JOURNAL OF COMPUTATIONAL PHYSICS 15, 81–97 (1974)

A Finite Difference Newton–Raphson Solution of the Atomic Hartree–Fock Problem

J. K. CAYFORD, W. R. FIMPLE, AND D. G. UNGER

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

Received May 30, 1973

The *SCF* iteration is coupled with a finite difference Newton–Raphson algorithm to solve the set of coupled second-order integrodifferential equations with split boundary conditions which constitutes the atomic HF problem. In the new method the two-point boundary conditions at $r = 0$ and $r = \infty$ as well as the Lagrange multipliers are incorporated into a large system of nonlinear algebraic equations which are solved by means of a generalized Newton–Raphson iteration which converges rapidly and efficiently. The need to estimate initial slopes of the radial functions and values of Lagrange multipliers has been completely eliminated. As an example a calculation of the $1S^2 2S$ open-shell configuration of Li is presented. Through the use of Richardson extrapolation an accuracy of nine significant figures has been achieved. The new method is easier to apply and more versatile than the conventional methods. Although only Li and Be have been attempted so far (each with complete success) the method can certainly handle very large systems.

1. INTRODUCTION

Although there have been many advances in numerical analysis in recent years, nearly all commonly used computer programs for solving the atomic Hartree–Fock (HF) equations [1–4] still basically employ the methods originally developed by Hartree [5]. While the HF radial integrodifferential equation with associated boundary conditions at $r = 0$ and ∞ is essentially a two-point boundary value problem, the conventional methods treat it as a sequence of initial value problems.

One of the most powerful numerical methods for solving two-point boundary-value problems is the finite difference Newton–Raphson algorithm, originally developed by Van Dine [6]. In the present work this algorithm is employed in conjunction with the well known self-consistent field (SCF) iteration to solve the atomic HF equations. The new method involves a number of distinct parts: (1) the HF integrodifferential equations with split boundary conditions are approximated by systems of finite-difference equations, (2) within a SCF iteration the system of finite difference equations associated with each radial function is solved separately

and sequentially by means of a generalized Newton–Raphson iteration (GNRI) [6], and (3) the usual SCF procedure is employed. The new method is most attractive in that it incorporates the split boundary conditions into the system of finite difference equations and treats the Lagrange multipliers as an integral part of the system rather than as adjuncts to the problem as in conventional methods.

The authors are aware of only one previous application of finite difference techniques to the calculation of atomic wavefunctions. This is the numerical solution of the two-dimensional S-limit Schrödinger equation for He by Winter, Diestler, and McKoy [7]. In this calculation very large matrices were diagonalized by conventional matrix eigenvalue techniques. From a practical point of view it is difficult to see how this approach could be extended to larger systems.

In the following section the HF problem is outlined, and in Section 3 the finite difference form of the atomic HF equations is derived. The basic algorithm to solve these equations is outlined in Section 4 and further illustrated in Section 5 by means of a specific example. Finally, in Section 6, some results are presented for the $1s^22s$ configuration of Li.

2. The Atomic Hartree–Fock Problem

If the total wavefunction of an N-electron atom is approximated by the antisymmetric product of N single-electron spin-orbitals

$$\Psi = \mathscr{A}(\phi_1(1) \cdots \phi_N(N)), \tag{1}$$

where

$$\phi_\alpha = R_{n_\alpha l_\alpha}(r)\, Y_{l_\alpha}^{m_\alpha}(\omega)\, \chi_{1/2}^{m_{s_\alpha}}, \qquad \alpha = 1,\ldots, N, \tag{2}$$

and further, if we define new radial functions,

$$P_\alpha(r) = rR_{n_\alpha l_\alpha}(r), \tag{3}$$

then the radial HF equations are a set of coupled integrodifferential equations of the following form.

$$\left[-\frac{1}{2}\frac{d^2}{dr^2} - \frac{Z}{r} + Y_\alpha(r) + \frac{l_\alpha(l_\alpha+1)}{2r^2} - \lambda_{\alpha\alpha}\right] P_\alpha(r)$$
$$= X_\alpha(r) + \sum_{\beta\neq\alpha} \lambda_{\alpha\beta}\, \Delta(\alpha,\beta)\, P_\beta(r), \qquad \alpha = 1,\ldots, N, \tag{4a}$$

where $\Delta(\alpha,\beta) \equiv \delta(l_\alpha l_\beta)\,\delta(m_\alpha m_\beta)\,\delta(m_{s_\alpha} m_{s_\beta})$, and for a single configuration,

$$Y_\alpha(r) \equiv \sum_\beta \sum_{\lambda=0}^{\infty} \sum_{\mu=-\lambda}^{+\lambda} \frac{4\pi}{2\lambda+1} B_{\alpha\alpha} B_{\beta\beta}^{*} \int_0^\infty P_\beta(r')^2 \frac{r_<^\lambda}{r_>^{\lambda+1}}\, dr'$$

and

$$X_\alpha(r) \equiv \sum_\beta \delta(m_{s_\alpha} m_{s_\beta}) \sum_{\lambda=0}^{\infty} \sum_{\mu=-\lambda}^{+\lambda} \frac{4\pi}{2\lambda+1} \mid B_{\alpha\beta} \mid^2 P_\beta(r) \int_0^\infty P_\alpha(r') P_\beta(r') \frac{r_<^\lambda}{r_>^{\lambda+1}} dr',$$

