Development of Iterative Techniques and Extrapolation Methods for Drazin Inverse Solution of Consistent or Inconsistent Singular Linear Systems

Avram Sidi

Computer Science Department
Technion – Israel Institute of Technology
Haifa 32000, Israel

Submitted by Abraham Berman

ABSTRACT

Consider the linear system of equations $Bx = f$, where B is an $N \times N$ singular matrix. In an earlier work by the author it was shown that iterative techniques coupled with standard vector extrapolation methods can be used to obtain or approximate a solution of this system when it is consistent. In the present work we expand on that approach to treat the case in which this system is in general inconsistent. Starting with Richardson's iterative method, we develop a family of new iterative techniques and vector extrapolation methods that enable us to obtain or approximate the Drazin inverse solution of this system whether the index of B is 1 or greater than 1. We show that the Drazin inverse solution can be constructed from a finite number of iterations, this number being at most $N + 2$. We also provide detailed convergence analyses of the new iterative techniques and vector extrapolation methods and give their precise rates of convergence.

1. INTRODUCTION

Consider the system of linear equations

$$Bx = f, \tag{1.1}$$

where B is a singular complex $N \times N$ matrix that is not necessarily Hermitian, and f is an N-dimensional complex vector. As is well known, if this system is consistent, then it has an infinity of solutions; otherwise, it has no solution.

LINEAR ALGEBRA AND ITS APPLICATIONS 167: 171–203 (1992)

© Elsevier Science Publishing Co., Inc., 1992
655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$05.00

In an earlier work by the author [9], it was shown that if the system in (1.1) is consistent, then both iterative techniques and standard vector extrapolation methods coupled with iterative techniques can be used to produce or approximate a solution of (1.1) under certain conditions. In particular, it was shown there that this may be possible when the index of B is 1, i.e., when the zero eigenvalues of B have only corresponding eigenvectors and no principal vectors, which is the case, for example, when B is diagonalizable. This may be possible also when the index of B is greater than 1, i.e., when the zero eigenvalues of B have corresponding principal vectors, but in this case a rather stringent condition is needed for the initial vector that is used in the iterative technique. In either case, the solution that is obtained turns out to be very closely related to the Drazin inverse solution of (1.1). For the Drazin inverse and its various applications we refer the reader to the books by Ben-Israel and Greville [1] and Campbell and Meyer [2].

Encouraged by the positive results of [9], in the present work we expand on that approach to give a rather thorough treatment of the problem of determining (or approximating) the Drazin inverse solution of (1.1). We do this under the most general conditions on B and f, namely, that the index of B is arbitrary, and that (1.1) may be consistent or inconsistent.

Roughly speaking, the Drazin inverse solution, which we denote throughout by s', is the unique vector that lies in the subspace $\mathscr{S}(B)$ spanned by the eigenvectors and principal vectors of B corresponding to its nonzero eigenvalues, and that satisfies the consistent system $Bx = f'$, where f' is that part of the vector f that lies in $\mathscr{S}(B)$. In case the matrix B has index 1 and the system in (1.1) is consistent, s' turns out to be a true solution. In case B is range-Hermitian, i.e., B and B^* have the same range, s' turns out to be the Moore-Penrose generalized inverse solution of the system in (1.1), whether the latter is consistent or not. (Recall that if B is diagonalizable, then its index is 1, and if B is normal, then it is range-Hermitian; see [1].) Trivially, when B is nonsingular, s' is the unique solution to (1.1). When (1.1) is consistent and thus has solutions, but the index of B is strictly greater than 1, the Drazin inverse solution of (1.1) does not necessarily satisfy (1.1). A rigorous and detailed discussion of all these points will be given in Section 2.

It is important to emphasize from the start that standard iterative techniques cannot be used in a straightforward manner to approximate s', the Drazin inverse solution of (1.1), in general. The reason for this is that the sequence of approximations generated this way may never converge to s' or may never converge at all, as follows from Theorem 3.1 of the present work. Similarly, standard vector extrapolation methods cannot be employed to obtain approximations to s'. Therefore, there is need for a totally new approach to develop techniques for approximating Drazin inverse solutions. The approach taken in this paper is based on standard iterative techniques, which, despite their possible divergence, contain valuable information on s' that can be extracted in various ways.

Throughout the present work we will consider Richardson's iterative method for (1.1) (see Varga [13, p. 141]) as our standard iterative technique. In this technique

$$x_{j+1} = x_j + \omega\left(f - Bx_j\right), \qquad j = 0, 1, \dots, \tag{1.2}$$

x_0 being an arbitrary initial vector. This is the iterative technique considered also in [9]. It is clear that (1.2) can be expressed in the matrix iterative form

$$x_{j+1} = Ax_j + b, \qquad j = 0, 1, \dots, \tag{1.3}$$

where

$$A = I - \omega B \quad \text{and} \quad b = \omega f. \tag{1.4}$$

As mentioned in [9], any other fixed point iterative technique for solving (1.1) is also of the form (1.3) with $A = M^{-1}Q$, where $B = M - Q$, and is a Richardson iterative method with $\omega = 1$ for the "preconditioned" linear system $M^{-1}Bx = M^{-1}f$. Thus there is no loss of generality in considering only Richardson's iterative method.

The plan of the present work is as follows: In the next section we review some of the results of [9], and introduce part of the notation that is used throughout the remainder. In Section 3 we analyze Richardson's iterative method for the singular system in (1.1), whether the latter is consistent or not. We subsequently modify the sequence of vectors $\{x_m\}_{m=0}^{\infty}$ obtained from it to produce another sequence $\{\hat{x}_m\}_{m=0}^{\infty}$ that converges to s', the Drazin inverse solution of (1.1), under certain conditions pertaining to the spectrum of B. These conditions are also needed when the system in (1.1) is singular and consistent, or even nonsingular. The main result of Section 3 is Theorem 3.2. In Section 4 we show how s' can be constructed from a finite number of the vectors x_j obtained from Richardson's iterative method. We also provide a precise count of these vectors and show that their number does not exceed $N + 2$. The main result in Section 4 is Theorem 4.2. Based on the construction of Section 4, in Section 5 we propose new extrapolation (or convergence acceleration) methods that produce approximations to s' with much larger convergence rates than the vectors in the sequence $\{\hat{x}_m\}_{m=0}^{\infty}$. A detailed convergence analysis of these methods is provided in Sections 6 and 7, the main result being Theorem 7.3.

We note that some of the techniques developed in the present work are similar to and extend those that were used in the papers by Sidi [7–9], Sidi and

Bridger [10], Sidi, Ford, and Smith [11], and Smith, Ford, and Sidi [12].

Use of iterative techniques for obtaining the Drazin inverse solution of a singular system has been considered in several papers. For example, Meyer and Plemmons [6] consider the Drazin inverse solution of a consistent singular system with a semiconvergent splitting of the matrix. In a recent work Eiermann, Marek, and Niethammer [3] consider the Drazin inverse solution of a singular system under the assumptions made in the present paper, and give a general framework for semiiterative methods appropriate for this purpose. Some of the material presented in Theorem 3.2 of the present work is closely related to [3]. We shall elaborate on this in Section 3.

Finally, we mention that the approach that we take in Section 4 for the construction of the Drazin inverse solution from Richardson's iterative method, and the extrapolation methods that follow in Sections 5–7, are completely new.

2. THEORETICAL PRELIMINARIES

Consider the linear system in (1.1) with B and f as described in the first paragraph of Section 1. There exists a nonsingular matrix V such that

$$V^{-1}BV = J = \begin{bmatrix} J_1 & & & \\ & J_2 & 0 & \\ & 0 & \ddots & \\ & & & J_\nu \end{bmatrix}, \tag{2.1}$$

where J_i are Jordan blocks of dimension r_i and have the form

$$J_i = \begin{bmatrix} \mu_i & 1 & & & 0 \\ & \mu_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & \mu_i \end{bmatrix}_{r_i \times r_i}, \qquad \mu_i \text{ an eigenvalue}. \tag{2.2}$$

If V has the columnwise partition

$$V = \left[v_{11} \,|\, v_{12} \,|\, \cdots \,|\, v_{1r_1} \,|\, v_{21} \,|\, v_{22} \,|\, \cdots \,|\, v_{2r_2} \,|\, \cdots \,|\, v_{\nu 1} \,|\, v_{\nu 2} \,|\, \cdots \,|\, v_{\nu r_\nu} \right], \tag{2.3}$$

then v_{i1} is the eigenvector corresponding to the eigenvalue μ_i, and in case $r_i > 1$, $v_{i2}, \dots, v_{ir_i}$ are the principal vectors (or the generalized eigenvectors) corresponding to μ_i. We actually have

$$\begin{aligned} Bv_{i1} &= \mu_i v_{i1}, \\ Bv_{ij} &= \mu_i v_{ij} + v_{i,j-1}, \quad j = 2, \dots, r_i, \qquad \text{for} \quad r_i > 1. \end{aligned} \tag{2.4}$$

Let us denote

$$\begin{aligned} \mathscr{S}(B) &= \operatorname{span}\{v_{ij}, 1 \leqslant j \leqslant r_i : \mu_i \neq 0\}, \\ \mathscr{N}(B) &= \operatorname{span}\{v_{i1} : \mu_i = 0\}, \qquad \text{null space of } B, \\ \mathscr{M}(B) &= \operatorname{span}\{v_{ij}, 2 \leqslant j \leqslant r_i : \mu_i = 0, r_i > 1\}, \\ \mathscr{T}(B) &= \mathscr{N}(B) \oplus \mathscr{M}(B). \end{aligned} \tag{2.5}$$

Obviously, the intersection of any two of the subspaces $\mathscr{S}(B)$, $\mathscr{N}(B)$, and $\mathscr{M}(B)$ consists of the zero vector only, and $\mathbf{C}^N = \mathscr{S}(B) \oplus \mathscr{N}(B) \oplus \mathscr{M}(B) = \mathscr{S}(B) \oplus \mathscr{T}(B)$. If $r_i = 1$ for all the eigenvalues $\mu_i = 0$, $\mathscr{M}(B)$ is defined to be the empty set, and $y \in \mathscr{M}(B)$ is interpreted as $y = 0$. This happens, for example, when B is diagonalizable.

