Theoret. Chim. Acta (Berl.) 39, 93—102 (1975)

The Convergence Properties of Direct Energy Minimisation with Respect to Linear Coefficients in the MC-LCAO-MO-SCF Approach

B. T. Sutcliffe

Department of Chemistry, University of York, England

Received February 26, 1975/April 25, 1975

It is shown that in the LCAO-MO-MC-SCF problem, if the molecular orbital orthonormality constraints are introduced in the manner suggested by Kari and Sutcliffe or indeed by any similar method then the Hessian of the problem with respect to the linear coefficients is singular. The nature of this singularity is analysed and it is shown that in general it is possible to remove it in a level-shifting-like scheme, but that only in certain special cases is this procedure likely to be quickly convergent.

Key words: MC approach—Direct energy minimisation within the MC approach

1. Introduction

In a recent paper [1] Sutcliffe produced arguments to account for the slow convergence of conjugate gradient and conjugate direction methods (for a survey see [3]) when used to optimise the energy with respect to linear coefficients in the LCAO-MO-SCF approach. The object of this short paper is to generalise the results obtained in the previous paper to the case of a general energy functional depending on a set of orbital coefficients. The algebra involved in deriving the results presented below is tedious in the extreme and we shall confine ourselves simply to presenting the main results.

It will be remembered that it is the object of most direct minimisation methods, when applied to an unconstrained function $f(\boldsymbol{x})$, to determine a direction $\boldsymbol{p}$ from any chosen point $\boldsymbol{x}=\boldsymbol{a}$ such that the new point

$$\hat{\boldsymbol{a}}=\boldsymbol{a}+\alpha\boldsymbol{p} \tag{1.1}$$

has the property that, for suitably chosen α,

$$f(\hat{\boldsymbol{a}})<f(\boldsymbol{a})\,. \tag{1.2}$$

In the limit $\hat{\boldsymbol{a}}$ should become the point at which the function has a minimum. It will also be remembered that if the function can be expanded in a Taylor series about the minimum then for any point $\boldsymbol{a}$ sufficiently close to the minimum, the direction of descent to the minimum may be given by:

$$\boldsymbol{p}=-\boldsymbol{H}^{-1}\boldsymbol{g}(\boldsymbol{a}) \tag{1.3}$$

where $\boldsymbol{H}$ is the matrix of second derivatives of the function evaluated at the minimum point (the Hessian at the minimum) and $\boldsymbol{g}(\boldsymbol{a})$ is the gradient of the function evaluated at the point $\boldsymbol{a}$. This formula requires, of course, that $\boldsymbol{H}$ be a positive definite matrix. In fact $f(\boldsymbol{x})$ may possess a minimum at a point even if $\boldsymbol{H}$ is only positive semi-definite and if this is the case, then special formulae must be developed for $\boldsymbol{p}$. It was shown in [1] that the Hessian at the minimum in the LCAO-MO-SCF closed shell and unrestricted problem was only positive semi-definite and some appropriate formulae were developed there for this case. It was further demonstrated in that paper that in such a case the quadratic convergence guarantees for conjugate gradient and conjugate direction methods can no longer be given, so that such methods may well prove to be unsatisfactory in practice.

2. Direct Minimisation of the Energy Functional

In the LCAO-MO approach the molecular orbitals of a problem $\boldsymbol{\phi}$ are given in terms of an m orbital basis $\boldsymbol{\eta}$, by the relation

$$\boldsymbol{\phi}=\boldsymbol{\eta}(\boldsymbol{T}\,\vdots\,\boldsymbol{T}_u) \tag{2.1}$$

where the matrix of coefficients $\boldsymbol{T}$ describes the n occupied and $\boldsymbol{T}_u$ the m-n unoccupied orbitals. We shall consider $\boldsymbol{\eta}$ to be a 1 by m matrix and $\boldsymbol{T}$ to be an m by n matrix with $n<m$. The basis $\boldsymbol{\eta}$ is of course arbitrary but usually one has in mind some fundamental basis, the fixed AO basis, in which the integrals of a problem are actually evaluated. We shall use the subscripts i, j etc. to denote the possibility of a range $(1, m)$ and r, s, etc. to denote the range $(1, n)$. The energy, E, in this approach can be regarded as a function of $\boldsymbol{T}$, but it is not an unconstrained function of $\boldsymbol{T}$ since we have the accessory orthonormality requirement