where

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

and $r_< = \min(r, r')$ and $r_> = \max(r, r')$. Because the radial functions are real, $\lambda_{\beta\alpha} = \lambda_{\alpha\beta}$. Equations (4) are to be solved subject to the following conditions:

boundary conditions:

$$P_\alpha(0) = P_\alpha(\infty) = 0, \qquad \alpha = 1,..., N, \tag{4b}$$

normalization conditions:

$$\int_0^\infty P_\alpha(r)^2 \, dr = 1, \qquad \alpha = 1,..., N, \tag{4c}$$

orthogonality conditions between radial functions associated with spin-orbitals with the same angular and spin quantum numbers:

$$\int_0^\infty P_\alpha(r) P_\beta(r) \, dr = 0 \quad \text{for all values of } \alpha \text{ and } \beta \text{ such that } \beta < \alpha \text{ and}$$
$$\Delta(\alpha, \beta) = 1. \tag{4d}$$

In the standard methods of solving the system of equations (4) the SCF procedure is employed, where each of the radial equations is numerically integrated in sequence ($\alpha = 1,..., N$), with the Coulomb and exchange integrals Y_α and X_α determined from previous SCF iterations. Each integration is performed as two initial-value problems. An outward integration is started at a point near the nucleus, while an inward integration is started at a large arbitrary radial value where it is matched to an exponential "tail." The inward and outward integrations are required to match in value and slope at some arbitrary intermediate point. This procedure requires estimates of the slope at the two starting points, quantities not specified in the boundary conditions, as well as estimates of the values of the multipliers $\lambda_{\alpha\beta}$. The standard methods for the most part differ in their specific procedures for translating a mismatch of the inward and outward integrations into better estimates of the slopes and multipliers.

3. Finite Difference Form

In the present paper it is shown how the system of equations (4) can be treated universally as true two-point boundary-value problems by means of finite-difference techniques, and the conditions (4b), (4c), and (4d) can be treated as an integral part of the system.

The first step is to translate (4) into finite difference form. But before proceeding it is very advantageous to make a change of radial variable suggested by Boys and Handy [8]

$$\rho = \frac{r}{1 + ar}, \qquad (a > 0). \tag{5}$$

The range of the new variable is finite (0 to 1/a) and can be spanned by a finite mesh of equally spaced points. Another advantage of the transformation (5) is that a constant density of mesh points in ρ corresponds to a greater density in r in the important region near the nucleus as opposed to the less important region far from the nucleus.

In transforming the HF equations the following substitution is made instead of that given by (3), as it has the advantage of still removing first derivatives when the equations are expressed in ρ

$$P_\alpha(\rho) = \rho R_{n_\alpha l_\alpha}(\rho). \tag{6}$$

Under the transformation defined by (5) and (6) the HF equations become the following:

$$\left[-\frac{1}{2}\frac{d^2}{d\rho^2} - \frac{Z}{\rho(1-a\rho)^3} + \frac{Y_\alpha(\rho)}{(1-a\rho)^4} + \frac{l_\alpha(l_\alpha+1)}{2\rho^2(1-a\rho)^2} - \frac{\lambda_{\alpha\alpha}}{(1-a\rho)^4}\right] P_\alpha(\rho)$$

$$= \frac{X_\alpha(\rho)}{(1-a\rho)^4} + \sum_{\beta\neq\alpha} \lambda_{\alpha\beta}\, \Delta(\alpha,\beta)(1-a\rho)^{-4} P_\beta(\rho), \qquad \alpha = 1,\ldots, N, \tag{7a}$$

where

$$Y_\alpha(\rho) \equiv \sum_{\beta,\lambda,\mu} \frac{4\pi}{2\lambda+1} B_{\alpha\alpha} B^*_{\beta\beta} \int_0^{1/a} P_\beta(\rho')^2\, g_\lambda(\rho,\rho')\, d\rho',$$

and

$$X_\alpha(\rho) \equiv \sum_{\beta,\lambda,\mu} \delta(m_{s_\alpha} m_{s_\beta}) \frac{4\pi}{2\lambda+1} |\, B_{\alpha\beta}\,|^2 P_\beta(\rho) \int_0^{1/a} P_\alpha(\rho')\, P_\beta(\rho')\, g_\lambda(\rho,\rho')\, d\rho',$$

where

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

and

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

The conditions (4b), (4c), and (4d) transform to

boundary conditions:

$$P_\alpha(0) = P_\alpha\left(\frac{1}{a}\right) = 0, \qquad \alpha = 1,\ldots, N, \tag{7b}$$

normalization conditions:

$$\int_0^{1/a} P_\alpha(\rho)^2(1 - a\rho)^{-4}\, d\rho = 1, \qquad \alpha = 1,\ldots, N, \tag{7c}$$

orthogonality conditions:

$$\int_0^{1/a} P_\alpha(\rho)\, P_\beta(\rho)(1 - a\rho)^{-4}\, d\rho = 0, \tag{7d}$$

for all values of α and β such that $\beta < \alpha$ and $\Delta(\alpha, \beta) = 1$.

