
ON THE ACCURACY OF CONVENTIONAL CALCULATIONS OF POTENTIAL ENERGY SURFACES

Ulf LANDSCHEID and Werner KUTZELNIGG

Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, 4630 Bochum, F.R.G.

Received November 16th, 1987

Accepted January 4th, 1988

Dedicated to Dr R. Zahradnik on the occasion of his 60th birthday.

It is demonstrated for the example of H_2^+ that with the conventional basis expansion method (in terms of Gaussians) with fixed non-linear parameters the error of a computed potential curve varies strongly with the geometry. If the non-linear parameters are optimized for the equilibrium geometry, much larger errors arise for other geometries. This is very pronounced if only one polarization function of a given l is added to the basis, and qualitatively the same problem arises for any new l . The dependence of the error on the distance may also be oscillatory. For the example of Li_2 in the SCF or an MC-SCF approximation essentially the same effects are observed. Here in addition substantial errors for large R may arise in the SCF curve if the basis is not able to describe the spurious ionic contributions.

If one computes a potential energy surface (PES) via pointwise approximate solution of the Schrödinger equation by the basis expansion method one cannot expect that the error is independent of the geometry, not even that it depends monotonically on changes in the geometry. This has been known since the early days of numerical quantum chemistry but is usually not checked in routine calculations of PES, except if one strives for very high accuracy¹. It is the purpose of this letter to draw attention to the fact that potential surfaces obtained with standard basis sets may be quite inaccurate even if the error at the equilibrium distance is satisfactorily small.

It is also a wisdom of the early days of numerical quantum chemistry that consistent accuracy for different geometries can only be achieved if the non-linear parameters of the basis are optimized for each geometry². This is, however, a very expensive procedure. It is usually cheaper to increase the size of the basis, but this may lead to an oscillatory dependence of the error on geometrical parameters, as we shall show here.

We shall, in this study, consider first the H_2^+ system in detail, mainly since the exact potential curve for this system is known and since the only errors that one makes are due to the choice of the basis. There is hence no interference with other sources of errors such as the problem to account for electron correlation etc..

H_2^+ , admittedly, is not in all respects typical for molecules currently investigated. On one hand H_2^+ is an ion, and long-range induction effects – which are not present in neutral molecules – are very important, we therefore get larger errors than we would e.g. get for H_2 . On the other hand the limit of non-interacting systems ($R \rightarrow \infty$) is – even for modest basis sets – described so well that the error for finite R is never smaller than that for $R = \infty$, i.e. that “counterpoise” corrections³ are not necessary.

For molecules larger than H_2^+ exact PES are not available, but a good measure of the accuracy of the calculation with a given basis set is the difference between this PES and that calculated with a somewhat larger basis. The simplest nontrivial example is that of Li_2 in the SCF approximation. It shows, in fact, similar effects as H_2^+ . Note that much larger variations of the error with the basis set are to be expected if correlation effects are included.

CALCULATIONS

H_2^+ with Various Basis Sets

We have started with calculations of H_2^+ with Gaussian basis sets consisting of s -functions (centered on the respective nuclei) only. The error $\Delta(R)$ of the calculated potential curve (PC) compared to the exact one⁴ is plotted in Figs 1a and 1b in different scale, for $0 \leq R \leq 10a_0$.*

For all but the largest basis sets the maximum error arises at $R = 0$, but this error is not very serious since the PC is very repulsive near $R = 0$ anyway. All s -only calculations show a local maximum of $\Delta(R)$ near $R = 2.5a_0$, slightly larger than the equilibrium distance $2a_0$. The size of this error appears to reach a limit of 12.5 millihartree (mH)** for a “saturated” s -basis, while for rather modest s -basis sets an error near 20 mH is obtained. This is a considerable fraction of the binding energy of 103 mH.

The error $\Delta(R)$ also exhibits a local minimum, somewhere between $R = 0$ and $R = 2a_0$, the magnitude of which depends on the size of the basis.

The effect of contracting the steepest lobes in the basis is negligible for $R > 3a_0$ and very small for $3a_0 > R > 1a_0$. Relaxation of the contraction reduces the error considerably only for very small R . This is to be expected, since for $R \rightarrow 0$ H_2^+ approaches the united atom He^+ and larger exponents of the basis functions are needed – which are present in the part of the basis that is usually contracted. It appears that for a “saturated” s -basis the error for $R \lesssim 0.5a_0$ becomes very small.

