Construction and analysis
of polynomial iterative methods
for non-hermitian systems of linear equations

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Mathematik der Universität Bielefeld

vorgelegt von

Jörg Liesen

November 1998

Acknowledgments

I thank my thesis advisor Prof. Ludwig Elsner for his careful guidance. He allowed me to pursue my own ideas, while constantly checking if I was on the right track. He always had time to talk and to listen, made numerous suggestions and gave important input. My work benefits strongly from his professional experience.

I am grateful to Prof. Paul Saylor, whom I met during my time as an M.S. student at the University of Illinois, for his support and friendship. This thesis gained a lot from his insight in iterative methods. Thanks, Paul, for being in close contact via email and for being my roommate at several occasions. (Please keep on writing trip reports!)

Thanks to Prof. Zdeněk Strakoš for caring so much about my progress since we met during IMPC '97 in Milovy, Czech Republic. In our frequent discussions he made numerous valuable suggestions (in particular for [50]). His work and ideas inspired much of my research. Thanks also for inviting me to Prague and for visiting Bielefeld.

Thanks to Tino Koch, coauthor of [48], for being a friend and enthusiastic colleague. Thanks to Dr. Reinhard Nabben for helpful comments and thorough reading of the manuscript.

My Ph.D. work was sponsored by the Department of Mathematics of the University of Bielefeld and by a grant from the Studienstiftung des deutschen Volkes.

1. Gutachter: Prof. Dr. Ludwig Elsner
2. Gutachter: Prof. Dr. Wolf-Jürgen Beyn
Datum der mündlichen Prüfung: 27. Januar 1999

Gedruckt auf alterungsbeständigem Papier nach ∞ ISO 9706

Contents

List of Figures

Basic notation and definitions

$\mathbf{R}$	real numbers
$\mathbf{C}$	complex numbers
$\hat{\mathbf{C}}$	extended complex plane, $\hat{\mathbf{C}} := \mathbf{C} \cup \{\infty\}$
$\mathbf{E}$	closed unit disk, $\mathbf{E} := \{z \,:\, \lvert z\rvert \leq 1\}$
$\mathbf{S}^C$	exterior of the set $\mathbf{S} \subset \hat{\mathbf{C}}$, $\mathbf{S}^C := \hat{\mathbf{C}} \setminus \mathbf{S}$
$\partial\mathbf{S}$	boundary of $\mathbf{S}$
$int(\mathbf{S})$	interior of $\mathbf{S}$
$z\mathbf{S}$	product of $z \in \mathbf{C}$ and $\mathbf{S} \subset \hat{\mathbf{C}}$, $z\mathbf{S} := \{zw \,:\, w \in \mathbf{S}\}$
$\lVert \cdot \rVert$	Euclidean vector and corresponding matrix norm
I	identity matrix (its size being determined by the context)
e_k	kth column of I
M^T	transpose of the matrix $M \in \mathbf{C}^{N\times N}$
M^H	hermitian transpose of M
$\Lambda(M)$	spectrum of M
$F(M)$	field of values of M, $F(M) := \{v^H M v \,:\, v \in \mathbf{C}^N,\ \lVert v\rVert = 1\}$
$\Lambda_\epsilon(M)$	ϵ-pseudospectrum of M, $\Lambda_\epsilon(M) := \{z \in \mathbf{C} \,:\, \lVert (zI - M)^{-1}\rVert \geq \epsilon^{-1}\}$, for all $\epsilon \geq 0$ [74]
$\kappa(M)$	condition number of M, $\kappa(M) := \lVert M\rVert \lVert M^{-1}\rVert$ for nonsingular M, $\kappa(M) := \infty$ otherwise
π_n	set of all polynomials of degree at most n with value one at the origin

1 Introduction

One of the most important tasks in scientific computing is to solve systems of linear equations,

$$Ax = b,\ A \in \mathbf{C}^{N\times N},\ b \in \mathbf{C}^N. \tag{1.1}$$

Such systems arise in many areas of science and engineering, often from discretizations of continuous problems. The coefficient matrix A typically has a special structure, and often most of its entries are zero. Such a matrix is called sparse.

Originally, direct methods, in particular Gaussian elimination, have been used for solving (1.1). These methods are based on a factorization of A. Since this requires $\mathcal{O}(N^3)$ floating point operations (flops), direct methods are only feasible if N is small. Furthermore, in many cases direct methods can only partially take advantage of sparsity. The factorization of A might even lead to fill-in, which means that the computed factors have more nonzero entries than A itself. Hence, for large N direct methods "are prohibitive both in terms of storage requirements and computing time" [29, p. 2].

To overcome the $\mathcal{O}(N^3)$ bottleneck and to exploit sparsity, dozens of iterative solution methods have been developed in the last two decades. Many of these methods are based on polynomials in A, and are therefore called *polynomial iterative methods (PIMs).* These methods are the topic of this thesis. One of the main characteristics of PIMs is that for each step only a few matrix-vector products are required. The computational effort therefore potentially is $\mathcal{O}(N^2)$ flops, even $\mathcal{O}(N)$ if A has very few nonzero entries. Because the solution of very large sparse linear systems is of growing importance in a variety of applications, PIMs have been gaining more and more attention and popularity in recent years.

1.1 What is a PIM?

Given an initial guess x_0 for the solution of (1.1), a PIM is an iterative method in which the iterates x_n are of the form

$$x_n = x_0 + q_n(A)r_0,\quad n = 1, 2, ..., \tag{1.2}$$

where $r_0 := b - Ax_0$ is the initial residual and the iteration polynomial $q_n(z)$ is of degree at most $n-1$. PIMs are also called Krylov subspace methods, because (1.2) is equivalent to

$$x_n \in x_0 + K_n(A, r_0),\quad n = 1, 2, ..., \tag{1.3}$$

where

$$K_n(A, r_0) := \mathrm{span}\{r_0, Ar_0, ..., A^{n-1}r_0\} \tag{1.4}$$

is the nth Krylov space generated by A and r_0.

The nth residual is given by

$$r_n := b - Ax_n = p_n(A)r_0,\quad p_n(z) := 1 - zq_n(z) \in \pi_n. \tag{1.5}$$

The polynomial $p_n(z)$, which is of degree at most n and satisfies $p_n(0) = 1$, is called the nth residual polynomial. If A is nonsingular, the nth error is

$$e_n := A^{-1}b - x_n = p_n(A)e_0. \tag{1.6}$$

Generally, the choice of the iterates or iteration polynomials should lead to small residuals or errors in a given norm specific to the method.

For hermitian systems, numerically stable and computationally cheap PIMs, in particular CG [42, 61] for hermitian positive definite and MINRES [59] for hermitian indefinite systems, are long known and well understood. Because of their favorable convergence properties, CG and MINRES are the methods of choice for the iterative solution of hermitian linear systems.

For non-hermitian systems the state-of-the-art is not as well developed. Starting in the mid-1980s, dozens of PIMs for non-hermitian systems have been proposed, but no clear method of choice crystallized. For non-hermitian systems an 'ideal' method, yielding optimal iterates with low work and storage requirements, cannot be constructed. As shown by Faber and Manteuffel, the solution of a non-hermitian linear system by a PIM must always be a compromise between optimality of the iterates and computational effort (see [23] for details). Many PIMs for non-hermitian systems generate iterates that are optimal in some norm, but these norms often depend on the initial guess and might be arbitrarily different from any standard norm (see [4] for details). In addition, the properties of most of the existing methods are not fully understood. Because of this unsatisfactory situation, the analysis of existing as well as the search for new PIMs for non-hermitian systems are currently actively pursued, and are of primary concern in this thesis.

1.2 Different types of PIMs

The existing PIMs can be characterized as either parameter-dependent (hybrid) or parameter-free.

Parameter-dependent (hybrid) PIMs. To make $\|r_n\|$ small, these PIMs try to make $\|p_n(A)\|$ small. This is usually done in two phases:

Phase I: Obtain information about the matrix A, for example an estimate of the spectrum, a pseudospectrum or the field of values of A.

Phase II: Iterate with polynomials that are constructed *explicitly* using the information from Phase I.

These methods are called parameter-dependent, because Phase II depends on parameters obtained in Phase I. Phase I typically requires a method to solve eigenvalue problems, while Phase II consists of the iterative solution of a linear system. Because of this mixture of ingredients, parameter-dependent PIMs are also called hybrid PIMs. We will use the latter name in the following.

Examples of hybrid PIMs are the Chebyshev-Krylov method by Elman, Saad and Saylor [18], which is based on Manteuffel's work [51, 52], and the Arnoldi-Faber method by Starke and Varga [69]. In Phase I, both methods compute an estimate of the spectrum of A, which is then surrounded by an ellipse (Chebyshev-Krylov) or polygon (Arnoldi-Faber). In Phase II, normalized Chebyshev or Faber polynomials on the set bounded by the ellipse or polygon, respectively, are used as residual polynomials.

A survey of the many variants of hybrid PIMs is given in [57]. Different strategies for obtaining information about A in Phase I are discussed in [53].

Parameter-free PIMs. These PIMs recursively generate bases of the Krylov spaces $K_n(A, r_0)$ and compute iterates $x_n \in x_0 + K_n(A, r_0)$ so that r_n satisfies certain properties. Typical examples of such properties are:

(a) Minimization property: $\|r_n\| = \min_{r \in r_0 + AK_n(A, r_0)} \|r\|$ (GMRES [66]).

(b) Ritz-Galerkin property: r_n is orthogonal to $K_n(A, r_0)$ (FOM [66]).

(c) Petrov-Galerkin property: r_n is orthogonal to some other n-dimensional subspace (BiCG [26], QMR [30]).

In contrast to hybrid PIMs, the iteration and residual polynomials are constructed *implicitly.* The relation (1.2) is satisfied via the equivalent relation (1.3). Surveys of parameter-free PIMs are given in [29, 31, 33, 70].

1.3 Organization and summary of our results

This thesis deals with the construction and analysis of polynomial iterative methods (PIMs) for systems of linear equations. We make no assumption on symmetry properties of A, and thus our focus is on PIMs for non-hermitian systems. Our main contributions are:

(a) The construction and analysis of a new hybrid PIM, named ABF method according to its three main ingredients, the Arnoldi process, 'bratwurst' shape sets and Faber polynomials.

(b) The derivation of new convergence results for GMRES.

(c) The derivation of a new polynomial preconditioner from the iteration polynomials of the ABF method.

In Section 2, *Background*, we present background facts from numerical analysis and approximation theory needed in later sections. One focus is the close relation between the theory of conformal maps and polynomial approximation in the complex plane:

If $\Omega \subset \mathbf{C}$ is compact, the fastest rate of convergence to zero of polynomials on Ω can be computed using the conformal map from the exterior of the closed unit disk onto the exterior of Ω. Moreover, this conformal map, called the exterior mapping function of Ω, can be used to construct polynomials that decrease to zero on Ω with the fastest possible rate.

These facts are key ingredients of our general strategy for the construction and analysis of PIMs:

Find some compact set Ω that contains a set associated with the given matrix A, e.g. the spectrum or a pseudospectrum of A. Such set Ω is referred to as inclusion set. Given the inclusion set Ω, use its exterior mapping function to analyze how fast polynomials can decrease to zero on Ω, and to construct polynomials explicitly, which decrease to zero quickly on Ω. Then apply these results to the construction and analysis of PIMs for a linear system with coefficient matrix A.

Note that $p_n(0) = 1$ holds for the residual polynomials of any PIM. Thus, the Maximum Modulus Theorem implies that $0 \notin \Omega$ must be assumed of any inclusion set Ω in our context. This restriction in many cases requires Ω to be non-convex to contain the spectrum or pseudospectra of a matrix. To find suitable non-convex inclusion sets is a major challenge in the construction and analysis of PIMs. We tackle this problem in Section 3.

In Section 3, *Inclusion sets generated by the conformal 'bratwurst' maps*, we derive conformal maps to be used as exterior mapping functions of 'bratwurst' shape inclusion sets. The 'bratwurst' shape sets are non-convex, and their boundary is shown to be of bounded rotation. The conformal maps are composed of elementary Möbius transformations and the Joukowsky map. We thus avoid the Schwarz-Christoffel approach for finding inclusion sets, which involves quantities defined implicitly through integral equations. With our approach, the transfinite diameter and asymptotic convergence factor of each of the 'bratwurst' shape sets are given by explicit and easily computable analytic formulas. Results of Section 3 are applied in Sections 4, 5 and 6.

In Section 4, *The hybrid ABF method for non-hermitian linear systems*, we construct a new hybrid PIM. The hybrid PIMs we found in the literature are typically either based on convex inclusion sets (e.g. the Chebyshev-Krylov method [18]), or require long-term recurrences for their iteration polynomials (e.g. the Faber-Arnoldi [69] method). In contrast, our new hybrid ABF method complements the existing methods by featuring non-convex inclusion sets *and* short-term recurrences.

The 'bratwurst' shape sets of Section 3 play an important role in the construction of the ABF method, as they are used as inclusion sets for sets associated with the given matrix A. Our initial observation is that the 'bratwurst' shape sets are closely related to ellipses. It can therefore be shown that the Faber polynomials for these sets are related to Chebyshev polynomials for ellipses. In particular, we show that the Faber polynomials satisfy three-term recurrences. This allows us to derive three-term recurrences for the iteration polynomials of our new method. In addition to its efficiency, our method is widely applicable, due to the non-convexity and variability of the inclusion sets it is based on.

Based on a convergence result for Phase II of our method, we show what information about the matrix A should be obtained in Phase I: If A is highly nonnormal, its spectrum might give misleading information for choosing the inclusion set. In such cases pseudospectra of A should be considered. Numerical examples illustrate our analytic results. The numerical examples also show that our method has the potential to solve linear systems with $\mathcal{O}(N^2)$ floating point operations, while work and storage requirements during the iteration are constant due to the three-term recurrences.

In Section 5, *Computable convergence bounds for GMRES*, we present an unconventional approach for analyzing the behavior of GMRES, arguably the most robust and most popular PIM for non-hermitian linear systems. We first generalize known results that relate the convergence behavior of GMRES applied to a given linear system on the one hand to the algorithm's behavior if applied to other, so-called GMRES-equivalent systems on the other hand. Considering GMRES-equivalent systems with coefficient matrices having their eigenvalues on the unit circle, we analyze the behavior of GMRES using Faber polynomials on circular arcs. We obtain two different new bounds on the residual norms.

Our bounds depend on the initial guess and are applicable even to singular linear systems, if a certain condition is satisfied. Thus they are conceptually different from and more widely applicable than many standard 'worst-case' results. We show in numerical examples that our bounds are capable of describing the actual convergence behavior of GMRES for the given linear system and initial guess. We also show that approximations to both our bounds are computable from information generated during the GMRES run. Using WZ-GMRES [77], a certain GMRES-implementation, we develop the new idea of predicting the performance of the algorithm during its run.

By-products of our use of WZ-GMRES are generalizations of known results on the numerical stability of WZ-GMRES. Our generalizations are also exact expressions for the GMRES residual norms in terms of the conditioning of a matrix produced during the GMRES run. Hence, they are additional GMRES convergence results incorporating the initial guess.

In Section 6, *The ABF polynomial preconditioner*, we introduce a polynomial preconditioner that is derived from the hybrid ABF method of Section 4. The preconditioning polynomials can be computed efficiently by a three-term recurrence. Moreover, the ABF polynomial preconditioner is of wide applicability, because it is based on the non-convex 'bratwurst' shape sets defined in Section 3. We show in numerical examples that the preconditioner can be successfully applied to improve the performance of GMRES.

2 Background

2.1 Krylov spaces and the Arnoldi process

Suppose that $A \in \mathbf{C}^{N\times N}$ and $v \in \mathbf{C}^N$. Then for $n = 1, 2, ...,$

$$K_n(A, v) := \operatorname{span}\{v, Av, ..., A^{n-1}v\} \tag{2.1}$$

is called the nth Krylov space generated by A and v. For technical reasons we define

$$K_0(A, v) := \{0\}. \tag{2.2}$$

The Krylov spaces form a nested sequence of subspaces of $\mathbf{C}^N$ with the following well-known property.

Lemma 2.1 *If $A \in \mathbf{C}^{N\times N}$ and $v \in \mathbf{C}^N$, then*

$$\dim K_N(A, v) = d \ \textit{for some} \ d, \ \ 0 \le d \le N,$$

is equivalent to

$$\begin{aligned} \dim K_j(A, v) &= j \ \ \textit{for} \ \ 0 \le j \le d \ \ \textit{and} \\ \dim K_j(A, v) &= d \ \textit{for} \ \ j \ge d + 1. \end{aligned}$$

It is an immediate consequence that $K_N(A, v) = K_d(A, v)$, where $d := \dim K_N(A, v)$. The construction of an orthonormal basis (with respect to the Euclidean inner product), $w_1, w_2, ..., w_d$, of $K_d(A, v)$ is called the Arnoldi process [2]. In many standard references, the orthogonalization performed by the Arnoldi process is implemented with the Classical or Modified Gram-Schmidt procedure, see e.g. [32]. As pointed out by Walker [76], it is also possible to use Householder transformations for this purpose. For a collection of different implementations of the Arnoldi process, we refer to [12].

In practice, d is not known, and the basis vectors are constructed recursively. After $n < d$ steps of the Arnoldi process, we have

$$AW_n = W_{n+1}H_n^e, \ \ W_n := [w_1, ..., w_n], \tag{2.3}$$

where

$$H_n^e = \begin{bmatrix} h_{11} & \cdots & h_{1n} \\ h_{21} & \ddots & \vdots \\ & \ddots & h_{nn} \\ & & h_{n+1,n} \end{bmatrix} \in \mathbf{C}^{(n+1)\times n}. \tag{2.4}$$

Lemma 2.1 implies that $h_{k+1,k} \neq 0$ for $1 \le k \le n$, i.e. H_n^e is unreduced. In exact arithmetic, the process has to terminate at step d with the final matrix relation

$$AW_d = W_d H_d, \ \ W_d \in \mathbf{C}^{N\times d}, \ \ H_d := W_d^H A W_d \in \mathbf{C}^{d\times d}. \tag{2.5}$$

The Arnoldi process plays an important role in the context of PIMs. Note that

$$H_n := W_n^H A W_n = \begin{bmatrix} h_{11} & \cdots & \cdots & h_{1n} \\ h_{21} & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & h_{n-1,n} & h_{nn} \end{bmatrix} \in \mathbf{C}^{n\times n}. \tag{2.6}$$

For small n, the extreme eigenvalues of H_n are often close approximations of the extreme eigenvalues of A [64]. Furthermore, H_n can be used to compute approximations of pseudospectra [73] and the field of values of A [53]. Hence, the Arnoldi process is a useful tool to obtain information about the matrix A in Phase I of hybrid PIMs (see e.g. [18, 53, 69] for such applications). In parameter-free PIMs, for example in FOM and GMRES [66], the Arnoldi process is used to generate the bases required to compute the iterates.

2.2 Exterior mapping functions and Faber polynomials

Major tools in the construction and analysis of PIMs are conformal maps and their application to polynomial approximation in the complex plane. Below, we survey the respective tools needed in this thesis. For more details, we refer to [10, 67, 78].

Definition 2.2 *We define the set $\mathcal{E}$ as the collection of all conformal maps $\Psi : \mathbf{E}^C \to \hat{\mathbf{C}}$, which have the following properties:*

(a) $\Psi(\infty) = \infty$.

(b) $\Psi'(\infty) > 0$.

(c) $\Psi(\mathbf{E}^C)$ *is simply connected in* $\hat{\mathbf{C}}$.

(d) $0 \in \Psi(\mathbf{E}^C)$.

Suppose that $\Omega \subset \mathbf{C}$ is compact and that $0 \notin \Omega$ and Ω^C is simply connected in $\hat{\mathbf{C}}$. Then the Riemann Mapping Theorem guarantees the existence of a unique conformal map Ψ,

$$\Psi : \mathbf{E}^C \to \Omega^C, \quad \Psi \in \mathcal{E}, \tag{2.7}$$

which we call the exterior mapping function of Ω. The positive real number

$$\Psi'(\infty) =: t > 0 \tag{2.8}$$

is called the transfinite diameter (or capacity) of Ω.

The exterior mapping function and its inverse can be expanded as

$$\Psi(w) = t\left(w + \alpha_0 + \frac{\alpha_1}{w} + \frac{\alpha_2}{w^2} + ...\right), \tag{2.9}$$

and

$$\Psi^{-1}(z) = \frac{z}{t} + \beta_0 + \frac{\beta_1}{z} + \frac{\beta_2}{z^2} + ... \tag{2.10}$$

The part of the expansion of $\left[\Psi^{-1}(z)\right]^n$, $n \geq 0$, which contains the nonnegative powers of z, is called the nth Faber polynomial $F_n(z)$ for Ω. Obviously, $F_0(z) \equiv 1$ and $F_n(z)$ is of exact degree n with leading term $(z/t)^n$.

With the definition $\mathcal{L}_R := \{\Psi(w) \ : \ |w| = R > 1\}$, a well-known formula for the computation of $F_n(z)$ is given by

$$F_n(z) = \frac{1}{2\pi i} \int\limits_{|w|=R} \frac{w^n \Psi'(w) dw}{\Psi(w) - z}, \quad n \geq 1, \quad \text{for all } \ z \in int(\mathcal{L}_R) \supset \Omega. \tag{2.11}$$

Originally introduced in the context of polynomial approximation of analytic functions in the complex plane [21], Faber polynomials have found applications in many areas of mathematics. Examples can be found in [10].

Suppose that $\partial\Omega$ is an analytic Jordan curve. Then, by the Carathéodory-Osgood Theorem, Ψ can be extended to a homeomorphism of $\mathbf{E}^C \cup \partial\mathbf{E}$ onto $\Omega^C \cup \partial\Omega$, and $\partial\Omega = \{\Psi(e^{i\beta}) \ : \ 0 \leq \beta \leq 2\pi\}$. Let $\theta(\beta)$ denote the angle between the positive real axis and the tangent of $\partial\Omega$ at the point $\Psi(e^{i\beta})$. Then the boundary rotation of Ω is defined as the total variation of $\theta(\beta)$ [58] (also cf. [49]), i.e. the Stieltjes integral

$$V(\Omega) := \int\limits_{\partial\Omega} |d\theta|. \tag{2.12}$$

If $V(\Omega) < \infty$, Ω is said to be of bounded boundary rotation.

The maximum norm of the Faber polynomials for Ω is uniformly bounded by $V(\Omega)$,

$$\max_{z\in\Omega} |F_n(z)| \leq \frac{V(\Omega)}{\pi}, \quad n \geq 0 \tag{2.13}$$

(see e.g. [69]).

2.3 Inclusion sets and asymptotic analysis

The explicit construction of polynomials for hybrid PIMs as well as the analysis of the speed of convergence of PIMs is often based on a so-called inclusion set $\Omega \subset \mathbf{C}$. What this set includes depends on the method and on properties of the given linear system. Here are some examples:

Example 2.3 $\Omega \supset \Lambda(A)$. Suppose that the matrix A in (1.1) is diagonalizable, $A = XDX^{-1}$, D diagonal. Then from (1.5) follows that

$$\frac{\|r_n\|}{\|r_0\|} \leq \kappa(X) \max_{\lambda\in\Lambda(A)} |p_n(\lambda)|. \tag{2.14}$$

Apparently, an upper bound on the rate of convergence of a PIM applied to (1.1) is computable, if it is known how fast its residual polynomials decrease to zero on $\Lambda(A)$. Furthermore, (2.14) suggests to use residual polynomials which decrease quickly on (an estimate of) $\Lambda(A)$ in Phase II of a hybrid PIM. However, analyzing and constructing polynomials on a discrete set is difficult, which leads to the introduction of a compact, connected set $\Omega \supset \Lambda(A)$, on which the analysis and construction is based.

Example 2.4 $\Omega \supseteq \Lambda_\epsilon(A) \supset \Lambda(A)$. If A is highly nonnormal, i.e. $\kappa(X)$ is very large, eigenvalue information might not be enough to explain or guarantee the convergence of a PIM, and a set larger than $\Lambda(A)$ might be considered as an inclusion set. This insight led Trefethen to the introduction of the ϵ-pseudospectrum $\Lambda_\epsilon(A)$ [74, p. 352]. Using the integral representation of the operator $p_n(A)$ (see e.g. [9, p. 201]),

$$p_n(A) = \frac{1}{2\pi i} \int_C p_n(z)\,(zI - A)^{-1}\,dz, \tag{2.15}$$

where C is any Jordan curve enclosing $\Lambda(A)$, and choosing $C = \partial\Lambda_\epsilon(A)$, from (1.5) follows that

$$\frac{\|r_n\|}{\|r_0\|} \leq \frac{L_\epsilon}{2\pi\epsilon} \max_{\lambda \in \Lambda_\epsilon(A)} |p_n(\lambda)|, \tag{2.16}$$

where L_ϵ denotes the arc length of $\partial\Lambda_\epsilon(A)$. Thus, the construction or analysis of PIMs might be based on $\Lambda_\epsilon(A)$ or, if $\Lambda_\epsilon(A)$ is not connected, on some compact, connected inclusion set $\Omega \supset \Lambda_\epsilon(A)$. But note that (2.16) yields a different bound for each $\epsilon > 0$. Moreover, ϵ appears in the numerator and denominator, so that implications of (2.16) often require a careful analysis.

Example 2.5 $\Omega = F(A) \supset \Lambda(A)$. Some authors base their analysis or construction of PIMs on the field of values $F(A)$. Since $F(A)$ is always compact and convex, $\Omega = F(A)$ might be chosen. We refer to [15, 53] for applications to hybrid PIMs and to [16, 20, 46] for applications to parameter-free PIMs.

Because of the constraint $p_n(0) = 1$, it generally has to be assumed that

$$0 \notin \Omega. \tag{2.17}$$

This assumption causes major practical difficulties. If A is indefinite, any convex set containing $\Lambda(A)$ will contain the origin as well, which often imposes strong restrictions on the applicability of PIMs or of convergence results. For example, the hybrid Chebyshev-Krylov method ($\partial\Omega$ is an ellipse) requires that the origin is outside the convex hull of $\Lambda(A)$, while the hybrid Arnoldi-Faber method ($\partial\Omega$ is a polygon symmetric to the real axis) requires that $\Lambda(A)$ does not contain real numbers λ_1 and λ_2 with $\lambda_1 < 0 < \lambda_2$. Examples of convergence results for PIMs that are of limited applicability are estimates of Eiermann [16] and Nachtigal et al. [56, Theorem 4] which are only useful if $0 \notin F(A)$ and $0 \notin \Lambda_\epsilon(A)$, respectively.

To study the asymptotic performance of PIMs, and to construct PIMs which converge well, we introduce the following concepts.

Definition 2.6 *(cf. e.g. [17, p. 1156]) Suppose that $\Omega \subset \mathbf{C}$ is compact. Then the asymptotic convergence factor of Ω is defined by*

$$R(\Omega) := \limsup_{n\to\infty} \left(\inf_{p_n \in \pi_n} \max_{z \in \Omega} |p_n(z)| \right)^{1/n}. \tag{2.18}$$

Furthermore, a sequence of polynomials $p_n \in \pi_n$, $n = 0, 1, 2, \ldots$, is called asymptotically optimal on Ω, if

$$\limsup_{n\to\infty} \left(\max_{z\in\Omega} |p_n(z)| \right)^{1/n} = R(\Omega). \tag{2.19}$$

If $0 \notin \Omega$, then for each $n \geq 0$ the infimum in (2.18) is uniquely attained (see e.g. [67, Chapter 5]). If additionally Ω^C is simply connected in $\hat{\mathbf{C}}$, then $0 < R(\Omega) < 1$ (see e.g. [78, Section 4.7]). Moreover, there exists a unique exterior mapping function $\Psi \in \mathcal{E}$ with $\Psi(\mathbf{E}^C) = \Omega^C$, and $R(\Omega)$ can be computed using the inverse Ψ^{-1}:

$$R(\Omega) = \frac{1}{|\Psi^{-1}(0)|} \tag{2.20}$$

(see e.g. [78, Section 4.7]).

If Ω is convex or of bounded boundary rotation, the normalized Faber polynomials for Ω,

$$\tilde{F}_n(z) := \frac{F_n(z)}{F_n(0)} \in \pi_n, \quad n \geq 0, \tag{2.21}$$

are asymptotically optimal (cf. [14, 69]).