We now proceed to translate the new system of equations (7) into finite difference form. The ρ 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 $h = 1/Ma$. The following notation will be used.

$$\rho_k = kh, \qquad k = 0,\ldots, M, \tag{8a}$$

$$P_\alpha(\rho_k) = P_\alpha^k, \qquad \alpha = 1,\ldots, N; \qquad k = 0,\ldots, M. \tag{8b}$$

The first-order approximation for the second derivative (central difference quotient) is used. Explicitly at mesh point k

$$\left.\frac{d^2P_\alpha}{d\rho^2}\right|_{\rho_k} \approx \frac{1}{h^2}(P_\alpha^{k-1} - 2P_\alpha^{k} + P_\alpha^{k+1}). \tag{9}$$

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

With these approximations the finite-difference HF equations are the following:

$$[D_\alpha^k - (1 - akh)^{-4}\lambda_{\alpha\alpha}]\, P_\alpha^k - \frac{1}{2h^2}(P_\alpha^{k-1} + P_\alpha^{k+1}) - X_\alpha^k(1 - akh)^{-4} - \sum_{\beta\neq\alpha} \lambda_{\alpha\beta}\, \Delta(\alpha, \beta)(1 - akh)^{-4}\, P_\beta^k = 0,$$

$$\alpha = 1,\ldots, N; \qquad k = 1,\ldots, M - 1, \tag{10a}$$

where

$$D_\alpha^k \equiv \frac{1}{h^2} - \frac{Z}{kh(1 - akh)^3} + \frac{l_\alpha(l_\alpha + 1)}{2k^2h^2(1 - akh)^2} + Y_\alpha^k(1 - akh)^{-4},$$

$$Y_\alpha^k \equiv \sum_{\beta,\lambda,\mu} \frac{4\pi}{2\lambda + 1} B_{\alpha\alpha}B_{\beta\beta}^* \sum_{j=1}^{M-1} (P_\beta^j)^2 g_\lambda^{j,k},$$

and

$$X_\alpha^k \equiv \sum_{\beta,\lambda,\mu} \delta(m_{s_\alpha}m_{s_\beta}) \frac{4\pi}{2\lambda + 1} |B_{\alpha\beta}|^2 P_\beta^k \sum_{j=1}^{M-1} P_\alpha^j P_\beta^j g_\lambda^{j,k},$$

where

$$g_\lambda^{j,k} \equiv \frac{p^\lambda}{(1 - aph)^\lambda} \frac{(1 - aqh)^{\lambda+1}}{q^{\lambda+1}} (1 - ajh)^{-4}$$

and $p \equiv \min(j, k)$, $q \equiv \max(j, k)$. In finite-difference form the conditions (7b–d) become

boundary conditions:

$$P_\alpha^0 = P_\alpha^M = 0, \qquad \alpha = 1,..., N, \tag{10b}$$

normalization conditions:

$$h \sum_{j=1}^{M-1} (1 - ajh)^{-4}(P_\alpha^j)^2 - 1 = 0, \qquad \alpha = 1,..., N, \tag{10c}$$

orthogonality conditions:

$$h \sum_{j=1}^{M-1} (1 - ajh)^{-4} P_\alpha^j P_\beta^j = 0 \quad \text{for all values of } \alpha \text{ and } \beta \text{ such that } \beta < \alpha \text{ and } \Delta(\alpha, \beta) = 1. \tag{10d}$$

The equations (10) are a set of algebraic equations in the unknowns P_α^k and $\lambda_{\alpha\beta}$, one equation for each unknown. It remains to describe a practical method for solving this system.

4. Method of Solution

The unknowns of the finite difference HF equations (10) are the $N(M + 1)$ values P_α^k and the Lagrange multipliers $\lambda_{\alpha\beta}$ for $\Delta(\alpha, \beta) = 1$. The boundary conditions (10b) can be substituted directly into (10a) for $\alpha = 1,..., N$ and $k = 1$ and

$M - 1$. In so doing we have reduced the order of our system of equations (and unknowns) by $2N$, and the two-point boundary conditions (10b) will be automatically satisfied. The normalization and orthogonality conditions (10c and 10d), however, still remain as part of the system.

At this point we make a radical departure from conventional methods of solution where the multipliers $\lambda_{\alpha\beta}$ are treated as parameters (eigenvalues in the diagonal representation), and reserved for special treatment. The simplicity of the present method stems largely from the fact that we treat all of the unknowns, the P_α^k and the $\lambda_{\alpha\beta}$, on equal footing. To emphasize this point and to facilitate the analysis, it is desirable to make a slight change in notation. We associate each multiplier $\lambda_{\alpha\beta}(\beta \leqslant \alpha)$ with the radial function P_α (i.e., with the set of unknowns P_α^k, $k = 1,\ldots, M - 1$), and to emphasize the equal footing we define $P_\alpha^M \equiv \lambda_{\alpha\alpha}$, $P_\alpha^{M+1} \equiv \lambda_{\alpha\nu}$,..., etc., one definition for the normalization condition and one for each of the orthogonality conditions for which $\Delta(\alpha, \gamma) = 1$, the number of these definitions depending upon the specific case. We note in passing that these definitions have nothing to do with the boundary conditions (10b). The P_α^M were discarded previously from the list of unknowns, and we merely are reusing the symbol here in a different context.

