JOURNAL OF COMPUTATIONAL PHYSICS 22, 435–450 (1976)

# A Finite-Difference Newton–Raphson Solution of the Multiconfiguration Hartree–Fock Problem

W. R. FIMPLE, B. J. MCKENZIE, AND S. P. WHITE

*Department of Physics, University of Canterbury, Christchurch, New Zealand*

Received February 6, 1976; revised April 15, 1976

The finite-difference Newton–Raphson algorithm coupled with a self-consistent field iteration, which recently has proved to be very successful in solving the atomic Hartree–Fock equations for a single configuration, has been extended to treat the multiconfiguration case. The problem is algebraicized through the introduction of finite-difference variables, treating the multipliers on normalization and orthogonality on an equal footing with the other variables, and the resulting large system of nonlinear algebraic equations is solved by means of a generalized Newton–Raphson iteration. Because of the particular ordering of the variables and equations and the coupled SCF iteration employed, the unit operation of the method involves the inversion of a partly block tridiagonal Jacobian matrix and can be solved rapidly by means of a partitioning. A solution of the $1s2s\,^1S$ excited state of helium is presented as an example.

## 1. INTRODUCTION

Recently a finite-difference Newton–Raphson algorithm (FDNRA), originally developed by Van Dine [1], has been successfully applied to the solution of the atomic Hartree–Fock (HF) problem [2]. The algorithm has also been used to solve the two-center electronic Schrodinger equation [3] and the factored HF equations for diatomic molecules [4].

In the present work the algorithm has been further extended to solve the atomic multiconfiguration Hartree–Fock (MCHF) equations. The pioneering computational work in this field has been done by Froese–Fischer [5, 6] using numerical integration techniques [7]. She reports convergence difficulties for configurations such as $1s2s\,^1S$ [6] and considers this case to be a critical test of any MCHF algorithm [8]. The difficulty is apparently associated with a strong orthogonality condition between two orbitals involving the same angular quantum numbers, both of which are in incomplete groups. This situation leads to significantly large values of off-diagonal multipliers.

Since one of the advantages of the FDNRA is the treatment of all multipliers on equal footing with the other variables, rather than as adjuncts to the problem,

it was thought that the finite-difference approach might be successful in solving the MCHF equations for the $1s2s^1S$ configuration of helium. Therefore, after first presenting the algorithm in general for atomic MCHF calculations, we apply it to this specific case as a first test for convergence.

In the following section the MCHF problem is outlined, and in Section 3 the system of equations is converted to finite-difference form. The method of solution, which is a generalization of that given in [2], is presented in Section 4 with further details in an Appendix, and the results of the $1s2s^1S$ test case are given in Section 5.

## 2. The MCHF Problem

Let the total wavefunction of an $N$-electron atom be approximated by

$$\Psi = \sum_{I=1}^{L} a_I \Phi_I , \tag{1}$$

where the

$$\Phi_I = \mathscr{A}(\phi_1{}^I(1) \cdots \phi_N{}^I(N)), \qquad I = 1,..., L, \tag{2}$$

are determinantal functions of $N$ spin–orbitals $\phi_\alpha{}^I$, $\alpha = 1,..., N$, and the $a_I$ are constant mixing coefficients. The mixing coefficients $a_I$ and the spin–orbitals $\phi_\alpha{}^I$ are determined variationally. The $a_I$ are generally not all independent. The ratios of certain ones of them must be held fixed in order to make (1) an eigenfunction of total angular momentum, depending upon the coupling scheme.

Following [2] we factorize the spin–orbitals into the form,

$$\phi_\alpha{}^I \equiv P_\alpha{}^I(\rho)\, Y_{l_\alpha{}^I}^{m_\alpha{}^I}(\omega)\, \chi_{1/2}^{m_{S_\alpha}^I}, \qquad \alpha = 1,..., N; \quad I = 1,..., L, \tag{3}$$

where

$$P_\alpha{}^I(\rho) \equiv \rho R_\alpha{}^I(r(\rho)), \tag{4}$$

with $R$ the usual radial function, $Y$ a spherical harmonic, and $\chi$ a spinor; and the radial variable $\rho$ is given in terms of the ordinary radius by the transformation

$$\rho \equiv r/(1 + ar) \qquad (a > 0). \tag{5}$$

As explained in [2] these variables are well suited to finite differences. In particular the transformation (5) maps the infinite range of $r$ into the finite region $0 \leqslant \rho \leqslant 1/a$.

In terms of the variables and orbitals defined by Eqs. (1) through (5) the MCHF equations are a set of coupled integrodifferential equations of the following form.

$$a_I \sum_J a_J \left[\left(-\frac{1}{2}\frac{d^2}{d\rho^2} - \frac{Z}{\rho(1-a\rho)^3} + \frac{l_\alpha^J(l_\alpha^J+1)}{2\rho^2(1-a\rho)^2}\right)\Delta(\alpha, I; \alpha, J)\, P_\alpha^J(\rho) + \frac{Y_\alpha^{IJ}(\rho)}{(1-a\rho)^4} P_\alpha^J(\rho) - \frac{X_\alpha^{IJ}(\rho)}{(1-a\rho)^4}\right] - \sum_{J,\beta} \frac{\lambda_{\alpha\beta}^{IJ}}{(1-a\rho)^4}\Delta(\alpha, I; \beta, J)\, P_\beta^J(\rho) = 0,$$

$$\alpha = 1,\ldots, N; \quad I = 1,\ldots, L, \quad \text{(6a)}$$

where $\Delta(\alpha, I; \beta, J) \equiv \delta(l_\alpha^I l_\beta^J)\,\delta(m_\alpha^I m_\beta^J)\,\delta(m_{S_\alpha}^I m_{S_\beta}^J)$,