Next we took basis sets with s and p functions. The results for basis sets with a moderately good s basis (5 lobes in the contraction 311) and a single p function

* $a_0 = 0.529177 \cdot 10^{-10}$ m.

** 1 mH = $2.6255 \cdot 10^3$ J mol⁻¹.

(as is often used in practice) are shown in Fig. 2. Both the size and the location of the maximum error depend sensitively on the chosen exponent. The error at the equilibrium distance $R = 2a_0$ is minimized by the exponent $\eta = 0.594$ to $\Delta \approx 1$ mH. However the error varies considerably in the neighbourhood of the minimum distance, it has a maximum of ≈ 5 mH near twice the equilibrium distance $R = 4a_0$ and it has a similar value at half the equilibrium distance. If one wants to minimize the error at large distances an exponent $\eta = 0.1$ is very good⁵, but for this exponent one has an error of ≈ 12 mH at the equilibrium distance.

By optimizing η as a function of R one can reduce the error in the whole region $1a_0 < R < 10a_0$ to ≈ 2 mH, but this is, of course, expensive and one prefers to use two or more p functions with fixed exponents. The errors from such calculations are shown in Fig. 3.

One can choose the two exponents such that the error is minimal near the equilibrium distance. This is achieved for $\eta_1 = 1.0$, $\eta_2 = 0.33$. The error is then 0.7 mH at $R = 2.0a_0$ (roughly half that obtained with a single p -function) but it is twice as large at $R = 5.0a_0$. One might expect that the combination $\eta = 0.594$ (good for $R = 2a_0$) and $\eta = 0.1$ (good for $R > 5a_0$) should be a good compromise between the η needed for small and large R . This basis set is, in fact, only slightly worse than the (1.0/0.33) set near $R = 2a_0$ and much better than the (1.0/0.33) set for $R > 5a_0$. However, the error is strongly dependent on R and has a steep maximum ($\Delta = 2$ mH near $R = 4a_0$).

With three p -sets one does better, but the error is still far from being a monotonic function of R , and the steep increase of the error for $R < 2a_0$ is unpleasant. However, as we have seen before, the error in this region can be reduced by augmenting the s -basis.

The error near the equilibrium distance can be reduced by adding a d -function with $\eta = 1.0$. Δ becomes 0.3 mH for a (311s, 3p, 1d) basis and 0.1 mH for a (3111s, 3p, 1d) basis. In either case the error has a local maximum of 0.75 or 0.45 mH respectively at $R = 4.5a_0$. In order to have an error $\lesssim 0.1$ mH consistently for all R one must optimize the exponent of the d -function as a function of R , or use two (or better three) d -functions with fixed exponents. The same game starts again if one includes an f -function.

On any level of accuracy that one wants to reach minimization of the error at the equilibrium distance automatically implies much larger errors at distances larger or shorter than R_0 .

Li₂ in the SCF-Approximation

An exact (restricted) Hartree-Fock potential curve of Li₂ is, of course, not known. We have therefore taken the results obtained with a (11s, 7p, 2d) basis as reference. From the fact that the error of a basis without d -functions with respect to this

