JOURNAL OF COMPUTATIONAL PHYSICS **16**, 160–166 (1974)

Stabilization of Solution of the Hartree–Fock Equations*

ROBERT D. COWAN AND JOSEPH B. MANN

University of California, Los Alamos Scientific Laboratory,
Los Alamos, New Mexico 87544

Received March 28, 1974

A very simple method is described for eliminating certain instabilities and other difficulties in the solution of the Hartree–Fock (HF) equations. The method consists of subtracting, from both the homogeneous and inhomogeneous portions of the differential equation, terms which exactly compensate each other in the self-consistent limit. These terms are chosen so as to largely cancel the original inhomogeneous term and thereby make the equation a nearly homogeneous one.

INTRODUCTION

The electronic structure of an N-electron atom is usually discussed in terms of the concept of an electron configuration

$$(n_1l_1)^{w_1}\,(n_2l_2)^{w_2}\cdots(n_ql_q)^{w_q}, \qquad \sum w_q = N, \tag{1}$$

using basis wavefunctions constructed from products of N one-electron functions

$$\varphi_{nlm}(r, \theta, \phi) = r^{-1}P_{nl}(r)\,Y_{lm}(\theta, \phi), \tag{2}$$

where Y_{lm} is a spherical harmonic. The radial functions $P_i(r) \equiv P_{n_il}(r)$, one for each subshell $(n_il_i)^{w_i}$, are usually found by solution of the Hartree–Fock equations, which consist of a set of q-coupled inhomogeneous differential equations, each of the form

$$[(d^2/dr^2) + f_i(r) - \epsilon_i]\,P_i(r) = g_i(r). \tag{3}$$

The eigenvalue ϵ_i (which is here defined to be positive for bound electrons) is determined by the requirements that

$$P_i(0) = P_i(\infty) = 0, \tag{4}$$

$$\text{number of nodes of } P_i(r) = n_i - l_i - 1, \tag{5}$$

---

* Work done under the auspices of the US Atomic Energy Commission.

Copyright © 1974 by Academic Press, Inc.
All rights of reproduction in any form reserved.

and

$$\| P_i \| \equiv \left[\int_0^\infty P_i^2 \, dr \right]^{1/2} = 1; \tag{6}$$

by convention, the arbitrary phase of $P_i(r)$ is chosen such that the "initial slope" is positive:

$$a_0 \equiv [P_i(r)/r^{l_i+1}]_{r=0} > 0. \tag{7}$$

The functions f_i and g_i are somewhat complicated integral functions of the $P_j(r)$, $1 \leqslant j \leqslant q$. Consequently, solution of the HF equations is usually carried out by a self-consistent iterative procedure of the following type.

(a) On the *m*th iteration cycle, a set of trial functions $P_j{}^m$(input) is chosen;

(b) Using these $P_j{}^m$, the functions $f_i(r)$ and $g_i(r)$ are computed (each i);

(c) For each of several values of ϵ_i the differential equation (3) is integrated numerically, until that value is found for which the function $P_i{}^m$(output) satisfies the conditions (4)–(7) (each i).

Steps (a) to (c) are repeated using new trial input functions, until each output function is identical with the corresponding input function within some suitable self-consistency criterion. The usual method of choosing new trial functions is to take

$$P_j^{m+1}(\text{input}) = cP_j{}^m(\text{output}) + (1 - c)\, P_j{}^m(\text{input}). \tag{8}$$

In almost all cases, this method works quite satisfactorily with some suitable value of c between about 0.3–0.7. However, for neutral and one- or twofold ionized atoms in the vicinity of the transition and rare-earth elements, the above iterative procedure diverges monotonically for excited d and/or f orbitals [1]. Several methods have been developed for stabilizing the iteration and obtaining self-consistent solutions [1, 2]. These involve either use of negative values of c (i.e., extrapolation rather than interpolation to obtain a new trial input function), or choice of some unnormalized integral of the differential equation for the output function $P_i{}^m$(output), i.e., temporary relaxation of the condition (6). In practice, each of these methods tends to be rather tricky and slow to converge, and to possess something less than universal applicability. We here describe a new method which is very simple and logical, and which appears to converge rapidly and straightforwardly in all cases.

