Numer. Math. 64, 213–240 (1993)

Numerische Mathematik

A hybrid Arnoldi-Faber iterative method for nonsymmetric systems of linear equations

Gerhard Starke and Richard S. Varga
Institute for Computational Mathematics, Kent State University, Kent, OH 44242, USA

Received August 14, 1991

Summary. We present here a new hybrid method for the iterative solution of large sparse nonsymmetric systems of linear equations, say of the form $Ax = b$, where $A \in \mathbb{R}^{N,N}$, with A nonsingular, and $b \in \mathbb{R}^N$ are given. This hybrid method begins with a limited number of steps of the Arnoldi method to obtain some information on the location of the spectrum of A, and then switches to a Richardson iterative method based on Faber polynomials. For a polygonal domain, the Faber polynomials can be constructed recursively from the parameters in the Schwarz-Christoffel mapping function. In four specific numerical examples of non-normal matrices, we show that this hybrid algorithm converges quite well and is approximately as fast or faster than the hybrid GMRES or restarted versions of the GMRES algorithm. It is, however, sensitive (as other hybrid methods also are) to the amount of information on the spectrum of A acquired during the first (Arnoldi) phase of this procedure.

Mathematics Subject Classification (1991): 65F10

1. Introduction

In recent years, a number of hybrid algorithms for the iterative solution of a large sparse real nonsymmetric and nonsingular system of linear equations,

$$Ax = b\,,$$

have been proposed. The underlying principle of all these algorithms is to start with an iterative method that requires no a priori information about the coefficient matrix, and then to switch to a parameter-dependent method, after some information (usually about the spectrum of the matrix) is obtained (see [26] for a survey of hybrid algorithms). The information about the location of the eigenvalues is

Correspondence to: R.S. Varga

obtained either by a modified power iteration (cf. e.g. Manteuffel [25]) or by the Arnoldi method (cf. e.g. Elman, Saad and Saylor [10]). In the second phase of the algorithm, Manteuffel as well as Elman, Saad and Saylor use Chebyshev iteration, but there are also other approaches, such as the one proposed by Saad [32] which applies a second-order Richardson iteration based on a polynomial that is optimal in the L^2-sense on some polygon constructed from the eigenvalue estimates. (A similar approach is used by Saylor and Smolarski [34].) The hybrid GMRES algorithm by Nachtigal et al. [26] avoids estimating eigenvalues and instead uses the polynomial implicitly constructed by GMRES for further iterations. Their philosophy in [26] for this approach is that ". . . if a matrix is nonsymmetric (more precisely, non-normal), any attempt to make use of its eigenvalues should be viewed with caution".

In this paper we present, in contrast, a new hybrid method which is directly based on eigenvalue estimates. It starts with a few steps of the Arnoldi method (producing GMRES iterates as a by-product), then constructs a polygon Ω in the complex plane which contains all eigenvalue estimates either as boundary points or as interior points, and then computes the Faber polynomials for this polygonal domain Ω from the Schwarz-Christoffel formula (produced by SCPACK; see [37]) for the conformal map from the exterior of the unit circle in the w-plane onto the exterior of this polygon Ω in the z-plane. The normalized Faber polynomial of some degree m is then applied to the last GMRES iterate, in the form of a Richardson iteration. Thus, in the terminology of [26], our hybrid method has the form

$$\text{Arnoldi/GMRES} \rightarrow \text{eigenvalue estimates} \rightarrow \text{polygon}$$

$$\rightarrow \text{Faber polynomials} \rightarrow \text{Richardson iteration}\ .$$

It can often be observed that the convergence behavior of a polynomial iterative method, when applied to a non-normal matrix A, does not only depend on the size of this polynomial on the spectrum, $\sigma(A)$ (which describes its *asymptotic* convergence rate) but on larger sets containing $\sigma(A)$. Two examples of such sets, which were investigated recently, are the ε-pseudo-spectra introduced by Trefethen in [39] and the field of values (cf. [4]). In this sense, it is important for the performance of our hybrid Arnoldi-Faber algorithm for non-normal matrices that the Faber polynomials are not only "nearly optimal" (in a very strong sense, as we shall see below) on the considered domain Ω itself, but also in neighborhoods of Ω determined from level sets of the conformal map, described above. In other words, the use of normalized Faber polynomials as residual polynomials implies that we take the eigenvalue estimates produced by the Arnoldi method for what they actually are: a very rough prediction of the location of the spectrum.

The idea to use hybrid algorithms arose naturally to overcome problems which were encountered for Krylov subspace methods: Since it is too expensive, in terms of vector operations and vector storage, to run the GMRES algorithm to completion, one has to restart after a certain number of steps (cf. [33]), and this can result in very slow convergence. Thus, it is desirable to acquire as much information about the linear system as possible using a small number of GMRES steps, and then to change to a parameter-dependent method based on this information.

It should be noted that our restriction here to matrices with real entries is basically for purposes of exposition. We could, in principle, also consider complex linear systems, although some of the constructive considerations in Sects. 2 and

4 would necessarily have to be modified. (We will comment on this in later sections and we plan to investigate this in detail in a forthcoming paper.) However, a necessary condition for our method to work is that the origin, $z = 0$, is not contained in the above polygon Ω. For real matrices, this implies the condition that the matrix A is not allowed to have both positive *and* negative real eigenvalues.

Historically, the notion of semi-iterative methods (and, in particular, the Chebyshev semi-iterative method) first appeared in [40] in 1957. Then, the popular cyclic Chebyshev semi-iterative method of Golub and Varga [17] appeared in 1961. The use of complex variable techniques in a semi-iterative setting was given in Kublanovskaja [23] in 1959, and later also in Niethammer and Varga [28] in 1983. The first specific use of Faber polynomials in a semi-iterative setting was studied in Eiermann et al. [7] in 1985. (For a more detailed discussion in 1989, see Eiermann [3].) An approach via minimal norm Nevanlinna-Pick interpolation lead Gutknecht [18] to the use of Faber polynomials as residual polynomials for semi-iterative methods in 1986. Faber polynomials were also used in 1988 by Farkova [13] (in Russian) for the construction of iterative methods (the English translation of this paper appeared in 1990). We note that since the Faber polynomials associated with an interval – and, more generally, ellipses – coincide with the classical Chebyshev polynomials, these polynomial iterations can be viewed as generalizations of the Chebyshev method. In this sense, our hybrid Arnoldi-Faber method can be viewed as a *generalization* of the hybrid Arnoldi-Chebyshev method by Elman et al. [10].

We start in Sect. 2 with an outline of the basic steps of our method. In Sect. 3, we present norm estimates for the Faber polynomials (normalized at the origin) associated with a Jordan domain Ω, of bounded boundary rotation, which does not contain the origin. These bounds, for the maximum norm of Faber polynomials on Ω and on corresponding level sets of the conformal mapping, imply that Faber polynomials are well suited for use as residual polynomials for non-normal matrix iterations. After that, we show in Sect. 4 how to explicitly construct the Faber polynomials for domains bounded by a polygon. Throughout this paper, the steps of the algorithm and the discussion of its effectiveness will be illustrated by the example of solving a linear system with the specific banded Toeplitz matrix

$$A = \begin{bmatrix} 1 & 1 & \frac{1}{2} & & & \\ & 1 & 1 & \frac{1}{2} & & \\ & & 1 & 1 & \frac{1}{2} & \\ & & & 1 & 1 & \frac{1}{2} \\ & & & & \ddots & \ddots & \ddots \end{bmatrix} \in \mathbb{R}^{1024,1024}, \tag{1.1}$$

which was also considered in the numerical experiments in [26]. For this and other examples (taken from [26] and from the discretization of convection-diffusion equations), the last section contains comparisons of the hybrid Arnoldi-Faber method with other hybrid methods (the hybrid GMRES method by Nachtigal et al. [26] and the hybrid Arnoldi-Chebyshev method by Elman et al. [10]), as well as variants of conjugate-gradient-like methods for nonsymmetric systems (restarted GMRES [33], CGS [36] and CGNR, the last method being the CG method applied to the normal equations).

2. Outline of the hybrid Arnoldi-Faber method

We begin this section with a short review of Arnoldi's method for computing estimates for the eigenvalues of a nonsymmetric matrix $A \in \mathbb{R}^{N,N}$ (cf. [1] and [31]). Starting with a vector $v_1 \in \mathbb{R}^N$ with $\|v_1\|_2 = 1$ (usually, this is the normalized initial residual $v_1 := r_0/\|r_0\|_2$ where $r_0 := b - Ax_0$, if considered in conjunction with solving the linear system $Ax = b$), an orthonormal basis for the Krylov subspace $K_m := \text{span}\{v_1, Av_1, \ldots, A^{m-1}v_1\}$ is constructed using the Gram-Schmidt process: For $j = 1, 2, \ldots, m$,

$$\text{(2.1)} \qquad \begin{cases} h_{i,j} := v_i^T A v_j, \quad i = 1, \ldots, j\ , \\ \hat{v}_{j+1} := Av_j - \sum_{i=1}^{j} h_{i,j} v_i\ , \\ h_{j+1,j} := \|\hat{v}_{j+1}\|_2, \qquad v_{j+1} := \hat{v}_{j+1}/h_{j+1,j}\ . \end{cases}$$

In practical implementations, it is usually more suitable to use a modified Gram-Schmidt process (cf. [33]). With $V_m := [v_1, \ldots, v_m]$ (so that $V_m \in \mathbb{R}^{N,m}$) and with $H_m := [h_{i,j}]_{1 \le i,j \le m}$ (so that $H_m \in \mathbb{R}^{m,m}$), (2.1) implies that

$$\text{(2.2)} \qquad AV_m = V_m H_m + \hat{v}_{m+1} e_m^T \qquad \text{and} \qquad V_m^T A V_m = H_m\ ,$$

(where $e_m := (0, \ldots, 0, 1)^T \in \mathbb{R}^m$). The eigenvalues $\{\lambda_i\}_{i=1}^m$ of the Hessenberg matrix H_m are then used as estimates for the eigenvalues of A. Recalling (cf. [21, p. 5]) that the field of values, $W(B)$, of a matrix $B \in \mathbb{C}^{N,N}$ is defined by

$$W(B) := \{x^H Bx : x \in \mathbb{C}^N \text{ and } x^H x = 1\}\ ,$$

we conclude from the second part of (2.2) that $W(H_m) \subseteq W(A)$, and also that $\{\lambda_i\}_{i=1}^m \subseteq W(A)$. For a normal matrix B, it is well-known that the field of values of B is the *convex hull* of its spectrum $\sigma(B)$, i.e. (cf. [21, p. 11]),

$$W(B) = Co(\sigma(B))\ .$$

In particular, if A is normal, then the Arnoldi process produces the eigenvalues $\{\lambda_i\}_{i=1}^m$ of H_m which necessarily lie in $Co(\sigma(A))$. In this case, the numbers $\{\lambda_i\}_{i=1}^m$ may be viewed as eigenvalue estimates of A. For non-normal matrices, $W(A)$ can be significantly *larger* than $Co(\sigma(A))$, which means that the numbers $\{\lambda_i\}_{i=1}^m$ may lie far outside of $Co(\sigma(A))$, as can be observed, for example, in Fig. 9.

