BIT
0000, Vol. 00, No. 0, pp. 000–000

0000-0000/00/0000-0000 $00.00

The worst-case GMRES for normal matrices

Jörg Liesen[1†] Petr Tichý[1‡]

September 9, 2003

[1] *Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany*
email: liesen@math.tu-berlin.de, tichy@math.tu-berlin.de.

Abstract.

We study the convergence of GMRES [14] for linear algebraic systems with normal matrices. In particular, we explore the standard bound based on a min-max approximation problem on the discrete set of the matrix eigenvalues. This bound is sharp, i.e. it is attainable by the GMRES residual norm [4, 8]. The question is how to evaluate or estimate the standard bound, and if it is possible to characterize the GMRES-related quantities for which this bound is attained (worst-case GMRES). In this paper we completely characterize the worst-case GMRES-related quantities in the next-to-last iteration step and evaluate the standard bound in terms of explicit polynomials involving the matrix eigenvalues. For a general iteration step, we develop a computable lower and upper bound on the standard bound. Our bounds allow to study the worst-case GMRES residual norm in dependence of the eigenvalue distribution. For hermitian matrices the lower bound is equal to the worst-case residual norm. In addition, numerical experiments show that the lower bound is generally very tight, and support our conjecture that it is to within a factor of $4/\pi$ of the actual worst-case residual norm. Since the worst-case residual norm in each step is to within a factor of the square root of the matrix size to what is considered an "average" residual norm, our results are of relevance beyond the worst case.

Key words: GMRES, evaluation of convergence, ideal GMRES, normal matrices, min-max problem

AMS subject classification: 15A06, 15A09, 15A18, 65F10, 65F15, 65F20

1 Introduction

Convergence analysis of GMRES [14] has been an active area of research since the algorithm's introduction, and numerous papers have been devoted to this subject, see e.g. [3, Chapter 3] and [10, Section 5.2] for surveys of results. When the system matrix is normal, the earliest upper bound on the GMRES residual

†The work of this author was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft.

‡The work of this author was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft and by the Grant Agency of Academy of Sciences of the Czech Republic under grant No. IAB1030306.

norms (henceforth called the "standard bound") represents a certain min-max approximation problem on the set of the matrix eigenvalues [14, Proposition 4]. Being independent of the initial residual, the standard bound is in fact a bound on the "worst-case" GMRES residual norms for the given system matrix. For normal matrices the standard bound has been shown to be sharp in the sense that for each GMRES iteration step there exists an initial residual (depending on the matrix and the iteration step) for which the bound is attained [4, 8]. In addition, for normal matrices the worst-case GMRES and the "average" GMRES behavior agree to within a factor of $n^{1/2}$ ($n =$ matrix size). By average behavior we here mean that GMRES is started with an initial residual having components in the matrix eigenvectors of approximately equal size (see Section 5 for details).

The sharpness of the standard bound and its closeness to the average case sometimes lead to the impression that the GMRES convergence behavior for normal matrices is fully understood. However, two major problems still remain open. First, the solution of the min-max approximation is unknown (except for special cases), and its known estimates based on only a few properties of the matrix (such as the condition number) are often misleading. Second, in many practical applications the initial residual is not "average", and a systematic study of the consequences for the GMRES convergence needs yet to be performed.

This paper is devoted to the first of the two problems, as its solution appears to be a prerequisite for studying the second. To this end it is of great interest to characterize the min-max approximation problem in terms of easily comprehensible expressions involving the matrix eigenvalues as well as to determine the initial residuals for which the standard bound is attained. Several results in this direction have been previously obtained in the literature. For (real) symmetric positive definite matrices, the initial residuals leading to the worst-case residual norm are completely characterized in [2, Section 2]. The analysis in [2] is based on classical results of approximation theory. In particular, in case of a symmetric positive definite matrix, the polynomial that solves the approximation problem on the matrix eigenvalues (i.e. the one for which the standard bound is attained) is the well-known min-max polynomial on a discrete set of real points (here the matrix eigenvalues). The result of [2] is derived in the context of the conjugate gradient method and can be applied in the GMRES context and to all complex hermitian matrices. A special case of this result (which in particular also assumes that the eigenvalues are real) is proved in [17] by solving a constraint optimization problem using Lagrange multipliers. The related paper [18] gives necessary and sufficient conditions on the eigenvalues of normal matrices so that there exists an initial residual for which GMRES stagnates throughout the iteration (called "complete stagnation" of GMRES). For any normal matrix satisfying these conditions the authors give formulas based on the matrix eigenvalues for all initial residuals that lead to complete stagnation [18, Theorem 3.1]. The complete stagnation obviously represents a special case of worst-case GMRES convergence behavior.

General bounds on the GMRES residual norms for normal matrices that depend on the matrix eigenvalues and the initial residual are derived in [7]. The main tool in this analysis is a factorization of the Krylov matrix. Using a similar starting point as in [7] we characterize the quantities in the next-to-last GMRES iteration step for normal matrices ($(n-1)$st step in case of an n by n matrix having n distinct eigenvalues) in terms of the initial residual and explicit polynomials involving the matrix eigenvalues. We give numerical illustrations of our analytic formulas that show how GMRES behaves for different eigenvalue distributions. Based on these results we completely characterize the worst-case GMRES quantities in the next-to-last iteration step. Then we analyze the worst-case GMRES residual norm in a general iteration step and develop a lower bound on this quantity. In case of hermitian matrices our results are the same as in [2, Section 2], but with a different proof. For the general (normal) case our results complement the existing literature. We prove that our lower bound is to within a factor of (at most) the order n to the actual worst-case residual norm. Furthermore, we conjecture that this bound is much more tight (namely to within a constant factor), and give supporting numerical evidence.

The paper is organized as follows. In Section 2 we develop the basic tools needed for our general analysis in Section 3. Numerical examples studying the closeness of the lower bound to the standard bound are given in Section 4, and a concluding discussion in Section 5 closes the paper.

Throughout the paper we assume exact arithmetic.

2 Basic concepts

In this section we define and develop the basic tools needed for our analysis. Let a linear system

$$Ax = b\,, \tag{2.1}$$

with a *nonsingular and normal* matrix $A \in \mathbb{C}^{n\times n}$ and $b \in \mathbb{C}^n$ be given. Furthermore, let $A = Q\Lambda Q^H$ be the eigendecomposition of A, where $Q^H Q = I$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and let $L = \{\lambda_1, \ldots, \lambda_n\}$ denote the set of all eigenvalues of A. To avoid unnecessary technical complications we will assume throughout this paper that *all eigenvalues of A are distinct.*

Suppose that we solve (2.1) with GMRES [14]. Starting from an initial guess x_0, this method computes the initial residual $r_0 = b - Ax_0$ and a sequence of iterates $x_1, x_2, \ldots$, so that the ith residual $r_i \equiv b - Ax_i$ satisfies

$$\|r_i\| = \|p_i(A)\, r_0\| = \min_{p\in\pi_i} \|p(A)\, r_0\|\,, \tag{2.2}$$

where π_i denotes the set of polynomials of degree at most i and with value one at the origin, and $\|\cdot\|$ denotes the 2-norm. We parameterize the initial residual r_0 by

$$r_0 = Q\,[\varrho_1, \ldots, \varrho_n]^T\,, \tag{2.3}$$

so that

$$r_i = p_i(A)\, r_0 = Q\,[p_i(\lambda_1)\varrho_1, \ldots, p_i(\lambda_n)\varrho_n]^T\,, \tag{2.4}$$

and (2.2) can be written in the form

$$\|r_i\| \;=\; \min_{p\in\pi_i} \left(\sum_{j=1}^{n} |p(\lambda_j)\,\varrho_j|^2 \right)^{1/2} . \tag{2.5}$$

It is well-known that for each GMRES iteration step i and each initial residual r_0 with at least $i+1$ nonzero coordinates ϱ_j, there exists a *unique* polynomial $p_i \in \pi_i$ that solves (2.5). This $p_i(\lambda)$ is called the ith GMRES polynomial.

