GMRES CONVERGENCE AND THE POLYNOMIAL NUMERICAL HULL FOR A JORDAN BLOCK

PETR TICHÝ[†] AND JÖRG LIESEN[†]

Abstract. Consider a system of linear algebraic equations with a nonsingular n by n matrix A. When solving this system with GMRES, the relative residual norm at the step k is bounded from above by the so called ideal GMRES approximation. This bound is sharp (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. In this paper we consider an n by n Jordan block J, and study the relation between ideal and worst-case GMRES as well as the problem of estimating the ideal GMRES approximations. Under some assumptions, we show that ideal and worst-case GMRES are identical at steps k and $n-k$ such that k divides n, and we derive explicit expressions for the $(n-k)$th ideal GMRES approximation. Furthermore, we extend previous results in the literature by proving new results about the radii of the polynomial numerical hulls of Jordan blocks. Using these, we discuss the tightness of the lower bound on the ideal GMRES approximation that is derived from the radius of the polynomial numerical hull of J.

Key words. Krylov subspace methods, GMRES convergence, polynomial numerical hull, Jordan block.

1. Introduction.

Suppose that we solve a linear system $Ax = b$ with the GMRES method [12]. Starting from an initial guess x_0, this method computes the initial residual $r_0 \equiv b - Ax_0$ and a sequence of iterates $x_1, x_2, \ldots$, so that the kth residual $r_k \equiv b - Ax_k$ satisfies

$$\|r_k\| \;=\; \min_{p\in\pi_k} \|p(A)\,r_0\|\,, \tag{1.1}$$

where π_k denotes the set of 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

$$r_k \in r_0 + A\mathcal{K}_k(A, r_0)\,, \qquad r_k \perp A\mathcal{K}_k(A, r_0)\,. \tag{1.2}$$

Here $\mathcal{K}_k(A, r_0) \equiv \text{span}\{r_0, Ar_0, \ldots A^{k-1}r_0\}$ is the kth Krylov subspace generated by A and r_0, and $\perp$ means orthogonality with respect to the Euclidean inner product. Without loss of generality we will consider in this paper that r_0 is a unit norm vector, i.e. $\|r_0\| = 1$.

The approximation problem (1.1) depends on the three input parameters A, r_0, and k. It turns out that it is very hard to analyze this problem in general. A common approach for investigating the GMRES convergence behavior is to bound (1.1) independently of r_0. Because of the submultiplicativity of the Euclidean norm, an upper bound on (1.1) is given by

$$\Phi_k^A \;\equiv\; \min_{p\in\pi_k} \|p(A)\|. \tag{1.3}$$

The problem (1.3) represents a matrix approximation problem and the value Φ_k^A is called the *ideal GMRES approximation* [6]. Clearly, Φ_k^A represents an upper bound

[†]Institute of Mathematics, Technical University of Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany. E-mail: `tichy@math.tu-berlin.de` and `liesen@math.tu-berlin.de`. This research was supported by the Emmy Noether-Programm of the Deutsche Forschungsgemeinschaft. The work of the first author was also supported by the Czech National Program of Research Information Society under project 1ET400300415.

on the *worst-case GMRES approximation*

$$\Psi_k^A \;\equiv\; \max_{\|r_0\|=1} \min_{p\in\pi_k} \|p(A)r_0\| \,. \tag{1.4}$$

The relation between ideal and worst-case GMRES has been investigated in several papers. The best known result is that $\Phi_k^A = \Psi_k^A$ for all k whenever A is normal [5, 9]. For nonnormal A, the situation is more complicated. Here some example matrices A are known for which $\Phi_k^A > \Psi_k^A$ [2, 15]. Despite the existence of these examples, it is still an open problem whether $\Phi_k^A = \Psi_k^A$ (or at least $\Phi_k^A \approx \Psi_k^A$) for larger classes of nonnormal matrices.

Another open problem in the context of (1.3) is how to determine or estimate the value of the ideal GMRES approximation Φ_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, 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 [11, p. 41]. Here $\mathcal{P}_k$ denotes the set of polynomials of degree at most k. The sets $\mathcal{H}_k(A)$ have been used to study the ideal GMRES behavior [2, 3, 4]. Based on the definition (1.5) it is not hard to see that these sets provide a lower bound on the ideal GMRES approximation [3],

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

Moreover, $\mathcal{H}_k(A)$ allows to identify when ideal GMRES fails to converge [2, 3],

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

In this paper we consider an n by n Jordan block J_λ with eigenvalue λ, and study the relation between ideal and worst-case GMRES as well as the problem of estimating the ideal GMRES approximations. We show that $\Phi_k^{J_\lambda} = \Psi_k^{J_\lambda}$ when k divides n. For $|\lambda| \geq 1$ and k dividing n, we derive explicit expressions for $\Phi_{n-k}^{J_\lambda}$ and prove that $\Phi_{n-k}^{J_\lambda} = \Psi_{n-k}^{J_\lambda}$. Furthermore, we extend the results of [1, 4] by proving new results about the radii of the polynomial numerical hulls of Jordan blocks. Using these, we discuss the closeness of the lower bound (1.6) in case of a Jordan block. We conclude that this bound is tight up to a constant for $k \leq n/2$ (in case n is even), but that it fails to characterize the ideal GMRES approximations for $k > n/2$. This proves that the lower bound (1.6) in case of a general nonnormal matrix cannot be expected to be tight for all k.

The paper is organized as follows. Section 2 summarizes relations between ideal and worst-case GMRES. In Section 3 we deal with the $(n-1)$st ideal and worst-case GMRES step for a Jordan block. Section 4 describes the structure behind the ideal GMRES convergence. As shown in Section 5, this structure can be used to translate results about steps 1 and $n-1$ into steps k and $n-k$, in case k divides n. In Section 6 we discuss the question how well the bound based on the polynomial numerical hull characterizes the ideal GMRES convergence, and Section 7 presents further discussion based on numerical experiments.

