Numer. Math. (2003) 93: 563–581
Digital Object Identifier (DOI) 10.1007/s002110100379

Numerische Mathematik

Orthogonal polynomials and semi-iterative methods for the Drazin-inverse solution of singular linear systems

Avram Sidi[1], **Yuliya Kanevsky**[2]

[1] Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel; e-mail: asidi@cs.technion.ac.il
[2] Mathematics Department, Technion – Israel Institute of Technology, Haifa 32000, Israel

Received June 28, 2000 / Revised version received May 23, 2001 /
Published online January 30, 2002 –

Summary. In this work we present a novel class of semi-iterative methods for the Drazin-inverse solution of singular linear systems, whether consistent or inconsistent. The matrices of these systems are allowed to have arbitrary index and arbitrary spectra in the complex plane. The methods we develop are based on orthogonal polynomials and can all be implemented by 4-term recursion relations independently of the index. We give all the computational details of the associated algorithms. We also give a complete convergence analysis for all methods.

Mathematics Subject Classification (1991): 15A06, 15A09, 30E10, 42C10, 65F10, 65F50

1 Introduction

Consider the linear system

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{C}^{N\times N}$ is singular and $\mathrm{ind}(A)$, the index of A, is arbitrary. We recall that $\mathrm{ind}(A)$ is the size of the largest Jordan block of A corresponding to its zero eigenvalue. Assume that the nonzero eigenvalues λ_j of A are in general complex and lie in an open half-plane of the complex λ-plane whose boundary is a straight line through the origin. In addition, allow (1.1) to be consistent or inconsistent.

Our purpose here is to develop a class of semi-iterative methods for the Drazin-inverse solution of (1.1), namely, the vector $A^{\mathrm{D}}b$, where A^{D} is the

Drazin inverse of the singular matrix A. In relation to the assumption that (1.1) can be consistent or inconsistent, we recall that, when (1.1) is consistent it has an infinity of solutions in the regular sense, but if $\text{ind}(A) > 1$ in addition, $A^{\text{D}}b$ may not always be a solution to (1.1) in the regular sense. For the Drazin inverse and its properties see, e.g., Ben-Israel and Greville [BG] or Campbell and Meyer [CM].

The problem of determining the Drazin-inverse solution of (1.1) by semi-iterative methods has drawn some attention recently. A general theory that gives necessary and sufficient conditions for convergence of semi-iterative methods has been given in Eiermann, Marek, and Niethammer [EMN]. Following [EMN], Hanke and Hochbruck [HH] gave a Chebyshev-like method for the special case in which $\text{ind}(A) = 1$ and the nonzero eigenvalues of A are real and positive and known to lie in the interval $[c - f, c + f]$ with $0 < f < c$. The approach of [HH] was extended by Climent, Neumann, and Sidi [CNS] to the case in which $\text{ind}(A) = a$, a being arbitrary, and, again, the nonzero eigenvalues of A are known to be real and positive and in $[c - f, c + f]$ with $0 < f < c$. Both [HH] and [CNS] contain a thorough convergence analysis for one of the methods they develop and that is related to Chebyshev polynomials. Both papers have been the main source of inspiration for the present work. Finally, we would like to mention the semi-iterative methods of Woźniakowski [W] and of Eiermann and Reichel [ER] that were developed for the case in which $\text{ind}(A) = 1$. As noted in [HH], both methods are based on the original Chebyshev acceleration.

The treatment of the case in which the spectrum of A is complex has been considered only recently in the M.Sc. thesis of Kanevsky [K]. The method of [K] is based largely on [CNS], but turns out to be too involved and expensive computationally. The approach we take to the general problem in this work is substantially different from those of [HH], [CNS], and [K] even though there are some similarities between them. In addition, the algorithms we derive here are entirely different from those of [HH], [CNS], and [K]. As far as computational costs are concerned, the methods in the present work have the same cost as those of [HH] and [CNS] but are much less expensive than that of [K].

The plan of this paper is as follows. In Sect. 2 we review the essential points concerning semi-iterative methods for the Drazin-inverse solution of singular systems. In Sect. 3 we give the construction of a sequence of polynomials that are known as "residual polynomials". When constructed appropriately, these polynomials enable us to implement the associated semi-iterative methods by 4-term recursion relations independently of the size of $\text{ind}(A)$. In Sect. 4 we give the computational details of one of the methods that is directly related to Chebyshev polynomials and that we denote DCA for short. In Sect. 5 we give the convergence theory of all the methods we develop in Sect. 3 in general and the method of Sect. 4 in particular. We are able to give more refined results for the method of Sect. 4. Finally, in Sect. 6 we illustrate the use of DCA and the convergence theory of Sect. 4 with a numerical example.

Our treatment makes extensive use of the theory of orthogonal polynomial expansions.

By going through the details of the developments of the next sections it becomes clear that everything will remain unchanged even if a is not necessarily $\text{ind}(A)$ but is an upper bound on $\text{ind}(A)$.

Before going on we mention that Drazin inverses in general and Drazin-inverse solutions of singular linear systems in particular arise in problems of statistics such as determining steady states of Markov chains. They arise also in the solution of problems in control theory and singular differential and difference equations. See, e.g., Campbell and Meyer [CM, Chapters 9,10] and Campbell [C].

2 General background and motivation

Beginning with an arbitrary initial vector x_0 and its residual vector $r_0 = b - Ax_0$, all semi-iterative methods generate the vectors $x_1, x_2, \ldots,$ through

(2.1) $$x_m = x_0 + q_{m-1}(A)r_0,$$

where $q_{m-1}(\lambda)$ is a polynomial in λ of degree at most $m-1$. Let us define

(2.2) $$p_m(\lambda) = 1 - \lambda q_{m-1}(\lambda).$$

We call $p_m(\lambda)$ the mth *residual polynomial* since

(2.3) $$r_m = b - Ax_m = p_m(A)r_0.$$

Note also that

(2.4) $$p_m(0) = 1.$$

As is shown in [EMN], necessary and sufficient conditions for the convergence of the sequence $\{x_m\}$ are

(2.5) $$\lim_{m\to\infty} p_m^{(i)}(0) = 0, \;\; i = 1, \ldots, a; \;\; a = \text{ind}(A),$$

and

(2.6) $$\lim_{m\to\infty} p_m^{(i)}(\lambda_j) = 0, \;\; i = 0, 1, \ldots, k_j - 1,$$

where λ_j are the nonzero eigenvalues of A and $k_j = \text{ind}(A - \lambda_j I)$. In [CNS] the conditions of (2.5) are satisfied by picking $p_m(\lambda)$ such that

(2.7) $$p_m^{(i)}(0) = 0, \;\; i = 1, \ldots, a, \;\; \text{for all } m = 0, 1, \ldots .$$

Polynomials $p_m(\lambda)$ satisfying (2.4) and (2.7) were considered previously in [HH] for the case $a = 1$.

