MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 118, APRIL 1972

On the Solution of Systems of Equations by the Epsilon Algorithm of Wynn

By E. Gekeler

Abstract. The ϵ-algorithm has been proposed by Wynn on a number of occasions as a convergence acceleration device for vector sequences; however, little is known concerning its effect upon systems of equations. In this paper, we prove that the algorithm applied to the Picard sequence $\mathbf{x}_{i+1} = F(\mathbf{x}_i)$ of an analytic function $F: \mathbf{R}^n \supset D \to \mathbf{R}^n$ provides a quadratically convergent iterative method; furthermore, no differentiation of F is needed. Some examples illustrate the numerical performance of this method and show that convergence can be obtained even when F is not contractive near the fixed point. A modification of the method is discussed and illustrated.

1. Introduction. The ϵ-algorithm is a nonlinear method for accelerating the convergence of sequences; in its simplest form, it is identical with the δ^2 transformation of Aitken [1]. The determinantal formulae upon which it is based were given by Jacobi [6], Schmidt [11], and Shanks [12]; Wynn [13] developed it and examined it thoroughly in connection with various sequences and series [14]–[17]. The ϵ-algorithm provides higher (integer) order methods for the computation of a fixed point of an analytic function $f: \mathbf{C} \supset D \to \mathbf{C}$ [4]. Using the generalized matrix inverse of Moore [8] and Penrose [9], the method has recently been applied to sequences of matrices and vectors as they arise, for example, in the solution of linear systems of equations [5], [7], [10], [18], [21], [22], [23]. Wynn points out that the algorithm also provides good results in the numerical solution of nonlinear systems [18], [19], [21], [22]. But, until now, nothing is known concerning convergence. In this paper, we examine the behaviour of the ϵ-algorithm when applied to the Picard sequence of an analytic function F: $\mathbf{R}^n \supset D \to \mathbf{R}^n$ with fixed point $\mathbf{z}$. With the help of a theorem of McLeod [7], we show that the algorithm, used in a manner similar to Steffensen's method, is a quadratically convergent iterative method for the computation of $\mathbf{z}$ (compare also Brezinski [2]*). Because of the complicated recursive relationships, the convergence considered is of local nature, and Landau symbols are used in the proof. A short discussion of numerical properties of the method follows at the end of the paper.

We use certain standard notations: $i \in \mathbf{N}$ means that i is a nonnegative integer; lower (upper) case bold face letters denote vectors (matrices); $||\mathbf{x}||$ is the Euclidean norm $(\mathbf{x}^*\mathbf{x})^{1/2}$ of the n-dimensional column vector $\mathbf{x} \in \mathbf{C}^n$; $O(||\mathbf{x}||^i)$ denotes a vector-valued function of the vector $\mathbf{x}$ whose norm remains bounded as $||\mathbf{x}|| \to 0$ after division by $||\mathbf{x}||^i$; $O\{||\mathbf{x}||^i\}$ denotes a real valued function with the same properties.

We also make use of the concept of an analytic function of a vector and of a vector-valued Taylor series. Let D be an open subset of $\mathbf{R}^n$, then $F: \mathbf{R}^n \supset D \to \mathbf{R}^n$ is called

Received October 12, 1970.

AMS 1969 *subject classifications.* Primary 6510, 6550.

Key words and phrases. ϵ-algorithm, acceleration of the convergence of sequences, quadratic convergent iterative method without differentiation, solution of equations.

* We are grateful to the referee for pointing out this paper.

analytic if, for every point $a \in D$, there is an open polycylinder $P = \{\mathbf{x} \in \mathbf{R}^n, |x_i - a_i| < r_i, 0 < r_i, 1 \leqq i \leqq n\} \subset D$, such that in P, $F(\mathbf{x})$ is equal to the sum of an absolutely summable power series in the n variables $x_i - a_i$ $(1 \leqq i \leqq n)$. An analytic function is indefinitely differentiable, and, if the segment joining $\mathbf{x}$ and $\mathbf{x} + \mathbf{y}$ is in D, we have, for $r \in \mathbf{N}$,

$$F(\mathbf{x} + \mathbf{y}) = F(\mathbf{x}) + \sum_{k=1}^{r-1} \frac{1}{k!} F^{(k)}(\mathbf{x}) \cdot \mathbf{y}^{(k)} + \left(\int_0^1 \frac{(1-t)^{r-1}}{(r-1)!} F^{(r)}(\mathbf{x} + t\mathbf{y})\, dt \right) \cdot \mathbf{y}^{(r)},$$

where $\mathbf{y}^{(k)}$ stands for $(\mathbf{y}, \mathbf{y}, \cdots, \mathbf{y})$ (k times). For further details, we refer to the famous book of Dieudonné [3].

