Theoret. chim Acta. (Berl.) 11, 135—144 (1968)

The Determination of SCF LCAO Solutions for Open Shell Configurations

DEREK H. SLEEMAN*

Electronic Computing Laboratory, The University, Leeds 2, England

Received November 21, 1967/March 28, 1968

A series of calculations has been carried out using the McWeeny and Roothaan Open Shell methods for determining LCAO-SCF solutions. The work reported here suggests that in order to obtain convergence to the absolute minimum with the McWeeny method it is often necessary to provide a very good initial approximation and that the rate of convergence could be improved by replacing the Steepest Descent by a more powerful minimization technique. Further the Roothaan method does not converge in all cases considered and it appears that Extrapolation techniques accelerate convergence but do not induce convergence in cases displaying oscillatory behaviour.

LCAO-SCF-Rechnungen an offenen Schalen wurden nach den Methoden von Roothaan und McWeeny durchgeführt. Um bei dem McWeenyschen Verfahren Konvergenz zum absoluten Minimum zu erhalten, ist oft eine schon sehr gute Anfangsnäherung nötig. Die Methode des steilsten Abstiegs ist keineswegs optimal. Extrapolationsmethoden können wohl eine Konvergenz beschleunigen, nicht aber das beim Roothaan-Verfahren manchmal auftretende Oszillieren unterdrücken.

On a exécuté des calculs LCAO SCF en couches ouvertes avec les méthodes de Roothaan et de McWeeny. Dans la méthode de McWeeny il est souvent nécessaire d'avoir une bonne approximation initiale, pour obtenir une convergence vers le minimum absolu. La méthode de descente le plus escarpé n'est pas optimale. Les méthodes d'extrapolation peuvent accélérer la convergence en effet, mais ils ne suppriment pas l'oscillation, qui arrive quelquefois dans la méthode de Roothaan.

I. Introduction

The aims of both the Roothaan [1] and McWeeny [2] Open Shell methods are:

a) to determine the *absolute* minimum of the total energy of the system and

b) to maintain orthonomality of the molecular orbitals.

Roothaan had a further aim that was:

c) to formulate a Hamiltonian such that *both* open and closed vectors were eigenvectors of it.

The reader is referred to the original papers for details of both methods, however to facilitate discussion some points about the Density Matrix method are given here.

In his Open Shell theory McWeeny[1] has distinguished between the different variational procedures which are possible when one has two or more physically distinct shells, and has considered the following possibilities:

a) The symmetrical case where all the shells are varied simultaneously, and where it is necessary to use a mean λ so as to reduce the complexity of the formulation (this is further discussed in Section II.1).

---

* Formerly: The Chemistry Department, King's College, London W.C. 2.

[1] As it stands the theory only applies to special cases of open shell systems. Sleeman [3] has suggested a modification which allows the general case to be considered.

b) The asymmetrical case where a shell (or degenerate shells) is varied and the others remain fixed. McWeeny has shown that in order to maintain idempotency of the modified D.M. $\delta \boldsymbol{R}_i$ must equal

$$-\lambda_i(1+\lambda_i^2 \boldsymbol{L}_i^2)^{-1}[\boldsymbol{L}_i+\lambda_i \boldsymbol{L}_i \boldsymbol{M}_i] \tag{1}$$

or if one assumes that $\lambda_i^2 \boldsymbol{L}_i^2 \to 0$ then

$$\delta \boldsymbol{R}_i = -\lambda_i[\boldsymbol{L}_i+\lambda_i \boldsymbol{L}_i \boldsymbol{M}_i] \tag{2}$$

where $\boldsymbol{R}_i$ and λ_i are the Density Matrix and step-length respectively for shell, i, $\boldsymbol{L}_i$ and $\boldsymbol{M}_i$ are matrices which depend upon various factors including the density matrix, the Hamiltonian and the occupation number for the shell (see [2] and [3] for details).