For convenience let us denote by Π_m the family of polynomials of degree at most m, and define

(2.8) $$\Pi_m^0 = \{p \in \Pi_m : p(0) = 1 \text{ and } p^{(i)}(0) = 0,\ i = 1, \ldots, a\}.$$

That is to say, Π_m^0 is the collection of all polynomials of degree at most m that satisfy (2.4) and (2.7). Note also that $p_m(\lambda) = 1$ is the only member of Π_m^0 for $m = 0, 1, \ldots, a$.

As in [CNS], in the present work too the polynomials $p_m(\lambda)$ are in Π_m^0 and, therefore, are necessarily of the form

(2.9) $$p_m(\lambda) = 1 - \lambda^{a+1} v(\lambda),\ \ v \in \Pi_{m-a-1},\ \ m \geq a + 1.$$

Now that we have decided to choose $p_m(\lambda)$ to be a polynomial in Π_m^0, we must make sure that the condition in (2.6) is satisfied, for without it we will not have a convergent method. We, therefore, dwell on this issue in the remainder of this section.

Let Ω be a closed domain in the complex λ-plane that contains only the nonzero eigenvalues of A, and define

(2.10) $$||f||_\Omega = \max_{\lambda \in \Omega} |f(\lambda)|.$$

Then, the sequence $\{p_m(\lambda)\}$ will satisfy (2.6) if it also satisfies

(2.11) $$\lim_{m \to \infty} ||p_m^{(i)}||_\Omega = 0,\ \ i = 0, 1, \ldots, \hat{k} - 1;\ \ \hat{k} = \max\{k_j : \lambda_j \neq 0\}.$$

But, by a result due to Pommerenke [P],

(2.12) $$||p_m'||_\Omega \leq K m^2 ||p_m||_\Omega,$$

for some constant $K > 0$ that is independent of m and p_m. Therefore, $\{p_m(\lambda)\}$ will satisfy (2.11) if it satisfies

(2.13) $$\lim_{m \to \infty} \left(m^{2\hat{k}-2} ||p_m||_\Omega \right) = 0.$$

Of course, if $||p_m||_\Omega = O(\mathrm{e}^{-\kappa m^\nu})$ as $m \to \infty$ for some $\kappa > 0$ and $\nu > 0$, then (2.13) and hence (2.6) will be satisfied. The polynomials of [HH] and [CNS] achieve precisely this. Keeping this in mind and drawing on the approach and results of Manteuffel [M] concerning Chebyshev acceleration for nonsingular systems whose matrices have complex spectra, we now propose the following course for the singular systems described in the first paragraph of Sect. 1.

Since the nonzero eigenvalues of the matrix A in (1.1) are assumed to lie in an open half-plane of the complex λ-plane whose boundary is a straight line through the origin, we can enclose these eigenvalues in a closed domain Ω that consists of an ellipse and its interior, and still ensure that the origin is neither in the interior of Ω nor on its boundary.

Let c, $c \pm f$, and ρ be, respectively, the center, foci, and sum of the semi-axes of this ellipse. Here c and f are in general complex and ρ is real and positive. Let us denote this ellipse by $\mathcal{E}(c, f, \rho)$. Thus, $\Omega = \text{int}\,\mathcal{E}(c, f, \rho) \cup \mathcal{E}(c, f, \rho)$. In the remainder of this work we shall take Ω to be exactly as described in this paragraph.

Recall that we want to have $\lim_{m\to\infty} ||p_m||_\Omega = 0$. The best way of achieving this is, obviously, by picking $p_m(\lambda)$ such that $||p_m||_\Omega = \min_{p\in\Pi_m^0} ||p||_\Omega$ for all m. However, $p_m(\lambda)$ determined this way, in general, do not lead to efficient recursive algorithms for $\{x_m\}$. If, however, we replace the maximum norm $||\cdot||_\Omega$ by an L_2-norm on Ω, we may be able to arrive at such a recursion relation. We now go a step further and replace the L_2-norm on Ω by an L_2-norm on the straight line segment joining the foci $c - f$ and $c + f$ of the ellipse $\mathcal{E}(c, f, \rho)$. Of course, the hope is that the polynomials $p_m(\lambda)$ obtained in this way will also satisfy (2.13).

An important point to mention concerning Ω is that, as the origin is neither in its interior nor on its boundary, the straight line segment joining the foci $c - f$ and $c + f$ of $\mathcal{E}(c, f, \rho)$ does not contain the origin either.

3 Construction of $p_m(\lambda)$ and x_m

Let $w(\lambda)$ be an admissible weight function on $[c-f, c+f]$, the straight line segment joining $c-f$ and $c+f$. That is to say, $w(\lambda) \geq 0$ for $\lambda \in [c-f, c+f]$. Define the inner product $(\cdot\,,\cdot)$ and norm $||\cdot||$ as in

$$(3.1)\quad (F, G) = \int_{c-f}^{c+f} w(\lambda)\overline{F(\lambda)}\,G(\lambda)\,|d\lambda| \ \text{ and } \ ||F|| = \sqrt{(F, F)}.$$

Here $|d\lambda|$ denotes the line element along $[c - f, c + f]$. Parameterizing $[c-f, c+f]$ by $\lambda = c+ft$, $t \in [-1, 1]$, what we mean by $\int_{c-f}^{c+f} H(\lambda)\,|d\lambda|$ is, of course,

$$\int_{c-f}^{c+f} H(\lambda)\,|d\lambda| = |f| \int_{-1}^{1} H(c + ft)\,dt.$$

In addition, we will need the inner product $(\cdot\,,\cdot)'$ and norm $||\cdot||'$ that are defined as in

$$(3.2)\qquad (F, G)' = (F, |\lambda|^{2a+2}G) \ \text{ and } \ ||F||' = \sqrt{(F, F)'},$$

and the polynomials $\phi_k(\lambda)$, $k = 0, 1, \ldots$, that are orthogonal with respect to $(\cdot\,,\cdot)'$, i.e., that satisfy

$$(3.3)\quad (\phi_k, \phi_l)' = \int_{c-f}^{c+f} w(\lambda)|\lambda|^{2a+2}\,\overline{\phi_k(\lambda)}\,\phi_l(\lambda)\,|d\lambda| = 0 \ \text{ if } \ k \neq l.$$

(Of course, $\phi_k(\lambda)$ is of degree exactly k for each k.) Obviously, we can pick the $\phi_k(\lambda)$ to be real on $[c-f, c+f]$. Equivalently, $\phi_k(c+ft)$ is real for t real.

