Theoret. Chim. Acta (Berl.) 42, 261–271 (1976)

THEORETICA CHIMICA ACTA

Determination of Molecular Properties by the Method of Moments

V. Criteria for "Good" Weight Functions and Application to Correlated Wave Functions

Tamás Szondy

Hungarian United Chemical Works, Computing Center, Erzsébet Királyné útja 1/c, H-1146 Budapest XIV, Hungary ★

Ede Kapuy

Quantum Theory Group, The Hungarian Academy of Sciences, Budafoki út 8, H-1111 Budapest XI, Hungary

Some precautions needed in the choice of weight functions when calculating wave functions by the method of moments are analysed. It appears that an important criterion for "good" weight functions is that the difference between the "total" and the "truncated" overlaps (both defined in the paper) be high.

The method of moments is applied to wave functions involving Hylleraas-type correlation factors using weight functions made up of products of single-particle orbitals. The aim of the calculations is partly to test the criteria for "good" weight functions, partly a preparation of more extended calculations of a similar type.

Key words: Method of moments

1. Introduction

In quantum mechanical calculations difficulties of integration often more fundamentally influence the choice of the analytical form of the variational wave functions than the physical considerations. The difficulties of integration, however, depend on the criteria by which the values of the variational parameters are determined. It appears that the use of the method of moments[1] as a criterion is a powerful tool for reducing difficulties of integration, thus opening the way of handling problems by physically better wave functions. The results obtained by the method of moments can, however, become unstable if some precautions are neglected. The aim of this paper is to analyse some of these precautions and test them on examples which, at the same time, prepare the application of the method of moments to variational wave functions made up of group orbitals and involving Hylleraas-type correlation factors between particles belonging to the same group [7–10]

★ The new name of the institute is: Computer Application Research and Development Center of the Chemical Industries.

[1] The idea of using the method of moments for reducing difficulties of integration seems to have emerged in [1]. A detailed summary of its basic principles has been given in [2]. Numerical results together with results of principle can be found in many papers, e.g. [3–6].

2. The Stability of Results Obtained by the Method of Moments

Let $\mathscr{H}$ be the Hamiltonian operator of a molecular system with eigenvalues E_i and eigenfunctions $U_i(\mathbf{x})$

$$(\mathscr{H} - E_i)U_i(\mathbf{x}) = 0; \qquad \langle U_i(\mathbf{x}) \mid U_i(\mathbf{x})\rangle = 1; \qquad E_i \leqslant E_{i+1}; \qquad (i = 0, 1, \ldots), \tag{1}$$

where $\mathbf{x}$ denotes the point in the configurational space. Be $u_i(\mathbf{x}, \mathbf{a})$ the variational wave function by which we want to approximate to $U_i(\mathbf{x})$ and which involves the variational parameters $\mathbf{a}$

$$u_i(\mathbf{x}, \mathbf{a}) = a_0\tilde{u}_i(\mathbf{x}, a_1, \ldots, a_m); \qquad \langle u_i(\mathbf{x}, \mathbf{a}) \mid u_i(\mathbf{x}, \mathbf{a})\rangle = 1, \tag{2}$$

i.e. we formally consider the normalization factor as a variational parameter. In order to determine the values of the variational parameters $\mathbf{a}$ and the approximation e_i to E_i by the method of moments we first choose a weight function generator

$$w_i(\mathbf{x}, \mathbf{b}) = b_0\tilde{w}_i(\mathbf{x}, b_1, \ldots, b_n); \qquad \langle w_i(\mathbf{x}, \mathbf{b}) \mid w_i(\mathbf{x}, \mathbf{b})\rangle = 1; \qquad n \geqslant m \tag{3}$$

depending on the point $\mathbf{b}$ of some parameter space other than $\mathbf{a}$. For simplicity we assume that both $u_i(\mathbf{x}, \mathbf{a})$ and $w_i(\mathbf{x}, \mathbf{b})$ satisfy the usual continuity and symmetry conditions, although if a symmetry operator commutes with $\mathscr{H}$ it is sufficient to assume that either u_i or w_i is an eigenfunction of it.

Let us introduce the notation

$$u_{ij}(\mathbf{x}, \mathbf{a}) = \partial u_i(\mathbf{x}, \mathbf{a})/\partial a_j; \qquad (j = 0, \ldots, m), \tag{4}$$

and define the weight functions by

$$w_{ik}(\mathbf{x}, \mathbf{b}) = \partial w_i(\mathbf{x}, \mathbf{b})/\partial b_k; \qquad (k = 0, \ldots, n). \tag{5}$$

It will be assumed that the functions u_{ij} form a linearly independent set, similarly the functions w_{ik}.