2. Relations between ideal and worst-case GMRES. If the matrix A is nonsingular and $\Phi_k^A > 0$, then the polynomial that solves the ideal GMRES approximation problem (1.3) is uniquely determined [6, Theorem 2]. This gives rise to the following definition.

DEFINITION 2.1. *Suppose that a nonsingular matrix A and a positive integer k are given for which $\Phi_k^A > 0$. Then the polynomial $\varphi_k \in \pi_k$ that satisfies*

$$\|\varphi_k(A)\| = \Phi_k^A = \min_{p\in\pi_k} \|p(A)\|$$

is called the kth ideal GMRES polynomial of A, and $\varphi_k(A)$ is called the kth ideal GMRES matrix of A.

In general it is an open problem which properties of A are necessary and sufficient so that $\Phi_k^A = \Psi_k^A$. In the following we will summarize the most important results for our context. We first present a lemma that characterizes the case $\Phi_k^A = \Psi_k^A$.

LEMMA 2.2. *Suppose that a nonsingular matrix A and a positive integer k are given for which $\Phi_k^A > 0$. Then $\Phi_k^A = \Psi_k^A$ if and only if there exist a unit norm vector r_0 and a polynomial $\psi \in \pi_k$, such that*

$$\psi(A)r_0 \perp A\mathcal{K}_k(A, r_0)\,, \tag{2.1}$$

and r_0 lies in the span of right singular vectors of $\psi(A)$ corresponding to its maximal singular value. If such ψ and r_0 exist, then $\psi = \varphi_k$.

Proof. If $\Phi_k^A = \Psi_k^A$, then there exists an unit norm vector r_0 and a GMRES polynomial $\psi \in \pi_k$ satisfying (2.1), cf. (1.2), and for the kth ideal GMRES polynomial φ_k of A,

$$\|\varphi_k(A)r_0\| \le \|\varphi_k(A)\| = \|\psi(A)r_0\|\,. \tag{2.2}$$

Since $\|\psi(A)r_0\|$ is minimal, the equality $\|\varphi_k(A)r_0\| = \|\psi(A)r_0\|$ holds. But this means that r_0 lies in the span of maximal right singular vectors of $\varphi_k(A)$, cf. (2.2). Moreover, since $\Psi_k^A > 0$, the kth GMRES polynomial is unique, cf. [6, Theorem 2]. Therefore $\varphi_k = \psi$, and hence r_0 lies in the span of maximal right singular vectors of $\psi(A)$.

Now assume that there exists a polynomial $\psi \in \pi_k$ and a unit norm vector r_0 such that (2.1) holds and r_0 lies in the span of maximal right singular vectors of $\psi(A)$. Then

$$\|\psi(A)\| = \|\psi(A)r_0\| = \min_{p\in\pi_k} \|p(A)r_0\| \le \|\varphi_k(A)\|. \tag{2.3}$$

Since φ_k is the ideal GMRES polynomial, $\|\psi(A)\| < \|\varphi_k(A)\|$ is impossible, and therefore equality holds in (2.3). In other words, $\Phi_k^A = \Psi_k^A$, and from uniqueness of φ_k it follows that $\psi = \varphi_k$. ☐

LEMMA 2.3. *For any nonsingular matrix A, $\Phi_1^A = \Psi_1^A$. If the kth ideal GMRES matrix $\varphi_k(A)$ of A has a simple maximal singular value, then $\Phi_k^A = \Psi_k^A$.*

Proof. The first statement is proven independently in [5] and [9], the second follows from [5, Lemma 2.4]. ☐

Faber, Joubert, Knill, and Manteuffel [2] prove that for an upper triangular Toeplitz matrix T,

$$\Phi_k^T = 1 \quad\Longleftrightarrow\quad \Psi_k^T = 1\,, \tag{2.4}$$

i.e. the ideal and worst-case GMRES approximations for upper triangular Toeplitz matrices are the same in case of stagnation. However, it is in general still an open problem, originally posed in [2, p. 722], whether the two approximations also coincide when ideal GMRES converges, i.e. when $\Phi_k^A < 1$.

The situation where $\Phi_k^A < 1$ can be identified using the polynomial numerical hull. Faber, Greenbaum, and Marshall [1] investigate the polynomial numerical hulls of an n by n Jordan block,

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

They show that 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 $1 > \varrho_{1,n} > \ldots > \varrho_{n-1,n} > 0$, and the radii are independent of the eigenvalue λ. In particular, Faber et al. [1] 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

$$\varrho_{1,n} = \cos\left(\tfrac{\pi}{n+1}\right)\,, \tag{2.6}$$

cf. [1, 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 [1, p. 238], that $\varrho_{n-1,n}$ is a solution of a certain nonlinear equation and can be bounded by

$$1 - \tfrac{\log(2n)}{n} \le \varrho_{n-1,n} \le 1 - \tfrac{\log(2n)}{n} + \tfrac{\log(\log(2n))}{n}\,. \tag{2.7}$$

Continuing this work, Greenbaum [4, p. 88] combines (1.6) and results of [1] to prove that for $k = 1, \ldots, n-1$,

$$|\lambda|^{-k} \;\ge\; \Phi_k^{J_\lambda} \;\ge\; \varrho_{k,n}^k |\lambda|^{-k} \quad \text{for} \quad |\lambda| \ge \varrho_{k,n}\,, \tag{2.8}$$

$$\Phi_k^{J_\lambda} \;=\; \Psi_k^{J_\lambda} \;=\; 1 \quad\Longleftrightarrow\quad |\lambda| \le \varrho_{k,n}\,. \tag{2.9}$$

