JOURNAL OF COMPUTATIONAL PHYSICS 67, 290–309 (1986)

Nonrelativistic Numerical MCSCF for Atoms

F. BIEGLER-KÖNIG AND JUERGEN HINZE

Fakultät für Chemie, Universität Bielefeld,
D 4800 Bielefeld, Federal Republic of Germany

Received November 1, 1985; revised March 5, 1986

We describe here the theory underlying a computer program we have developed for the calculation of nonrelativistic numerical electronic MCSCF wavefunctions for atoms. The formulation and program are based on earlier developments and programs by Froese-Fischer [*The Hartree–Fock Method for Atoms* (Wiley, New York, 1977)] with some significant modifications and extensions to be described here explicitly. © 1986 Academic Press, Inc.

I. INTRODUCTION

The direct numerical determination of the electronic wavefunctions for atoms has a long history and the underlying theory is well developed already [2]. In the central field case it is sufficient to carry out a one-dimensional numerical integration of the SCF or MCSCF equations, which can be done rapidly and with high numerical accuracy, using the finite difference method [3]. If the finite difference integration of the differential equations and the required quadratures are implemented with a consistent truncation error, the Richardson extrapolation [4] provides an effective control of the numerical accuracy achieved. This and the greater computational efficiency make the numerical determination of electronic wavefunctions for atoms superior to the basis-function expansion alternative. A drawback, however, is the difficulty to obtain convergence in solving the general MCSCF equations. The quadratically convergent algorithms developed for basis-function calculations [5] cannot be extended and implemented easily to the direct numerical determination of the orbitals.

In the following we will give a brief formulation of the MCSCF theory as it applies to the numerical determination of atomic wavefunctions. Several procedures of increasing computational complexity to obtain convergence in the solution of the nonlinear MCSCF problem will be described; procedures we have implemented and used effectively. In Section IV we will present some numerical details concerning the integration grid, the consideration of the boundary conditions and the evaluation of the electron interaction functions. We will make some remarks about the selection of the initial shell-functions and conclude with the results of some illustrative example calculations.

0021-9991/86 $3.00
Copyright © 1986 by Academic Press, Inc.
All rights of reproduction in any form reserved.

II. MCSCF-THEORY FOR ATOMS

The total wavefunction of a state K, which is an eigenfunction of the operators L^2, L_z, S^2, and S_z is written as

$$\Psi_K = \sum_{I=1}^{N} \Phi_I C_{IK}, \tag{1}$$

where the configuration state functions (CSFs) Φ_I are mutually orthonormal minimal superpositions of Slater determinants (SDs) obtained via the appropriate coupling of the angular momenta, such that the CSFs are eigenfunctions of L^2, L_z, S^2, and S_z with all the same eigenvalues as desired for the state considered. The SDs are constructed from spin orbitals

$$\psi_{\lambda im\sigma} = P_{\lambda i}(r)/r\, Y_{\lambda m}(\theta, \varphi)\, \sigma(s) \tag{2}$$

with the shell-functions $P_{\lambda i}(r)$ the same for the $(2\lambda + 1)$ components of m and the α and β spin components of a shell. To keep the notation simple we will use a single index, i, for the shell functions, dropping the index λ, which signifies the angular momentum $l = \lambda(i)$ for the shell-function P_i.

The unknowns for the determination of the state-functions Ψ_K are now the expansion coefficients C_{IK} and the shell-functions P_i. The former are obtained as the solutions of the matrix eigenvalue equation

$$HC = CE, \tag{3}$$

where E is diagonal with the eigenvalues E_K and C is the matrix with the eigenvectors C_K as columns. Equation (3) obtains, provided

$$\langle \Phi_I \,|\, \Phi_J \rangle = \delta_{IJ} \tag{4}$$

as demanded for the construction of the CSFs. This is achieved readily if the spinorbitals are chosen orthonormal

$$\langle \Psi_{\lambda im\sigma} \mid \Psi_{\lambda' jm'\sigma'} \rangle = \delta_{\lambda\lambda'}\, \delta_{\sigma\sigma'}\, \delta_{ij} \tag{5}$$

which demands that shell-functions of the same l-value are maintained orthonormal, i.e.,

$$\int_0^\infty P_i(r)\, P_j(r)\, dr = \delta_{ij} \qquad \text{for} \quad \lambda(i) = \lambda(j). \tag{6}$$

Note that we have restricted the shell-functions to be real, which may be done without loss of generality.

The energy matrix elements in (3) are obtained explicitly as

$$\begin{aligned} H_{IJ} &= \langle \Phi_I | \mathscr{H} | \Phi_J \rangle \\ &= \sum_{ij} \Big[\gamma_{ij}^{IJ} \langle P_i | h[\lambda(i)] | P_j \rangle + 1/2 \sum_{klv} \Gamma_{ijklv}^{IJ} \\ &\quad \times \Big\langle P_i \langle P_k | \frac{r_<^{v}}{r_>^{v+1}} | P_l \rangle P_j \Big\rangle \Big], \end{aligned} \tag{7}$$

where the one electron operator

$$h[\lambda] = -1/2 \frac{d^2}{dr^2} - \frac{z}{r} + \frac{\lambda(\lambda+1)}{2r^2}$$

depends explicitly on the λ-value of the shell.

The structure factors γ and Γ are the transition matrix elements of the first and second order reduced density matrix in the space of the shell-functions. They are determined analytically through the angular momentum coupling construction of the CSFs and an integration over the spin-functions and spherical harmonics, specifying the spin-orbitals together with spherical harmonics of the Laplace expansion of the $1/r_{12}$ operator [6]. Note that the γ_{ij}'s will be zero unless $\lambda(i) = \lambda(j)$.