The method of moments determines the values of the $\mathbf{a}$ and the approximate eigenvalue e_i from the requirement minimize

$$\sum_{k=0}^{m} (\mathrm{Re}\langle w_{ik}(\mathbf{x}, \mathbf{b}) \mid \mathscr{H} - e_i \mid u_i(\mathbf{x}, \mathbf{a})\rangle)^2 \Big/ \sum_{k=0}^{m} (\mathrm{Re}\langle w_{ik}(\mathbf{x}, \mathbf{b}) \mid u_i(\mathbf{x}, \mathbf{a})\rangle)^2 \tag{6}$$

as a function of $\mathbf{a}$ and e_i keeping $\mathbf{b}$ fixed. In the important special case $m = n$ this means the solving of the set of equations

$$\mathrm{Re}\langle w_{ik}(\mathbf{x}, \mathbf{b}) \mid \mathscr{H} - e_i \mid u_i(\mathbf{x}, \mathbf{a})\rangle = 0; \qquad (k = 0, \ldots, n = m). \tag{7}$$

Although (6) may have practical advantages even in the case $m = n$ (e.g. if the parameters $a_1, \ldots, a_m$ are linear it leads to a symmetrical eigenvalue problem) we shall consider only (7) as the analysis of (6) is more involved and probably leads to similar results.

Eqs. (7) become equivalent with the method of energy variation if $w_{ik}(\mathbf{x}, \mathbf{b}) \equiv u_{ik}(\mathbf{x}, \mathbf{a}_e)$, where $\mathbf{a}_e$ denotes the point of the parameter space $\mathbf{a}$ yielded by the method of energy variation, i.e. $\langle u_i(\mathbf{x}, \mathbf{a}) \mid \mathscr{H} - e_i \mid u_i(\mathbf{x}, \mathbf{a})\rangle$ = stationary. The practical advantage of the method of moments results from the fact that we can apply also other weight functions and use the freedom in their choice to reduce difficulties of integration. Evidently, however, the method of moments will only casually yield good results if this freedom is misused in such a way that the roots of (6) or (7) become unstable under small changes

of the weight functions. We thus have to answer two questions: a) Under which circumstances do Eqs. (7) become "ill-conditioned", i.e. cause "unfortunate" small changes of the weight functions large changes in their roots; b) under which circumstances will the roots of Eqs. (7) be good approximations to presumably "good" values. As our aim is not to improve but to simplify the most frequently used variational method, the method of energy variation, we tacitly shall assume that the "good" values of the parameters $\mathbf{a}$ and the energy e_i are $\mathbf{a}_e$ and e_{ie} yielded by the method of energy variation.

The analysis can be simplified by the fact that as Eqs. (7) are linear in the weight functions we can replace the weight functions by any linearly independent linear combinations.

$$v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a})) = \sum_{k=0}^{m} c_{ijk}(\mathbf{a}) w_{ik}(\mathbf{x}, \mathbf{b}); \qquad (j = 0, \ldots, m), \tag{8}$$

say, such which all but the 0th are orthogonal to $u_i(\mathbf{x}, \mathbf{a})$

$$\langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a})) \mid u_i(\mathbf{x}, \mathbf{a})\rangle = 0; \qquad (j = 1, \ldots, m), \tag{9}$$

and in order to make the definition of the coefficients $c_{ijk}(\mathbf{a})$ unique we assume that the new weight functions $v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}))$ have a maximum overlap with the corresponding derivatives of the wave function

$$\langle v_{ij} - u_{ij} \mid v_{ij} - u_{ij}\rangle = \text{minimum}; \qquad (j = 0, \ldots, m). \tag{10}$$

The value e_i in the last m equations of (7) can then be replaced by any other value, E_i, say

$$\mathrm{Re}\langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a})) \mid \mathscr{H} - E_i \mid u_i(\mathbf{x}, \mathbf{a})\rangle = 0; \qquad (j = 1, \ldots, m). \tag{11}$$

(This means that we can formally decouple the determination of the parameters and the energy.)

Let us denote the roots of (11) by $\mathbf{a}_m$. Writing instead of $\mathbf{a}_m$ the "good" values $\mathbf{a}_e$ into (11) the right-hand sides will be equal to some (generally non-zero) values z_j

$$\mathrm{Re}\langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}_e)) \mid \mathscr{H} - E_i \mid u_i(\mathbf{x}, \mathbf{a}_e)\rangle = z_j; \qquad (j = 1, \ldots, m). \tag{12}$$

Qualitatively the smaller the absolute values of the z_j's and the less sensitive the roots of (11) to small changes in the right-hand side the better approximation $\mathbf{a}_m$ to $\mathbf{a}_e$.

