MATHEMATICS OF COMPUTATION, VOLUME 28, NUMBER 127, JULY 1974, PAGES 731-741

The Solution of Systems of Equations Using the ϵ-Algorithm, and an Application to Boundary-Value Problems

By Claude Brezinski and Alain C. Rieu

Abstract. In this paper, the authors describe the properties of an algorithm to solve systems of nonlinear equations. The algorithm does not use any derivatives. The convergence of this algorithm is quadratic under quite mild conditions. This method can also be used to solve systems of linear equations with infinitely many solutions. The second part of the paper is devoted to the application of this algorithm to the solution of multipoint boundary-value problems for differential equations. A theorem of convergence is proved and various numerical examples are given.

1. Introduction. One of the simpler ideas for solving a multipoint boundary-value problem associated with a system of p simultaneous first order differential equations in p dependent variables is to reduce the given problem to an initial-value problem. If $\mathbf{v}$ is a vector of initial values, the boundary conditions may be written in the form $\mathbf{v} = F(\mathbf{v})$. In this paper, we describe a quadratically convergent iterative process, based upon repeated use of the vector ε-algorithm of Wynn [18], for solving this equation and hence solving the given boundary-value problem.

In the following section, we briefly review some known results concerning the use of the vector ε-algorithm in the solution of systems of equations and, by application of a theorem of McLeod [11], show how this algorithm may under certain conditions be used to determine solutions of a set of linear equations $\mathbf{Bs} = \mathbf{b}$ when $\mathbf{B}$ is a real singular square matrix. The use of the vector ε-algorithm to solve systems of the form $(\mathbf{I} - \mathbf{A})\mathbf{s} = \mathbf{b}$ where $\mathbf{A}$ is a real singular square matrix is also studied. This result is in turn used to generalise an algorithm, recently and independently discovered by Brezinski [2],[3] and Gekeler [6], for solving general systems of p equations in p unknowns.

The application of the algorithm to nonlinear multipoint boundary-value problems is studied in the last section, in which we also give some numerical examples.

2. The Solution of Systems of Equations.

2.1 *The Linear Case.* We set the theory in $\mathbf{R}^p$ $(1 \leq p < \infty)$. The vector ε-algorithm is a nonlinear sequence to sequence transformation: vectors $\{\boldsymbol{\varepsilon}_k^{(q)}\}$ are constructed from the initial values

$$\boldsymbol{\varepsilon}_{-1}^{(q)} = \mathbf{0} \qquad (q = 1, 2, \ldots), \quad \boldsymbol{\varepsilon}_0^{(q)} = \mathbf{s}_q \qquad (q = 0, 1, \ldots),$$

Received January 12, 1973, revised May 29, 1973.

AMS (MOS) subject classifications (1970). Primary 65L10, 65N35, 65B05, 65H10.

Key words and phrases. Solution of systems of equations, algorithm with quadratic convergence, boundary-value problem, differential equations.

$\{\mathbf{s}_q\}$ being a prescribed sequence of vectors in $\mathbf{R}^p$, by use of the relationships

$$\boldsymbol{\varepsilon}_{k+1}^{(q)} = \boldsymbol{\varepsilon}_{k-1}^{(q+1)} + (\boldsymbol{\varepsilon}_k^{(q+1)} - \boldsymbol{\varepsilon}_k^{(q)})^{-1} \qquad (k, q = 0, 1, \ldots),$$

where the inverse of the nonzero vector $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ is defined to be $\mathbf{x}^{-1} = \{\sum_{i=1}^{p} x_i^2\}^{-1}\mathbf{x}$.

It was shown by J. B. McLeod [11] that if the vectors $\{\mathbf{s}_q\}$ satisfy a linear recursion of the form

$$(1) \qquad \sum_{\nu=0}^{m} c_\nu \mathbf{s}_{q+\nu} = \left\{\sum_{\nu=0}^{m} c_\nu\right\}\mathbf{s} \qquad (q = 0, 1, \ldots),$$

where the $\{c_\nu\}$ are real numbers with $\sum_{\nu=0}^{m} c_\nu \neq 0$, $\mathbf{s}$ is a constant vector, and the vectors $\{\boldsymbol{\varepsilon}_{2m}^{(q)}\}$ can be constructed in the sense that at no stage in the determination of these vectors $\boldsymbol{\varepsilon}_k^{(q+1)} = \boldsymbol{\varepsilon}_k^{(q)}$, then $\boldsymbol{\varepsilon}_{2m}^{(q)} = \mathbf{s}$ $(q = 0, 1 \cdots)$ identically. (If, in this case, $\lim_{q=\infty} \mathbf{s}_q$ exists, its value is, of course, $\mathbf{s}$.)

The above algorithm is also connected [19] with the solution of a set of linear equations, expressed in the form

$$(2) \qquad \mathbf{Bs} = \mathbf{b}$$

where $\mathbf{B}$ is a real $p \times p$ matrix and $\mathbf{b} \in \mathbf{R}^p$. We set $\mathbf{B} = \mathbf{I} - \mathbf{A}$ so that $\mathbf{s} = \mathbf{As} + \mathbf{b}$, prescribe $\mathbf{s}_0$, and determine the sequence $\{\mathbf{s}_q\}$ by use of the recursion

$$(3) \qquad \mathbf{s}_{q+1} = \mathbf{A}\mathbf{s}_q + \mathbf{b} \qquad (q = 0, 1, \ldots).$$

It is easily verified that

$$(4) \qquad \mathbf{s}_q = \mathbf{A}^q\mathbf{s}_0 + \sum_{i=0}^{q-1} \mathbf{A}^i\mathbf{b} \qquad (q = 0, 1, \ldots).$$