In the revised notation equations (10a), (10c), and (10d) are a system of equations of the form

$$f_\alpha^k(P_1^1, P_1^2,\ldots; P_2^1, P_2^2,\ldots;\ldots; P_\alpha^1, P_\alpha^2,\ldots;\ldots; P_N^1, P_N^2,\ldots) = 0,$$
$$\alpha = 1,\ldots, N; \qquad k = 1,\ldots, M,\ldots, \tag{11}$$

where the equations $f_\alpha^{k<M} = 0$ are the finite difference equations (10a) for radial function α at each internal mesh point k, and the equations $f_\alpha^{k\geqslant M} = 0$ are the normalization and orthogonality conditions associated with radial function α. An orthogonality condition between two radial functions, β and γ say, is only included once in the set (11) and is associated with the radial function $\delta = \max(\beta, \gamma)$. The number of equations and the number of unknowns associated with each value of α in (11) will depend upon the specific case, and it is not easy to write down a general expression in terms of N and M. In every case, however, the number of equations is equal to the number of unknowns.

The equations (11) are a set of nonlinear algebraic equations in the variables P_α^k. The problem is to find the roots of these equations, i.e., to determine a value for each of the independent variables P_α^k such that each of the functions f_α^k is zero within a given tolerance. In order to solve this problem we introduce two complimentary iterative procedures: (1) the well-known SCF iteration and (2) a generalized Newton–Raphson iteration (GNRI).

We proceed as follows. We order the equations (11) in increasing values of α and k, grouping all equations with the same value of α together in subsets. Usually

we start with the innermost shell ($\alpha = 1$) and proceed outward, although the optimum ordering of radial functions with respect to SCF convergence varies from case to case. Within a SCF iteration we solve each of the subsets of equations (11) associated with each radial function (value of α) separately and in a sequence $\alpha = 1, 2,..., N$ by means of the GNRI to be described. In the usual SCF manner, for a given value of α in the sequence, the unknowns $P_\beta{}^j(\beta \neq \alpha; j = 1,..., M-1,...)$ 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 β in the present SCF iteration, whichever has occurred later. The dependence of the functions $f_\alpha{}^k$ on the variables $P_\beta{}^j(\beta \neq \alpha)$ is due to Coulomb and exchange integrals, off-diagonal multiplier terms, and orthogonality conditions. The functions $f_\alpha^{k<M}$, corresponding to the finite difference equations (10a), also depend upon the variables $P_\alpha{}^j$ in two ways: explicitly ($j = k-1, k, k+1, M, M+1,...$) and implicitly through the exchange integral $X_\alpha{}^k(j = 1,..., M-1)$. It is most advantageous to treat the $P_\alpha{}^j$ in the exchange integral as constants in the same manner as the $P_\beta{}^j(\beta \neq \alpha)$ while permitting the explicit dependence to vary within a SCF iteration. The $P_\alpha{}^j$ appearing in the functions $f_\alpha^{k\geqslant M}$, corresponding to normalization and orthogonality conditions (10c) and (10d), are also allowed to vary. This SCF procedure is similar to those commonly employed, except that the description is in terms of the finite difference variables.

We now address ourselves to the task of solving the subsets of equations (11) for each value of α. Due to the SCF procedure just outlined we now only consider the functional dependence of the $f_\alpha{}^k$ on the variables $P_\alpha{}^j(j = 1,..., M...)$ with the other variables treated as given constants. The solution is obtained by means of a generalized Newton–Raphson iteration. Let $\mathbf{P}_\alpha^{(n)} \equiv (P_\alpha{}^1, P_\alpha{}^2,...)^{(n)}$ be a solution vector of the values of the variables $P_\alpha{}^k$ evaluated at the nth GNRI. Also we define $\mathbf{F}_\alpha^{(n)} \equiv (f_\alpha{}^1, f_\alpha{}^2,...)^{(n)}$ to be a vector of the function values $f_\alpha{}^k(\mathbf{P}_\alpha^{(n)})$ evaluated at the nth GNRI. At the $(n+1)$th iteration the solution vector is given in terms of values at the nth iteration by

$$\mathbf{P}_\alpha^{(n+1)} = \mathbf{P}_\alpha^{(n)} - (J^{(n)})^{-1}\mathbf{F}_\alpha^{(n)}, \tag{12}$$

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

$$J_{lm}^{(n)} = \left[\frac{\partial f_\alpha{}^l}{\partial P_\alpha{}^m}\right]^{(n)}. \tag{13}$$

See [6] for a derivation of the iteration (12). Kantorovich and Akilov [9] give the general conditions under which such an iteration is convergent.

The iteration (12) is repeated until

$$\max(|f_\alpha{}^k(\mathbf{P}_\alpha^{(n)})|) < \text{tolerance}, \tag{14}$$

since when $\mathbf{F}^{(n)} = 0$ the problem is solved exactly for the given subset α. The starting vector $\mathbf{P}_\alpha^{(0)} = (P_\alpha{}^1, P_\alpha{}^2, \ldots)^{(0)}$ contains the respective values of the variables $P_\alpha{}^k$ as determined in the previous SCF iteration. The question of appropriate starting vectors $\mathbf{P}_\alpha^{(0)}$ in the first SCF iteration will be deferred until we discuss a specific case.

