Theoret. Chim. Acta (Berl.) 39, 161—172 (1975)

# A Least-Squares Method Applied to the Hydrogen Molecule

Inge Røeggen*
Theoretical Chemistry Department, University of Oxford

John P. Huggett and Edward A. G. Armour
Department of Mathematics, University of Nottingham

Received March 13, 1975

A least-squares approach is proposed to remedy some of the weaknesses of the method of moments. It is shown by an application to the hydrogen molecule that the least-squares procedure yields reliable results in cases where the method of moments fails. The prediction of equilibrium internuclear distances by the least-squares method is also found to be more reliable than the corresponding results obtained by the method of moments.

*Key words:* Bivariational procedures – Method of moments versus a least squares method

## 1. Introduction

There exists a vast variety of methods devoted to the problem of getting good approximations to the eigenvalues and the eigenfunctions of the Schrödinger equation for a molecular system. Most of these methods or algorithms are based on the standard Ritz variational principle. The advantage of the Ritz method is that the eigenvalue problem contains only straightforward algebraic expressions when linear trial functions are used, and that it yields simultaneous upper bounds on both ground and excited states energies. However, further progress in the application of this method seems to be hampered by difficulties in integral evaluations. When explicit correlation terms are introduced into the trial functions, one is posed with the problem of evaluating very difficult multi-dimensional integrals for the matrix elements of the Hamiltonian in question. To surmount this obstacle some bivariational procedures have been devised, among which the method of moments [1, 3] seems to be quite promising.

The attractiveness of the method of moments is due to the possibility of simplification of the integral computation. This can be achieved by using a very effective basis set on the right, say correlated wave functions, and then use such functions on the left which lead to relatively simple integrals. The choice of the set on the left can also be made in a way leading to a wave function with an improved accuracy in physically important regions of configuration space. Contrary to the standard Ritz procedure, the approximate eigenvalue obtained by the method of moments has no upper-bound character. This might be considered as a

---

* Present address: University of Tromsø, Institute of Mathematical and Physical Sciences, P.O. Box 953, 9001 Tromsø, Norway.

disadvantage, but this lack of upper-bound character yields the possibility of obtaining some "random sampling type" error estimates [2]. A more serious objection, however, is that the method is not entirely reliable. Schwartz [4] and Armour [5] have shown that calculations may fail, partly or completely, due to numerical instability if the imbalance between the functions used on the right and the left is too great. To improve the method's reliability, two proposals have been made. The first one, which has been discussed by Hegyi *et al.* [2], is to determine criteria for selecting an appropriate set of functions to use on the left in conjunction with a given set on the right. The second possibility is to average the results obtained in a large number of method of moment calculations. In this paper we shall advocate a third approach, a generalization of the method of moments to a least-squares problem.

In the next section we shall review the relevant features of the method of moments, reformulate it into a least-squares problem, and discuss some theoretical aspects of this least-squares method. In the third section the proposed method is applied to the hydrogen molecule as a test example.

## 2. A Generalization of the Method of Moments

The dissymmetric equations characterizing the method of moments can be obtained by the following considerations. Let $\{\chi_i, i=1, \ldots, n\}$ and $\{\phi_i, i=1, \ldots, n\}$ be two nonidentical sets of expansion functions. Approximations to the eigenfunctions of the Schrödinger equation

$$\mathscr{H}\Psi = E\Psi \tag{1}$$

are assumed to be of the form

$$\begin{aligned} X &= \textstyle\sum_{i=1}^{n} a_i \chi_i \\ \Phi &= \textstyle\sum_{i=1}^{n} b_i \phi_i \,. \end{aligned} \tag{2}$$

The coefficients $\{b_i\}$ and the approximate eigenvalue $W$ are determined by requiring that the projection of $(\mathscr{H}-W)\Phi$ onto the subspace spanned by $\{\chi_i\}$ should be zero, *i.e.*

$$\langle \chi_k | (\mathscr{H}-W) | \textstyle\sum_{i=1}^{n} b_i \phi_i \rangle = 0 \,, \qquad k=1, \ldots, n \tag{3}$$

which is equivalent to

$$\textstyle\sum_{i=1}^{n} (H_{ki} - WS_{ki}) b_i = 0 \,, \qquad k=1, \ldots, n \tag{4}$$

where

$$\begin{aligned} H_{ki} &= \langle \chi_k | \mathscr{H} | \phi_i \rangle \\ S_{ki} &= \langle \chi_k | \phi_i \rangle \,. \end{aligned} \tag{5}$$

By exactly the same type of argument one arrives at the equations for the coefficients $\{a_i\}$:

$$\textstyle\sum_{k=1}^{n} a_k^* (H_{ki} - WS_{ki}) = 0 \,, \qquad i=1, \ldots, n \,. \tag{6}$$

The two sets of Eqs. (4) and (6) may be called respectively the direct and the adjoint bivariational equations. Since the secular determinants associated with these two sets of equations differ only by an interchange of rows and columns, the two sets of equations must have the same eigenvalues. However, the eigenvectors of the direct and the adjoint equations corresponding to the same eigenvalue are in general different. Because of the non-symmetric character of the matrices involved, complex roots may also appear. If this occurs in the neighbourhood of the energy values one is interesting in, one of the basis sets must be changed.