$$Y_\alpha^{IJ}(\rho) \equiv \delta(m_{S_\alpha}^I m_{S_\alpha}^J) \sum_\beta \delta(m_{S_\beta}^I m_{S_\beta}^J) \sum_{\lambda,\mu} (4\pi/(2\lambda+1))\, B_{\alpha\alpha}^{IJ} B_{\beta\beta}^{JI*} \cdot \int_0^{1/a} P_\beta^I(\rho')\, P_\beta^J(\rho')\, g_\lambda(\rho, \rho')\, d\rho',$$

and

$$X_\alpha^{IJ}(\rho) \equiv \sum_\beta \delta(m_{S_\alpha}^I m_{S_\beta}^J)\,\delta(m_{S_\alpha}^J m_{S_\beta}^I) \sum_{\lambda,\mu} (4\pi/(2\lambda+1))\, B_{\alpha\beta}^{IJ} B_{\alpha\beta}^{JI*} P_\beta^J(\rho) \cdot \int_0^{1/a} P_\beta^I(\rho')\, P_\alpha^J(\rho')\, g_\lambda(\rho, \rho')\, d\rho',$$

where

$$B_{\alpha\beta}^{IJ} \equiv \int_0^{4\pi} Y_{l_\alpha^I}^{m_\alpha^I}(\omega)^*\, Y_\lambda^\mu(\omega)^*\, Y_{l_\beta^J}^{m_\beta^J}(\omega)\, d\omega,$$

and

$$g_\lambda(\rho, \rho') \equiv (\rho_<^\lambda/(1-a\rho_<)^\lambda)((1-a\rho_>)^{\lambda+1}/\rho_>^{\lambda+1})(1-a\rho')^{-4},$$

and

$$\rho_< = \min(\rho, \rho') \qquad \text{and} \qquad \rho_> = \max(\rho, \rho').$$

For given values of the coefficients $a_I$, Eqs. (6a) are to be solved subject to the following conditions.

Boundary conditions:

$$P_\alpha^I(0) = P_\alpha^I(1/a) = 0, \qquad \alpha = 1,\ldots, N, \quad I = 1,\ldots, L, \quad \text{(6b)}$$

normalization conditions:

$$\int_0^{1/a} P_\alpha^I(\rho)\, P_\alpha^J(\rho)(1-a\rho)^{-4}\, d\rho - 1 = 0,$$
$$\alpha = 1,\ldots, N, \quad I = 1,\ldots, L, \quad J = 1,\ldots, L; \quad \text{(6c)}$$

orthogonality conditions:

$$\int_0^{1/a} P_\alpha^I(\rho)\, P_\beta^J(\rho)(1-a\rho)^{-4}\, d\rho = 0 \qquad \text{for all values of } \alpha, I, \beta, \quad \text{(6d)}$$

and $J$ such that $\beta < \alpha$ and $\Delta(\alpha, I; \beta, J) = 1$.

The variational equations for the mixing coefficients can be put in the form of a matrix eigenvalue equation. Since the $a_I$'s generally are not all independent it is convenient to define a new set of independent coefficients. We reexpand the total wavefunction (1) in an equivalent set of vector coupled linear combinations of determinants

$$\Psi = \sum_{K=1}^{L'} a_K' \Phi_K' \qquad (L' \leqslant L), \tag{7}$$

where $\Phi_K' = \sum_J C_J^K \Phi_J$. The $C_J^K$ are vector coupling coefficients and the sum on $J$ is over a particular subset of the $L$ determinants, depending upon the specific case. The coefficients $a_I$ are obtained from the independent coefficients $a_K'$ by simply multiplying by the appropriate vector coupling coefficient.

The variational equations for the $a_K$'s are

$$\sum_K (H_{JK} - \Lambda G_{JK})\, a_K' = 0, \qquad J = 1,..., M', \tag{8a}$$

where

$$H_{JK} = \langle \Phi_J' \mid H \mid \Phi_K' \rangle,$$
$$G_{JK} = \langle \Phi_J' \mid \Phi_K' \rangle,$$

and $\Lambda$ is the total energy and the Lagrange multiplier associated with the normalization constraint

$$\sum_{J,K} a_J' G_{JK} a_K' - 1 = 0. \tag{8b}$$

The orbitals, determined by solving (6), depend upon the values of the expansion coefficients, while the latter, determined by a solution of (8), depend upon the orbitals through the matrix elements $H_{JK}$ and $G_{JK}$. These two sets of equations must be solved in a self-consistent manner.

## 3. Finite Difference Form

We now algebraicize the system of equations (6) by introducing a finite-difference approximation. The $\rho$ axis between $\rho = 0$ and $\rho = 1/a$ is divided by a mesh of $M - 1$ evenly spaced internal points, so that the interval between points is $h = 1/Ma$. We adopt the notation

$$\rho_k = kh, \qquad k = 0,..., M, \tag{9}$$

$$P_\alpha^{Ik} = P_\alpha^I(\rho_k), \qquad \alpha = 1,..., N; \quad I = 1,..., L; \quad k = 0,..., M. \tag{10}$$

As in [2], a first-order approximation for the second derivative (central difference quotient) is used. Explicitly at mesh point $k$,

$$(d^2P_\alpha^I/d\rho^2)|_{\rho_k} \approx (1/h^2)(P_\alpha^{I(k-1)} - 2P_\alpha^{Ik} + P_\alpha^{I(k+1)}). \tag{11}$$

The trapezoidal rule is used to approximate the integrals, which is entirely consistent with the first-order approximation of the second derivative.

