Pramāṇa – J. Phys., Vol. 25, No. 3, September 1985, pp. 281–290. © Printed in India.

On the direct determination of constrained pure state one-electron density matrices: Part I. A new theoretical model for closed-shell systems

K K DAS, PRIYOTOSH KHAN† and S P BHATTACHARYYA*
Department of Physical Chemistry, Indian Association for the Cultivation of Science, Jadavpur, Calcutta 700 032, India
†Department of Mathematics, Behala College, Parnasree, Calcutta 700 062, India

MS received 1 February 1985; revised 8 July 1985

Abstract. A new method based on the penalty-function way of satisfying equality constraints is proposed for the determination of constrained pure state one-electron density matrices for closed-shell many-electron systems. The algorithm suggested can handle many constraints simultaneously. Certain interesting features of the proposed algorithm are discussed with numerical examples.

Keywords. Constrained variational calculation; direct determination electron density; penalty function method; pure state density; closed shell systems.

PACS No. 31·15

1. Introduction

At the lowest level of approximation, the ground-state electronic structure of many-electron atoms and molecules is described by a single-determinant wave function constructed from a set of orthonormal one-electron orbitals, which have been determined by invoking the variational principle (Roothaan 1951; Hartree 1928; Fock 1930; Slater 1930). These orbitals determine the one-electron density in terms of which the atomic and molecular properties can be calculated. The importance and usefulness of one-electron density in the context of quantum chemical concepts and calculations can hardly be overestimated (Bamzai and Deb 1981, and references cited therein). Instead of determining the wavefunction first and then the density one may adopt a slightly different mode of approach. It is possible to treat the elements of the one-electron density matrix (to be determined) as the basic variables and try to determine the density itself directly through a variational procedure. McWeeny's steepest descent technique (McWeeny 1956, 1957) belongs to the second category of methods. One may note here that the density or the one-electron orbitals can be determined by applying the variational principle to an energy functional that incorporates the orthonormality constraints in one way or the other. Since the wavefunction so obtained is correct only to first order, the magnitude of error in the calculated values of properties (observables) other than energy (i.e. expectation values of operators which do not commute with $\hat{H}$) may well be quite high. One may try to improve the situation by resorting to the application of the variational principle to still more modified energy functionals which

---

* To whom correspondence should be addressed.

incorporate in them certain additional constraints over and above those implied by the required orthonormality of the one-electron orbitals. In other words, the one-electron orbitals must not only be orthonormal and minimize energy but also minimize the mean-square deviations of a chosen set of calculated one-electron properties from their exact (theoretical or experimental) values. Mukherjee and Karplus (1963) were the first to attempt a calculation of this kind. Further extension was made by Rasiel and Whitman (1965). Later Byers-Brown (1966) and Chong (1973) analysed the constrained variational problem thoroughly which revealed the possibility of useful applications of the method. However, the problem has never been solved in a generalized fashion. Mukherjee and Karplus (1963), for example, enforced the constraints in a brute-force manner and for most of the systems of chemical interest, the technique adopted is inappropriate. Recently, interest in constrained variational problems seems to have been revived (Westhaus 1983). While Westhaus proposed a generalization of the density functional formalism using constrained variations of the expectation value $\tilde{\varepsilon} = \langle \Psi | \hat{H} | \Psi \rangle$, Zeiss and Whitehead 1983 tried to shape the one-electron density to reproduce certain one-electron properties. The present paper deals with a general and practicable solution of the multiply constrained variational problem by adopting what is currently known as the method of penalty functions for enforcing relevant equality constraints (Fiacco 1970; Fiacco and McCormik 1968; Mukherjee 1975). The problem is formulated in two steps: first, the case of energy minimization with a single constraint imposed on the orbitals (or equivalently on the one-electron density) is formulated and then the more general case of minimization with any number of constraints is dealt with. Only the theoretical model for closed-shell states is presented in what follows. The general formulation for restricted and unrestricted open-shell states will be presented separately (Das and Bhattacharyya 1985).