$$\boldsymbol{T}^T\boldsymbol{S}\boldsymbol{T}=\boldsymbol{1}_n \tag{2.2}$$

where $\boldsymbol{S}$ is the overlap matrix in the basis $\boldsymbol{\eta}$ and where we assume real orbitals and coefficients. As Fletcher [3] was the first to show these constraints may be removed by writing

$$\boldsymbol{T}=\boldsymbol{Y}\boldsymbol{U} \tag{2.3}$$

with $\boldsymbol{U}$ chosen so that

$$\boldsymbol{U}\boldsymbol{U}^T=(\boldsymbol{Y}^T\boldsymbol{S}\boldsymbol{Y})^{-1}. \tag{2.4}$$

The energy E may then be treated as a function of the unconstrained variables $\boldsymbol{Y}$. This is possible as long as $(\boldsymbol{Y}^T\boldsymbol{S}\boldsymbol{Y})$ is non-singular.

Using the chain rule it is a straightforward matter to establish that the first and second derivatives of E with respect to the elements of $\boldsymbol{Y}$ are given by

$$\partial E/\partial Y_{ir}=(\boldsymbol{W}\boldsymbol{U}^T)_{ir}+\sum_{juv} Y_{ju}V^{(1)}_{uv,ir}W_{jv} \tag{2.5}$$

and

$$\begin{aligned}\partial^2 E/\partial Y_{js}\partial Y_{ir} = &\sum_u (W_{iu}V^{(1)}_{ru,js} + W_{ju}V^{(1)}_{su,ir})\\ &+ \sum_{uv} U_{su}K_{ju,iv}U_{rv}\\ &+ \sum_{kuv} Y_{ku}V^{(2)}_{js,uv,ir}W_{kv}\\ &+ \sum_{ktuv} (U_{rt}Y_{kv}V^{(1)}_{vu,js}K_{ku,it} + U_{su}Y_{kv}V^{(1)}_{vt,ir}K_{ju,kt})\\ &+ \sum_{kltuvw} Y_{kw}Y_{lv}V^{(1)}_{wt,js}V^{(1)}_{vu,ir}K_{kt,lu} \end{aligned} \tag{2.6}$$

where

$$W_{ir} = \partial E/\partial T_{ir}\,; \quad K_{js,ir} = \partial^2 E/\partial T_{js}\partial T_{ir}\,,$$

$$V^{(1)}_{uv,ir} = \partial U_{uv}/\partial Y_{ir}\,; \quad V^{(2)}_{js,uv,ir} = \partial^2 U_{uv}/\partial Y_{js}\partial Y_{ir}\,.$$

We may imagine that we have chosen our basis $\boldsymbol{\eta}$ in such a way that

$$\boldsymbol{S} = \boldsymbol{I}_m, \quad \boldsymbol{T} = \boldsymbol{Y} = \begin{pmatrix} \boldsymbol{I}_n \\ \cdots \\ \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{U} = \boldsymbol{I}_n \tag{2.7}$$

that is to say we choose our basis to be orthonormal and also to be such as to have as its first n members the occupied molecular orbitals. This is a basis of the kind that Hillier and Saunders [4] call a *trial molecular orbital* (TMO) basis. In the TMO basis the above formulae simplify considerably and yield

$$\partial E/\partial Y_{ir} = W_{ir} + \sum_{uv} W_{uv}V^{(1)}_{uv,ir}\,, \tag{2.8}$$

$$\begin{aligned}\partial^2 E/\partial Y_{js}\partial Y_{ir} = &K_{js,ir} + \sum_u (W_{iu}V^{(1)}_{ru,js} + W_{ju}V^{(1)}_{su,ir})\\ &+ \sum_{uv} (W_{uv}V^{(2)}_{js,uv,ir} + V^{(1)}_{uv,js}K_{uv,ir} + V^{(1)}_{uv,ir}K_{js,uv})\\ &+ \sum_{twuv} V^{(1)}_{vw,ir}K_{ut,vw}V^{(1)}_{ut,js}\,. \end{aligned} \tag{2.9}$$