If $\mathbf{I} - \mathbf{A} = \mathbf{B}$ is nonsingular,

$$(5) \qquad \mathbf{s}_q = \mathbf{A}^q\{\mathbf{s}_0 - \mathbf{B}^{-1}\mathbf{b}\} + \mathbf{B}^{-1}\mathbf{b} \qquad (q = 0, 1, \ldots).$$

If $\sum_{\nu=0}^{m} c_\nu \lambda^\nu$ is the minimal polynomial of $\mathbf{A}$ with respect to the vector $\mathbf{s}_0 - \mathbf{B}^{-1}\mathbf{b}$, then

$$\left\{\sum_{\nu=0}^{m} c_\nu \mathbf{A}^\nu\right\}\{\mathbf{s}_0 - \mathbf{B}^{-1}\mathbf{b}\} = 0 \qquad (q = 0, 1, \ldots),$$

i.e.,

$$(6) \qquad \sum_{\nu=0}^{m} c_\nu\{\mathbf{s}_{q+\nu} - \mathbf{B}^{-1}\mathbf{b}\} = 0 \qquad (q = 0, 1, \ldots).$$

Setting $\mathbf{s} = \mathbf{B}^{-1}\mathbf{b}$, we see that recursions (1) and (6) are equivalent. Hence, if the vectors $\{\mathbf{s}_q\}$ are produced either by direct use of recursion (3), or by an equivalent process, and the vectors $\{\boldsymbol{\varepsilon}_{2m}^{(q)}\}$ can be constructed, then these vectors (in particular, the first one $\boldsymbol{\varepsilon}_{2m}^{(0)}$) are equal to the solution of Eq. (2).

Our first contribution is to study the case where the matrix $\mathbf{A}$ is singular; let τ be the multiplicity of the root $\lambda = 0$ for the minimal polynomial of $\mathbf{A}$ with respect to the vector $\mathbf{s}_0 - \mathbf{B}^{-1}\mathbf{b}$ (it is possible that $\tau = 0$, of course). In this case, the coefficients c_ν $(\nu = 0, 1, \cdots, \tau - 1)$ of this minimal polynomial are zero. Formula (6) yields the reduced recursion

$$\sum_{\nu=0}^{m-\tau} \hat{c}_\nu \mathbf{s}_{\tau+q+\nu} = \left\{\sum_{\nu=0}^{m-\tau} \hat{c}_\nu\right\}\mathbf{s} \qquad (q = 0, 1, \ldots)$$

where $\hat{c}_\nu = c_{\tau+\nu}$ $(\nu = 0, 1, \ldots, m - \tau)$, and we derive the result that if no breakdown occurs in their construction, the vectors $\boldsymbol{\varepsilon}_{2(m-\tau)}^{(q+\tau)}$ $(q = 0, 1, \ldots)$ are all equal to the solution of Eq. (2). For the sake of completeness, we mention that if $\tau > 1$, the vector $\boldsymbol{\varepsilon}_{2m}^{(0)}$ cannot be constructed: the vectors $\boldsymbol{\varepsilon}_{2(m-\tau)}^{(q)}$ $(q = \tau, \tau + 1, \ldots, 2\tau)$ are equal, the vectors $\boldsymbol{\varepsilon}_{2(m-\tau)+1}^{(q)}$ $(q = \tau, \tau + 1, \ldots, 2\tau - 1)$ are infinite, and thereafter the computations required for the construction of the vector $\boldsymbol{\varepsilon}_{2m}^{(0)}$ break down. If $\tau = 1$, we have $\boldsymbol{\varepsilon}_{2m}^{(0)} = \boldsymbol{\varepsilon}_{2m-2}^{(1)} = \mathbf{s}$ as required.

Our second contribution is to show that even when the matrix $\mathbf{B}$ of Eq. (2) is singular, and if a solution exists, then under certain conditions the vector ε-algorithm may be used to construct solutions.

THEOREM 1. *Let* (2) *be a set of linear algebraic equations in* $\mathbf{R}^p$ *and possess a solution* $\mathbf{s} \in \mathbf{R}^p$. *Having prescribed the initial vector* $\mathbf{s}_0 \in \mathbf{R}^p$, *let the minimal polynomial of* $\mathbf{B}$ *with respect to the vector* $\mathbf{s}_0 - \mathbf{s}$ *be of degree* m *and possess the roots* $\lambda = 0$ *with multiplicity* 1 *and* $\lambda = 1$ *with multiplicity* τ; *let vectors* $\{\mathbf{s}_q\}$ *be constructed by means of recursion* (3), *where* $\mathbf{A} = \mathbf{I} - \mathbf{B}$.

If vectors $\{\boldsymbol{\varepsilon}_{2(m-\tau)-2}^{(q+\tau)}\}$ *can be constructed by application of the* ε-*algorithm to the sequence* $\{\mathbf{s}_q\}$, *then* $\boldsymbol{\varepsilon}_{2(m-\tau)-2}^{(q+\tau)} = \boldsymbol{\epsilon}$ $(q = 0, 1, \cdots)$ *where* $\boldsymbol{\varepsilon}$ *is a solution of Eq.* (2).

Proof. We deal first with the case $\tau = 0$.