Using the extended Hessenberg matrix $\tilde{H}_m := [h_{i,j}] \in \mathbb{R}^{m+1,m}$, the first equation in (2.2) can be written as $AV_m = V_{m+1}\tilde{H}_m$. This gives rise to the GMRES algorithm by Saad and Schultz [33], where the least squares problem, namely,

$$\min_{v \in K_m} \|b - A(x_0 + v)\|_2 = \min_{v \in K_m} \|r_0 - Av\|_2 = \min_{y \in \mathbb{R}^m} \|\beta v_1 - AV_m y\|_2$$

$$= \min_{y \in \mathbb{R}^m} \|\beta V_{m+1} e_1 - V_{m+1}\tilde{H}_m y\|_2 = \min_{y \in \mathbb{R}^m} \|\beta e_1 - \tilde{H}_m y\|_2$$

with $\beta := \|r_0\|_2$, is solved by a QR decomposition of the Hessenberg matrix $\tilde{H}_m$ which is updated from step to step.

We next describe a technique for constructing a polygon Ω in $\mathbb{C}$, symmetric with respect to the real axis, arising from the Arnoldi eigenvalue estimates $\Lambda := \{\lambda_j\}_{j=1}^m$. For this, we assume that all the real eigenvalue estimates in Λ are positive. (If not, i.e., $\Lambda \cap \mathbb{R} \subseteq \{x \in \mathbb{R} : x < 0\}$, then, of course, one has to consider the corresponding linear system with the matrix $-A$.) Define Ω^+ to be the *convex hull* of the set Λ^+ given by

$$\Lambda^+ := \{\lambda_j \in \Lambda : \operatorname{Im}\lambda_j \geqq 0\} \cup \max\{\operatorname{Re}\lambda_j : \lambda_j \in \Lambda\} ,$$

and then define the polygon Ω as the union of Ω^+ and its reflection in the real axis, i.e.,

$$\Omega := \{z \in \mathbb{C} : z \in \Omega^+ \text{ or } \bar{z} \in \Omega^+\} .$$

The construction of Ω^+ begins with

$$\zeta_1 := \max\{\operatorname{Re}\lambda_j : \lambda_j \in \Lambda\} > 0 ,$$

and this is defined to be the first vertex of Ω^+, with $\arg\zeta_1 = 0$. Inductively, having the i-th vertex ζ_i, the next vertex of Ω^+ is the point ζ_{i+1} of Λ, for which the turning angle $\arg(\zeta_{i+1} - \zeta_i) - \arg(\zeta_i - \zeta_{i-1}) \geqq 0$ is smallest (where we define $\arg(\zeta_1 - \zeta_0) := \pi/2$). In this way, Ω^+ is constructed. (If ζ_1 is the only real number contained in Λ^+, then ζ_1 is also the only real vertex of Ω^+, and we obtain a "butterfly-shaped" polygon Ω as in Fig. 9 for Example 6.3. One could also *add* the real number $\min\{\operatorname{Re}\lambda_j : \lambda_j \in \Lambda\}$ to Γ if it is positive. Other possibilities are clearly possible, such as connecting conjugate complex pairs of Λ, having a small imaginary part in modulus, by a vertical line segment to define Ω^+.)

For the example matrix (1.1), the eigenvalue estimates of A after 16 Arnoldi steps and the resulting polygon are shown in Fig. 1. Note that the exact eigenvalue 1 of the matrix in (1.1) (marked by a dot) lies outside the polygon Ω.

If Λ is such that it contains real numbers λ_1 and λ_2 with $\lambda_1 < 0 < \lambda_2$, then Ω necessarily contains the origin, $z = 0$, and in this case, our hybrid Arnoldi-Faber method *cannot* be used (as is similarly the case with many other known iterative methods). We henceforth assume, in what is to follow, that

$$0 \notin \Omega .$$

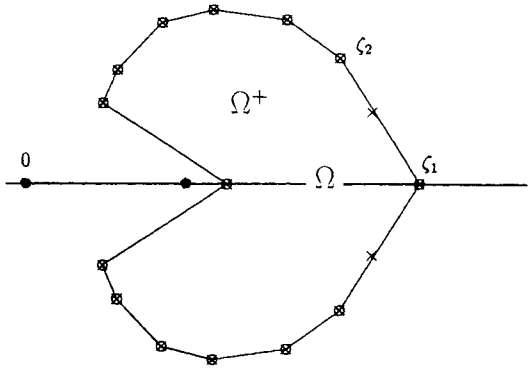

Fig. 1. Eigenvalue estimates after 16 Arnoldi steps

Note that this assumption does not rule out the possibility that $\min\{\mathrm{Re}\lambda_j: 1 \leqq j \leqq m\} < 0 < \max\{\mathrm{Re}\lambda_j: 1 \leqq j \leqq m\}$ (see, e.g., Fig. 9 in Sect. 6).

For the general case where A is a nonsingular *complex* matrix, recall that we have to be able to find a polygon Ω, containing all the Arnoldi eigenvalue estimates $\{\lambda_i\}_{i=1}^m$ but excluding the origin. This is possible, e.g., if there exist $r > 0$ and θ real such that the strip-like set $S(r, \theta)$, defined by

$$S(r, \theta) := \{z \in \mathbb{C}: |z - \tau e^{i\theta}| \leqq r \text{ for all } \tau \geqq 0\} ,$$

does not contain any of the eigenvalue estimates $\{\lambda_i\}_{i=1}^m$. We plan to consider such complex extensions in a subsequent paper.

In the second phase of our hybrid algorithm, we will use an iterative method based on Faber polynomials. How the Faber polynomials for polygonal domains can be constructively determined and why they are very useful for polynomial matrix iterations, will be explained in detail in Sects. 3 and 4.

A large number of the well-known iterative methods, e.g., GMRES, Bi-CG, CGS, Chebyshev and Richardson iteration, etc., for the solution of a linear system $Ax = b$ (where $A \in \mathbb{R}^{N,N}$ is assumed to be nonsingular) can be written as

$$x_m = x_0 + q_{m-1}(A)r_0 , \tag{2.3}$$

where q_{m-1}, a polynomial of degree $m-1$ (written $q_{m-1} \in \Pi_{m-1}$), is called the *iteration polynomial*. If we define the associated error vectors and residuals by $e_m := A^{-1}b - x_m$ and $r_m := b - Ax_m$, respectively, we obtain

$$r_m = b - A(x_0 + q_{m-1}(A)r_0) = (I - Aq_{m-1}(A))r_0 ,$$

and

$$e_m = A^{-1}r_m = A^{-1}(I - Aq_{m-1}(A))r_0 = (I - Aq_{m-1}(A))e_0 .$$

With the *residual polynomial* $p_m(z) := 1 - zq_{m-1}(z)$, so that $p_m(0) = 1$, these two equations above can be written as

$$r_m = p_m(A)r_0 \qquad \text{and} \qquad e_m = p_m(A)e_0 . \tag{2.4}$$

From (2.4), we immediately obtain (for any vector norm on $\mathbb{C}^N$)

$$\|r_m\| \leqq \|p_m(A)\| \, \|r_0\| \qquad \text{and} \qquad \|e_m\| \leqq \|p_m(A)\| \, \|e_0\| \quad (m \geqq 1) ,$$

and thus, our aim is to choose, for any $m \geqq 1$, the polynomial p_m in Π_m such that $\|p_m(A)\|$ is as small as possible among all residual polynomials, i.e., polynomials of degree m with $p_m(0) = 1$.

In an asymptotic sense as $m \to \infty$, the norm of the error reduction is connected only to the modulus of p_m on the spectrum, $\sigma(A)$, of A, in that

$$\overline{\lim_{m \to \infty}} \, \|p_m(A)\|^{1/m} = \overline{\lim_{m \to \infty}} \left(\max_{\lambda \in \sigma(A)} |p_m(\lambda)| \right)^{1/m} \tag{2.5}$$

for any matrix norm (cf. [6]). For special norms (e.g., the spectral norm), rigorous bounds for $\|p_m(A)\|$ can be established in terms of the maximum of p_m on the ε-pseudo-spectrum (cf. [39]) or on the field of values (cf. [4]) of A. Since the polygonal domain Ω is considered as an approximation of the spectrum of A, our aim is to find a polynomial p_m in Π_m with $p_m(0) = 1$ such that $|p_m(z)|$ is small on Ω.

We will see below that Faber polynomials (suitably normalized) possess this property. Moreover, Faber polynomials are also small on neighborhoods of Ω determined by the conformal map from the exterior of the unit circle onto the exterior of this polygon Ω. This is the reason why we expect the hybrid Arnoldi-Faber method to perform well for non-normal matrices. In addition, these Faber polynomials can be constructed quite easily for polygonal domains, provided that the number of vertices of Ω is moderate.

3. Some properties of Faber polynomials

In this paper, we have previously claimed that Faber polynomials have many desirable properties. In this section, we justify these claims.

In what is to follow, Ω will denote a compact set in $\mathbb{C}$ with $0 \notin \Omega$, where it is assumed that Ω is bounded by a Jordan curve. (Specifically, any polygonal Ω with $0 \notin \Omega$, which is created from eigenvalue estimates of the matrix A, is a concrete example of such a compact set!) Then $\bar{\mathbb{C}} \setminus \Omega$, the complement of Ω with respect to the extended complex plane, is simply connected. This implies, by the Riemann Mapping Theorem, that there is a unique conformal mapping $\Phi(z) = w$ from $\bar{\mathbb{C}} \setminus \Omega$ in the z-plane onto the exterior of the unit disk in the w-plane, with $\Phi(\infty) = \infty$ and with $\Phi'(\infty) > 0$. In a neighborhood of ∞, the Laurent expansion of $\Phi(z)$ is

$$\Phi(z) = dz + d_0 + \frac{d_1}{z} + \frac{d_2}{z^2} + \cdots , \tag{3.1}$$

where $c := 1/d > 0$ is the *logarithmic capacity* of Ω. We remark that as $0 \notin \Omega$ and as Φ maps $\bar{\mathbb{C}} \setminus \Omega$ onto $\{w \in \mathbb{C}: |w| > 1\}$, then

$$|\Phi(0)| > 1 , \tag{3.2}$$

which is of importance to us below. The Faber polynomial of degree m, with respect to Ω, is then defined as the *polynomial part* of the Laurent expansion for $(\Phi(z))^m$, i.e., $F_0(z) := 1$ and if

$$(\Phi(z))^m = \sum_{j=-m}^{\infty} d_{j,m} z^{-j} ,$$

then

$$(\Phi(z))^m =: F_m(z) + \sum_{j=1}^{\infty} d_{j,m} z^{-j} \quad (m \geqq 1) \tag{3.3}$$

(cf. [12] or [15, Ch. I, Sect. 6]). The (unique) inverse mapping of Φ will be denoted by Ψ, i.e., Ψ maps $\{w \in \mathbb{C}: |w| > 1\}$ conformally onto $\bar{\mathbb{C}} \setminus \Omega$. The Laurent expansion of $\Psi(w)$,

$$\Psi(w) = cw + c_0 + \frac{c_1}{w} + \frac{c_2}{w^2} + \cdots , \tag{3.4}$$

is then valid for all $|w| > 1$. With the coefficients $c, c_0, c_1, \ldots$ in (3.4), the Faber polynomials can then be directly computed recursively from

$$F_0(z) = 1, \qquad F_1(z) = (z - c_0)/c , \tag{3.5}$$

$$F_{m+1}(z) = [zF_m(z) - (c_0 F_m(z) + \cdots + c_m F_0(z)) - mc_m]/c \quad (m \geqq 1) ,$$

(cf. [20, p. 512]).

Since the boundary of Ω is assumed to be a Jordan curve, it is known that $\Psi(w)$ has a continuous extension from $\{w \in \mathbb{C}: |w| > 1\}$ to $\{w \in \mathbb{C}: |w| \geqq 1\}$. If, in addition, the boundary curve of Ω is rectifiable, the *total boundary rotation* of Ω is defined by

$$V(\Omega) := \int_0^{2\pi} |d_t \arg(\Psi(e^{it}) - \Psi(e^{i\theta}))| , \tag{3.6}$$

where $\theta \in [0, 2\pi)$. We will only consider domains Ω for which $V(\Omega)$ is finite, and such domains are said to be of *bounded boundary rotation*. If Ω is convex, then $V(\Omega) = 2\pi$. For a domain bounded by a polygon, $V(\Omega)$ is just the sum of the absolute values of the turning angles (as defined in Sect. 2), so that *any* polygonal Ω is necessarily of bounded boundary rotation. (For a detailed discussion of the above concepts, we recommend Gaier [15, Ch. I, Sect. 6].)