With these approximations the finite-difference MCHF equations are

$$a_I \sum_J a_J[(D_\alpha^{Jk} + Y_\alpha^{IJk}(1 - akh)^{-4})\, P_\alpha^{Jk} - (1/2h^2)(P_\alpha^{J(k-1)} + P_\alpha^{J(k+1)}) - X_\alpha^{IJk}(1 - akh)^{-4}] - \sum_{J,\beta} \lambda_{\alpha\beta}^{IJ}\, \Delta(\alpha, I; \beta, J)(1 - akh)^{-4}\, P_\beta^{Jk} = 0,$$

$$\alpha = 1,..., N; \quad I = 1,..., L; \quad k = 1,..., M - 1, \tag{12a}$$

where

$$D_\alpha^{Jk} \equiv \left[\frac{1}{h^2} - \frac{Z}{kh(1 - akh)^3} + \frac{l_\alpha^J(l_\alpha^J + 1)}{2k^2h^2(1 - akh)^2}\right] \Delta(\alpha, I; \alpha, J),$$

$$Y_\alpha^{IJk} \equiv \delta(m_{S_\alpha}^I m_{S_\alpha}^J) \sum_\beta \delta(m_{\zeta_\beta}^I m_{\zeta_\beta}^J) \sum_{\lambda,\mu} (4\pi/(2\lambda + 1))\, B_{\alpha\alpha}^{IJ} B_{\beta\beta}^{JI*} \sum_{j=1}^{L-1} P_\beta^{Ij} P_\beta^{Jj} g_\lambda^{jk},$$

and

$$X_\alpha^{IJk} \equiv \sum_\beta \delta(m_{S_\alpha}^I m_{S\beta}^J)\, \delta(m_{S_\alpha}^J m_{S\beta}^I) \sum_{\lambda,\mu} (4\pi/(2\lambda + 1))\, B_{\alpha\beta}^{IJ} B_{\alpha\beta}^{JI*} P_\beta^{Jk} \sum_{j=1}^{L-1} P_\beta^{Ij} P_\alpha^{Jj} g_\lambda^{jk},$$

where

$$g_\lambda^{jk} = (p^\lambda/(1 - aph)^\lambda)((1 - aqh)^{\lambda+1}/q^{\lambda+1})(1 - ajh)^{-4},$$

and $p \equiv \min(j, k)$, $q \equiv \max(j, k)$. In terms of finite-difference variables, conditions (6b)–(d) become

boundary conditions:

$$P_\alpha^{I0} = P_\alpha^{IM} = 0, \qquad \alpha = 1,..., N; \quad I = 1,..., L; \tag{12b}$$

normalization conditions:

$$h \sum_{j=1}^{M-1} P_\alpha^{Ij} P_\alpha^{Jj}(1 - ajh)^{-4} - 1 = 0, \qquad \alpha = 1,..., N; \quad I = 1,..., L; \quad J = 1,..., L; \tag{12c}$$

orthogonality conditions:

$$h \sum_{j=1}^{M-1} P_\alpha^{Ij} P_\beta^{Jj}(1 - ajh)^{-4} = 0 \qquad \text{for all values of } \alpha, I, \beta, \tag{12d}$$

and $J$ such that $\beta < \alpha$ and $\Delta(\alpha, I; \beta, J) = 1$.

For given values of the coefficients $a_I$, Eqs. (12) are a set of nonlinear algebraic equations in the unknowns $P_\alpha^{Ik}$ and $\lambda_{\alpha\beta}^{IJ}$, one equation corresponding to each unknown. In the next section we give a practical method for solving this system of algebraic equations.

## 4. Method of Solution

The sets of Eqs. (8) and (12) are solved alternately in a self-consistent manner. We first consider the solution of the finite-difference MCHF equations (Eqs. (12)) for given values of the coefficients $a_I$. The unknowns of (12) are the $NL(M+1)$ values $P_\alpha^{Ik}$ and the Lagrange multipliers $\lambda_{\alpha\beta}^{IJ}$ for $\Delta(\alpha, I; \beta, J) = 1$. As a trivial first step one can substitute the boundary conditions (12b) into (12a) for $\alpha = 1,..., N$; $I = 1,..., L$, and $k = 0$ and $k = M - 1$. In so doing we have reduced the order of our system of equations and unknowns by $2LN$, and the boundary conditions (12b) will be automatically satisfied.

As emphasized in [2] it is important to treat all of the unknowns, the multipliers included, on an equal footing. With this point in mind, and also, as we shall see, in order to make the calculation tractable, we define a vector $\mathbf{U}_\alpha$ for each value of $\alpha$ with components,

$$U_\alpha{}^K \equiv P_\alpha^{Ik}, \qquad \alpha = 1,..., N; \quad I = 1,..., L; \quad k = 1,..., M-1, \tag{13a}$$

where

$$\begin{aligned} K &\equiv (k-1)L + I, \\ U_\alpha^{K+(L-1)M} &= \lambda_{\alpha\alpha}^{IJ}, \qquad \alpha = 1,..., N; \quad I, J = 1,..., L, \end{aligned} \tag{13b}$$

where

$$\begin{aligned} K &\equiv (J-1)L + I, \\ U_\alpha^{K+(L-1)M+M^2} &= \lambda_{\alpha\beta}^{IJ}, \qquad \alpha, \beta = 1,..., N, (\beta < \alpha); \quad I, J = 1,..., L, \end{aligned} \tag{13c}$$

where $K$ is simply increased by one for each required Lagrange multiplier $\lambda_{\alpha\beta}^{IJ}$ corresponding to an orthogonality condition (12d) for which $\Delta(\alpha, I; \beta, J) = 1$, the order being unimportant. It is important to note that the vector $\mathbf{U}_\alpha$ contains consecutively $P_\alpha^{11}, P_\alpha^{21},..., P_\alpha^{L1}; P_\alpha^{12}, P_\alpha^{22},..., P_\alpha^{L2};... ; P_\alpha^{1(M-1)}, P_\alpha^{2(M-1)},..., P_\alpha^{L(M-1)}$, i.e., the $P_\alpha$ from each determinant at the first mesh point, followed by the same quantities at the second mesh point, etc. These are followed in the vector by the multipliers which are treated the same as the other variables with no special considerations.