2. Penalty function method

Since the method of penalty functions has hardly been used for tackling constrained variational problems frequently encountered in quantum chemistry (see however Mukherjee 1975) it would perhaps be useful to present a brief introduction to this method before we take up the actual problem at our disposal. Although rather unknown in the domain of quantum chemistry, the method of penalty functions is one of the simplest and most well-known methods of solving the problems of mathematical programming. The basic idea of the method consists in approximately reducing a constrained minimization problem to an unconstrained one by replacing the original function (or functional) to be minimized by an auxiliary one. The choice of auxiliary function is made in such a way that it *coincides with the original function in the admissible domain* of the variables but increases steeply outside it. We may now formulate the basic scheme in a little more concrete terms.

Let us suppose that $F(x)$ is the function to be minimized subject to the constraints that

$$f_i(x) \leqslant 0 \qquad i = 1, 2 \ldots n;$$

we also introduce the following definitions:

$$g_0(t) = \begin{cases} t^2, & t \geqslant 0, \\ 0, & t < 0. \end{cases}$$

We may now define a function $h(x, \beta)$ (the penalty function) as follows:

$$h(x, \beta) = \beta \sum_{i=1}^{n} g_0[f_i(x)].$$

It is clear that

$$h(x, \beta) = 0 \quad \text{if} \quad x \in D,$$

where $D \equiv \{x: f_i(x) \leqslant 0,\ i = 1, 2, \dots n\}$.
If $x \notin D$, $h(x, \beta) > 0$ and also $h(x, \beta) \to +\infty$ as $\beta \to \infty$. With these definitions the subsidiary (auxiliary) minimization problem then reduces to the unconstrained minimization of the function $\tilde{F}(x)$ where,

$$\tilde{F}(x, \beta) = F(x) + h(x, \beta).$$

It is obvious that the construction of $h(x, \beta)$ as suggested here is not unique. Many alternative definitions are possible for $h(x, \beta)$ or the auxiliary function $\tilde{F}(x, \beta)$. It is sufficient for $h(x, \beta)$ or $\tilde{F}(x, \beta)$ to have certain specific features for achieving convergence [Laenberger 1971]. For our purpose, however, the Morrison function type of choice for $\tilde{F}(x, \beta)$ appears to be the best (Morrison 1968). This involves a slightly different construction of $\tilde{F}(x)$ in that one now assumes

$$\tilde{F}(x, \beta, F_L) = [F(x) - F_L(x)]^2 + \beta[f(x) - b]^2,$$

where $F_L(x)$ stands for the lower bound to the constrained value of $F(x)$ and the constraint condition on x is $f(x) = b$. Instead of dealing with functions, however, we shall be dealing with functionals. Nevertheless the basic structure remains the same.

3. Theory

3.1 *Single constraint*

Let P be the trial one-electron spatial density matrix ($n \times n$) for the ground state of our $2m$ electron system in a discrete orthonormal basis. The constraints on P are

$$P^2 = P, \tag{1a}$$

and

$$\operatorname{Tr} P = n. \tag{1b}$$

Then the unconstrained energy functional E is given by

$$E = 2\operatorname{Tr} hP + \operatorname{Tr} PG(P), \tag{2}$$

where

$$h_{pq} = \langle \phi_p | h | \phi_q \rangle,$$

and

$$G_{pq} = 2J_{pq}(P) - K_{pq}(P),$$

$$J_{pq} = \sum_{s,r} P_{rs} \langle \phi_p \phi_s | g | \phi_q \phi_r \rangle, \quad K_{pq} = \sum_{s,r} P_{rs} \langle \phi_p \phi_s | g | \phi_r \phi_q \rangle,$$

where g stands for r_{12}^{-1}. The functional containing the constraint (1) on P incorporated in it can be easily written as

$$E_c = E + \beta \operatorname{Tr} (P^2 - P)^2, \tag{3}$$

where β is a penalty weighting factor having the dimension of energy. Let us assume that E_0^{HF} is the exact Hartree-Fock energy of our system. We may now define a modified constrained energy functional as follows:

$$E_c = (E - E_0^{\mathrm{HF}})^2 + \beta_1 \operatorname{Tr}(P^2 - P)^2, \tag{4}$$