The relationship of $X$ and $\Phi$ to the corresponding exact eigenfunction $\Psi$, has been discussed by Boys [6]. Boys demonstrated that

$$W=E+\mu^{+}\mu(W^{11}+\mu W^{12}+\mu^{+}W^{21}+\dots)\,, \tag{7}$$

$$\Phi=\Psi+\mu(\Phi^{01}+\mu^{+}\Phi^{11}+\mu\Phi^{02}+\dots)\,, \tag{8}$$

$$X=\Psi+\mu^{+}(X^{10}+\mu X^{11}+\mu^{+}X^{20}+\dots)\,, \tag{9}$$

where $\Psi$ and $E$ are the exact eigenfunction and eigenvalue, and the eigenfunction is assumed to be real. The parameters $\mu$ and $\mu^{+}$ are unknown which are defined to be the least-squares errors in fitting $\Psi$ with linear combinations of the expansion functions $\{\phi_i\}$ and $\{\chi_i\}$; i.e.

$$\mu=\{\langle\Psi-\textstyle\sum_{i=1}^{n} b'_i\phi_i|\Psi-\sum_{i=1}^{n} b'_i\phi_i\rangle\}^{1/2}\,, \tag{10}$$

$$\mu^{+}=\{\langle\Psi-\textstyle\sum_{i=1}^{n} a'_i\chi_i|\Psi-\sum_{i=1}^{n} a'_i\chi_i\rangle\}^{1/2}\,. \tag{11}$$

The parameters $W^{11}$, $\Phi^{01}$, $X^{10}$, etc. in Eqs. (7–9) are Taylor series coefficients. Provided that these coefficients are not unduly large, and the functions $\{\phi_i\}$ used on the right are capable of fitting $\Psi$ very accurately, it can be expected that W and $\Phi$ will be very accurate even though the functions $\{\chi_i\}$ used on the left are only capable of fitting $\Psi$ with considerably less accuracy.

Unfortunately, the method of moments equations are akward to solve and their solutions may prove to be numerically unstable, particularly, if the difference between the basis sets used on the right and left is considerable. The numerical problems due to the imbalance between the basis sets can be avoided by casting Eqs. (4) and (6) into a least-squares problem and by using a larger basis set onto which $(\mathscr{H}-W)\Phi$ [or $(\mathscr{H}-W)X$] is projected, i.e.

$$\begin{aligned}\sigma&=\textstyle\sum_{k=1}^{m\geq n} w_k|\langle\chi_k|(\mathscr{H}-W)|\Phi\rangle|^2\\&=\textstyle\sum_{k=1}^{m\geq n} w_k|\langle\chi_k|(\mathscr{H}-W)|\sum_{i=1}^{n} b_i\phi_i\rangle|^2\end{aligned} \tag{12}$$

where $\{w_k\}$ are weight factors. The functions $\{\chi_k\}$ used on the left are assumed to be normalized. The coefficients $\{b_i\}$ and the eigenvalue $W$ can then be determined by minimizing the functional $\sigma$ when the coefficients are subjected to a given normalization condition.

The only further requirement in the least-squares (LSQ) method beyond the method of moments are overlap integrals of the form $\langle\chi_k|\chi_j\rangle$. The evaluation of these integrals represents no difficulties except in calculations on many electron systems in which the $\chi_i$'s contain non-separable functions of the interelectronic distances. Difficulties of this type can be avoided by applying the LSQ-method in the context of the transcorrelated method [7]. The key point is that the LSQ-method is numerically much more stable than the method of moments; there is no necessity to invert the overlap matrix $\boldsymbol{S}$ in Eq. (4), or carry out some other equivalent procedure, which may lead to difficulties if $\boldsymbol{S}$ is nearly singular. Thus an increased stability and reliability is obtained at the cost of a few more integrals to be calculated since $m$ is usually greater than $n$. In the case $m=n$, the LSQ approach is equivalent to the method of moments. However, even in this case it is reasonable to expect that the LSQ procedure is numerically more stable than the conventional routines for solving the dissymmetric equations.