The upper bound in (2.8) can be replaced by 1 if $|\lambda| \le 1$. The lower bound in (2.8) is a special case of the general lower bound (1.6) on the ideal GMRES approximation based on the polynomial numerical hull. The tightness of this lower bound is examined in Section 6 below. For example, combining the first assertion in Lemma 2.3, (2.6), and (2.8) shows that for $|\lambda| \ge \cos\left(\frac{\pi}{n+1}\right)$,

$$|\lambda|^{-1} \;\ge\; \Phi_1^{J_\lambda} \;=\; \Psi_1^{J_\lambda} \;\ge\; \cos\left(\tfrac{\pi}{n+1}\right)\, |\lambda|^{-1}\,.$$

Using previous results, if $\lambda = 0$, then the polynomial numerical hull of J_λ of each degree contains the origin, which implies that both ideal and worst-case GMRES completely stagnate. Hence of interest in our context is only the *nonsingular case*, i.e. $\lambda \neq 0$. Moreover, each $\lambda \in \mathbb{C}$ can be written as $\lambda = |\lambda| e^{\mathbf{i}\alpha}$, and it holds that

$$J_\lambda \;=\; e^{\mathbf{i}\alpha} U J_{|\lambda|} U^H\,, \qquad U \equiv \operatorname{diag}(e^{\mathbf{i}\alpha}, e^{\mathbf{i}2\alpha}, \ldots, e^{\mathbf{i}n\alpha})\,. \tag{2.10}$$

Therefore, to investigate ideal and worst-case GMRES, it suffices to concentrate only on real and positive λ. All results can be then easily extended to all complex λ using the unitary similarity transformation defined by (2.10).

Throughout the paper, we will use the backward identity I_n^B and the matrix $I_n^\pm$ defined by

$$I_n^B \equiv \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} \in \mathbb{R}^{n\times n}, \quad I_n^\pm \equiv \operatorname{diag}(1,-1,\ldots,(-1)^{n-1})\,. \tag{2.11}$$

3. The next-to-last ideal and worst-case GMRES approximations. Consider the $(n-1)$st ideal and worst-case GMRES approximations for an n by n Jordan block J_λ, $\lambda > 0$. Our main result, stated in Theorem 3.4 below, is that $\Phi_{n-1}^{J_\lambda} = \Psi_{n-1}^{J_\lambda}$ for $\lambda \geq 1$, and we also give an explicit expression for $\Phi_{n-1}^{J_\lambda}$ in terms of the eigenvalue λ. The proof of this result will make use of three technical lemmas. The first lemma is a slight reformulation of [10, Corollary 2.2].

LEMMA 3.1. *Consider an n by n Jordan block J_λ, a vector $r_0 = [\rho_1,\ldots,\rho_n]^T$ with $\rho_n \neq 0$, and let χ be the unique solution of the linear system*

$$\begin{pmatrix} \rho_1 & \cdots & \rho_{n-1} & \rho_n \\ \vdots & \iddots & \iddots & \\ \rho_{n-1} & \iddots & & \\ \rho_n & & & \end{pmatrix} \chi = I_n^\pm \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}. \tag{3.1}$$

Then the $(n-1)$st GMRES residual r_{n-1} for J_λ, and the initial residual r_0 satisfies $\|r_{n-1}\|^{-2} r_{n-1} = \chi$.

LEMMA 3.2. *Let $\lambda > 0$ be given and let $r_0 \in \mathbb{R}^n$ be the unit norm vector*

$$r_0 \equiv (-1)^{n-1}\|\xi\|^{-1}\, I_n^B \xi\,, \tag{3.2}$$

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

$$\xi_{i+1} = \lambda^{\frac{n-1}{2}-i}\frac{(-1)^i}{4^i}\binom{2i}{i}, \qquad i = 0,\ldots,n-1. \tag{3.3}$$

Then the $(n-1)$st GMRES residual r_{n-1} for the n by n Jordan block J_λ and the initial residual r_0 is given by $r_{n-1} = \|\xi\|^{-3}\,\xi$ and hence

$$\|r_{n-1}\| = \|\xi\|^{-2} = \frac{1}{\lambda^{n-1}}\left[\sum_{i=0}^{n-1}(4\lambda)^{-2i}\binom{2i}{i}^2\right]^{-1}. \tag{3.4}$$

Proof. Since the last component of $r_0 = (-1)^{n-1}\|\xi\|^{-1}[\xi_n,\ldots,\xi_1]^T$ is nonzero, Lemma 3.1 implies that the $(n-1)$st GMRES residual for J_λ and r_0 satisfies $\|r_{n-1}\|^{-2}r_{n-1} = \chi$, where χ is the unique solution of

$$\frac{(-1)^{n-1}}{\|\xi\|}\begin{pmatrix} \xi_n & \cdots & \xi_2 & \xi_1 \\ \vdots & \iddots & \iddots & \\ \xi_2 & \iddots & & \\ \xi_1 & & & \end{pmatrix} \chi = I_n^\pm \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}. \tag{3.5}$$

Using the definition (3.3), 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. [13, p. 44]. The n previous equations can be written in matrix form as

$$\begin{pmatrix} \xi_n & \dots & \xi_2 & \xi_1 \\ \vdots & \iddots & \iddots & \\ \xi_2 & \iddots & & \\ \xi_1 & & & \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = (-1)^{n-1} I_n^{\pm} \begin{pmatrix} 1 \\ \lambda \\ \vdots \\ \lambda^{n-1} \end{pmatrix}. \tag{3.6}$$

A comparison of (3.6) and (3.5) shows that the solution of (3.5) is $\chi = \|\xi\|\xi$. Now $\chi = \|r_{n-1}\|^{-2} r_{n-1}$ implies that r_{n-1} indeed is of the form $r_{n-1} = \|\xi\|^{-3}\,\xi$. A straightforward computation shows that $\|r_{n-1}\|$ is given by (3.4). ☐

Lemma 3.3. *Let $\lambda > 0$ be given and let $\xi^+ \equiv I_n^{\pm}\xi$, where the vector ξ is defined as in Lemma 3.2. Then there exists an uniquely determined Hankel matrix H of the form*

$$H = \begin{pmatrix} h_n & \dots & h_2 & h_1 \\ \vdots & \iddots & \iddots & \\ h_2 & \iddots & & \\ h_1 & & & \end{pmatrix}, \quad \text{such that} \quad \xi^+ = H\xi^+. \tag{3.7}$$

If $\lambda \geq 1$, the matrix 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 $\xi^+ = [\xi_1^+, \dots, \xi_n^+]^T$ are positive.