NEW METHOD

In order to understand this new method we first consider Fig. 1, taken from Griffin, Cowan and Andrew [1]. The solid curve represents the norm of that integral $P_i(r)$ of the differential equation (3) which satisfies (4), as a function of the energy parameter ϵ_i. This curve has singularities at the eigenvalues $\epsilon_H{}^1$, $\epsilon_H{}^2$,..., of the homogeneous equation obtained by setting $g_i(r) \equiv 0$ in (3). Let us suppose that we were seeking an integral with a single node, for example, a $3p$, $4d$ or $5f$ function. If the ordinate scale is such that unit norm occurs along the line C, then the desired integral corresponds either to the point C_1 or to the point $C_1{}'$ (whichever integral has positive initial slope). In either case, the value of ϵ lies close to the eigenvalue $\epsilon_H{}^2$ for the one-node integral of the homogeneous equation; it may be inferred that the effect of the exchange term $q_i(r)$ in (3) is small and that the integral $P_i{}^m$(output) is much like the integral of the homogeneous equation. If, on the other hand, unit norm corresponds to the horizontal line B, then one of the single-node integrals ($B_1{}'$) occurs close to $\epsilon_H{}^2$, but the other (B_1) occurs closer to the eigenvalue $\epsilon_H{}^1$ of the *zero*-node integral of the homogeneous equation. In some cases with unit norm at C, the variation with ϵ in the form of $P(r)$ is quite different[1] from that corresponding to Fig. 1, and the single-node normalized integral with positive initial slope occurs at the point marked C_0.

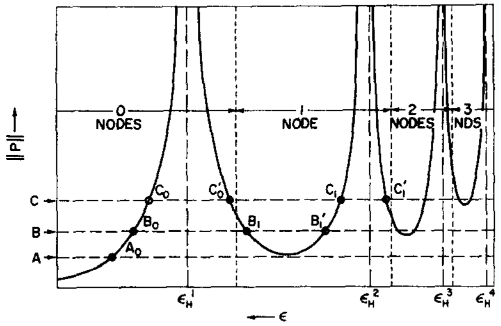

FIG. 1. Typical norm curve for those integrals $P(r)$ of the HF Eq. (3) which satisfy the boundary conditions (4). Singularities appear at energies $\epsilon_H{}^k$ ($k = 1, 2, 3, 4,...$) that correspond to integrals of the associated homogeneous equation having 0, 1, 2, 3,..., nodes, respectively.

Whenever the self-consistent integral corresponds to one of the cases B_1 or C_0, it is evident that the effect of the exchange term $g_i(r)$ in (3) is very large, and that the desired single-node integral of (3) is rather different from the single-node integral of the homogeneous equation. These are the cases which are unstable in the sense described above, as can be seen from the plausibility argument given by Griffin *et al.* [1]: Suppose that the exchange term $g_i(r)$ is scaled down by some constant factor which is made to gradually approach zero, so that the inhomoge-

---

[1] Compare the lower half of Fig. 5 of [1] (which is, however, difficult to read clearly; the node count changes by unity at points slightly to the left of each singularity).

neous equation approaches the homogeneous equation as a limit. It may be seen from (3) that, for any ϵ, the integral $P_i(r)$ of the inhomogeneous equation is scaled down by this same factor; thus each section of the norm curve decreases everywhere and approaches the limiting form of a square well (i.e., $\| P \| = 0$ everywhere except at $\epsilon = \epsilon_H{}^k$). In any of the cases C_1, C_1' or B_1', the lowering of the norm curve forces ϵ toward $\epsilon_H{}^2$, and the normalized integral of the inhomogeneous equation approaches the correct (single-node) integral of the homogeneous equation. However, in the cases B_1, C_0 and C_0', ϵ is forced toward $\epsilon_H{}^1$, and the normalized integral of (3) approaches the wrong (zero-node) integral of the homogeneous equation.