Similar to [7, 17, 18], we start with a factorization of the Krylov matrix,

$$K_{i+1} \;\equiv\; [r_0, Ar_0, \dots, A^i r_0] \tag{2.6}$$

for some i, $0 \le i \le n-1$. We denote $D \equiv \mathrm{diag}(\varrho_1, \dots, \varrho_n)$, and

$$V_{i+1} \;\equiv\; \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^i \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^i \end{bmatrix}. \tag{2.7}$$

Then $K_{i+1} = QDV_{i+1}$, and the Moore-Penrose generalized inverse of K_{i+1} is given by $K_{i+1}^{+} = (DV_{i+1})^{+}Q^H$. If $\mathrm{rank}(D) \ge i+1$, then K_{i+1} has full column rank, and GMRES does not terminate before the step $i+1$. In this case, as shown in [7, Theorem 2.1], see also [11, Theorem 2.1], the ith GMRES residual satisfies

$$\begin{aligned} r_i \;&=\; \|r_i\|^2\,(K_{i+1}^{+})^H\, e_1 \\ &=\; \|r_i\|^2\, Q\, \left[(DV_{i+1})^{+}\right]^H\, e_1\,, \end{aligned} \tag{2.8}$$

where $e_1 = [1, 0, \dots, 0]^T$. Comparing (2.4) and (2.8) shows that

$$p_i(\lambda_j)\,\varrho_j \;=\; \|r_i\|^2\, \left[(DV_{i+1})^{+}\right]^H_{j1}\,, \qquad j \;=\; 1, \dots, n\,, \tag{2.9}$$

where $\left[(DV_{i+1})^{+}\right]^H_{j1}$ denotes the jth entry in the first column of $\left[(DV_{i+1})^{+}\right]^H$. Note that (2.9) gives the complete correspondence between the ith GMRES polynomial, the ith GMRES residual norm, the coordinates of r_0 in the eigenvectors of A, and the eigenvalues of A. To understand fully the behavior of GMRES for normal matrices it would be desirable to have a general formula for the entries in the first column of $\left[(DV_{i+1})^{+}\right]^H$. However, such a formula is for a general value of i unknown. In the following subsection we will study the special case $i = n-1$, in which (2.8)–(2.9) can be significantly simplified.

2.1 The $(n-1)$st GMRES step

Without loss of generality we restrict our analysis in this subsection to vectors r_0 with nonzero coordinates ϱ_j, $j = 1, \dots, n$. In case $d \ge 1$ coordinates ϱ_j are zero, the corresponding eigencomponents do not play any role for GMRES, and hence the formulas for $i = n-1$ derived below will hold for $i = n-d-1$. Since

$\varrho_j \neq 0$ for all j, GMRES terminates, i.e. computes the solution x, exactly in step n, and its residual norms satisfy

$$\|r_0\| \;\geq\; \|r_1\| \;\geq\; \ldots \;\geq\; \|r_{n-1}\| \;>\; \|r_n\| \;=\; 0\,. \tag{2.10}$$

In the step $i = n-1$, the Vandermonde matrix V_n is square and invertible (all eigenvalues are distinct). Then $[(DV_n)^+]^H = D^{-H}V_n^{-H}$, and (2.8) is equivalent to

$$r_{n-1} \;=\; \|r_{n-1}\|^2\, QD^{-H}V_n^{-H}e_1\,. \tag{2.11}$$

Formulas for the entries of an inverse Vandermonde matrix are well known, see e.g. [6, Chapter 21.1.]. In general, the jth entry in the mth column of the matrix V_n^{-T} is the coefficient of the jth Lagrange polynomial,

$$l_j(\lambda) \;\equiv\; \prod_{\substack{k=1\\k\neq j}}^{n} \frac{\lambda_k - \lambda}{\lambda_k - \lambda_j}\,, \tag{2.12}$$

corresponding to λ^{m-1}, $m = 1, \ldots, n$. Hence the first column of V_n^{-H} is given by the complex conjugates of the constant terms of the $l_j(\lambda)$, i.e.

$$V_n^{-H}\,e_1 \;=\; [l_1(0), \ldots, l_n(0)]^H \;=\; \left[\prod_{\substack{k=1\\k\neq 1}}^{n} \frac{\lambda_k}{\lambda_k - \lambda_1}, \ldots, \prod_{\substack{k=1\\k\neq n}}^{n} \frac{\lambda_k}{\lambda_k - \lambda_n}\right]^H\,. \tag{2.13}$$

The following theorem explains how the $(n-1)$st GMRES residual and iteration polynomial depend on the eigenvalue distribution of A (represented by the values $l_j(0)$) and how on the initial residual r_0 (represented by the coordinates ϱ_j).

THEOREM 2.1. *Suppose that GMRES is applied to the system* (2.1) *with the normal matrix $A \in \mathbb{C}^{n\times n}$ having n distinct eigenvalues, and that r_0 is parameterized by* (2.3) *with $\varrho_j \neq 0$ for all j. Then the norm of $(n-1)$st GMRES residual r_{n-1} satisfies*

$$\|r_{n-1}\| \;=\; \left(\sum_{j=1}^{n} \left|\frac{l_j(0)}{\varrho_j}\right|^2\right)^{-1/2}\,, \tag{2.14}$$

and the $(n-1)$st GMRES polynomial $p_{n-1}(\lambda)$ has the form

$$p_{n-1}(\lambda) \;=\; \|r_{n-1}\|^2 \sum_{j=1}^{n} \frac{\overline{l_j(0)}}{|\varrho_j|^2}\, l_j(\lambda). \tag{2.15}$$

PROOF. Inserting (2.13) into (2.11) yields

$$r_{n-1} \;=\; \|r_{n-1}\|^2\, Q\, \left[l_1(0)\varrho_1^{-1}, \ldots, l_n(0)\varrho_n^{-1}\right]^H\,, \tag{2.16}$$

from which (2.14) follows immediately by taking norms. Next, using the property $l_j(\lambda_k) = \delta_{jk}$, the polynomial $p_{n-1}(\lambda)$ can be written as a linear combination of the Lagrange polynomials,

$$p_{n-1}(\lambda) \; = \; \sum_{j=1}^{n} p_{n-1}(\lambda_j)\, l_j(\lambda)\,. \tag{2.17}$$

Equating (2.4) for $i = n-1$ with (2.16) shows that

$$p_{n-1}(\lambda_j) \; = \; \|r_{n-1}\|^2 \, \frac{\overline{l_j(0)}}{|\varrho_j|^2}\,, \qquad j = 1, \ldots, n\,, \tag{2.18}$$

which, inserted into (2.17), shows (2.15). $\square$

Theorem 2.1 gives formulas for the $(n-1)$st GMRES residual and polynomial in terms of the eigenvalues of A and the coordinates of r_0 in the eigenvectors of A. The influences of both quantities are well separated in (2.14) and (2.15), so that these formulas answer all questions about the $(n-1)$st step of GMRES applied to normal matrices.

