RIACS

Research Institute for Advanced Computer Science
NASA Ames Research Center

On Polynomial Preconditioning for Indefinite Hermitian Matrices

Roland Freund

IN-61
DATE OVERRIDE
43043
p 34

August 1989

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report 89.32

NASA Cooperative Agreement Number NCC 2-387

On Polynomial Preconditioning for Indefinite Hermitian Matrices

Roland Freund[†]

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report 89.32
August 1989

We are concerned with the minimal residual method combined with polynomial preconditioning for solving large linear systems $Ax = b$ with indefinite Hermitian coefficient matrices A. The standard approach for choosing the polynomial preconditioner leads to preconditioned systems which are postive definite. Here, we investigate a different strategy which leaves the preconditioned coefficient matrix indefinite. More precisely, the polynomial preconditioner is designed to cluster the positive, resp. negative eigenvalues of A around 1, resp. around some negative constant. In particular, it is shown that such indefinite polynomial preconditioners can be obtained as the optimal solutions of a certain two-parameter family of Chebyshev approximation problems. We establish some basic results for these approximation problems and sketch a Remez type algorithm for their numerical solution. The problem of selecting the parameters such that the resulting indefinite polynomial preconditioner speeds up the convergence of minimal residual method optimally is also addressed. For this task, we propose an approach based on the concept of asymptotic convergence factors. Finally, some numerical examples of indefinite polynomial preconditioners are given.

Subject Classification: AMS(MOS): 65F10, 41A10, 65D15.

Key words: Indefinite Hermitian matrices, minimal residual method, polynomial preconditioning, asymptotic convergence factor, Remez algorithm.

[†]Institut fur Angewandte, Mathematik und Statistik, Universitat Wurzburg, D-8700 Wurzburg, F.R.G., and the Research Institute for Advanced Computer Science, NASA Ames Research Center, Moffett Field, CA 94035. Work reported herein supported by Cooperative Agreement NCC2-387 between the National Aeronautics and Space Administration (NASA) and the Universities Space Research Association (USRA).

1. Introduction

Conjugate gradient type algorithms combined with preconditioning are among the most effective iterative procedures for solving large sparse nonsingular linear systems

$$Ax = b. \tag{1}$$

In recent years, polynomial preconditioning has attracted much interest. The technique consists of selecting a polynomial s of small degree and then applying a conjugate gradient type method to one of the two linear systems

$$s(A)Ax = s(A)b \tag{2}$$

(left preconditioning), or

$$As(A)y = b, \quad x = s(A)y \tag{3}$$

(right preconditioning). Remark, that (2) and (3) are both equivalent to the original linear system (1). Moreover, the systems (2) and (3) have the same coefficient matrix $s(A)A = As(A)$. Clearly, the polynomial s should be chosen such that the conjugate gradient iteration for (2) resp. (3) converges as fast as possible.

For the case of Hermitian positive definite A, the idea goes back to Rutishauser [24] who proposed polynomial preconditioning in the fifties as a remedy for roundoff in the classical conjugate gradient (CG hereafter) algorithm of Hestenes and Stiefel [17]. The recent revival [18] of Rutishauser's method and the general interest in polynomial preconditioning is mainly motivated by the attractive features of this technique for vector and parallel computers (see [25] for a survey).

In this note, we are concerned with polynomial preconditioning for linear systems (1) with Hermitian, but indefinite coefficient matrices A. An obvious strategy for the design of the preconditioner is to choose s such that $s(A)A$ is as close as possible to the identity matrix I. This approach was studied in detail by Ashby [2] and Ashby, Manteuffel, and Saylor [3]. Note that the resulting preconditioned system (2) resp. (3) is then Hermitian positive definite and thus can be solved by the standard CG algorithm.

The purpose of this paper is to document our study of a second preconditioning strategy which, in contrast to the first approach, leaves the preconditioned matrix $s(A)A$ indefinite. Roughly speaking, s is chosen such that $s(A)A$ is as close as possible to I on the positive part of the spectrum of A and as close as possible to μI, where $\mu \in \mathbb{R}$ is some negative constant, on the negative part of the spectrum of A. In particular, we will show how polynomials s of this type can be obtained as solutions of a family of Chebyshev approximation problems depending on two paramaters, namely μ and a weight factor $w \in \mathbb{R}$. The question of how to choose these parameters in order to speed up the convergence of the iteration as much as possible will also be addressed. Finally, note that, since the resulting matrix $s(A)A$ is now indefinite, the standard CG algorithm is no longer suitable for solving (2) resp. (3), and we use the minimal residual (MR hereafter) method instead.

The paper is organized as follows. In Section 2, we recall some basic facts about the MR algorithm. In Section 3, matrices with spectrum symmetric to the origin are considered, and it is shown that, in certain situations, the MR method is equivalent to the CG algorithm applied to the normal equations. In Section 4, we are concerned with the computation of the asymptotic convergence factor for the MR method based on the knowledge of two intervals which contain all eigenvalues of A. In Section 5, a two-parameter family of Chebyshev approximation problems is introduced, and some basic properties are listed. In Section 6, we consider indefinite polynomial preconditioners and show that there is an intimate connection with the class of approximation problems investigated in the previous section. A Remez type algorithm for the numerical solution of these problems is described in Section 7. Some numerical examples of indefinite polynomial preconditioners and their associated asymptotic convergence factors are presented in Section 8. Finally, we draw our conclusions in Section 9.

Throughout this paper, A is assumed to be a nonsingular Hermitian, but indefinite $N \times N$ matrix. $\sigma(A)$ denotes the spectrum of A, and $\|x\|_2 = \sqrt{x^H x}$ is the Euclidian norm of $x \in \mathbf{C}^N$. Moreover, the notation Π_n will be used for the set of all complex polynomials of degree at most n. Finally, we denote by $\Pi_n^{(r)}$ the subclass which consists of all real polynomials in Π_n.

2. The minimal residual algorithm. Error bounds

Let $x_0 \in \mathbf{C}^N$ be any initial guess for the true solution $A^{-1}b$ of (1), and let $r_0 = b - Ax_0$ be the corresponding residual vector. Moreover, we denote by

$$K_n := \operatorname{span}\{r_0, Ar_0, A^2 r_0, \dots, A^{n-1} r_0\} \tag{4}$$

the nth Krylov subspace of $\mathbf{C}^N$ generated by r_0 and A. Starting from x_0, the MR method generates a sequence of approximations x_n, $n = 1, 2, \dots$, to $A^{-1}b$ which are uniquely defined by the minimal residual property

$$\|b - Ax_n\|_2 = \min_{x \in x_0 + K_n} \|b - Ax\|_2 \ , \quad x_n \in x_0 + K_n. \tag{5}$$

For Hermitian positive definite matrices A, the MR algorithm was introduced by Stiefel [26] as a variant of the classical CG method. However, the algorithm given in [26] may break down (see e.g. [5,10]) for indefinite Hermitian matrices. A stable implementation (algorithm MINRES in [22]) of the MR approach for indefinite Hermitian matrices was first devised by Paige and Saunders [22]. The main ingredient of MINRES is the celebrated Lanzcos algorithm [20].

Algorithm 1 (Lanczos).

```
0) Set $v_0 = 0$, $\beta_1 = \|r_0\|_2$, $v = r_0$.
For $n = 1, 2, \dots$
1) If $\beta_n = 0$, stop.
   Otherwise, compute
2) $v_n = v/\beta_n$, $\alpha_n = v_n^H A v_n$,
```

$$v = Av_n - \alpha_n v_n - \beta_n v_{n-1},\ \beta_{n+1} = \|v\|_2\ .$$

In the following proposition, some basic facts about the Lanczos algorithm and its connection with the MR method are listed. We refer the reader to [13, pp. 325] and [22] for proofs. The notations

$$V_n := (v_1, v_2, \ldots, v_n) \quad \text{and} \quad S_n := \begin{pmatrix} \alpha_1 & \beta_2 & 0 & \cdots & 0 \\ \beta_2 & \alpha_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \beta_n \\ 0 & \cdots & 0 & \beta_n & \alpha_n \\ 0 & \cdots & \cdots & 0 & \beta_{n+1} \end{pmatrix} \tag{6}$$

are used. Note that S_n is an $(n+1) \times n$ matrix.

Proposition 1.
a) In exact arithmetic, Algorithm 1 stops for $n = m + 1$ where $m := \dim K_N$.
b) The termination index m is equal to the minimal number of components in any expansion of r_0 into orthonormal eigenvectors of A, i.e.

$$r_0 = \sum_{j=1}^{m} \rho_j u_j,$$
$$\text{where } \rho_j > 0,\ Au_j = \lambda^{(j)} u_j,\ \lambda^{(1)} < \lambda^{(2)} < \cdots < \lambda^{(m)},\ u_j^H u_k = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{if } j \neq k \end{cases}\ . \tag{7}$$

c) For $n = 1, 2, \ldots, m$ the nth iterate x_n of the MR method is given by $x_n = x_0 + V_n y_n$ where y_n is the solution of the linear system

$$S_n^H S_n y = \beta_1 S_n^H \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{8}$$

Moreover, in exact arithmetic $x_m = A^{-1}b$.

In MINRES, the MR iterates are computed via solving the linear system (8). This can be done very efficiently using a QR decomposition of S_n. Furthermore, such a factorization of S_n is readily obtained from the QR decomposition of S_{n-1} in the previous step (see [22] for details). The resulting algorithm can be stated as follows.

Algorithm 2 (MINRES implementation [22] of the MR method).

```
0) Choose $x_0 \in \mathbf{C}^N$ and set $v = b - Ax_0$, $v_0 = p_0 = p_{-1} = 0$,
     $\beta_1 = \bar{\eta}_1 = \|v\|_2$, $c_0 = c_{-1} = 1$, $s_0 = s_{-1} = 0$.
For $n = 1, 2, \ldots$
```

```
1) If $\beta_n = 0$, stop: $x_{n-1}$ solves $Ax = b$.
   Otherwise, compute
2) $v_n = v/\beta_n$, $\alpha_n = v_n^H A v_n$,
   $v = Av_n - \alpha_n v_n - \beta_n v_{n-1}$, $\beta_{n+1} = ||v||_2$,
3) $\epsilon_n = s_{n-2}\beta_n$, $\delta_n = s_{n-1}\alpha_n + c_{n-1}c_{n-2}\beta_n$,
   $\tilde{\gamma}_n = c_{n-1}\alpha_n - s_{n-1}c_{n-2}\beta_n$,
   $\gamma_n = \sqrt{\tilde{\gamma}_n^2 + \beta_{n+1}^2}$, $c_n = \tilde{\gamma}_n/\gamma_n$, $s_n = \beta_{n+1}/\gamma_n$,
4) $p_n = (v_n - \delta_n p_{n-1} - \epsilon_n p_{n-2})/\gamma_n$,
   $x_n = x_{n-1} + \eta_n p_n$ with $\eta_n = c_n\tilde{\eta}_n$,
   $\tilde{\eta}_{n+1} = -s_n\tilde{\eta}_n$.
```

Remark 1. The finite termination property of the Lanczos algorithm does no longer hold in the presence of roundoff error (see e.g. [13, pp. 332]), and the stopping criterion stated in Algorithm 2 is not useful in practice. Instead, one should terminate the iteration as soon as the norm $||r_n||_2$ of the residual vector $r_n = b - Ax_n$ is sufficiently reduced. As Paige and Saunders [22] have pointed out, $||r_n||_2$ can be obtained without computing the vector r_n itself by using the identity $||r_n||_2 = \beta_1 s_1 s_2 \cdots s_n$.

Remark 2. Other numerically stable implementations of the MR method for Hermitian indefinite matrices were devised by Fletcher [9] and Chandra [5]. See also [10, 28] for further properties of the MR approach. Finally, we note that — as is typical for conjugate gradient type algorithms — there is an intimate connection between the MR method and orthogonal polynomials (see [11]).

For the choice of a suitable preconditioner for a conjugate gradient type algorithm, it is crucial to have error bounds for its iterates. Next, we state such estimates for the MR method. For this purpose, some information on the location of the eigenvalues of A is necessary. Here and in the sequel, we assume that two intervals $[a, b]$ and $[c, d]$ are known such that

$$\sigma(A) \subset [a, b] \cup [c, d] \quad \text{where} \quad c < d < 0 < a < b. \tag{9}$$

Note that, ideally, b resp. c would be the largest resp. smallest eigenvalue of A and a resp. d the smallest positive resp. largest negative eigenvalue.

