SOME THEORETICAL RESULTS DERIVED FROM POLYNOMIAL NUMERICAL HULLS OF JORDAN BLOCKS

ANNE GREENBAUM*

Abstract. The polynomial numerical hull of degree k for a square matrix A is a set in the complex plane designed to give useful information about the norms of functions of the matrix; it is defined as

$$\{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \text{ for all polynomials } p \text{ of degree } k \text{ or less}\}.$$

In a previous paper [V. Faber, A. Greenbaum, and D. Marshall, *The polynomial numerical hulls of Jordan blocks and related matrices*, *Lin. Alg. Appl.*, 374 (2003), pp. 231–246] analytic expressions were derived for the polynomial numerical hulls of Jordan blocks. In this paper, we explore some consequences of these results. We derive lower bounds on the norms of functions of Jordan blocks and triangular Toeplitz matrices that approach equalities as the matrix size approaches infinity. We demonstrate that even for moderate size matrices these bounds give fairly good estimates of the behavior of matrix powers, the matrix exponential, and the resolvent norm. We give new estimates of the convergence rate of the GMRES algorithm applied to a Jordan block. We also derive a new estimate for the field of values of a general Toeplitz matrix.

Key words. polynomial numerical hull, field of values, Toeplitz matrix

AMS subject classifications. 15A60, 65F15, 65F35

1. Introduction.

The *polynomial numerical hull of degree* k for an n by n matrix A was introduced by Nevanlinna in [15, 16] and further studied by Greenbaum in [8]. It is a set designed to give more information than the spectrum alone can provide about the behavior of the matrix under the action of polynomials and other functions. It is defined as

$$\mathcal{H}_k(A) = \{z \in \mathbf{C} : \ \|p(A)\| \geq |p(z)| \ \ \forall p \in \mathcal{P}_k\}, \tag{1}$$

where k is a positive integer and $\mathcal{P}_k$ denotes the set of polynomials of degree k or less. In this paper $\|\cdot\|$ will always denote the 2-norm for vectors and the associated spectral norm for matrices: $\|B\| = \max_{\|v\|_2=1} \|Bv\|_2$.

While it is clear that $\mathcal{H}_k(A)$ provides a convenient lower bound on the norms of polynomials of degree k or less in A; i.e., $\|p(A)\| \geq \max_{z \in \mathcal{H}_k(A)} |p(z)|$, it also may provide estimates of the norms of other functions such as e^{tA} or $(\zeta I - A)^{-1}$, where $t \geq 0$ and $\zeta \in \mathbf{C}$ are parameters. Since any primary matrix function $f(A)$ can be written as a polynomial of degree at most $n-1$ in A [11], we can write

$$\|f(A)\| = \|p_f(A)\| \geq \max_{z \in \mathcal{H}_{n-1}(A)} |p_f(z)|. \tag{2}$$

The polynomial p_f is the one that matches f at the eigenvalues of A, and, if an eigenvalue λ_i corresponds to a Jordan block of size $m_i > 1$, then the first $m_i - 1$ derivatives of p_f also match those of f at λ_i. If $f(A)$ can be well approximated by a lower degree polynomial, say, $p_k(A)$, then $\|f(A)\|$ also can be related to the maximum value of p_k on $\mathcal{H}_k(A)$.

An equivalent definition of the polynomial numerical hull is [8]:

$$\mathcal{H}_k(A) = \{z \in \mathbf{C} : \ \min_{\substack{p \in \mathcal{P}_k \\ p(z)=1}} \|p(A)\| = 1\}. \tag{3}$$

*University of Washington, Mathematics Dept., Box 354350, Seattle, WA 98195. This work was supported in part by NSF grant DMS-0208353.

That is, if one considers polynomials of the form $p(A) = I + \sum_{j=1}^{k} c_j (A - zI)^j$, then if $z \in \mathcal{H}_k(A)$ then the coefficients that minimize $\|p(A)\|$ are just $c_1 = \ldots = c_k = 0$, while if $z \notin \mathcal{H}_k(A)$, then there are coefficients that make $\|p(A)\| < 1$. This establishes a close connection between the polynomial numerical hull and the ideal GMRES algorithm [9], whose convergence after k steps is measured by the quantity

$$\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(A)\|. \tag{4}$$

This quantity is less than 1, and we say that ideal GMRES(k) converges, if and only if $0 \notin \mathcal{H}_k(A)$.

