

## Energy continuity in multi-reference CI calculations

James S. Wright and Richard J. Williams

Ottawa-Carleton Institute for Research and Graduate Studies in Chemistry, Carleton University  
Campus, Ottawa, Canada K1S 5B6

Configuration interaction calculations based on changing numbers of reference configurations at different geometries have a theoretical inconsistency which can affect the continuity of a calculated potential surface. As the number of reference configurations is increased for adjacent points on a potential curve (e.g. to describe molecular dissociation), the CI space can increase by large quantum jumps. Using the MRD-CI method of Buenker and Peyerimhoff and coworkers, we give several criteria which help to ensure energy continuity across these changes in CI space, and demonstrate these criteria for the hydrogen fluoride potential curve.

**Key words:** Potential surface—*ab initio*—configuration interaction

### 1. Introduction

Some recent studies of molecular potential energy curves have addressed the question of “size consistency” in an *ab initio* calculation [1–4]. With respect to a diatomic molecule  $AB$ , this may be defined as the consistency between calculations of the total energy of  $E(\text{atom } A) + E(\text{atom } B)$  compared to  $E(\text{molecule } AB, \text{ large } R)$ . The reason for a possible inconsistency is due to the differing size of the CI space in any scheme involving only single and double excitations. Using a multi-reference configuration-interaction method [5, 6] involving single and double-excitations from the reference configurations (MRD-CI), Buenker et al. [6] studied the magnitude of this inconsistency for calculations on  $N_2$  [2] and on  $S_2$  and  $SO$  [3]. In general, errors were dependent on basis set size and decreased to less than 0.1 eV by use of the full-CI estimate of quadruple excitations.

A closely related question occurs in another context in multi-reference CI calculations. As the bond distance changes, the number of reference configurations

needed to form a starting point for the CI also changes. In HF, for example, from 1 to 3 configurations are needed, where only one is required near  $R_e$  but three are required for  $R > 3.0$  bohr. In an MRD-CI calculation the CI space increases in rather large quantum jumps at each increase in the number of reference configurations, since single and double excitations are allowed from each reference configuration (hereafter abbreviated to each "Main"). To what extent will the computed energy be discontinuous in the region of a change in the number of Mains? This addresses the question of "size consistency" of another sort, which shall be termed "configuration consistency".

An obvious solution seen at various times in the literature [7, 8] is to include all Mains along the entire curve which have been shown to be important in any region, i.e. to use 3 Mains for the entire HF curve. A similar approach was used by Kraemer et al. [9] in a study of  $\text{H}_2\text{O}$ , where (the same) 10 Mains were used throughout. However, on dissociation surfaces for other triatomics and for polyatomics this approach can lead to a very large CI space and can easily become impractical. The purpose of the present paper is to explore the conditions under which energy continuity is maintained through regions of configuration inconsistency.

## 2. Main configurations

Consider the HF molecule in its ground state ( $X^1\Sigma^+$ ) dissociating to ground state products  $\text{H}(^2S)$  and  $\text{F}(^2P)$ . The potential energy at various internuclear distances  $R$  was calculated using the MRD-CI method, as described in the literature [5, 6]. A Gaussian basis set of ( $9s5p1d/4s2p1d$ ) orbitals on fluorine and ( $4s1p/2s1p$ ) on hydrogen was employed, along with  $s$ - and  $p$ -bond functions at the midpoint of the bond [10]. The  $1s$  core of fluorine was frozen in the CI; all other orbitals were allowed to participate. The selection threshold  $T$  was varied over the range 0–5 microhartree so that a maximum of 2000 configurations was retained in the final wave function. The extrapolation procedure described previously [5] was used when needed to estimate the  $T=0$  CI energy (denoted  $E_{\text{CI}}$ ) and the multi-reference Davidson correction [2] was added to this value to obtain the full-CI energy estimate (denoted  $F_{\text{CI}}$ ). This approach [10] led to a binding energy of 6.17 eV at  $R = R_e$ , in good agreement with the experimental value of 6.12 eV [11].

At  $R = R_e = 1.733$  bohr, the dominant configuration in the CI wave function is the SCF configuration  $(1\sigma)^2(2\sigma)^2(3\sigma)^2(1\pi)^4$ , abbreviated  $M1$ . This configuration gives an adequate description of the inner wall of the potential over the range 1.14–2.2 bohr (at  $R < 1.14$  bohr the molecule is no longer bound with respect to its atoms) and forms the basis of the CI which follows. Near  $R = 2.2$  bohr a second Main ( $M2$ ) becomes significant with configuration  $(1\sigma)^2(2\sigma)^2(4\sigma)^2(1\pi)^4$ , representing excitation into the antibonding  $\sigma_{\text{HF}}^*$  orbital. A third Main ( $M3$ ), consisting of  $(1\sigma)^2(2\sigma)^2(4\sigma)^1(5\sigma)^1(1\pi)^4$ , also contributes to a small extent, and then fades away at large  $R$ . At  $R > 3.0$  bohr a fourth Main ( $M4$ ) enters, having open shell configuration  $(1\sigma)^2(2\sigma)^2(3\sigma)^1(4\sigma)^1(1\pi)^4$ , as reported also by Buenker

et al. [12] for HCl and HF. The additional Mains  $M2$ ,  $M3$  and  $M4$  are built from the SCFMO's of  $M1$ .

