SIAM J. NUMER. ANAL.
Vol. 23, No. 1, February 1986

013

ACCELERATION OF CONVERGENCE OF VECTOR SEQUENCES*

AVRAM SIDI†, WILLIAM F. FORD‡ AND DAVID A. SMITH§

Abstract. A general approach to the construction of convergence acceleration methods for vector sequences is proposed. Using this approach, one can generate some known methods, such as the minimal polynomial extrapolation, the reduced rank extrapolation, and the topological epsilon algorithm, and also some new ones. Some of the new methods are easier to implement than the known methods and are observed to have similar numerical properties. The convergence analysis of these new methods is carried out, and it is shown that they are especially suitable for accelerating the convergence of vector sequences that are obtained when one solves linear systems of equations iteratively. A stability analysis is also given, and numerical examples are provided. The convergence and stability properties of the topological epsilon algorithm are likewise given.

1. Introduction. In a recent work by the present authors [8] a survey of convergence acceleration methods for sequences of vectors is given, and five of these methods are tested and compared numerically using a process that has been termed *cycling*: the minimal polynomial extrapolation (MPE) of Cabay and Jackson [2], the reduced rank extrapolation (RRE) of Eddy [3] and Mešina [4] (the equivalence of the methods proposed in [3] and [4] is shown in [8]), the scalar epsilon algorithm (SEA) of Wynn [9], the vector epsilon algorithm (VEA) of Wynn [10], and the topological epsilon algorithm (TEA) of Brezinski, see [1, pp. 172–205]. One of the conclusions of this survey is that the MPE and RRE have about the same properties, and in general, have better convergence than others, in the sense that the MPE and the RRE achieve a given level of accuracy with fewer vectors than the SEA, VEA, and TEA. VEA and SEA are also similar in performance, except that the latter is more prone to numerical instability problems. However, TEA, while interesting from a theoretical point of view, appears to be not as effective as either VEA or SEA; see [8] for further details.

All of the methods above have the following important properties:

(1) It is observed numerically that in many instances they accelerate the convergence of a slowly converging vector sequence and they make a diverging sequence converge to an "anti-limit" that has an immediate interpretation.

(2) They depend solely on the given vector sequence whose convergence is being accelerated; they do not depend on how the vector sequence is generated.

(3) Their implementation is straightforward.

For more details and an extensive bibliography see [8].

It turns out that the implementation of the MPE and RRE requires the least-squares solution of an overdetermined and in general inconsistent set of linear equations, the number of the equations in this set being equal to the dimension of the vectors in the given sequence. For many practical problems, the dimension of these vectors may be finite but very large; consequently, one may have to store a large rectangular matrix in memory, making the MPE and RRE somewhat expensive in both storage and time.

* Received by the editors June 23, 1984, and in revised form February 15, 1985. Paper presented at the SIAM 30th Anniversary Meeting, Stanford, California, July 1982.

† Computer Science Department, Technion-Israel Institute of Technology, Haifa 32000, Israel. The work of this author was done partly while he was a National Research Council-NASA Research Associate and partly under NASA contract NAS3-23606.

‡ Computer Services Division, National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio 44135.

§ Department of Mathematics, Duke University, Durham, North Carolina 27706. The work of this author was supported by NASA grant NSG-3160.

Therefore, it would be desirable to have methods as efficient as MPE and RRE but less demanding in storage and time.

In the next section a general framework for deriving convergence acceleration methods for vector sequences—vector accelerators for short—is proposed. Within this framework one can derive several methods, some old (MPE, RRE, and TEA) and some new. It turns out that one of the new methods is very similar to the MPE and RRE, but does not require the use of the least-squares method, requires very little storage, and, at least numerically, is as efficient as MPE. The convergence analysis of this method, which we shall call the modified MPE (MMPE), is carried out in § 3 for a class of vector sequences that includes those arising from the iterative solution of systems of linear equations. We prove that this method is a bona fide convergence acceleration method, and also provide its rate of acceleration. The stability properties of this method are taken up in § 4. In § 5 the convergence and stability properties of the TEA are analyzed using the techniques of §§ 3 and 4. Finally in § 6 we test the MMPE on some examples numerically, and compare it with the MPE. On the basis of this comparison one could conclude that the MPE and MMPE have similar performances, and this is indeed the case as has been shown by the first author in a recent work [7] (this issue, pp. 197–209).

2. Development of vector accelerators. In this section we shall develop a general framework within which one can derive vector accelerators of different kinds. We shall motivate this development in the way Shanks [6] motivates his development of the e_k-transformation for scalar sequences.

2.1. The Shanks transformation. Shanks starts with a scalar sequence X_m, $m = 0, 1, \cdots$, that has the property

$$X_m \sim S + \sum_{i=1}^{\infty} a_i \lambda_i^m \quad \text{as } m \to \infty, \tag{2.1}$$

where S, a_i, and λ_i are constants independent of m, $\lambda_i \neq 1$, $i = 1, 2, \cdots$, $\lambda_i \neq \lambda_j$ for all $i \neq j$, and $|\lambda_1| \geqq |\lambda_2| \geqq \cdots$. In (2.1) S is $\lim_{m\to\infty} X_m$ if all $|\lambda_i| < 1$; otherwise, S is called the *anti-limit* of the sequence X_m, $m = 0, 1, \cdots$. As one way of approximating S, Shanks proposes to solve the set of $2k+1$ nonlinear equations

$$X_m = S_{n,k} + \sum_{i=1}^{k} \tilde{a}_i \tilde{\lambda}_i^m, \qquad n \leqq m \leqq n+2k, \tag{2.2}$$

for $S_{n,k}$, which is taken to be an approximation to S, with $\tilde{a}_i$, $\tilde{\lambda}_i$, $i = 1, \cdots, k$, being the rest of the unknowns. The solution $S_{n,k}$ turns out to have the following determinant representation:

$$S_{n,k} = \frac{\begin{vmatrix} X_n & X_{n+1} & \cdots & X_{n+k} \\ \Delta X_n & \Delta X_{n+1} & \cdots & \Delta X_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta X_{n+k-1} & \Delta X_{n+k} & \cdots & \Delta X_{n+2k-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \Delta X_n & \Delta X_{n+1} & \cdots & \Delta X_{n+k} \\ \vdots & \vdots & & \vdots \\ \Delta X_{n+k-1} & \Delta X_{n+k} & \cdots & \Delta X_{n+2k-1} \end{vmatrix}}, \tag{2.3}$$

where Δ is the forward difference operator defined by $\Delta b_i = b_{i+1} - b_i$, $\Delta^p b_i = \Delta(\Delta^{p-1} b_i)$, $p \geqq 2$, provided the determinant in the denominator of (2.3) is nonzero.

Two equivalent formulations follow from (2.3), independent of any reference to the nonlinear equations in (2.2).

(a) With $S_{n,k}$ as given in (2.3), there are k parameters β_i, $i=0,1,\cdots,k-1$, which solve the system of $k+1$ linear equations

$$X_m = S_{n,k} + \sum_{i=0}^{k-1} \beta_i \Delta X_{m+i}, \qquad n \leqq m \leqq n+k, \tag{2.4}$$

as can be verified by solving (2.4) for $S_{n,k}$ by Cramer's rule. By taking the differences of the equations in (2.4), we see that the β_i satisfy

$$\Delta X_m = \sum_{i=0}^{k-1} \beta_i \Delta^2 X_{m+i}, \qquad n \leqq m \leqq n+k-1, \tag{2.5}$$

and that, once the β_i have been determined from (2.5), $S_{n,k}$ can be computed from one of the equations in (2.4), say that for which $m=n$.

(b) From (2.4) it follows that $S_{n,k}$, along with the parameters γ_i, $i=0,1,\cdots,k$, satisfies the system of linear equations

$$S_{n,k} = \sum_{i=0}^{k} \gamma_i X_{m+i}, \qquad n \leqq m \leqq n+k, \tag{2.6}$$

subject to

$$\sum_{i=0}^{k} \gamma_i = 1. \tag{2.7}$$

By taking the differences of the equations in (2.6), we see that the γ_i satisfy

$$0 = \sum_{i=0}^{k} \gamma_i \Delta X_{m+i}, \qquad n \leqq m \leqq n+k-1, \tag{2.8}$$

subject to (2.7), see also [1, pp. 52–53]. Once the γ_i have been determined from (2.7) and (2.8), $S_{n,k}$ can be computed from one of the equations in (2.6). Furthermore, if $\gamma_k \neq 0$, then (2.7) and (2.8) are equivalent to

$$\sum_{i=0}^{k-1} c_i \Delta X_{m+i} = -\Delta X_{m+k}, \qquad n \leqq m \leqq n+k-1, \tag{2.9}$$

where

$$\gamma_i = \frac{c_i}{\sum_{j=0}^{k} c_j}, \qquad 0 \leqq i \leqq k, \quad c_k = 1, \tag{2.10}$$

provided $\sum_{j=0}^{k} c_j \neq 0$.