If the set on the left is an orthonormal basis and the weight factors $\{w_k\}$ are all set equal to 1, then it is obvious that $\sigma$ is an approximation to the variance.

$$\begin{aligned}\sigma &= \textstyle\sum_{k=1}^{m\geqq n} |\langle\chi_k|\mathscr{H}-W|\Phi\rangle|^2\\ &= \textstyle\sum_{k=1}^{m\geqq n} \langle\Phi|\mathscr{H}-W|\chi_k\rangle\langle\chi_k|\mathscr{H}-W|\Phi\rangle\\ &\leqq \textstyle\sum_{k=1}^{\infty} \langle\Phi|\mathscr{H}-W|\chi_k\rangle\langle\chi_k|\mathscr{H}-W|\Phi\rangle\\ &= \langle\Phi|(\mathscr{H}-W)^2|\Phi\rangle = \sigma_{\text{var}}\,. \end{aligned} \tag{13}$$

Since

$$\frac{\partial\sigma_{\text{var}}}{\partial W}=0 \tag{14}$$

at the minimum for $\sigma_{\text{var}}$, one obtains

$$W=\frac{\langle\Phi|\mathscr{H}|\Phi\rangle}{\langle\Phi|\Phi\rangle} \tag{15}$$

in the case $\Phi$ and $W$ is determined by minimizing the variance. This last result should indicate that for a given $n$ and increasing $m$, the energy value obtained by minimizing $\sigma$ will converge towards the value given by the Rayleigh quotient. There is, however, no argument to suggest that the convergence is monotonic.

The preceding arguments might lead one to believe that with this particular choice of weight factors and an orthonormal basis on the left, that there is an upper bound character in the LSQ-method when $m>n$. In general this is certainly not the case, which the following counter example will show.

Suppose that the basis sets $\{\phi_i, i=1, \dots, m\}$ and $\{\chi_k, k=1, \dots, m\}$ span the vector spaces $\Lambda$ and $\Omega$, respectively. A method of moments calculation based on these sets yields an eigenvalue $W$ and a corresponding normalized eigenvector $\Phi$. Suppose further that $W$ is less than the exact energy value. A new set $\{\eta_i, i=1, \dots, n\}$ spanning a proper subspace $\Gamma$ of $\Lambda$, i.e. the dimension $n$ is less than the dimension $m$ of $\Lambda$, is chosen such that

$$\eta_1=\Phi\,. \tag{16}$$

An LSQ-calculation with a basis set on the right spanning $\Gamma$ and $\{\chi_k, k=1, \dots, m\}$ on the left will obviously give $W$ and $\Phi$ as a solution. This particular LSQ-calculation yields an energy below the exact value irrespective of how much greater $m$ is than $n$.

Furthermore, if

$$\eta_1(x)=\cos x\Phi+\sin x\Phi' \tag{17}$$

where $\Phi'$ is a function orthogonal to $\Phi$, i.e.

$$\begin{aligned} &\langle\Phi|\Phi'\rangle = 0 \\ &\langle\Phi'|\Phi'\rangle = 1 \end{aligned} \tag{18}$$

is used instead of $\eta$, and $W(x)$ is continuous in some interval containing the point $x=0$, then there exists a $\delta>0$ such that $W(x)$ is less than the exact value for $|x|<\delta$.

On the other hand, the LSQ eigenvalue $W$ is an upper bound if the set on the right spanning $\Gamma$ is a subspace of $\Omega$, generated by the set $\{X_k, k=1, \ldots, m\}$ used on the left. By assuming that $\{\Psi_k, k=1, \ldots, m\}$ is an orthonormal set spanning $\Omega$, one obtains the following value for $W$ at the minimum of the LSQ-functional

$$W = \frac{1}{2}\frac{\sum_{k=1}^{m}\{\langle\Phi|\mathscr{H}|\Psi_k\rangle\langle\Psi_k|\Phi\rangle + \langle\Phi|\Psi_k\rangle\langle\Psi_k|\mathscr{H}|\Phi\rangle\}}{\sum_{k=1}^{m}\langle\Phi|\Psi_k\rangle\langle\Psi_k|\Phi\rangle}. \tag{19}$$

Since $\Gamma \subset \Omega$,

$$|\Phi\rangle = \sum_{k=1}^{m}|\Psi_k\rangle\langle\Psi_k|\Phi\rangle \tag{20}$$

By combining Eqs. (19) and (20) one arrives at

$$W = \frac{\langle\Phi|\mathscr{H}|\Phi\rangle}{\langle\Phi|\Phi\rangle}. \tag{21}$$

Since $\Phi$ is obtained by the LSQ-procedure, the value $W$ given Eq. (21) is above the result which coult be obtained by minimizing the Rayleigh-quotient.

The LSQ-method presented has been used with success in connection with scattering problems [2] Resently it has also been suggested by Røeggen [8] that the LSQ-method may be useful in dealing with large configuration interaction expansions. Results are presented in the next section which indicate that it also has an important role as a more reliable alternative to the method of moments.