We similarly define a function vector $\mathbf{F}_\alpha(\mathbf{U}_1,..., \mathbf{U}_N)$ with components $f_\alpha{}^K(\mathbf{U}_1,..., \mathbf{U}_N)$. In terms of these definitions the finite-difference MCHF Eqs. (12) can be written as

$$f_\alpha{}^K(\mathbf{U}_1,..., \mathbf{U}_N) = 0, \qquad \alpha = 1,..., N; \quad K = 1,..., L(L+M-1),... \tag{14}$$

where the limit on $K$ depends upon the number of orthogonality conditions in (12d). For a given $\alpha$, (14) corresponds to (12a) with $K = (k-1)L + I$, to (12c) with $K = (J + M - 2)L + I$, and finally to (12d) for $K > (M + L - 1)L$.

We now wish to determine values for all the components of the **U** vectors such that the set of nonlinear algebraic equations (14) are satisfied within a given tolerance. As in [2] we introduce two complementary iterative procedures: (1) the well-known SCF procedure and (2) a generalized Newton–Raphson iteration (GNRI).

Within an SCF iteration we solve in sequence the subsets of Eqs. (14) for $\alpha = 1,..., N$, by means of the GNRI to be described. In the usual SCF manner for a given value of $\alpha$ in the sequence, the unknowns $\mathbf{U}_\beta (\beta \neq \alpha)$ in the functions $f_\alpha^K$ are treated as constants and are fixed at their respective values as determined in the preceding SCF iteration or in the solution of the subset $\beta$ in the present SCF iteration, whichever has occurred later. The unknowns $P_\alpha^{Ik}$ occurring in the exchange integrals are given similar treatment.

Within an SCF iteration the GNRI is used to obtain a solution of the subsets of Eqs. (14) corresponding to each value of $\alpha$. If we denote the solution and function vectors evaluated at the $n$th GNRI by $\mathbf{U}_\alpha^{(n)}$ and $\mathbf{F}_\alpha^{(n)}$, respectively, then at the $(n+1)$th iteration the solution vector is given by

$$\mathbf{U}_\alpha^{(n+1)} = \mathbf{U}_\alpha^{(n)} - (J^{(n)})^{-1} \mathbf{F}_\alpha^{(n)}, \tag{15}$$

where the matrix elements of the Jacobian matrix $J^{(n)}$ are given by

$$J_{KK'}^{(n)} = [\partial f_\alpha^K / \partial U_\alpha^{K'}]^{(n)}. \tag{16}$$

The iteration (12) is repeated until

$$\max(|f_\alpha^K(\mathbf{U}_\alpha^{(n)})|) < \text{tolerance},$$

since when $\mathbf{F}_\alpha^{(n)} = 0$ the problem is solved exactly for the given subset $\alpha$.

The particular ordering of the elements in the $\mathbf{U}_\alpha$ vector given by (13) and the elements in the function vector $\mathbf{F}_\alpha$ given by (14) lead to Jacobian matrices (16) which are in a special nearly block tridiagonal form. Consequently Eq. (15) may be rapidly solved by means of a partitioning (see Appendix).

Within a general SCF iteration the starting vectors $\mathbf{U}_\alpha^{(0)}$ contain the respective values of the variables $U_\alpha^K$ as determined in the previous SCF iteration. On the first SCF iteration after obtaining new expansion coefficients from a solution of (8), the starting vectors $\mathbf{U}_\alpha^{(0)}$ should be set respectively equal to final $\mathbf{U}_\alpha$ vectors corresponding to the old expansion coefficients.

Turning our attention briefly to the solution of (8) for new values of the expansion coefficients corresponding to a given set of determinantal functions $\Phi_J$, the first

step is the evaluation of the matrix elements $H_{JK}$ and $G_{JK}$ in (8) by means of numerical quadrature. In accordance with our evaluation of Coulomb and exchange integrals, we also use trapezoidal rule here. Having evaluated the matrix elements, the solution of (8) is simply a symmetric matrix eigenvalue problem for which standard algorithms are available.

## 5. The $1s2s^1S$ State of Helium

As mentioned previously we have chosen the $1s2s^1S$ state of helium as a test case because the strong orthogonality condition between the $1s$ and $2s$ orbitals is reported to cause convergence difficulties. Also it is the most simple case for which more than one derivative term appears in (6a).

We consider the mixing of the three configurations $1s2s$, $1s^2$, and $2s^2$ to compose the $1s2s^1S$ state. We have restricted the number of configurations to these three in the interest of simplicity, since we are only testing the finite-difference algorithm and not doing a definitive calculation. Furthermore, we restrict the $1s$ functions in $1s^2$ to be the same as in $1s2s$ and similarly for the $2s$ functions in $2s^2$. Therefore we only have two radial functions instead of $N \times L = 2 \times 3 = 6$ which we would have without these restrictions.

