being studied.

Infrared spectroscopy has proven to be a very valuable experimental tool in the characterization of complexes. By studying the shifts of intramolecular vibrational frequencies due to complex formation one can obtain information about the structure of the complex formed. The weakness of such an approach is that the interpretation of the experimental results is often difficult.

Quantum chemical *ab initio* calculations provide a powerful tool for directing theoretical prediction of the structure of a formed molecular complex. The major

weakness of this approach is that for systems with a moderate complexity the calculations needed are often so computer demanding that accurate predictions are not possible today. For such systems one often ends up in a situation where several structures are equally stable within the numerical accuracy of the computations.

To obtain a better predictive capacity, it can often be fruitful to combine theoretical and experimental techniques and calculate the experimentally observed spectroscopic data directly and compare the theoretical spectrum with the experimental one. By such an approach, it is often possible to obtain results of high reliability with a moderate theoretical and experimental effort. As an extra bonus, it is possible to use the theoretical calculations to obtain an understanding of why a given vibrational frequency is shifted with a precise wavenumber. This requires, however, that the calculations are accurate and reflect real physical interactions. It is well known that the study of intermolecular interactions is haunted by a methodological error known as the Basis Set Superposition Error [1] (BSSE), which originates from that the quality of the basis set used in the quantum chemical calculations varies with the geometry of the system. To remedy this problem Boys and Bernardi [2] suggested that not only the complex should be studied with the basis set of the interacting molecules, but that also the non-interacting molecules should be studied with this basis set and that the interaction energy should be calculated as the difference between these energies obtained with the same basis set. This method is named the Counterpoise (CP) method and is generally accepted as the best method for correcting this error[1]. It has been shown that this error affects both the system interaction energy as well as the structure of the complex formed and the electrostatic properties of the interacting molecules. The influence of this type of error on vibrational frequencies has, however, not been studied in a systematic way, despite the fact that some authors [3, 4] have actually used CP corrected potential energy surfaces to calculate the

Contribution to the Björn Honorary Issue

Correspondence to: J. M. Hermida-Ramón  
e-mail: ramon@signe.teokem.lu.se

vibrational frequencies. The purpose of this work is to gain some insight into this problem.

### Method and computational details

In order to carry out the project outlined above, a suitable system must be found. The following three requirements should preferably be fulfilled:

1. The system should be so small that a basis set that is effectively BSSE free can be used for the system.
2. The intermolecular interaction between the molecules in the system should be so strong that significant effects can be observed.
3. It should be easy to perform geometry optimization on the system.

The HF dimer in a linear conformation was chosen as a test case which fulfills these requirements. All calculations were done at SCF level and several four different basis sets were studied. We have chosen not to use a method describing the correlation of the electron motion for several reasons.

1. The basis set requirements for correlated calculations are much larger than for SCF calculations. This would significantly increase the computational resources needed for the work.
2. Systems that are of real interest are normally so large that a SCF procedure is preferred in order to keep the computational costs at a reasonable level. In the future this is likely to change.
3. The use of a correlated method would have made the analysis more difficult and is not likely to change the observed behavior. In particular, we noted that problems with near linear dependencies in the one particle basis set are likely to occur and to create severe problems for the analysis.

All calculations were performed with the MOLCAS.5.2 package [5]. Four different basis sets were investigated. They were all constructed from the same primitive ANO [6] type basis sets —14s,9p,4d,3f,1g basis set was used for fluorine and a 8s,4p,3d,1f basis set was used for hydrogen.

1. The smallest basis set is contracted to [4s3p1d] for fluorine and to [3s1p] for hydrogen. This basis set was named small (S).
2. The second smallest basis set contracted to [5s4p2d] for fluorine and to [4s2p] for hydrogen and was named medium (M).
3. The second biggest basis set uses a very weak contraction to [7s7p4d3f1g] for fluorine and [7s4p3d1f] for hydrogen. This basis set was named large (L).
4. In the biggest basis set, the primitive basis set was augmented with one extra g function for fluorine (exponent 0.9) and one extra f primitive for hydrogen (exponent 0.6). This primitive basis was contracted to [9s8p4d3f2g] for fluoride and to [7s4p3d2f] for

hydrogen. This means that the dfg basis functions for fluorine and the pdf basis functions were left uncontracted. The extra flexibility for the fluorine sp basis set was obtained by allowing the s-exponents 0.29086165 and 0.07810237 as well as the p-exponent 0.07810237 to be varied independently. This basis set is named extended (E).