## 3. Application to $H_2$

The program used to perform the work was developed from the program for the method of moments calculations reported in Ref. [5]. The sets of expansion functions are defined in terms of the confocal elliptical coordinates $\lambda$, $\mu$, and $\phi$. If the nuclei A and B are at a distance $R$ apart, then the coordinates $\lambda$ and $\mu$ of any point are given by

$$\lambda = (r_A + r_B)/R \tag{22}$$

$$\mu = (r_A - r_B)/R \tag{23}$$

where $r_A$, for example, is the distance of the point from nucleus A. The third coordinate $\phi$ is the azimuthal angle. The sets of expansion functions, $\{\chi_i\}$ and $\{\phi_i\}$, were chosen from the set of functions of the form:

$$\frac{1}{2\pi}(\lambda_1^m\lambda_2^n\mu_1^j\mu_2^k + \lambda_1^n\lambda_2^m\mu_1^k\mu_2^j)\varrho_{12}^p \exp\{-\delta(\lambda_1+\lambda_2)\} \tag{24}$$

Table 1. Basis functions used, characterized by indices defined in connection with Eq. (24)

| i | $m_i$ | $n_i$ | $j_i$ | $k_i$ | $p_i$ |
|---|---|---|---|---|---|
| 1 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 0 | 2 | 0 |
| 3 | 0 | 0 | 1 | 1 | 0 |
| 4 | 1 | 0 | 1 | 1 | 0 |
| 5 | 1 | 0 | 0 | 2 | 0 |
| 6 | 0 | 0 | 2 | 2 | 0 |
| 7 | 1 | 0 | 0 | 0 | 0 |
| 8 | 2 | 0 | 0 | 0 | 0 |
| 9 | 1 | 1 | 2 | 0 | 0 |
| 10 | 1 | 1 | 0 | 0 | 0 |
| 11 | 2 | 1 | 0 | 0 | 0 |
| 12 | 2 | 0 | 1 | 1 | 0 |
| 13 | 1 | 1 | 1 | 1 | 0 |
| 14 | 1 | 1 | 2 | 0 | 0 |
| 15 | 2 | 1 | 1 | 1 | 0 |
| 16 | 1 | 1 | 2 | 2 | 0 |
| 17 | 2 | 2 | 1 | 1 | 0 |
| 18 | 0 | 0 | 0 | 0 | 1 |
| 19 | 0 | 0 | 0 | 0 | 2 |
| 20 | 2 | 0 | 0 | 0 | 1 |
| 21 | 0 | 0 | 1 | 1 | 1 |
| 22 | 1 | 0 | 0 | 0 | 1 |
| 23 | 1 | 0 | 0 | 0 | 2 |

where

$$\varrho_{12} = \frac{2}{R} r_{12} \tag{25}$$

and the numbers $m, n, j, k$, and $p$ have non-negative integer values. Since the calculation is on the ground state which is a singlet of $\Sigma_g^+$ symmetry, the condition $(-1)^{j+k} = 1$ must be satisfied. A list of the basis set from which $\{\chi_i\}$ and $\{\phi_i\}$ were chosen, is given in Table 1. The value of $\delta$ is set equal to 0.95 for all functions.

The evaluation of the relevant matrix elements $H_{ki}$ and $S_{ki}$ is described in Ref. [5]. The matrices $\boldsymbol{H}$ and $\boldsymbol{S}$ are all real.

When the basis sets $\{\chi_k, k=1, \ldots, m\}$ and $\{\phi_i, i=1, \ldots, n\}$ have been chosen, the program first evaluated the matrices $\boldsymbol{H}$ and $\boldsymbol{S}$, Eq. (5), and then transformed to representations $\boldsymbol{H}'$ and $\boldsymbol{S}'$ over orthonormal bases $\{\eta_k, k=1, \ldots, m\}$ and $\{\theta_i, i=1, \ldots, n\}$ derived from $\{\chi_k\}$ and $\{\phi_i\}$, respectively. The program then minimized the functional

$$\sigma = \sum_{k=1}^{m \geqq n} \langle \Phi | \mathscr{H} - \boldsymbol{W} | \eta_k | \boldsymbol{H} - W | \Phi \rangle \tag{26}$$

with respect to $\Phi$ and $W$. Since

$$\Phi = \sum_{i=1}^{n} b_i \theta_i \,, \tag{27}$$

the LSQ-functional has the following form in matrix notation

$$\begin{aligned}\sigma &= \boldsymbol{b}^T(\boldsymbol{H}' - W\boldsymbol{S}')^T(\boldsymbol{H}' - W\boldsymbol{S}')\boldsymbol{b} \\ &= \boldsymbol{b}^T\boldsymbol{A}(W)\boldsymbol{b}\,.\end{aligned} \tag{28}$$

