ON WORST-CASE GMRES, IDEAL GMRES, AND THE POLYNOMIAL NUMERICAL HULL OF A JORDAN BLOCK

PETR TICHÝ*, JÖRG LIESEN†, AND VANCE FABER‡

Abstract. When solving a linear algebraic system $Ax = b$ with GMRES, the relative residual norm at each step is bounded from above by the so-called ideal GMRES approximation. This worst-case bound is sharp (i.e. it is attainable by the relative GMRES residual norm) in case of a normal matrix A, but it need not characterize the worst-case GMRES behavior if A is nonnormal. Characterizing the tightness of this bound for nonnormal matrices A represents an important and largely open problem in the convergence analysis of Krylov subspace methods. In this paper we address this problem in case A is a single Jordan block. We study the relation between ideal and worst-case GMRES as well as the problem of estimating the ideal GMRES approximation. Furthermore, we prove new results about the radii of the polynomial numerical hulls of Jordan blocks. Using these, we discuss the closeness of the lower bound on the ideal GMRES approximation that is derived from the radius of the polynomial numerical hull.

Key words. GMRES convergence, ideal GMRES, polynomial numerical hull, Jordan block.

AMS subject classifications. 65F10, 65F35, 49K35.

1. Introduction.

Let a nonsingular matrix $A \in \mathbb{C}^{n\times n}$ and a vector $b \in \mathbb{C}^n$ be given. Suppose that we apply the GMRES method [14] with initial guess $x_0 = 0$ (chosen here for convenience and without loss of generality) to the linear system $Ax = b$. Then this method computes a sequence of iterates $x_1, x_2, \ldots$, so that the kth residual $r_k \equiv b - Ax_k$ satisfies

$$\|r_k\| \;=\; \min_{p\in\pi_k} \|p(A)\, b\| \,. \tag{1.1}$$

Here π_k denotes the set of (complex) polynomials of degree at most k and with value one at the origin, and $\|\cdot\|$ denotes the Euclidean norm. The residual r_k is uniquely determined by the minimization condition (1.1) and satisfies the equivalent orthogonality condition

$$r_k \in b + A\mathcal{K}_k(A, b)\,, \qquad r_k \perp A\mathcal{K}_k(A, b)\,. \tag{1.2}$$

Here $\mathcal{K}_k(A, b) \equiv \operatorname{span}\{b, Ab, \ldots A^{k-1}b\}$ is the kth Krylov subspace generated by A and b, and $\perp$ means orthogonality with respect to the Euclidean inner product. Without loss of generality we will consider that b is a unit norm vector, i.e. $\|b\| = 1$.

A common approach for investigating the GMRES convergence behavior is to bound (1.1) independently of b, and thus to study the algorithm's worst-case behavior. In particular, for each iteration step k one may analyze the *worst-case GMRES approximation*

$$\psi_k(A) \;\equiv\; \max_{\|v\|=1} \min_{p\in\pi_k} \|p(A)v\|\,. \tag{1.3}$$

*Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, 18207 Prague, Czech Republic, email: tichy@cs.cas.cz. The work of this author was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft and by the Czech National Program of Research "Information Society" under project 1ET400300415.

†Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, email: liesen@math.tu-berlin.de. This work was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft.

‡BD Biosciences - Bioimaging Systems, email: vance_faber@bd.com.

The quantity $\psi_k(A)$ is attainable by the GMRES residual norm in the following sense: For a given matrix A and every GMRES step k, there exists a unit norm initial vector b, for which the resulting kth GMRES residual norm is equal to $\psi_k(A)$. It should be noted, however, that for a given nonnormal matrix A and integer k the quantity $\psi_k(A)$ typically is very hard to compute. In fact, we are unaware of any efficient algorithm for performing this computation.

Using the submultiplicativity of the Euclidean norm (or by changing the order of maximization and minimization in (1.3)), we can easily find the following upper bound on (1.3),

$$\psi_k(A) \;\leq\; \min_{p\in\pi_k} \|p(A)\| \;\equiv\; \varphi_k(A)\,. \tag{1.4}$$

The quantity $\varphi_k(A)$, called the kth *ideal GMRES approximation*, has been introduced by Greenbaum and Trefethen [7]. They argue that it is important to investigate this quantity to improve the understanding of GMRES (and matrix iterations in general) particularly in the nonnormal case, since the ideal GMRES approximation "disentangles the matrix essence of the [GMRES] process from the distracting effects of the initial vector", see [7, p. 362].

Before continuing this line of thought we have to stress a subtle point: In case $A \in \mathbb{R}^{n\times n}$ it is customary (and we will follow this custom) to assume that $b \in \mathbb{R}^n$, and to consider the approximation problem (1.3) only for $v \in \mathbb{R}^n$. In this (real) case, the values $\psi_k(A)$ and $\varphi_k(A)$ are both attained by real polynomials $p \in \pi_k$. For the worst-case GMRES approximation $\psi_k(A)$ this fact is obvious, while for the ideal GMRES approximation $\varphi_k(A)$ this has been shown in [10, Theorem 3.1].

After the 1994 paper [7], several studies have been devoted to the problem of characterizing the relation between $\psi_k(A)$ and $\varphi_k(A)$, and in particular the tightness of the inequality (1.4). The best known result is that (1.4) is an equality, i.e. $\psi_k(A) = \varphi_k(A)$ for all $k \geq 0$, whenever A is normal [6, 11]. In addition, (1.4) is an equality for arbitrary A and $k = 1$ [6, 11], for triangular Toeplitz matrices when the right hand side of (1.4) equals one [3], and also when the matrix $p_*^{(k)}(A)$ that solves the ideal GMRES approximation problem (1.4) has a simple maximal singular value [6, Lemma 2.4]. On the other hand, some examples of nonnormal matrices have been constructed, for which (1.4) is a sharp inequality [3, 17]. Despite the existence of these counterexamples, it is still an open question whether (1.4) is an equality (or at least tight inequality) for larger classes of nonnormal matrices.

Another open problem in the context of (1.4) is how to determine or estimate the value of the ideal GMRES approximation $\varphi_k(A)$ in general. A possible approach that is still under development is to associate the matrix A with some set in the complex plane and to relate the norm of the matrix polynomial to the maximum norm of the polynomial on this set. An appropriate set, designed to give useful information about the norm of functions of a matrix A, is the *polynomial numerical hull of degree k*,

$$\mathcal{H}_k(A) \;\equiv\; \{z \in \mathbb{C} :\; \|p(A)\| \geq |p(z)| \;\text{ for all } p \in \mathcal{P}_k\}\,, \tag{1.5}$$

introduced by Nevanlinna [13, p. 41]. Here $\mathcal{P}_k$ denotes the set of (complex) polynomials of degree at most k. Based on the definition (1.5) it is not hard to see that these sets provide a lower bound on the ideal GMRES approximation [4],

$$\min_{p\in\pi_k} \max_{z\in\mathcal{H}_k(A)} |p(z)| \;\leq\; \varphi_k(A)\,. \tag{1.6}$$

Moreover, $\mathcal{H}_k(A)$ allows us to identify when ideal GMRES fails to converge [3, 4],

$$\varphi_k(A) = 1 \quad \Longleftrightarrow \quad 0 \in \mathcal{H}_k(A)\,. \tag{1.7}$$

While polynomial numerical hulls appear to be a valuable tool, their determination or computation represents a difficult open problem even for simple classes of nonnormal matrices.

In summary, the investigation of worst-case and ideal GMRES as well as the polynomial numerical hulls for nonnormal matrices is at its very beginning. We believe that in this situation it is helpful to study relatively simple nonnormal matrices, for which explicit solutions of some of the open problems can be derived. Continuing the work started in [2] and [5], we here consider A being an $n \times n$ Jordan block J_λ with eigenvalue $\lambda \in \mathbb{C}$.

When experimenting with the MATLAB software SDPT3 [18] and some Jordan blocks J_λ of small size ($n = 20$, say), we observed numerically that $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$ for $0 \leq k \leq n$. This led us to conjecture that

$$\psi_k(J_\lambda) = \varphi_k(J_\lambda) \quad \text{for all } \lambda, n \text{ and } 0 \leq k \leq n.$$

At first sight, proving this conjecture looks not too difficult; after all, one just has to deal with a single Jordan block. However, it turns out that the approximation problems behind the quantities $\psi_k(A)$ and $\varphi_k(A)$ as well as the exact determination of $\mathcal{H}_k(A)$ are highly nontrivial even in case $A = J_\lambda$. When trying to prove our conjecture we found that numerous cases need to be distinguished, and in the end we were unable to prove all of them. Nevertheless, we believe that the work presented here has been worthwhile. In particular, it uncovered a previously unknown structure behind the worst-case and ideal GMRES approximation problems in case $A = J_\lambda$, it extended the recent results of [2, 5] on the polynomial numerical hulls of Jordan blocks, and it led to new results about the bound (1.6).

Since the presentation below is rather technical, we give a detailed overview of the sections and the corresponding results in this paper:

- In Section 2 we summarize known results on worst-case and ideal GMRES as well as the polynomial numerical hull.
- In Section 3 we show that $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$ for $0 \leq k < n/2$ and whenever $|\lambda|$ is outside a small interval on the positive real line.
- In Section 4 we study the structure of the polynomials that solve the ideal GMRES approximation problem, i.e. the polynomials for which the value $\varphi_k(J_\lambda)$ is attained. This allows us to show that $\varphi_k(J_\lambda) = \psi_k(J_\lambda)$ for all λ in case k divides n. Moreover, we establish a relationship between the radii of polynomial numerical hulls of J_λ.
- In Section 5 we analyze the quantities $\psi_{n-1}(J_\lambda)$ and $\varphi_{n-1}(J_\lambda)$. This allows us to show that $\varphi_{n-k}(J_\lambda) = \psi_{n-k}(J_\lambda)$ whenever $|\lambda| \geq 1$ and k divides n.
- Finally, in Section 6 we apply results of the previous sections to analyze the closeness of the bound (1.6) on the kth ideal GMRES approximation. We are unaware that any theoretical results in this direction have been obtained previously.

2. Notation and theoretical background.

The following result collects a number of basic results concerning the quantities $\psi_k(A)$ and $\varphi_k(A)$. These results are either easy to verify, or they have been published in [10, Theorem 3.1] or [3, Proposition 2.1].