By the standard technique, expressing the Krylov subspace (4) $K_n = \{q(A)r_0 \mid q \in \Pi_{n-1}\}$ in terms of polynomials and using the expansion (7) of r_0, one readily deduces from (5) the following result.

Theorem 1. *For* $n = 1, 2 \ldots$:

$$\frac{||b - Ax_n||_2}{||b - Ax_0||_2} \leq E_n(a, b, c, d) \tag{10}$$

where $E_n(a, b, c, d)$ *is the optimal value of the approximation problem*

$$E_n(a, b, c, d) := \min_{p \in \Pi_n^{(r)} : p(0)=1} \; \max_{\lambda \in [a,b] \cup [c,d]} |p(\lambda)| \,. \tag{11}$$

Note that the outlined derivation of the bound (10), actually leads to the complex version of (11) with Π_n instead of $\Pi_n^{(r)}$. Standard results (e.g. [21]) from approximation theory guarantee that there always exists a unique optimal polynomial p_n^* for this complex approximation problem. Moreover, it is easily verified (cf. [21, Theorem 27]) that p_n^* is real, and therefore it is sufficient to consider only polynomials $p \in \Pi_n^{(r)}$ in (11).

Unfortunately, the solution of (11) is explicitly known only for special cases. For example, it is well known (see e.g. [2]) that for intervals of equal length $b - a = d - c$ the optimal polynomials are suitably transformed Chebyshev polynomials. The solution of (11) is also known for a variety of other parameters a, b, c, d, and can be found in the classical work of Achieser [1] (see also Peherstorfer [23, Section 5]). For the general case, there is no closed expression for the optimal value $E_n(a, b, c, d)$ of (11). However, it is known that for $n \to \infty$ this quantity behaves like κ^n where $\kappa = \kappa(a, b, c, d) \in (0, 1)$. More precisely, it holds

$$\lim_{n\to\infty} (E_n(a,b,c,d))^{1/n} =: \kappa(a,b,c,d) \quad \text{and} \quad 0 < \kappa(a,b,c,d) < 1 \tag{12}$$

(see Eiermann, Niethammer, and Varga [8], where this result is established for more general sets in the complex plane). $\kappa(a, b, c, d)$ is usually called the *asymptotic convergence factor*. In Section 4, we will derive an explicit formula for κ in terms of elliptic integrals. Based on this representation, κ can be very easily computed numerically.

3. A remark on matrices with symmetric spectrum

The simplest way to obtain a conjugate gradient type method for linear systems $Ax = b$ with arbitrary nonsingular coefficient matrix A, is to apply the standard CG algorithm to the Hermitian positive definte normal equations $A^H Ax = A^H b$. The drawback of this approach is that the condition number of $A^H A$ is the square of that of the original matrix A with the consequence that the resulting iteration will, in general, converge very slowly (see e.g. [27]). However, there are situations where working with the original system offers only little advantage over solving the normal equations or where the two approaches are even equivalent. Roughly speaking, this is the case if A has many eigenvalues in the right as well as in the left halfplane of $\mathbb{C}$ and/or if $\sigma(A)$ exhibits certain symmetries.

In this section, we are concerned with indefinite Hermitian matrices A with such a symmetrical spectrum. Since $A = A^H$, the normal equations corresponding to the original system (1) assume the form

$$A^2 x = Ab. \tag{13}$$

Next, we apply the standard CG algorithm [17] to (13) with $x_0 \in \mathbb{C}^N$ as initial guess. The resulting procedure — referred to as CGNE method in the sequel — generates a sequence of iterates x_k^{CGNE}, $k = 1, 2, \ldots$ which are characterized by the minimization property

$$\|b - Ax_k^{CGNE}\|_2 = \min_{x \in x_0 + K_k^2} \|b - Ax\|_2 \,, \quad x_k \in x_0 + K_k^2. \tag{14}$$

Here

$$K_k^2 := \operatorname{span}\{Ar_0, A^2(Ar_0), A^4(Ar_0), \ldots, A^{2(k-1)}(Ar_0)\}$$

is the kth Krylov subspace generated by Ar_0 and A^2. As in the previous section, we denote by x_n, $n = 1, 2, \ldots$, the iterates produced by the MR algorithm applied to the original system $Ax = b$. It is assumed that the MR and CGNE methods are both started with the same initial guess x_0.

It turns out that the MR and CGNE approaches are equivalent whenever the eigenvalues of A are symmetric to the origin, i.e.

$$\lambda \in \sigma(A) \quad \text{implies} \quad -\lambda \in \sigma(A), \tag{15}$$

and the starting residual r_0 has a "symmetric" expansion into eigenvectors of A. More precisely, we have the following

Theorem 2. *Let m, ρ_j, $\lambda^{(j)}$, $j = 1, \ldots, m$, be defined by the expansion (7) of r_0 and assume that $l := m/2 \in \mathbb{N}$ and that*

$$\lambda^{(j)} = -\lambda^{(m+1-j)}, \quad \rho_j = \rho_{m+1-j}, \quad j = 1, 2, \ldots, l, \tag{16}$$

holds. Then, for $k = 1, 2, \ldots, l$

$$x_{2k} = x_{2k+1} = x_k^{CGNE}.$$

Proof: Let $u_1, \ldots, u_m$ be the eigenvectors of A which occur in the expansion (7) of r_0. It is convenient to introduce the following notation. A vector $v \in \mathbf{C}^N$ is called *even*, resp. *odd*, if

$$v = \sum_{j=1}^{m} \xi_j u_j, \quad \text{with} \quad \xi_j = \xi_{m+1-j}, \quad \text{resp.} \quad \xi_j = -\xi_{m+1-j} \quad \text{for} \quad j = 1, 2, \ldots, l.$$

Obviously, the following properties hold:
(i) For any $\gamma \in \mathbf{C}$, γv is even, resp. odd, if v is even, resp. odd.
(ii) Av is even, resp. odd, if v is odd, resp. even.
(iii) $v^H A v = 0$ for any even or odd v.
Next consider Algorithm 1 and let v_n, $n = 1, \ldots, m$, be the Lanczos vectors. Clearly, r_0 and therefore also v_1 are even vectors. Using (i)–(iii), it follows by induction that v_n is even, resp. odd, if its index n is odd, resp. even, and that

$$\alpha_n = 0 \quad \text{and} \quad \beta_{n+1} v_{n+1} = A v_n - \beta_n v_{n-1} \quad \text{for} \quad n = 1, 2, \ldots, m. \tag{17}$$

The first identity in (17) and the definition of S_n in (6) imply that the linear system (8) has the following structure:

$$\begin{pmatrix} \times & 0 & \times & 0 & \cdots & 0 \\ 0 & \times & 0 & \times & \ddots & \vdots \\ \times & 0 & \times & 0 & \ddots & 0 \\ 0 & \times & 0 & \ddots & \ddots & \times \\ \vdots & \ddots & \ddots & \ddots & \times & 0 \\ 0 & \cdots & 0 & \times & 0 & \times \end{pmatrix} y = \begin{pmatrix} 0 \\ \times \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{18}$$

Here, a "×" indicates a possible nonzero entry. By reordering the equations and the unknowns, a system of the type (18) can be transformed into one with the block structure

$$\begin{pmatrix} \times & 0 \\ 0 & \times \end{pmatrix} \begin{pmatrix} y^{(1)} \\ y^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ \times \end{pmatrix}$$

where $y^{(1)}$ resp. $y^{(2)}$ contains all components of y with odd resp. even index. Hence $y^{(1)} = 0$ and, from part c) of Proposition 1, we deduce that

$$x_{2k+1} = x_{2k} \in x_0 + \operatorname{span}\{v_2, v_4, v_6, \ldots, v_{2k}\}. \tag{19}$$

On the other hand, (17) implies that the subspace on the right-hand side of (19) is just the Krylov space K_k^2 (note that $Ar_0 = \beta_2\beta_1 v_2$). Thus $x_{2k} \in x_0 + K_k^2$, and, in view of (5) and (14), it follows that $x_{2k} = x_k^{CGNE}$. This concludes the proof of the theorem. ■

Remark 3. The eigenvalues of Hermitian matrices of the type

$$A = \begin{pmatrix} 0 & B \\ B^H & 0 \end{pmatrix}, \quad \text{where } B \text{ is any } p \times q \text{ matrix,}$$

always fulfill the symmetry condition (15) (see e.g. [13, pp. 285]). Moreover, it is easily verified that the remaining part of the assumption (16) is guaranteed if the starting residual is of the form

$$r_0 = \begin{pmatrix} u \\ 0 \end{pmatrix}, \quad u \in \mathbf{C}^p \quad \text{or} \quad r_0 = \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v \in \mathbf{C}^q.$$

Remark 4. In [10], it is shown that the assumption (16) implies a similar equivalence between CG applied to the "inner" normal equations $A^2y = b$, $Ay = x$ and two other conjugate gradient type algorithms for $Ax = b$, namely SYMMLQ [22] and Fridman's method (see e.g. [28, 10]).

4. Computation of the asymptotic convergence factor for two intervals

In this section, we are concerned with the actual computation of the asymptotic convergence factor $\kappa(a, b, c, d)$ defined in (12). As Eiermann, Li, and Varga [7] pointed out, asymptotic convergence factors — not only for the union of two real intervals, but for more general compact sets $\Omega \subset \mathbf{C}$ — can be expressed in terms of the Green's function $G(\lambda; \infty)$ (see e.g. [29, pp. 65]) for $\Omega^c := \hat{\mathbf{C}} \setminus \Omega$ with pole at infinity. Note that the existence of $G(\lambda; \infty)$ is guaranteed, if Ω^c is of finite connectivity; moreover, the Green's function is then uniquely defined by the following three properties:
(i) $G(\cdot; \infty)$ is a real harmonic function on $\mathbf{C} \setminus \Omega$.
(ii) There is a $r_0 \in \mathbb{R}$ such that $G(\lambda; \infty) - \log|\lambda|$ is harmonic for all $\lambda \in \hat{\mathbf{C}}$ with $|\lambda| \geq r_0$.
(iii) $\lim_{\lambda \to \lambda_0} G(\lambda; \infty) = 0$ for all $\lambda_0 \in \partial\Omega$.

For $\Omega = [a,b] \cup [c,d]$, the set Ω^c is doubly connected, and by applying the results from [7, Section 3] it follows that

$$\kappa(a,b,c,d) = \exp\bigl(-G(0;\infty)\bigr). \tag{20}$$

Next, we use the connection (20) with the Green's function to derive a representation of $\kappa(a,b,c,d)$ in terms of elliptic integrals.

First, let $\Omega^c \subset \hat{\mathbf{C}}$ be any doubly connected region with $\infty \in \Omega^c$. Suppose we know a conformal mapping

$$f : A_r \to \Omega^c, \quad \text{with} \quad f(A_r) = \Omega^c, \tag{21}$$

of some annulus

$$A_r := \{z \in \mathbf{C} \mid r < |z| < 1\}, \quad \text{with} \quad 0 < r < 1, \tag{22}$$

onto Ω^c. Moreover, it is assumed that

$$\tau := f^{-1}(\infty) \quad \text{satisfies} \quad r < \tau < 1. \tag{23}$$

Note that (23) can always be achieved by a simple rotation in the z-plane. By means of f, the problem of finding the Green's function for Ω^c can be reduced to that of determining the Green's function $G_r(z;\tau)$ for the annulus A_r with pole at τ. More precisely, the following identity

$$G(\lambda;\infty) = G_r(f^{-1}(\lambda);\tau), \quad \lambda \in \Omega^c. \tag{24}$$

holds (e.g. [16, p. 259]). However, there are explicit representations for $G_r(z;\tau)$. Here we will use the following formula (see [16, pp. 259]):

$$G_r(z;\tau) = \Bigl(\frac{\log|z|}{\log r} - 1\Bigr)\log\tau - \log\left|\frac{\sum_{j=-\infty}^{\infty} \; r^{j^2}\bigl(-z/(r\tau)\bigr)^j}{\sum_{j=-\infty}^{\infty} \; r^{j^2}(-\tau z/r)^j}\right| . \tag{25}$$