Note that the relation (2.14) implies the upper bound

$$\|r_{n-1}\| \; \leq \; \min_{1 \leq j \leq n} \left| \frac{\varrho_j}{l_j(0)} \right| \,. \tag{2.19}$$

The same upper bound follows from [7, Theorem 4.1] with $i = n-1$.

Example 2.1. For numerical illustration we compute the values $|l_j(0)|$ for four different real eigenvalue distributions. Each dot in Figure 2.1 represents a data point $(\lambda_j, |l_j(0)|)$.

For the top left figure we use uniformly distributed eigenvalues in the interval $[1/20, 1]$, i.e. $\lambda_j = j/20$, for $j = 1, \ldots, 20$. We see that $|l_{10}(0)| \approx 10^5$ is the largest of the values $|l_j(0)|$. Then (2.19) implies that for any normal matrix having such eigenvalues, the GMRES residual norm in the next-to-last step will be of order 10^{-5} or smaller (note that $0 < |\varrho_j| < 1$ by assumption).

For the top right figure we use the eigenvalues of the 20 by 20 prolate matrix generated by the MATLAB command `A=gallery('prolate',20)`. Prolate matrices arise in signal processing. They are symmetric, extremely ill conditioned (here: $\lambda_1 \approx 1.76 * 10^{-14}$, $\lambda_{20} = 1 - \lambda_1$, condition number $\approx 5.69 * 10^{13}$), and their eigenvalues form two clusters that are symmetric about a certain point (here: symmetric about 0.5); see [16] for more information. In our example the cluster close to zero causes severe trouble for GMRES. None of the values $|l_j(0)|$ is larger than one, which typically (i.e., unless a very peculiar distribution of the coefficients ϱ_j is constructed) will lead to almost complete stagnation until the very last step, cf. (2.14). This represents a counterexample for the frequent assertion that in case of k (here: $k = 2$) eigenvalue clusters GMRES will essentially need only k steps for a significant reduction of the residual norm. In fact, the location of the clusters relative to the origin and relative to each other is

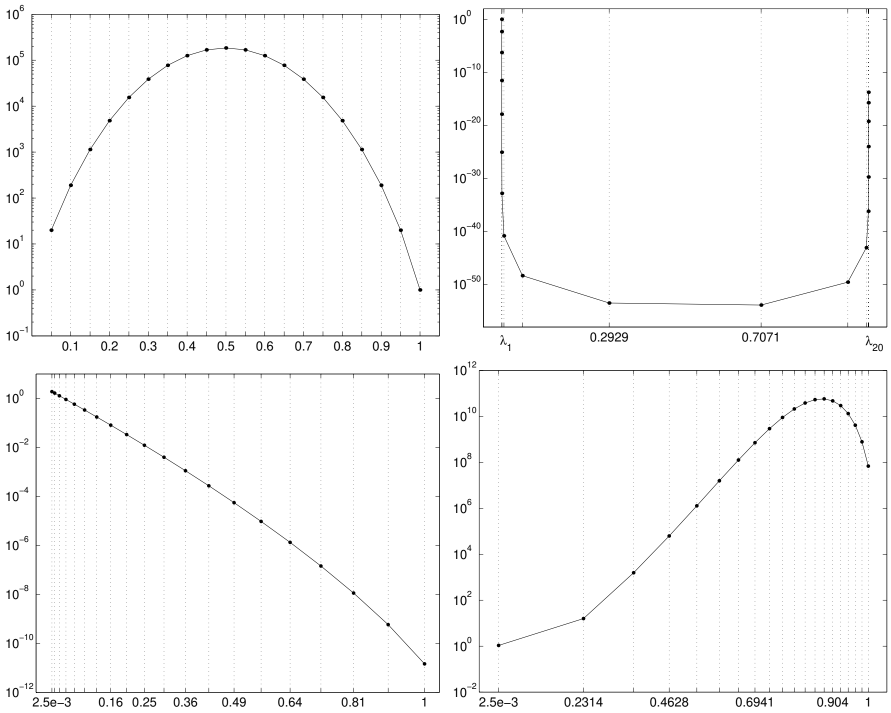

Figure 2.1: The values $|l_j(0)|$ for example eigenvalue distributions.

of great importance for the GMRES performance. This is also demonstrated in the two further examples.

The bottom left and bottom right figures show the values $|l_j(0)|$ for the eigenvalue distributions

$$\begin{aligned} \lambda_j^{(0)} &= j^2/400\,, \quad j = 1, \ldots, 20\,, \quad \text{and} \\ \lambda_j^{(1)} &= \log(j)/\log(20)\,, \quad j = 2, \ldots, 20\,, \quad \lambda_1^{(1)} \;=\; 1/400\,, \end{aligned}$$

having clusters close to zero and one, respectively. Each normal matrix having either the $\lambda_j^{(0)}$ or the $\lambda_j^{(1)}$ as its eigenvalues has the (moderate) condition number 400. Nevertheless, the GMRES residual norms in the next-to-last step for the two eigenvalue sets may differ by several orders of magnitude. While the value of (2.14) for the eigenvalues $\lambda_j^{(0)}$ is typically close to one, it is typically of order 10^{-10} for the eigenvalues $\lambda_j^{(1)}$. This is a numerical illustration why the convergence bounds for GMRES and other Krylov subspace methods such as CG and MINRES that are based on the condition number *only* (see [3, Chapter 3.1] for an overview), can provide misleading information about the actual convergence behavior.

3 Worst-case residual norm

In this section we study the worst-case GMRES residual norms for normal matrices. By "worst-case" we mean, for a given matrix A, the maximally attainable GMRES residual norm in every iteration step i. To make our notion precise we introduce the following definition.

DEFINITION 3.1. *An ith worst-case GMRES residual r_i^w for $A \in \mathbb{C}^{n\times n}$ is a GMRES residual that satisfies*

$$\|r_i^w\| \;=\; \max_{\|r_0\|=1} \min_{p\in\pi_i} \|p(A)\, r_0\|\,, \qquad i = 1,\ldots,n-1. \tag{3.1}$$

A few remarks concerning our definition are in place. First, the restriction that $\|r_0\| = 1$ in (3.1) is made for convenience only. If we would drop this restriction, then the right hand side of (3.1) and all subsequent formulas based on (3.1) must be multiplied by $\|r_0\|$.

Second, as indicated by the wording of the definition, worst-case residuals are not unique. For example, when r_0^i yields a certain ith worst-case residual r_i^w for a given matrix A, then for all $|\alpha| = 1$, $\alpha\, r_0^i$ yields, for the same A, the ith GMRES residual $\alpha\, r_i^w$. Obviously, $\|r_i^w\| = \|\alpha\, r_i^w\|$, so that all vectors $\alpha\, r_i^w$ are ith worst-case residuals for A.