The nth equation in $\xi^+ = H\xi^+$ 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 $n-1, \dots, 1$ it is clear that the entries $h_2, \dots, h_n$ of H are uniquely determined.

To show the remaining part of the lemma, we will first prove by induction that for $\lambda \geq 1$, H is nonnegative with $h_i > 0$, $i = 1, \dots, n$. We already know that $h_1 > 0$. Now suppose that $h_1 > 0, \dots, h_j > 0$ for some $j \geq 1$. The $(n-j)$th equation in $\xi^+ = H\xi^+$ 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. (3.3), 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,

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

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, H is nonnegative and $\xi^+ > 0$ is an eigenvector of H corresponding to the eigenvalue 1. Therefore, 1 must be an eigenvalue of maximum modulus [8, Corollary 8.1.30., p. 493]. Moreover, since $H^2 > 0$, H is primitive, cf. [8, Theorem 8.5.2., p. 516], and there exists only one eigenvalue of maximum modulus. ∎

We now can state and prove the main result of this section.

THEOREM 3.4. *Consider an n by n Jordan block J_λ with $\lambda \geq 1$. Then the unit norm initial residual r_0 defined in (3.2)–(3.3) solves the worst-case GMRES approximation problem (1.4) for J_λ and $k = n-1$, and*

$$\Phi_{n-1}^{J_\lambda} \;=\; \Psi_{n-1}^{J_\lambda} \;=\; \frac{1}{\lambda^{n-1}}\left[\sum_{i=0}^{n-1}(4\lambda)^{-2i}\binom{2i}{i}^2\right]^{-1}. \tag{3.8}$$

Proof. Consider the $(n-1)$st GMRES residual r_{n-1} for J_λ and the initial residual r_0 defined in (3.2)–(3.3), and denote by p_{n-1} the corresponding GMRES polynomial, i.e.

$$r_{n-1} \;=\; p_{n-1}(J_\lambda)\, r_0\,. \tag{3.9}$$

Using (3.4), $\|r_{n-1}\|$ is equal to the rightmost expression in (3.8). To prove the assertion it suffices to show that r_0 is a maximal right singular vector of the matrix $p_{n-1}(J_\lambda)$, cf. Lemma 2.2. 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 (2.11), 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

$$p_{n-1}(J_\lambda) \;=\; U\,(D\hat{I}_n^\pm)\,(\hat{I}_n^\pm U^T I_n^B) \;=\; U\,S\,(\hat{I}_n^\pm U^T I_n^B)\,, \tag{3.10}$$

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

Substituting (3.2), (3.4) and (3.10) into (3.9), we obtain

$$\xi = (-1)^{n-1}\|\xi\|^2 US\hat{I}_n^\pm U^T\xi. \tag{3.11}$$

Similarly as in Lemma 3.3, denote $\xi^+ \equiv I_n^\pm\xi > 0$. Multiplying both sides of (3.11) from the left by $I_n^\pm$ we receive

$$\begin{aligned} \xi^+ \;=\; H\xi^+\,, \quad H &\equiv (-1)^{n-1}\|\xi\|^2(I_n^\pm U)\,S\hat{I}_n^\pm\,(I_n^\pm U)^T \\ &= (-1)^{n-1}\|\xi\|^2(I_n^\pm p_{n-1}(J_\lambda)I_n^B I_n^\pm)\,. \end{aligned} \tag{3.12}$$

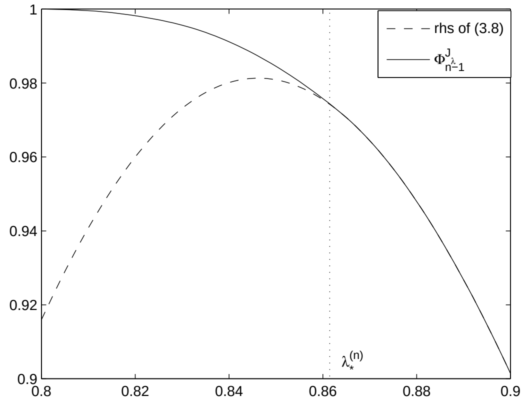

FIG. 3.1. *The right hand side of (3.8) and $\Phi^{J_\lambda}_{n-1}$ plotted as a function of λ.*

Since $p_{n-1}(J_\lambda)$ is an upper triangular Toeplitz matrix, the expression (3.12) shows that H is a Hankel matrix of the form (3.7). Considering the eigenvalue decomposition $H = Q\Lambda Q^T$ it is easy to see that

$$Q = I_n^{\pm} U, \qquad \Lambda = (-1)^{n-1}\|\xi\|^2 S\hat{I}_n^{\pm}. \tag{3.13}$$

Therefore, the modulus of any eigenvalue of H is a $\|\xi\|^2$-multiple of some singular value of $p_{n-1}(J_\lambda)$. Consequently, ξ^+ in (3.12) is an eigenvector corresponding to the eigenvalue of maximum modulus of H if and only if r_0 is a right singular vector corresponding to the maximal singular value of $p_{n-1}(J_\lambda)$. By Lemma 3.3, H has only one eigenvalue of maximum modulus, and ξ^+ is the corresponding eigenvector. Hence r_0 is the maximal right singular vector of $p_{n-1}(J_\lambda)$, which completes the proof. □

