

NASA-TM-83443

NASA Technical Memorandum 83443

NASA-TM-83443 19830023119

Convergence and Stability Properties of Minimal Polynominal and Reduced Rank Extrapolation Algorithms

Avram Sidi
*Lewis Research Center
Cleveland, Ohio*

July 1983

LIBRARY COPY

SEP 13 1983

LANGLEY RESEARCH CENTER
LIBRARY, NASA
HAMPTON, VIRGINIA

NASA

40 1 1 RN/NASA-TN-83443

DISPLAY 40/2/1

83N31390*# ISSUE 19 PAGE 3204 CATEGORY 64 RPT#: NASA-TN-83443

E-1659 NAS 1.15:83443 83/07/00 27 PAGES UNCLASSIFIED DOCUMENT

UTTL: Convergence and stability properties of minimal polynomial and reduced rank extrapolation algorithms

AUTH: A/SIDI, A. FAA: A/(Technion - Israel Inst. of Tech., Haifa)

CORP: National Aeronautics and Space Administration, Lewis Research Center, Cleveland, Ohio. AVAIL.NTIS SAP: HC A03/MF A01

MAJS: /*ALGORITHMS/*COMPUTATION/*CONVERGENCE/*EXTRAPOLATION/*NUMERICAL STABILITY
/*PROBLEM SOLVING/*VECTORS (MATHEMATICS)

MINS: / ASYMPTOTIC SERIES/ LINEAR EQUATIONS/ OPTIMIZATION/ POLYNOMIALS/ SERIES
EXPANSION

ABA: M.G.

ABS: The minimal polynomial and reduced rank extrapolation algorithms are two acceleration of convergence methods for sequences of vectors. In a recent survey these methods were tested and compared with the scalar, vector, topological epsilon algorithms, and were observed to be more efficient than the latter. It was also observed that the two methods have similar convergence properties. The convergence and stability properties of these methods are analyzed and the performance of the acceleration methods when applied to a class of vector sequences that includes those sequences obtained from systems of linear equations by using matrix iterative methods is discussed.

ENTER:

CONVERGENCE AND STABILITY PROPERTIES OF MINIMAL POLYNOMIAL
AND REDUCED RANK EXTRAPOLATION ALGORITHMS

Avram Sidi*
National Aeronautics and Space Administration
Lewis Research Center
Cleveland, Ohio 44135

ABSTRACT

The minimal polynomial and reduced rank extrapolation algorithms are two acceleration of convergence methods for sequences of vectors. In a recent survey these methods were tested and compared with the scalar, vector, and topological epsilon algorithms, and were observed to be more efficient than the latter. It was also observed that the two methods have similar convergence properties. The purpose of the present work is to analyze the convergence and stability properties of these methods, and to show that they are bona fide acceleration methods when applied to a class of vector sequences that includes those sequences obtained from systems of linear equations by using matrix iterative methods.

1. INTRODUCTION

The minimal polynomial extrapolation (MPE) and the reduced rank extrapolation (RRE) algorithms are two methods that have been devised for accelerating the convergence of sequences of vectors. In a recent survey carried out by D. A. Smith, W. F. Ford, and A. Sidi (unpublished) these two methods were tested and compared with the scalar, vector, and topological epsilon algorithms. It was observed numerically that the MPE and the RRE have similar convergence properties and are more efficient than the three epsilon algorithms. In the present work we analyze the convergence and stability properties of the

*Computer Science Department, Technion - Israel Institute of Technology, Haifa, Israel. Work partly done as a National Research Council - NASA Research Associate.

N83-31390#

MPE and the RRE, and show that they are bona fide convergence acceleration methods when applied to a family of sequences that includes those sequences obtained from systems of linear equations by using matrix iterative methods.

In a recent work [3] a general framework for deriving convergence acceleration methods for vector sequences has been proposed. Within this framework one can derive several methods, some old (including the MPE, RRE, and the topological epsilon algorithm), and some new (including a method that has been designated the modified MPE). The approach of [3] is formulated in general normed linear spaces of finite or infinite dimension. In the present work we use the formulations of the MPE and the RRE as they are given in [3]. We also make the assumption that the normed linear space in which the vector sequence is defined is an inner product space, with the norm being induced by the inner product.

The plan of the present paper is as follows: In Section 2 we specify the vector sequences whose convergence we are seeking to accelerate, describe the MPE and the RRE as they were formulated in [3], deriving at the same time determinant representations for them. We shall use Section 2 to also introduce much of the notation that we use in the remainder of this work. In Section 3 we analyze the convergence properties of both methods and obtain actual rates of acceleration for them. In Section 4 we analyze their stability properties. The results of Sections 3 and 4 are helpful in explaining some of the numerical results obtained from the MPE and the RRE. The techniques used in the present work are similar in nature to those developed and used in [3] in the analysis of the modified MPE and the topological epsilon algorithm. The analysis in the present work, however, is considerably heavier due to the extreme nonlinearity of the MPE and the RRE. Surprisingly, all the conclusions that were drawn for the modified MPE hold for MPE and RRE.

2. NOTATION AND DESCRIPTION OF ALGORITHMS

Let B be an inner product space defined over the field of complex numbers. In this work we shall adopt the following convention for the homogeneity property of the inner product. For $y, z \in B$ and α, β complex numbers, the inner product (\cdot, \cdot) is defined such that $(\alpha y, \beta z) = \bar{\alpha}\beta(y, z)$. The norm of a vector $x \in B$ will be defined by $\|x\| = \sqrt{(x, x)}$.

Let us consider a sequence of vectors $x_i, i = 0, 1, \dots$, in B . We shall assume that

$$(2.1) \quad x_m \sim s + \sum_{i=1}^{\infty} v_i \lambda_i^m \text{ as } m \rightarrow \infty,$$

where s and $v_i, i = 1, 2, \dots$, are vectors in B , and $\lambda_i, i = 1, 2, \dots$, are scalars, such that $\lambda_i \neq 1, i = 1, 2, \dots, \lambda_i \neq \lambda_j$ if $i \neq j, |\lambda_1| \geq |\lambda_2| \geq \dots$, and that there can be only a finite number of λ_i whose moduli are equal. Without loss of generality, we assume in (2.1) that $v_i \neq 0, \lambda_i \neq 0$ for all $i \geq 0$. The meaning of (2.1) 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 $m \geq m_0$ the vector

$$\tilde{v}_N(m) = \left(x_m - s - \sum_{i=1}^{N-1} v_i \lambda_i^m \right) / \lambda_N^m$$

satisfies

$$(2.1a) \quad \left\| \tilde{v}_N(m) \right\| \leq K.$$

If $|\lambda_1| < 1$, then $\lim_{m \rightarrow \infty} x_m$ exists and is simply s . If $|\lambda_1| \geq 1$, then $\lim_{m \rightarrow \infty} x_m$ does not exist, and s is said to be the anti-limit of the sequence $x_m, m = 0, 1, \dots$. Our problem is to find a good approximation to s , whether it be the limit or the anti-limit of the sequence, from a relatively small number of the vectors $x_i, i = 0, 1, \dots$.