The formulae derived above also apply to degenerate sub-shells provided the summation is over *all* the sub-shells. Further, although McWeeny originally derived his open-shell theory for an orthogonal basis, the same formulation can be used with a non-orthogonal basis provided the integrals are transferred to the orthogonal basis. Löwdin [4] has shown that the transformations are:

$$\begin{aligned} \boldsymbol{h}_{(\text{orthog})} &= \boldsymbol{S}^{-\frac{1}{2}} \boldsymbol{h}_{(\text{non-orthog})} \boldsymbol{S}^{-\frac{1}{2}}, \\ \boldsymbol{R}_{(\text{orthog})} &= \boldsymbol{S}^{+\frac{1}{2}} \boldsymbol{R}_{(\text{non-orthog})} \boldsymbol{S}^{+\frac{1}{2}}, \end{aligned} \tag{3}$$

where $\boldsymbol{S}$ is the overlap matrix between the non-orthogonal basis-functions.

The computer programs for the iterative procedures were designed to be as general as possible and therefore we programmed the generalized form of McWeeny's asymmetrical method[2] and incorporated the transformation discussed above. For similar reasons these same transformations were incorporated into the program for the Roothaan's Open Shell method,

Slater Type Orbitals were chosen as basis functions for both the methods and it was found necessary to use an approximation to calculate some of the two-electron integrals. In the original programs the Mulliken [5] approximation was used.

II. Discussion of Results

The systems which were investigated initially were the OH^+, $OH^{\cdot}$ and OH^- configurations. The results which were obtained by both the McWeeny and Roothaan methods are summarized in Figs. 1, 2 and 3 respectively[3]. For these configurations the McWeeny method has always converged to a minimum, however slowly, and the Roothaan method converged very rapidly for certain cases. Prior to a detailed analysis of the convergence, the following points are suggested as possible reason for the convergence problems:

a) Relevant to both Methods

1. Rounding errors introduced in the iterative procedure.
2. The choice of initial Density Matrices,

b) Relevant to the Density Matrix Method

The higher-order terms which have been neglected are not negligible.

---

[2] Initially with the approximated form for $\delta \boldsymbol{R}_i$ (Eq. 2).

[3] In general, results of calculations will be given in Table 3 as final total energies and the number of cycles required for convergence. Further information can be obtained from the author on request.

II.1 Comment on the Analysis

With both methods, but more particularly with the Roothaan method, difficulties were experienced due to rounding-off errors. The peculiar difficulty noted with the Diagonalization method was due to the inaccuracies introduced by the actual Diagonalization procedure and this was overcome by increasing the precision of this procedure. The common difficulties were due to the inaccuracies of the basic atomic integrals and to small rounding errors introduced at each cycle by the iterative procedures themselves. This first difficulty was overcome by increasing the number of terms in the various analytical expressions for the integrals and the second was solved by setting small elements to zero at each iteration.

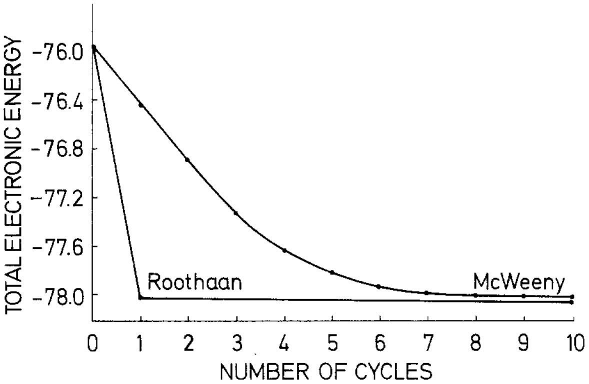

Fig. 1. Comparison of the rates of convergence with the McWeeny and Roothaan methods for the OH^+ configuration

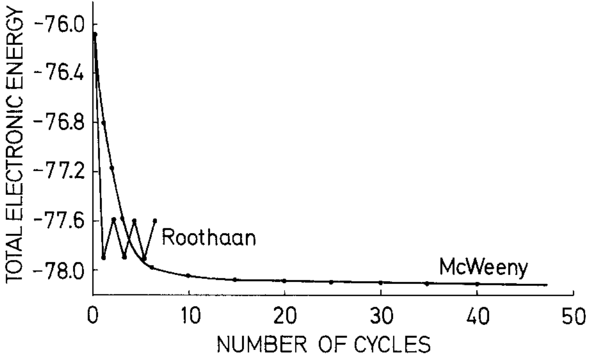

Fig. 2. Comparison of the behaviour observed with the McWeeny and Roothaan methods for the $OH^{\cdot}$ configuration

The choice of the initial density matrices is very important in the case of an open-shell as there may well be more than one configuration to which the procedure could converge. However, it is extremely difficult to choose initial density matrices for an open-shell system and it may well be necessary to carry out a calculation on the associated closed-shell system: open-shell calculations would then be carried out for each of a series of appropriate trial density matrices, and it would seem likely that the lowest total energy for the series would correspond to the ground state of the open-shell configuration.