We recall that the integer $\max\{r_i : \mu_i = 0\}$ is the index of B, and we shall denote it by d throughout.

We now state Theorem 2.1 of [9], which provides us with a necessary and sufficient condition for the consistency of the system in (1.1) exclusively in terms of the eigenvectors and principal vectors of B:

THEOREM 2.1. *When B is singular, the system in* (1.1) *has a solution if and only if f can be expanded in terms of the columns of the matrix V, excluding the vectors v_{ir_i} corresponding to the zero eigenvalues μ_i. If a solution s exists, then it is of the form $s = s' + s'' + s'''$, where $s' \in \mathscr{S}(B)$ and $s'' \in \mathscr{M}(B)$ and they are uniquely determined, and $s''' \in \mathscr{N}(B)$ and is nonunique. In fact, if $f = f' + f''$, where $f' \in \mathscr{S}(B)$ and $f'' \in \mathscr{M}(B)$, then $Bs' = f'$ and $Bs'' = f''$.* [*Recall also that $s'' = 0$ when $\mathscr{M}(B)$ is the empty set.*]

REMARKS.

(1) Using the Jordan canonical form of the matrix B^D, the Drazin inverse of B, it is easy to verify that s' is, in fact, $B^D f$, the Drazin inverse solution of (1.1). (For the Jordan canonical form of B^D, see, e.g., [3].)

(2) When d, the index of B, is 1, the Drazin inverse B^D is also called the group inverse and is denoted by $B^{\#}$. In case B is range-Hermitian, we have $d = 1$, and B^D coincides with B^+, the Moore-Penrose generalized inverse of B, so that $s' = B^+ f$. A normal matrix is a range-Hermitian matrix.

Before we end this section, we note that the eigenvalues of the iteration matrix A in (1.3) and (1.4) are $\lambda_i = 1 - \omega\mu_i$, and the corresponding eigenvectors are $\hat{v}_{i1} = v_{i1}$, while the corresponding principal vectors are $\hat{v}_{ij} = (-\omega)^{-j+1} v_{ij}$, $2 \leqslant j \leqslant r_i$, for $r_i > 1$, exactly in the sense of (2.4). Corresponding to $\mu_i = 0$, we have $\lambda_i = 1$; hence $A\hat{v}_{i1} = \hat{v}_{i1}$. Also,

$$A^m \hat{v}_{ij} = \sum_{l=1}^{j} \binom{m}{j-l} \lambda_i^{m-j+l} \hat{v}_{il}. \tag{2.6}$$

All this is mentioned in [9].

Finally, the following lemma will be of utmost importance in the next sections.

LEMMA 2.2. *Let $\phi(z)$ be a scalar or vector valued polynomial of degree d. Then, for any ζ,*

$$\phi(\zeta) - \phi(0) = -\sum_{i=1}^{d} \left[\Delta^i \phi(\zeta)\right] \binom{-\zeta}{i}, \tag{2.7}$$

where $\Delta\phi(\zeta) = \phi(\zeta + 1) - \phi(\zeta)$, $\Delta^2\phi(\zeta) = \Delta(\Delta\phi(\zeta))$, etc.

Proof. By Newton's interpolation formula at the points $\zeta, \zeta + 1, \dots, \zeta + d$, we have

$$\phi(z) = \phi(\zeta) + \sum_{i=1}^{d} \left[\Delta^i \phi(\zeta)\right] \binom{z-\zeta}{i}. \tag{2.8}$$

The result follows by setting $z = 0$ in (2.8). ∎

3. RICHARDSON'S ITERATIVE METHOD FOR SINGULAR SYSTEMS

Consider Richardson's iterative method as described in (1.2)–(1.4) for the system (1.1), assuming that this system is not necessarily consistent. Our purpose now is to first analyze the nature of the sequence $\{x_m\}_{m=0}^{\infty}$ obtained by employing this method, and then to propose a modification. This modification results in a convergent sequence that has a useful limit, namely s', the Drazin inverse solution, under certain conditions pertaining to the spectrum of B.

THEOREM 3.1. *Let $f = f' + \tilde{f}$ and $x_0 = x_0' + \tilde{x}_0$ with $f', x_0' \in \mathcal{S}(B)$ and $\tilde{f}, \tilde{x}_0 \in \mathcal{T}(B)$. Let s' be the unique solution of $Bx = f'$ that lies in $\mathcal{S}(B)$: see Theorem 2.1. Then the sequence $\{x_m\}_{m=0}^{\infty}$ generated by Richardson's iterative method satisfies*

$$x_m - s' = A^m(x_0' - s') + T(m), \tag{3.1}$$

where $T(m)$ is a polynomial in m of degree at most $d = \max\{r_i : \mu_i = 0\}$, the index of B, with vector coefficients in $\mathcal{T}(B)$, and is such that $T(0) = \tilde{x}_0$. [The exact degree τ of $T(m)$ depends on $\tilde{f}$ and $\tilde{x}_0$, and can be deduced form (3.5)–(3.10) in the proof below. In case (1.1) is inconsistent, $\tau = d$, in general.]

Proof. From (1.2)–(1.4), $f = f' + \tilde{f}$, and $Bs' = f'$, it follows that

$$x_{j+1} - s' = A(x_j - s') + \omega\tilde{f}, \qquad j = 0, 1, \dots . \tag{3.2}$$

By induction, (3.2) implies

$$x_m - s' = A^m(x_0 - s') + R(m), \tag{3.3}$$

with

$$R(m) = \left(\sum_{q=0}^{m-1} A^q\right)\omega\tilde{f}, \qquad m = 1, 2, \dots . \tag{3.4}$$

Now $\tilde{f} \in \mathcal{T}(B)$ means that

$$\omega\tilde{f} = \sum_i{}' \sum_{j=1}^{r_i} \delta_{ij}\hat{v}_{ij}, \tag{3.5}$$

where $\sum_i'$ stands for summation over those values of i for which $\mu_i = 0$. Substituting (3.5) in (3.4), and invoking (2.6), we have

$$R(m) = \sum_i{}' \sum_{j=1}^{r_i} \delta_{ij} \sum_{l=1}^{j} \hat{v}_{il} \sum_{q=0}^{m-1} \binom{q}{j-l}, \tag{3.6}$$

which, by the identity

$$\sum_{k=0}^{p} \binom{k}{i} = \binom{p+1}{i+1},$$

becomes

$$R(m) = \sum_i{}' \sum_{j=1}^{r_i} \delta_{ij} \sum_{l=1}^{j} \hat{v}_{il} \binom{m}{j-l+1} = \sum_i{}' \sum_{p=1}^{r_i} \left(\sum_{j=p}^{r_i} \delta_{ij} \hat{v}_{i,j-p+1} \right) \binom{m}{p}. \tag{3.7}$$

From the fact that

$$\binom{m}{i} = \frac{m(m-1)\dots(m-i+1)}{i!}$$

is a polynomial in m of degree exactly i it follows that $R(m)$ is a polynomial in m of degree at most $d = \max\{r_i : \mu_i = 0\}$ with vector coefficients in $\mathscr{T}(B)$, and is such that $R(0) = 0$. Also, when the system in (1.1) is inconsistent, then $\delta_{ir_i} \neq 0$ for some i in (3.5). If all of these δ_{ir_i} are nonzero when (1.1) is inconsistent, then the exact degree of $R(m)$ is exactly d.

Invoking now $x_0 = x_0' + \tilde{x}_0$ on the right hand side of (3.3), we obtain (3.1) with

$$T(m) = A^m \tilde{x}_0 + R(m). \tag{3.8}$$

Again, by $\tilde{x}_0 \in \mathscr{T}(B)$, we have

$$\tilde{x}_0 = \sum_i{}' \sum_{j=1}^{r_i} \varepsilon_{ij} \hat{v}_{ij}. \tag{3.9}$$

Thus, by (2.6), after some manipulation,

$$A^m \tilde{x}_0 = \sum_i{}' \sum_{p=0}^{r_i - 1} \left(\sum_{j=p+1}^{r_i} \varepsilon_{ij} \hat{v}_{i,j-p} \right) \binom{m}{p}. \tag{3.10}$$

As is seen, $A^m \tilde{x}_0$ is a polynomial in m of degree at most $d - 1$, with vector coefficients in $\mathscr{T}(B)$. Also, letting $m = 0$ in (3.8), we obtain $T(0) = \tilde{x}_0$. This completes the proof of the theorem. ■

Theorem 2.2 in [9], which is an extension of a known result on Richardson's iterative method, states that when all the nonzero eigenvalues of the matrix B are in the same open half of the complex plane containing the origin on its boundary, the sequence $\{x_m\}_{m=0}^{\infty}$ converges for some (complex) ω if and only if the system in (1.1) has a solution s, and $x_0 - s \in \mathscr{S}(B) \oplus \mathscr{N}(B)$. In Theorem 3.2 below we provide a further extension of this result. Specifically, we first propose a modification of Richardson's iterative method, and then show that it converges to s' starting from *any* initial vector, whether the singular system in (1.1) is consistent or not.