Hence, the asymptotic convergence factor as well as asymptotically optimal polynomials for the inclusion set Ω can be derived using the exterior mapping function of Ω. The construction and analysis of PIMs can therefore be treated in a unified framework using conformal maps.

3 Inclusion sets generated by the conformal 'bratwurst' maps

This section deals with the derivation of exterior mapping functions of 'bratwurst' shape sets in the complex plane. Geometrically, a 'bratwurst' shape set is the result of deforming an ellipse with foci on the real axis, by conformally mapping the real axis onto the unit circle. It turns out that such sets are well suited to serve as general purpose inclusion sets. The exterior mapping functions are compositions of elementary Möbius transformations and the Joukowsky map. Therefore, quantities associated with the 'bratwurst' shape sets, in particular the transfinite diameter and the asymptotic convergence factor, are easily computable.

In Section 3.1, we derive the exterior mapping functions. We then formally define the class of 'bratwurst' shape sets and show properties of these sets in Section 3.2. Section 3.3 contains several numerical examples.

The derivation of the exterior mapping functions is a joint work with Tino Koch and first appeared in [48].

3.1 Derivation of the maps

To explain our motivation for deriving exterior mapping functions of 'bratwurst' shape sets, let us briefly review the inclusion sets that have been used in the literature on PIMs for non-hermitian linear systems. These can be roughly divided into two classes.

The first class consists of convex sets with analytic boundaries and very simple geometry, for example disks [66, Theorem 5] and sets bounded by ellipses [24, 25]. The geometry of these sets allows an easy computation of asymptotic convergence factors and of asymptotically optimal polynomials. However, the convexity leads to limited applicability of the resulting theorems and methods because of the general constraint that the origin has to be outside the inclusion set.

The second class consists of non-convex sets with more complicated geometry and non-analytic boundaries, for example annular sectors [8, 55] and (non-convex) sets bounded by polygons [53, 69]. While the non-convexity of these sets leads to broader applicability, the non-analytic boundary makes the computation of the exterior mapping function difficult. Often, the exterior mapping function is determined by the Schwarz-Christoffel formula, and thus the computation of asymptotic convergence factors and asymptotically optimal polynomials involves quantities that are defined implicitly through integral equations. The latter can only be solved numerically, which is expensive and sometimes susceptible to round-off errors. We refer to [41, Section 16.10] for fundamental numerical difficulties with Schwarz-Christoffel integrals and to [55, p. 12] for particular difficulties encountered with annular sectors.

Our approach leads to inclusion sets that combine the best of both worlds, namely broad applicability and inexpensive computability of associated quantities. In particular, our sets are non-convex and (in cases of non-empty interiors) have an analytic boundary.

Theorem 3.1 *Suppose that $\lambda_m \in \partial\mathbf{E}$ and $\phi \in (0, 2\pi)$ are given. Define*

$$\epsilon_{max} := \tan\frac{\phi}{4}\left(1+\tan\frac{\phi}{8}\right) \quad \textit{and} \quad P := \tan\frac{\phi}{4} + \left(\cos\frac{\phi}{4}\right)^{-1}. \tag{3.1}$$

Then for each $\epsilon \in [0, \epsilon_{max})$, the function

$$\Psi_\epsilon(w) = \frac{(w-\lambda_m N_\epsilon)(w-\lambda_m M_\epsilon)}{(N_\epsilon - M_\epsilon)w + \lambda_m(N_\epsilon M_\epsilon - 1)}, \tag{3.2}$$

where

$$N_\epsilon = \frac{1}{2}\left(\frac{P}{1+\epsilon} + \frac{1+\epsilon}{P}\right) \quad \textit{and} \quad M_\epsilon = \frac{(1+\epsilon)^2 - 1}{2\tan\frac{\phi}{4}(1+\epsilon)}, \tag{3.3}$$

is an element of $\mathcal{E}$. Furthermore, if we define $\Omega_\epsilon := \Psi_\epsilon(\mathbf{E}^C)^C$, then $\lambda_m \notin \Omega_\epsilon$ and $\left\{\lambda_m e^{i\beta} \,:\, \frac{\phi}{2} \le \beta \le 2\pi - \frac{\phi}{2}\right\} \subset \Omega_\epsilon$. The transfinite diameter and asymptotic convergence factor of Ω_ϵ are $t(\Omega_\epsilon) = 1/(N_\epsilon - M_\epsilon)$ and $R(\Omega_\epsilon) = 1/N_\epsilon$, respectively.

Proof. Our constructive proof makes extended use of Möbius transformations

$$\psi(w) = \frac{aw+b}{cw+d}, \quad a, b, c, d \in \mathbf{C}, \quad ad - bc \neq 0.$$

We exhibit the function Ψ_ϵ with the stated properties as the composition of three relatively simple conformal maps,

$$\Psi_\epsilon = \psi_2 \circ J \circ \psi_1.$$

We first give an overview of the construction:

Suppose that $\epsilon \geq 0$. Then ψ_1 is a certain Möbius transformation that maps $\mathbf{E}^C$ onto $((1+\epsilon)\mathbf{E})^C$. We next apply the Joukowsky map

$$J(w) = \frac{1}{2}\left(w + w^{-1}\right). \tag{3.4}$$

For each $r \geq 1$, J maps the set $|w| > r$ conformally onto the exterior of an ellipse with foci ± 1 and semiaxes $\left(r \pm r^{-1}\right)/2$ (see e.g. [54, Chapter 10]). Hence, $(J \circ \psi_1)(\mathbf{E}^C)$ is the exterior of the ellipse with foci ± 1 and semiaxes $\left((1+\epsilon) \pm (1+\epsilon)^{-1}\right)/2$. Finally, the Möbius transformation ψ_2 maps the real axis onto the unit circle such that the image of $(J \circ \psi_1)(\mathbf{E}^C)$ under ψ_2 is Ω_ϵ^C, and Ω_ϵ has the stated property.

In detail, since ψ_2 is conformal it transforms one of the halfplanes

$$\mathbf{H_U} := \{w \,:\, Im(w) > 0\} \;\text{ or }\; \mathbf{H_L} := \{w \,:\, Im(w) < 0\}$$

to $int(\mathbf{E})$. Let $\mathcal{C}(\Omega^1, \Omega^2)$ denote the set of all conformal maps from Ω^1 to Ω^2. If $\tau_1,\, \tau_2 \in \mathcal{C}(\mathbf{H_U}, int(\mathbf{E}))$, then $\tau_2 \circ \tau_1^{-1} \in \mathcal{C}(int(\mathbf{E}), int(\mathbf{E}))$. It is a classical result in complex function theory that each such mapping is of the form

$$f(w) = e^{i\theta}\frac{w+c}{\bar{c}w+1}, \quad |c| < 1, \quad \theta \in \mathbf{R} \tag{3.5}$$

(see e.g. [44, Chapter 15]). Using the Cayley map $\tau_C(w) = (w-i)/(w+i) \in \mathcal{C}(\mathbf{H_U}, int(\mathbf{E}))$, a direct calculation yields

$$\mathcal{C}(\mathbf{H_U}, int(\mathbf{E})) = \left\{ e^{i\theta} \frac{w-s}{w-\bar{s}} \ : \ \theta \in \mathbf{R},\ s \in \mathbf{H_U} \right\}. \tag{3.6}$$

The function $\psi(w) = -w \in \mathcal{C}(\mathbf{H_U}, \mathbf{H_L})$, shows that

$$\mathcal{C}(\mathbf{H_L}, int(\mathbf{E})) = \left\{ e^{i\theta} \frac{w+s}{w+\bar{s}} \ : \ \theta \in \mathbf{R},\ s \in \mathbf{H_U} \right\}. \tag{3.7}$$

For our purpose we fix $\psi_2 \in \mathcal{C}(\mathbf{H_L}, int(\mathbf{E}))$. To guarantee that $\Omega_\epsilon := \Psi_\epsilon(\mathbf{E}^C)^C$ satisfies $\lambda_m \notin \Omega_\epsilon$ and $\left\{ \lambda_m e^{i\beta} \ : \ \frac{\phi}{2} \leq \beta \leq 2\pi - \frac{\phi}{2} \right\} \subseteq \Omega_\epsilon$, we set

$$\psi_2(-1) = \lambda_m e^{-i\phi/2}, \quad \psi_2(1) = \lambda_m e^{i\phi/2}, \quad \psi_2(\infty) = \lambda_m. \tag{3.8}$$

These properties yield

$$\psi_2(w) = \lambda_m \frac{w + i \tan \frac{\phi}{4}}{w - i \tan \frac{\phi}{4}}. \tag{3.9}$$

Observe that $\psi_2 \left(i \tan \frac{\phi}{4} \right) = \infty$. To have $\Psi_\epsilon \in \mathcal{E}$, we have to satisfy $(a) - (d)$ of Definition 2.2. For (a) and (d), we require

$$\Psi_\epsilon(\infty) = \infty \ \text{ and } \ 0 \in \Psi_\epsilon(\mathbf{E}^C). \tag{3.10}$$

We define

$$P_\infty := J^{-1} \left(i \tan \frac{\phi}{4} \right)$$

by taking the branch of the inverse Joukowsky map, $J^{-1}(z) = z + (z^2-1)^{1/2}$, that induces the greatest possible modulus $|P_\infty|$, and compute

$$P_\infty = i \left(\tan \frac{\phi}{4} + \left(\cos \frac{\phi}{4} \right)^{-1} \right). \tag{3.11}$$

The preimages of ∞ and 0 under ψ_2 are complex conjugates and so they are under $\psi_2 \circ J$, i.e.,

$$(\psi_2 \circ J)(P_\infty) = \infty \ \text{ and } \ (\psi_2 \circ J)(\overline{P_\infty}) = 0. \tag{3.12}$$

To actually satisfy both conditions in (3.10), which now are closely related by P_∞ and $\overline{P_\infty}$, we require $P_\infty \in \psi_1(\mathbf{E}^C) = \{ w \ : \ |w| > 1 + \epsilon \}$. This holds if ϵ satisfies

$$0 \leq \epsilon < |P_\infty| - 1 = \tan \frac{\phi}{4} \left(1 + \tan \frac{\phi}{8} \right) =: \epsilon_{max}. \tag{3.13}$$

The still unknown function ψ_1 can be written as

$$\psi_1(w) = (1+\epsilon)\rho(w),$$

where $\rho \in \mathcal{C}(int(\mathbf{E}), int(\mathbf{E}))$ and hence is of the form (3.5). Note that ρ has to satisfy $\rho(\infty) = P_\infty/(1+\epsilon)$, which implies $c = e^{-i\theta}(1+\epsilon)/\overline{P_\infty}$. We choose

$e^{i\theta} = -i\lambda_m^{-1}$. As shown below, this choice leads to a conformal map Ψ_ϵ that satisfies the condition (b) of Definition 2.2, i.e. $\Psi'_\epsilon(\infty) > 0$. Our choice leads to the following analytic terms for ψ_1 and ψ_1^{-1}:

$$\psi_1(w) = (1+\epsilon)\frac{iP_\infty w + \lambda_m(1+\epsilon)}{i(1+\epsilon)w + \lambda_m \overline{P_\infty}} \quad \text{and} \quad \psi_1^{-1}(w) = \frac{i\lambda_m}{1+\epsilon}\frac{\overline{P_\infty}w - (1+\epsilon)^2}{w + \overline{P_\infty}}.$$

So far, we have constructed

$$\Psi_\epsilon(w) = (\psi_2 \circ J \circ \psi_1)(w) = \lambda_m \frac{\psi_1^2(w) + 2i\tan\frac{\phi}{4}\psi_1(w) + 1}{\psi_1^2(w) - 2i\tan\frac{\phi}{4}\psi_1(w) + 1}. \tag{3.14}$$

According to (3.14), Ψ_ϵ is a rational function of w, the denominator and numerator of which are polynomials of degree at most 2. To simplify (3.14), we ask for the preimages of 0 and ∞ under Ψ_ϵ or, in other words, for the roots of the numerator and denominator. From (3.12) and the fact that $J(w) = J(w^{-1})$ for all $w \in \hat{\mathbf{C}}$ follows that

$$\begin{aligned} 0 &= \Psi_\epsilon\left(\psi_1^{-1}(\overline{P_\infty})\right) = \Psi_\epsilon\left(\psi_1^{-1}\left(\frac{1}{\overline{P_\infty}}\right)\right) \quad \text{and} \\ \infty &= \Psi_\epsilon\left(\psi_1^{-1}(P_\infty)\right) = \Psi_\epsilon\left(\psi_1^{-1}\left(\frac{1}{P_\infty}\right)\right) = \Psi_\epsilon(\infty). \end{aligned}$$

Hence, for the zeros of Ψ_ϵ we have

$$N_\epsilon := \frac{1}{\lambda_m}\psi_1^{-1}(\overline{P_\infty}) \quad \text{and} \quad M_\epsilon := \frac{1}{\lambda_m}\psi_1^{-1}(1/\overline{P_\infty}). \tag{3.15}$$

An easy calculation shows that N_ϵ and M_ϵ are as given in (3.3). Clearly, only one of these zeros lies in $\mathbf{E}^C$. In fact, for all $\epsilon \in [0, \epsilon_{max})$ we have

$$0 \le M_\epsilon < 1 < N_\epsilon. \tag{3.16}$$

Similarly, since ∞ is a fixedpoint of Ψ_ϵ, the second preimage of ∞ lies in $int(\mathbf{E})$. It can be expressed as a surprisingly simple rational function of the zeros,

$$\psi_1^{-1}\left(\frac{1}{P_\infty}\right) = \frac{\lambda_m P}{1+\epsilon}\,\frac{(1+\epsilon)^2+1}{P^2+1} = \lambda_m \frac{1 - N_\epsilon M_\epsilon}{N_\epsilon - M_\epsilon}, \tag{3.17}$$

where $P := |P_\infty|$. We use (3.15) and (3.17) to transform (3.14) into

$$\Psi_\epsilon(w) = \frac{(w - \lambda_m N_\epsilon)(w - \lambda_m M_\epsilon)}{(N_\epsilon - M_\epsilon)w + \lambda_m(N_\epsilon M_\epsilon - 1)}.$$

We can easily check that

$$\Psi'_\epsilon(\infty) = \lim_{w\to\infty}\frac{\Psi_\epsilon(w)}{w} = \frac{1}{N_\epsilon - M_\epsilon} > 0$$

(cf.(3.16)). Thus, Ψ_ϵ satisfies (b) of Definition 2.2 and $t(\Omega_\epsilon) = 1/(N_\epsilon - M_\epsilon)$. Furthermore, referring to (c) of Definition 2.2, we remark that $\Psi_\epsilon(\mathbf{E}^C)$ obviously is simply connected in $\hat{\mathbf{C}}$. To complete the proof we apply (2.20), which shows that $R(\Omega_\epsilon) = 1/N_\epsilon$. □

Example 3.2 (Circular arc) Suppose that $\lambda_m \in \partial\mathbf{E}$ and $\phi \in (0, 2\pi)$. If we choose $\epsilon = 0$, then $N_\epsilon = \left(\cos\frac{\phi}{4}\right)^{-1}$, $M_\epsilon = 0$ and Ψ_ϵ in (3.2) reduces to

$$\Psi_\epsilon(w) = \frac{w(w\cos\frac{\phi}{4} - \lambda_m)}{w - \lambda_m \cos\frac{\phi}{4}}. \tag{3.18}$$

This function maps $\mathbf{E}^C$ onto the exterior of the circular arc $\{\lambda_m e^{i\beta} \;:\; \frac{\phi}{2} \leq \beta \leq 2\pi - \frac{\phi}{2}\}$. Allowing for the difference in notation, Ψ_ϵ in (3.18) is the same as the mapping function derived by Coleman and Myers for the same purpose (cf. [8, Formula (2.19)]). In light of the uniqueness guaranteed by the Riemann Mapping Theorem, this is certainly no surprise.

3.2 Definition and properties of the 'bratwurst' shape sets

We now use the exterior mapping functions of Theorem 3.1 to define a class of sets which are well suited to serve as general purpose inclusion sets.

Definition 3.3 *Suppose that $\lambda_m \in \partial\mathbf{E}$ and $\phi \in (0, 2\pi)$ are given. Define the corresponding ϵ_{max} as in (3.1). Using Ψ_ϵ as in (3.2), we define the class of sets $\mathbf{B}(\lambda_m, \phi)$ by*

$$\mathbf{B}(\lambda_m, \phi) := \left\{ s\Omega_\epsilon \;:\; \Omega_\epsilon = \left(\Psi_\epsilon(\mathbf{E}^C)\right)^C,\ \epsilon \in [0, \epsilon_{max}),\ s \in (0, \infty) \right\}. \tag{3.19}$$

Lemma 3.4 *For each $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ the following hold:*

(a) $0 \notin s\Omega_\epsilon$ *and* $\lambda_m \notin s\Omega_\epsilon$.

(b) $(s\Omega_\epsilon)^C$ *is simply connected in* $\hat{\mathbf{C}}$.

(c) The transfinite diameter of $s\Omega_\epsilon$ *is* $t(s\Omega_\epsilon) = st(\Omega_\epsilon) = s/(N_\epsilon - M_\epsilon)$.

(d) $R(s\Omega_\epsilon) = R(\Omega_\epsilon) = 1/N_\epsilon$.

(e) If $\Omega_\epsilon = \left(\Psi_\epsilon(\mathbf{E}^C)\right)^C$, *then* $\partial(s\Omega_\epsilon) = s\Psi_\epsilon(\partial\mathbf{E})$ *and, for* $\epsilon \in (0, \epsilon_{max})$, $\partial(s\Omega_\epsilon)$ *is an analytic Jordan curve.*

Proof. If we define $\psi(w) := sw$, then $s\Omega_\epsilon = \left((\psi \circ \Psi_\epsilon)(\mathbf{E}^C)\right)^C$. Now the assertions $(a) - (d)$ are immediate consequences of the properties of $\psi \circ \Psi_\epsilon$.

To show (e), we can assume without loss of generality that $s = 1$. First suppose that $\epsilon = 0$. Then Ω_ϵ is a circular arc and $\Omega_\epsilon = \partial(\Omega_\epsilon) = \Psi_\epsilon(\partial\mathbf{E})$ (cf. Example 3.2). Now suppose that $\epsilon \in (0, \epsilon_{max})$. Then from the construction it is easy to see that Ψ_ϵ can be extended holomorphically into a full neighborhood of $\partial\mathbf{E}$. To see this, consider that $\Psi_\epsilon(\infty) = \infty$ and that the second preimage of ∞ lies in $int(\mathbf{E})$. Hence, Ψ_ϵ is analytic on $\mathbf{E}^C \cup \mathbf{A}$, where $\mathbf{A} := \{z \;:\; r < |z| \leq 1\}$, for some $r < 1$. This shows in particular that $\partial(\Omega_\epsilon) = \Psi_\epsilon(\partial\mathbf{E})$ is an analytic Jordan curve. $\square$

As expected, $t(s\Omega_\epsilon)$ and $R(\Omega_\epsilon)$ are independent of λ_m, i.e. invariant under rotation. Moreover, $R(\Omega_\epsilon)$ is invariant under scaling. The simplicity of our approach demonstrates the dependence of the transfinite diameter and asymptotic convergence factor on the parameters determining the inclusion set, particularly on ϵ and ϕ. We briefly analyze this dependence for $R(\Omega_\epsilon)$: First note that the modulus of P in (3.1) is strictly increasing with increasing angle ϕ. Hence, for each fixed ϵ, $R(\Omega_\epsilon)$ strictly decreases with increasing ϕ. The same is true for fixed ϕ and decreasing ϵ. In this case, as ϵ ranges from ϵ_{max} to zero, $R(\Omega_\epsilon)$ ranges from one to $\cos\frac{\phi}{4}$.

Lemma 3.5 *Each $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ is of bounded boundary rotation.*

Proof. Suppose that some $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$, $\epsilon \in (0, \epsilon_{max})$, is given. Clearly, $V(s\Omega_\epsilon) = V(\Omega_\epsilon)$ for all s and hence we can assume without loss of generality that $s = 1$. We have shown in Lemma 3.4 that $\Psi_\epsilon(\partial\mathbf{E}) = \partial\Omega_\epsilon$ is an analytic Jordan curve. Therefore, the boundary rotation of Ω_ϵ is given by

$$V(\Omega_\epsilon) = \int_0^{2\pi} \left| 1 + Re\left(\frac{e^{it}\Psi_\epsilon''(e^{it})}{\Psi_\epsilon'(e^{it})}\right)\right| dt$$

(cf. e.g. [58]), where we denote

$$\Psi_\epsilon'(e^{it}) = \frac{d}{dw}\Psi_\epsilon(w)\bigg|_{z=e^{it}}.$$

But since Ψ_ϵ is conformal, $\Psi_\epsilon'(w) \neq 0$ on $\partial\mathbf{E}$, so that $Re(w\Psi_\epsilon''(w)/\Psi_\epsilon'(w))$ defines a continuous function on $\partial\mathbf{E}$. We easily conclude that $V(\Omega_\epsilon) < \infty$, i.e. Ω_ϵ is of bounded boundary rotation. To complete the proof, we remark that for $\epsilon = 0$, Ω_ϵ is a circular arc and hence obviously of bounded boundary rotation. A formula for $V(\Omega_\epsilon)$, $\epsilon = 0$, is given in [8, p. 200]. □

We conclude from Lemma 3.5, that the normalized Faber polynomials (cf. (2.21)) for $s\Omega_\epsilon$ are asymptotically optimal. Hence they are well suited to serve as residual polynomials in Phase II of a hybrid PIM. We make use of this fact in Section 4.

How do we determine an inclusion set $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ for a given bounded set $\mathbf{S} \subset \mathbf{C}$ not containing the origin? Ideally, we would like to compute the inclusion set with the smallest asymptotic convergence factor at minimal cost. However, due to non-convexity and numerous degrees of freedom of the elements of $\mathbf{B}(\lambda_m, \phi)$, both constraints are difficult to satisfy simultaneously, except for special cases such as $\mathbf{S} \subset \partial\mathbf{E}$. Below we outline a strategy that produced satisfactory results in our applications, while compromising between computational cost and optimality of the asymptotic convergence factor:

Recall that each element of $\mathbf{B}(\lambda_m, \phi)$ depends on four parameters, λ_m, ϕ, s and ϵ. We first determine ϕ, λ_m and s. Using the arguments of the elements of $\mathbf{S}$, we compute the largest $\phi \in (0, 2\pi)$ and corresponding $\lambda_m \in \partial\mathbf{E}$ such that

$$\left\{ r\lambda_m e^{i\beta} \; : \; -\frac{\phi}{2} < \beta < \frac{\phi}{2}, \; r \geq 0 \right\} \cap \mathbf{S} = \emptyset.$$

We call ϕ and λ_m the largest gap in $\mathbf{S}$ and the center of that gap, respectively. Next, observe that $s\Omega_\epsilon$ intersects the interior and exterior of $s\mathbf{E}$. Therefore $\mathbf{S}$ should have non-empty intersections with the interior and exterior of $s\mathbf{E}$ as well. If $\mathbf{S}$ is contained in an annulus with inner radius r_1 and outer radius r_2, we choose $s = \sqrt{r_1 r_2}$.

Given λ_m, ϕ and s, we compute the unique smallest $\epsilon \geq 0$ so that $(s\psi_2)^{-1}(\mathbf{S})$ (cf. (3.8)) is contained in the ellipse with foci ± 1 and semiaxes $(r \pm r^{-1})/2$, $r = 1+\epsilon$. If $\epsilon < \epsilon_{max}$, where ϵ_{max} as in (3.1) depends on ϕ, $\mathbf{S} \subset s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ is satisfied. Otherwise, no 'bratwurst' shape inclusion set exists for the current choice of parameters. For improvement, we might decrease the gap and compute the new corresponding ϵ. If the results are still unsatisfactory, s needs to be adjusted. Example 3.8 shows applications of this strategy.

3.3 Numerical examples

Example 3.6 We show what typical elements of $\mathbf{B}(\lambda_m, \phi)$ look like and how the transfinite diameter and asymptotic convergence factor change with changing parameter ϵ. Suppose that $\lambda_m = -1$ and $\phi = \pi/2$. Then $\epsilon_{max} \approx 0.4966$. Figure 1 shows $\partial\Omega_\epsilon$ for $\Omega_\epsilon \in \mathbf{B}(-1, \pi/2)$ and $\epsilon = 0$ (solid), 0.05 (dotted), 0.1 (dashed), 0.15 (dashdot); λ_m and the origin are indicated as '$*$'. The corresponding transfinite diameters and asymptotic convergence factors were computed using (c) and (d) of Lemma 3.4 and are listed in the following table:

ϵ	0	0.05	0.1	0.15
$t(\Omega_\epsilon)$	0.9239	1.0575	1.2235	1.4360
$R(\Omega_\epsilon)$	0.9239	0.9403	0.9544	0.9633

Example 3.7 We compare elements of $\mathbf{B}(\lambda_m, \phi)$ with annular sectors

$$\mathbf{S}(\lambda_m, \phi, h) := \left\{ r\lambda_m e^{i\beta} \; : \; \frac{\phi}{2} \leq \beta \leq 2\pi - \frac{\phi}{2}, \;\; h \leq r \leq 1 \right\},$$

which have been proposed in the literature as general purpose inclusion sets. Using Gaussian quadrature to evaluate exterior mapping functions of annular sectors derived by Coleman and Myers [8] with the Schwarz-Christoffel approach, Modersitzki and Opfer compute transfinite diameters of $\mathbf{S}(1, \phi, 1/2)$ for some values of ϕ. To obtain a high accuracy, the effort for this computation is "quite high" [55, p.15]. We scale corresponding elements of $\mathbf{B}(1, \phi)$ by $s = 3/4$, choose a suitable ϵ and compute the transfinite diameter and asymptotic convergence factor using Lemma 3.4. These computations can be performed with only a few floating point operations. The particular choices are $\epsilon = 0.1$ and $\epsilon = 0.2$ for $\phi = \pi/2$ and $\phi = \pi$, respectively.

ϵ, ϕ	$R(\frac{3}{4}\Omega_\epsilon)$	$t(\frac{3}{4}\Omega_\epsilon)$	$t(\mathbf{S}(1, \phi, 1/2))$
0.1, $\pi/2$	0.9544	0.9176	0.9382
0.2, π	0.7972	0.7002	0.7587

The results about $\mathbf{S}(1, \phi, 1/2)$ are taken from [55, Table 7.1].

Plots of the inclusion sets ($\partial(\frac{3}{4}\Omega_\epsilon)$ dashed, $\partial\mathbf{S}(1, \phi, 1/2)$ solid) are shown in in Figures 2 ($\epsilon = 0.1$, $\phi = \pi/2$) and 3 ($\epsilon = 0.2$, $\phi = \pi$).

Example 3.8 We describe applications of our strategy for finding a 'bratwurst' shape inclusion set for a given matrix spectrum as outlined in Section 3.2. Information on the test matrices is given in Appendix A.

For our first application we use the matrix A=GRCAR250. We compute $\Lambda(A)$ with MATLAB's `eig` function. The spectrum is located in an annulus with inner radius $r_1 \approx 1.6137$ and outer radius $r_2 \approx 2.2643$, and we compute $s = \sqrt{r_1 r_2} \approx 1.9115$. From the arguments of the elements of $\Lambda(A)$ we compute the largest gap in $\Lambda(A)$, $\phi_L \approx 1.02\,\pi$, and the center of that gap, $\lambda_m = -1$. Next, for $j\,\phi_L$, $j = 1, 0.95, 0.9, \ldots$, ϵ_j is determined as the smallest parameter so that the ellipse with foci ± 1 and semiaxes $r \pm r^{-1}$, $r = 1 + \epsilon_j$, contains $(s\psi_2)^{-1}(\Lambda(A))$, where

$$(s\psi_2)^{-1}(z) = i \left(\tan \frac{j\,\phi_L}{4} \right) \frac{z + s\lambda_m}{z - s\lambda_m}$$

(cf. (3.8)). For each j, the asymptotic convergence factor of $s\Omega_{\epsilon_j} \in \mathbf{B}(\lambda_m, j\,\phi_L)$ is computed. This loop generally terminates when either $j = 0$, or the asymptotic convergence factor does not decrease for two consecutive steps. In our example, the loop terminates for $j = 0.85$. The final inclusion set parameters are $\phi \approx 0.87\,\pi$, $\lambda_m = -1$, $s \approx 1.9115$ and $\epsilon \approx 0.2273$. The corresponding set has asymptotic convergence factor 0.8696 and transfinite diameter 2.1301. Its boundary is shown in Figure 4; the pluses represent $\Lambda(A)$, λ_m and the origin are indicated as '$*$'.