Our aim is to find a polynomial p_m in Π_m with $p_m(0) = 1$ which is uniformly small on Ω, i.e., we wish to solve the minimization problem

$$\gamma_m(\Omega) := \inf \left\{ \max_{z \in \Omega} |p_m(z)| : p_m \in \Pi_m \text{ with } p_m(0) = 1 \right\} \quad (m \geqq 0) . \tag{3.7}$$

It is known (cf. [30, Theorem 3]) that there always exists a $\hat{p}_m(z) \in \Pi_m$ with $\hat{p}_m(0) = 1$ for which

$$\gamma_m(\Omega) = \max_{z \in \Omega} |\hat{p}_m(z)| \quad (m \geqq 0) . \tag{3.8}$$

While the existence of these polynomials $\{\hat{p}_m(z)\}_{m=0}^{\infty}$, which satisfy (3.8), is guaranteed for general Ω, the exact solution $\hat{p}_m$ for *each* $m \geqq 0$ of (3.8) is only known in rather special cases of Ω (such as circles, ellipses, etc.). However, it is sufficient for our purposes to find polynomials which are "nearly optimal" for (3.7). The result of the following theorem is that, at least for convex Ω, the suitably normalized Faber polynomials possess this property.

Theorem 3.1. *Let Ω be convex with $0 \notin \Omega$. Then, the normalized Faber polynomials $\tilde{F}_m(z) := F_m(z)/F_m(0)$, associated with Ω, satisfy*

$$\gamma_m(\Omega) \leqq \max_{z \in \Omega} |\tilde{F}_m(z)| < \frac{2}{|\Phi(0)|^m - 1} \leqq \frac{2}{1 - \dfrac{1}{|\Phi(0)|^m}} \gamma_m(\Omega) , \tag{3.9}$$

for all $m \geqq 1$.

Proof. From Pommerenke [29, Lemma 1], the Faber polynomials $F_m(z)$, with respect to Ω, can be represented as

$$F_m(\Psi(e^{i\theta})) = \frac{1}{\pi} \int_0^{2\pi} e^{imt} d_t \arg(\Psi(e^{it}) - \Psi(e^{i\theta})) .$$

Since $\Psi(e^{i\theta}) \in \partial\Omega$ for each $0 \leqq \theta < 2\pi$, we see, on using the maximum principle for analytic functions, that

$$\max_{z \in \Omega} |F_m(z)| \leqq \frac{1}{\pi} \int_0^{2\pi} |d_t \arg(\Psi(e^{it}) - \Psi(e^{i\theta}))| = \frac{V(\Omega)}{\pi} = 2 , \tag{3.10}$$

the last inequality following from the hypothesis that Ω is convex. Moreover, it is known from Kövari and Pommerenke [22, Theorem 2] that

$$|F_m(\Psi(w)) - w^m| < 1 \quad \text{for } |w| \geqq 1 \text{ and } m \geqq 0\,, \tag{3.11}$$

provided that the convex set Ω is not an interval. In the case that Ω is an interval, it is known (cf. [35, Sect. 2.1.1]) that $F_m(\Psi(w)) = w^m + w^{-m}$ for all $m \geqq 1$; whence,

$$|F_m(\Psi(w)) - w^m| = |(w^m + w^{-m}) - w^m| = |w|^{-m} < 1 \tag{3.12}$$

$$\text{for } |w| > 1 \text{ and } m \geqq 1\,,$$

which is similar to the result of (3.11). The reverse triangle inequality applied to (3.11) and (3.12), respectively, gives

$$|F_m(\Psi(w))| > |w|^m - 1 \quad \text{for } |w| > 1 \text{ and } m \geqq 0\,,$$

and

$$|F_m(\Psi(w))| > |w|^m - |w|^{-m} \quad \text{for } |w| > 1 \text{ and } m \geqq 1\,.$$

On choosing $\hat{w} := \Phi(0)$, so that (cf. (3.2)) $|\hat{w}| > 1$, and on using the fact that $\Psi(\Phi(z)) \equiv z$ for all $z \in \bar{\mathbb{C}} \setminus \Omega$, the above inequalities reduce to

$$|F_m(0)| > |\Phi(0)|^m - 1 \quad (m \geqq 0)\,, \tag{3.13}$$

and

$$|F_m(0)| > |\Phi(0)|^m - \frac{1}{|\Phi(0)|^m} > |\Phi(0)|^m - 1 \quad (m \geqq 1)\,, \tag{3.14}$$

which also shows (since $F_0(z) \equiv 1$) that $F_m(0) \neq 0$ for any $m \geqq 0$. Since by definition $\tilde{F}_m(z) := F_m(z)/F_m(0)$, it directly follows from (3.7) and (3.10)–(3.14) that the first and second inequalities of (3.9) of Theorem 3.1 are thus valid.

For the final inequality of (3.9), it suffices to show that

$$\gamma_m(\Omega) \geqq \frac{1}{|\Phi(0)|^m} \quad (m \geqq 1)\,. \tag{3.15}$$

This is a direct consequence of Bernstein's Lemma (cf. [41, Sect. 4.6]) as pointed out in Eiermann et al. [5, p. 160] in a different normalization (i.e., $p_m(1) = 1$ instead of $p_m(0) = 1$), in fact for more general domains Ω from the class $\mathcal{M}$, where

$$\mathcal{M} := \{\Omega \subseteq \mathbb{C}\colon \Omega \text{ is compact, } 0 \notin \Omega\,, \tag{3.16}$$
$$\Omega \text{ has no isolated points, and } \bar{\mathbb{C}} \setminus \Omega \text{ is of finite connectivity}\}\,.$$

A different proof of (3.15) was given by Gutknecht [18] who considered our minimization problem (3.7) in the larger space $\mathcal{P}_m(\Omega)$, consisting of all functions which are analytic and bounded on compact subsets of $(\bar{\mathbb{C}} \setminus \Omega) \cup \partial\Omega$ and which have a pole of order at most m at the point at infinity. Obviously, Π_m is a subspace of $\mathcal{P}_m(\Omega)$, and we have

$$\gamma_m(\Omega) \geqq \min\left\{\max_{z \in \partial\Omega} |P_m(z)|\colon P_m \in \mathcal{P}_m(\Omega),\ P_m(0) = 1\right\} =: \hat{\gamma}_m(\Omega)\,. \tag{3.17}$$

The solution of the above minimax problem is given by

$$\hat{\gamma}_m(\Omega) = \max_{z \in \partial\Omega} |\tilde{P}(z)| ,$$

where

$$\tilde{P}_m(z) := (\Phi(z)/\Phi(0))^m \quad (m \geqq 1) .$$

Since $|\Phi(z)| = 1$ for any $z \in \partial\Omega$, this yields

$$\gamma_m(\Omega) \geqq \hat{\gamma}_m(\Omega) = \frac{1}{|\Phi(0)|^m} \quad (m \geqq 1) , \tag{3.18}$$

which establishes the first inequality of (3.9). We remark that the statement that $\tilde{P}_m(z) := (\Phi(z)/\Phi(0))^m$ is a solution of the minimax problem (3.17) can be shown by reducing the minimization problem in $\mathscr{P}_m(\Omega)$ to a Carathéodory-Fejér problem [2] (see also [8]). In principle, this could also be shown directly by adapting Trefethen's proof of the near-circularity criterion [38, Theorem 1], which is based on Rouché's Theorem. □

Theorem 3.1 states, since $|\Phi(0)| > 1$ from (3.2), that the uniform norms on Ω of the normalized Faber polynomials F_m are, as $m \to \infty$, only about twice as large as the exact solutions of the minimax problem of (3.7) on Ω when Ω is convex. However, Ω is only an estimate for the spectrum, $\sigma(A)$, of the matrix A, and the convergence behavior of polynomial iterations for non-normal matrices does not depend only on the size of the polynomial on the spectrum itself but also on its size in neighborhoods of $\sigma(A)$, i.e., the ε-pseudo-spectrum (cf. [39]) or the field of values (cf. [4]) of A. Therefore, it is important that the polynomials $\tilde{F}_m$ are also nearly optimal (in the same sense as above) in neighborhoods of Ω, i.e., on the sets Ω_ρ bounded by the *level curves*

$$\Gamma_\rho := \{z = \Psi(w) : |w| = \rho\} \quad (\rho \geqq 1) \tag{3.19}$$

of the conformal map Ψ. The (compact) set Ω_ρ, defined as the set of all $z \in \mathbb{C}$ which lie interior or on Γ_ρ, is called the *level set* for ρ, where $\rho \geqq 1$.

The following lemma, concerning convexity of the *level sets* Ω_ρ, is useful for our subsequent considerations.

Lemma 3.2. *For each Ω, there is a unique real number $\hat{\rho}$, with $1 \leqq \hat{\rho} < \infty$, such that Ω_ρ is convex for all $\rho \geqq \hat{\rho}$, and Ω_ρ is not convex for any $\rho < \hat{\rho}$.*

Proof. For a fixed $\rho \geqq 1$, we first show that Ω_ρ is convex if and only if

$$\operatorname{Re}\left(1 + w\frac{\psi''(w)}{\psi'(w)}\right) \geqq 0 \quad \text{for all } |w| = \rho . \tag{3.20}$$

This is very similar to the condition for a univalent function to map the unit disk onto a convex set (cf. [27, p. 224]).

For the proof of (3.20), we note that Ω_ρ is convex if and only if the tangent direction

$$\arg\left(\frac{d}{d\vartheta}\Psi(\rho e^{i\vartheta})\Big|_{\vartheta=\Theta}\right) = \arg\left(i\rho e^{i\Theta}\Psi'(\rho e^{i\Theta})\right)$$

$$= \arg\left(e^{i\Theta}\Psi'(\rho e^{i\Theta})\right) + \frac{\pi}{2} = \operatorname{Im}\left(\ln(e^{i\Theta}\Psi'(\rho e^{i\Theta}))\right) + \frac{\pi}{2}$$

is non-decreasing as a function of Θ, for $0 \leqq \Theta \leqq 2\pi$. This is equivalent to

$$0 \leqq \operatorname{Im} \frac{d}{d\Theta} \ln(e^{i\Theta}\Psi'(\rho e^{i\Theta})) = \operatorname{Im} \frac{ie^{i\Theta}(\Psi'(\rho e^{i\Theta}) + \rho e^{i\Theta}\Psi''(\rho e^{i\Theta})}{e^{i\Theta}\Psi'(\rho e^{i\Theta})}$$

$$= \operatorname{Re}\left(1 + \rho e^{i\Theta} \frac{\Psi''(\rho e^{i\Theta})}{\Psi'(\rho e^{i\Theta})}\right), \quad \text{for all } 0 \leqq \Theta \leqq 2\pi,$$

which is (3.20). Next, it follows from (3.4) that

$$1 + w\frac{\Psi''(w)}{\Psi'(w)} = 1 + O\left(\frac{1}{w^2}\right) \quad \text{as } |w| \to \infty, \tag{3.21}$$

so that Ω_ρ is necessarily convex for every $\rho \geqq 1$ sufficiently large. It thus suffices to show that if $\Omega_{\tilde\rho}$ is convex for some $\tilde\rho \geqq 1$, then so is Ω_ρ for any $\rho \geqq \tilde\rho$. Assuming that $\Omega_{\tilde\rho}$ is convex, (3.20) implies that

$$\min_{|w|=\tilde\rho} \operatorname{Re}\left(1 + w\frac{\Psi''(w)}{\Psi'(w)}\right) \geqq 0.$$