LEMMA 2.1. *Let $A \in \mathbb{C}^{n\times n}$ be a matrix with minimal polynomial degree $d(A)$. Then the following hold:*

1. *$\psi_k(A)$ and $\varphi_k(A)$ are both nonincreasing in k.*
2. *$\psi_0(A) = \varphi_0(A) = 1$.*
3. *$0 < \psi_k(A) \le \varphi_k(A)$ for $1 \le k \le d(A) - 1$.*
4. *If A is nonsingular, then $\psi_k(A) = \varphi_k(A) = 0$ for all $k \ge d(A)$.*
5. *If A is singular, then $\psi_k(A) = \varphi_k(A) = 1$ for all $k \ge 0$.*

The previous theorem shows that to investigate the relation between worst-case and ideal GMRES, one only has to consider nonsingular matrices A and positive integers $k < d(A)$. In this case $\varphi_k(A) > 0$, and the polynomial that solves the ideal GMRES approximation problem (1.4) is uniquely determined [7, Theorem 2]. This gives rise to the following definition.

DEFINITION 2.2. *For a nonsingular matrix $A \in \mathbb{C}^{n\times n}$, and a positive integer $k < d(A)$, the uniquely determined polynomial $p_*^{(k)} \in \pi_k$ that satisfies*

$$\|p_*^{(k)}(A)\| = \varphi_k(A) = \min_{p\in\pi_k} \|p(A)\|,$$

is called the kth ideal GMRES polynomial of A, and the matrix $p_*^{(k)}(A)$ is called the kth ideal GMRES residual matrix of A.
The matrix A is called ideal of degree k, when $\varphi_k(A) = \psi_k(A)$, and A is called ideal, when $\varphi_k(A) = \psi_k(A)$ for $k = 1, \ldots, d(A) - 1$.

We point out that if A is ideal of some degree k, then this does not necessarily imply that A is ideal of any other degree. In fact, it would be interesting to characterize necessary and sufficient conditions on A that allow one to conclude from idealness of some degree to idealness of other degrees.

In general it is an open problem which properties of A are necessary and sufficient for A to be ideal. Below we summarize the most important results for our context. Proofs of all of these statements can be found in [6, 11].

LEMMA 2.3. *Any nonsingular matrix $A \in \mathbb{C}^{n\times n}$ is ideal of degree $k = 1$. Moreover, the following hold:*

1. *If A is normal, then A is ideal.*
2. *If $p_*^{(k)}(A)$ has a simple maximal singular value, then A is ideal of degree k.*

Let us discuss the condition in the second item. If A is a normal matrix with (distinct) eigenvalues $\lambda_1, \ldots, \lambda_{d(A)}$, then the ideal GMRES approximation problem is a (scalar) min-max problem on the set of the eigenvalues,

$$\varphi_k(A) \;=\; \min_{p\in\pi_k} \|p(A)\| \;=\; \min_{p\in\pi_k} \max_{\lambda_i} |p(\lambda_i)|\,.$$

It is well known that the corresponding min-max polynomial of degree k attains its maximum value on at least $k+1$ of the eigenvalues, see, e.g., [1, Chapter 3, §4]. Hence in this case the multiplicity of the maximal singular value of $p_*^{(k)}(A)$ is at least $k+1$. Since any normal matrix is ideal, we see that the condition in the second item is *not necessary*.

This fact has already been noted, and explained by a similar argument, by Greenbaum and Trefethen [7]. Based on some numerical observations, they consider the case in which $p_*^{(k)}(A)$ for a nonnormal matrix A has a simple maximal singular value the “generic case”, see [7, p. 366]. However, we believe that the situation of $p_*^{(k)}(A)$

having a multiple maximal singular value can be quite frequent also for nonnormal A. For a clear example see Fig. 4.1 below, which shows that for the 20×20 Jordan block J_λ with $\lambda = 1$, only 9 out of 19 matrices $p_*^{(k)}(J_\lambda)$ have a simple maximal singular value.

We denote the maximal singular value of a matrix B by $\sigma_{max}(B)$, and we define the linear space

$$\Sigma(B) \equiv \operatorname{span}\{v : v \text{ is a right singular vector of } B \text{ corresponding to } \sigma_{max}(B)\}\,.$$

We use such spaces in the next result, which gives a further characterization of the case $\psi_k(A) = \varphi_k(A)$. This result can be found in a more general form in [3, Lemma 2.16], but we formulate and prove it here independently of [3].

LEMMA 2.4. *Suppose that a nonsingular matrix A and a positive integer $k < d(A)$ are given. Then $\psi_k(A) = \varphi_k(A)$ if and only if there exist a polynomial $q \in \pi_k$ and a unit norm vector $b \in \Sigma(q(A))$, such that*

$$q(A)b \perp A\mathcal{K}_k(A, b)\,. \tag{2.1}$$

If such q and b exist, then $q = p_*^{(k)}$.

Proof. If $\psi_k(A) = \varphi_k(A)$, then there exist a unit norm vector b and a polynomial $q \in \pi_k$ satisfying (2.1), cf. (1.2), such that $\|p_*^{(k)}(A)\| = \|q(A)b\|$. Since $\|p_*^{(k)}(A)b\| \leq \|p_*^{(k)}(A)\|$ and $\|q(A)b\|$ is minimal,

$$\|p_*^{(k)}(A)b\| = \|p_*^{(k)}(A)\| = \|q(A)b\|. \tag{2.2}$$

But this means that $b \in \Sigma(p_*^{(k)}(A))$. Moreover, since $1 \leq k \leq d(A) - 1$, we know that $\psi_k(A) > 0$ by Lemma 2.1, and thus the kth GMRES polynomial is unique, cf. [7, Theorem 2]. Therefore $p_*^{(k)} = q$, and hence $b \in \Sigma(q(A))$.

Now assume that there exist a polynomial $q \in \pi_k$ and a unit norm vector b such that (2.1) holds and $b \in \Sigma(q(A))$. Then

$$\|q(A)\| = \|q(A)b\| = \min_{p \in \pi_k} \|p(A)b\| \leq \|p_*^{(k)}(A)\|. \tag{2.3}$$

Since $p_*^{(k)}$ is the ideal GMRES polynomial, $\|q(A)\| < \|p_*^{(k)}(A)\|$ is impossible, and therefore equality holds in (2.3). But then $\psi_k(A) = \varphi_k(A)$, and from uniqueness of $p_*^{(k)}$ it follows that $q = p_*^{(k)}$. ☐

In [3], the k-dimensional generalized field of values of A,

$$F(\{A^i\}_{i=1}^k) \equiv \left\{ \begin{pmatrix} v^*Av \\ \vdots \\ v^*A^kv \end{pmatrix} : v^*v = 1 \right\},$$

is used to characterize when worst-case or ideal GMRES do not converge.

THEOREM 2.5. *For a nonsingular matrix $A \in \mathbb{C}^{n\times n}$ the following hold:*
1. *$\psi_k(A) = 1$ if and only if $0 \in F(\{A^i\}_{i=1}^k)$.*
2. *$\varphi_k(A) = 1$ if and only if $0 \in \operatorname{cvx}(F(\{A^i\}_{i=1}^k))$, the convex hull of $F(\{A^i\}_{i=1}^k)$.*

Note that when $A \in \mathbb{R}^{n\times n}$ is real, one can take the *real* k-dimensional generalized field of values A (defined over $v \in \mathbb{R}^n$, $v^Tv = 1$).

The k-dimensional generalized field of values of any triangular Toeplitz matrix $T \in \mathbb{C}^{n\times n}$ is a convex set [3], and, therefore,

$$\psi_k(T) = 1 \iff \varphi_k(T) = 1\,, \tag{2.4}$$

i.e. T is ideal of degree k in case of stagnation. However, it is in general still an open problem, originally posed in [3, p. 722], whether T is ideal of degree k when ideal GMRES converges, i.e. when $\varphi_k(T) < 1$.

In this paper we concentrate on an $n \times n$ Jordan block

$$J_\lambda = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \equiv \lambda I_n + E_n\,. \tag{2.5}$$

Apart from the identity matrix I_n and the shift E_n, we will use the backward identity I_n^B and the diagonal matrix $I_n^\pm$ defined by

$$I_n^B \equiv \begin{pmatrix} & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & \end{pmatrix} \in \mathbb{R}^{n\times n}\,, \quad I_n^\pm \equiv \operatorname{diag}(1, -1, \ldots, (-1)^{n-1})\,. \tag{2.6}$$

As explained above, the singular case ($\lambda = 0$) is uninteresting, so we only consider the *nonsingular case*, i.e. $\lambda \neq 0$. Each $\lambda \in \mathbb{C}$ can be written as $\lambda = |\lambda| e^{\mathbf{i}\alpha}$, and it holds that

$$J_\lambda = e^{\mathbf{i}\alpha} U J_{|\lambda|} U^H\,, \qquad U \equiv \operatorname{diag}(e^{\mathbf{i}\alpha}, e^{\mathbf{i}2\alpha}, \ldots, e^{\mathbf{i}n\alpha})\,. \tag{2.7}$$

Since J_λ is unitarily similar to $J_{|\lambda|}$, and the values of the approximation problems we deal with are unitarily invariant, it suffices to consider *real and positive* λ. All results can be easily extended to all $\lambda \in \mathbb{C}$ using the unitary similarity transformation defined by (2.7). Since $d(J_\lambda) = n$, we will consider $k = 1, \ldots, n-1$, so that $0 < \psi_k(J_\lambda) \leq \varphi_k(J_\lambda)$, and the corresponding ideal GMRES polynomials are well defined in the sense of Definition 2.2.