The effective structure factors of a specific state, K, are determined once Eq. (3) is solved as

$$\gamma_{ij}^{K} = \sum_{IJ} C_{IK} C_{JK} \gamma_{ij}^{IJ} \quad \text{and} \quad \Gamma_{ijklv}^{K} = \sum_{IJ} C_{IK} C_{JK} \Gamma_{ijklv}^{IJ}. \tag{9}$$

They would enter the energy functional to be minimized for the determination of the orbitals optimal for the characterization of state K. If state K is not the lowest state of a particular symmetry it is advisable to calculate this excited state together with all the lower states of same symmetry, thus maintaining an upper bound for the excited state. In this case it is also appropriate to define and use suitably averaged structure factors [7]

$$\gamma_{ij} = \sum_{K} \omega_K \gamma_{IJ}^{K} \quad \text{and} \quad \Gamma_{ijklv} = \sum_{K} \omega_K \Gamma_{ijklv}^{K} \tag{10}$$

with $\sum_K \omega_K = 1$ to avoid that the orbitals, once they are optimized for an excited excited state alone, would give in the CI calculation for this state a lower CI energy than for the state, which should in actuality be lower. The weights used in Eq. (10) are free except for their normalization. Deficiencies due to this averaging can be rectified through the inclusion of extra CSFs in the CI-expansion. Note that this averaging of the structure factors is also advantageous, if a common orbital set is desired in the characterization of several states. Such a demand would greatly facilitate the later calculation of transition properties between such states. We will deal in the following with these general, averaged structure factors.

The energy functional to be minimized with respect to a change in the shell-functions is

$$\begin{aligned} L\{E\} &= \sum_K \omega_K E_K - \sum_{ij} \varepsilon_{ij}(\langle P_i \mid P_j \rangle - \delta_{ij}) \\ &= \sum_{ij} \gamma_{ij} \langle P_i \mid h[\lambda(i)] \mid P_j \rangle + 1/2 \sum_{ijkl\nu} \Gamma_{ijkl\nu} \\ &\quad \times \left\langle P_i \langle P_k \mid \frac{r_<^\nu}{r_>^{\nu+1}} \mid P_l \rangle P_j \right\rangle - \sum_{ij} \varepsilon_{ij}(\langle P_i \mid P_j \rangle - \delta_{ij}) \end{aligned} \tag{11}$$

with the restrictive conditions of orbital orthonormality introduced using the Lagrange multipliers ε_{ij}. Note that some of the off-diagonal Lagrange multipliers will be zero, i.e., if $\lambda(i) \neq \lambda(j)$, or may be chosen to be zero (a) between closed shells and (b) between shells among which a full CI expansion is used.

Extremalization of the energy functional, Eq. (11) with respect to a change of the shell-function P_i leads to the standard MCSCF–Fock equations

$$1/2\nabla_i L\{E\} = \sum_j \left[\gamma_{ij} h + \sum_{ijkl\nu} \Gamma_{ijkl\nu} U_{kl\nu} - 1/2(\varepsilon_{ij} + \varepsilon_{ji}) \right] P_j = 0 \tag{12}$$

with the two-electron potentials

$$U^{(r)}_{kl\nu} = \langle P_k \mid \frac{r_<^\nu}{r_>^{\nu+1}} \mid P_l \rangle. \tag{13}$$

The gradient operator ∇_i denotes here the vector of partial derivatives with respect to the independent variables in P_i. In a numerical representation of the shell functions P_i, the independent variables would be the values of P_i at the grid points.

Equations (12), coupled one-dimensional second order differential equations, may be solved using standard finite difference methods [3], an appropriate discretization of r, and recognizing that at the solution

$$\varepsilon_{ij} = \sum_n \langle P_i \mid \gamma_{jn} h + \sum_{kl\nu} \Gamma_{jnkl\nu} U_{kl\nu} \mid P_n \rangle. \tag{14}$$

However, even with the structure factors fixed, Eqs. (12) depend explicitly on the solution functions P, because of the two-electron potential; thus an iteration technique is indicated. For the solution functions we must have $\varepsilon_{ij} = \varepsilon_{ji}$ [8]. Note also that the structure factors depend implicitely on the shell-functions, the solutions of Eqs. (12).

III. Solving of the MCSCF Problem

A. *The Standard Procedure*

Quadratically convergent algorithms to solve the MCSCF problem [5] vary the Lagrange functional of the energy, Eq. (11), as a total, including the dependency of the structure factors on the CI coefficients, which depend on the shell-functions. In the standard MCSCF methods these dependencies are considered decoupled. An outer iterative CI-cycle solves the regular equation (3) with fixed shell-funtions, thus determining the effective structure factors. The inner iterative SCF cycle solves the general Fock equations (12), with the structure factors fixed to obtain new shell-functions. To accomplish this a minor innermost iterative cycle is needed to solve Eqs. (12), with structure factors and two particle potential fixed. This we accomplish by solving for the m shell functions

$$(D_i - \varepsilon_{ii})\, P_i = X_i \tag{15}$$

with

$$D_i = \gamma_{ii} h + \sum_{klv} \Gamma_{iiklv} U_{klv} \tag{16}$$

and

$$X_i = \sum_{j=i} \left[1/2(\varepsilon_{ij} + \varepsilon_{ij}) - \gamma_{ij} h - \sum_{klv} \Gamma_{ijklv} U_{klv} \right] P_j \tag{17}$$

using inverse iteration recomputing ε_{ii} in each step.

In total we have a threefold nested iterative cycle of the following structure:

(1) Obtain a set of starting functions P_i,

(2) orthonormalize the P_i's,

(3) compute the two-particle potentials U_{klv}.

Start CI-cycle:

(4) Construct H and solve $HC = CE$.

(5) If E_K desired is a minimum and the inner selfconsistency conditions are satisfied, we are done, else:

(6) Compute the effective structure factors.