It has been proved by Wynn [11] that $S_{n,k}$, when computed from a sequence X_m, $m=0,1,\cdots$, that is of the form given in (2.1), converges to S as $n \to \infty$ (k fixed), under certain conditions on the λ_i, faster than X_n itself. Wynn actually gives rates of convergence for $S_{n,k}$ for $n \to \infty$.

2.2. Derivation of vector accelerators. Let us now consider a sequence of vectors, x_m, $m=0,1,\cdots$, in a general normed vector space B, satisfying

$$x_m \sim s + \sum_{i=1}^{\infty} v_i \lambda_i^m \quad \text{as } m \to \infty, \tag{2.11}$$

where s and v_i are vectors in B, and λ_i are scalars, independent of m, $\lambda_i \neq 1$, $i=$

$1, 2, \cdots, \lambda_i \neq \lambda_j$ for all $i \neq j$, and $|\lambda_1| \geqq |\lambda_2| \geqq \cdots$. We also assume that there can be only a finite number of λ_i having the same modulus. The meaning of (2.11) is that, for any integer $N > 0$, there exist a positive constant K and a positive integer m_0 that depend only on N, such that for every $m \geqq m_0$

$$\left\| x_m - s - \sum_{i=1}^{N-1} v_i \lambda_i^m \right\| \leqq K |\lambda_N|^m, \tag{2.11a}$$

with $\|\cdot\|$ being the norm associated with the vector space B.

A simple example of such a sequence is that produced by a matrix iterative technique for solving the linear system of equations

$$x = Ax + b, \tag{2.12}$$

where A is a nondefective $M \times M$ matrix, and b and x are M-dimensional column vectors. If s is the solution to (2.12), and for given x_0, the vectors x_m are generated by

$$x_{m+1} = Ax_m + b, \qquad m = 0, 1, \cdots, \tag{2.13}$$

then

$$x_m = s + \sum_{i=1}^{M'} \alpha_i v_i \lambda_i^m, \qquad m = 0, 1, \cdots, \tag{2.14}$$

where α_i are scalars, λ_i and v_i are the eigenvalues and corresponding eigenvectors of the matrix A, and $M' \leqq M$ is the number of the distinct eigenvalues.

The condition stated in (2.11) is analogous to that stated in (2.1) for scalar sequences. Since, as shown in [11], the Shanks transformation accelerates the convergence of scalar sequences satisfying (2.1), we expect that its extensions to vector sequences, through the formulations (a) and (b) following (2.3), may also produce acceleration of convergence for vector sequences satisfying (2.11). The extensions of the two formulations can be achieved as follows:

Approach (a). In equations (2.5) replace X_j by x_j, and "solve" in some sense the resulting *overdetermined* (and in general inconsistent) system

$$\Delta x_m = \sum_{i=0}^{k-1} \beta_i \Delta^2 x_{m+i}, \qquad n \leqq m \leqq n+k-1, \tag{2.15}$$

for the β_i. Once the β_i have been determined, compute the approximation $s_{n,k}$ to s by

$$s_{n,k} = x_n - \sum_{i=0}^{k-1} \beta_i \Delta X_{n+i}, \tag{2.16}$$

which is obtained by replacing $S_{n,k}$ and X_j in (2.4) by $s_{n,k}$ and x_j, respectively, and considering $m = n$. In general $m \geqq n$ can be considered.

Approach (b). In equations (2.9) replace X_j by x_j, and "solve" in some sense the resulting *overdetermined* (and in general inconsistent) system

$$\sum_{i=0}^{k-1} c_i \Delta x_{m+i} = -\Delta x_{m+k}, \qquad n \leqq m \leqq n+k-1, \tag{2.17}$$

for the c_i. Once the c_i, $i = 0, 1, \cdots, k-1$, have been determined, set

$$c_k = 1, \qquad \gamma_i = \frac{c_i}{\sum_{j=0}^{k} c_j}, \qquad 0 \leqq j \leqq k, \tag{2.18}$$

assuming $\sum_{j=0}^{k} c_j \neq 0$. Finally compute the approximation $s_{n,k}$ to s by

$$s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i}, \tag{2.19}$$

which is obtained by replacing $S_{n,k}$ and X_j in (2.6) by $s_{n,k}$ and x_j, respectively, and considering $m = n$. In general $m \geqq n$ can be considered.

We see that for both approaches, we need to "solve" an overdetermined and, in general, inconsistent system of equations of the form

$$\sum_{i=0}^{k-1} d_i w_{m+i} = \tilde{w}_m, \qquad n \leqq m \leqq n+k-1, \tag{2.20}$$

where w_j and $\tilde{w}_j$ are vectors in B, and d_i are unknown scalars. If r, the dimension of B, is greater than k, then even one of the equations in (2.20) gives rise to an overdetermined system of r equations. We can, however, propose various ways for obtaining a set of d_i that solves (2.20) in some sense. In what follows, we give three such methods, with the understanding that other methods can also be proposed.

Method (1). Assuming $r > k$, consider only one of the equations in (2.20), namely that with $m = n$, and solve for the d_i that minimize some norm of the vector $\Delta = \sum_{i=0}^{k-1} d_i w_{n+i} - \tilde{w}_n$. Depending on the norm being used, different acceleration methods will be obtained. For example, if r is finite and the (weighted) l_p norms are used with $p = 1, 2, \infty$, then the determination of the d_i becomes relatively easy. For $p = 2$ the solution can be achieved by using any one of the least-squares packages available, and for $p = 1$ and $p = \infty$ the minimization problems can be solved by using linear programming techniques; see the review paper by Rabinowitz [5]. The l_2 norm with equal weights gives rise to RRE for Approach (a) and MPE for Approach (b). The rest of the acceleration methods have not appeared in the literature before.

Method (2). Assuming $r > k$, consider only one of the equations in (2.20), namely that with $m = n$, and obtain the d_i by solving the system of k equations

$$\sum_{i=0}^{k-1} d_i Q_j(w_{n+i}) = Q_j(\tilde{w}_n), \qquad j = 1, \cdots, k, \tag{2.21}$$

where Q_j are linearly independent bounded linear functionals on the space B. When B is an inner product space, we can take $Q_j(y) = (q_j, y)$, where q_j are vectors in B and $(\cdot, \cdot)$ is the inner product associated with B. If r is finite, and the vector q_j is chosen to be the i_jth unit vector, that is, $(q_j, z) = i_j$th component of z, then the method above is equivalent to demanding that only k out of the r equations be satisfied, namely those corresponding to the i_jth components, $j = 1, \cdots, k$. Obviously such an acceleration method demands less storage and time for its implementation than methods like the MPE and RRE. It is not difficult to see that Approaches (a) and (b) both give the same acceleration method, which has not been given in the literature before. Due to its similarity to the MPE, we shall call this method the *modified MPE* (MMPE). In §§ 3 and 4 we shall analyze the convergence and stability properties of this method in detail.

Method (3). Consider all the equations in (2.20) and obtain the d_i by solving the system of k equations

$$\sum_{i=0}^{k-1} d_i Q(w_{m+i}) = Q(\tilde{w}_m), \qquad n \leqq m \leqq n+k-1, \tag{2.22}$$

where Q is a bounded linear functional on the space B. In this case Approaches (a) and (b) give the same method, and this method is nothing but the TEA. In § 5 we shall analyze the convergence and stability properties of this method in detail. By comparing (2.21) and (2.22), we see that for $k=1$ the MMPE and the TEA are identical when we choose $Q_1 = Q$.

Finally we note that all of the methods obtained as above are nonlinear in the x_i.

3. Convergence analysis of MMPE. As in § 2.2, we start with a sequence of vectors x_i, $i=0, 1, \cdots$, in a normed vector space B with norm $\|\cdot\|$, that has a limit or anti-limit s. We write $u_i = \Delta x_i = x_{i+1} - x_i$, $i = 0, 1, \cdots$. Then the MMPE, as obtained from Approach (a) or (b) in conjunction with Method (2) (see § 2.2), can be summarized (and reformulated) as follows: By Approach (b), the approximation $s_{n,k}$ to s is given as

$$s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i}, \tag{3.1}$$

where γ_i are obtained from

$$\begin{aligned} &\sum_{i=0}^{k} \gamma_i = 1, \\ &\sum_{i=0}^{k} \gamma_i Q_j(u_{n+i}) = 0, \qquad 1 \leqq j \leqq k. \end{aligned} \tag{3.2}$$

When $\gamma_k \neq 0$, equations (3.2) are equivalent to (2.18) and (2.21), in which $d_i = c_i$, $w_{n+i} = u_{n+i}$, $0 \leqq i \leqq k-1$, and $\tilde{w}_n = -u_{n+k}$, as can be verified by inspection.

