Theor Chim Acta (1989) 75: 425-432

Theoretica Chimica Acta

Orbitals that minimize the variance

D. Filberg and H. Kleindienst

Institut für Physikalische Chemie I, Universität Düsseldorf, D-4000 Düsseldorf 1, Federal Republic of Germany

(Received July 4, revised December 30, 1988/Accepted January 5, 1989)

A MCSCF-method for minimization of the variance yielding an optimal basisset of variance minimizing orbitals in analogy to Löwdin's natural orbitals for minimal energy is developed.

Key words: Lower bounds — MCSCF

1. Introduction

Minimizing the Rayleigh quotient

$$R[\Psi]=\frac{\langle\Psi|H\Psi\rangle}{\langle\Psi|\Psi\rangle}\geq E_0 \tag{1}$$

an upper bound $E_r=\langle\Psi_r|H\Psi_r\rangle$, $\|\Psi_r\|=1$ to the groundstate energy E_0 is obtained. Minimization of

$$R[\lambda,\Psi]=\frac{\langle\Psi|(H-\lambda)^2\Psi\rangle}{\langle\Psi|\Psi\rangle}=V\geq V_0 \tag{2}$$

leads to a minimal variance $\tilde{V}$, an optimal function $\tilde{\Psi}$ and a real number $\tilde{\lambda}$ with

$$\tilde{\lambda}=\langle\tilde{\Psi}|H\tilde{\Psi}\rangle\geq E_0, \qquad \|\tilde{\Psi}\|=1 \tag{3}$$

(see [1, 2]). Then, the lower bound $\tilde{E}_l=\tilde{\lambda}-\sqrt{\tilde{V}}\leq E_0$ is much better than $E_l=\lambda_r-\sqrt{V_r}\leq\tilde{E}_l\leq E_0$ with $V_r=R[\lambda_r,\Psi_r]$. This is true for an extended basis as for small basis much more. In [3] the analogon was stated using Temple's formula.

Considering the original eigenvalue problem $H\Psi-E\Psi=0$ with $\Psi\in D_H\subset L^2$ (D_H domain of H, $\|\Psi\|=1$), one has to keep in mind that the minimization of $\|H\Psi-E\Psi\|$ and not of $\langle H\Psi|\Psi\rangle$ is the adequate problem estimating accurate

energies and wavefunctions in a finite dimensional vectorspace. Therefore, there are good reasons for the calculation of orbitals that minimize the variance (VMOs) instead of the energy.

For the three-body system we were successfully using an ansatz with correlated wavefunctions dependent on distance coordinates [6, 7] but for the general many body problem this ansatz is unsuitable. Therefore, the following question arises: what is the situation like using the orbital picture? Avoiding an "explosion" of the basis dimension and the diagonalisation of almost singular matrices we took advantage of the MCSCF-method, which is successful in the calculation of upper bounds. Here, a CI-ansatz of the form

$$\Psi = \sum_{I}^{K} c_I \Phi_I \tag{4}$$

is used with the Φ_I being Slater-determinants built up by the orbitals ψ_i.

It is possible to determine suitable VMOs if the following three problems can be solved, (i) the transfer of the MCSCF-method to the minimization of V, (ii) the obtained calculus leads to a convergent procedure, (iii) sufficient small values for the variance can be obtained. It is shown that the answer to (i) and (ii) is positive, but the problem (iii) cannot be solved without highspeed supercomputers.

2. Spinorbital expansion of the variance

Using the Born-Oppenheimer approximation, the atomic hamiltonian H of a system with n electrons and the nuclear charge Z is given by

$$H = \sum_{i}^{n} h_i + \sum_{i<j}^{n} g_{ij} \tag{5}$$

with

$$h_i = -\frac{1}{2}\Delta_i - \frac{Z}{r_i}, \tag{6}$$

$$g_{ij} = \frac{1}{r_{ij}}. \tag{7}$$

Actually, the derivation uses the symmetry $g_{ij} = g_{ji}$ only.