A related set defined in [8] is

$$\mathcal{F}_k(A) = \{q^* A q : \ q^* q = 1 \ \text{ and } q^* A^j q = (q^* A q)^j, \quad j = 1, \ldots, k\}. \tag{5}$$

For any matrix A, $\mathcal{F}_k(A) \subset \mathcal{H}_k(A)$, and it was argued in [8], based on results in [5], that if A is a normal matrix or a triangular Toeplitz matrix, then these two sets are identical. The precise class of matrices for which these two sets are identical is not known.

A number of simple properties of polynomial numerical hulls were derived in [4, 8, 15, 16]. Here we list several of these for future reference:

THEOREM 1.1.

(i) *$\mathcal{H}_k(\cdot)$ is invariant under unitary similarity transformations.*

(ii) *For scalars $\alpha, \beta \in \mathbf{C}$, $\mathcal{H}_k(\alpha I + \beta A) = \alpha + \beta \mathcal{H}_k(A)$.*

(iii) *If $\mathcal{F}(A)$ denotes the* field of values *($\mathcal{F}(A) = \{q^* A q : \ q^* q = 1\}$) and $\sigma(A)$ the spectrum, then $\mathcal{F}(A) = \mathcal{H}_1(A) \supset \mathcal{H}_2(A) \supset \ldots \supset \mathcal{H}_m(A) = \mathcal{H}_{m+1}(A) = \ldots = \sigma(A)$, where m is the degree of the minimal polynomial of A.*

(iv) *If $\mathcal{H}_k(A) \supset \Omega$, then $\mathcal{H}_k(A) \supset pco_k(\Omega)$, where*

$$pco_k(\Omega) = \{\zeta \in \mathbf{C} : \ |p(\zeta)| \le \max_{z \in \Omega} |p(z)| \quad \forall p \in \mathcal{P}_k\}.$$

Proof.

(i) If Q is a unitary matrix and p any polynomial, then $p(Q^* A Q) = Q^* p(A) Q \Rightarrow \|p(Q^* A Q)\| = \|p(A)\| \Rightarrow \mathcal{H}_k(Q^* A Q) = \mathcal{H}_k(A)$.

(ii) If $\beta = 0$, then it is clear that $\mathcal{H}_k(\alpha I) = \{\alpha\}$, so assume $\beta \neq 0$. For any polynomial $p \in \mathcal{P}_k$, define $q \in \mathcal{P}_k$ by $q(\alpha + \beta z) = p(z)$, or, $q(z) = p((z - \alpha)/\beta)$. Clearly, every $q \in \mathcal{P}_k$ can be written in this form for some $p \in \mathcal{P}_k$, and then $p(A) = q(\alpha I + \beta A)$. It follows that $\zeta \in \mathcal{H}_k(A)$ if and only if $\|p(A)\| \ge |p(\zeta)| \ \forall p \in \mathcal{P}_k$ if and only if $\|q(\alpha I + \beta A)\| \ge |q(\alpha + \beta \zeta)| \ \forall q \in \mathcal{P}_k$ if and only if $\alpha + \beta \zeta \in \mathcal{H}_k(\alpha I + \beta A)$.

(iii) It follows from definition (1) that, for every k, $\mathcal{H}_k(A) \supset \sigma(A)$, since if $(\mathbf{v}, \lambda)$ is an eigenpair of A then $p(A)\mathbf{v} = p(\lambda)\mathbf{v} \ \Rightarrow \ \|p(A)\| \ge |p(\lambda)|$. For $k \ge m$, $\mathcal{H}_k(A) = \sigma(A)$, since if p is the minimal polynomial of A then $\|p(A)\| = 0$ but $p(z) = 0$ only at the eigenvalues of A. The inclusions $\mathcal{H}_j(A) \supset \mathcal{H}_{j+1}(A)$ are clear from definition (1). For a proof that $\mathcal{H}_1(A) = \mathcal{F}(A)$, see [8] or [15].