The unit operation of the method is the solution of the GNRI equation (12). Because of the SCF procedure and the particular ordering and grouping of Eqs. (11) just described, the Jacobian matrices (13) are always in a special nearly tridiagonal form which, by means of a partitioning, enables a rapid solution of (12) (see Appendix).

The details of this SCF–GNRI algorithm will be made clearer by application to a specific case.

5. Application to the $1s^2 2s$ Configuration of Li

This case has been chosen because it is the simplest example which illustrates the algorithm in full. Let P_1 and P_2 be the two 1s radial functions and P_3 the 2s radial function. Usually P_1 and P_2 are taken to be equal, but there is actually no physical or mathematical reason for this. Since the spin-orbitals 1 and 2 are automatically orthogonal because of spin, and since we assume the spin for 3 to be the same as that for 1, only an orthogonality condition on P_1 and P_3 is required. Let $\lambda_{11} = P_1{}^M$, $\lambda_{22} = P_2{}^M$, $\lambda_{33} = P_3{}^M$, and $\lambda_{31} = P_3{}^{M+1}$. To notationally separate those variables which are to take part in the present GNRI from those whose values are to be held constant, let us denote the latter by Q instead of P. So, for those variables to be held constant, $P_\alpha{}^k \to Q_\alpha{}^k$.

Writing out the system of equations (10) for this special case ($Z = 3$), including the Coulomb and exchange integrals $Y_\alpha{}^k$ and $X_\alpha{}^k$ explicitly, we have the following:

$$f_1{}^k = \Big[\frac{1}{h^2} - \frac{3}{kh(1 - akh)^3} + (1 - akh)^{-4} \sum_{j=1}^{M-1} ((Q_2{}^j)^2 + (Q_3{}^j)^2)(1 - ajh)^{-4} \times \frac{(1 - aqh)}{q} - (1 - akh)^{-4} P_1{}^M\Big] P_1{}^k - \frac{1}{2h^2}(P_1^{k-1} + P_1^{k+1}) - Q_3{}^k(1 - akh)^{-4} \sum_{j=1}^{M-1} Q_1{}^j Q_3{}^j (1 - ajh)^{-4} \frac{(1 - aqh)}{q} - Q_3^{M+1}(1 - akh)^{-4} Q_3{}^k = 0, \qquad k = 1, \ldots, M - 1, \tag{15}$$

$$f_1{}^M = \sum_{j=1}^{M-1} (1 - ajh)^{-4}(P_1{}^j)^2 - 1/h = 0, \tag{16}$$

$$f_2^k = \left[\frac{1}{h^2} - \frac{3}{kh(1-akh)^3} + (1-akh)^{-4}\sum_{j=1}^{M-1}((Q_1^j)^2 + (Q_3^j)^2)\right.$$

$$\left.\times (1-ajh)^{-4}\frac{(1-aqh)}{q} - (1-akh)^{-4}P_2^M\right]P_2^k$$

$$-\frac{1}{2h^2}(P_2^{k-1} + P_2^{k+1}) = 0, \qquad k = 1,\ldots, M-1, \tag{17}$$

$$f_2^M = \sum_{j=1}^{M-1}(1-ajh)^{-4}(P_2^j)^2 - 1/h = 0, \tag{18}$$

$$f_3^k = \left[\frac{1}{h^2} - \frac{3}{kh(1-akh)^3} + (1-akh)^{-4}\sum_{j=1}^{M-1}((Q_1^j)^2 + (Q_2^j)^2)\right.$$

$$\left.\times (1-ajh)^{-4}\frac{(1-aqh)}{q} - (1-akh)^{-4}P_3^M\right]P_3^k$$

$$-\frac{1}{2h^2}(P_3^{k-1} + P_3^{k+1})$$

$$-Q_1^k(1-akh)^{-4}\sum_{j=1}^{M-1}Q_3^jQ_1^j(1-ajh)^{-4}\frac{(1-aqh)}{q}$$

$$-P_3^{M+1}(1-akh)^{-4}Q_1^k = 0, \qquad k = 1,\ldots, M-1, \tag{19}$$

$$f_3^M = \sum_{j=1}^{M-1}(1-ajh)^{-4}(P_3^j)^2 - \frac{1}{h} = 0, \tag{20}$$

$$f_3^{M+1} = \sum_{j=1}^{M-1}(1-ajh)^{-4}Q_1^jP_3^j = 0, \tag{21}$$

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

A SCF iteration consists of first using the GNRI to solve (15) and (16) with all the Q's held constant at their respective values from the previous SCF iteration. Next, (17) and (18) are solved in the same manner. Finally (19), (20), and (21) are solved, but this time the Q's in the last terms of Eqs. (19) and (21) are held constant at their respective values from the solution of (15) and (16) in the present SCF iteration instead of the preceding SCF iteration. This procedure results in all spin-

orbitals being orthonormal after every SCF iteration and seems to aid convergence. The Jacobian matrix for the system of Eqs. (19), (20), and (21) is given in the appendix.