From now on, let $\Omega := [a,b] \cup [c,d]$. So far, we have shown that, by means of (20), (24), and (25), the desired quantity $\kappa(a,b,c,d)$ can be expressed in terms of $f^{-1}(0)$ where f is a conformal mapping satisfying (21)–(23). Such functions f are explicitly known (see e.g. Kobers's dictionary of conformal representations [19, pp. 191]) and are of the form

$$f(z) = f_{k,\tau}(z) := \frac{a-b}{2}\,\frac{\operatorname{sn}^2\bigl(\frac{K'}{\pi}\log z;k\bigr) + \operatorname{sn}^2\bigl(\frac{K'}{\pi}\log\tau;k\bigr)}{\operatorname{sn}^2\bigl(\frac{K'}{\pi}\log z;k\bigr) - \operatorname{sn}^2\bigl(\frac{K'}{\pi}\log\tau;k\bigr)} + \frac{a+b}{2} . \tag{26}$$

Here, $w = \operatorname{sn}(u;k)$ is the Jacobian elliptic function (see e.g. [14, pp. 904]) defined — via its inverse $u = \operatorname{sn}^{-1}(w;k)$ — by

$$u = \operatorname{sn}^{-1}(w;k) = \int_0^w \frac{d\xi}{\sqrt{(1-\xi^2)(1-k^2\xi^2)}} . \tag{27}$$

The real number k is a parameter (the *modulus* of sn) with $k \in [0,1]$. The number K' in (26) is not a free parameter, but depends on k:

$$K' = K'(k) := \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1-(1-k^2)\sin^2\varphi}} \quad \Bigl(= \operatorname{sn}^{-1}(1;\sqrt{1-k^2}\Bigr) . \tag{28}$$

Similarly, we set

$$K = K(k) := \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}} \quad \left(= \mathrm{sn}^{-1}(1;k)\right) . \tag{29}$$

Note that $\mathrm{sn}(u;k)$ is a doubly-periodic meromorphic function with periods $4K$ and $2iK'$ and poles at the points $2mK + (2n+1)iK'$, $m,n \in \mathbb{Z}$. Finally, we remark that the branch of the logarithm in (26) is chosen such that

$$\log z = \log|z| + i \arg z \ , \quad -\pi < \arg z \leq \pi.$$

Using standard techniques from complex analysis, it is readily verified that the function (26) indeed maps an annulus A_r of the type (22) conformally onto the complement of two disjoint real intervals. Here, the inner radius r of A_r is given by

$$r = r(k) := \exp\left(\frac{-\pi K(k)}{K'(k)}\right) . \tag{30}$$

Moreover, the image of the outer boundary $|z| = 1$ of A_r under f is just the interval $[a,b]$. Hence, it only remains to adjust the two free parameters k and τ in (26) such that the inner boundary $|z| = r$ of A_r is mapped onto $[c,d]$. This requirement leads to the two equations

$$\frac{a-b}{2}\,\frac{1+\mathrm{sn}^2(M;k)}{1-\mathrm{sn}^2(M;k)} + \frac{a+b}{2} = c, \tag{31a}$$

$$\frac{a-b}{2}\,\frac{1/k^2+\mathrm{sn}^2(M;k)}{1/k^2-\mathrm{sn}^2(M;k)} + \frac{a+b}{2} = d, \tag{31b}$$

where we have set

$$M = M(k,\tau) := \frac{K'(k)\log\tau}{\pi} . \tag{32}$$

By solving first (31a) for $\mathrm{sn}^2(M;k)$ and, subsequently, (31b) for k^2, we obtain

$$\mathrm{sn}(M;k) = -\sqrt{\frac{a-c}{b-c}} \quad \text{and} \quad k = \sqrt{\frac{(a-d)(b-c)}{(a-c)(b-d)}} . \tag{33}$$

Note that, by (23) and (32), $M < 0$, and thus, in view of (27), also $\mathrm{sn}(M;k) < 0$. By (33) and (32), the two free parameters k and τ in (26) and the function f are now uniquely determined. By (20) and (24), we have

$$\kappa(a,b,c,d) = \exp\bigl(-G_r(z_0;\tau)\bigr) \quad \text{where} \quad z_0 := f^{-1}(0). \tag{34}$$

Therefore, it remains to determine the solution z_0 of $f(z) = 0$. To this end, we set

$$u_0 = \frac{K'}{\pi}\log z_0 \quad \text{or, equivalently,} \quad z_0 = \exp\left(\frac{\pi u_0}{K'}\right) . \tag{35}$$

Using (26) and the first relation in (33), it follows that u_0 is the solution of

$$\operatorname{sn}(u_0; k) = \sqrt{\frac{b}{a}}\ \operatorname{sn}(M; k) = -\sqrt{\frac{b(a-c)}{a(b-c)}}\ . \tag{36}$$

Next, recall (e.g. [14, p. 914]) the identity

$$\operatorname{sn}(v + iK'; k) \equiv \frac{1}{k \operatorname{sn}(v; k)}\ . \tag{37}$$

By means of (37), (35), and (33), we deduce from (36) that

$$u_0 = v_0 + iK', \quad v_0 := -\operatorname{sn}^{-1}\Big(\sqrt{\frac{a(b-d)}{b(a-d)}}; k\Big)\ , \quad \text{and} \quad z_0 = -\exp\Big(\frac{\pi v_0}{K'}\Big)\ . \tag{38}$$

Finally, using (34), (25) (with $z = z_0$), (30), (32), and (38), one arrives at the formula

$$\kappa(a,b,c,d) = \exp\Big(\Big(1 + \frac{v_0}{K}\Big)\frac{\pi M}{K'}\Big)\ \frac{\sum_{j=-\infty}^{\infty}\ \exp\Big(-\pi\frac{K}{K'}j^2 + \frac{\pi}{K'}(v_0 - M + K)j\Big)}{\sum_{j=-\infty}^{\infty}\ \exp\Big(-\pi\frac{K}{K'}j^2 + \frac{\pi}{K'}(v_0 + M + K)j\Big)}\ . \tag{39}$$

For the numerical evaluation of (39) it is advantageous to rewrite (39). To this end, let

$$\theta(z, \lambda) := \sum_{j=-\infty}^{\infty}\ \exp(-\pi\lambda j^2 + 2izj) \tag{40}$$

be one of the theta functions (see e.g. [14, p. 921]). By means of (40), it follows from (39) that

$$\kappa(a,b,c,d) = \exp\Big(\Big(1 + \frac{v_0}{K}\Big)\frac{\pi M}{K'}\Big)\ \frac{\theta(z_1, \lambda_0)}{\theta(z_2, \lambda_0)} \tag{41}$$

where

$$\lambda_0 = \frac{K}{K'}, \quad z_1 = \frac{\pi}{2iK'}(v_0 - M + K), \quad z_2 = \frac{\pi}{2iK'}(v_0 + M + K).$$

A straightforward computation, using Jacobi's identity (see e.g. [15, p.272])

$$\theta(z, \lambda) \equiv \frac{1}{\sqrt{\lambda}}\ \exp\Big(\frac{-z^2}{\pi\lambda}\Big)\ \theta\Big(\frac{z}{i\lambda}, \frac{1}{\lambda}\Big)$$

with $\lambda = \lambda_0$ and $z = z_1$ resp. $z = z_2$, shows that the representation (41) is equivalent to the final formula (43) stated in the following theorem. Furthermore, by the variable transformation $\xi = w/\sqrt{t+1}$ in (27), we have expressed the elliptic integrals, which occur in (28), (29), (33), and (38), in terms of the standard form

$$R_F(x,y,z) := \frac{1}{2}\int_0^{\infty} \frac{dt}{\sqrt{(t+x)(t+y)(t+z)}}\ , \quad x, y, z \geq 0, \tag{42}$$

of the ellpitic integral of the first kind.

Theorem 3. *Let $c < d < 0 < a < b$. Then:*

$$\kappa(a,b,c,d) = \frac{\vartheta_4\big(\pi(v_0 - M)/(2K), q\big)}{\vartheta_4\big(\pi(v_0 + M)/(2K), q\big)}, \quad \vartheta_4(\psi, q) := 1 + 2\sum_{j=1}^{\infty} (-1)^j q^{j^2} \cos(2\psi j), \tag{43}$$

where

$$\begin{gathered} q = \exp\Big(\frac{-\pi K'}{K}\Big), \quad k = \sqrt{\frac{(a-d)(b-c)}{(a-c)(b-d)}}, \quad K = R_F(1, 0, 1-k^2), \\ K' = R_F(1, 0, k^2), \quad M = -\sqrt{\frac{a-c}{b-c}}\, R_F\Big(1, \frac{b-a}{b-c}, \frac{b-a}{b-d}\Big), \\ v_0 = -\sqrt{\frac{a(b-d)}{b(a-d)}}\, R_F\Big(1, \frac{d(a-b)}{b(a-d)}, \frac{c(a-b)}{b(a-c)}\Big). \end{gathered} \tag{44}$$

The following corollary will follow readily from the representation (43) and (44) of the asymptotic convergence factor κ.

Corollary 1.
a) $\kappa(a,b,c,d)$ is a continuous function on $\{(a,b,c,d) \in \mathbb{R}^4 \mid c < d < 0 < a < b\}$.
b) Let $\{a_n\}_{n\in\mathbb{N}}$, $\{b_n\}_{n\in\mathbb{N}}$, $\{c_n\}_{n\in\mathbb{N}}$, and $\{d_n\}_{n\in\mathbb{N}}$, be given convergent sequences with limits a, b, c, and d, respectively. Moreover, assume that $c_n < d_n < 0 < a_n < b_n$ for all $n \in \mathbb{N}$ and that $c < d \leq 0 \leq a < b$. If $a = 0$ and/or $d = 0$, then

$$\lim_{n\to\infty} \kappa(a_n, b_n, c_n, d_n) = 1.$$

Proof: First, note that all the operations in (43) and (44) are continuous as long as $c < d < 0 < a < b$ holds, and part a) is obviously true. We now turn to the proof of part b). Let κ_n, $q^{(n)}$, $k^{(n)}, \ldots, v_0^{(n)}$ denote the quantities in (43) and (44) evaluated at a_n, b_n, c_n, d_n. We need to check their behavior for $n \to \infty$. There are three cases, namely
(i) $d = 0 < a$,
(ii) $d < 0 = a$,
(iii) $d = 0 = a$.
In the cases (i) and (ii), the sequences $q^{(n)}$, $k^{(n)}$, $K^{(n)}, \ldots, v_0^{(n)}$ converge for $n \to \infty$ to finite limits q, k, $K, \ldots, v_0$, respectively, and $K > 0$. Furthermore, $v_0 = -K$ in case (i) and $v_0 = 0$ in case (ii). Therefore, in view of (43), κ_n converges to 1. Finally, consider the case (iii). Here k_n converges to $k = 0$. By (29), (28), and (44), it follows that

$$\lim_{n\to\infty} K_n = 0, \quad \lim_{n\to\infty} K'_n = \infty, \quad \text{and} \quad \lim_{n\to\infty} q_n = 0.$$

Using the definition of the theta function in (43), we deduce

$$\lim_{n\to\infty} \vartheta_4(\psi_n, q_n) = 1 \quad \text{for all} \quad \psi_n \in \mathbb{R},$$

and hence, by (43), $\lim_{n\to\infty} \kappa_n = 1$. This concludes the proof of the corollary. ■

Remark 5. The theta functions ϑ_4 and θ in (43) and (40), respectively, are connected through

$$\vartheta_4(\psi, q) \equiv \theta(\psi + \pi/2, -(\log q)/\pi), \quad \psi \in \mathbb{R}, \quad 0 < q < 1.$$

There are whole books filled with the numerous properties and idenities which hold for theta functions. In the sequel, we will make use of the relations

$$\vartheta_4(\pi/2, q) = \sqrt{2K/\pi} \tag{45}$$

and

$$\begin{aligned} \vartheta_4(\psi, q) &= \prod_{j=1}^{\infty} \left(1 - 2q^{2j-1}\cos(2\psi) + q^{2(2j-1)}\right)\left(1 - q^{2j}\right) \\ &\geq \vartheta_4(0, q) = \left(\frac{2K}{\pi}\sqrt{1-k^2}\right)^{1/2} \quad \text{for all } \psi \in \mathbb{R} \end{aligned} \tag{46}$$

(see e.g. [14, pp. 921]). Here, q is defined in (44) with $K = K(k)$ and $K' = K'(k)$ given by (29) and (28).