2. Picard Sequences. We consider some iterative schemes for determining a fixed point $\mathbf{z}$ of the equation $\mathbf{x} = F(\mathbf{x})$. If $\mathbf{s}_p$ $(p \in \mathbf{N}, 0 \leqq p)$ is near $\mathbf{z}$, we have, using a Taylor expansion for $F(\mathbf{z})$,

$$\mathbf{z} = F(\mathbf{s}_p) + F'(\mathbf{s}_p)(\mathbf{z} - \mathbf{s}_p) + O(||\mathbf{z} - \mathbf{s}_p||^2). \tag{1}$$

Thus, when using the simple iteration scheme

$$\mathbf{s}_{p+1} = F(\mathbf{s}_p) \qquad (0 \leqq p), \tag{2}$$

we have

$$\mathbf{z} - \mathbf{s}_{p+1} = F'(\mathbf{s}_p)(\mathbf{z} - \mathbf{s}_p) + O(||\mathbf{z} - \mathbf{s}_p||^2).$$

Hence, the simple scheme (2) is, in general, at best linearly convergent; whether it converges or not depends upon the magnitudes of the eigenvalues of the Jacobian matrices $F'(\mathbf{s}_p)$ $(0 \leqq p)$ in the neighbourhood of $\mathbf{z}$. We can, however, devise a quadratically convergent scheme based upon the solution of the linear system

$$\hat{\mathbf{s}}_{p+1} = F(\hat{\mathbf{s}}_p) + F'(\hat{\mathbf{s}}_p)(\hat{\mathbf{s}}_{p+1} - \hat{\mathbf{s}}_p) \qquad (0 \leqq p)$$

or

$$(\mathbf{I} - F'(\hat{\mathbf{s}}_p))\hat{\mathbf{s}}_{p+1} = F(\hat{\mathbf{s}}_p) - F'(\hat{\mathbf{s}}_p)\hat{\mathbf{s}}_p \qquad (0 \leqq p) \tag{3}$$

for $\hat{\mathbf{s}}_{p+1}$. For, replacing $\mathbf{s}_p$ in formula (1) by $\hat{\mathbf{s}}_p$, we now have

$$\mathbf{z} - \hat{\mathbf{s}}_{p+1} = F'(\hat{\mathbf{s}}_p)(\mathbf{z} - \hat{\mathbf{s}}_{p+1}) + O(||\mathbf{z} - \hat{\mathbf{s}}_p||^2) \qquad (0 \leqq p),$$

i.e.,

$$(\mathbf{I} - F'(\hat{\mathbf{s}}_p))(\mathbf{z} - \hat{\mathbf{s}}_{p+1}) = O(||\mathbf{z} - \hat{\mathbf{s}}_p||^2) \qquad (0 \leqq p)$$

or, again subject to certain assumptions concerning the eigenvalues of $F'(\mathbf{x})$ in the neighbourhood of $\mathbf{z}$,

$$\mathbf{z} - \hat{\mathbf{s}}_{p+1} = O(||\mathbf{z} - \hat{\mathbf{s}}_p||^2) \qquad (0 \leqq p).$$

The second scheme, although yielding quadratic convergence, involves evaluation of a Jacobian matrix and the solution of a linear system at each stage. However, by use of the ϵ-algorithm one can, as we shall show, obtain quadratic convergence without the computation of the derivatives occurring in the Jacobian matrix, and without the solution of a linear system.

3. The Algorithm. The ϵ-algorithm [13], [22] is a computational procedure in which successive columns of an array $(\epsilon_q^{(p)})_{0 \leqq p, 0 \leqq q}$ with row index p are obtained by use of the formula

$$\epsilon_{q+1}^{(p)} = \epsilon_{q-1}^{(p+1)} + (\epsilon_q^{(p+1)} - \epsilon_q^{(p)})^{-1} \qquad (0 \leqq p, 0 \leqq q), \tag{4}$$

starting from the initial conditions

$$\epsilon_{-1}^{(p)} = 0, \qquad \epsilon_0^{(p)} = s_p \qquad (0 \leqq p). \tag{5}$$

If the inverse of a nonzero vector $\mathbf{x} \in \mathbf{C}^n$ is defined, by [8], [9],

$$\mathbf{x}^{-1} = (\mathbf{x}^*\mathbf{x})^{-1}\bar{\mathbf{x}}, \tag{6}$$

then we can apply the algorithm to sequences $\{\mathbf{s}_p\}_{0 \leqq p}$ of vectors and have the fundamental theorem [7], [23] which we need later:

THEOREM 1. *Let* $\{\mathbf{s}_p\}_{0 \leqq p}$ *be a sequence of vectors with complex coefficients which satisfy the irreducible linear recursion*

$$\sum_{r=0}^{m} c_r \mathbf{s}_{p+r} = \left(\sum_{r=0}^{m} c_r\right)\mathbf{s} \qquad (0 \leqq p), \tag{7}$$

where $\mathbf{s}$ *is fixed and*

$$\sum_{r=0}^{m} c_r \neq 0, \qquad c_r \in \mathbf{R}. \tag{8}$$

If then the elements of the array $(\boldsymbol{\varepsilon}_q^{(p)})$ *are determined by using* (4), (5), *and* (6), *and if all* $\boldsymbol{\varepsilon}_q^{(p)}$ *with* $p + q \leqq 2m$ *exist, then*

$$\boldsymbol{\varepsilon}_{2m}^{(0)} = \mathbf{s}.$$

Following a conjecture of Wynn [24] and Greville [5], Theorem 1 remains true if relations (7), (8) hold for complex scalars only, but this has not yet been proved. In conclusion, we get