Start SCF-cycle

(7) Compute $\nabla_i E$ for $i = 1, m$ and

$$\varepsilon_{ij} + \varepsilon_{ji} = 1/2(\langle P_i \mid \nabla_j E \rangle + \langle \nabla_i E \mid P_j \rangle) \qquad \text{for} \quad j = 1, m.$$

(8) Compute D_i, X_i, and $\nabla_i L(E)$ for $i = 1, m$.

(9) If all $\|\nabla_i L(E)\|$ are small enough, the SCF cycle is converged, go to step (4), else:

Start inverse iteration cycle for $i = 1, m$:

(10) Obtain new P_i's from $P_i = (D_i - \varepsilon_{ii})^{-1} X_i$.

(11) Normalize P_i's and compute $\varepsilon_{ii} = \langle P_i \mid D_i P_i - X_i \rangle$.

(12) If the change in P_i or ε_{ii} was too large, go to step (10), else:

End inverse iteration:

(13) Orthonormalize new P_i's.

(14) Compute the two particle potentials U_{klv} and go to step (7).

End SCF cycle;
End CI cycle.

This iterative process is by far not universally convergent; without modification it will converge only for the most benign SCF calculations. A significant improvement of the convergence can be achieved with the following linear search technique after step (14).

Let us denote the quantities used in SCF iteration n in steps (7) and (8) as P^n, E^n and $\nabla L(E^n)$ and the resulting vector in step (13) as $\bar{P}^{n+1}$. Note: the vectors P^n, $\bar{P}^{n+1}$, and $\nabla L\{E\}$ used here are the combination of all shell-function vectors into one single supervector. We define a linear search direction by

$$d = P^n - \bar{P}^{n+1} \tag{18}$$

and calculate

$$E(P^n + \lambda d) = \min_{\mu \in \mathbb{R}} E(P^n + \mu d), \tag{19}$$

where $(P^n + \mu d)$ is to be understood as the set of vectors obtained after adding μd to P^n and orthonormalizing. The new shell-functions are obtained as

$$P^{n+1} = (P^n + \lambda d). \tag{20}$$

For the minimum search in Eq. (19) to determine λ we fit, following the Davidon algorithm [9], a cubic polynomial to the data points $E(P^n)$, $E(P^n + \mu d)$ and the two directional derivatives $d^T \cdot \nabla L\{E(P^n)\}$ and $d^T \cdot \nabla L\{E(P^n + \mu d)\}$. In this linear search the steps (2)–(4), (7), and (8) have to be carried out a few times, often just twice.

This procedure is akin to the damping process used by Froese-Fischer [1], however, here with an optimal damping coefficient λ determined. It leads in almost all cases rapidly to convergence. To be sure, it should be combined with an effective dynamical control of the convergence threshold, in order not to carry out unnecessary inner iteration cycles, while the outer iterations are still far from convergence. It is also advisable to permit in the initial iterations the selective optimization of individual orbitals while others are held frozen.

It is conceivable that the following two strategems might yield an additional improvement in the convergence:

(a) Use of a multidimensional nonlinear optimization [9] in the determination of the optimal damping factors with

$$d_i = P_i^n - \bar{P}_i^{n+1} \tag{21}$$

and

$$P_n^{n+1} = (P_i^n + \lambda_i d_i), \tag{22}$$

where the optimal λ_i's are determined by the nonlinear minimization of $E(P_i^n + \mu_i d_i)$.

(b) Unitary transformations among the new shell-functions in an SCF iteration, such as to obtain a symmetric Lagrange multiplier matrix [1, 10].

We have not yet implemented these two options into our program, because our experience with the simpler linear search outlined above has been sufficiently good. To be sure, there are situations, where the coupling of the CI-mixing coefficients and the orbitals is too strong and our improved standard procedure will converge only slowly or not at all. In this case a quadratically convergent process, including explicitly the CI-orbital coupling is called for.

B. *A Quadratically Convergent Numerical* MCSCF *Procedure*

The classical quadratically convergent algorithm to solve a nonlinear problem is Newton's method, which leads to a set of linear equations for the independent variables with the Hessian, the matrix of all partial second order derivatives as the coefficient matrix and the vector of first order partial derivatives as right-hand side. This procedure derived and used effectively in MCSCF calculations with basis-functions [5] cannot be used efficiently, according to our experience, in numerical MCSCF calculations for two reasons:

(i) The Hessian matrix of the dimension ($N =$ number of shell-functions × number of grid-points) will be exceedingly large and full, not banded or sparse.

(ii) This large linear system, which has to be solved iteratively [11] in each Newton-step, is nearly singular and becomes especially unstable if some shells are only weakly occupied.

The numerical instability and the amount of computational effort required to apply Newton's method directly nullifies the computational advantages of the direct numerical determination of the shell-functions.

As an alternative we have implemented a gradient method of quasi-Newton type [9] to minimize the energy directly. Such methods satisfy so-called "quadratic termination conditions," i.e., they solve, in the absence of rounding errors, a quadratic minimization problem in at most N (number of unknowns) iterations. The particular algorithm we have implemented is the (complementary Davidon–Fletcher–Powell) CDFP-method [9] where in each step an improved approximation to the inverse of the Hessian matrix is obtained by adding to the old approximation a particular symmetric rank-two matrix.

The detailed structure of this algorithm for the numerical MCSCF is as follows:

(1) Obtain a set of starting functions P^0.

(2) Orthonormalize P_i's and compute E^0 and $\nabla L\{E^0\}$.

(3) Set $k=0$ and $H^0=\mathbf{1}$ $(N \times N)$.