Beyond  $R = 4.0$  a further complication develops in that  $M1$  no longer represents the dominant configuration, which then becomes  $M4$ . At this stage it is necessary to switch to the open shell SCF represented by  $M4$ . As is common in highly stretched configurations, SCF convergence problems at this point forced us to use the open shell triplet SCF. However, the CI was correctly the singlet  $M4$ , and the other two important configurations based on the SCF for  $M4$  turned out to be  $M1$  and  $M2$ , so that for the purposes of this paper it is sufficient to discuss the set of Mains  $M1$ ,  $M2$ ,  $M3$  and  $M4$ .

### 3. Energy continuity

We now examine in detail the regions near 2.2, 3.0 and 4.0 bohr where incremental changes occur. Results near  $R = 2.2$  bohr are shown in Table 1. In this table,  $T$  is the selection threshold value, which is varied to keep the matrix diagonalization manageable, and which is small enough so that it should not affect any arguments which follow. Use of  $M1$  leads to 971 configurations generated, whereas  $M1 + M2$  leads to 1802 (configuration inconsistency) and  $M1 + M2 + M3$  yields 3107 configurations. The squared contribution of each Main is indicated by  $C_i^2$ . At  $R < 1.9264$  only  $M1$  contributes appreciably to the CI wave function, the other (squared) coefficients being only 0.002. Nevertheless the  $E_{CI}$  energy is lowered by 0.03 eV on inclusion of a second Main and a further 0.01 eV on inclusion of the third. This behavior is to be expected [13]; expansion of the CI space should

**Table 1.** CI specifications for HF near 2 bohr

$R$ (bohr)	Mains <sup>a</sup>	$T^b$	$\frac{\# \text{Gener.}}{\# \text{Select}}$	$E_{CI}$ (eV) <sup>c</sup>	$F_{CI}$ (eV) <sup>d</sup>	$M1$	$C_i^2$ $M2$	$M3$
1.9264	$M1, (M2), (M3)$	0	971/971	-5.758	-5.889	0.955	(0.002)	(0.002)
1.9264	$M1, M2, (M3)$	0	1802/1802	-5.788	-5.905	0.954	0.003	(0.002)
1.9264	$M1, M2, M3$	1	3107/1058 <sup>e</sup>	-5.807	-5.904	0.954	0.003	0.003
2.0980	$M1, (M2), (M3)$	0	971/971	-5.222	-5.376	0.952	(0.005)	(0.004)
2.0980	$M1, M2, (M3)$	0	1802/1802	-5.269	-5.394	0.949	0.006	(0.004)
2.0980	$M1, M2, M3$	1	3107/1153	-5.295	-5.392	0.949	0.006	0.004
2.2348	$M1, (M2), (M3)$	0	971/971	-4.717	-4.893	0.948	(0.007)	(0.004)
2.2348	$M1, M2, (M3)$	0	1802/1802	-4.783	-4.913	0.944	0.010	(0.004)
2.2348	$M1, M2, M3$	1	3107/1223	-4.813	-4.911	0.944	0.010	0.005
2.3600	$M1, (M2), (M3)$	0	971/971	-4.234	-4.434	0.944	(0.010)	(0.004)
2.3600	$M1, M2, (M3)$	0	1802/1802	-4.320	-4.454	0.939	0.014	(0.005)
2.3600	$M1, M2, M3$	1	3107/1261	-4.352	-4.452	0.938	0.014	0.005

<sup>a</sup> Bracketed Mains are not explicitly included but contribute to the final wave function by the indicated amount.