Formula (4) again holds. Since $\mathbf{B}$ is singular, formula (5) cannot be used; nevertheless, a solution $\mathbf{s}$ exists; we have $\mathbf{b} = (\mathbf{I} - \mathbf{A})\mathbf{s}$ and, from formula (4),

$$\mathbf{s}_q = \mathbf{A}^q\mathbf{s}_0 + \mathbf{s} - \mathbf{A}^q\mathbf{s} \qquad (q = 0, 1, \ldots),$$

$$\mathbf{s}_q - \mathbf{s}_{q+1} = (\mathbf{I} - \mathbf{A})\mathbf{A}^q(\mathbf{s}_0 - \mathbf{s}) \qquad (q = 0, 1, \ldots).$$

Under the stated assumptions concerning $\mathbf{B}$, the minimal polynomial of the matrix $\mathbf{A}$ with respect to the vector $\mathbf{s}_0 - \mathbf{s}$ has the form $\sum_{\nu=0}^m c'_\nu \lambda^\nu$, where

$$\sum_{\nu=0}^{m} c'_\nu \lambda^\nu = (1 - \lambda) \sum_{\nu=0}^{m-1} \gamma_\nu \lambda^\nu \qquad \left(\gamma = \sum_{\nu=0}^{m-1} \gamma_\nu \neq 0\right).$$

Hence

$$\sum_{\nu=0}^{m-1} \gamma_\nu \mathbf{s}_{q+\nu} - \sum_{\nu=0}^{m-1} \gamma_\nu \mathbf{s}_{q+\nu+1} = \left\{(\mathbf{I} - \mathbf{A})\mathbf{A}^q \sum_{\nu=0}^{m-1} \gamma_\nu \mathbf{A}^\nu\right\}(\mathbf{s}_0 - \mathbf{s}) = \mathbf{0} \qquad (q = 0, 1, \ldots).$$

Thus the sums occurring on the left-hand side of this relationship are constant: we have

$$\sum_{\nu=0}^{m-1} \gamma_\nu \mathbf{s}_{q+\nu} = \gamma \boldsymbol{\varepsilon} \qquad (q = 0, 1, \ldots) \tag{7}$$

say. The fundamental result for the vector ε-algorithm quoted above tells us that if the vectors concerned can be computed, then $\boldsymbol{\varepsilon}_{2m-2}^{(q)} = \boldsymbol{\varepsilon}$ $(q = 0, 1, \ldots)$.

Furthermore, setting $q = 0$ in relationship (7),

$$(\mathbf{I}-\mathbf{A})\boldsymbol{\varepsilon} = \gamma^{-1}(\mathbf{I}-\mathbf{A}) \sum_{\nu=0}^{m-1} \gamma_\nu\{\mathbf{A}^\nu(\mathbf{s}_0-\mathbf{s})+\mathbf{s}\} = (\mathbf{I}-\mathbf{A})\mathbf{s} = \mathbf{b}.$$

Thus $\boldsymbol{\varepsilon}$ is a solution of Eq. (2) as stated. If $\tau > 0$ then Eq. (7) can be written with $m-\tau$ instead of m and for $q = \tau, \tau+1, \ldots$. Thus $\boldsymbol{\varepsilon}_{2(m-\tau)-2}^{(q+\tau)} = \boldsymbol{\varepsilon}$ $(q = 0, 1, \ldots)$ which ends the proof of the theorem.

If $\hat{\boldsymbol{\varepsilon}}$ and $\tilde{\boldsymbol{\varepsilon}}$ are two distinct vectors produced in the above manner from two initial vectors $\hat{\mathbf{s}}_0$ and $\tilde{\mathbf{s}}_0$, then $(1+\alpha)\hat{\boldsymbol{\varepsilon}} - \alpha\tilde{\boldsymbol{\varepsilon}}$ where $-\infty < \alpha < +\infty$ is also a solution of Eq. (2).

Some other results concerning the vector ε-algorithm can be found in a paper by Brezinski [4].

2.2 *The Nonlinear Case.* Let F be a mapping of $\mathbf{R}^p$ into $\mathbf{R}^p$ which is differentiable in a neighbourhood $\mathbf{D}$ of $\mathbf{v} \in \mathbf{R}^p$ such that $\mathbf{v} = F(\mathbf{v})$. We propose the following algorithm for finding $\mathbf{v}$. Select $\mathbf{v}_0 \in \mathbf{D}$. For $n = 0, 1, \ldots$:

(i) Let the matrix $\mathbf{I} - F'(\mathbf{v})$ (the prime here and later denotes the Fréchet derivative [5] with respect to the indicated argument) be nonsingular. Let the minimal polynomial of $F'(\mathbf{v})$ with respect to the vector $\mathbf{v}_n - \mathbf{v}$ be of degree m, and possess the root $\lambda = 0$ with multiplicity τ. Set $Q = m - \tau$ and $s_0 = \mathbf{v}_n$.

(ii) Compute the vectors $\mathbf{s}_q$ $(q = 1, 2, \ldots, 2Q+\tau)$ by use of the recursion $\mathbf{s}_{q+1} = F(\mathbf{s}_q)$ $(q = 0, 1, \ldots, 2Q+\tau-1)$.

(iii) Apply the vector ε-algorithm to the sequence $\mathbf{s}_q$ $(q = \tau, \tau+1, \ldots, 2Q+\tau)$ and set $\mathbf{v}_{n+1} = \boldsymbol{\varepsilon}_{2Q}^{(\tau)}$.

THEOREM 2. *If the above algorithm can be applied, the sequence* $\{\mathbf{v}_n\}$ *converges at least quadratically to the fixed point* $\mathbf{v}$ *of the mapping* F, *in the sense that* $\|\mathbf{v}_{n+1} - \mathbf{v}\| = O(\|\mathbf{v}_n - \mathbf{v}\|^2)$, *for every* $\mathbf{v}_0 \in D' \subseteq D$.