Iteration cycle:

(4) Define the search direction.

$$d^k = -H^k \cdot \nabla L\{E^k\}$$

(5) Find the λ which minimizes $E(P^k + \lambda\, d^k)$; if no minimum can be found then

(a) if $k=0$ we are done,

(b) if $k \neq 0$ go to (3), restart; else:

(6) Set $P^{k+1} = (P^k + \lambda\, d^k)$ orthonormalized compute E^{k+1} and $\nabla L\{E^{k+1}\}$.

(7) If $\|\nabla L\{E^{k+1}\}\|$ is small enough we are done; else:

(8) With $\sigma = \lambda\, d^k$ compute

$$\begin{aligned}
\gamma &= \nabla L\{E^{k+1}\} - \nabla L\{E^k\},\\
h &= H^k \cdot \gamma,\\
b &= (1 + \gamma^T \cdot h/\gamma^T \cdot \sigma) \cdot \sigma - h,\\
s &= \sigma/\gamma^T \cdot \sigma\\
H^{k+1} &= H^k + s^T b - h^T s,
\end{aligned}$$

set $k = k+1$ and go to 4.

The computation of E in step (2) and (6) implies the solution of the eigenvalue problem, Eq. (3). Naturally the computations of E and $\nabla L\{E\}$ are to be performed with orthonormalized shell-functions only, using the procedures as outlined in steps (3), (4), and (6)–(8) of the first algorithm. Some advantages of the CDFP-method are

(i) it always converges to a minimum (if existent), this may be a local minimum though;

(ii) no second derivatives are needed;

(iii) no ill-conditioned linear system must be solved;

(iv) the update of H^k requires only $\mathcal{O}(N^2)$ operations and is, since H^k is positive definite, numerically stable.

The major drawback of this method is that it requires much more computeroperations, assuming N iterations, $\mathcal{O}(N^3)$, and store, $\mathcal{O}(N^2)$, with $N =$ number of shell functions × grid points, than the algorithm given for the standard MCSCF method. Providing good starting functions and freezing some already well-determined orbitals initially can reduce the computational demand significantly. Frequently we have found convergence in much less than N cycles.

IV. NUMERICAL CONSIDERATIONS

In this section we describe in detail the numerical procedures implemented to carry out the computations of the energy and its gradients needed in the algorithms outlined above.

A. *Choice of the Grid*

In the finite difference method an equally spaced grid to represent the functions and operators is desirable. However, an equidistant discretization of r directly is not suitable because

(i) the kinetic energy operator will be inappropriate near the origin,

(ii) errors will be caused by the cutoff at some finite r,

(iii) the desired functions $P_i(r)$ will in general be rapidly varying for small r and slowly varying as r gets large.

We decided not to use a logarithmic grid [1, 2], which requires an analytical origin expansion, but rather to map the infinite interval of $r[0, \infty]$ onto a finite one with $\rho[0, 1]$ using the rational transformation

$$\rho = \frac{r}{r+b} \tag{23}$$

with $b > 0$ a real parameter, which controls the distribution of points in r. For n equidistant gridpoints we obtain

$$\rho_p = \frac{p}{n+1}, \qquad r_p = \frac{bp}{n-p+1} \qquad \text{for} \quad p = 1, 2, ..., n \tag{24}$$

with the boundary points

$$\rho_0 = r_0 = 0,$$

$$\rho_{n+1} = 1, \qquad r_{n+1} = \infty$$

and the step-size

$$h = \frac{1}{n+1}. \tag{25}$$

The resulting Jacobi determinant of the coordinate transformation gives

$$dr = \frac{b}{(1-\rho)^2}\, d\rho. \tag{26}$$

It is useful to define transformed radial functions in the new rational coordinate as

$$Q_i(\rho) = \frac{b^{1/2}}{(1-\rho)}\, P_i(\rho). \tag{27}$$

Using this and the volume element, Eq. (26), we obtain for the matrix elements of an arbitrary operator $\mathcal{O}$ the form

$$\begin{aligned} \langle P_i |\, \mathcal{O} \,| P_j \rangle_r &= \left\langle \frac{1-\rho}{b^{1/2}} Q_i \middle| \frac{b}{(1-\rho)^2}\, \mathcal{O} \middle| \frac{1-\rho}{b^{1/2}} Q_j \right\rangle_\rho \\ &= \langle Q_i |\, \frac{1}{1-\rho}\, \mathcal{O}(1-\rho) \,| Q_j \rangle_\rho. \end{aligned} \tag{28}$$

For the required normalization, two- and one-electron integrals, we obtain with the transformed shell-functions

$$\langle P_i | P_j \rangle_r = \langle Q_i | Q_j \rangle_\rho, \tag{29}$$

$$\langle P_i |\, U_{kl\nu}(r) \,| P_j \rangle_r = \langle Q_i |\, U_{kl\nu}(\rho) \,| Q_j \rangle_\rho \tag{30}$$

with

$$U_{kl\nu}(\rho) = \langle Q_k(\rho') |\, \left(\frac{b\rho_<}{1-\rho_<} \right) \Big/ \left(\frac{b\rho_>}{1-\rho_>} \right)^{\nu+1} | Q_l(\rho') \rangle_{\rho'} \tag{31}$$

and

$$\langle P_i |\, \mathscr{h}(r) \,| P_j \rangle_r = \langle Q_i |\, \mathscr{h}(\rho) \,| Q_j \rangle_\rho \tag{32}$$

with

$$\mathscr{h}(\rho) = -\frac{(1-\rho)^2}{2b^2} \frac{d^2}{d\rho^2} (1-\rho)^2 - \frac{Z(1-\rho)}{b\rho} + \frac{\lambda(\lambda+1)(1-\rho^2)}{2b^2\rho^2} \tag{33}$$