As mentioned in the Introduction, the polynomial numerical hull (1.5) appears to be useful in the analysis of ideal GMRES. As shown in [2], for each $k = 1, \ldots, n-1$, $\mathcal{H}_k(J_\lambda)$ is a circle around the eigenvalue λ with some radius $\varrho_{k,n}$, where

$$0 < \varrho_{n-1,n} < \cdots < \varrho_{1,n} < 1,$$

and $\varrho_{k,n}$ is independent of the eigenvalue λ. In particular, the authors of [2] concentrate on determining the radii $\varrho_{1,n}$ and $\varrho_{n-1,n}$. Since $\mathcal{H}_1(J_\lambda)$ is equal to the field of values of J_λ, it holds that

$$\varrho_{1,n} = \cos\left(\tfrac{\pi}{n+1}\right)\,, \tag{2.8}$$

cf. [2, p. 235]. The problem of determining $\varrho_{n-1,n}$ is equivalent to a classical problem in complex approximation theory, closely related to the Carathéodory-Fejér interpolation problem. Using this connection it is shown in [2, p. 238], that $\varrho_{n-1,n}$ is a solution of a certain nonlinear problem and can be bounded by

$$1 - \tfrac{\log(2n)}{n} \leq \varrho_{n-1,n} \leq 1 - \tfrac{\log(2n)}{n} + \tfrac{\log(\log(2n))}{n}\,. \tag{2.9}$$

Continuing this work, Greenbaum [5, p. 88] combines (1.6), (2.4) and results of [2] to prove that for $k = 1, \dots, n-1$,

$$\varrho_{k,n}^k \lambda^{-k} \le \varphi_k(J_\lambda) \;\le\; \lambda^{-k} \quad \text{for} \quad \lambda \ge \varrho_{k,n}\,, \tag{2.10}$$

$$\psi_k(J_\lambda) \;=\; \varphi_k(J_\lambda) \;=\; 1 \quad\Longleftrightarrow\quad \lambda \le \varrho_{k,n}\,. \tag{2.11}$$

The upper bound on $\varphi_k(J_\lambda)$ in (2.10) can be replaced by 1 if $\lambda \le 1$. The lower bound in (2.10) is a special case of the general lower bound (1.6) on the ideal GMRES approximation based on the polynomial numerical hull. The closeness of this lower bound is examined in Section 6 below.

We point out that the lower bound on $\varrho_{n-1,n}$ in (2.9) approaches 1 as $n \to \infty$. Hence the equivalence (2.11) implies that for each λ with $0 < |\lambda| < 1$, there exists a positive integer $n = n_\lambda$ such that for the $n \times n$ Jordan block J_λ, $\psi_{n-1}(J_\lambda) = \varphi_{n-1}(J_\lambda) = 1$. In other words, both worst-case and ideal GMRES stagnate completely for each Jordan block J_λ corresponding to an eigenvalue λ inside the unit circle, provided that J_λ is sufficiently large. The more interesting cases are therefore the Jordan blocks J_λ with $|\lambda| \ge 1$.

3. Worst-case and ideal GMRES for $k < n/2$. In this section we show that if $|\lambda|$ is outside a small interval around one, then J_λ is ideal of degree k for $1 \le k < n/2$. We start with a general characterization of the radius $\varrho_{k,n}$ of the polynomial numerical hull of degree k of J_λ.

LEMMA 3.1. *A positive real number ϱ satisfies $\varrho \le \varrho_{k,n}$ if and only if there exists a real unit norm vector b such that*

$$b^T E_n^j b = (-\varrho)^j, \qquad j = 1, \dots, k. \tag{3.1}$$

Proof. A positive real number ϱ satisfies $\varrho \le \varrho_{k,n}$ if and only if an $n \times n$ Jordan block J_ϱ satisfies $\psi_k(J_\varrho) = \varphi_k(J_\varrho) = 1$, cf. (2.11). This is equivalent with the existence of a real unit norm vector b such that

$$b \perp J_\varrho \mathcal{K}_k(J_\varrho, b) \;=\; \mathcal{K}_k(E_n, J_\varrho b) \;=\; \mathcal{K}_k(E_n, \varrho\, b + E_n b). \tag{3.2}$$

But the orthogonality of b to the space $\mathcal{K}_k(E_n, \varrho\, b + E_n b)$ means that

$$0 \;=\; b^T\,(\varrho\, E_n^{j-1} b + E_n^j b), \quad j = 1, \dots, k,$$

which can be written in the equivalent form (3.1). ∎

THEOREM 3.2. *An $n \times n$ Jordan block J_λ with $\lambda > 0$ is ideal of degree k with the kth ideal GMRES polynomial given by*

$$q(z) = (1 - \lambda^{-1} z)^k \tag{3.3}$$

if and only if $1 \le k < n/2$ *and* $\lambda \;\ge\; \varrho_{k,n-k}^{-1}$.

Proof. Since

$$q(J_\lambda) = (-1)^k \lambda^{-k} E_n^k,$$

each $w \in \Sigma(q(J_\lambda))$ has to be of the form

$$w = (0, \dots, 0, b_1, \dots, b_{n-k})^T, \tag{3.4}$$

and hence

$$q(J_\lambda)w = (-1)^k \lambda^{-k} (b_1, \ldots, b_{n-k}, 0, \ldots, 0)^T.$$

Using Lemma 2.4 and the previous observation, J_λ is ideal of degree k and q is both the worst-case and ideal GMRES polynomial if and only if there exists a unit norm vector w of the form (3.4) such that

$$q(J_\lambda)w \perp J_\lambda \mathcal{K}_k(J_\lambda, w) = \mathcal{K}_k(E_n, \lambda w + E_n w), \tag{3.5}$$

i.e.

$$\lambda\, w^T (E_n^{j-1})^T E_n^k w + w^T (E_n^j)^T E_n^k w = 0, \qquad j = 1, \ldots, k. \tag{3.6}$$

Since w has the special form (3.4), it holds that $w^T (E_n^j)^T E_n^k w = b^T E_{n-k}^{k-j} b$, where $b \equiv (b_1, \ldots, b_{n-k})^T$. Then, (3.6) is equivalent to

$$\lambda^{-1} b^T E_{n-k}^{k-j} b + b^T E_{n-k}^{k-j+1} b = 0, \qquad j = 1, \ldots, k. \tag{3.7}$$

Writing the equations (3.7) in the reverse order (for $j = k, \ldots, 1$), we obtain

$$\lambda^{-1}\, b^T E_{n-k}^{j-1} b + b^T E_{n-k}^{j} b = 0, \qquad j = 1, \ldots, k, \tag{3.8}$$

or, equivalently,

$$b^T E_{n-k}^{j} b = (-\lambda^{-1})^j, \qquad j = 1, \ldots, k. \tag{3.9}$$

Clearly, if $k \geq n/2$, then E_{n-k}^k is the zero matrix. In this case at least one of the conditions in (3.9) takes the form $0 = (-\lambda)^k$, and the system (3.9) does not have a solution for any positive λ. On the other hand, for $1 \leq k < n/2$, the system (3.9) has a solution if and only if $\lambda^{-1} \leq \varrho_{k,n-k}$, cf. Lemma 3.1, which completes the proof. ∎

We summarize what we have seen so far in the following corollary.

Corollary 3.3. *For an $n \times n$ Jordan block J_λ with eigenvalue $\lambda \in \mathbb{C}$, and $1 \leq k < n/2$ the following hold:*
1. *If $|\lambda| \leq \varrho_{k,n}$, then J_λ is ideal of degree k with $\psi_k(J_\lambda) = \varphi_k(J_\lambda) = 1$.*
2. *If $|\lambda| \geq \varrho_{k,n-k}^{-1}$, then J_λ is ideal of degree k with $\psi_k(J_\lambda) = \varphi_k(J_\lambda) = \lambda^{-k}$.*

The first item already was shown in (2.11), the second follows from Theorem 3.2. In summary, for $1 \leq k < n/2$ and $|\lambda| \geq 0$, we completely understand the situation except for the cases

$$\varrho_{k,n} \;<\; |\lambda| \;<\; \varrho_{k,n-k}^{-1}\,. \tag{3.10}$$

The lower bound in (3.10) is bounded from below by $1/2$, and it approaches 1 for $n \to \infty$, while the upper bound in (3.10) is bounded from above by 2.

4. Structure of the ideal GMRES residual matrices for a Jordan block. In this section we analyze the special structure of the ideal GMRES residual matrices for a Jordan block, which we originally discovered numerically when experimenting with the semidefinite programming package SDPT3 [18]. Since the development below is quite technical, we start with a high-level description of a simple example.

Consider the 6×6 Jordan block J_λ with $\lambda = 1$. As shown below, its second, third and fourth ideal GMRES residual matrices are upper triangular Toeplitz matrices of the form

$$\underbrace{\begin{pmatrix} \bullet & \circ & \bullet & & & \\ & \bullet & \circ & \bullet & & \\ & & \bullet & \circ & \bullet & \\ & & & \bullet & \circ & \bullet \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{p_*^{(2)}(J_1)}, \quad \underbrace{\begin{pmatrix} \bullet & \circ & \circ & \bullet & & \\ & \bullet & \circ & \circ & \bullet & \\ & & \bullet & \circ & \circ & \bullet \\ & & & \bullet & \circ & \circ \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{p_*^{(3)}(J_1)}, \quad \underbrace{\begin{pmatrix} \bullet & \circ & \bullet & \circ & \bullet & \\ & \bullet & \circ & \bullet & \circ & \bullet \\ & & \bullet & \circ & \bullet & \circ \\ & & & \bullet & \circ & \bullet \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{p_*^{(4)}(J_1)},$$

where "$\bullet$" stands for a nonzero entry and "$\circ$" represents a zero entry. It is easy to see that there exist permutation matrices P_2, P_3 and P_4 that transform $p_*^{(2)}(J_1)$, $p_*^{(3)}(J_1)$ and $p_*^{(4)}(J_1)$ into block diagonal matrices with upper triangular Toeplitz blocks,

$$\underbrace{\begin{pmatrix} \bullet & \bullet & & & & \\ & \bullet & \bullet & & & \\ & & \bullet & & & \\ & & & \bullet & \bullet & \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_2^T p_*^{(2)}(J_1) P_2}, \quad \underbrace{\begin{pmatrix} \bullet & \bullet & & & & \\ & \bullet & & & & \\ & & \bullet & \bullet & & \\ & & & \bullet & & \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_3^T p_*^{(3)}(J_1) P_3}, \quad \underbrace{\begin{pmatrix} \bullet & \bullet & \bullet & & & \\ & \bullet & \bullet & & & \\ & & \bullet & & & \\ & & & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_4^T p_*^{(4)}(J_1) P_4}.$$

Since the transformation $p_*^{(k)}(J_1) \to P_k^T p_*^{(k)}(J_1) P_k$ is orthogonal, and all diagonal blocks of $P_k^T p_*^{(k)}(J_1) P_k$ are equal, the ideal GMRES approximation $\|p_*^{(k)}(J_1)\|$ equals the norm of any diagonal block of $P_k^T p_*^{(k)}(J_1) P_k$.

These observations are the key to analyzing the kth and $(n-k)$th ideal GMRES approximations for J_1 and, more generally, for any Jordan block J_λ, when k divides n. The following lemma formalizes the just described orthogonal transformation and shows the connection between the singular value decompositions of $p_*^{(k)}(J_\lambda)$ and of a diagonal block of $P_k^T p_*^{(k)}(J_\lambda) P_k$.