Let the coefficients corresponding to the configurations $1s2s$, $1s^2$, and $2s^2$ be denoted by $a_1$, $a_2$, and $a_3$, respectively. We write the finite-difference equations in terms of the components of a U vector defined as follows. Denoting the $1s$ and $2s$ functions by subscripts 1 and 2,

$$\begin{aligned} U^{2k-1} &\equiv P_1(\rho_k) \equiv P_1{}^k, \\ U^{2k} &\equiv P_2{}^k, \qquad k = 1,\ldots, M-1, \\ U^{2M-1} &\equiv \lambda_{11}, \\ U^{2M} &\equiv \lambda_{22}, \\ U^{2M+1} &\equiv \lambda_{12}. \end{aligned} \tag{17}$$

In terms of these variables the system of finite-difference equations may be written as

$$\begin{aligned} f^{2k-1} = {} & [(a_1{}^2 + 2a_2{}^2)((1/h^2) - (2/kh(1 - akh)^3)) + 2a_2{}^2 Y_{11}^k \\ & + a_1{}^2 Y_{22}^k + 2^{3/2} a_1 a_2 Y_{12}^k - (U^{2M-1}/(1 - akh)^4)]\, U^{2k-1} \\ & - (1/2h^2)(a_1{}^2 + 2a_2{}^2)(U^{2k-3} + U^{2k+1}) \\ & + [2^{1/2} a_1 (a_2 + a_3)((1/h^2) - (2/kh(1 - akh)^3)) + 2^{1/2} a_1 a_3 Y_{22}^k \\ & + 2^{1/2} a_1 a_2 Y_{11}^k + (a_1{}^2 + 2a_2 a_3)\, Y_{12}^k - (U^{2M+1}/(1 - akh)^4)]\, U^{2k} \\ & - 2^{-1/2}(a_1/h_2)(a_2 + a_3)(U^{2k-2} + U^{2k+2}) = 0, \qquad k = 1,\ldots, M-1; \end{aligned} \tag{18a}$$

$$f^{2k} = [(a_1^2 + 2a_3^2)((1/h^2) - (2/kh(1 - akh)^3)) + 2a_3^2 Y_{22}^k + a_1^2 Y_{11}^k + 2^{3/2} a_1 a_3 Y_{12}^k - (U^{2M}/(1 - akh)^4)]\, U^{2k} - (1/2h^2)(a_1^2 + 2a_3^2)(U^{2k-2} + U^{2k+2}) + [2^{1/2} a_1 (a_2 + a_3)((1/h^2) - (2/kh(1 - akh)^3)) + 2^{1/2} a_1 a_2 Y_{11}^k + 2^{1/2} a_1 a_3 Y_{22}^k + (a_1^2 + 2a_2 a_3)\, Y_{12}^k - (U^{2M+1}/(1 - akh)^4)]\, U^{2k-1} - 2^{-1/2}(a_1/h^2)(a_2 + a_3)(U^{2k-3} + U^{2k+1}) = 0, \qquad k = 1,\ldots, M - 1; \quad (18b)$$

$$f^{2M-1} = h \sum_{j=1}^{M-1} (1 - ajh)^{-4}\, U^{2j-1} U^{2j-1} - 1 = 0; \quad (18c)$$

$$f^{2M} = h \sum_{j=1}^{M-1} (1 - ajh)^{-4}\, U^{2j} U^{2j} - 1 = 0; \quad (18d)$$

$$f^{2M+1} = h \sum_{j=1}^{M-1} (1 - ajh)^{-4}\, U^{2j-1} U^{2j} = 0; \quad (18e)$$

where

$$Y_{\alpha\beta}^k = (1 - akh)^{-4} \sum_{j=1}^{M-1} (1 - ajh)^{-4} ((1 - aqh)/q)\, U^{2j+\alpha-2} U^{2j+\beta-2}$$

and $q = \max(j, k)$.

Since we only have two orbitals we use the GNRI to determine them simultaneously which gives us a Jacobian matrix containing $2 \times 2$ subblocks (see Appendix). We commence the calculation with only the principal $1s2s$ configuration included (i.e., $a_1 = 1$, $a_2 = a_3 = 0$). As in [2] we start with hydrogenic ($Z = 2$) $1s$ and $2s$ functions in finite-difference form which approximately satisfy (18) with all interactions set to zero. A tracking procedure is employed where the interaction terms in (18) are multiplied by a parameter $\epsilon$, and $\epsilon$ is varied from 0 to 1 in steps of 0.1, the converged U vector for one step serving as a starting approximation for the next.

The calculation was carried out with 100 mesh points. Weak Newton–Raphson and SCF tolerances of $10^{-4}$ and $10^{-2}$ respectively were set during tracking. The final single configuration $1s2s$ solution at $\epsilon = 1.0$ was obtained with tolerances of $10^{-7}$ and $10^{-5}$. The total energy of $-2.170141$ a.u., the values of the multipliers ($\lambda_{11} = -1.71655$, $\lambda_{22} = -0.186837$, and $\lambda_{12} = 0.150860$), and the $1s$ and $2s$ functions themselves all agree to four figures with the previously published results of Cohen and Kelly [9] for the $1s2s$ single-configuration. Complete agreement is not expected since the finite-difference results have not been extrapolated to the $h \to 0$ limit [2]. These authors used the computer program of Froese-Fischer. The calculated energy is actually below the experimental energy for the $1s2s\,^1S$