Example. Let A be a nondefective $M \times M$ (complex) matrix and b , an M -dimensional (complex) vector, and consider the solution of the linear system of equations

$$(2.2) \quad x = Ax + b.$$

Let $\lambda_1, \dots, \lambda_M$ and v_1, \dots, v_M be the eigenvalues and corresponding eigenvectors of A . Assume also that 1 is not an eigenvalue of A so that (2.2) has a unique solution, which we shall denote by s . For a given x_0 , we generate the sequence $x_j, j = 1, 2, \dots$, by the matrix iterative method

$$(2.3) \quad x_{j+1} = Ax_j + b, \quad j = 0, 1, \dots$$

Let $x_0 - s = \sum_{i=1}^M \alpha_i v_i$ for some scalars α_i . Then

$$(2.4) \quad x_m = s + \sum_{i=1}^M \alpha_i v_i \lambda_i^m, \quad m = 0, 1, \dots$$

As is known if $\alpha_1 \neq 0$, which will be the case for the given x_0 in general, then $\lim_{m \rightarrow \infty} x_m = s$ provided $|\lambda_1| < 1$, otherwise s is the anti-limit.

Let us denote $u_i = \Delta x_i = x_{i+1} - x_i, i = 0, 1, \dots$, and $w_i = \Delta u_i = \Delta^2 x_i, i = 0, 1, \dots$

The MPE. Let k be an integer less than or equal to the dimension of the space B . The approximation $s_{n,k}$ to s is given by

$$(2.5) \quad s_{n,k} = \sum_{j=0}^k \gamma_j x_{n+j},$$

where the γ_j are obtained from

$$(2.6) \quad r_j = \frac{c_j}{\sum_{i=0}^k c_i}, \quad j = 0, 1, \dots, k,$$

with $c_k = 1$, and c_0, \dots, c_{k-1} being determined as the solution to the minimization problem

$$(2.7) \quad c_0, \dots, c_{k-1} \quad \left\| \sum_{i=0}^{k-1} c_i u_{n+i} + u_{n+k} \right\|,$$

provided $\sum_{i=0}^k c_i \neq 0$.

For the example above, (2.7) is equivalent to the least squares solution of the overdetermined system of M equations

$$(2.8) \quad U_{n,k} c = -u_{n+k},$$

where $U_{n,k}$ is the $M \times k$ matrix

$$(2.9) \quad U_{n,k} = (u_n, u_{n+1}, \dots, u_{n+k-1}),$$

and c is the column vector $(c_0, c_1, \dots, c_{k-1})^T$, when $(y, z) = y^* z = \sum_{i=1}^M \bar{y}^i z^i$,

where $y = (y^1, \dots, y^M)^T$ and $z = (z^1, \dots, z^M)^T$, and y^* is the Hermitean conjugate of y . This is the way the MPE was developed originally in [1].

We shall now give a determinant expression for $s_{n,k}$ that will be of use in the remainder of this work. Since $\|y\| = \sqrt{(y, y)}$ for any vector y in B , the c_i that solve the minimization problem in (2.7) satisfy the normal equations

$$(2.10) \quad \sum_{j=0}^{k-1} (u_{n+i}, u_{n+j}) c_j = -(u_{n+i}, u_{n+k}), \quad 0 \leq i \leq k-1.$$

Consequently, the γ_j that are defined by (2.6) satisfy the equations

$$(2.11) \quad \begin{cases} \sum_{j=0}^k \gamma_j = 1 \\ \sum_{j=0}^k (u_{n+i}, u_{n+j}) \gamma_j = 0, \quad 0 \leq i \leq k-1, \end{cases}$$

provided these equations have a solution. Assuming that the determinant of the matrix of equations (2.11) is nonzero, and using Cramer's rule, we can write the solution of (2.11) as

$$(2.12) \quad \gamma_j = \frac{N_j}{\sum_{i=0}^k N_i} = \frac{N_j}{D(1, \dots, 1)}, \quad 0 \leq j \leq k,$$

where N_j is the cofactor of σ_j in the determinant

$$(2.13) \quad D(\sigma_0, \dots, \sigma_k) = \begin{vmatrix} \sigma_0 & \sigma_1 & \cdot & \cdot & \cdot & \sigma_k \\ u_{0,0} & u_{0,1} & \cdot & \cdot & \cdot & u_{0,k} \\ u_{1,0} & u_{1,1} & \cdot & \cdot & \cdot & u_{1,k} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ u_{k-1,0} & u_{k-1,1} & \cdot & \cdot & \cdot & u_{k-1,k} \end{vmatrix}$$

with $u_{i,j} = (u_{n+i}, u_{n+j})$, $i, j \geq 0$. In the first row of the determinant $D(\sigma_0, \dots, \sigma_k)$ we also allow $\sigma_0, \dots, \sigma_k$ to be vectors in B , in which case (2.13) is to be interpreted as

$$(2.14) \quad D(\sigma_0, \dots, \sigma_k) = \sum_{i=0}^k \sigma_i N_i.$$

Combining (2.5), (2.12), and (2.14), we can now express $s_{n,k}$ as

$$(2.15) \quad s_{n,k} = \frac{D(x_n, \dots, x_{n+k})}{D(1, \dots, 1)}.$$

The RRE. Let k be an integer less than or equal to the dimension of the space B . The approximation $s_{n,k}$ to s is given by

$$(2.16) \quad s_{n,k} = x_n + \sum_{i=0}^{k-1} q_i u_{n+i},$$

where the q_i are determined as the solution of the minimization problem

$$(2.17) \quad q_0, \dots, q_{k-1} \left\| u_n + \sum_{i=0}^{k-1} q_i w_{n+i} \right\|.$$

For the example above, (2.17) is equivalent to the least squares solution of the overdetermined system of equations

$$(2.18) \quad w_{n,k} q = -u_n,$$

where $w_{n,k}$ is the $M \times k$ matrix

$$(2.19) \quad w_{n,k} = (w_n, w_{n+1}, \dots, w_{n+k-1}),$$

and q is the column vector $q = (q_0, q_1, \dots, q_{k-1})^T$, when $(y, z) = \sum_{i=1}^M \bar{y}_i z_i$,

where y and z are again $y = (y^1, \dots, y^M)^T$ and $z = (z^1, \dots, z^M)^T$.