Third, for each normal matrix $A \in \mathbb{C}^{n\times n}$ (with n distinct eigenvalues) and each GMRES iteration step $i = 1,\ldots,n-1$, there exists an ith worst-case residual r_i^w. The reasoning goes as follows. Assuming that $\|r_0\| = 1$, the standard upper bound on the GMRES residual norms [14, Proposition 4] follows easily from (2.2),

$$\|r_i\| \;\le\; \min_{p\in\pi_i} \|p(A)\| \;=\; \min_{p\in\pi_i} \max_{\lambda_j\in L} |p(\lambda_j)|. \tag{3.2}$$

The quantity $\min_{p\in\pi_i} \|p(A)\|$ (called the "ideal GMRES" approximation [5]) is independent of r_0 and thus represents an upper bound on the worst-case GMRES residual norm for the matrix A in step i. As shown independently in [4] and [8], for each normal matrix A and each step i, there exists an initial residual r_0^i so that equality holds in (3.2). Clearly, the ith GMRES residual corresponding to r_0^i is an ith worst-case GMRES residual for A in the sense of Definition 3.1.

Fourth, except for special cases, there exists no *single* initial residual that leads to a worst-case GMRES residual norm $\|r_i^w\|$ in *every* step i. Typically the convergence curve produced by the norms $\|r_i^w\|$ is an artifact, and does not correspond to an actual GMRES run with a given matrix and initial residual.

Fifth, since we assume that all eigenvalues are distinct, it holds $\|r_i^w\| > 0$ for $i = 0,\ldots,n-1$. Therefore, the initial residual r_0^i corresponding to r_i^w has at least $i+1$ nonzero coordinates in the eigenvector basis.

For each subset S of the eigenvalues of A, $S \subseteq L$, we denote

$$M_i^S \;\equiv\; \min_{p\in\pi_i} \max_{\lambda_j\in S} |p(\lambda_j)|. \tag{3.3}$$

The result of [4, 8], which will play an important role in our further development, can in this notation be phrased as follows: For each normal matrix $A \in \mathbb{C}^{n\times n}$ (with n distinct eigenvalues) and each $i = 1, \ldots, n-1$, there exists a worst-case GMRES residual r_i^w with

$$\|r_i^w\| \;=\; M_i^L\,. \tag{3.4}$$

As outlined in the Introduction it is of great interest to find explicit formulas for the polynomials that achieve the min-max value M_i^L, and to identify the properties of the initial residuals r_0^i that yield a worst-case GMRES residual in step i. In the following we will address these questions. We will first consider the iteration step $i = n-1$, and then the case of a general iteration step i.

3.1 Worst case in step $n-1$

The following result completely characterizes the worst-case GMRES in the next-to-last iteration step.

THEOREM 3.1. *For a given normal matrix $A \in \mathbb{C}^{n\times n}$ with n distinct eigenvalues the unit norm initial residual r_0^{n-1} yields an $(n-1)$st worst-case GMRES residual if and only if the coordinates of r_0^{n-1} in the eigenvectors of A satisfy*

$$|\varrho_j^{n-1}|^2 \;=\; \frac{|l_j(0)|}{\sum\limits_{k=1}^{n}|l_k(0)|}\,, \qquad j = 1, \ldots, n\,. \tag{3.5}$$

The norm of the $(n-1)$st worst-case GMRES residual r_{n-1}^w is given by

$$\|r_{n-1}^w\| \;=\; \left(\sum_{k=1}^{n}|l_k(0)|\right)^{-1}, \tag{3.6}$$

and the corresponding worst-case GMRES polynomial $p_{n-1}^w(\lambda)$ has the form

$$p_{n-1}^w(\lambda) \;=\; \|r_{n-1}^w\|\sum_{j=1}^{n}\frac{\overline{l_j(0)}}{|l_j(0)|}\;l_j(\lambda)\,. \tag{3.7}$$

Moreover,

$$|p_{n-1}^w(\lambda_j)| \;=\; \|r_{n-1}^w\| \;=\; M_{n-1}^L\,, \qquad j = 1, \ldots, n\,, \tag{3.8}$$

where L denotes the set of eigenvalues of A.

PROOF. To find an $(n-1)$st worst-case GMRES residual we need to maximize the GMRES residual norm given by (2.14) under the constraint that the initial residual has unit norm. This is equivalent to solving the following constraint minimization problem for the coordinates of the initial residual in the eigenvectors of A,

$$\min_{\varrho_1^{n-1}\neq 0,\ldots,\varrho_n^{n-1}\neq 0}\;\sum_{j=1}^{n}\frac{|l_j(0)|^2}{|\varrho_j^{n-1}|^2}\,, \qquad \text{where} \quad \sum_{j=1}^{n}|\varrho_j^{n-1}|^2 \;=\; 1\,.$$

According to Cauchy's inequality,

$$\sum_{j=1}^{n} \left| \frac{l_j(0)}{\varrho_j^{n-1}} \right|^2 \;=\; \sum_{j=1}^{n} \left| \frac{l_j(0)}{\varrho_j^{n-1}} \right|^2 \, \sum_{j=1}^{n} |\varrho_j^{n-1}|^2 \;\geq\; \left(\sum_{j=1}^{n} |l_j(0)| \right)^2 ,$$

with equality if and only if

$$\xi \left| \frac{l_j(0)}{\varrho_j^{n-1}} \right| \;=\; |\varrho_j^{n-1}| \quad \Leftrightarrow \quad \xi \, |l_j(0)| \;=\; |\varrho_j^{n-1}|^2 ,$$

for all $j = 1, \ldots, n$ and some real ξ. The number ξ is determined from

$$\xi \sum_{k=1}^{n} |l_k(0)| \;=\; \sum_{k=1}^{n} |\varrho_k^{n-1}|^2 \;=\; 1 \quad \Rightarrow \quad \xi \;=\; \left(\sum_{k=1}^{n} |l_k(0)| \right)^{-1} .$$

Hence $|\varrho_j^{n-1}|^2$ satisfies (3.5) and the norm of the corresponding worst-case residual is given by (3.6).

Next, if we substitute $|\varrho_j^{n-1}|^2$ in the form (3.5) into (2.15) and use the fact that $|\varrho_j^{n-1}|^2 = |l_j(0)| \, \|r_{n-1}^w\|$, then we obtain the worst-case polynomial (3.7). Finally, since $l_j(\lambda_k) = \delta_{jk}$, the worst-case polynomial has at every eigenvalue the same absolute value as shown in the first equality in (3.8), and the second equality in (3.8) follows from (3.4) with $i = n - 1$. □

Remark 3.1. *Note that the theorem gives, besides the GMRES context, the explicit solution for a general polynomial approximation problem in the complex plane. In particular,* (3.6) *can be derived with some effort from the results of [13, Section 3]. It can be shown that* (3.6) *is equivalent to*

$$M_i^L = \frac{|\det V_n|}{\sum\limits_{j=1}^{n} |\det V_n^{(j)}|} ,$$

where $V_n^{(j)}$ *denotes the* $(n-1)$*-by-*$(n-1)$ *matrix resulting from deletion of the first column and* j*th row of* V_n*. In our notation, this corresponds to the formula given in [13, Remark 3, p. 692]. However, we are unaware that* (3.5) *or* (3.7) *have been found before.*

Remark 3.2. *We point out that the* $(n-1)$*st worst-case GMRES polynomial* $p_{n-1}^w(\lambda)$ *as given in* (3.7) *is uniquely determined, since it depends only on the uniquely determined quantities* $\|r_{n-1}^w\|$ *and* $l_j(\lambda)$, $j = 1, \ldots, n$*. In particular,* $p_{n-1}^w(\lambda)$ *is independent of the choice of the coordinates* (3.5)*, so that all initial residuals with coordinates* (3.5) *lead to the same* $(n-1)$*st worst-case GMRES polynomial.*

Theorem 3.1 generalizes the results of [2, Section 2] (for $i = n - 1$) from hermitian to all normal matrices. In addition, the theorem allows to give new proofs for a number of known results. We present two examples:

1. *Complete stagnation of GMRES.* The question we ask is whether for a given normal matrix A there exists a unit norm vector r_0^{n-1} such that GMRES completely stagnates, i.e.

$$1 \;=\; \|r_0^{n-1}\| \;=\; \|r_{n-1}^w\|\,.$$

   Using Theorem 3.1 and the uniqueness of the $(n-1)$st worst-case GMRES polynomial it is easy to see that in case of complete stagnation this polynomial is given by $p_{n-1}^w(\lambda) \equiv 1$. Then (3.7) implies

$$p_{n-1}^w(\lambda_j) \;=\; \frac{\overline{l_j(0)}}{|l_j(0)|} \;=\; 1\,, \qquad j = 1, \ldots, n.$$

   In other words, complete stagnation can occur only if all $l_j(0)$, $j = 1, \ldots, n$, are real and positive. Using other means this result was previously derived in [18, Theorem 3.1].