Proof. For the case in which $\tau = 0$, the above result has been proved by Gekeler [6] and Brezinski [2],[3]. The basic idea of this proof is as follows. During the computation of a sequence $\{\mathbf{s}_q\}$, we have, since F is Fréchet differentiable,

$$\mathbf{s}_{q+1} - \mathbf{v} = F'(\mathbf{v})(\mathbf{s}_q - \mathbf{v}) + O[\|\mathbf{s}_q - \mathbf{v}\|^2] \tag{8}$$

($O[x]$ denotes a vector whose norm is of the same order as the real number x). Let $\sum_{\nu=0}^m c_\nu \lambda^\nu$ be the minimal polynomial of $F'(\mathbf{v})$ with respect to $\mathbf{s}_0 - \mathbf{v}$. Then

$$\begin{aligned}\sum_{\nu=0}^{m} c_\nu(\mathbf{s}_{q+\nu} - \mathbf{v}) &= \left\{\sum_{\nu=0}^{m} c_\nu F'(\mathbf{v})^\nu\right\}(\mathbf{s}_0 - \mathbf{v}) \\ &\quad + \sum_{\nu=0}^{m} c_\nu O[\|\mathbf{s}_0 - \mathbf{v}\|^2] \qquad (q = 0, 1, \ldots, m)\end{aligned} \tag{9}$$

or, since the first sum on the right-hand side of this relationship represents the zero mapping, and $\mathbf{s}_0 = \mathbf{v}_n$,

$$\sum_{\nu=0}^{m} c_\nu(\mathbf{s}_{q+\nu} - \mathbf{v} - O[\|\mathbf{v}_n - \mathbf{v}\|^2]) = \mathbf{0}. \tag{10}$$

This condition is very similar to relationship (1) upon which McLeod's result is based. We now use the continuity of the vector ε-algorithm, and have

$$\mathbf{v}_{n+1} = \boldsymbol{\varepsilon}_{2m}^{(0)} = \mathbf{v} + O[\|\mathbf{v}_n - \mathbf{v}\|^2].$$

If at any stage $\tau > 0$ in place of $\tau = 0$, the generalization of McLeod's result introduced in the proof of Theorem 1 can be used in the above analysis.

For mappings of a very special kind, convergence of even higher order than the second can be demonstrated.

THEOREM 3. *If for all* $\mathbf{s}$ *in the neighbourhood* $\mathbf{D}$ *described above,* $F(\mathbf{s}) - \mathbf{v} - F'(\mathbf{v})(\mathbf{s} - \mathbf{v}) = O[\|\mathbf{s} - \mathbf{v}\|^a]$, *where* $a \geqq 2$ *is a positive integer, then for the vectors* $\{\mathbf{v}_n\}$ *generated by the above algorithm*

$$\|\mathbf{v}_{n+1} - \mathbf{v}\| = O(\|\mathbf{v}_n - \mathbf{v}\|^a). \tag{11}$$

If O *may be replaced by* o *in the above condition, a corresponding change may be made in formula* (11).

Proof. Replace 2 by a in formulas (8)–(10), and thereafter O by o.

Theorem 3 is a generalization of a result due to Ostrowski [12] for $p = 1$.

We offer some remarks. Concerning the way in which the algorithm is implemented:

(1) In practice, m and τ are unknown. In general, the computations of each stage are based on the assumption that $m = p$, $\tau = 0$, so that $\mathbf{v}_{n+1} = \boldsymbol{\varepsilon}_{2p}^{(0)}$. If, in fact, $m < p$ and $\tau > 0$, then the vectors $\boldsymbol{\varepsilon}_{2(m-\tau)}^{(q+\tau)}$ $(q = 0, 1, \ldots)$ are all equal (they are encountered during the attempt to construct $\boldsymbol{\varepsilon}_{2p}^{(0)}$), and we take $\mathbf{v}_{n+1} = \boldsymbol{\varepsilon}_{2(m-\tau)}^{(\tau)}$.

(2) Instead of the vector ε-algorithm, the scalar ε-algorithm might be applied to each component of the vector $\{\mathbf{s}_q\}$. However, organisation of the computation is a little easier using the vector $\boldsymbol{\varepsilon}$-algorithm, and numerical experience suggests that this mode of application is more stable.

Concerning the theory of the algorithm:

(3) If $p = 1$, the algorithm reduces to Steffensen's method, of which it may be considered to be a generalization in p dimensions, and whose properties are shared by the algorithm.

(4) It is impossible, without either calculating derivatives or inverting matrices, to construct a quadratically convergent process each of whose iterations requires less than $2m$ evaluations of F. In this sense, the proposed algorithm is optimal. Ulm [17] has proposed an extension of Steffensen's method which is of quadratic order but is not optimal in the sense that it requires more computations of F and matrix inversions. Henrici [9] has given a quadratic method which requires only $p + 1$ evaluations of F and a matrix inversion; however, numerical experience (as described by Ortega and Rheinboldt [13]) reveals that this method in unstable.

(5) If the mapping F has a unique fixed point $\mathbf{v}$, and the algorithm converges, then naturally the vector sequence produced converges to $\mathbf{v}$. If F has many fixed points (as, for example, described in Theorem 1), then the sequence may well converge to a fixed point which is not the vector $\mathbf{v}$ occurring in the description of the algorithm.

(6) F is not assumed to be a contraction mapping.