β_1 being the penalty weight factor of appropriate dimension. If we now treat the elements of P as our basic variational parameters, then by setting $\delta E_c = 0$ we can arrive at the equation for the iterative determination of P. Note that, for minimum of E_c the two necessary conditions are

$$\begin{aligned} &E = E_0^{\mathrm{HF}}, \\ \text{and} \qquad &P^2 = P, \end{aligned} \tag{5}$$

so that at the minimization point, the trial density matrix must satisfy the idempotency constraint (1) as also must the trial energy E merge with the exact Hartree-Fock (HF) energy as closely as possible. We have presumed so far that E_0^{HF} is known. The more realistic case is when E_0^{HF} is not known to start with. In that case, a workable algorithm for the variational determination of P can be based on a slightly different functional E_C^L where

$$E_C^L = (E - E_L)^2 + \beta_1 \operatorname{Tr}(P^2 - P)^2, \tag{6}$$

E_L is a lower bound to E_0^{HF}, the actual value of which can be updated at any stage of the iterative process by some well-defined procedure (to be defined later). The variational condition on E_C^L requires that

$$\delta E_C^L = 2(E - E_L)\delta E + \beta_1 \operatorname{Tr}(4P^3 - 6P^2 + 2P)\delta P = 0 \tag{7}$$

Since,

$$\begin{aligned} \delta E &= 2\operatorname{Tr}(h)\delta P + 2\operatorname{Tr} G(P)\,\delta P, \\ &= 2\operatorname{Tr}(h + G(P))\delta P, \end{aligned}$$

equation (7) becomes

$$4(E - E_L)\operatorname{Tr}\{[h + G(P)]\}\delta P + \beta_1 \operatorname{Tr}\{(4P^3 - 6P^2 + 2P)\}\,\delta P = 0,$$

which further simplifies to

$$\operatorname{Tr}\left[\{(E - E_L)[h + G(P)]\} + \frac{\beta_1}{2}(2P^3 - 3P^2 + P)\right]\delta P = 0. \tag{8}$$

Since the variations in the density variables δP are arbitrary and independent, (8) implies that

$$(E - E_L)[h + G(P)] + (2P^3 - 3P^2 + P)\frac{\beta_1}{2} = 0,$$

which rearranges to

$$2(E - E_L)[h + G(P)]\beta_1^{-1} + (2P^3 - 3P^2 + P) = 0. \tag{9}$$

It is not difficult to see that (9) immediately leads to an iterative (self-consistent) scheme for the determination of the orthonormality-constrained density matrix P that satisfy our variational condition on energy also. To do this, however, we must have a means of estimating E_L locally. Assuming that E^i_L is a current estimate of the lower bound to the ground state energy (currently being approximated by E^i_L) to be determined, we can rewrite (9) easily as follows:

$$\overline{P}_i = 3P_i^2 - 2P_i^3 - 2[h+G(P)]\,(E_i - E^i_L)\beta_1^{-1}. \tag{9a}$$

We may now note that the two sides of (9a) will not be equal unless we have got the appropriately constrained density matrix (i.e. a correct solution of (9)). Otherwise the right side of (9a) merely furnishes a new estimate of $P = \overline{P}_i$ leading us therefore to the following iterative scheme.

$$P_i \rightarrow P_{i+1} = 3P_i^2 - 2P_i^3 - 2[h+G(P)]\,(E_i - E^i_L)\beta_1^{-1}, \tag{10}$$

with

$$E^{i+1}_L = E^i_L + a\{\mathrm{Tr}\,(A_i^\dagger A_i)\}^{1/2}, \tag{11}$$

where

$$A_i = P_i + 2P_i^3 - 3P_i^2 + 2[h+G(P_i)]\,(E - E^i_L)\beta_1^{-1}, \tag{12}$$

and a is a parameter satisfying the condition $0 < a \leqslant 1$ and E^i is the constrained energy at the ith iteration. Clearly, the iterations terminate only when

$$(E^{i+1}_L - E^i_L) = 0,$$

implying that

$$E^{i+1}_L = E^i_L,$$

a condition that can be fulfilled only if

$$\mathrm{Tr}\,(A_i^\dagger A_i) = 0,$$

which means that appropriately constrained density matrix P has been found. In course of iterations $\mathrm{Tr}\,(A_i^\dagger A_i)$ is expected to decrease gradually as (i) $|E_i - E^i_L| \rightarrow 0$ and (ii) as idempotency constraint on P is satisfied more and more exactly (note that if $P_i^2 = P_i$, $P_i + 2P_i^3 - 3P_i^2 = 0$).