TABLE I

History of Coefficient Tracking Procedure[a]

| N | Coefficient values in (18) | | | Energy Expectation | Values from diagonalization | | | Energy from diagonalization |
|---|---|---|---|---|---|---|---|---|
| | $a_1$ | $a_2$ | $a_3$ | | $a_1$ | $a_2$ | $a_3$ | |
| | 1.000000 | 0.000000 | 0.000000 | −2.170141 | 0.922940 | 0.356911 | −0.144211 | −2.119310 |
| 5 | 0.996879 | 0.073198 | −0.029576 | −2.153041 | 0.940180 | 0.316767 | −0.125382 | −2.128706 |
| 5 | 0.991202 | 0.122904 | −0.049132 | −2.149069 | 0.965069 | 0.242597 | −0.098935 | −2.143979 |
| 5 | 0.987352 | 0.147083 | −0.059189 | −2.145242 | 0.962964 | 0.250874 | −0.098808 | −2.142309 |
| 3 | 0.981344 | 0.178084 | −0.072460 | −2.142111 | 0.970888 | 0.222818 | −0.087916 | −2.141206 |
| 3 | 0.978113 | 0.193047 | −0.077633 | −2.140714 | 0.980584 | 0.181416 | −0.074458 | −2.140644 |
| 3 | 0.978953 | 0.189173 | −0.076576 | −2.143760 | 0.976793 | 0.198129 | −0.081365 | −2.143747 |
| 2[b] | 0.977886 | 0.193654 | −0.078971 | −2.143758 | 0.978122 | 0.192624 | −0.078565 | −2.143757 |
| 2 | 0.978005 | 0.193139 | −0.078768 | −2.143757 | 0.977994 | 0.193185 | −0.078785 | −2.143757 |
| 2 | 0.977999 | 0.193162 | −0.078777 | −2.143758 | 0.978001 | 0.193155 | −0.078774 | −2.143758 |
| 2 | 0.978000 | 0.193158 | −0.078775 | −2.143758 | 0.978001 | 0.193154 | −0.078773 | −2.143758 |

[a] Energy values in atomic units.

[b] SCF and GNRI tolerances decreased to $10^{-5}$ and $10^{-7}$, respectively.

state of helium (−2.146915 a.u., [10]). This result does not violate the variational principle because we are dealing with an excited state.

We next consider the addition of the $1s^2$ and $2s^2$ configurations. After calculating the six independent matrix elements of the Hamiltonian using the $1s$ and $2s$ functions just obtained, we diagonalize the symmetric matrix, using Jacobi's procedure, and obtain new values of the coefficients ($a_1 = 0.922940$, $a_2 = 0.356911$, and $a_3 = -0.144211$). We also obtain three energy eigenvalues (−2.772667, −2.119310, and −0.670193 a.u.) which, by the Hylleraas Undheim theorem, are three upper bounds on the energy levels of the three lowest-lying $^1S$ states.

The next problem is to make the coefficients self-consistent. During this exercise we set our tolerances once again to the tracking values. We put the new values of the coefficients into the finite-difference equations (Eqs. (18)), with the interactions at full strength, and tried to solve the system using the single configuration $1s$ and $2s$ functions for the starting **U** vector. But the GNRI failed to converge because the new values of the coefficients are too far removed from the single-configuration values of 1, 0, and 0.

To overcome this problem the great circle on the unit sphere in coefficient space between the old and the new coefficient values was divided into $N$ equal segments. The values at the end of the first segment were put back into (18) instead of those predicted by the diagonalization. Thus the values of the coefficients were changed in the proper direction in coefficient space while always maintaining overall normalization. After every solution of (18) obtaining new orbitals, the Hamiltonian matrix was again diagonalized, providing a new segmented great circle and new values of the coefficients. The value of $N$ was initially taken to be 5 and was reduced to 2 as the values became more self-consistent. A history of this procedure is given in Table I. In retrospect it seems that the convergent process could have been accelerated by decreasing $N$ faster and also tightening the tolerances sooner.

The final $1s$ and $2s$ functions and the values of the multipliers are given in Table II where every fifth value is tabulated. The small negative tail on the $1s$ function is a characteristic of the $1s2s\,^1S$ state. It also appears in the single configuration function. It will also be noticed that the off-diagonal Lagrange multiplier $\lambda_{12}$, which was of comparable magnitude to $\lambda_{22}$ in the single configuration case, has now been considerably reduced, meaning that the orthogonality constraint is weaker for the multiconfiguration case. This fact is reflected in an increased rate of GNRI convergence.

Finally one could repeat the calculation with different numbers of mesh points, and use Richardson extrapolation [11] to determine values of energy and coefficients corresponding to the differential system. Since we are only testing the method we have elected not to do this.

It has been pointed out that the MCHF equations which we have just solved using finite differences do not have a unique solution [12]. One can perform an

TABLE II

Final Converged U Vector[a,b]

| 1s | | 2s | |
|---|---|---|---|
| $K$ | $U^K$ | $K$ | $U^K$ |
| 1 | 0.553571$D$ − 01 | 2 | 0.238985$D$ − 03 |
| 11 | 0.298310$D$ + 00 | 12 | 0.133867$D$ − 02 |
| 21 | 0.485323$D$ + 00 | 22 | 0.241384$D$ − 02 |
| 31 | 0.617635$D$ + 00 | 32 | 0.362489$D$ − 02 |
| 41 | 0.697418$D$ + 00 | 42 | 0.505870$D$ − 02 |
| 51 | 0.727978$D$ + 00 | 52 | 0.666419$D$ − 02 |
| 61 | 0.713964$D$ + 00 | 62 | 0.817982$D$ − 02 |
| 71 | 0.661562$D$ + 00 | 72 | 0.906313$D$ − 02 |
| 81 | 0.578634$D$ + 00 | 82 | 0.844262$D$ − 02 |
| 91 | 0.474738$D$ + 00 | 92 | 0.512793$D$ − 02 |
| 101 | 0.360898$D$ + 00 | 102 | −0.226348$D$ − 02 |
| 111 | 0.248982$D$ + 00 | 112 | −0.149784$D$ − 01 |
| 121 | 0.150500$D$ + 00 | 122 | −0.334954$D$ − 01 |
| 131 | 0.746810$D$ − 01 | 132 | −0.565105$D$ − 01 |
| 141 | 0.260531$D$ − 01 | 142 | −0.796142$D$ − 01 |
| 151 | 0.250450$D$ − 02 | 152 | −0.942121$D$ − 01 |
| 161 | −0.407637$D$ − 02 | 162 | −0.884892$D$ − 01 |
| 171 | −0.310922$D$ − 02 | 172 | −0.550961$D$ − 01 |
| 181 | −0.684900$D$ − 03 | 182 | −0.120368$D$ − 01 |
| 191 | −0.115753$D$ − 05 | 192 | −0.203427$D$ − 04 |
| 197 | −0.931191$D$ − 13 | 198 | −0.163649$D$ − 11 |
| 199 | −0.174990$D$ + 01 | 200 | −0.144334$D$ + 00 |
| 201 | 0.684257$D$ − 01 | | |

[a] See Eq. (17).