By means of Theorem 3, the asymptotic convergence factor $\kappa(a, b, c, d)$ can be very easily computed numerically. For the calculation of the integrals R_F of the type (42), which occur in (44), there are standard algorithms. For the numerical examples presented in Section 8, we have used a procedure due to Carlson [4, Algorithm 1]. Finally, in (43), an infinite series needs to be computed twice. In the following, let $\psi \in \mathbb{R}$ and $J \in \mathbb{N}$. Moreover, suppressing the parameter q and the index 4, we set

$$\vartheta(\psi) := \vartheta_4(\psi, q) \quad \text{and} \quad \vartheta^{(J)}(\psi) := 1 + +2\sum_{j=1}^{J} (-1)^j q^{j^2} \cos(2\psi j)\ . \tag{47}$$

If J is chosen large enough, the finite series $\vartheta^{(J)}(\psi)$ will yield a sufficiently accurate approximation to $\vartheta(\psi)$. We now derive a formula for such an integer J. Using (43), (47), (45), and the fact that $0 < q < 1$, one obtains

$$\begin{aligned} |\vartheta(\psi) - \vartheta^{(J)}(\psi)| &= 2\left|\sum_{j=J+1}^{\infty} (-1)^j q^{j^2} \cos(2\psi j)\right| \\ &\leq 2\sum_{j=J+1}^{\infty} q^{j^2} = 2q^{(J+1)^2}\sum_{j=0}^{\infty} q^{j^2+2j(J+1)} \leq 2q^{(J+1)^2}\sum_{j=0}^{\infty} q^{j^2} \\ &= q^{(J+1)^2}\left(1 + \vartheta_4(\pi/2, q)\right) = q^{(J+1)^2}\left(1 + \sqrt{2K/\pi}\right)\ . \end{aligned} \tag{48}$$

With (47), (46), and (48), we arrive at the estimate

$$\left|\frac{\vartheta(\psi) - \vartheta^{(J)}(\psi)}{\vartheta(\psi)}\right| \leq q^{(J+1)}\ \frac{1 + \left(\pi/(2K)\right)^{1/2}}{\left(1-k^2\right)^{1/4}}. \tag{49}$$

From (49), it follows that the truncated series $\vartheta^{(J)}(\psi)$ approximates $\vartheta(\psi)$ with a relative error

$$\left|\frac{\vartheta(\psi)-\vartheta^{(J)}(\psi)}{\vartheta(\psi)}\right| \leq \epsilon,$$

if J is chosen as

$$J := [t] \quad \text{where} \quad s := \left|\frac{\log\bigl(\epsilon(1-k^2)^{1/4}\bigr)+\log\bigl(1+(\pi/(2K))^{1/2}\bigr)}{\log q}\right|^{1/2} .$$

Here, as usual, $[t]$ denotes the integer part of $t \in \mathbb{R}$.

We conclude this section by stating the following proposition which follows as a special case of a more general result due to Eiermann, Li, and Varga [7, Proposition 3]. This monotonicity property of the asymptotic convergence factor κ will be used in Section 6.

Proposition 2. *Let $c \leq c' < d' \leq d < 0 < a \leq a' < b' \leq b$ be given and assume that at least one of the inequalities "$\leq$" is strict. Then:*

$$\kappa(a',b',c',d') < \kappa(a,b,c,d).$$

5. A family of Chebyshev approximation problems

As we will see in the next section, the task of finding an optimal polynomial preconditioner for $Ax = b$ leads to a family of Chebyshev approximation problems. In this section, some results for such approximation problems are presented.

In the following, it is assumed that $S := [a,b] \cup [c,d]$ is the union of a positive and negative interval with arbitrary, but fixed endpoints $c < d < 0 < a < b$. Moreover, $l \in \mathbb{N}$ always denotes a positive integer. Finally, set

$$\Gamma := \{(\mu, w) \in \mathbb{R} \times \mathbb{R} \mid w > 0\}.$$

We will study the following family of approximation problems depending on the two parameters $(\mu, w) \in \Gamma$:

$$\gamma_l(\mu,w) := \min_{s\in\Pi_{l-1}^{(r)}} \|f-\lambda s\|_\omega , \quad \|f-\lambda s\|_\omega := \max_{\lambda\in S} \bigl|\omega(\lambda)\bigl(f(\lambda)-\lambda s(\lambda)\bigr)\bigr|, \tag{50}$$

where

$$\omega(\lambda) = \begin{cases} 1 & \text{if } \lambda > 0 \\ w & \text{if } \lambda < 0 \end{cases}, \quad f(\lambda) = \begin{cases} 1 & \text{if } \lambda > 0 \\ \mu & \text{if } \lambda < 0 \end{cases} . \tag{51}$$

Remark 6. For the special case $\mu = w = 1$, (50) reduces to the approximation problem (11) (with n replaced by l) which arose in Section 2 in connection with error bounds for the MR method.

(50) is a linear Chebyshev approximation problem: We seek to approximate $f(\lambda)$ by polynomials of the form $\lambda s(\lambda) \in \Pi_l^{(r)}$ in the weighted uniform norm $\|\cdot\|_\omega$. Note that $0 \notin S$, and this guarantees that Haar's condition is satisfied. Standard results (see e.g. [21]) from approximation theory show that there always exists a unique optimal polynomial for (50) which is characterized by an equioscillation property. We summarize these results for (50) in the following

Proposition 3. *Let $l \in \mathbb{N}$ and $(\mu, w) \in \Gamma$. Then:*
a) There exists a unique optimal polynomial $s_l^\star(\lambda; \mu, w) \in \Pi_{l-1}^{(r)}$ for (50).
b) $s \in \Pi_{l-1}^{(r)}$ is the optimal polynomial for (50) if, and only if, there exist $l+1$ extremal points

$$c \leq \lambda_0 < \lambda_1 < \cdots < \lambda_{k_{neg}-1} \leq d, \quad a \leq \lambda_{k_{neg}} < \lambda_{k_{neg}+1} < \cdots < \lambda_l \leq b \tag{52}$$

of $\omega(\lambda) - \lambda s(\lambda)$ and a number $y \in \mathbb{R}$ such that

$$\omega(\lambda_j(f(\lambda_j) - \lambda_j s(\lambda_j)) = \begin{cases} (-1)^j y & \text{for } j = 0, 1, \ldots, k_{neg} - 1 \\ (-1)^{j-1} y & \text{for } j = k_{neg}, k_{neg}+1, \ldots, l \end{cases} . \tag{53}$$

Moreover, if s is optimal, then $\gamma_l(\mu, w) = |y|$.

Here, a point $\lambda^\star \in S$ is called an *extremal point* of $f - \lambda s(\lambda)$ if

$$\left|w(\lambda^\star)(f(\lambda^\star) - \lambda^\star s(\lambda^\star))\right| = \|f - \lambda s\|_\omega .$$

The following corollary is a simple consequence of part b) of Proposition 3.

Corollary 2. *Let $s_l^\star(\lambda; \mu, w)$ be the optimal polynomial of (50).*
a) $s_l^\star \equiv 0$ if, and only if, $l = 1$ and $w = 1/\mu$.
b) Unless $s_l^\star \not\equiv 0$, there are at least $l+1$ and at most $l+3$ extremal points of $f - \lambda s_l^\star$ in S. Moreover, at most $l-1$ of these extremal points are contained in the interior $(a, b) \cup (c, d)$ of S.

Proof: By using (52) and (53), one readily verifies part a).
We now turn to part b). First, note that, by part b) of Proposition 3, $f - \lambda s_l^\star$ has at least $l+1$ extremal points in S. Next, recall (cf. (51)) that f is constant for $\lambda > 0$ and $\lambda < 0$, respectively. Hence

$$\left(f(\lambda) - \lambda s_l^\star(\lambda; \mu, w)\right)' = -\left(\lambda s_l^\star(\lambda; \mu, w)\right)' =: p(\lambda) \quad \text{for all } \lambda \neq 0. \tag{54}$$

Now assume that $s_l^\star \not\equiv 0$. Then p is a polynomial of degree not exceeding $l-1$ and $p \not\equiv 0$. This shows that p has at most $l-1$ zeros. On the other hand, in view of (54), $p(\lambda_j) = 0$ for all extremal points $\lambda_j \in S \setminus \{a, b, c, d\}$ of $f - \lambda s_l^\star$, and thus there are at most $l-1$ such "inner" extremal points λ_j. Therefore, altogether, there can not be more than $l - 1 + 4 = l + 3$ extremal points in S. ■

In the next section, we will also make use of the fact that the optimal value of (50) depends continuously on the parameters μ and w.

Lemma. *Let $l \in \mathbb{N}$. Then, the optimal value $\gamma_l(\mu, w)$ of (50) is a continuous function of $(\mu, w) \in \Gamma$.*

We remark that, for w fixed, it follows from a standard result (see e.g. [30, Lemma 13.1]) in Chebyshev approximation theory that $\gamma_l(\mu, w)$ is a continuous function of μ. The proof given in [30] is easily adapted to the family of approximation problems (50).

Proof of the Lemma: Let μ_1, $\mu_2 \in \mathbb{R}$, w_1 ,$w_2 > 0$ be arbitrary, and denote by f_1, f_2, ω_1, ω_2 the associated functions (51). Furthermore, assume that $l \in \mathbb{N}$ is fixed, and let $s_1^\star$ and $s_2^\star$ be the optimal polynomials of (50) corresponding to (μ_1, w_1) and (μ_2, w_2), respectively. By using the optimality of $s_1^\star$, the triangle inequality, and the obvious fact that $||\cdot||_{\omega_1} \leq \max\{1, w_1/w_2\}||\cdot||_{\omega_2}$, we obtain the estimates

$$\begin{aligned}\gamma_l(\mu_1, w_1) &= ||f_1 - \lambda s_1^\star||_{\omega_1} \leq ||f_1 - \lambda s_2^\star||_{\omega_1} \\ &\leq ||f_1 - f_2||_{\omega_1} + ||f_2 - \lambda s_2^\star||_{\omega_1} \\ &\leq w_1\,|\mu_1 - \mu_2| + \max\{1, w_1/w_2\}\ \gamma_l(\mu_2, w_2).\end{aligned} \tag{55}$$

With $\gamma_l(\mu_2, w_2) \leq ||f_2||_{\omega_2} = \max\{1, |\mu_2|w_2\}$, it follows from (55) that

$$\gamma_l(\mu_1, w_1) - \gamma_l(\mu_2, w_2) \leq w_1\,|\mu_1 - \mu_2| + \max\{0, (w_1 - w_2)/w_2\}\ \max\{1, |\mu_2|w_2\}. \tag{56}$$

Obviously, we may exchange the parameters (μ_1, w_1) and (μ_2, w_2) in (56). Therefore, (56) leads to the inequality

$$\begin{aligned}&|\gamma_l(\mu_1, w_1) - \gamma_l(\mu_2, w_2)| \\ &\leq |\mu_1 - \mu_2|\max\{w_1, w_2\} + |w_1 - w_2|\max\{1/w_1, 1/w_2\}\ \max\{1, |\mu_1|w_1, |\mu_2|w_2\}\end{aligned}$$

which implies the continuity of $\gamma_l(\mu, w)$. ■

Remark 7. In general, $\gamma_l(\mu, w)$ is not differentiable. Typically, differentiability gets lost when the number $k_{neg} = k_{neg}(\mu, w)$ of negative extremal points in (52) and (53) changes. The following example illustrates this behavior. Let $l = 2$, $S = [1,3] \cup [-2,-1]$, and $\mu = -2$ be fixed. It is straightforward to verify, by means of part b) of Proposition 3, that the best polynomial $s^\star(\lambda; w)$ and corresponding optimal value $\gamma(w)$ of (50) are given by

$$s^\star(\lambda; w) = \frac{2(4-\lambda)}{7}, \quad \gamma(w) = \frac{1}{7}, \quad \text{if} \quad 0 < w \leq 0.1,$$

and by

$$s^\star(\lambda; w) = \frac{2(2 - \lambda/\xi)}{\xi + 2 - 1/\xi}, \ \gamma(w) = \frac{\xi - 2 + 1/\xi}{\xi + 2 - 1/\xi}, \ \text{with } \xi = 1 + 2\sqrt{\frac{3w}{1+2w}}, \ \text{if } 0.1 \leq w \leq w_0.$$

Here w_0 is the unique root of $4w^2 - 188w + 49 = 0$ in the interval $(0,1)$. Moreover, the extremal points are $1,2,3$, if $0 < w < 0.1$, $-2,1,2,3$, if $w = 0.1$, and $-2,1,\xi$, if $0.1 < w \leq w_0$. Obviously, $\gamma(w)$ is a differentiable function of w for $0 < w < w_0$, $w \neq 0.1$, but, since

$$\lim_{w \to 0.1-0} \gamma'(w) = 0 \quad \text{and} \quad \lim_{w \to 0.1+0} \gamma'(w) = 100/147,$$

$\gamma(w)$ is not differentiable for $w = 0.1$.