The peculiar phenomenon displayed by the Diagonalization method is oscillation. Oscillation in supposedly converging functions (as in Fig. 2) is well known in various branches of Mathematics and an extrapolation method has been suggested by Aitken [6].

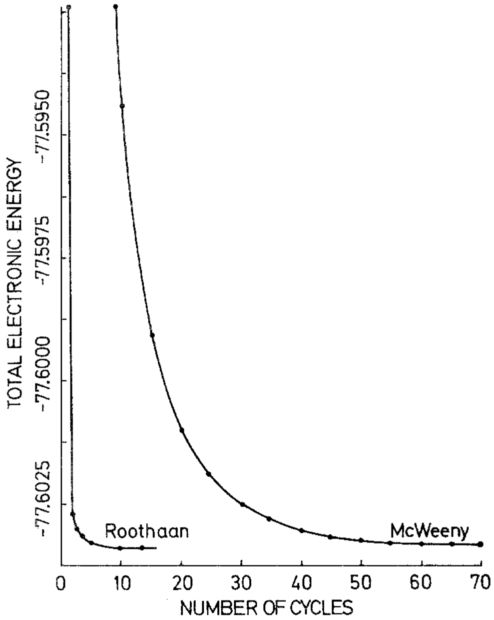

Fig. 3. Comparison of the rates of convergence with the McWeeny and Roothaan methods for the OH^- configuration

Roothaan and Bagus [7] have adapted it for use in this context, the relationship which they give is:

$$t_p = t_{i+2,p} + \frac{(t_{i+2,p} - t_{i+1,p})^2}{(2t_{i+1,p} - t_{i,p} - t_{i+2,p})} \tag{4}$$

where $\boldsymbol{t}$ is a vector representing all m occupied orbitals, p is a particular element of this "supervector" and i is the SCF cycle number. Before extrapolation it is necessary to standardize the sign of the largest element of each of the vectors, and after extrapolation an ortho-normalization of the vectors is necessary.

Nesbet [8] pointed out that some of the difficulties experienced above could be overcome by applying extrapolation to the D.M. directly; using a similar notation to that given above, we have:

$$\boldsymbol{R} = \boldsymbol{R}_{i+2} + \frac{(\boldsymbol{R}_{i+2} - \boldsymbol{R}_{i+1})^2}{(2\boldsymbol{R}_{i+1} - \boldsymbol{R}_i - \boldsymbol{R}_{i+2})}. \tag{5}$$

After the extrapolation procedure the vectors must once more be orthognormalized; however, this is simpler here as it is only necessary to make the D.M. idempotent.

Eqs. (4) and (5) have been re-arranged, following Csizmadia [9] so that limits may be applied to the extrapolation. Obviously a zero denominator must be avoided as must any change which would be large compared with the elements being extrapolated. Unfortunately, the limits of extrapolation have to be determined empirically – which is rather unsatisfactory.

For either method the extrapolation procedure is carried out every 3 cycles and the result is used as input for the next iteration. For the OH˙ radical oscillation is not overcome by either extrapolation procedure; in fact, the behaviour recorded is very similar to that of Fig. 2. However, the rate of convergence of OH^- has been considerably improved by extrapolation (see Table 3).

In order to see whether these observations were generally applicable the HCN (bend and straight) and HCOO˙ [4] configurations were considered. In view of the lack of convergence with the Roothaan method for the closed-shell HCN configurations and further convergence with the open-shell HCOO˙ it would seem that the Diagonalization method converges very rapidly for *some* closed – and open-shell configurations; however, the conditions are yet to be understood.