The problem is now to determine the elements of the matrices $V^{(1)}$ and $V^{(2)}$. By straightforward but rather tedious manipulation it may be shown that condition (2.4) yields in the TMO basis, the following results

$$V^{(1)}_{vu,ir} + V^{(1)}_{uv,ir} = -(\delta_{iv}\delta_{ru} + \delta_{iu}\delta_{rv}) \tag{2.10}$$

and

$$\begin{aligned}V^{(2)}_{js,vu,ir} + V^{(2)}_{js,uv,ir} = &-\delta_{ij}(\delta_{ru}\delta_{sv} + \delta_{rv}\delta_{su})\\ &-(V^{(1)}_{su,ir}\delta_{jv} + V^{(1)}_{sv,ir}\delta_{ju} + V^{(1)}_{rv,js}\delta_{iu} + V^{(1)}_{ru,js}\delta_{iv})\\ &-\sum_{t'} (V^{(1)}_{t'v,js}\delta_{ru}\delta_{it'} + V^{(1)}_{t'u,js}\delta_{rv}\delta_{it'}\\ &+ V^{(1)}_{t'u,ir}\delta_{jt'}\delta_{sv} + V^{(1)}_{t'v,ir}\delta_{jt'}\delta_{su}\\ &+ V^{(1)}_{t'u,ir}V^{(1)}_{t'v,js} + V^{(1)}_{t'v,ir}V^{(1)}_{t'u,js})\,. \end{aligned} \tag{2.11}$$

Clearly it is not possible to make these results completely explicit without specifying a form for $\boldsymbol{U}$ and determining its derivatives. To return to an arbitary basis

however, it is easy to see that we can always write

$$V^{(1)}_{vu,ir} = -\omega_{vu}(V^{(1)}_{vu,ir} + V^{(1)}_{uv,ir}), \tag{2.12}$$

$$V^{(2)}_{js,vu,ir} = -\omega_{vu}(V^{(2)}_{js,vu,ir} + V^{(2)}_{js,uv,ir}) \tag{2.13}$$

for suitably chosen ω_{vu}.

Thus if $\boldsymbol{U}$ is chosen $(\boldsymbol{Y}^T\boldsymbol{S}\boldsymbol{Y})^{-\frac{1}{2}}$ corresponding to symmetric orthogonalisation the consequent requirement that $d\boldsymbol{U}$ shall also be symmetric implies that $\omega_{vu} = \frac{1}{2}$ for all vu. The same result can be shown to hold for the common variant of this method in which $\boldsymbol{U}$ is not necessarily symmetric. If $\boldsymbol{U}$ is chosen to be an upper triangle, corresponding to Schmidt orthogonalisation it follows then that $\omega_{vu} = 1$, $v < u, \omega_{uu} = \frac{1}{2}$; $\omega_{vu} = 0$, $v > u$. Whatever the method adopted, however the ω_{vu} have the following properties

$$\omega_{vu} + \omega_{uv} = 1, \quad \omega_{uu} = \tfrac{1}{2}. \tag{2.14}$$

These can be considered to hold even if U_{uv} is identically zero (as it might be through symmetry) so that $V^{(1)}_{uv,ir} = 0$ and so on.

Using the forms for $V^{(1)}$ and $V^{(2)}$ (2.8) and (2.9) may be re-written after extensive manipulation

$$\partial E/\partial Y_{ir} = W_{ir}(1-\Delta_i) + \Delta_i\omega_{ri}(W_{ir} - W_{ri}), \tag{2.15}$$

$$\begin{aligned}
\partial^2 E/\partial Y_{js}\partial Y_{ir} = {} & (1-\omega_{ir}\Delta_i)(1-\omega_{js}\Delta_j)K_{js,ir} \\
& + \Delta_i\Delta_j\omega_{ri}\omega_{sj}K_{sj,ri} - \omega_{sj}\Delta_j(1-\omega_{ir}\Delta_i)K_{sj,ir} \\
& - \omega_{ri}\Delta_i(1-\omega_{js}\Delta_j)K_{js,ri} \\
& + \delta_{ij}(\Delta_i\Delta_j(1-\omega_{si}\omega_{ri}) - 1)(W_{sr}\omega_{sr} + W_{rs}\omega_{rs}) \\
& + \delta_{rs}(\Delta_i\Delta_j(1-\omega_{ir}\omega_{jr})(W_{ji}\omega_{ji} + W_{ij}\omega_{ij}) \\
& - (W_{ij}\omega_{rj}\Delta_j + W_{ji}\omega_{ri}\Delta_i)) \\
& + \delta_{is}(\Delta_j(1-\omega_{js}\omega_{rs})(W_{jr}\omega_{jr} + W_{rj}\omega_{rj}) - W_{jr}\omega_{sr}) \\
& + \delta_{jr}(\Delta_i(1-\omega_{ir}\omega_{sr})(W_{is}\omega_{is} + W_{si}\omega_{si}) - W_{is}\omega_{rs})
\end{aligned} \tag{2.16}$$