In order to obtain practically more useful, although still rather qualitative, statements let us subtract the equations $\mathrm{Re}\langle u_{ij}(\mathbf{x}, \mathbf{a}_e) \mid \mathscr{H} - E_i \mid u_i(\mathbf{x}, \mathbf{a}_e)\rangle = 0$; $(j = 1, \ldots, m)$ (obtained in a straightforward way from the basic equations of the method of energy variation) from Eqs. (11). By the Schwarz inequality we obtain the following upper bounds:

$$|z_j|^2 \leqslant \langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}_e)) - u_{ij}(\mathbf{x}, \mathbf{a}_e) \mid v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}_e)) - u_{ij}(\mathbf{x}, \mathbf{a}_e)\rangle$$

$$\times \langle u_i(\mathbf{x}, \mathbf{a}_e) \mid (\mathscr{H} - E_i)^2 \mid u_i(\mathbf{x}, \mathbf{a}_e)\rangle. \tag{13}$$

From the second term of the right-hand side of (13) it can be seen that the better approximation $u_i(\mathbf{x}, \mathbf{a}_e)$ to the exact wave function the smaller the $|z_j|$'s. Thus the better can the variational wave function approximate to the exact one the more stable the method of moments. This conclusion is very satisfactory, although entirely qualitative.

From the first term of the right-hand side of (13) it follows that in order to obtain small $|z_j|$'s the weight functions $v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}))$ must have a possibly high overlap with the corresponding derivatives of the wave function

$$1 - [(\mathrm{Re}\langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a})) \mid u_{ij}(\mathbf{x}, \mathbf{a})\rangle)^2 / (\langle v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_i(\mathbf{a})) \mid v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}))\rangle \times \langle u_{ij}(\mathbf{x}, \mathbf{a}) \mid u_{ij}(\mathbf{x}, \mathbf{a})\rangle)]^{1/N} \ll 1. \tag{14}$$

The Nth root appearing in (14) (where N denotes the number of particles of the system) attempts to make the criterion (14), as far as possible, independent of the number of particles.

In any case, the more complicated functions enter (12) the more, in general, the right-hand side of (13) overestimates the left-hand side. This is another very satisfactory qualitative statement.

Investigating the left-hand side of (12) it can be seen that the criterion (14) itself does not guarantee the stability of the roots. In order to verify this let us for a minute assume that (contradicting our basic assumptions) the set of w_{ik}'s is linearly dependent. This can still be consistent with Eqs. (9), but in any case it means that the set of v_{ij}'s is either linearly dependent or consists of less than m functions. Consequently the roots of (12) become ill-determined. It can thus be expected that if the set of w_{ik}'s is close to being linearly dependent the roots of Eqs. (12) will be ill-conditioned. For us the most important consequence of this is that their roots will be very sensitive to "unfortunate" small changes of the weight functions.

By $v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a})) \approx u_{ij}(\mathbf{x}, \mathbf{a})$, following from (14), the approximate linear dependence of the weight functions can be a consequence of an approximate linear dependence of the $u_{ij}(\mathbf{x}, \mathbf{a})$'s. This also causes serious trouble in the method of energy variation, making the results sensitive to round-off errors. The situation in the case of the method of moments is worse as the results become sensitive to small changes of the weight functions. However, as the v_{ij}'s are generally only rough approximations to the corresponding u_{ij}'s their approximate linear dependence can also be a consequence of an "unfortunate" choice of the weight functions. If the set of u_{ij}'s is near to being linearly dependent the best we can do is a probable contraction of the "almost redundant" variational parameters in u_i. If the approximate linear dependence of the weight functions is due to their "unfortunate" choice one can attempt to choose better weight functions. In any case, we shall attempt to give a practically useful semi-quantitative criterion which, at least, can signal the danger.

Let us define auxiliary weight functions $v_{ij}^{(t)}$ by

$$v_{ij}^{(t)}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j^{(t)}(\mathbf{a})) = \sum_{\substack{k=0 \\ k \neq j}}^{m} c_{ijk}^{(t)}(\mathbf{a}) w_{ik}(\mathbf{x}, \mathbf{a}); \qquad (j = 1, \ldots, m) \tag{15}$$

which also satisfy Eqs. (9)–(10). The weight functions $v_{ij}^{(t)}$ and quantities associated with them will be referred to as "truncated". If we find that the overlap between $v_{ij}^{(t)}$ and u_{ij} differs only insignificantly from the overlap between v_{ij} and u_{ij}, i.e. the criterion

$$1 - \text{the left-hand side of (14)} \ll 1 - [(\mathrm{Re}\langle v_{ij}^{(t)}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j^{(t)}(\mathbf{a})) \mid u_{ij}(\mathbf{x}, \mathbf{a})\rangle)^2 / (\langle v_{ij}^{(t)}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j^{(t)}(\mathbf{a})) \mid v_{ij}^{(t)}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j^{(t)}(\mathbf{a}))\rangle \langle u_{ij}(\mathbf{x}, \mathbf{a}) \mid u_{ij}(\mathbf{x}, \mathbf{a})\rangle)]^{1/N} \tag{16}$$

is violated, then this is a warning that the function w_{ij} is (at least in the subspace spanned by the u_{ij}'s) near to a linear combination of the other weight functions and the problem is probably ill-conditioned.