Expressing the variance V in a spinorbital-basis $\{\psi_i\}$, first of all the operator

$$F^2 = (H-\lambda)^2 = H^2 - 2\lambda H + \lambda^2 \tag{8}$$

has to be written as a sum of one-, two-, three-, and fourelectron operators. Using (5), relation (8) leads to

$$F^2 = \sum_{i,j}^{n} h_i h_j + \sum_{i<j}^{n} \sum_{k}^{n} [g_{ij}h_k + h_k g_{ij}] + \sum_{i<j}^{n} \sum_{k<l}^{n} g_{ij}g_{kl} - 2\lambda \sum_{i}^{n} h_i - 2\lambda \sum_{i<j}^{n} g_{ij} + \lambda^2 \tag{9}$$

resulting in

$$F^2 = \lambda^2 + \sum_{i}^{n} a_i(\lambda) + \sum_{i<j}^{n} b_{ij}(\lambda) + \sum_{i<j}^{n} \sum_{k\neq i,j}^{n} c_{ij,k} + \sum_{i<j}^{n} \sum_{k<l(k,l\neq i,j)}^{n} d_{ij,kl} \tag{10}$$

after rearrangement. Thereby, the following abbreviations are used:

$$a_i(\lambda) = h_i(h_i - 2\lambda), \tag{11}$$

$$b_{ij}(\lambda) = 2h_ih_j + (h_i + h_j)g_{ij} + g_{ij}(h_i + h_j) + g_{ij}^2 - 2\lambda g_{ij}, \tag{12}$$

$$c_{ij,k} = 2g_{ij}h_k + g_{ij}(g_{ik} + g_{jk}), \tag{13}$$

$$d_{ij,kl} = g_{ij}g_{kl}. \tag{14}$$

Furthermore, Löwdin's [8] generalized spindensity matrices (SDM) of the order k with McWeeny's [9] normalization are introduced:

$$\Gamma^{(k)}(\tau_1', \ldots, \tau_k', \tau_1, \ldots, \tau_k),$$
$$k!\binom{n}{k} \int \Psi^*(\tau_1', \ldots, \tau_k', \tau_{k+1}, \ldots, \tau_n)\Psi(\tau_1, \ldots, \tau_n)\, d\tau_{k+1} \cdot \ldots \cdot d\tau_n. \tag{15}$$

The SDM can be expanded in a m-dimensional, orthonormal spinorbital-basis $\{\psi_i\}$, as shown for the case of second order:

$$\Gamma^{(2)}(\tau_1'\tau_2', \tau_1\tau_2) = \sum_{p,t}^{m} \sum_{q,u}^{m} \Gamma^{(2)}_{pq,tu} \psi_p^*(\tau_1')\psi_q^*(\tau_2')\psi_t(\tau_1)\psi_u(\tau_2) \tag{16}$$

with the $\Gamma^{(2)}_{pq,tu}$ standing for the elements of the SDM in this basis. The evaluation of the SDM elements up to the order four can be done with the expressions based on the creation- and annihilation-operators as given, e.g. by H.-J. Werner and W. Meyer [10, Appendix C].

Using the SDM-elements of order one to four and the relation (10), the variance can be written as

$$\begin{aligned} V = \langle \Psi | F^2\Psi \rangle = \lambda^2 &+ \sum_{p,t}^{m} \Gamma^{(1)}_{p,t} \langle p | a_1 t \rangle + \tfrac{1}{2} \sum_{p,t}^{m} \sum_{q,u}^{m} \Gamma^{(2)}_{pq,tu} \langle pq | b_{12} tu \rangle \\ &+ \tfrac{1}{2} \sum_{p,t}^{m} \sum_{q,u}^{m} \sum_{r,v}^{m} \Gamma^{(3)}_{pqr,tuv} \langle pqr | c_{12,3} tuv \rangle \\ &+ \tfrac{1}{4} \sum_{p,t}^{m} \sum_{q,u}^{m} \sum_{r,v}^{m} \sum_{s,w}^{m} \Gamma^{(4)}_{pqrs,tuvw} \langle pqrs | d_{12,34} tuvw \rangle \end{aligned} \tag{17}$$

which is comprehensed to

$$V = \lambda^2 + \sum_{p,t}^{m} \sum_{q,u}^{m} \sum_{r,v}^{m} \sum_{s,w}^{m} \langle pqrs | Gtuvw \rangle \tag{18}$$

with

$$G = G_{pqrs,tuvw} = \frac{\Gamma^{(1)}_{p,t}}{m^3} a_1 + \frac{\Gamma^{(2)}_{pq,tu}}{2m^2} b_{12} + \frac{\Gamma^{(3)}_{pqr,tuv}}{2m} c_{12,3} + \frac{\Gamma^{(4)}_{pqrs,tuvw}}{4} d_{12,34}. \tag{19}$$

In order to achieve (19), the orthonormality of the spinorbitals ψ_i, represented by their indices, was used. Furthermore, we took advantage of the fact that all the operators a_i $(i=2,\ldots,n)$, e.g., give the same contribution as a_1 (b_{ij}, $c_{ij,k}$, and $d_{ij,kl}$ analogously).