where Δ_i is 1 if $i \leqq n$ and 0 if $i > n$.

Since at a solution point $\partial E/\partial Y_{ir}$ vanishes we can immediately infer from (2.15) that at a solution point

$$\begin{aligned}
W_{ir} &= W_{ri} \quad i \leqq n \\
W_{ir} &= 0 \quad i > n.
\end{aligned} \tag{2.17}$$

If we now examine the rrth column of the Hessian matrix we find after rather extensive manipulation that

$$\begin{aligned}
\partial^2 E/\partial Y_{js}\partial Y_{rr} &= (\omega_{sj}\omega_{js}\delta_{rj} + \omega^2_{sj}\delta_{rs})(W_{sj} - W_{js}), \quad j \leqq n \\
&= -\delta_{rs}W_{jr}, \quad j > n
\end{aligned} \tag{2.18}$$

with a precisely similar result for the ssth row. From (2.18) and (2.17) it follows that the Hessian at the minimum contains n zero rows and columns, each zero row and column intersecting at (rr, rr)th element. Thus we have established quite

generally that any energy functional in which the orthogonality constraints are applied according to (2.3) and (2.4) possesses only a semi-definite Hessian with respect to the unconstrained variables. It is clear that the number of zero roots of the Hessian is at least n, and if we make the assumption that a minimum exists then by hypothesis the Hessian must be *positive* semi-definite.

3. The Choice of a Descent Direction for a Positive Semi-Definite Hessian

If we expand the energy about the minimum (denoted by $E(\mathbf{0})$) then we obtain, to second order.

$$E(\delta \boldsymbol{Y}) = E(\mathbf{0}) + \tfrac{1}{2} \sum_{js,ir} (E^{(2)}(\mathbf{0}))_{js,ir} \delta Y_{js} \delta Y_{ir} \tag{3.1}$$

where $E^{(2)}(\mathbf{0})$ denotes the Hessian at the minimum. We can also invent the function $\bar{E}(\delta \boldsymbol{Y})$ by the definition

$$\bar{E}(\delta \boldsymbol{Y}) = E(\mathbf{0}) + \tfrac{1}{2} \sum_{js,ir} (\lambda_{ir} \delta_{js,ir} + E^{(2)}(\mathbf{0})_{js,ir}) \delta Y_{ir} \delta Y_{js} \tag{3.2}$$

where the λ_{ir} are arbitrary positive constants, and where obviously

$$\bar{E}(\mathbf{0}) = E(\mathbf{0}) . \tag{3.3}$$

Thus the two functions have the same minimum and to minimise one is to minimise the other. Now we are at liberty to choose the λ_{ir} in such a way as to make the Hessian of (3.2) positive definite, and hence a direction of descent is now certainly given by

$$\boldsymbol{p} = -(\boldsymbol{\lambda} + \boldsymbol{E}^{(2)}(\mathbf{0}))^{-1} \boldsymbol{g} . \tag{3.4}$$

Clearly this is a simple extension of the Goldfeld-Quandt-Trotter [5] method, and reduces to this method when $\boldsymbol{\lambda}$ is chosen as a multiple of the unit matrix. If we choose $\boldsymbol{\lambda}$ in such a way as to make the now non-singular Hessian heavily diagonal dominant so that we can expand the inverse about the diagonal and write to any desired accuracy

$$\boldsymbol{p} = -\boldsymbol{e}^{-1} \boldsymbol{g} \tag{3.5}$$

where e is a diagonal matrix with elements

$$e_{ir} = \lambda_{ir} + (E^{(2)}(0))_{ir,ir} . \tag{3.6}$$

It follows at once therefore that $\boldsymbol{p}$ can be written as a rectangular matrix $\boldsymbol{P}$ with elements

$$P_{ir} = -E^{(1)}(\delta \boldsymbol{Y})_{ir} / e_{ir} \tag{3.7}$$

where $E^{(1)}(\delta \boldsymbol{Y})$ is the matrix of first derivatives, given generally by (2.5), evaluated at the point $\mathbf{0} + \delta \boldsymbol{Y}$.