We denote the scalars $Q_j(u_m)$ by $u_{m,j}$ for $1 \leqq j \leqq k$ and $m \geqq 0$, and we define $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ to be the determinant

$$D(\sigma_0, \sigma_1, \cdots, \sigma_k) = \begin{vmatrix} \sigma_0 & \sigma_1 & \cdots & \sigma_k \\ u_{n,1} & u_{n+1,1} & \cdots & u_{n+k,1} \\ u_{n,2} & u_{n+1,2} & \cdots & u_{n+k,2} \\ \vdots & \vdots & & \vdots \\ u_{n,k} & u_{n+1,k} & \cdots & u_{n+k,k} \end{vmatrix}, \tag{3.3}$$

when σ_i are scalars. Let N_i be the cofactor of σ_i in the first row expansion of this determinant. Then

$$D(\sigma_0, \sigma_1, \cdots, \sigma_k) = \sum_{i=0}^{k} \sigma_i N_i. \tag{3.4}$$

When σ_i are vectors, we again let $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ be defined by the determinant in (3.3), and take (3.4) as the *interpretation* of this determinant. Thus $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ is a scalar (or vector) if the σ_i are scalars (or vectors).

Solving the system in (3.2) by Cramer's rule, we obtain

$$\gamma_i = \frac{N_i}{D(1, 1, \cdots, 1)} = \frac{N_i}{\sum_{j=0}^{k} N_j}, \qquad 0 \leqq i \leqq k, \tag{3.5}$$

provided $D(1, 1, \cdots, 1) \neq 0$; in what follows, we assume that this is so.

By (3.5), (3.1), and (3.4), we can finally express $s_{n,k}$ as

$$s_{n,k} = \frac{D(x_n, x_{n+1}, \cdots, x_{n+k})}{D(1, 1, \cdots, 1)}. \tag{3.6}$$

LEMMA 3.1. *The error in $s_{n,k}$ can be expressed as*

$$s_{n,k}-s=\frac{D(x_n-s, x_{n+1}-s, \cdots, x_{n+k}-s)}{D(1,1,\cdots,1)}. \tag{3.7}$$

Proof. (3.7) follows easily from (3.4) to (3.6). □

In the sequel we shall assume that the vectors x_m satisfy (2.11). Without loss of generality we shall also assume that $\lambda_i \neq 0$ and $v_i \neq 0$ for all $i \geqq 1$. Then

$$u_m \sim \sum_{i=1}^{\infty} z_i \lambda_i^m \quad \text{as } m \to \infty, \tag{3.8}$$

where $z_i=(\lambda_i-1)v_i$, $i=1,2,\cdots$. Since $\lambda_i \neq 1$ and $v_i \neq 0$ for all $i \geqq 1$, we have $z_i \neq 0$ for all $i \geqq 1$. In addition, by (2.11), for any operator T in the dual space of B,

$$T(u_m) \sim \sum_{i=1}^{\infty} T(z_i)\lambda_i^m \quad \text{as } m \to \infty. \tag{3.9}$$

Consequently

$$u_{m,j} \sim \sum_{i=1}^{\infty} z_{i,j}\lambda_i^m \quad \text{as } m \to \infty, \tag{3.10}$$

where $z_{i,j}=Q_j(z_i)$, $i \geqq 1$, $1 \leqq j \leqq k$.

Note that when the sequence x_m, $m=0,1,\cdots$, is generated by the matrix iterative method described in § 2.2, the summations over i in (3.8), (3.9), and (3.10) extend as far as M', which is a finite number; therefore, (3.9) and hence (3.10) hold automatically for this case, and $\sim$ is replaced by $=$.

The following theorem is the main result of this section.

THEOREM 3.2. *Define $Q_j(v_i)=v_{i,j}$, $i \geqq 1$, $1 \leqq j \leqq k$, and let*

$$F=\begin{vmatrix} v_{1,1} & v_{2,1} & \cdots & v_{k,1} \\ v_{1,2} & v_{2,2} & \cdots & v_{k,2} \\ \vdots & \vdots & & \vdots \\ v_{1,k} & v_{2,k} & \cdots & v_{k,k} \end{vmatrix} \neq 0. \tag{3.11}$$

Assume that the v_i are linearly independent, and that the λ_i satisfy

$$|\lambda_1| \geqq \cdots \geqq |\lambda_k| > |\lambda_{k+1}| \geqq |\lambda_{k+2}| \geqq \cdots. \tag{3.12}$$

Then, for all sufficiently large n, $D(1,1,\cdots,1) \neq 0$; hence, $s_{n,k}$, as given in (3.6), exists. Furthermore,

$$s_{n,k}-s=\Gamma(n)\lambda_{k+1}^n[1+o(1)] \quad \text{as } n \to \infty, \tag{3.13}$$

where the vector $\Gamma(n)$ is nonzero and bounded for all sufficiently large n. If, in addition, $|\lambda_{k+1}|>|\lambda_{k+2}|$, then

$$\Gamma(n)=\frac{1}{F}\begin{vmatrix} v_1 & v_2 & \cdots & v_{k+1} \\ z_{1,1} & z_{2,1} & \cdots & z_{k+1,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{k+1,2} \\ \vdots & \vdots & & \vdots \\ z_{1,k} & z_{2,k} & \cdots & z_{k+1,k} \end{vmatrix} \prod_{i=1}^{k} \frac{(\lambda_{k+1}-\lambda_i)}{(\lambda_i-1)^2}. \tag{3.14}$$

Proof. For simplicity of notation we shall sometimes denote $G_n=D(x_n-s,\cdots,x_{n+k}-s)$ and $H_n=D(1,\cdots,1)$, and we shall shorten "$\alpha_n \sim \beta_n$ as $n \to \infty$" to "$\alpha_n \sim \beta_n$."

By (3.3) and (3.10) we have

$$H_n \sim \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^n & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+1} & \cdots & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+k} \\ \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^n & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+1} & \cdots & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+k} \\ \vdots & \vdots & & \vdots \\ \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^n & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+1} & \cdots & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+k} \end{vmatrix}. \tag{3.15}$$

We are allowed to write (3.15) since $D(1, 1, \cdots, 1)$, being a determinant, is the sum of a finite number ($k!$) of products of $u_{i,j}$, and its asymptotic expansion as $n \to \infty$ is the sum of the products of the asymptotic expansions of the respective $u_{i,j}$. By the multilinearity property of determinants, (3.15) is equivalent to

$$H_n \sim \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} \left(\prod_{p=1}^{k} z_{i_p,p} \right) \left(\prod_{p=1}^{k} \lambda_{i_p}^n \right) V(1, \lambda_{i_1}, \cdots, \lambda_{i_k}), \tag{3.16}$$

where $V(\xi_0, \xi_1, \cdots, \xi_k)$ is the Vandermonde determinant defined by

$$V(\xi_0, \xi_1, \cdots, \xi_k) = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ \xi_0 & \xi_1 & \cdots & \xi_k \\ \xi_0^2 & \xi_1^2 & \cdots & \xi_k^2 \\ \vdots & \vdots & & \vdots \\ \xi_0^k & \xi_1^k & \cdots & \xi_k^k \end{vmatrix}. \tag{3.17}$$

Since $V(\xi_0, \cdots, \xi_k)$ is odd under an interchange of the indices $0, 1, \cdots, k$, by Lemma A.1 given in the appendix to this work, (3.16) can be expressed as

$$H_n \sim \sum_{1 \leqq i_1 < i_2 < \cdots < i_k}^{\infty} \begin{vmatrix} z_{i_1,1} & z_{i_2,1} & \cdots & z_{i_k,1} \\ z_{i_1,2} & z_{i_2,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_1,k} & z_{i_2,k} & \cdots & z_{i_k,k} \end{vmatrix} V(1, \lambda_{i_1}, \cdots, \lambda_{i_k}) \left(\prod_{p=1}^{k} \lambda_{i_p}^n \right). \tag{3.18}$$

By (3.12), the most dominant term in the summation on the right side of (3.18) as $n \to \infty$ is that for which $i_1 = 1$, $i_2 = 2, \cdots, i_k = k$, provided that the determinant

$$\tilde{F} = \begin{vmatrix} z_{1,1} & z_{2,1} & \cdots & z_{k,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{k,2} \\ \vdots & \vdots & & \vdots \\ z_{1,k} & z_{2,k} & \cdots & z_{k,k} \end{vmatrix} \tag{3.19}$$

is nonzero. But since $z_i = (\lambda_i - 1)v_i$, $i \geqq 1$, we have

$$\tilde{F} = \left[\prod_{i=1}^{k} (\lambda_i - 1) \right] F, \tag{3.20}$$

where F is as defined in (3.11). Since $F \neq 0$ by assumption, $\tilde{F} \neq 0$ too; hence the first part of the theorem follows, with

$$D(1, \cdots, 1) = \left[\prod_{i=1}^{k} (\lambda_i - 1) \right] F \left(\prod_{p=1}^{k} \lambda_p^n \right) V(1, \lambda_1, \cdots, \lambda_k)[1 + o(1)] \quad \text{as } n \to \infty. \tag{3.21}$$