Due to the conventions (15) and (4) made above, the SDM like their elements are dependent on the coefficients c_I:

$$\Gamma^{(k)} = \sum_{I,J}^{K} c_I c_J {}^{IJ}\Gamma^{(k)} \tag{20}$$

with

$${}^{IJ}\Gamma^{(k)} = k!\binom{n}{k} \int \Phi_I^* \Phi_J \, d\tau_{k+1} \ldots d\tau_n. \tag{21}$$

Obviously, (19) can be written as

$$G = \sum_{I,J}^{K} c_I c_J {}^{IJ}G \tag{22}$$

yielding for the variance:

$$V = \lambda^2 + \sum_{I,J}^{K} c_I c_J \sum_{p,t}^{m} \sum_{q,u}^{m} \sum_{r,v}^{m} \sum_{s,w}^{m} \langle pqrs \,|\, {}^{IJ}Gtuvw \rangle = \lambda^2 + \sum_{I,J}^{K} c_I c_J \langle {}^{IJ}G \rangle \tag{23}$$

3. The MCSCF-equations

Varying the c_I-coefficients as well as the spinorbital basis $\{\psi_i\}$, the method of orbital rotation, known from the MCSCF procedures [11–13], can be used. According to this method, a unitary matrix U is determined with

$$\tilde{\psi} = U^T \psi. \tag{24}$$

U transforms the given m-dimensional, orthonormal basisset $\{\psi_i\}$ into the $\tilde{m}$-dimensional $(m \geq \tilde{m})$ set of the VMO. Inserting (24) into (23), the MCSCF minimization of the variance leads to a set of stationary conditions for the c_I and the U_{ij} (27, 28), which have to be solved under the following constraints:

$$Z_{ij} \doteq (U^T U)_{ij} - \delta_{ij} = 0 \qquad (i,j=1,\ldots,\tilde{m}), \tag{25}$$

$$T \doteq \sum_{I}^{K} c_I^2 - 1 = 0, \tag{26}$$

taken into account by the Lagrangian multipliers λ_{ij} $(i,j=1,\ldots,\tilde{m})$ and η:

$$\frac{\partial}{\partial U_{\mu\nu}} \left[V - \sum_{i,j}^{\tilde{m}} \lambda_{ij} \{ (U^T U)_{ij} - \delta_{ij} \} \right] = 2A_{\mu\nu} - 2(U\Lambda)_{\mu\nu} \equiv 0, \tag{27}$$

$$(\mu = 1,\ldots,m;\ \nu = 1,\ldots,\tilde{m}),$$

$$\frac{\partial}{\partial c_\mu}\left[V-\eta\left(\sum_I^K c_I^2-1\right)\right]=2\sum_I^K c_I\langle {}^{I\mu}G\rangle-2\eta c_\mu\equiv 0 \qquad (\mu=1,\ldots,K). \tag{28}$$

Λ is the symmetric matrix of the λ_{ij} and $A_{\mu\nu}$ is given by:

$$A_{\mu\nu}=\frac{1}{2}\frac{\partial V}{\partial U_{\mu\nu}}$$

$$=\sum_{I,J}^K c_I c_J \sum_{p,\ldots,w}^{\tilde m}\langle(\delta_{p\nu}\mu\tilde q\tilde r\tilde s+\delta_{q\nu}\tilde p\mu\tilde r\tilde s+\delta_{r\nu}\tilde p\tilde q\mu\tilde s+\delta_{s\nu}\tilde p\tilde q\tilde r\mu)|{}^{IJ}G\tilde t\tilde u\tilde v\tilde w\rangle. \tag{29}$$