B. *Discretization of the One-Electron Operator*

Using the discretization all functions become vectors of length n and the Fock equations turn into a set of nonlinear algebraic equations, with the operators in general banded $n \times n$ matrices. The Coulomb and centrifugal potential parts in $h(\rho)$ Eq. (33), become diagonal matrices, and for the kinetic energy part we choose a five-point formula to approximate the second derivative. Using $f(\rho) = (1-\rho)^2 Q(\rho$ we get

$$\frac{d^2}{d\rho^2} f(\rho_p) = \frac{1}{12h^2} [-f(p-2) + 16f(p-1) - 30f(p) \\ + 16f(p+1) - f(p+2)] + \mathcal{O}(h^4). \tag{34}$$

This formula applies directly for $p > 2$ and also for $p = 2$ because we have $f(0) = 0$. The points at the large ρ boundary $p = n-1$ and $p = n$ require no special attention, since for $p = n+1$ the function and all its derivatives are zero. Only the point $p = 1$ needs to be considered separately, if the order of the approximation is to be retained through h^4. One possibility would be to use for $f''(\rho_1)$ a forward difference formula of $\mathcal{O}(h^4)$. This would lead to a matrix Δ for the $d^2/d\rho^2$ operator, which is pentadiagonal except in the first row. It is then appropriate to use the symmetric form $(\Delta + \Delta^T)/2$ because of the hermiticity of the problem considered here. We, however, decided to use a different approach, which retains the symmetry and pentadiagonal structure of Δ. For this we use the asymptotic behavior of the radial functions Q at the origin

$$\frac{d^{\lambda+2}}{d\rho^{\lambda+2}} Q(0) = \frac{\lambda+2}{\lambda+1} [(\lambda+1(\lambda+2) - Zb] \frac{d^{\lambda+1}}{d\rho^{\lambda+1}} Q(0) \tag{35}$$

derived in Appendix 1. This leads, see Appendix 2, to the approximation

$$\frac{d^2}{d\rho^2} f(\rho_1) = \frac{1}{12h^2} [(a_\lambda - 30) f_{(1)} + 16 f_{(2)} - f_{(3)}] \tag{36}$$

with

$$a_\lambda = (-1)^\lambda \frac{(1+h)^2}{(1-h)^2} \frac{1 - h[(\lambda+2) - bZ/(\lambda+1)]}{1 + h[(\lambda+2) - bZ/(\lambda+1)]}. \tag{37}$$

For all integrations a McLaurin corrected trapezoidal rule is used, which is for analytic functions of fourth order

$$\int^1 f(\rho)\, d\rho = h \sum_{p=1}^{n} f(\rho_p) + \frac{h^2}{12} f'(0) - f'(1) + \mathcal{O}(h^4). \tag{38}$$

Here the derivatives $f'(0)$ fand $f'(1)$ vanish, since we have

$$Q(0) = Q(1) = 0$$

and

$$(\mathscr{k}Q)(0)=(\mathscr{k}Q)(1)=0$$

thus

$$\frac{d}{d\rho}\{Q\mathscr{k}Q\}_0=\frac{d}{d\rho}\{Q\mathscr{k}Q\}_1=0.$$

C. *The Two-Electron Potentials*

For the calculation of the two-electron potentials it is appropriate to derive and solve a differential equation. In the rational coordinate ρ we have

$$\begin{aligned}U_{kl\nu}(\rho)&=\int_0^\rho Q_k(\tau)\frac{(b\tau)^\nu(1-\rho)^{\nu+1}}{(b\rho)^{\nu+1}(1-\tau)^\nu}Q_l(\tau)\,d\tau\\&\quad+\int_\rho^1 Q_k(\tau)\frac{(b\rho)^\nu(1-\tau)^{\nu+1}}{(b\tau)^{\nu+1}(1-\rho)^\nu}Q_l(\tau)\,d\tau\\&=\frac{1}{b\rho}Y_{kl\nu}(\rho)\end{aligned}\tag{39}$$

with

$$\begin{aligned}Y_{kl\nu}(\rho)&=\int_0^\rho Q_k(\tau)\frac{\tau^\nu(1-\rho)^{\nu+1}}{\rho^\nu(1-\tau)^\nu}Q_l(\tau)\,d\tau\\&\quad+\int_\rho^1 Q_k(\tau)\frac{\rho^{\nu+1}(1-\tau)^{\nu+1}}{\tau^{\nu+1}(1-\rho)^\nu}Q_l(\tau)\,d\tau.\end{aligned}\tag{40}$$

Defining

$$Z_{kl\nu}(\rho)=\int_0^\rho Q_k(\tau)\frac{\tau^\nu(1-\rho)^{\nu+1}}{\rho^\nu(1-\tau)^\nu}Q_l(\tau)\,d\tau\tag{41}$$

and taking the derivatives of $Y_{kl\nu}$ and $Z_{kl\nu}$ with respect to ρ, leads to an initial value problem of two first order differential equations [1] of the form

$$\frac{d}{d\rho}Z_{kl\nu}(\rho)=-\frac{\nu+\rho}{\rho(1-\rho)}Z_{kl\nu}(\rho)+(1-\rho)\,Q_k(\rho)\,Q_l(\rho)\tag{42a}$$

and

$$\frac{d}{d\rho}Y_{kl\nu}(\rho)=\frac{\nu+1-\rho}{\rho(1-\rho)}Y_{kl\nu}(\rho)-\frac{2\nu+1}{\rho(1-\rho)}Z_{kl\nu}(P)\tag{42b}$$

with

$$Z_{kl\nu}(0) = Y_{kl\nu}(0) = 0$$

and

$$Z_{kl\nu}(1) = Y_{kl\nu}(1) = 0.$$

Taking the derivative of Eq. (42b) and substituting $Z_{kl\nu}$ gives an equivalent two-point boundary value problem