COROLLARY. *Let* $\mathbf{z}$ *be the unique solution of the linear system* $\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{c}$ *with real coefficients and let* m *be the degree of the minimal polynomial of the matrix* $\mathbf{A}$ *for* $\mathbf{y} = \mathbf{x}_0 - \mathbf{z}$. *If the* ϵ*-algorithm is applied to the Picard sequence* $\{\mathbf{x}_p; \mathbf{x}_{p+1} = \mathbf{A}\mathbf{x}_p + \mathbf{c}\}_{0 \leqq p}$ *and if all* $\boldsymbol{\varepsilon}_q^{(p)}$ *with* $p + q \leqq 2m$ *exist, then*

$$\boldsymbol{\varepsilon}_{2m}^{(0)} = \mathbf{z}.$$

Proof. Let $p(x) = \sum_{r=0}^{m} a_r x^r$ be the minimal polynomial of $\mathbf{A}$ for $\mathbf{y}$, then

$$\sum_{r=0}^{m} a_r \mathbf{x}_{p+r} = \left(\sum_{r=0}^{m} a_r\right)\mathbf{z} + \left(\sum_{r=0}^{m} a_r \mathbf{A}^{p+r}\right)\mathbf{y} = \left(\sum_{r=0}^{m} a_r\right)\mathbf{z},$$

because $\mathbf{x}_p = \mathbf{z} + \mathbf{A}^p\mathbf{y}$ holds. By assumption, we have $\sum_{r=0}^{m} a_r \neq 0$, since 1 is not eigenvalue of $\mathbf{A}$ (the equation $\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{c}$ has a unique solution), and the Corollary results from Theorem 1.

4. The Application of the Epsilon Algorithm to Picard Sequences. The general strategy adopted in deriving our main result is this: we first consider the behaviour of the vectors $\tilde{\varepsilon}_q^{(p)}$ ($p + q \leqq 2n$) derived by means of the ϵ-algorithm from the sequence $\tilde{s}_p = z + A^p y$ $(0 \leqq p)$, where $y, z \in R^n$ and A is a real $n \times n$ matrix, for small values of $||y||$ (we know from the above Corollary that, subject to certain conditions, $\tilde{\varepsilon}_{2n}^{(0)} = z$). We then consider the behaviour of corresponding vectors derived from the sequence $s_p = \tilde{s}_p + \delta_p$, where $\delta_p = O(||y||^2)$ $(0 \leqq p)$. Finally, we use these results with $A = F'(z)$ and

$$s_{p+1} = F(s_p) = z + F'(z)(s_p - z) + O(||s_p - z||^2) \qquad (0 \leqq p)$$

to examine the behaviour of the vectors $\varepsilon_q^{(p)}$ produced from this iterative scheme when s_0 is near a fixed point z and, in particular, to show that repeated use of the vector $\varepsilon_{2n}^{(0)}$ in place of s_0 results in a quadratically convergent process for determining the fixed point in question. *In the sequel, let $Q_m(A) \subset R^n$ be the set of vectors* x *for which m is the degree of the minimal polynomial of* A.

LEMMA 1. *For a given* z, *let $\tilde{\varepsilon}_q^{(p)}$ be the vectors obtained by means of the ϵ-algorithm from the sequence $\{\tilde{s}_p; \tilde{s}_p = z + A^p y\}_{0 \leqq p}$. If there is a neighbourhood U of* 0 *such that all $\tilde{\varepsilon}_q^{(p)}$ with $p + q \leqq 2m$ exist for all $y \in U \cap Q_m(A)$, then*

$$\tilde{\varepsilon}_q^{(p)} = z + O(||y||), \qquad q \text{ even},$$

$$\tilde{\varepsilon}_q^{(p)} = O(||y||^{-1}), \qquad q \text{ odd},$$

for $y \in Q_m(A)$ and $p + q \leqq 2m$.

Proof. Let $m > 0$, $p \leqq 2m - q$, and $\Delta_p \tilde{\varepsilon}_q^{(p)} = \tilde{\varepsilon}_q^{(p+1)} - \tilde{\varepsilon}_q^{(p)}$. For $q = 1$, we get $\Delta_p \tilde{\varepsilon}_0^{(p)} = A^p(A - I)y = B_p y$, and $B_p y \neq 0$ for $y \in Q_m(A)$, by assumption. Hence,

$$\begin{aligned} ||\tilde{\varepsilon}_1^{(p)}|| &= ||(y^* B_p^* B_p y)^{-1} B_p y|| \\ &= \frac{1}{||y||} \frac{y^* y}{y^* B_p^* B_p y} \frac{1}{||y||} ||B_p y|| \leqq \frac{1}{||y||} \frac{||B_p||}{\lambda_{\min}}, \end{aligned}$$

where $0 < \lambda_{\min}$ is the smallest eigenvalue of $B_p^* B_p$. Let now $k \in N$, $k < m$, $y \in Q_m(A)$, and let the statement be true for all $q \leqq 2k$. By assumption, we have $\Delta_p \tilde{\varepsilon}_{2k}^{(p)} = O(||y||) \neq 0$, thus