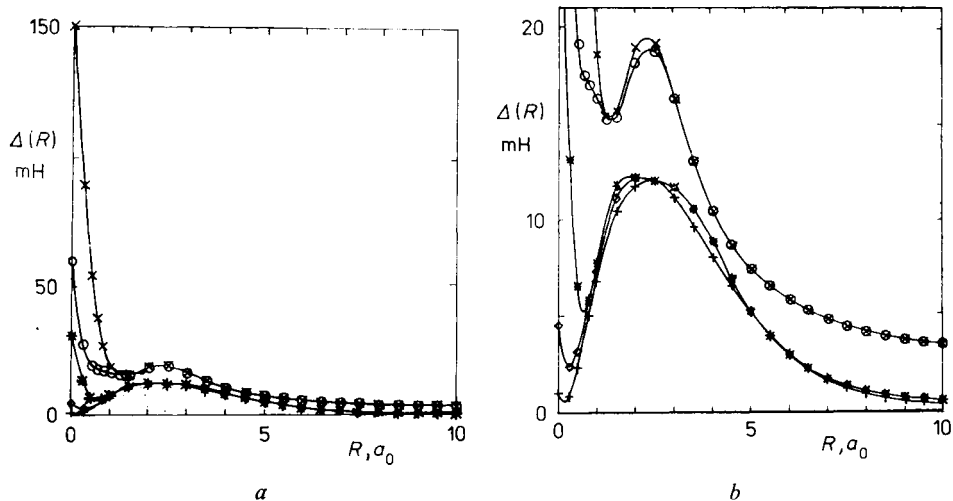


FIG. 1

H_2^+ . Error $\Delta(R)$ in millihartree of calculations with various "s-only" Huzinaga Gaussian basis sets. R in a_0 . \times (21), \circ (111), $*$ (311), \square (11111), $+$ (31111); **1a** and **1b** in different scale

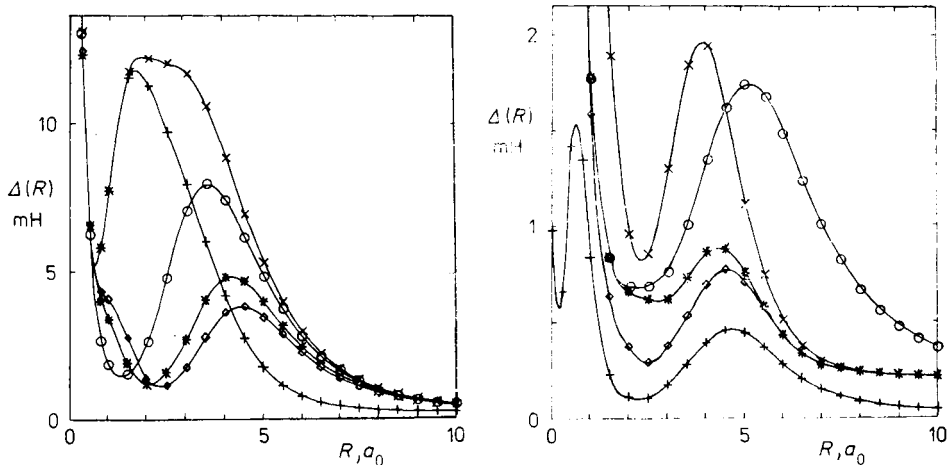


FIG. 2

As Fig. 1, but with basis sets, including one p -function. The s -part is always (311). η_p : \times no p ; \circ 1.0; $*$ 0.594; \square 0.5; $+$ 0.1

FIG. 3

As Fig. 1, but basis sets including two or three p -functions and no or one d -function. The s -part is (311), except for $+$ where it is (31111). η_p : \times 0.594, 0.1; \circ 1.0, 0.33; $*$ 1.0, 0.33, 0.1; \square 1.0, 0.33, 0.1; $+$ 1.0, 0.33, 0.1; η_d 1.0

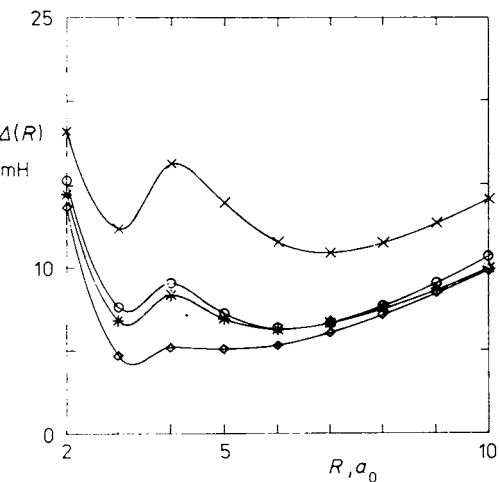


FIG. 4

Li_2 . Error $\Delta(R)$ in millihartree of SCF calculations with various basis sets as compared to a calculation with an $(11s, 7p, 2d)$ basis. The Huzinaga s -basis always (31111) ; \times no p ; \odot $(1p)$; $*$ $(2p)$; \square $(3p)$, exponents in the text