6. Indefinite polynomial preconditioners

In this section, we return to the polynomial preconditioned MR method for solving the linear system (1), $Ax = b$. In particular, the question of how to choose an appropriate polynomial s for the preconditioned systems (2) resp. (3) is addressed.

As in Section 2, it is always assumed that A is a given indefinite Hermitian $N \times N$ matrix and that $a, b, c, d \in \mathbb{R}$ are known such that

$$\sigma(A) \subset S := [a, b] \cup [c, d] \quad \text{where} \quad c < d < 0 < a < b \tag{57}$$

(cf. (9)). In this paper, we will not consider the problem of how to actually obtain such bounds. The reader is referred to [3,11] where some results regarding this question can be found.

First, we note that the coefficient matrix $As(A)$ of the preconditioned systems (2) or (3) is Hermitian if, and only if, s is a real polynomial. Therefore, in the following, it is always assumed that $s \in \Pi_{l-1}^{(r)}$ where $l \in \mathbb{N}$ is an arbitrary, but fixed integer. Furthermore, in order to guarantee that $As(A)$ is nonsingular, we require that $s(\lambda) \neq 0$ for all $\lambda \in S$. Since s is continuous and in view of (57), this condition implies that there are essentially two different cases: Either

$$\lambda s(\lambda) > 0 \quad \text{for all } \lambda \in S, \tag{58}$$

or

$$\lambda s(\lambda) > 0 \quad \text{for all } \lambda \in [a, b], \quad \text{and} \quad \lambda s(\lambda) < 0 \quad \text{for all } \lambda \in [c, d]. \tag{59}$$

Clearly, also the two cases with reversed inequalities may occur, but these can always be reduced to (58) resp. (59) by replacing s by $-s$.

If (58) is satisfied, then, by (57), the preconditioned matrix $As(A)$ is positive definite. For the case (58), the standard strategy for the choice of the polynomial s is to require that $\lambda s(\lambda)$ approximates the constant function 1 as close as possible on S. Here, closeness is measured in the uniform norm on S, i.e. s is given as the optimal solution of the approximation problem (50), (51) with $\mu = w = 1$. This case was studied in detail by Ashby [2] and Ashby, Manteuffel, and Saylor [3].

If (59) holds, then, in view of (57), the preconditioned system remains indefinite, and we will use the

Definition 1. *A polynomial $s \in \Pi_{l-1}^{(r)}$ is called an indefinite polynomial preconditioner for $Ax = b$ if (59) is satisfied.*

In the following, we will investigate indefinite polynomial preconditioners and, in particular, develop a strategy for an optimal choice of s.

From now on, it is always assumed that s satisfies (59). The criterion for selecting the preconditioner which we will propose here is based only on properties of the coefficient matrix $As(A)$, and hence is the same for left and right polynomial preconditioning (2) and (3). For simplicity, we will consider only the approach (3) in the sequel. More precisely, let $x_0 \in \mathbf{C}^N$ be any initial guess for the solution of $Ax = b$, and let y_n, $n = 1, 2, \ldots$, be the sequence of iterates generated by the MR method (Algorithm 2) applied to

$$As(A)y = b - Ax_0 \ (=: r_0), \quad \text{with starting vector} \quad y_0 := 0. \tag{60}$$

The iterates and residual vectors corresponding to the original system $Ax = b$ are then given by

$$x_n = x_0 + s(A)y_n \quad \text{and} \quad r_n = b - Ax_n = r_0 - As(A)y_n, \tag{61}$$

respectively. Notice that only the iterates y_n are updated at each step. The corresponding approximate solution x_n of $Ax = b$ needs to be computed only once, namely in the very last step of the algorithm. Furthermore, we remark that, in view of (61), the residual vectors of y_n(with respect to (60)) and of x_n (with respect to $Ax = b$) are identical. This is a slight advantage of right polynomial preconditioning over the approach (2).

Next, using the results from Section 2, we state error bounds for the preconditioned MR method. Setting

$$\bar{a} := \min_{\lambda \in [a,b]} \lambda s(\lambda), \quad \bar{b} := \max_{\lambda \in [a,b]} \lambda s(\lambda), \quad \bar{c} := \min_{\lambda \in [c,d]} \lambda s(\lambda), \quad \bar{d} := \max_{\lambda \in [c,d]} \lambda s(\lambda), \tag{62}$$

it follows from (57) and (59) that

$$\sigma(As(A)) \subset \bar{S} := [\bar{a}, \bar{b}] \cup [\bar{c}, \bar{d}] \quad \text{and} \quad \bar{c} < \bar{d} < 0 < \bar{a} < \bar{b}. \tag{63}$$

Obviously, the numbers defined in (62) depend on s, and we will indicate this, if necessary, by writing $\bar{a}(s)$, $\bar{b}(s)$, $\bar{c}(s)$, $\bar{d}(s)$. Then, in view of (63), Theorem 1 yields the estimates

$$\frac{\|b - Ax_n\|_2}{\|b - Ax_0\|_2} \leq E_n(\bar{a}, \bar{b}, \bar{c}, \bar{d}), \quad n = 1, 2, \ldots . \tag{64}$$

Furthermore, by (12), the error bound in (64) behaves like

$$E_n(\bar{a}, \bar{b}, \bar{c}, \bar{d}) \approx \left(\kappa(\bar{a}, \bar{b}, \bar{c}, \bar{d})\right)^n, \quad \text{for } n \text{ large.} \tag{65}$$

Therefore, (64) and (65) suggest the following notion of an optimal indefinite polynomial preconditioner.

Definition 2. *An indefinite polynomial preconditioner $s^* \in \Pi_{l-1}^{(r)}$ is called optimal if*

$$\kappa(s^*) \leq \kappa(s) \tag{66}$$

for all indefinite polynomial preconditioners $s \in \Pi_{l-1}^{(r)}$. Here, and in the sequel,

$$\kappa(s) := \kappa\big(\bar{a}(s), \bar{b}(s), \bar{c}(s), \bar{d}(s)\big).$$

Finally, we get to the promised connection between indefinite polynomial preconditioners and the family of approximation problems (50). Let $(\mu, w) \in \Gamma$ and $s^*(\lambda) := s_l^*(\lambda; \mu, w)$

the corresponding optimal polynomial for (50). First, we characterize those cases where $s^\star$ yields an indefinite preconditioner. With (50) and (51), it follows that the numbers (62) associated with $s^\star$ are

$$\begin{aligned} \bar{a}(s^\star) &= 1 - \gamma_l(\mu, w), \quad \bar{b}(s^\star) = 1 + \gamma_l(\mu, w), \\ \bar{c}(s^\star) &= \mu - \frac{\gamma_l(\mu, w)}{w}, \quad \bar{d}(s^\star) = \mu + \frac{\gamma_l(\mu, w)}{w}. \end{aligned} \tag{67}$$

In view of (59), (62), and (67), $s^\star$ is an indefinite polynomial preconditioner if, and only if, $(\mu, w) \in \Gamma_l$. Here, we have set

$$\Gamma_l := \{(\mu, w) \in \mathbb{R} \times \mathbb{R} \mid w > 0, \ \gamma_l(\mu, w) < 1, \text{ and } \mu < -\gamma_l(\mu, w)/w\}. \tag{68}$$

Moreover, by (67), if $s^\star$ is an indefinite polynomial preconditioner, then

$$\kappa(s^\star) = g_l(\mu, w) := \kappa\big(1 - \gamma_l(\mu, w), 1 + \gamma_l(\mu, w), \mu - \frac{\gamma_l(\mu, w)}{w}, \mu + \frac{\gamma_l(\mu, w)}{w}\big). \tag{69}$$

Notice that $g_l(\mu, w)$ is a well-defined function for $(\mu, w) \in \Gamma_l$.

After all these preliminaries, we can now state the main result of this section in the following form.

Theorem 4. *Let $l \in \mathbb{N}$.*
a) Let $s \in \Pi_{l-1}^{(r)}$ be an indefinite polynomial preconditioner, $\bar{a}$, $\bar{b}$, $\bar{c}$, $\bar{d}$ the corresponding numbers defined in (62), and set

$$\mu_1 = \frac{\bar{d} + \bar{c}}{\bar{b} + \bar{a}} \quad \text{and} \quad w_1 = \frac{\bar{b} - \bar{a}}{\bar{d} - \bar{c}}. \tag{70}$$

Then, the optimal polynomial $s^\star(\lambda) := s_l^\star(\lambda; \mu_1, w_1)$ of (50) with parameters μ_1 and w_1 is an indefinite polynomial preconditioner and, unless $s^\star \equiv s$,

$$\kappa(s^\star) < \kappa(s). \tag{71}$$

b) There exist parameters μ_0 and w_0 such that

$$g_l(\mu_0, w_0) = \min_{(\mu, w) \in \Gamma_l} g_l(\mu, w), \quad (\mu_0, w_0) \in \Gamma_l. \tag{72}$$

c) Let μ_0 and w_0 satisfy (72). Then, the optimal polynomial $s_l^\star(\lambda; \mu_0, w_0)$ of the approximation problem (50) with parameters μ_0 and w_0 is an optimal indefinite polynomial preconditioner.

Proof: First, we prove part a). Let $s \in \Pi_{l-1}^{(r)}$ be an indefinite polynomial preconditioner, and hence, by (63), $\bar{c} < \bar{d} < 0 < \bar{a} < \bar{b}$. Moreover, by replacing s by $(2/(\bar{a} + \bar{b}))s$, we may assume without loss of generality that

$$\bar{a} + \bar{b} = 2. \tag{73}$$

Note that this does not change the asymptotic convergence factor $\kappa(s)$ associated with s. Indeed, it is easily verified that $\kappa(s) = \kappa(\alpha s)$ for all $\alpha \in \mathbb{R} \setminus \{0\}$ and all indefinite polynomial preconditioners s. Now, by using (62), (70), and (73), we obtain

$$\max_{\lambda \in [a,b]} |1 - \lambda s(\lambda)| = \max_{\lambda \in [c,d]} \left| w_1 \left(\mu_1 - \lambda s(\lambda) \right) \right| = \frac{\bar{b} - \bar{a}}{2}. \tag{74}$$

In view of (50) and (51), we conclude from (74) that

$$\gamma := \gamma_l(\mu_1, w_1) \leq \frac{\bar{b} - \bar{a}}{2} \quad \text{with} \quad \text{“ = ”} \quad \text{holding iff} \quad s \equiv s^\star. \tag{75}$$

With (67), (70), and (73), it follows from (75) that

$$\bar{c}(s) \leq \bar{c}(s^\star) < \bar{d}(s^\star) \leq \bar{d}(s) < 0 < \bar{a}(s) \leq \bar{a}(s^\star) < \bar{b}(s^\star) \leq \bar{b}(s) \tag{76}$$

where, unless $s \equiv s^\star$, at least one of the inequalities "$\leq$" is strict. In particular, (76) shows that $s^\star$ is an indefinite polynomial preconditioner. Moreover, by Proposition 2, (76) implies (71).

We now turn to the proof of part b). In view of (69), (68), part a) of Corollary 1 (see Section 4), and the Lemma proved in Section 5, the function $g_l(\mu, w)$ is continuous on Γ_l. Furthermore, by (12), (68), and (69),

$$g_l(\mu, w) < 1 \quad \text{for all} \quad (\mu, w) \in \Gamma_l. \tag{77}$$

Next, remark that, by (68), the boundary $\partial\Gamma_l$ of Γ_l is given by

$$\partial\Gamma_l = \{(\mu, w) \in \mathbb{R} \times \mathbb{R} \mid w > 0,\ 1 - \gamma_l(\mu, w) = 0,\ \text{and/or}\ \mu + \gamma_l(\mu, w)/w = 0\}. \tag{78}$$

By means of (69), (78), and part b) of Corollary, we conclude that

$$\lim_{(\mu,w) \to (\tilde{\mu},\tilde{w}),\ (\mu,w) \in \Gamma_l} g_l(\mu, w) = 1 \quad \text{for all} \quad (\tilde{\mu}, \tilde{w}) \in \partial\Gamma_l. \tag{79}$$

From (77), (79), and the continuity of g_l, it follows that g_l attains its minimum on Γ_l, i.e. (72) holds true.

Finally, in view of (69), (66), and (71), the statement of part c) is an immediate consequence of part a) of this theorem. ■

By means of part c) of Theorem 4, an optimal indefinite polynomial preconditioner can be constructed via the numerical solution of approximation problems of the form (50). In the next section, we sketch an algorithm for this task.