2. *Ideal GMRES approximation.* The proofs of (3.4) in [4, 8] are based on intricate constructions. For the special case $i = n-1$ we now give a simple proof of (3.4), i.e. that

$$\max_{\|r_0\|=1} \min_{p\in\pi_i} \|p(A)r_0\| \;=\; \min_{p\in\pi_i} \|p(A)\|$$

   holds for all normal matrices A. As in (3.6) and (3.7), let r_{n-1}^w and $p_{n-1}^w(\lambda)$ denote an $(n-1)$st worst-case GMRES residual and polynomial for A, respectively. Then

$$\begin{aligned}
\min_{p\in\pi_{n-1}} \|p(A)\| \;&=\; \min_{p\in\pi_{n-1}} \max_{\lambda_j\in L} |p(\lambda_j)| \\
&\leq\; \max_{\lambda_j\in L} |p_{n-1}^w(\lambda_j)| \\
&=\; \|r_{n-1}^w\| \\
&=\; \max_{\|r_0\|=1} \min_{p\in\pi_{n-1}} \|p(A)r_0\| \\
&\leq\; \min_{p\in\pi_{n-1}} \|p(A)\|\,,
\end{aligned}$$

   so that equality must hold throughout. Note that for the last inequality we have used the standard bound (3.2).

3.2 Worst case in a general step i

We next attempt to characterize the worst-case GMRES in a general iteration step $i < n-1$. To this end we derive a lower bound on the min-max value

$$M_i^L \;=\; \min_{p\in\pi_i} \max_{\lambda_j\in L} |p(\lambda_j)|\,.$$

We use the simple fact that

$$M_i^L \;\geq\; M_i^S \tag{3.9}$$

holds for any set $S \subseteq L$. For any set $S \subseteq L$ containing $i+1$ distinct elements there exists a normal $(i+1)$-by-$(i+1)$ matrix with the spectrum S. To this matrix we can apply Theorem 3.1 which completely characterizes the worst-case GMRES in step i, and in particular shows that

$$M_i^S \;=\; \left(\sum_{k=1}^{i+1} |l_k^S(0)| \right)^{-1}, \tag{3.10}$$

where $l_k^S(\lambda)$, $k = 1, \ldots, i+1$, denotes the kth Lagrange polynomial corresponding to the elements in the set S. Using (3.9) and (3.10) it is easy to see that for the given matrix A with the spectrum L,

$$M_i^L \;\geq\; \max_{\substack{S \subseteq L \\ |S|=i+1}} M_i^S \;=\; \max_{\substack{S \subseteq L \\ |S|=i+1}} \left(\sum_{k=1}^{i+1} |l_k^S(0)| \right)^{-1}. \tag{3.11}$$

The natural question arises how close is the lower bound (3.11). In the following we will discuss this question and distinguish between two situations: Either all eigenvalues of A are real, or A has at least one non-real eigenvalue. The first case covers symmetric and hermitian matrices, the second case all other normal matrices.

3.2.1 All eigenvalues are real: (3.11) *is an equality*

When all eigenvalues forming the set L are real, then it follows from a classical result of approximation theory that (3.11) is an equality for $i = 1, \ldots, n-1$. This means that for each $i = 1, \ldots, n-1$ there exists a set $\widehat{S} \subseteq L$ with $|\widehat{S}| = i+1$, such that

$$M_i^L \;=\; M_i^{\widehat{S}} \;=\; \left(\sum_{k=1}^{i+1} |l_k^{\widehat{S}}(0)| \right)^{-1},$$

see e.g. [1, Theorem 2.4 and Corollary 2.5]. In this case Theorem 3.1 can be applied to the given (symmetric or hermitian) matrix A. In particular, (3.5) shows that the coordinates of r_0^i satisfy

$$|\varrho_j^i|^2 \;=\; \frac{|l_j^{\widehat{S}}(0)|}{\sum_{k=1}^{i+1} |l_k^{\widehat{S}}(0)|} \quad \text{if} \quad \lambda_j \in \widehat{S}\,, \qquad \varrho_j^i \;=\; 0 \quad \text{if} \quad \lambda_j \notin \widehat{S}\,.$$

Since r_0^i has only $i+1$ nonzero coordinates in the eigenvectors of A, GMRES will for this initial residual have the worst-case residual norm in the step i, but then terminate in the subsequent step $i+1$. Using a different approach, these results have been previously derived for symmetric positive definite matrices in the context of the conjugate gradient method [2].

3.2.2 At least one non-real eigenvalue: (3.11) *may be strict*

When L contains at least one non-real eigenvalue, then (3.11) may be strict for $i = 1, \ldots, n-2$. In fact, the smallest set $S \subseteq L$ for which $M_i^L = M_i^S$ might contain as many as $2i+1$ distinct elements in the general complex case, see e.g. [1, Corollary 2.5]. For $|S| > i+1$, however, the results of Theorem 3.1 cannot be used, and we are unable to express M_i^S in terms of explicit polynomials. Still, the inequality (3.11) represents a lower bound for M_i^L. Furthermore, we can find an upper bound for M_i^L using an approach similar to the proof of [7, Theorem 4.1].

THEOREM 3.2. *For any set L of n distinct complex points it holds*

$$M_i^L \;\leq\; \sqrt{(i+1)(n-i)}\, \max_{\substack{S \subseteq L \\ |S| = i+1}} M_i^S\,, \quad i = 1, \ldots, n-2\,. \tag{3.12}$$