The normalization constraint

$$\boldsymbol{b}^T\boldsymbol{b} = 1 \tag{29}$$

is dealt with by the method of Lagrange multipliers, i.e. the functional

$$\sigma_\mu = \sigma - \mu(\boldsymbol{b}^T\boldsymbol{b} - 1)$$

is considered. By requiring that $\sigma_\mu$ should be stationary with respect to the variations in the parameter $W$ and the coefficients $\{b_i\}$, one obtains the following equations

$$W = \frac{1}{2}\,\frac{\boldsymbol{b}^T(\boldsymbol{S}'^T\boldsymbol{H}' + \boldsymbol{H}'^T\boldsymbol{S}')\boldsymbol{b}}{\boldsymbol{b}^T\boldsymbol{S}'^T\boldsymbol{S}'\boldsymbol{b}} \tag{30}$$

and

$$(\boldsymbol{A}(W) - \mu\boldsymbol{I})\boldsymbol{b} = 0 \tag{31}$$

where $\boldsymbol{I}$ is the identity matrix. Equations (30) and (31) are solved by iterations. A starting value is chosen for $E$ which is in the region to which $W$ is expected to converge. Then Eq. (31) is solved with this value of $W$, for the lowest eigenvalue $\mu$ and associated normalized eigenvector, $\boldsymbol{b}$, of $A(W)$. A new value of $W$ is then determined according to Eq. (30). The procedure is repeated until convergence is obtained.

The convergence criterion used in all cases with one exception to be specified later, was that the iterations were continued until the change in the energy value, $W$, produced by successive iterations was less than $10^{-6}$ a.u. The accompanying change in $\sigma$ was recorded in each case and was found always to be much less that $10^{-6}$.

The basis sets to be employed on the left in the calculations are

Set A: $1 \rightarrow 17$
Set B: $1 \rightarrow 5, 7, 8, 10, 18 \rightarrow 22, 6, 11, 23, 9$
Set C: $1 \rightarrow 8, 18, 19, 9 \rightarrow 15$

where the numbering is according to Table 1. When less than the maximum number of basis functions on the left is used, say $m$ functions, then these functions are the first $m$ members of the set considered when the numbering is defined by the listing of functions above. For all calculations the same set is used on the right, namely the first 10 functions of set B. It should be noticed that the set A is qualitatively different from B and C since it does not include functions depending on the electronic distance $r_{12}$.

The first problem considered was the question of the convergence properties of the minimization technique adopted. A series of calculations using set A on the left and with $m = n = 10$ was performed with different initial values of $W$

Table 2. LSQ-calculations with basis set A on the left

| No. of fns. on l.h.s. | $\sigma$ $x10^6$ | $W$ (a.u.) | Binding energy (eV) | No. of iterations |
|---|---|---|---|---|
| 10 | 0 | - 1.177852 | 4.8394 | 27 |
| 11 | 24 | - 1.177876 | 4.8400 | 8 |
| 12 | 85 | - 1.188232 | 5.1218 | 25 |
| 13 | 131 | - 1.178203 | 4.8489 | 5 |
| 14 | 263 | - 1.178168 | 4.8480 | 3 |
| 15 | 275 | - 1.178607 | 4.8599 | 4 |
| 16 | 276 | - 1.178613 | 4.8601 | 2 |
| 17 | 334 | - 1.178880 | 4.8673 | 4 |

Initial value of $W = -1.15$ a.u.
Internuclear distance $R = 1.40$ a.u.
Convergence criterion: $\delta W < 10^{-6}$
The quantity $\delta W$ is the difference in $W$ values obtained in the final and the preceding iterations for the number of functions on the left.

ranging from $-1.15$ a.u. to $-1.50$ a.u. The maximum discrepancy in the energy at the minimum was found to be $5 \cdot 10^{-6}$ a.u. This discrepancy can probably be attributed to round-off errors (MM-calculation failed in this case), and we conclude that the minimization technique used is reliable. The number of iterations of course, depends on how far the initial value is away from the value of the energy at the minimum. In our work we used only direct iterations. But the number of necessary iterations can certainly be reduced by applying an extrapolation technique to the iteration sequence of $W$-values obtained, say Pade approximants of type II [9]. As far as the ground state is concerned, a satisfactory choice of initial value can be based on the self-consistent field result and a rough estimate of the correlation energy.

In Table 2 LSQ-calculations with basis set A on the left is presented. The initial value of $W = -1.15$ a.u. and the internuclear distance $R = 1.40$ a.u. This is a case where there is an imbalance between the two basis sets employed since the set on the right includes functions depending on $r_{12}$ while the set on the left does not. It might therefore not be surprising that all the energy values obtained when $m$ is ranging from 10 to 17, are below the "exact" one which is $-1.17444$ a.u. [10].