LEMMA 4.1. *Let n and k be positive integers, $n > k$, and let d be their greatest common divisor. Define $m \equiv n/d$ and $\ell = k/d$. Consider the $m \times m$ upper triangular Toeplitz matrix B,*

$$B \equiv \sum_{j=0}^{\ell} b_j E_m^j, \quad \textit{and let} \quad B = USV^T \tag{4.1}$$

be its singular value decomposition. Then the singular value decomposition of the $n\times n$ matrix G,

$$G \equiv \sum_{j=0}^{\ell} b_j E_n^{jd} \quad \textit{is given by} \quad G = (U \otimes I_d)(S \otimes I_d)(V \otimes I_d)^T . \tag{4.2}$$

Proof. Define the $n \times n$ matrix P by $P \equiv [I_m \otimes e_1, \ldots, I_m \otimes e_d]$, then

$$P^T G P = I_d \otimes B = I_d \otimes (USV^T) = (I_d \otimes U)(I_d \otimes S)(I_d \otimes V)^T ,$$

and hence

$$\begin{aligned} G &= P(I_d \otimes U)(I_d \otimes S)(I_d \otimes V)^T P^T \\ &= [P(I_d \otimes U)P^T]\,[P(I_d \otimes S)P^T]\,[P(I_d \otimes V)P^T]^T \\ &= (U \otimes I_d)(S \otimes I_d)(V \otimes I_d)^T . \end{aligned}$$

In the last equation we have used [8, Corollary 4.3.10]. □

As outlined above, our strategy is as follows: Having an ideal GMRES residual matrix G of the special form (4.2), we can find a permutation matrix P such that $P^T G P = I \otimes B$ (where I and B have the appropriate sizes), and then investigate the norm and properties of G through the norm and properties of the block B.

THEOREM 4.2. *Let n and k be positive integers, $n > k$, and let d be their greatest common divisor. Let $\lambda > 0$ and define $m \equiv n/d$, $\ell \equiv k/d$,*

$$J_\lambda \equiv \lambda I_n + E_n\,, \qquad J_\mu \equiv \mu I_m + E_m\,, \qquad \mu \equiv \lambda^d.$$

Suppose that the ℓth ideal GMRES polynomial $p_*^{(\ell)}$ of J_μ is of the form

$$p_*^{(\ell)}(z) = \sum_{j=0}^{\ell} c_j(\mu - z)^j\,. \tag{4.3}$$

If J_μ is ideal of degree ℓ, then J_λ is ideal of degree k, and

$$\psi_\ell(J_\mu) = \varphi_\ell(J_\mu) \quad = \quad \psi_k(J_\lambda) = \varphi_k(J_\lambda).$$

Moreover, the kth ideal GMRES polynomial $p_*^{(k)}$ of J_λ is given by

$$p_*^{(k)}(z) = \sum_{j=0}^{\ell} c_j(\lambda - z)^{jd}\,. \tag{4.4}$$

Proof. Given the ℓth ideal GMRES polynomial $p_*^{(\ell)} \in \pi_\ell$ of J_μ as in (4.3), we define the polynomial

$$q(z) \;\equiv\; \sum_{j=0}^{\ell} c_j(\lambda - z)^{jd} \;\in\; \pi_k\,. \tag{4.5}$$

Our goal is to show that this polynomial q, which is equal to $p_*^{(k)}$ in (4.4), is the kth ideal GMRES polynomial of J_λ. We will show this by constructing a unit norm vector $b \in \Sigma(q(J_\lambda))$, such that the condition (2.1) is satisfied.

From

$$p_*^{(\ell)}(J_\mu) \;=\; \sum_{j=0}^{\ell} c_j(-E_m)^j, \qquad q(J_\lambda) \;=\; \sum_{j=0}^{\ell} c_j(-E_n)^{jd}\,, \tag{4.6}$$

we see that the matrices $p_*^{(\ell)}(J_\mu)$ and $q(J_\lambda)$ have a similar structure as the matrices B and G, respectively, in Lemma 4.1 (up to the sign in case d is even).

By assumption, $\psi_\ell(J_\mu) = \varphi_\ell(J_\mu) > 0$, and hence by Lemma 2.4 there exists a unit norm vector $w \in \Sigma(p_*^{(\ell)}(J_\mu))$, such that

$$p_*^{(\ell)}(J_\mu)w \;\perp\; J_\mu \mathcal{K}_\ell(J_\mu, w)\,. \tag{4.7}$$

Define $S_\mu \in \mathbb{R}^{m\times m}$, $v \in \mathbb{R}^m$, and $B \in \mathbb{R}^{m\times m}$ by

$$S_\mu \;\equiv\; \begin{cases} J_\mu\,, \\ I_m^{\pm}\, J_\mu\, I_m^{\pm}, \end{cases} \qquad v \equiv \begin{cases} w\,, & \text{if } d \text{ is odd,} \\ I_m^{\pm}\, w, & \text{if } d \text{ is even,} \end{cases} \tag{4.8}$$

$$B \;\equiv\; p_*^{(\ell)}(S_\mu). \tag{4.9}$$

Then it easily follows that

$$Bv \;\perp\; S_\mu \mathcal{K}_\ell(S_\mu, v)\,, \tag{4.10}$$

and $v \in \Sigma(B)$. Since B is a Toeplitz matrix, the matrix $I_m^B B$ is symmetric, and hence unitarily diagonalizable, $I_m^B B = V\Lambda V^T$. Therefore, there exists a diagonal matrix $\hat{I}_m^{\pm}$ having entries 1 and -1 on its diagonal, such that

$$B \;=\; (I_m^B V \hat{I}_m^{\pm})\,(\hat{I}_m^{\pm}\Lambda)\,V^T$$

is the singular value decomposition of B. If $z \in \Sigma(B)$ is a right singular vector, then the corresponding left singular vector is given either by $I_m^B z$ or by $-I_m^B z$. Since $v \in \Sigma(B)$, we can decompose this vector as $v = v^+ + v^-$. Here v^+ resp. v^- are the orthogonal projections of v onto the space spanned by right singular vectors $z \in \Sigma(B)$ with the corresponding left singular vector equal to $I_m^B z$ resp. $-I_m^B z$.

Denoting by δ the maximal singular value of $p_*^{(\ell)}(J_\mu)$,

$$Bv = \delta\, I_m^B (v^+ - v^-)\,, \quad \text{and} \quad \delta \equiv \|p_*^{(\ell)}(J_\mu)\| = \|B\| = \|q(J_\lambda)\|\,, \tag{4.11}$$

where we have applied Lemma 4.1 to obtain the last equality.

Since $v \in \Sigma(B)$, Lemma 4.1 implies that $v \otimes e_j \in \Sigma(q(J_\lambda))$, where e_j denotes the jth standard basis vector for $j = 1, \ldots, d$. Now define $e_\lambda \equiv [1, -\lambda, \ldots, (-\lambda)^{d-1}]^T$, and

$$b \;\equiv\; \gamma \sum_{j=1}^{d} (-\lambda)^{j-1}\, v \otimes e_j \;=\; \gamma\,(v \otimes e_\lambda)\,, \tag{4.12}$$

where γ is chosen so that $\|b\| = 1$. Clearly, $b \in \Sigma(q(J_\lambda))$, and b can be decomposed as

$$b = b^+ + b^-, \qquad b^+ \equiv \gamma\,(v^+ \otimes e_\lambda), \qquad b^- \equiv \gamma\,(v^- \otimes e_\lambda)\,,$$

with $q(J_\lambda)b^+ = \delta I_n^B b^+$, $q(J_\lambda)b^- = -\delta I_n^B b^-$. Hence, using the first expression in (4.11),

$$\begin{aligned} q(J_\lambda)b &= \gamma q(J_\lambda)\,(b^+ + b^-) = \gamma\delta I_n^B (b^+ - b^-) \\ &= \gamma\delta I_n^B\,((v^+ - v^-) \otimes e_\lambda) = \gamma\delta\,((I_m^B (v^+ - v^-)) \otimes (I_d^B e_\lambda)) \\ &= \gamma\,((Bv) \otimes (I_d^B e_\lambda))\,. \end{aligned} \tag{4.13}$$

We next show that

$$q(J_\lambda)b \;\perp\; J_\lambda^j b\,, \quad j = 1, \ldots, k\,, \tag{4.14}$$

i.e. that q is a GMRES polynomial for J_λ and the initial vector b. Since

$$\operatorname{span}\{J_\lambda b, \ldots, J_\lambda^k b\} \;=\; \operatorname{span}\{E_n^0 J_\lambda b, \ldots, E_n^{k-1} J_\lambda b\}\,, \tag{4.15}$$

the relation (4.14) holds if and only if

$$q(J_\lambda)b \;\perp\; E_n^j J_\lambda b\,, \quad j = 0, \ldots, k-1\,. \tag{4.16}$$

Let us decompose the index j as

$$j = sd + t, \qquad s = 0, \ldots, l-1, \qquad t = 0, \ldots, d-1. \tag{4.17}$$

An elementary computation shows that

$$J_\lambda b \;=\; \gamma\, J_\lambda (v \otimes e_\lambda) = \;\gamma\, ((S_\mu v) \otimes e_d)\,.$$

Multiplication of $J_\lambda b$ from the left by E_n^j shifts all entries of $J_\lambda b$ upwards by j positions. Using (4.17), $E_n^j J_\lambda b$ can be written as

$$E_n^j J_\lambda b \;=\; \gamma\, E_n^{sd}((S_\mu v) \otimes e_{d-t}) = \gamma\, ((E_m^s S_\mu v) \otimes e_{d-t})\,. \tag{4.18}$$

Now from (4.13) and (4.18) we obtain

$$\begin{aligned}(q(J_\lambda)b)^T (E_n^j J_\lambda b) &= \gamma^2\, ((Bv) \otimes (I_d^B e_\lambda))^T\, ((E_m^s S_\mu v) \otimes e_{d-t}) \\ &= \gamma^2\, [(Bv)^T E_m^s S_\mu v]\, [e_\lambda^T I_d^B e_{d-t}].\end{aligned}$$

Similar as in (4.15), $E_m^s S_\mu v \in S_\mu \mathcal{K}_\ell(S_\mu, v)$ for $s = 0, \ldots, l-1$. Since Bv is orthogonal to $S_\mu \mathcal{K}_\ell(S_\mu, v)$, cf. (4.10), it holds that $(Bv)^T E_m^s S_\mu v = 0$ for $s = 0, \ldots, l-1$. In other words, we just proved (4.14).