The geometry optimizations on the linear form of the HF dimer and of the monomer were done analytically. The standard procedure to evaluate vibrational frequencies from a potential energy surface is to use the harmonic approximation. In this approximation the vibrational frequencies only depend on the harmonic force constants. Due to technical problems, two different strategies were followed to calculate these force constants and the corresponding vibrational frequencies. For the two smaller basis sets labeled small and medium it was possible to use the analytical Hessian construction scheme available in the MOLCAS package. However, due to near linear dependence in the two biggest basis sets, this was not possible for the two largest basis sets. In stead, we were forced to calculate the Hessian by use of a numerical approach. This approach was checked against the analytical approach for one of the small basis sets where no problem with linear dependencies were expected. The two procedures yielded identical frequencies. The numerically fitted polynomials are presented in the Appendix.

For each basis set six different Hessians were evaluated. They were

1. The Hessian for the unperturbed HF molecule with the basis set of the HF molecule.
2. One Hessian for the HF molecule with the geometry of the donor in the complex and one Hessian for the acceptor geometry. The basis set of the HF molecule was used for these two Hessians.
3. One Hessian for the HF molecule with the geometry of the donor in the complex and one Hessian for the acceptor geometry. The composite basis set of the donor and acceptor was used in the calculations to obtain these two Hessians.
4. One Hessian for the optimized complex structure was calculated with the composite basis set.

The Hessians 1, 2 and 4 have been analyzed using Wilson's method for G and F matrices the so called GF [7] theory to yield frequencies. For a definition of the G and F matrices see Ref. [7]. The analysis of the Hessian for the donor or acceptor according to 3, obtained using the composite basis set of the donor and the acceptor is somewhat more complicated. The origin of the problem is that no mass can be assigned in a natural way to the sites where the basis functions of the ghost atoms are located, and that the Hessian matrix formally depends on the degrees of freedom also for the Ghost atoms. Here we solved this problem by assigning a very large mass ( $10^5$ ) to these sites. We used a similar trick when we analyzed the Hessian for the real complexes. This means

actually three different sets of frequencies were obtained from the Hessian of the complexes. One where the actual masses were used, one where the masses of the donor molecule atoms were given a very large value and one where the masses of the acceptor were given these large values. The advantage of this trick is that we can identify the so called kinematic coupling occurring when two vibrational frequencies are close.

Due to the simple structure of the system studied only two frequencies corresponding to intramolecular vibrations exist. The frequencies obtained were analyzed in the following way. The difference between the frequencies obtained for the optimized monomer structure and for the monomers studied at the structure in the complex is only due to the change in geometry since the same basis set is used for both calculations. It is then possible to calculate the differences between the frequencies obtained for the monomers at the dimer geometry with the monomer and also with the composite basis sets. This difference is the true BSSE contribution to the vibrational frequencies. Unfortunately this difference depends slightly on the masses assigned to the sites with ghost basis functions. Test calculations we performed show, however, that this effect is very small and we estimate it by assigning very large masses ( $10^5$  a.u.) to the sites with ghost basis functions. Calculations with these very large masses and the Hessians 3 and 4 will be denoted by  $\text{dimer} + \infty + \text{CP}$  and  $\text{dimer} + \infty$ , respectively.

## Results

In Table 1 we present data obtained at the Hartree Fock level of approximation together with the corresponding geometry parameters for the HF dimer. The amount of data is fairly large and probably deserves some explanation. At the top of the Table (line 1) we give the energies obtained for the optimized complex (a linear structure is assumed). This is followed (line 2) by the energies obtained for the optimized free HF molecule.

The donor and acceptor energies that then follow (lines 3 and 4) are the energies obtained for the donor and acceptor for the geometries of these systems in the complex. The basis set of the HF monomer is used. If the basis set of the composite system is used and one optimizes the HF monomer geometry one obtains the next two entries (lines 5 and 6) labeled in the Table donor energy CP and acceptor energy CP. Note that the energies for the HF donor and acceptor molecules with the geometry in the complex are not given in the Table. They can however be calculated from the BSSE for the donor and acceptor given further down in the Table. The structural information that then follows is structured in the following way. First we specify the bond length of the free HF molecule (line 7). This is followed by the optimized bond lengths obtained for the free HF donor and acceptor in the composite basis set (lines 8 and 9). The geometrical information section is completed by giving the structures for the donor and acceptor in the complex (lines 10 and 11). Finally, we give information relevant to the complex formation. First we present the BSSE for the donor, acceptor and for the entire system (lines 12–14). This is followed by the complex formation energy and the BSSE corrected complex formation energy line (15 and 16). The relaxation energies for the donor and acceptor are then given (line 17) and at the very end of the Table we present the interaction energy corrected for the both the BSSE and the relaxation energy (line 18).