We now would like to pick the residual polynomial $p_m(\lambda)$ to satisfy

$$(3.4)\qquad ||p_m|| = \min_{p\in\Pi_m^0} ||p||.$$

Theorem 3.1 *The minimization problem in* (3.4) *admits a unique solution* $p_m \in \Pi_m^0$ *that is given as*

$$(3.5)\quad p_m(\lambda) = 1 - \lambda^{a+1}v_m(\lambda) \ \textit{ with } \ v_m(\lambda) = \sum_{k=0}^{m-a-1} \delta_k\phi_k(\lambda),$$

where δ_k *are independent of* m *and are given by*

$$(3.6)\qquad \delta_k = \frac{(\phi_k, \lambda^{-a-1})'}{(\phi_k, \phi_k)'}, \quad k = 0, 1, \ldots\,.$$

In addition, $(p_m, \lambda^{a+j}) = 0$*,* $j = 1, \ldots, m-a$*.*

Proof. Using the fact that every $p \in \Pi_m^0$ is of the form $p(\lambda) = 1-\lambda^{a+1}v(\lambda)$, $v \in \Pi_{m-a-1}$, we first have

$$(3.7)\qquad \begin{aligned} ||p||^2 &= \int_{c-f}^{c+f} w(\lambda)\left|1-\lambda^{a+1}v(\lambda)\right|^2\,|d\lambda| \\ &= \int_{c-f}^{c+f} w(\lambda)|\lambda|^{2a+2}\left|\lambda^{-a-1}-v(\lambda)\right|^2\,|d\lambda| \\ &= \left(||\lambda^{-a-1}-v||'\right)^2. \end{aligned}$$

Substituting (3.7) in (3.4), we have

$$(3.8)\qquad \min_{p\in\Pi_m^0} ||p|| = \min_{v\in\Pi_{m-a-1}} ||\lambda^{-a-1}-v||'.$$

That is to say, we are actually looking for a best polynomial approximation $v(\lambda)$ in Π_{m-a-1} to the function λ^{-a-1} that is an L_2 function in the norm $||\cdot||'$. (Note that λ^{-a-1} is analytic on $[c-f, c+f]$ since the origin is not contained in $[c-f, c+f]$.) We know that there exists a unique polynomial $v \in \Pi_{m-a-1}$, v_m say, that solves (3.8), and $v_m(\lambda) = \sum_{k=0}^{m-a-1} \delta_k\phi_k(\lambda)$, with δ_k being independent of m and given by (3.6). We leave the rest to the reader. □

Since $\phi_k(\lambda)$ are orthogonal polynomials, they satisfy a 3-term recursion relation of the form

$$\phi_{k+1}(\lambda) = (\alpha_k\lambda + \beta_k)\phi_k(\lambda) + \gamma_k\phi_{k-1}(\lambda), \quad k = 1, 2, \ldots, \tag{3.9}$$

for some scalars α_k, β_k, and γ_k. We will now use this to obtain a recursion relation among the $p_m(\lambda)$ that will enable us to construct the sequence $\{x_m\}$ in a very efficient manner.

Let us define the polynomials $u_m(\lambda)$ by

$$u_m(\lambda) = \frac{p_m(\lambda) - p_{m+1}(\lambda)}{\lambda}, \quad m \geq a. \tag{3.10}$$

It is clear from Theorem 3.1 that

$$u_m(\lambda) = \lambda^a[v_{m+1}(\lambda) - v_m(\lambda)] = \lambda^a\delta_{m-a}\phi_{m-a}(\lambda). \tag{3.11}$$

Theorem 3.2 *The $u_m(\lambda)$ satisfy the 3-term recursion relation*

$$u_m(\lambda) = (\omega_m\lambda + \mu_m)u_{m-1}(\lambda) + \nu_m u_{m-2}(\lambda), \quad m \geq a + 2, \tag{3.12}$$

where

$$\omega_m = \alpha_{m-a-1}\frac{\delta_{m-a}}{\delta_{m-a-1}}, \quad \mu_m = \beta_{m-a-1}\frac{\delta_{m-a}}{\delta_{m-a-1}}, \quad \textit{and}$$
$$\nu_m = \gamma_{m-a-1}\frac{\delta_{m-a}}{\delta_{m-a-2}}. \tag{3.13}$$

Proof. The proof can be done by substituting (3.11) in (3.9). □

With the results of Theorems 3.1 and 3.2 available to us, we now go on to the recursive computation of the x_m.

Theorem 3.3 *The sequence $\{x_m\}$ defined in Sect. 1 can be computed from the 4-term recursion relation*

$$x_{m+1} = x_m + \omega_m A(x_m - x_{m-1}) + \mu_m(x_m - x_{m-1})$$
$$+\nu_m(x_{m-1} - x_{m-2}), \quad m \geq a + 2, \tag{3.14}$$

with the initial conditions

$$x_a = x_0, \quad x_{a+1} = x_a + \delta_0\phi_0(A)A^a r_0, \quad \textit{and}$$
$$x_{a+2} = x_{a+1} + \delta_1\phi_1(A)A^a r_0. \tag{3.15}$$

Proof. From (2.1), (2.2), and (3.10), we first note that

$$x_{m+1} - x_m = [q_m(A) - q_{m-1}(A)]r_0 = u_m(A)r_0. \tag{3.16}$$

Invoking now Theorem 3.2 in (3.16), the result in (3.14) follows. As for the initial conditions in (3.15), they follow from $x_0 = x_1 = \cdots = x_a$ and from

$$x_m = x_0 + \sum_{k=0}^{m-a-1} \delta_k \phi_k(A) A^a r_0, \quad m \geq a+1, \tag{3.17}$$

which is a consequence of

$$q_{m-1}(\lambda) = \lambda^a v_m(\lambda) = \sum_{k=0}^{m-a-1} \delta_k \phi_k(\lambda) \lambda^a. \tag{3.18}$$

□

Looking at the expressions that define the constants α_k, β_k, γ_k, and δ_k, we realize that when c and f are arbitrary complex numbers, they will be complex in general. In one case of interest to us in solving (1.1) a simplification takes place if c is real and f is purely imaginary and the weight function $w(\lambda)$ is chosen suitably.

Theorem 3.4 *Let c be real and f be purely imaginary, and pick $w(\lambda)$ to have the symmetry property*

$$w(c - ft) = w(c + ft), \quad 0 \leq t \leq 1. \tag{3.19}$$

Then δ_k is real for even k and purely imaginary for odd k. In addition, α_k and β_k are purely imaginary and γ_k is real, while ω_m, μ_m, and ν_m all turn out to be real.