From (3.21), it follows from the minimum principle for harmonic functions that

$$\min_{|w|\geqq\tilde\rho} \operatorname{Re}\left(1 + w\frac{\Psi''(w)}{\Psi'(w)}\right) \geqq 0.$$

Then, for any $\rho \geqq \tilde\rho$, it follows that $\bar{\mathbb{C}}\backslash\Omega_\rho \subseteq \bar{\mathbb{C}}\backslash\Omega_{\tilde\rho}$, and the minimum principle for harmonic functions again gives

$$0 \leqq \min_{|w|\geqq\tilde\rho} \operatorname{Re}\left(1 + w\frac{\Psi''(w)}{\Psi'(w)}\right) \leqq \min_{|w|\geqq\rho} \operatorname{Re}\left(1 + w\frac{\Psi''(w)}{\Psi'(w)}\right)$$

$$\leqq \min_{|w|=\rho} \operatorname{Re}\left(1 + w\frac{\Psi''(w)}{\Psi'(w)}\right),$$

so that Ω_ρ is also convex. □

Theorem 3.3. *For some $\tilde\rho$ with $1 \leqq \tilde\rho < |\Phi(0)|$, let $\Omega_{\tilde\rho}$ be convex. Then, the normalized Faber polynomials $\tilde F_m(z)$, associated with Ω, satisfy*

$$\gamma_m(\Omega_\rho) \leqq \max_{z\in\Omega_\rho} |\tilde F_m(z)| < \frac{2\rho^m}{|\Phi(0)|^m - \rho^m} \leqq \frac{2}{1 - \left(\frac{\rho}{|\Phi(0)|}\right)^m} \gamma_m(\Omega_\rho), \tag{3.22}$$

for all $\tilde\rho \leqq \rho < |\Phi(0)|$ and all $m \geqq 1$.

Proof. The proof is very similar to that of Theorem 3.1. Fix a ρ in the interval $[\tilde\rho, |\Phi(0)|)$. The conformal map Φ_ρ from $\bar{\mathbb{C}}\backslash\Omega_\rho$ onto $\{w\in\mathbb{C}: |w| > 1\}$, with $\Phi_\rho(\infty) = \infty$ and $\Phi_\rho'(\infty) > 0$, is given by $\Phi_\rho(z) = \Phi(z)/\rho$. Thus, the Faber polynomials $F_{m,\rho}$, associated with Ω_ρ, are given by $F_{m,\rho}(z) = F_m(z)/\rho^m$. Since $\Omega_{\tilde\rho}$ is convex, Lemma 3.2 tells us that Ω_ρ also is and we obtain, in analogy to (3.10) and (3.13),

$$\max_{z\in\Omega_\rho} |F_m(z)| \leqq 2\rho^m \tag{3.23}$$

and

$$|F_m(0)| > |\Phi(0)|^m - \rho^m \,. \tag{3.24}$$

Using the same idea as in the proof of Theorem 3.1, it follows that

$$\gamma_m(\Omega_\rho) \geqq \left(\frac{\rho}{|\Phi(0)|}\right)^m , \tag{3.25}$$

which implies the second part of (3.22). □

Example 3.1. Consider the special case where the compact set Ω (with more than one point) is given by an ellipse (or an interval). It is well-known that, under these circumstances, the conformal mapping Ψ from $\{w \in \mathbb{C}: |w| > 1\}$ onto $\bar{\mathbb{C}} \setminus \Omega$ has the form

$$\Psi(w) = cw + c_0 + \frac{c_1}{w} \quad (c \cdot c_1 \neq 0) \,, \tag{3.26}$$

where the foci of the ellipse are given by $c_0 \pm 2\sqrt{cc_1}$. Thus, the recursive definition (3.5) for the Faber polynomials gives

$$F_0(z) = 1, \qquad F_1(z) = (z - c_0)/c \,,$$

$$F_2(z) = (z - c_0)^2/c^2 - 2c_1/c \,,$$

$$F_{m+1}(z) = [(z - c_0)F_m(z) - c_1 F_{m-1}(z))]/c \quad (m \geqq 2) \,.$$

Comparing this to the well-known recurrence relation for the Chebyshev polynomials, namely

$$T_0(z) = 1, \qquad T_1(z) = z \,,$$

$$T_{m+1}(z) = 2zT_m(z) - T_{m-1}(z), \quad (m \geqq 1) \,,$$

we obtain

$$F_m(z) = 2\left(\frac{c_1}{c}\right)^{m/2} T_m\left(\frac{z - c_0}{2\sqrt{cc_1}}\right) \quad (m \geqq 1) \,. \tag{3.27}$$

This means that the Faber polynomials for ellipses coincide with (properly scaled) Chebyshev polynomials and, consequently, matrix iterations based on Faber polynomials can be viewed as generalizations of the Chebyshev method. In this sense, our hybrid Arnoldi-Faber method can be viewed as a generalization of the hybrid Arnoldi-Chebyshev method by Elman et al. [10].

It is known for ellipses that the associated Faber polynomials are not only nearly optimal but that they actually solve the minimization problem (3.7) in many cases. In particular, when Ω is a real interval, the solutions for (3.7) are given by scaled Chebyshev polynomials. The same is true if, roughly speaking, the origin is not too close to the ellipse (cf. [14]).

One of the big advantages of using Faber polynomials is that, for the corresponding polynomial iterative method to converge, it is *not* necessary for the origin to be outside the convex hull of the spectrum (as is necessary, for example, for the

Chebyshev method). It is therefore useful to extend the results of Theorems 3.1 and 3.3 to non-convex regions.

Theorem 3.4. *Let Ω be of bounded boundary rotation with $0 \notin \Omega$, and let $V(\Omega)$ denote the total boundary rotation (cf. (3.6)) of Ω. Then, the normalized Faber polynomials, associated with Ω, satisfy*

$$\gamma_m(\Omega) \leqq \max_{z \in \Omega} |\tilde{F}_m(z)| \leqq \frac{V(\Omega)/\pi}{|\Phi(0)|^m - (1 + V(\Omega)/\pi)} \tag{3.28}$$

$$\leqq \frac{V(\Omega)/\pi}{1 - \dfrac{(1 + V(\Omega)/\pi)}{|\Phi(0)|^m}} \gamma_m(\Omega)$$

for all $m > \ln\left(1 + \dfrac{V(\Omega)}{\pi}\right) \Big/ (\ln|\Phi(0)|)$. *Moreover, if* $1 \leqq \rho < |\Phi(0)|$, *then*

$$\gamma_m(\Omega_\rho) \leqq \max_{z \in \Omega_\rho} |\tilde{F}_m(z)| \leqq \frac{\dfrac{V(\Omega_\rho)}{\pi}}{1 - \left(1 + \dfrac{V(\Omega_\rho)}{\pi}\right)\left(\dfrac{\rho}{|\Phi(0)|}\right)^m} \gamma_m(\Omega_\rho) \tag{3.29}$$

is valid for all $m > \ln\left(1 + \dfrac{V(\Omega)}{\pi}\right) \Big/ (\ln|\Phi(0)| - \ln\rho)$.

Proof. The proof of Theorem 3.1 can be adapted to prove (3.28) if we replace (3.10) and (3.13) by

$$\max_{z \in \Omega} |F_m(z)| \leqq \frac{V(\Omega)}{\pi}$$

and

$$|F_m(0)| \geqq |\Phi(0)|^m - \left(1 + \frac{V(\Omega)}{\pi}\right),$$

the second of these inequalities following from the (relatively crude) estimate

$$\max_{z \in \bar{\mathbb{C}} \setminus \Omega} |F_m(z) - (\Phi(z))^m| \leqq \max_{z \in \partial\Omega} |F_m(z) - (\Phi(z))^m|$$

$$\leqq \max_{z \in \partial\Omega} (|F_m(z)| + |\Phi(z)|^m) \leqq \frac{V(\Omega)}{\pi} + 1 .$$

In the same manner, the proof of Theorem 3.3 using

$$\max_{z \in \Omega_\rho} |F_m(z)| \leqq \frac{V(\Omega_\rho)}{\pi} \rho^m$$

and

$$|F_m(0)| \geqq |\Phi(0)|^m - \left(1 + \frac{V(\Omega_\rho)}{\pi}\right)$$

in place of (3.23) and (3.24), implies (3.29). The restriction in m in the statement of Theorem 3.4 is necessary to ensure that the denominator of the fraction in (3.28) is positive. □

The inequalities of Theorem 3.4 are not as sharp as those of Theorems 3.1 and 3.3. For example, (3.28) gives

$$\gamma_m(\Omega) \leqq \max_{z \in \Omega} |\tilde{F}_m(z)| \leqq \frac{2}{1 - \frac{3}{|\Phi(0)|^m}} \gamma_m(\Omega)$$

for convex Ω, while (3.9) gives

$$\gamma_m(\Omega) \leqq \max_{z \in \Omega} |\tilde{F}_m(z)| \leqq \frac{2}{1 - \frac{1}{|\Phi(0)|^m}} \gamma_m(\Omega) .$$

However, both bounds state, for large m, that the normalized Faber polynomials for Ω are only off by a factor of about $V(\Omega)/\pi$ from the minimum, $\gamma_m(\Omega)$, in (3.7).

Recalling our technique from Sect. 2 for constructing a polygon Ω which is symmetric with respect to the real axis, one can show by geometric considerations that

$$V(\Omega) < 4\pi$$

in all cases. Moreover, since $V(\Omega)$ is the sum of the absolute values of the turning angles (as defined in Sect. 2), we note that $V(\Omega)$ can be explicitly calculated in the course of determining the polygon Ω.

The theoretical significance of Theorems 3.1, 3.3 and 3.4 can be utilized as follows. Suppose that the compact set Ω, having bounded boundary rotation with $0 \notin \Omega$, is such that there is a ρ, with $1 \leqq \rho < |\Phi(0)|$ such that the level set Ω_ρ contains the spectrum, $\sigma(A)$, of the given matrix A. Then, the normalized Faber polynomials $F_m(z)$, when used as residual polynomials (cf. (2.4)), are necessarily convergent from (2.5). This will be used in the numerical results of Sect. 6.

4. The construction of Faber polynomials for polygons

Our main motivation for using Faber polynomials as residual polynomials in the second phase of our iteration procedure is that they can be constructively computed easily from the geometrical parameters of Ω, i.e., from the location of the edges and turning angles of the polygon Ω. To be precise, we compute the parameters in the Schwarz-Christoffel representation of the conformal map Ψ from the exterior of the unit circle in the w-plane onto $\bar{\mathbb{C}} \backslash \Omega$ in the z-plane, using Trefethen's SCPACK [37], and then we compute the Faber polynomials recursively from these parameters.

Since we essentially consider real matrices A here, the polygonal domain Ω is symmetric with respect to the real line. Let $\zeta_1, \zeta_2, \ldots, \zeta_p$ denote the vertices of the polygon in the upper half-plane (in counter-clockwise order, where we assume that ζ_1 and ζ_p are real), and let $\beta_1\pi, \beta_2\pi, \ldots, \beta_p\pi$ denote the corresponding exterior turning angles (see the polygon on the right in Fig. 2).