Summarizing, q is the kth GMRES polynomial for the matrix J_λ and the initial vector $b \in \Sigma(q(J_\lambda))$. Using Lemma 2.4, it holds that $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$ and, therefore, q is the kth ideal GMRES polynomial of J_λ. Moreover, Lemma 4.1 implies that the ideal GMRES residual matrices $p_*^{(\ell)}(J_\mu)$ and $p_*^{(k)}(J_\lambda)$ have the same norm and thus $\psi_\ell(J_\mu) = \varphi_\ell(J_\mu) = \psi_k(J_\lambda) = \varphi_k(J_\lambda)$. □

Note that the integers ℓ and m defined in Theorem 4.2 are relatively prime. The assertion of this theorem is quite tricky, so some explanation is appropriate. Suppose we know that an $m \times m$ Jordan block J_μ is ideal of degree ℓ, where ℓ and m are relatively prime. Then by Theorem 4.2, an $n \times n$ Jordan block J_λ is ideal of degree k, where $n \equiv dm$, $k \equiv d\ell$, $\lambda \equiv \mu^d$, and d is *any* positive integer. Therefore, to prove that any Jordan block is ideal, it would be sufficient to show that any Jordan block is ideal of degree k whenever k and the size of the Jordan block are relatively prime; all the other cases are then covered by Theorem 4.2. In other words, Theorem 4.2 reduces the question of idealness of Jordan blocks to block sizes n and steps k, where k and n are relatively prime.

Example 1. Consider the 20×20 Jordan block J_1. In Fig 4.1 we plot the multiplicity of the maximal singular value of $p_*^{(k)}(J_1)$ for $k = 1, \ldots, 19$. Apparently, the multiplicity is equal to the greatest common divisor of n and k. In particular, at steps k such that k and n are relatively prime, the maximal singular value of $p_*^{(k)}(J_1)$ is simple. (The same phenomenon can be observed numerically also for other choices of n.) By the second item in Lemma 2.3, J_1 is ideal of degree k in the steps where k and n are relatively prime. Then Theorem 4.2 implies that J_1 is ideal. □

Theorem 4.2 also allows us to prove the following result about the radii of the polynomial numerical hulls of Jordan blocks.

Theorem 4.3. *Let n and k be positive integers, $n > k$, and let d be their greatest common divisor. Define $m \equiv n/d$, $\ell \equiv k/d$. Then the radius $\varrho_{k,n}$ of the kth polynomial numerical hull of an $n \times n$ Jordan block satisfies*

$$\varrho_{k,n} \;=\; \varrho_{\ell,m}^{1/d}\,. \tag{4.19}$$

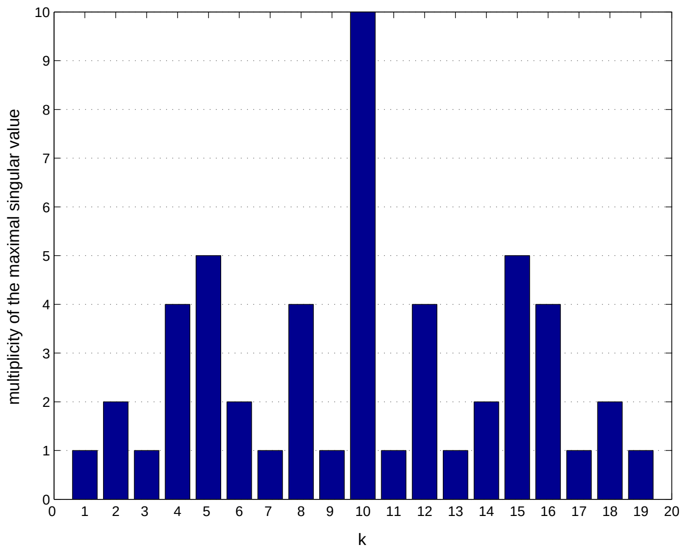

FIG. 4.1. *Multiplicity of the maximal singular value of* $p_*^{(k)}(J_1)$ *for the* 20×20 *Jordan block* J_1 *and* $k = 1, \ldots, 19$.

Proof. Let $\lambda > 0$ and consider Jordan blocks

$$J_\lambda \equiv \lambda I_n + E_n\,, \qquad J_\mu \equiv \mu I_m + E_m\,, \qquad \mu \equiv \lambda^d.$$

We prove the following equivalence

$$\mu \leq \varrho_{\ell,m} \overset{\mathbf{A}}{\iff} \psi_\ell(J_\mu) = \varphi_\ell(J_\mu) = 1 \overset{\mathbf{B}}{\iff} \psi_k(J_\lambda) = \varphi_k(J_\lambda) = 1 \overset{\mathbf{C}}{\iff} \lambda \leq \varrho_{k,n}.$$

The equivalences **A** and **C** follow from (2.11), so we only have to prove the equivalence **B**. From Theorem 4.2,

$$\psi_\ell(J_\mu) = \varphi_\ell(J_\mu) = 1 \quad \Longrightarrow \quad \psi_k(J_\lambda) = \varphi_k(J_\lambda) = 1.$$

On the other hand, suppose that $\psi_k(J_\lambda) = \varphi_k(J_\lambda) = 1$. Consider the polynomial $p_*^{(\ell)}$ of the form (4.3). Then, similarly as in the proof of Theorem 4.2, the polynomial q defined by (4.5) satisfies $q \in \pi_k$ and $\|q(J_\lambda)\| = \|p_*^{(\ell)}(J_\mu)\|$, cf. (4.11). Now if $\varphi_k(J_\mu) = \|p_*^{(\ell)}(J_\mu)\| < 1$, then $\|q(J_\lambda)\| < 1 = \varphi_k(J_\lambda)$, which contradicts the optimality property of the kth ideal GMRES polynomial $p_*^{(k)}$ of J_λ. Therefore $\varphi_k(J_\mu) = 1$, which implies that $\psi_k(J_\mu) = 1$, cf. (2.4), and thus **B** must hold.

Consequently, for each $\lambda > 0$, $\lambda^d \leq \varrho_{\ell,m} \iff \lambda \leq \varrho_{k,n}$, which implies (4.19). ◻

COROLLARY 4.4. *Consider an* $n \times n$ *Jordan block* J_λ *with* $\lambda > 0$. *Let* $k < n$ *be a positive integer dividing* n. *Then* $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$, *and if* $\lambda \geq \varrho_{k,n}$, *then*

$$\lambda^{-k} \cos\left(\frac{\pi}{n/k+1}\right) \;\leq\; \varphi_k(J_\lambda) \;\leq\; \lambda^{-k}\,. \tag{4.20}$$

The k*th ideal GMRES polynomial* $p_*^{(k)}$ *of* J_λ *is of the form*

$$p_*^{(k)}(z) = c_0 + c_1\,(\lambda - z)^k\,, \tag{4.21}$$

where c_0 and c_1 are the coefficients of the first ideal GMRES polynomial (4.3) of the $\frac{n}{k} \times \frac{n}{k}$ Jordan block J_{λ^k}. Moreover, it holds that

$$\varrho_{k,n} = \varrho_{1,n/k}^{1/k} = \left[\cos\left(\tfrac{\pi}{n/k+1}\right)\right]^{1/k}. \tag{4.22}$$

Proof. All results follow from Theorem 4.2 and Theorem 4.3. If k divides n, then $d = k$ is their greatest common divisor, and $m = n/k$, $\ell = 1$. For $\ell = 1$, the assumption $\psi_\ell(J_\mu) = \varphi_\ell(J_\mu) > 0$ in Theorem 4.2 is satisfied and therefore $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$. In (4.22) we used Theorem 4.3 and the explicit form of the radius $\varrho_{1,n/k}$, cf. (2.8). The bound (4.20) is just the bound (2.10), where for $\varrho_{k,n}$ we substituted its exact value on the right hand side of (4.22). ▯

If k divides n, then $p_*^{(k)}(z) = 1$ for $\lambda \le \varrho_{n,k}$ and $p_*^{(k)}(z) = c_0 + c_1(\lambda - z)^k$ for $\lambda > \varrho_{n,k}$. For $\lambda \ge \varrho_{k,n-k}^{-1}$ we know that $c_0 = 0$ and $c_1 = \lambda^{-k}$, cf. Theorem 3.2. Moreover, since $\psi_k(J_\lambda) = \varphi_k(J_\lambda)$, it follows from (2.11) and Theorem 3.2 that $c_0 \neq 0$ and $c_1 \neq 0$ whenever $\varrho_{k,n} < \lambda < \varrho_{k,n-k}^{-1}$. Then, from the form of the kth ideal GMRES polynomial (4.21) it is easy to see that the k roots of $p_*^{(k)}$ are uniformly distributed on the circle around λ with radius $|c_0/c_1|^{1/k}$.

EXAMPLE 2. Consider an $n \times n$ Jordan block J_λ with $\lambda > 0$, n even and $k = n/2$. This gives $d = n/2$, $m = 2$, $\ell = 1$, and $\mu = \lambda^{n/2}$ in Theorem 4.2. Since for the 2×2 Jordan block J_μ, $\psi_1(J_\mu) = \varphi_1(J_\mu) > 0$, Theorem 4.2 implies that $\psi_1(J_\mu) = \varphi_1(J_\mu) = \psi_{n/2}(J_\lambda) = \varphi_{n/2}(J_\lambda)$. Theorem 4.3 shows that

$$\varrho_{n/2,n} = \varrho_{1,2}^{1/k} = 2^{-2/n}. \tag{4.23}$$

Moreover, by a direct computation of the first ideal GMRES approximation for the 2×2 Jordan block J_μ with $\mu = \lambda^{n/2}$, we obtain that for $\lambda \ge 2^{-2/n}$,

$$c_0 = \frac{2}{4\lambda^n + 1}, \qquad c_1 = \frac{1}{\lambda^{n/2}} \frac{4\lambda^n - 1}{4\lambda^n + 1}, \qquad \varphi_{n/2}(J_\lambda) = \frac{4\lambda^{n/2}}{4\lambda^n + 1}. \tag{4.24}$$

Using (2.10) and the fact that $\varrho_{k,n}^k \ge \varrho_{n/2,n}^k = 2^{-2k/n} \ge 2^{-1}$ for $k \le n/2$, we get the bound

$$\frac{1}{2}\lambda^{-k} \le \varphi_k(J_\lambda) \le \lambda^{-k}, \qquad k \le n/2. \tag{4.25}$$

▯

5. The next-to-last worst-case and ideal GMRES approximations. In this section we consider the $(n-1)$st worst-case and ideal GMRES approximations for an $n \times n$ Jordan block J_λ with $\lambda > 0$. Our main result, stated in Theorem 5.5 below, is that $\psi_{n-1}(J_\lambda) = \varphi_{n-1}(J_\lambda)$ for $\lambda \ge 1$. We also give an explicit expression for $\varphi_{n-1}(J_\lambda)$ in terms of the eigenvalue λ. The proof of this result will make use of three technical lemmas. To simplify the notation, we define the vector

$$e_\lambda^{(n)} \equiv I_n^{\pm}[1, \lambda, \ldots, \lambda^{n-1}]^T$$

and the Hankel matrix

$$H(v_1, \ldots, v_n) \equiv \begin{pmatrix} v_1 & v_2 & \cdots & v_n \\ v_2 & & \iddots & \\ \vdots & \iddots & & \\ v_n & & & \end{pmatrix}. \tag{5.1}$$

The first lemma is a slight reformulation of [12, Corollary 2.2].