The data presented in Table 1 clearly indicates that the two largest basis sets yield effectively the same geometrical and energetic information. The apparent inconsistency that the BSSE-corrected complex formation energy is larger for the largest basis set than the non corrected interaction energy is due to that the relaxation energy for the monomers are not included in this estimate. The observed behavior is thus due to the relaxation energies being larger than the actual BSSE. The so called corrected complex energies (labeled with corr) in Table 1

**Table 1.** Total energies (+200) in a.u. and geometries in a.u. for the linear HF dimer and energies for the monomers (+100) obtained with the different basis sets. The Basis set superposition error as well as the complex formation energies are given in kcal/mol. Distances are given in Bohr

	Small	Medium	Large	Extended
Complex energy	-0.1388665	-0.1427161	-0.1468007	-0.1468160
Energy for free HF	-0.0667063	-0.0687461	-0.0708666	-0.0708783
Donor energy	-0.0667031	-0.0687410	-0.0708578	-0.0708695
Acceptor energy	-0.0667023	-0.0687438	-0.0708638	-0.0708756
Donor energy CP	-0.0671152	-0.0690727	-0.0708649	-0.0708727
Acceptor energy CP	-0.0668572	-0.0687921	-0.0708718	-0.0708799
HF-dist free molecule	1.7028	1.6973	1.6954	1.6954
HF-dist donor CP	1.7012	1.6964	nq	nq
HF-dist acceptor CP	1.7026	1.6973	nq	nq
HF-dist donor	1.7057	1.7010	1.7004	1.7004
HF-dist acceptor	1.7061	1.6998	1.6982	1.6982
BSSE donor	0.2540	0.2034	0.0045	0.0020
BSSE acceptor	0.0945	0.0288	0.0040	0.0027
BSSE total	0.3485	0.2322	0.0085	0.0047
Complex energy	-3.4225	-3.2780	-3.1800	-3.1749
Complex energy (CP)	-3.0785	-3.0506	-3.1778	-3.1774
Donor relax energy	0.0020	0.0032	0.0055	0.0055
Acceptor relax energy	0.0025	0.0015	0.0017	0.0017
Corr complex energy	-3.0745	-3.0459	-3.1706	-3.1702

**Table 2.** Frequencies obtained with the small basis set (S) (see text). BSSE = 0.3485 kcal/mol. Bond distances (Bohr): donor 1.7057; acceptor 1.7061; optimized monomer 1.7028

	$\nu_{\text{Donor}}$	$\nu_{\text{Acceptor}}$
Dimer	4433.0	4456.7
Dimer + $\infty$	4437.1	4452.1
Monomer (dim. geom.)	4449.4	4452.4
Dimer + $\infty$ + CP	4442.9	4450.2
BSSE	6.5	2.2
Opt dimer + $\infty$ + CP	4481.7	4479.7
Optimized monomer	4477.7	4477.7

**Table 3.** Frequencies obtained with the medium basis set (M) (see text). BSSE = 0.23217 kcal/mol. Bond distances (Bohr): donor 1.7010; acceptor 1.6998; optimized monomer 1.6973

	$\nu_{\text{Donor}}$	$\nu_{\text{Acceptor}}$
Dimer	4422.8	4447.2
Dimer + $\infty$	4425.0	4444.5
Monomer (dim. geom.)	4436.4	4446.6
Dimer + $\infty$ + CP	4433.2	4445.7
BSSE	3.2	0.9
Opt dimer + $\infty$ + CP	4467.8	4472.9
Optimized monomer	4468.5	4468.5

**Table 4.** Frequencies obtained with the large basis set (L) (see text). BSSE = 0.00952 kcal/mol. Bond distances (Bohr): donor 1.7004; acceptor 1.6982; optimized monomer 1.6954

	$\nu_{\text{Donor}}$	$\nu_{\text{Acceptor}}$
Dimer	4420.9	4451.6
Dimer + $\infty$	4423.0	4449.6
Monomer (dim. geom.)	4432.5	4450.9
Dimer + $\infty$ + CP	4432.5	4450.9
BSSE	0.0	0.0
Optimized monomer	4474.8	4474.8

are the interaction energies obtained when both the correction for the BSSE and the relaxation is applied. The results presented in Table 1 show that the BSSE error is largest for the donor when the S or the M basis set are used. When the two larger basis sets are used the BSSE error is very small and almost equal for the donor and the acceptor. In Table 1, we also present data showing the geometries obtained for the donor and acceptor when a geometry optimization is performed in the composite basis for that molecule. In order to prevent the studied molecule from moving into the empty basis functions of the ghost system, the optimization was performed with the constraints that the distance between the two ghost atoms are fixed and that the distance from the center of mass for the molecule studied to the ghost atoms are fixed. This optimization is only performed for the two smallest basis sets since we did not expect to find any effect for the bigger ones. This is supported by the fact that the two larger basis sets yields the same geometry for the HF monomer. As can be seen from Table 1,