THEOREM 3.2. *Assume that*

$$\arg \mu_i \in \left(\theta - \frac{\pi}{2}, \theta + \frac{\pi}{2}\right) \quad \textit{for} \quad \mu_i \neq 0, \quad \textit{some } \theta,$$

$$0 < |\omega| < \frac{2\cos\alpha}{\rho(B)}, \quad \alpha = \max\{|\arg \mu_i - \theta| : \mu_i \neq 0\}; \qquad \arg\omega = -\theta, \tag{3.11}$$

where $\rho(B) = \max_i |\mu_i|$ *is the spectral radius of* B. *Let* $\{x_m\}_{m=0}^{\infty}$ *be the sequence obtained from* (1.2)–(1.4). *Let* $d = \max\{r_i : \mu_i = 0\}$, *the index of* B, *as before. Now set*

$$\hat{x}_m = x_m + \sum_{i=1}^{d} \left(\Delta^i x_m\right)\binom{-m}{i}$$

$$= x_m + \sum_{i=1}^{d} \frac{(-1)^i\left(\Delta^i x_m\right)}{i!} \prod_{j=0}^{i-1}(m+j), \tag{3.12}$$

where $\Delta x_m = x_{m+1} - x_m$, $\Delta^2 x_m = \Delta(\Delta x_m)$, *etc. Then the sequence* $\{\hat{x}_m\}_{m=0}^{\infty}$ *converges to* $s' + \tilde{x}_0$, *where* s' *and* $\tilde{x}_0$ *are exactly as described in Theorem* 3.1. *In fact,*

$$\hat{x}_m = s' + \tilde{x}_0 + O\left(m^{d+h-1}\tilde{\rho}^m\right) \quad \textit{as} \quad m \to \infty, \tag{3.13}$$

where $\tilde{\rho} = \max\{|\lambda_i| : \lambda_i \neq 1\}$ *and* $h = \max\{r_i : |\lambda_i| = \tilde{\rho}\}$. *Here* $\lambda_i = 1 - \omega\mu_i$ *are, as before, eigenvalues of* A.

Proof. First, Theorem 3.1 applies, and (3.1) holds, and $T(m)$ is a polynomial in m of degree at most d, with vector coefficients in $\mathscr{T}(B)$. Thus, by Lemma 2.2,

$$T(m) - T(0) = -\sum_{i=1}^{d}\left[\Delta^i T(m)\right]\binom{-m}{i}. \tag{3.14}$$

But $T(0) = \tilde{x}_0$ and $T(m) = x_m - s' - e_m$, where we have denoted $e_m = A^m(x_0' - s')$ for short. Therefore, (3.14) becomes

$$x_m - s' - e_m - \tilde{x}_0 = -\sum_{i=1}^{d}\left(\Delta^i x_m - \Delta^i e_m\right)\binom{-m}{i}. \tag{3.15}$$

Invoking (3.12) in(3.15), we have

$$\hat{x}_m - s' - \tilde{x}_0 = e_m + \sum_{i=1}^{d}\left(\Delta^i e_m\right)\binom{-m}{i}. \tag{3.16}$$

Now $x_0' - s' \in \mathscr{S}(B)$, and $\mathscr{S}(B)$ is the subspace spanned by the eigenvectors and principal vectors v_{ij} of B corresponding to the nonzero eigenvalues μ_i, or equivalently, the subspace spanned by the eigenvectors and principal vectors $\hat{v}_{ij}$ of A corersponding to eigenvalues λ_i that are not unity. Also, $\lambda_i \neq 1$ implies $|\lambda_i| < 1$ under the assumptions made in the statement of the theorem; see [9, Theorem 2.2]. Consequently,

$$e_m = A^m(x_0' - s') = O\left(m^{h-1}\tilde{\rho}^m\right) \quad \text{as} \quad m \to \infty. \tag{3.17}$$

Substituting (3.17) on the right hand side of (3.16), we obtain (3.13), thus proving the theorem. ■

REMARKS.

(1) If there is no θ for which $\arg \mu_i \in (\theta - \pi/2, \theta + \pi/2)$, all $\mu_i \neq 0$, then there exists no complex ω for which $\{\hat{x}_m\}_{m=0}^{\infty}$ converges. Such a situation arises, for example, when B has both positive and negative eigenvalues.

(2) When $\operatorname{Re}\mu_i > 0$ for $\mu_i \neq 0$, we can take $\theta = 0$; thus we can choose ω to be positive real, as can be seen from (3.11).

(3) By picking $x_0 = 0$, or $x_0 = B^d\xi$ for an arbitrary vector ξ, we can cause $\tilde{x}_0 = 0$ everywhere in Theorem 3.2. This results in $\lim_{m\to\infty}\hat{x}_m = s'$.

(4) When $d = 1$, (3.1) and (3.12) take on very simple forms, namely

$$x_m - s' = A^m(x_0' - s') + \tilde{x}_0 + m\omega\tilde{f} \tag{3.18}$$

and

$$\hat{x}_m = x_m - m\,\Delta x_m, \tag{3.19}$$

respectively.

In view of remarks (3) and (4) above, we have the following very useful result for the case in which B is a range-Hermitian matrix:

THEOREM 3.3. *If the matrix B in Theorem 3.2 is range-Hermitian, and if $x_0 = 0$, or if $x_0 = B\xi$ for an arbitrary vector ξ, then the sequence $\{\hat{x}_m\}_{m=0}^{\infty}$, where $\hat{x}_m = x_m - m\,\Delta x_m$, $m = 0, 1, \ldots$, converges to $s' = B^+ f$; in fact,*

$$\hat{x}_m = B^+ f + O(m\tilde{\rho}^m) \qquad \text{as} \quad m \to \infty, \tag{3.20}$$

with $\tilde{\rho}$ as defined following (3.13).

Proof. Observe that for this case $r_i = 1$ for all i, so that $d = 1$, $h = 0$, $\tilde{x}_0 = 0$, and $s' = B^+ f$ in Theorem 3.2. The details are left to the reader. ■

REMARK. If x_0 is picked arbitrarily in Theorem 3.3, then $\tilde{x}_0$ is not necessarily zero and should be added to the right hand side of (3.20).

3.1. *A Multipoint Iterative Technique for $\{\hat{x}_m\}$*

We have not been able to find a stationary fixed point method that generates the vectors $\hat{x}_m$. A stationary multipoint iterative technique does exist, however, and we turn to this in Theorem 3.4 below.

THEOREM 3.4. *The vectors $\hat{x}_m$ of Theorem 3.2 can be computed recursively from*

$$\hat{x}_{m+d+1} = \sum_{j=0}^{d} (-1)^{d-j} \binom{d+1}{j} A^{d+1-j} \hat{x}_{m+j} + \omega^{d+1} B^d f, \qquad m \geqslant 0, \tag{3.21}$$

with $\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_d$ as given in (3.12).

Proof. We shall prove (3.21) by verification. First, we note that

$$A^q x_m = x_{m+q} - \sum_{l=0}^{q-1} A^l b, \qquad \text{all } m, q. \tag{3.22}$$

Let us set

$$Q = \sum_{j=0}^{d+1} (-1)^{d+1-j} \binom{d+1}{j} A^{d+1-j} \hat{x}_{m+j}. \tag{3.23}$$

Substituting (3.12) in (3.23), and invoking (3.22), we obtain

$$Q = \sum_{j=0}^{d+1} (-1)^{d+1-j} \binom{d+1}{j} \sum_{i=0}^{d} \left[\Delta^i \left(x_{m+d+1} - \sum_{l=0}^{d-j} A^l b \right) \right] \binom{-m-j}{i}, \tag{3.24}$$

where the summation on l is zero for $j = d + 1$. Now

$$\sum_{j=0}^{d+1} (-1)^{d+1-j} \binom{d+1}{j} \sum_{i=0}^{d} \left(\Delta^i x_{m+d+1} \right) \binom{-m-j}{i}$$

$$= \sum_{i=0}^{d} \left(\Delta^i x_{m+d+1} \right) \Delta^{d+1} \binom{-m}{i}, \tag{3.25}$$

and this vanishes, since

$$\Delta^{d+1} \binom{-m}{i} = 0 \qquad \text{for} \quad 0 \leqslant i \leqslant d.$$

As a result of this, and by the facts that $\Delta^i(\sum_{l=0}^{d-j} A^l b) = 0$ for $i \geqslant 1$ and $\sum_{l=0}^{d-j} A^l b = 0$ for $j = d + 1$, (3.24) becomes

$$Q = \sum_{j=0}^{d} (-1)^{d-j} \binom{d+1}{j} \sum_{l=0}^{d-j} A^l b$$

$$= \sum_{l=0}^{d} A^l b \sum_{j=0}^{d-l} (-1)^{d-j} \binom{d+1}{j}$$

$$= \sum_{l=0}^{d} (-1)^l \binom{d}{l} A^l b, \tag{3.26}$$

where we have used the identity

$$\sum_{j=0}^{q}(-1)^j\binom{d+1}{j} = (-1)^q\binom{d}{q}.$$

It is now easy to see that

$$Q = (I - A)^d b = \omega^{d+1}B^d f. \tag{3.27}$$

(3.21) now follows. ■

It is worth mentioning that the result given in (3.21) follows from (1.1)–(1.4) and (3.12), whether (1.1) is consistent or not, i.e., it is an identity.