$$\frac{d^2}{d\rho^2} Y_{kl\nu}(\rho) = \frac{\nu(\nu+1)}{\rho^2(1-\rho)^2} Y_{kl\nu}(\rho) - \frac{2\nu+1}{\rho} Q_k(\rho)\, Q_l(\rho) \tag{43}$$

with

$$Y_{kl\nu}(0) = Y_{kl\nu}(1) = 0.$$

Because of the special form and stability of the differential equation (43) it is far more efficient and accurate to solve the boundary value problem. Since first derivatives are absent in Eq. (43) we can use Numerov's method [12]. Defining $f_p = \nu(\nu+1)/[\rho_p(1-\rho_p)]^2$, $g_p = Q_k(\rho_p)\, Q_l(\rho_p)$, and $Y_p = Y_{kl\nu}(\rho_p)$ we get

$$\begin{aligned}(2\nu+1)^{-1}&[(f_{p-1}/12 - 1/h^2)\, Y_{p-1} + (10 f_p/12 + 2/h^2)\, Y_p + (f_{p+1}/12 - 1/h^2)\, Y_{p+1}] \\ &= (g_{p-1} + 10 g_p + g_{p+1})/12 + \mathcal{O}(h^4)\end{aligned} \tag{44}$$

with the boundary values

$$f_0 y_0 = f_{n+1} y_{n+1} = y_0 = y_{n+1} = g_0 = g_{n+1} = 0.$$

To compute $y_1, y_2, \ldots, y_n$ a tridiagonal system

$$\begin{bmatrix} a_1 & b_2 & & 0 \\ b_1 & a_1 & & \\ & & a_{n-1} & b_n \\ 0 & & b_{n-1} & a_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \tag{45}$$

must be solved with

$$a_p = (10 h^2 f_p/12 + 2)/(2\nu+1),$$

$$b_p = (h^2 f_p/12 - 1)/(2\nu+1) \qquad \text{for} \quad p = 1, \ldots, n$$

and

$$d_p = (g_{p-1} + 10 g_p + g_{p+1})\, h^2/12.$$

We note that the coefficient matrix does not depend on the radial functions. Hence it is possible to perform the LU decomposition of the tridiagonal matrices for all ν-values only once before starting the integration. The computation of the two-electron potentials reduces then just to the back-substitution part of solving the linear system.

D. *Estimation of the Numerical Error*

We have been careful to implement all procedures for the evaluation of the energy with a truncation error of $\mathcal{O}(h^4)$. The error of $\mathcal{O}(h^2)$ in the determination of ∇E at the origin does not spoil this, see Appendix 2. Thus an extrapolation [4] is possible once calculations of the same system have been performed with different step-sizes using

$$E_{\text{ex}}(0) = E(h) + Ah^t \tag{46}$$

or even the ε-algorithm [13]. In our implementation t should be between 3 and 4; our experience indicates $t \cong 3.8$. An estimate of the numerical error is given by

$$\delta(E) = \frac{E_{\text{ex}} - E(h)}{E_{\text{ex}}}; \tag{47}$$

thus we have in this numerical process a rigorous control on the accuracy.

V. The Starting Functions

The shell-functions to start the MCSCF iterations can be obtained frequently from some prior calculations of a similar problem. If no such results are available we obtain them by integrating the atomic system using the model-potential suggested by Garvey, Jackman, and Green [14]. This requires simply the numerical integration of the one-dimensional differential eigenvalue equation

$$\left[-1/2 \frac{d^2}{dr^2} + (VG_i(r) - \varepsilon_i) \right] P_i(r) = 0. \tag{48}$$

With $VG_i(r)$ the model potential for the shell-function P_i as suggested by Green *et al.*, i.e.,

$$VG_i(r) = -\frac{z}{r} + \frac{\lambda(\lambda+1)}{2r^2} + (N_e - 1)\{1 - [\eta(e^{-\zeta r} - 1)/ + 1]^{-1}\}/r \tag{49}$$

with N_e the number of electrons of the atom and the constants η and ζ, which have been chosen to minimize the energy tabulated [14].

The integration is carried out again using the rational coordinate ρ as described above and Numerov's method. It is our experience that the shell-functions thus obtained approximate the standard H–F shell-functions rather well.

TABLE I

Atom	Configurations	State	n	$-E$	$-E/T$	$\|\nabla E\|$	ΔE
H^-	$1s^2$	$^1S(1s^2)$	80	0.4879579635	1.000286	—	
			160	0.4879319646	1.000064	0.395×10^{-4}	
			ext.	0.4879302314	—	—	0.36×10^{-5}
	$1s^2+2s^2+3s^2+4s^2$	$^1S(1s^2)$	80	0.5274973468	1.000058	0.202×10^{-4}	
	$+2p^2+3p^2+4p^2$		160	0.5274902706	1.000007	0.872×10^{-5}	
	$+3d^2+4d^2+4f^2$		ext	0.5274897989			0.89×10^{-6}
He	$1s^2$	$^1S(1s^2)$	100	2.861690957	1.000027	0.281×10^{-5}	
			200	2.861680759	1.000002	0.166×10^{-5}	
			ext	2.861680079			0.24×10^{-6}
	$1s^2+2s^2+2p^2$	$^1S(1s^2)$	80	2.898171440	1.000219	0.172×10^{-4}	
			160	2.898036676	1.000024	0.606×10^{-5}	
			ext	2.898027689			0.31×10^{-5}
	$1s^2+2s^2+3s^2+4s^2$	$^1S(1s^2)$	80	2.902936625	1.000109	0.223×10^{-4}	
	$+2p^2+3p^2+4p^2$		160	2.902911247	1.000027	0.224×10^{-5}	
	$+3d^2+4d^2+4f^2$		ext	2.902909555			0.58×10^{-6}
	$1s^2+1s2s+2s^2+1s3s$	$^1S(1s^2)$	90	2.900069056	1.000036	0.203×10^{-4}	
	$+2s3s+3s^2+2p^2$		180	2.900052557	1.000005	0.189×10^{-2}	
	$+2p3p+3p^2$		ext	2.900051457			0.38×10^{-6}