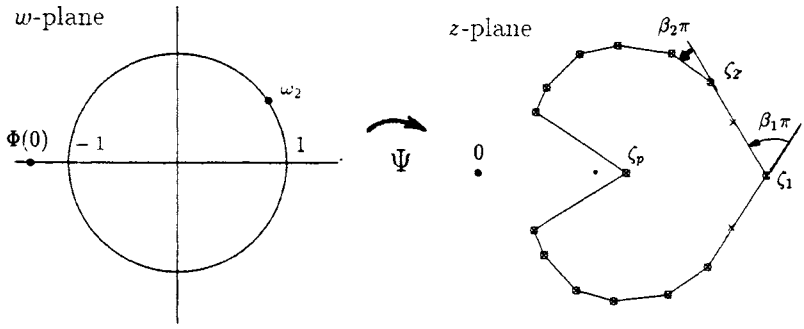

Fig. 2. Ψ maps the exterior of the unit circle onto $\bar{\mathbb{C}}\setminus\Omega$

The conformal map Ψ, mapping the exterior of the unit circle onto $\bar{\mathbb{C}}\setminus\Omega$ and satisfying $\Psi(\infty)=\infty$ and $\Psi'(\infty)>0$, is given by the Schwarz-Christoffel formula

$$\Psi(w)=\Psi(w_0)+\int_{w_0}^{w}\Psi'(\omega)d\omega\,,$$

with

$$\Psi'(w)=c\left(1-\frac{1}{w}\right)^{\beta_1}\prod_{j=2}^{p-1}\left(1-2\frac{\mathrm{Re}\,\omega_j}{w}+\frac{1}{w^2}\right)^{\beta_j}\left(1+\frac{1}{w}\right)^{\beta_p} \tag{4.1}$$

(cf. [19, Sect. 5.12]), where the points $\{\omega_j\}_{j=2}^{p-1}$ on the upper half of the unit circle are the pre-images of the vertices of Ω in the upper half-plane, i.e., $\Psi(\omega_j)=\zeta_j$, $j=2,\ldots,p-1$ and $\Psi(1)=\zeta_1$, $\Psi(-1)=\zeta_p$. The parameters $\omega_2,\ldots,\omega_{p-1}$ and c (where c is the capacity of Ω from (3.4)) are unknown and have to be computed numerically. Usually, SCPACK can only handle *interior* mapping functions, i.e., the Schwarz-Christoffel transformation is from the open unit disk onto the interior of a (possibly unbounded) polygon. The package can be modified so that it computes *exterior* mapping functions as indicated in Trefethen [37, p. 101] and also Ellacott [9, Sect. 3.2], but for polygons symmetric with respect to the real axis, we prefer to use the following "trick" used by Li ([24, Sect. 4.2]) to obtain Ψ from an interior mapping function.

Let $\tilde{\Psi}$ denote the ("interior") conformal map from the open unit disk onto the unbounded region in the upper half-plane bounded by the polygon with vertices $\zeta_1,\zeta_2,\ldots,\zeta_p,\infty$ such that $\tilde{\Psi}(0)=\zeta_0$ with an arbitrarily chosen $\zeta_0\in\{z\in\mathbb{C}:\mathrm{Im}\,z>0\}\setminus\Omega$ and $\tilde{\Psi}'(0)>0$. This function is given by

$$\tilde{\Psi}(\tilde{w})=\zeta_0+\int_0^{\tilde{w}}\tilde{\Psi}'(\omega)d\omega$$

with

$$\tilde{\Psi}'(\tilde{w})=\tilde{c}\left(1-\frac{\tilde{w}}{\tilde{\omega}_1}\right)^{(\beta_1-1)/2}\left(\prod_{j=2}^{p-1}\left(1-\frac{\tilde{w}}{\tilde{\omega}_j}\right)^{\beta_j}\right)\times\left(1-\frac{\tilde{w}}{\tilde{\omega}_p}\right)^{(\beta_p-1)/2}\left(1-\frac{\tilde{w}}{\tilde{\omega}_{p+1}}\right)^{-2}, \tag{4.2}$$

where SCPACK numerically determines the parameters $\tilde{\omega}_1, \ldots, \tilde{\omega}_{p+1}$ and $\tilde{c}$. Note that $\tilde{\Psi}(\tilde{\omega}_j) = \zeta_j$ for $j = 1, \ldots, p$ and $\tilde{\Psi}(\tilde{\omega}_{p+1}) = \infty$. The rational transformation

$$\tilde{w} = g(w) := \frac{\tilde{\omega}_1(\tilde{\omega}_{p+1} - \tilde{\omega}_p)(w+1)^2 + \tilde{\omega}_p(\tilde{\omega}_1 - \tilde{\omega}_{p+1})(w-1)^2}{(\tilde{\omega}_{p+1} - \tilde{\omega}_p)(w+1)^2 + (\tilde{\omega}_1 - \tilde{\omega}_{p+1})(w-1)^2} \tag{4.3}$$

maps $\{w \in \mathbb{C}: \operatorname{Im} w > 0, |w| > 1\}$ conformally onto $\{\tilde{w} \in \mathbb{C}: |\tilde{w}| < 1\}$ in such a way that $g(1) = \tilde{\omega}_1$, $g(-1) = \tilde{\omega}_p$ and $g(\infty) = \tilde{\omega}_{p+1}$. This can be shown in the following way. First, $v = -1/w$ maps $\{w \in \mathbb{C}: \operatorname{Im} w > 0, |w| > 1\}$ conformally onto $\{v \in \mathbb{C}: \operatorname{Im} v > 0, |v| < 1\}$, then this domain is mapped onto the first quadrant by $\hat{v} = (1+v)/(1-v)$. By $\tilde{v} = \hat{v}^2$, we have a conformal mapping from the first quadrant onto the upper half-plane, and, finally, the upper half-plane is mapped onto the unit circle by $\hat{w} = (\mathrm{i} - \tilde{v})/(\mathrm{i} + \tilde{v})$. Thus, the composition of these four conformal transformations,

$$\hat{w} = \tilde{g}(w) = \frac{\mathrm{i}(w+1)^2 - (w-1)^2}{\mathrm{i}(w+1)^2 + (w-1)^2},$$

maps $\{w \in \mathbb{C}: \operatorname{Im} w > 0, |w| > 1\}$ conformally onto $\hat{w} \in \mathbb{C}: |\hat{w}| < 1\}$ (see also [24, Proof of Theorem 12]). However, we have $\tilde{g}(1) = 1$, $\tilde{g}(-1) = -1$ and $\tilde{g}(\infty) = \mathrm{i}$, so we have to apply a mapping from the open unit disk onto itself such that the points 1, -1 and i are mapped onto $\tilde{\omega}_1$, $\tilde{\omega}_p$ and $\tilde{\omega}_{p+1}$, respectively. This leads to the mapping function g of (4.3).

By the reflection principle (cf. [19, pp. 389]), $\tilde{\Psi}(g(w))$ conformally maps the exterior of the unit circle onto $\bar{\mathbb{C}} \backslash \Omega$. Thus, $\Psi(w) = \tilde{\Psi}(g(w))$. Solving (4.3) with respect to w leads to

$$w + \frac{1}{w} = 2\,\frac{(\tilde{\omega}_1 + \tilde{\omega}_p - 2\tilde{\omega}_{p+1})\tilde{\omega} + \tilde{\omega}_1\tilde{\omega}_{p+1} + \tilde{\omega}_p\tilde{\omega}_{p+1} - 2\tilde{\omega}_1\tilde{\omega}_p}{(\tilde{\omega}_1 - \tilde{\omega}_p)(\tilde{\omega} - \tilde{\omega}_{p+1})},$$

so that we finally obtain $\omega_2, \ldots, \omega_{p-1}$ as the points on the upper half of the unit circle satisfying

$$\operatorname{Re}\omega_j = \frac{(\tilde{\omega}_1 + \tilde{\omega}_p - 2\tilde{\omega}_{p+1})\tilde{\omega}_j + \tilde{\omega}_1\tilde{\omega}_{p+1} + \tilde{\omega}_p\tilde{\omega}_{p+1} - 2\tilde{\omega}_1\tilde{\omega}_p}{(\tilde{\omega}_1 - \tilde{\omega}_p)(\tilde{\omega}_j - \tilde{\omega}_{p+1})}. \tag{4.4}$$

From $\Psi'(w) = \tilde{\Psi}'(g(w))g'(w)$, we obtain the capacity of Ω, on using (4.2) and (4.3) and on letting $w \to \infty$, as

$$c = \left|\frac{\tilde{c}(\tilde{\omega}_p - \tilde{\omega}_1)}{4(\tilde{\omega}_{p+1} - \tilde{\omega}_p)(\tilde{\omega}_{p+1} - \tilde{\omega}_1)}\right| \cdot \prod_{j=2}^{p-1} |\tilde{\omega}_{p+1} - \tilde{\omega}_j|^{\beta_j} \tag{4.5}$$

(cf. [24. Theorem 12]).

At this point, we recall that the Faber polynomials satisfy the recurrence relations of (3.5):

$$F_0(z) = 1, \qquad F_1(z) = (z - c_0)/c,$$

$$F_{m+1}(z) = [zF_m(z) - (c_0F_m(z) + \cdots + c_{m-1}F_1(z)) - (m+1)c_m]/c \quad (m \geqq 1),$$

where c, c_0, c_1, ... are the coefficients in the Laurent series representation (3.4) of $\Psi(w)$:

$$\Psi(w) = cw + c_0 + \frac{c_1}{w} + \frac{c_2}{w^2} \dots . \tag{4.6}$$

If we differentiate (4.6) with respect to w, we obtain

$$\Psi'(w) = c - \frac{c_1}{w^2} - \frac{2c_2}{w^3} - \cdots . \tag{4.7}$$

The coefficients of this series can be computed from the formula (4.1). We have

$$\left(1 - \frac{1}{w}\right)^{\beta_1} = \sum_{k=0}^{\infty} \binom{\beta_1}{k} (-1)^k \frac{1}{w^k} =: \sum_{k=0}^{\infty} \gamma_k^{(1)} \frac{1}{w^k} ,$$

$$\left(1 - 2\frac{\mathrm{Re}\omega_j}{w} + \frac{1}{w^2}\right)^{\beta_j} = \sum_{k=0}^{\infty} \left(\sum_{l=0}^{k} \binom{\beta_j}{l} \binom{\beta_j}{k-l} \omega_j^l \bar{\omega}_j^{k-l} \right) (-1)^k \frac{1}{w^k} =: \sum_{k=0}^{\infty} \gamma_k^{(j)} \frac{1}{w^k}$$

for $j = 2, \dots, p-1$, and

$$\left(1 + \frac{1}{w}\right)^{\beta_p} = \sum_{k=0}^{\infty} \binom{\beta_p}{k} \frac{1}{w^k} =: \sum_{k=0}^{\infty} \gamma_k^{(p)} \frac{1}{w^k} .$$

Note that, for $j = 2, \dots, p-1$,

$$\gamma_k^{(j)} = \begin{cases} -2 \displaystyle\sum_{l=0}^{(k-1)/2} \binom{\beta_j}{l} \binom{\beta_j}{k-l} \mathrm{Re}[\omega_j^{k-2l}], & k \text{ odd} ; \\ 2 \displaystyle\sum_{l=0}^{(k/2)-1} \binom{\beta_j}{l} \binom{\beta_j}{k-l} \mathrm{Re}[\omega_j^{k-2l}] + \binom{\beta_j}{k/2}^2, & k \text{ even} , \end{cases}$$

so that we can compute these numbers using only real arithmetic.