3.2. *Connection with Semiiterative Methods*

Finally, we mention that, after drawing the proper analogy, and some tedious algebra, it can be shown that there exists a semiiterative method of the type discussed in [3] that gives the sequences $\{\hat{x}_m\}_{m=0}^{\infty}$. The paper [3] introduces a general theory for semiiterative methods that can be used for approximating the Drazin inverse solution. This theory provides a set of conditions that need to be satisfied by any semiiterative method, but does not specify the method uniquely. It is very interesting that the approach and techniques used in the present work, which are entirely different than those employed in [3], should produce a semiiterative technique of the type given in [3]. We add for the sake of completeness that, in the notation of [3], the polynomials $p_k(z)$ associated with the semiiterative method that produces the sequence $\{\hat{x}_m\}_{m=0}^{\infty}$ of the present work are given as

$$p_k(z) = z^{k-q}\sum_{i=0}^{q}\binom{-k+q}{i}(z-1)^i.$$

Thus, these $p_k(z)$, in addition to satisfying the necessary conditions of [3], namely, $p_k(1) = 1$, $p_k^{(j)}(1) = 0$, $1 \leqslant j \leqslant q$, satisfy also $p_k^{(j)}(0) = 0$, $0 \leqslant j \leqslant k - q - 1$, and are uniquely defined by them. Note that in our notation $q = d$ and $k = m + d$.

4. RICHARDSON'S ITERATIVE METHOD AND EXACT CONSTRUCTION OF s'

In the previous section we proposed a modification of Richardson's iterative method that is useful for the singular linear system in (1.1), whether this system is consistent or not. We showed that the sequence of vectors $\{\hat{x}_m\}_{m=0}^{\infty}$ obtained from this modification converges to the Drazin inverse solution s', the unique solution of the consistent system $Bx = f'$ that is in $\mathscr{S}(B)$, for some ω, under certain conditions concerning the spectrum of B. We also provided the precise rate of convergence of this sequence.

In the present section we show that s' can be constructed from a finite number of the vectors x_i obtained by Richardson's iterative method. We shall actually carry out this construction in all detail.

A concept that has been very useful in earlier work (see Sidi [8, 9] and Smith, Ford, and Sidi [12]) and will be of utmost importance in the present work is that of the *minimal polynomial of a matrix with respect to a vector*. We will call the polynomial $P(\lambda) = \sum_{i=0}^{k} c_i \lambda^i$, $c_k = 1$, the minimal polynomial of the matrix A with respect to the vector u if

$$P(A)u \equiv \left(\sum_{i=0}^{k} c_i A^i \right) u = 0$$

and

$$k = \min\left\{ p : \left(\sum_{i=0}^{p} \beta_i A^i \right) u = 0,\ \beta_p = 1 \right\}.$$

It is known that $P(\lambda)$ exists and is unique. It is also known that if $R(\lambda)$ is another polynomial for which $R(A)u = 0$, then $P(\lambda)$ divides $R(\lambda)$. Consequently, $P(\lambda)$ divides the minimal polynomial of A, which in turn divides the characteristic polynomial of A.

LEMMA 4.1. *Consider the singular linear system in* (1.1), *which may or may not be consistent. Let the sequence* $\{x_m\}_{m=0}^{\infty}$ *be generated by Richardson's iterative method as in* (1.2)–(1.4). *Denote by* $P(\lambda)$ *the minimal polynomial of* A *with respect to the vector* $e_n = A^n(x_0' - s')$, *with* x_0' *as in Theorem* 3.2, *and let* k_0 *be the degree of* $P(\lambda)$. *Then* $P(1) \neq 0$, *and*

$$k_0 \leqslant \sum_{\substack{i=1 \\ \mu_i \neq 0}}^{\nu} r_i = N - \sum_{\substack{i=1 \\ \mu_i = 0}}^{\nu} r_i \equiv \bar{N} \leqslant \operatorname{rank} B \leqslant N - 1. \qquad (4.1)$$

($\bar{N}$ = rank B *when* $\mathcal{M}(B)$ *is empty.) In addition,* $P(\lambda)$ *is also the minimal polynomial of* A *with respect to the vector* $\Delta^{\tau+1}x_n$, *where* τ *is the exact degree of the vector valued polynomial* $T(m)$ *in* (3.1); *hence* $\tau \leqslant d$. *Finally,* $P(\lambda)$ *can be constructed from the vectors* $x_n, x_{n+1}, \ldots, x_{n+k_0+\tau+1}$, *i.e., from at most* $\bar{N} + d + 2 \leqslant N + 2$ *vectors.*

Proof. Since $x_0' - s' \in \mathcal{S}(B)$, and $\mathcal{S}(B)$ is the subspace spanned by the eigenvectors and principal vectors $\hat{v}_{ij}$ of A belonging to eigenvalues not equal to unity, $P(\lambda)$ is not divisible by $\lambda - 1$; hence $P(1) \neq 0$, and (4.1) holds.

Letting $m = n$ in (3.1), and applying the operator $\Delta^{\tau+1}$ to both sides, we obtain

$$\Delta^{\tau+1}x_n = \Delta^{\tau+1}e_n, \tag{4.2}$$

since $T(m)$ is a polynomial in m of degree τ, and thus $\Delta^{\tau+1}T(m) = 0$. If $Q(\lambda)$ is the minimal polynomial of A with respect to $\Delta^{\tau+1}e_n$, and hence with respect to $\Delta^{\tau+1}x_n$, we must have

$$0 = Q(A)\,\Delta^{\tau+1}x_n = Q(A)\,\Delta^{\tau+1}e_n = Q(A)(A - I)^{\tau+1}e_n, \tag{4.3}$$

by $\Delta e_n = (A - I)e_n$. Therefore, $P(\lambda)$ divides $Q(\lambda)(\lambda - 1)^{\tau+1}$. However, $\lambda - 1$ is not a factor of $P(\lambda)$, so that $P(\lambda)$, in fact, divides $Q(\lambda)$. Similarly, $P(A)e_n = 0$ implies

$$\begin{aligned} 0 &= (A - I)^{\tau+1}P(A)e_n = P(A)(A - I)^{\tau+1}e_n \\ &= P(A)\,\Delta^{\tau+1}e_n = P(A)\,\Delta^{\tau+1}x_n, \end{aligned} \tag{4.4}$$

which, in turn, implies that $Q(\lambda)$ divides $P(\lambda)$. Consequently, $Q(\lambda) \equiv P(\lambda)$.

Let us write

$$P(\lambda) = \sum_{i=0}^{k_0} c_i\lambda^i, \qquad c_{k_0} = 1. \tag{4.5}$$

Then, from (4.2) and (4.4),

$$P(A)\,\Delta^{\tau+1}x_n = \sum_{i=0}^{k_0} c_i\,\Delta^{\tau+1}x_{n+i} = 0, \tag{4.6}$$

i.e,. the coefficients $c_0, c_1, \ldots, c_{k_0-1}$ of $P(\lambda)$ satisfy the set of overdetermined equations

$$\left[\Delta^{\tau+1}x_n \;\middle|\; \Delta^{\tau+1}x_{n+1} \;\middle|\; \cdots \;\middle|\; \Delta^{\tau+1}x_{n+k_0-1}\right]\begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{k_0-1} \end{bmatrix} = -\Delta^{\tau+1}x_{n+k_0}. \quad (4.7)$$

Since $P(\lambda)$ exists and is unique, (4.7) has a unique solution for c_i, $0 \leqslant i \leqslant k_0 - 1$. Thus, $P(\lambda)$ can actually be determined from the vectors $x_n, x_{n+1}, \ldots, x_{n+k_0+\tau+1}$ by solving (4.7). This completes the proof. ■

REMARK. Note that $P(\lambda)$, the minimal polynomial of A with respect to e_n, can be determined exactly as above if τ is replaced by any integer greater than τ throughout. In particular, we can replace τ by its maximum value possible, namely d. The reason for this is that $P(\lambda)$ is also the minimal polynomial of A with respect to $\Delta^q x_n$ for *any* integer $q \geqslant \tau + 1$, as $\Delta^q x_n = \Delta^q e_n$ for *all* $q \geqslant \tau + 1$; cf. (4.2). This is an important observation, and will be used later.