For the proof of the second part we proceed similarly. By (2.11), (3.3), and (3.10), we have

$$G_n \sim \begin{vmatrix} \sum_{i_0=1}^{\infty} v_{i_0}\lambda_{i_0}^{n} & \sum_{i_0=1}^{\infty} v_{i_0}\lambda_{i_0}^{n+1} & \cdots & \sum_{i_0=1}^{\infty} v_{i_0}\lambda_{i_0}^{n+k} \\ \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n} & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+1} & \cdots & \sum_{i_1=1}^{\infty} z_{i_1,1}\lambda_{i_1}^{n+k} \\ \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n} & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+1} & \cdots & \sum_{i_2=1}^{\infty} z_{i_2,2}\lambda_{i_2}^{n+k} \\ \vdots & \vdots & & \vdots \\ \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n} & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+1} & \cdots & \sum_{i_k=1}^{\infty} z_{i_k,k}\lambda_{i_k}^{n+k} \end{vmatrix}. \tag{3.22}$$

Again from the multilinearity property of determinants

$$G_n \sim \sum_{i_0=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} v_{i_0}\left(\prod_{p=1}^{k} z_{i_p,p}\right)\left(\prod_{p=0}^{k} \lambda_{i_p}^{n}\right) V(\lambda_{i_0}, \lambda_{i_1}, \cdots, \lambda_{i_k}). \tag{3.23}$$

By Lemma A.1 given in the appendix to this work, (3.23) can be expressed as

$$G_n \sim \sum_{1 \leqq i_0 < i_1 < \cdots < i_k} \begin{vmatrix} v_{i_0} & v_{i_1} & \cdots & v_{i_k} \\ z_{i_0,1} & z_{i_1,1} & \cdots & z_{i_k,1} \\ z_{i_0,2} & z_{i_1,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_0,k} & z_{i_1,k} & \cdots & z_{i_k,k} \end{vmatrix} V(\lambda_{i_0}, \lambda_{i_1}, \cdots, \lambda_{i_k})\left(\prod_{p=0}^{k} \lambda_{i_p}^{n}\right). \tag{3.24}$$

By the assumptions made following (2.11), there is a finite number of λ_i with moduli equal to $|\lambda_{k+1}|$. Let $|\lambda_{k+1}| = \cdots = |\lambda_{k+r}| > |\lambda_{k+r+1}|$. From this and (3.12), it follows that the dominant term on the right side of (3.24) is the sum of those terms with indices $i_0 = 1$, $i_1 = 2, \cdots, i_{k-1} = k$, $i_k = k+l$, $l = 1, 2, \cdots, r$, that is,

$$\begin{aligned} D(x_n - s, \cdots, x_{n+k} - s) = \left(\prod_{p=1}^{k} \lambda_p^{n}\right) \sum_{l=1}^{r} \lambda_{k+l}^{n} V(\lambda_1, \cdots, \lambda_k, \lambda_{k+l}) \\ \cdot \begin{vmatrix} v_1 & \cdots & v_k & v_{k+l} \\ z_{1,1} & \cdots & z_{k,1} & z_{k+l,1} \\ z_{1,2} & \cdots & z_{k,2} & z_{k+l,2} \\ \vdots & & \vdots & \vdots \\ z_{1,k} & \cdots & z_{k,k} & z_{k+l,k} \end{vmatrix} [1 + o(1)] \quad \text{as } n \to \infty. \end{aligned} \tag{3.25}$$

Now the cofactor of v_{k+l} in the determinant in (3.25) is $\tilde{F}$, which is nonzero since $F \neq 0$. Therefore, for n sufficiently large, the coefficients of $v_{k+1}, \cdots, v_{k+r}$ are nonzero. Since we also assumed that the v_i are linearly independent, the summation in (3.25) is never zero. Combining (3.21) and (3.25) in (3.7) results in (3.13). If $|\lambda_{k+1}| > |\lambda_{k+2}|$, then $r = 1$. In this case (3.14) follows from (3.21), (3.25), and the fact that

$$V(\xi_0, \xi_1, \cdots, \xi_k) = \prod_{0 \leqq i < j \leqq k} (\xi_j - \xi_i). \tag{3.26}$$

□

Note. The condition on the determinant given in (3.11) has some interesting implications when the normed vector space B is a complete inner product space. In this case, for each Q_j in the dual space of B, there exists a unique vector q_j in B, such that $Q_j(z)=(q_j, z)$ for every z in B, where $(\cdot,\cdot)$ is the inner product associated with B. Then (3.11) becomes

$$F=\begin{vmatrix}(q_1,v_1) & (q_1,v_2) & \cdots & (q_1,v_k)\\ (q_2,v_1) & (q_2,v_2) & \cdots & (q_2,v_k)\\ \vdots & \vdots & & \vdots\\ (q_k,v_1) & (q_k,v_2) & \cdots & (q_k,v_k)\end{vmatrix}\neq 0. \tag{3.27}$$

One of the consequences of (3.27) is that both sets of vectors $Q^k=\{q_1,\cdots,q_k\}$ and $V^k=\{v_1,\cdots,v_k\}$ have to be linearly independent. Another consequence is that the intersection of the subspace span Q^k with that orthogonal to span V^k must be $\{0\}$.

The asymptotic error analysis of the MMPE as given in Theorem 3.2 leads one to the following important conclusions:

(1) Under the conditions stated in the theorem, the MMPE is a bona fide vector accelerator in the sense that

$$\frac{\|s_{n,k}-s\|}{\|x_{n+k+1}-s\|}=O\left[\left(\frac{\lambda_{k+1}}{\lambda_1}\right)^n\right] \quad \text{as } n\to\infty. \tag{3.28}$$

This means that if $x_n\to s$ as $n\to\infty$, that is, $|\lambda_1|<1$, then $s_{n,k}\to s$ as $n\to\infty$, and more quickly. Also if $\lim_{m\to\infty} x_m$ does not exist, that is, $|\lambda_1|\geqq 1$, then $s_{n,k}\to s$ as $n\to\infty$, provided that $|\lambda_{k+1}|<1$. The reason that we write x_{n+k+1} in (3.28) is that $s_{n,k}$ in the MMPE makes use of the $k+2$ vectors $x_n, x_{n+1},\cdots,x_{n+k+1}$.

(2) The result in (3.13) shows that when the MMPE is applied to a vector sequence generated by using the matrix iterative method described in § 2.2, with the notation therein, it will be especially effective when the iteration matrix A has a small number of large eigenvalues (k – many when $s_{n,k}$ is being used) that are well separated from the small eigenvalues.

(3) By inspection of $\Gamma(n)$ in (3.13) and (3.14), it follows that a loss of accuracy will take place in $s_{n,k}$ when $\lambda_1,\cdots,\lambda_k$ are close to 1, since $\|\Gamma(n)\|$ becomes large in this case. When the vector sequence is obtained by solving the linear system of equations given in (2.12) by the iterative technique in (2.13), this means that if A has large eigenvalues near 1, there will be a loss of accuracy in $s_{n,k}$. In fact eigenvalues near 1 would cause the system in (2.12) to be nearly singular.

4. Stability of MMPE. Let us denote the γ_j of the previous section by $\gamma_j^{(n,k)}$. Then the propagation of errors introduced in the x_m will be controlled, to some extent, by $\sum_{j=0}^k|\gamma_j^{(n,k)}|$; the larger this quantity, the worse the error propagation is expected to be. With this in mind, we say that $s_{n,k}$ is *asymptotically stable* if

$$\sup_n \sum_{j=0}^k |\gamma_j^{(n,k)}|<\infty. \tag{4.1}$$

Since $\sum_{j=0}^k \gamma_j^{(n,k)}=1$ by (3.2), then $\sum_{j=0}^k|\gamma_j^{(n,k)}|\geqq 1$, so that the most ideal situation is that in which $\gamma_j^{(n,k)}\geqq 0$, $0\leqq j\leqq k$, for n sufficiently large. The following theorem shows that for the type of sequences considered in Theorem 3.2, $s_{n,k}$ as obtained from MMPE is asymptotically stable, and that $\gamma_j^{(n,k)}>0$, $0\leqq j\leqq 1$, for sufficiently large n, whenever λ_i, $1\leqq i\leqq k$, are real and negative.