3.2 *Multiple constraints*

Let us suppose that the set of Hermitian operators $\{\hat{B}_k\}$ (which may not commute with the many-electron hamiltonian $\hat{H}$, see Mukherjee and Karplus 1963) represents a set of observables of our system. Let b_k^0's represent the exact (theoretical or experimental) values of these observables for our system in the ground state. If we now demand that the variational criterion for the orthonormality constrained wavefunction or the one-electron density matrix should not only aim at minimising the energy, but also minimize the mean square deviation of the calculated values of the observables from their experimental counterparts at the cost of obtaining a slightly inferior value of E, we should not use E^L_C but use a modified functional (E_{MC}) which incorporates in it all the additional constraints. One must not fail to note that we are attempting a minimization in the least square sense. This is particularly important since $\hat{H}$ and a set of noncommuting Hermitian operators $\{\hat{B}_k\}$ cannot have simultaneous eigenstates. Thus, with a trial one-electron density P, we can define a functional E_{MC} (MC stands for multiple constraint) as follows (β_1 and λ_k's are penalty weight factors of appropriate dimensions)

$$E_{\mathrm{MC}} = (E - E_0^{\mathrm{HF}})^2 + \beta_1\,\mathrm{Tr}\,(P^2 - P)^2 + \sum_k \{\mathrm{Tr}\,(PB_k) - b_k^0\}^2\,\lambda_k \tag{13}$$

clearly, $E_{\mathrm{MC}} = 0$ only if $E = E_0^{\mathrm{HF}}$, $P^2 = P$

and $\sum_k \{\mathrm{Tr}\,(PB_k) - b_k^0\}^2 = 0.$

Again, for practical applications E_0^{HF} will not be known to start with so that we replace it by a suitable estimate of a lower-bound to E_0^{HF} (say E_L) and redefine the functional as follows:

$$E_{\mathrm{MC}}^L = (E - E_L)^2 + \beta_1 \,\mathrm{Tr}\,(P^2 - P)^2 + \sum_k [\mathrm{Tr}\,(PB_k) - b_k^0]^2 \lambda_k. \tag{14}$$

If an arbitrary variation δP is now introduced in P, the first order change in the constrained functional will be given by

$$\begin{aligned}\delta E_{\mathrm{MC}}^L &= 2(E_i - E_L^i)\,\delta E_i + \beta_1 \,\mathrm{Tr}\,\{(4P_i^3 - 6P_i^2 + 2P_i)\delta P_i\}\\ &\quad + \sum_k [2\,(\mathrm{Tr}\,(P_i B_k) - b_k)\,\mathrm{Tr}\,(B_k\,\delta P_i)]\,\lambda_k\\ &= 4(E_i - E_L^i)\mathrm{Tr}\,\{h + G(P_i))\delta P_i\}\\ &\quad + \beta_1 \,\mathrm{Tr}\,\{(4P_i^3 - 6P_i^2 + 2P_i)\delta P_i\}\\ &\quad + \sum_k 2d_k^i \,\mathrm{Tr}\,(B_k\,\delta P_i)\lambda_k,\end{aligned} \tag{15}$$

where d_k^i represents the deviation of the calculated value of the kth property from its exact or experimental value at the ith iterative stage, P_i is the trial one-electron density matrix at the ith iterative stage and E_i is the corresponding trial energy while E_L^i is the lower bound to E estimated at the ith stage of the process. At the stationary point, $\delta E_{\mathrm{MC}}^L = 0$ for arbitrary variation δP in P and this condition leads to an iterative scheme for the self-consistent determination of P. Thus since $\delta E_{\mathrm{MC}}^L = 0$ and δP is arbitrary, we have from (15)

$$2(E_i - E_L^i)(h + G(P_i)) + (2P_i^3 - 3P_i^2 + P_i)\beta_1 + \sum_k \lambda_k d_k^i B_k = 0. \tag{16}$$