Comparing (4.1) with (4.7) leads to

$$1 - \sum_{k=2}^{m} \frac{(k-1)c_{k-1}}{w^k} = \prod_{j=1}^{p} \left(1 + \sum_{k=1}^{m} \frac{\gamma_k^{(j)}}{w^k}\right) + O\left(\frac{1}{w^{m+1}}\right), \quad \text{as } w \to \infty . \tag{4.8}$$

The product of the Laurent polynomials on the right hand side in (4.8) can be expressed as a product of lower triangular Toeplitz matrices. Hence, on comparing corresponding powers of w in (4.8), we obtain, in matrix form, the following computable expression for the unknowns $\{c_j\}_{j=1}^{m-1}$:

$$\begin{bmatrix} 1 \\ 0 \\ -c_1 \\ -2c_2 \\ \vdots \\ -(m-1)c_{m-1} \end{bmatrix} = \prod_{j=1}^{p} \begin{bmatrix} 1 & & & \\ \gamma_1^{(j)} & \ddots & & \\ \vdots & \ddots & \ddots & \\ \gamma_m^{(j)} & \dots & \gamma_1^{(j)} & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} . \tag{4.9}$$

With (4.5) and (4.9), only c_0, the constant term in (4.6), is yet to be determined. We could compute c_0 by one evaluation of the Schwarz-Christoffel integral, but for our

purposes, it is sufficient to use the following simple procedure. We can express c_0 from (4.6) as

$$c_0 = \Psi(w) - \left(cw + \frac{c_1}{w} + \cdots + \frac{c_{m-1}}{w^{m-1}} + \cdots \right). \tag{4.10}$$

Now, our polygon Ω consists, in our construction, of 2 real vertices ζ_1 and ζ_p and $p-2$ vertices $\{\zeta_k\}_{k=2}^{p-1}$ in $\mathrm{Im} z > 0$ and $p-2$ reflected vertices in $\mathrm{Im}\, z < 0$, giving a total of $2p-2$ vertices in all for Ω. Since $\Psi(\omega_j) = \zeta_j$, $j = 1, \ldots, 2p-2$, where $|\omega_j| = 1$ for all $1 \leqq j \leqq 2p-2$, then evaluating (4.10) at the point ω_j leads to

$$c_0 = \zeta_j - \left(c\omega_j + \frac{c_1}{\omega_j} + \cdots + \frac{c_{m-1}}{\omega_j^{m-1}} + \cdots \right) \quad (1 \leqq j \leqq 2p-2)\,. \tag{4.11}$$

On neglecting the terms involving $\omega_j^{-m}, \omega_j^{-m-1}, \ldots$, in (4.11) and on adding the $2p-2$ equations of (4.11), we have

$$\tilde{c}_0 := \frac{1}{2p-2}\left[\sum_{j=1}^{2p-2} \zeta_j - \left(c \sum_{j=1}^{2p-2} \omega_j + c_1 \sum_{j=1}^{2p-2} \frac{1}{\omega^j} + \cdots + c_{m-1} \sum_{j=1}^{2p-2} \frac{1}{\omega_j^{m-1}} \right) \right], \tag{4.12}$$

which gives our approximation, $\tilde{c}_0$, to c_0 (which is dependent on m). Roughly speaking, this amounts, for a fixed m, to moving the image of the unit circle in the w-plane, under the mapping

$$cw + \frac{c_1}{w} + \cdots \frac{c_{m-1}}{w^{m-1}}$$

so that it approximately fits the given polygon Ω in the z-plane. Note that, since Ω is symmetric with respect to the real axis, (4.12) can be rewritten as

$$\tilde{c}_0 = \frac{1}{p}\left[\frac{\zeta_1 + \zeta_p}{2} + \sum_{j=2}^{p-1} \mathrm{Re}\, \zeta_j - \sum_{k=1}^{m-1} \frac{1 + (-1)^k}{2} c_k - \sum_{j=2}^{p-1} \left(c\, \mathrm{Re}\, \omega_j + \sum_{k=1}^{m-1} c_k\, \mathrm{Re}\, \omega_j^k \right) \right] \tag{4.13}$$

which can be evaluated using only real arithmetic. To derive (4.13) from (4.12), we only used the fact that $\zeta_{2p-j} = \bar{\zeta}_j$ and $\omega_{2p-j} = \bar{\omega}_j$ for $j = 2, \ldots, p-1$.

5. Implementational details

In our computations, it will always be assumed that m (the degree of some iteration polynomial, number of GMRES steps, etc.) is *small* compared to N (the dimension of the system, i.e., $A \in \mathbb{R}^{N,N}$). We will therefore count the number of vector operations, i.e., N scalar multiplications and N additions, as a measure for the complexity of the algorithm, as in Nachtigal et al. [26]. Let l denote the average number of elements per row in the matrix A, so that a matrix-vector multiplication costs l vector operations.

According to estimates in [33], the number of vector operations for m steps of the GMRES method requires

$$\sum_{j=1}^{m} (l + 2j + 1) + m = m(l + m + 3)$$

vector operations. The fact that the computing time grows like m^2, rather than m, is one of the reasons why the GMRES algorithm is generally restarted, in cases of slow convergence.

There are several possibilities for the practical computation of (cf. (2.3))

$$x_{\text{new}} = x_{\text{old}} + q_{m-1}(A) r_{\text{old}} . \tag{5.1}$$

It might seem that the natural way to implement a polynomial iterative method based on Faber polynomials is to make use of its recurrence relations (3.5) during the iteration, rather than to compute the coefficients of the polynomial $\tilde{F}_m(z)$ itself. This leads to

$$x_1 := \frac{1}{c_0} r_{\text{old}} + x_{\text{old}} , \tag{5.2}$$

$$x_j := -\frac{1}{c} \frac{F_{j-1}(0)}{F_j(0)} r_{j-1} - \frac{c_0}{c} \frac{F_{j-1}(0)}{F_j(0)} x_{j-1} - \cdots$$

$$- \frac{c_{j-2}}{c} \frac{F_1(0)}{F_j(0)} x_1 - j \frac{c_{j-1}}{c} \frac{1}{F_j(0)} x_{\text{old}} ,$$

for $j = 2, \ldots, m$ with $x_{\text{new}} := x_m$ (cf. [3, Lemma 1]). Note that, since the recurrence relations for the Faber polynomials do not have a finite length, the number of vector operations is given by

$$\sum_{j=1}^{m} (l + j) = ml + \frac{m(m+1)}{2} .$$

In other words, the work grows quadratically with the size of the polynomial (which is basically the same as in GMRES). There is also the fact that up to m preceding vectors must be stored during the iteration process. However, since the same must be done during the Arnoldi process, this is not really a drawback. Here, we choose to compute the coefficients of the polynomial $\tilde{F}_m(z)$, and then to apply the iteration polynomial $q_{m-1}(z) := (1 - \tilde{F}_m(z))/z =: \alpha_{0,m} + \alpha_{1,m} z + \cdots + \alpha_{m-1,m} z^{m-1}$ in a Horner-type iteration (cf. [11]):

$$w_0 = \alpha_{m-1,m} r_{\text{old}} ,$$

$$w_j = A w_{j-1} + \alpha_{m-1-j,m} r_{\text{old}}, \quad j = 1, \ldots, m-1 ,$$

$$x_{\text{new}} = x_{\text{old}} + w_{m-1} .$$

This reduces the number of vector operations to $m(l + 1)$. Although this approach could cause instabilities, due to large residuals in intermediate steps of the Horner scheme, no significant difference between these two implementations in our computations were observed in our numerical experiments.

Since we assumed that N, the dimension of the given matrix A, is much larger than m, the number of Arnoldi steps (which is also an upper bound for the number

of vertices in the polygon), then the work in specifically computing the parameters in the Schwarz-Christoffel formula (4.1) and in constructing the Faber polynomials using (3.5) is small compared to one iteration step and this will not be included in the number of operations in the examples of Sect. 6.

6. Numerical experiments

In this section, we compare our hybrid Arnoldi-Faber algorithm with other hybrid methods (the hybrid GMRES method by Nachtigal et al. [26] and the hybrid Arnoldi-Chebyshev method by Elman et al. [10]), as well as with variants of CG-like methods for nonsymmetric systems (restarted GMRES [33], CGS [36] and CGNR, where CGNR is the CG method applied to the normal equations). Examples 6.1 and 6.3 are taken from [26]; these matrices were artificially constructed to illustrate specific properties of the algorithms, such as, for example, the effect of non-normality on their performance. Example 6.2 is a discretized non-self-adjoint elliptic boundary value problem which can be viewed as a model problem for convection-diffusion equations.

Example 6.1. Let us consider again the banded Toeplitz matrix (1.1). In Fig. 1, we plotted the polygon that was constructed from the eigenvalue estimates after 16 Arnoldi steps, by the procedure described in Sect. 2. Here, the right hand side $\mathbf{b}$ and the initial guess x_0 were chosen as random vectors. The convergence behavior (more specifically, the norm of the residual in terms of the amount of work as described in Sect. 5) for the different methods is shown in Fig. 3. We compared Arnoldi-Faber (16), i.e., the hybrid method that results from using the Faber polynomials with respect to the polygon described above, to Hybrid GMRES(16), the hybrid GMRES algorithm by Nachtigal et al., GMRES(16), the restarted version of GMRES of order 16, CGS and CGNR.

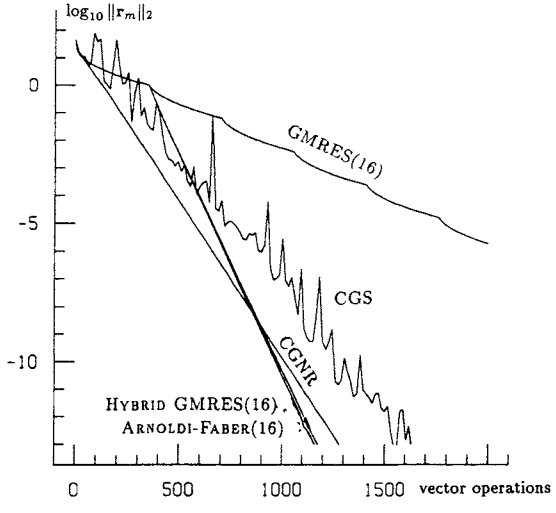

Fig. 3. Example 6.1: convergence behavior

We see that Arnoldi-Faber (16) and Hybrid GMRES(16), which are virtually indistinguishable, converge rapidly and smoothly, as CGNR also does. CGS shows a typical erratic convergence behavior, and restarted GMRES is much more slowly convergent.

Example 6.2. Consider the boundary value problem

$$-\Delta u + \tau u_x = f(x, y), \quad (x, y) \in S, \tag{6.1}$$

$$u(x, y) = g(x, y), \quad (x, y) \in \partial S$$

on the unit squares $S := (0, 1) \times (0, 1)$ with the boundary ∂S. Here, the functions $f: S \cup \partial S \to \mathbb{R}$ and $g: \partial S \to \mathbb{R}$ are assumed to be continuous. Discretizing (6.1) using central differences with meshsize $h = \dfrac{1}{n+1}$ leads to a linear system of equations for n^2 unknowns with a special block structure. If we set $u_{i,j} := u(x_j, y_i)$ with $x_j := jh$, $j = 1, \ldots, n$, $y_i := ih$, $i = 1, \ldots, n$, we obtain a linear system of equations for

$$x = (u_{1,1}; \ldots; u_{1,n}; \ldots; u_{n,1}; \ldots; u_{n,n})^{\mathrm{T}},$$

where the coefficient matrix $A \in \mathbb{R}^{N,N}$ is given by the Kronecker sum

$$B \otimes I_n + I_n \otimes C = \begin{bmatrix} 2I_n + C & -I_n & & \\ -I_n & 2I_n + C & \ddots & \\ & \ddots & \ddots & -I_n \\ & & -I_n & 2I_n + C \end{bmatrix}. \tag{6.2}$$

with

$$B = \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix} \quad \text{and}$$

$$C = \begin{bmatrix} 2 & -1+\mu & & \\ -1-\mu & \ddots & \ddots & \\ & \ddots & \ddots & -1+\mu \\ & & -1-\mu & 2 \end{bmatrix}$$

where $\mu := \tau(h/2)$. In all our numerical experiments for the boundary value problem (6.1), we choose $n = 32$, so that $N = n^2 = 1024$.