This is the way the RRE was developed originally in [2].

For this case too $s_{n,k}$ can be expressed as the quotient of two determinants, and we turn to this now. Again by the fact that $\|y\| = \sqrt{(y,y)}$ for any vector y in B , the q_i that solve the minimization problem in (2.17) satisfy the normal equations

$$(2.20) \quad \sum_{j=0}^{k-1} (w_{n+i}, w_{n+j}) q_j = -(w_{n+i}, u_n), \quad 0 \leq i \leq k-1.$$

Substituting $w_{n+j} = u_{n+j+1} - u_{n+j}$ on the left hand side of (2.20), and rearranging, we obtain

$$(2.21) \quad (w_{n+i}, u_n)(1 - q_0) + \sum_{j=1}^{k-1} (w_{n+i}, u_{n+j})(q_{j-1} - q_j) + (w_{n+i}, u_{n+k})q_{k-1} = 0,$$

$$0 \leq i \leq k-1.$$

Let us define

$$(2.22) \quad \gamma_0 = 1 - q_0, \quad \gamma_k = q_{k-1}, \quad \gamma_j = q_{j-1} - q_j, \quad 1 \leq j \leq k-1.$$

It is easily verified that

$$(2.23) \quad \sum_{j=0}^k \gamma_j = 1,$$

so that (2.22) and (2.23) establish a one-to-one correspondence between the q_i , $0 \leq i \leq k-1$, and the γ_j , $0 \leq j \leq k$. Consequently, the linear system of equations (2.20) for the q_j is equivalent to the linear system

$$(2.24) \quad \left\{ \begin{array}{l} \sum_{j=0}^k \gamma_j = 1 \\ \sum_{j=0}^k (w_{n+i}, u_{n+j}) \gamma_j = 0, \quad 0 \leq i \leq k-1, \end{array} \right.$$

for the γ_j , by (2.21), (2.22), and (2.23). Similarly, substituting $u_{n+i} = x_{n+i+1} - x_{n+i}$ on the right hand side of (2.16), rearranging, and invoking (2.22), we see that $s_{n,k}$ for RRE (as for MPE) can be expressed in the form (2.5). Assuming now that the determinant of the matrix of equations (2.24) is nonzero, and using Cramer's rule, we can express the γ_j that solve (2.24) exactly as in (2.12), where N_j is the cofactor of σ_j in the determinant $D(\sigma_0, \dots, \sigma_k)$ given in (2.13) with $u_{i,j} \equiv (w_{n+i}, u_{n+j})$, $i, j \geq 0$. Again when $\sigma_0, \dots, \sigma_k$ are vectors in B , $D(\sigma_0, \dots, \sigma_k)$ is to be interpreted as in (2.14). Consequently $s_{n,k}$, also for RRE, can be expressed as in (2.15).

Before closing this section we shall state a result that will be of use in the remainder of this work.

Lemma 2.1. Let i_0, i_1, \dots, i_k be integers greater than or equal to 1, and assume that the scalars v_{i_0, \dots, i_k} are odd under an interchange of any two indices i_0, \dots, i_k . Let $\sigma_i, i \geq 1$, be scalars (or vectors) and let $t_{i,j}, i \geq 1, 1 \leq j \leq k$, be scalars. Define

$$(2.25) \quad I_{k,N} = \sum_{i_0=1}^N \cdots \sum_{i_k=1}^N \sigma_{i_0} \left(\prod_{p=1}^k t_{i_p,p} \right) v_{i_0, \dots, i_k}$$

and

$$(2.26) \quad J_{k,N} = \sum_{1 \leq i_0 < i_1 < \dots < i_k \leq N} \begin{vmatrix} \sigma_{i_0} & \sigma_{i_1} & \cdot & \cdot & \cdot & \sigma_{i_k} \\ t_{i_0,1} & t_{i_1,1} & \cdot & \cdot & \cdot & t_{i_k,1} \\ t_{i_0,2} & t_{i_1,2} & \cdot & \cdot & \cdot & t_{i_k,2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ t_{i_0,k} & t_{i_1,k} & \cdot & \cdot & \cdot & t_{i_k,k} \end{vmatrix} v_{i_0, \dots, i_k}$$

where the determinant in (2.26) is to be interpreted in the same way as $D(\sigma_0, \dots, \sigma_k)$ in (2.13). Then

$$(2.27) \quad I_{k,N} = J_{k,N}.$$

For a proof of Lemma 2.1, see [3, Appendix].

3. CONVERGENCE ANALYSIS

We have seen in the previous section that $s_{n,k}$, for both the MPE and the RRE, is given by

$$(3.1) \quad s_{n,k} = \frac{D(x_n, \dots, x_{n+k})}{D(1, \dots, 1)},$$

where $D(\sigma_0, \dots, \sigma_k)$ is defined by (2.13) with $u_{i,j} = (u_{n+i}, u_{n+j})$, $i, j \geq 0$, for MPE, and $u_{i,j} = (w_{n+i}, u_{n+j})$, $i, j \geq 0$, for RRE. Subtracting s from both sides of (3.1), and making use of (2.14), we obtain the error formula

$$(3.2) \quad s_{n,k} - s = \frac{D(x_n - s, \dots, x_{n+k} - s)}{D(1, \dots, 1)}.$$

Under the assumption (2.1), we have

$$(3.3) \quad u_m \sim \sum_{i=1}^{\infty} v_i (\lambda_i - 1) \lambda_i^m \quad \text{as } m \rightarrow \infty,$$

and consequently,

$$(3.4) \quad w_m \sim \sum_{i=1}^{\infty} v_i (\lambda_i - 1)^2 \lambda_i^m \quad \text{as } m \rightarrow \infty.$$

From (3.3) and (3.4) we see that

$$(3.5) \quad u_{p,q} = (u_{n+p}, u_{n+q}) \sim \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (v_i, v_j) (\bar{\lambda}_i - 1) (\lambda_j - 1) \bar{\lambda}_i^{-n+p} \lambda_j^{n+q}$$

as $n \rightarrow \infty$,

for MPE, and

$$(3.6) \quad u_{p,q} = (w_{n+p}, u_{n+q}) \sim \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (v_i, v_j) (\bar{\lambda}_i - 1)^2 (\lambda_j - 1) \bar{\lambda}_i^{-n+p} \lambda_j^{n+q}$$

as $n \rightarrow \infty$,

for RRE. (3.5) and (3.6) can be rewritten in the condensed and unified form

$$(3.7) \quad u_{p,q} \sim \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} z_{ij} \bar{\lambda}_i^{-n+p} \lambda_j^{n+q} \quad \text{as } n \rightarrow \infty,$$

where we have defined

$$(3.8) \quad z_{ij} = (v_i, v_j) (\bar{\lambda}_i - 1) (\lambda_j - 1), \quad i, j \geq 1,$$

for MPE, and

$$(3.9) \quad z_{ij} = (v_i, v_j) (\bar{\lambda}_i - 1)^2 (\lambda_j - 1), \quad i, j \geq 1,$$

for RRE.