$$(\Delta_p \tilde{\varepsilon}_{2k}^{(p)})^* (\Delta_p \tilde{\varepsilon}_{2k}^{(p)}) = O\{||y||^2\},$$

$$\begin{aligned} \tilde{\varepsilon}_{2k+1}^{(p)} &= \tilde{\varepsilon}_{2k-1}^{(p+1)} + [(\Delta_p \tilde{\varepsilon}_{2k}^{(p)})^* (\Delta_p \tilde{\varepsilon}_{2k}^{(p)})]^{-1} \Delta_p \tilde{\varepsilon}_{2k}^{(p)} \\ &= O(||y||^{-1}) + O\{||y||^{-2}\} O(||y||) = O(||y||^{-1}). \end{aligned}$$

$\Delta_p \tilde{\varepsilon}_{2k+1}^{(p)} \neq 0$, since, by assumption, all $\tilde{\varepsilon}_q^{(p)}$ which contribute to $\tilde{\varepsilon}_{2m}^{(0)}$ exist. Therefore,

$$(\Delta_p \tilde{\varepsilon}_{2k+1}^{(p)})^* (\Delta_p \tilde{\varepsilon}_{2k+1}^{(p)}) = O\{||y||^{-2}\},$$

and

$$\begin{aligned} \tilde{\varepsilon}_{2k+2}^{(p)} &= \tilde{\varepsilon}_{2k}^{(p+1)} + [(\Delta_p \tilde{\varepsilon}_{2k+1}^{(p)})^* (\Delta_p \tilde{\varepsilon}_{2k+1}^{(p)})]^{-1} \Delta_p \tilde{\varepsilon}_{2k+1}^{(p)} \\ &= z + O(||y||) + O\{||y||^2\} O(||y||^{-1}) = z + O(||y||), \end{aligned}$$

and the assertion of the lemma follows by induction.

LEMMA 2. *Let $\{\delta_p\}_{0 \leqq p}$ be a sequence of analytic functions $\delta_p(y) = O(||y||^2)$. For a given* z, *let $\varepsilon_q^{(p)}$ be the vectors obtained by means of the ϵ-algorithm from the sequence*

$\{s_p;\ s_p = z + A^p y + \delta_p(y)\}_{0 \leq p}$. *If there is a neighbourhood* U *of* 0 *such that all* $\varepsilon_q^{(p)}$, $\bar{\varepsilon}_q^{(p)}$ *with* $p + q \leqq 2m$ *exist for all* $y \in U \cap Q_m(A)$, *then*

$$\varepsilon_q^{(p)} = \bar{\varepsilon}_q^{(p)} + O(\|y\|^2), \qquad q \text{ even},$$

$$\varepsilon_q^{(p)} = \bar{\varepsilon}_q^{(p)} + O(1), \qquad q \text{ odd},$$

for $y \in Q_m(A)$ *and* $p + q \leqq 2m$.

Proof. Let $m > 0$ and $p \leqq 2m - q$. For $q = 1$, we have $\Delta_p\varepsilon_0^{(p)} = \Delta_p\bar{\varepsilon}_0^{(p)} + O(\|y\|^2) \neq 0$ and $\Delta_p\bar{\varepsilon}_0^{(p)} \neq 0$ for $y \in Q_m(A)$, by assumption. Then

$$\begin{aligned}(\Delta_p\varepsilon_0^{(p)})^*(\Delta_p\varepsilon_0^{(p)}) &= (\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)}) + 2(\Delta_p\bar{\varepsilon}_0^{(p)})^*O(\|y\|^2) + O\{\|y\|^4\}\\ &= (\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})\left[1 + 2\frac{(\Delta_p\bar{\varepsilon}_0^{(p)})^*O(\|y\|^2)}{(\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})} + \frac{O\{\|y\|^4\}}{(\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})}\right].\end{aligned}$$

$\Delta_p\bar{\varepsilon}_0^{(p)} = O(\|y\|)$ and hence,

$$(\Delta_p\varepsilon_0^{(p)})^*(\Delta_p\varepsilon_0^{(p)}) = (\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})(1 + O\{\|y\|\}).$$

Since $\Delta_p\varepsilon_0^{(p)}$ is an analytic function, we get

$$[(\Delta_p\varepsilon_0^{(p)})^*(\Delta_p\varepsilon_0^{(p)})]^{-1} = [(\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})]^{-1}[1 + O\{\|y\|\}]$$

and

$$\begin{aligned}\varepsilon_1^{(p)} &= \bar{\varepsilon}_1^{(p)} + [(\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})]^{-1}O\{\|y\|\}\Delta_p\bar{\varepsilon}_0^{(p)}\\ &\quad + [(\Delta_p\bar{\varepsilon}_0^{(p)})^*(\Delta_p\bar{\varepsilon}_0^{(p)})]^{-1}[1 + O\{\|y\|\}]O(\|y\|^2)\\ &= \bar{\varepsilon}_1^{(p)} + O(1).\end{aligned}$$