Rewriting (16) slightly, we have the iterative scheme we are looking for

$$P_{i+1} = 3P_i^2 - 2P_i^3 - \left[\sum_k \bar{\lambda}_k d_k^i B_k + 2(E_i - E_L^i)\,[h + G(P_i)]\beta_1^{-1}\right] \tag{17}$$

where

$$\bar{\lambda}_k = \lambda_k/\beta_1,$$

$$E_L^{i+1} = E_L^i + a\{\mathrm{Tr}\,(A_i^\dagger A_i)\}^{1/2}, \tag{18}$$

$$A_i = 2(E_i - E_L^i)\,(h + G(P_i))\beta_1^{-1} + 2P_i^3 - 3P_i^2 + P_i + \sum_k \bar{\lambda}_k d_k^i B_k, \tag{19}$$

and a is a damping factor satisfying $0 < a \leqslant 1$.

4. Results and discussion

The two schemes outlined in §3 for the self-consistent determination of singly or multiply constrained one-electron density matrix of an atomic or molecular system are quite general. If the only constraint imposed on P is the orthonormality constraint, the iterative scheme based on (9) should lead to the Hartree-Fock density itself. However, one must carefully study the convergence behaviour of the iterative sequence generated by (10) and (17). To test the numerical viability of the algorithms suggested we will present here the results of some exploratory calculations. The calculations have been performed at the CNDO/2 level of approximation (Pople and Segal 1965). Since our main purpose here is to show that the proposed algorithms are indeed workable, the use of CNDO/2 level of approximation is of no major consequence. Apart from providing evidence of workability these results also shed some light on the convergence behaviour of these algorithm and should provide a guideline for similar investigation in an ab initio framework.

4.1 *The case of a single constraint*

In figure 1 we have displayed the convergence profile of the proposed algorithm (cf equation (10)). The molecule used is CO with the experimental C–O bond length. The iterative procedure was started with a trial non-idempotent one-electron density matrix $P_0 (P_0^2 \neq P_0)$ which led to an energy $E^0 = -36{\cdot}797114$ a.u. much below the actual Hartree-Fock energy within the CNDO/2 approximations. The value of the Morrison-function (functional in the present case) at the start was 0·406376 and the starting value of the lower bound to the constrained energy (E_L^0) was set equal to $E_L^0 = -37{\cdot}296020$ a.u. The monotonic but slow decrease in the magnitude of the Morrison-functional is clearly evident from the figure. Although the figure displays the progress of the minimization only upto 67 iterations, the iterations were continued well beyond 67 iterations. At 67th iterations, $E = -36{\cdot}33791404$, $E_L = -36{\cdot}338096$ a.u. which are already close to HF result. In 92 iterations the functional decreased to a value as low as

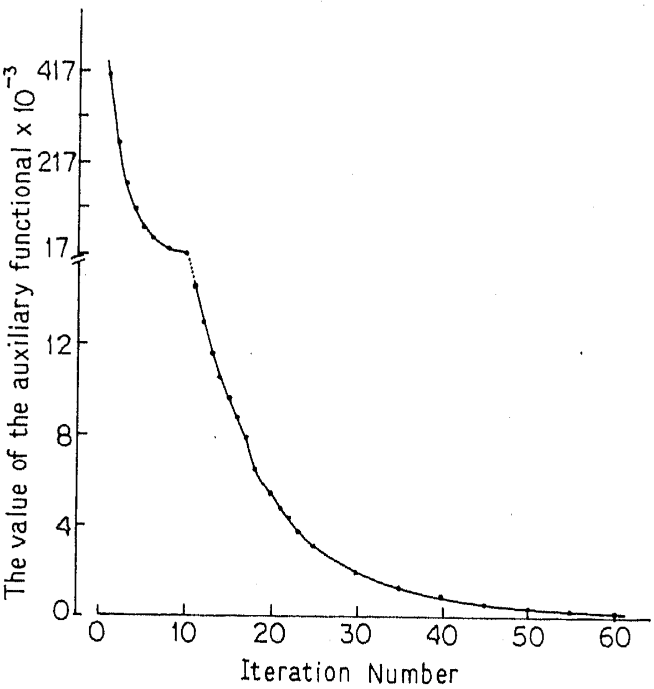

Figure 1. Actual value of the auxiliary (constrained) functional shown as a function of number of iterations. The system is carbon monoxide molecule at equilibrium geometry.