LEMMA 5.1. *Consider the linear algebraic system $J_\lambda x = b$, with an $n \times n$ Jordan block J_λ, and a right hand side vector $b = [b_1, \ldots, b_n]^T$ such that $b_n \neq 0$. If $x_0 = 0$, then the $(n-1)$st GMRES residual r_{n-1} is uniquely determined by the linear system*

$$\|r_{n-1}\|^{-2} \, H(b_1, \ldots, b_n) \, r_{n-1} \; = \; e_\lambda^{(n)} . \tag{5.2}$$

LEMMA 5.2. *Let $\lambda > 0$ be given and let $b \in \mathbb{R}^n$ be the unit norm vector*

$$b \; \equiv \; (-1)^{n-1} \|\xi\|^{-1} \, I_n^B \xi \, , \tag{5.3}$$

where $\xi = [\xi_1, \ldots, \xi_n]^T$ has the components

$$\xi_{i+1} \; = \; \lambda^{\frac{n-1}{2} - i} \frac{(-1)^i}{4^i} \binom{2i}{i}, \qquad i = 0, \ldots, n-1. \tag{5.4}$$

Then the $(n-1)$st GMRES residual r_{n-1} for the $n \times n$ Jordan block J_λ and the initial vector b is given by $r_{n-1} = \|\xi\|^{-3} \, \xi$, and hence

$$\|r_{n-1}\| = \|\xi\|^{-2} = \frac{1}{\lambda^{n-1}} \left[\sum_{i=0}^{n-1} (4\lambda)^{-2i} \binom{2i}{i}^2 \right]^{-1} . \tag{5.5}$$

Proof. Since the last component of $b = (-1)^{n-1} \|\xi\|^{-1} [\xi_n, \ldots, \xi_1]^T$ is nonzero, Lemma 5.1 implies that the $(n-1)$st GMRES residual for J_λ and b satisfies

$$\frac{(-1)^{n-1}}{\|\xi\| \, \|r_{n-1}\|^2} \, \widehat{H} \, r_{n-1} \; = \; e_\lambda^{(n)} \, , \tag{5.6}$$

where $\widehat{H} = H(\xi_n, \ldots, \xi_1)$. Using the definition (5.4), the numbers ξ_{i+1} satisfy for $j = 0, \ldots, n-1$,

$$\sum_{i=0}^{j} \xi_{i+1} \xi_{j-i+1} = \frac{(-1)^j}{4^j} \lambda^{n-j-1} \sum_{i=0}^{j} \binom{2i}{i} \binom{2(j-i)}{j-i} = (-1)^j \lambda^{n-j-1}.$$

In the last equality we use the fact that the sum of the products of the given binomial coefficients is equal to 4^j, see e.g. [15, p. 44]. The n previous equations can be written in matrix form as

$$\widehat{H} \, \xi \; = \; (-1)^{n-1} e_\lambda^{(n)} . \tag{5.7}$$

A comparison of (5.7) and (5.6) shows that $\xi = \|r_{n-1}\|^{-2} r_{n-1} \|\xi\|^{-1}$ and, therefore, $\|\xi\|^{-2} = \|r_{n-1}\|$. Finally, $r_{n-1} = \xi \, \|\xi\| \, \|r_{n-1}\|^2 = \xi \, \|\xi\|^{-3}$. A straightforward computation shows that $\|r_{n-1}\|$ is given by (5.5). ∎

REMARK 5.3. *It is not hard to check that ξ_{i+1} defined in (5.4) can be computed by the recurrence*

$$\xi_1 = \lambda^{\frac{n-1}{2}}, \qquad \xi_{i+1} = -\xi_i \lambda^{-1} \frac{2i-1}{2i}, \qquad i = 1, \ldots, n-1. \tag{5.8}$$

LEMMA 5.4. *Let $\lambda > 0$ be given and let $\xi^+ \equiv I_n^{\pm}\xi$, where the vector ξ is defined as in Lemma 5.2. Then there exists an uniquely determined Hankel matrix $\hat{H}$ such that*

$$\xi^+ = \hat{H}\,\xi^+. \tag{5.9}$$

If $\lambda \geq 1$, the matrix $\hat{H}$ is primitive and has only one eigenvalue of maximum modulus. This eigenvalue is equal to 1, and ξ^+ is the corresponding eigenvector.

Proof. First note that since the entries of ξ alternate in sign and $\xi_1 > 0$, all components of $\xi^+ = [\xi_1^+, \ldots, \xi_n^+]^T$ are positive. We are now going to construct the Hankel matrix $\hat{H}$ of the form $\hat{H} = H(h_n, \ldots, h_1)$.

The nth equation in $\xi^+ = \hat{H}\xi^+$ is $h_1\xi_1^+ = \xi_n^+$, i.e. $h_1 = \xi_n^+/\xi_1^+$. Therefore, h_1 is well-defined and positive. Considering the equations $n-1, \ldots, 1$ it is clear that the entries $h_2, \ldots, h_n$ of $\hat{H}$ are uniquely determined.

To show the remaining part of the lemma, we will first prove by induction that for $\lambda \geq 1$, $\hat{H}$ is nonnegative with $h_i > 0$, $i = 1, \ldots, n$. We already know that $h_1 > 0$. Now suppose that $h_1 > 0, \ldots, h_j > 0$ for some $j \geq 1$. The $(n-j)$th equation in $\xi^+ = \hat{H}\xi^+$ is of the form

$$\xi_{n-j}^+ = h_{j+1}\xi_1^+ + \sum_{i=2}^{j+1} h_{j-i+2}\xi_i^+ = h_{j+1}\xi_1^+ + \sum_{i=1}^{j} h_{j-i+1}\xi_{i+1}^+.$$

Using the definitions of ξ_{i+1}^+ and ξ_{i+1}, cf. (5.4) and (5.8), it holds that

$$\xi_{i+1}^+ = \lambda^{-1}\left(\xi_i^+ - \frac{\xi_i^+}{2i}\right)$$

and, therefore,

$$\begin{aligned}\xi_{n-j}^+ &= h_{j+1}\xi_1^+ + \lambda^{-1}\sum_{i=1}^{j} h_{j-i+1}\xi_i^+ - \lambda^{-1}\sum_{i=1}^{j} h_{j-i+1}\frac{\xi_i^+}{2i}\\ &= h_{j+1}\xi_1^+ + \lambda^{-1}\xi_{n-j+1}^+ - \lambda^{-1}\sum_{i=1}^{j} h_{j-i+1}\frac{\xi_i^+}{2i}.\end{aligned}$$

Finally,

$$h_{j+1} = (\xi_1^+)^{-1}\left(\xi_{n-j}^+ - \lambda^{-1}\xi_{n-j+1}^+ + \left[\lambda^{-1}\sum_{i=1}^{j} h_{j-i+1}\frac{\xi_i^+}{2i}\right]\right). \tag{5.10}$$

The term in the square brackets is positive according to the induction hypothesis. Moreover, since the sequence $\xi_1^+, \xi_2^+, \ldots$ is decreasing for $\lambda \geq 1$, it holds that $\xi_{n-j}^+ > \lambda^{-1}\xi_{n-j+1}^+$, i.e. $h_{j+1} > 0$.

Summarizing, $\hat{H}$ is nonnegative and $\xi^+ > 0$ is an eigenvector of $\hat{H}$ corresponding to the eigenvalue 1. Therefore, 1 must be an eigenvalue of maximum modulus [9, Corollary 8.1.30.]. Moreover, since $\hat{H}^2 > 0$, $\hat{H}$ is primitive, cf. [9, Theorem 8.5.2.], and there exists only one eigenvalue of maximum modulus. □

Now we can state and prove the main result of this section.

THEOREM 5.5. *Consider an $n \times n$ Jordan block J_λ with $\lambda \geq 1$. Then the unit norm vector b defined in (5.3)–(5.4) solves the worst-case GMRES approximation problem (1.3) for J_λ and $k = n-1$, and it holds that*

$$\psi_{n-1}(J_\lambda) \;=\; \varphi_{n-1}(J_\lambda) \;=\; \frac{1}{\lambda^{n-1}} \left[\sum_{i=0}^{n-1} (4\lambda)^{-2i} \binom{2i}{i}^2 \right]^{-1} . \tag{5.11}$$

Proof. Consider the $(n-1)$st GMRES residual r_{n-1} for J_λ and the initial vector b defined in (5.3)–(5.4), and denote by p_{n-1} the corresponding GMRES polynomial, i.e.

$$r_{n-1} \;=\; p_{n-1}(J_\lambda)\, b\,. \tag{5.12}$$

Using (5.5), $\|r_{n-1}\|$ is equal to the rightmost expression in (5.11). To prove the assertion it suffices to show that b is a maximal right singular vector of the matrix $p_{n-1}(J_\lambda)$, cf. Lemma 2.4. Since $p_{n-1}(J_\lambda)$ is an upper triangular Toeplitz matrix, the matrix $p_{n-1}(J_\lambda)I_n^{B}$, where I_n^{B} is defined in (2.6), is symmetric, and hence unitarily diagonalizable. Denote its eigendecomposition by $p_{n-1}(J_\lambda)I_n^{B} = UDU^T$, where D is a nonsingular real diagonal matrix, and $U^TU = UU^T = I_n$. Given D, there exists a (uniquely determined) diagonal matrix $\hat{I}_n^{\pm}$ having entries 1 or -1 on its diagonal such that $S \equiv D\hat{I}_n^{\pm}$ is a real diagonal matrix with positive diagonal entries. Then

$$p_{n-1}(J_\lambda) \;=\; U\,(D\hat{I}_n^{\pm})\,(\hat{I}_n^{\pm}U^T I_n^{B}) \;=\; U\,S\,(\hat{I}_n^{\pm}U^T I_n^{B})\,, \tag{5.13}$$

and the rightmost expression is the singular value decomposition of $p_{n-1}(J_\lambda)$.