It should be noticed that it is not at once clear that we may use the simplified formulae (2.15) and (2.16) in (3.7). If we choose the exact solution basis together with some appropriate orthogonal extension as our TMO basis then we can certainly use (2.16) for $\boldsymbol{E}^{(2)}(\mathbf{0})$. However, since, by hypothesis the current point is not a point at which, in this basis, $\boldsymbol{S}$, $\boldsymbol{T}$, $\boldsymbol{Y}$, and $\boldsymbol{U}$ can be expressed in the form

(2.7), then (2.15) cannot consistently be used for $E^{(1)}(\delta \boldsymbol{Y})$. *Mutatis mutandi* if we use the current basis as the TMO basis, we may use (2.15) but not (2.16).

In practice of course we shall be restricted to using the current basis as the TMO basis, and we shall evaluate both $\boldsymbol{E}^{(1)}$ and $\boldsymbol{E}^{(2)}$ at the current point, relying on the constancy of the Hessian in the quadratic region, to justify this procedure eventually. On this basis then, using both (2.15) and (2.16), the matrix $\boldsymbol{P}$ has elements

$$\begin{aligned} P_{ir} &= W_{ir}/(W_{rr} - K_{ir,ir} - \lambda_{ir}) \qquad i > n \\ &= 0 \qquad\qquad\qquad\qquad\quad i = r \\ &= \omega_{ri}(W_{ir} - W_{ri})/(\omega_{ri}^2(W_{rr} + W_{ri} + 2K_{ir,ri} - K_{ir,ir} - K_{ri,ri}) - \lambda_{ir}) \qquad i \leqq n\,. \end{aligned} \tag{3.8}$$

The up-dated $\boldsymbol{Y}$ matrix is then

$$\boldsymbol{Y} \rightarrow \hat{\boldsymbol{Y}} = \boldsymbol{Y} + \alpha \boldsymbol{P} \tag{3.9}$$

where α is chosen to minimise $E(\boldsymbol{Y})$.

The overlap matrix specified by $\hat{\boldsymbol{Y}}$ is

$$\hat{\boldsymbol{Y}}^T \hat{\boldsymbol{Y}} = \boldsymbol{I}_n + \alpha(\boldsymbol{P}_t + \boldsymbol{P}_t^T) + \alpha^2 \boldsymbol{P}^T \boldsymbol{P} \tag{3.10}$$

and the overlap matrix between the new vectors and the currently unoccupied vectors is $\alpha \boldsymbol{P}_b^T$. Here $\boldsymbol{P}_t$ and $\boldsymbol{P}_b$ represent the first n and the last m-n rows of $\boldsymbol{P}$ respectively. Thus this new $\boldsymbol{Y}$ does not at once specify the occupied portion of a new TMO basis. A new matrix $\boldsymbol{U}$ must therefore be chosen according to (2.3) and (2.4), with the proviso that the orthogonalising process used in constructing $\boldsymbol{U}$ is compatible with the choice of ω_{ri} in (3.8). Furthermore Eqs. (2.5) and (2.6) must be used in any future up-dating using (3.8). This makes the method extremely cumbersome to use in practice.

An alternative approach would be to construct a new matrix $\boldsymbol{T}$ according to (2.3) and (2.4), to orthogonalise the presently unoccupied basis to this $\boldsymbol{T}$ and then to treat the resulting complete matrix as specifying a new TMO basis. In this new basis of course it is possible to use (2.15) and (2.16) in evaluating (3.8), but it has the disadvantage that an effectively new problem is posed at each descent and thus it is impossible to give any kind of finite convergence guarantees. However one might consider this an appropriate strategy when $(\boldsymbol{P}_t + \boldsymbol{P}_t^T)$ vanished, in which case the constraints would be up-dated only to second order, and this one might expect the effect on Y of orthonormalisation to be rather small.