In the previous theorem we use the assumption $\lambda \geq 1$. It is natural to ask, what is the relation between ideal and worst-case GMRES for $\varrho_{n-1,n} < \lambda < 1$ and whether the right hand side of (3.8) still characterizes these quantities. Our numerical experiments predict that $\Phi^{J_\lambda}_{n-1} = \Psi^{J_\lambda}_{n-1}$ also for λ between $\varrho_{n-1,n}$ and 1. However, for each integer n there seems to exist a $\lambda_*^{(n)}$, $\varrho_{n-1,n} < \lambda_*^{(n)} < 1$, such that $\Psi^{J_\lambda}_{n-1}$ is not equal to the right hand side of (3.8) for $\lambda < \lambda_*^{(n)}$. In other words, the right hand side of (3.8) does not characterize the ideal and worst-case GMRES approximation for all $\lambda \geq \varrho_{n-1,n}$. This situation is demonstrated in Fig. 3.1. We consider $n = 10$ so that $\varrho_{n-1,n} \approx 0.8$. By the dashed line we plot the right hand side of (3.8) and by the solid line the ideal GMRES approximation $\Phi^{J_\lambda}_{n-1}$ as a function of λ.

Also note that the lower bound on $\varrho_{n-1,n}$ in (2.7) approaches 1 for $n \to \infty$, the equivalence (2.9) implies that for each λ with $0 < \lambda < 1$, there exists a positive integer n_λ such that for the n_λ by n_λ Jordan block J_λ, $\Phi^{J_\lambda}_{n-1} = \Psi^{J_\lambda}_{n-1} = 1$. In other words, both ideal and worst-case 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.

4. Structure of the ideal GMRES matrices for a Jordan block. In the following, we will translate the results for the 1st resp. $(n-1)$st ideal GMRES approximation to the kth resp. $(n-k)$th ideal GMRES approximation, where k divides

n. To this end we will use the special structure of the ideal GMRES matrices, which we originally discovered numerically (for our experiments we use MATLAB 6.5 Release 13 and the semidefinite programming package SDPT3 [16]). Since the development below is quite technical, we will start with a simple example. Consider a 6 by 6 Jordan block J_λ. Then its second, third and fourth ideal GMRES matrices are upper triangular Toeplitz matrices of the form

$$\underbrace{\begin{pmatrix} \bullet & \circ & \bullet & & & \\ & \bullet & \circ & \bullet & & \\ & & \bullet & \circ & \bullet & \\ & & & \bullet & \circ & \bullet \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{\varphi_2(J_\lambda)}, \quad \underbrace{\begin{pmatrix} \bullet & \circ & \circ & \bullet & & \\ & \bullet & \circ & \circ & \bullet & \\ & & \bullet & \circ & \circ & \bullet \\ & & & \bullet & \circ & \circ \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{\varphi_3(J_\lambda)}, \quad \underbrace{\begin{pmatrix} \bullet & \circ & \bullet & \circ & \bullet & \\ & \bullet & \circ & \bullet & \circ & \bullet \\ & & \bullet & \circ & \bullet & \circ \\ & & & \bullet & \circ & \bullet \\ & & & & \bullet & \circ \\ & & & & & \bullet \end{pmatrix}}_{\varphi_4(J_\lambda)},$$

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 $\varphi_2(J_\lambda)$, $\varphi_3(J_\lambda)$ and $\varphi_4(J_\lambda)$ into block diagonal matrices with upper triangular Toeplitz blocks,

$$\underbrace{\begin{pmatrix} \bullet & \bullet & & & & \\ & \bullet & \bullet & & & \\ & & \bullet & & & \\ & & & \bullet & \bullet & \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_2^T\varphi_2(J_\lambda)P_2}, \quad \underbrace{\begin{pmatrix} \bullet & \bullet & & & & \\ & \bullet & & & & \\ & & \bullet & \bullet & & \\ & & & \bullet & & \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_3^T\varphi_3(J_\lambda)P_3}, \quad \underbrace{\begin{pmatrix} \bullet & \bullet & \bullet & & & \\ & \bullet & \bullet & & & \\ & & \bullet & & & \\ & & & \bullet & \bullet & \bullet \\ & & & & \bullet & \bullet \\ & & & & & \bullet \end{pmatrix}}_{P_4^T\varphi_4(J_\lambda)P_4}.$$

Since the transformation $\varphi_k(J_\lambda) \to P_k^T\varphi_k(J_\lambda)P_k$ is orthogonal, and all diagonal blocks of $P_k^T\varphi_k(J_\lambda)P_k$ are equal, the ideal GMRES approximation $\Phi_k^{J_\lambda} = \|\varphi_k(J_\lambda)\|$ equals the norm of any diagonal block of $P_k^T\varphi_k(J_\lambda)P_k$. These observations are the key to analyzing the kth and $(n-k)$th ideal GMRES approximations for 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 $\varphi_k(J_\lambda)$ and of a diagonal block of $P_k^T\varphi_k(J_\lambda)P_k$.

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 by m upper triangular Toeplitz matrix B,*

$$B \equiv \sum_{i=0}^{\ell} b_i E_m^i, \quad \text{and let} \quad B = USV^T \tag{4.1}$$

be its singular value decomposition. Then the singular value decomposition of the n by n matrix G,

$$G \equiv \sum_{i=0}^{\ell} b_i E_n^{id} \quad \text{is given by} \quad G = (U \otimes I_d)(S \otimes I_d)(V \otimes I_d)^T. \tag{4.2}$$

Proof. Define the n by 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 [7, Corollary 4.3.10]. □

As described by the example of the 6 by 6 Jordan block above and by Lemma 4.1, our strategy is as follows: Having an ideal GMRES 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.

Lemma 4.2. *Let n and k be positive integers, $n > k$, and let d be their greatest common divisor. Let $\lambda > 0$ be given 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 φ_ℓ of J_μ is of the form

$$\varphi_\ell(z) = \sum_{i=0}^{\ell} c_i (\mu - z)^i \,. \tag{4.3}$$

If $\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} > 0$, then $\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} = \Phi_k^{J_\lambda} = \Psi_k^{J_\lambda}$, and the kth ideal GMRES polynomial φ_k of J_λ is given by

$$\varphi_k(z) = \sum_{i=0}^{\ell} c_i (\lambda - z)^{id} \,. \tag{4.4}$$

Proof. Let us define the polynomial

$$\psi(z) \;\equiv\; \sum_{i=0}^{\ell} c_i (\lambda - z)^{id} \,. \tag{4.5}$$

By assumption, $\varphi_\ell \in \pi_\ell$, which implies that $\psi \in \pi_k$.