Proof. We analyze first the δ_k by studying the expression in (3.6). Obviously, $(\phi_k, \phi_k)' = (||\phi_k||')^2 > 0$. Next,

$$\begin{aligned}(\phi_k, \lambda^{-a-1})' &= \int_{c-f}^{c+f} w(\lambda) \overline{\phi_k(\lambda)\, \lambda^{a+1}}\, |d\lambda| \\ &= |f| \sum_{j=0}^{a+1} \binom{a+1}{j} c^{a+1-j} \\ &\quad \times \int_{-1}^{1} w(c+ft)\phi_k(c+ft)(-ft)^j\, dt,\end{aligned} \tag{3.20}$$

where we have made use of the fact that $\overline{\lambda} = c - ft$ when $\lambda = c + ft$ and of the fact that $\phi_k(\lambda)$ is real for $\lambda \in [c-f, c+f]$. By (3.19), $\phi_k(c+ft)$ is an even or odd function of t if k is even or odd, respectively. As a result, in (3.20), the terms with $j + k$ odd vanish, which, by the fact that f is purely imaginary, implies that $(\phi_k, \lambda^{-a-1})'$ is real when k is even, and purely imaginary or zero when k is odd. This completes the proof of the assertion on the δ_k. We next take a look at the α_k, β_k, and γ_k. Recall that $\psi_k(t) \equiv \phi_k(c + ft)$ are real polynomials for t real and, as already shown, $\psi_k(t)$ is even or odd depending on whether k is even or odd respectively. As a result, the $\psi_k(t)$ satisfy the 3-term recursion relation

$$\psi_{k+1}(t) = \hat{\alpha}_k t \psi_k(t) + \hat{\gamma}_k \psi_{k-1}(t) \tag{3.21}$$

with $\hat{\alpha}_k$ and $\hat{\gamma}_k$ real. Letting now $\lambda = c + ft$ in (3.9), and comparing with (3.21), we see that

$$\hat{\alpha}_k = f\alpha_k, \;\; 0 = c\,\alpha_k + \beta_k, \;\; \text{and} \;\; \hat{\gamma}_k = \gamma_k.$$

Consequently, α_k and β_k are purely imaginary, while γ_k is real. The assertion about ω_m, μ_m, and ν_m now follows by invoking (3.13). This completes the proof. □

Theorem 3.4 has an important implication with regard to the actual computation of the vectors x_m. Suppose that the matrix A in (1.1) is real. This means that the nonzero eigenvalues of A are either real or they come in complex conjugate pairs. If the nonzero eigenvalues of A all have positive (or negative) real parts, then they can all be contained on or in the interior of an ellipse $\mathcal{E}(c, f, \rho)$ whose center c is real positive (or real negative) and whose foci are $c \pm f$, f being either real or purely imaginary, such that the origin is neither on $\mathcal{E}(c, f, \rho)$ nor in its interior. In case f is real, everything turns out to be real naturally. In case f is purely imaginary, we can cause the ω_m, μ_m, and ν_m to be real by *choosing* $w(\lambda)$ to satisfy (3.19). The global implication of all this is that if $Ax = b$ is a real system, the computation of the x_m through (3.14) can be carried out in *real* arithmetic.

3.1 The general algorithm

Based on the developments above, we now have the following algorithm for computing the x_m.

Algorithm 3.1.

Step 1. Pick an admissible weight function $w(\lambda)$ on $[c - f, c + f]$, and compute the coefficients α_k, β_k, and γ_k in the 3-term recursion relation of the corresponding orthogonal polynomials $\phi_k(\lambda)$ with respect to the inner product $(\cdot\,,\cdot)'$. (See (3.2) and (3.9).)
Step 2. Compute the coefficients δ_k in the expansion (3.5) by (3.6).
Step 3. Compute the coefficients ω_m, μ_m, and ν_m via (3.13).
Step 4. Compute the x_m via (3.14) and (3.15).

3.2 Computational remarks

It is clear from (3.13) that in order to be able to use the recursion relation in (3.14) and (3.15) we need (i) the coefficients α_k, β_k, and γ_k in the recursion relation of (3.9) among the $\phi_k(\lambda)$, and (ii) the δ_k defined in (3.5) and (3.6). In some cases these are available or can be computed without much difficulty.

In this respect the following observation concerning the δ_k is useful. By the fact that $\phi_k(\lambda) \equiv \psi_k(t)$ is real when $\lambda \in [c-f, c+f]$ (hence $t = (\lambda - c)/f \in [-1, 1]$) we can write

$$(3.22)\quad (\phi_k, \lambda^{-a-1})' = |f| \int_{-1}^{1} \hat{w}(c+ft)\psi_k(t)(c+ft)^{-a-1}\, dt = -f^{-a-1}\frac{1}{a!}\frac{d^a}{dz^a}\Lambda_k(z),$$

where $\hat{w}(\lambda) \equiv w(\lambda)|\lambda|^{2a+2}$, $z = -c/f$, and $\Lambda_k(z)$ is the function of the second kind corresponding to $\psi_k(t)$, namely,

$$(3.23)\quad \Lambda_k(z) = |f| \int_{-1}^{1} \hat{w}(c+ft)\frac{\psi_k(t)}{z-t}\, dt.$$

It is known that if the $\psi_k(t)$ satisfy the recursion relation

$$(3.24)\quad \psi_{k+1}(t) = (\hat{\alpha}_k t + \hat{\beta}_k)\psi_k(t) + \hat{\gamma}_k\psi_{k-1}(t), \quad k \geq 1,$$

then the $\Lambda_k(z)$ satisfy

$$(3.25)\quad \Lambda_{k+1}(z) = (\hat{\alpha}_k z + \hat{\beta}_k)\Lambda_k(z) + \hat{\gamma}_k\Lambda_{k-1}(z), \quad k \geq 1,$$

the difference being only in the initial conditions $\psi_0(t)$ and $\psi_1(t)$ in (3.24) and $\Lambda_0(z)$ and $\Lambda_1(z)$ in (3.25), respectively. As for the derivatives of the $\Lambda_k(z)$ these can be computed by differentiating the recursion relation in (3.25) and using the appropriate initial conditions, namely, $\frac{d^i}{dz^i}\Lambda_0(z)$, $\frac{d^i}{dz^i}\Lambda_1(z)$, $i = 1, \ldots, a$.

The procedure above becomes especially suitable for those choices of $w(\lambda)$ that give $\phi_k(\lambda)$ in terms of the classical orthogonal polynomials. For example, if we choose

$$(3.26)\quad w(\lambda) = |\lambda|^{-2a-2}|\lambda - c - f|^{\mu}|\lambda - c + f|^{\nu}, \quad \mu, \nu > -1,$$

then $\phi_k(\lambda) = P_k^{(\mu,\nu)}((\lambda - c)/f)$, where $P_k^{(\mu,\nu)}(t)$ are the Jacobi polynomials, and their 3-term recursion relation is simple and available. We can cause (3.19) to be satisfied and hence Theorem 3.4 to hold if we choose $\mu = \nu = \rho - \frac{1}{2}$ in (3.26). In this case $\phi_k(\lambda) = C_k^{(\rho)}((\lambda - c)/f)$, where

$C_k^{(\rho)}(t)$ are the ultraspherical (Gegenbauer) polynomials. In the next section we look at the most obvious special case of this for which $\rho = 0$ and $\psi_k(t)$ are the Chebyshev polynomials of the first kind, although other special cases such as that of Chebyshev polynomials of the second kind ($\rho = 1$) and that of Legendre polynomials ($\rho = \frac{1}{2}$) can be considered with the same ease.