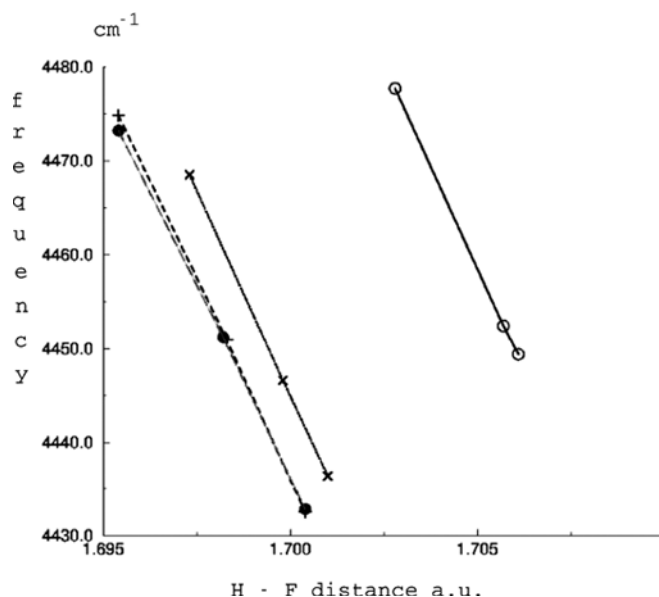
the ghost basis functions tend to shorten the HF distance bringing it closer to the HF bond length obtained for the free HF molecule with the larger basis sets.

In Tables 2, 3, 4, 5 we present the calculated frequencies obtained for the different basis sets. The structure of the data in Tables 2, 3, 4, 5 is as follows. First we give the frequencies calculated for the complex (line 1). This is followed by frequencies calculated with the Hessian of the complex but where the masses of the non-studied molecule are given very large values (line 2). The difference between the complex frequencies and these numbers can be interpreted as a kinematic coupling. The third entry in the Table gives the frequencies obtained for the monomers with the monomer basis set but at the geometry in the complex. This is followed by frequencies calculated for the same geometry but with the composite basis set and very large masses (line 4). The contribution to the vibrational frequencies from the BSSE is presented in line 5. In the following line, in Tables 2 and 3, we give the frequencies calculated with the composite basis set at the geometry optimized with the composite basis set. This line only occurs in Tables 2 and 3. At the end of the tables, we present the frequencies obtained for the optimized monomer. Six observations can be made from the data in Tables 2, 3, 4, 5.

1. The frequencies calculated with the two largest basis sets are almost identical and this indicates that the basis set limit results at the Hartree-Fock level of approximation is reached
2. For a given basis set the vibrational frequency depends mainly on the bond length. This is illustrated in Fig. 1 where the calculated vibrational frequencies for the HF monomer are shown as a function of the H-F bond length for each of the four studied basis sets. Figure 1 indicates that the shifts varies almost linearly with the HF bond length and that the frequency changes with the bond length in almost the same way for all studied basis sets. This indicates that the anharmonicity of the potential functions are also well described with the smallest basis sets.
3. A kinematic coupling, given by the difference between the dimer frequencies and the dimer +  $\infty$  frequencies, lowers the donor frequency and raises the acceptor frequency by 2 to 4 wavenumbers.
4. Large BSSE corresponds to large basis set superposition error contributions to the vibrational frequencies.

**Table 5.** Frequencies obtained with the extended basis set (E) (see text). BSSE = 0.00467 kcal/mol. Bond distances (Bohr): donor 1.7004; acceptor 1.6982; optimized monomer 1.6954

	$\nu_{\text{Donor}}$	$\nu_{\text{Acceptor}}$
Dimer	4420.7	4451.6
Dimer + $\infty$	4422.9	4449.5
Monomer (dim. geom.)	4431.9	4450.3
Dimer + $\infty$ + CP	4431.9	4450.3
BSSE	0.0	0.0
Optimized monomer	4474.2	4474.2



**Fig. 1.** Calculated vibrational frequencies for the different basis sets as a function of H-F distance. The symbols used are  $\circ$  small basis set,  $\times$  medium basis set,  $+$  large basis set and  $\bullet$  extended basis set