Continued

TABLE 1—*Continued*

Atom	Configurations	State	n	$-E$	$-E/T$	$\Vert\nabla E\Vert$	ΔE
		$^1S(1s2s)$	90	2.145908580	1.001111	0.184×10^{-2}	
			180	2.145624240	1.000058	0.114×10^{-1}	
			ext	2.145605284			0.88×10^{-5}
		$^1S(1s3s)$	180	2.061181799	1.001017	0.250×10^{-2}	
			360	2.060989641	1.000005	0.674×10^{-4}	
			ext	2.060976831			0.62×10^{-5}
	$1s2p+1s4p+2s3p+5p3d$	$^1P(1s2p)$	120	2.122612836	1.000025	0.636×10^{-4}	
			240	2.122589126	0.999974	0.211×10^{-3}	
			ext	2.122587545			0.74×10^{-6}
	$1s2p+1s3p+2s4p+5p3d$	$^1P(1s3p)$	120	2.054819734	1.000099	0.340×10^{0}	
			240	2.054770354	0.999991	0.184×10^{-4}	
			ext	2.054767062			0.16×10^{-5}
Ne	$1s^2 2s^2 2p^6$	^1S	80	128.5486032	1.000128	0.309×10^{-3}	
			160	128.5472151	1.000090	0.660×10^{-3}	
			ext	128.5471226			0.72×10^{-6}
	$2p^5(3p+4p+5p)$	$^3D(3p)$	240	127.9260491	1.000036	0.943×10^{-2}	0.47×10^{-7}
		$^3D(4p)$	240	127.8678044	1.000033	0.394×10^{-2}	0.42×10^{-6}
		$^3D(5p)$	240	127.8487913	1.000095	0.371×10^{-2}	0.60×10^{-5}
Hg	$6s^2$	^1S	320	18408.99024	1.000047	0.883×10^{-3}	
	$6s6p$	^3P	320	18408.91345	1.000085	0.421×10^{-2}	
	$6s6p$	^1P	320	18408.81360	1.000082	0.941×10^{-1}	
	$6s7s$	^1S	320	18408.81349	1.000115	0.203×10^{-1}	
	$6s7p$	^1P	320	18408.79167	1.000114	0.169×10^{-2}	
	$6s6d$	^1D	320	18408.79136	1.000117	0.617×10^{-2}	
	Hg^+6s	^2S	320	18408.66667	1.000115	0.881×10^{-3}	

To obtain starting shell-functions for purely correlating orbitals, which are not present in the standard H–F configuration, we again use Eq. (48) with the Green potential VG_i modified, such as to yield orbitals with the appropriate node structure but sufficiently localized in the space where the shells to be correlated are localized.

VI. Some Examples and Conclusion

We conclude the description of our numerical MCSCF program for atoms by presenting a few selected results of calculations for ground and excited states. These are not to represent systematic and defintive studies but just to indicate the workings and accuracy of our program. The results are collected in Table I. Inspection of these results reveals a few points:

(i) To obtain well correlated wavefunctions a large number of configurations is necessary, this holds in particular if excited states are desired.

(ii) Results obtained with just one grid-size should be used with caution; as the Richardson extrapolation indicates, the finite difference calculation with too few points may feign too good a result.

We have described and implemented a numerical MCSCF program of great generality and a few novel numerical aspects. The effectiveness of the program has been demonstrated. To be sure, for calculations with extensive configuration interaction, possible with this program, the calculation of the structure factors not detailed here, needs to be and has been automated also.

We believe this program represents a valuable alternative to the well documented and widely used programs of C. Froese–Fischer [2] with the following advantageous characteristics:

(a) Any combination of configuration can be included into the MCSCF functions.

(b) the SCF convergence is improved through the automatic and optimal determination of the damping factor λ.

(c) Even in the most difficult cases convergence can be obtained using the quadratically convergent nonlinear optimization, though this is computationally expensive.

Appendix 1: The Radial Functions near the Origin

To investigate the behavior of Q_i near the origin we use the power series expansion

$$Q(\rho) = \sum_{i=0}^{\infty} a_i \rho^{i+\alpha}. \tag{A1}$$

Substituting this into the Fock equation, Eq. (12), yields

$$\frac{1}{b^2}(1-\rho)\left((1-\rho)\frac{d^2}{d\rho^2}(1-\rho)^2+\frac{2bZ}{\rho}-\lambda(\lambda+1)\frac{(1-\rho)}{\rho^2}\right)\sum_{i=0} a_i\rho^{i+\alpha}=\mathcal{O}(\rho^\alpha) \tag{A2}$$

with $\lambda=\lambda(i)$ the l-value of the shell Q_i studied.

Rearranging gives

$$\begin{aligned} &a_0[\alpha(\alpha-1)-\lambda(\lambda+1)]\,\rho^{\alpha-2}\\ &\quad+\{2bZ+a_0[\lambda(\lambda+1)-\lambda(\lambda-1)-2\lambda(\lambda+1)]\\ &\quad+a_1[\alpha(\alpha+1)-\lambda(\lambda+1)]\}\,\rho^{\alpha-1}=\mathcal{O}(\rho^\alpha) \end{aligned} \tag{A3}$$

and equating the coefficients of $\rho^{\alpha-2}$ and $\rho^{\alpha-1}$ equal to zero results in

$$\alpha=\lambda+1 \tag{A4}$$

and

$$a_1=a_0[(\lambda+1)(\lambda+2)-bZ]/(\lambda+1). \tag{A5}$$

Hence we obtain

$$\frac{d^{\lambda+2}}{d\rho^{\lambda+2}}Q(0)=\frac{\lambda+2}{\lambda+1}[(\lambda+1)(\lambda+2)-bZ]\frac{d^{\lambda+1}}{d\rho^{\lambda+1}}Q(0) \tag{A6}$$

with the derivatives of order less than $(\lambda+1)$ all zero.