(iv) This follows directly from definition (1).

□

In [4] analytic expressions were derived for the polynomial numerical hulls of a Jordan block. It was shown that the hulls of degrees 1 through $n-1$ for an n by n

Jordan block are disks about the eigenvalue with radii ranging between $1 - O(1/n^2)$ and $1 - O(\log n/n)$. These results were used to derive fairly tight inner and outer bounds on the polynomial numerical hulls of banded triangular Toeplitz matrices, using the fact that a triangular Toeplitz matrix is just a polynomial in the Jordan block with eigenvalue zero.

In this paper we explore some consequences of these results. We derive lower bounds on the norms of functions of Jordan blocks and triangular Toeplitz matrices that approach equalities as the matrix size approaches infinity. We demonstrate with numerical examples that these bounds provide fairly good estimates of the norms of powers, exponentials, and resolvents of such matrices, even for moderate size matrices. We also derive a new estimate of the field of values of a general Toeplitz matrix.

2. Norms of Functions of a Matrix. The following theorem relates the norm of a function $f(A)$ to the size of the polynomial p_f defined in (2) on $\mathcal{H}_{m-1}(A)$, where m is the degree of the minimal polynomial.

THEOREM 2.1. *Let A have minimal polynomial*

$$q_A(z) = \prod_{j=1}^{d} (z - \lambda_j)^{\ell_j},$$

where $\lambda_1, \ldots, \lambda_d$ are distinct and each $\ell_j \geq 1$. Let $m = deg(q_A(z)) = \sum_{j=1}^{d} \ell_j$. Let $f(z)$ be a scalar valued function, whose domain includes $\lambda_1, \ldots, \lambda_d$. For each λ_j with $\ell_j > 1$, assume that λ_j is in the interior of the domain of $f(z)$ and that $f(z)$ is $\ell_j - 1$ times differentiable at λ_j. Let $f(A)$ be the primary matrix function associated with the stem function $f(z)$. Then

$$\|f(A)\| \geq \max_{z \in \mathcal{H}_{m-1}(A)} \left| \sum_{j=1}^{d} \left(\prod_{\substack{i=1 \\ i \neq j}}^{d} (z - \lambda_i)^{\ell_i} \right) \left(\sum_{i=0}^{\ell_j - 1} \frac{1}{i!} \varphi_j^{(i)}(\lambda_j)(z - \lambda_j)^i \right) \right|, \tag{6}$$

where $\varphi_j(z) = f(z) / \prod_{\substack{i=1 \\ i \neq j}}^{d} (z - \lambda_i)^{\ell_i}$.

Proof. See [11, p. 391 and 412–413]. This follows from the fact that $f(A) = p_f(A)$, where $p_f(z)$ is the polynomial of degree $m - 1$ that matches f at each point λ_j, and whose derivatives of order $1, \ldots, \ell_j - 1$ match those of f at each point λ_j with $\ell_j > 1$. The expression on the right-hand side of (6) is the Hermite-Lagrange interpolation formula for such a polynomial. $\square$

The following corollaries show how to apply Theorem 2.1 to different types of matrices.

COROLLARY 2.2. *If A is diagonalizable and has distinct eigenvalues $\lambda_1, \ldots, \lambda_m$, and if $f(z)$ is a scalar valued function whose domain includes $\lambda_1, \ldots, \lambda_m$, then*

$$\|f(A)\| \geq \max_{z \in \mathcal{H}_{m-1}(A)} \left| \sum_{j=1}^{m} f(\lambda_j) \prod_{\substack{i=1 \\ i \neq j}}^{m} \frac{z - \lambda_i}{\lambda_j - \lambda_i} \right|. \tag{7}$$

COROLLARY 2.3. *If A is similar to an n by n Jordan block $J(\lambda)$ with eigenvalue λ, and if $f(z)$ is a scalar valued function that is $n-1$ times differentiable at λ, then*

$$\|f(A)\| \geq \max_{z \in \mathcal{H}_{n-1}(A)} \left| \sum_{i=0}^{n-1} \frac{f^{(i)}(\lambda)}{i!}(z-\lambda)^i \right|. \tag{8}$$