Obviously, more elaborate and probably more complicated criteria could be derived but it can hardly be believed that a weight function $\mathbf{w}_{ij}$ which can give no significant contribu-

tion to the overlap with its "own" derivative u_{ij} will be a "good" weight function. Consequently it is not sure that the use of more elaborate criteria pays off.

The calculation of the coefficients $c_{ijk}(\mathbf{a})$ means the solving of the set of linear equations

$$\sum_{k=0}^{m} c_{ijk}\langle w_{il} \mid w_{ik}\rangle = \langle w_{il} \mid u_{ij}\rangle; \qquad (l = 0, \ldots, m; j = 1, \ldots, m) \tag{17}$$

and a set of similar equations with $k, l \neq j$ in the truncated case. In practice, the linear dependence can also be signaled by the fact that Eqs. (17) are ill-conditioned.

The error in the approximate eigenvalue e_i can be estimated by similar tools. From the 0th equation of (7) it follows that

$$e_{im} = \mathrm{Re}[\langle w_{i0}(\mathbf{x}, \mathbf{b}) \mid \mathscr{H} \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle / \langle w_{i0}(\mathbf{x}, \mathbf{b}) \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle] \tag{18}$$

By the orthogonality relations (9) we can replace w_{i0} in (18) by v_{i0}. Making use of the identity $e_{im} = E_i - (E_i - e_{im})$ the error in e_{im} can be written as

$$e_{im} - E_i = \mathrm{Re}[\langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid \mathscr{H} - E_i \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle / \langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle]. \tag{19}$$

If we apply the Schwarz inequality to the numerator at the right-hand side of (19) the upper bound

$$\mid e_{im} - E_i \mid \leqslant \mid \mathrm{Re}\langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle \mid^{-1} \times$$

$$\times [\langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m))\rangle \langle u_i(\mathbf{x}, \mathbf{a}_m) \mid (\mathscr{H} - E_i)^2 \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle]^{1/2} \tag{20}$$

of the error is obtained. For the ground state by the identity $\mathscr{H} - E_0 = (\mathscr{H} - E_0)^{1/2} (\mathscr{H} - E_0)^{1/2}$ (20) can also be written as

$$\mid e_{0m} - E_0 \mid \leqslant \mid \mathrm{Re}\langle v_{00}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid u_0(\mathbf{x}, \mathbf{a}_m)\rangle \mid^{-1} \times$$

$$[\langle v_{00}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid \mathscr{H} - E_0 \mid v_{00}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m))\rangle \langle u_0(\mathbf{x}, \mathbf{a}_m) \mid \mathscr{H} - E_0 \mid u_0(\mathbf{x}, \mathbf{a}_m)\rangle]^{1/2}. \tag{21}$$

An elementary analysis of (20)–(21) yields the result that, similar to the case of the method of energy variation, the error in e_{im} contains only terms proportional to the square of the error in u_i provided that the criterion

$$[(\mathrm{Re}\langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle)^2 / (\langle v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m)) \mid v_{i0}(\mathbf{x}, \mathbf{b}, \mathbf{c}_0(\mathbf{a}_m))\rangle \times$$

$$\times \langle u_i(\mathbf{x}, \mathbf{a}_m) \mid u_i(\mathbf{x}, \mathbf{a}_m)\rangle)]^{1/N} \approx 1 \tag{22}$$

is not severely violated. Attention is again called to the fact that the more complicated functions make up the integrands the more, in general, the Schwarz inequality overestimates the integrals.

Essentially equivalent results for the error in the energy have been derived in [2] and [3] by different tools.