4 The Drazin-Chebyshev acceleration (DCA)

Let us choose

$$w(\lambda) = |\lambda|^{-2a-2} \left(|\lambda - c - f|\,|\lambda - c + f|\right)^{-1/2}. \tag{4.1}$$

With this $w(\lambda)$, the inner product $(\cdot\,,\cdot)'$ becomes

$$\begin{aligned}(F, G)' &= \int_{c-f}^{c+f} \frac{\overline{F(\lambda)}G(\lambda)}{\sqrt{|\lambda - c - f|\,|\lambda - c + f|}}\,|d\lambda| \\ &= \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}}\,\overline{F(c+ft)}G(c+ft)\,dt.\end{aligned} \tag{4.2}$$

Then $\phi_k(\lambda) = T_k((\lambda - c)/f)$, where $T_k(t)$ is the kth Chebyshev polynomial of the first kind, and satisfies the recursion relation

$$T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t), \quad k \geq 1, \tag{4.3}$$

with the initial conditions

$$T_0(t) = 1 \text{ and } T_1(t) = t. \tag{4.4}$$

Therefore, α_k, β_k, and γ_k are readily available through

$$\alpha_k = \frac{2}{f}, \quad \beta_k = -\frac{2c}{f}, \quad \text{and} \quad \gamma_k = -1. \tag{4.5}$$

For determining the δ_k we need $(\phi_k, \phi_k)'$ and $(\phi_k, \lambda^{-a-1})'$. First,

$$(\phi_k, \phi_k)' = \int_{-1}^{1} \frac{[T_k(t)]^2}{\sqrt{1-t^2}}\,dt = \begin{cases} \pi \text{ if } k = 0, \\ \frac{\pi}{2} \text{ if } k \neq 0. \end{cases} \tag{4.6}$$

Next,

$$\begin{aligned}(\phi_k, \lambda^{-a-1})' &= \int_{-1}^{1} \frac{T_k(t)}{\sqrt{1-t^2}}\,(c+ft)^{-a-1}\,dt \\ &= \frac{f^{-a-1}}{a!}\frac{d^a}{dz^a}\left\{\int_{-1}^{1} \frac{T_k(t)}{\sqrt{1-t^2}}\,(t-z)^{-1}\,dt\right\}; \\ z &= -\frac{c}{f} \notin [-1, 1].\end{aligned} \tag{4.7}$$

But (see [L], p. 299)

(4.8)
$$\int_{-1}^{1} \frac{T_k(t)}{\sqrt{1-t^2}}\,(z-t)^{-1}\,dt = \frac{2\pi}{q-q^{-1}}\,q^{-k},\qquad k=0,1,\ldots;\ \ z\notin[-1,1],$$

where $q \equiv q(z)$ satisfies

(4.9)
$$q^2 - 2zq + 1 = 0 \ \text{ and } \ |q| > 1.$$

Combining (4.6)–(4.8), we thus have

(4.10)
$$\delta_k = \begin{cases} -2\dfrac{f^{-a-1}}{a!}\dfrac{d^a}{dz^a}\left(\dfrac{1}{q-q^{-1}}\right) & \text{if } k=0,\\ -4\dfrac{f^{-a-1}}{a!}\dfrac{d^a}{dz^a}\left(\dfrac{q^{-k}}{q-q^{-1}}\right) & \text{if } k\neq 0. \end{cases}$$

That is to say, the δ_k are also readily available. Note that as a is normally a small integer, the computation of δ_k is quite simple. All we need for this is $\dfrac{d^a}{dz^a}H_k(q(z))$, where $H_k(q) = q^{-k}/(q-q^{-1})$. For instance, when $a=1$ we have $\dfrac{d}{dz}H_k(q(z)) = H_k'(q(z))q'(z)$, where $q'(z)$ can be obtained by differentiating $q^2-2zq+1=0$. Specifically, $q'(z) = q/(q-z)$. We can compute the higher-order derivatives of $H_k(q(z))$ similarly.

In Theorem 5.2 we give the polynomials $p_m(\lambda)$ for $w(\lambda)$ as in (4.1) in closed form.

Let us now look at two special cases:

(i) c and f are both real, and $0 < f < c$.
In this case $z = -c/f < -1$, hence

$$q = z - \sqrt{z^2-1} = -\left(\frac{c}{f} + \sqrt{\left(\frac{c}{f}\right)^2 - 1}\right),$$

and everything is real.

(ii) c is real positive and f is purely imaginary, $f = \mathrm{i}|f|$.
In this case $z = -c/f = \mathrm{i}c/|f|$, hence

$$q = z + \sqrt{z^2-1} = \mathrm{i}\left(\frac{c}{|f|} + \sqrt{\left(\frac{c}{|f|}\right)^2 + 1}\right),$$

where we have taken $\sqrt{-1} = +\mathrm{i}$. That is, both z and q are purely imaginary.

For convenience we give here the algorithm resulting from the choice of $w(\lambda)$ as in (4.1).

Algorithm 4.1. (Drazin-Chebyshev Acceleration (DCA))

Input: (i) a, c, f.
(ii) A procedure for computing the δ_k by (4.10) with $z = -c/f$ and q defined as in (4.9).
(iii) A procedure for computing Av for any given vector v.

Step 1. Pick an arbitrary vector x_0 and compute $r_0 = b - Ax_0$ and $u = A^a r_0$.

Step 2. Set $x_a = x_0$ and compute $\Delta_a \equiv x_{a+1} - x_a = \delta_0 u$ and set $x_{a+1} = x_a + \Delta_a$ and $\Delta_{a+1} \equiv x_{a+2} - x_{a+1} = \dfrac{\delta_1}{f}(Au - cu)$ and set $x_{a+2} = x_{a+1} + \Delta_{a+1}$.

Step 3. For $m = a+2, a+3, \ldots$, until convergence compute
$\Delta_m = \dfrac{2}{f}\dfrac{\delta_{m-a}}{\delta_{m-a-1}} A\Delta_{m-1} - \dfrac{2c}{f}\dfrac{\delta_{m-a}}{\delta_{m-a-1}}\Delta_{m-1} - \dfrac{\delta_{m-a}}{\delta_{m-a-2}}\Delta_{m-2}$
and set $x_{m+1} = x_m + \Delta_m$.

Before closing this section, we would like to note that the weight function considered in [CNS] (with real positive c and f) is

$$w(\lambda) = \lambda^{-a}[(\lambda - c + f)(c + f - \lambda)]^{-1/2}, \ \ 0 < f < c,$$

that is different from the one in (4.1). (The case with $a = 1$ is already in [HH], as mentioned earlier.) Surprisingly, the solution for $p_m(\lambda)$ with this weight function can be expressed in terms of Chebyshev polynomials in a relatively simple way. The analogous weight function for complex c and f would be

$$w(\lambda) = |\lambda|^{-a}(|\lambda - c - f|\,|\lambda - c + f|)^{-1/2}.$$

It would be interesting to know whether we can express the corresponding $p_m(\lambda)$ in terms of Chebyshev polynomials, as was done in [CNS] for $0 < f < c$. This seems to be an intriguing mathematical problem whose solution is not immediate, and our efforts so far have not been successful.