On the other hand, all the energy values in Table 2 with one exception, $m = 12$, are close to the "exact" value. However, which of these results obtained with different $m$ values, is the best one in the meaning closest to the exact result, can certainly not be decided upon by considering the sequence of $W$ values. On the other hand, the safest choice will in most cases be to choose the result corresponding to the maximum value of $m$ since the functional $\sigma$ in this case represents the best approximation to the variance. Another point pertaining to Table 2 which we would like to emphasize is the result for $m = n = 10$. In this case the method of moments calculation fails completely while the LSQ-method yields a very reasonable result and thereby supports the assumption about the LSQ-method's greater stability.

Table 3. LSQ-calculations with basis set B on the left

| No. of fns. on l.h.s. | $\sigma$ x$10^6$ | $W$ (a.u.) | Binding energy (eV) | No. of iterations |
|---|---|---|---|---|
| 11 | 75 | - 1.173923 | 4.7324 | 2 |
| 12 | 150 | - 1.173927 | 4.7326 | 2 |
| 13 | 253 | - 1.173931 | 4.7327 | 2 |
| 14 | 858 | - 1.173933 | 4.7327 | 2 |
| 15 | 858 | - 1.173933 | 4.7327 | 2 |
| 16 | 923 | - 1.173930 | 4.7326 | 2 |
| 17 | 998 | - 1.173925 | 4.7325 | 2 |

The method of moments calculation which is variational in this case, yields $W = -1.173943$ a.u., and this value is taken as the initial value for the LSQ-calculations. Internuclear distance $R = 1.40$ a.u. Convergence criterion: $\delta\sigma < 10^{-6}$. The quantity $\delta\sigma$ is the difference in $\sigma$ values obtained in the final and the preceding iterations for the number of functions on the left.

Table 4. LSQ energy values (a.u.) for varying internuclear distance using set A on the left

| $R$ | No. of functions on l.h.s. | | | |
|---|---|---|---|---|
| (a.u.) | 14 | 15 | 16 | 17 |
| 1.30 | - 1.177576 | - 1.177693 | - 1.177712 | - 1.177921 |
| 1.31 | - 1.177836 | - 1.177989 | - 1.178007 | - 1.178221 |
| 1.32 | - 1.178047 | - 1.178234 | - 1.178251 | - 1.178471 |
| 1.33 | - 1.178209 | - 1.178431 | - 1.178446 | - 1.178672 |
| 1.34 | - 1.178236 | - 1.178581 | - 1.178595 | - 1.178827 |
| 1.35 | - 1.178399 | - 1.178667 | - 1.178699 | - 1.178937 |
| 1.36 | - 1.178430 | - 1.178750 | - 1.178761 | - 1.179004 |
| 1.37 | - 1.178420 | - 1.178772 | - 1.178781 | - 1.179030 |
| 1.38 | - 1.178373 | - 1.178754 | - 1.178762 | - 1.179017 |
| 1.39 | - 1.178288 | - 1.178698 | - 1.178706 | - 1.178966 |
| 1.40 | - 1.178168 | - 1.178606 | - 1.178613 | - 1.178879 |
| 1.41 | - 1.178015 | - 1.178480 | - 1.178486 | - 1.178757 |
| 1.42 | - 1.177830 | - 1.178320 | - 1.178326 | - 1.178601 |
| 1.43 | - 1.177613 | - 1.178128 | - 1.178133 | - 1.178414 |
| 1.44 | - 1.177367 | - 1.177905 | - 1.177910 | - 1.178196 |
| 1.45 | - 1.177092 | - 1.177652 | - 1.177657 | - 1.177948 |

Convergence criterion: $\delta W < 10^{-6}$ a.u.

There is a peculiarity in Table 2 which has to be mentioned, namely the result for $m = 12$. This result deviates strongly from the main stream of results given in the table. If the absolute difference in $W$-values obtained with $m = 12$ and any other $m$ value is compared with the differences in energy values for any pair of $m$'s not involving $m = 12$, then differences involving $m = 12$ are found to be an order of magnitude greater than those not involving $m = 12$. A first guess of an explanation of this surprising result might be that the result is due to an error of some type connected with the basis function number 12 of set A. However, this possi-