Note that when the sequence $x_j, j = 0, 1, \dots$, is generated by a matrix iterative process as described in the example of Section 2, then in (3.7) the upper limits on the summations on the right hand side are replaced by M , and \sim is replaced by $=$.

In Lemma 3.1 and Lemma 3.2 below we derive the asymptotic expansions of $D(x_n - s, \dots, x_{n+k} - s)$ and $D(1, \dots, 1)$, respectively, for $n \rightarrow \infty$, assuming (2.1). In Theorem 3.1 we give the main result of this paper that shows that

both the MPE and RRE are true acceleration methods when applied to sequences of vectors satisfying (2.1), in the sense that $\lim_{n \rightarrow \infty} \|s_{n,k} - s\| / \|x_{n+k+1} - s\| = 0$.

In what follows we shall denote asymptotic relations of the form " $\alpha_n \sim \beta_n$ " as " $n \rightarrow \infty$ " by " $\alpha_n \sim \beta_n$ " for short. Also we shall let $\sum_i \equiv \sum_{i=1}^{\infty}$,

$$\sum_{i,j} \equiv \sum_{i=1}^{\infty} \sum_{j=1}^{\infty}, \text{ and } \sum_{1 \leq i_1 < i_2 < \dots < i_k} \equiv \sum_{i_1=1}^{\infty} \sum_{i_2=i_1+1}^{\infty} \dots \sum_{i_k=i_{k-1}+1}^{\infty}, \text{ for}$$

simplicity.

Lemma 3.1. Let the sequence of vectors x_i , $i = 0, 1, 2, \dots$, be as described in the previous section. For i_p, j_p positive integers, define

$$(3.10) \quad R_{\substack{i_1, \dots, i_k \\ j_0, j_1, \dots, j_k}} = \begin{vmatrix} v_{j_0} & v_{j_1} & \dots & v_{j_k} \\ z_{i_1 j_0} & z_{i_1 j_1} & \dots & z_{i_1 j_k} \\ z_{i_2 j_0} & z_{i_2 j_1} & \dots & z_{i_2 j_k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ z_{i_k j_0} & z_{i_k j_1} & \dots & z_{i_k j_k} \end{vmatrix},$$

where the interpretation of this determinant is like that of $D(\sigma_0, \sigma_1, \dots, \sigma_k)$ in (2.13) and (2.14). Let $V(\xi_0, \xi_1, \dots, \xi_k)$ be the Vandermonde determinant

$$(3.11) \quad V(\xi_0, \xi_1, \dots, \xi_k) = \begin{vmatrix} 1 & \xi_0 & \dots & \xi_0^k \\ 1 & \xi_1 & \dots & \xi_1^k \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 1 & \xi_k & \dots & \xi_k^k \end{vmatrix} = \prod_{0 \leq i < j \leq k} (\xi_j - \xi_i).$$

Then $D(x_n - s, \dots, x_{n+k} - s)$ has the asymptotic expansion given by

$$(3.12) \quad D(x_n - s, \dots, x_{n+k} - s) \sim \sum_{1 \leq j_0 < j_1 < \dots < j_k} \left(\prod_{p=0}^k \lambda_{j_p}^n \right) \\
\times V(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k}) \sum_{1 \leq i_1 < i_2 < \dots < i_k} \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^n \right) \\
\times V(\bar{\lambda}_{i_1}, \dots, \bar{\lambda}_{i_k}) R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k}.$$

Proof. Let us denote $D(x_n - s, \dots, x_{n+k} - s)$ by P_n for short. By (2.13), (2.1), and (3.7), we have

$$(3.13) \quad P_n \sim \begin{vmatrix} \sum_{j_0} v_{j_0} \lambda_{j_0}^n & \sum_{j_0} v_{j_0} \lambda_{j_0}^{n+1} & \dots & \sum_{j_0} v_{j_0} \lambda_{j_0}^{n+k} \\ \sum_{i_1, j_1} z_{i_1, j_1} \bar{\lambda}_{i_1}^n \lambda_{j_1}^n & \sum_{i_1, j_1} z_{i_1, j_1} \bar{\lambda}_{i_1}^n \lambda_{j_1}^{n+1} & \dots & \sum_{i_1, j_1} z_{i_1, j_1} \bar{\lambda}_{i_1}^n \lambda_{j_1}^{n+k} \\ \sum_{i_2, j_2} z_{i_2, j_2} \bar{\lambda}_{i_2}^{n+1} \lambda_{j_2}^n & \sum_{i_2, j_2} z_{i_2, j_2} \bar{\lambda}_{i_2}^{n+1} \lambda_{j_2}^{n+1} & \dots & \sum_{i_2, j_2} z_{i_2, j_2} \bar{\lambda}_{i_2}^{n+1} \lambda_{j_2}^{n+k} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i_k, j_k} z_{i_k, j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^n & \sum_{i_k, j_k} z_{i_k, j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^{n+1} & \dots & \sum_{i_k, j_k} z_{i_k, j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^{n+k} \end{vmatrix}.$$

By the multilinearity property of determinants we can reexpress (3.13) in the form

$$(3.14) \quad P_n \sim \sum_{j_0} \sum_{i_1, j_1} \cdots \sum_{i_k, j_k} v_{j_0} \lambda_{j_0}^n \left(\prod_{p=1}^k z_{i_p j_p} \lambda_{i_p}^{-n+p-1} \lambda_{j_p}^n \right) \\ \times v \left(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k} \right).$$

By changing the order of summation in (3.14), we have

$$(3.15) \quad P_n \sim \sum_{i_1} \cdots \sum_{i_k} \left(\prod_{p=1}^k \lambda_{i_p}^{-n+p-1} \right) \sum_{j_0} \sum_{j_1} \cdots \sum_{j_k} v_{j_0} \left(\prod_{p=1}^k z_{i_p j_p} \right) \\ \times \left(\prod_{p=0}^k \lambda_{j_p}^n \right) v \left(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k} \right).$$