It remains only to define a starting vector for the algorithm. If the terms involving the Coulomb and exchange integrals as well as the off-diagonal multipliers are ignored the system of finite difference equations (15), (17), and (19) corresponds physically to three noninteracting electrons in the field of the bare nucleus. The solution corresponds to hydrogenic wavefunctions ($Z = 3$) expressed as a function of ρ. For the starting vector, therefore, the P_1^k and P_2^k ($k = 1,..., M-1$) are both set equal to the lowest hydrogenic eigenfunctions in finite difference form and P_1^M and P_2^M equal to the corresponding eigenvalue. Similarly the $P_3^k (k = 1,..., M)$ are set equal to the first excited state. The starting values for the off-diagonal multipliers $P_\alpha^{k>M}$ are zero.

In the event that difficulty is encountered in converging from the hydrogenic solutions, one can multiply the ignored terms by a tracking parameter ϵ and increase ϵ in several steps from 0 to 1 with the converged solution for one value of ϵ serving as the starting vector for the next. In the Li calculation which follows one such intermediate step was required ($\epsilon = 0, 0.5, 1.0$). Also solutions for one atom can be tracked into solutions for a larger atom by means of turning on the additional Coulomb and exchange interactions in a number of steps with the additional orbitals starting from appropriate hydrogenic functions.

6. Results for the $1s^2 2s$ Configuration of Li

Continuing with the Li example we present results for the total HF energy and the Lagrange multipliers. The result for the HF energy appears in Table I where the Richardson h^2-extrapolation procedure [10] has been carried out to determine the limiting value of the HF energy as $h \to 0$. The various extrapolants are arranged in a Neville table [11] where the accuracy of the extrapolants improves from left to right, and consequently the element on the far right of the table is the most accurate approximation. The other elements are useful in giving an estimate of the accuracy. The results for λ_{11}, λ_{22}, and λ_{33} are similarly presented.

For these double-precision calculations the tolerance for the GNRI (see Eq. (14)) was set at 10^{-11}. Quadratic convergence is obtained. In a typical sequence of generalized Newton–Raphson iterations the convergence criterion (14) takes on values of the order of 10^0, 10^{-3}, 10^{-7}, and 10^{-12}. The SCF convergence criterion used was that the largest absolute change in any of the P_α^k from one SCF iteration to the next be less than 10^{-10}.

With 105 intervals in the mesh the converged value of λ_{31} turns out to be $-0.5306684623 \times 10^{-10}$. Since this value is of the order of the tolerance of the

TABLE I

Neville Tables for Hartree–Fock Energy and λ Multipliers for Li

M = number of intervals in mesh	h = mesh[b] spacing = 1/(aM)	0th order extrapolant	1st order extrapolant	2nd order extrapolant	3rd order extrapolant
		HF energy[a]			
45	0.02002002002	−7.432260308			
			−7.432753021		
65	0.01386001386	−7.432516868		−7.432750941	
			−7.432751524		−7.432750918
85	0.01059883413	−7.432614303		−7.432750922	
			−7.432751153		
105	0.00858000858	−7.432661471			
		λ_{11}			
45	0.02002002002	−2.484760980			
			−2.486676948		
65	0.01386001386	−2.485758644		−2.486675785	
			−2.486676111		−2.486675610
85	0.01059883413	−2.486139599		−2.486675642	
			−2.486675822		
105	0.00858000858	−2.486324420			
		λ_{22}			
45	0.02002002002	−2.466734377			
			−2.468701023		
65	0.01386001386	−2.467758430		−2.468699916	
			−2.468700227		−2.468699720
85	0.01059883413	−2.468149487		−2.468699755	
			−2.468699936		
105	0.00858000858	−2.468339211			
		λ_{33}			
45	0.02002002002	−0.1969356494			
			−0.1963669527		
65	0.01386001386	−0.1966395233		−0.1963672406	
			−0.1963671599		−0.1963672172
85	0.01059883413	−0.1965264312		−0.1963672212	
			−0.1963671979		
105	0.00858000858	−0.1964715481			

[a] *HF* energy is in a.u.

[b] $a = 1.11$ (see Eq. (5)).

calculation, and since roundoff error is also of this order, an accurate h^2-extrapolation is not possible.

It is well known that with the two 1s radial functions P_1 and P_2 not constrained to be equal, the HF equations are invariant under a unitary transformation among spin-orbitals of the same spin (in this case P_1 and P_3) and furthermore, since the λ matrix is Hermitian, there exists a representation in which it is diagonal. It is clear that within the given tolerance the algorithm has converged upon this diagonal solution. That this is certainly the case is evident from the fact that when the set of finite difference equations (15)–(21) are satisfied within the tolerance of 10^{-10} we have an accurate solution of the HF equations. But as a check a second calculation was performed where λ_{31} was assumed to be zero. For this calculation the term containing P_3^{M+1} in (19) (P_3^{M+1} corresponds to λ_{31}), as well as the orthogonality condition between P_1 and P_3, Eq. (21), were eliminated from the system. P_1 and P_3 still turn out to be orthogonal without the constraint (21) because of the Hermitian property of the HF equations in the diagonal representation. The results of the two calculations are identical within the given tolerance.

The fact that the solution of the HF equations is unique only within a unitary transformation can lead to numerical instabilities with conventional methods of solution. For this reason one is usually content to assume the diagonal representation (even in cases where it does not legitimately exist, e.g. open shells with two electrons of opposite spin in each orbital). We note that in the present calculation no such numerical instabilities occurred, even though the solution was not unique.