Table 5. LSQ binding energies ($eV$) for varying internuclear distance using set A on the left

| $R$ (a.u.) | No. of functions on l.h.s. 14 | 15 | 16 | 17 |
|---|---|---|---|---|
| 1.30 | 4.8318 | 4.8350 | 4.8356 | 4.8412 |
| 1.31 | 4.8389 | 4.8431 | 4.8436 | 4.8494 |
| 1.32 | 4.8447 | 4.8498 | 4.8502 | 4.8562 |
| 1.33 | 4.8491 | 4.8551 | 4.8555 | 4.8617 |
| 1.34 | 4.8522 | 4.8592 | 4.8596 | 4.8659 |
| 1.35 | 4.8542 | 4.8621 | 4.8624 | 4.8689 |
| 1.36 | 4.8551 | 4.8638 | 4.8641 | 4.8707 |
| 1.37 | 4.8548 | 4.8644 | 4.8646 | 4.8714 |
| 1.38 | 4.8535 | 4.8639 | 4.8641 | 4.8710 |
| 1.39 | 4.8512 | 4.8624 | 4.8626 | 4.8697 |
| 1.40 | 4.8480 | 4.8599 | 4.8601 | 4.8673 |
| 1.41 | 4.8438 | 4.8564 | 4.8566 | 4.8640 |
| 1.42 | 4.8387 | 4.8521 | 4.8522 | 4.8597 |
| 1.43 | 4.8329 | 4.8468 | 4.8470 | 4.8546 |
| 1.44 | 4.8262 | 4.8408 | 4.8409 | 4.8487 |
| 1.45 | 4.8187 | 4.8339 | 4.8340 | 4.8420 |

Convergence criterion: $\delta W < 10^{-6}$ a.u.

Table 6. LSQ energy values (a.u.) for varying internuclear distance using set C on the left

| $R$ (a.u.) | No. of functions on l.h.s. 14 | 15 | 16 | 17 |
|---|---|---|---|---|
| 1.30 | - 1.171549 | - 1.171479 | - 1.171469 | - 1.171438 |
| 1.31 | - 1.172008 | - 1.171942 | - 1.171932 | - 1.171902 |
| 1.32 | - 1.172413 | - 1.172351 | - 1.172341 | - 1.172312 |
| 1.33 | - 1.172767 | - 1.172709 | - 1.172698 | - 1.172669 |
| 1.34 | - 1.173070 | - 1.173016 | - 1.173004 | - 1.172976 |
| 1.35 | - 1.173325 | - 1.173274 | - 1.173262 | - 1.173235 |
| 1.36 | - 1.173534 | - 1.173485 | - 1.173472 | - 1.173446 |
| 1.37 | - 1.173696 | - 1.173650 | - 1.173637 | - 1.173612 |
| 1.38 | - 1.173815 | - 1.173771 | - 1.173758 | - 1.173733 |
| 1.39 | - 1.173891 | - 1.173850 | - 1.173836 | - 1.173811 |
| 1.40 | - 1.173927 | - 1.173887 | - 1.173873 | - 1.173849 |
| 1.41 | - 1.173922 | - 1.173884 | - 1.173870 | - 1.173846 |
| 1.42 | - 1.173879 | - 1.173842 | - 1.173828 | - 1.173804 |
| 1.43 | - 1.173799 | - 1.173763 | - 1.173748 | - 1.173725 |
| 1.44 | - 1.173682 | - 1.173648 | - 1.173632 | - 1.173610 |
| 1.45 | - 1.173531 | - 1.173497 | - 1.173482 | - 1.173459 |

Convergence criterion: $\delta W < 10^{-6}$ a.u.

bility can be disregarded since the actual function is also included in the sets with $m > 12$ and the corresponding functionals do not display the same peculiarities. The most reasonable explanation seems to be that the functional with $m = 12$ has a strange behaviour around the minimum, perhaps in the form of two close

Table 7. LSQ binding energies (eV) for varying internuclear distance using set C on the left

| $R$ | No. of functions on l.h.s. | | | |
|---|---|---|---|---|
| (a.u.) | 14 | 15 | 16 | 17 |
| 1.30 | 4.6678 | 4.6659 | 4.6657 | 4.6648 |
| 1.31 | 4.6803 | 4.6785 | 4.6783 | 4.6774 |
| 1.32 | 4.6914 | 4.6897 | 4.6894 | 4.6886 |
| 1.33 | 4.7010 | 4.6994 | 4.6991 | 4.6983 |
| 1.34 | 4.7092 | 4.7078 | 4.7074 | 4.7067 |
| 1.35 | 4.7162 | 4.7148 | 4.7145 | 4.7137 |
| 1.36 | 4.7218 | 4.7205 | 4.7202 | 4.7195 |
| 1.37 | 4.7263 | 4.7250 | 4.7247 | 4.7240 |
| 1.38 | 4.7295 | 4.7283 | 4.7280 | 4.7273 |
| 1.39 | 4.7316 | 4.7305 | 4.7301 | 4.7294 |
| 1.40 | 4.7325 | 4.7315 | 4.7311 | 4.7304 |
| 1.41 | 4.7324 | 4.7314 | 4.7310 | 4.7303 |
| 1.42 | 4.7312 | 4.7302 | 4.7298 | 4.7292 |
| 1.43 | 4.7291 | 4.7281 | 4.7277 | 4.7271 |
| 1.44 | 4.7259 | 4.7250 | 4.7245 | 4.7239 |
| 1.45 | 4.7218 | 4.7209 | 4.7204 | 4.7198 |