Substituting (5.3), (5.5) and (5.13) into (5.12), we obtain

$$\xi = (-1)^{n-1}\|\xi\|^2 US\hat{I}_n^{\pm}U^T\xi. \tag{5.14}$$

Similarly as in Lemma 5.4, denote $\xi^+ \equiv I_n^{\pm}\xi > 0$. Multiplying both sides of (5.14) from the left by $I_n^{\pm}$ we receive

$$\xi^+ \;=\; \hat{H}\xi^+\,, \quad \hat{H} \equiv (-1)^{n-1}\|\xi\|^2 (I_n^{\pm}U)\,S\hat{I}_n^{\pm}\,(I_n^{\pm}U)^T = (-1)^{n-1}\|\xi\|^2 (I_n^{\pm}p_{n-1}(J_\lambda)I_n^{B}I_n^{\pm})\,. \tag{5.15}$$

Since $p_{n-1}(J_\lambda)$ is an upper triangular Toeplitz matrix, the expression (5.15) shows that $\hat{H}$ is a Hankel matrix. Considering the eigenvalue decomposition $\hat{H} = Q\Lambda Q^T$ it is easy to see that

$$Q = I_n^{\pm}U, \qquad \Lambda = (-1)^{n-1}\|\xi\|^2 S\hat{I}_n^{\pm}. \tag{5.16}$$

Therefore, the modulus of any eigenvalue of $\hat{H}$ is a $\|\xi\|^2$-multiple of some singular value of $p_{n-1}(J_\lambda)$. Consequently, ξ^+ in (5.15) is an eigenvector corresponding to the eigenvalue of maximum modulus of $\hat{H}$ if and only if b is a right singular vector corresponding to the maximal singular value of $p_{n-1}(J_\lambda)$. By Lemma 5.4, $\hat{H}$ has only one eigenvalue of maximum modulus, and ξ^+ is the corresponding eigenvector. Hence b is the maximal right singular vector of $p_{n-1}(J_\lambda)$. ∎

In the previous theorem we use the assumption $\lambda \geq 1$. It is natural to ask about the relation between worst-case and ideal GMRES for $\varrho_{n-1,n} < \lambda < 1$, and whether for such λ the right hand side of (5.11) still characterizes $\psi_{n-1}(J_\lambda)$ and $\varphi_{n-1}(J_\lambda)$. While

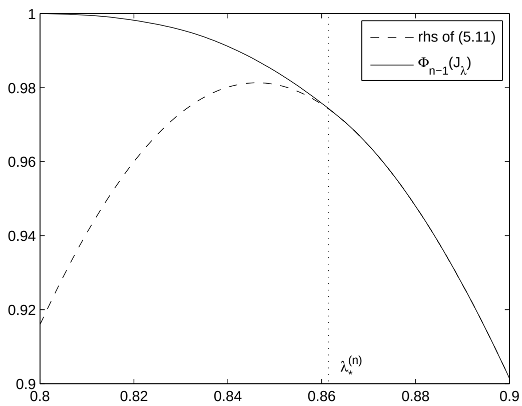

FIG. 5.1. *The right hand side of (5.11) and $\varphi_{n-1}(J_\lambda)$ plotted as a function of λ.*

our numerical experiments predict that $\psi_{n-1}(J_\lambda) = \varphi_{n-1}(J_\lambda)$ also for $\varrho_{n-1,n} < \lambda < 1$, for each integer n there seems to exist a $\lambda_*^{(n)}$, $\varrho_{n-1,n} < \lambda_*^{(n)} < 1$, such that $\varphi_{n-1}(J_\lambda)$ is larger than the right hand side of (5.11) whenever $\lambda < \lambda_*^{(n)}$. In other words, the right hand side of (5.11) does *not* characterize $\psi_{n-1}(J_\lambda)$ and $\varphi_{n-1}(J_\lambda)$ for all $\lambda \geq \varrho_{n-1,n}$. This situation is demonstrated in Fig. 5.1, which shows a numerical experiment with $n = 10$, giving $\varrho_{n-1,n} \approx 0.8$. The dashed line shows the right hand side of (5.11), and the solid line shows the ideal GMRES approximation $\varphi_{n-1}(J_\lambda)$, both as a function of λ.

In the following corollary we combine results of Theorem 4.2, Theorem 4.3 and Theorem 5.5.

COROLLARY 5.6. *Consider an $n \times n$ Jordan block J_λ with $\lambda \geq 1$. If $k < n$ is a positive integer dividing n, then*

$$\psi_{n-k}(J_\lambda) \;=\; \varphi_{n-k}(J_\lambda) \;=\; \frac{1}{\lambda^{n-k}} \left[\sum_{i=0}^{n/k-1} \lambda^{-2ki} 4^{-2i} \binom{2i}{i}^2 \right]^{-1} . \tag{5.17}$$

and

$$\varrho_{n-k,n} = \varrho_{n/k-1,n/k}^{1/k} . \tag{5.18}$$

Proof. The parameters in Theorem 4.2 and Theorem 4.3 are given by $d = k$, $m = n/k$, $\ell = m - 1$ and $\mu = \lambda^k$. Applying Theorem 5.5 to the $m \times m$ Jordan block J_μ we see that $\varphi_{m-1}(J_\mu) = \psi_{m-1}(J_\mu)$, and this quantity is positive. Hence the assumption of Theorem 4.2 is satisfied. Therefore, $\varphi_{m-1}(J_\mu) = \varphi_{n-k}(J_\lambda) = \psi_{n-k}(J_\lambda)$. The value of $\varphi_{m-1}(J_\mu)$ (and also of $\varphi_{n-k}(J_\lambda)$) is given by (5.11), where n and λ have to be replaced by m and λ^k, respectively. □

For example, if $n \geq 4$ is even and $k = 2$, then $m = n/2$ and (5.18) means that $\varrho_{n-2,n} = \varrho_{m-1,m}^{1/2}$. Using a completely different and highly nontrivial proof technique based on complex analysis, the same result has been obtained in [2, p. 241]. Tight bounds on $\varrho_{m-1,m}$ are given by (2.9). Note that for n even and $k = n/2$, it can be easily checked that the rightmost expression in (5.17) agrees with the rightmost expression in (4.24).

6. Polynomial numerical hulls and the ideal GMRES convergence. In [4, Section 3], some numerical examples with nonnormal matrices A of (small) size n are given, for which

$$\varphi_{n-1}(A) \;\le\; C \min_{p\in\pi_{n-1}} \max_{z\in\mathcal{H}_{n-1}(A)} |p(z)|\,,$$

where C is a moderate size constant. It is not shown, however, whether the constant depends on n, or how close the bound (1.6) may be for a general nonnormal matrix A. As we are unaware of any such results in the literature, we here study this question using our above results for an $n\times n$ Jordan block J_λ. We concentrate on the case $\lambda = 1$. We need the following lemma, which can be proven by a straightforward computation; see also [16].

LEMMA 6.1. *The singular value decomposition of the $n\times n$ Jordan block J_1 is given by $J_1 = USV^T$, where*

$$V \;=\; \{v_{ij}\}_{i,j=1}^n\,, \qquad v_{ij} \;=\; \tfrac{2}{\sqrt{2n+1}} \sin\left(\tfrac{2i-1}{2n+1} j\pi\right), \tag{6.1}$$

$$U \;=\; \{u_{ij}\}_{i,j=1}^n\,, \qquad u_{ij} \;=\; \tfrac{2}{\sqrt{2n+1}} \sin\left(\tfrac{2i}{2n+1} j\pi\right), \tag{6.2}$$

$$S \;=\; \mathrm{diag}(\sigma_i)\,, \qquad \sigma_i \;=\; 2\cos\left(\tfrac{i\pi}{2n+1}\right), \quad i = 1,\dots,n. \tag{6.3}$$

THEOREM 6.2. *Consider the $n\times n$ Jordan block J_1, and let $k<n$ be a positive integer dividing n. Then the ideal GMRES approximations $\varphi_k(J_1)$ and $\varphi_{n-k}(J_1)$ are bounded by*

$$\cos\left(\tfrac{\pi}{2n/k}\right) \;\le\; \varphi_k(J_1) \;\le\; \cos\left(\tfrac{\pi}{2n/k+1}\right), \tag{6.4}$$

$$\left[1+\tfrac{1}{2}\log(n/k)\right]^{-1} \;\le\; \varphi_{n-k}(J_1) \;\le\; \left[1+\tfrac{1}{4}\log(n/k)\right]^{-1}. \tag{6.5}$$

Proof. We first prove (6.4). In the notation of Theorem 4.2, $m \equiv n/k$ and $\ell = 1$. Denote by J the $m\times m$ Jordan block with the eigenvalue one. Since $\psi_1(J) = \varphi_1(J) > 0$, Theorem 4.2 implies that $\varphi_k(J_1) = \varphi_1(J)$. It therefore suffices to bound $\|p_*^{(1)}(J)\|$.