Since $\left(\prod_{p=0}^k \lambda_{j_p}^n \right) v \left(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k} \right)$ is odd under an interchange of any two of the indices j_p , Lemma 2.1 can be applied, and we obtain

$$(3.16) \quad P_n \sim \sum_{i_1} \cdots \sum_{i_k} \left(\prod_{p=1}^k \lambda_{i_p}^{-n+p-1} \right) \sum_{1 \leq j_0 < j_1 < \cdots < j_k} R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k} \\ \times \left(\prod_{p=0}^k \lambda_{j_p}^n \right) v \left(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k} \right).$$

Interchanging the order of the summations in (3.16), we have

$$(3.17) \quad P_n \sim \sum_{1 \leq j_0 < j_1 < \dots < j_k} \left(\prod_{p=0}^k \lambda_{j_p}^n \right) v(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k}) \\ \times \sum_{i_1} \dots \sum_{i_k} \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{p-1} \right) \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{-n} \right) R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k}.$$

Let now $N_{j_q}^{i_1, \dots, i_k}$ be the cofactor of v_{j_q} in the determinant expression for $R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k}$. Then

$$(3.18) \quad R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k} = \sum_{q=0}^k v_{j_q} N_{j_q}^{i_1, \dots, i_k}.$$

Observing from (3.10) that the $N_{j_q}^{i_1, \dots, i_k}$ are odd under an interchange of any two of the indices i_1, \dots, i_k , we can apply Lemma 2.1 again to the multiple sum with respect to the indices i_1, \dots, i_k , with $R_{j_0, j_1, \dots, j_k}^{i_1, \dots, i_k}$ replaced by (3.18). The terms that are odd in the indices i_1, \dots, i_k now are

$\left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{-n} \right) N_{j_q}^{i_1, \dots, i_k}$. By Lemma 2.1 then

$$(3.19) \quad P_n \sim \sum_{1 \leq j_0 < j_1 < \dots < j_k} \left(\prod_{p=0}^k \lambda_{j_p}^n \right) v(\lambda_{j_0}, \lambda_{j_1}, \dots, \lambda_{j_k}) \\ \times \sum_{q=0}^k v_{j_q} \sum_{i_1} \dots \sum_{i_k} \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{-n} \right) N_{j_q}^{i_1, \dots, i_k} v(\bar{\lambda}_{i_1}, \dots, \bar{\lambda}_{i_k}),$$

which, by invoking (3.18), can be reexpressed as (3.12), thus completing the proof. \square

Lemma 3.2. Let the sequence of vectors $x_i, i = 0, 1, 2, \dots$, be as described in the previous section. For i_p, j_p positive integers, define

$$(3.20) \quad S_{\substack{i_1, \dots, i_k \\ j_1, \dots, j_k}} = \begin{vmatrix} z_{i_1 j_1} & z_{i_1 j_2} & \cdot & \cdot & \cdot & z_{i_1 j_k} \\ z_{i_2 j_1} & z_{i_2 j_2} & \cdot & \cdot & \cdot & z_{i_2 j_k} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ z_{i_k j_1} & z_{i_k j_2} & \cdot & \cdot & \cdot & z_{i_k j_k} \end{vmatrix}.$$

Then $D(1, \dots, 1)$ has the asymptotic expansion given by

$$(3.21) \quad D(1, \dots, 1) \sim \sum_{1 \leq j_1 < j_2 < \dots < j_k} \left(\prod_{p=1}^k \lambda_{j_p}^n \right) v(1, \lambda_{j_1}, \dots, \lambda_{j_k}) \\ \times \sum_{1 \leq i_1 < i_2 < \dots < i_k} \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{-n} \right) v(\bar{\lambda}_{i_1}, \dots, \bar{\lambda}_{i_k}) S_{\substack{i_1, \dots, i_k \\ j_1, \dots, j_k}}.$$

Proof. Let us denote $D(1, \dots, 1)$ by Q_n for short. From (2.13) and (3.7), we have

$$(3.22) \quad Q_n \sim \begin{vmatrix} 1 & 1 & \cdot & \cdot & \cdot & 1 \\ \sum_{i_1, j_1} z_{i_1 j_1} \bar{\lambda}_{i_1}^{-n} \lambda_{j_1}^n & \sum_{i_1, j_1} z_{i_1 j_1} \bar{\lambda}_{i_1}^{-n} \lambda_{j_1}^{n+1} & \cdot & \cdot & \cdot & \sum_{i_1, j_1} z_{i_1 j_1} \bar{\lambda}_{i_1}^{-n} \lambda_{j_1}^{n+k} \\ \sum_{i_2, j_2} z_{i_2 j_2} \bar{\lambda}_{i_2}^{-n+1} \lambda_{j_2}^n & \sum_{i_2, j_2} z_{i_2 j_2} \bar{\lambda}_{i_2}^{-n+1} \lambda_{j_2}^{n+1} & \cdot & \cdot & \cdot & \sum_{i_2, j_2} z_{i_2 j_2} \bar{\lambda}_{i_2}^{-n+1} \lambda_{j_2}^{n+k} \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \sum_{i_k, j_k} z_{i_k j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^n & \sum_{i_k, j_k} z_{i_k j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^{n+1} & \cdot & \cdot & \cdot & \sum_{i_k, j_k} z_{i_k j_k} \bar{\lambda}_{i_k}^{-n+k-1} \lambda_{j_k}^{n+k} \end{vmatrix}.$$

Again by the multilinearity property of determinants

$$(3.23) \quad Q_n \sim \sum_{i_1, j_1} \cdots \sum_{i_k, j_k} \left(\prod_{p=1}^k z_{i_p j_p} \lambda_{i_p}^{-n+p-1} \lambda_{j_p}^n \right) v(1, \lambda_{j_1}, \dots, \lambda_{j_k}).$$

By changing the order of summation in (3.23) and observing that

$$\left(\prod_{p=1}^k \lambda_{j_p}^n \right) v(1, \lambda_{j_1}, \dots, \lambda_{j_k})$$

is odd under an interchange of any two of the

indices j_p , we can apply Lemma 2.1 to the summation over j_1, \dots, j_k . The result is

$$(3.24) \quad Q_n \sim \sum_{i_1} \cdots \sum_{i_k} \left(\prod_{p=1}^k \lambda_{i_p}^{-n+p-1} \right) \\ \times \sum_{1 \leq j_1 < j_2 < \dots < j_k} S_{j_1, \dots, j_k}^{i_1, \dots, i_k} \left(\prod_{p=1}^k \lambda_{j_p}^n \right) v(1, \lambda_{j_1}, \dots, \lambda_{j_k}).$$

Changing the order of summation again, and observing that $\left(\prod_{p=1}^k \lambda_{i_p}^{-n} \right) S_{j_1, \dots, j_k}^{i_1, \dots, i_k}$

is odd under an interchange of any two of the indices i_p , we can now apply Lemma 2.1 to the summation over i_1, \dots, i_k . As a result, we obtain (3.21).