7. A Remez type algorithm

The standard tool for the numerical solution of general real linear Chebyshev approximation problems is the method of Remez (see e.g. [21, pp. 105] or [30, pp. 163]). For the case $\mu = w = 1$, de Boor and Rice [6] devised a Remez type procedure for the approximation problem (50) which exploits the special structure of (50). In this section, we outline an extension of their algorithm for the general family (50).

Let $l \in \mathbb{N}$, $\mu \in \mathbb{R}$, and $w > 0$ be given. We are concerned with the approximation problem (50) where the functions f and ω are defined in (51) and $S = [a,b] \cup [c,d]$ is the union of two intervals with endpoints $c < d < 0 < a < b$. In the following, let $s \in \Pi_{l-1}^{(r)}$ be any candidate for the optimal polynomial s^* of (50). It will be convenient to introduce the so-called residual polynomial

$$p(\lambda) = p(\lambda; s) := 1 - \lambda s(\lambda) \tag{80}$$

corresponding to s. Note that, by (51) and (80),

$$\omega(\lambda)\big(f(\lambda) - \lambda s(\lambda)\big) = \begin{cases} p(\lambda) & \text{if } \lambda > 0 \\ w\big(\mu - 1 + p(\lambda)\big) & \text{if } \lambda < 0 \end{cases} . \tag{81}$$

The Remez type procedure for the numerical solution of (50) is based on the equioscillation property stated in part b) of Proposition 3: We seek a polynomial $s \in \Pi_{l-1}^{(r)}$ with $l+1$ extremal points (52) such that (53) holds for some number $y \in \mathbb{R}$. For any $k_{neg} = 0, 1, \ldots, l$, denote by

$$\Lambda_{k_{neg}} := \{\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_l) \mid c \le \lambda_0 < \cdots < \lambda_{k_{neg}-1} \le d,\ a \le \lambda_{k_{neg}} < \cdots < \lambda_l \le b\}$$

the set of all possible λ_j for which (52) holds. By means of $\Lambda_{k_{neg}}$, we can parametrize all the polynomials s which fulfill (53).

Proposition 4. *To each $\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_l) \in \Lambda_{k_{neg}}$, $k_{neg} = 0, 1, \ldots, l$, there is a unique polynomial $s(\cdot;\Lambda) \in \Pi_{l-1}^{(r)}$ and a unique number $y(\Lambda) \in \mathbb{R}$ such that (53) holds true. Moreover, $s(\cdot;\Lambda)$ is defined via*

$$\begin{aligned} 1 - \lambda s(\lambda;\Lambda) &:= \sum_{j=0}^{k_{neg}-1} (1 - \mu + (-1)^j y/w) L_j(\lambda) + \sum_{j=k_{neg}}^{l} (-1)^{j-1} y L_j(\lambda), \\ L_j(\lambda) &:= \prod_{\substack{i=0 \\ i \ne j}}^{l} \frac{\lambda - \lambda_i}{\lambda_j - \lambda_i}, \end{aligned} \tag{82}$$

and $y = y(\Lambda)$ is given by

$$y := \frac{1 + (\mu - 1)\sum_{j=0}^{k_{neg}-1} L_j(0)}{(1/w)\sum_{j=0}^{k_{neg}-1}(-1)^j L_j(0) + \sum_{j=k_{neg}}^{l}(-1)^{j-1} L_j(0)}. \tag{83}$$

Proof: By means of (81), we rewrite (53) in the form

$$p(\lambda_j) = \begin{cases} 1 - \mu + (-1)^j y/w & \text{for } j = 0, 1, \ldots, k_{neg} - 1 \\ (-1)^{j-1} y & \text{for } j = k_{neg}, k_{neg} + 1, \ldots, l \end{cases} \tag{84}$$

where p is the residual polynomial (80) corresponding to $s(\cdot;\Lambda)$. Here $y \in \mathbb{R}$ is still a free parameter. By the Lagrange interpolation formula, for any fixed $y \in \mathbb{R}$, there is a unique polynomial $p \in \Pi_l^{(r)}$ which satisfies (84), and p is given by (82). It remains to determine y. For this purpose, we remark that p is the residual polynomial (80) of $s(\cdot;\Lambda) \in \Pi_{l-1}^{(r)}$ iff $p(0) = 1$. Using the Lagrange representation (82) of p, it is readily verified that this condition is equivalent to the formula for y in (83). Finally, we notice that — as is easily checked by means of (52) and the definition of L_j in (82) — all the terms in the sums of the numerator of the right-hand side of (83) have the same sign, namely $(-1)^{k_{neg}-1}$. Hence this numerator is never zero and y is well defined by (83). ■

In the sequel, we will use the notation

$$\epsilon(\lambda;\Lambda) := \omega(\lambda)\big(f(\lambda) - \lambda s(\lambda;\Lambda)\big) \tag{85}$$

for the error function corresponding to $s(\cdot;\Lambda)$. Now, in view of Proposition 4 and part b) of Proposition 3, the approximation problem (50) is reduced to the task of finding the unique vector $\Lambda^\star \in \Lambda_{k^\star_{neg}}$, with $k^\star_{neg} \in \{0, 1, \ldots, l\}$, whose components $\lambda_j^\star$ are indeed extremal points of $f(\lambda) - \lambda s(\lambda;\Lambda^\star)$. By (53) and (85), this last requirement is equivalent to

$$\big(\ |\epsilon(\lambda_j^\star;\Lambda^\star)|\ =\ \big)\quad |y(\Lambda^\star)| = \max_{\lambda \in S} |\epsilon(\lambda;\Lambda^\star)|. \tag{86}$$

Furthermore, note that for all $\Lambda = (\lambda_0, \ldots, \lambda_l) \in \Lambda_{k_{neg}}$ and $k_{neg} \in \{0, 1, \ldots, l\}$

$$\big(\ |\epsilon(\lambda_j;\Lambda)|\ =\ \big)\quad |y(\Lambda)| < |y(\Lambda^\star)| = \gamma_l(\mu, w) < \max_{\lambda \in S} |\epsilon(\lambda;\Lambda)|, \quad \text{if} \quad \Lambda \neq \Lambda^\star. \tag{87}$$

The optimal vector $\Lambda^\star$ which satisfies (86) can be computed by a Remez type iteration. We now describe a typical step of such a procedure. After, say n, iterations, the algorithm has generated an approximation $\Lambda := \Lambda^{(n)} \in \Lambda_{k_{neg}}$, where $k_{neg} := k_{neg}^{(n)}$, to $\Lambda^\star$. Unless $\Lambda = \Lambda^\star$, one constructs the next iterate $\tilde{\Lambda} := \Lambda^{(n+1)}$ as follows. In view of (86) and (87), the choice of the elements $\tilde{\lambda}_j$ of $\tilde{\Lambda}$ aims at fulfilling

$$|\epsilon(\tilde{\lambda}_j;\tilde{\Lambda})| \ \approx\ \max_{\lambda \in S} |\epsilon(\lambda;\tilde{\Lambda})|, \quad j = 0, 1, \ldots, l,$$

as good as possible. In order to achieve this, one first computes λ_j' which correspond to some approximate local maxima of $|\epsilon(\lambda;\Lambda)|$ near λ_j under the additional constraint that

$$\Lambda' := (\lambda_0', \lambda_1', \ldots, \lambda_l') \in \Lambda_{k_{neg}}.$$

Next, it is checked whether any of the endpoints $\lambda_e \in \{a, b, c, d\}$ of S satisfies $|\epsilon(\lambda_e; \Lambda)| > |y(\Lambda)|$ and can be exchanged with one of the elements, say λ'_{j_0} of Λ' such that

$$\tilde{\Lambda} := \Lambda' \cup \{\lambda_e\} \setminus \{\lambda'_{j_0}\} \in \Lambda_{\tilde{k}_{neg}}.$$

This procedure guarantees that $\tilde{\Lambda}$ approximates (86) better than Λ in the sense (cf. (87)) that

$$|y(\Lambda)| < |y(\tilde{\Lambda})| \leq |y(\Lambda^*)|.$$

The iterates $\Lambda^{(n)}$ of such a Remez procedure can be expected to converge quadratically to Λ^*. Here, we will not give a convergence proof and, instead, refer to the approximation theory literature (see [21, 30] and the references therein) for a more general study of Remez type methods.

The outlined Remez procedure for the approximation problem (50) can be summarized as follows.

Algorithm 3 (Sketch of a Remez procedure for (50)).

```
0) Choose $k_{neg} \in \{0, 1, \dots, l\}$ (e.g. $k_{neg} = [l(d-c)/(b-a+d-c)]$)
   and $\Lambda = (\lambda_0, \dots, \lambda_l) \in \Lambda_{k_{neg}}$.
Repeat steps 1) through 4) until convergence:
1) Using (83) and (82), compute $y := y(\Lambda)$ and some representation (e.g. Newton
   interpolation) of the residual polynomial $p$ corresponding to $s(\cdot; \Lambda)$.
2) For $j = 0, 1, \dots, l$:
   Set
        $\epsilon_j(\lambda) := (\operatorname{sign} \epsilon(\lambda_j; \Lambda))\epsilon(\lambda; \Lambda)$        (88)
   and compute $\lambda'_j$ such that:
   (i) $\epsilon_j(\lambda'_j)$ approximates some local maximum of $\epsilon_j(\lambda)$ near $\lambda_j$,
   (ii) $\Lambda' := (\lambda'_0, \lambda'_1, \dots, \lambda'_l) \in \Lambda_{k_{neg}}$.
3) Compute $\tilde{\Lambda} := (\tilde{\lambda}_0, \tilde{\lambda}_1, \dots, \tilde{\lambda}_l)$ as follows:
   (i) Compute $\eta_0 = \epsilon(c; \Lambda)$ and $\eta_1 = \epsilon(\lambda_0; \Lambda)$. If $k_{neg} = 0$, set $\eta_1 = -\eta_1$.
       If $\eta_0 \eta_1 < -y^2$:
       Set $\tilde{\lambda}_0 = c$, $\tilde{\lambda}_j = \lambda'_{j-1}$ for $j = 1, \dots, l$, $\tilde{k}_{neg} = k_{neg} + 1$,
       and go to 4).
   (ii) Compute $\eta_0 = \epsilon(d; \Lambda)$ and $\eta_1 = \epsilon(\lambda_{k_{neg}}; \Lambda)$. If $k_{neg} = 0$, set $\eta_1 = -\eta_1$.
       If $\eta_0 \eta_1 < -y^2$:
       Set $\tilde{\lambda}_j = \lambda'_j$ for $j = 0, \dots, k_{neg} - 1$, $\lambda_{k_{neg}} = d$,
       $\tilde{k}_{neg} = k_{neg} + 1$, $\tilde{\lambda}_j = \lambda'_j$ for $j = \tilde{k}_{neg}, \dots, l$,
       and go to 4).
   (iii) Compute $\eta_0 = \epsilon(a; \Lambda)$, and,
       if $k_{neg} \leq l$, $\eta_1 = \epsilon(\lambda_{k_{neg}}; \Lambda)$, resp., if $k_{neg} = l + 1$, $\eta_1 = -\epsilon(\lambda_l; \Lambda)$.
       If $\eta_0 \eta_1 < -y^2$:
       Set $\tilde{\lambda}_j = \lambda'_j$ for $j = 0, \dots, k_{neg} - 2$, $\tilde{k}_{neg} = k_{neg} - 1$,
       $\tilde{\lambda}_{\tilde{k}_{neg}} = a$, $\tilde{\lambda}_j = \lambda'_j$ for $j = \tilde{k}_{neg} + 1, \dots, l$,
       and go to 4).
```

```
    (iv) Compute $\eta_0 = \epsilon(b;\Lambda)$, and $\eta_1 = \epsilon(\lambda_l;\Lambda)$. If $k_{neg} = l+1$, set $\eta_1 = -\eta_1$.
        If $\eta_0\eta_1 < -y^2$:
        Set $\tilde{\lambda}_j = \lambda'_{j+1}$ for $j = 0,\ldots,l-1$,
        $\tilde{\lambda}_l = b$, $\tilde{k}_{neg} = k_{neg} - 1$,
        and go to 4).
    (v) Set $\tilde{\Lambda} := \Lambda'$, $\tilde{k}_{neg} = k_{neg}$.
4) Set $\Lambda := \tilde{\Lambda}$, $k_{neg} := \max\{\min\{\tilde{k}_{neg}, l\}, 0\}$, and go to 2).
```

Remark 8. Practical procedures for computing the approximate local maxima in step 2) of Algorithm 3 can be found in [12]. For the numerical examples presented in the next section, we have used quadratic interpolation. E.g. for an interior point $\lambda_j \in (a,b) \cup (c,d)$ one proceeds as follows. Let $\xi_0 := \lambda_j$ and $\xi := \lambda_{j+1}$ (resp. λ_{j-1}) if $\epsilon'_j(\xi_0) > 0$ (resp. < 0). Then, set $\xi_1 = \xi$ except for the case that $j = l$ (resp. $j = 0$) or for the case that ξ and ξ_1 are not contained in the same interval of $S = [a,b] \cup [c,d]$. In both cases, we choose ξ_1 as the endpoint of S which lies between ξ_0 and ξ. Next, the function ϵ_j is interpolated by the quadratic q defined by $q(\xi_0) = \epsilon_j(\xi_0)$, $q(\xi_1) = \epsilon_j(\xi_1)$, and $q'(\xi_0) = \epsilon'_j(\xi_0)$. If q attains its maximum, say at ξ^*, we repeat the whole process a second time with $\xi_1 := \xi^*$. The new resulting ξ^*, if it exists, is our canditate for λ'_j. If one of the two quadratic interpolations fails or if the resulting Λ' would not satisfy (52), a crude local search for the maximum near λ_j is applied, based on simply evaluating $\epsilon_j(\lambda)$ for several $\lambda \approx \lambda_j$ and determining the largest value.