5 Convergence analysis

In order to be able to carry out the error analysis, we need some new notation.

We shall denote by $\hat{\mathcal{S}}$ the direct sum of the invariant subspaces of A corresponding to its nonzero eigenvalues, and by $\tilde{\mathcal{S}}$, its invariant subspace corresponding to its zero eigenvalue. Thus, $\hat{\mathcal{S}}$ is $\mathcal{R}(A^a)$, the range of A^a, and $\tilde{\mathcal{S}}$ is $\mathcal{N}(A^a)$, the null space of A^a. Every vector in $\mathbb{C}^N$ can be expressed as the sum of two unique vectors, one in $\hat{\mathcal{S}}$ and the other in $\tilde{\mathcal{S}}$.

If we write $b = \hat{b} + \tilde{b}$, where $\hat{b} \in \hat{\mathcal{S}}$ and $\tilde{b} \in \tilde{\mathcal{S}}$, then $A^{\mathrm{D}}b$, the Drazin-inverse solution of $Ax = b$, is the unique vector in $\hat{\mathcal{S}}$ that satisfies the consistent system $Ax = \hat{b}$. Similarly, let us write $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x}_0 \in \hat{\mathcal{S}}$ and $\tilde{x}_0 \in \tilde{\mathcal{S}}$. Then from Theorem 4.1 in [CNS] we have

$$x_m - A^{\mathrm{D}}b - \tilde{x}_0 = p_m(A)(\hat{x}_0 - A^{\mathrm{D}}b), \tag{5.1}$$

so that the convergence of $\{x_m\}$ hinges on the convergence of $p_m(A)(\hat{x}_0 - A^{\mathrm{D}}b)$ to zero. Continuing exactly as is done following Theorem 4.1 in [CNS], and using (2.12), we obtain

$$||p_m(A)(\hat{x}_0 - A^{\mathrm{D}}b)|| \le K m^{2(\hat{k}-1)} ||p_m||_{\Omega}; \quad \hat{k} = \max\{k_j : \lambda_j \in \sigma(A)\backslash\{0\}\}, \tag{5.2}$$

where $|| \cdot ||_{\Omega}$ is as defined in (2.10), $K > 0$ is some constant that depends only on A and $\hat{x}_0$, and $\sigma(A)$ stands for the spectrum of A. Therefore, the rate of convergence of x_m to $(A^{\mathrm{D}}b + \tilde{x}_0)$ is determined by the rate of convergence of $||p_m||_{\Omega}$ to zero. Below we shall give a thorough analysis of the convergence of the sequence $\{||p_m||_{\Omega}\}$.

In Chapter 12 of Szegő [Sz] the following definition is given: A function $h(\theta)$ is said to belong to the class G if it is defined, nonnegative, and measurable on $[-\pi, \pi]$ and the integrals $\int_{-\pi}^{\pi} h(\theta)\, d\theta$ and $\int_{-\pi}^{\pi} |\log h(\theta)|\, d\theta$ converge, the first integral being positive.

Theorem 5.1 *Let $w(\lambda)$ be such that $h(\theta) \equiv w(c + f\cos\theta)|\sin\theta|$ is in the class G. Then*

$$\limsup_{m\to\infty} \left(||p_m||_{\Omega}\right)^{1/m} \le \frac{\rho}{\rho_0}, \tag{5.3}$$

where ρ_0 is the sum of the semi-axes of the ellipse whose center and foci are c and $c \pm f$ and that passes through the origin. Consequently,

$$\limsup_{m\to\infty} \left(||x_m - A^{\mathrm{D}}b - \tilde{x}_0||\right)^{1/m} \le \frac{\rho}{\rho_0}. \tag{5.4}$$

Remark. Obviously, $\mathcal{E}(c, f, \rho)$ and $\mathcal{E}(c, f, \rho_0)$ are confocal ellipses, the former being completely in the interior of the latter and $\rho < \rho_0$.

Proof. From Theorem 3.1 we know that $p_m(\lambda) = 1 - \lambda^{a+1} v_m(\lambda)$ so that

$$|p_m(\lambda)| = |\lambda^{a+1}|\, |\lambda^{-a-1} - v_m(\lambda)|. \tag{5.5}$$

Upon invoking the maximum modulus principle, (5.5) gives

$$||p_m||_{\Omega} \le C \left(\max_{\lambda \in \mathcal{E}(c,f,\rho)} |\lambda^{-a-1} - v_m(\lambda)| \right); \quad C = \max_{\lambda \in \mathcal{E}(c,f,\rho)} |\lambda^{a+1}|. \tag{5.6}$$

Again from Theorem 3.1 we know that $v_m(\lambda)$ is the $(m-a-1)$st partial sum of the orthogonal polynomial expansion $\sum_{k=0}^{\infty} \delta_k \phi_k(\lambda)$ of λ^{-a-1} with respect to the inner product $(\cdot\,,\cdot)'$.

For convenience we now switch to the variable t, $t = (\lambda - c)/f$. In this variable we have $\phi_k(\lambda) = \phi_k(c+ft) \equiv \psi_k(t)$, $k = 0, 1, \ldots$, and $\lambda^{-a-1} = (c+ft)^{-a-1} \equiv X(t)$. The ellipses $\mathcal{E}(c, f, \rho)$ and $\mathcal{E}(c, f, \rho_0)$ in the λ-plane are mapped to the ellipses $\mathcal{E}(0, 1, \rho/|f|)$ and $\mathcal{E}(0, 1, \rho_0/|f|)$, respectively in the t-plane. Thus

(5.7)

$$\max_{\lambda \in \mathcal{E}(c,f,\rho)} |\lambda^{-a-1} - v_m(\lambda)| = \max_{t \in \mathcal{E}(0,1,\rho/|f|)} \left| X(t) - \sum_{k=0}^{m-a-1} \delta_k \psi_k(t) \right|.$$

Now since $h(\theta)$ is in the class G, so is $\hat{h}(\theta) = h(\theta)|c + f\cos\theta|^{2a+2}$. Also, since λ^{-a-1} is analytic in $\operatorname{int}\mathcal{E}(c, f, \rho_0)$, $X(t)$ is analytic in $\operatorname{int}\mathcal{E}(0, 1, \rho_0/|f|)$. Therefore, from Theorem 12.7.3 in [Sz], $\sum_{k=0}^{\infty} \delta_k \psi_k(t)$ converges to $X(t)$ uniformly in every compact subset in the interior of $\mathcal{E}(0, 1, \rho_0/|f|)$, and we have

$$X(t) - \sum_{k=0}^{m-a-1} \delta_k \psi_k(t) = \sum_{k=m-a}^{\infty} \delta_k \psi_k(t) \quad \text{for all } t \text{ in the interior of } \mathcal{E}(0,1,\rho_0/|f|). \tag{5.8}$$

