Comparison of Fundamental and Harmonic Frequencies of First-Row Closed-Shell Diatomics Calculated Using Full ab Initio Methods and Composite Methods

Ryan P. A. Bettens*

Department of Chemistry, National University of Singapore, 3 Science Drive 3, Singapore 117543

Received: July 3, 2003; In Final Form: December 22, 2003

The fundamental frequencies of vibration of 12 first-row closed-shell diatomics have been predicted using both full and composite levels of CCSD(T) ab initio theory. For a given CPU time budget, composite levels of theory were found to predict harmonic frequencies significantly better than full ab initio methods. However, little improvement was obtained in the computation of the anharmonic correction with composite methods, its being already well predicted at the CCSD(T) small basis set level. It was found that for a given CPU time budget the most accurate fundamental frequencies are obtained by performing a calculation of the harmonic frequencies using a composite method where the optimal choice of basis sets involved a larger cc-pV*X*Z basis set only 1 greater in the valence designation compared with the smaller basis set. An anharmonic correction computed using a small basis set or a low-level composite method could then be added to these harmonic frequencies. The implication of these findings is that accurate fundamental frequencies can be computed cheaply and efficiently by first computing the harmonic frequencies using an accurate composite method and then correcting these frequencies with an anharmonic correction obtained by solving the nuclear Schrödinger equation by some means on a potential energy surface generated using a CCSD(T) small basis set or a low-level composite method.

Introduction

Composite methods have enjoyed significant success in recent years in predicting energies of molecules (see ref 1 and references therein). Such methods approximate high levels of ab initio theory by performing a series of lower-level calculations. However, little work has been done regarding the accuracy of these methods for the determination of other molecular properties. Collins and co-workers have successfully utilized composite methods in the construction of potential energy surfaces for reactive systems. (See, for instance, the reaction of H + H_2O.[2]) Such methods have also very recently been used in a bound-state problem for the determination of zero-point energies and ground-state rotational constants.[3] However, no work has appeared that utilizes composite methods for the calculation of fundamental frequencies.

The computation of vibrational frequencies has seen much interest in recent years,[4] with frequencies determined on average to within 8 cm^{-1} using CCSD(T) and large basis sets.[5] However, the CPU time of CCSD(T) calculations scales as $O(N^7)$, where N is the number of basis functions. Thus, if the success of composite methods for computing energies can be carried over into the calculation of frequencies, then significantly larger molecular systems can be studied with high accuracy. However, to obtain such high accuracy, at the very least core−valence correlation should be incorporated within the CCSD(T) calculation.

It is not the purpose of the work presented here to evaluate the accuracy of a given ab initio calculation with respect to experiment but rather to investigate the accuracy of composite methods in predicting fundamental and harmonic frequencies determined from a full high-level ab initio calculation. Here, the harmonic and fundamental frequencies of 12 closed-shell first-row diatomics are studied using both full and composite ab initio methods. Diatomics serve as simple test cases because the fundamentals can be calculated simply and quickly via variational methods. This paper is set out as follows. The computational details follow this Introduction, after which the results are presented and discussed. The final section provides the conclusion.