In fact it has been shown in [1] that in the case of a simple closed shell or unrestricted LCAO-MO-SCF procedure $\boldsymbol{P}_t$ itself vanishes, so that this approach should be effective there[1]. In fact in this case our equations (3.8) become simple generalisations of the Hillier and Saunders *level shifting method* [4] with λ_{ir} equivalent to their level shifting parameter and α equivalent to to their damping factor (for an open shell generalisation of this see Guest and Saunders [6]). Level shifting methods are known to be extremely effective in these cases.

[1] P_t vanishes in this case because one is free to choose $W_{ir} = W_{ri}$. There is a misprint in Eq. (5.8) of [1] the matrix printed $\boldsymbol{Q}$ should be printed $\boldsymbol{0}$, the n by n null matrix.

In general, however, we can see that if $\omega_{ri}=\omega_{ir}$ and we choose $\lambda_{ir}=\lambda_{ri}$, then $\boldsymbol{P}_t$ becomes a skew-symmetric matrix, and the first order terms vanish from (3.10) automatically. Thus we are free to make the first order terms vanish if we adopt a symmetric orthogonalisation procedure in constructing $\boldsymbol{U}$, but not if we adopt a Schmidt procedure. In the Schmidt procedure $\boldsymbol{P}_t$ is a lower triangle with a zero diagonal; so that to first order $\boldsymbol{Y}$ is a lower triangle with a unit diagonal.

Apart from these general considerations and the requirement that the inverse on the right of (3.4) exists there appears to be nothing that can easily be said about the choice of $\boldsymbol{\lambda}$. If we choose $\boldsymbol{\lambda}$ to be a large constant multiple of the unit matrix then the Eqs. (3.8) become effectively the steepest descent equations. Since an ordinary steepest descent procedure is known in general, to have a very poor rate of final convergence, we would avoid such a choice of the λ_{ir}. On the whole it would anyway be good tactics to choose the λ_{ir} as small as possible, at least in the quadratic region of the energy. To see this we notice that it is easy to show that if we choose α by minimising $\bar{E}$ (which is a quadratic function) along $\boldsymbol{P}$ chosen by (3.4), then α will always be unity. This is a general result for quadratic functions (see [1]) in so far as Eqs. (3.8) are good approximations to (3.4) this result will hold if we choose our direction according to (3.8). In practice we shall find α by minimising E, and the same result will hold under the same conditions, in the quadratic region, only if the λ_{ir} are zero. Thus if we can keep the λ_{ir} small while (3.8) remains a good approximation to the inverse Hessian, we can be confident that α will be approximately unity and thus we can avoid the need for a time-consuming linear search to minimise E along $\boldsymbol{P}$.

In fact if there are n and only n zero roots it should be sufficient to ensure the positive definiteness of the right-hand side of (3.4), to choose $\lambda_{ir}=\delta_{ir}\lambda$ where λ is a positive constant. If this is possible and (3.5) remains a good approximation then α will be as close as is possible to 1. In practice however it may be preferable to choose $\lambda_{ir}=\lambda$ for $i>n$ and λ' for $i<n$, choosing λ' large at the beginning of the solution process to minimise occupied orbital mixing.

4. The Realisation of the Method in the LCAO-MC-SCF Case

If we write the energy (in atomic units) for the conventional fixed nucleus problem in the LCAO-MC-SCF case as (see [7], p. 142, see also [8]).

$$E=\sum_{rs}\sum_{ij} T^*_{ir}h_{ij}T_{js}P^{(1)}_{sr}+\tfrac{1}{2}\sum_{rstu}\sum_{ukl} T^*_{ir}T^*_{js}g_{ijkl}T_{kt}T_{lu}P^{(2)}_{tu,rs} \tag{4.1}$$

where

$$h_{ij}=\langle\eta|h|\eta_j\rangle,\quad g_{ijkl}=\langle\eta_i\eta_j|1/r_{12}|\eta_k\eta_e\rangle \tag{4.2}$$

with $\boldsymbol{P}^{(1)}$ and $\boldsymbol{P}^{(2)}$ denoting the one and two particle density matrices in the MO basis.