It might be asked if the solution is not unique why did the algorithm converge on the particular diagonal solution? We note that the GNRI works on one radial function at a time, holding the others constant, and therefore at each stage in a SCF iteration the solution is unique. Also the results of one SCF iteration uniquely determine the results of the next (within tolerance). Thus, the only question is whether the lack of a unique solution causes the SCF sequence to diverge, and we see at least in the present calculation that it does not. Which one of the non-unique solutions is finally converged upon is determined by the starting approximation. In the present calculation λ_{31} was started at zero, and within the given tolerance it remained there.

The method certainly does not depend upon including the nondiagonal multipliers and the associated orthogonality conditions in the system of finite difference equations. They have been included in the Li example only to present the algorithm in full generality. All that is required is that the number of finite-difference equations must equal the number of unknowns P_a^k.

Finally we give an indication of the computing time requirements for the method. On an IBM 360/44 computer in double precision for Li with 85 intervals in the mesh, 0.36 sec were required per spin-orbital per SCF iteration. Starting from

hydrogenic orbitals 17 SCF iterations were required to achieve convergence to a tolerance of 10^{-10}, independent of the number of mesh points in the range 45–105. It is difficult to compare with other methods since this information is not available in the literature for our particular example.

It should be noted that once a solution has been obtained for one mesh spacing, solutions for other mesh spacings can be obtained rapidly in one or two iterations from a linearly interpolated starting vector, so the Richardson extrapolation process does not require very much extra computing time. Also the use of acceleration parameters to improve SCF convergence can be incorporated into the method.

Besides the calculation of the $1s^22s$ configuration of Li presented, calculations of the excited $1s^22p$, $1s^23s$, $1s^23p$, and $1s^23d$ configurations have been performed. Also a calculation of the closed-shell $1s^22s^2$ configuration of Be in the diagonal representation has been performed.

7. Conclusions

The GNRI in conjunction with the SCF iteration is a relatively simple and efficient method for obtaining accurate numerical solutions of the atomic HF equations. The computer storage requirements are no greater than conventional methods. Although calculations have been performed to date only for Li and Be, there is no reason why the algorithm could not compete favorably with conventional methods in handling large atoms.

The chief advantage of the present finite difference Newton–Raphson algorithm over conventional methods lies in the treatment of the split boundary conditions. Where conventional methods incorporate rather artful procedures for translating a mismatch in the inward and outward integrations into new estimates of the values of multipliers and initial slopes, the present algorithm includes the split boundary conditions and the multipliers in the system of finite-difference equations, and the problem is solved universally rather than by a combersome process of fitting together the various parts. Consequently the uninitiated should find the present algorithm much easier to apply.

Appendix

As an illustration of the form of the Jacobian matrix encountered using SCF techniques we give the example of the Jacobian matrix for the $2s$ orbital of Li as derived from Eqs. (19), (20), and (21).

$$J_{kj} = \frac{\partial f_3^k}{\partial P_3^j}, \tag{A-1}$$

$$= +\frac{1}{h_2} - \frac{3}{kh(1-akh)^3} + (1-akh)^{-4} \sum_{l=1}^{M-1} ((Q_1^l)^2 + (Q_2^l)^2)$$

$$\times (1-alh)^{-4} \frac{(1-aqh)}{q} - (1-akh)^{-1} P_3^M, \qquad j = k = 1,\ldots, M-1,$$

$$= -\frac{1}{2h^2}, \qquad |k-j| = 1 \qquad j,k = 1,..,M-1,$$

$$= -(1-akh)^{-4} P_3^k, \qquad j = M, \quad k = 1,\ldots, M-1,$$

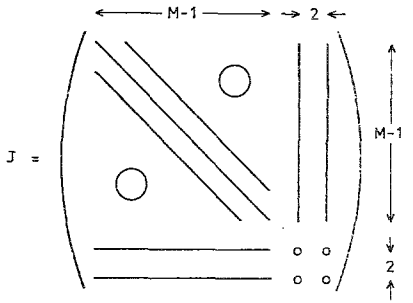
$$= 2(1-ajh)^{-4} P_3^j, \qquad k = M, \quad j = 1,\ldots, M-1,$$

$$= -(1-akh)^{-4} Q_1^k, \qquad j = M+1, \quad k = 1,\ldots, M-1,$$

$$= (1-ajh)^{-4} Q_1^j, \qquad k = M+1, \quad j = 1,\ldots, M-1,$$

$$= 0, \qquad \text{for all other } k, j \text{ not included in any of the above}$$

In matrix form J is the following,



(A-2)

All Jacobian matrices encountered are of this form, though the number of nonzero columns and rows changes, depending on the number of orthogonality conditions associated with the spin-orbital under consideration.

We will outline the rapid solution of the GNRI equation (12) that is possible if J is of the form given by (A-2).