This completes the proof. □

Theorem 3.1. Let the sequence of vectors x_i , $i = 0, 1, 2, \dots$, be as described in the previous section. If, in addition, the v_i are linearly independent, and

$$(3.25) \quad |\lambda_1| \geq \dots \geq |\lambda_k| > |\lambda_{k+1}| \geq |\lambda_{k+2}| \geq \dots$$

then, for all sufficiently large n , $s_{n,k}$ exists, and

$$(3.26) \quad s_{n,k} - s = F(n)\lambda_{k+1}^n [1 + o(1)] \text{ as } n \rightarrow \infty,$$

where the vector $F(n)$ satisfies $\|F(n)\| \leq \tilde{K}$, for some constant $\tilde{K} > 0$ independent of n , and for all sufficiently large n . If, in addition,

$$(3.27) \quad |\lambda_{k+1}| > |\lambda_{k+2}|,$$

then

$$(3.28) \quad F(n) = \frac{R_{1,\dots,k+1}^{1,\dots,k}}{S_{1,\dots,k}^{1,\dots,k}} \prod_{i=1}^k \left(\frac{\lambda_{k+1} - \lambda_i}{\lambda_i - 1} \right) [1 + o(1)] \text{ as } n \rightarrow \infty.$$

Proof. From (3.20) and (3.8) (or (3.9)),

$$(3.29) \quad S_{1,\dots,k}^{1,\dots,k} = H G(v_1, \dots, v_k),$$

where $H = \left| \prod_{i=1}^k (\lambda_i - 1) \right|^2$ for MPE (or $H = \prod_{i=1}^k (\bar{\lambda}_i - 1)^2 (\lambda_i - 1)$ for RRE),

and $G(v_1, \dots, v_k)$ is the Gram determinant of the vectors v_1, \dots, v_k , given by

$$(3.30) \quad G(v_1, \dots, v_k) = \begin{vmatrix} (v_1, v_1) & (v_1, v_2) & \cdot & \cdot & \cdot & (v_1, v_k) \\ (v_2, v_1) & (v_2, v_2) & \cdot & \cdot & \cdot & (v_2, v_k) \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ (v_k, v_1) & (v_k, v_2) & \cdot & \cdot & \cdot & (v_k, v_k) \end{vmatrix}.$$

Since the v_i are linearly independent, $G(v_1, \dots, v_k)$ is nonzero. Also, since $\lambda_i \neq 1$ for all i , H is nonzero. Consequently, $S_{1,\dots,k}^{1,\dots,k}$ is nonzero. Next, since $\lambda_i \neq \lambda_j$ for $i \neq j$, and $\lambda_i \neq 1$ for all i , $V(\bar{\lambda}_1, \dots, \bar{\lambda}_k)$ and

$V(1, \lambda_1, \dots, \lambda_k)$ are nonzero on account of (3.11). Finally, we observe that, due to (3.25), the dominant term in the multiple sum on the right hand side of (3.21) in Lemma 3.2 would be that for which $i_p = p, j_p = p, p = 1, \dots, k$, provided $V(1, \lambda_1, \dots, \lambda_k)$ and $S_{1, \dots, k}^{1, \dots, k}$ are nonzero, which we have already proved to be so. Consequently,

$$(3.31) \quad D(1, \dots, 1) = \left| \prod_{p=1}^k \lambda_p^n \right|^2 V(1, \lambda_1, \dots, \lambda_k) V(\bar{\lambda}_1, \dots, \bar{\lambda}_k) \\ \times S_{1, \dots, k}^{1, \dots, k} [1 + o(1)] \text{ as } n \rightarrow \infty.$$

This also implies that, for all sufficiently large n , $D(1, \dots, 1) \neq 0$, hence $s_{n,k}$ exists. This completes the proof of the first part of the theorem.

To prove (3.26), we should analyze the behavior of $D(x_n - s, \dots, x_{n+k} - s)$ for $n \rightarrow \infty$. By our assumptions made following (2.1), there is only a finite number of λ_i whose moduli are $|\lambda_{k+1}|$. Let $|\lambda_{k+1}| = \dots = |\lambda_{k+r}| > |\lambda_{k+r+1}|$. From this and (3.25), it follows that the dominant term on the right hand side of (3.12) - provided it is nonzero - is the sum of those terms with indices $i_p = p, p = 1, \dots, k, j_{p-1} = p, p = 1, \dots, k, j_k = k + \ell, \ell = 1, \dots, r$, namely

$$(3.32) \quad D(x_n - s, \dots, x_{n+k} - s) = \left| \prod_{p=1}^k \lambda_p^n \right|^2 V(\bar{\lambda}_1, \dots, \bar{\lambda}_k) \\ \times \sum_{\ell=1}^r \lambda_{k+\ell}^n V(\lambda_1, \dots, \lambda_k, \lambda_{k+\ell}) R_{1, \dots, k, k+\ell}^{1, \dots, k} [1 + o(1)] \text{ as } n \rightarrow \infty.$$

From (3.10) and (3.20),

$$(3.33) \quad R_{1, \dots, k, k+\ell}^{1, \dots, k} = \sum_{i=1}^k \beta_{\ell i} v_i + (-1)^k S_{1, \dots, k}^{1, \dots, k} v_{k+\ell},$$

so that the summation on the right hand side of (3.32) becomes

$$(3.34) \quad \sum_{i=1}^k v_i \left[\sum_{\ell=1}^r \lambda_{k+\ell}^n V(\lambda_1, \dots, \lambda_k, \lambda_{k+\ell}) \beta_{\ell i} \right] \\ + (-1)^k S_{1, \dots, k}^{1, \dots, k} \sum_{\ell=1}^r \lambda_{k+\ell}^n V(\lambda_1, \dots, \lambda_k, \lambda_{k+\ell}) v_{k+\ell}.$$

Since $S_{1, \dots, k}^{1, \dots, k} \neq 0$, $V(\lambda_1, \dots, \lambda_k, \lambda_{k+\ell}) \neq 0$, and the vectors v_i are linearly independent, the second summation in (3.34) is never zero. This proves (3.32). Combining (3.31) and (3.32), (3.26) follows. If (3.27) holds, then $r = 1$. In this case (3.28) follows from (3.31), (3.32), and (3.11). This completes the proof of the theorem. □