Let now $k \in \mathbf{N}$, $k < m$, $y \in Q_m(A)$, and let the statement be true for all $q \leqq 2k$. By assumption, we have $\Delta_p\varepsilon_{2k}^{(p)} = \Delta_p\bar{\varepsilon}_{2k}^{(p)} + O(\|y\|^2) \neq 0$ and $\Delta_p\bar{\varepsilon}_{2k}^{(p)} \neq 0$. According to the proof for $q = 1$, we get, by use of Lemma 1,

$$[(\Delta_p\varepsilon_{2k}^{(p)})^*(\Delta_p\varepsilon_{2k}^{(p)})]^{-1}\ \Delta_p\varepsilon_{2k}^{(p)} = [(\Delta_p\bar{\varepsilon}_{2k}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k}^{(p)})]^{-1}\ \Delta_p\bar{\varepsilon}_{2k}^{(p)} + O(1)$$

and hence,

$$\varepsilon_{2k+1}^{(p)} = \bar{\varepsilon}_{2k+1}^{(p)} + O(1).$$

$\Delta_p\varepsilon_{2k+1}^{(p)} = \Delta_p\bar{\varepsilon}_{2k+1}^{(p)} + O(1)$ and $\Delta_p\bar{\varepsilon}_{2k+1}^{(p)}$ are equally supposed to be different from zero and, therefore, we get, by use of Lemma 1,

$$\begin{aligned}&(\Delta_p\varepsilon_{2k+1}^{(p)})^*(\Delta_p\varepsilon_{2k+1}^{(p)})\\ &\quad = (\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})\left[1 + 2\frac{(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*O(1)}{(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})} + \frac{O\{1\}}{(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})}\right].\\ &\quad = (\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})(1 + O\{\|y\|\}).\end{aligned}$$

$$\begin{aligned}\varepsilon_{2k+2}^{(p)} &= \bar{\varepsilon}_{2k}^{(p+1)} + O(\|y\|^2)\\ &\quad + [(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})]^{-1}[1 + O\{\|y\|\}][\Delta_p\bar{\varepsilon}_{2k+1}^{(p)} + O(1)]\\ &= \bar{\varepsilon}_{2k+2}^{(p+1)} + O(\|y\|^2) + [(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})]^{-1}O\{\|y\|\}\ \Delta_p\bar{\varepsilon}_{2k+1}^{(p)}\\ &\quad + [(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})^*(\Delta_p\bar{\varepsilon}_{2k+1}^{(p)})]^{-1}[1 + O\{\|y\|\}]O(1)\\ &= \bar{\varepsilon}_{2k+2}^{(p)} + O(\|y\|^2).\end{aligned}$$

In conclusion, we have the following result:

THEOREM 2. *Let* $F\colon \mathbf{R}^n \supset D \to \mathbf{R}^n$ *be an analytic function with fixed point* $\mathbf{z} \in \mathring{D}$ *and let* $Q_m(F'(\mathbf{z})) \subset \mathbf{R}^n$ *be the set of vectors* $\mathbf{x}$ *for which* m *is the degree of the minimal polynomial of* $F'(\mathbf{z})$. *Further, let* $\boldsymbol{\varepsilon}_q^{(p)}$ *and* $\tilde{\boldsymbol{\varepsilon}}_q^{(p)}$ *be the vectors obtained by means of the* ϵ*-algorithm from the sequences*

$$\{\mathbf{s}_p;\ \mathbf{s}_{p+1} = F(\mathbf{s}_p)\}_{0\leq p}, \quad \textit{and} \quad \{\tilde{\mathbf{s}}_p;\ \tilde{\mathbf{s}}_p = \mathbf{z} + (F'(\mathbf{z}))^p\,(\mathbf{s}_0 - \mathbf{z})\}_{0\leq p},$$

respectively. Assume that

(i) 1 *is not an eigenvalue of* $F'(\mathbf{z})$,

(ii) *the vectors* $\boldsymbol{\varepsilon}_q^{(p)}$, $\tilde{\boldsymbol{\varepsilon}}_q^{(p)}$, $p + q \leqq 2m$, *exist for all* $\mathbf{s}_0$ *sufficiently close to* $\mathbf{z}$ *with* $\mathbf{s}_0 - \mathbf{z} \in Q_m(F'(\mathbf{z}))$.

Set

$$\boldsymbol{\varepsilon}_{2m}^{(0)} = G(\mathbf{s}_0, \cdots, \mathbf{s}_{2m}) = H_F(\mathbf{s}_0), \tag{9}$$

then the computational procedure

$$\mathbf{x}_{i+1} = H_F(\mathbf{x}_i) \qquad (0 \leqq i)$$

is, for $\mathbf{x}_0$ *sufficiently close to* $\mathbf{z}$ *and* $\mathbf{x}_0 - \mathbf{z} \in Q_m(F'(\mathbf{z}))$, *a quadratically convergent iterative method for the computation of* $\mathbf{z}$.