We will now construct a unit norm vector r_0 lying in the span of maximal right singular vectors of $\psi(J_\lambda)$, such that the condition (2.1) is satisfied. According to Lemma 2.2, this means that ψ (which is equal to φ_k in (4.4)) is the kth ideal GMRES polynomial of J_λ.

From

$$\varphi_\ell(J_\mu) \;=\; \sum_{i=0}^{\ell} c_i (-E_m)^i, \qquad \psi(J_\lambda) \;=\; \sum_{i=0}^{\ell} c_i (-E_n)^{id} \,, \tag{4.6}$$

we see that the matrices $\varphi_\ell(J_\mu)$ and $\psi(J_\lambda)$ 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, $\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} > 0$, and hence by Lemma 2.2 there exists a unit norm vector w in the span of the maximal right singular vectors of $\varphi_\ell(J_\mu)$, such that

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

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

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

$$B \;\equiv\; \varphi_\ell(S_\mu). \tag{4.9}$$

Then it easily follows that

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

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. In other words, when v is a right singular vector of the Toeplitz matrix B, then the corresponding left singular vector is of the form $\pm I_m^{B} v$.

Denoting by δ the maximal singular value of $\varphi_\ell(J_\mu)$,

$$Bv = \pm\,\delta\, I_m^{B} v\,, \quad \text{and} \quad \delta \equiv \|\varphi_\ell(J_\mu)\| = \|B\| = \|\psi(J_\lambda)\|\,, \tag{4.11}$$

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

Since v lies in the span of the right singular vectors of B corresponding to δ, the vectors $v \otimes e_i$, where e_i denotes the ith standard basis vector for $i = 1, \ldots, d$, lie in the span of the right singular vectors of $\psi(J_\lambda)$ corresponding to δ, cf. Lemma 4.1. Now define $e_\lambda \equiv [1, -\lambda, \ldots, (-\lambda)^{d-1}]^T$, and

$$r_0 \;\equiv\; \gamma \sum_{i=1}^{d} (-\lambda)^{i-1}\, v \otimes e_i \;=\; \gamma\,(v \otimes e_\lambda)\,, \tag{4.12}$$

where γ is chosen so that $\|r_0\| = 1$. Clearly, r_0 lies in the span of the right singular vectors of the Toeplitz matrix $\psi(J_\lambda)$ corresponding to δ. Then $\pm I_n^{B} r_0$ lies in the span of the corresponding left singular vectors. Together with the first expression in (4.11) this yields

$$\begin{aligned} \psi(J_\lambda) r_0 &= \gamma \psi(J_\lambda)\,(v \otimes e_\lambda) \\ &= \pm\gamma\delta I_n^{B}\,(v \otimes e_\lambda) = \pm\gamma\delta\,((I_m^{B} v) \otimes (I_d^{B} e_\lambda)) \\ &= \pm\gamma\,((Bv) \otimes (I_d^{B} e_\lambda))\,. \end{aligned} \tag{4.13}$$

We next show that

$$\psi(J_\lambda) r_0 \;\perp\; J_\lambda^i r_0\,, \quad i = 1, \ldots, k\,, \tag{4.14}$$

i.e. that ψ is a GMRES polynomial for J_λ and the initial residual r_0. Since

$$\operatorname{span}\{J_\lambda r_0, \ldots, J_\lambda^k r_0\} \;=\; \operatorname{span}\{E_n^0 J_\lambda r_0, \ldots, E_n^{k-1} J_\lambda r_0\}\,, \tag{4.15}$$

the relation (4.14) holds if and only if

$$\psi(J_\lambda)r_0 \;\perp\; E_n^i J_\lambda r_0\,, \quad i = 0,\ldots,k-1\,. \tag{4.16}$$

Let us decompose the index i as

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

An elementary computation shows that

$$J_\lambda r_0 \;=\; \gamma\, J_\lambda(v\otimes e_\lambda) = \;\gamma\,((S_\mu v)\otimes e_d)\,.$$

Multiplication of $J_\lambda r_0$ from the left by E_n^i shifts all entries of $J_\lambda r_0$ upwards by i positions. Using (4.17), $E_n^i J_\lambda r_0$ can be written as

$$E_n^i J_\lambda r_0 \;=\; \gamma\, E_n^{sd}((S_\mu v)\otimes e_{d-q}) = \gamma\,((E_m^s S_\mu v)\otimes e_{d-q})\,. \tag{4.18}$$

Now from (4.13) and (4.18) we obtain

$$\begin{aligned}(\psi(J_\lambda)r_0)^T(E_n^i J_\lambda r_0) &= \pm\gamma^2\,((Bv)\otimes(I_d^B e_\lambda))^T\,((E_m^s S_\mu v)\otimes e_{d-q})\\ &= \pm\,\gamma^2\,[(Bv)^T E_m^s S_\mu v]\,[e_\lambda^T I_d^B e_{d-q}].\end{aligned}$$

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, ψ is the kth GMRES polynomial for the matrix J_λ and the initial residual r_0 that lies in the span of right singular vectors corresponding to the maximal singular value of $\psi(J_\lambda)$. Using Lemma 2.2, it holds that $\Phi_k^{J_\lambda} = \Psi_k^{J_\lambda}$ and, therefore, ψ is the kth ideal GMRES polynomial of J_λ. Moreover, Lemma 4.1 implies that the ideal GMRES matrices $\varphi_\ell(J_\mu)$ and $\varphi_k(J_\lambda)$ have the same norm and thus $\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} = \Phi_k^{J_\lambda} = \Psi_k^{J_\lambda}$. $\square$