PROOF. Consider any normal matrix $A \in \mathbb{C}^{n \times n}$ having n distinct eigenvalues forming the set L. Let r_0^i denote an initial residual that yields an ith worst-case GMRES residual r_i^w and let ϱ_j^i, $j = 1, \ldots, n$, denote the coordinates of r_0^i in the eigenvectors of A. The min-max value M_i^L can be written, according to (3.4) and (2.8), in the form

$$M_i^L \;=\; \|r_i^w\| \;=\; \|\, e_1^H (D_i V_{i+1})^+ \,\|^{-1},$$

where $D_i \equiv \operatorname{diag}(\varrho_1^i, \ldots, \varrho_n^i)$. Now consider $i+1$ rows of $D_i V_{i+1}$ that form a square matrix U of order $i+1$ such that $|\det(U)|$ is maximal. Then, as in the proof of [7, Theorem 4.1],

$$\|r_i^w\| \leq \sqrt{(i+1)(n-i)}\; \|U^{-H} e_1\|^{-1}\,. \tag{3.13}$$

The matrix U is defined by some $i+1$ eigenvalues and by corresponding coordinates ϱ_j^i. Denote the set of eigenvalues that define U by $\hat{S} = \{\lambda_1^{\hat{S}}, \ldots, \lambda_{i+1}^{\hat{S}}\}$ and the corresponding (nonzero) coordinates by $\varrho_1^{\hat{S}}, \ldots, \varrho_{i+1}^{\hat{S}}$. Using (2.13), $\|r_0^i\| = 1$, and Cauchy's inequality we obtain

$$\|U^{-H} e_1\|^2 \;=\; \sum_{j=1}^{i+1} \left| \frac{l_j^{\hat{S}}(0)}{\varrho_j^{\hat{S}}} \right|^2 \;\geq\; \sum_{j=1}^{i+1} \left| \frac{l_j^{\hat{S}}(0)}{\varrho_j^{\hat{S}}} \right|^2 \sum_{j=1}^{i+1} |\varrho_j^{\hat{S}}|^2 \;\geq\; \left(\sum_{j=1}^{i+1} |l_j^{\hat{S}}(0)| \right)^2,$$

i.e.

$$\|U^{-H} e_1\|^{-1} \;\leq\; \left(\sum_{j=1}^{i+1} |l_j^{\hat{S}}(0)| \right)^{-1} \;=\; M_i^{\hat{S}}\,. \tag{3.14}$$

Thus we have found a set $\widehat{S} \subseteq L$, $|\widehat{S}| = i+1$, such that

$$\|r_i^w\| \;\leq\; \sqrt{(i+1)(n-i)}\; M_i^{\hat{S}}. \tag{3.15}$$

Substituting in (3.15) for $M_i^{\hat{S}}$ the maximum of M_i^S over all subsets $S \subseteq L$, $|S| = i+1$, we obtain (3.12). $\square$

Our numerical experiments with various spectra (see Section 4) show that the lower bound (3.11) is very tight and that the upper bound (3.12) represents an overestimation. In particular, we *conjecture* that there exists a small constant $C > 1$ such that

$$\max_{\substack{S\subseteq L\\|S|=i+1}} M_i^S \;\le\; M_i^L \;\le\; C \max_{\substack{S\subseteq L\\|S|=i+1}} M_i^S\,, \quad i = 1,\ldots,n-2\,, \tag{3.16}$$

holds for *all* sets L containing n distinct complex numbers (for $i = n-1$, (3.16) obviously holds with $C = 1$). In our numerical tests the ratio

$$\frac{M_i^L}{\max\limits_{\substack{S\subseteq L\\|S|=i+1}} M_i^S} \tag{3.17}$$

was maximal for sets L containing n numbers uniformly distributed on the unit circle. On such sets of points, (3.17) for $i = n-2$ converges from below to $4/\pi$ as $n \to \infty$. Hence $C = 4/\pi$ is the smallest constant for which (3.16) can hold for all sets L with $|L| = n$, cf. the Appendix. On the other hand, we were unable to find a set L for which the ratio (3.17) was larger than $4/\pi$.

4 Numerical experiments

We now study the worst-case GMRES residual norms, our lower bound (3.11), and our conjecture (3.16) with $C = 4/\pi$ for four different eigenvalue sets L. In the left part of Figures 4.1– 4.4 we plot the worst-case GMRES residual norms $\|r_i^w\|$ (bold line), and the values

$$\max_{\substack{S\subseteq L\\|S|=i+1}} M_i^S \qquad \text{(solid line)}\,,$$

$$\frac{4}{\pi}\max_{\substack{S\subseteq L\\|S|=i+1}} M_i^S \qquad \text{(dashed line)}\,.$$

Our conjecture is that the dashed curve is an upper bound on the worst-case GMRES residual norm in every step. The right part of each figure shows the corresponding eigenvalue distributions. In the step i, we compute the values M_i^S for all subsets $S \subseteq L$, $|S| = i+1$, and determine our bounds from their maximum. This computation is quite expensive, so we consider only small sets of points ($n = 18$). The worst-case GMRES residual norm in every step is computed using the function `cheby0` of the semidefinite programming package SDPT3 [15]. Although this function may fail to converge when $\|r_i^w\|$ becomes very small (see below for details), it is the most reliable function we know for this type of computation. All experiments are performed in Matlab 6.5 Release 13 on an AMD Athlon XP 2100+ personal computer with machine precision $\varepsilon \sim 10^{-16}$.

Roots of unity. In the first numerical experiment we consider the eigenvalue set L consisting of the 18th roots of unity, i.e.

$$\lambda_k \;=\; e^{\mathbf{i}\frac{2k\pi}{18}}, \quad k = 1,\ldots,18\,. \tag{4.1}$$

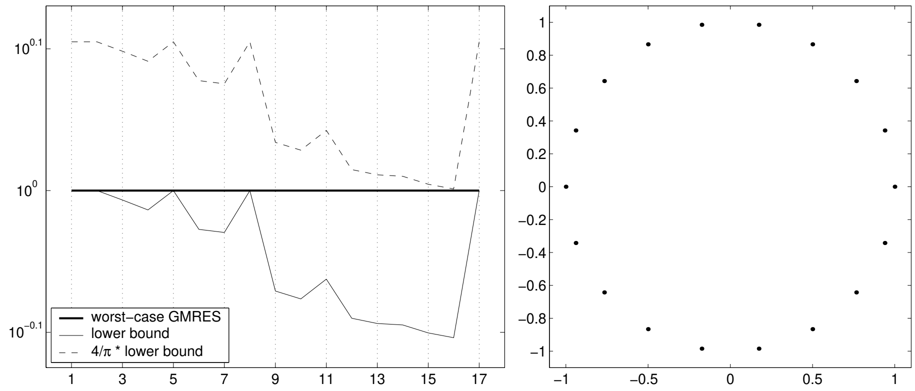

Figure 4.1: Worst-case GMRES and our bounds for roots of unity.

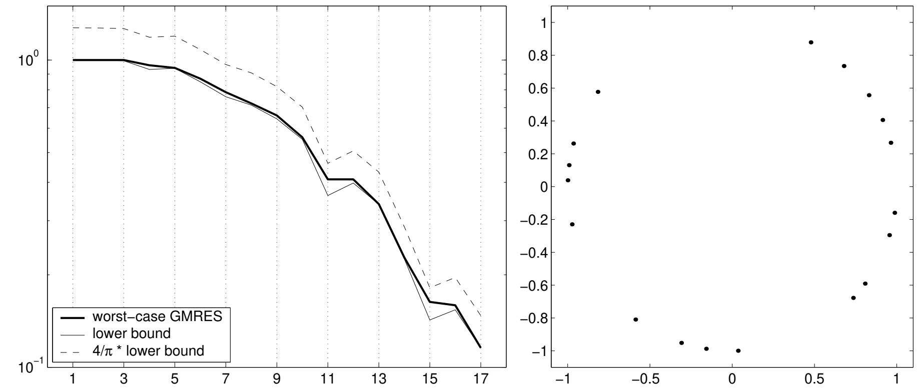

Figure 4.2: Worst-case GMRES and our bounds for random eigenvalues on the unit circle.