It is then a straightforward matter to show (assuming real orbitals and coefficients) that

$$\boldsymbol{W}=2(\boldsymbol{hTP}^{(1)}+\boldsymbol{Z}) \tag{4.3}$$

with

$$Z_{ir}=\sum_{stukjl} T^*_{js} g_{ijkl} T_{kt} T_{lu} P^{(2)}_{tu,rs} \tag{4.4}$$

and that

$$K_{js,ir}=2(h_{ij}P^{(1)}_{sr}+\sum_{tu}(g_{ijtu}(P^{(2)}_{tu,rs}+P^{(2)}_{ts,ru})+g_{itju}P^{(2)}_{su,rt}) \tag{4.5}$$

where g_{ijtu} denotes the integral

$$\langle \eta_i \eta_j | 1/r_{12} | \phi_t \phi_u \rangle . \tag{4.6}$$

It is again an easy matter to show that at a minimum, subject to constraints,

$$(\boldsymbol{hTP}^{(1)}+\boldsymbol{Z})=\boldsymbol{ST\varepsilon} \tag{4.7}$$

where ε is an n by n symmetric matrix of Lagrange Multipliers.

If we consider for a moment the simple closed shell problem then it is the case that

$$\boldsymbol{Z}=2\boldsymbol{GT} \tag{4.8}$$

where $\boldsymbol{G}$ is the usual electron interaction matrix $2\boldsymbol{J}-\boldsymbol{K}$, and that therefore

$$\boldsymbol{W}=4\boldsymbol{fT} \tag{4.9}$$

where $\boldsymbol{f}$ is the usual Fock matrix $\boldsymbol{h}+\boldsymbol{G}$.

In this case we know that in any real basis, $\boldsymbol{f}$ is symmetric and we further know that we can always find a basis (the basis of canonical MO's) in which, at any stage, $\boldsymbol{f}$ can be expressed as

$$\boldsymbol{f}=\left(\begin{array}{c|c} \varepsilon_0 & \bar{f}^T \\ \hline \bar{f} & \varepsilon_u \end{array}\right) \tag{4.10}$$

where ε_0 and ε_u are diagonal matrices of the occupied and unoccupied orbital energies respectively. This basis is also of course a TMO basis as defined in (2.7) so that (4.9) simplifies to

$$\boldsymbol{W}=\begin{pmatrix} \varepsilon_0 \\ \hline \bar{f} \end{pmatrix}. \tag{4.11}$$

It follows from (4.8) and (4.5) that in any orthogonal basis we may write

$$K_{js,ir}=4f_{ij}\delta_{rs}+16g_{ijrs}-4g_{ijsr}-4g_{isjr} . \tag{4.12}$$

In the TMO basis in this case it follows from (2.15) and (2.16) that

$$\begin{aligned} E^{(1)}_{ir} &= 0 \qquad i \leqq n \\ &= 4\bar{f}_{ir} \qquad i > n \end{aligned} \tag{4.13}$$

$$\begin{aligned} E^{(2)}_{js,ir} &= 0 \quad i \text{ or } j \leqq n \\ &= 4\delta_{ij}\delta_{rs}(\varepsilon_i-\varepsilon_r)+16g_{ijrs}-4g_{ijsr}-4g_{isjr} \quad i \text{ and } j>n . \end{aligned} \tag{4.14}$$

If one had a sufficiently good TMO basis so that the occupied and unoccupied orbitals were stable then one would usually expect ε_j to be positive, as it is drawn

from ε_u and ε_r to be negative since it is drawn from ε_0. In consequence one would expect the first term of the non-zero part of (4.14) to be positive and much larger than the subsequent terms. The Hessian would in this case be strongly diagonal dominant. In any case $\varepsilon_r < \varepsilon_i$, by hypothesis, so that the leading term must at least be positive. For the two electron part the diagonal term is positive unless $\langle ir|1/r_{12}|ir\rangle > 4\langle ii|1/r_{12}|rr\rangle$, and one might well anticipate that the two electron terms on the diagonal would in themselves, be greater than their off-diagonal counterparts. Thus it is perhaps not too much to say that the choice of a canonical basis forces the non-zero part of the Hessian to be naturally diagonal-dominant with the consequent possibility that the λ_{ir} here may be chosen as small as is convenient. Furthermore this choice keeps, to first order, the orthonormality properties of the updated matrix. Last, but by no means least, we see that the factorization (4.8) of $\boldsymbol{Z}$, which makes an eigen-value problem possible, also implies that at least a portion (and hopefully the most significant portion) of the two-electron part of the second derivative matrix can be written in terms of the two-electron parts of the gradient matrix. This means that we need little extra work to construct the denominators in (3.8). We can therefore think of the properties of the canonical MO's in this case as the paradigm for any TMO's to be used in a level-shifting method for solving the general MC-SCF problem.