As an example, consider an n by 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 Lemma 4.2. Since for the 2 by 2 Jordan block J_μ, $\Psi_1^{J_\mu} = \Phi_1^{J_\mu} > 0$, Lemma 4.2 implies that $\Phi_1^{J_\mu} = \Psi_1^{J_\mu} = \Phi_{n/2}^{J_\lambda} = \Psi_{n/2}^{J_\lambda}$. Moreover, a direct computation of the first ideal GMRES approximation for the 2 by 2 Jordan block J_μ, $\mu = \lambda^{n/2}$, shows that for $\lambda \geq 2^{-2/n}$,

$$\Phi_{n/2}^{J_\lambda} \;=\; \Phi_1^{J_\mu} \;=\; \frac{4\lambda^{n/2}}{4\lambda^n + 1}. \tag{4.19}$$

Lemma 4.2 also allows 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 by n Jordan block satisfies*

$$\varrho_{k,n} \;=\; \varrho_{\ell,m}^{1/d}\,. \tag{4.20}$$

*We thank Anne Greenbaum for motivating us to investigate the radii $\varrho_{k,n}$ using our theory of ideal and worst-case GMRES for a Jordan block.

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

$$\mu \le \varrho_{\ell,m} \overset{\mathbf{A}}{\Longleftrightarrow} \Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} = 1 \overset{\mathbf{B}}{\Longleftrightarrow} \Phi_k^{J_\lambda} = \Psi_k^{J_\lambda} = 1 \overset{\mathbf{C}}{\Longleftrightarrow} \lambda \le \varrho_{k,n}.$$

The equivalences **A** and **C** follow from (2.9), so we only have to prove the equivalence **B**. From Lemma 4.2,

$$\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} = 1 \quad \Longrightarrow \quad \Phi_k^{J_\lambda} = \Psi_k^{J_\lambda} = 1.$$

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

Consequently, for each $\lambda > 0$, $\lambda^d \le \varrho_{\ell,m} \Longleftrightarrow \lambda \le \varrho_{k,n}$, which implies (4.20). ☐

For example, if n is even, then (4.20) for $k = n/2$, $d = k$, $m = 2$, $\ell = 1$ shows that $\varrho_{n/2,n} = \varrho_{1,2}^{1/k} = 2^{-2/n}$, cf. (2.6). The explicit value of $\varrho_{n/2,n}$ can be used to obtain bounds on $\Phi_k^{J_\lambda}$ for $k \le n/2$. Suppose that $\lambda \ge \varrho_{n/2,n}$. Using (2.8) 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 obtain

$$\lambda^{-k} \;\ge\; \Phi_k^{J_\lambda} \;\ge\; \frac{1}{2}\lambda^{-k}, \qquad k \;\le\; n/2\,. \tag{4.21}$$

We will next use our above results to study the kth and $(n-k)$th ideal and worst-case GMRES approximation for J_λ in case k divides n.

5. Results for k and $n-k$ in case k divides n. First consider positive integers k and n, such that $k < n$ divides n. Then $d = k$ is their greatest common divisor, and $m = n/k$, $\ell = 1$ in Theorem 4.3. Using the explicit form of the radius $\varrho_{1,n/k}$, cf. (2.6), the relation (4.20) implies

$$\varrho_{k,n} = \varrho_{1,n/k}^{1/k} = \left[\cos\left(\tfrac{\pi}{n/k+1}\right)\right]^{1/k}. \tag{5.1}$$

Since for $\ell = 1$ it holds that $\Phi_\ell^{J_\mu} = \Psi_\ell^{J_\mu} > 0$, the assumption of Lemma 4.2 is always satisfied when the positive integer $k < n$ divides n, so that we can apply the lemma directly.

THEOREM 5.1. *Consider an n by n Jordan block J_λ with $\lambda > 0$. Let $k < n$ be a positive integer dividing n. Then $\Phi_k^{J_\lambda} = \Psi_k^{J_\lambda}$, and if $\lambda \ge \varrho_{k,n}$,*

$$\lambda^{-k} \;\ge\; \Phi_k^{J_\lambda} \;\ge\; \lambda^{-k}\cos\left(\tfrac{\pi}{n/k+1}\right). \tag{5.2}$$

The kth ideal GMRES polynomial φ_k of J_λ can be written in the form

$$\varphi_k(z) = c_0 + c_1\,(\lambda - z)^k\,, \tag{5.3}$$

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

Proof. All results follow from Lemma 4.2 and Theorem 4.3. The bound (5.2) is just the bound (2.8), where for $\varrho_{k,n}$ we substituted its exact value on the right hand side of (5.1). ∎

From the form of the kth ideal GMRES polynomial (5.3) it is easy to see that for $\lambda > \varrho_{n,k}$, we have $c_1 \neq 0$, and the k roots of φ_k are uniformly distributed on the circle around λ with radius $|c_0/c_1|^{1/k}$. If $\lambda \leq \varrho_{n,k}$ then $\varphi_k(z) = 1$.

Now consider the case $n-k$ such that $k < n$ divides n. Then the greatest common divisor of $n-k$ and n is k, and the parameters d, m and ℓ from Lemma 4.2 are given by $d = k$, $m = n/k$, $\ell = m-1$. Using Theorem 4.3 we obtain

$$\varrho_{n-k,n} = \varrho_{m-1,m}^{1/k}\,. \tag{5.4}$$

For example, if $n \geq 4$ is even and $k = 2$, then $m = n/2$ and (5.4) 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 is obtained by Faber et al. in [1, p. 241]. Tight bounds on $\varrho_{m-1,m}$ are given by (2.7).

In the following theorem we combine results of Lemma 4.2 and Theorem 3.4.