In this case worst-case GMRES completely stagnates, cf. [18], which is confirmed by the bold line in Figure 4.1. The lower bound (3.11) closely approximates the worst-case residual norm, and the lower bound multiplied by $4/\pi$ represents an upper bound. As shown in the Appendix, see also [12], the lower bound approaches $\pi/4$ from above in the step $i = n - 2$ (here: $i = 16$) when $n \to \infty$. Hence in this step the lower bound multiplied by $4/\pi$ is proven to be a (sharp) upper bound on the worst-case GMRES. The tightness of this bound, even for the moderate $n = 18$, is clearly visible in Figure 4.1.

Random eigenvalues on the unit circle. For random eigenvalues on the unit circle (cf. the right part of Figure 4.2), the worst-case GMRES residual norms do not stagnate completely, but still converge very slowly (decreasing only about one order of magnitude until the next-to-last step). The lower bound (3.11) is very close to the worst-case residual norm, only in the iteration steps 4, 7, 11 and

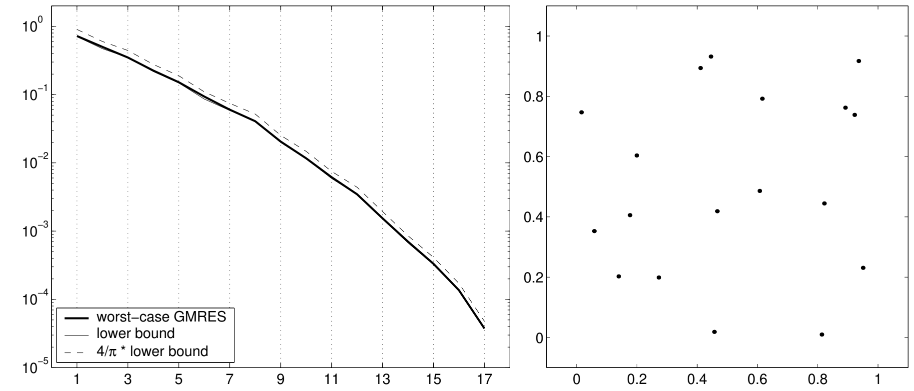

Figure 4.3: Worst-case GMRES and our bounds for random eigenvalues in the region $[0, 1] \times \mathbf{i}\,[0, 1]$.

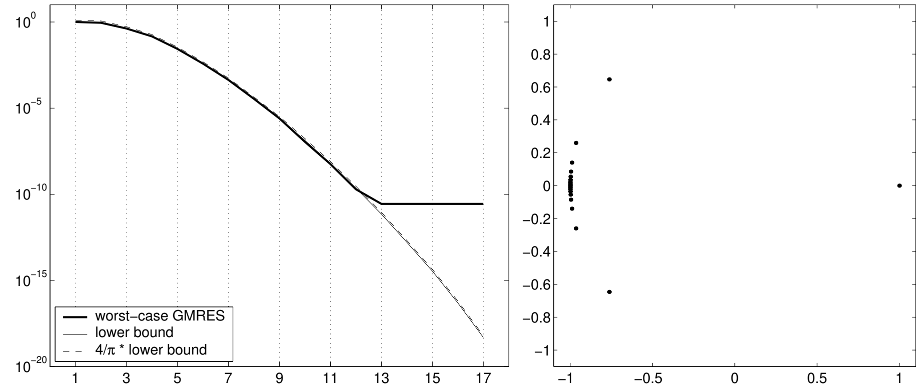

Figure 4.4: Worst-case GMRES and our bounds for the Helmert matrix.

15 they differ slightly. As above, the lower bound multiplied by $4/\pi$ represents an upper bound.

Random eigenvalues in the region $[0, 1] \times \mathbf{i}\,[0, 1]$. In this case the convergence of the worst-case residual norms is moderately fast; they decrease about 4 orders of magnitude until the next-to-last step, cf. Figure 4.3. Again the lower bound (3.11) is a good estimate (bold and soild line almost coincide), and the dashed line represents an upper bound.

Helmert matrix. For the last experiment we use the Helmert matrix generated by the Matlab command `gallery('orthog',18,4)`. Helmert matrices occur in a number of practical problems, for example in applied statistics [9]. Our matrix is orthogonal, and the eigenvalues cluster around -1, see the right part of Figure 4.4. The worst-case GMRES residual norm decreases quickly throughout the iteration. Until the 12th step the worst-case curve and the lower bound almost coincide, and the lower bound multiplied by $4/\pi$ represents an upper bound. However, when the worst-case residual norm drops below the level of 10^{-10} (iteration step 13 and beyond), the function `cheby0` apparently has reached its final level of accuracy and henceforth stagnates. Such stagnation (sometimes divergence) can be generally observed when the final accuracy level is reached, but we are unaware of an analysis how this level depends on the problem parameters.

In summary, the numerical experiments demonstrate that our lower bound (3.11) is very tight. Moreover, in all experiments the lower bound multiplied by $4/\pi$ represents an upper bound on the worst-case GMRES residual norms, which supports that our conjecture (3.16) with $C = 4/\pi$ is true. Note that in all experiments the bound (3.12), which contains a factor between $n^{1/2}$ and n, represents an overestimation.

5 Concluding discussion

We conclude the paper with a discussion of our results and starting points for further work.

1. Interpretation of the lower bound (3.11). Recall that the worst-case GMRES residual norm in step i is equal to the min-max value M_i^L. This value represents the solution of an ith degree polynomial approximation problem on n distinct eigenvalues forming the set L. We bound this value from below by the same approximation problem, but on subsets of L containing exactly $i+1$ eigenvalues. The solution of each "reduced"problem (polynomial of degree i on $i+1$ distinct points) is given in Theorem 3.1.

For illustration of this process we consider the set L consisting of the nth roots of unity, cf. Figure 4.1 for $n = 18$. As shown in [18, Theorem 3.4], worst-case GMRES completely stagnates in this case, i.e. $M_{n-1}^L = 1$. Intuitively, for each $i < n-1$ there exists a subset $\widehat{S} \subset L$, $|\widehat{S}| = i+1$, that closely resembles the $(i+1)$st roots of unity. For such a set the min-max value $M_i^{\widehat{S}}$ is close to one (in orders of magnitude), which is why the lower bound (3.11) is very tight. In particular, whenever $n \bmod (i+1) = 0$, there exists a set $\widehat{S} \subset L$ consisting of exactly the $(i+1)$st roots of unity. For all such iteration steps i, the lower bound (3.11) is an equality, cf. $i = 1, 2, 5, 8$ for $n = 18$ in Figure 4.1.

Note that here, and for general sets L, M_i^S is close to M_i^L only for special sets $S \subset L$ with $|S| = i+1$. Analyzing the structure of these sets based on the eigenvalue distribution of A is a topic we plan to pursue in our future work.

2. Worst case vs. average (unbiased) case. Due to orthogonality of the eigenvectors of A, initial residuals with (approximately) equal components in all eigenvectors are often considered the "average" case, see e.g. [5, Section 7]. We prefer to call them "unbiased" since they are not biased towards a certain eigenvector direction. For simplicity, consider any unbiased unit norm initial residual r_0^u with eigenvector components of *equal* size, i.e. $|\varrho_j^u| = n^{-1/2}$, $j = 1, \ldots, n$. Then (2.5) and (3.4) show that the ith GMRES residual r_i^u corresponding to r_0^u satisfies

$$\begin{aligned}\|r_i^w\| \;\geq\; \|r_i^u\| \;&=\; n^{-1/2} \min_{p\in\pi_i} \left(\sum_{j=1}^{n} |p(\lambda_j)|^2 \right)^{-1/2} \\ &\geq\; n^{-1/2} \min_{p\in\pi_i} \max_{1\leq j\leq n} |p(\lambda_j)| \\ &=\; n^{-1/2}\, \|r_i^w\|\,.\end{aligned}$$

Since the unbiased (average) and the worst case GMRES residual norms agree up to a factor of $n^{1/2}$, our results are relevant beyond the specific analysis of the worst-case GMRES.