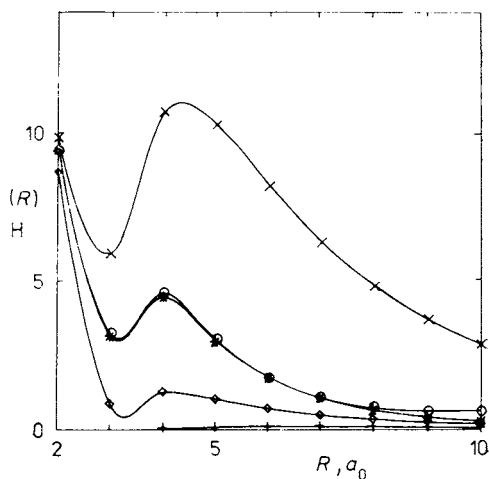


FIG. 5

As Fig. 4, but with an s -basis $(3, 8 \times 1)$; \times no p ; \odot $(1p)$; $*$ $(2p)$; \square $(3p)$; $+$ $(7p)$

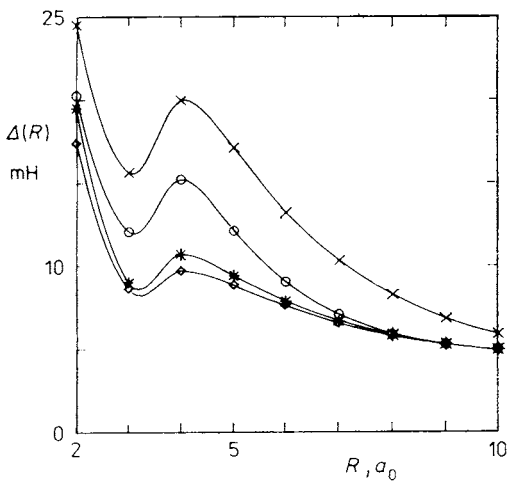


FIG. 6

As Fig. 4, but in a three-configuration MC-SCF (only σ -MOs) and different exponents for $(2p)$: 0.1107, 0.0443

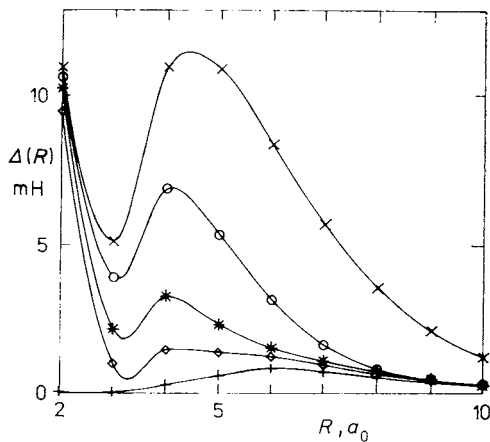


FIG. 7

As Fig. 5, but in a three-configuration MC-SCF (only σ -MOs) and exponents as on Fig. 6

reference is smaller than 0.2 mH (see Fig. 5) in the whole region of distances, we conclude that the errors with respect to the exact PC would not differ much from those with respect to our reference.

In Fig. 4 the errors from calculations with a $7s$ -basis and p -sets of various size are plotted. With s -function only we find a (local) minimum $\Delta(3a_0) = 12.5$ mH, a (local) maximum $\Delta(4a_0) = 16.5$ mH and another flat local minimum $\Delta(7a_0) = 11$ mH.

One p -function (with $\eta = 0.07$), leads to a lowering of the error by 5–7 mH, almost independent of R . The results with two p -sets ($\eta_1 = 0.07$, $\eta_2 = 0.0292$) do not differ much from those with one p -set. With three d sets ($\eta_1 = 0.175$; $\eta_2 = 0.07$; $\eta_3 = 0.0292$) the local maximum seems to disappear and the error is rather independent of R in the region $3a_0 \leq R \leq 7a_0$, although its size (≈ 5 mH) is rather large (the binding energy is 39 mH).

Increasing of the size of the s -basis from $7s$ to $11s$ (contraction $3,4 \times 1$ and $3,8 \times 1$ respectively) has a noticeable effect for large R . While with $7s$ the error increases with R for $R \geq 7a_0$, it decreases with the $9s$ basis and appears to approach zero for large enough R (to test this is not so easy due to SCF convergence problems at these large distances).

The (local) maximum error is, on going from $7s$ to $11s$, only reduced from 16 to 11 mH. The effect of the addition of 1, 2 or 3 p -sets is similar as with the $5s$ basis. With $11s$ and $3p$ the error is for $3 \leq R \leq 10a_0$ of the order 1 mH.