Consequently, by the fact that $\mathcal{E}(0, 1, \rho/|f|) \subset \operatorname{int}\mathcal{E}(0, 1, \rho_0/|f|)$, there holds

$$\begin{aligned} \max_{t \in \mathcal{E}(0,1,\rho/|f|)} \left| X(t) - \sum_{k=0}^{m-a-1} \delta_k \psi_k(t) \right| \\ \leq \sum_{k=m-a}^{\infty} |\delta_k| \left(\max_{t \in \mathcal{E}(0,1,\rho/|f|)} |\psi_k(t)| \right). \end{aligned} \tag{5.9}$$

Let us pick the $\phi_k(\lambda)$, hence $\psi_k(t)$, to be orthonormal, i.e.,

$$\begin{aligned} (\phi_k, \phi_l)' &= |f| \int_{-1}^{1} w(c+ft)|c+ft|^{2a+2} \psi_k(t)\psi_l(t)\,dt \\ &= \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{if } k \neq l. \end{cases} \end{aligned} \tag{5.10}$$

Then, from Theorem 12.1.2 in [Sz], there exists a unique function $U(\zeta)$, analytic for $|\zeta| > 1$ with $U(\infty) > 0$, such that, when t is in the complex plane cut along the real interval $[-1, 1]$, and ζ is that root of $t = \frac{1}{2}(\zeta + \zeta^{-1})$ satisfying $|\zeta| > 1$, there holds

$$\psi_k(t) \sim U(\zeta)\zeta^k \quad \text{as } k \to \infty. \tag{5.11}$$

Furthermore, (5.11) holds uniformly for $|\zeta| \geq R > 1$. Note also that when λ is on $\mathcal{E}(c, f, \rho)$, t is on $\mathcal{E}(0, 1, \rho/|f|)$ and the corresponding ζ satisfies $|\zeta| = \rho/|f|$. As a result,

$$\max_{t \in \mathcal{E}(0,1,\rho/|f|)} |\psi_k(t)| \leq L \left(\frac{\rho}{|f|}\right)^k \quad \text{for all } k. \tag{5.12}$$

Here $L > 0$ is some constant independent of k. Next, from the proof of Theorem 12.7.3 in [Sz], we have $\limsup_{k\to\infty} |\delta_k|^{1/k} \leq \rho_0/|f|$, so that, with $\epsilon > 0$ but arbitrarily close to zero,

$$|\delta_k| \leq \left(\frac{|f|}{\rho_0} + \epsilon\right)^k \quad \text{for all sufficiently large } k. \tag{5.13}$$

Combining (5.12) and (5.13) in (5.9), we obtain, for all sufficiently large m,

(5.14)
$$\max_{t \in \mathcal{E}(0,1,\rho/|f|)} \left| X(t) - \sum_{k=0}^{m-a-1} \delta_k \psi_k(t) \right| \leq L \sum_{k=m-a}^{\infty} \left[\left(\frac{|f|}{\rho_0} + \epsilon\right) \frac{\rho}{|f|}\right]^k,$$

from which (5.3) follows. Combining (5.3) with (5.1) and (5.2), (5.4) follows. □

A sharper result can be obtained for the special case of DCA treated in Sect. 4. This we do in the next theorem where we also give a closed-form expression for the corresponding polynomials $p_m(\lambda)$.

Theorem 5.2 *Let $w(\lambda)$ be as in* (4.1). *Then*

(5.15)
$$p_m(\lambda) = 2\left(\frac{\lambda}{f}\right)^{a+1} \frac{1}{a!} \frac{d^a}{dz^a} \left\{ \frac{q^{-m+a}}{q - q^{-1}} \, \frac{qT_{m-a}(t) - T_{m-a-1}(t)}{t - z} \right\},$$

where $t = (\lambda - c)/f$, $z = -c/f$, and q is as in (4.9), *and*

$$||p_m||_\Omega = O\left(m^a (\rho/\rho_0)^m\right) \quad \text{as } m \to \infty. \tag{5.16}$$

Consequently,

$$x_m - A^{\mathrm{D}}b - \tilde{x}_0 = O\left(m^{a+2(\hat{k}-1)} (\rho/\rho_0)^m\right) \quad \text{as } m \to \infty. \tag{5.17}$$

Proof. We start with

$$(5.18)\qquad p_m(\lambda) = \lambda^{a+1}\left[\lambda^{-a-1} - \sum_{k=0}^{m-a-1} \delta_k \phi_k(\lambda)\right] = \lambda^{a+1}\left[(c+ft)^{-a-1} - \sum_{k=0}^{m-a-1} \delta_k T_k(t)\right].$$

Substituting (4.10) and

$$(5.19)\qquad (c+ft)^{-a-1} = \frac{f^{-a-1}}{a!}\frac{d^a}{dz^a}(t-z)^{-1}, \quad z = -\frac{c}{f},$$

in (5.18), we obtain

$$(5.20)\qquad p_m(\lambda) = \left(\frac{\lambda}{f}\right)^{a+1}\frac{1}{a!}\frac{d^a}{dz^a}Q$$

with Q defined by

$$(5.21)\qquad Q = \frac{1}{t-z} + \frac{4}{q-q^{-1}}\sum_{k=0}^{m-a-1}{}' q^{-k}T_k(t),$$

where $\sum_{k=0}^{n}{}' d_k \equiv \frac{1}{2}d_0 + \sum_{k=1}^{n} d_k$. Now when $1 - 2t\sigma + \sigma^2 \neq 0$,

$$(5.22)\qquad \sum_{k=0}^{s-1}{}' \sigma^k T_k(t) = \frac{1-\sigma^2}{2(1-2t\sigma+\sigma^2)} - \sigma^s\,\frac{T_s(t) - \sigma T_{s-1}(t)}{1-2t\sigma+\sigma^2}.$$

By (5.22) and the fact that $q^2 + 1 = 2zq$, we have

$$(5.23)\qquad \sum_{k=0}^{m-a-1}{}' q^{-k}T_k(t) = \frac{q-q^{-1}}{4(z-t)} - q^{-m+a}\,\frac{qT_{m-a}(t) - T_{m-a-1}(t)}{2(z-t)},$$

which, upon substituting in (5.21), gives

$$(5.24)\qquad Q = \frac{2q^{-m+a}}{q-q^{-1}}\,\frac{qT_{m-a}(t) - T_{m-a-1}(t)}{t-z}.$$