In practical applications the initial residual may, for example, be biased towards the eigenvalue distribution of A. Often such biased initial residuals result from choosing a nonzero initial guess x_0. The biased case depends strongly on the specific application, and a general analysis is beyond the scope of this paper.

Acknowledgment

We thank Prof. Ludwig Elsner for pointing out a simplification in a previous version of our proof of Theorem 3.1.

Appendix

PROPOSITION A.1. *The smallest constant C for which* (3.16) *can hold for all sets L containing n distinct complex numbers is $C = 4/\pi$.*

PROOF. We will show that for the set $L = \{\lambda_1, \ldots, \lambda_n\}$ defined by

$$\lambda_k \;=\; e^{\mathbf{i}\frac{2k\pi}{n}}, \quad k = 1, \ldots, n\,, \tag{A.1}$$

the ratio (3.17) for $i = n-2$ converges from below to $C = 4/\pi$ as $n \to \infty$.

Note that all sets $S \subset L$ with $|S| = n-1$ can be obtained by rotation of the set $L - \{\lambda_1\}$. Therefore

$$\max_{\substack{S\subseteq L\\ |S|=n-1}} M_{n-2}^{S} \;=\; M_{n-2}^{L-\{\lambda_1\}} \;=\; \left(\sum_{k=2}^{n} \prod_{\substack{j=2\\ j\neq k}}^{n} \frac{|\lambda_j|}{|\lambda_j - \lambda_k|} \right)^{-1}.$$

Substituting $|\lambda_j| = 1$ for all j, and

$$|\lambda_j - \lambda_k| \;=\; |e^{\mathbf{i}\frac{2j\pi}{n}} - e^{\mathbf{i}\frac{2k\pi}{n}}| \;=\; 2\,\sin\left(\frac{|j-k|\pi}{n} \right),$$

shows that

$$M_{n-2}^{L-\{\lambda_1\}} \;=\; \left(\sum_{k=1}^{n-1} \frac{1}{2^{n-2}} \prod_{\substack{j=1\\ j\neq k}}^{n-1} \frac{1}{\sin\left(\frac{j\pi}{n}\right)} \right)^{-1} = 2^{n-2} \left(\frac{\sum_{k=1}^{n-1} \sin\left(\frac{k\pi}{n}\right)}{\prod_{j=1}^{n-1} \sin\left(\frac{j\pi}{n}\right)} \right)^{-1}.$$

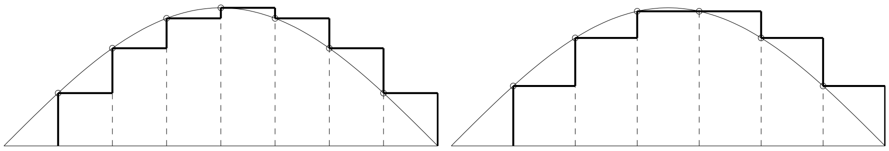

Figure A.1: The approximation of the integral for n even (left part) and n odd (right part).

Using the standard formula

$$\prod_{j=1}^{n-1} \sin\left(\frac{j\pi}{n}\right) = \frac{n}{2^{n-1}},$$

we obtain

$$M_{n-2}^{L-\{\lambda_1\}} = \left[\frac{2}{n}\sum_{k=1}^{n-1}\sin\left(\frac{k\pi}{n}\right)\right]^{-1} = \frac{\pi}{2}\left[\frac{\pi}{n}\sum_{k=1}^{n-1}\sin\left(\frac{k\pi}{n}\right)\right]^{-1}. \tag{A.2}$$

Note that the expression on the right hand side of (A.2) is an approximation of an integral,

$$\frac{\pi}{n}\sum_{k=1}^{n-1}\sin\left(\frac{k\pi}{n}\right) < \int_0^{\pi}\sin(x)\,dx = 2, \qquad \lim_{n\to\infty}\left[\frac{\pi}{n}\sum_{k=1}^{n-1}\sin\left(\frac{k\pi}{n}\right)\right] = 2,$$

see Figure A.1 for a numerical illustration. Therefore,

$$M_{n-2}^{L-\{\lambda_1\}} > \frac{\pi}{4}, \qquad \lim_{n\to\infty} M_{n-2}^{L-\{\lambda_1\}} = \frac{\pi}{4}.$$

As shown in [18], complete stagnation of GMRES can occur for normal matrices having the spectrum L, and hence $M_i^L = 1$ for $i = 1, \ldots, n-1$. Therefore,

$$M_{n-2}^{L} < \frac{4}{\pi}\max_{\substack{S\subseteq L\\ |S|=n-1}} M_i^S, \qquad \lim_{n\to\infty}\left[\frac{4}{\pi}\max_{\substack{S\subseteq L\\ |S|=n-1}} M_i^S\right] = M_{n-2}^{L},$$

which completes the proof. A similar result can be shown for other i; see [12] for more details. $\square$

REFERENCES

1. R. A. DeVore and G. G. Lorentz, *Constructive approximation*, vol. 303 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 1993.
2. A. Greenbaum, *Comparison of splittings used with the conjugate gradient algorithm*, Numer. Math., 33 (1979), pp. 181–193.

3. ———, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.
4. A. Greenbaum and L. Gurvits, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 348–358.
5. A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.
6. N. J. Higham, *Accuracy and stability of numerical algorithms*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996.
7. I. Ipsen, *Expressions and bounds for the GMRES residual*, BIT, 40 (2000), pp. 524–535.
8. W. Joubert, *A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439.
9. H. Lancaster, *The Helmert matrices*, Amer. Math. Monthly, 72 (1965), pp. 4–12.
10. J. Liesen, *Construction and analysis of polynomial iterative methods for non-hermitian systems of linear equations*, PhD thesis, Fakultät für Mathematik der Universität Bielefeld, November 1998.
11. J. Liesen, M. Rozložník, and Z. Strakoš, *Least squares residuals and minimal residual methods*, SIAM J. Sci. Comput., 23 (2002), pp. 1503–1525.
12. J. Liesen and P. Tichý, *A min-max problem on roots of unity*, tech. report, Department of Mathematics, Technical University of Berlin, Germany, 2003.
13. T. J. Rivlin and H. S. Shapiro, *A unified approach to certain problems of approximation and minimization*, J. Soc. Indust. Appl. Math., 9 (1961), pp. 670–699.
14. Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
15. K. Toh, M. Todd, and R. Tütüncü, *SDPT3 – a Matlab software package for semidefinite programming, version 2.1*, the software is available on the site `http://www.math.nus.edu.sg/~mattohkc/sdpt3.html`, June 2001.
16. J. M. Varah, *The prolate matrix*, Linear Algebra Appl., 187 (1993), pp. 269–278.
17. I. Zavorin, *Spectral factorization of the Krylov matrix and convergence of GMRES*, Tech. Report CS-TR-4309, Computer Science Department, University of Maryland, 2001.
18. I. Zavorin, D. O'Leary, and H. Elman, *Complete stagnation of GMRES*, Linear Algebra Appl., 367 (2003), pp. 165–183.