In a second application, we consider Λ(PDE225). The same computation as for GRCAR250 yields $\lambda_m = -1$, $\phi \approx 1.27\,\pi$, $s \approx 2.4261$ and $\epsilon \approx 1.2821$. The asymptotic convergence factor and transfinite diameter of the computed inclusion set are 0.926 and 5.0, respectively. Its boundary and Λ(PDE225) are shown in Figure 5; λ_m and the origin are indicated as '$*$'.

Note that it would be difficult to include Λ(GRCAR250) or Λ(PDE225) in a set bounded by a circle or an ellipse not containing the origin.

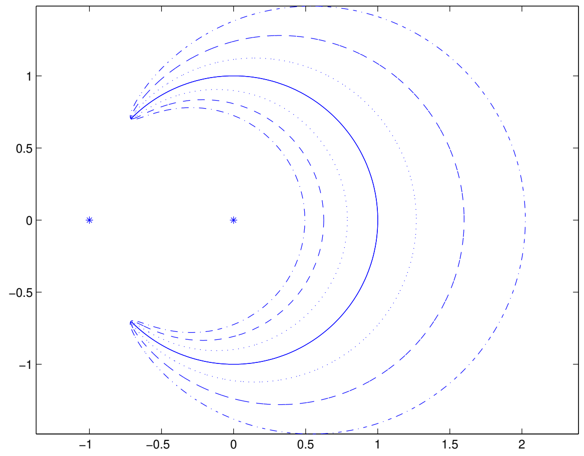

Figure 1: 'Bratwurst' shape inclusion sets for different ϵ.

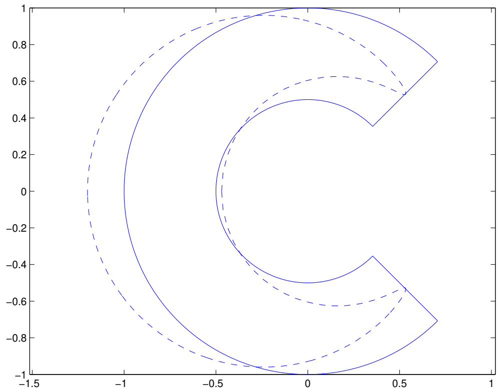

Figure 2: 'Bratwurst' shape inclusion set vs. annular sector (I).

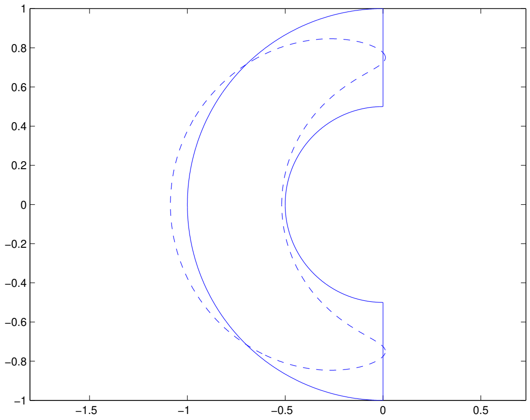

Figure 3: 'Bratwurst' shape inclusion set vs. annular sector (II).

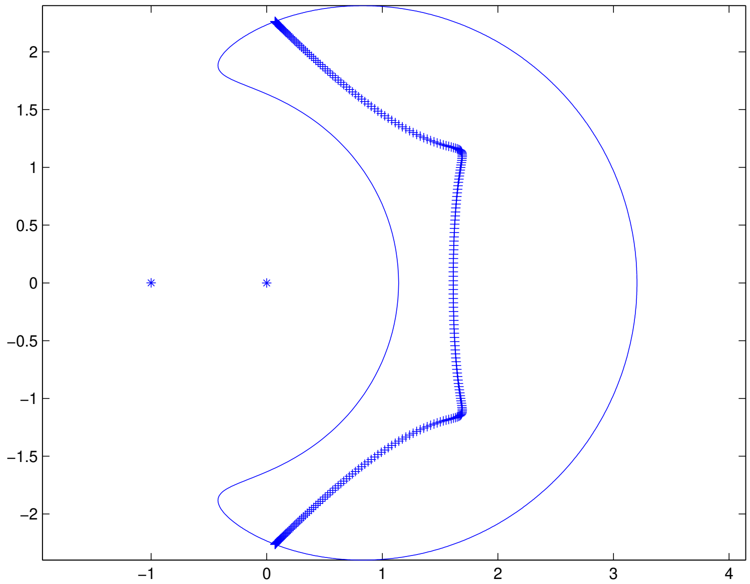

Figure 4: Inclusion of Λ(GRCAR250).

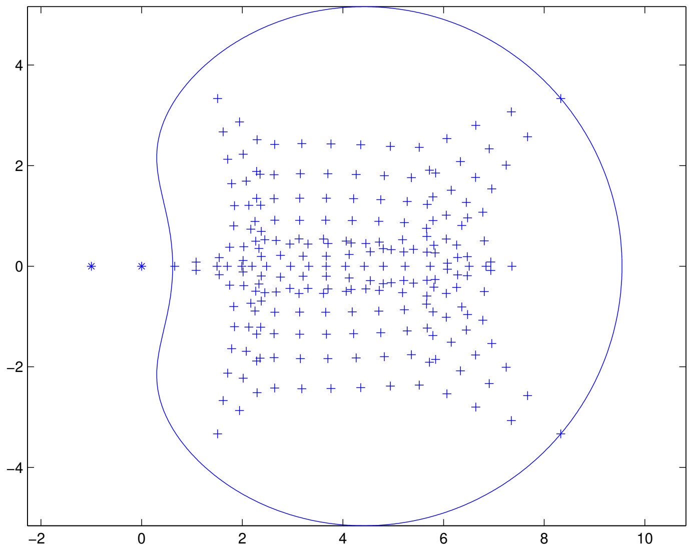

Figure 5: Inclusion of Λ(PDE225).

4 The hybrid ABF method for non-hermitian linear systems

This section contains the construction, analysis and application of a new hybrid PIM for non-hermitian linear systems. The two phases of our method applied to a linear system (1.1) can be described as follows:

Phase I: If a priori spectral information about A is available, determine a suitable inclusion set $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ from that information. Otherwise, run $n << N$ steps of the Arnoldi process, which yields the matrix H_n (cf. (2.6)). Compute $\Lambda(H_n) \approx \Lambda(A)$, or, if it is known that A is highly nonnormal, use H_n to approximate a pseudospectrum of A. Then determine an inclusion set of $\Lambda(H_n)$ or of the approximate pseudospectrum.

Phase II: Iterate with polynomials, so that the residual polynomials are normalized Faber polynomials for $s\Omega_\epsilon$. If the iteration does not converge well, go back to Phase I and change the inclusion set (with or without obtaining more information about A first).

According to its three main ingredients, we refer to our method as hybrid Arnoldi-Bratwurst-Faber method, abbreviated by the acronym ABF method.

The application of the Arnoldi process in Phase I of a hybrid PIM is a well-explored subject (cf. Section 2). Furthermore, the elements of $\mathbf{B}(\lambda_m, \phi)$ as well as a strategy for finding a suitable element are discussed in Section 3. Hence, the main work in the construction of the ABF method now is required by Phase II, i.e. by deriving a formula for the iteration polynomials.

To motivate the reader to follow us through this derivation, we list upfront the main advantages of the ABF method:

- The iteration polynomials satisfy a three-term recurrence.
- The computation of each iteration polynomial requires only one matrix-vector product plus a few $\mathcal{O}(N)$ operations, and no inner product.
- r_n can be obtained economically, without explicitly computing x_n.
- The ABF method is applicable to a wide range of nonsingular linear systems. In particular, the 'bratwurst' shape sets have the potential to include spectra or pseudospectra of indefinite matrices.

The recurrences for Phase II are developed in Sections 4.1 and 4.2. Section 4.3 contains a listing of the complete algorithm of Phase II, implementational details and a convergence theorem. We furthermore comment on how to choose the inclusion set. Numerical experiments are described in Section 4.4.

4.1 Faber polynomials for the inclusion sets

We begin the construction of Phase II of the ABF method by deriving a recurrence for the Faber polynomials for $s\Omega_\epsilon$. In this derivation we are led by the following idea:

The Faber polynomials for sets bounded by ellipses coincide with the classical Chebyshev polynomials (see e.g. [67, Chapter 2]), which are given by

$$C_n(z) = \frac{w^n + w^{-n}}{2}, \quad z = \frac{w + w^{-1}}{2}, \quad n \geq 0. \tag{4.1}$$

It is well known, that the Chebyshev polynomials can be generated by a three-term recurrence of the form

$$C_0(z) \equiv 1, \;\; C_1(z) = z, \;\; C_{n+1}(z) = 2zC_n(z) - C_{n-1}(z), \;\; n \geq 1. \tag{4.2}$$

Because $\partial(s\Omega_\epsilon)$ is the image of an ellipse under a certain Möbius transformation, the recurrence (4.2) should be the key ingredient in a recurrence for the Faber polynomials for $s\Omega_\epsilon$.

We state the result of our investigation in the following theorem.

Theorem 4.1 *Suppose that some $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ is given as in Definition 3.3. Define the corresponding N_ϵ and M_ϵ as in (3.3), and let Ψ_ϵ as in (3.2) denote the exterior mapping function of Ω_ϵ. Define*

$$S_\epsilon := \frac{N_\epsilon M_\epsilon - 1}{N_\epsilon - M_\epsilon}, \quad t_\epsilon := \frac{s}{N_\epsilon - M_\epsilon}, \quad Q_\epsilon := (S_\epsilon + N_\epsilon)(S_\epsilon + M_\epsilon), \tag{4.3}$$

and

$$V(z) := \lambda_m^2 N_\epsilon M_\epsilon - \lambda_m t_\epsilon^{-1} S_\epsilon z, \quad \text{for} \;\; z \in \mathbf{C}. \tag{4.4}$$

Let z_0 denote the zero of $V(z)$ and let $C_n(z)$ denote the nth Chebyshev polynomial as in (4.1). Then for $n \geq 1$, the Faber polynomials for $s\Omega_\epsilon$ satisfy

$$F_n(z) = 2(V(z))^{n/2} C_n\left(\frac{(V(z))^{-1/2}(\lambda_m(N_\epsilon + M_\epsilon) + t_\epsilon^{-1} z)}{2}\right) - (-\lambda_m S_\epsilon)^n \tag{4.5}$$

for $z \in \mathbf{C} \setminus \{z_0\}$ and $F_n(z_0) = \lambda_m^n (S_\epsilon^{-1} Q_\epsilon - S_\epsilon)^n - (-\lambda_m S_\epsilon)^n$.

Proof. First note that

$$\Psi(w) := s\Psi_\epsilon(w) = t_\epsilon \frac{(w - \lambda_m N_\epsilon)(w - \lambda_m M_\epsilon)}{w + \lambda_m S_\epsilon}$$

is the exterior mapping function of $s\Omega_\epsilon$. It is easy to show that

$$\Psi(w) = t_\epsilon \left[w - \lambda_m(N_\epsilon + M_\epsilon + S_\epsilon) + \frac{\lambda_m^2 Q_\epsilon}{w + \lambda_m S_\epsilon} \right] \tag{4.6}$$

and

$$\Psi'(w) = t_\epsilon \left[1 - \frac{\lambda_m^2 Q_\epsilon}{(w + \lambda_m S_\epsilon)^2} \right]. \tag{4.7}$$

An elementary calculation furthermore shows that

$$S_\epsilon = -\frac{\cos\frac{\phi}{4}}{2} \left(\frac{1 + (1+\epsilon)^2}{1+\epsilon} \right),$$

and hence, for $\epsilon \in [0, \epsilon_{max})$, where ϵ_{max} is defined as in (3.1),

$$0 < \cos \frac{\phi}{4} \le |\lambda_m S_\epsilon| < 1. \tag{4.8}$$

Now choose some real number $R > 1$ and define $\mathcal{L}_R := \{s\Psi_\epsilon(w) \ : \ |w| = R\}$. Then (4.6), (4.7) and (4.8) show that for each fixed $z \in int(\mathcal{L}_R)$ the function

$$f(w) := \frac{w^n \Psi'(w)}{\Psi(w) - z} \tag{4.9}$$

is analytic in $\mathbf{C} \setminus R\mathbf{E}$.

We use (2.11) to compute the Faber polynomials for $s\Omega_\epsilon$. For all $n \ge 1$ and $z \in int(\mathcal{L}_R)$, this yields

$$\begin{aligned} F_n(z) &= \frac{1}{2\pi i} \int\limits_{|w|=R} f(w)dw \\ &= \frac{1}{2\pi i} \int\limits_{|w|=R} \frac{w^n(1 - \lambda_m^2 Q_\epsilon (w + \lambda_m S_\epsilon)^{-2})dw}{[w - \lambda_m(N_\epsilon + M_\epsilon + S_\epsilon) + \lambda_m^2 Q_\epsilon (w + \lambda_m S_\epsilon)^{-1}] - t_\epsilon^{-1} z} \\ &= \frac{1}{2\pi i} \int\limits_{|u-\lambda_m S_\epsilon|=R} \frac{u^{-1}(u - \lambda_m S_\epsilon)^n (u^2 - \lambda_m^2 Q_\epsilon) du}{u^2 - [\lambda_m(N_\epsilon + M_\epsilon + 2S_\epsilon) + t_\epsilon^{-1} z]u + \lambda_m^2 Q_\epsilon}. \end{aligned} \tag{4.10}$$

In the last equation, we have used the variable transformation $w + \lambda_m S_\epsilon =: u$, $dw = du$, and we have multiplied the numerator and denominator of the integrand by u^2. Since $f(w)$ is analytic in $\mathbf{C} \setminus R\mathbf{E}$, the three finite singularities of the integrand $f(u - \lambda_m S_\epsilon)$ are inside the circle $|u - \lambda_m S_\epsilon| = R$. One singularity is $u = 0$, the other two are the zeros of the quadratic term in u,

$$u^2 - [\lambda_m(N_\epsilon + M_\epsilon + 2S_\epsilon) + t_\epsilon^{-1} z]u + \lambda_m^2 Q_\epsilon =: (u - \zeta_1)(u - \zeta_2). \tag{4.11}$$

Using (4.11) in (4.10), and applying the Residue Theorem yields

$$\begin{aligned} F_n(z) &= \frac{1}{2\pi i} \int\limits_{|u-\lambda_m S_\epsilon|=R} f(u - \lambda_m S_\epsilon) du \\ &= Res(f(u - \lambda_m S_\epsilon); \zeta_1) + Res(f(u - \lambda_m S_\epsilon); \zeta_2) \\ &\quad + Res(f(u - \lambda_m S_\epsilon); 0). \end{aligned} \tag{4.12}$$

In the computation of the residues we use that $\lambda_m^2 Q_\epsilon = \zeta_1 \zeta_2$ and get

$$\begin{aligned} Res(f(u - \lambda_m S_\epsilon); \zeta_1) &= \lim_{u \to \zeta_1} (u - \zeta_1) \frac{(u - \lambda_m S_\epsilon)^n (u^2 - \lambda_m^2 Q_\epsilon)}{u(u - \zeta_1)(u - \zeta_2)} \\ &= \frac{(\zeta_1 - \lambda_m S_\epsilon)^n (\zeta_1^2 - \lambda_m^2 Q_\epsilon)}{\zeta_1(\zeta_1 - \zeta_2)} = (\zeta_1 - \lambda_m S_\epsilon)^n. \end{aligned}$$

The two other residues are computed similarly, so that (4.12) leads to

$$F_n(z) = (\zeta_1 - \lambda_m S_\epsilon)^n + (\zeta_2 - \lambda_m S_\epsilon)^n - (-\lambda_m S_\epsilon)^n. \tag{4.13}$$

Note that

$$(\zeta_1 - \lambda_m S_\epsilon)(\zeta_2 - \lambda_m S_\epsilon) = \lambda_m^2 N_\epsilon M_\epsilon - \lambda_m t_\epsilon^{-1} S_\epsilon z =: V(z). \tag{4.14}$$

Since $\lambda_m t_\epsilon^{-1} S_\epsilon \neq 0$, $V(z)$ is a polynomial of exact degree one with zero $z_0 := \lambda_m t_\epsilon S_\epsilon^{-1} N_\epsilon M_\epsilon$. For $z \neq z_0$, we define

$$\tilde{\zeta}_1 := (V(z))^{-1/2}(\zeta_1 - \lambda_m S_\epsilon) \text{ and } \tilde{\zeta}_2 := (V(z))^{-1/2}(\zeta_2 - \lambda_m S_\epsilon). \tag{4.15}$$

Then $\tilde{\zeta}_1\tilde{\zeta}_2 = 1$, i.e. $\tilde{\zeta}_2 = \tilde{\zeta}_1^{-1}$.

We apply this in (4.13) and get

$$\begin{aligned} F_n(z) &= \left((V(z))^{1/2}\tilde{\zeta}_1\right)^n + \left((V(z))^{1/2}\tilde{\zeta}_1^{-1}\right)^n - (-\lambda_m S_\epsilon)^n \\ &= (V(z))^{n/2}\left(\tilde{\zeta}_1^{\,n} + \tilde{\zeta}_1^{-n}\right) - (-\lambda_m S_\epsilon)^n \\ &= 2(V(z))^{n/2} C_n\left(\frac{\tilde{\zeta}_1 + \tilde{\zeta}_1^{-1}}{2}\right) - (-\lambda_m S_\epsilon)^n, \end{aligned} \tag{4.16}$$

for all $z \in (int(\mathcal{L}_R) \setminus \{z_0\})$ and $n \geq 1$. In the last equation we employed the nth Chebyshev polynomial, specifically

$$2C_n(z) = \tilde{\zeta}_1^n + \tilde{\zeta}_1^{-n}, \text{ where } z = \frac{\tilde{\zeta}_1 + \tilde{\zeta}_1^{-1}}{2}.$$

Because $R > 1$ can be chosen arbitrarily, (4.16) in fact holds for all $z \in \mathbf{C} \setminus \{z_0\}$ and $n \geq 1$. To simplify (4.16), we compute

$$\begin{aligned} \tilde{\zeta}_1 + \tilde{\zeta}_1^{-1} &= \tilde{\zeta}_1 + \tilde{\zeta}_2 = (V(z))^{-1/2}(\zeta_1 + \zeta_2 - 2\lambda_m S_\epsilon) \\ &= (V(z))^{-1/2}(\lambda_m(N_\epsilon + M_\epsilon) + t_\epsilon^{-1} z), \end{aligned}$$

which shows (4.5).

We finally have to compute $F_n(z_0)$. Inserting the definition of z_0 into (4.11) gives

$$u^2 - \lambda_m(S_\epsilon^{-1} Q_\epsilon + S_\epsilon) + \lambda_m^2 Q_\epsilon = (u - \zeta_1)(u - \zeta_2),$$

and therefore $\zeta_1 = \lambda_m S_\epsilon^{-1} Q_\epsilon$, $\zeta_2 = \lambda_m S_\epsilon$. Applying this to (4.13) shows that $F_n(z_0)$ has the stated form. □

As expected, the Faber polynomials for $s\Omega_\epsilon$ are closely related to the classical Chebyshev polynomials for sets bounded by ellipses. Theorem 4.1 reveals this relation and is thus interesting from a theoretical point of view. But the relation (4.5) is certainly not well suited for the iterative solution of linear systems, since it involves square roots and negative exponents. We therefore proceed to derive a simpler, more efficiently computable recurrence for the $F_n(z)$.

We first apply $C_0(z) \equiv 1$ and $C_1(z) = z$ to (4.5) and get

$$F_0(z) \equiv 1 \text{ and } F_1(z) = t_\epsilon^{-1} z + \lambda_m(N_\epsilon + M_\epsilon + S_\epsilon), \tag{4.17}$$

where $F_0(z) \equiv 1$ holds by definition of the Faber polynomials (cf. Section 2.2). Now denote $y := (V(z))^{-1/2}(\lambda_m(N_\epsilon + M_\epsilon) + t_\epsilon^{-1} z)/2$, then (4.5) can be written as

$$F_n(z) = 2(V(z))^{n/2} C_n(y) - (-\lambda_m S_\epsilon)^n.$$

With the three-term recurrence (4.2), we get

$$\begin{aligned}
F_{n+1}(z) + (-\lambda_m S_\epsilon)^{n+1} &= 2(V(z))^{(n+1)/2} C_{n+1}(y) \\
&= 2(V(z))^{(n+1)/2} \left[2yC_n(y) - C_{n-1}(y)\right] \\
&= 2\Big[(V(z))^{1/2} y \left(2(V(z))^{n/2} C_n(y)\right) \\
&\quad -\frac{V(z)}{2}\left(2(V(z))^{(n-1)/2} C_{n-1}(y)\right)\Big] \\
&= \left(F_1(z) + (-\lambda_m S_\epsilon)\right)\left(F_n(z) + (-\lambda_m S_\epsilon)^n\right) \\
&\quad -V(z)\left(F_{n-1}(z) + (-\lambda_m S_\epsilon)^{n-1}\right). \qquad (4.18)
\end{aligned}$$

(4.17) and (4.18) constitute a three-term recurrence for computing the Faber polynomials for $s\Omega_\epsilon$, not involving negative exponents or square roots. Clearly, this three-term recurrence holds for all $z \in \mathbf{C}$, and no separate formula for $F_n(z_0)$ is required. In particular, we have proved the following lemma.

Lemma 4.2 *In the notation of Theorem 4.1, let $F_n(z)$ denote the nth Faber polynomial for $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$. Then the polynomials*

$$\hat{F}_n(z) := F_n(z) + (-\lambda_m S_\epsilon)^n, \quad n \geq 0, \qquad (4.19)$$

satisfy

$$\hat{F}_0(z) \equiv 2, \quad \hat{F}_1(z) = t_\epsilon^{-1} z + \lambda_m(N_\epsilon + M_\epsilon), \qquad (4.20)$$

and

$$\hat{F}_{n+1}(z) = \hat{F}_1(z)\hat{F}_n(z) - V(z)\hat{F}_{n-1}(z), \quad n \geq 1. \qquad (4.21)$$

We next derive an upper bound for the maximum norm of the normalized Faber polynomials (2.21) for $s\Omega_\epsilon$. The proof of Theorem 4.1 shows that

$$F_n(0) = (\zeta_1 - \lambda_m S_\epsilon)^n + (\zeta_2 - \lambda_m S_\epsilon)^n - (-\lambda_m S_\epsilon)^n \qquad (4.22)$$

(cf. (4.13)), where

$$u^2 - [\lambda_m(N_\epsilon + M_\epsilon + 2S_\epsilon)]u + \lambda_m^2 Q_\epsilon =: (u - \zeta_1)(u - \zeta_2) \qquad (4.23)$$

(cf. (4.11)). Since $Q_\epsilon = (S_\epsilon + N_\epsilon)(S_\epsilon + M_\epsilon)$, it is easy to see that $\zeta_1 = \lambda_m(S_\epsilon + N_\epsilon)$, $\zeta_2 = \lambda_m(S_\epsilon + M_\epsilon)$, and thus

$$F_n(0) = \lambda_m^n \left(N_\epsilon^n + M_\epsilon^n - (-S_\epsilon)^n\right). \qquad (4.24)$$

From (3.16) and (4.8) follows that $|F_n(0)| = N_\epsilon^n + M_\epsilon^n - (-S_\epsilon)^n$ for $n \geq 0$. Using this and (2.13) leads to the following lemma.

Lemma 4.3 *In the notation of Theorem 4.1, let $F_n(z)$ denote the nth Faber polynomial for $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$. Then*

$$\max_{z \in s\Omega_\epsilon} \left|\tilde{F}_n(z)\right| = \max_{z \in s\Omega_\epsilon} \left|\frac{F_n(z)}{F_n(0)}\right| \leq \frac{V(s\Omega_\epsilon)}{\pi\left(N_\epsilon^n + M_\epsilon^n - (-S_\epsilon)^n\right)}, \qquad (4.25)$$

where $V(s\Omega_\epsilon)$ denotes the boundary rotation of $s\Omega_\epsilon$.

We remark that Lemma 4.3 and Lemma 3.4 (d) yield

$$\limsup_{n\to\infty}\left(\max_{z\in s\Omega_\epsilon}|\tilde{F}_n(z)|\right)^{1/n}=\frac{1}{N_\epsilon}=R(s\Omega_\epsilon),$$

which confirms that the normalized Faber polynomials are asymptotically optimal for $s\Omega_\epsilon$.

4.2 Iteration polynomials for the ABF method

In Phase II of the ABF method, we use iteration polynomials $q_n(z)$, so that the residual polynomials $p_n(z)$ are the normalized Faber polynomials (2.21) for $s\Omega_\epsilon$, i.e.

$$p_n(z)=1-zq_n(z)=\tilde{F}_n(z). \tag{4.26}$$

The iteration polynomials therefore satisfy

$$q_n(z)=\frac{1-\tilde{F}_n(z)}{z}. \tag{4.27}$$

To derive a recurrence for the iteration polynomials, we start with a recurrence for $F_n(0)=:\rho_n$. (4.17) and (4.18) yield

$$\rho_0=1,\quad \rho_1=\lambda_m(N_\epsilon+M_\epsilon+S_\epsilon), \tag{4.28}$$

and, for $n\geq 2$,

$$\begin{aligned}\rho_n &= (\rho_1+(-\lambda_m S_\epsilon))(\rho_{n-1}+(-\lambda_m S_\epsilon)^{n-1})\\ &\quad -\lambda_m^2 N_\epsilon M_\epsilon(\rho_{n-2}+(-\lambda_m S_\epsilon)^{n-2})-(-\lambda_m S_\epsilon)^n\\ &= \lambda_m(N_\epsilon+M_\epsilon)\rho_{n-1}-\lambda_m^2 N_\epsilon M_\epsilon\rho_{n-2}\\ &\quad -(-\lambda_m S_\epsilon)^{n-2}\lambda_m^2(S_\epsilon+N_\epsilon)(S_\epsilon+M_\epsilon).\end{aligned} \tag{4.29}$$

It can be seen from (4.17), that

$$q_0(z)\equiv 0,\ \text{ and }\ q_1(z)\equiv -\frac{1}{t_\epsilon\rho_1}. \tag{4.30}$$

Furthermore, for $n\geq 2$,

$$\begin{aligned}zq_n(z) &= 1-\tilde{F}_n(z)=\frac{1}{\rho_n}(\rho_n-F_n(z))\\ &= \frac{1}{\rho_n}\Big[\rho_{n-1}\left(\rho_1+(-\lambda_m S_\epsilon)\right)\left(1-\tilde{F}_{n-1}(z)\right)\\ &\quad -\lambda_m^2 N_\epsilon M_\epsilon\rho_{n-2}\left(1-\tilde{F}_{n-2}(z)\right)\\ &\quad -zt_\epsilon^{-1}\left(\hat{F}_{n-1}(z)+\lambda_m S_\epsilon\hat{F}_{n-2}(z)\right)\Big],\end{aligned} \tag{4.31}$$

where the polynomials $\hat{F}_n(z)$ are defined as in (4.19).

We use the relation $1 - \tilde{F}_n(z) = zq_n(z)$ and divide (4.31) by z. Hence, for $n \geq 2$,

$$\begin{aligned} q_n(z) &= \frac{1}{\rho_n}\Big[\lambda_m(N_\epsilon + M_\epsilon)\rho_{n-1}q_{n-1}(z) - \lambda_m^2 N_\epsilon M_\epsilon \rho_{n-2}q_{n-2}(z) \\ &\quad - t_\epsilon^{-1}\left(\hat{F}_{n-1}(z) + \lambda_m S_\epsilon \hat{F}_{n-2}(z)\right)\Big]. \end{aligned} \tag{4.32}$$

(4.30) and (4.32) are a three-term recurrence for the iteration polynomials $q_n(z)$. Note that r_n can be computed without explicitly computing x_n,

$$r_n = \tilde{F}_n(A)r_0 = \frac{1}{\rho_n}F_n(A)r_0 = \frac{1}{\rho_n}(\hat{F}_n(A)r_0 - (-\lambda_m S_\epsilon)^n r_0). \tag{4.33}$$

Hence, the residual norms in Phase II of the ABF method are obtainable without explicitly forming $b - Ax_n$.