Again the extrapolation procedures did not induce convergence in the case of the HCN configurations and thus it seems we can deduce that extrapolation accelerates convergence but does not help in case of divergence.

With the McWeeny method the problem which has to be investigated is the slow rate of convergence. As has already been mentioned the initial computer program written incorporated the generalized asymmetrical approach (with the approximation for λ). With this program it was noticed that frequently the changes made to the Density Matrices for the OH calculations were such that idempotency was not maintained rigorously (and thus idempotency had to be restored at each cycle). Further, when a larger configuration, HCOO˙ was considered for large steplengths the modified Density Matrix was found to be outside "the radius of convergence" of the idempotency routine (i.e. the values of the matrix elements diverged). McWeeny [10] has indicated that reducing the step-lengths merely reduced *the rate of convergence* and so it was decided to limit the maximum value of λ permissible. Table 1 summarises the results obtained from one trial D.M. for HCOO˙:

Table 1

Maximum size of λ permitted	Behaviour of iterative procedure
0.1, 0.25	Begins to descend, however λ for 2nd cycle becomes negative
0.50	The energy oscillates between two values: eventually the $\boldsymbol{R}$ formed is outside the radius of convergence of the idempotency routine
0.70	Descends monotonically
0.75	Begins to descend: $\boldsymbol{R}$ for 2nd cycle is outside the radius of convergence of the idempotency routine
3.9485 (unlimited)	The modified $\boldsymbol{R}$ is outside the radius of convergence of the idempotency routine

---

[4] In the basis set used, the 1*s*-orbital was made orthogonal to the 2*s*-orbital. (The 1*s*-orbital is commonly said to be factorized out.)

The above results are capable of a number of interpretations: however it is clear that it is necessary to choose λ empirically so that the D.M. in *just* within the limits of the idempotency routine and that the choice of step-length might well be *crucial* in determining the actual configuration to which the method eventually converged. As this procedure seemed to be a very unsatisfactory one it was decided to include the full expression for λ as given in Eq. (1). (The approximation assumes that $\lambda_i^2 L_i^2 \to 0$, and although this may well be a very good approximation when self consistency is being approached it is *not* usually true at the beginning of an iterative procedure.)

When the full expression was included in the program it was found that the initial descents, for the OH, HCN, and HCOO systems, were steeper but that as the iterative procedure continued the descents were identical; however, idempotency was now maintained rigorously to second order,

Further, the observations noted in this section indicate very strongly that the symmetrical treatment suggested by McWeeny (see [2]), where a *mean* step-length is used, may lead to the wrong energy surface and that the procedure suggested here may be the only workable one.

II.2. Investigation of the Convergence Problem with the $HCOO^{\cdot}$ *System*

An inspection of the results for the $HCOO^{\cdot}$ system (see Table 3) will show that the two methods have converged to *different* total energies corresponding to totally different occupations of the molecular orbitals. In case the difficulty encountered was peculiar to the Open-shell problem it was decided to carry out the comparable calculation for the (factorized-out) $HCOO^-$ configuration. As the same difficulties were encountered a thorough investigation of the problem was made:

a) Cross-Check on Programming

One great problem with programs of this complexity is to be absolutely sure they are "bug-free" and so further extensive cross-checks were made at this point between the programs described above and their closed-shell counterparts. These checks reaffimed our belief in the open-shell programs.

b) Compatibility of Final States

The initial density matrix used in both the $HCOO^-$ calculations contained less than the maximum possible number of π-electrons. The final density matrix from the Roothaan method incorporated the correct number, whereas the D.M, from the McWeeny calculation contained the initial number. This is because a re-arrangement of vectors is always possible with the Roothaan method, but no such re-arrangement is possible for the McWeeny method with non-interacting shells. The trial D.M. for the McWeeny method was modified so as to contain 6 π-electrons and the results obtained for $HCOO^-$ are summarized in Fig. 4. Similarly, the initial Density Matrices for $HCOO^{\cdot}$ were chosen to be the same and the McWeeny method gave the original π-orbital as the open one, and the Roothaan method gave a σ-orbital.

c) Local Minima and Choice of Initial D.M.