(7) Since, in the next section, we make use of Runge-Kutta procedures in the evaluation of $F(\mathbf{s}_q)$, we mention that Alt [1] has used the vector ε-algorithm to devise an implicit Runge-Kutta scheme in which the step size may be changed at will without fear of the effects of instability. We also mention a recent paper by Wynn [20] in which repeated use of the ε-algorithm is applied to integration schemes suitable for the numerical solution of differential equations, and a further paper [21] in which algorithms related to the ε-algorithm are used for the evaluation of singular and highly oscillatory integrals.

3. Application to Multipoint Boundary-Value Problems. Let the components of the vector $\mathbf{y}(t) = (y_1(t), y_2(t), \ldots, y_p(t)) \in \mathbf{R}^p$ be functions of the real variable t, those of $f\{t, \mathbf{y}(t)\} \in \mathbf{R}^p$ be functions of t and possibly all components of $\mathbf{y}(t)$, and those of $h_i\{\mathbf{y}(t_i)\} \in \mathbf{R}^p$ be functions of the fixed numbers $y_1(t_i), \ldots, y_p(t_i)$ ($i = 1, 2, \ldots, r < \infty$) where, without loss of generality, we take $t_1, t_2, \ldots, t_r$ to be an increasing sequence of real numbers. We consider the multipoint boundary-value problem

$$d\mathbf{y}(t)/dt = f\{t, \mathbf{y}(t)\}, \tag{12}$$

$$\sum_{i=1}^{r} h_i\{\mathbf{y}(t_i)\} = \mathbf{c}, \tag{13}$$

where $\mathbf{c} \in \mathbf{R}^p$ is constant. Define $\mathbf{y}(t; \mathbf{u})$ to be the solution of Eq. (12) for which $\mathbf{y}(t_1) = \mathbf{u}$. The function $\mathbf{y}(t; \mathbf{v})$ satisfies the above multipoint boundary-value problem if and only if

$$h_1\{\mathbf{v}\} + \sum_{i=2}^{r} h_i\{\mathbf{y}(t_i; \mathbf{v})\} - \mathbf{c} = 0 \tag{14}$$

(Eq. (14) is merely Eq. (13) rearranged). Shooting methods operate by solving Eq. (14) and, in this way, reducing a boundary-value problem to a Cauchy problem. We use the algorithm of the preceding section to assist in the solution of Eq. (14), written in the form $\mathbf{v} = F(\mathbf{v})$, where

$$F(\mathbf{v}) = \mathbf{v} + h_1\{\mathbf{v}\} + \sum_{i=2}^{r} h_i\{\mathbf{y}(t_i; \mathbf{v})\} - \mathbf{c}. \tag{15}$$

We now have

$$F'(\mathbf{v}) = \mathbf{I} + \frac{\partial h_1\{\mathbf{v}\}}{\partial \mathbf{v}} + \sum_{i=2}^{r} \frac{\partial h_i\{\mathbf{y}(t_i; \mathbf{v})\}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}(t_i; \mathbf{v})}{\partial \mathbf{v}}. \tag{16}$$

In this special case, step (ii) of our algorithm becomes:

(ii) For $q = 0, 1, \ldots, 2Q + \tau - 1$,

(a) solve Eq. (12) numerically over the range $t_1 \leq t \leq t_r$ by use of a Runge-Kutta or other suitable process to obtain the solution $\mathbf{y}\{t, \mathbf{s}_q\}$;

(b) compute $h_1\{\mathbf{s}_q\}$; obtain $\mathbf{y}\{t_i; \mathbf{s}_q\}$ ($i = 1, 2, \ldots, r$) (if necessary by interpolation) and hence compute $h_i\{\mathbf{y}(t_i; \mathbf{s}_q)\}$ ($i = 2, 3, \ldots, r$);

(c) compute

$$\mathbf{s}_{q+1} = F(\mathbf{s}_q) = \mathbf{s}_q + h_1\{\mathbf{s}_q\} + \sum_{i=2}^{r} h_i\{\mathbf{y}(t_i; \mathbf{s}_q)\} - \mathbf{c}. \tag{17}$$

Naturally, the results of Theorems 2 and 3 can be applied to the special case of our algorithm under consideration in this section. Furthermore, remark (1) of the preceding section concerning the value of Q and τ to be used at each stage still holds.

Theorem 2, which ensures quadratic convergence, cannot be applied in cases in which $\mathbf{I} - F'(\mathbf{v})$ is singular. (In the very special case in which $F(\mathbf{v}) = \mathbf{A}\mathbf{v} + \mathbf{b}$, where $\mathbf{A}$ is a square matrix and $\mathbf{I} - F'(\mathbf{v}) = \mathbf{I} - \mathbf{A}$ has the eigenvalue $\lambda = 0$ with multiplicity 1, Theorem 1 can, of course be applied.) It is not an elementary matter to verify whether or not this condition holds for the operators $F'(\mathbf{v})$ considered in this section; we therefore introduce some assumptions which are sufficient to ensure that it is satisfied.

THEOREM 4. (i) *Let $\partial f(t,\mathbf{y})/\partial \mathbf{y}$ exist and be uniformly bounded, i.e., let a constant k exist such that*

$$\left\| \frac{\partial f(t,\mathbf{y})}{\partial \mathbf{y}} \right\| < k \quad \forall (t,\mathbf{y}) \in [t_1, t_r] \times \mathbf{R}^p. \tag{18}$$

(ii) *Let the functions $h_i\{\mathbf{y}(t_i)\}$ in Eq.* (13) *have the form $\mathbf{H}_i \mathbf{y}(t_i)$, where $\mathbf{H}_i$ is a linear operator $(i = 1, 2, \ldots, r)$ with $\mathbf{S}_0 = \sum_{i=1}^r \mathbf{H}_i$ being invertible.*

(iii) *Let*

$$|t_r - t_1| < k^{-1} \ln\left\{1 + \left(\sum_{i=2}^r \|\mathbf{H}_i\|\right)^{-1} \|\mathbf{S}_0^{-1}\|^{-1}\right\}. \tag{19}$$

Then the multipoint boundary-value problem of Eqs. (12) *and* (13) *has a unique solution and, if application of the vector ε-algorithm during stage* (iii) *does not break down, the algorithm described above produces a sequence of vectors $\{\mathbf{v}_n\}$ converging quadratically to* $\mathbf{v}$.