The criteria (14), (16) and (22) must obviously be satisfied in the neighbourhood of $\mathbf{a} = \mathbf{a}_m$. As we do not know $\mathbf{a}_m$ at the beginning of the calculations we often have to proceed as follows. We choose some starting values of the parameters $\mathbf{b}$ and calculate $\mathbf{a}_m$. Then, if the requirements $v_{ij}(\mathbf{x}, \mathbf{b}, \mathbf{c}_j(\mathbf{a}_m)) \approx u_{ij}(\mathbf{x}, \mathbf{a}_m)$ (or, in practice, the requirements $w_{ij}(\mathbf{x}, \mathbf{b}) \approx u_{ij}(\mathbf{x}, \mathbf{a}_m)$) are not fulfilled we choose new values of the $\mathbf{b}$'s and iterate.

This procedure is very similar to the self-consistent field procedure and can be carried out parallel with it. If both the parameters **a** and **b** are linear such an iterative procedure is, in general, unnecessary.

It is often possible to carry out the procedure outlined in the preceding paragraph in the following way. We construct the weight function generator (3) by making some systematic small changes in the variational wave function (2) in order to simplify the integrals and then formally replace the parameters a_i by parameters b_i. Thus we can consider both $u_i(\mathbf{x}, \mathbf{a})$ and $w_i(\mathbf{x}, \mathbf{b})$ as approximate wave functions and can try to determine the parameters in both of them by the method of moments using each other also as weight function generators. This means that we determine both u_i and w_i from the requirement

$$\mathrm{Re}[\langle w_i(\mathbf{x}, \mathbf{b}) \mid \mathscr{H} \mid u_i(\mathbf{x}, \mathbf{a})\rangle / \langle w_i(\mathbf{x}, \mathbf{b}) \mid u_i(\mathbf{x}, \mathbf{a})\rangle] = \text{stationary}. \tag{23}$$

The advantages of such a procedure are obvious but it must be handled with great care as it can yield worse weight functions than an explicit adjusting of the v_{ij}'s to the u_{ij}'s. Probably in most cases the best method is to start with (23) and if some tests signal deterioration of the weight functions we switch on explicit adjusting.

3. Applications

The calculations presented below form part of the investigations to apply the method of moments for the determination of correlated wave functions. Most of the methods applied for actual calculations are using the single-particle function expansions which severely restrict the rate of convergence (natural orbital expansion [7], pseudo-natural orbital expansion [8], extended separated pair theory [9]). The rate of convergence can partly be improved by an explicit inclusion of the interparticle co-ordinates into the trial functions [10]. This approach, however, is associated with severe difficulties of integration if the number of particles is not extremely low.

The variant of the method of moments used by Boys and Handy applies a particular form of weight functions which, to a certain extent, facilitates the evaluation of the integrals [3]. The method proposed in the present paper uses weight functions which can be decomposed into products of single-particle functions, and wave functions constructed from group orbitals, taking into account by interparticle co-ordinates the correlation of particles belonging to the same group. It can easily be verified that in this way the difficulties of integration can significantly be reduced as in the worst integrals the co-ordinates of particles belonging to two groups are inseparably coupled, while the method of energy variation leads to integrals inseparably involving the co-ordinates of all particles. The method is particularly suitable for wave functions of the cluster expansion type.

In order to test the applicability of the method of moments in the case of such weight and wave functions simple test calculations have been made for the helium atom and the negative hydrogen ion. In both cases three different approximations have been made, the wave functions $u_0(\mathbf{x}, \mathbf{a})$ being

$$\begin{aligned} &\text{Appr. A:} \quad a_0(\exp(-a_1 r_1 - a_1 r_2) \\ &\qquad + a_2 \exp(-a_1 r_1 - a_1 r_2) r_{12}) \\ &\text{Appr. B:} \quad a_0(\exp(-a_1 r_1 - a_1 r_2) \\ &\qquad + a_2 \exp(-a_1 r_1 - a_1 r_2) r_{12} \end{aligned} \tag{24}$$

$$+ a_3 \exp(-a_1 r_1 - a_1 r_2) r_1 r_2$$
$$+ a_4 \exp(-a_1 r_1 - a_1 r_2)(r_1^2 + r_2^2)) \quad (25)$$

Appr. C: $a_0(\exp(-a_1 r_1 - a_1 r_2)$
$$+ a_2 \exp(-a_1 r_1 - a_1 r_2) r_{12}$$
$$+ a_3 \exp(-a_1 r_1 - a_1 r_2) r_1 r_2$$
$$+ a_4 \exp(-a_1 r_1 - a_1 r_2)(r_1^2 + r_2^2)$$
$$+ a_5 \exp(-a_1 r_1 - a_1 r_2)(r_1 + r_2)$$
$$+ a_6 \exp(-a_1 r_1 - a_1 r_2) r_{12}^2) \quad (26)$$

Here r_1, r_2 and r_{12} denote the distances of electron 1 and 2 from the fixed nucleus and each other, respectively.