THEOREM 4.1. *Under the conditions stated in Theorem* 3.2, $s_{n,k}$ *is asymptotically stable.*

Proof. By (3.5), it is sufficient to show that $\gamma_j^{(n,k)}$, $0 \leqq j \leqq k$, stay bounded for $n \to \infty$, which in turn guarantees (4.1). Now

$$N_j = (-1)^j \begin{vmatrix} u_{n,1} & \cdots & u_{n+j-1,1} & u_{n+j+1,1} & \cdots & u_{n+k,1} \\ u_{n,2} & \cdots & u_{n+j-1,2} & u_{n+j+1,2} & \cdots & u_{n+k,2} \\ \vdots & & \vdots & \vdots & & \vdots \\ u_{n,k} & \cdots & u_{n+j-1,k} & u_{n+j+1,k} & \cdots & u_{n+k,k} \end{vmatrix}. \tag{4.2}$$

Substituting the asymptotic expansions of $u_{m,j}$ as given in (3.10), and using the multilinearity property of determinants, we obtain

$$N_j \sim \sum_{i_1=1}^{\infty} \cdots \sum_{i_k=1}^{\infty} \left(\prod_{p=1}^{k} z_{i_p,p} \right) \left(\prod_{p=1}^{k} \lambda_{i_p}^n \right) C_j(\lambda_{i_1}, \cdots, \lambda_{i_k}) \quad \text{as } n \to \infty, \tag{4.3}$$

where

$$C_j(\xi_1, \cdots, \xi_k) = (-1)^j \begin{vmatrix} 1 & \xi_1 & \cdots & \xi_1^{j-1} & \xi_1^{j+1} & \cdots & \xi_1^k \\ 1 & \xi_2 & \cdots & \xi_2^{j-1} & \xi_2^{j+1} & \cdots & \xi_2^k \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 1 & \xi_k & \cdots & \xi_k^{j-1} & \xi_k^{j+1} & \cdots & \xi_k^k \end{vmatrix}. \tag{4.4}$$

Since $C_j(\xi_1, \cdots, \xi_k)$ is odd under an interchange of the indices $1, \cdots, k$, Lemma A.1 in the appendix again applies, and we have

$$N_j \sim \sum_{1 \leqq i_1 < i_2 < \cdots < i_k}^{\infty} \begin{vmatrix} z_{i_1,1} & z_{i_2,1} & \cdots & z_{i_k,1} \\ z_{i_1,2} & z_{i_2,2} & \cdots & z_{i_k,2} \\ \vdots & \vdots & & \vdots \\ z_{i_1,k} & z_{i_2,k} & \cdots & z_{i_k,k} \end{vmatrix} C_j(\lambda_{i_1}, \cdots, \lambda_{i_k}) \left(\prod_{p=1}^{k} \lambda_{i_p}^n \right). \tag{4.5}$$

By (3.11), (3.19), (3.20), and (3.12)

$$N_j = \left(\prod_{p=1}^{k} \lambda_p^n \right) \tilde{F}[C_j(\lambda_1, \cdots, \lambda_k) + o(1)] \quad \text{as } n \to \infty. \tag{4.6}$$

Combining (4.6) and (3.21) in (3.5), and using (3.20), we obtain

$$\gamma_j^{(n,k)} = \frac{C_j(\lambda_1, \cdots, \lambda_k)}{V(1, \lambda_1, \cdots, \lambda_k)} + o(1) \quad \text{as } n \to \infty. \tag{4.7}$$

Obviously, (4.7) also implies that $|\gamma_j^{(n,k)}| < \infty$ for sufficiently large n. This then proves (4.1). □

Inspection of (4.4) reveals that $C_j(\lambda_1, \cdots, \lambda_k)$ is the cofactor of λ^j in the first row of

$$V(\lambda, \lambda_1, \cdots, \lambda_k) = \begin{vmatrix} 1 & \lambda & \cdots & \lambda^k \\ 1 & \lambda_1 & \cdots & \lambda_1^k \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_k & \cdots & \lambda_k^k \end{vmatrix}, \tag{4.8}$$

that is,

$$V(\lambda, \lambda_1, \cdots, \lambda_k) = \sum_{j=0}^{k} C_j(\lambda_1, \cdots, \lambda_k) \lambda^j. \tag{4.9}$$

Combining (4.7) and (4.9), we obtain the following interesting result:

$$\lim_{n\to\infty} \sum_{j=0}^{k} \gamma_j^{(n,k)} \lambda^j = \frac{V(\lambda, \lambda_1, \cdots, \lambda_k)}{V(1, \lambda_1, \cdots, \lambda_k)}. \tag{4.10}$$

Invoking (3.26), (4.10) finally becomes

$$\lim_{n\to\infty} \sum_{j=0}^{k} \gamma_j^{(n,k)} \lambda^j = \prod_{i=1}^{k} \left(\frac{\lambda_i - \lambda}{\lambda_i - 1}\right). \tag{4.11}$$

Since $V(\lambda, \lambda_1, \cdots, \lambda_k)$ is a polynomial of degree k in λ and vanishes when $\lambda = \lambda_i$, $1 \leqq i \leqq k$, we have

$$V(\lambda, \lambda_1, \cdots, \lambda_k) = C_k(\lambda_1, \cdots, \lambda_k) \prod_{p=1}^{k} (\lambda - \lambda_p). \tag{4.12}$$

Upon expanding the product on the right side of (4.12), and comparing with (4.9), we get

$$C_{k-j}(\lambda_1, \cdots, \lambda_k) = (-1)^j C_k(\lambda_1, \cdots, \lambda_k) \sum_{1 \leqq i_1 < \cdots < i_j \leqq k} \lambda_{i_1} \cdots \lambda_{i_j}, \qquad 1 \leqq j \leqq k. \tag{4.13}$$

It is obvious from (4.13) that if λ_i, $1 \leqq i \leqq k$, are real and negative, then $\operatorname{sgn} C_{k-j}(\lambda_1, \cdots, \lambda_k) = \operatorname{sgn} C_k(\lambda_1, \cdots, \lambda_k)$ for $0 \leqq j \leqq k$, and this implies that $\gamma_j^{(n,k)} > 0$, $0 \leqq j \leqq k$, for n sufficiently large. If λ_i, $1 \leqq i \leqq k$, are real and positive, then (4.13) implies that $\gamma_j^{(n,k)} \gamma_{j+1}^{(n,k)} < 0$, $0 \leqq j \leqq k-1$.

5. Convergence and stability of TEA. In this section we shall consider the convergence and stability properties of the TEA, which is obtained from Approach (a) or (b) in conjunction with Method (3) of § 2.2. The TEA can be summarized as follows: The approximation $s_{n,k}$ to s is given by

$$s_{n,k} = \sum_{i=0}^{k} \gamma_i x_{n+i}, \tag{5.1}$$

where γ_i are obtained from the equations

$$\begin{aligned} &\sum_{i=0}^{k} \gamma_i = 1, \\ &\sum_{i=0}^{k} \gamma_i Q(u_{m+i}) = 0, \qquad n \leqq m \leqq n+k-1, \end{aligned} \tag{5.2}$$

with $u_m = \Delta x_m = x_{m+1} - x_m$, $m \geqq 0$, as before. When $\gamma_k \neq 0$, equations (5.2) are equivalent to (2.18) and (2.22), in which $d_i = c_i$, $w_{m+i} = u_{m+i}$, $0 \leqq i \leqq k-1$, and $\tilde{w}_m = -u_{m+k}$, as can be verified by inspection.

We now write the equation in (5.2) in the form

$$\begin{aligned} &\sum_{i=0}^{k} \gamma_i = 1, \\ &\sum_{i=0}^{k} \gamma_i u_{n+i,j} = 0, \qquad 1 \leqq j \leqq k, \end{aligned} \tag{5.3}$$

where this time $u_{m,j} = Q(u_{m+j-1})$, $m \geqq 0$, $1 \leqq j \leqq k$. Defining $D(\sigma_0, \sigma_1, \cdots, \sigma_k)$ as in (3.3) but with the $u_{m,j}$ of § 3 replaced by the new $u_{m,j}$, we see that $s_{n,k}$ for TEA is given exactly by (3.6), as can be verified by Cramer's rule. Let us also define $z_{i,j} = Q(z_i)\lambda_i^{j-1} = Q(v_i)(\lambda_i - 1)\lambda_i^{j-1}$, $i \geqq 1$, $1 \leqq j \leqq k$, where the z_i are defined in § 3. Then under the assumptions preceding (3.10), (3.10) holds with the $z_{i,j}$ of § 3 replaced by the new $z_{i,j}$.