Computational Details

Quantum chemical calculations were performed at the following levels of theory: CCSD(T)/cc-pV*X*Z, where X = D, T, Q, and 5 and at the CCSD(T)/6-311++G(2df,2pd) and CCSD(T)/6-311G(d,p) levels. In the case of hydrogen fluoride, aug-cc-pV*X*Z were also utilized. The calculations were carried out using the MOLPRO 2000.1 suite of programs.[6] The composite energies were based on the ad hoc expression $E_{\mathrm{L/S}} = E[\mathrm{CCSD(T)/S}] + E[\mathrm{MP2/L}] - E[\mathrm{MP2/S}]$,(1) where S and L refer to small and large basis sets, respectively. $E_{\mathrm{L/S}}$ is an approximation to the E[CCSD(T)/L] level of theory. In this work, the L basis sets were chosen as cc-pV*X*Z, where X = 5, Q, T, and the 6-311++G(2df,2pd) levels. The S basis sets were cc-pV*X*Z, where X = D and T and 6-311G(d,p). Equation 1 is analogous to the expression used by Pople et al. in their G3XMP2 method.[1] The shorthand designations used in this work to describe the above calculations are 5, Q, ++, T, 311, D for the CCSD(T)/cc-pV5Z, CCSD(T)/cc-pVQZ, CCSD(T)/6-311++G(2df,2pd), CCSD(T)/cc-pVTZ, CCSD(T)/6-311G(d,p), and CCSD(T)/cc-pVDZ, respectively. A composite level of theory is designated as L/S, where L and S are shorthand notations of the above basis sets (i.e., 5, Q, ++, T, 311, and D). Note that core−valence correlation was not taken into account in this work.

The PES for the diatomics was determined by incrementing the bond lengths by small amounts 40 times from some compressed value relative to r_e. This produced a total of 41

---

* E-mail: chmbrpa@nus.edu.sg. Phone: +65 6874 2846. Fax: +65 6779 1691.

10.1021/jp035924y CCC: $27.50

TABLE 1: Experimental and Experimental − Calculated[a] Fundamental Frequencies in cm^{-1}

molecule	exptl[b]	5	Q	++	T	311	D
H_2	4159	−4	−3	−4	−7	−18	14
LiH	1360	14	13	14	10	−3	32
BH	2269	5	6	6	14	21	22
HF	3959	−15	−27	−43	−43	−84	−8
Li_2	346	5	5	8	6	4	10
LiF	894	10	3	12	0	−50	−63
BeO	1464	11	14	25	29	55	111
BF	1379	7	4	4	−2	27	79
C_2	1827	−7	−4	5	7	22	25
CO	2143	4	5	9	15	5	25
N_2	2330	−2	2	16	12	18	20
F_2	894	−9	−4	13	−1	148	136
mad[c]	−	8	7	13	12	38	37
rms[d]	−	9	10	17	17	56	61

[a] See text for a description of the level of theory. [b] All experimental frequencies are from ref 7. [c] Mean absolute deviation. [d] Root mean square.

energies that ranged in value on either side of r_e up to about the energy of the 10th vibrational state. An *n*th order polynomial was then least-squares fit to the energies, whereby *n* was as small as possible but gave no residual greater than 1 cm^{-1}. The order of the required polynomial ranged from 5th at some levels for CO to 12th at some levels for HF and Li_2. The polynomial was checked for sensible results between the 41 points. The fitted polynomials were then transformed into Hermite polynomials, and the 1D anharmonic oscillator was solved variationally in the harmonic oscillator basis to obtain convergence to 1 cm^{-1} in the first excited state energy. This usually required about eight basis functions. Atomic masses[7] were used for all calculations.

Results and Discussion

Table 1 shows the experimental[8] and experimental minus calculated fundamental frequencies for the molecules studied in this work at the CCSD(T)/B level, where B represents each of the basis sets used in this work. Although it is not the current focus of this work to achieve good agreement between experiment and theory, this comparison has been made to provide the reader with an indication of the level of agreement that can be achieved without including core-correlation or relativistic effects. The basis sets in Table 1 have been arranged in order of decreasing quality. Not surprisingly, an inspection of the root-mean-square (rms) and mean absolute deviations (mad) clearly indicates a systematic variation in the errors between experiment and theory, with the higher levels being more accurate than the lower levels. There are, however, some interesting points to note. First, there is no significant difference between the results obtained at the CCSD(T)/cc-pVTZ and the CCSD(T)/6-311++G(2df,2pd) levels. Apparently, the additional diffuse functions included in the latter basis set have little impact on the stretching frequencies because the 6-311G(2df,2pd) and the cc-pVTZ basis sets are of similar quality.