Proof. By the corollary and Lemma 2, we have

$$H_F(\mathbf{x}_0) = \boldsymbol{\varepsilon}_{2m}^{(0)} = \mathbf{z} + O(\|\mathbf{x}_0 - \mathbf{z}\|^2)$$

for $\mathbf{x}_0 - \mathbf{z} \in Q_m(F'(\mathbf{z}))$.

5. A Modification of the Method. When a system of equations $\mathbf{x} = F(\mathbf{x})$ of order n is to be solved by the ϵ-algorithm, the way of doing this is normally to put $m = n$. Then, we need, for each step of iteration, $4n^3 + 2n^2$ multiplications, $2n^2 + n$ divisions, $6n^3 - n^2$ additions/subtractions and the computation of $\mathbf{s}_p = F(\mathbf{s}_{p-1})$ for $1 \leqq p \leqq 2n$. The computation of the vectors $\mathbf{s}_p$ rather quickly produces a characteristic overflow if the eigenvalues of the Jacobian matrix $F'(\mathbf{x})$ are greater in absolute value than unity near the fixed point $\mathbf{z}$. This disadvantage can possibly be eliminated by replacing the Picard sequence $\mathbf{s}_{p+1} = F(\mathbf{s}_p)$ by

$$\mathbf{s}_{p+1} = F_\alpha(\mathbf{s}_p) = (1 - \alpha)\mathbf{s}_{p-1} + \alpha F(\mathbf{s}_p) \qquad (0 \leqq p)$$

with a suitable α, $0 < \alpha < 1$; in this way, the rate of growth of the components of the vectors $\mathbf{s}_p$ is reduced. If we have, for example, $\rho(F'(\mathbf{z})) = 2$ for the spectral radius ρ of $F'(\mathbf{z})$, we get $\rho(F_\alpha'(\mathbf{z})) = 3/2$ for $\alpha = 1/2$. Those eigenvalues λ of $F'(\mathbf{z})$ for which $|\lambda| < 1$ are thereby increased, but they remain smaller than one in absolute value. Apart from this, convergence is slow if the eigenvalues of $F'(\mathbf{x})$ approach one near $\mathbf{z}$.

The rounding errors affect the computation severely. Perhaps, it is possible that the numerical properties can be improved if a modification proposed by Wynn [20] is applied. If the eigenvalues λ of $F'(\mathbf{x})$ with $|\lambda| < 1$ predominate, we can indicate a modification of the method, by giving up the (theoretic) quadratic convergence, which considerably reduces the amount of work. To achieve this, we replace $2m$ by $2[(m + 1)/2]$ in (9) and obtain for the basic formula of the algorithm

$$\boldsymbol{\varepsilon}_n^{(0)} = G(\mathbf{s}_0, \cdots, \mathbf{s}_n) = H_F^*(\mathbf{s}_0) \tag{9*}$$

in the case $m = n$ even. We need now, per step of iteration, only

$$(n^3 + 8n^2 - 4n)/8$$

multiplications/divisions,

$$(6n^3 - 2n^2)/8$$

additions/subtractions and the computation of $\mathbf{s}_p = F(\mathbf{s}_{p-1})$ for $1 \leqq p \leqq n$.

6. Numerical Examples. Let F: $\mathbf{R}^4 \to \mathbf{R}^4$. In order to illustrate the method of Theorem 2 and its modifications, we consider some systems of quadratic equations $\mathbf{x} = F(\mathbf{x})$ with fixed point $\mathbf{z} = (1, 1, 1, 1)^T$:

$$F(\mathbf{x}) = \mathbf{z} + F'(\mathbf{z})(\mathbf{x} - \mathbf{z}) + \tfrac{1}{2}F''(\mathbf{z})(\mathbf{x} - \mathbf{z})^{(2)}. \tag{10}$$

For the Taylor series (10), we write briefly

$$F(\mathbf{x}) = \mathbf{z} + \mathbf{A}(\mathbf{x} - \mathbf{z}) + Q(\mathbf{x} - \mathbf{z}) \tag{11}$$

and choose for $\mathbf{A}$ (linear) and Q various mappings. The fixed point $\mathbf{z}$ of the systems given in that manner is computed by means of single-precision arithmetic with ten decimal digits. In detail, let $P^{(i)}(\mathbf{x}) = (p_1^{(i)}(\mathbf{x}), \cdots, p_4^{(i)}(\mathbf{x}))^T$ and

$$\begin{aligned}
p_1^{(1)}(\mathbf{x}) &= -(x_1^2 + x_1x_4)/2, & p_1^{(2)}(\mathbf{x}) &= -x_1^2/4,\\
p_2^{(1)}(\mathbf{x}) &= -x_2^2/2, & p_2^{(2)}(\mathbf{x}) &= -x_2^2/4,\\
p_3^{(1)}(\mathbf{x}) &= -x_3^2/2, & p_3^{(2)}(\mathbf{x}) &= -x_3^2/4,\\
p_4^{(1)}(\mathbf{x}) &= -(x_4x_1 + x_4^2)/2, & p_4^{(2)}(\mathbf{x}) &= -x_4^2/4.
\end{aligned}$$