THEOREM 5.2. *Consider an n by n Jordan block J_λ with $\lambda \geq 1$. If $k < n$ is a positive integer dividing n, then*

$$\Phi_{n-k}^{J_\lambda} \;=\; \Psi_{n-k}^{J_\lambda} \;=\; \frac{1}{\lambda^{n-k}} \left[\sum_{i=0}^{n/k-1} \lambda^{-2ki} 4^{-2i} \binom{2i}{i}^2 \right]^{-1}. \tag{5.5}$$

Proof. The parameters in Lemma 4.2 are given by $d = k$, $m = n/k$, $\ell = m-1$ and $\mu = \lambda^k$. Applying Theorem 3.4 to the m by m Jordan block J_μ we see that $\Phi_{m-1}^{J_\mu} = \Psi_{m-1}^{J_\mu}$, and this quantity is positive. Hence the assumption of Lemma 4.2 is satisfied. Therefore, $\Phi_{m-1}^{J_\mu} = \Phi_{n-k}^{J_\lambda} = \Psi_{n-k}^{J_\lambda}$. The value of $\Phi_{m-1}^{J_\mu}$ (and also of $\Phi_{n-k}^{J_\lambda}$) is given by (3.8), where n and λ have to be replaced by m and λ^k, respectively. ∎

Note that for n even and $k = n/2$, it can be easily checked that the rightmost expression in (5.5) agrees with the rightmost expression in (4.19).

6. Polynomial numerical hulls and the ideal GMRES convergence. Here we are interested in the question how closely the lower bound (1.6), which in case of an n by n Jordan block J_λ with $\lambda \geq \varrho_{k,n}$ is identical to the lower bound in (2.8), approximates the ideal GMRES approximation. To study this question, we concentrate on the n by n Jordan block J_λ with $\lambda = 1$. We need the following lemma, which can be proven by a straightforward computation; see also [14].

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

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

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

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

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

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

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

Proof. We first prove (6.4). In the notation of Lemma 4.2, $m \equiv n/k$ and $\ell = 1$. Denote by J the m by m Jordan block with the eigenvalue one. Since $\Phi_1^J = \Psi_1^J > 0$, Lemma 4.2 implies that $\Phi_k^{J_1} = \Phi_1^J$. It therefore suffices to bound $\|\varphi_1(J)\|$.

The upper bound in (6.4) follows from

$$\|\varphi_1(J)\| \leq \|I - \frac{1}{2}J\| = \frac{1}{2}\|J\| = \cos\left(\tfrac{\pi}{2m+1}\right),$$

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 by m matrix $T_{\omega,m}$,

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

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

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

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,

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

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

We next prove (6.5). Using (5.5), the value of $\Phi^{J_1}_{n-k}$ is given by

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

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 $\Phi^{J_\lambda}_{n-k}$,

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

Using these inequalities and (6.6) we obtain (6.5). $\square$

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:

$$\Phi^{J_1}_k \sim \cos\left(\tfrac{\pi}{2n/k+1}\right), \qquad \text{for } k \leq n/2,\ k \text{ divides } n, \tag{6.8}$$

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

The convergence bound based on the polynomial numerical hull (i.e. (1.6), which is the lower bound in (2.8) in case of a Jordan block), is $\Phi^{J_1}_k \geq \varrho^k_{k,n}$. For k dividing n, we know $\varrho_{k,n}$ explicitly, and this lower bound can be evaluated, cf. (5.2). For other k, one can use the explicit value of $\varrho_{n/2,n}$ resp. the lower bound on $\varrho_{n-1,n}$, cf. (4.21) resp. [4, p. 88], giving

$$\Phi^{J_1}_k \geq 2^{-2k/n} \geq \frac{1}{2}, \qquad \text{for } k = 1, \ldots, n/2, \tag{6.10}$$

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

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

$$\Phi_{n-1}^{J_1} \;\sim\; [1+\log n]^{-1}\,,$$

while the lower bound (6.11) yields

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

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 $\Phi_{n-1}^{J_1}$ 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 $\Phi_{n-1}^{J_1}$ based on the upper bound on $\varrho_{n-1,n}$, cf. (2.7), 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. Further discussion. We have shown that for k dividing n and $\lambda\in\mathbb{C}$, $\Psi_k^{J_\lambda}=\Phi_k^{J_\lambda}$, and $\Psi_{n-k}^{J_\lambda}=\Phi_{n-k}^{J_\lambda}$ if $|\lambda|\geq 1$. Our numerical experiments suggest that indeed $\Psi_k^{J_\lambda}=\Phi_k^{J_\lambda}$ for $\lambda\in\mathbb{C}$ and *each* positive integer k. To prove this result, it would be sufficient to show that $\Psi_k^{J_\lambda}=\Phi_k^{J_\lambda}$ if k and n are relatively prime (for other k one could then apply Lemma 4.2). Numerical observations show that if k and n are relatively prime and $|\lambda|\geq 1$, then the ideal GMRES matrix $\varphi_k(J_\lambda)$ has a simple maximal singular value, which implies that $\Psi_k^{J_\lambda}=\Phi_k^{J_\lambda}$, cf. Lemma 2.3.

Acknowledgments. We thank Anne Greenbaum, Jurjen Duintjer Tebbens and two anonymous referees for helpful comments and suggestions.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] ———, *Some theoretical results derived from polynomial numerical hulls of Jordan blocks*, Electron. Trans. Numer. Anal., 18 (2004), pp. 81–90 (electronic).

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

[6] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrixapproximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.

[7] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1994. Corrected reprint of the 1991 original.

[8] ———, *Matrix analysis*, Cambridge University Press, Cambridge, 1999. Corrected reprint of the 1985 original.

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

[10] J. Liesen and Z. Strakoš, *Convergence of GMRES for tridiagonal Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 233–251 (electronic).

[11] O. Nevanlinna, *Convergence of iterations for linear equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1993.

[12] 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.

[13] 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.

[14] 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.

[15] K. C. Toh, *GMRES vs. ideal GMRES*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 30–36.

[16] 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.