THEOREM 4.2. *Let the sequence* $\{x_m\}_{m=0}^{\infty}$, *the integers* τ, n, k_0, *and the polynomial* $P(\lambda)$ *be as in Lemma* 4.1. *With* $\tilde{x}_0$ *as in Theorem* 3.2, *the vector* $s' + \tilde{x}_0$ *can be constructed from the vectors* x_i, $n \leqslant i \leqslant n + k_0 + \tau + 1$, *i.e., from at most* $\bar{N} + d + 2 \leqslant N + 2$ *vectors, in the following way: Define*

$$S_m = \frac{\sum_{i=0}^{k_0} c_i x_{m+i}}{\sum_{i=0}^{k_0} c_i}, \qquad m = n, n+1, n+2, \ldots, \quad (4.8)$$

and let

$$\beta_q(m) = \frac{1}{P(1)}\frac{1}{q!}\frac{d^q}{d\lambda^q}\left[P(\lambda)\lambda^m\right]\bigg|_{\lambda=1} = \frac{\sum_{j=0}^{k_0} c_j \binom{m+j}{q}}{\sum_{j=0}^{k_0} c_j}, \qquad q = 0, 1, \ldots . \quad (4.9)$$

(Thus $\beta_0(m) = 1$.) Let $\tilde{\beta}_i(m)$ be defined by the Maclaurin expansion

$$\left(\sum_{i=0}^{\tau} \beta_i(m) z^i\right)^{-1} = \sum_{i=0}^{\infty} \tilde{\beta}_i(m) z^i. \tag{4.10}$$

(Thus $\tilde{\beta}_0(m) = 1$.) Then

$$s' + \tilde{x}_0 = S_m + \sum_{i=1}^{\tau}\left[\binom{-m}{i} - \sum_{q=1}^{i} \tilde{\beta}_{i-q}(m)\beta_q(0)\right]\Delta^i S_m,$$
$$m = n, n+1, \ldots . \tag{4.11}$$

As before, by letting $x_0 = 0$, or $x_0 = B^d\xi$, where ξ is an arbitrary vector, we can cause $\tilde{x}_0 = 0$ in (4.11).

Proof. From (3.1) and (4.5) we have

$$\sum_{j=0}^{k_0} c_j(x_{m+j} - s') = \sum_{j=0}^{k_0} c_j A^{m+j}(x_0' - s') + \sum_{j=0}^{k_0} c_j T(m+j). \tag{4.12}$$

The first summation on the right hand side of (4.12) is simply $P(A)e_m$. For $m = n, n+1, \ldots,$ we have $e_m = A^{m-n}e_n$. Thus $P(A)e_m = A^{m-n}P(A)e_n = 0$. Invoking this, $\sum_{j=0}^{k_0} c_j = P(1) \neq 0$ (which follows from Lemma 4.1), and (4.8) in (4.12), we obtain

$$S_m - s' = \frac{\sum_{j=0}^{k_0} c_j T(m+j)}{\sum_{j=0}^{k_0} c_j} \equiv \hat{T}(m), \qquad m = n, n+1, \ldots . \tag{4.13}$$

Obviously, $\hat{T}(m)$ is a polynomial in m of degree exactly τ, with vector coefficients in $\mathscr{T}(B)$. By Lemma 2.2,

$$\hat{T}(m) - \hat{T}(0) = -\sum_{i=1}^{\tau}\left[\Delta^i \hat{T}(m)\right]\binom{-m}{i}, \qquad \text{all } m. \tag{4.14}$$

But, by (4.13),

$$\Delta^i \hat{T}(m) = \Delta^i S_m, \qquad m = n, n+1, \ldots, \quad i = 1, 2, \ldots . \tag{4.15}$$

Combining (4.14) and (4.15) in (4.13), we have

$$s' = S_m + \sum_{i=1}^{\tau}(\Delta^i S_m)\binom{-m}{i} - \hat{T}(0), \qquad m = n, n+1, \ldots . \tag{4.16}$$

In order to determine $\hat{T}(0)$ we proceed as follows: Since $T(\eta)$ is a polynomial in η of degree τ with vector coefficients, it can be expressed in the form

$$T(\eta) = \sum_{i=0}^{\tau} a_i \binom{\eta}{i}. \tag{4.17}$$

Consequently, we can write

$$\hat{T}(m) = \frac{\sum_{j=0}^{k_0} c_j T(m+j)}{\sum_{j=0}^{k_0} c_j} = \sum_{i=0}^{\tau} \beta_i(m) a_i, \tag{4.18}$$

where we have invoked (4.9). Substituting (4.18) in (4.13), and applying Δ^q to both sides of (4.13), we have

$$\Delta^q \hat{T}(m) = \Delta^q S_m = \sum_{i=q}^{\tau} \beta_{i-q}(m) a_i, \qquad m = n, n+1, \ldots, \qquad q = 1, 2, \ldots, \tau. \tag{4.19}$$

Here we have made use of

$$\Delta^q \beta_i(m) = \beta_{i-q}(m), \qquad q = 0, 1, \ldots, \tag{4.20}$$

which can be obtained by employing

$$\Delta^q \binom{m}{i} = \binom{m}{i-q}, \qquad q = 0, 1, \ldots. \tag{4.21}$$

(4.21), in turn, can be obtained by induction from the identity

$$\Delta \binom{m}{i} = \binom{m+1}{i} - \binom{m}{i} = \binom{m}{i-1}. \tag{4.22}$$

Let us now write (4.19) in the matrix form

$$\begin{bmatrix} \beta_0(m) & \beta_1(m) & \beta_2(m) & \dots & \beta_{\tau-1}(m) \\ & \beta_0(m) & \beta_1(m) & \dots & \beta_{\tau-2}(m) \\ & & \beta_0(m) & \dots & \beta_{\tau-3}(m) \\ & \bigcirc & & & \vdots \\ & & & & \beta_0(m) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_\tau \end{bmatrix}$$

$$= \begin{bmatrix} \Delta S_m \\ \Delta^2 S_m \\ \Delta^3 S_m \\ \vdots \\ \Delta^\tau S_m \end{bmatrix}, \qquad m = n, n+1, \dots . \tag{4.23}$$

The matrix in (4.23) is a semicirculant matrix, and its inverse is the semicirculant matrix

$$\begin{bmatrix} \tilde\beta_0(m) & \tilde\beta_1(m) & \tilde\beta_2(m) & \dots & \tilde\beta_{\tau-1}(m) \\ & \tilde\beta_0(m) & \tilde\beta_1(m) & \dots & \tilde\beta_{\tau-2}(m) \\ & & \tilde\beta_0(m) & \dots & \tilde\beta_{\tau-3}(m) \\ & \bigcirc & & & \vdots \\ & & & & \tilde\beta_0(m) \end{bmatrix},$$

with $\tilde\beta_i(m)$ as defined in (4.10); see, e.g, [5, pp. 14–16]. Consequently,

$$a_q = \sum_{i=q}^{\tau} \tilde\beta_{i-q}(m)\,\Delta^i S_m, \qquad m = n, n+1, \dots . \tag{4.24}$$

Substituting now (4.24) in (4.18), letting $m = 0$ there, and using the facts that $\beta_0(m) = 1$ and $T(0) = a_0 = \tilde{x}_0$, we obtain

$$\hat{T}(0) = \tilde{x}_0 + \sum_{i=1}^{\tau}\left(\sum_{q=1}^{i} \tilde\beta_{i-q}(m)\beta_q(0)\right)\Delta^i S_m, \qquad m = n, n+1, \dots . \tag{4.25}$$

(4.11) can now be obtained by substituting (4.25) in (4.16).

Setting $m = n$ in (4.11), we see that if the scalars c_i are known, then $s' + \tilde{x}_0$ is determined from the vectors x_i, $n \leqslant i \leqslant n + k_0 + \tau$. The c_i, on the other hand, are determined from the vectors x_i, $n \leqslant i \leqslant n + k_0 + \tau + 1$, as shown in Lemma 4.1. Consequently, only the vectors x_i, $n \leqslant i \leqslant n + k_0 + \tau + 1$, are needed for constructing $s' + \tilde{x}_0$.

The rest of the proof is trivial, and we shall omit it. ■

We note that the $\tilde{\beta}_i(m)$ that are defined by (4.10) can be computed recursively from

$$\begin{aligned} \tilde{\beta}_0(m) &= 1, \\ \tilde{\beta}_i(m) &= -\sum_{q=1}^{\tau} \tilde{\beta}_{i-q}(m)\beta_q(m), \qquad i = 1,2,\ldots,\tau. \end{aligned} \tag{4.26}$$

We also note that the remark following the proof of Lemma 4.1 applies in Theorem 4.2 as well. That is to say, the integer τ can be replaced by d throughout. [This would mean, in particular, that $a_i = 0$ for $\tau < i \leqslant d$ in(4.18), (4.19), and (4.24).] This observation will be of use later.

The construction above assumes an especially simple form in case $\mathcal{M}(B)$ is the empty set. In this case $d = 1$. Consequently, if the system in (1.1) is inconsistent, we have $\tau = 1$. Using this, (4.11) becomes

$$S_n - \left(n + \frac{P'(1)}{P(1)} \right) \Delta S_n = s' + \tilde{x}_0. \tag{4.27}$$

We restate this result separately for the case in which B is a range-Hermitian matrix.

THEOREM 4.3. *Consider the singular and not necessarily consistent system in* (1.1), *and let the matrix B there be range-Hermitian. Let $\{x_m\}_{m=0}^{\infty}$ be the sequence generated by Richardson's iterative method* (1.2)–(1.4), *with $x_0 = 0$, or $x_0 = B\xi$, ξ being chosen arbitrarily. Let $P(\lambda) = \sum_{i=1}^{k_0} c_i \lambda^i$, $c_{k_0} = 1$, be the minimal polynomial of A with respect to the vector $\Delta^2 x_n$, and define S_m as in* (4.8). *Then*

$$S_n - \left(n + \frac{P'(1)}{P(1)} \right) \Delta S_n = B^+ f, \tag{4.28}$$

and the vectors that are used in the construction of $B^+ f$ are $x_n, x_{n+1}, \ldots, x_{n+k_0+2}$.

In the terminology of Krylov subspace methods Theorem 4.3 implies that, starting with $x_0 = 0$ or $x_0 = B\xi$, $B^+ f$ is obtained in exactly $k_0 + 2$ steps of Richardson's iterative method, which, by (4.1), implies at most $N + 1$ of these iterations.