If one follows the proof of Theorem 3.2, one realizes that (3.18) and (3.24), which form the most important parts of it, are consequences of (2.11) and (3.10). Consequently (3.18) and (3.24) remain true for the TEA provided the $z_{i,j}$ of § 3 are replaced by those of the present section. Starting with these observations, we now prove the following theorem:

THEOREM 5.1. *Assume that*

$$Q(v_i)\neq 0, \qquad 1\leqq i\leqq k, \tag{5.4}$$

and that all the conditions of Theorem 3.2 (*with the exception of* (3.11)) *are satisfied. Then, for all sufficiently large n,* $\sum_{j=0}^{k} N_j\neq 0$; *hence* $s_{n,k}$ *as given in* (3.6) *exists. Furthermore,*

$$s_{n,k}-s=\Lambda(n)\lambda_{k+1}^{n}[1+o(1)] \quad \text{as } n\to\infty, \tag{5.5}$$

where the vector $\Lambda(n)$ *is nonzero and bounded for all sufficiently large n. If, in addition,* $|\lambda_{k+1}|>|\lambda_{k+2}|$, *then*

$$\Lambda(n)=\begin{vmatrix} v_1 & v_2 & \cdots & v_{k+1}\\ z_{1,1} & z_{2,1} & \cdots & z_{k+1,1}\\ z_{1,2} & z_{2,2} & \cdots & z_{k+1,2}\\ \vdots & \vdots & & \vdots\\ z_{1,k} & z_{2,k} & \cdots & z_{k+1,k}\end{vmatrix}\frac{\prod_{i=1}^{k}(\lambda_{k+1}-\lambda_i)}{V(\lambda_1,\cdots,\lambda_k)[\prod_{i=1}^{k}(\lambda_i-1)^2][\prod_{i=1}^{k}Q(v_i)]}. \tag{5.6}$$

Proof. The proof of this theorem proceeds along the same line as that of Theorem 3.2, using the additional relation

$$\begin{vmatrix} z_{1,1} & z_{2,1} & \cdots & z_{k,1}\\ z_{1,2} & z_{2,2} & \cdots & z_{k,2}\\ \vdots & \vdots & & \vdots\\ z_{1,k} & z_{2,k} & \cdots & z_{k,k}\end{vmatrix}=V(\lambda_1,\cdots,\lambda_k)\left[\prod_{i=1}^{k}Q(v_i)\right]. \tag{5.7}$$ □

Note. When the normed vector space B is a Hilbert space, the condition (5.4) has the following implication: Let q be the unique vector in B for which $Q(z)=(q,z)$ for every z in B. Then (5.4) implies that q cannot be orthogonal to any of the vectors v_i, $1\leqq i\leqq k$.

As a result of the asymptotic error analysis of the TEA given in Theorem 5.1, we can draw conclusions that are identical to those about the MMPE given at the end of § 3. We find that (3.28) is replaced by

$$\frac{\|s_{n,k}-s\|}{\|x_{n+2k}-s\|}=O\left[\left(\frac{\lambda_{k+1}}{\lambda_1}\right)^n\right] \quad \text{as } n\to\infty, \tag{5.8}$$

since $s_{n,k}$ for TEA is formed by taking into account the $2k+1$ vectors $x_n, x_{n+1},\cdots,x_{n+2k}$, instead of the $k+2$ vectors $x_n, x_{n+1},\cdots,x_{n+k+1}$ used to form $s_{n,k}$ for the MMPE. This in turn implies that the MMPE is a more economical vector accelerator than the TEA, since it attains the same rate of acceleration as the TEA while using approximately half the number of vectors.

Finally the stability properties of the TEA are very similar to those of the MMPE as is stated in the following theorem.

THEOREM 5.2. *Under the conditions stated in Theorem* 5.1 $s_{n,k}$ *is asymptotically stable. Furthermore,* (4.7) *and* (4.11) *hold too.*

Proof. Similar to that of Theorem 4.1. □

The conclusions that were drawn from the stability analysis of the MMPE in § 4, are true for the TEA too, as some analysis reveals.

6. Numerical examples. In § 3 we analyzed the convergence properties of $s_{n,k}$ for the MMPE as $n \to \infty$, and derived an asymptotic error estimate for it, obtaining at the same time its rate of convergence. In this section we apply the MMPE and the MPE to three vector sequences obtained as iterative approximations to linear systems of equations. The numerical results verify the conclusions of the asymptotic error analysis of § 3. They also indicate that the MMPE and the MPE have very similar performances.

In all the examples below the MPE is implemented by solving the generally overdetermined system

$$\sum_{i=0}^{k-1} c_i u_{n+i} = -u_{n+k},$$

for the c_i using the method of least squares and then setting $\gamma_i = c_i / \sum_{j=0}^{k} c_j$, $0 \leqq i \leqq k$, with $c_k = 1$, in (2.17). As before $u_j = x_{j+1} - x_j$, $j = 0, 1, \cdots$. The MMPE, on the other hand, is implemented by solving the linear system of k equations

$$\sum_{i=0}^{k-1} c_i u_{n+i,j} = -u_{n+k,j}, \qquad 1 \leqq j \leqq k,$$

where $u_{m,j}$ denotes the jth component of the vector u_m. That is to say, we pick the linear functional Q_j in (2.17) to be the projection operator onto the subspace spanned by the jth unit vector, $j = 1, \cdots, k$. We then set $\gamma_i = c_i / \sum_{j=0}^{k} c_j$, $0 \leqq i \leqq k$, with $c_k = 1$, in (2.17).

Example 1. The vectors x_i are obtained by setting $x_0 = 0$ and $x_{i+1} = Ax_i + b$, where A is the iteration matrix associated with the Gauss-Seidel method for the system of linear equations $Cx = d$, where

$$C = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & -3 & 1 & 5 \\ 3 & 1 & 6 & -2 \\ 4 & 5 & -2 & -1 \end{bmatrix}$$

and d and/or b are determined by requiring that the solution s to $Cx = d$ be the vector with all its entries equal to 1. The eigenvalues of A are approximately $\lambda_1 = \bar{\lambda}_2 = -2.3500 \pm 2.0506i$, $\lambda_3 = -0.0228$, and $\lambda_4 = 0$. Therefore, $\lim_{i\to\infty} x_i$ does not exist. In Table 1 we give the errors $\|s_{n,k} - s\|_\infty$ computed in the l_∞ norm for $k = 2$ and $0 \leqq n \leqq 5$, both for the MMPE and the MPE.

TABLE 1

l_∞ norms of the errors $s_{n,2} - s$ for Example 1, computed using MMPE and MPE, for $k = 2$ and $0 \leqq n \leqq 5$

	MMPE	MPE
n	$\|s_{n,2} - s\|_\infty$	$\|s_{n,2} - s\|_\infty$
0	6×10^{-1}	1×10^{0}
1	8×10^{-3}	7×10^{-3}
2	2×10^{-4}	2×10^{-4}
3	4×10^{-6}	4×10^{-6}
4	1×10^{-7}	9×10^{-8}
5	9×10^{-10}	9×10^{-10}

[The numbers have been rounded to one significant decimal digit. The base iterations x_j diverge.]

Example 2. The vectors x_i are obtained by setting $x_0=0$ and $x_{i+1}=Ax_i+b$, $i=0,1,\cdots$, where

$$A=0.06\cdot\begin{bmatrix} 5&2&1&1&&&&&&& \\ 2&6&3&1&1&&&&&& \\ 1&3&6&3&1&1&&&&& \\ 1&1&3&6&3&1&1&&&& \\ &1&1&3&6&3&1&1&&& \\ &&1&1&3&6&3&1&1&& \\ &&&1&1&3&6&3&1&1& \\ &&&&1&1&3&6&3&1&1 \\ &&&&&1&1&3&6&3&1 \\ &&&&&&1&1&3&6&2 \\ &&&&&&&1&1&2&5 \end{bmatrix}$$

and b is determined by requiring that the solution s of the system $x=Ax+b$ be the vector with all its entries equal to 1. The eigenvalues of A are all real and in $(0,1)$, and are approximately $\lambda_1=0.8965$, $\lambda_2=0.7318$, $\lambda_3=0.5297$, $\lambda_4=0.3600,\cdots,\lambda_{11}=0.0313$, in decreasing order. Since $\lambda_1<1$, the sequence x_i, $i=0,1,2,\cdots$, converges.