Evidently, the self-consistent iteration can be stabilized by modifying the differential equation so as to make the effect of the inhomogeneous term small. Thus, let us consider the equation

$$[(d^2/dr^2) + f_i(r) - h_i(r) - \epsilon_i]\, P_i(r) = q_i(r) - h_i(r)\, P_i(\text{input}). \tag{9}$$

When self-consistency has been reached, $P_i(r)$ will be essentially identical with the trial input function $P_i(\text{input})$ and so (9) will reduce to the desired equation (3). Meanwhile, if $h_i(r)$ is chosen such that the right-hand side of (9) has nearly everywhere a magnitude much smaller than $| g_i(r)|$, then the inhomogeneous equation will be quite close to its associated homogeneous equation and the required solution of (9) will always correspond to one of the cases C_1, C_1' or B_1' in Fig. 1. That is, the presence of the term $-h_i(r)$ on the left-hand side of (9) has shifted the positions of the eigenvalues $\epsilon_H{}^k$ of the homogeneous equation such that the value of ϵ_i for the appropriate normalized single-node integral of (9) necessarily lies close to $\epsilon_H{}^2$.

A possible choice for the function $h_i(r)$ is a suitable modification [3, 4] of the Slater $\rho^{1/3}$ exchange term. However, a choice which is both simpler and more foolproof (especially when starting the iteration with poor trial functions, or when making calculations for a specific LS term rather than for the configuration-average energy) is to use simply

$$h_i(r) = kg_i(r)/P_i(\text{input}), \tag{10}$$

where $k = 0.95$ appears to be satisfactory in all cases. The only complications arise from the singularities in h_i that exist at the positions of the nodes of $P_i(\text{input})$. These singularities can be avoided by replacing the denominator of (10) with a function equal to $P_i(\text{input})$, except that its magnitude at any point r is never allowed to become smaller than (say) 0.01 times the largest magnitude of $P_i(\text{input})$ for any $r_1 \leqslant r$; an example is shown in Fig. 2. This modification still leaves a discontinuity in h_i at the position of each node of $P_i(\text{input})$. In practice, these discontinuities seem to have only negligible effect on the integral $P_i(r)$ obtained by numerical integration of (9); however, if desired, $h_i(r)$ can be smoothed by linear interpolation over a range of (say) five integration-mesh intervals centered about each discontinuity.

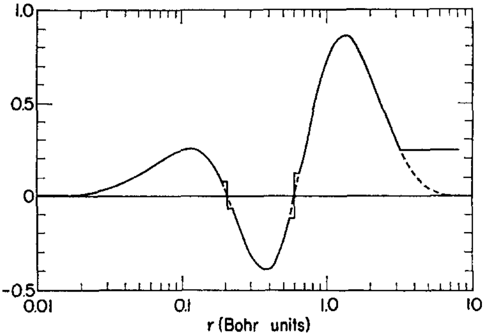

FIG. 2. Dashed curve: A typical radial function $P_{nl}(r)$ having two nodes ($n = l + 3$). Solid curve: Modified function for use in the denominator of (10); for clarity of illustration, the cutoff has been made at 0.3 (rather than 0.01) times the magnitude of the preceding antinode.

APPLICATIONS

The method represented by Eqs. (9) and (10), with the above-mentioned modification of the denominator of (10), and with or without the further smoothing of h_i, is almost trivially easy to code into any existing computer program for the numerical self-consistent-field solution of the HF equations. We have applied this procedure to the outermost, singly occupied orbital in a wide variety of configurations, including not only cases where the self-consistent solution of (3) corresponds to the point B_1 in Fig. 1 ("type 2" in the classification of Griffin *et al.*) and to C_0 (type 3), but also to cases B_1' and C_1 (type 1) and C_1' (type 0),[2] where this procedure is not usually necessary. In all cases, this procedure converted the problem to one of type 1 or type 0. The self-consistent iteration therefore converged smoothly and rapidly, with the use of (8) and $c \simeq 0.3$ or 0.4,[3] and computing times were thus essentially the same as for normally stable cases; examples are given in

---

[2] Since exchange effects increase the strength of electron binding and therefore the magnitude of ϵ_i, one would not normally expect solutions with $|\epsilon_i|$ less than the corresponding $|\epsilon_H{}^k|$. However, type 0 solutions do occur when making calculations for a specific LS term (e.g., for the d electron of $pd\,{}^1P$) rather than for the configuration average. (The node-count arrangement is like that of Fig. 5 of [1]; see footnote 1).

[3] For La I $6s5d5f$ the new procedure is required for both $5d$ and $5f$, and c had to be cut to about 0.02.

Table I. The advantage of the new method, however, lies not so much in relatively minor savings in computing time as in providing an easily programmed, universal method which practically eliminates frustrating convergence failures.