Starting with a random right hand side $\boldsymbol{b}$ and a random initial guess x_0, we carried out 16 steps of the Arnoldi process and obtained the eigenvalue estimates on the left halves of Figs. 4 (for $\mu = 2$) and 5 (for $\mu = 4$). The dots (filling a rectangle) are the exact eigenvalues of A. The right halves of Figs. 4 and 5 show the convergence behavior of Arnoldi-Faber(16), Arnoldi-Chebyshev (16), Hybrid GMRES (16) and GMRES(16). Again, the Arnoldi-Faber and the hybrid GMRES method both perform much better than restarted GMRES. Note that neither

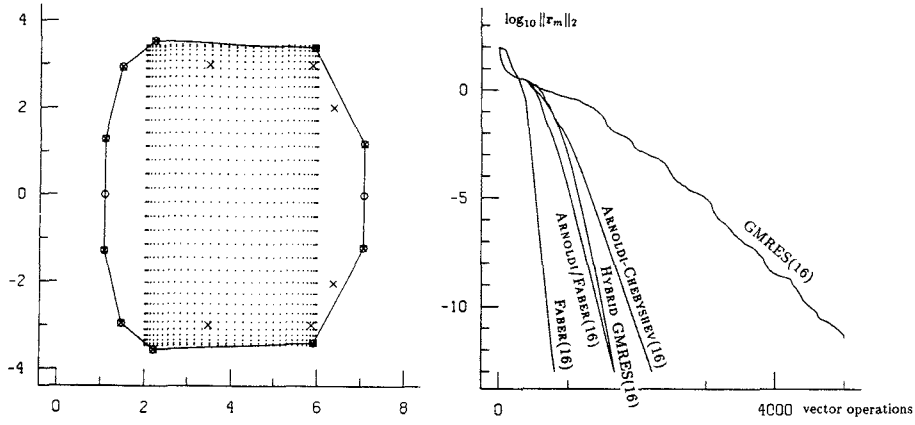

Fig. 4. Example 6.2 ($\mu = 2$): eigenvalue estimates (*left*) and convergence behavior (*right*)

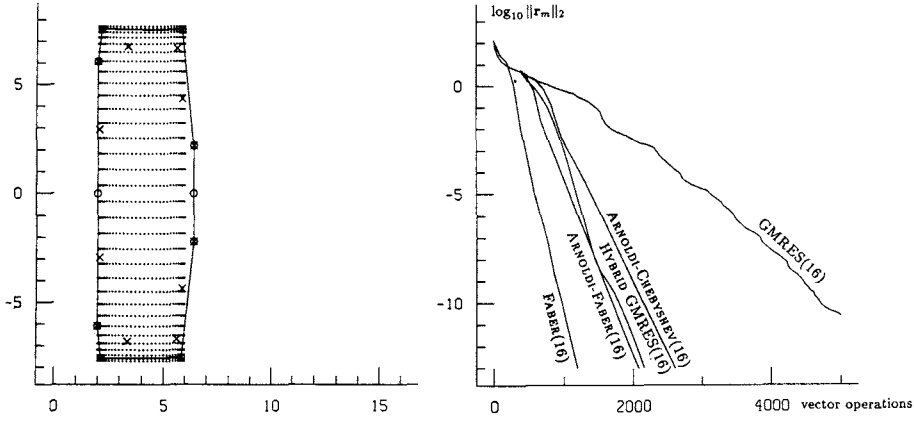

Fig. 5. Example 6.2 ($\mu = 4$): eigenvalue estimates (*left*) and convergence behavior (*right*)

CGNR nor CGS showed any sign of convergence here, and, for this reason, these methods are not included in Figs. 4 and 5. In this example, the hybrid Arnoldi-Chebyshev method also converges quite rapidly. That the asymptotic rate of convergence for the Arnoldi-Chebyshev method is smaller than the one for Arnoldi-Faber is a direct consequence of the comparison theorem in Niethammer and Varga [28, Theorem 3] (cf. also [5, Proposition 3]) and the fact that any ellipse enclosing the eigenvalue estimates has to be larger than the corresponding polygon.

At this point, it should be noted that, in many applications, a priori information on the location of the spectrum of A might be available from the physical or engineering origins of the problem. This information can then be used directly to construct the Faber polynomials without appealing to the Arnoldi process. For example, if we use the Faber polynomial of degree 16 with respect to the rectangle defined by the convex hull of the spectrum for the above example (denoted by Faber (16) in Figs. 4 and 5), the convergence is much faster than any of the other Krylov subspace methods!

In Figs. 3–5, the error curves for the hybrid GMRES and the Arnoldi-Faber method look very similar. This is a phenomenon that we encountered in many different examples: Whenever we start out with random vectors (which can be considered as being equally rich in all the eigenvector components of the given matrix A), the Arnoldi-Faber and the hybrid GMRES method showed very similar convergence behavior. However, there are situations, where the right hand side and also the starting guess, cannot be considered as being random and, instead, have a very regular structure. For linear systems arising from discretized elliptic boundary value problems, for example, the right hand side consists of the discrete values of a (possibly smooth) function and, as a starting guess, one might want to use a simple function matching the boundary values. We therefore ran Example 6.2 again with the right hand side vectors $\boldsymbol{b} = (-1, 1, \ldots, -1, 1)^T$ (Fig. 6) and $\boldsymbol{b} = (1, 1, \ldots, 1, 1)^T$ (Figs 7, 8). As our starting guess, we choose $\boldsymbol{x}_0 = (0, \ldots, 0)^T$ in both examples.

On the right in Fig. 6, we see that the hybrid GMRES algorithm has become much slower than our Arnoldi-Faber method. A more interesting situation occurs for the polygonal domains in Fig. 7: Starting the Arnoldi process with $\boldsymbol{v}_1 := (1, 1, \ldots, 1, 1)^T/32$ does not give satisfactory estimates for the spectrum of A any more. In fact, after 16 steps, the information about A is still so poor that both methods, Arnoldi-Faber(16) and Hybrid GMRES(16) do not converge. After 32 Arnoldi steps, the eigenvalue estimates are sufficient for Arnoldi-Faber to converge, while the hybrid GMRES algorithm still diverges (see Fig. 8). Finally, Hybrid GMRES(40) converges but needs about twice as many iterations as Arnoldi-Faber(40). Again, neither CGNR nor CGS converged at all, and are therefore also excluded from Figs. 6 and 7.

We made the same observation in the other examples of this section: When we start with a vector having a regular structure, Arnoldi-Faber usually converges faster than Hybrid GMRES. Our explanation for this phenomenon is that, since the vector $\boldsymbol{v}_1$ is rich in only some of the eigenvector components (and possibly does not even contain some of the others), the GMRES polynomial overemphasizes part of the spectrum. The Faber polynomial, however, which is based on eigenvalue

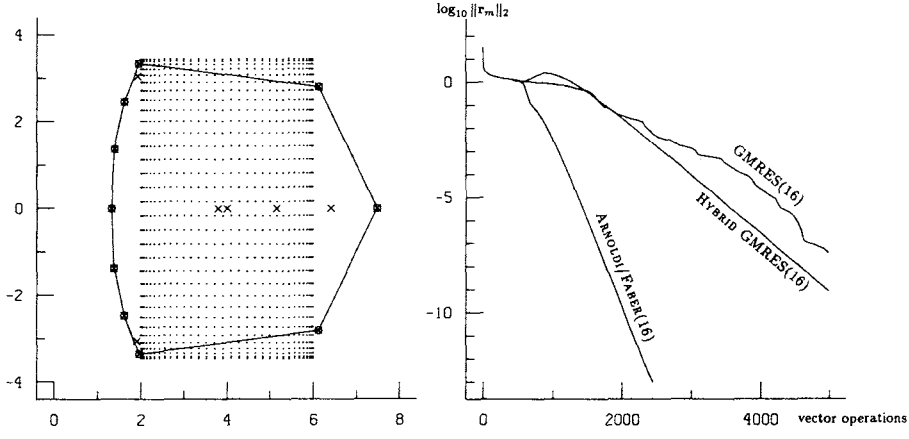

Fig. 6. Example 6.2 ($\mu = 2$, $\boldsymbol{b} = (-1, 1, \ldots, -1, 1)^T$): eigenvalue estimates (*left*) and convergence behavior (*right*)

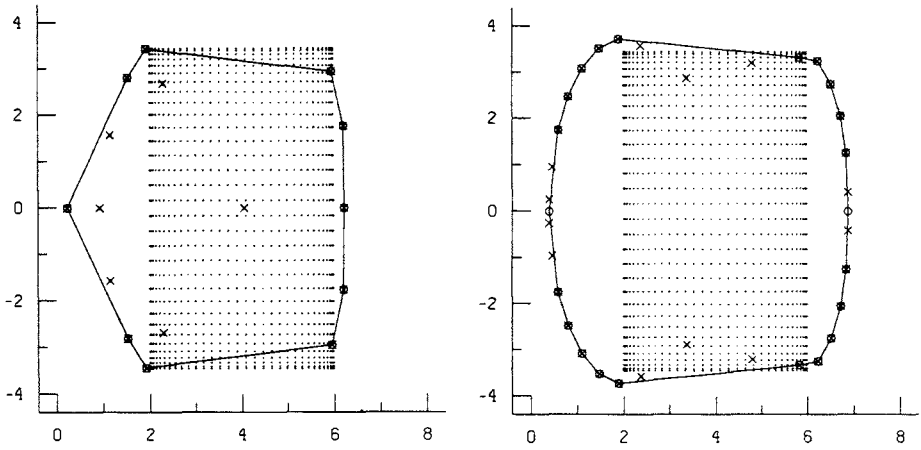

Fig. 7. Example 6.2 ($\mu = 2, \boldsymbol{b} = (1, 1, \ldots, 1, 1)^T$): eigenvalue estimates after 16 (*left*) and 32 (*right*) Arnoldi steps

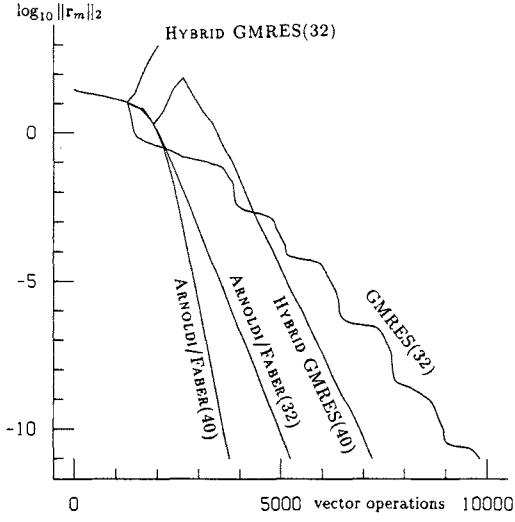

Fig. 8. Example 6.2 ($\mu = 2, \boldsymbol{b} = (1, 1, \ldots, 1, 1)^T$): convergence behavior

estimates, is small on the entire spectrum (as long as it is approximated by the Arnoldi estimates).