Second, there is no great difference between the results obtained at the CCSD(T)/cc-pVDZ and CCSD(T)/6-311G(d,p) levels. Again, apparently the additional single valence basis function included in the later basis set has little impact on the stretching frequencies compared to the cc-pVDZ basis set because the 6-31G(d,p) basis set is of similar quality to the cc-pVDZ basis set.

Third, the agreement between theory and experiment appears to be comparatively bad in the case of HF and to a lesser extent in the cases of BeO and LiH. HF (and BH) was the focus of the recent study of Larsen et al.[9] These authors, among other things, looked at predicted fundamental frequencies at the CCSD(T)/cc-pVDZ level of theory. The fundamental reported in the current work at this level agrees with the result of Larsen et al. However, the reasonable agreement with experiment at this level is fortuitous, as clearly indicated in Table 1. Increasing the basis set to the 6-311G(d,p) level results in a deviation of −84 cm^{-1}. Convergence to near the experimental result is also noted to be slow, with a deviation of −15 cm^{-1} even at the CCSD(T)/cc-pV5Z level. Interestingly, augmenting the cc-pV*X*Z basis sets alleviates this situation significantly. The difference between experiment and theory becomes −7, −7, −1, and −27 cm^{-1} at the CCSD(T)/aug-cc-pV*X*Z levels where *X* is 5, Q, T, and D, respectively. Unfortunately, augmented basis sets do not yet exist for Be or Li, so a similar comparison cannot be made.

TABLE 2: Convergence of the Calculated[a] Fundamental Frequencies in cm^{-1}

molecule	5	5 − Q[b]	Q	Q − T[b]	T	T − D[b]	D
H_2	4163	1	4162	−4	4166	21	4145
LiH	1346	−1	1347	−3	1350	22	1328
BH	2264	1	2263	8	2255	8	2247
HF	3974	−12	3986	−16	4002	35	3967
Li_2	341	0	341	1	340	4	336
LiF	884	−7	891	−3	894	−63	957
BeO	1453	3	1450	15	1435	82	1353
BF	1372	−3	1375	−5	1380	80	1300
C_2	1834	3	1831	11	1820	18	1802
CO	2139	1	2138	10	2128	10	2118
N_2	2332	4	2328	10	2318	8	2310
F_2	903	5	898	3	895	137	758

[a] See text for a description of the level of theory. [b] The notation *X* − *Y* represents the fundamental frequency at level of theory *X* minus the fundamental frequency at level of theory *Y*.

To make meaningful comparisons between lower and higher levels of theory, the convergence of the fundamental frequencies should be established. Table 2 provides the differences between fundamental frequencies computed for the same molecule at progressively higher levels of theory. It is clear that, apart from HF and to a lesser extent LiF, convergence has approximately been achieved. Thus, for the rest of this discussion the results pertaining to HF and LiF have been excluded.

The harmonic and anharmonic contributions to the fundamental frequencies were computed. The anharmonic contribution to the calculated fundamental frequency was deduced by taking the difference between the calculated fundamental frequency and the calculated harmonic frequency. Table 3 provides the differences between the harmonic frequency calculated at various levels of theory and the harmonic frequency calculated at the CCSD(T)/cc-pV5Z level for the 10 remaining diatomics in this study. Table 3 also provides the same comparison for the anharmonic contribution. Only the mean absolute deviation (mad) and root mean square (rms) are given for brevity.