5. DEVELOPMENT OF NEW EXTRAPOLATION METHODS FOR s'

In the previous section we showed how s' can be constructed from an appropriate number of the vectors $\{x_m\}_{m=0}^{\infty}$ that are generated by Richardson's iterative method. Based on this development, in the present section we propose a class of extrapolation methods that produce approximations to s'. These approximations will be shown to have better convergence properties than the sequence $\{\hat{x}_m\}_{m=0}^{\infty}$. We note that the input needed for the extrapolation methods of this section is the sequence $\{x_j\}_{j=0}^{\infty}$ obtained as in (1.2) and the integer $d = \max\{r_i : \mu_i = 0\}$, namely, the index of B, or any integer greater than d when d is not known precisely. (When the matrix B is diagonalizable, we know immediately that $d = 1$.)

5.1. *The General Extrapolation Method*

Pick the integers $n \geqslant 0$ and $k \geqslant 1$.

Step 1.

(i) Determine the scalars γ_j, $0 \leqslant j \leqslant k$, by "solving" the overdetermined and, in general, inconsistent system of linear equations

$$\sum_{j=0}^{k} \gamma_j \Delta^{d+1} x_{m+j} = 0, \qquad n \leqslant m \leqslant n+k,$$
$$\text{subject to} \quad \sum_{j=0}^{k} \gamma_j = 1. \tag{5.1}$$

Various "solution" methods will be explained below. More are possible, however.

(ii) With γ_j determined as above, compute the scalars $\beta_q(n)$ and $\beta_q(0)$ from

$$\beta_q(m) = \sum_{j=0}^{k} \gamma_j \binom{m+j}{q}. \tag{5.2}$$

[Thus, $\beta_0(m) = 1$.] Next, compute the scalars $\tilde{\beta}_i(n)$ from

$$\tilde{\beta}_0(n) = 1,$$
$$\tilde{\beta}_i(n) = -\sum_{q=1}^{i} \tilde{\beta}_{i-q}(n)\beta_q(n), \qquad i = 1,2,\ldots,d-1. \tag{5.3}$$

Step 2. Compute the vectors S_m, $n \leqslant m \leqslant n + d$, by

$$S_m = \sum_{j=0}^{k} \gamma_j x_{m+j}. \tag{5.4}$$

Step 3. Compute the vector $Z_{n,k}$, the approximation to $s' + \tilde{x}_0$, from

$$Z_{n,k} = S_n + \sum_{i=1}^{d} \left[\binom{-n}{i} - \sum_{q=1}^{i} \tilde{\beta}_{i-q}(n) \beta_q(0) \right] \Delta^i S_n. \tag{5.5}$$

(As mentioned before, by picking $x_0 = 0$ or $x_0 = B^d \xi$, where ξ is an arbitrary vector, we can cause $\tilde{x}_0 = 0$, so that $Z_{n,k}$ becomes an approximation to s' only.)

For the case in which $d = 1$ the extrapolation method above takes on an especially simple form: $Z_{n,k}$ is now given by

$$Z_{n,k} = S_n - \left(n + \sum_{j=0}^{k} j\gamma_j \right) \Delta S_n. \tag{5.6}$$

5.2. *Methods for the Solution of (5.1)*

Method 1. Determine the scalars $c_0, c_1, \ldots, c_{k-1}$ as the solution to the problem

$$\left\| \sum_{j=0}^{k-1} c_j \, \Delta^{d+1} x_{n+j} + \Delta^{d+1} x_{n+k} \right\| = \min, \tag{5.7}$$

where $\|z\|$ is an arbitrary norm on $\mathbf{C}^N$. Following this, set $c_k = 1$, and compute the γ_j from

$$\gamma_j = \frac{c_j}{\sum_{i=0}^{k} c_i}, \qquad j = 0, 1, \ldots, k. \tag{5.8}$$

Method 2. Determine $\gamma_0, \gamma_1, \ldots, \gamma_k$ as the solution to the problem

$$\left\| \sum_{j=0}^{k} \gamma_j \, \Delta^{d+1} x_{n+j} \right\| = \min$$

$$\text{subject to} \quad \sum_{j=0}^{k} \gamma_j = 1, \tag{5.9}$$

where $\| z \|$ is an arbitrary norm on $\mathbf{C}^N$. By defining the scalars $\delta_0, \delta_1, \ldots, \delta_{k-1}$ by

$$1 - \delta_0 = \gamma_0; \qquad \delta_{i-1} - \delta_i = \gamma_i, \quad 1 \leqslant i \leqslant k-2; \qquad \delta_{k-1} = \gamma_k, \tag{5.10}$$

we can rewrite the minimization problem of (5.9) in the equivalent form

$$\left\| \Delta^{d+1} x_n + \sum_{i=0}^{k-1} \delta_i \, \Delta^{d+2} x_{n+i} \right\| = \min. \tag{5.11}$$

In methods 1 and 2 above, (weighted) l_p norms can be employed. In particular, if the l_1 or l_∞ norm is used in (5.7) and (5.11), then the solution can be obtained by linear programming techniques. If the l_2 norm is used, then the solution can be achieved by appropriate least squares techniques or even from the normal equations.

Method 3. Pick k linearly independent vectors $g_1, \ldots, g_k$, and determine the scalars $c_0, c_1, \ldots, c_{k-1}$ as the solution to the k linear equations

$$\sum_{j=0}^{k-1} \left(g_i, \Delta^{d+1} x_{n+j} \right) c_j = -\left(g_i, \Delta^{d+1} x_{n+k} \right), \qquad 1 \leqslant i \leqslant k. \tag{5.12}$$

Following this, set $c_k = 1$, and compute the γ_j from (5.8).

Method 4. Pick a nonzero vector g, and determine the scalars $c_0, c_1, \ldots, c_{k-1}$ as the solution to the k linear equations

$$\sum_{j=0}^{k-1} \left(g, \Delta^{d+1} x_{n+i+j} \right) = -\left(g, \Delta^{d+1} x_{n+i+k} \right), \qquad 0 \leqslant i \leqslant k-1. \tag{5.13}$$

Following this, set $c_k = 1$, and compute the γ_j from (5.8).

In methods 3 and 4 above, the inner product (y, z) is general. Also, as mentioned in [8, 10, 11], if the vectors x_i are in an arbitary normed linear space, then the inner products $(g_i, \Delta^{d+1} x_{n+j})$ in (5.12) and $(g, \Delta^{d+1} x_{n+i+j})$ in (5.13) can be replaced by $Q_i(\Delta^{d+1} x_{n+j})$ and $Q(\Delta^{d+1} x_{n+i+j})$ respectively, where Q_i and Q are bounded linear functionals on the space, and $Q_1, \ldots, Q_k$ are also linearly independent.

Above we have followed closely the developments of [11, Section 2.2]. We note that methods 1 and 2 with the most general l_2 norm are analogous to those that produce MPE and RRE respectively. Similarly, methods 3 and 4 are analogous to those that produce MMPE and TEA respectively. For details and references pertaining to these methods, see, for example, [7, 8, 10, 11].

Now the fact that in all of methods 1–4 the γ_j can be determined as solutions to a system of $k+1$ linear equations suggests that it might be possible to express $Z_{n,k}$ in determinantal form and that, as a result, there might exist recursion relations amongst the variuos $Z_{n,k}$, similar to those derived in Ford and Sidi [4]. It is proposed to tackle this problem in a future publication.

It is important to emphasize that the scalars γ_j, $\beta_i(m)$, and $\tilde{\beta}_i(m)$, as well as the vectors S_m, that are needed in the general extrapolation algorithm described above all depend on n and k. The analysis of their precise nature as functions of n and k is the subject of the next sections.

5.3. *Connection with Lemma 4.1 and Theorem 4.2*

We note that the steps of the general extrapolation method that lead to $Z_{n,k}$ are very similar to those of Lemma 4.1 and Theorem 4.2 that produce $s' + \tilde{x}_0$. It is worthwhile analyzing their differences closely, as this will be very helpful later.

(1) In Lemma 4.1 and Theorem 4.2, k_0 is the exact degree of the minimal polynomial of A with respect to $e_n = A^n(x_0' - s')$ or, equivalently, with respect to $\Delta^{\tau+1}x_n$, where τ is the exact degree of the vector valued polynomial $T(m)$ in (3.1). In the general extrapolation method k is an arbitrary positive integer that is at our disposal, and the integer τ has been replaced by d. This has been done because τ depends on f, the right hand side of (1.1), and on x_0, which is picked arbitrarily, but d depends only on the matrix B, and thus is fixed. Furthermore, $\tau \leqslant d$ always. If we require the exact degree of $T(m)$ in our method, then we cannot have a practical method of extrapolation.