4.3 Phase II of the ABF method

A complete listing of Phase II of the ABF method is given below in Algorithm 4.4. A MATLAB implementation of this algorithm is shown in Appendix B.

Algorithm 4.4 (Phase II of the ABF method)

Input:

- *Linear system parameters:* $A \in \mathbf{C}^{N\times N}$, $b \in \mathbf{C}^N$, $x_0 \in \mathbf{C}^N$
- *Inclusion set parameters:* $\lambda_m \in \partial\mathbf{E}$, $\phi \in (0, 2\pi)$, $\epsilon \in [0, \epsilon_{max})$, $s > 0$

Initialize:

$P := \tan\frac{\phi}{4} + \left(\cos\frac{\phi}{4}\right)^{-1}$, $N_\epsilon := \frac{1}{2}\left(\frac{P}{1+\epsilon} + \frac{1+\epsilon}{P}\right)$,

$M_\epsilon := \frac{(1+\epsilon)^2 - 1}{2\tan\frac{\phi}{4}(1+\epsilon)}$, $S_\epsilon := \frac{N_\epsilon M_\epsilon - 1}{N_\epsilon - M_\epsilon}$, $t_\epsilon^{-1} := \frac{N_\epsilon - M_\epsilon}{s}$,

$c_1 := \lambda_m(N_\epsilon + M_\epsilon)$, $c_2 := \lambda_m^2 N_\epsilon M_\epsilon$, $c_3 := -\lambda_m S_\epsilon$,

$c_4 := \lambda_m^2(S_\epsilon + N_\epsilon)(S_\epsilon + M_\epsilon)$,

$\rho_0 := 1$, $\rho_1 := c_1 - c_3$,

$r_0 := b - Ax_0$,

$\hat{F}_0(A)r_0 := 2r_0$, $\hat{F}_1(A)r_0 := t_\epsilon^{-1}Ar_0 + c_1 r_0$,

$q_0(A)r_0 := 0$, $q_1(A)r_0 := -t_\epsilon^{-1}\rho_1^{-1}r_0$

First step:

$x_1 = x_0 + q_1(A)r_0$

$\|r_1\| = \|\hat{F}_1(A)r_0 - c_3 r_0\| \,/\, |\rho_1|$

$n = 1$

Main loop:

While $\|r_n\| > tol$

$$n = n+1 \tag{4.34}$$

$$\rho_n = c_1\rho_{n-1} - c_2\rho_{n-2} - c_3^{n-2}c_4 \tag{4.35}$$

$$v_n = t_\epsilon^{-1}\left(\hat{F}_{n-1}(A)r_0 - c_3\hat{F}_{n-2}(A)r_0\right) \tag{4.36}$$

$$\hat{F}_n(A)r_0 = Av_n + c_1\hat{F}_{n-1}(A)r_0 - c_2\hat{F}_{n-2}r_0 \tag{4.37}$$

$$q_n(A)r_0 = \frac{1}{\rho_n}\left(c_1\rho_{n-1}q_{n-1}(A)r_0 - c_2\rho_{n-2}q_{n-2}(A)r_0 - v_n\right) \tag{4.38}$$

$$\|r_n\| = \|\hat{F}_n(A)r_0 - c_3^n r_0\| \;/\; |\rho_n| \tag{4.39}$$

End

Final step:

$x_n = x_0 + q_n(A)r_0$

In the main loop we introduce v_n to avoid evaluating $t_\epsilon^{-1}(\hat{F}_{n-1}(A)r_0 - c_3\hat{F}_{n-2}(A)r_0)$ twice (cf. (4.36) - (4.38)). With the use of v_n, the storage requirements are constantly five vectors, $\hat{F}_{n-2}(A)r_0$, $\hat{F}_{n-1}(A)r_0$, v_n, $q_{n-2}(A)r_0$ and $q_{n-1}(A)r_0$, plus A, x_0, r_0 and a few scalars. If storage rather than floating point operations (flops) is a concern, it is possible to work without v_n and instead perform the additional operations.

The only matrix-vector product in the main loop is Av_n in (4.37). Note that every PIM requires at least one matrix-vector product with A per step, for if otherwise the degree of the iteration polynomial cannot increase. Additionally, the main loop requires 8 scalar-vector products and 6 vector additions. If $NORM$ denotes the number of flops to compute the Euclidean vector norm, typically $2N$ plus one square root, the effort of one step of the main loop is approximately $2N^2 + 14N + NORM$ flops.

We point out that, apart from the computation of $\|r_n\|$, Phase II of the ABF method does not require reduction operations such as inner products. In particular, the type of the required operations allows an efficient implementation of Algorithm 4.4 on a parallel machine (see e.g. [65, Chapter 11] for details on parallel implementations).

Because $0 < |\lambda_m S_\epsilon| < 1$ (cf. (4.8)), forming $c_3^n = (-\lambda_m S_\epsilon)^n$ for large n does not lead to computational overflow problems. Furthermore, $c_3^n \to 0$ as $n \to \infty$. Hence, if $|c_3^n|$ for large n is smaller than the machine precision, the operations involving c_3^n do not have to be performed. This can be used to increase the efficiency of a computer implementation.

To derive convergence results for the application of Phase II of the ABF method to a linear system (1.1), we obviously have to make assumptions on the information obtained in Phase I. Specifically, we have to assume that the inclusion set $s\Omega_\epsilon$ determined in Phase I is an inclusion set of $\Lambda(A)$. Otherwise there is no guarantee for the convergence of the ABF method. With this assumption, we can easily derive the following convergence theorem.

Theorem 4.5 *Suppose that Phase II of the ABF method is applied to the linear system (1.1) with initial guess x_0, and that $\Lambda(A) \subset s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$. Then the residuals satisfy $r_n = \tilde{F}_n(A) r_0$, $n \geq 0$, and their norms are bounded by*

$$\frac{\|r_n\|}{\|r_0\|} \leq \frac{V(s\Omega_\epsilon) L(s\Omega_\epsilon)}{2\pi^2 (N_\epsilon^n + M_\epsilon^n - (-S_\epsilon)^n)} \max_{z \in \partial(s\Omega_\epsilon)} \| (zI - A)^{-1} \|, \tag{4.40}$$

where $V(s\Omega_\epsilon)$ and $L(s\Omega_\epsilon)$ denote the rotation and arc length of $\partial(s\Omega_\epsilon)$, respectively.

Furthermore, if A is diagonalizable, $A = XDX^{-1}$, D diagonal, then

$$\frac{\|r_n\|}{\|r_0\|} \leq \frac{\kappa(X) V(s\Omega_\epsilon)}{\pi (N_\epsilon^n + M_\epsilon^n - (-S_\epsilon)^n)}. \tag{4.41}$$

Proof. Applying Lemma 4.3 to (2.15) and (2.14) shows (4.40) and (4.41), respectively. □

Both (4.40) and (4.41) show that the residuals produced by Phase II of the ABF method satisfy

$$\|r_n\| \to 0 \text{ for } n \to \infty,$$

as long as $\Lambda(A) \subset s\Omega_\epsilon$. Furthermore, Theorem 4.5 shows that the performance of Phase II is influenced by the sensitivity of the eigenvalues of A. A closer analysis of (4.40) and (4.41) reveals which information about A should be obtained in Phase I to determine the inclusion set $s\Omega_\epsilon$:

If A has well-conditioned eigenvalues, then both (4.40) and (4.41) indicate that the speed of convergence of Phase II is mainly determined by the asymptotic convergence factor $1/N_\epsilon$. In this case N_ϵ should be maximized, i.e. $s\Omega_\epsilon$ should be chosen tightly around $\Lambda(A)$. Thus it is desirable that Phase I yields a close approximation of $\Lambda(A)$. If no a priori information about $\Lambda(A)$ is available, we use the Arnoldi process and base the choice of $s\Omega_\epsilon$ on $\Lambda(H_n)$.

On the other hand, if the resolvent norms are large in neighborhoods of $\Lambda(A)$, then $\partial(s\Omega_\epsilon)$ should avoid these neighborhoods, i.e. $s\Omega_\epsilon$ should overestimate $\Lambda(A)$. In this case it is desirable that Phase I reveals the regions in which the resolvent norms are large. Therefore, if it is known that A is highly nonnormal, an approximation of its pseudospectrum should be considered, typically $\Lambda_{\tilde{\epsilon}}(H_n)$ for small $\tilde{\epsilon} > 0$. (We here use $\tilde{\epsilon}$ for the parameter corresponding to the pseudospectrum to avoid confusion with the parameter ϵ corresponding to the inclusion set.) In light of the high cost of computing $\Lambda_{\tilde{\epsilon}}(H_n)$ compared to computing $\Lambda(H_n)$, a major difficulty lies in the choice of $\tilde{\epsilon}$. If $\tilde{\epsilon}$ is too large, $\Lambda_{\tilde{\epsilon}}(H_n)$ might contain the origin and thus might be useless for determining the inclusion set. If $\tilde{\epsilon}$ is too small, $\Lambda_{\tilde{\epsilon}}(H_n)$ might not yield enough information about the spectral sensitivity of A. We know of no method for choosing the 'correct' value of $\tilde{\epsilon}$ a priori.

Often, the number n of Arnoldi steps as well as the $\tilde{\epsilon}$ for computing $\Lambda_{\tilde{\epsilon}}(H_n)$ can only be determined dynamically. If Phase II does not converge well for the current choice of parameters, Phase I has to be repeated to find a new inclusion set. Such 'adaptive steps' are essential ingredients of most hybrid PIMs.

4.4 Applications of the ABF method

In numerical examples we show a typical application of the ABF method and study the influence of the inclusion set on the convergence of Phase II of the method in case of sensitive eigenvalues. Information on the test matrices is given in Appendix A. The MATLAB program of Phase II used in these experiments is listed in Appendix B.

Example 4.6 Hybrid PIMs are typically used if several linear systems have to be solved, which have either the same matrix and different right-hand sides, or which have different right-hand sides but matrices with similar spectra (or pseudospectra, fields of values). In these situations, the information obtained in Phase I has to be computed only once and can be used several times.

We give an example using the PDEN matrices, which all originate from the same partial differential equation. Our experiment shows that the approximation of the spectrum of A=PDE225 can be applied successfully to solve systems with PDEN for different dimensions N.

In Phase I, we apply 25 steps of the Arnoldi process to A and the vector $v = A[1,...,1]^T$. With the choice of the starting vector as the product of A and the vector of all ones, we intend to amplify in v the components of the eigenvectors corresponding to the largest eigenvalues of A. Heuristically, this leads to faster convergence of the Arnoldi process for the extreme eigenvalues of A, which is useful for finding an inclusion set of $\Lambda(A)$.

The Arnoldi algorithm yields the matrix H_{25}, an orthogonal projection of A onto the Krylov space $K_{25}(A, v)$. We next compute $\Lambda(H_{25})$ with MATLAB's `eig` function. Using the strategy outlined in Section 3.2 (also cf. Example 3.8), we compute an inclusion set of $\Lambda(H_{25})$. The computed parameters are $\lambda_m = -1$, $\phi \approx 1.27\,\pi$, $s \approx 2.544$ and $\epsilon \approx 1.2316$. To guarantee that $\Lambda(H_{25})$ is contained in the interior of the inclusion set, we increase ϵ by 5%, i.e. we define $\epsilon_1 := 1.05\,\epsilon$ and use

$$\mathbf{S}_1 := s\Omega_{\epsilon_1} \in \mathbf{B}(\lambda_m, \phi).$$

To study the influence of inclusion sets on the performance of Phase II we use a second set,

$$\mathbf{S}_2 := s\Omega_{\epsilon_2} \in \mathbf{B}(\lambda_m, \phi),$$

where $\epsilon_2 := 1.1\,\epsilon$. Figure 6 shows $\Lambda(H_{25})$ (pluses), $\partial\mathbf{S}_1$ (solid) and $\partial\mathbf{S}_2$ (dashed). Using Lemma 3.4 we compute

$$R(\mathbf{S}_1) \approx 0.928 < R(\mathbf{S}_2) \approx 0.9369.$$

We next apply Phase II based on the two inclusion sets to linear systems with the matrices PDE225, PDE900 and PDE2961. For all three matrices the right-hand side is the vector of all ones of appropriate dimension, $x_0 = 0$, and the termination criterion is $\|r_n\|/\|r_0\| < 10^{-8}$. The resulting curves of residual norms are shown in Figures 7-9 (solid curve for $\mathbf{S}_1$ and dashed curve for $\mathbf{S}_2$). In all cases the convergence is faster when $\mathbf{S}_1$ is used, which corresponds to the asymptotic convergence factors.

The flop counts for the application of Phase II to the three linear systems are as follows:

	flops/N^2 for $\mathbf{S}_1$	flops/N^2 for $\mathbf{S}_2$
PDE225	34.5	36.9
PDE900	31.2	33.4
PDE2961	40.2	42.9

For both $\mathbf{S}_1$ and $\mathbf{S}_2$, Phase II of the ABF method solves the three linear systems with $\mathcal{O}(N^2)$ flops, and the constants involved in the $\mathcal{O}$-terms are small compared to N^2. The effort of Phase I, which is run only once, is negligible.

Example 4.7 We study the influence of high non-normality of the coefficient matrix on the behavior of Phase II of the ABF method.

We consider the matrix of [69, Example 6.3], namely the shifted GRCAR matrix

$$A = \begin{bmatrix} \frac{9}{10} & 1 & 1 & 1 & & & \\ -1 & \frac{9}{10} & 1 & 1 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & -1 & \frac{9}{10} & 1 & 1 & 1 \\ & & & -1 & \frac{9}{10} & 1 & 1 \\ & & & & -1 & \frac{9}{10} & 1 \\ & & & & & -1 & \frac{9}{10} \end{bmatrix} \in \mathbf{R}^{1024\times 1024}, \tag{4.42}$$

which has highly sensitive eigenvalues (cf. Appendix A). Starke and Varga [69] use this matrix to show the dependence of their hybrid Arnoldi-Faber method on the accuracy of the Arnoldi eigenvalue estimates. They note that their method diverges when the inclusion set is derived from 24 Arnoldi steps, while their method converges well when 48 Arnoldi steps are performed. We further analyze this observation and demonstrate the general difficulty of finding 'optimal' inclusion sets for hybrid PIMs.

We apply 48 steps of the Arnoldi process to A and $v = A[1, ..., 1]^T$ and compute the eigenvalues of the resulting matrix H_{48} with MATLAB's `eig` function. Note that the Arnoldi eigenvalues (circles in Figure 10) largely overestimate the actual eigenvalues (pluses in Figure 10). For applications of Phase II, we consider inclusion sets $2\Omega_\epsilon \in \mathbf{B}(-1, 4\pi/7)$. Figure 10 shows $\partial(2\Omega_\epsilon)$ for $\epsilon = 0.25$ (solid), 0.24 (dashed), 0.23 (dashdot), 0.22 (dotted).

$\Lambda(A)$ and all but two of the eigenvalues of H_{48} are obviously contained in each of the inclusion sets. Figure 11, a close-up of the small box in Figure 10, reveals that $\Lambda(H_{48})$ is only contained in $2\Omega_{0.25}$ and $2\Omega_{0.24}$.

We next apply Phase II of the ABF method to A, $b = [1, ..., 1]^T$, $x_0 = 0$ and the four inclusion sets. The iteration is terminated when $\|r_n\|/\|r_0\| < 10^{-8}$. The convergence curves are shown in Figure 12 (solid, dashed, dashdot and dotted for $\epsilon = 0.25$, 0.24, 0.23 and 0.22, respectively). Clearly, minor changes of the inclusion sets in this example lead to major differences in the behavior of Phase II.

At first sight, the convergence behavior seems to depend on the inclusion of $\Lambda(H_{48})$ by $2\Omega_\epsilon$. For $\epsilon \in \{0.25, 0.24\}$, we have $\Lambda(H_{48}) \subset 2\Omega_\epsilon$ and Phase II converges quickly. Moreover, the slight difference in the speed of convergence is well reflected by the asymptotic convergence factors

$$R(2\Omega_{0.24}) \approx 0.9696 < R(2\Omega_{0.25}) \approx 0.9715.$$

On the other hand, $\Lambda(H_{48}) \not\subset 2\Omega_\epsilon$ for $\epsilon \in \{0.23, 0.22\}$. In these cases, the convergence behavior is drastically different. In spite of smaller asymptotic convergence factors compared to $\epsilon \in \{0.25, 0.24\}$,

$$R(2\Omega_{0.22}) \approx 0.9657 < R(2\Omega_{0.23}) \approx 0.9677,$$

Phase II needs more steps to converge. For $\epsilon = 0.22$, the residual norms even increase for about 750 consecutive steps, before the speed of convergence dramatically increases.

But whether $2\Omega_\epsilon$ contains $\Lambda(H_{48})$ or not is by no means a plausible explanation for these drastic differences. For a closer analysis, we introduce the symbol of the Toeplitz matrix A,

$$f_A(z) = -z^{-1} + \frac{9}{10} + z + z^2 + z^3.$$

As shown in [60], $f_A(\partial\mathbf{E})$, the spectrum of the (infinite dimensional) Toeplitz operator associated with A, is roughly the same as $\Lambda_\epsilon(A)$ for small ϵ. Nachtigal et al. [57] call the region bounded by $f_A(\partial\mathbf{E})$ the *effective spectrum* of A. Now consider that the normalized Faber polynomials satisfy

$$\tilde{F}_n(A) = \frac{1}{2\pi i} \int_C \tilde{F}_n(z)(zI - A)^{-1} dz,$$

where C is any Jordan curve enclosing $\Lambda(A)$. Hence, for $\|\tilde{F}_n(A)\|$ to be small, $|\tilde{F}_n(z)|$ should be small in the region bounded by $f_A(\partial\mathbf{E})$. This, however, is not satisfied for $\epsilon \in \{0.23, 0.22\}$, while this is satisfied for $\epsilon \in \{0.25, 0.24\}$. Figure 14 shows contour plots of $|\tilde{F}_n(z)|$ for $2\Omega_{0.22}$ (left column), $2\Omega_{0.25}$ (right column) and $n = 25$ (first plot in each column), $n = 100$ (second plot in each column). The third plot in each column shows a close-up of the small box in the second plot. Plotted are the contour lines for

$$|\tilde{F}_n(z)| = 2^j, \quad j = 0, 1, ..., 10.$$

In each of the plots in Figure 14, the dashed line is $f_A(\partial\mathbf{E})$.

We note a striking difference: For $\epsilon = 0.25$, $f_A(\partial\mathbf{E})$ is fully contained in the interior of the innermost contour, i.e. inside the region in which $|\tilde{F}_n(z)| < 1$. (The same holds for $\epsilon = 0.24$.) For $\epsilon = 0.22$, the regions of large resolvent norms and large $|\tilde{F}_n(z)|$ partly overlap. (The same holds for $\epsilon = 0.23$.) This explains why Phase II converges slower for $\epsilon \in \{0.23, 0.22\}$ than for $\epsilon \in \{0.25, 0.24\}$.

Note that the differences in the convergence behavior cannot be explained simply by the inclusion of $f_A(\partial\mathbf{E})$ by $2\Omega_\epsilon$. None of the inclusion sets completely contains the effective spectrum of A (cf. Figure 13).

We are not able to explain analytically the dramatic change in the speed of convergence of Phase II for $\epsilon \in \{0.23, 0.22\}$. To do so, more exact information about $\|\tilde{F}_n(A)\|$ from the resolvent norms must be inferred. But this is a difficult task in general: As shown in [38],

$$\|(zI - A)^{-1}\| = \|(zI - B)^{-1}\| \quad \text{for all} \;\; z \in \mathbf{C}$$

and two complex square matrices A and B does *not* imply that

$$\|p(A)\| = \|p(B)\| \;\; \text{for each polynomial} \;\; p(z).$$

Hence, information about the resolvent norms (or of pseudospectra) of a matrix A is in general not sufficient to determine the norm of a polynomial in A.

We summarize the main results of this experiment:

If A is highly nonnormal, the behavior of Phase II of the ABF method might be very sensitive to small changes of the inclusion set, and thus the determination of the set becomes a difficult matter. In such cases, Arnoldi eigenvalue estimates are an unreliable tool for this determination. Evidently, small changes in the starting vector or round-off errors might lead to slightly perturbed eigenvalue estimates. We believe that all hybrid PIMs depending on spectral information must suffer from this nuisance, which might be overcome by considering pseudospectra in Phase I.

Remark 4.8 For all our choices of ϵ, Phase II of the ABF method clearly outperforms a modified Gram-Schmidt implementation of (full) GMRES (cf. e.g. [65, Algorithm 6.9]) for the linear system of Example 4.7:

	flops/N^2
Phase II, $\epsilon = 0.25$	25.1
Phase II, $\epsilon = 0.24$	24.4
Phase II, $\epsilon = 0.23$	26.8
Phase II, $\epsilon = 0.22$	30.7
(full) MGS-GMRES	198.8

We also point out that in case of Toeplitz matrices a priori spectral information for choosing the inclusion set is generally available from the effective spectrum. The boundary of that set is computable at almost no cost.

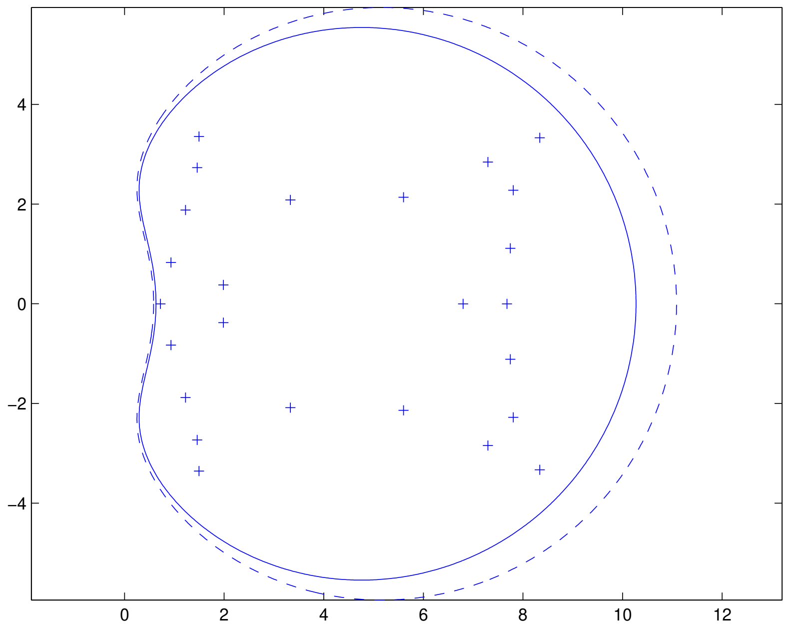

Figure 6: PDE225: Inclusion of approximate eigenvalues by sets $\mathbf{S}_1$ and $\mathbf{S}_2$.

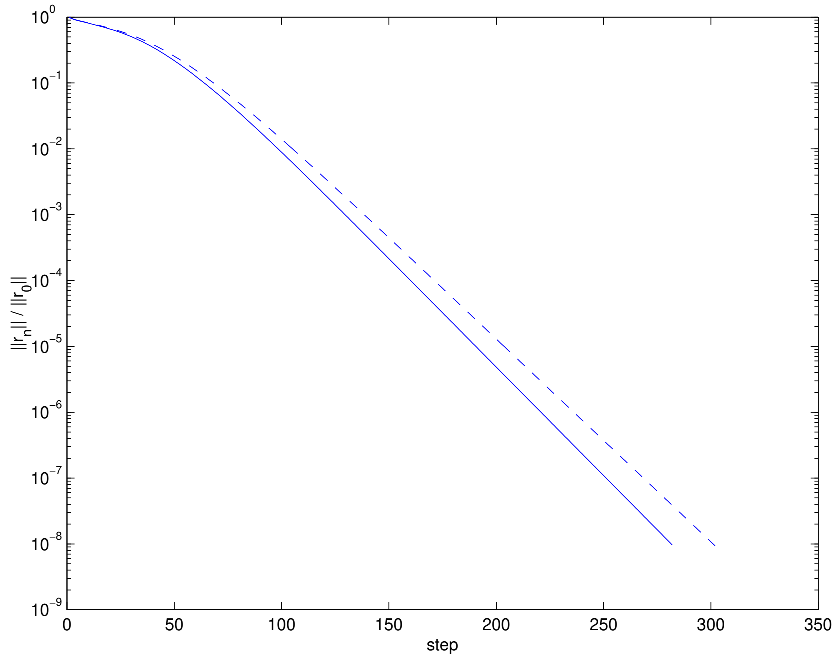

Figure 7: PDE225: Convergence of Phase II of the ABF method based on inclusion sets $\mathbf{S}_1$ and $\mathbf{S}_2$.

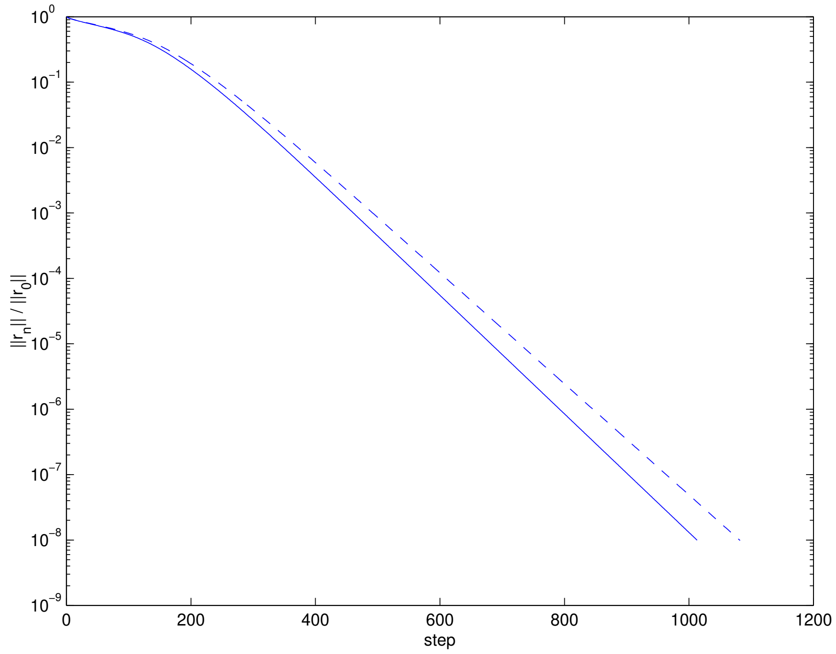

Figure 8: PDE900: Convergence of Phase II of the ABF method based on inclusion sets $\mathbf{S}_1$ and $\mathbf{S}_2$.

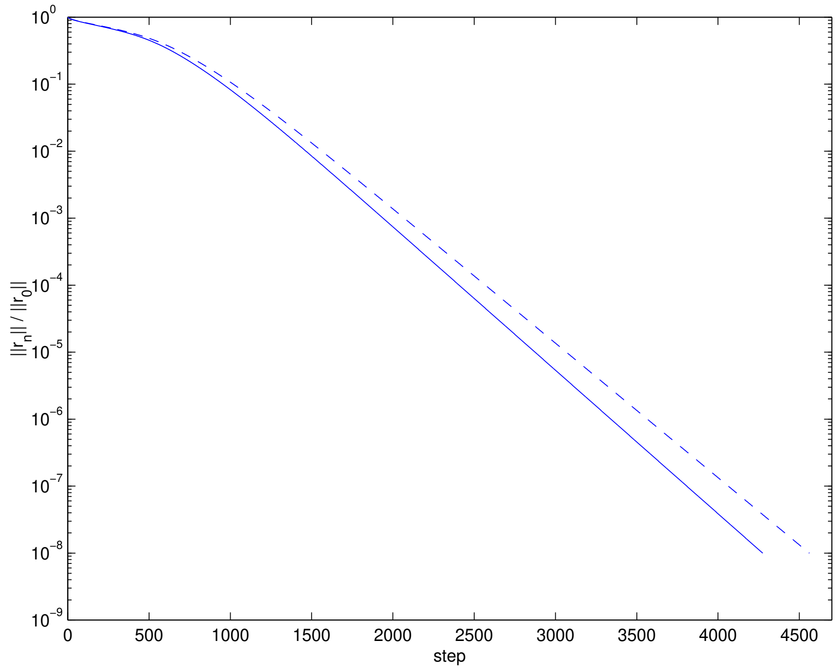

Figure 9: PDE2961: Convergence of Phase II of the ABF method based on inclusion sets $\mathbf{S}_1$ and $\mathbf{S}_2$.