- The extra basis functions of the ghost basis set lowers the calculated vibrational frequencies and brings them closer to the basis set limit results for the relevant geometry. Larger shifts are found for HF molecules where large BSSE energies are found. The effect is larger for the donor molecule, since for the donor the light H atom, which is the particle mainly moving during the vibration, experiences the influence of the ghost basis set from the acceptor molecule. The effect varies from 6 wavenumbers to 0 from the smallest to the largest basis set.
- When the geometry of the donor and acceptor is optimized in the composite basis set, this results in a geometry closer to the large basis set geometries. Thus one obtains shorter bond lengths and higher frequencies.

## Conclusions

Calculations were done with four different basis sets on the two HF molecules in the linear HF dimer to study which effects influence the frequency shifts.

It was found that the main source for the shifts is that induced by the change on bond length through the anharmonicity of the potential energy surface. The kinematic coupling gives rise to a small but significant effect and lowers the donor frequency and raises the acceptor frequency by a few wavenumbers.

The BSSE introduces a red shift for the studied frequency. The red shift brings the vibrational frequency obtained with the studied basis set closer to the basis set limit results and can be related to the BSSE energy.

*Acknowledgements.* J. M. H.-R. wishes to thank the "Ministerio de Educación, Cultura y Deporte" for the award of a research grant.

## Appendix 1

Energy calculations to obtain the frequencies numerically were carried out. The intramolecular geometry of both monomers and the intermolecular geometry were modified and the energy calculated (energy results will be provided on request to the authors). These energies were fitted to quadratic equations. Considering  $r_1$  as the acceptor bond distance,  $r_2$  as the donor bond distance and  $r_3$  as the intermolecular distance, the quadratic equations for the extended basis set are:

$$E = -200.146816 + 0.358558r_1^2 + 0.355497r_2^2 + 0.002844r_3^2 - 0.001604r_1r_2 - 0.001061r_1r_3 + 0.004024r_2r_3$$

for the dimer.

$$E = -100.070880 + 0.358635r_1^2 + 0.000000r_2^2 + 0.000001r_3^2 + 0.000000r_1r_2 + 0.000001r_1r_3 - 0.000001r_2r_3$$

for the acceptor with the dimer  $+\infty$  + CP conditions.

$$E = -100.070873 + 0.000001r_1^2 + 0.355680r_2^2 + 0.000001r_3^2 + 0.000000r_1r_2 + 0.000000r_1r_3 - 0.000002r_2r_3$$

for the donor with the dimer  $+\infty$  + CP conditions.

$$E = -100.070876 + 0.358643r_1^2$$

for the isolated acceptor.

$$E = -100.070870 + 0.355682r_2^2$$

for the isolated donor.

The last four equations (the equations for the monomers) have linear terms, but they have no influence on the value of the frequencies if an harmonic approximation is used.

Finally, the equation for the optimized monomer is:

$$E = -100.070878 + 0.362505r^2$$

where  $r$  is the intramolecular bond distance.

From these equations, the frequencies of the E basis set were obtained using a harmonic approximation. Similar equations were obtained for the M and L basis set. These equations give the same frequencies as those provided by the MOLCAS5.2. program for the M basis set and the ones shown in Table 4 for the L basis set.

## References

- van Duijneveldt FB (1997) In: Scheiner S (ed) Molecular interactions. Wiley, New York, p 81
- Boys EB, Bernardi F (1970) Mol Phys 19: 553
- Simon S, Duran M, Dannenberg JJ (1996) J Chem Phys 105: 11024
- Daza MC, Dobado JA, Molina-Molina J, Salvador P, Duran M, Villaveces JL (1999) J Chem Phys 110: 11806A

5. Andersson K, Baryz M, Bernhardsson A, Blomerg MRA, Boussard P, Cooper DL, Fleig T, Fülcher MP, Hess B, Karström G, Lindh R, Malmqvist P.-Å, Neogrady P, Olsen J, Ross BO, Sadlej AJ, Schimmelpfenning B, Schtz M, Seijo L, Serrano L, Seigbahn PE, Thorsteinsson T, Veryazov V, Wahlgren U, Widmark P.-O (2001) MOLCAS, version 5.2. Department of Theoretical Chemistry, Lund University, Sweden
6. Widmark P.-O, Malmqvist P.-Å, Ross BO (1990) *Theor Chim Acta* 77: 291
7. Wilson EB, Decius JC, Cross PC (1955) In: *Molecular vibrations*. McGraw-Hill, New York