TABLE I

Computing Times[a] to Self Consistency (seconds) for a CDC 7600

Configuration	Old Method	New Method
Ti I $3d^2 4s4p$	12	11
Ti I $3d^2 4s4d$	18[b]	11
La I $4f6s^2$	12	11
La I $6s^2 5f$	16[b]	11
Os I $5d^6 6s6p$	54	50
Os I $5d^6 6s5f$	58[b]	37

[a] Times are for 6 SCF iteration cycles for all Ti runs, and are for 5 SCF cycles for all La and Os runs. Times would be about five times longer for a CDC 6600 and fifteen times longer for an IBM 7094 II.

[b] Using method M2 of Griffin *et al.* [1] for the final orbital.

The price paid for this convenience is that the rapid changes in $h_i(r)$ near the zeros of $P_i(r)$ introduce minor numerical inaccuracies in integration of the differential equation. However, to all intents and purposes (i.e., to at least the seventh decimal place) the final function $P_i(r)$ was independent of the value of k used in (10), of the cutoff value (0.01 or 0.1) used in modifying the denominator of (10), and of whether or not the discontinuities in $h_i(r)$ were smoothed out, and was identical with the function obtained by the methods of Griffin *et al.* [1].

For some d orbitals where the solution is of type 2 when using (3), a value of k as small as 0.3 is sufficient to convert the solution of (9)–(10) to type 1 and thereby make the new method convergent. However, for f orbitals where the solution of (3) is of type 3, k cannot be smaller than about 0.9; this is because the effect of h_i in (9) must be large enough to produce the "collapse" of the integral of the homogeneous equation and the corresponding unit decrease in the effective quantum number

$$n_k^* = (\epsilon_H^k)^{-1/2},$$

and this collapse depends critically on the depth of the effective potential energy function $h_i - f_i$ in the inner-well region [5].

For simplicity, we suggest that a value $k \simeq 0.95$ be used in all cases, even when a smaller value would suffice. This has the minor disadvantage of always producing a norm curve which has a very small magnitude except in the regions $\epsilon_i \cong \epsilon_H{}^k$, where the curve turns up very sharply and steeply towards ∞. However, this introduces only minor complications into the iteration on ϵ_i to find a normalized function $P_i(r)$, especially if a hyperbolic relation

$$P_i(r) \propto (\epsilon_i - \epsilon_H{}^k)^{-1}$$

is assumed for the prediction of a new trial value for ϵ_i.

The method (9)–(10) is valuable also for cases (including p orbitals, such as $3p$ in the configuration Na I $2p^6 3p$) where the self-consistent solution of (3) is of the stable type 1 or 0, but where the initial estimate of the trial function P_i(input) is so poor as to result (on the first cycle) in a norm curve having a minimum greater than unity [1], as when unit norm occurs along the line A in Fig. 1. Routine use of (9) on at least the first cycle or two assures that the norm curve will always lie low enough that there exists a normalized output function with the correct number of nodes and positive initial slope.

We have also found this new method to be useful in solution of the relativistic Dirac–Hartree–Fock equations [6]. It has made possible solution for certain excited d and f orbitals where standard methods failed completely, presumably for exactly the same reasons as in the HF problem, though we have not investigated this in detail.

ACKNOWLEDGMENT

We are indebted to D. C. Griffin, whose modification of C. Froese Fischer's well-known HF program was used for the present investigation.

REFERENCES

1. D. C. GRIFFIN, R. D. COWAN AND K. L. ANDREW, *Phys. Rev. A*3 (1971), 1233.
2. C. FROESE FISCHER, *J. Computational Phys.* **10** (1972), 211; *Comput. Phys. Commun.* **4** (1972), 107.
3. R. D. COWAN, *Phys. Rev.* **163** (1967), 54.
4. I. LINDGREN, *Int. J. Quant. Chem.* **S5** (1971), 411; A. ROSÉN AND I. LINDGREN, *Phys. Scrip.* **6** (1972), 109.
5. D. C. GRIFFIN, K. L. ANDREW AND R. D. COWAN, *Phys. Rev.* **177** (1969), 62.
6. J. B. MANN AND J. T. WABER, *Atomic Data* **5** (1973), 201.