(2) The γ_j in the general extrapolation method are analogous to $c_j/\sum_{i=0}^{k_0} c_i$ in Lemma 4.1 and Theorem 4.2. We recall that the equations that the c_j of Lemma 4.1 and Theorem 4.2 satisfy, namely,

$$\sum_{j=0}^{k_0-1} c_j \Delta^{\tau+1} x_{n+j} = -\Delta^{\tau+1} x_{n+k} \tag{5.14}$$

[cf. (4.6)], are consistent and have a unique solution, despite the fact that they are overdetermined. With $k < k_0$, however, the equations (5.1) that are "solved" in order to determine the γ_j are inconsistent for each value of m, $n \leqslant m \leqslant n + k$. In particular, the equations

$$\sum_{j=0}^{k-1} c_j \Delta^{d+1} x_{n+j} = -\Delta^{d+1} x_{n+k}, \tag{5.15}$$

obtained by setting $m = n$ in (5.1), which are "solved" in methods 1 and 3 [cf. (5.7) and (5.12)], are inconsistent, and thus have no solution in the usual sense. This assertion can be proved by contradiction. For, if we assume that there exist $c_0, c_1, \dots, c_{k-1}$ satisfying (5.15), then, by $\Delta^{q+1} x_{n+i} = A^i \Delta^{q+1} x_n$ for $q \geqslant \tau$, we have $(\sum_{j=0}^{k} c_j A^j)\, \Delta^{d+1} x_n = 0$, with $c_k = 1$. But this implies that the polynomial $\sum_{i=0}^{k} c_i \lambda^i$ annihilates $\Delta^{d+1} x_n$, and yet has a smaller degree than the minimal polynomial with respect to $\Delta^{d+1} x_n$, which is impossible.

(3) $\hat{T}(0)$ in Theorem 4.2 is a truly constant vector, i.e., the sum

$$\sum_{i=1}^{\tau} \left(\sum_{q=1}^{i} \tilde{\beta}_{i-q}(m) \beta_q(0) \right) \Delta^i S_m, \qquad m = n, n+1, \dots,$$

in (4.25) is independent of m. The analogous sum

$$\sum_{i=1}^{d} \left(\sum_{q=1}^{i} \tilde{\beta}_{i-q}(m) \beta_q(0) \right) \Delta^i S_m$$

that appears in (5.5) with $m = n$ does not have this property when $k < k_0$.

(4) If $k = k_0$ in the general extrapolation method, then $Z_{n,k_0} = s' + \tilde{x}_0$

(a) always if the γ_j are obtained using method 1 or method 2;
(b) if the γ_j are uniquely determined using method 3 or method 4.

We leave the verification of these assertions to the reader.

6. PRELIMINARY RESULTS FOR THE CONVERGENCE ANALYSIS OF $Z_{n,k}$

In this section we derive a closed form expression for the error $Z_{n,k} - (s' + \tilde{x}_0)$ in terms of the quantities that are computed for the general extrapolation method. This expression turns out to be very suitable for investigating the behavior of the error for $n \to \infty$ with k fixed. The error for fixed n and increasing k will be the subject of a future publication.

THEOREM 6.1. *Denote*

$$E_m = \sum_{j=0}^{k} \gamma_j e_{m+j}, \qquad all\ m. \tag{6.1}$$

Then

$$Z_{n,k} - (s' + \tilde{x}_0) = E_n + \sum_{i=1}^{d} \left[\binom{-n}{i} - \sum_{q=1}^{i} \tilde{\beta}_{i-q}(n)\beta_q(0) \right] \Delta^i E_n. \tag{6.2}$$

Proof. Replacing m in (3.1) by $m + j$, multiplying both sides of (3.1) by γ_j, and summing over j, we obtain

$$S_m - E_m - s' = \sum_{j=0}^{k} \gamma_j T(m+j) \equiv \hat{T}(m), \qquad \text{all } m. \tag{6.3}$$

Since $T(\eta)$ is a polynomial in η of degree at most d, it can be written in the form

$$T(\eta) = \sum_{i=0}^{d} a_i \binom{\eta}{i}, \qquad T(0) = a_0 = \tilde{x}_0. \tag{6.4}$$

Consequently,

$$\hat{T}(m) = \sum_{i=0}^{d} \beta_i(m) a_i. \tag{6.5}$$

Applying now the technique employed in the proof of Theorem 4.2, we obtain

$$\hat{T}(m) - \hat{T}(0) = -\sum_{i=1}^{d} \binom{-m}{i} \Delta^i \hat{T}(m), \qquad \text{all } m, \tag{6.6}$$

with

$$\hat{T}(0) = \tilde{x}_0 + \sum_{i=1}^{d} \left(\sum_{q=1}^{i} \tilde{\beta}_{i-q}(m)\beta_q(0) \right) \Delta^i \hat{T}(m), \qquad \text{all } m. \tag{6.7}$$

Combining (6.6) and (6.7) in (6.3), and invoking

$$\Delta^i \hat{T}(m) = \Delta^i (S_m - E_m), \qquad \text{all } m, \quad i \geq 1, \tag{6.8}$$

which follows from (6.3), we obtain

$$s' + \tilde{x}_0 = S_m - E_m + \sum_{i=1}^{d}\left[\binom{-m}{i} - \sum_{q=1}^{i}\tilde{\beta}_{i-q}(m)\beta_q(0)\right]\Delta^i(S_m - E_m),$$

$$\text{all } m. \tag{6.9}$$

Subtracting (6.9) with $m = n$ from (5.5), we obtain (6.2). ∎

The following result is a modification of Theorem 3.2 of [9] that is suitable for our analysis.

THEOREM 6.2. *For m sufficiently large,* x_m *satisfies*

$$x_m - (s' + \tilde{x}_0) - T(m) = e_m = \sum_{i=1}^{\bar{\nu}} P_i(m)\sigma_i^m, \qquad \textit{some} \quad \bar{\nu} \leqslant \nu. \tag{6.10}$$

Here σ_i *are complex numbers satisfying*

$$\sigma_i \neq 0, \quad \sigma_i \neq 1; \qquad \sigma_i \neq \sigma_j \quad \textit{if } i \neq j, \tag{6.11}$$

and are ordered so that

$$|\sigma_1| \geqslant |\sigma_2| \geqslant |\sigma_3| \geqslant \cdots. \tag{6.12}$$

$P_i(m)$ *are polynomials in m with vector coefficients, given in the form*

$$P_i(m) = \sum_{j=0}^{p_i} y_{ij}\binom{m}{j}, \qquad \textit{some integer} \quad p_i \geqslant 0, \tag{6.13}$$

where the vectors y_{ij}, $j = 0, 1, \ldots, p_i$, $i = 1, 2, \ldots, \bar{\nu}$, *form a linearly independent set. (Specifically, the* σ_i *are distinct nonzero eigenvalues of A corresponding to the nonzero eigenvalues* μ_q *of B, and* y_{ij} *are some linearly independent combinations of the eigenvectors and principal vectors corresponding to those eigenvalues* $\lambda_q = 1 - \omega\mu_q$ *of A that are all equal to* σ_i*. As a result,* $p_i + 1 \leqslant \max\{r_q : \lambda_q = 1 - \omega\mu_q = \sigma_i\}$*. Also, m sufficiently large means* $m \geqslant \max\{r_q : 1 - \omega\mu_q = 0\}$*.)*

The proof of Theorem 6.2 can be achieved by recalling that $e_m = A^m(x_0' - s')$, and that $x_0' - s' \in \mathcal{S}(B)$, and following the discussion of [10, Section 2]. Note that $p_i = 0$ for all i in case B is a diagonalizable matrix.

COROLLARY 6.3. *For m sufficiently large, and $q \geqslant d+1$, $\Delta^q x_m$ has the expansion*

$$\Delta^q x_m = \sum_{i=1}^{\bar{\nu}} \hat{P}_i(m)\sigma_i^m, \tag{6.14}$$

where $\hat{P}_i(m)$ are polynomials with vector coefficients, given in the form

$$\hat{P}_i(m) = \sum_{j=0}^{p_i} \hat{y}_{ij}\binom{m}{j}. \tag{6.15}$$

Here, for each pair of integers i and j, $\hat{y}_{ij}$ is a linear combination of y_{il}, $0 \leqslant l \leqslant p_i$, and $\hat{y}_{ip_i} = (\sigma_i - 1)^q y_{ip_i} \neq 0$, so that $\hat{P}_i(m)$ is of degree p_i exactly. In addition, the vectors $\hat{y}_{ij}$, $j = 0, 1, \ldots, p_i$, $i = 1, \ldots, \bar{\nu}$, form a linearly independent set.

Proof. Applying Δ^q with $q \geqslant \mathrm{d}+1$ to both sides of (6.10), we obtain

$$\Delta^q x_m = \Delta^q e_m = \sum_{i=1}^{\bar{\nu}} \Delta^q\left[P_i(m)\sigma_i^m\right]. \tag{6.16}$$

The result follows by actually evaluating $\Delta^q[P_i(m)\sigma_i^m]$, and recalling that $\sigma_i \neq 1$. The details are left to the reader. ■

In the remainder of this work we shall analyze four special cases of the general extrapolation method. We recall that the only difference between the various extrapolation methods stems from the procedures employed in the determination of the γ_i. In view of this the four methods are those in which the γ_i are determined from

(1) method 1 with the general l_2 norm,
(2) method 2 with the general l_2 norm,
(3) method 3 with the general inner product,
(4) method 4 with the general inner product.

For these methods the γ_i are determined from the linear systems of equations

$$\begin{aligned} &\sum_{j=0}^{k} u_{ij}\gamma_j = 0, \qquad i = 0, 1, \ldots, k-1, \\ &\sum_{j=0}^{k} \gamma_j = 1, \end{aligned} \tag{6.17}$$

where

$$u_{ij} = \left(\Delta^{d+1}x_{n+i}, \Delta^{d+1}x_{n+j}\right) \quad \text{for method 1,}$$
$$u_{ij} = \left(\Delta^{d+2}x_{n+i}, \Delta^{d+1}x_{n+j}\right) \quad \text{for method 2,}$$
$$u_{ij} = \left(g_{i+1}, \Delta^{d+1}x_{n+j}\right) \quad \text{for method 3,} \tag{6.18}$$
$$u_{ij} = \left(g, \Delta^{d+1}x_{n+i+j}\right) \quad \text{for method 4.}$$

For methods 1 and 2 the equations in (6.17) are obtained from the normal equations associated with the least squares problems in (5.7) and (5.11) respectively.