APPENDIX 2. DISCRETIZATION AT THE BOUNDARY $\rho=0$

For the five-point difference formula, Eq. (34) at the point $Q(h)$ we need the fictitious point $Q(-h)$. To this end we expand the radial function Q with angular quantum number λ around $\rho=0$ and obtain

$$Q(h)=Q(0)+\frac{h^{\lambda+1}}{(\lambda+1)!}Q(0)^{(\lambda+1)}+\frac{h^{\lambda+2}}{(\lambda+2)!}Q(0)^{(\lambda+2)}+\mathcal{O}(h^{\lambda+3}) \tag{A7}$$

and

$$(-1)^\lambda Q(-h)=(-1)^\lambda Q(0)-\frac{h^{\lambda+1}}{(\lambda+1)!}Q(0)^{(\lambda+1)}+\frac{h^{\lambda+2}}{(\lambda+2)!}Q(0)^{(\lambda+2)}+\mathcal{O}(h^{\lambda+3})$$

with the notation $Q(0)^{(\lambda+1)}$ the $(\lambda+1)$ th derivative of Q with respect to ρ at $\rho=0$. Adding and subtracting the Eqs. (A7) gives

$$Q(0)^{(\lambda+1)}=\frac{(\lambda+1)!}{2h^{\lambda+1}}[Q(h)-(-1)^\lambda Q(-h)]+\mathcal{O}(h^2) \tag{A8}$$

and

$$Q(0)^{(\lambda+2)}=\frac{(\lambda+2)!}{2h^{\lambda+2}}\,[Q(h)+(-1)^{\lambda}\,Q(-h)]+\mathcal{O}(h^2).$$

Combining this with Eq. (A6) yields

$$Q(-h)=\frac{h[(\lambda+2)-bZ/(\lambda+1)]-1}{h[(\lambda+2)-bZ/(\lambda+1)]+1}\,(-1)^{\lambda}\,Q(h)+\mathcal{O}(h^{\lambda+4}). \qquad \text{(A9)}$$

The element for $p=1$ required in Eq. (34) is

$$\begin{aligned} f(-1)&=-(1+h)^2\,Q(-1)\\ &=-(1+h)^2\frac{h[(\lambda+2)-bZ/(\lambda+1)]-1}{h[(\lambda+2)-bZ/(\lambda+1)]+1}\,(-1)\,Q(1)+\mathcal{O}(h^{\lambda+4})\\ &=-\frac{(1+h)^2}{(1-h)^2}\frac{h[(\lambda+2)-bZ/(\lambda+1)]-1}{h[(\lambda+2)-bZ/(\lambda+1)]+1}\,(-1)^{\lambda}\,f(1)+\mathcal{O}(h^{\lambda+4}). \qquad \text{(A10)}\end{aligned}$$

As this enters Eq. (34) multiplied with $1/12h^2$ to yield Eq. (36) the error term will be $\mathcal{O}(h^{\lambda+2})$. However, in an integral this single point will always be multiplied with $Q(1)$, which itself is $\mathcal{O}(h^{\lambda+1})$ to give an error of order $\mathcal{O}(h^{2\lambda+3})$ multiplied by h, thus maintaining the total error term of $\mathcal{O}(h^4)$.

ACKNOWLEDGMENTS

We are grateful for the financial support of the "Fonds der Chemischen Industrie" and the "Deutsche Forschungsgemeinschaft" (DFG), through the SFB-Grant 216 "Polarization and Correlation in Atomic Collision Complexes."

REFERENCES

1. C. Froese-Fischer, *The Hartree Fock Method for Atoms* (Wiley, New York, 1977).
2. D. R. Hartree, *The Calculation of Atomic Structures* (Wiley, New York, 1957); F. Herman and S. Skillman, *Atomic Structure Calculation* (Prentice–Hall, Englewood Cliffs, NJ, 1963); C. Froese-Fischer, *Comput. Phys. Commun.* **2**, 124 (1971), *J. Comput. Phys.* **10**, 211 (1972); **27**, 221 (1978).
3. L. Collatz, *The Numerical Treatment of Differential Equations* (Springer, New York, 1966).
4. L. F. Richardson, *Philos. Trans. Roy. Soc. London Ser. A* **210**, 307 (1960).
5. H. F. Werner and W. Meyer, *J. Chem. Phys.* **73**, 2342 (1980); **74**, 5794 (1981), and others.
6. E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra* (Cambridge Univ. Press, Cambridge, 1970).
7. K. K. Docken and J. Hinze, *J. Chem. Phys.* **57**, 4928 (1972).
8. J. Hinze, *J. Chem. Phys.* **59**, 6424 (1973).
9. G. R. Walsh, *Methods of Optimization* (Wiley, New York, 1975).

10. G. DAS AND A. C. WAHL, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Plenum, New York, 1977), p. 79.
11. L. A. HAGEMAN AND D. M. YOUNG, *Applied Iterative Methods* (Academic Press, New York, 1981).
12. B. NUMEROV, *Publ. Observatoire Central Astrophys. Russ.* **2**, 188 (1933).
13. G. A. BAKER AND P. R. GRAVES-MORRIS, "Padé Approximation Basic Theory," Part I, Vol. 13, *Encyclopedia of Mathematics and its Applications*, edited by G. C. Rota (Addison-Wesley, New York, 1981).
14. R. H. GARVEY, C. H. JACKMAN, AND A. E. S. GREEN, *Phys. Rev. A* **12**, 1144 (1975).