The result in (5.15) now follows. To prove (5.16), we start by observing that $||p_m||_\Omega$ is achieved on $\mathcal{E}(c,f,\rho)$ in the λ-plane and on $\mathcal{E}(0,1,\rho/|f|)$ in the t-plane. We next observe that the operator $\dfrac{d^a}{dz^a}$ does not affect $T_{m-a}(t)$ and $T_{m-a-1}(t)$ as these are independent of z. Now when $t \in \mathcal{E}(0,1,\rho/|f|)$, $|T_k(t)| \sim \frac{1}{2}(\rho/|f|)^k$ as $k \to \infty$. Since $\lambda = 0$ when $t = z$,

Table 1.

number of iterations	max. error on $\mathcal{E}_1$	max. error on $\mathcal{E}_2$	max. error on $\mathcal{E}_3$
5	3.3	0.5	8.7×10^{-2}
10	0.4	9.6×10^{-3}	2.8×10^{-5}
15	3.0×10^{-2}	9.1×10^{-5}	4.6×10^{-9}
20	1.7×10^{-3}	6.7×10^{-7}	5.8×10^{-13}
25	7.7×10^{-5}	4.3×10^{-9}	6.4×10^{-17}
30	3.2×10^{-6}	2.5×10^{-11}	
35	1.2×10^{-7}	1.4×10^{-13}	
40	4.3×10^{-9}	7.4×10^{-16}	
45	1.4×10^{-10}		
50	4.4×10^{-12}		
55	1.4×10^{-13}		
60	5.4×10^{-15}		
65	1.9×10^{-16}		

$z \in \mathcal{E}(0,1,\rho_0/|f|)$, and since $\frac{1}{2}(q+q^{-1}) = z$, we have $|q| = \rho_0/|f|$. Finally, $\frac{d^a}{dz^a}(q^{-m}) = O(m^a q^{-m})$ as $m \to \infty$. Combining all these in (5.15) the proof of (5.16) can now be completed. The proof of (5.17) is achieved by combining (5.16) with (5.1) and (5.2). □

6 Numerical example

We now illustrate the theory of the previous section with a numerical example. In this example we take A to be a real $N \times N$ block-diagonal matrix, where $N = 2(n_1+n_2+n_3)+5$, whose first $n_1+n_2+n_3$ blocks are of the form $\begin{pmatrix} a_k^{(i)} & b_k^{(i)} \\ -b_k^{(i)} & a_k^{(i)} \end{pmatrix}$ while its last two blocks are

$$\begin{pmatrix} 0 & \epsilon_1 \\ 0 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, A has $n_1+n_2+n_3$ pairs of complex conjugate eigenvalues of the form $a_k^{(i)} \pm \mathrm{i}\, b_k^{(i)}$. We distribute these eigenvalues in the complex plane as follows: The first n_1 pairs of them are located on the ellipse $\mathcal{E}_1 = \mathcal{E}(11, \mathrm{i}\sqrt{11}, 11)$, the next n_2 pairs of them are located on the ellipse $\mathcal{E}_2 = \mathcal{E}(11, \mathrm{i}\sqrt{11}, 3+2\sqrt{5})$ that is in the interior of $\mathcal{E}_1$, and the last n_3 pairs of the eigenvalues are situated on the degenerate ellipse $\mathcal{E}_3 = \mathcal{E}(11, \mathrm{i}\sqrt{11}, \sqrt{11})$ in the interior of $\mathcal{E}_2$. Thus all three ellipses are confocal with center at $c = 11$ and foci at $c \pm f = 11 \pm \mathrm{i}\sqrt{11}$ and, therefore, have their semi-major axes perpendicular to the real axis. We choose $a_k^{(i)} = 11 + \alpha_i \cos(k-1)\theta_i$, and $b_k^{(i)} = \beta_i \sin(k-1)\theta_i$, $k = 1,\ldots,n_i$, where α_i and β_i are the semi-axes of the corresponding ellipse $\mathcal{E}_i$, and $\theta_i = \pi/(n_i - 1)$, $i = 1, 2, 3$. Next, we take the vector $\hat{b}$ to be the product $A\hat{x}$, where $\hat{x} = (\underbrace{1 \ldots 1}_{2(n_1+n_2+n_3)} \underbrace{0 \ldots 0}_{5})^{\mathrm{T}}$.

Then we add to $\hat{b}$ a vector $\tilde{b}$ whose first $2(n_1 + n_2 + n_3)$ entries are zero and the last 5 are arbitrary. We take $x_0 = 0$ so that the solution will be purely $A^{\mathrm{D}}b = \hat{x}$.

In Table 1 we present some of the results we obtained by applying DCA to the system $Ax = b = \hat{b} + \tilde{b}$ with $a = 2$ and $n_1 = 10$, $n_2 = 5$, and $n_3 = 5$.

By the theory developed in the previous sections the errors in the first $2n_1$ components of the solution converge to zero at the rate of $m^2(0.489)^m$, those in the next $2n_2$ components at the rate of $m^2(0.332)^m$, and those in the next $2n_3$ at the rate of $m^2(0.147)^m$. The last 5 components are zero in all x_m, $m = a, a + 1, \ldots$, as is clear from (5.1).

References

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

[C] S.L. Campbell: Singular Systems of Differential Equations. Pitman, London, 1980

[CM] S.L. Campbell, C.D. Meyer, Jr: Generalized Inverses of Linear Transformations. Pitman, London, 1979

[CNS] J.-J. Climent, M. Neumann, A. Sidi: A semi-iterative method for real spectrum singular linear systems with an arbitrary index. J. Comp. Appl. Math. **87**, 21–38 (1997)

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

[ER] M. Eiermann, L. Reichel: On the application of orthogonal polynomials to the iterative solution of singular linear systems of equations, in Vector and Parallel Computing, J. Dongarra, I. Duff, P. Gaffney, S. McKee (eds.), Ellis Horwood, Chichester, 1989, 285–297

[HH] M. Hanke, M. Hochbruck: A Chebyshev–like semiiteration for inconsistent linear systems. Electr. Trans. Numer. Anal. **1**, 89–103 (1993)

[K] Y. Kanevsky: Semi-iterative methods for Drazin-inverse solution of singular linear systems. M.Sc. thesis, Technion-Israel Institute of Technology, August 1999. (Supervisor: A. Sidi)

[L] Y.L. Luke: The Special Functions and Their Approximations. Vol. I, Academic Press, New York, 1969

[M] T.A. Manteuffel: The Tchebychev iteration for nonsymmetric linear systems. Numer. Math. **28**, 307–327 (1977)

[P] Ch. Pommerenke: On the derivative of a polynomial. Michigan Math. J. **6**, 373–375 (1959)

[Si] A. Sidi: A unified approach to Krylov subspace methods for the Drazin-inverse solution of singular nonsymmetric systems. Linear Algebra Appl. **298**, 99–113 (1999)

[Sz] G. Szegő: Orthogonal Polynomials. 4th edition, American Mathematical Society, Providence, RI, 1975

[W] H. Woźniakowski: Numerical stability of the Chebyshev method for the solution of large linear systems. Numer. Math. **28**, 191–209 (1977)