Proof. Under the conditions of the theorem, the multipoint boundary-value problem of Eqs. (12) and (13) has a unique solution [7].

In the special case being considered, $F'(\mathbf{v}) - \mathbf{I} = \mathbf{S}$, where

$$\mathbf{S} = \mathbf{H}_1 + \sum_{i=2}^r \mathbf{H}_i \, \partial \mathbf{y}(t_i; \mathbf{v})/\partial \mathbf{v},$$

so that

$$\mathbf{S} - \mathbf{S}_0 = \sum_{i=2}^r \mathbf{H}_i\{\partial \mathbf{y}(t_i; \mathbf{v})/\partial \mathbf{v} - \mathbf{I}\}$$

and, using operator norms,

$$\|\mathbf{S} - \mathbf{S}_0\| \le \sum_{i=2}^r \|\mathbf{H}_i\| \, \|\partial \mathbf{y}(t_i; \mathbf{v})/\partial \mathbf{v} - \mathbf{I}\|.$$

The operator $\mathbf{T}(t) = \partial \mathbf{y}\{t; \mathbf{v}\}/\partial \mathbf{v} - \mathbf{I}$ satisfies the equations

$$\frac{d\mathbf{T}(t)}{dt} = \frac{\partial f\{t, \mathbf{y}(t; \mathbf{v})\}}{\partial \mathbf{y}} \mathbf{T}(t) + \frac{\partial f\{t, \mathbf{y}(t; \mathbf{v})\}}{\partial \mathbf{y}}, \quad \mathbf{T}(t_1) = 0,$$

and, in view of relationship (18), $\|\mathbf{T}(t)\| \le (e^{k|t-t_1|} - 1)$. Hence

$$\|\mathbf{S} - \mathbf{S}_0\| \le \sum_{i=2}^r \|\mathbf{H}_i\| (e^{k(t-t_i)} - 1) \le (e^{k|t_r - t_1|} - 1) \sum_{i=2}^r \|\mathbf{H}_i\|.$$

Since $|t_r - t_1|$ satisfies inequality (19), we have

$$(e^{k|t_r - t_1|} - 1) \sum_{i=2}^r \|\mathbf{H}_i\| < \|\mathbf{S}_0^{-1}\|^{-1}.$$

Hence $\|\mathbf{S} - \mathbf{S}_0\| < \|\mathbf{S}_0^{-1}\|^{-1}$ and, from a classical result in the theory of linear operators [16] it follows that $\mathbf{S}$ is invertible. The result of the theorem now follows from Theorem 3.

The conditions of Theorem 4 are rather restrictive; they exclude such cases as

Dirichlet conditions for two-point boundary-value problems. But, in practice, we have found our algorithm to be effective for these problems, and also on larger intervals than those obeying inequality (19).

Examples. (1) Our first example is set in $\mathbf{R}^3$ and, in the notation of Eq. (13), $r = 3$ also, with $t_1 = 0$, $t_2 = \frac{1}{2}$, $t_3 = 1$. In this example, Eqs. (12) and (13) may be written in component form as

$$(20)\quad \frac{dy_1(t)}{dt} = y_2(t) - y_3(t),\quad \frac{dy_2(t)}{dt} = y_1(t)^2 + y_2(t),\quad \frac{dy_3(t)}{dt} = y_1(t)^2 + y_3(t),$$

and

$$y_1(0) = 1,\quad y_2(\tfrac{1}{2}) = e - 4,\quad y_3(1) = -4 - e - e^2,$$

respectively. Thus, in the notation of Theorem 4, we have

$$\mathbf{H}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad \mathbf{H}_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\quad \mathbf{H}_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},\quad \mathbf{c} = \begin{bmatrix} 1 \\ e-4 \\ -4-e-e^2 \end{bmatrix}.$$

Starting with the initial vector $\mathbf{v}_0 = \mathbf{0}$, carrying out the numerical integration of Eqs. (20) as described in step (iia) of the algorithm by means of a Runge-Kutta process with absolute error stipulated to be less than $0.5_{10}-9$, and setting $m = 3$, $\tau = 0$ in step (iii), we find that

$$\begin{aligned}
\mathbf{v}_1 &= (0.921\,428\,452, -0.951\,029\,784, -0.100\,612\,601),\\
\mathbf{v}_2 &= (0.999\,698\,042, -1.002\,885\,341, -7.49_{10}-4),\\
\mathbf{v}_3 &= (0.999\,999\,542, -0.999\,997\,855, -3.70_{10}-7),\\
\mathbf{v}_4 &= (0.999\,999\,995, -0.999\,999\,996, -1.10_{10}-10).
\end{aligned}$$

The next iteration yields $\mathbf{v}_5 = \mathbf{v}_4$: we have solved the discretized problem to the same accuracy as that prescribed for the numerical integration procedure.