The weight function generators have been obtained in all three cases by a) replacing $\exp(-a_1 r_1 - a_1 r_2) r_{12}$ by

Appr. 0: $\exp(-c_0 a_1 r_1 - c_0 a_1 r_2) r_{12}^0$ (27)

Appr. 2: $\exp(-c_2 a_1 r_1 - c_2 a_1 r_2) r_{12}^2$ (28)

in the second terms of (24)-(26) and b) formally writing instead of a_i's, b_i's. The constants c_0 and c_2 have been chosen in several different ways and are to be discussed later. By $r_{12}^2 = r_1^2 + r_2^2 - 2(x_1 x_2 + y_1 y_2 + z_1 z_2)$ the weight functions can be decomposed into products of single-particle functions.

The six approximations belonging to the wave functions (24)-(26) and the weight function generators obtained by the replacements (27)-(28) will be denoted in an obvious way by A0, A2, B0, B2, and C0, C2.

We first investigate the behaviour of the total and truncated overlaps belonging to the substitutions (27)-(28) as functions of the constants c_0, respectively c_2. The values listed in Table 1 belong to $a_1 = 1$ but the trivial behaviour of the integrals under the change of a common scale factor ensures that the results are also similar for $a_1 \neq 1$. It is also clear that the investigations can be limited to v_{02}, the other v_{0j}'s will hardly cause any surprise. In order to blow up the differences and thus make the results more easy to compare the exponents $1/N$ in the criteria (14)-(16) have been disregarded.

The data listed in Table 1 suggest the following: a) The approximations "2" are probably better than the approximations "0". b) It is not surprising if the stability and consequently the accuracy of results in the approximation "0" decrease if the number of variational parameters increase; the approximation C0 is probably exceptionally "malevolent" as the round-off errors completely overshadow the difference between the total and the truncated overlaps and, in addition, the best value of c_0 is very close to 1.0 which may cause trouble in the numerical stability of the eigenvalue problem.

This "malevolent" behaviour of the approximations "0" could hardly be predicted if we studied only the overlaps between the weight functions and the corresponding derivatives of the wave function. The maximum overlap between the function $\exp(-r_1 - r_2) r_{12}$ and the functions $\exp(-c_0 r_1 - c_0 r_2) r_{12}^0$ and $\exp(-c_2 r_1 - c_2 r_2) r_{12}^2$ is, namely, 0.9204 and 0.9628 respectively. This can indicate a moderate superiority of the approximations "2" but by no means an extreme one.

Table 1. Truncated and total overlap values belonging to the weight functions v_{02} in the approximations A0–C2

Appr.:	A0	B0	C0		A2	B2	C2
	Trunc.[a]	Trunc.[a]	Trunc.[a]		Trunc.[a]	Trunc.[a]	Trunc.[a]
	0.2025	0.0722	0.0095		0.2025	0.0722	0.0095
c_0	Total[b]	Total[b]	Total[b]	c_2	Total[b]	Total[b]	Total[b]
1.10	0.0706	0.0634	0.0095	0.90	0.0385	0.0374	0.0061
1.05	0.0698	0.0634	0.0095	0.95	0.0274	0.0272	0.0055
0.95	0.0699	0.0634	0.0095	1.15	0.0119	0.0117	0.0043
0.90	0.0708	0.0634	0.0095	1.10	0.0080	0.0072	0.0038
0.85	0.0726	0.0636	0.0095	1.15	0.0071	0.0045	0.0033
0.80	0.0752	0.0661	0.0095	1.20	0.0094	0.0031	0.0029
0.75	0.0789	0.0722	0.0095	1.25	0.0150	0.0026	0.0026
0.70	0.0837	0.0698	0.0095	1.30	0.0241	0.0028	0.0024
0.65	0.0897	0.0684	0.0095	1.35	0.0366	0.0035	0.0023
0.60	0.0971	0.0680	0.0095	1.40	0.0522	0.0046	0.0022
0.55	0.1059	0.0681	0.0095	1.45	0.0702	0.0059	0.0022
0.50	0.1162	0.0683	0.0095	1.50	0.0900	0.0075	0.0023

[a] The right-hand side of Eq. (16) without the exponent $1/N$.
[b] The left-hand side of Eq. (14) without the exponent $1/N$.

The results of the calculations on the ground state of the helium atom and the negative hydrogen ion are listed in Table 2. All results are given in atomic units. Where no entries are given the results were either meaningless or could not be calculated because of numerical instability or insufficient convergence of the iterations. For the negative hydrogen ion no reliable expectation value of the operator $1/r_{12}$ could be found in the literature.