At first glance, it may appear that the bound in Theorem 2.1 is a discontinuous function of the matrix entries, since it depends on the eigenvalues and their algebraic multiplicities; yet an arbitrarily small change in the matrix entries may change completely the multiplicities of the eigenvalues. For f sufficiently smooth, however, i.e., for f $(n-1)$-times continuously differentiable on a convex set containing $\lambda_1, \ldots, \lambda_d$ in its relative interior or for f analytic on a simply connected open set containing $\lambda_1, \ldots, \lambda_d$, this is not the case. In this case the interpolation polynomial for f can be represented by a Newton formula using divided differences, and it can be shown that the coefficients of this formula are continuous functions of the interpolation points $\lambda_1, \ldots, \lambda_d$ [11, p. 395].

Using knowledge of the norms of infinite Toeplitz matrices, one can obtain upper bounds on $\|f(J(\lambda))\|$ to go with the lower bound of Corollary 2.3 applied to $A = J(\lambda)$. By definition, $f(J(\lambda))$ is the triangular Toeplitz matrix whose diagonals are the values $f^{(i)}(\lambda)/i!$. The norm of an infinite triangular Toeplitz matrix with these diagonals is

$$\max_{|z-\lambda| \leq 1} \left| \sum_{i=0}^{n-1} \frac{f^{(i)}(\lambda)}{i!}(z-\lambda)^i \right|. \tag{9}$$

Since the finite matrix $f(J(\lambda))$ can be thought of as the restriction of the infinite matrix to a finite dimensional subspace, its norm is less than or equal to that of the infinite matrix, and therefore expression (9) gives an upper bound on $\|f(J(\lambda))\|$. It follows that if $\mathcal{H}_{n-1}(J(\lambda))$ approaches a disk of radius 1 about λ as $n \to \infty$ (which will be shown to be the case in the next section), then the inequality (8) for $A = J(\lambda)$ approaches an equality as $n \to \infty$.

Note also that if T is a triangular Toeplitz matrix with diagonals a_0, a_1, …, a_{n-1} then $T = \sum_{j=0}^{n-1} a_j J^j$, where $J = J(0)$ is the Jordan block with eigenvalue zero. Norms of functions of T can be estimated by applying Corollary 2.3 to T directly, or they can be estimated by applying Corollary 2.3 to J, using the fact that $f(T) = f(s(J))$, where $s(z) = \sum_{j=0}^{n-1} a_j z^j$ is the *symbol* of the Toeplitz matrix. We will use the latter approach.

3. The Polynomial Numerical Hulls of Jordan Blocks. The following theorem was established in [4], where part of it was shown to be essentially equivalent to much earlier results proved by Goluzin [6], [7, Theorem 6, pp. 522–523] and by Schur and Szegö [20].

THEOREM 3.1. *The polynomial numerical hull of any degree $k < n$ for an n by n Jordan block with eigenvalue λ is a disk about λ. For $k = 1$, the radius of the disk is $r_{1,n} = \cos\frac{\pi}{n+1}$. For $k = n-1$, the radius $r_{n-1,n}$ is the positive root of*

$$2r^n + r - 1 = 0, \tag{10}$$

when n is even, and is greater than or equal to the positive root of this equation when n is odd. For n odd, $r_{n-1,n}$ is the largest value of r that satisfies

$$1 - r - 2r^n + 2r^n \frac{[1-\cos(d/(n-1))] + r[1-\cos((\pi-d)/n)]}{1+r} \geq 0 \quad \forall d.$$

In either case,

$$r_{n-1,n} = 1 - \frac{\log(2n)}{n} + \frac{\log(\log(2n))}{n} - \frac{\epsilon_n}{n}, \tag{11}$$

where $\epsilon_n > 0$ and $\lim_{n\to\infty} \epsilon_n = 0$, and

$$r_{n-1,n} > 1 - \frac{\log(2n)}{n} \quad \forall n \geq 2. \tag{12}$$

It follows from this theorem that all of the hulls of degrees 1 through $n-1$ for an n by n Jordan block are disks about the eigenvalue with radii between $1 - O(1/n^2)$ and $1 - O(\log n/n)$.