The above boundary-value problem has the unique solution

$$y_1(t) = 2 - e^t,\quad y_2(t) = -4 - (4t-2)e^t + e^{2t},$$

$$y_3(t) = -4 - (4t-3)e^t + e^{2t}.$$

Thus $(y_1(0), y_2(0), y_3(0)) = (1, -1, 0)$, the vector computed to the prescribed accuracy by our process. If the recursion $\mathbf{s}_{q+1} = F(\mathbf{s}_q)$ is left to look after itself, it diverges. The application of the vector ε-algorithm to the 24 first of these iterations fails to approach the solution.

(2) In the second example, $p = 4$ and $r = 2$, with $t_1 = 0$, $t_2 = 1$; Eq. (12) has the form

$$(21)\qquad d\mathbf{y}(t)/dt = -\mathbf{y}(t);$$

Eq. (13) has the form described in clause (iii) of Theorem 4, with

$$\mathbf{H}_1 = \begin{bmatrix} 0 & e^{-1} & 0 & 0 \\ e^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{H}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

We first remark that there is, in this very simple case, no need to go through the numerical integration process of step (iia) of the algorithm: Eq. (21) can be solved exactly by inspection. Secondly, in this case Eq. (15) takes the form

$$\mathbf{v} = \mathbf{v} - \mathbf{G}\mathbf{v} + \mathbf{c} \tag{22}$$

where $\mathbf{G} = \mathbf{H}_1 + e^{-1}\mathbf{H}_2$. It can immediately be seen that the solutions to our example satisfy the conditions

$$y_1(0) + y_2(0) = e, \quad y_3(0) = y_4(0) = e, \tag{23}$$

and are infinite in number. Equation (22) is purely algebraic; its solution may be studied with the help of Theorem 1. The minimal polynomial of $\mathbf{A} = \mathbf{I} - \mathbf{G}$ is $(1 - e^{-1} - \lambda)(1 - 2e^{-1} - \lambda)(1 - \lambda)$; it is of degree 3 and has 1 as a single root. It follows from Theorem 1 that for any initial vector $\mathbf{v}_0$, the vector $\boldsymbol{\varepsilon}_4^{(0)} = \mathbf{g} = (g_1, g_2, g_3, g_4)$ (if not a precedessor) derived by means of the ε-algorithm from the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_8$ produced by the iteration $\mathbf{v}_{q+1} = \mathbf{v}_q - \mathbf{G}\mathbf{v}_q + \mathbf{c}$ will be such that in terms of Eq. (21), $\mathbf{y}(t; \mathbf{g})$ satisfies the boundary conditions (23) and, in particular, $g_1 + g_2 = e$, $g_3 + g_4 = e$. Starting with $\mathbf{v}_0 = (-2, 1, 3, 1)$, we find, to the accuracy indicated, that

$$\boldsymbol{\varepsilon}_4^{(0)} = (0.85914\ 091422,\ 1.85914\ 091422,\ 2.71828\ 182845,\ 2.71828\ 182845).$$

(3) The last example is a rather more difficult nonlinear two-point boundary-value problem [10],[14] set in $\mathbf{R}^5$, and $t_1 = 0$, $t_2 = T$ where T is a prescribed positive real number. Equation (12) has the form

$$\begin{aligned}
&\frac{dy_1(t)}{dt} = y_2(t), \quad \frac{dy_2(t)}{dt} = y_3(t), \\
&\frac{dy_3(t)}{dt} = -1.55y_1(t)y_3(t) + 0.1y_2(t)^2 + 0.2y_2(t) - y_4(t)^2 + 1.0, \\
&\frac{dy_4(t)}{dt} = y_5(t), \\
&\frac{dy_5(t)}{dt} = -1.55y_1(t)y_5(t) + 0.2y_4(t) + 1.1y_2(t)y_4(t) - 0.2
\end{aligned} \tag{24}$$

and Eq. (13) may be written as

$$y_1(0) = y_2(0) = y_4(0) = 0, \quad y_2(T) = 0, \quad y_4(T) = 1.$$

This example is derived from a boundary layer problem. The solutions of the differential equations (24) are very sensitive to small variations of the initial conditions, and for large values of T it is impossible to solve the above boundary-value problem directly by shooting methods.

Roberts and Shipman [15] have attacked the above problem by means of a continuation method, solving the problem over the interval $[0, T_j]$ with $y_2(T_j) = 0$, $y_4(T_j) = 1$ at each stage, where $0 < T_1 < T_2 < \cdots < T_{j'} = T$ $(j' < \infty)$. We use our algorithm in the same way. Each problem, solved over the interval $[0, T_j]$, results in the determination of missing initial values $y_3(0)$, $y_5(0)$; these are used in conjunction with the remaining fixed initial values as an initial guess upon which to base the construction of a solution over the interval $[0, T_{j+1}]$. Table 1 gives these successive missing values for successive values of T_j and the error

$$\left\| h_1(\mathbf{v}) + \sum_{i=2}^{r} h_i\{\mathbf{y}(t_i; \mathbf{v})\} - \mathbf{c} \right\|^2 .$$

During the computation over the early intervals $[0, T_j]$ agreement to only three or so decimal figures is required; agreement to a larger number of decimal figures is reserved for the later intervals. As a point of detail, we mention that in this case a parameter δ was introduced into Eq. (17) which becomes

$$\mathbf{s}_{q+1} = \mathbf{s}_q + \delta\left[h_1(\mathbf{s}_q) + \sum_{i=2}^{r} h_i\{\mathbf{y}(t_i; \mathbf{s}_q)\} - \mathbf{c}\right].$$

δ being taken small enough to suppress the effects of the explosive growth of $\mathbf{y}(t; \mathbf{s}_q)$ (values of δ between 10^{-8} and 10^{-12} were used); thus, we can avoid too small increments of T_j at each stage and save considerable computing time. The true solution of this problem is unknown and it is impossible to say how many decimal figures are exact.