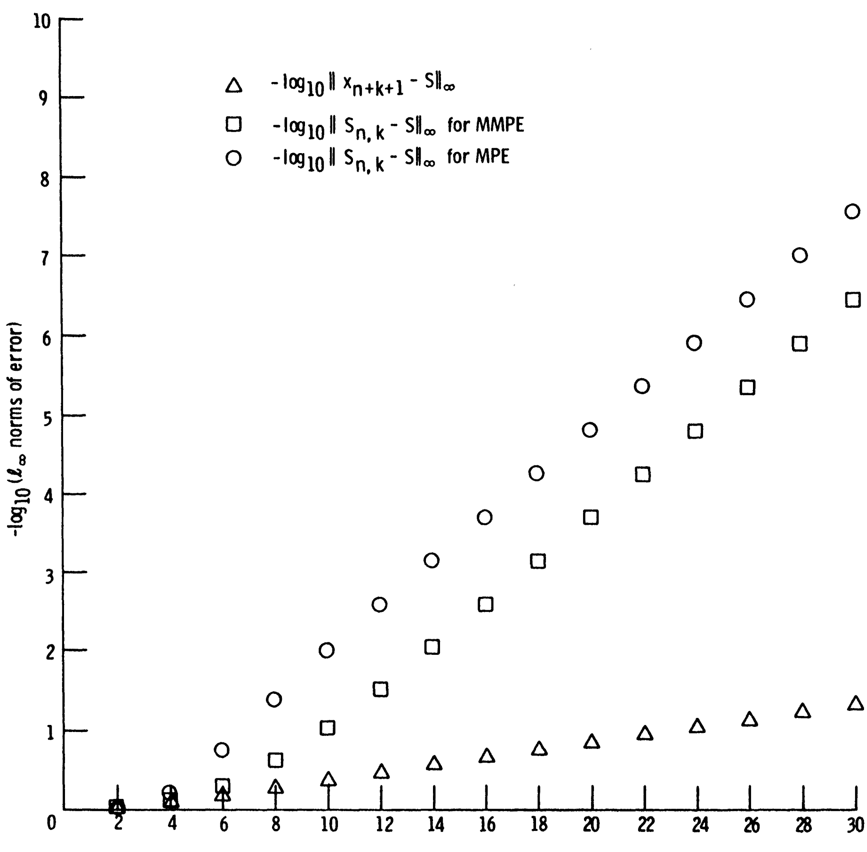

FIG. 1. *Results for Example* 2 *taking* $k=1$.

In Figs. 1 and 2 we give the results of the computations for $\|s_{n,k}-s\|_\infty$ using both the MMPE and the MPE with $k=1$ and $k=2$, respectively. The figures also include $\|x_{n+k+1}-s\|_\infty$.

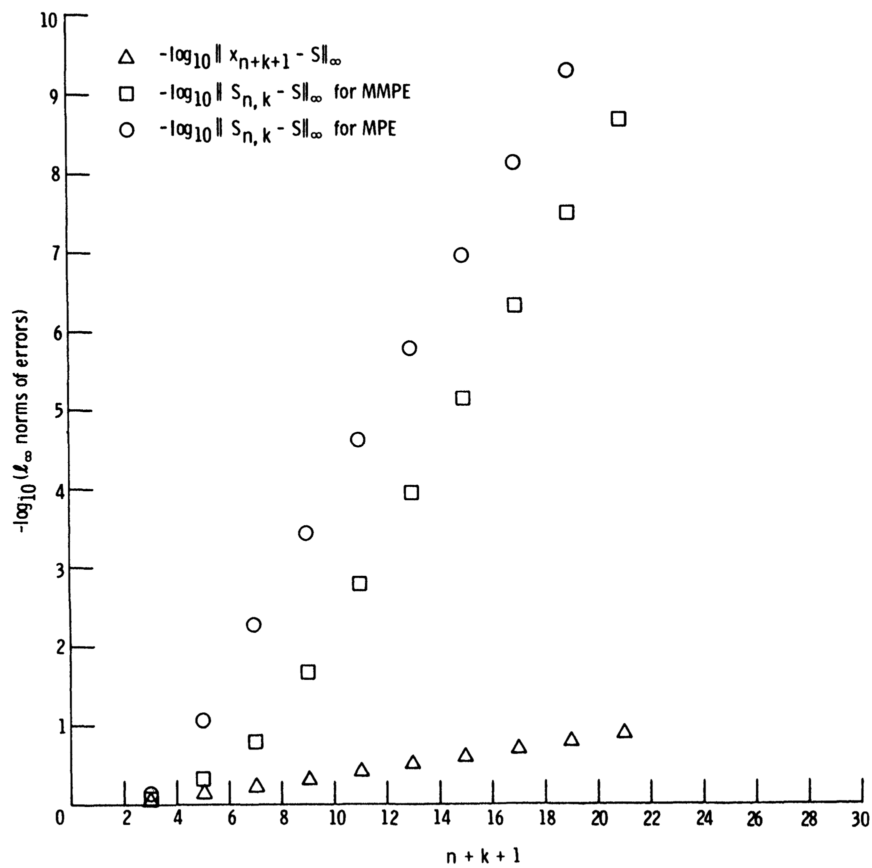

FIG. 2. *Results for Example 2 taking* $k=2$.

Example 3. The vectors x_i are obtained by setting $x_0=0$ and $x_{i+1}=Ax_i+b$, $i=0,1,\cdots$, where

$$A=0.04\cdot\begin{bmatrix} 12 & 11 & & & & & \\ 11 & 11 & 10 & & & & \\ 10 & 10 & 10 & 9 & & & \\ \vdots & \vdots & \vdots & \vdots & & & \\ 2 & 2 & 2 & 2 & \cdots & 2 & 1 \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}$$

and b is determined by requiring that the solution s of the system $x=Ax+b$ be the vector with all its entries equal to 1. The eigenvalues of A are all real and are approximately $\lambda_1=1.2892$, $\lambda_2=0.8080$, $\lambda_3=0.4924$, $\lambda_4=0.2785,\cdots,\lambda_{12}=0.0012$, in decreasing order. Since $\lambda_1>1$, the sequence x_i, $i=0,1,2,\cdots$, diverges. In Figs. 3 and 4 we give the results of the computations for $\|s_{n,k}-s\|_\infty$ using both the MMPE and the MPE, with $k=2$ and $k=3$, respectively.

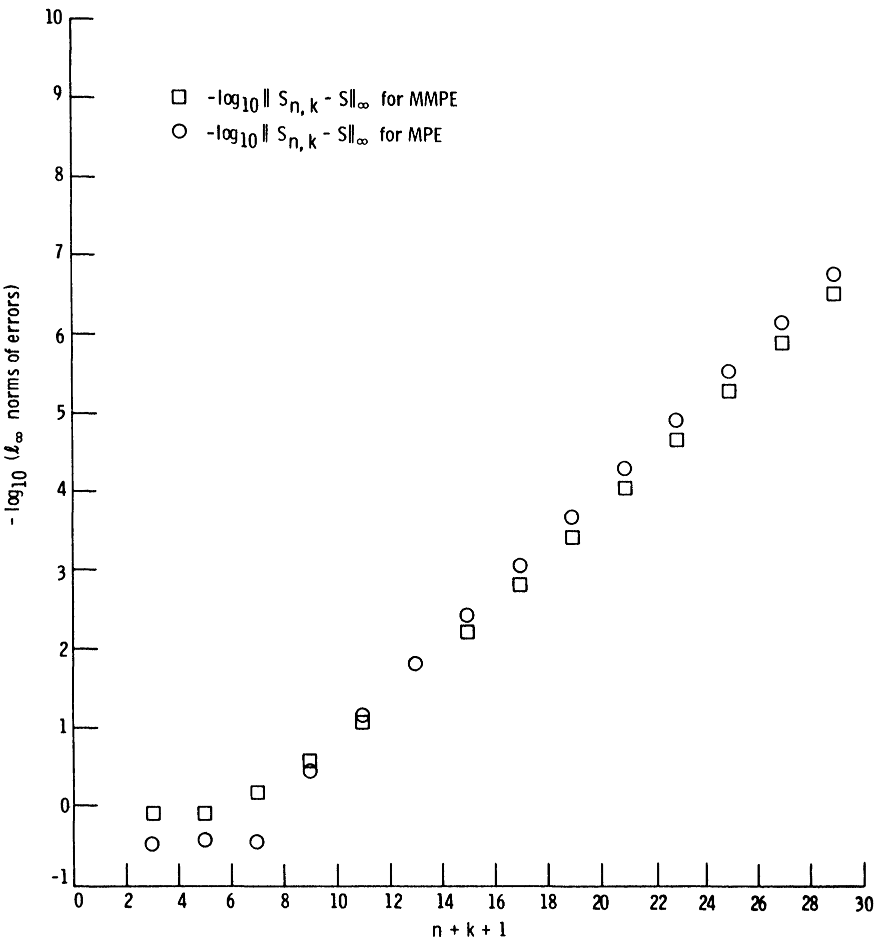

FIG. 3. *Results for Example* 3 *taking* $k=2$. *The base iterations* x_j *diverge.*

Appendix.