Since one may argue that the SCF approximation is quite unphysical for large R , one may as well conclude that the SCF-basis errors for large R are rather meaningless. We have therefore compared the same basis sets in a MC-SCF calculation, consisting of two σ_g^2 and one σ_u^2 configurations for the valence electrons. The results are in Figs 6 and 7. For the basis sets with a large s -part (Figs 5 and 7) the pattern is qualitatively very similar in the SCF and the MC-SCF case. For the basis sets with a small s -part (Figs 4 and 6) the errors in the MC-SCF calculation for all discussed basis sets become small for $R \rightarrow \infty$ and reach the limit of the error for two Li atoms (4.6 mH), but in the SCF calculations the error increases with R . The explanation is probably the following one.

For large R the SCF wave function is a mixture of "covalent" and an "ionic" structure. The latter, describing the system $\text{Li}^+ - \text{Li}^-$ requires basis functions with very small η (so-called "diffuse functions"). These are not present in the basis sets with small s -part, but are available in the larger basis. So the error in Fig. 4 for large R is due to the inability of the basis to describe the (physically spurious) ionic contributions in the SCF wave function.

CONCLUSION

The results of this study indicate clearly that calculations without polarization functions or the inclusion of a single polarization function (with fixed exponent η) of

a given angular momentum l necessarily lead to a substantial variation of the error of a computed potential curve (or hypersurface) with changes of the geometry. If one optimizes the exponent η of this polarization function for the equilibrium geometry one gets rather large errors for other geometries. Individual optimization of η for each computed geometry is better, but usually too time-consuming. Use of more than one polarization function reduces the error, but may lead to a strong oscillatory behaviour of the error on the bond distance. With other choices of the exponents than those reported here one can even get more spectacular oscillations.

The problem how to achieve consistent accuracy over a whole potential surface without an intolerable amount of computational effort cannot be regarded as solved, at least as far traditional basis expansion methods are concerned.

It should be noted that the errors to which we have here paid attention result from the fact that a polarization function with a fixed η has a "maximum overlapping power" at some $R(\eta)$. They have nothing to do with the well-known basis superposition errors³, which come from the fact that a "molecular basis" describes a single atom better than a purely "atomic basis". These errors also vary with the geometry, but can be corrected by the counterpoise method. There are further errors due to the inability of gaussian wave functions to take care of the nuclear cusp condition⁶. In the case of H_2^+ these errors are much smaller than those considered here.

If there are strong variations of the error with the geometry one might expect wrong equilibrium geometries and force constants. However, if the non-linear parameters are optimized for the equilibrium geometry, the variations of the error in immediate neighborhood of the equilibrium geometry are small, such that the effect on equilibrium geometries and force constants is quite small. In fact the R_e and ω_e values obtained with the basis sets of this study look rather reasonable and do not display the erratic behaviour of $A(R)$. Conversely good equilibrium geometries and force constants do not imply a good description of the potential surfaces for other geometries (and this independently of the fact that often molecules are in well-closed shell states only in their equilibrium geometry and close to it).

There are other (and possibly more serious) sources of errors in potential hypersurfaces, mainly those due to our inability to account for 100% of the correlation energy and the fact that the correlation energy is often very sensitive to changes in the geometry. The observation that even for the simplest possible molecule, namely H_2^+ , standard quantum chemical methods are hardly capable of guaranteeing a consistent accuracy along the potential curve, should certainly damp our optimism concerning the quantum chemical methods in current use.

The authors thank V. Staemmler for useful comments and Ulrich Meier for his MC-SCF program.

REFERENCES

1. Meyer W., Botschwina P., Burten P.: *J. Chem. Phys.* *84*, 891 (1986).
2. Hoyland J. R.: *J. Chem. Phys.* *40*, 3540 (1964).
3. Boys S. F., Bernardi F.: *Mol. Phys.* *19*, 553 (1970).
4. Wind H.: *J. Chem. Phys.* *42*, 2371 (1965).
5. Kochanski E.: *Theor. Chim. Acta* *39*, 339 (1975).
6. Klopper W., Kutzelnigg W.: *J. Mol. Struct. (THEOCHEM)* *135*, 339 (1986).