The results of Table 1 agree to six figures with those of Roberts and Shipman. These authors give a justification [15], based on the Kantorovich sufficiency theorem

TABLE 1

T_j	$y_3(0)$	$y_5(0)$	*error*
3, 5	− .9738661361858899	.6470687025823338	.17D − 1
4	− .9691814158187645	.6525290853740247	.36D − 5
4, 5	− .9666762206420572	.6531111218305619	.70D − 7
5	− .9662973199428564	.6529987946964353	.29D − 11
5, 5	− .9662891737064793	.6529317756489144	.74D − 11
6	− .9663041351251466	.6529134552438886	.50D − 12
6, 5	− .9663099896872114	.6529099441374421	.16D − 8
7	− .9663114878332717	.6529095123304221	.23D − 9
8	− .9663118133931346	.6529095593410232	.58D − 12
8, 5	− .9663118098329179	.6529095729051251	.10D − 14
9	− .9663118055095461	.6529095769575010	.22D − 15
10	− .9663118030841837	.6529095779273979	.54D − 15
10, 5	− .9663118029686369	.6529095778846625	.77D − 8
11, 2	− .9663118029692275	.6529095778464226	.34D − 12

for Newton's method, for the continuation method when used in conjunction with quasi-linearisation [8]. Unfortunately, a similar justification for the continuation method used in conjunction with our method has not yet been achieved.

In conclusion we remark that the boundary conditions (13) are not the most general that can be treated by the method described above.

Acknowledgement. The authors wish to thank the referee for many valuable suggestions and for a more elegant proof of Theorem 2.

89 Rue Daguerre
75014–Paris, France

9 Rue Trouillet-Derel
92600–Asnieres, France

1. R. ALT, *Méthodes A-Stables pour l'Intégration des Systèmes Différentiels Mal Conditionnés*, Thèse, 3ième cycle, Paris.

2. C. BREZINSKI, "Application de l'ε-algorithme à la résolution des systèmes non linéaires," *C. R. Acad. Sci. Paris Sér. A-B*, v. 271, 1970, pp. A1174-A1177. MR **42** #7046.

3. C. BREZINSKI, "Sur un algorithme de résolution des systèmes non linéaires," *C. R. Acad. Sci. Paris Sér. A-B*, v. 272, 1971, pp. A145-A148. MR **43** #1410.

4. C. BREZINSKI, "Some results in the theory of the vector ε-algorithm," *Linear Algebra.* v.8, 1974, pp. 77–86.

5. J. DIEUDONNÉ, *Fondaments de l'Analyse Moderne*, Pure and Appl. Math., vol. 10, Academic Press, New York and London, 1964. MR **22** #11074.

6. E. GEKELER, "On the solution of systems of equations by the epsilon algorithm of Wynn," *Math. Comp.*, v. 26, 1972, pp. 427-436.

7. M. A. GOLDBERG, "A generalized invariant imbedding equation," *J. Math. Anal. Appl.*, v. 33, 1971, pp. 518-528. MR **44** #7095a.

8. T. R. GOODMAN & G. N. LANCE, "The numerical integration of two-point boundary value problems," *MTAC*, v. 10, 1956, pp. 82-86. MR **18**, 420.

9. P. HENRICI, *Elements of Numerical Analysis*, Wiley, New York, 1964. MR **29** #4173.

10. J. F. HOLT, "Numerical solution of nonlinear two-point boundary problems by finite difference methods," *Comm. ACM*, v. 7, 1964, pp. 366-373. MR **29** #5387.

11. J. B. MCLEOD, "A note on the ε-algorithm," *Computing* (*Arch. Elektron.Rechnen*), v. 7, 1971, pp. 17-24. MR **44** #7693.

12. A. M. OSTROWSKI, *Solution of Equations and Systems of Equations*, 2nd ed., Pure and Appl. Math., vol. 9, Academic Press, New York, 1966. MR **35** #7575.

13. J. M. ORTEGA & W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970. MR 42 #8686.

14. S. M. ROBERTS & J. S. SHIPMAN, "Continuation in shooting methods for two-point boundary value problems," *J. Math. Anal. Appl.*, v. 18, 1967, pp. 45-58. MR **34** #7029.

15. S. M. ROBERTS & J. S. SHIPMAN, "Justification for the continuation method in two-point boundary value problems," *J. Math. Anal. Appl.*, v. 21, 1968, pp. 23-30. MR **36** #3511.

16. L. SCHWARTZ, *Topologie Générale et Analyse Fonctionnelle*, Hermann, Paris, 1970.

17. S. YU. ULM, "Extension of Steffensen's method for solving operator equations," *U.S.S.R. Comput. Math. and Math. Phys.*, v. 4, 1964, pp. 159-165.

18. P. WYNN, "Acceleration techniques for iterated vector and matrix problems," *Math. Comp.*, v. 16, 1962, pp. 301-322. MR **26** #3176.

19. P. WYNN, *Upon a Conjecture Concerning a Method for Solving Linear Equations, and Certain Other Matters*, University of Wisconsin, Math. Res. Center Report No. 626, 1966.

20. P. WYNN, "Accélération de la convergence de séries d'opérateurs en analyse numérique," *C. R. Acad. Sci. Paris Sér. A-B*, v. 276, 1973, pp. A803–A806.

21. P. WYNN, "Upon some continuous prediction algorithms. I, II," *Calcolo* 1972; Parts 3 and 4,v. 9, pp. 197–234 and pp. 235–278.