[b] Number of mesh points $M = 100$; $a = 1.0$ (see Eq. (5)).

orthogonal transformation mixing the 1$s$ and 2$s$ radial functions (without altering the total energy), and obtain a new solution of the equations with different values of the coefficients. The same is true for the finite-difference system. In such circumstances one might be concerned about getting an instability in the sequential solutions for the radial functions and the coefficients. It is satisfying to note from Table I that no such instability occurred, as the convergence is nearly monotonic. Nonuniqueness in the form of the solution is not a problem with the FDNRA.

## 6. Conclusions

The finite-difference Newton–Raphson algorithm has been extended to treat the multiconfiguration atomic Hartree–Fock equations in their full generality. Although the test case was a relatively simple one from the point of view of number of configurations and orbitals, nevertheless it is a case which has caused difficulties in the past. We do not see any fundamental problems which would preclude the treatment of more complicated cases with the present algorithm.

One final note: It might seem to be advantageous to include the expansion coefficients in the U vector and use the GNRI to solve for them along with the orbitals and multipliers. This procedure has been tried and it ruins the quadratic convergence of the GNRI. It still converges, but only at a rate comparable to the SCF convergence. The reason is that one cannot construct a Jacobian matrix which includes the partial derivatives of the variational equations with respect to the coefficients and which is also compatible with the absence of partial derivatives with respect to variables in the Coulomb and exchange integrals. One could only do this in a consistent manner if one abandoned the SCF iteration completely, but then the Jacobian matrix would be completely full and the procedure becomes impractical.

## Appendix

Due to the particular ordering of the $\mathbf{U}_\alpha$ and $\mathbf{F}_\alpha$ vectors defined in Section 4 the Jacobian matrices $J^{(n)}$ encountered in (15) are all of the partly block tridiagonal form shown in Fig. 1. The submatrix $A$ is an $L(M-1) \times L(M-1)$ block

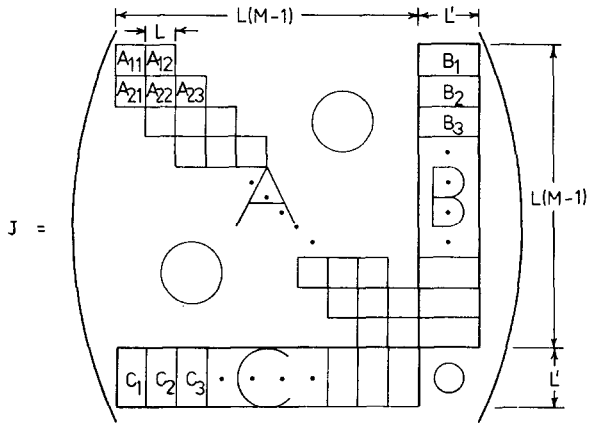


Figure 1

tridiagonal square matrix with nonzero $L \times L$ square subblocks $A_{i,i-1}$, $A_{ii}$, and $A_{i,i+1}$ ($i = 1,..., L(M-1)$). $B$ is a general $L(M-1) \times L'$ matrix and $C$ is a general $L' \times L(M-1)$ matrix, where $L'$ is the total number of normalization and orthogonality constraints associated with a given value of $\alpha$. It is useful to partition the $B$ and $C$ matrices into general $L \times L'$ and $L' \times L$ submatrices, respectively, and label them $B_i$ and $C_i$ ($i = 1,..., M-1$).

The following algorithm for the solution of (12) is a generalization of the one given in [2] for the special case $L = 1$. First, let us rewrite (12) as

$$J^{(n)}\Delta_\alpha = \mathbf{F}_\alpha^{(n)}, \tag{A1}$$

where

$$\Delta_\alpha \equiv \mathbf{U}_\alpha^{(n)} - \mathbf{U}^{(n+1)}. \tag{A2}$$

Once (A1) is solved for $\Delta_\alpha$, the vector $\mathbf{U}_\alpha^{(n+1)}$ is directly obtained since $\mathbf{U}_\alpha^{(n)}$ is already known. Furthermore, we define

$$\Delta_\alpha = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}_\alpha \tag{A3}$$

and

$$\mathbf{F}_\alpha^{(n)} = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{pmatrix}_\alpha, \tag{A4}$$

where $\Delta_1$ and $\mathbf{F}_1$ are vectors of length $L(M-1)$ and $\Delta_2$ and $\mathbf{F}_2$ are vectors of length $L'$. Equation (A1) can be written as