Furthermore, let

$$\begin{aligned}
\mathbf{D}_1 &= (0.9, 0.8, 0.7, 0.6),\\
\mathbf{D}_2 &= (1.5, 0.8, 0.7, 0.6),\\
\mathbf{D}_3 &= (2.0, 0.8, 0.7, 0.6)
\end{aligned}$$

be diagonal matrices and

$$\mathbf{U}_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1\\ 1 & 1 & -1 & -1\\ 1 & -1 & 1 & -1\\ 1 & -1 & -1 & 1 \end{pmatrix}, \qquad \mathbf{U}_2 = \begin{pmatrix} 1 & 1 & 1 & 1\\ 1 & 2 & 3 & 4\\ 1 & 3 & 6 & 10\\ 1 & 4 & 10 & 20 \end{pmatrix}.$$

We remark that $\mathbf{U}_1$ is orthogonal, whereas $\mathbf{U}_2$ is the ill-conditioned Pascal matrix of order four having an integer-valued inverse. It should be pointed out that

$$\left(\frac{\partial P^{(j)}(\mathbf{x} - \mathbf{z})}{\partial \mathbf{x}}\right)\Bigg|_{\mathbf{x}=\mathbf{z}} = 0 \quad (\text{Matrix}) \qquad (j = 1, 2);$$

hence, choosing $Q = P^{(j)}$ in eq. (11), we get, indeed, $F'(\mathbf{z}) = \mathbf{A}$. Now, if $\mathbf{A} = \mathbf{U}_m\mathbf{D}_l\mathbf{U}_m^{-1}$ ($l = 1, 2, 3$; $m = 1, 2$), then $\mathbf{D}_l$ is the matrix of eigenvalues and $\mathbf{U}_m$ is the matrix of eigenvectors of $F'(\mathbf{z})$.

In Examples I–VI, $\mathbf{z}$ is computed by the method proposed in Theorem 2.

I	2.0	$1.2\cdot10^{-2}$	$1.0\cdot10^{-5}$					
II	$7.4\cdot10^{-1}$	$6.6\cdot10^{-1}$	$4.5\cdot10^{-1}$	$1.4\cdot10^{-1}$	$6.8\cdot10^{-2}$	$8.4\cdot10^{-3}$	$7.5\cdot10^{-5}$	$2.5\cdot10^{-8}$
III	$6.0\cdot10^{-1}$	$5.4\cdot10^{-5}$						
IV	0.9	$8.2\cdot10^{-2}$	$2.7\cdot10^{-6}$					
V	2.0	$9.9\cdot10^{-1}$	$2.8\cdot10^{-6}$					
VI	1.8	$1.5\cdot10^{-1}$	$3.7\cdot10^{-2}$	$1.1\cdot10^{-2}$	$4.4\cdot10^{-3}$	$2.2\cdot10^{-3}$	$1.1\cdot10^{-3}$	$3.6\cdot10^{-4}$
VII	1.9	$8.6\cdot10^{-2}$	$5.5\cdot10^{-3}$	$5.0\cdot10^{-5}$	$5.2\cdot10^{-8}$			
VIII	$7.9\cdot10^{-1}$	$6.5\cdot10^{-1}$	$4.3\cdot10^{-1}$	$1.3\cdot10^{-1}$	$4.6\cdot10^{-2}$	$1.6\cdot10^{-3}$	$5.1\cdot10^{-5}$	$4.8\cdot10^{-8}$
IX	$6.0\cdot10^{-1}$	$6.0\cdot10^{-3}$	$4.0\cdot10^{-6}$					
X	$8.9\cdot10^{-1}$	$1.1\cdot10^{-1}$	$3.2\cdot10^{-4}$	$2.4\cdot10^{-7}$				
XI	$3.8\cdot10^{-1}$	$5.1\cdot10^{-1}$	$1.1\cdot10^{-1}$	$3.8\cdot10^{-4}$	$1.1\cdot10^{-6}$			
XII	1.5	$3.2\cdot10^{-1}$	$1.1\cdot10^{-1}$	$4.6\cdot10^{-2}$	$2.1\cdot10^{-2}$	$1.0\cdot10^{-2}$	$9.5\cdot10^{-4}$	$6.5\cdot10^{-4}$

Example I: $F'(\mathrm{z}) = \mathbf{U}_1\mathbf{D}_1\mathbf{U}_1^{-1}$, $Q = P^{(1)}$, initial vector $\mathbf{x}_0 = 2\mathrm{z}$;
Example II: $F'(\mathrm{z}) = \mathbf{U}_1\mathbf{D}_2\mathbf{U}_1^{-1}$, $Q = P^{(1)}$, $\mathbf{x}_0 = \mathbf{0}$;
Example III: as Example II but using $\mathbf{x}_0 = 2\mathrm{z}$;
Example IV: $F'(\mathrm{z}) = \mathbf{U}_2\mathbf{D}_2\mathbf{U}_2^{-1}$, $Q = P^{(2)}$, $\mathbf{x}_0 = 0.5\mathrm{z}$;
Example V: as Example IV but using $\mathbf{x}_0 = 1.5\mathrm{z}$;
Example VI: $F'(\mathrm{z}) = \mathbf{U}_1\mathbf{D}_3\mathbf{U}_1^{-1}$, $Q = P^{(1)}$, $\mathbf{x}_0 = 2\mathrm{z}$, using the modified Picard sequence $\mathrm{s}_{p+1} = F_\alpha(\mathrm{s}_p)$ with $\alpha = 1/2$.