10^{-6}. The energy E^i and the lower bound E^i_L merged with one another. The energy E so calculated tally nicely with the HF energy of the system under CNDO/2 approximations.

4.2 *The case of more than one constraint*

The exploratory calculations under this heading have been carried out on lithium hydride molecule ($r_{\text{Li-H}} = 2{\cdot}3428$ a.u.). In addition to the idempotency constraint, one external constraint has been imposed on P in that we demand

$$\mathrm{Tr}\,\{P\hat{B}_k\} = b_k^0,$$

where $\hat{B}_k = e \cdot \mathbf{r}$ and b_k^0 is the constrained expectation value of $\hat{B}_k$ which we want the constrained density P to reproduce. Tr $\{P\hat{B}_k\}$ with P taken to the HF density of Li-H under CNDO approximations was found to be equal to 2·62 D. Figure 2(a) shows the convergence behaviour of the algorithm where the constraining value of the observable b_k^0 was chosen equal to 2·75 D while figure 2(b) represents an identical plot with $b_k^0 = 2{\cdot}8$ D. In each case, there is oscillation in the functional value to start with when P is highly non-idempotent. The oscillatory behaviour is seen to be replaced by a smooth and monotonically decreasing tendency after the first few iterations. Although smooth, the approach to convergence is rather slow and we feel that some kind of damping or

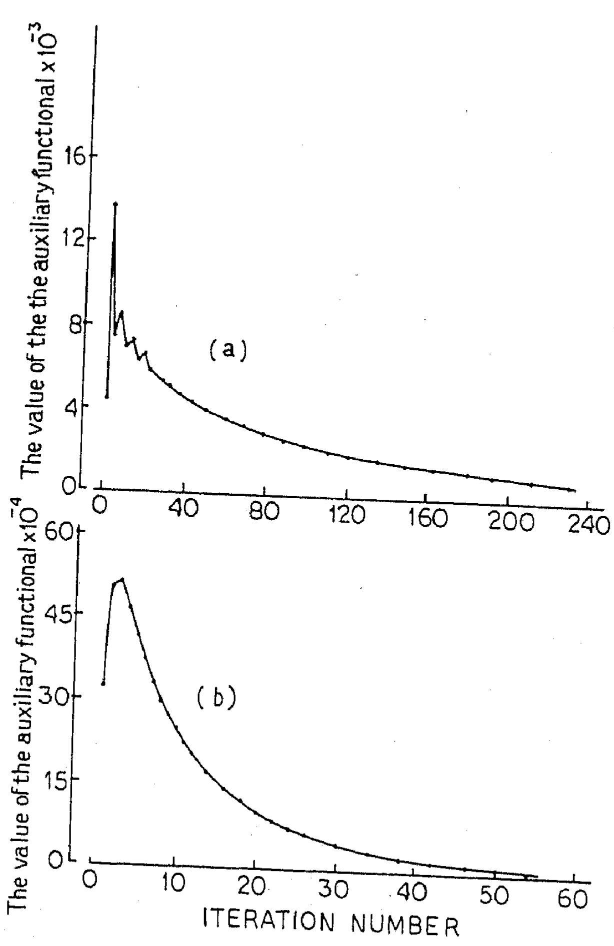

Figure 2(a). Minimization of the auxiliary functional shown as function of the number of iterations in the multiply constrained case. The system of LiH molecule with $b_k^0 = 2{\cdot}75$ D (see text for details. (**b**) The same plot with $b_k^0 = 2{\cdot}80$ D.

interpolation technique has to be introduced to accelerate the approach to convergence. This is being studied at present. As expected, the constrained density $\tilde{P}$ leads to energy higher than the Hartree-Fock energy. The sacrifice in energy (ΔE) being higher, higher the value of $|b_k^0 - b_k^{\mathrm{HF}}|$ (Byers Brown 1966).