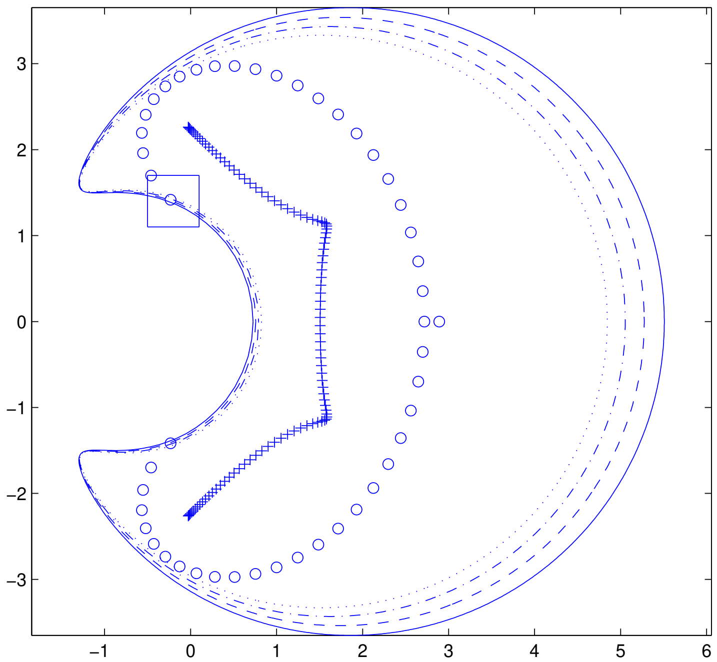

Figure 10: Shifted GRCAR matrix: Eigenvalues, approximate eigenvalues and four different inclusion sets.

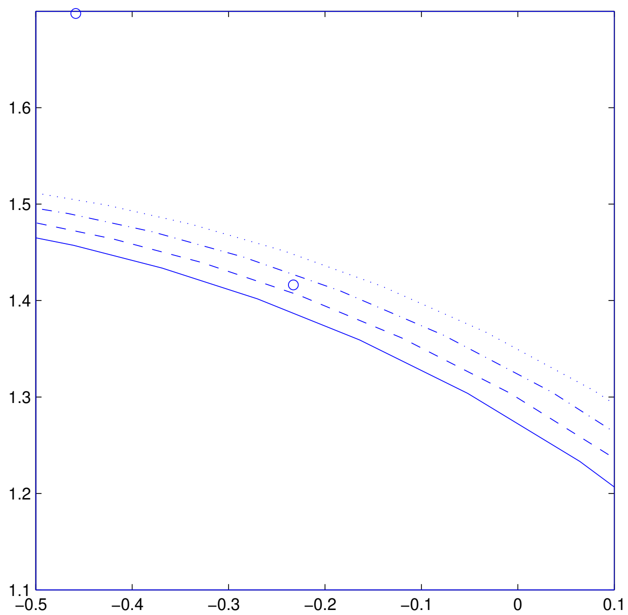

Figure 11: Close-up of the small box in Figure 10.

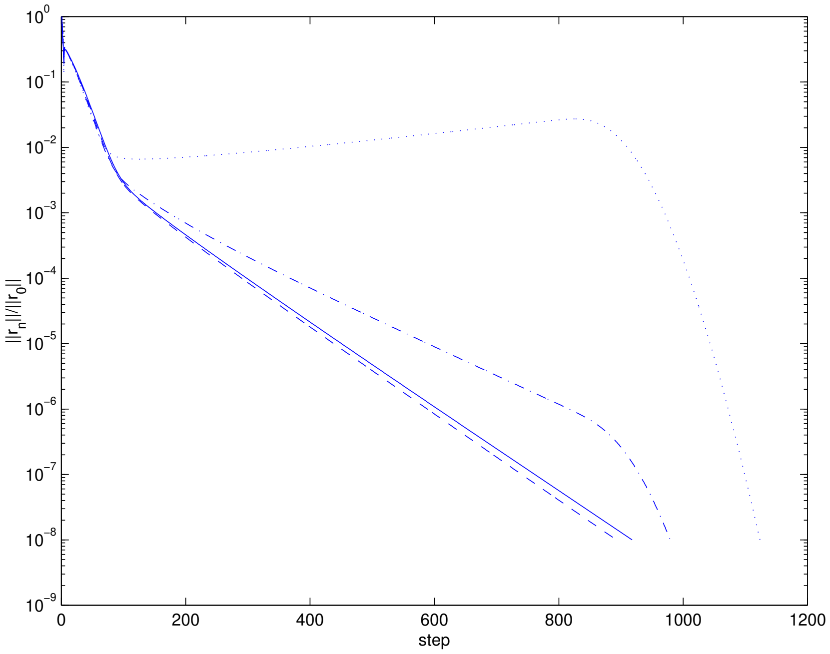

Figure 12: Shifted GRCAR matrix: Convergence of Phase II of the ABF method based on the four inclusion sets.

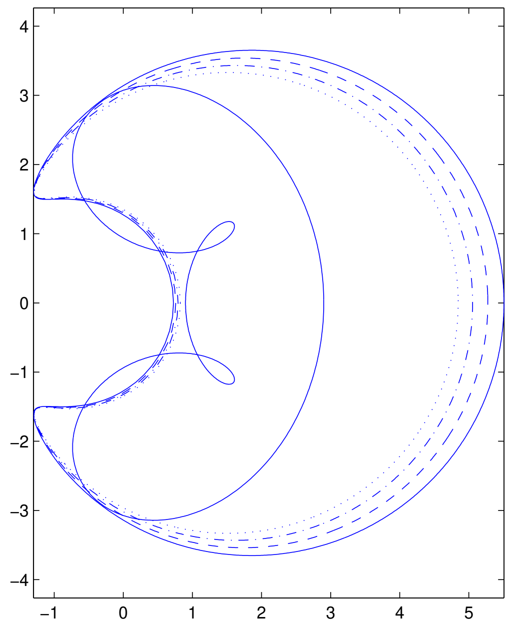

Figure 13: Shifted GRCAR matrix: Effective spectrum and the four inclusion sets.

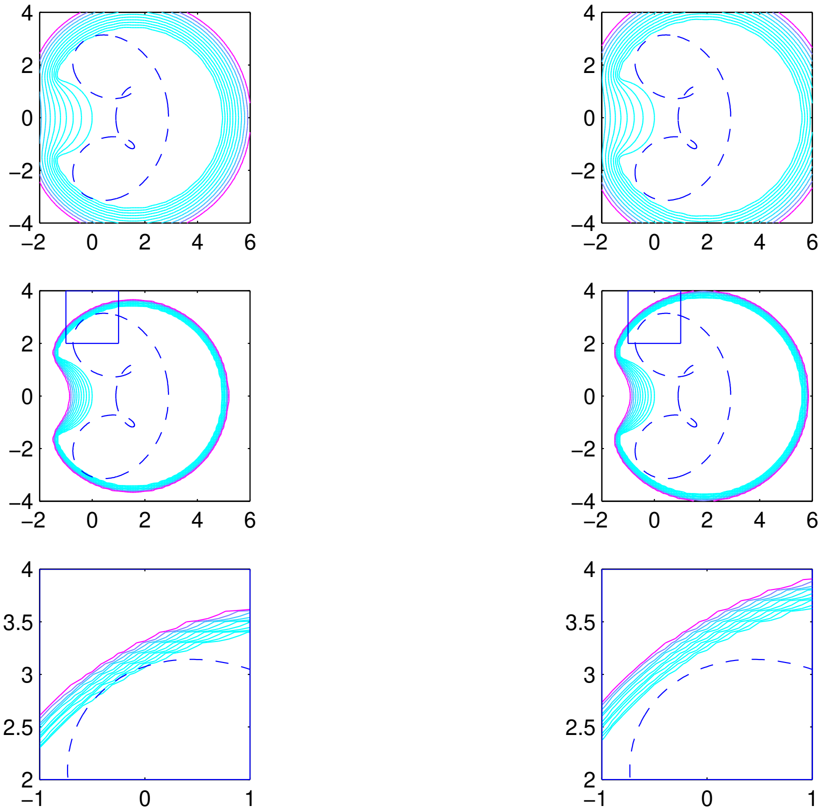

Figure 14: Shifted GRCAR matrix: Effective spectrum and contour plots of $|\tilde{F}_n(z)|$, $n = 25, 100$, for two different inclusion sets.

5 Computable convergence bounds for GMRES

In this section we analyze GMRES [66], arguably the most robust and most popular parameter-free PIM. The GMRES residuals satisfy the minimization property

$$\|r_n\| = \min_{r \in r_0 + AK_n(A,r_0)} \|r\| = \min_{p_n \in \pi_n} \|p_n(A) r_0\|. \tag{5.1}$$

GMRES might thus be viewed as the optimal PIM, yielding the smallest possible Euclidean residual norms. But this optimality comes not without cost, as work and storage requirements during the iteration grow linearly with the number of steps.

It is important to understand how GMRES works, and in recent years the convergence analysis of the algorithm has been an active field of research. Among others, Saad and Schultz [66], Nachtigal, Reddy and Trefethen [56], Starke [68] and Eiermann [16] derived upper bounds on the residual norms. However, none of these bounds leads to a complete characterization of the algorithm's convergence behavior. Some bounds are only applicable to a restricted class of nonsingular linear systems, others include factors that are not computable. All of these bounds are independent of the initial guess. As demonstrated in examples by Faber et al. [22] and Toh [72] this might lead to an over-estimation even of the algorithm's worst-case behavior. It is generally agreed, that the convergence analysis of GMRES needs further work (see e.g. [70, p. 187]).

Based on characterizations of matrices that generate the same GMRES residuals, so-called GMRES-equivalent matrices, we derive new convergence results. Our results differ from those mentioned above in three main ways:

- Our bounds on the algorithm's residual norms depend on the initial guess.
- Our bounds give computable reduction factors for the residual norms for any nonsingular linear system and are applicable even to the singular case, if a certain condition is satisfied.
- Approximations to all factors in our bounds can be obtained during the GMRES run. These approximations lead to a heuristic which often allows us to predict after only a few steps, how GMRES will perform in the iteration.

In Section 5.1, we introduce the GMRES algorithm and its mathematical properties. We then briefly review several convergence results in Section 5.2, before generalizing known characterizations of GMRES-equivalent matrices in Section 5.3. Section 5.4 includes the derivation of our new bounds. In Section 5.5, we explain how to obtain approximations to the factors in our bounds from information generated during the GMRES run. We also devise heuristics for the behavior of these approximations and for the prediction of the performance of the GMRES algorithm. Finally, Section 5.6 includes numerical examples.

Major parts of Section 5 first appeared in [50].

5.1 The mathematical properties of GMRES

We introduce the GMRES algorithm and state several of its properties. Note that the linear system (1.1) might be singular, if not assumed otherwise, and that here we are not concerned with implementational details.

Definition 5.1 *Suppose that x_0 is an initial guess for the solution of a linear system (1.1). For $n = 1, 2, ...,$ let*

$$x_n \in x_0 + K_n(A, r_0) \tag{5.1}$$

be a vector, such that

$$r_n \in r_0 + AK_n(A, r_0) \tag{5.2}$$

satisfies

$$\|r_n\| = \min_{r \in r_0 + AK_n(A, r_0)} \|r\|, \tag{5.3}$$

then the vectors x_n and r_n are called nth GMRES iterate and residual, respectively.

As a comparison with [66, Formulas (4) and (8)] shows, our definition of the GMRES iterates and residuals is mathematically equivalent to the original derivation by Saad and Schultz. The next well-known result (see e.g. [9, p. 8]) justifies Definition 5.1.

Lemma 5.2 *Suppose that $A \in \mathbf{C}^{N \times N}$ and $r_0 \in \mathbf{C}^N$. Then for every $n = 1, 2, ...,$ there exists a unique vector $r_n \in r_0 + AK_n(A, r_0)$ that satisfies (5.3).*

Thus, for $n = 1, 2, ...,$ the GMRES residuals are uniquely defined. In case of a nonsingular linear system, uniqueness of the residuals does result in uniqueness of the iterates. Definition 5.1 leads to the following algorithm, which does, however, not require uniqueness of the iterates.

Algorithm 5.3 (GMRES)

Input: $A \in \mathbf{C}^{N \times N}$, $b \in \mathbf{C}^N$, $x_0 \in \mathbf{C}^N$

Initialize: $r_0 := b - Ax_0$, $n = 0$

While $r_n \neq 0$

$n = n + 1$

Construct $x_n \in x_0 + K_n(A, r_0)$ *so that*

$$\|r_n\| = \|b - Ax_n\| = \min_{r \in r_0 + AK_n(A, r_0)} \|r\|$$

End While

If $r_n = 0$ for some $n \geq 0$, or, equivalently, x_n is a solution of (1.1), we say that GMRES terminates at step n. Before analyzing the termination properties of the algorithm, we point out an equivalent way to define the GMRES residuals.

Lemma 5.4 *Let $S \subset \mathbf{C}^N$ be a subspace and suppose that $y_0 \in \mathbf{C}^N$ and $y_n \in y_0 + S$. Then $\|y_n\| = \min_{y \in y_0 + S} \|y\|$ if and only if $y_n \perp S$.*

From this classical result (see e.g. [9, p. 9]) it follows that our definition of the GMRES residuals via (5.2) and (5.3) is equivalent to defining them by (5.2) combined with

$$r_n \perp AK_n(A, r_0). \tag{5.4}$$

The following lemma is a direct consequence of (5.3).

Lemma 5.5 *The GMRES algorithm applied to a linear system (1.1) and initial guess x_0 terminates at step $d \geq 0$ if and only if d is the smallest integer for which $r_0 \in AK_d(A, r_0)$.*

To further analyze this situation, we consider Lemma 2.1 with $v = Ar_0$, which shows that the spaces $AK_n(A, r_0)$ coincide for $n \geq \dim AK_N(A, r_0)$. The following is an immediate consequence.

Corollary 5.6 *Suppose that a linear system (1.1) and an initial guess x_0 are given and define $d := \dim AK_N(A, r_0)$. Then for the GMRES residuals r_n, $n \geq d$, $r_n = r_d$ follows.*

Thus, either GMRES terminates at some step $n \leq d := \dim AK_N(A, r_0) \leq N$, which requires $r_0 \in AK_N(A, r_0) = AK_d(A, r_0)$, or the algorithm fails to terminate, but stagnates after step d. The next theorem can be shown easily using Lemma 2.1 with $v = Ar_0$.

Theorem 5.7 (cf. [66, Proposition 2]) *Suppose that the matrix A in (1.1) is nonsingular. Then for any initial guess x_0, GMRES applied to (1.1) terminates at step $d \geq 0$ if and only if $\dim AK_N(A, r_0) = d$.*

While GMRES must fail to terminate for inconsistent linear systems, for consistent singular systems the question of termination depends on the initial guess. To derive a reasonable bound on the residual norms, one has to assume that GMRES will terminate, which is equivalent to assuming $r_0 \in AK_d(A, r_0) = AK_N(A, r_0)$. In the derivation of our new bounds we will always make this assumption. For further analysis of the application of GMRES to singular systems, in particular the case $r_0 \notin AK_d(A, r_0)$, we refer to [6].

5.2 A review of GMRES convergence results

Let us briefly review the most important convergence results for GMRES. Further analysis of some of these results can be found, for example, in [34, Section 3.2]. We suppose that a linear system (1.1) and an initial guess x_0 are given.

A typical starting point for the derivation of bounds on the GMRES residual norms is the minimization property

$$\|r_n\| = \min_{p_n \in \pi_n} \|p_n(A) r_0\|. \tag{5.5}$$

This shows two important property of GMRES: First, the residual polynomial $p_n(z)$ is constructed to minimize the Euclidean norm of the vector $p_n(A)r_0$ and not of the matrix $p_n(A)$. The behavior of GMRES might therefore be very different from the behavior of Phase II of any hybrid PIM. The latter usually attempt to minimize $\|p_n(A)\|$. Second, (5.5) implies that the residual norms are monotonically decreasing, which is, for example, not guaranteed for Phase II of the ABF method.

Now suppose that A is diagonalizable, $X^{-1}AX = D$, D diagonal. Then for the GMRES residuals it follows from (5.5) that

$$\frac{\|r_n\|}{\|r_0\|} \le \kappa(X) \min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(A)} |p_n(\lambda)|. \tag{5.6}$$

This bound was the first convergence result for GMRES [66, Proposition 4].

If A is normal, then $\kappa(X) = 1$ and (5.6) is sharp in the sense that for each n, there exists an initial guess for which equality holds [35, 47]. This suggests that if A is normal, the convergence of GMRES is mainly determined by $\Lambda(A)$ and can be analyzed by studying $\mu_n(\Omega)$, where

$$\mu_n(\Omega) := \inf_{p_n \in \pi_n} \max_{z \in \Omega} |p_n(z)|, \tag{5.7}$$

and $\Omega \subset \mathbf{C}$ is an inclusion set of $\Lambda(A)$. But note that the eigenvalues of a non-hermitian matrix can be anywhere in the complex plane, while $\mu_n(\Omega)$ is only known for line segments (cf. [27]). In some cases, upper bounds on $\mu_n(\Omega)$ can be derived. One example is Lemma 4.3, which can be readily applied to the convergence analysis of GMRES.

For diagonalizable but nonnormal A, $\kappa(X)$ might be very large, and the right-hand side of (5.6) might be a large overestimate of the actual residual norm. It is furthermore not clear, how the conditioning of the *eigenvectors* of A does influence the convergence behavior of GMRES applied to a *linear system* with A.

The sharpness results from [35, 47] can lead to confusion. They show that the convergence behavior of GMRES is better understood if A is normal, provided that the solution to the corresponding approximation problem (5.7) is known. They show neither that the algorithm converges faster for normal matrices nor that the degree of normality of a matrix has a direct influence on the convergence behavior. In fact, Greenbaum and Strakoš [37, Section 3.1] show that any nonincreasing curve of residual norms,

$$\|r_0\| \ge \|r_1\| \ge \dots \ge \|r_{N-1}\| > \|r_N\| = 0,$$

can be produced by GMRES applied to a normal, even unitary $N \times N$ matrix.

A bound on the GMRES residual norms can also be derived using pseudospectra of A. With $C = \partial\Lambda_\epsilon(A)$ in (2.15), from (5.5) it follows that

$$\frac{\|r_n\|}{\|r_0\|} \le \frac{L_\epsilon}{2\pi\epsilon} \inf_{p_n \in \pi_n} \max_{\lambda \in \Lambda_\epsilon(A)} |p_n(\lambda)|, \tag{5.8}$$

where L_ϵ denotes the arc length of $\partial\Lambda_\epsilon(A)$ [56, Theorem 4]. Note that (5.8) holds for all linear systems, while (5.6) requires A to be diagonalizable. But both (5.6) and (5.8) yield no reduction factor if A is singular. Generally, it is very costly to compute $\Lambda_\epsilon(A)$, in particular when the cost of this computation is compared to the cost of a single GMRES run. Just like (5.6), the bound

(5.8) sometimes can not describe the rate of convergence of GMRES, because no bound on $\mu_n(\Lambda_\epsilon(A))$ is known. We also note that it is possible to construct examples in which the right-hand side of (5.8) overestimates the left-hand side considerably (see e.g. [37, Section 1]).

The integral representation of operators is also used by Campbell et al. [7]. By constructing polynomials which are related to the minimal polynomial of A, bounds on the residual norms are derived from (5.5). These bounds relate the rate of convergence of GMRES to the clustering of $\Lambda(A)$. The results are of a qualitative nature, because the bounds on the residual norms include asymptotic error constants which are difficult to quantify and which are potentially unbounded.

Another bound on GMRES residual norms is based on the field of values $F(A)$. Defining $\nu(A) := \min_{z\in F(A)} |z|$, Starke [68, Theorem 3.2] shows that

$$\frac{\|r_n\|}{\|r_0\|} \leq (1-\nu(A)\nu(A^{-1}))^{n/2}. \tag{5.9}$$

This bound gives a reduction factor for the residual norms, provided that $0 \notin F(A)$ and $0 \notin F(A^{-1})$, which excludes indefinite matrices. It is possible to obtain approximations to the right-hand side of (5.9) during the GMRES run by solving small eigenproblems. Thus (5.9) is interesting from a practical point of view. However, while $F(A)$ in many cases provides useful information about the matrix A, the concept has its limitations in this context. For example, suppose that A is hermitian positive definite and denote $\kappa := \kappa(A)$. Then $\nu(A)\nu(A^{-1}) = \kappa^{-1}$. But if $\kappa > 1$, then

$$\left(1-\kappa^{-1}\right)^{1/2} \;>\; \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}.$$

Hence, for each $\kappa > 1$, there exists an index n_0, such that for all $n \geq n_0$, the bound (5.9) is weaker than the standard bound (see e.g. [34, p. 51])

$$\frac{\|r_n\|}{\|r_0\|} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^n, \tag{5.10}$$

which is derived from (5.6) using translated Chebyshev polynomials.

There exists a second bound based on the field of values. Using the asymptotic convergence factor $R(F(A))$, Eiermann [16] shows that

$$\frac{\|r_n\|}{\|r_0\|} \leq 2c_n[R(F(A))]^n, \tag{5.11}$$

where the c_n are some positive constants satisfying $c_n < 2/(1-[R(F(A))]^n)$. As is the case for (5.9), the bound (5.11) requires that $0 \notin F(A)$. In most cases it is expensive to compute $R(F(A))$. Also note that if $R(F(A)) \approx 1$, then the constants c_n might be very large. With regard to this, we point out that in all examples of a recent study by Ernst [20], computed asymptotic convergence factors $R(F(A))$ range between 0.995 and 0.9999, although GMRES shows very different types of convergence behavior.

Note that none of the bounds presented above depends on r_0. One problem with these bounds surfaces, when we consider unitary similarity transformations of A. Since $\Lambda(A)$, $\Lambda_\epsilon(A)$ and $F(A)$ are invariant under such transformations, all of the above bounds are the same for the linear systems (1.1), and (1.1) with A replaced by $U^H AU$ for any unitary matrix U. It is obvious that the solutions to these two linear systems are generally different. But also the behavior of GMRES applied to these systems with the same initial guess might be completely different:

Let u_1 be an eigenvector of A corresponding to a nonzero eigenvalue λ. We complete u_1 by unit vectors $u_2, ..., u_N$ to an orthonormal basis of $\mathbf{C}^N$ and denote $U_1 := [u_1, u_2, ..., u_N]$. Let $U_2 \in \mathbf{C}^{N\times N}$ be a unitary matrix for which $U_2 b = \|b\| e_1$. We define $U := U_1 U_2$, then the application of GMRES to $U^H AUx = b$ with $x_0 = 0$ yields

$$\|r_1\| = \min_{p_1\in\pi_1} \|p_1(U^H AU)b\| = \|b\| \min_{\alpha\in\mathbf{C}} \|e_1 - \alpha\lambda e_1\| = 0,$$

i.e. GMRES terminates at step one. On the other hand, the application of GMRES to $Ax = b$ with $x_0 = 0$ might need many more steps to terminate. Being invariant under unitary similarity transformations of A, all of the above bounds can lead to a misprediction of the actual behavior of GMRES.

We remark that not considering the dependence of GMRES on the initial residual can also lead to a misprediction of the worst-case behavior. Since GMRES minimizes $\|p_n(A)r_0\|$ for a given r_0, 'worst-case GMRES' is the maximization of $\min_{p_n\in\pi_n} \|p_n(A)r_0\|$ over all r_0 with norm one. On the other hand, (5.6), (5.8), (5.11) and the bounds in [7] are derived from the 'ideal GMRES' [39] approximation problem

$$\min_{p_n\in\pi_n} \|p_n(A)\|.$$

It is clear that

$$\max_{\|r_0\|=1} \min_{p_n\in\pi_n} \|p_n(A)r_0\| \le \min_{p_n\in\pi_n} \|p_n(A)\|, \tag{5.12}$$

which means that 'ideal GMRES' forms an upper envelope for all possible actual GMRES convergence curves. However, there are examples for which this envelope is never attained. Faber et al. [22] and Toh [72] constructed matrices for which for a certain n, the left-hand side of (5.12) is strictly less than the right-hand side. For linear systems with these matrices, bounds based on 'ideal GMRES' will overestimate the worst-case behavior. In the example of [72], the ratio of left- to right-hand side can be made arbitrarily small, which means that the over-estimation is arbitrarily large. However, the practical implications of such results are not yet clear.

5.3 GMRES-equivalent matrices

For a given matrix $A \in \mathbf{C}^{N\times N}$ and vector $r_0 \in \mathbf{C}^N$, for which $r_0 \in AK_N(A, r_0)$, we now characterize the matrices $B \in \mathbf{C}^{N\times N}$ that satisfy

$$AK_n(A, r_0) = BK_n(B, r_0), \quad 1 \le n \le \dim AK_N(A, r_0).$$

Since A might be singular and we allow $\dim AK_N(A, r_0) < N$, we obtain generalizations of results of Greenbaum and Strakoš [37, Section 2] (also cf. [1, 36]). In the following, by $r_n^{M,r}$ we denote the nth residual, when GMRES is applied to a linear system with the matrix M and the initial residual r.

General assumptions. A linear system (1.1) and initial guess x_0 are given such that $2 \leq \dim AK_N(A, r_0) =: d \leq N$ and $r_0 \in AK_d(A, r_0)$. The latter assumption is always satisfied when A is nonsingular (cf. Theorem 5.7). By assuming $d \geq 2$ we exclude the trivial cases of GMRES termination at the steps zero and one. Furthermore, let $w_1, w_2, ..., w_d$ be orthonormal vectors that satisfy

$$\text{span}\{w_1, w_2, ..., w_n\} = AK_n(A, r_0), \quad 1 \leq n \leq d. \tag{5.13}$$

Denoting $W_d := [w_1, w_2, ..., w_d] \in \mathbf{C}^{N \times d}$, we have

$$r_0 = W_d h, \quad \text{where } h := [\eta_1, ..., \eta_d]^T, \quad \text{and } \eta_n := w_n^H r_0, \quad 1 \leq n \leq d. \tag{5.14}$$

It is easy to see that

$$AW_d = W_d H, \tag{5.15}$$

where $H \in \mathbf{C}^{d \times d}$ is an unreduced upper Hessenberg matrix. If H was reduced, then for some n, $1 \leq n \leq d-1$, Aw_n would be a linear combination of $w_1, ..., w_n$. This contradicts the dimension assumption on $AK_N(A, r_0)$.

Lemma 5.8 *The matrix H in (5.15) is nonsingular. Furthermore, if we define*

$$\hat{H} := \begin{bmatrix} 0 & \dots & 0 & 1/\eta_d \\ 1 & & & -\eta_1/\eta_d \\ & \ddots & & \vdots \\ & & 1 & -\eta_{d-1}/\eta_d \end{bmatrix}, \tag{5.16}$$

then

$$H = \tilde{R}\hat{H}, \tag{5.17}$$

where $\tilde{R} \in \mathbf{C}^{d \times d}$ is some nonsingular and upper triangular matrix.