In order to establish whether the McWeeny method was sensitive to the initial approximation, various Density matrices obtained from the Roothaan method were used as starting points for the McWeeny method. When the Density

Matrix from the first iteration of the Roothaan method, indicated in Fig. 4 by C, was used as an initial approximation the McWeeny method still converged to the result indicated by A. However, when the second Roothaan D.M., D, was used the McWeeny method converged to the same result as that obtained by the Roothaan method (i.e. B). N.B. C is "above" A and D "below" it.

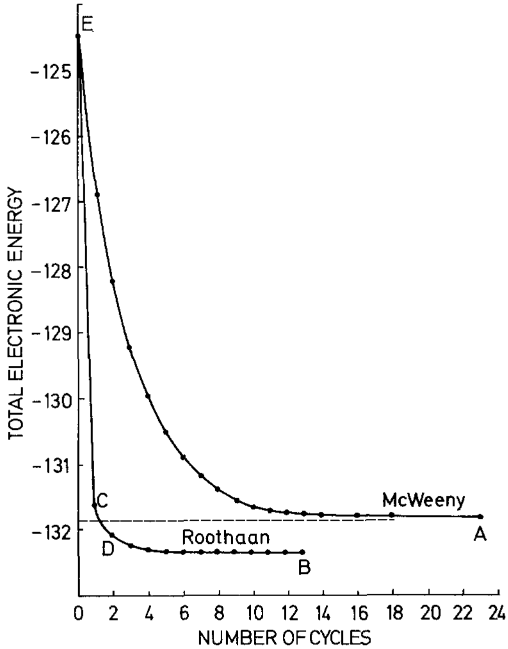

Fig. 4. Comparison of the convergence obtained with the McWeeny and Roothaan methods for the factorized-out HCOO⁻ configurations

Secondly, when the D.M. corresponding to the minimum total energy obtained by the McWeeny method, A, was used as the initial D.M. for the Roothaan method, this continued to descend until it reached its former value, B, as indicated in Fig. 4. These results are summarized in Table 2.

Table 2

Initial state	Final state obtained by the Roothaan method	Final state obtained by the McWeeny method
E	B	A
C	B	A
D	B	B
A	B	A

All letters refer to points on Fig. 4.

Interpretation of the Results

From the first piece of evidence we might suggest that the McWeeny method is sensitive to the initial Density Matrix chosen and that in the cases where the initial D.M. is a long way from the solution the iterative procedure may well converge to a local minimum. However, the second piece of evidence suggests that the solution obtained is *not* a local minimum at all but that the rate of convergence of the McWeeny method in the region of A is such as to *appear* to all the usual criteria, that the method has converged.

Nevertheless, the above results suggest that to obtain convergence to the absolute minimum with the McWeeny method it is necessary to choose a very good initial approximation,

A Further Analysis of the Convergence Difficulties of the McWeeny Method

At a number of points in his derivation of λ McWeeny has truncated the series and has only included terms up to second order. As we have already noted the resultant solution is, in fact, very dependent upon the λ chosen and therefore these approximations may well be of considerable importance.

Table 3. *The table summarises all the configurations and methods which have been considered. Total energies which are consistent to less than 1 in the n^{th} place are quoted to $n-1$ decimal places. (N.B. Due to the various problems of convergence this does not necessarily mean that they are self consistent to the same accuracy)*

Configuration	Number of cycles				Total electronic Energy a.u.
	McWeeny	Roothaan			
		un-extrapolated	Extrapolation I	Extrapolation II	
OH^{+} [a]	68	10	—	—	− 78.01900
$OH^{\cdot}$ [a]	65	no convergence	no convergence	no convergence	− 78.13081
OH^{-}	80	14	8	—	− 77.60349
HCN (straight)	60	no convergence	—	—	− 116.37
HCN (bent)	60	no convergence	no convergence	—	− 114.4
$HCOO^{\cdot}$ (factorized out)[a]	32				− 128.3008
		7	—	—	− 129.81273
$HCOO^{-}$ (factorized out)	23				− 131.830
		13	—	—	− 132.25938
BN [b]	70	no convergence	—	—	− 93.04671

[a] Open-shell configurations.

[b] Using the Rüdenberg approximation.