The Examples VII–XII are the same as Examples I–VI, respectively, but z is computed using formula (9*) instead of (9).

The above table contains in column i ($1 \leqq i \leqq 8$) the values $\|\mathbf{x}_i - \mathbf{x}_{i-1}\|$ (compare Theorem 2) with rounded mantissae; values for which $\|\mathrm{z} - \mathbf{x}_i\| < 5.0 \cdot 10^{-9}$ (the process has then terminated) are omitted. Generally speaking, we have found that the algorithm produces better results if the Jacobian matrix of the given system $\mathbf{x} = F(\mathbf{x})$ is symmetric. Finally, it should be mentioned that it seems to be impossible at the moment to say more about the error than that it is of quadratic order.

Acknowledgement. The author is indebted to the referee for many suggestions used in the preparation of this paper.

Universität Mannheim
Lehrstuhl für Mathematik IV
Mannheim 6800
Germany

1. A. C. Aitken, "On Bernoulli's numerical solution of algebraic equations," *Proc. Roy. Soc. Edinburgh Sect. A*, v. 46, 1926, pp. 289–305.
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. J. Dieudonné, *Foundations of Modern Analysis*, Pure and Appl. Math., vol. **10**, Academic Press, New York-London, 1960. MR **22** #11074.
4. E. Gekeler, "Über den ε-Algorithmus von Wynn," *Z. Angew. Math. Mech.* (To appear.)
5. T. N. E. Greville, *On Some Conjectures of P. Wynn Concerning the ε-Algorithm*, University of Wisconsin Math. Res. Center Report #877, 1968.
6. C. G. J. Jacobi, "Über die Darstellung einer Reihe gegebener Werte durch eine gebrochene rationale Funktion," *J. Reine Angew. Math.*, v. 30, 1846, pp. 127–156.
7. J. B. McLeod, *A Fundamental Result in the Theory of the ε-Algorithm*, University of Wisconsin Math. Res. Center Report #685, 1966.
8. E. H. Moore, "On the reciprocal of the general algebraic matrix," *Bull. Amer. Math. Soc.*, v. 26, 1920, pp. 394–395. (Abstract.)
9. R. Penrose, "A generalized inverse for matrices," *Proc. Cambridge Philos. Soc.*, v. 51, 1955, pp. 406–413. MR **16**, 1082.
10. L. D. Pyle, "A generalized inverse ε-algorithm for constructing intersection projection matrices, with applications," *Numer. Math.*, v. 10, 1967, pp. 86–102. MR **36** #2296.
11. R. J. Schmidt, "On the numerical solution of linear simultaneous equations by an iterative method," *Philos. Mag.*, v. (7) 32, 1941, pp. 369–383. MR **3**, 276.
12. D. Shanks, "Non-linear transformations of divergent and slowly convergent sequences," *J. Mathematical Phys.*, v. 34, 1955, pp. 1–42. MR **16**, 961.
13. P. Wynn, "On a device for computing the $e_m(S_n)$ transformation," *MTAC*, v. 10, 1956, pp. 91–96. MR **18**, 801.
14. P. Wynn, "On a procrustean technique for the numerical transformation of slowly convergent sequences and series," *Proc., Cambridge Philos. Soc.*, v. 52, 1956, pp. 663–671. MR **18**, 478.
P. Wynn, "The rational approximation of functions which are formally defined by a power series expansion," *Math. Comp.*, v. 14, 1960, pp. 147–186. MR **22** #7244.
16. P. Wynn, "On repeated application of the ε-algorithm," *Chiffres*, v. 4, 1961, pp. 19–22. MR **26** #6639.

17. P. Wynn, "A comparison between the numerical performances of the Euler transformation and the epsilon algorithm," *Chiffres*, v. 4, 1961, pp. 23–29.

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, *Acceleration Techniques in Numerical Analysis, With Particular Reference to Problems in One Independent Variable*, Proc. IFIP Congress 1962, North-Holland, Amsterdam, 1963, pp. 149–156.

20. P. Wynn, "Singular rules for certain non-linear algorithms," *Nordisk Tidskr. Informationsbehandling*, v. 3, 1963, pp. 175–195. MR **29** #4219.

21. P. Wynn, "Continued fractions whose coefficients obey a non-commutative law of multiplication," *Arch. Rational Mech. Anal.*, v. 12, 1963, pp. 273–312. MR **26** #2766.

22. P. Wynn, "General purpose vector epsilon algorithm ALGOL procedures," *Numer. Math.*, v. 6, 1964, pp. 22–36. MR **29** #4220.

23. P. Wynn, *Upon a Conjecture Concerning a Method for Solving Linear Equations, and Certain Other Matters*, University of Wisconsin, Math. Res. Center Report #626, 1966.

24. P. Wynn, "Vector continued fractions," *Linear Algebra Appl.*, v. 1, 1968, pp. 357–395. MR **38** #176.