It is clear by examining Table 3 that a substantial improvement in the accuracy of the harmonic frequencies is achieved by using composite methods compared with the CCSD(T)/S frequencies. For example, the mad and rms harmonic frequencies are improved from 47 and 64 cm^{-1} to 14 and 19 cm^{-1}, respectively, when using harmonic frequencies computed using the T/D composite method compared with harmonic frequencies computed at the D level. It is also noted that little improvement is obtained here by increasing the L basis set beyond L = T. The same situation is observed with S = 311. Harmonic frequencies at the T/311 level are significantly better than at the 311 level, but little, if any, improvement is obtained when L is made larger than T. A less pronounced, but significant,

TABLE 3: Agreement between CCSD(T)/cc-pV5Z Harmonic Frequencies and Anharmonic Corrections (in cm^{-1}) and Lower-Level Full ab Initio Calculations and Composite Levels of Theory

	harmonic		anharmonic	
comparison[a]	mad[b]	rms[c]	mad[b]	rms[c]
5-D	47	64	3	4
5-T/D	14	19	1	2
5-++/D	15	22	2	4
5-Q/D	12	19	1	2
5-5/D	12	19	2	3
5-311	32	54	1	2
5-T/311	13	18	2	2
5-++/311	14	20	2	2
5-Q/311	11	17	2	2
5-5/311	11	16	2	2
5-T	9	11	1	2
5-++/T	9	12	1	2
5-Q/T	6	8	1	2
5-5/T	3	5	3	6

[a] See text for a description of the level of theory. The notation $X - Y$ represents the harmonic frequency or anharmonic correction at level of theory X minus that found for level of theory Y. [b] Mean absolute deviation. [c] Root mean square.

improvement is achieved in the harmonic frequencies by using the Q/T composite method rather than simply the T level of theory to compute the harmonic frequencies. Note that it was remarked earlier that the ++ level is little better than the T level of theory, which is born out here. Likewise, the 311 level is little better than the D level, a result also illustrated in Table 3.

Interestingly, the anharmonic contribution is always reasonably well predicted, even at the D level of theory. A small improvement is observed using the T/D composite method rather than the D level of theory, but little more improvement is gained by increasing the level of theory or composite method. This can possibly be understood if one considers the comparative sizes of the harmonic frequencies and the anharmonic contribution—the anharmonic contribution is about a factor of 10 smaller than the harmonic frequency. Thus, on the basis of this observation and the findings born out in Table 3, one could justify the use of a less accurate level of theory to compute the anharmonic correction and add this correction to an accurate calculation of the harmonic frequency.

Table 4 illustrates the accuracy of the above approach. Only the Dunning basis sets have been employed in this table, and the results from a harmonic frequency calculation using the 5/Q composite method have also been included. Fundamental frequencies have been computed using varying levels of theory for the harmonic frequency calculation and the D level and T/D method for the anharmonic correction. The notation used to describe this type of calculation is *A*,*B*, where *A* is the level of theory, or composite method, used to compute the harmonic frequency and *B* is the level of theory, or composite method, used to compute the anharmonic correction. These fundamentals are then compared with the fundamental frequencies computed at the CCSD(T)/cc-pV5Z level. It is immediately seen that a significant systematic improvement is achieved in the calculated fundamentals for both of the *A*,D and *A*,T/D series, with excellent agreement using the 5/Q,D and 5/Q,T/D procedures.

TABLE 4: Agreement between CCSD(T)/cc-pV5Z Fundamental Frequencies (in cm^{-1}) and Fundamentals Computed Using Separate Levels of Theory for the Harmonic and Anharmonic Contributions

comparison[a]	mad[b]	rms[c]
5-D,D	36	43
5-T/D,D	15	19
5-T,D	11	11
5-Q/T,D	7	9
5-Q,D	4	5
5-5/Q,D	3	3
5-5,D	3	4
5-T/D,T/D	13	19
5-T,T/D	9	10
5-Q/T,T/D	6	8
5-Q,T/D	3	3
5-5/Q,T/D	2	2
5-5,T/D	1	2

[a] See text for a description of the level of theory. The notation *A*,*B* represents a fundamental frequency computed using the harmonic frequency at level *A*, and the anharmonic correction at level *B*. [b] Mean absolute deviation. [c] Root mean square.