Proof. From $r_0 = W_d h$ and the definition of $\hat{H}$ it follows that

$$A[r_0, W_{d-1}] = AW_d \begin{bmatrix} \eta_1 & 1 & & \\ \eta_2 & & \ddots & \\ \vdots & & & 1 \\ \eta_d & 0 & \dots & 0 \end{bmatrix} = AW_d \hat{H}^{-1}. \tag{5.18}$$

From (5.13) it follows that $w_1 = e^{i\alpha} Ar_0 / \|Ar_0\|$ for some $\alpha \in \mathbf{R}$. Without loss of generality we assume $\alpha = 0$. We let H_{d-1} denote the first $d-1$ columns of the matrix H to get

$$A[r_0, W_{d-1}] = [Ar_0, AW_{d-1}] = W_d[\|Ar_0\| e_1, H_{d-1}].$$

With $W_d^H A W_d = H$ and (5.18), this yields

$$H = [\|Ar_0\|e_1, H_{d-1}]\hat{H}.$$

Since H is unreduced, $[\|Ar_0\|e_1, H_{d-1}]$ is nonsingular and upper triangular. In combination with nonsingularity of $\hat{H}$ it also follows that H is nonsingular. □

Definition 5.9 *Suppose that $A \in \mathbf{C}^{N\times N}$ and $r_0 \in \mathbf{C}^N$. Then $B \in \mathbf{C}^{N\times N}$ is called GMRES-equivalent to A with respect to r_0, if*

$$\dim BK_N(B, r_0) = d \text{ and } AK_n(A, r_0) = BK_n(B, r_0), \;\; 1 \le n \le d,$$

where $d := \dim AK_N(A, r_0)$.

Clearly, if B is GMRES-equivalent to A with respect to r_0, then

$$r_n^{A,r_0} = r_n^{B,r_0} \;\text{ for }\; 1 \le n \le d.$$

The following result is a generalization of [37, Theorems 2.1 and 2.2].

Theorem 5.10 *Under our general assumptions, for a matrix $B \in \mathbf{C}^{N\times N}$ the following three statements are equivalent.*

- *(a) B is GMRES-equivalent to A with respect to r_0.*
- *(b) B satisfies $BW_d = W_d\tilde{R}\hat{H}$, where $\hat{H}$ is defined in (5.16) and $\tilde{R} \in \mathbf{C}^{d\times d}$ is a nonsingular upper triangular matrix.*
- *(c) B satisfies $BW_d = W_d\bar{R}H$, where H is defined in (5.15) and $\bar{R} \in \mathbf{C}^{d\times d}$ is a nonsingular upper triangular matrix.*

Proof. The equivalence of (b) and (c) is a direct consequence of Lemma 5.8. We thus only have to show that (a) and (b) are equivalent. Suppose that (a) holds. Since $r_0 = W_d h$, we have

$$B[r_0, W_{d-1}] = BW_d\hat{H}^{-1}. \tag{5.19}$$

Furthermore, from (5.13) and the GMRES-equivalence of A and B follows that

$$B[r_0, W_{d-1}] = W_d\tilde{R}, \tag{5.20}$$

for some upper triangular matrix $\tilde{R}$. Now (5.19) and (5.20) yield $BW_d = W_d\tilde{R}\hat{H}$. To show that (b) holds, we have to show that the matrix $\tilde{R} = (r_{ij})$ in (5.20) must be nonsingular. Since $\dim BK_N(B, r_0) = d$, we have $B^n r_0 \neq 0$ for $1 \le n \le d$. In particular, $Br_0 \neq 0$ and hence $r_{11} \neq 0$. Moreover,

$$w_n = \sum_{k=1}^{n} \alpha_k B^k r_0, \;\; \text{for } \; 1 \le n \le d,$$

and some complex numbers, $\alpha_1, ..., \alpha_n$, not all zero. Therefore,

$$Bw_n = \sum_{k=1}^{n} \alpha_k B^{k+1} r_0 \neq 0, \;\; \text{for } \; 1 \le n \le d-1.$$

Now suppose that $r_{nn} = 0$ for some n, $2 \le n \le d$. Then the nth column of (5.20) is

$$Bw_{n-1} = \sum_{k=1}^{n-1} r_{kn} w_n \neq 0.$$

But from this follows $\dim BK_n(B, r_0) \le n-1$, which contradicts our assumption on the dimension of $BK_N(B, r_0)$. Thus, $r_{nn} \neq 0$ for $1 \le n \le d$, i.e. $\tilde{R}$ is nonsingular.

Conversely, if (b) holds, then $B^n r_0 = B^n W_d h = r_{11} W_d (\tilde{R}\hat{H})^{n-1} e_1$. Since $\tilde{R}$ is nonsingular and upper triangular, $AK_n(A, r_0) = \operatorname{span}\{w_1, w_2, ..., w_n\} = BK_n(B, r_0)$ for $1 \le n \le d$. Finally we note that $B^{d+1} r_0 \in BK_d(B, r_0)$, so that $\dim BK_N(B, r_0) = d$. □

By varying $\tilde{R}$ and $\bar{R}$ in Theorem 5.10, we obtain a whole class of matrices which are GMRES-equivalent to the given matrix A with respect to the initial residual $r_0 = W_d h$. We obtain a useful corollary.

Corollary 5.11 *Suppose that, in the notation of Theorem 5.10,*

$$B = W_d \bar{R} H W_d^H \quad \text{and} \quad \hat{B} = W_d \tilde{R} \hat{H} W_d^H,$$

where $\bar{R}$ and $\tilde{R}$ are arbitrary nonsingular upper triangular matrices. Then B and $\hat{B}$ are both GMRES-equivalent to A with respect to r_0, and in particular,

$$\|r_n^{B,r_0}\| = \|r_n^{\hat{B},r_0}\| = \|r_n^{A,r_0}\|, \quad 1 \le n \le d.$$

We remark that if $d < N$, then the matrices B and $\hat{B}$ in Corollary 5.11 must be singular, regardless of A being singular or not.

5.4 New bounds on GMRES residual norms

The results of this section are derived under the general assumptions stated at the beginning of Section 5.3. Using Theorem 5.10, we can relate the convergence of GMRES for the given linear system to the convergence for some other system, *which is easier to analyze.* Our starting point for the derivation of the new convergence bounds therefore is the construction of suitable matrices that are GMRES-equivalent to A with respect to the given r_0.

As stated in Corollary 5.11, for arbitrary nonsingular upper triangular matrices $\bar{R}, \tilde{R} \in \mathbf{C}^{d \times d}$, the matrices

$$W_d \bar{R} H W_d^H \quad \text{and} \quad W_d \tilde{R} \hat{H} W_d^H \tag{5.21}$$

are both GMRES-equivalent to A with respect to r_0. We consider the following decompositions of H and $\hat{H}$,

$$H = QR, \quad \text{and} \quad \hat{H} = \hat{R}\hat{Q}. \tag{5.22}$$

The matrices Q and $\hat{Q}$ are unitary and the matrices R and $\hat{R}$ are nonsingular and upper triangular. All four are of size $d \times d$. These two decompositions always exist, due to nonsingularity of H and $\hat{H}$ (see e.g. [32]).

Two other similar decompositions are possible, $H = R_1Q_1$ and $\hat{H} = Q_2R_2$. We are particularly interested in the spectra of the Q-factors of these decompositions. Since Q_2 is simply a unitary shift matrix and the results derived for $\Lambda(\hat{Q})$ also hold for $\Lambda(Q_1)$ (cf. (5.17)), we do not consider these two other decompositions.

For later use we note that $W_d^H A W_d = QR$ yields

$$\kappa(R) \leq \kappa(A), \tag{5.23}$$

with equality if $d = N$. Often $d = N$ is a reasonable assumption, in particular when the 'generic' case of an initial residual without special properties is to be analyzed.

Choosing $\bar{R} = R^{-1}$ and $\tilde{R} = \hat{R}^{-1}$ in (5.21), and using (5.22), we define the matrices

$$B := W_d R^{-1} Q R W_d^H \quad \text{and} \quad \hat{B} := W_d \hat{Q} W_d^H, \tag{5.24}$$

which are both GMRES-equivalent to A with respect to r_0.

Remark 5.12 Greenbaum and Strakoš [37, Section 3.1] consider the RQ-decomposition of the matrix H. They point out that if $d = N$, then the matrix $\hat{B} = W_N \hat{Q} W_N^H$ is unitary. Thus, whenever A and r_0 are such that $d = N$, there exists a unitary matrix which is GMRES-equivalent to A with respect to r_0.

Theorem 5.13 *Suppose that B and $\hat{B}$ are defined as in (5.24). Then*

$$\frac{\|r_n\|}{\|r_0\|} \leq \kappa(R) \min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(Q)} |p_n(\lambda)|, \tag{5.25}$$

and

$$\frac{\|r_n\|}{\|r_0\|} \leq \min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(\hat{Q})} |p_n(\lambda)|, \tag{5.26}$$

for $1 \leq n \leq d$.

Proof. Corollary 5.11 yields for $1 \leq n \leq d$,

$$\begin{aligned} \|r_n\| \equiv \|r_n^{A,r_0}\| &= \|r_n^{B,r_0}\| = \min_{p_n \in \pi_n} \|W_d R^{-1} p_n(Q) R W_d^H r_0\| \\ &\leq \|R^{-1}\| \, \|R\| \min_{p_n \in \pi_n} \|p_n(Q)\| \, \|r_0\|, \end{aligned}$$

from which (5.25) follows. The proof of (5.26) is similar. □

To study the approximation problem

$$\min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(U)} |p_n(\lambda)|, \quad U \text{ unitary}, \tag{5.27}$$

we need the following definition.

Definition 5.14 *Suppose that the spectrum of the $d \times d$ unitary matrix U is given by $\Lambda(U) := \{e^{i\beta_j} \,:\, 0 \leq \beta_1 \leq \beta_2 \leq \cdots \leq \beta_d < 2\pi\}$ and let $\beta_{d+1} := 2\pi + \beta_1$. We define the d gaps in $\Lambda(U)$ by*

$$g_j := \beta_{j+1} - \beta_j, \quad 1 \leq j \leq d. \tag{5.28}$$

For $1 \leq j \leq d$, the complex number $e^{i(\beta_{j+1}+\beta_j)/2}$ is called the center of the gap g_j. $\phi := \max\{g_1, g_2, ..., g_d\}$ denotes the largest gap in $\Lambda(U)$.

Note that for every unitary matrix U and every $\alpha \in \mathbf{R}$, we have

$$\min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(U)} |p_n(\lambda)| = \min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(e^{i\alpha} U)} |p_n(\lambda)|. \tag{5.29}$$

Therefore, in the study of (5.27) we can assume without loss of generality, that one is the center of the largest gap ϕ in $\Lambda(U)$. Now suppose that $\phi \in (0, 2\pi)$, or, equivalently, that $\Lambda(U)$ contains at least two distinct points. Then $\Lambda(U) \subset \Omega_\epsilon \in \mathbf{B}(1, \phi)$ for each $\epsilon \in [0, \epsilon_{max})$, and

$$\min_{p_n \in \pi_n} \max_{\lambda \in \Lambda(U)} |p_n(\lambda)| \leq \min_{p_n \in \pi_n} \max_{z \in \Omega_\epsilon} |p_n(z)|, \quad \Omega_\epsilon \in \mathbf{B}(1, \phi). \tag{5.30}$$

The tightest fit around $\Lambda(U)$ is achieved by choosing $\epsilon = 0$, i.e.,

$$\Lambda(U) \subset \Omega_0 := \left\{ e^{i\beta} \ : \ \frac{\phi}{2} \leq \beta \leq 2\pi - \frac{\phi}{2} \right\} \in \mathbf{B}(1, \phi). \tag{5.31}$$

We now apply Theorem 4.1 to derive an upper bound on the maximum norm of the normalized Faber polynomials for Ω_0. In the notation of Theorem 4.1, the parameters corresponding to $s = 1$, $\lambda_m = 1$, $\phi \in (0, 2\pi)$ and $\epsilon = 0$ are

$$N_0 = \left(\cos \frac{\phi}{4}\right)^{-1} =: \gamma, \quad M_0 = 0,$$

$$S_0 = -\frac{1}{N_0} = -\gamma^{-1}, \quad t_\epsilon = \frac{1}{N_0} = \gamma^{-1}, \quad Q_0 = S_0(N_0 + S_0).$$

Furthermore, $V(z) = z$, and thus for $n \geq 1$, the nth Faber polynomial for Ω_0 satisfies

$$F_n(z) = 2 z^{n/2} C_n \left(\gamma \frac{z^{-1/2} + z^{1/2}}{2} \right) - \gamma^{-n}, \quad z \in \mathbf{C} \setminus \{0\}, \tag{5.32}$$

and

$$F_n(0) = \gamma^n - \gamma^{-n}. \tag{5.33}$$

Remark 5.15 He [40] derives a different expression for the Faber polynomials for circular arcs, not revealing the relation to the Chebyshev polynomials. This relation, however, is a key ingredient for estimating the maximum norm of the normalized Faber polynomials for Ω_0.

For $z = e^{i\beta}$,

$$\frac{z^{-1/2} + z^{1/2}}{2} = \cos \frac{\beta}{2},$$

which, inserted into (5.32), leads to

$$F_n\left(e^{i\beta}\right) = 2 e^{in\beta/2} C_n \left(\gamma \cos \frac{\beta}{2} \right) - \gamma^{-n}. \tag{5.34}$$

If $\beta \in [\phi/2, 2\pi - \phi/2]$, then $\gamma \cos \frac{\beta}{2} \in [-1, 1]$. Using the well-known bound

$$|C_n(z)| \leq 1, \quad n \geq 0, \quad z \in [-1, 1]$$

(see e.g. [34, Chapter 3]), we see from (5.34) that

$$|F_n(z)| \leq 2 + \gamma^{-n}, \quad z \in \Omega_0. \tag{5.35}$$

Combining (5.35) and (5.33) gives the following lemma.

Lemma 5.16 *Suppose that $\phi \in (0, 2\pi)$ and that $F_n(z)$ is the nth Faber polynomial for $\Omega_0 \in \mathbf{B}(1, \phi)$. Then*

$$\max_{z\in\Omega_0} |\tilde{F}_n(z)| = \max_{z\in\Omega_0} \left|\frac{F_n(z)}{F_n(0)}\right| \leq \frac{2+\gamma^{-n}}{\gamma^n - \gamma^{-n}}, \quad n \geq 1,$$

where $\gamma := \left(\cos \frac{\phi}{4}\right)^{-1}$.

As Lemma 5.16 shows, the normalized Faber polynomials converge to zero on Ω_0, with the speed of convergence being directly related to the largest gap ϕ: The larger ϕ, the faster the convergence and vice versa.

We now combine (5.23), (5.30), Theorem 5.13 and Lemma 5.16 for our main convergence result. Note that from our general assumption $d \geq 2$ follows that $\Lambda(Q)$ and $\Lambda(\hat{Q})$ both contain at least two distinct points, so that Lemma 5.16 is applicable.

Theorem 5.17 *Suppose that a linear system (1.1) and an initial guess x_0 are given, such that $2 \leq d := \dim AK_N(A, r_0) \leq N$ and $r_0 \in AK_d(A, r_0)$. Define the matrices Q, R and $\hat{Q}$ as in (5.22), and let ϕ and $\hat{\phi}$ denote the largest gaps in $\Lambda(Q)$ and $\Lambda(\hat{Q})$, respectively. Define $\gamma := \left(\cos \frac{\phi}{4}\right)^{-1}$ and $\hat{\gamma} := \left(\cos \frac{\hat{\phi}}{4}\right)^{-1}$, then the GMRES residuals for (1.1) and x_0 satisfy*

$$\frac{\|r_n\|}{\|r_0\|} \leq \kappa(R) \frac{2+\gamma^{-n}}{\gamma^n - \gamma^{-n}}, \tag{5.36}$$

and

$$\frac{\|r_n\|}{\|r_0\|} \leq \frac{2+\hat{\gamma}^{-n}}{\hat{\gamma}^n - \hat{\gamma}^{-n}}, \tag{5.37}$$

for $1 \leq n \leq d$. Furthermore, $\kappa(R) \leq \kappa(A)$, with equality if $d = N$.

To analyze Theorem 5.17 and its implications, we first point out that γ^{-1} and $\hat{\gamma}^{-1}$, the rates of reduction of the bounds (5.36) and (5.37), *depend on the initial guess*, and that both rates lie strictly between zero and one. We have not found an algebraic relation between the two rates. As we will show by numerical examples, γ^{-1} and $\hat{\gamma}^{-1}$ might be of different sizes and have the potential to describe the rate of convergence of GMRES for the given initial residual.

At first sight, (5.37) seems to be more useful than (5.36), since it does not include the constant $\kappa(R)$ (or $\kappa(A)$). However, it is possible to implement GMRES in a way so that approximations to (5.36) are easily computable in each step of the algorithm. Moreover, these approximations lead to a heuristic which often allows us to early predict the behavior of GMRES in later stages of the computation. The situation is different for (5.37). Computing approximations to this bound during the GMRES run is sometimes not possible and is always more expensive. Furthermore, we have not found a heuristic for the behavior of these approximations.

From our point of view, the appearance of the constant $\kappa(R)$ (or $\kappa(A)$) in an upper bound on GMRES residual norms is not a nuisance, but rather indicates the algorithm's dependence on the conditioning of the coefficient matrix. Ill-conditioning of A often leads to slow convergence of GMRES. Since well-conditioning does not imply fast GMRES convergence, it is generally difficult to strengthen this statement. The worst possible GMRES behavior, no progress until step N, can even occur if A is perfectly conditioned. Just consider an $N \times N$ unitary matrix, having the first N roots of unity as its eigenvalues (cf. [56, Example C]).

5.5 Approximate bounds predict the behavior of GMRES

Note that if close approximations to the right-hand sides of (5.36) and (5.37) are available in early stages of the computation, these potentially *predict* the behavior of GMRES in later stages. It is therefore useful to compute such approximations during the GMRES run.

5.5.1 WZ-GMRES

To obtain the approximations, we consider an implementation of GMRES first proposed by Walker and Zhou [77]. In the following we will refer to this implementation as WZ-GMRES.

The original implementation of GMRES [66] is based on the Arnoldi process applied to the matrix A and the vector $r_0/\|r_0\|$. However, following [77], we will use

$$w_1 := Ar_0/\|Ar_0\| \tag{5.38}$$

as the starting vector for the Arnoldi process. The resulting matrix equations are given in (2.3)-(2.5). Recall that for our given linear system and initial guess we have assumed $\dim AK_N(A, r_0) = d$. Then, since

$$K_n(A, w_1) = AK_n(A, r_0), \quad 1 \leq n \leq d,$$

the application of the Arnoldi process to A and w_1 yields orthonormal vectors $w_1, w_2, ..., w_d$, that satisfy (5.13). Using these vectors, (5.1) can be rewritten as

$$x_n = x_0 + [r_0/\|r_0\|, W_{n-1}]y_n \tag{5.39}$$

for some $y_n \in \mathbf{C}^n$, $1 \leq n \leq d$. For $n = 1$, W_0 is simply an empty matrix. To implement GMRES we use the orthogonality relation (5.4) in the computation of y_n. Together with (2.3) and (5.39), this yields

$$0 = W_n^H r_n = W_n^H r_0 - W_n^H A[r_0/\|r_0\|, W_{n-1}]y_n,$$

which is equivalent to

$$[(\|Ar_0\|/\|r_0\|)e_1, H_{n-1}^e]y_n = W_n^H r_0. \tag{5.40}$$

Thus, y_n is the solution of an upper triangular system. Since H_{n-1}^e is unreduced, y_n is well defined. It is important to note that r_n can be obtained without explicitly computing x_n [63],

$$r_n = r_{n-1} - (w_n^H r_0)w_n. \tag{5.41}$$

The upper triangular solve for y_n therefore has to be performed only if the residual norm computed from (5.41) is small enough. Clearly, in exact arithmetic, the termination of the Arnoldi process at step d is equivalent to the termination of WZ-GMRES at step d. WZ-GMRES is a complete and economical implementation of the GMRES algorithm.

5.5.2 Numerical stability of WZ-GMRES

Before we proceed with the derivation of approximations to the bounds (5.36) and (5.37), we comment on the numerical stability of WZ-GMRES. Note that the matrix

$$V_n := [r_0/\|r_0\|, W_{n-1}]\,, \tag{5.42}$$

used in (5.39) to compute the iterate x_n does not have orthogonal columns. This is a significant difference of WZ-GMRES from the original implementation of GMRES (cf. [66, Formula (8)]). However, assuming exact arithmetic, Walker and Zhou [77, Lemma 3.1] show that

$$\kappa(V_n) \le 2\frac{\|r_0\|}{\|r_{n-1}\|}, \quad n \ge 1. \tag{5.43}$$

Thus, V_n can only be ill conditioned if the scaled residual norms have decreased considerably. We next derive an exact expression for $\kappa(V_n)$.

Lemma 5.18 *Using the above notation,*

$$\kappa(V_n) = \frac{\|r_0\| + \left(\|r_0\|^2 - \|r_{n-1}\|^2\right)^{1/2}}{\|r_{n-1}\|}, \quad 1 \le n \le d. \tag{5.44}$$

Proof. The assertion is obvious for $n = 1$, and thus we assume that $2 \le n \le d$. Using (5.14), we get

$$V_n = W_n \begin{bmatrix} \eta_1/\|r_0\| & 1 & & \\ \eta_2/\|r_0\| & & \ddots & \\ \vdots & & & 1 \\ \eta_d/\|r_0\| & 0 & \dots & 0 \end{bmatrix},$$

where $\eta_k = w_k^H r_0$ for $1 \le k \le d$. We apply the orthogonality relation (5.4), $r_k \perp w_k \in AK_k(A, r_0)$, to (5.41), yielding

$$\|r_{k-1}\|^2 = \|r_k + \eta_k w_k\|^2 = \|r_k\|^2 + |\eta_k|^2,$$

i.e.,

$$|\eta_k|^2 = \|r_{k-1}\|^2 - \|r_k\|^2, \quad 1 \le k \le d. \tag{5.45}$$

Now note that $\kappa(V_n) = (\lambda_{max}/\lambda_{min})^{1/2}$, where λ_{max} and λ_{min} are the largest and smallest eigenvalues of

$$V_n^H V_n = \begin{bmatrix} 1 & u^H \\ u & I \end{bmatrix}, \tag{5.46}$$

where $u = [\eta_1/\|r_0\|, ..., \eta_{n-1}/\|r_0\|]^T$. We consider two cases:

Case 1: GMRES stagnates from step 1 to step $n-1$, i.e.

$$\|r_0\| = \|r_1\| = ... = \|r_{n-1}\|.$$

This yields $V_n^H V_n = I$, and (5.44) is satisfied with $\kappa(V_n) = 1$.

Case 2: GMRES does not stagnate from step 1 to step $n-1$, i.e.

$$\|r_0\| \neq \|r_{n-1}\|.$$

Then

$$\|u\| = \frac{\left(\|r_0\|^2 - \|r_{n-1}\|^2\right)^{1/2}}{\|r_0\|} \neq 0,$$

and there exists a unitary matrix Q with $Qu = \|u\|e_1$. Hence,

$$\begin{bmatrix} 1 & \\ & Q \end{bmatrix} \begin{bmatrix} 1 & u^H \\ u & I \end{bmatrix} \begin{bmatrix} 1 & \\ & Q^H \end{bmatrix} = \begin{bmatrix} 1 & \|u\| & \\ \|u\| & 1 & \\ & & I \end{bmatrix},$$

and thus

$$\Lambda(V_n^H V_n) = \{1 - \|u\|, 1, 1 + \|u\|\},$$

from which (5.44) follows. □

Lemma 5.18 shows that, in exact arithmetic, V_n is ill conditioned if and only if $\|r_{n-1}\|$ is small compared to $\|r_0\|$. Thus, even in finite precision arithmetic, numerical difficulties in (5.39) are unlikely to occur before the ratio $\|r_{n-1}\|/\|r_0\|$ is small. Moreover, ill-conditioning of V_n is necessary and sufficient for fast convergence of GMRES! Surprising results of this type recently have been of great interest (see e.g. [45, 62] and the references therein).

As noted in [77], the scaling of the first column of V_n is important for the numerical stability of WZ-GMRES. This observation can be made precise using the following result.

Lemma 5.19 *Using the above notation, define* $\tilde{V}_n := [r_0, W_{n-1}]$. *Then*

$$\kappa(\tilde{V}_n) = \frac{\alpha + \left(\alpha^2 - 4\|r_{n-1}\|^2\right)^{1/2}}{2\|r_{n-1}\|}, \quad 1 \le n \le d, \tag{5.47}$$

where $\alpha := 1 + \|r_0\|^2$.

Proof. The proof is similar to the proof of Lemma 5.18. We thus only state that the largest and smallest eigenvalues of $\tilde{V}_n^H \tilde{V}_n$ are

$$\lambda_{max} = \frac{1}{2}\left(\alpha + \left(\alpha^2 - 4\|r_{n-1}\|^2\right)^{1/2}\right),$$
$$\lambda_{min} = \frac{1}{2}\left(\alpha - \left(\alpha^2 - 4\|r_{n-1}\|^2\right)^{1/2}\right).$$

□

An easy computation shows that

$$\kappa(\tilde{V}_n) - \kappa(V_n) \geq \frac{(\|r_0\| - 1)^2}{\|r_{n-1}\|}. \tag{5.48}$$

Hence, if $\|r_0\| \neq 1$, then with decreasing residual norms, $\kappa(\tilde{V}_n)$ grows much faster than $\kappa(V_n)$. We remark that scaling the first column of V_n to have unit length yields the optimal condition number of all possible scalings (cf. e.g. [11, Section 2]).

For further details about WZ-GMRES we refer to [63, 77].

5.5.3 Computing approximations to (5.36)

If (5.38) holds, then the vectors $w_1, w_2, ..., w_d$ produced by the Arnoldi process satisfy (5.13). Consequently, the matrix H_d in (2.5) can be identified with the matrix H in (5.22). We can therefore compute the numbers ϕ and $\kappa(R)$ used in (5.36) from the QR-decomposition of H_d. Thus, the right-hand side of (5.36) is computable from information generated during the run of WZ-GMRES. Moreover, because of the recursive nature of the Arnoldi process, the matrix H_n^e is the leading $(n+1) \times n$ submatrix of H_d. For $n = 1, 2, ..., d-1$, we can therefore decompose

$$H_n^e = Q_n \begin{bmatrix} R_n \\ 0 \end{bmatrix}, \quad Q_n \in \mathbf{C}^{(n+1)\times(n+1)}, \quad R_n \in \mathbf{C}^{n \times n}, \tag{5.49}$$

and use ϕ_n, the largest gap in $\Lambda(Q_n)$, as an approximation to ϕ and $\kappa(R_n)$ as an approximation to $\kappa(R)$. Note that because H_n^e is unreduced, R_n is nonsingular and (5.49) is well defined.

The decomposition (5.49) is particularly easy to compute. It is well known, that because of the special structure of H_n^e, the matrix Q_n can be derived as a product of n Givens transformations, $Q_n = G_1 \cdot .... \cdot G_n$, where

$$G_j = \begin{bmatrix} I_{j-1} & & & \\ & c_j & -\overline{s}_j & \\ & s_j & c_j & \\ & & & I_{n-j} \end{bmatrix} \in \mathbf{C}^{(n+1)\times(n+1)}, \quad 1 \leq j \leq n. \tag{5.50}$$

Suppose that for some n, $2 \leq n \leq d-1$,

$$Q_{n-1}^H H_{n-1}^e = \begin{bmatrix} R_{n-1} \\ 0 \end{bmatrix}, \quad Q_{n-1}^H = G_{n-1}^H \cdot ... \cdot G_1^H. \tag{5.51}$$

We define

$$h^n := [h_{1n}, ..., h_{nn}]^T, [\rho_1, ..., \rho_n]^T := Q_{n-1}^H h^n, \quad \text{and} \quad \nu_n := [\rho_1, ..., \rho_{n-1}]^T.$$

Then

$$\begin{aligned} Q_n^H H_n^e &= G_n^H \begin{bmatrix} Q_{n-1}^H & \\ & 1 \end{bmatrix} \begin{bmatrix} H_{n-1}^e & h^n \\ & h_{n+1,n} \end{bmatrix} \\ &= G_n^H \begin{bmatrix} R_{n-1} & \nu_n \\ & \rho_n \\ & h_{n+1,n} \end{bmatrix}. \end{aligned} \tag{5.52}$$

If $\rho_n = |\rho_n| e^{i\delta_n}$, we define $\omega_n := (|\rho_n|^2 + |h_{n+1,n}|^2)^{1/2}$ and choose

$$c_n = \frac{|\rho_n|}{\omega_n}, \tag{5.53}$$

$$s_n = \frac{e^{-i\delta_n} h_{n+1,n}}{\omega_n}, \tag{5.54}$$

to give (5.52) the required upper triangular form. Hence, to compute Q_n and R_n from Q_{n-1} and R_{n-1}, respectively, only one matrix-vector product and a few flops to compute c_n and s_n are required.