LEMMA A.1. *Let* $i_0, i_1, \cdots, i_k$ *be integers greater than or equal to* 1, *and assume that the scalars* $v_{i_0,\cdots,i_k}$ *are odd under an interchange of any two indices* $i_0, \cdots, i_k$. *Let* σ_i, $i \geqq 1$, *be scalars (or vectors), and let* $t_{i,j}$, $i \geqq 1$, $1 \leqq j \leqq k$, *be scalars. Define*

$$I_{k,N} = \sum_{i_0=1}^{N} \cdots \sum_{i_k=1}^{N} \sigma_{i_0}\left(\prod_{p=1}^{k} t_{i_p,p}\right) v_{i_0,\cdots,i_k} \tag{A.1}$$

and

$$J_{k,N} = \sum_{1 \leqq i_0 < i_1 < \cdots < i_k \leqq N} \begin{vmatrix} \sigma_{i_0} & \sigma_{i_1} & \cdots & \sigma_{i_k} \\ t_{i_0,1} & t_{i_1,1} & \cdots & t_{i_k,1} \\ t_{i_0,2} & t_{i_1,2} & \cdots & t_{i_k,2} \\ \vdots & \vdots & & \vdots \\ t_{i_0,k} & t_{i_1,k} & \cdots & t_{i_k,k} \end{vmatrix} v_{i_0,\cdots,i_k}, \tag{A.2}$$

where the determinant in (A.2) *is to be interpreted in the same way as* $D(\sigma_0, \cdots, \sigma_k)$ *in* (3.3). *Then*

$$I_{k,N} = J_{k,N}. \tag{A.3}$$

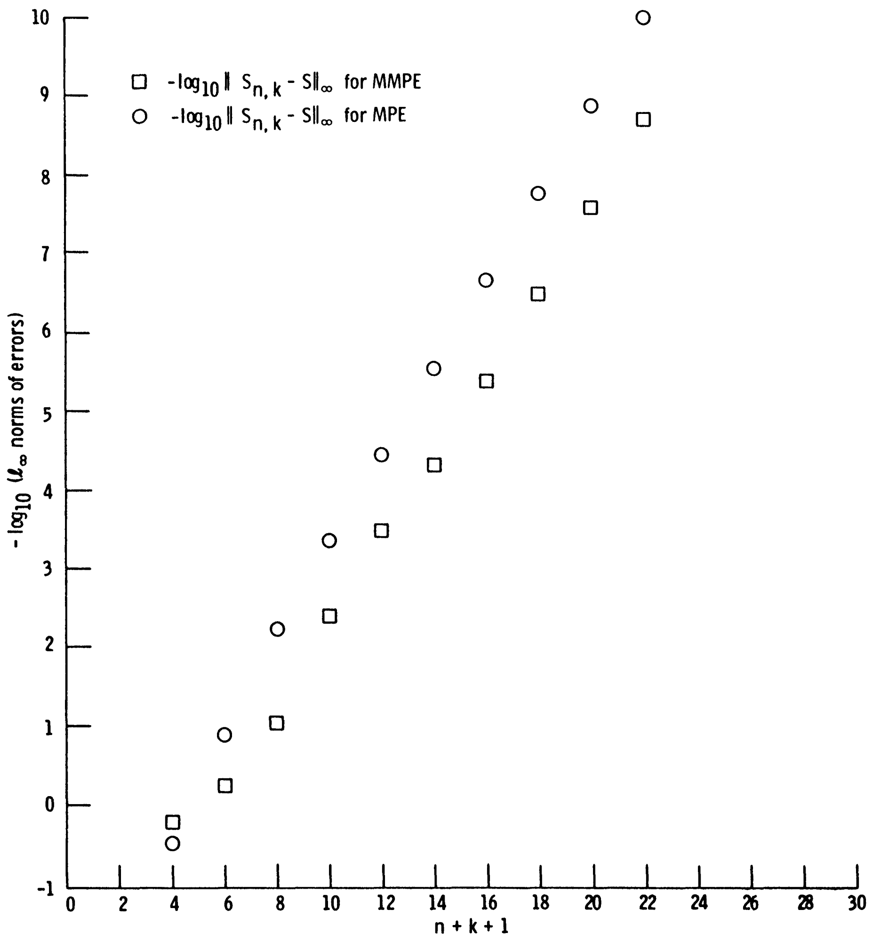

FIG. 4. *Results for Example* 3 *taking* $k=3$. *The base iterations* x_j *diverge.*

Proof. Let Σ_k be the set of all permutations π of $\{0,1,\cdots,k\}$. Then by the definition of determinants

$$J_{k,N}=\sum_{1\leqq i_0<i_1<\cdots<i_k\leqq N}\sum_{\pi\in\Sigma_k}(\operatorname{sgn}\pi)\sigma_{i_{\pi(0)}}\left(\prod_{p=1}^{k}t_{i_{\pi(p)},p}\right)v_{i_0,\cdots,i_k}. \tag{A.4}$$

The notation $\pi(p)$, where $0\leqq p\leqq k$, designates the image of π as a function operating on the index set. Now $\pi^{-1}\pi(p)=p$, $0\leqq p\leqq k$, and $\operatorname{sgn}\pi^{-1}=\operatorname{sgn}\pi$ for any permutation $\pi\in\Sigma_k$. Hence

$$J_{k,N}=\sum_{1\leqq i_0<i_1<\cdots<i_k\leqq N}\sum_{\pi\in\Sigma_k}(\operatorname{sgn}\pi^{-1})\sigma_{i_{\pi(0)}}\left(\prod_{p=1}^{k}t_{i_{\pi(p)},p}\right)v_{i_{\pi^{-1}\pi(0)},\cdots,i_{\pi^{-1}\pi(p)}}. \tag{A.5}$$

By the oddness of $v_{i_0,\cdots,i_k}$, we have

$$(\operatorname{sgn}\pi^{-1})v_{i_{\pi^{-1}\pi(0)},\cdots,i_{\pi^{-1}\pi(p)}}=v_{i_{\pi(0)},\cdots,i_{\pi(p)}}. \tag{A.6}$$

Substituting (A.6) in (A.5), we obtain

$$J_{k,N}=\sum_{1\leqq i_0<i_1<\cdots<i_k\leqq N}\sum_{\pi\in\Sigma_k}\sigma_{i_{\pi(0)}}\left(\prod_{p=1}^{k}t_{i_{\pi(p)},p}\right)v_{i_{\pi(0)},\cdots,i_{\pi(p)}}. \tag{A.7}$$

Since $v_{i_0,\cdots,i_k}$ is odd under interchange of the indices $i_0, \cdots, i_k$, it vanishes when any two of these indices are equal. Using this fact in (A.1), we see that $I_{k,N}$ is just the sum over all permutations of the *distinct* indices $i_0, \cdots, i_k$. The result now follows by comparison with (A.7).

Note that when the σ_i are scalars, (A.3) remains true also for the case in which the $v_{i_0,\cdots,i_k}$ are vectors. When the σ_i *and* $v_{i_0,\cdots,i_k}$ are vectors, (A.3) still holds provided $\sigma_{i_0} v_{i_0,\cdots,i_k}$ is interpreted as a direct (tensor) product.

Acknowledgments. The computations for this paper were carried out at the Computer Services Division of NASA Lewis Research Center. The authors are pleased to thank S. D. Silliman of Cleveland State University for performing the computations reported in § 6.

REFERENCES

[1] C. BREZINSKI, *Accélération de la convergence en analyse numérique*, Lecture Notes in Mathematics 584, Springer-Verlag, Berlin, 1977.

[2] S. CABAY AND L. W. JACKSON, *A polynomial extrapolation method for finding limits and antilimits of vector sequences*, this Journal, 13 (1976), pp. 734–752.

[3] R. P. EDDY, *Extrapolating to the limit of a vector sequence*, in Information Linkage Between Applied Mathematics and Industry, P. C. C. Wang, ed., Academic Press, New York, 1979, pp. 387–396.

[4] M. MEŠINA, *Convergence acceleration for the iterative solution of the equations* $X = AX + f$, Comp. Meth. Appl. Mech. Eng., 10 (1977), pp. 165–173.

[5] P. RABINOWITZ, *Applications of linear programming to numerical analysis*, SIAM Rev., 10 (1968), pp. 121–159.

[6] D. SHANKS, *Non-linear transformations of divergent and slowly convergent sequences*, J. Math. and Phys., 34 (1955), pp. 1–42.

[7] A. SIDI, *Convergence and stability properties of minimal polynomial and reduced rank extrapolation algorithms*, NASA Technical Memorandum 83443, July 1983, this issue, pp. 197–209.

[8] D. A. SMITH, W. F. FORD AND A. SIDI, *Extrapolation methods for vector sequences*, Duke Univ. preprint, February 1984, submitted to SIAM Rev.

[9] P. WYNN, *On a device for computing the* $e_m(S_n)$ *transformation*, MTAC, 10 (1956), pp. 91–96.

[10] ———, *Acceleration techniques for iterated vector and matrix problems*, Math. Comp., 16 (1962), pp. 301–322.

[11] ———, *On the convergence and stability of the epsilon-algorithm*, this Journal, 3 (1966), pp. 91–122.