Concluding Remarks

In terms of the fundamental frequencies of vibration of the 12 diatomics studied in this work at the CCSD(T) level, the 6-311G(d,p) basis set was found to be essentially the same quality as the cc-pVDZ basis set. This was also found to be the case with the 6-311++G(2df,2pd) and cc-pVTZ basis sets. The simple composite method considered in this work involved performing only two calculations at each data point using the CCSD(T)/S and MP2/L levels of theory (results at the MP2/S level are also required but come as part of the CCSD(T)/S calculation), where S and L represent small and large basis sets, respectively.

It was found that this composite method substantially improved the predicted harmonic frequencies over those at the CCSD(T)/S level. The optimal choice of S and L was when L was only 1 greater in the valence designation compared with the S basis set. That is, when S = cc-pVDZ, L = cc-pVTZ, or when S = cc-pVTZ, L = cc-pVQZ. Increasing the L basis set further did not result in any significant improvement of the predicted frequencies. The composite level is meant to approximate an ab initio calculation at the CCSD(T)/L level. This approximation was found to be reasonably accurate when S and L basis sets were different by only 1 in their valence designations.

The anharmonic correction to the harmonic frequency was found to be well reproduced even at the CCSD(T)/cc-pVDZ level, with only a slight improvement using the S = cc-pVDZ, L = cc-pVTZ composite method. Thus, it appeared justifiable to compute fundamental frequencies by first calculating the harmonic frequency at a high level of theory and then correcting this frequency using an anharmonic correction at a much lower level of theory. This procedure was applied and found to be extremely accurate.

On the basis of the results presented here, at least for stretching modes, accurate fundamental frequencies can be obtained at relatively low computational expense by first computing harmonic frequencies using a composite method that is as high as possible and then applying an anharmonic correction using the S = cc-pVDZ, L = cc-pVTZ composite method. This is significant because one could hypothesize that the latter composite method could be used to generate a potential energy surface for a polyatomic system and the anharmonic correction could be extracted and applied to a high-level composite method harmonic frequency calculation to obtain cheap, but accurate, fundamental frequencies.

Acknowledgment. I acknowledge the financial support of the National University of Singapore's Faculty Research Grant.

References and Notes

(1) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.

(2) Bettens, R. P. A.; Collins, M. A.; Jordan, M. J. T.; Zhang, D. H. *J. Chem. Phys.* **2000**, *112*, 10162.

(3) Bettens, R. P. A. *J. Am. Chem. Soc.* **2003**, *125*, 584.

(4) See the special issues of *Spectrochim. Acta, Part A* **1997**, *53* and *Spectrochim. Acta, Part A* **1999**, *55*.

(5) Preface in *Spectrochimica Acta, Part A* **1997**, *53*, vii.

(6) MOLPRO 2000.1 is a package of ab initio programs written by H.-J. Werner and P. J. Knowles, with contributions from R. D. Amos, A. Bernhardsson, A. Berning, P. Celani, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, C. Hampel, G. Hetzer, T. Korona, R. Lindh, A. W. Lloyd, S. J. McNicholas, F. R. Manby, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, R. Pitzer, G. Rauhut, M. Schütz, H. Stoll, A. J. Stone, R. Tarroni, and T. Thorsteinsson.

(7) Mills, I.; Cvitas, T.; Homann, K.; Kallay, N.; Kuchitsu, K. *Quantities, Units and Symbols in Physical Chemistry*; Blackwell Scientific Publications: Oxford, England, 1988.

(8) Huber, K. P.; Herzberg, G. *Constants of Diatomic Molecules* (data prepared by J. W. Gallagher and R. D. Johnson, III); In NIST Chemistry WebBook, NIST Standard Reference Database Number 69; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg MD, July 2001; (http://webbook.nist.gov).

(9) Larsen, H.; Olsen, J.; Jørgensen, P.; Gauss, J. *Chem. Phys. Lett.* **2001**, *342*, 200.