4.3 *Certain general characteristics of the algorithms*

(i) The choice of penalty weight factors (β_1 or $\bar{\lambda}_k$) appears to be highly critical. A proper choice of β_1 (or $\bar{\lambda}_k$, as the case may be) may force convergence on an otherwise diverging sequence. Generally, one has to start with a small value of β_1 or $\bar{\lambda}_k$ which are made to increase progressively as the density gradually approaches the constrained region of the density variable.

(ii) A good estimate of E_L^0 is very essential. In the multiply constrained case one starts with the HF density in general. Since the constrained energy $E_c > E_0^{\mathrm{HF}}$, E_L^0 is in reality an upper bound to E_c and can be set equal to $(E_{\mathrm{HF}}^0 + \sigma)$ where $\sigma \approx 0{\cdot}1$ a.u. The constant a in (11) or (18) is crucially important in controlling the overall behaviour of the corresponding iterative sequences. In the multiply constrained case $-1 \leqslant a < 0$ (E_L is an upper bound) while in the orthonormality constrained case $0 < a \leqslant 1$.

(iii) It is possible to diversify the algorithm further and use a hybrid algorithm that uses both the Lagrangian multiplier and penalty function ways of satisfying equality constraints. Such a scheme is being numerically tested at present. The comparative features of this hybrid algorithm and the purely penalty function-based alternative techniques will be presented separately (Das and Bhattacharyya 1985).

(iv) It is possible to minimize the constrained functionals directly by adopting the steepest descent technique for example, or by adopting the method of conjugate gradients (Mukherjee 1975).

(v) The extension of these algorithms to the unrestricted open-shell case is straightforward. For the restricted open-shell, however, there are certain additional complications. The algorithms for handling the open-shell problem will be presented shortly (Das and Bhattacharyya 1985).

Acknowledgements

The authors sincerely thank the Department of Science and Technology, New Delhi for a research grant. The authors wish to thank Dr D Mukherjee and Dr (Mrs) Bharati Bhattacharyya for useful discussions.

References

Bamzai A S and Deb B M 1981 *Rev. Mod. Phys.* **53** 593

Byers Brown W 1966 *J. Chem. Phys.* **44** 567

Chong D P 1973 *J. Chem. Phys.* **59** 1721

Das K K and Bhattacharyya S P 1985 *Chem. Phys. Lett.* (Communicated)

Fiacco A V 1970 *J. Opt. Theory. Appl.* **6** 252

Fiacco A V and McCormik G P 1968 Nonlinear programming: Sequential unconstrained optimization technique (New York: Wiley)

Fock V 1930 *Z. Phys.* **61** 126
Hartree D R 1928 *Proc. Cambridge Philos. Soc.* **24** 89
Laenberger D 1971 *J. Opt. Theory Appl.* **7** 39
McWeeny R 1956 *Proc. R. Soc. (London)* **A235** 496
McWeeny R 1957 *Proc. R. Soc. (London)* **A241** 239
Morrison D D 1968 *SIAM J. Numer. Anal.* **5** 83
Mukherjee A and Karplus M 1963 *J. Chem. Phys.* **38** 44
Mukherjee D 1975 *Int. J. Quant. Chem.* **9** 943
Pople J A, Santry D P and Segal G A 1965 *J. Chem. Phys.* **43** S129
Pople J A and Segal G A 1965 *J. Chem. Phys.* **43** S136
Rasiel Y and Whitman D R 1965 *J. Chem. Phys.* **42** 1965
Roothaan C C J 1951 *Rev. Mod. Phys.* **23** 69
Slater J C 1930 *Phys. Rev.* **35** 210
Westhaus P 1983 *J. Chem. Phys.* **78** 6833
Zeiss G and Whitehead M A 1983 *Int. J. Quant. Chem.* **24** 651