In view of (6.17) we have

$$\sum_{i=0}^{k} \gamma_i \lambda^i = \frac{D(1, \lambda, \ldots, \lambda^k)}{D(1, 1, \ldots, 1)}, \tag{6.19}$$

where

$$D(y_0, y_1, \ldots, y_k) = \begin{vmatrix} y_0 & y_1 & \cdots & y_k \\ u_{00} & u_{01} & \cdots & u_{0k} \\ u_{10} & u_{11} & \cdots & u_{1k} \\ \vdots & \vdots & & \vdots \\ u_{k-1,0} & u_{k-1,1} & \cdots & u_{k-1,k} \end{vmatrix}, \tag{6.20}$$

i.e., γ_i are the coefficients of the polynomial $D(1, \lambda, \ldots, \lambda^k)/D(1, 1, \ldots, 1)$.

Similarly, in view of (6.1) we have

$$E_m = \frac{D(e_m, e_{m+1}, \ldots, e_{m+k})}{D(1, 1, \ldots, 1)}. \tag{6.21}$$

Note that the first row of the determinant $D(e_m, \ldots, e_{m+k})$ in (6.21) is composed of vectors. When $y_0, y_1, \ldots, y_k$ in $D(y_0, \ldots, y_k)$ are vectors, the latter is interpreted to be its expansion with respect to its first row. Consequently, $D(y_0, \ldots, y_k)$ becomes a vector in this case.

All of these developments are based on [7, 11].

7. CONVERGENCE OF $Z_{n,k}$ FOR k FIXED AND $n \to \infty$

By Theorem 6.1 concerning the error $Z_{n,k} - (s' + \tilde{x}_0)$, it is obvious that analyzing E_n and the $\beta_i(0)$ and $\tilde{\beta}_i(n)$, for k fixed and $n \to \infty$, is what needs to be done.

We recall that we are dealing with the four special cases of the general extrapolation method that are described following the proof of Corollary 6.3. For all four methods, we assume that the distinct eigenvalues σ_i that are present in (6.10) satisfy

$$|\sigma_1| \geqslant |\sigma_2| \geqslant \cdots \geqslant |\sigma_t| > |\sigma_{t+1}| \geqslant \cdots \tag{7.1}$$

for some integer t, and that

$$k = \sum_{j=1}^{t} (p_j + 1). \tag{7.2}$$

We furthermore assume that

$$\begin{vmatrix} (g_1, y_{10}) & \cdots & (g_1, y_{1p_1}) & (g_1, y_{20}) & \cdots & (g_1, y_{2p_2}) & \cdots & (g_1, y_{t0}) & \cdots & (g_1, y_{tp_t}) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ (g_k, y_{10}) & \cdots & (g_k, y_{1p_1}) & (g_k, y_{20}) & \cdots & (g_k, y_{2p_2}) & \cdots & (g_k, y_{t0}) & \cdots & (g_k, y_{tp_t}) \end{vmatrix} \neq 0 \tag{7.3}$$

for method 3, and

$$\prod_{j=1}^{t} (g, y_{jp_j}) \neq 0 \tag{7.4}$$

for method 4. No additional assumptions are needed for methods 1 and 2.

THEOREM 7.1. *Under the conditions stated above,*

$$E_n = O(n^{\bar{p}} |\sigma_{t+1}|^n) \qquad as \quad n \to \infty, \tag{7.5}$$

where $\bar{p}$ is a nonnegative integer given by

$$\bar{p} = \max\{p_i : |\sigma_i| = |\sigma_{t+1}|\}, \tag{7.6}$$

and also, for all λ,

$$\sum_{i=0}^{k} \gamma_i \lambda^i = \prod_{i=1}^{t} \left(\frac{\lambda - \sigma_i}{1 - \sigma_i} \right)^{p_i + 1} + O\left(n^{\alpha} \left| \frac{\sigma_{t+1}}{\sigma_t} \right|^n \right) \quad as \quad n \to \infty, \tag{7.7}$$

where α is some nonnegative integer.

Proof. For the proof of (7.5) we substitute (6.10) and (6.14) and (6.15) in (6.25), and employ the technique that was used in the proof of Theorem 3.1 of [10]. Similarly, for the proof of (7.7) we substitute (6.14) and (6.15) in(6.19), and employ the technique that was used in the proof of Theorem 3.2 of [10]. We shall omit the details. ∎

THEOREM 7.2. *Let us define the scalars δ_i by*

$$\prod_{i=1}^{t} \left(\frac{\lambda - \sigma_i}{1 - \sigma_i} \right)^{p_i + 1} \equiv \sum_{i=0}^{k} \delta_i \lambda^i. \tag{7.8}$$

Then, for $0 \leqslant q \leqslant d$,

$$\beta_q(0) = \sum_{j=0}^{k} \delta_j \binom{j}{q} + O\left(n^{\alpha} \left| \frac{\sigma_{t+1}}{\sigma_t} \right|^n \right) \quad as \quad n \to \infty \tag{7.9}$$

$$= O(1) \quad as \quad n \to \infty,$$

$$\beta_q(n) = \sum_{j=0}^{k} \delta_j \binom{n+j}{q} + O\left(n^{\alpha+q} \left| \frac{\sigma_{t+1}}{\sigma_t} \right|^n \right) \quad as \quad n \to \infty \tag{7.10}$$

$$= O(n^q) \quad as \quad n \to \infty,$$

and, for $0 \leqslant q \leqslant d - 1$,

$$\tilde{\beta}_q(n) = O(n^q) \quad as \quad n \to \infty. \tag{7.11}$$

Proof. From (7.7) in Theorem 7.1 it follows that

$$\gamma_j = \delta_j + O\left(n^{\alpha}\left|\frac{\sigma_{t+1}}{\sigma_t}\right|^n\right) \qquad as \quad n \to \infty. \tag{7.12}$$

Combining (5.2) and (7.12), we obtain (7.9) and (7.10). Equation (7.11) is obtained by combining (5.3) and (7.10), and by using induction. ■

Finally, Theorem 7.3 below is our main convergence result pertaining to $Z_{n,k}$ for $n \to \infty$.

THEOREM 7.3. *The error in* $Z_{n,k}$ *satisfies*

$$Z_{n,k} - \left(s' + \tilde{x}_0\right) = O\left(n^{\bar{p}+d} |\sigma_{t+1}|^n\right) \qquad as \quad n \to \infty. \tag{7.13}$$

Proof. The proof of (7.13) can be achieved by combining (7.5), (7.9), and (7.11) in (6.2), and noting that $\Delta^i E_n = O(n^{\bar{p}} |\sigma_{t+1}|^n)$ as $n \to \infty$, for $i \geqslant 1$ as well as $i = 0$, because $\sigma_j \neq 1$ for all j. We leave out the details. ■

The author gratefully acknowledges the very helpful conversations that he has had with Professor Moshe Israeli during the preparation of this paper. The author would also like to thank Professors Gene H. Golub, Ivo Marek, and Wilhelm Niethammer and Dr. Martin Hanke for drawing his attention to the literature on the Drazin inverse and semiiterative methods for singular systems. This research was supported by the Fund for the Promotion of Research at the Technion.

REFERENCES

1 A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, Wiley, New York, 1974.

2 S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations*, Pitman, London, 1979.

3 M. Eiermann, I. Marek, and W. Niethammer, On the solution of singular linear systems of algebraic equations by semi-iterative methods, *Numer. Math.*, 53:265–283 (1988).

4 W. F. Ford and A. Sidi, Recursive algorithms for vector extrapolation methods, *Appl. Numer. Math.* 4:477–489 (1988).

5 P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, Wiley, New York, 1974.
6 C. D. Meyer, Jr. and R. J. Plemmons, Convergent powers of a matrix with applications to iterative methods for singular linear systems, *SIAM J. Numer. Anal.* 14:699–705 (1977).
7 A. Sidi, Convergence and stability properties of minimal polynomial and reduced rank extrapolation algorithms, *SIAM J. Numer. Anal.*, 23:197–209 (1986); originally appeared as NASA TM-83443, July 1983.
8 A. Sidi, Extrapolation vs. projection methods for linear systems of equations, *J. Comput. Appl. Math.* 22:71–88 (1988).
9 A. Sidi, Application of vector extrapolation methods to consistent singular linear systems, *Appl. Numer. Math.* 6:487–500 (1989/90).
10 A. Sidi and J. Bridger, Convergence and stability analyses for some vector extrapolation methods in the presence of defective iteration matrices, *J. Comput. Appl. Math.* 22:35–61 (1988).
11 A. Sidi, W. F. Ford, and D. A. Smith, Acceleration of convergence of vector sequences, *SIAM J. Numer. Anal.* 23:178–196 (1986); originally appeared as NASA TP-2193, Dec. 1983.
12 D. A. Smith, W. F. Ford, and A. Sidi, Extrapolation methods for vector sequences, *SIAM Rev.*, 29:199–233 (1987); Erratum: Correction to "Extrapolation methods for vector sequences," *SIAM Rev.* 30:623–624 (1988).
13 R. S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, New York, 1962.

Received 16 September 1991; final manuscript accepted 26 November 1991