The columns of Table 2 correspond to the following calculations:

1. The best values of c_0, respectively c_2, taken from Table 1. Because of numerical stability problems the value $c_0 = 0.97$ may be slightly inaccurate.
2. The best values of c_0, respectively c_2, calculated from the maximum overlap between $\exp(-r_1-r_2)r_{12}$ and $\exp(-c_0r_1-c_0r_2)r_{12}^0$ respectively $\exp(-c_2r_1-c_2r_2)r_{12}^2$.
3. Arbitrarily chosen values of c_0 and c_2.
4. The values obtained by the method of energy variation.
5. The "exact" values [11].

Table 2 contains the expectation values of the Hamiltonian operator and the operators $\delta^{(3)}(r_{12})$, $r_1^2 + r_2^2$, and $1/r_{12}$.

4. Discussion

The results presented in Table 2 need little comment. They fulfil (perhaps to an unexpected extent) the predictions following from the criteria (14), (16) and (22). We have the impression that the results may be a bit "too beautiful" and do not want to overestimate the power of the criteria. In any case, the results together with the results of many other papers indicate that the method of moments, if used with precaution, can yield reliable results comparable with those obtained by the method of energy variation.

It may, perhaps, be useful to call attention to the following fact. The maximum overlap

Table 2. Results of the calculations on the helium atom and the negative hydrogen ion. The explanation is given in the text

Helium Atom

Appr. A0	Method of Moments			En. Var.	Exact
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	−2.82633	−2.82524	−2.91970	−2.89112	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.22542	0.21119	0.08006	0.11991	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	2.05774	2.08061	2.19278	2.15366	2.3870
$\langle 1/r_{12} \rangle$ =	1.09758	1.07933	0.92613	0.97429	0.94582
Appr. B0					
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	−2.89877	–	−2.89096	−2.90268	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.12264	–	0.13742	0.11539	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	2.37371	–	2.42442	2.35390	2.3870
$\langle 1/r_{12} \rangle$ =	0.95369	–	0.96424	0.94802	0.94582
Appr. C0					
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	–	–	–	−2.90333	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	–	–	–	0.11196	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	–	–	–	2.38647	2.3870
$\langle 1/r_{12} \rangle$ =	–	–	–	0.94608	0.94582
Appr. A2					
c_2 =	1.13	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−2.89004	−2.88982	−2.89020	−2.89112	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.12160	0.12195	0.12135	0.11991	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	2.16050	2.15875	2.16119	2.15366	2.3870
$\langle 1/r_{12} \rangle$ =	0.97509	0.97571	0.97475	0.97429	0.94582
Appr. B2					
c_2 =	1.27	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−2.90118	−2.90117	−2.90124	−2.90268	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.11826	0.11828	0.11826	0.11539	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	2.36403	2.36440	2.36715	2.35390	2.3870
$\langle 1/r_{12} \rangle$ =	0.94991	0.94988	0.94943	0.94802	0.94582
Appr. C2					
c_2 =	1.42	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−2.90148	−2.90077	−2.90064	−2.90333	−2.90372
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.11789	0.12048	0.12105	0.11196	0.10635
$\langle r_1^2 + r_2^2 \rangle$ =	2.37093	2.36218	2.37033	2.38647	2.3870
$\langle 1/r_{12} \rangle$ =	0.94930	0.95065	0.95031	0.94608	0.94582

Negative Hydrogen Ion

Appr. A0	Method of moments			En. Var.	Exact
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	−0.47428	−0.47440	−0.47510	−0.50878	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.01249	0.01245	0.01226	0.00389	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	12.74054	12.74470	12.76577	12.65480	23.827
$\langle 1/r_{12} \rangle$ =	0.42648	0.42621	0.42480	0.37311	–

Table 2–*continued*

Appr. B0	Method of moments			En. Var.	Exact
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	−0.51365	–	−0.45746	−0.52566	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.00584	–	0.02655	0.00320	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	17.49715	–	15.44325	18.26688	23.827
$\langle 1/r_{12} \rangle$ =	0.34767	–	0.39464	0.32447	–
Appr. C0					
c_0 =	0.97	0.75	0.50	–	–
$\langle \mathscr{H} \rangle$ =	–	–	–	−0.52646	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	–	–	–	0.00336	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	–	–	–	19.20641	23.827
$\langle 1/r_{12} \rangle$ =	–	–	–	0.32362	–
Appr. A2					
c_2 =	1.13	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−0.50974	−0.50903	−0.51006	−0.50878	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.00371	0.00384	0.00365	0.00389	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	12.72990	12.64700	12.76120	12.65480	23.827
$\langle 1/r_{12} \rangle$ =	0.37097	0.37286	0.37018	0.37311	–
Appr. B2					
c_2 =	1.27	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−0.52529	−0.52530	−0.52553	−0.52566	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.00327	0.00327	0.00322	0.00320	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	18.39629	18.41183	18.46714	18.26688	23.827
$\langle 1/r_{12} \rangle$ =	0.32433	0.32421	0.32339	0.32447	–
Appr. C2					
c_2 =	1.42	1.25	1.083	–	–
$\langle \mathscr{H} \rangle$ =	−0.52662	−0.52647	−0.52634	−0.52646	−0.52775
$\langle \delta^{(3)}(r_{12}) \rangle$ =	0.00330	0.00335	0.00340	0.00336	0.00274
$\langle r_1^2 + r_2^2 \rangle$ =	19.17686	19.14109	19.15572	19.20641	23.827
$\langle 1/r_{12} \rangle$ =	0.32375	0.32392	0.32383	0.32362	–