<sup>b</sup> In microhartree

<sup>c</sup> Relative to atomic energies:  $F(-99.52544 \text{ hartree})$  and  $H(-0.4976 \text{ hartree})$

<sup>d</sup> Relative to atomic energies:  $F(-99.52953 \text{ hartree})$  and  $H(-0.4976 \text{ hartree})$

<sup>e</sup> Estimated extrapolation uncertainty to  $T=0$  is 0.01 eV in this case

cause a lowering in the value of  $E_{CI}$ . The  $F_{CI}$  estimate is more stable and shows a variation of  $<0.02$  eV as both Mains are added. There is very little change in the coefficients for the three cases, which show agreement to within 0.001. Data at  $R = 2.0980$  are particularly revealing. Here the  $M1$  calculation gives an  $F_{CI}$  energy of  $-5.376$  eV relative to atoms, and  $C_i^2$  values of 0.952, 0.005 and 0.004 for  $M1$ ,  $M2$  and  $M3$ , respectively. The  $M1, M2$  calculation gives an energy of  $-5.394$  eV and  $C_i^2$  values of 0.949, 0.006 and 0.004. The change in the first coefficient suggests that  $M2$  was needed. The constancy (and small value) of the third coefficient suggests that  $M3$  is not needed, and inclusion of  $M3$  has only a very small effect on the  $F_{CI}$  energy. These trends are repeated at 2.2348 and 2.3600 bohr, where  $M2$  becomes more significant. Based on these data, it is clear that  $F_{CI}$  energies are preferred, and that inclusion of additional Mains should be conditional on *a*) the magnitude of their contribution, and *b*) the change in  $C_i^2$  values when more Mains are added.

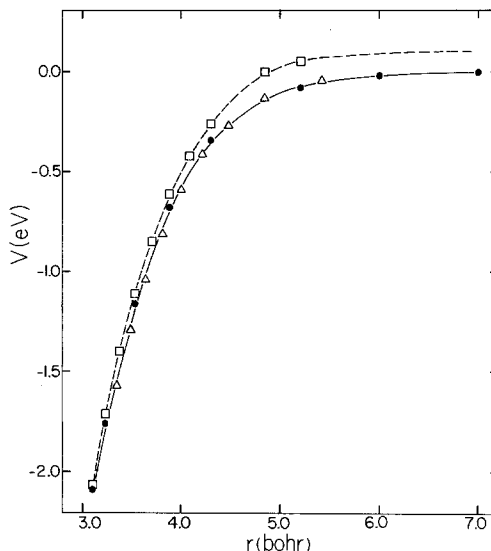
Results near  $R = 3.0$  bohr, where the switch from two to three Mains occurs, are shown in Table 2. Here the number of generated configurations is 1802 ( $M1 + M2$ ) or 2771 ( $M1 + M2 + M4$ ). At  $R = 2.8409$  the second Main is important with  $C_2^2 = 0.042$  and the third begins to appear, with  $C_3^2 = 0.005$ . Inclusion of  $M4$  explicitly again causes only a small change in coefficients and a change in the  $F_{CI}$  energy of less than 0.01 eV. At  $R = 2.9665, 3.0971$  and  $3.2340$  bohr the coefficient  $C_4^2$  increases progressively from 0.007 to 0.012 (two-Main calculation) and the  $F_{CI}$  energy difference for the two-Main vs. three-Main calculation shows a progressively growing deviation from 0.02 to 0.06 eV. (In this case the  $E_{CI}$  shows smaller changes). Thus to maintain continuity in the  $F_{CI}$  calculation to within 0.02 eV, the third Main must be included when its (squared) coefficient exceeds 0.005.

At  $R = 4$  bohr it was mentioned earlier that  $M4$  becomes dominant. If  $M1$  is nevertheless chosen as the basis of the CI, then the potential curve dissociates to an energy which is 0.1 eV above that of the separated atoms (an apparent "size inconsistency"). Switching to  $M4$  at  $R \geq 4$  bohr removes this problem, and dissociation occurs to within 0.01 eV of the separated atom limit. However,

**Table 2.** CI specifications for HF near 3 bohr

$R$ (bohr)	Mains <sup>a</sup>	$T^b$	$\frac{\# \text{ Gener.}}{\# \text{ Select}}$	$E_{CI}$ (eV) <sup>c</sup>	$F_{CI}$ (eV) <sup>d</sup>	$M1$	$C_i^2$ $M2$	$M4$
2.8409	$M1, M2, (M4)$	0	1802/1802	-2.658	-2.809	0.907	0.042	(0.005)
2.8409	$M1, M2, M4$	2	2771/1207	-2.664	-2.783	0.907	0.042	0.005
2.9665	$M1, M2, (M4)$	0	1802/1802	-2.285	-2.443	0.895	0.052	(0.007)
2.9665	$M1, M2, M4$	2	2771/1256	-2.292	-2.408	0.894	0.053	0.008
3.0971	$M1, M2, (M4)$	0	1802/1802	-1.929	-2.098	0.880	0.065	(0.009)
3.0971	$M1, M2, M4$	5	2771/1050	-1.947	-2.059	0.879	0.065	0.010
3.2340	$M1, M2, (M4)$	0	1802/1802	-1.590	-1.774	0.863	0.079	(0.012)
3.2340	$M1, M2, M4$	5	2771/1052 <sup>e</sup>	-1.608	-1.714	0.860	0.080	0.015

<sup>a, b, c, d, e</sup> as in Table 1



**Fig. 1.** Potential curve for HF (ground state) in the region 3–7 bohr: ( $\Delta$ ) experimental (RKR) points; ( $\bullet$ ) *ab initio* points based on configuration *M4*; ( $\square$ ) *ab initio* points based on configuration *M1*

continuity must again be demonstrated where there is a change in the SCF from a closed to an open-shell configuration, since this will affect the CI which follows.