To compute ϕ_n we have to solve an $n \times n$ unitary eigenproblem. For the efficient computation of $\kappa(R_n)$ we can use the method of incremental condition estimation [5]. With this method, the computation of $\kappa(R_n)$ costs only $\mathcal{O}(n)$ operations, provided $\kappa(R_{n-1})$ is known.

5.5.4 Computing approximations to (5.37)

Suppose that we apply WZ-GMRES to (1.1) and x_0 and use the Arnoldi vectors $w_1, w_2, ..., w_d$ to set up the matrix $\hat{H}$ as defined in (5.16). In particular, $\hat{H}$ can be obtained at no additional cost, if the updated residuals are computed using (5.41). The bound (5.37) depends only on the largest gap $\hat{\phi}$ in $\Lambda(\hat{Q})$, where $\hat{Q}$ is defined by the RQ-decomposition $\hat{H} = \hat{R}\hat{Q}$ (cf. (5.22)).

From (5.41) it follows that

$$|\eta_n| = |w_n^H r_0| = \|r_{n-1} - r_n\|, \quad \text{for} \quad 1 \le n \le d.$$

Hence, if GMRES does not stagnate in step n, $2 \le n \le d-1$, then $\eta_n \ne 0$, and the decomposition

$$\hat{H}_n := \begin{bmatrix} 0 & \dots & 0 & 1/\eta_n \\ 1 & & & -\eta_1/\eta_n \\ & \ddots & & \vdots \\ & & 1 & -\eta_{n-1}/\eta_n \end{bmatrix} = \hat{R}_n \hat{Q}_n \quad \in \mathbf{C}^{n \times n} \tag{5.55}$$

is defined. In this case, $\hat{\phi}_n$, the largest gap in $\Lambda(\hat{Q}_n)$, can be used as an approximation to $\hat{\phi}$. However, we have no easy way to update the decomposition (5.55). $\hat{R}_n$ and $\hat{Q}_n$ have to be recomputed from scratch in each step. As shown above, this is drastically different for the decomposition (5.49).

Approximations $\hat{\phi}_n$ can also be computed from the RQ-decomposition of the matrix $H_n := W_n^H A W_n$ (cf. (2.6)). In exact arithmetic, stagnation of WZ-GMRES in step n leads to singularity of H_n. Hence $\hat{\phi}_n$ is not well defined in this case. But on the contrary to (5.55), an approximation $\hat{\phi}_n$ might still be computable. This is particularly true in finite precision arithmetic.

5.5.5 Behavior of the approximations and prediction of the performance of GMRES

Q_n in (5.49) is obtained from Q_{n-1} by a low rank modification. In particular, if $q_1, ..., q_n$ denote the columns of Q_{n-1}, then

$$Q_n = \begin{bmatrix} q_1 & \dots & q_{n-1} & c_n q_n & -\overline{s}_n q_n \\ 0 & \dots & 0 & s_n & c_n \end{bmatrix}.$$

By (5.53), $c_n \neq -1$, so that we can readily apply results of Elsner and He [19, Section 5]. We partition

$$Q_n = \begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix}, \quad U_{11} = [q_1, ..., q_{n-1}, c_n q_n] \in \mathbf{C}^{n\times n}, \quad U_{22} = c_n,$$

to get, in the notation of [19, Theorem 5.3],

$$U_n := U_{11} - U_{12}(I + U_{22})^{-1} U_{21} = Q_{n-1}.$$

From [19, Theorem 5.3] follows, if

$$\begin{aligned} \Lambda(Q_{n-1}) &= \{e^{i\beta_j} \; : \; 0 \le \beta_1 \le ... \le \beta_n < 2\pi\} \text{ and} \\ \Lambda(Q_n) &= \{e^{i\tau_j} \; : \; 0 \le \tau_1 \le ... \le \tau_{n+1} < 2\pi\}, \end{aligned}$$

then

$$\tau_j \le \beta_j \le \tau_{j+1}, \text{ for } 1 \le j \le n.$$

Because of this interlacing property, we can expect a 'smooth' behavior of the eigenvalues of the matrices Q_n and, in particular, of their largest gaps. This is not necessarily the case for the spectra of the matrices $\hat{Q}_n$ in (5.55), since we know of no easy algebraic relation between them. We remark that often for small n, ϕ_n and $\hat{\phi}_n$ are very close to ϕ and $\hat{\phi}$, respectively. We found no clear explanation for these observations.

The matrix R_n in (5.49) is the leading principal submatrix of the matrix R in (5.22). Thus, we have $\kappa(R_{n-1}) \le \kappa(R_n) \le \kappa(R)$ for $n = 2, ..., d-1$. Monitoring the numbers $\kappa(R_n)$ is a useful check on the numerical accuracy of the WZ-GMRES iterates. As (5.40) shows, the conditioning of R_{n-1} determines the accuracy of the vector y_n. In the derivation of our bounds we have assumed $r_0 \in AK_d(A, r_0)$ for some $d \le N$. From this assumption follows that $\kappa(R_n) \le \kappa(R) < \infty$. But for ill-conditioned (or singular) A, $\kappa(R)$ and thus $\kappa(R_n)$ might still be very large. In such case, it might be useful to terminate WZ-GMRES early.

We point out that the bounds (5.36) and (5.37) are independent of each other and might be of different use. On the one hand, if GMRES stagnates early in the iteration, then $\hat{\phi}$ must be small, since (5.37) involves no large constant. Thus, $\hat{\phi}$ and, if defined, the $\hat{\phi}_n$ cannot predict if GMRES will converge faster in later phases of the iteration. On the other hand, early stagnation of GMRES does not imply that ϕ is small, since (5.36) involves the potentially large constant $\kappa(R)$ (or $\kappa(A)$). This has interesting practical consequences. Suppose that GMRES stagnates early in the iteration and that the ϕ_n remain constant during this phase of stagnation. Then they are possibly close approximations to ϕ and if they are large, they predict that the algorithm will converge well later in the iteration. We are led to the following heuristic, which is confirmed by numerical experiments in Section 5.6.

Heuristic 5.20 *Suppose that for small n, the ϕ_n stagnate on a (relatively) high level. Then they are often close approximations to ϕ and their size predicts that GMRES will converge well in later phases of the iteration. On the other hand, a fast decrease of the ϕ_n to zero predicts that GMRES will converge slowly over the complete course of the iteration.*

5.6 Numerical experiments

We consider three example matrices, BWM200, PDE225 and FS1836. Information on these matrices is given in Appendix A.

We apply a Householder version [76] of WZ-GMRES to linear systems with these matrices. We compute the approximations ϕ_n as described in Section 6.2 and the $\hat{\phi}_n$ from the RQ-decomposition of $H_n = W_n^H A W_n$ (cf. Section 6.3). To check the numerical accuracy of our computations, we compare the norms of the directly computed residuals $(b - Ax_n)$ and the 'updated' residuals (cf. (5.41)).

Before analyzing our results we point out that the bounds (5.36) and (5.37), in particular the approximations ϕ_n and $\hat{\phi}_n$, are sensitive to the scaling of the diagonal entries of the R-factors in the QR- and RQ-decompositions. The results of this section are derived with the following choices:

- The QR-decompositions (5.49) are computed using Givens transformations with nonnegative cosines, i.e. $c_n \geq 0$ for all n (cf. (5.53)).
- The RQ-decompositions $H_n = \hat{R}_n \hat{Q}_n$ are computed so that the $\hat{R}_n$ have positive diagonal entries.

Of all tested scalings, these choices typically lead to the largest gaps ϕ and $\hat{\phi}$ and thus yield the smallest upper bounds (5.36) and (5.37).

Example 5.21 We use A=BWM200, $b = [1, ..., 1]^T$ and $x_0 = 0$. We terminate WZ-GMRES at step d when the updated scaled residual norms satisfy

$$\|r_d\|/\|r_0\| < 10^{-15}.$$

Since the residual norm has almost dropped to the machine precision, $\mathbf{u} \approx 10^{-16}$, we are justified to identify d with $\dim AK_N(A, r_0)$. Hence, we consider the last computed approximations, ϕ_d, $\hat{\phi}_d$ and $\kappa(R_d)$ as equal to the respective quantities used in Theorem 5.17.

Figure 15 shows the computed (solid) and updated (dashed) scaled residual norms produced by WZ-GMRES. These two curves are the same, except for the last few steps, when the algorithm has reached its final accuracy level. We also plot the upper bounds (5.36) (pluses) and (5.37) (dashdot). We stress that the actual right-hand sides of (5.36) and (5.37) are plotted, and not just the asymptotic convergence factors γ^{-1} and $\hat{\gamma}^{-1}$. Figures 16 and 17 show the approximations ϕ_n (solid), $\hat{\phi}_n$ (dashed) and $\kappa(R_n)$, respectively, during the run of the algorithm. The final quantities are listed in the following table.

	N	d	ϕ	$\hat{\phi}$	$\kappa(R)$	$\kappa(A)$
BWM200	200	99	2.9629	2.9625	3.4	3.4

For this example, (5.36) and (5.37) describe the actual convergence curve very well. Due to the constant $\kappa(R)$, (5.36) is slightly weaker than (5.37). Figure 16 shows that the ϕ_n (solid) converge 'smoothly' to their final level, while the $\hat{\phi}_n$ (dashed) show an erratic behavior. Only after step 85 the two approximations are about the same. With respect to Heuristic 5.20, we remark that the smooth behavior of the ϕ_n and their early stagnation at a high level predicts that GMRES will not stagnate in later phases of the iteration. This is indeed the case. Finally, Figure 17 shows that $\kappa(R_n) \approx \kappa(R) = \kappa(A)$ after only a few steps, so that close approximations to the right-hand side of (5.36) are obtainable early in the iteration.

Example 5.22 We use A=PDE225, $b = A[1, ..., 1]^T$, $x_0 = 0$ and the same termination criterion as for BWM200. Figures 18, 19 and 20 have the same notation as Figures 15, 16 and 17, respectively.

	N	d	ϕ	$\hat{\phi}$	$\kappa(R)$	$\kappa(A)$
PDE225	225	97	1.9675	1.4250	35.1	39.1

GMRES converges slowly for the first 30 steps, so that ϕ is potentially larger than $\hat{\phi}$. This heuristic is confirmed by the actual values. Also, in accordance with Heuristic 5.20, the increase in the speed of convergence of GMRES is predicted by the large size of the ϕ_n early in the iteration. However, since our bounds decrease with the constant rates γ^{-1} and $\hat{\gamma}^{-1}$, they are generally not capable to approximate closely a curve with a dramatically changing slope. Hence, the bounds in this example cannot be as tight as for BWM200. As Figures 19 and 20 show, close approximations to both our bounds are available after 40 steps.

Example 5.23 We use A=FS1836, two different right-hand sides,

$$b^{(1)} = [1, ..., 1]^T, \;\; b^{(2)} = A[1, ..., 1]^T,$$

and $x_0 = 0$ in both cases. A is ill conditioned, $\kappa(A) \approx 10^{11}$, and the performance of GMRES for a linear system with A is very sensitive to the initial residual.

Figure 21 shows the convergence curves for $b^{(1)}$ and $b^{(2)}$ (dashed/solid and pluses/dashdot for computed/updated residual norms, respectively).

For $b^{(1)}$, fast convergence sets in after 25 steps of total stagnation. There is no stagnation for $b^{(2)}$. Due to ill-conditioning of A, in both cases a finite precision accuracy of little more than $\mathbf{u} * \kappa(A)$ is achievable by WZ-GMRES. We terminate WZ-GMRES shortly after we notice a considerable difference between the computed and updated residuals. The quantities ϕ_n, $\hat{\phi}_n$ and $\kappa(R_n)$

at the termination step n are listed in the table below. After termination of WZ-GMRES we continue the Arnoldi process to compute ϕ and $\hat{\phi}$. For both right-hand sides, these are close to the last computed approximations. Figure 22 shows the gaps ϕ_n and $\hat{\phi}_n$ (solid and dashed for $b^{(1)}$, dashdot and pluses for $b^{(2)}$, respectively) during the iteration. Figure 23 shows the condition numbers $\kappa(R_n)$ (solid for $b^{(1)}$, dashdot for $b^{(2)}$), which for both right-hand sides quickly approach $\kappa(A)$.

	n	ϕ_n	ϕ	$\hat{\phi}_n$	$\hat{\phi}$	$\kappa(R_n)$
FS1836 ($b^{(1)}$)	50	2.2044	2.1931	0.2497	0.2497	1.1e+11
FS1836 ($b^{(2)}$)	25	1.3674	1.2548	0.8892	0.8909	3.3e+09

Our bounds, particularly (5.36), overestimate the actual convergence curves, and hence they are not included in Figure 21. Still our theory provides useful information:

WZ-GMRES stagnates for $b^{(1)}$, so $\hat{\phi}$ should be (and is) small. On the other hand, ϕ for $b^{(1)}$ is large and the behavior of the ϕ_n predicts for small n that the algorithm eventually will stop to stagnate. If interpreted as 'rates of convergence', $\hat{\gamma}^{-1} \approx 0.99$ describes the first phase of the iteration, and $\gamma^{-1} \approx 0.85$ describes the second phase. For $b^{(2)}$ there is no stagnation and ϕ and $\hat{\phi}$ are closer to each other. Note that the speed of decrease of the scaled residual norms for $b^{(2)}$ is slower than for $b^{(1)}$ after the stagnation. This corresponds to the fact that $\gamma^{-1} \approx 0.95$ and $\hat{\gamma}^{-1} \approx 0.97$ for $b^{(2)}$ are both smaller than γ^{-1} for $b^{(1)}$.

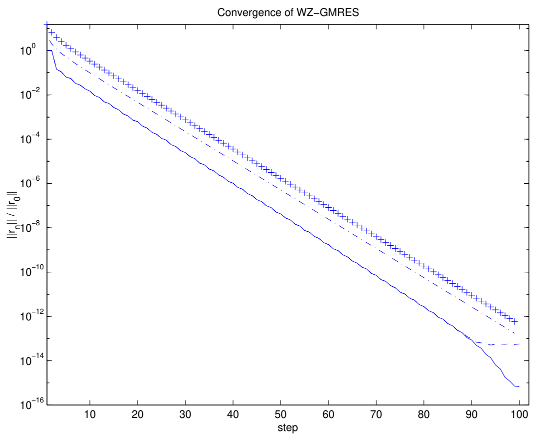

Figure 15: BWM200: GMRES convergence curves and convergence bounds.

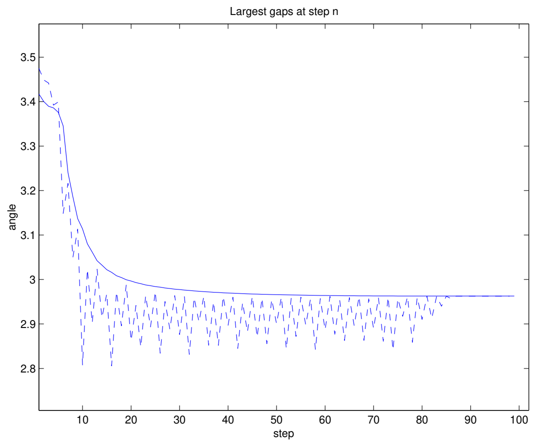

Figure 16: BWM200: The gaps ϕ_n and $\hat{\phi}_n$.

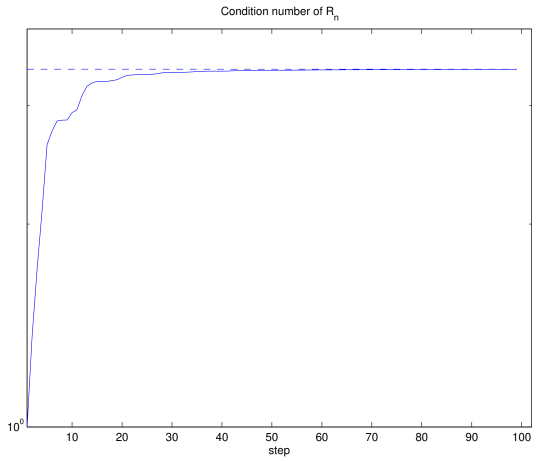

Figure 17: BWM200: $\kappa(R_n)$.

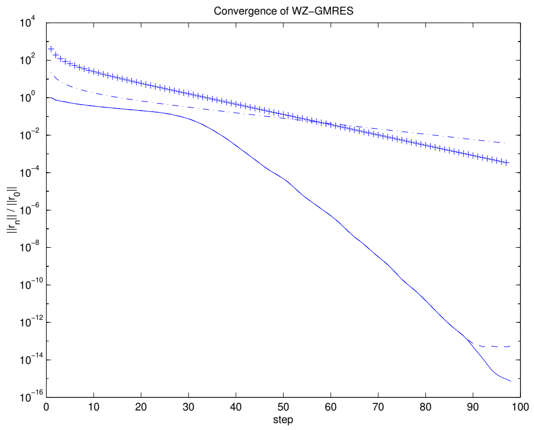

Figure 18: PDE225: GMRES convergence curves and convergence bounds.

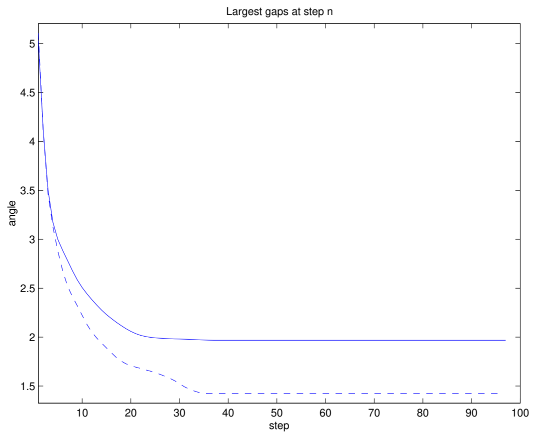

Figure 19: PDE225: The gaps ϕ_n and $\hat{\phi}_n$.

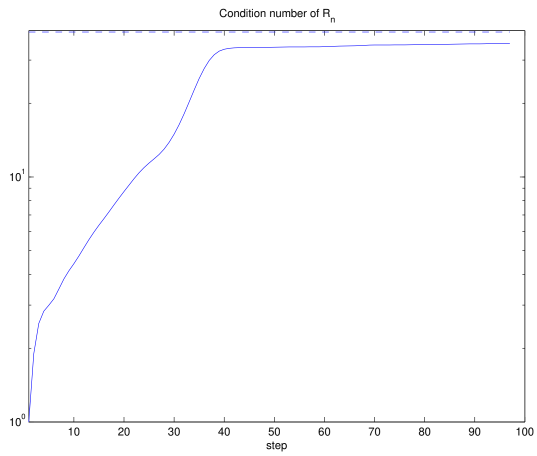

Figure 20: PDE225: $\kappa(R_n)$.

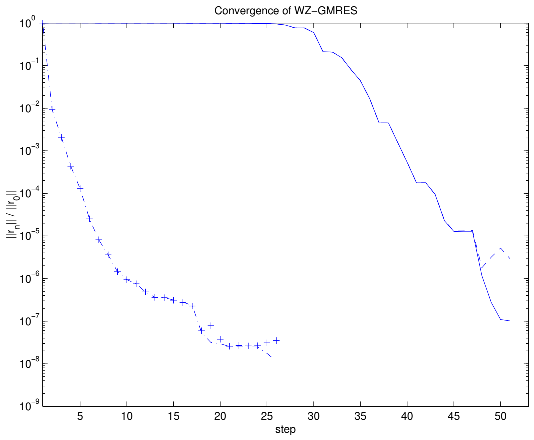

Figure 21: FS1836: GMRES convergence for $b^{(1)}$ and $b^{(2)}$.

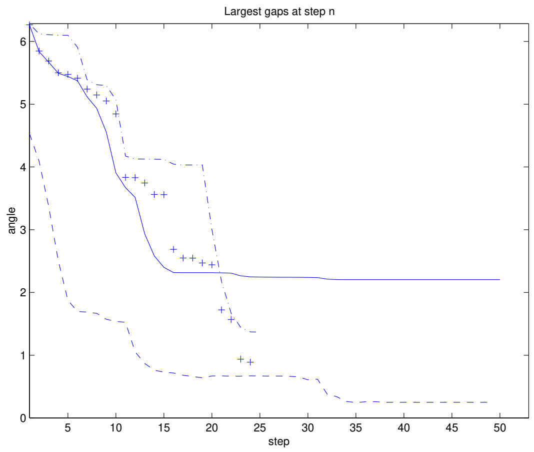

Figure 22: FS1836: The gaps ϕ_n and $\hat{\phi}_n$ for $b^{(1)}$ and $b^{(2)}$.

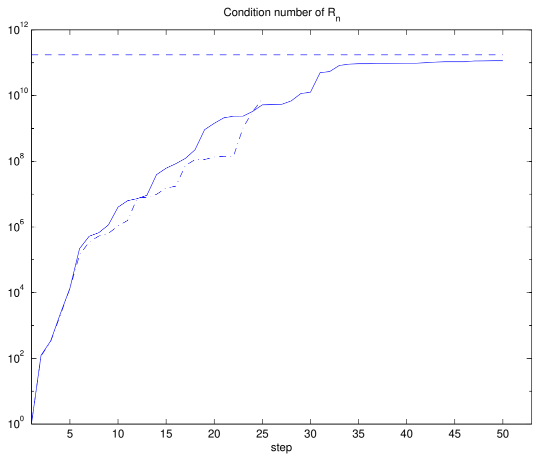

Figure 23: FS1836: $\kappa(R_n)$ for $b^{(1)}$ and $b^{(2)}$.

6 The ABF polynomial preconditioner

In this section we introduce a polynomial preconditioner which is derived from the iteration polynomials of the ABF method. Our preconditioner is characterized by the following main properties:

- The preconditioning polynomials are the iteration polynomials of the ABF method, and thus can be generated with a three-term recurrence.
- The preconditioner is widely applicable, due to the variability of the 'bratwurst' shape sets.

In Section 6.1, we briefly review the basics of (polynomial) preconditioning. Details about the ABF polynomial preconditioner and an algorithm for polynomial preconditioned GMRES are given in Section 6.2. Numerical experiments with this preconditioned method are described in Section 6.3.

6.1 Basics of (polynomial) preconditioning

Preconditioning the linear system (1.1) means to replace it by a modified system,

$$M^{-1}Ax = M^{-1}b \quad \text{or} \quad AM^{-1}\hat{x} = b,\ x = M^{-1}\hat{x},$$

called left or right preconditioned system, respectively. The choice of the preconditioner M^{-1} should make the preconditioned system easier to solve than (1.1). The strategies to achieve this goal are so numerous and diverse that even experts in the field view the choice of a preconditioner as "a combination of art and science" [65, p. 265]. Among others, the choice of the preconditioner depends on the given linear system, the computer architecture and the PIM applied to solve the preconditioned system.

A polynomial preconditioner is nothing but a polynomial in A,

$$M^{-1} = P(A).$$

Since A and $P(A)$ commute, the left and right polynomial preconditioned system have the same coefficient matrix. Attractive polynomial preconditioners are potentially given by any PIM because of the following:

The kth residual polynomial satisfies

$$p_k(z) = 1 - zq_k(z),$$

where $q_k(z)$ is the corresponding kth iteration polynomial of degree at most $k-1$. If the PIM converges well, i.e. $p_k(z) \to 0$ on some compact inclusion set $\Omega \supset \Lambda(A)$, we have

$$zq_k(z) \to 1 \text{ on } \Omega.$$

Therefore, if $q_k(A)$ is chosen as preconditioner, the spectrum of the preconditioned matrix becomes clustered in a neighborhood of the point one. Such clustering typically leads to fast convergence of PIMs.

Hybrid PIMs are particularly attractive for the development of polynomial preconditioners, because explicit formulas for their iteration polynomials are known. A review of several polynomial preconditioners that have been developed this way can be found in [65, Chapter 12.3].

6.2 ABF polynomial preconditioned GMRES

The idea of the ABF polynomial preconditioner is to use

$$M^{-1} = q_k(A),$$

where $q_k(z)$ is the kth iteration polynomial of the ABF method as defined in Section 4.2. Before these polynomials can be constructed, we need an approximation of $\Lambda(A)$ as the basis for the choice of an inclusion set $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$. After the inclusion set parameters are determined, the polynomials can be generated with a three-term recurrence and any PIM might be applied to solve the preconditioned system.

Our main interest here is to apply the ABF polynomial preconditioner to GMRES. We adapt the preconditioned GMRES scheme of [65, Algorithm 12.2] in the following algorithm:

Algorithm 6.1 (ABF polynomial preconditioned GMRES)

Start or restart:
Compute the current residual vector $r := b - Ax$.

Adaptive GMRES step:
Run m_1 *steps of GMRES for solving* $Ad = r$.
Update x *by* $x := x + d$.

Parameters for the ABF polynomial preconditioner:
Compute $\Lambda(H_{m_1})$.
Compute an inclusion set $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ *for* $\Lambda(H_{m_1})$.
Compute $q_k(z)$*, the* k*th iteration polynomial of the ABF method.*

Preconditioned iteration:
Compute the current residual $r = b - Ax$.
Run m_2 *steps of GMRES applied to* $Aq_k(A)d = q_k(A)r$.
Update x *by* $x := x + d$.
Test for convergence, possibly restart.

The matrix H_{m_1} of Arnoldi coefficients is automatically generated during the 'Adaptive GMRES step' of Algorithm 6.1. Hence, the required approximation of $\Lambda(A)$ can be obtained efficiently. The degree of the preconditioning polynomial is typically very small compared to the system dimension N and depends on the performance of the 'Preconditioned iteration' of Algorithm 6.1.

We point out that polynomial preconditioners require a careful implementation. For a vector v, $Aq_k(A)v$ needs to be computed with a sequence of matrix-vector products. The $\mathcal{O}(N^3)$ operation $A * q_k(A)$ usually reduces the efficiency of a preconditioned PIM.

As indicated in Section 6.1, the success of a polynomial preconditioner derived from iteration polynomials of a certain PIM depends on the performance of the PIM itself for the given linear system. In case of the ABF polynomial preconditioner,

$$\Lambda(A) \subset s\Omega_\epsilon$$

should be satisfied by the inclusion set. Making this assumption, we can prove the following result on the clustering of $\Lambda(Aq_k(A))$.

Lemma 6.2 *Suppose that $A \in \mathbf{C}^{N\times N}$ is diagonalizable and that $q_k(z)$ is the kth iteration polynomial of the ABF method corresponding to the inclusion set $s\Omega_\epsilon \in \mathbf{B}(\lambda_m, \phi)$ of $\Lambda(A)$. Then, in the notation of Theorem 4.1,*

$$|1-\lambda| \;\leq\; \frac{V(s\Omega_\epsilon)}{\pi\left(N_\epsilon^k + M_\epsilon^k - (-S_\epsilon)^k\right)},$$

holds for every $\lambda \in \Lambda(Aq_k(A))$.

Proof. If $A = XDX^{-1}$ is the eigendecomposition of A, then $Aq_k(A) = XDq_k(D)X^{-1}$ is the eigendecomposition of $Aq_k(A)$. The assertion now follows by applying Lemma 4.3 to $Aq_k(A) = XDq_k(D)X^{-1}$. □

We see that in case of diagonalizable A and $\Lambda(A) \subset s\Omega_\epsilon$, the ABF polynomial preconditioner is guaranteed to induce a clustering of the spectrum of the preconditioned matrix in a neighborhood of the point one.

6.3 Numerical experiments

In numerical experiments we demonstrate that the use of the ABF polynomial preconditioner can lead to a significant decrease of the number of steps and flops required by GMRES.

For simplicity of notation, $k = 0$ here means that no preconditioner is used. Information on the test matrices is given in Appendix A.

Example 6.3 We study the clustering of $\Lambda(Aq_k(A))$ and the computational effort of preconditioned GMRES for changing degree of the preconditioning polynomial.