The asymptotic error analysis of the MPE and RRE as given in Theorem 3.1, leads us to the following important conclusions:

(1) Under the conditions stated in the theorem, the MPE and RRE are bonafide acceleration methods in the sense that

$$(3.35) \quad \frac{\|s_{n,k} - s\|}{\|x_{n+k+1} - s\|} = O \left[\left(\frac{\lambda_{k+1}}{\lambda_1} \right)^n \right] \text{ as } n \rightarrow \infty.$$

This means that if $x_n \rightarrow s$ as $n \rightarrow \infty$, i.e., $|\lambda_1| < 1$, then $s_{n,k} \rightarrow s$ as $n \rightarrow \infty$, and more quickly. Also if $\lim_{m \rightarrow \infty} x_m$ does not exist, i.e., $|\lambda_1| \geq 1$, then $s_{n,k} \rightarrow s$ as $n \rightarrow \infty$, provided that $|\lambda_{k+1}| < 1$. The reason that we write x_{n+k+1} in (3.35) is that $s_{n,k}$ in both the MPE and RRE is computed from $x_n, x_{n+1}, \dots, x_{n+k+1}$.

(2) When the MPE and RRE are being applied to a vector sequence generated by using a matrix iterative method, they will be especially effective when the iteration matrix 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) As can be seen from (3.28), a loss of accuracy will take place in $s_{n,k}$ when $\lambda_1, \dots, \lambda_k$ are close to 1. For sequences of vectors obtained from the iterative solution of linear systems of equations, this means that the matrix of the system is nearly singular.

These conclusions are the same as those for the modified MPE, which has been defined and analyzed in [3].

Finally note that the results of this section (and of Section 4) will not change if the MPE and RRE are replaced by any other method giving rise to (3.1), with $D(\sigma_0, \dots, \sigma_k)$ defined as in (2.13), as long as the $u_{i,j}$ in (2.13) satisfy (3.7) with $z_{ij} = (v_i, v_j)\mu_i v_j$, where μ_i and v_j are fixed nonzero constants.

4. STABILITY ANALYSIS

Let us denote γ_j in (2.5) by $\gamma_j^{(n,k)}$ for both MPE and RRE. We say that $s_{n,k}$ is asymptotically stable if

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

Roughly speaking, this means that if errors are introduced in the vectors x_m ,

then the error in $s_{n,k}$ stays bounded as $n \rightarrow \infty$. Since $\sum_{j=0}^k \gamma_j^{(n,k)} = 1$, the

most ideal situation is one in which $\gamma_j^{(n,k)} \geq 0$, $0 \leq j \leq k$, for all suffi-

ciently large n , so that $\sum_{j=0}^k \left| \gamma_j^{(n,k)} \right| = \sum_{j=0}^k \gamma_j^{(n,k)} = 1$.

The following theorem gives the stability properties of both the MPE and RRE when they are applied to vector sequences satisfying the conditions of Theorem 3.1. $s_{n,k}$ denotes either that obtained from MPE or RRE.

Theorem 4.1. For vector sequences satisfying the conditions stated in Theorem 3.1, $s_{n,k}$ is asymptotically stable. Actually, the following are true:

(1) The $\gamma_q^{(n,k)}$ satisfy

$$(4.2) \quad \gamma_q^{(n,k)} = \frac{C_q(\lambda_1, \dots, \lambda_k)}{V(1, \lambda_1, \dots, \lambda_k)} [1 + o(1)] \quad \text{as } n \rightarrow \infty,$$

where

$$(4.3) \quad C_q(\xi_1, \dots, \xi_k) = (-1)^q \begin{vmatrix} 1 & \xi_1 & \dots & \xi_1^{q-1} & \xi_1^{q+1} & \dots & \xi_1^k \\ 1 & \xi_2 & \dots & \xi_2^{q-1} & \xi_2^{q+1} & \dots & \xi_2^k \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & \xi_k & \dots & \xi_k^{q-1} & \xi_k^{q+1} & \dots & \xi_k^k \end{vmatrix},$$

thus

$$(4.4) \quad \sum_{q=0}^k \left| \gamma_q^{(n,k)} \right| = \frac{\sum_{q=0}^k |C_q(\lambda_1, \dots, \lambda_k)|}{|V(1, \lambda_1, \dots, \lambda_k)|} [1 + o(1)] \quad \text{as } n \rightarrow \infty,$$

i.e., (4.1) holds.

(2) If $\lambda_1, \dots, \lambda_k$ are real and negative, then

$$(4.5) \quad \gamma_q^{(n,k)} > 0, \quad 0 \leq q \leq k, \quad \text{for } n \text{ sufficiently large,}$$

thus

$$(4.6) \quad \sum_{q=0}^k \left| \gamma_q^{(n,k)} \right| = 1 + o(1) \quad \text{as } n \rightarrow \infty.$$

(3) If $\lambda_1, \dots, \lambda_k$ are real and positive, then

$$(4.7) \quad \gamma_q^{(n,k)} \gamma_{q+1}^{(n,k)} < 0, \quad 0 \leq q \leq k-1, \quad \text{for } n \text{ sufficiently large,}$$

thus

$$(4.8) \quad \sum_{q=0}^k \left| \gamma_q^{(n,k)} \right| = \left| \prod_{i=1}^k \left(\frac{\lambda_i + 1}{\lambda_i - 1} \right) \right| [1 + o(1)] \quad \text{as } n \rightarrow \infty.$$

(4) For any $\lambda_1, \dots, \lambda_k$,

$$(4.9) \quad \lim_{n \rightarrow \infty} \sum_{q=0}^k \gamma_q^{(n,k)} \lambda^q = \prod_{i=1}^k \left(\frac{\lambda_i - \lambda}{\lambda_i - 1} \right).$$

Proof. Since $\gamma_q^{(n,k)}$ is given by (2.12), and we already know the asymptotic behavior of $D(1, \dots, 1)$ from Lemma 3.2, it is sufficient to analyze N_q asymptotically. By deleting the first row and the $(q+1)$ st column of the determinant in (3.22), after some manipulation, we have

$$(4.10) \quad N_q \sim \sum_{i_1, j_1} \dots \sum_{i_k, j_k} \left(\prod_{p=1}^k z_{i_p j_p} \frac{\lambda_{i_p}^{-n+p-1} \lambda_{j_p}^n}{\lambda_{i_p} \lambda_{j_p}} \right) C_q(\lambda_{j_1}, \dots, \lambda_{j_k}).$$