Remark 9. In view of (81), the error function (85), ϵ, and the residual polynomial p defined in (82) are connected through

$$\epsilon(\lambda;\Lambda) = \begin{cases} p(\lambda) & \text{if } \lambda > 0 \\ w\big(\mu - 1 + p(\lambda)\big) & \text{if } \lambda < 0 \end{cases}. \tag{89}$$

In particular, the function (88), ϵ_j, and its derivative can easily be computed via (89) and some representation of the polynomial p.

8. Numerical examples

Based on the connection with the family of approximation problems (50), we have computed indefinite polynomial preconditioners in a number of cases. For the solution of (50), the Remez procedure described in the previous section was used. Optimal indefinite polynomial preconditioners were computed by solving the unconstrained optimization problem (72) numerically. Recall (cf. Remark 7 in Section 5) that the function g_l in (72) is continuous, but only piecewise differentiable. The numerical evaluation of asymptotic convergence factors was done as outlined in Section 4.

In the sequel, we present the results of a typical example. The set S is given by

$$S := [a,b] \cup [c,d] \quad \text{with} \quad a = 0.01,\ b = 0.99,\ c = -0.59,\ d = -0.1. \tag{90}$$

The asymptotic convergence factor, which corresponds to no preconditioning, is

$$\kappa(a,b,c,d) = 0.9590\ldots .$$

For $l = 2,3,\ldots,10$, we have computed indefinite polynomial preconditioners via solving (50) with the following parameters. The first choice

$$\mu_{-1} = -1 \quad \text{and} \quad w_{-1} = 1 \tag{91}$$

aims at clustering the positive and negative eigenvalues of $As(A)$ uniformly around 1 and -1, respectively. The resulting asymptotic convergence factor is denoted by κ_{-1} in the table below. A second obvious strategy is to choose the parameters in (50) such that the two intervals (63), $\bar{S} := [\bar{a},\bar{b}] \cup [\bar{c},\bar{d}]$ containing the eigenvalues of the preconditioned coefficient matrix $As(A)$ have the same relative length and position as the original intervals $[a,b] \cup [c,d]$, i.e.

$$\frac{b+a}{d+c} = \frac{\bar{b}+\bar{a}}{\bar{d}+\bar{c}} \quad \text{and} \quad \frac{b-a}{d-c} = \frac{\bar{b}-\bar{a}}{\bar{d}-\bar{c}}. \tag{92}$$

It is readily verified that (92), is fulfilled for the parameters

$$\mu_1 = \frac{d+c}{b+a} \quad \text{and} \quad w_1 = \frac{b-a}{d-c}$$

(cf. part a) of Theorem 4). The resulting asymptotic convergence factor for this choice will be denoted by κ_1. Note that for the example (90) considered here

$$\mu_1 = -0.69 \quad \text{and} \quad w_1 = 2. \tag{93}$$

Finally, via part c) of Theorem 4, we have also computed the optimal asymptotic convergence factor κ_{opt} and the corresponding parameters μ_{opt}, w_{opt} of (50). The following table lists the results for the three different strategies.

Table 1

These results are quite typical for the numerical experiments which we have performed. In particular, they show that the simple strategy (91) leads to very poor convergence rates, in particular as l increases. The second strategy leads to better results, but is still by far inferior to the best possible choice. Also notice that the optimal parameters μ_{opt} and w_{opt} exhibit a rather erratic behavior as l increases.

The following two plots show, for two cases, the polynomials $\lambda s_{10}^{\star}(\lambda)$ corresponding to the indefinite preconditioned coefficient matrix $As(A)$. Here $s_{10}^{\star}(\lambda)$ denotes the optimal polynomial of (50) with $l = 10$. For Figure 1, the parameters (93), $\mu_1 = -0.69$ and $w_1 = 2$, were used. Figure 2 corresponds again to $\mu = \mu_1$, but with increased weight $w = 10$.

Figure 1

Figure 2

Finally, the last two plots show the surface of the function $1/g_l(\mu, w)$ (cf. (69)) whose maximum, in view of part c) of Theorem 4, corresponds to an optimal indefinite polynomial preconditioner. For the plots we have set $g_l(\mu, w) = 1$ if $(\mu, w) \not\in \Gamma_l$ (cf. (68)). In both cases, the left corner is the point $(\mu, w) = (0, 0)$. The axis pointing towards the reader is the μ-axis. Figure 3 displays the results for $l = 3$ and Figure 4 for $l = 10$.

Figure 3

Figure 4

9. Conclusions

We have investigated polynomial preconditioners for Hermitian indefinite linear systems which lead to indefinite preconditioned coefficient matrices. In particular, it was shown that such polynomials can be obtained via the solution of a two-parameter family of Chebyshev approximation problems. Based on the concept of asymptotic convergence factors, we have introduced the notion of an optimal indefinite polynomial preconditioner. In order to actually compute such an optimal preconditioner, one needs to minimize a continuous, but only piecewise differentiable function of two variables. Moreover, each evaluation of this function requires the solution of an approximation problem of the type (50). A Remez type procedure for the numerical solution of (50) was outlined. A few numerical examples of indefinite polynomial preconditioners were presented. In a forthcoming technical report, we will report on numerical tests for the minimal residual algorithm combined with the indefinite preconditioners developed in this paper and compare this approach with other preconditioning strategies for indefinite Hermitian matrices.

References

[1] Achieser, N.I.: Über einige Funktionen, welche in zwei gegebenen Intervallen am wenigsten von Null abweichen. *Bull. Acad. Sci. URSS VII Série* **9**, 1163–1202 (1932)

[2] Ashby, S.F.: Polynomial preconditioning for conjugate gradient methods. Ph. D. Thesis, Department of Computer Science, Report 1355, University of Illinois at Urbana-Champaign, December 1987

[3] Ashby, S.F., Manteuffel, T.A., Saylor, P.E.: Adaptive polynomial preconditioning for Hermitian indefinite linear systems. Technical Report UCRL-100970, Lawrence Livermore National Laboratory, April 1989

[4] Carlson, B.C.: Computing elliptic integrals by duplication. *Numer. Math.* **33**, 1–16 (1979)

[5] Chandra, R.: Conjugate gradient methods for partial differential equations. Ph. D. Thesis, Computer Science Department, Research Report 129 Yale University, January 1978

[6] de Boor, C., Rice, J.R.: Extremal polynomials with application to Richardson iteration for indefinite linear systems. *SIAM J. Sci. Stat. Comput.* **3**, 47–57 (1989)

[7] Eiermann, M., Li, X., Varga, R.S.: On hybrid semi-iterative methods. *SIAM J. Numer. Anal.* **26**, 152–168 (1989)

[8] Eiermann, M., Niethammer, W., Varga, R.S.: A study of semiiterative methods for nonsymmetric systems of linear equations. *Numer. Math.* **47**, 505–533 (1985)

[9] Fletcher, R.: Conjugate gradient methods for indefinite systems. In: Numerical Analysis Dundee 1975 (G.A. Watson, ed.), pp. 73-89. Lecture Notes in Mathematics 506. Berlin, Heidelberg, New York: Springer 1976

[10] Freund, R.: Über einige cg-ähnliche Verfahren zur Lösung linearer Gleichungssysteme. Doctoral Thesis, Universität Würzburg, F.R. of Germany, May 1983

[11] Freund, R.: Pseudo Ritz values for indefinite Hermitian matrices. Technical Report TR 89.33, RIACS, NASA Ames Research Center, August 1989

[12] Golub, G.H., Smith, L.B.: Chebyshev approximation of continuous functions by a Chebyshev systems of functions. *Collected Algorithm from CACM* **414**, P1–P10 (1971)

[13] Golub, G.H., Van Loan, C.F.: Matrix computations. Baltimore: The Johns Hopkins University Press 1983

[14] Gradshteyn, I.S., Ryzhik, I.M.: Table of integrals, series, and products. San Diego: Academic Press 1980

[15] Henrici, P.: Applied and computational complex analysis Vol. 2. New York, London, Sydney, Toronto: Wiley 1977

[16] Henrici, P.: Applied and computational complex analysis Vol. 3. New York, London, Sydney, Toronto: Wiley 1986

[17] Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards* **49**, 409-436 (1952)

[18] Johnson, O.G., Micchelli, C.A., Paul, G.: Polynomial preconditioners for conjugate gradient calculations. *SIAM J. Numer. Anal.* **20**, 362-376 (1983)

[19] Kober, H.: Dictionary of conformal representations. New York: Dover 1957

[20] Lanczos, C.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. *J. Res. Nat. Bur. Standards* **45**, 255-282 (1950)

[21] Meinardus, G.: Approximation of functions: Theory and numerical methods. Berlin, Heidelberg, New York: Springer 1967

[22] Paige, C.C., Saunders, M.A.: Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.* **12**, 617-629 (1975)

[23] Peherstorfer, F.: Orthogonal polynomials in L^1-approximation. *J. Approx. Theory* **52**, 241–268 (1988)

[24] Rutishauser, H.: Theory of gradient methods. In: Refined iterative methods for computation of the solution and the eigenvalues of self-adjoint boundary value problems, pp. 24-49. Mitteilungen aus dem Institut für Angewandte Mathematik an der ETH Zürich **8**. Basel: Birkhäuser 1959

[25] Saad, Y.: Krylov subspace methods on supercomputers. *SIAM J. Sci. Stat. Comput.*, to appear

[26] Stiefel, E.: Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme. *Comment. Math. Helv.* **29**, 157–179 (1955)

[27] Stoer, J.: Solution of large linear systems of equations by conjugate gradient type methods. In: Mathematical programming – The state of the art (A. Bachem, M. Grötschel, and B. Korte, eds.), pp. 540-565. Berlin, Heidelberg, New York, Tokyo: Springer 1983

[28] Stoer, J., Freund, R.: On the solution of large indefinite systems of linear equations by conjugate gradient algorithms. In: Computing methods in applied sciences and engineering V (R. Glowinski and J.L. Lions, eds.), pp. 35-53. Amsterdam: North Holland 1982

[29] Walsh, J.L.: Interpolation and approximation by rational functions in the complex domain. Amer. Math. Soc. Colloq. Publ. Vol. XX, 5th edition. Providence, R. I.: Amer. Math. Soc. 1969

[30] Werner, H.: Vorlesung über Approximationstheorie. Lecture Notes in Mathematics 14. Berlin, Heidelberg, New York: Springer 1966

l	2	3	4	5	6	7	8	9	10
κ_{-1}	0.986	0.974	0.962	0.957	0.937	0.937	0.922	0.915	0.906
κ_1	0.959	0.936	0.932	0.918	0.908	0.902	0.886	0.885	0.869
κ_{opt}	0.948	0.905	0.874	0.859	0.821	0.820	0.776	0.752	0.734
μ_{opt}	-1.92	-0.68	-2.37	-0.68	-1.92	-1.96	-1.68	-3.78	-1.60
w_{opt}	0.65	3.40	0.74	5.80	1.45	1.40	2.70	0.76	4.40