Convergence criterion: $\delta W < 10^{-6}$ a.u.

minima, and that this abnormal character of the functional is changed when more basis functions are added. The occurence of functionals that have this peculiar behaviour might be considered as a weakness of the proposed LSQ-procedure. In our calculations the case $m = 12$ associated with set A, is the only one with this particular behaviour, and that might indicate that this problem occurs rather seldom. Furthermore, anomalous functionals of this type can easily be discovered, isolated and disregarded by performing a set of calculations with a different number of basis functions on the left. The additional work involved in performing such a sequence of calculations is negligible since the minimization of one of these functionals yields extremely good initial values to be used in the succeeding calculations.

In Table 3 one of our preliminary LSQ calculations is presented. For these calculations a weaker convergence criterion was used since the iterative procedure was terminated when $\delta\sigma < 10^{-6}$. However, comparison calculations using the stronger convergence criterion, $\delta W < 10^{-6}$, indicate that this is unlikely to affect significantly the accuracy of the $W$ values obtained. The calculation illustrates one of the theoretical points discussed in Section 2. The basis set on the right is in this case a proper subset of the basis set used on the left. Theoretically we then expect the LSQ-calculations to yield energy values above the variational result obtained with the basis on the right. All the results in Table 3 are in agreement with this theoretical prediction.

Finally, a set of calculations of the energy and the binding energy for varying internuclear distance is presented. The results are given in Tables 4–7. Two different basis sets are used on the left, set A and set C. In the former case the position for the predicted minimum lies in the range 1.36–1.38 a.u. For increasing *m* the minimum is shifted to the right. As for the second set, consistently better values are obtained for the equlibrium internuclear distance, ranging from 1.40 to 1.41 a.u. The predicted position of the minimum changes in this case very little with respect to the number of basis functions used on the left. The results in Tables 4–7 should therefore indicate that when the basis sets do not differ qualitatively too much, as in the calculations involving set C, the LSQ procedure will predict reliable internuclear distances. And even if this condition is not satisfied, as in the calculation involving set A, improved results compared with a method of moments calculation might be expected.

## 4. Conclusions

In an evaluation of the LSQ-procedure we would like to emphasize that the proposed method has essentially the same advantage as the method of moments, namely the possibility of simplification in the calculation of integrals. However, on the basis of this work it seems reasonable to claim greater reliability of the LSQ-method compared with the method of moments since the former yields good results where the latter fails completely. In addition, the LSQ method seems to be more reliable in predicting equilibrium distances. But the relative merits of the two methods considered can only be settled completely when additional calculations are performed. It is the opinion of the authors that in this context the LSQ-method has not had the attention it deserves.

In this work we have restricted ourselves to the calculation of the energy. An equally important and interesting aspect which has to be studied, is the question of how to calculate expectation values within this LSQ-framework. One approach could be the following: After determining the right hand wave function by minimizing the LSQ-functional, a second function can be obtained by projecting the r.h. function onto the subspace spanned by the basis set used on the left. These two approximate wave functions can then be used in the standard method of moments expression for expectation values. But a test of this approach might be the theme for another study.

*Acknowledgement.* We wish to thank the S.R.C. for a maintenance grant for one of us (J.P.H.).

## References

1. Szondy, E., Szondy, T.: Acta. Phys. Hung. **20**, 253 (1966)
2. Hegyi, M. G., Mezei, M., Szondy, T.: Theoret. Chim. Acta. (Berl.) **15**, 273, 283 (1969); **21**, 168 (1971)
3. Ladányi, K., Lengyel, V., Szondy, T.: Theoret. Chim. Acta. (Berl.) **21**, 176 (1971)
4. Schwartz, C.: J. Comp. Phys. **2**, 90 (1967)
5. Armour, A. E. G.: Mol. Phys. **25**, 993 (1973)
6. Boys, S. F.: Proc. R. Soc. **A 309**, 195 (1969)
7. Boys, S. F., Handy, N.: Proc. Roy. Soc. **A 310**, 43, 63 (1969)
8. Røeggen, I.: Manuscript submitted to Chem. Phys. Letters
9. Røeggen, I.: Chem. Phys. Letters **22**, 140 (1973)
10. Kolos, W., Roothaan, C. C. J.: Rev. Mod. Phys. **32**, 219 (1960)

Dr. I. Røeggen
University of Tromsø
Institute of Mathematical and Physical Sciences
P.O. Box 953
9001 Tromsø, Norway