In order to eliminate the unknown λ_{ij}, (27) is written as a matrix equation:

$$A-U\Lambda=0. \tag{30}$$

This leads to

$$B\doteq A-UA^TU\equiv 0 \tag{31}$$

equivalent to

$$B_{\mu\nu}\doteq A_{\mu\nu}-(UA^TU)_{\mu\nu}\equiv 0 \qquad (\mu=1,\ldots,m;\ \nu=1,\ldots,\tilde m). \tag{32}$$

As follows from the constraints (25, 26), just $\tilde m\cdot m-(\tilde m/2)(\tilde m+1)$ elements U_{ij} and $K-1$ coefficients c_I are independent between $-1\le U_{ij},\ c_I\le 1$. Therefore, the variation of V yields only a corresponding number of independent equations instead of the $K+\tilde m\cdot m$ (28, 32). The missing relations are given by the constraints (25, 26). Substitution of the redundant equations in (28, 32) results in the following nonlinear system in the unknown U_{ij} and c_I:

$$B_{\mu\nu}=A_{\mu\nu}-(UA^TU)_{\mu\nu}=0 \qquad (\mu=\nu+1,\ldots,m;\ \nu=1,\ldots,\tilde m), \tag{33}$$

$$Z_{\mu\nu}=(U^TU)_{\mu\nu}-\delta_{\mu\nu}=0 \qquad (\mu=1,\ldots,\nu;\ \nu=1,\ldots,\tilde m), \tag{34}$$

$$Y_\mu=\sum_I^K c_I\langle {}^{I\mu}G\rangle-\eta c_\mu=0 \qquad (\mu=1,\ldots,K-1), \tag{35}$$

$$T=\sum_I^K c_I^2-1=0. \tag{36}$$

The Lagrangian multiplier η is given by

$$\eta=V-\lambda^2 \tag{37}$$

obtained from (35) using (26) and (23). The iterative solution of the nonlinear equations (33 ... 36) is accomplished using the Newton-Raphson method with a special damping scheme, which guaranteed quadratic convergence [14]. Thus problem (ii) was solved.

4. Calculations

The method described above was tested for the simple case of the helium atom. Here, the spinfree wavefunction Ψ for the singlet-state can be written

$$\Psi(1,2)=\sum_{I}^{K} d_I \chi_I(1)\chi_I(2) \tag{38}$$

with the NOs χ_I from Löwdin [15, 16]. The application of the ansatz (38) leads to numerous simplifications for the variance and the equations derived above [14]. The MCSCF minimization of the variance starts with the coefficients d_I and NO χ_I, which are estimated for minimum energy by the method of Ahlrichs and Driessler [17]. Our intention was to develop an expression for the wavefunction in analogy to (38), with the VMO $\tilde{\chi}_I$ and the corresponding $\tilde{d}_I$ for minimum variance, e.g.

$$\tilde{\Psi}(1,2)=\sum_{I}^{K} \tilde{d}_I \tilde{\chi}_I(1)\tilde{\chi}_I(2). \tag{39}$$

The calculations were performed for two different cases: (a) the radial limit case with $H=h_1+h_2+(1/r_>)$, (b) the groundstate for the He-atom with $H=h_1+h_2+1/r_{12}$.

Case (a) gives an example of a CI-expansion of excellent convergence. Here, the minimization of the variance was performed in the Hilbert space spanned by the radial functions $\Phi(r)$. Using a sixteen-dimensional AO-basis

$$\Phi_n = r^{n-1}\, e^{-\alpha r} \qquad (n=1,\ldots,16) \tag{40}$$

leads to the energy and variance values presented in Table 1. Here, K VMOs are generated from the first 12 NOs. The value $\tilde{E}(K=8)$ is in good agreement with the energy -2.8790286 a.u. given by Silverstone et al. [18], but the variance, however, is not very small (a value of ca. 10^{-8} a.u. might be called "very small").

In contrast to the radial limit, the CI-expansion of the helium ground-state is slowly convergent. For the calculations in case (b) the real STO-basis $\Phi_{n,l,m}=r^{n-1}\, e^{-\alpha r} P_l^{|m|}(\cos\vartheta)\, T_m(\varphi)$ with $T_m(\varphi)=\sin(|m|\varphi)(m<0)$ or $T_m(\varphi)=\cos(|m|\varphi)(m\geq 0)$ and $n=1,2,\ldots;\ l=0,1,\ldots,n;\ m=-l,\ldots,l;\ \alpha=2.59$ was

Table 1. Energy and variance values in the radial-limit case in a.u.