Table 1

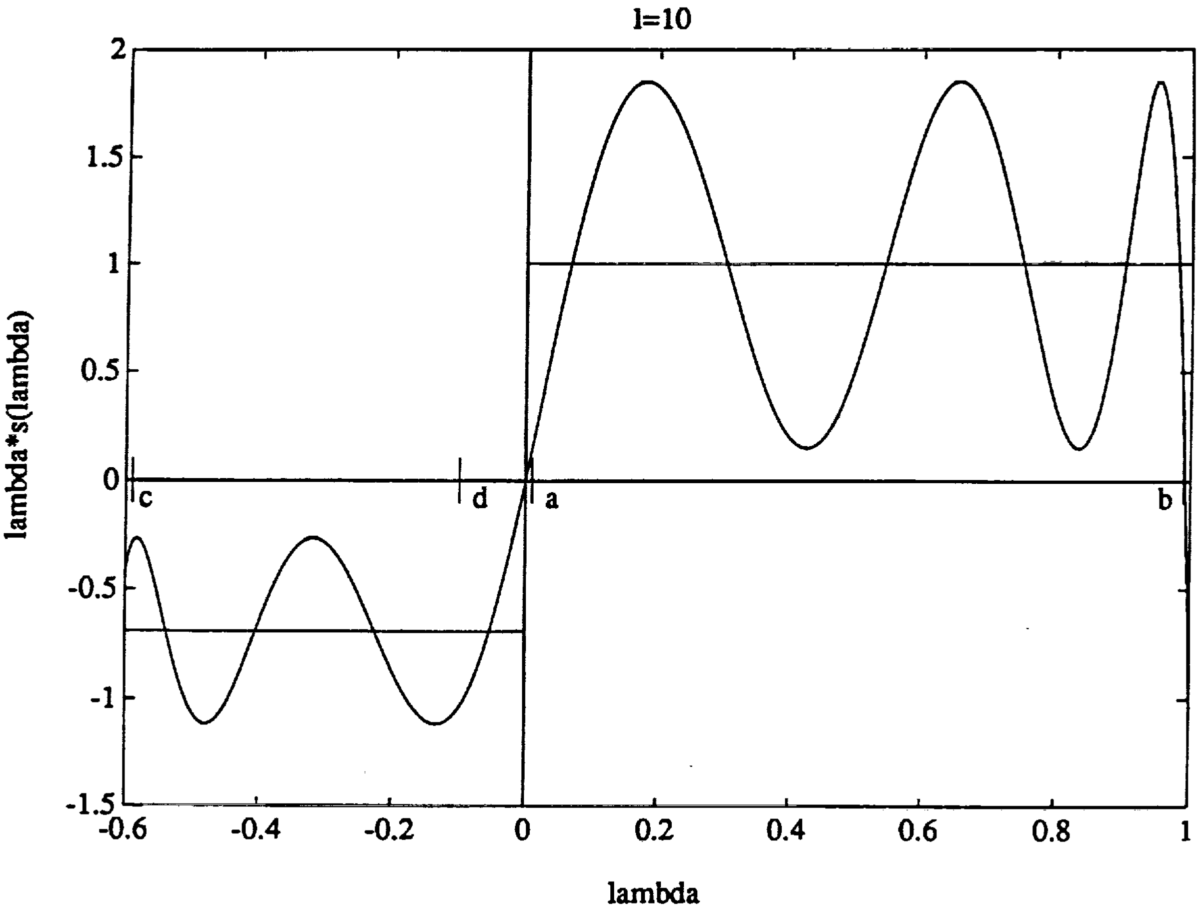
l=10
2
1.5
1
0.5
0
-0.5
-1
-1.5
lambda*s(lambda)
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
lambda
c
d
a
b

Figure 1

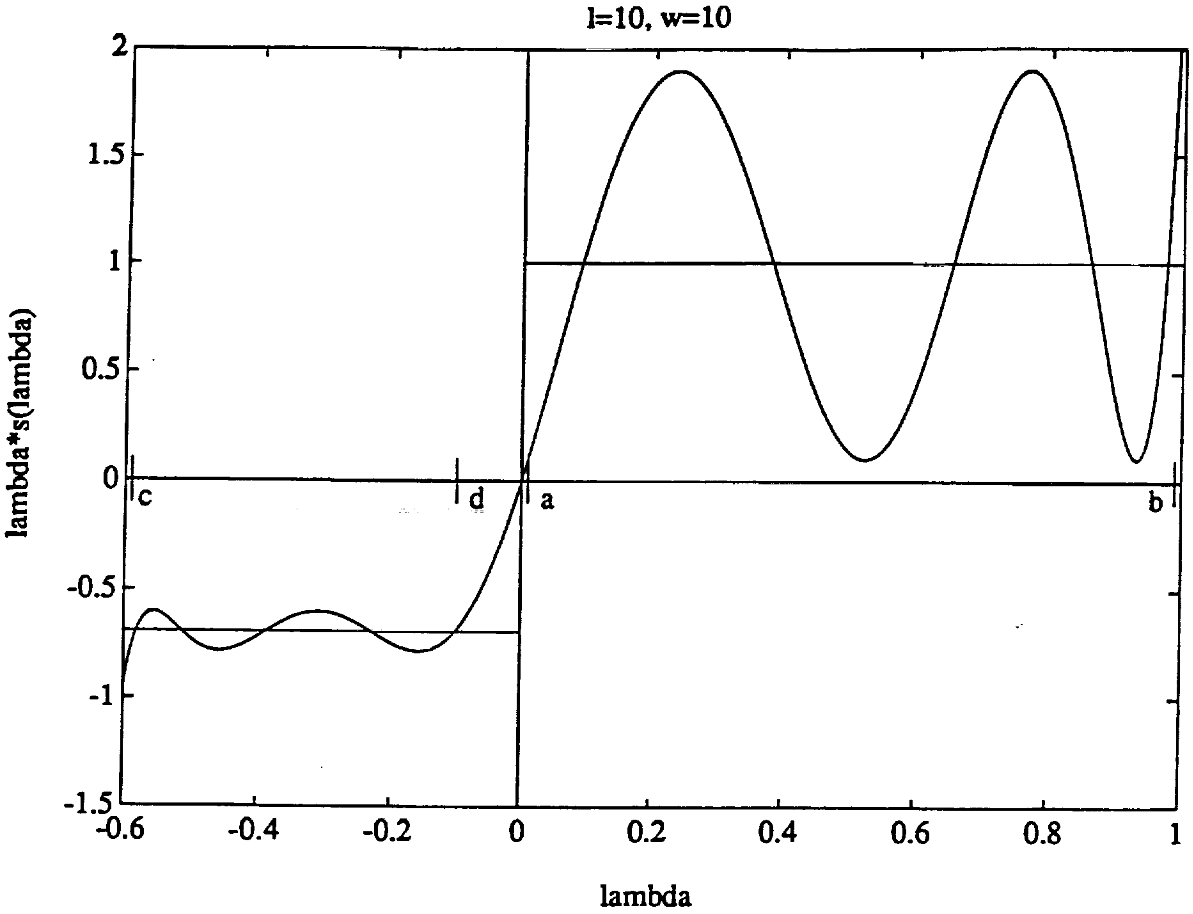
l=10, w=10
2
1.5
1
0.5
0
-0.5
-1
-1.5
lambda*s(lambda)
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
lambda
c
d
a
b

Figure 2

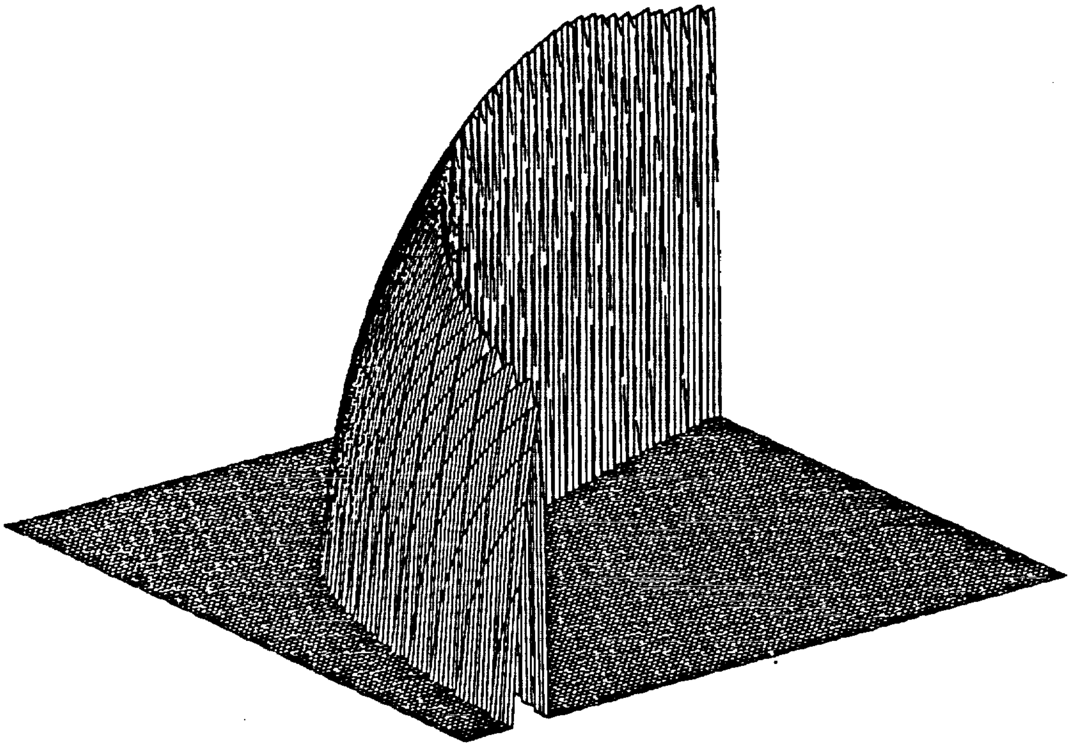

Figure 3

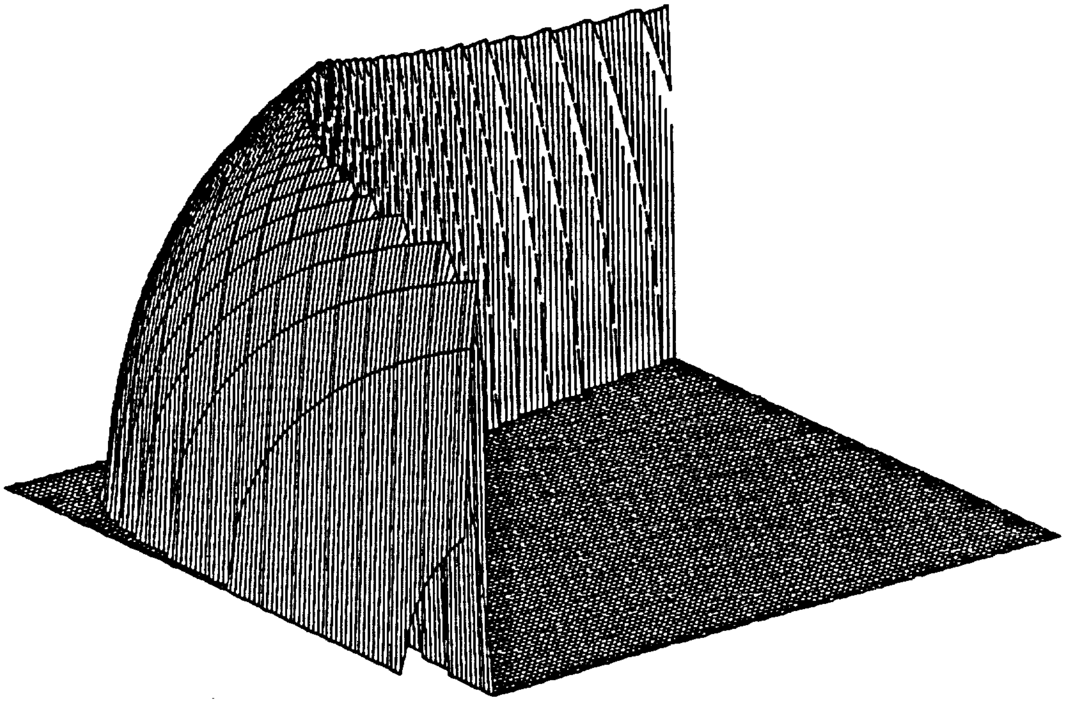

Figure 4