It is clear that one can always go some way to achieving this paradigm situation in the general MC-SCF case when, through some particular choice of the forms of the matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$, one can express (4.7) in the form of an eigen-value equation or as a set of coupled eigen-value equations, since in this way a set of orbitals with quasi-canonical properties can be achieved. Thus one can derive formulas from the formulas (4.3), (4.5), and (3.8) given here, which are slight generalisations of Guest and Saunders' [6] level-shifting method for the usual open-shell SCF problem, and of Wood and Veillard's [9] formulas for level-shifting in the Clementi-Veillard MC-SCF problem. An analysis similar to that given for the chosen shell problem above makes it plausible that level-shifting methods should work very well in both these kinds of problem.

In the general case however nothing so helpful can be said. If one wishes to use a level-shifting method, then clearly one should start from a TMO basis in which $(W_{ir} - W_{ri})$ is close to zero for $i \leqq n$ and in any updating, symmetrically orthogonalise the occupied orbitals. This prescription is of course equivalent to saying that a TMO basis should be chosen in which the occupied TMO's do not mix very strongly, and it is well known that this is difficult to achieve in practice, as an initial guess. Furthermore there is, in this case, no simple relationship between the matrices $\boldsymbol{K}$ and $\boldsymbol{W}$ so that to apply (3.8) a significant proportion of the matrix $\boldsymbol{K}$ must be constructed at each iteration.

It thus seems that in the most general case, even a level shifting method may well be only slowly convergent. Furthermore, if we balance the time taken to compute the elements of $\boldsymbol{K}$ against the number of extra iterations that would have to be performed in a steepest-descents procedure, it is not at all clear that a steepest descents procedure in the TMO basis might not be just as economical, particularly in the early stages of the minimising procedure. Of course one could not be sure of this without detailed and extensive numerical trials.

5. Summary

As was stated in the introduction to this paper, if in a minimisation problem, the Hessian is not positive definite, then quadratic convergence guarantees cannot be given for direct minimisation methods. In particular they cannot be given for the conjugate gradient method. This means that the convergence of direct methods in this context must be regarded as uncertain. The convergence is not bound to be bad, but it may well be, since it cannot be proved to be good.

We have shown here that in general the Hessian for the MC-SCF problem is singular and that in consequence we may expect trouble if a direct minimisation scheme is used to minimise the energy. We have shown that a variant of the Goldfeld-Quandt-Trotter scheme is a possible way out of this difficulty, but it does not seem likely that it will be a panacea. Numerical investigations on this topic are currently being undertaken.

Acknowledgement. The author would like to thank Professor J. C. Ballhausen and Professor J. P. Dahl very much indeed for their hospitality at the H. C. Ørsted Institute and the Technical University of Denmark, where this work was begun.

References

1. Sutcliffe, B. T.: Theoret. Chim. Acta (Berl.) **33**, 201 (1974)
2. Garton, D., Sutcliffe, B. T.: Chem. Soc. Specialist Periodical Repts. 1, 34 (1974)
3. Fletcher, R.: Mol. Phys. **19**, 55 (1970)
4. Hillier, I., Saunders, V.: Proc. Roy. Soc. A **320**, 161 (1970/71)
5. Goldfeld, S. M., Quandt, R. E., Trotter, H. F.: Econometrica **34**, 541 (1966)
6. Guest, M., Saunders, V.: Mol. Phys. **28**, 819 (1974)
7. McWeeny, R., Sutcliffe, B. T.: Methods of molecular quantum mechanics. New York: Academic Press 1969
8. Kari, R., Sutcliffe, B. T.: Intern. J. Quantum Chem. **35**, 625 (1971)
9. Veillard, A., Wood, M.: Mol. Phys. **26**, 595 (1973)

Dr. B. T. Sutcliffe
Department of Chemistry
University of York
Heslington, York, YO1 5DD, England