Fig. 1 shows an expanded view of the potential curve in the region beyond 3 bohr.  $F_{CI}$  values based on *M1* (dashed line) are seen to dissociate above the atomic limit.  $F_{CI}$  values based on *M4* are seen to merge smoothly, however, into the potential curve from below and also to show the correct dissociative behavior, since the open shell SCF goes correctly to the separated-atom limit. The potential curves based on *M1* or *M4* blend smoothly at  $R = 3.2$  bohr, which was therefore chosen as the place to switch to the open-shell configuration. Agreement with the experimental (*RKR*) points is very satisfactory.

#### 4. Conclusions

To correctly describe molecular potential curves in a CI calculation, it is necessary to use multiple reference configurations. It would be highly desirable to preserve energy continuity over the entire curve or surface without including all relevant Mains over the entire space. This can be accomplished to a high degree of precision by including only those Mains which contribute significantly to the CI wave function. In the present case of HF, to obtain an  $F_{CI}$  energy stable to 0.02 eV or better, explicit use of additional Mains need only be invoked when their (squared) contribution to the CI exceeds 0.005. Incrementing the number of Mains should then take place in this region. This procedure ensures that energy continuity is maintained while at the same time minimizing the number of reference configurations genuinely required in the calculation. Continuity can also be established in a similar way even when the MO's forming the CI are changed, as described here and shown in Fig. 1.

Experience gained with other diatomics has shown that these conclusions are generally true, but that the inclusion threshold may vary to some extent (and depends on the desired precision), so that continuity checks should be performed routinely. In general, energy continuity at either the SCF or CI level is not a guarantee of accuracy of the resulting potential surface. Our own calculations show, for example, that the SCF potential curve for HF is continuous though far from accurate. Continuity across changes in the size of the CI space should be considered to be a necessary but not sufficient condition to ensure accuracy at the CI level of treatment. In addition, AO basis set errors in many cases may be appreciably greater than discontinuity errors. The improvement to be realized through continuity will be limited by the magnitude of such errors. Nevertheless, these criteria do result in an improvement of the energy at the CI stage for a given AO basis. When calculations on potential surfaces with multiple dissociation channels are undertaken, these procedures should allow continuous surfaces to be generated with a minimum of computational effort.

*Acknowledgement.* This research was supported by the Natural Sciences and Engineering Research Council of Canada and by the University of Ottawa Computer Center. We thank Dr. R. J. Buenker for the use of his CI package.

## References

1. Raffanetti, R. C., Hsu, K., Shavitt, I.: *Theoret. Chim. Acta (Berl.)* **45**, 33 (1977)
2. Butscher, W., Shih, S. K., Buenker, R. J., Peyerimhoff, S. D.: *Chem. Phys. Letters* **52**, 457 (1977)
3. Theodorakopoulos, G., Peyerimhoff, S. D., Buenker, R. J.: *Chem. Phys. Letters* **81**, 413 (1981)
4. Swope, W. C., Lee, Y. P., Schaefer III, H. F.: *J. Chem. Phys.* **70**, 947 (1979)
5. Buenker, R. J., Peyerimhoff, S. D.: *Theoret. Chim. Acta (Berl.)* **35**, 33 (1974)
6. Buenker, R. J., Peyerimhoff, S. D., Butscher, W.: *Mol. Phys.* **35**, 771 (1978)
7. Segal, G. A., Wolf, K.: *Chem. Phys.* **56**, 321 (1981)
8. Grein, F., Kapur, A.: *J. Chem. Phys.* **77**, 415 (1982)
9. Kraemer, W. P., Roos, B. O., Siegbahn, P. E. M.: *Chem. Phys.* **69**, 305 (1982)
10. Wright, J. S., Williams, R. J.: *J. Chem. Phys.*, **78**, 5264 (1983)
11. Huber, K. P., Herzberg, G.: *Molecular spectra and molecular structure*, Vol. 4. New York: Van Nostrand Reinhold 1979
12. Bettendorf, M., Peyerimhoff, S. D., Buenker, R. J.: *Chem. Phys.* **66**, 261 (1982)
13. Buenker, R. J., Shih, S. K., Peyerimhoff, S. D.: *Chem. Phys.* **36**, 97 (1979)

Received August 8, 1983/October 28, 1983