K	$-E_r^{a}$	V_r^{b}	$-\tilde{E}^{c}$	$\tilde{V}^{d}$
2	2.877925	2.731E−2	2.877464	2.247E−2
4	2.878980	4.883E−3	2.878929	2.885E−3
6	2.879021	1.626E−3	2.879012	8.227E−4
8	2.879026	7.413E−4	2.879024	5.982E−4

[a] Energy values obtained by the Ritz minimization

[b] Variance values calculated from wavefunctions obtained by the Ritz minimization

[c] Energy values calculated from wavefunctions obtained by the variance minimization

[d] Variance values obtained by the variance minimization

Table 2. Energy and variance values for the He groundstate in a.u.

K	$-E_r^a$	V_r^b	$-\tilde{E}^c$	$\tilde{V}^d$
2	2.877505	4.511E−1	2.824531	3.645E−1
4	2.890651	3.360E−1	2.864241	2.750E−1
6	2.897797	2.620E−1	2.880824	2.150E−1
8	2.898891	2.361E−1	2.884663	1.959E−1

[a–d] See Table 1

used, including the sixteen orbitals 1*s*, 2*s*, 2*p*, 3*s*, 3*p*, 3*d*, 4*s*, 5*s*. The results of the MCSCF minimization of the variance - obtained with the corresponding sixteen-dimensional NO-basis - are listed in Table 2. The integrals needed for this calculation are well known except those containing $1/r_{12}^2$, which are given in [19–21].

5. Discussion

In the present paper it is shown that the formalism of MCSCF can be adapted to the minimization of the variance. Furthermore, the procedure developed is of quadratic convergence.

Remark. The convergence is a crucial point, because, e.g. the extension of the fast and elegant procedure of Ahlrichs and Driessler [17] to our problem failed. No convergence could be achieved for VMOs.

However, the main problem of the method derived is the computational effort, which turned out to be immense. Rapid increasing CPU-times prevent an extension to basis sets much greater than used in our calculations on a Siemens 7580 S. Therefore, an application of the method to more complicated systems yielding small variances seems not to be practicable. Such calculations may be possible on a supercomputer, but for our opinion an analogous calculus based on correlated wavefunctions will be more advantageous.

Recently, Kutzelnigg [4] calculated a good upper bound $E_r = -2.9037240$ a.u. for the groundstate energy E_0 of the He-atom with correlated wavefunctions using a relative small basis (ca. 80 configurations). This leads to a much better variance value of $V_r \simeq 10^{-4}$ [5].

References

1. Kleindienst H, Altmann W (1976) Int J Quant Chem 10:873
2. Kleindienst H, Baumgarten E, Altmann W (1976) Int J Quant Chem 10:887
3. Kleindienst H, Müller W (1980) Theor Chim Acta 56:183
4. Kutzelnigg W (1985) Theor Chim Acta 68:445
5. Wolpers M (1988) Varianzminimierung mit r_{12}-abhängigen Wellenfunktionen. Diplomarbeit, Universität Düsseldorf
6. Kleindienst H, Pauli G, Müller W (1981) Theor Chim Acta 59:451
7. Kleindienst H, Hoppe D (1986) Theor Chim Acta 70:221

8. Löwdin PO (1955) Phys Rev 97:1474
9. McWeeny R (1960) Rev Mod Phys 32(2):225
10. Werner HJ, Meyer W (1980) J Chem Phys 73:2342
11. Hinze J (1973) J Chem Phys 59(12):6424
12. Knowles PJ, Werner HJ (1985) Chem Phys Lett 115(3):259
13. Werner HJ, Knowles PJ (1985) J Chem Phys 82(11):5053
14. Filberg D (1988) Varianzminimierende Orbitale. Dissertation, Universität Düsseldorf
15. Löwdin PO (1955) Phys Rev 97:1474
16. Löwdin PO, Shull H (1956) Phys Rev 101:1730
17. Ahlrichs R, Driessler F (1975) Theor Chim Acta 36:275
18. Silverstone HJ, Carroll DP, Silver DM (1978) J Chem Phys 68:616
19. Pauli G, Kleindienst H (1984) Theor Chim Acta 64:481
20. Matcha RL, Daiker KC (1974) J Math Phys 15:114
21. Steinborn EO, Filter E (1975) Theor Chim Acta 38:261