between $\exp(-r_1-r_2)r_{12}$ and $\exp(-c_0r_1-c_0r_2)r_{12}^0$ is obtained at $c_0 = 0.75$. Consequently from simple overlap considerations one would prefer this value. However, in the approximation B0 the difference between the total and the truncated overlap, just at $c_0 = 0.75$, is almost zero, although the total overlap itself is not small. Now, in our calculations we were unable to make the iterations convergent in approximation B0 with $c_0 = 0.75$. On the other hand at $c_0 = 0.97$ the iterations smoothly converged although the left-hand side base is almost degenerate.

It is been tested on a few examples how the results in the approximations "0" behave if more weight functions are applied than the minimum, i.e. we calculate by Eq. (6). The additional weight functions had the form $\exp(-c_0'r_1-c_0''r_2)$ $(c_0' \neq c_0'')$. In all calculations five additional weight functions have been applied with various values of c_0' and c_0'' more or less arbitrarily distributed around c_0. The accuracy and stability of the results definitely improved but did not reach the preciseness of the approximations "2".

It has also been investigated how the results behave if simple weight factors as $1/r_1$ are introduced into the criteria (14)–(16). In most cases the expectation values of those quan-

tities which are sensitive to the shape of the wave function in regions preferred by the weight factor slightly improved, but the differences were small and not convincing.

The calculations have been carried out with a 28-bit precision.

Acknowledgements. The authors are much indebted to Mr. M. Preisich and Prof. Dr. A. Kónya for continuous interest and support of their work.

Our best thanks are due to Mrs. Zs. Szabó and Mr. P. Váczi for valuable discussions and other kinds of help.

The calculations have been carried out on the GIER computer of the Computing Center of the United Hungarian Chemical Works.

References

1. Szondy, T.: Z. Naturforsch. **17a**, 552 (1962)
2. Szondy, É., Szondy, T.: Acta Phys. Hung. **20**, 253 (1966)
3. Boys, F. S.: Proc. Roy. Soc. (Lond.) **A309**, 195 (1969); Boys, F. S., Handy, N. C.: Proc. Roy. Soc. (Lond.) **A309**, 209 (1969); **A310**, 43 (1969); **A310**, 63 (1969)
4. Hegyi, M. G., Mezei, M., Szondy, T.: Theoret. Chim. Acta (Berl.) **15**, 273 (1969); **15**, 283 (1969); **21**, 168 (1971); Ladányi, K., Lengyel, V., Szondy, T.: Theoret. Chim. Acta (Berl.) **21**, 176 (1971)
5. Armour, E. A. G.: Mol. Phys. **24**, 163 (1972); **24**, 181 (1972); **25**, 993 (1973); Handy, N. C.: J. Chem. Phys. **51**, 3205 (1969); Handy, N. C., Epstein, S. T.: J. Chem. Phys. **53**, 1392 (1970)
6. Huggett, J. P., Armour, E. A. G.: Theoret. Chim. Acta (Berl.) **39**, 161 (1975)
7. Davidson, E. R.: Advan. Quantum Chem. **6**, 235 (1972) and Refs. therein
8. Meyer, W.: J. Chem. Phys. **58**, 1017 (1973) and Refs. therein
9. Kapuy, E.: Theoret. Chim. Acta (Berl.) **6**, 281 (1966); **12**, 397 (1968); Robb, M. A., Csizmadia, I. G.: Intern. J. Quantum Chem. **4**, 365 (1970); **5**, 605 (1971); **6**, 367 (1972)
10. Szász, L.: Phys. Rev. **126**, 169 (1962) and Refs. therein
11. Pekeris, C. L.: Phys. Rev. **115**, 1216 (1959); **126**, 143 (1962); **126**, 1470 (1962); Frankowski, K., Pekeris, C. L.: Phys. Rev. **146**, 46 (1966)

Received May 6, 1976