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix}\begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}_\alpha = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{pmatrix}_\alpha, \tag{A5}$$

or equivalently it can be expressed as two matrix equations

$$A\Delta_1 + B\Delta_2 = \mathbf{F}_1 \tag{A6}$$

and

$$C\Delta_1 = \mathbf{F}_2. \tag{A7}$$

From (A6) and (A7) we obtain directly that

$$CA^{-1}B\Delta_2 = CA^{-1}\mathbf{F}_1 - \mathbf{F}_2 \tag{A8}$$

and

$$A\Delta_1 = \mathbf{F}_1 - BA\Delta_2. \tag{A9}$$

Equations (A8) and (A9) are solved for $\Delta_1$ and $\Delta_2$ in the following way. First, we define $B_{ij}$ ($j = 1,..., L'$) to be the column vectors of the submatrix $B_i$ ($i = 1,..., M-1$). Similarly, let $C_{ji}$ ($j = 1,..., L'$) denote the row vectors of the

submatrix $C_i$ ($i = 1,..., M-1$). We obtain the $L(M-1) \times L'$ matrix $A^{-1}B$ by defining the $L \times L$ matrices [11]

$$W_1 \equiv A_{11}^{-1}A_{12}\,; \qquad W_i \equiv (A_{ii} - A_{ii-1}W_{i-1})^{-1}A_{ii+1}\,, \qquad i = 2,..., M-2, \quad \text{(A10)}$$

and length $L$ column vectors

$$G_{1j} \equiv A_{11}^{-1}B_{1j}\,; \qquad G_{ij} \equiv (A_{ii} - A_{ii-1}W_{i-1})^{-1}(B_{ij} - A_{ii-1}G_{i-1j}),$$
$$i = 2,..., M-1; \quad j = 1,..., L', \quad \text{(A11)}$$

where Cayley matrix multiplication is understood. These quantities make possible the recursive computation of the elements of the $L(M-1) \times L'$ matrix, $A^{-1}B$.

$$(A^{-1}B)_{M-1,j} = G_{M-1,j}\,; \qquad (A^{-1}B)_{ij} = G_{ij} - W_i(A^{-1}B)_{i+1,j}\,,$$
$$i = M-2, M-3,..., 1; \quad j = 1,..., L'. \quad \text{(A12)}$$

It is important to realize, however, that the quantities $(A^{-1}B)_{ij}$ in (A12) are not the matrix elements themselves but are rather column vectors of length $L$ whose components are the desired matrix elements. $A^{-1}B$ has the same dimensions as $B$, and can be partitioned into submatrices, each of $L$ rows, $(A^{-1}B)_i$ ($i = 1,..., M-1$), in the same manner that $B$ was partitioned. The subscript $i$ in $(A^{-1}B)_{ij}$ indicates the submatrix and the $j$ indicates the column vector in that submatrix.

The $L' \times L'$ matrix $CA^{-1}B$ is now obtained.

$$(CA^{-1}B)_{jk} = \sum_{i=1}^{M-1} C_{ji}(A^{-1}B)_{ik}\,, \quad \text{(A13)}$$

where, again, Cayley multiplication between the row vector $C_{ji}$ and column vector $(A^{-1}B)_{ik}$ in each term in the summation is understood.

In a similar manner, by partitioning the vector $\mathbf{F}_1$ in (A8) into subvectors of length $L$, the $L'$-dimensional column vector $CA^{-1}\mathbf{F}_1$ can be obtained. Then by inverting the $L' \times L'$ matrix $CA^{-1}B$ and premultiplying (A8) by $(CA^{-1}B)^{-1}$, we obtain $\boldsymbol{\Delta}_2$. Finally, using the same techniques once again, we obtain $\boldsymbol{\Delta}_1$ $[= A^{-1}(\mathbf{F}_1 - B\boldsymbol{\Delta}_2)]$, and the problem is solved.

Note that the matrix $A$ is not actually inverted at any stage of the solution. Also the number of calculations and storage locations for this algorithm are both proportional to $L[(3L + 2L')(M-1) - 2L]$ (i.e., $M$ to the first power). Thus, it is clear how the partitions defined in Fig. 1 greatly reduce the time and storage requirements, for without them $J$ would have to be treated as a general square matrix of dimension $L(M-1) + L'$. The time required for its inversion would then be proportional to $[L(M-1) + L']^3$, and $[L(M-1) + L']^2$ storage locations would be required.

## REFERENCES

1. C. P. VAN DINE, W. R. FIMPLE, AND T. N. EDELBAUM, Application of a Finite-Difference Newton–Raphson Algorithm to the Problem of Low-Thrust Trajectory Optimization, *in* "Methods in Astrodynamics and Celestial Mechanics," Progress in Astronautics and Aeronautics, Vol. 17 (R. L. Duncombe and V. G. Szebehely, Eds.), Academic Press, New York/London, 1966.
2. J. K. CAYFORD, W. R. FIMPLE, AND D. G. UNGER, *J. Computational Phys.* **15** (1974), 81.
3. J. K. CAYFORD, W. R. FIMPLE, D. G. UNGER, AND S. P. WHITE, *J. Computational Phys.* **16** (1974), 259.
4. W. R. FIMPLE AND S. P. WHITE, *Int. J. Quant. Chem.* **9** (1975), 301.
5. C. FROESE-FISCHER, *Computer Phys. Commun.* **4** (1972), 107.
6. C. FROESE-FISCHER, *Computer Phys. Commun.* **1** (1969), 151.
7. C. FROESE-FISCHER, *Computer Phys. Commun.* **2** (1971), 124.
8. C. FROESE-FISCHER, private communication.
9. M. COHEN AND P. S. KELLY, *Canad. J. Phys.* **43** (1965), 1867.
10. C. E. MOORE, "Atomic Energy Levels," Vol. I, U.S. Dept. of Commerce, *Nat. Bur. Stand. U.S. Circ.* 467, U.S. Government Printing Office, Washington, D.C.
11. L. F. RICHARDSON, *Philos. Trans. Roy. Soc. London Ser. A* **226** (1927), 299; Z. KOPAL, "Numerical Analysis," Section V–C, Wiley, New York, 1961.
12. C. FROESE-FISCHER, *Canad. J. Phys.* **51** (1973), 1238.