However, quite apart from these approximations the McWeeny formulation as it stands is dependent upon the Steepest Descent being a workable method for determining the minima of multi-variable general functions from, as often as not, poor initial approximations. There is substantial evidence, see for instance Fletcher and Powell [11], that this is *not* the case. There are a number of alternative methods available for function minimization including the Conjugate-Gradient method [12] and one due to Davidon[5]. In conjunction with Fletcher the problem of determining SCF LCAO solution for closed-shells was formulated such that the Conjugate-Gradients method could be used. The convergence[6] obtained with the H_2 and LiH molecules was considerably better than that obtained with

---

[5] Discussion of and reference to this method are given in Ref. [11].

[6] A detailed report of this work is in course of preparation.

the Steepest Descent method and only marginally worse than the very steep descent obtained with the Roothaan method.

A further problem, the eigenvalue problem, which the Steepest Descent method is known not to be able to deal with satisfactorily has recently been solved using the Conjugate-Gradient method [13]. This is further evidence that the difficulties in minimization are caused by inefficiency of the method of Steepest Descent rather than by any ill-behaviour of the function to be minimized, and suggests that this may well be a significant contributory factor to the problem reported here of slow convergence with the OH and HCN systems and to the more serious problems encountered with the larger HCOO [7] system.

Conclusion and Suggestions for Further Investigations

1. Care should be taken to avoid rounding errors from being introduced into iterative procedures.
2. For both the iterative methods, but more particularly for the McWeeny method, it is very important that a good initial density matrix should be chosen.
3. For the McWeeny method it would be very interesting to see how the convergence would be improved by removing the remaining approximations,
4. Following the encouraging initial investigations reported above it would be of great interest to consider larger configurations using a more powerful minimization technique (e.g. Conjugate-Gradients) and to eventually extend this approach to consider open-shells. However, the difficulty of applying a minimization technique to these problems is in the application of the auxiliary conditions.
5. It would be very informative were it possible to formulate the precise conditions under which the Roothaan Diagonalization method would converge.

Acknowledgements. The author would like to express his gratitude to Dr. T. E. Peacock [8] under whose direction a part of this work was carried out, to Dr. M. P. Melrose [8] both for valuable discussion and use of certain programs and routines, to Dr. R. Fletcher [9] for discussions on the Numerical Analytical aspects of the work and to the University of London Institute of Computer Science for time on its "Mercury" and "Atlas" computers.

References

1. Roothaan, C. C. J.: Rev. modern Physics **32**, 179 (1960).
2. McWeeny, R.: Proc. Roy. Soc. (London) A **241**, 239 (1957).
3. Sleeman, D. H.: Univ. of London. Ph. D. Thesis 1965.
4. Löwdin, P. O.: J. chem. Physics **18**, 365 (1950).
5. Mulliken, R. S.: J. Chim. physique **46**, 497 (1949).
6. See Hatree, D. R.: "Numerical Analysis". Clarendon Press 1958, or any standard Numerical Analysis text.
7. Roothaan, C. C. J., and P. S. Bagus: Methods in Computational Physics, Academic Press, **2**, 47 (1963).

---

[7] The results obtained for the BN molecule using the Rüdenberg [14] approximation were very similar to those for OH˙ reported above. These observations substantiate the remarks made above and provide some further evidence that the behaviour noted is a characteristic of the iterative methods themselves.

[8] King's College, London.

[9] The University of Leeds.

8. Nesbet, R. K.: Quarterly Progress Reports, Solid State and Molecular Theory Group, M.I.T. **18** (1955).
9. Csizmadia, I. G.: Private communication (1964).
10. McWeeny, R.: Physic. Rev. **114**, 1528 (1959).
11. Fletcher, R., and M. J. D. Powell: Computer Journal **6**, 163 (1963).
12. —, and C. M. Reeves: Computer Journal **7**, 149 (1964).
13. Bradbury, W. W., and R. Fletcher: Numerische Mathematik **9**, 259 (1966).
14. Rüdenberg, K.: J. chem. Physics **19**, 1433 (1951).

Dr. Derek H. Sleeman
Electronic Computing Laboratory
The University of Leeds
Leeds 2, England