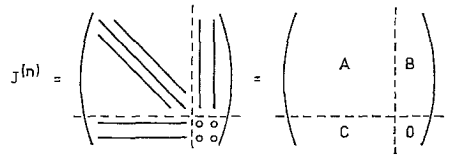
First, omitting the subscript α, let us rewrite Eq. (12) as

$$J^{(n)}\Delta = \mathbf{F}^{(n)}, \tag{A-3}$$

where the vector Δ is defined by

$$\Delta \equiv \mathbf{P}^{(n)} - \mathbf{P}^{(n+1)}.$$

Once (A-3) is solved for Δ, the vector $\mathbf{P}^{(n+1)}$ is directly obtained since $\mathbf{P}^{(n)}$ is already known. Next, we perform the following partitions.



(A-4)

where A is an $(M-1) \times (M-1)$ tridiagonal matrix, B is a general $(M-1) \times 2$ matrix and C is a general $2 \times (M-1)$ matrix. Furthermore, we let

$$\Delta = \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix} \tag{A-5}$$

and

$$\mathbf{F}^{(n)} = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{pmatrix}, \tag{A-6}$$

where Δ_1 and $\mathbf{F}_1$ are vectors of length $M-1$ while Δ_2 and $\mathbf{F}_2$ are vectors of length 2. Equation A-3 now becomes the following:

$$\begin{pmatrix} A & B \\ C & 0 \end{pmatrix} \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix} = \begin{pmatrix} \mathbf{F}_1 \\ \mathbf{F}_2 \end{pmatrix} \tag{A-7}$$

or equivalently it can be expressed as two matrix equations as follows:

$$C\Delta_1 = \mathbf{F}_2 , \tag{A-8}$$

$$A\Delta_1 + B\Delta_2 = \mathbf{F}_1 . \tag{A-9}$$

From (A-8) and (A-9) we obtain directly that

$$CA^{-1}B\Delta_2 = CA^{-1}\mathbf{F}_1 - \mathbf{F}_2 \tag{A-10}$$

and

$$A\Delta_1 = \mathbf{F}_1 - B\Delta_2 . \tag{A-11}$$

Equations (A-10) and (A-11) are solved in the following way. First, we obtain $A^{-1}B$ by defining the following relationships [12]:

$$W_1 \equiv A_{11}^{-1}A_{12} ; \quad W_j \equiv (A_{jj} - A_{jj-1}W_{j-1})^{-1} A_{jj+1} , \quad j = 2,..., M-1, \tag{A-12}$$

$$G_{1\alpha} \equiv A_{11}^{-1}B_{1\alpha} ; \quad G_{j\alpha} \equiv (A_{jj} - A_{jj-1}W_{j-1})^{-1}(B_{j\alpha} - A_{jj-1}G_{j-1\alpha}),$$
$$j = 2,..., M-1; \quad \alpha = 1, 2, \tag{A-13}$$

which make possible the recursive computation of the components of $A^{-1}B$.

$$(A^{-1}B)_{M-1,\alpha} = G_{M-1,\alpha}\,; \qquad (A^{-1}B)_{j,\alpha} = G_{j\alpha} - W_j(A^{-1}B)_{j+1,\alpha}\,,$$
$$j = M-2,\ldots,1; \qquad \alpha = 1, 2. \qquad \text{(A-14)}$$

In a similar fashion $A^{-1}\mathbf{F}_1$ in (A-10) is also obtained. By inverting the 2×2 matrix $CA^{-1}B$ and multiplying (A-10) through by $(CA^{-1}B)^{-1}$, we obtain Δ_2. Finally by the same technique that we used to obtain $A^{-1}B$, we obtain $\Delta_1[= A^{-1}(\mathbf{F}_1 - B\Delta_2)]$, and the problem is solved. Note that the matrix A is not actually inverted at any stage during the solution. Also the number of calculations and storage locations for this method of solving Eq. (12) are both proportional to $M + 1$. Thus, it is clear how the above partitions greatly reduce the time and storage requirements, for without them J would have to be treated as a general square matrix of dimension $M + 1$. The time required for its inversion would then be proportional to $(M + 1)^3$, and $(M + 1)^2$ storage locations would be required.

References

1. C. Froese, *Canad. J. Phys.* **41** (1963), 1895.
2. F. C. Smith and W. R. Johnson, *Phys. Rev.* **160** (1967), 136.
3. D. F. Mayers and F. O'Brien, *J. Phys. B.* **1** (1968), 145.
4. J. B. Mann and J. T. Waber, *J. Chem. Phys.* **53** (1970), 2397.
5. D. R. Hartree, "The Calculation of Atomic Structures," Wiley, New York, 1957.
6. 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. Szebeheley, Eds.), Academic Press, New York and London, 1966.
7. N. W. Winter, D. Diestler, and Vincent McKoy, *J. Chem. Phys.* **48** (1968), 1879.
8. S. F. Boys and N. C. Handy, *Proc. Roy. Soc. A* **310** (1969), 63.
9. L. V. Kantorovich and G. P. Akilov, "Functional Analysis in Normed Spaces," Pergamon Press, New York, 1964.
10. L. F. Richardson, *Philos. Trans. Roy. Soc. London Ser. A* **226** (1927), 299; Z. Kopal, "Numerical Analysis," Section V-C, Wiley, New York, 1961.
11. E. H. Neville, *J. Indian Math. Soc.* **20** (1934), 87.
12. R. S. Varga, "Matrix Iterative Analysis," p. 196, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1962.