Example 6.3. We include this final example to show that the Arnoldi-Faber method can also work well for matrices where some of the eigenvalues have negative real parts. On the other hand, the same example points out the limitation of our method, i.e., unstable behavior or even divergence if the information on the location of the spectrum obtained from the Arnoldi process is not sufficient. We

consider the matrix

$$A = \begin{bmatrix} \frac{9}{10} & 1 & 1 & 1 & & & & \\ -1 & \frac{9}{10} & 1 & 1 & 1 & & & \\ & -1 & \frac{9}{10} & 1 & 1 & 1 & & \\ & & -1 & \frac{9}{10} & 1 & 1 & 1 & \\ & & & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \in \mathbb{R}^{1024,1024} \tag{6.3}$$

which is the "Grcar matrix" from Example 2 in [26], shifted to the left by 0.1. As it can be seen in Fig. 9, there is a small part of the spectrum, $\sigma(A)$, of A in (6.3) that is contained in the left half-plane (the smallest real part in $\sigma(A)$ is about -0.03). We note that this means that the hybrid Arnoldi-Chebyshev method, applied to this problem, cannot be convergent! In this example, the right hand side and the starting guess are again random vectors.

Figure 10 shows that the hybrid GMRES algorithm and Arnoldi-Faber(48) outperform the restarted GMRES versions. However, since Arnoldi-Faber(24) diverges, the same graph also shows the dependence of the hybrid Arnoldi-Faber method on the accuracy of the Arnoldi eigenvalue estimates. On the other hand, it is surprising that the Arnoldi-Faber method converges at all, since, in this example, part of the spectrum of A is not even contained in the polygon (cf. Fig. 9). We remark that the matrix of (6.3) has a relatively small condition number, so that CGNR outperforms all the methods included in Fig. 10 (cf. [26, Fig. 8.2b]).

It should be noted that the good performance of the hybrid schemes depend heavily on the sparsity of the matrix, i.e., the fact that matrix-vector multiplications are cheap. One might look at this from the point of view that the hybrid methods actually do not converge faster than restarted GMRES (in terms of the number of iterations), but that they are much cheaper.

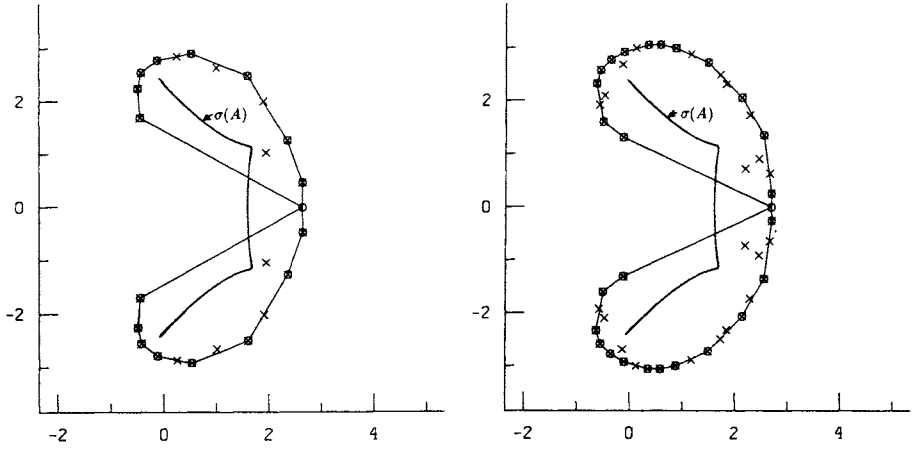

Fig. 9. Example 6.3: eigenvalue estimates after 24 (*left*) and 48 (*right*) Arnoldi steps

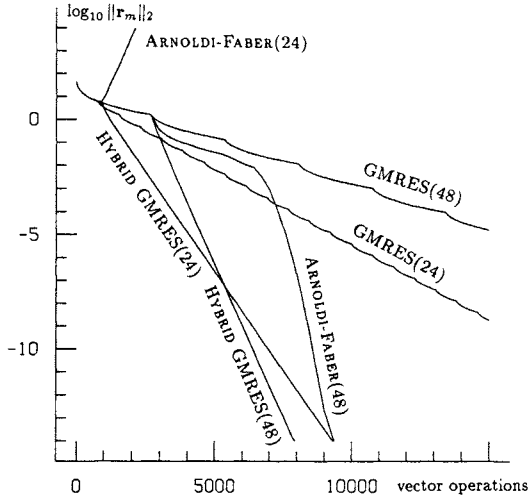

Fig. 10. Example 6.3: convergence behavior

7. Conclusions

Our limited numerical experiments indicate that our new hybrid Arnoldi-Faber method, although based on eigenvalue estimates, performs quite well for non-normal matrices. If the Krylov subspace is based on a random vector, i.e., a vector presumably rich in all the eigenvector components of the matrix A, then usually, both the Arnoldi-Faber method and the hybrid GMRES algorithm by Nachtigal, Reichel and Trefethen converge much faster than restarted versions of GMRES. However, if the right hand side of the linear system and the initial guess possess som'e regular structure, we observed that Arnoldi-Faber can also outperform hybrid GMRES. The performance of our method does not seem to depend directly on the non-normality of the matrix. However, it is quite *sensitive* to the amount of information obtained about the location of the spectrum obtained during the Arnoldi phase which is, on the other hand, dependent on the distance from normality of A. Of course, more experimentation (including more realistic problems) should be carried out on all these competing methods so that more information is gained as to which of these competing methods is most reliable under general circumstances. We plan to report on more extensive numerical experimentation of our hybrid Arnoldi-Faber method.

Acknowledgements. We would like to thank Michael Eiermann, Martin Gutknecht and Nick Trefethen for many helpful comments which helped to improve the presentation of this paper. This work, done while the first author was visiting the Institute for Computational Mathematics at Kent State University during the academic year 1990/91, was supported by the Alexander von Humboldt-Stiftung. The first author would like to thank the second author for the hospitality he received during this period.

References

1. Arnoldi, W.E. (1951): The principle of minimized iterations in the solution of the matrix eigenvalue problem. Quart. Appl. Math. **9**, 17–29
2. Carathéodory, C., Fejér, L. (1911): Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz. Rend. Circ. Mat. Palermo **32**, 218–239
3. Eiermann, M. (1989): On semiiterative methods generated by Faber polynomials. Numer. Math. **56**, 139–156
4. Eiermann, M. (1992): Fields of values and iterative methods. Linear Algebra Appl. (to appear)
5. Eiermann, M., Li, X., Varga, R.S. (1989): On hybrid semi-iterative methods. SIAM J. Numer. Anal. **26**, 152–168
6. Eiermann, M., Niethammer, W. (1983): On the construction of semiiterative methods. SIAM J. Numer. Anal. **20**, 1153–1160
7. Eiermann, M., Niethammer, W., Varga, R.S. (1985): A study of semiiterative methods for non-symmetric systems of linear equations. Numer. Math. **47**, 505–533
8. Eiermann, M., Starke, G. (1990): The near-best solution of a polynomial minimization problem by the Carathéodory-Fejér method. Constr. Approx. **6**, 303–319
9. Ellacott, S.W. (1983): Computation of Faber series with application to numerical polynomial approximation in the complex plane. Math. Comp. **40**, 575–587
10. Elman, H.C., Saad, Y., Saylor, P.E. (1986): A hybrid Chebyshev Krylov subspace algorithm for solving nonsymmetric systems of linear equations. SIAM J. Sci. Stat. Comput. **7**, 840–855
11. Elman, H.C., Streit, R.L. (1986): Polynomial iteration for nonsymmetric indefinite linear systems. In: Numerical Analysis, Lecture Notes in Mathematics 1230, pp. 103–117. Springer, Berlin Heidelberg New York
12. Faber, G. (1903): Über polynomische Entwicklungen. Math. Ann. **57**, 389–408
13. Farkova, N.A. (1988): The use of Faber polynomials to solve systems of linear equations. U.S.S.R. Comput. Maths. Math. Phys. **28**, 22–32
14. Fischer, B., Freund, R.W. (1991): Chebyshev polynomials are not always optimal. J. Approx. Theory **65**, 261–272
15. Gaier, D. (1987): Lectures on complex approximation. Birkhäuser, Boston Basel Stuttgart
16. Golub, G.H., Loan, C.F.V. (1989): Matrix computations, 2nd edn. Johns Hopkins University Press, Baltimore London
17. Golub, G.H., Varga, R.S. (1961): Chebyshev semiiterative methods, successive overrelaxation iterative methods, and second order Richardson iterative methods. Numer. Math. **3**, 147–168
18. Gutknecht, M.H. (1986): An iterative method for solving linear equations based on minimum norm Pick-Nevanlinna interpolation. In: C.K.C. et al., eds. Approximation theory V. pp. 371–374. New York, Academic Press
19. Henrici, P. (1974): Applied and computational complex analysis I. Wiley, New York London Sydney Toronto
20. Henrici, P. (1986): Applied and computational complex analysis III. Wiley, New York London Sydney Toronto
21. Horn, R.A., Johnson, C.R. (1991): Topics in matrix analysis. Cambridge University Press, Cambridge New York Port Chester Melbourne Sydney
22. Kövari, T., Pommerenke, C. (1967): On Faber polynomials and Faber expansions. Math. Z. **99**, 193–206
23. Kublanovskaja, V.N. (1959): Applications of analytic continuation in numerical analysis by means of change of variables. Trudy Mat. Inst. Steklov **53**, 145–185
24. Li, X. (1989): An adaptive method for solving nonsymmetric linear systems involving application of SCPACK. PhD thesis, Kent State University
25. Manteuffel, T.A. (1978): Adaptive procedure for estimating parameters for the nonsymmetric Tchebychev iteration. Numer. Math. **31**, 183–208
26. Nachtigal, N.M., Reichel, L., Trefethen, L.N. (1992): A hybrid GMRES algorithm for nonsymmetric matrix iterations. SIAM J. Matrix Anal. Appl. **13**, 796–825
27. Nehari, Z. (1952): Conformal mapping. McGraw-Hill, New York
28. Niethammer, W., Varga, R.S. (1983): The analysis of k-step iterative methods for linear systems from summability theory. Numer. Math. **41**, 177–206
29. Pommerenke, C. (1965): Konforme Abbildung und Fekete-Punkte. Math. Z. **89**, 422–438

30. Rivlin, T.J., Shapiro, H.S. (1961): A unified approach to certain problems of approximation and minimization. J. Soc. Indust. Appl. Math. **9**, 670–699
31. Saad, Y. (1980): Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices. Linear Algebra Appl. **34**, 269–295
32. Saad, Y. (1987): Least squares polynomials in the complex plane and their use for solving nonsymmetric linear systems. SIAM J. Numer. Anal. **24**, 155–169
33. Saad, Y., Schultz, M.H. (1986): GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. **7**, 856–869
34. Saylor, P.E., Smolarski, D.C. (1991): Implementation of an adaptive algorithm for Richardson's method. Linear Algebra Appl. **154–156**, 615–646
35. Smirnov, V.I., Lebedev, N.A. (1968): Functions of a complex variable: Constructive theory. M.I.T. Press, Cambridge, MA
36. Sonneveld, P. (1989): CGS, a fast Lanczos-type solver for nonsymmetric linear systems. SIAM J. Sci. Stat. Comput. **10**, 36–52
37. Trefethen, L.N. (1980): Numerical computation of the Schwarz-Christoffel transformation. SIAM J. Sci. Stat. Comput. **1**, 82–102
38. Trefethen, L.N. (1981): Near-circularity of the error curve in complex Chebyshev approximation. J. Approx. Theory **31**, 344–367
39. Trefethen, L.N. (1990): Approximation theory and numerical linear algebra. In: Algorithms for approximation II, 336–360. Chapman & Hall, London
40. Varga, R.S. (1957): A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials. J. Soc. Indust. Appl. Math. **5**, 39–46
41. Walsh, J.L. (1956): Interpolation and approximation by rational functions in the complex domain, 2nd edn. American Mathematical Society, Rhode Island