Observe that $\left(\prod_{p=1}^k \lambda_{j_p}^n \right) C_q(\lambda_{j_1}, \dots, \lambda_{j_k})$ is odd under an interchange of any two

of the indices j_p . Consequently,

$$(4.11) \quad N_q \sim \sum_{1 \leq j_1 < j_2 < \dots < j_k} \left(\prod_{p=1}^k \lambda_{j_p}^n \right) C_q(\lambda_{j_1}, \dots, \lambda_{j_k}) \\ \times \sum_{1 \leq i_1 < i_2 < \dots < i_k} \left(\prod_{p=1}^k \bar{\lambda}_{i_p}^{-n} \right) V(\bar{\lambda}_{i_1}, \dots, \bar{\lambda}_{i_k}) S_{j_1, \dots, j_k}^{i_1, \dots, i_k}$$

follows from (4.10) in exactly the same way (3.21) follows from (3.23). Invoking (3.25),

$$(4.12) \quad N_q = \left| \prod_{p=1}^k \lambda_p^n \right|^2 V(\bar{\lambda}_1, \dots, \bar{\lambda}_k) C_q(\lambda_1, \dots, \lambda_k) S_{1, \dots, k}^{1, \dots, k} [1 + o(1)]$$

as $n \rightarrow \infty$,

follows from (4.11) in the same way (3.31) follows from (3.21). Combining (4.12) and (3.31) in (2.12), (4.2) follows. (4.4) follows directly from (4.2). This proves (1).

Note that $V(\lambda, \lambda_1, \dots, \lambda_k)$ is a polynomial of degree k in λ . From (3.11) and (4.3) we have

$$(4.13) \quad V(\lambda, \lambda_1, \dots, \lambda_k) = \sum_{q=0}^k C_q(\lambda_1, \dots, \lambda_k) \lambda^q.$$

Since $V(\lambda, \lambda_1, \dots, \lambda_k)$ has $\lambda_1, \dots, \lambda_k$ as its only zeros, we also have

$$(4.14) \quad V(\lambda, \lambda_1, \dots, \lambda_k) = C_k(\lambda_1, \dots, \lambda_k) \prod_{i=1}^k (\lambda - \lambda_i).$$

If $\lambda_1, \dots, \lambda_k$ are real and negative, then (4.13) and (4.14) imply that $C_q(\lambda_1, \dots, \lambda_k)$, $0 \leq q \leq k$, are all of the same sign. This, along with (4.2),

implies (4.5). (4.6) is a direct consequence of (4.5) and (2.11). This proves (2).

If $\lambda_1, \dots, \lambda_k$ are real and positive, then (4.13) and (4.14) imply that $C_q(\lambda_1, \dots, \lambda_k)C_{q+1}(\lambda_1, \dots, \lambda_k) < 0$, $0 \leq q \leq k - 1$. This, along with (4.2), implies (4.7) and

$$(4.15) \quad \sum_{q=0}^k \left| \gamma_q^{(n,k)} \right| = \frac{\left| \sum_{q=0}^k (-1)^q C_q(\lambda_1, \dots, \lambda_k) \right|}{|V(1, \lambda_1, \dots, \lambda_k)|} [1 + o(1)] \text{ as } n \rightarrow \infty$$

$$= \left| \frac{V(-1, \lambda_1, \dots, \lambda_k)}{V(1, \lambda_1, \dots, \lambda_k)} \right| [1 + o(1)] \text{ as } n \rightarrow \infty,$$

which, by (3.11), reduces to (4.8). This proves (3).

Multiplying both sides of (4.2) by λ^q , summing over q from 0 to k , and finally making use of (4.13) and (3.11), (4.9) follows, thus proving (4).

□

REFERENCES

- [1] S. Cabay and L. W. Jackson, A polynomial extrapolation method for finding limits and antilimits of vector sequences, SIAM, J. Numer. Anal., 13 (1976), pp. 734-752.
- [2] R. P. Eddy, Extrapolating to the limit of a vector sequence, in Information Linkage between Applied Mathematics and Industry, (P.C.C. Wang, ed.), New York; Academic Press, 1979, pp. 387-396.
- [3] A. Sidi, W. F. Ford, and D. A. Smith, Acceleration of convergence of vector sequences. NASA TP-2193, 1983.

| | | | | | |
|---|--|--|--|---|------------|
| 1. Report No. NASA TM-83443 | | 2. Government Accession No. | | 3. Recipient's Catalog No. | |
| 4. Title and Subtitle Convergence and Stability Properties of Minimal Polynomial and Reduced Rank Extrapolation Algorithms | | | | 5. Report Date July 1983 | |
| | | | | 6. Performing Organization Code None | |
| 7. Author(s) Avram Sidi | | | | 8. Performing Organization Report No. E-1659 | |
| | | | | 10. Work Unit No. | |
| 9. Performing Organization Name and Address National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135 | | | | 11. Contract or Grant No. | |
| | | | | 13. Type of Report and Period Covered Technical Memorandum | |
| 12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546 | | | | 14. Sponsoring Agency Code | |
| | | | | | |
| 15. Supplementary Notes Avram Sidi, Computer Science Department, Technion - Israel Institute of Technology, Haifa, Israel. Work partly done as a National Research Council - NASA Research Associate at Lewis Research Center. | | | | | |
| 16. Abstract The minimal polynomial and reduced rank extrapolation algorithms are two acceleration of convergence methods for sequences of vectors. In a recent survey these methods were tested and compared with the scalar, vector, topological epsilon algorithms, and were observed to be more efficient than the latter. It was also observed that the two methods have similar convergence properties. The purpose of the present work is to analyze the convergence and stability properties of these methods, and to show that they are bona fide acceleration methods when applied to a class of vector sequences that includes those sequences obtained from systems of linear equations by using matrix iterative methods. | | | | | |
| 17. Key Words (Suggested by Author(s)) Vector sequences Acceleration of convergence Minimal polynomial extrapolation Reduced rank extrapolation | | | 18. Distribution Statement Unclassified - unlimited STAR Category 64 | | |
| 19. Security Classif. (of this report) Unclassified | | 20. Security Classif. (of this page) Unclassified | | 21. No. of pages | 22. Price* |

National Aeronautics and
Space Administration

Washington, D.C.
20546

Official Business

Penalty for Private Use, \$300

SPECIAL FOURTH CLASS MAIL
BOOK



Postage and Fees Paid
National Aeronautics and
Space Administration
NASA-451

NASA

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return