Our example linear system is given by A=PDE225 and $b = [1, ..., 1]^T$. For simplicity, we only perform the part 'Preconditioned iteration' of Algorithm 6.1, using a modified Gram-Schmidt implementation of GMRES (cf. e.g. [65, Algorithm 6.9]) with $x_0 = [0, ..., 0]^T$. The polynomial preconditioner $q_k(A)$ is based on the inclusion set $\mathbf{S}_1$ from Example 4.6.

In the following table, we list the number of flops (divided by N^2) and steps required until $\|r_n\|/\|r_0\| < 10^{-8}$, where $r_n = b - Ax_n$ is the computed GMRES residual. The flop counts include the *computation and application* of the preconditioners. These are computed using a sequence of matrix-vector products.

k	flops/N^2	steps
0	48.4	66
4	18.1	29
5	16.3	24
6	14.9	20

The best results are achieved by the preconditioner $M^{-1} = q_6(A)$, which reduces the number of flops *and* steps by roughly 70% compared to unpreconditioned GMRES.

Figure 24 shows the curves of the updated scaled residual norms for the preconditioned GMRES iteration and $k = 0$ (solid), 4 (dashed), 5 (dashdot), 6 (dotted).

The reason for the significant reduction of GMRES steps mainly lies in the clustering of the spectra of the preconditioned matrices. Figure 25 shows $\Lambda(Aq_k(A))$ for $k = 0, 4, 5, 6$. Obviously, even for small values of k, the preconditioned spectra lie in close neighborhoods of the point one.

Example 6.4 We study the influence of the inclusion set on the performance of the preconditioned iteration. We use the linear system given by the shifted GRCAR matrix as in (4.42) and $b = [1, ..., 1]^T$. With the same GMRES implementation, termination criterion and initial guess as in Example 6.3, we perform the part 'Preconditioned iteration' of Algorithm 6.1. We use two polynomial preconditioners $q_4(A)$ of degree 3, which are derived from the inclusion sets

$$2\Omega_\epsilon \in \mathbf{B}(-1, 4\pi/7) \quad \text{for} \quad \epsilon \in \{0.25, 0.22\}.$$

The flop and step counts for these experiments are listed below:

	flops/N^2	steps
$k = 0$	198.8	300
$k = 4,\ \epsilon = 0.25$	43.2	128
$k = 4,\ \epsilon = 0.22$	115.4	210

Figure 26 shows the curves of the updated scaled residual norms for the preconditioned iteration with $k = 0$ (solid), $k = 4 \ / \ \epsilon = 0.25$ (dashed) and $k = 4 \ / \ \epsilon = 0.22$ (dashdot). Using the ABF polynomial preconditioner leads to significant improvements of the performance of GMRES. For $\epsilon = 0.25$, the preconditioned version even achieves a flop reduction of about 80%.

Note that a small difference in ϵ induces a large difference in the behavior of the preconditioned iteration. This is comparable to our observations of the behavior of Phase II of the ABF method (cf. Example 4.7). The differences here are also caused by the non-normality of A, and they correspond to the performance of Phase II for the given linear system: The better Phase II converges, the better is the performance of the preconditioned iteration.

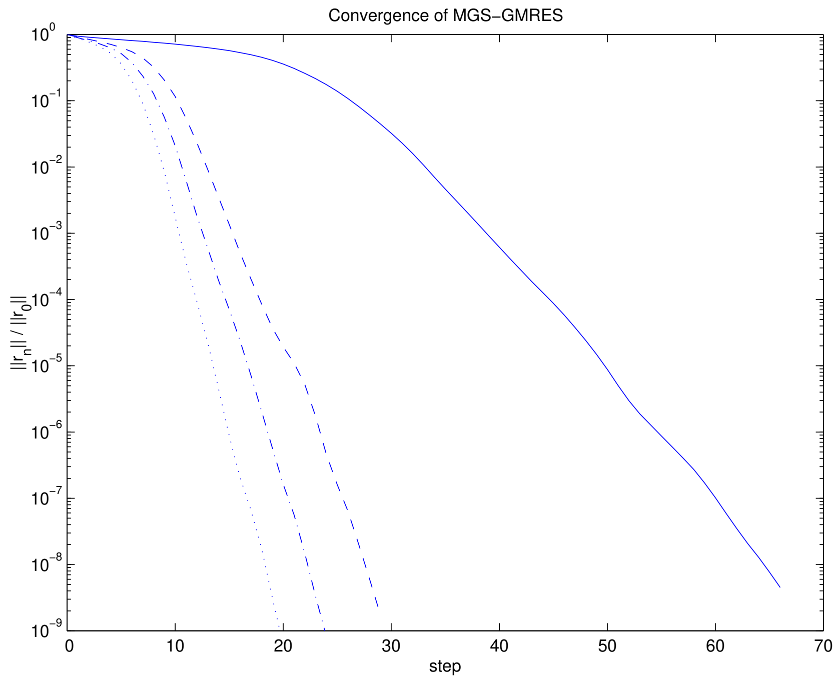

Figure 24: PDE225: Convergence of ABF polynomial preconditioned GMRES.

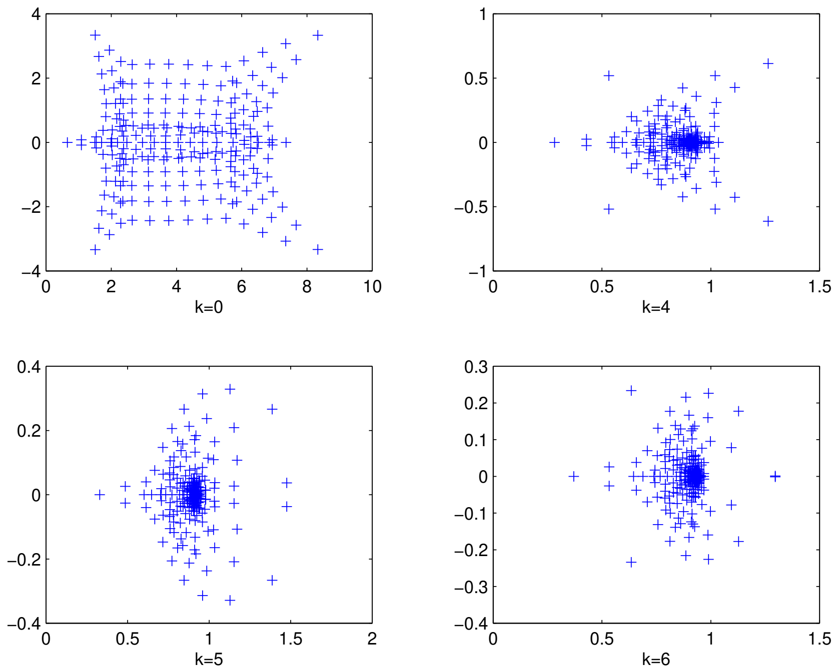

Figure 25: PDE225: Spectra of preconditioned matrices.

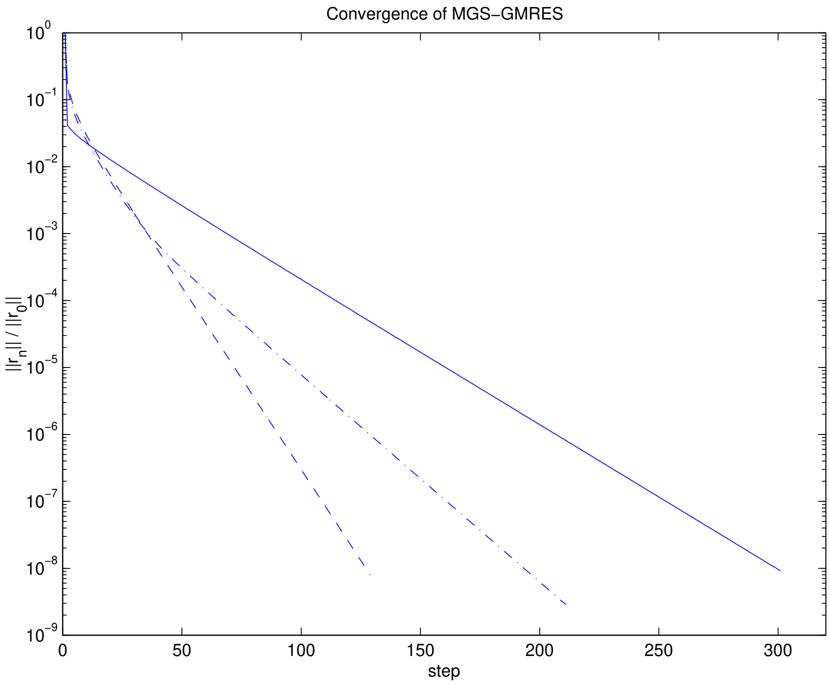

Figure 26: Shifted GRCAR matrix: Convergence of ABF polynomial preconditioned GMRES.

7 Concluding remarks

This thesis contributes to the development of the rapidly growing field of polynomial iterative methods (PIMs) for systems of linear equations. Our main contributions are the construction and analysis of the ABF method, a new hybrid PIM, and new convergence results for GMRES [66], a well-known parameter-free PIM. We furthermore derive a new polynomial preconditioner from the iteration polynomials of the ABF method.

Our results of Section 4, *The hybrid ABF method for non-hermitian linear systems*, demonstrate that it is possible to construct hybrid PIMs featuring non-convex inclusion sets *and* short-term recurrences. An important observation is that the 'bratwurst' shape sets defined in Section 3, *Inclusion sets generated by the conformal 'bratwurst' maps*, are Möbius-equivalent to ellipses. We can therefore show that the Faber polynomials for the 'bratwurst' shape sets are related to the Chebyshev polynomials for sets bounded by ellipses. Hence, the Faber polynomials satisfy a three-term recurrence. From this relation we derive efficient three-term recurrences for the iteration polynomials of our new hybrid method.

The 'bratwurst' shape sets are particular instances of conformal deformations of ellipses by Möbius transformations which are attractive for our purpose. We believe that further hybrid PIMs featuring short-term recurrences can be constructed by considering other conformal (Möbius) deformations of ellipses. In particular, the shape of the inclusion sets for further hybrid PIMs might be 'custom-made' for the needs of specific applications. This approach is certainly an alternative to the Schwarz-Christoffel formula.

Our numerical experiments with the ABF method and the ABF polynomial preconditioner are evidence of the influence of non-normality on the performance of PIMs. Much more research is required until this influence is fully understood. Work in these directions is currently actively pursued (cf. e.g. Trefethen's recent research proposal [75]).

The GMRES algorithm minimizes $\|p_n(A)r_0\|$, and thus the initial guess influences the algorithm's behavior. In Section 5, *Computable convergence bounds for GMRES*, we incorporate this influence in convergence bounds. Our results are conceptually different from standard 'worst-case' results, which are based on the 'ideal GMRES' [39] approximation problem $\min_{p_n \in \pi_n} \|p_n(A)\|$.

Unlike the results obtainable from 'ideal GMRES', our bounds are also applicable to singular linear systems, if a certain condition is satisfied. In numerical experiments we show that our bounds are capable of describing the actual convergence behavior of GMRES for the given linear system and initial guess. We furthermore show that approximations to our bounds are computable from information generated during the run of WZ-GMRES [77], a particular GMRES implementation. We are thus able to develop a heuristic for predicting the performance of the algorithm during its run. This new approach might prove useful for other iterative methods as well.

A Model problems and test matrices

In Appendix A, we present information on the test matrices used in our numerical experiments. All numerical experiments reported in this thesis were performed using the software MATLAB 5.0 [71].

GRCAR. The matrix GRCARN is a nonsymmetric Toeplitz matrix of the form

$$\mathrm{GRCAR}N = \begin{bmatrix} 1 & 1 & 1 & 1 & & & \\ -1 & 1 & 1 & 1 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & -1 & 1 & 1 & 1 & 1 \\ & & & \ddots & \ddots & \ddots & 1 \\ & & & & \ddots & \ddots & 1 \\ & & & & & -1 & 1 \end{bmatrix} \in \mathbf{R}^{N \times N}, \tag{A.1}$$

which can be set up in MATLAB with the command `gallery('grcar',N,3)`. The GRCAR matrices are standard examples in the context of PIMs (see e.g. [28, 53, 57]). They are interesting because of the high sensitivity and special distribution of their eigenvalues, e.g. see Figure 4, in which the pluses represent $\Lambda(\mathrm{GRCAR}250)$. While the GRCAR matrices are well conditioned, their eigenvector matrices are very ill conditioned. Some typical data are shown in the following table:

N	$\kappa(\mathrm{GRCAR}N)$	$nnz(\mathrm{GRCAR}N)$	$\kappa(X_N)$
250	3.62	1243	9.31e+33
500	3.63	2493	2.47e+36

$\kappa(X_N)$ and $nnz(\mathrm{GRCAR}N)$ denote the condition number of the eigenvector matrix and the number of nonzero entries of GRCARN, respectively. The condition numbers were computed using MATLAB's `cond` function.

PDE. The matrix PDEN is an $N \times N$, real and nonsymmetric matrix derived from the finite difference discretization of a two-dimensional elliptic partial differential operator. We downloaded three such matrices (N=255, 900 and 2961) from the Matrix Market web site at

`http://math.nist.gov/MatrixMarket/data/NEP/matpde/matpde.html`.

Matrix statistics:

N	$\kappa(\mathrm{PDE}N)$	$nnz(\mathrm{PDE}N)$	$\kappa(X_N)$
225	39.06	1065	2.80e+06
900	152.56	4380	2.88e+15

The notation is the same as in the corresponding table for the GRCAR matrices. $\kappa(X_{900})$ was estimated using MATLAB's `condest` function. For an example spectrum of a PDEN matrix, see Figure 5, in which the pluses represent $\Lambda(\mathrm{PDE}225)$. More information about the PDEN matrices can be found in the user's guide for the NEP matrix collection [3], which is available from the Matrix Market web site.

Other matrices. The other matrices used in numerical experiments are FS1836 from the Harwell-Boeing matrix collection [13] and BWM200[1] from the NEP matrix collection. Both are real, nonsymmetric and were downloaded from the Matrix Market web site.

Matrix statistics:

A	N	$\kappa(A)$	$nnz(A)$	$\kappa(X_A)$
FS1836	183	1.74e+11	1000	5.62e+17
BWM200	200	3.39	796	2.23

B MATLAB code for Phase II of the ABF method

In Appendix B we list the MATLAB program of Algorithm 4.4 which was used in the experiments reported in Section 4. The program terminates either if the number of iterations of the main loop exceeds `MAXIT`, or if the computed residual $r_n = b - Ax_n$ satisfies $\|r_n\|/\|r_0\|$ <`tol`. The parameters `MAXIT` and `tol` have to be provided by the user. The vectors x_n and r_n are only computed explicitly, if the scaled residual norm computed using (4.33) is less than `tol`.

```
function [uprnorm,xn,flag] = phasetwo(A,b,x0,lm,phi,epsi,s,MAXIT,tol)
%
% Phase II of the ABF method.
%
% Input:  Linear system parameters:  A, b, x0
% Inclusion set parameters:
% lm - (guess for) center of largest gap in Lambda(A)
% phi - (guess for) largest gap in Lambda(A)
% epsi - parameter for shape of the inclusion set
% s - scaling of the inclusion set (positive)
% Other parameters:
% MAXIT - maximal number of iterations of the main loop
% tol - required accuracy of the solution
%
% Output:
% uprnorm - vector with updated residual norms
% xn - the final approximation
% flag - equals 1 if ||b-A*xn|| / ||r0|| < tol, 0 otherwise
%

flag = 0; pot = 1;
```

[1] On July 1, 1998, several matrix files in Matrix Market, including BWM200, were found to contain corrupt data in the sense that the matrices do not correspond to their underlying partial differential equations. The corrupt files were replaced by updated versions. The file of BWM200 used in our experiments was downloaded before July 1, 1998, and thus does not represent the reaction-diffusion model it should come from. However, we are not interested in properties of this model and thus the outcome of our experiments is unaffected by the change of matrix files in Matrix Market.

```
% Initialize
 P = tan(phi/4)+1/cos(phi/4); et = 1+epsi;
 N = (P/et+et/P)/2; M = (et^2-1)/(2*tan(phi/4)*et);
 S = (N*M-1)/(N-M); tinv = (N-M)/s;
 c1 = lm*(N+M); c2 = lm^2*N*M; c3 = -lm*S;
 c4 = lm^2*(S+N)*(S+M);
 rold = 1; rnew = c1 - c3;
 r0 = b-A*x0; nr0 = norm(r0); uprnorm(1,1) = 1;
 Fold = 2*r0; Fnew = tinv*(A*r0) + c1*r0;
 qold = zeros(length(A),1); qnew = -(tinv/rnew)*r0;

% Step 1
 xn = x0 + qnew;

% Main loop
for n = 2:MAXIT
 uprnorm(n,1) = norm(Fnew-c4*pot*r0)/(abs(rnew)*nr0);
 if uprnorm(n,1) < tol
  % Compute xn and rn only if uprnorm is small enough
  xn = x0 + qnew;
  if norm(b-A*xn)/nr0 < tol
   flag = 1;
  end
 end
 if flag == 1
  break
 end
 zwi = rnew;
 rnew = c1*rnew - c2*rold - c4*pot;
 rvold = rold;
 rold = zwi;
 vn = tinv*(Fnew - c3*Fold);
 zwi = Fnew;
 Fnew = A*vn + c1*Fnew - c2*Fold;
 Fold = zwi;
 zwi = qnew;
 qnew = (c1*rold*qnew - c2*rvold*qold - vn)/rnew;
 qold = zwi;
 pot = pot*c3;
end

% Final step
if flag ~= 1
  xn = x0 + qnew;
end
```

References

[1] M. Arioli, V. Pták, and Z. Strakoš, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.

[2] W. Arnoldi, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. of Appl. Math., 9 (1951), pp. 17–29.

[3] Z. Bai, D. Day, J. Demmel, and J. Dongarra, *A test matrix collection for non-hermitian eigenvalue problems, Release 1.0*, tech. report, 1996.

[4] T. Barth and T. Manteuffel, *Variable metric conjugate gradient methods*, in Advances in Numerical Methods for Large Sparse Sets of Linear Equations, Number 10, Matrix Analysis and Parallel Computing, PCG 94, M. Natori and T. Nodera, eds., 1994, pp. 165–188.

[5] C. Bischof, *Incremental condition estimation*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 312–322.

[6] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.

[7] S. L. Campbell, I. C. F. Ipsen, C. T. Kelley, and C. D. Meyer, *GMRES and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.

[8] J. P. Coleman and N. J. Myers, *The Faber polynomials for annular sectors*, Math. of Comput., 64 (1995), pp. 181–203.

[9] J. B. Conway, *A course in functional analysis*, Springer-Verlag, New York, 2nd ed., 1990.

[10] J. Curtiss, *Faber polynomials and the Faber series*, American Math. Monthly, 78 (1971), pp. 577–596.

[11] J. Demmel, *The condition number of equivalence transformations that block diagonalize matrix pencils*, SIAM J. Numer. Anal., 20 (1983), pp. 599–610.

[12] J. Drkošová, A. Greenbaum, M. Rozložník, and Z. Strakoš, *Numerical stability of GMRES*, BIT, 35 (1995), pp. 309–330.

[13] I. S. Duff, R. G. Grimes, and J. G. Lewis, *User's guide for the Harwell-Boeing sparse matrix collection (Release I)*, Tech. Report TR/PA/92/86, CERFACS, 1992.

[14] M. Eiermann, *On semiiterative methods generated by Faber polynomials*, Numer. Math., 56 (1989), pp. 139–156.

[15] ——, *Fields of values and iterative methods*, Lin. Alg. Appl., 180 (1993), pp. 167–198.

[16] ——, *Field of values and iterative methods*, in preparation, (1996).

[17] M. Eiermann and W. Niethammer, *On the construction of semi-iterative methods*, SIAM J. Numer. Anal., 20 (1983), pp. 1153–1160.

[18] H. C. Elman, Y. Saad, and P. E. Saylor, *The hybrid Chebyshev Krylov subspace algorithm for solving nonsymmetric systems of linear equations*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 840–855.

[19] L. Elsner and C. He, *Perturbation and interlace theorems for the unitary eigenvalue problem*, Lin. Alg. Appl., 188-189 (1993), pp. 207–229.

[20] O. Ernst, *Residual-minimizing Krylov subspace methods for stabilized discretizations of convection-diffusion equations*, submitted to SIAM J. Matrix Anal. Appl., (1998).

[21] G. Faber, *Über polynomische Entwickelungen*, Math. Ann., 57 (1903), pp. 398–408.

[22] V. Faber, W. Joubert, E. Knill, and T. Manteuffel, *Minimal residual method stronger than polynomial preconditioning*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 707–729.

[23] V. Faber and T. Manteuffel, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.

[24] B. Fischer and R. Freund, *On the constrained Chebyshev approximation on ellipses*, J. Approx. Theory, 62 (1990), pp. 297–315.

[25] ——, *Chebyshev polynomials are not always optimal*, J. Approx. Theory, 65 (1991), pp. 261–272.

[26] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Proceedings of the 6th Biennial Dundee Conference on Numerical Analysis, G. Watson, ed., vol. 506 of Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1976, pp. 73–89.

[27] R. Freund and S. Ruscheweyh, *On a class of Chebyshev approximation problems which arise in connection with a conjugate gradient type method*, Numer. Math., 48 (1986), pp. 525–542.

[28] R. W. Freund, *Quasi-kernel polynomials and their use in non-Hermitian matrix iterations*, J. Comput. Appl. Math., 43 (1992), pp. 135–158.

[29] R. W. Freund, G. H. Golub, and N. M. Nachtigal, *Iterative solution of linear systems*, Acta Numerica, 1 (1992), pp. 57–100.

[30] R. W. Freund and N. M. Nachtigal, *QMR: A quasi-minimal residual method for non-hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.

[31] G. Golub and H. van der Vorst, *Closer to the solution: Iterative linear solvers*, in The State of the Art in Numerical Analysis, I. Duff and G.A.Watson, eds., Clarendon Press, Oxford, 1997, pp. 63–92.

[32] G. H. Golub and C. F. van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore, 3rd ed., 1996.

[33] A. Greenbaum, *Krylov subspace approximations to the solution of a linear system*, in Linear and nonlinear conjugate gradient-related methods (Seattle, WA, 1995), SIAM, Philadelphia, PA, 1996, pp. 83–91.

[34] ——, *Iterative methods for solving linear systems*, SIAM, Philadelphia, 1997.

[35] A. Greenbaum and L. Gurvits, *MAX-MIN properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 348–358.

[36] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 465–469.

[37] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, G. H. Golub, A. Greenbaum, and M. Luskin, eds., vol. 60 of The IMA volumes in mathematics and its applications, Springer-Verlag, New York, 1994, pp. 95–118.

[38] A. Greenbaum and L. N. Trefethen, *Do the pseudospectra of a matrix determine its behavior?*, Tech. Report TR 93-1371, Cornell CS Dept., 1993.

[39] ——, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[40] M. He, *The Faber polynomials for circular arcs*, in Mathematics of Computation 1943–1993: a half-century of computational mathematics (Vancouver, BC, 1993), vol. 48 of Proc. Sympos. Appl. Math., Amer. Math. Soc., Providence, RI, 1994, pp. 301–304.

[41] P. Henrici, *Applied and computational complex analysis*, vol. 3, Wiley, New York, 1986.

[42] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436 (1953).

[43] N. J. Higham, *Handbook of writing for the mathematical sciences*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1993.

[44] E. Hille, *Analytic function theory. Vol. II*, Ginn and Co., Boston, Mass.-New York-Toronto, Ont., 1962.

[45] I. Ipsen, *A different approach to bounding the minimal residual norm in Krylov methods*, submitted to SIAM J. Numer. Anal., (1998).

[46] ——, *A note on the field of values of non-normal matrices*, submitted to Linear Alg. Appl., (1998).

[47] W. Joubert, *A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439.

[48] T. Koch and J. Liesen, *The conformal bratwurst map and polynomial iterative methods*, Preprint 98-078, Sonderforschungsbereich 343, Universität Bielefeld, 1998.

[49] T. Kövari and C. Pommerenke, *On Faber polynomials and Faber expansions*, Math. Z., 99 (1967), pp. 193–206.

[50] J. Liesen, *Computable convergence bounds for GMRES*, Preprint 98-043, Sonderforschungsbereich 343, Universität Bielefeld, 1998.

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

[52] ——, *Adaptive procedure for estimating parameters for the nonsymmetric Tchebychev iteration*, Numer. Math., 31 (1978/79), pp. 183–208.

[53] T. A. Manteuffel and G. Starke, *On hybrid iterative methods for nonsymmetric systems of linear equations*, Numer. Math., 73 (1996), pp. 489–506.

[54] A. I. Markushevich, *Theory of functions of a complex variable. Vol. I*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1965.

[55] J. Modersitzki and G. Opfer, *Faber versus minimal polynomials on annular sectors*, Preprint Reihe A 88, Hamburger Beitr. zur Angew. Math., 1994.

[56] N. M. Nachtigal, S. Reddy, and L. N. Trefethen, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 778–795.

[57] N. M. Nachtigal, L. Reichel, and L. N. Trefethen, *A hybrid GMRES algorithm for nonsymmetric linear systems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 796–825.

[58] V. Paatero, *Über die konforme Abbildung von Gebieten deren Ränder von beschränkter Drehung sind*, Ann. Acad. Sci. Fenn. Ser. A., 33 (1931), pp. 1–77.

[59] C. Paige and M. Saunders, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.

[60] L. Reichel and L. N. Trefethen, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices*, Linear Algebra Appl., 162/164 (1992), pp. 153–185.

[61] J. K. Reid, *On the method of conjugate gradients for the solution of large sparse systems of linear equations*, in Large sparse sets of linear equations (Proc. Conf., St. Catherine's Coll., Oxford, 1970), Academic Press, London, 1971, pp. 231–254.

[62] M. Rozložník, *Numerical stability of the GMRES method*, PhD thesis, Institute of Computer Science, Academy of Sciences of the Czech Republic, 1997.

[63] M. Rozložník and Z. Strakoš, *Variants of the residual minimizing Krylov space methods*, in Proceedings of the XI-th Summer School on Software and Algorithms of Numerical Mathematics, I. Marek, ed., 1995, pp. 208–225.

[64] Y. Saad, *Variations on Arnoldi's method for computing eigenelements of large unsymmetric matrices*, Lin. Alg. Appl., 34 (1980), pp. 269–295.

[65] ——, *Iterative methods for solving sparse linear systems*, PWS Publishing Company, Boston, 1996.

[66] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.

[67] V. Smirnov and N. Lebedev, *Functions of a complex variable - Constructive theory*, The M.I.T. Press, Cambridge, MA, 1968.

[68] G. Starke, *Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems*, Numer. Math., 78 (1997), pp. 103–117.

[69] G. Starke and R. S. Varga, *A hybrid Arnoldi-Faber method for nonsymmetric systems of linear equations*, Numer. Math., 64 (1993), pp. 213–240.

[70] Z. Strakoš, *Convergence and numerical behavior of the Krylov space methods*, in Proceedings of the NATO ASI Institute "Algorithms for Large Sparse Linear Algebraic Systems: The State of the Art and Applications in Science and Engineering" (invited lectures), Kluiver Academic, 1998, pp. 175–196.

[71] The MathWorks, *Using MATLAB (December 1996)*, The MathWorks Inc., Natick, 1996.

[72] K.-C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.

[73] K.-C. Toh and L. N. Trefethen, *Calculation of pseudospectra by the Arnoldi iteration*, SIAM J. Sci. Comput., 17 (1996), pp. 1–15.

[74] L. N. Trefethen, *Approximation theory and numerical linear algebra*, in Algorithms for approximation II, J. Mason and M. Cox, eds., Chapman and Hall, London, 1990, pp. 336–360.

[75] ——, *The Oxford eigenvalue project - Case for support*, tech. report, Oxford University Computing Laboratory, 1998.

[76] H. F. Walker, *Implementation of the GMRES method using Householder transformations*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 152–163.

[77] H. F. Walker and L. Zhou, *A simpler GMRES*, Numer. Lin. Alg. Appl., 1 (1994), pp. 571–581.

[78] J. Walsh, *Interpolation and approximation by rational functions in the complex domain*, American Mathematical Society, Providence, RI, 5th ed., 1969.