The upper bound in (6.4) follows from

$$\|p_*^{(1)}(J)\| \le \|I - \frac{1}{2}J\| = \frac{1}{2}\|J\| = \cos\left(\tfrac{\pi}{2m+1}\right),$$

where $\|J\| = \sigma_1(J)$ is known, cf. Lemma 6.1. For $\omega\in\mathbb{R}$, define the polynomial

$$p_\omega(z) \equiv 1 - \omega z.$$

The norm of $p_\omega(J)$ is the square root of the maximal eigenvalue of

$$p_\omega(J)^T p_\omega(J) = \begin{pmatrix} \gamma_\omega & -\beta_\omega & & \\ -\beta_\omega & \alpha_\omega & \ddots & \\ & \ddots & \ddots & -\beta_\omega \\ & & -\beta_\omega & \alpha_\omega \end{pmatrix},$$

where $\alpha_\omega \equiv \omega^2 + (1-\omega)^2$, $\beta_\omega \equiv (1-\omega)\,\omega$, $\gamma_\omega \equiv (1-\omega)^2$. Next, define the $m\times m$ matrix $T_{\omega,m}$,

$$T_{\omega,m} \equiv \mathrm{tridiag}(-\beta_\omega, \alpha_\omega, -\beta_\omega).$$

Denote the characteristic polynomials of $p_\omega(J)^T p_\omega(J)$ and $T_{\omega,m}$ by

$$\eta_{\omega,m}(z) \equiv \det(zI_m - p_\omega(J)^T p_\omega(J)), \quad \tau_{\omega,m}(z) \equiv \det(zI_m - T_{\omega,n}).$$

It is not hard to see that

$$\eta_{\omega,m}(z) = \tau_{\omega,m}(z) + \omega^2 \tau_{\omega,m-1}(z).$$

Using results of classical polynomial theory, the roots of the polynomials $\tau_{\omega,m}$ and $\tau_{\omega,m-1}$ interlace. Therefore, the maximal root of $\eta_{\omega,m}$ (equal to $\|p_\omega(J)\|^2$) must lay between the maximal roots of $\tau_{\omega,m}$ and $\tau_{\omega,m-1}$ (between the maximal eigenvalues of $T_{\omega,m}$ and $T_{\omega,m-1}$). It is well known that the eigenvalues of $T_{\omega,m-1}$ are given by

$$\lambda^{(j)}_{\omega,m-1} = \alpha_\omega - 2\beta_\omega \cos\left(\tfrac{j\pi}{m}\right), \quad j = 1, \ldots, m-1\,.$$

Considering these eigenvalues as a function of ω, and taking derivatives with respect to ω, shows that the minimum is obtained for $\omega = 1/2$. Therefore,

$$\|p_\omega(J)\|^2 \;\geq\; \max_j \lambda^{(j)}_{\frac{1}{2},m-1} = \frac{1}{2} + \frac{1}{2}\cos\left(\tfrac{\pi}{m}\right) = \cos^2\left(\tfrac{\pi}{2m}\right).$$

Taking square roots, we obtain the lower bound in (6.4).

We next prove (6.5). Using (5.17), the value of $\varphi_{n-k}(J_1)$ is given by

$$\varphi_{n-k}(J_1) = \left[\sum_{i=0}^{m-1} \vartheta_{i+1}\right]^{-1}, \quad \vartheta_{i+1} \equiv \frac{1}{4^{2i}}\binom{2i}{i}^2. \tag{6.6}$$

We first prove that for $j \geq 2$ it holds that

$$\frac{1}{4(j-1)} \leq \vartheta_j \leq \frac{1}{2j}. \tag{6.7}$$

For $j = 2$, $\vartheta_2 = \frac{1}{4}$ and (6.7) holds. Suppose that (6.7) is satisfied for some $j \geq 2$. We show that this inequality holds also for $j+1$. For ϑ_{j+1} we obtain

$$\begin{aligned}\vartheta_{j+1} &= \left(1 - \frac{1}{2j}\right)^2 \vartheta_j \leq \frac{1}{2j}\left(1 - \frac{1}{2j}\right)^2 \frac{j+1}{j+1} \\ &= \frac{1}{2(j+1)}\left(1 - \frac{3}{4j^2} + \frac{1}{4j^3}\right) \leq \frac{1}{2(j+1)}.\end{aligned}$$

Similarly,

$$\vartheta_{j+1} \geq \frac{1}{4(j-1)}\left(1 - \frac{1}{2j}\right)^2 \frac{4j}{4j} = \frac{1}{4j}\left(1 + \frac{1}{4j^2} + \frac{1}{4j^2(j-1)}\right) \geq \frac{1}{4j}\,,$$

and (6.7) holds. Now, we can find upper and lower bounds on $\varphi_{n-k}(J_\lambda)$,

$$\begin{aligned}\sum_{i=0}^{m-1} \vartheta_{i+1} &= 1 + \sum_{j=2}^{m} \vartheta_j \leq 1 + \frac{1}{2}\sum_{j=2}^{m} \frac{1}{j} \leq 1 + \frac{1}{2}\int_1^m x^{-1}\,dx = 1 + \frac{1}{2}\log(m), \\ \sum_{i=0}^{m-1} \vartheta_{i+1} &= 1 + \sum_{j=2}^{m} \vartheta_j \geq 1 + \frac{1}{4}\sum_{j=2}^{m} \frac{1}{j-1} \geq 1 + \frac{1}{4}\int_1^m x^{-1}\,dx = 1 + \frac{1}{4}\log(m).\end{aligned}$$

Using these inequalities and (6.6) we obtain (6.5). □

For simplicity, let us assume that n is even. The bounds (6.4) and (6.5) predict that the convergence of ideal GMRES for J_1 has two phases:

$$\varphi_k(J_1) \;\sim\; \cos\left(\tfrac{\pi}{2n/k+1}\right), \qquad \text{for } k \le n/2\,,\ k \text{ divides } n\,, \tag{6.8}$$

$$\varphi_{n-k}(J_1) \;\sim\; \left[1+\log(n/k)\right]^{-1}, \qquad \text{for } n-k > n/2\,,\ k \text{ divides } n\,. \tag{6.9}$$

The convergence bound based on the polynomial numerical hull, i.e. (1.6), which is the lower bound in (2.10) in case of a Jordan block, is $\varphi_k(J_1) \ge \varrho_{k,n}^k$. For k dividing n we know $\varrho_{k,n}$ explicitly, and this lower bound can be evaluated, cf. (4.20). For other k one can use the explicit value of $\varrho_{n/2,n}$ resp. the lower bound on $\varrho_{n-1,n}$, cf. (4.25) resp. [5, p. 88], giving

$$\frac{1}{2} \;\le\; 2^{-2k/n} \le \varphi_k(J_1)\,, \qquad \text{for } k = 1, \ldots, n/2\,, \tag{6.10}$$

$$\left[1-\tfrac{\log(2n)}{n}\right]^k \le \varphi_k(J_1)\,, \qquad \text{for } k = n/2+1, \ldots, n-1\,. \tag{6.11}$$

Comparing (6.10) and (6.8) shows that the lower bound in (6.10) is a tight approximation of the actual ideal GMRES approximations. Hence the polynomial numerical hull of J_1 gives good information about the first phase of the ideal GMRES convergence. However, the information is less reliable in the second phase. In particular, consider the ideal GMRES approximation for $n-1$. Then (6.9) shows that

$$\varphi_{n-1}(J_1) \;\sim\; \left[1+\log n\right]^{-1},$$

while the lower bound (6.11) yields

$$\left[1-\tfrac{\log(2n)}{n}\right]^{n-1} \;\le\; \varphi_{n-1}(J_1)\,.$$

A real analysis exercise shows that

$$\lim_{n\to\infty} 2n\left[1-\tfrac{\log(2n)}{n}\right]^{n-1} \;=\; 1\,.$$

Hence for large n and $k = n-1$, the value on the right hand side of the lower bound (6.11) is of order $\mathcal{O}(1/n)$, while the actual ideal GMRES approximation $\varphi_{n-1}(J_1)$ is of order $\mathcal{O}(1/\log(n))$. Note that since

$$\lim_{n\to\infty} \tfrac{2n}{\log(n)}\left[1-\tfrac{\log(2n)}{n}+\tfrac{\log(\log(2n))}{n}\right]^{n-1} \;=\; 1\,,$$

an approximation of $\varphi_{n-1}(J_1)$ based on the upper bound on $\varrho_{n-1,n}$, cf. (2.9), also would fail to predict the correct order of magnitude of the ideal GMRES approximation.

As shown by this example, the bound (1.6) on the kth ideal GMRES approximation for a general nonnormal matrix A based on the polynomial numerical hull of A of degree k, cannot be expected to be tight for all k.

7. Concluding remarks. Motivated by the (in general) open question of how to characterize the convergence of the GMRES method in the nonnormal case, we have studied the behavior of worst-case and ideal GMRES for an $n \times n$ Jordan block J_λ with eigenvalue $\lambda \in \mathbb{C}$. We conjecture that any such J_λ is ideal. We have shown in this paper that J_λ is ideal of degree k if any of the following conditions is satisfied:

1. $|\lambda| \leq \varrho_{k,n}$,
2. k divides n,
3. $k < n/2$ and $|\lambda| \geq \varrho_{k,n-k}^{-1}$,
4. $k \geq n/2$, $n-k$ divides n and $|\lambda| \geq 1$.

Apart from studying the idealness of J_λ, we have extended the results of [2, 5] by proving new results about the radii of the polynomial numerical hulls of Jordan blocks. Using these, we discussed the closeness of (1.6), i.e. the lower bound on ideal GMRES based on polynomial numerical hull.

Acknowledgments. We thank Anne Greenbaum for many helpful comments, and in particular for pointing out that Theorem 4.2 can be used for investigating the radii of polynomial numerical hulls of Jordan blocks. We also thank Jurjen Duintjer Tebbens for helpful comments and suggestions.

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] V. Faber, A. Greenbaum, and D. E. Marshall, *The polynomial numerical hulls of Jordan blocks and related matrices*, Linear Algebra Appl., 374 (2003), pp. 231–246.
[3] 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.
[4] A. Greenbaum, *Generalizations of the field of values useful in the study of polynomial functions of a matrix*, Linear Algebra Appl., 347 (2002), pp. 233–249.
[5] ———, *Some theoretical results derived from polynomial numerical hulls of Jordan blocks*, Electron. Trans. Numer. Anal., 18 (2004), pp. 81–90 (electronic).
[6] A. Greenbaum and L. Gurvits, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 348–358.
[7] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrixapproximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.
[8] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.
[9] ———, *Matrix analysis*, Cambridge University Press, Cambridge, 1999. Corrected reprint of the 1985 original.
[10] W. Joubert, *On the convergence behavior of the restarted GMRES algorithm for solving nonsymmetric linear systems*, Numer. Linear Algebra Appl., 1 (1994), pp. 427–447.
[11] ———, *A robust GMRES-based adaptive polynomial preconditioning algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439.
[12] J. Liesen and Z. Strakoš, *Convergence of GMRES for tridiagonal Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 233–251 (electronic).
[13] O. Nevanlinna, *Convergence of iterations for linear equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993.
[14] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetriclinear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[15] R. P. Stanley, *Enumerative combinatorics. Vol. 1*, vol. 49 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1997. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.
[16] P. Tichý and J. Liesen, *Worst-case and ideal GMRES for a Jordan block*, Preprint 19-2005, Institute of Mathematics, Technical University of Berlin, 2005.
[17] K. C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.

[18] K. C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3 – a Matlab software package for semidefinite programming, version 2.3. Interior point methods.* June 2001.