4. Examples. Using Theorem 3.1 together with Corollary 2.3, we can now give good lower bounds on the norms of functions of Jordan blocks. According to the arguments after (9), these lower bounds approach exact expressions as the matrix size approaches infinity, and the following examples show that the bounds can be quite good even for moderate size values of n. In most cases, these bounds are not the best known; by carefully studying a specific function one often can improve upon estimates derived in this general setting. Still, as will be demonstrated, the estimates derived from polynomial numerical hulls are often close to optimal for matrices of this type.

Let $J(\lambda)$ denote the n by n Jordan block with eigenvalue λ. Then

$$\|J(\lambda)^k\| \geq \begin{cases} (|\lambda| + r_{k,n})^k \geq (|\lambda| + 1 - \log(2n)/n)^k, & k < n \\ \sum_{j=0}^{n-1} \begin{pmatrix} k \\ j \end{pmatrix} |\lambda|^{k-j} r_{n-1,n}^j \geq & \\ \quad \sum_{j=0}^{n-1} \begin{pmatrix} k \\ j \end{pmatrix} |\lambda|^{k-j} (1 - \log(2n)/n)^j, & k \geq n \end{cases}. \tag{13}$$

Figure 1a shows a plot of $\|J(\lambda)^k\|$ and the lower bound (13) for a 50 by 50 Jordan block with eigenvalue $\lambda = -.7$. As can be seen from the figure, the estimate (13) is quite close to the actual value of $\|J(\lambda)^k\|$, although slightly sharper estimates can be obtained by other means; see, for example, [1].

Similarly, one can give a lower bound on the norm of the exponential of a Jordan block:

$$\|e^{tJ(\lambda)}\| \geq e^{t\Re(\lambda)} \left| \sum_{\ell=0}^{n-1} \frac{t^\ell}{\ell!} r_{n-1,n}^\ell \right| \geq e^{t\Re(\lambda)} \left| \sum_{\ell=0}^{n-1} \frac{t^\ell}{\ell!} (1 - \log(2n)/n)^\ell \right|. \tag{14}$$

This estimate is plotted in Figure 1b, along with the actual value $\|e^{tJ(\lambda)}\|$, again for a 50 by 50 Jordan block with eigenvalue $\lambda = -.7$. This estimate is less precise than the one for matrix powers but still gives a good idea of the actual behavior of the matrix exponential. The difference between the curves becomes smaller as the matrix size increases.

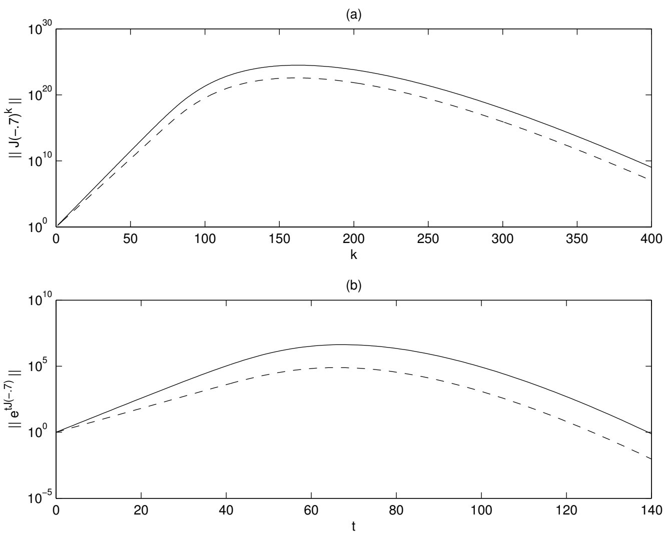

FIG. 1. *Norms of functions of a Jordan block of size $n = 50$ with eigenvalue $\lambda = -.7$. (a) Norm of $J(\lambda)^k$ (solid) and lower bound (13) (dashed), (b) Norm of $e^{tJ(\lambda)}$ (solid) and lower bound (14) (dashed).*

Polynomial numerical hulls also can be used to gain information about the resolvent norm, $\|(\zeta I - A)^{-1}\|$, for values of ζ throughout the complex plane. The ϵ-pseudospectrum of a matrix A [3] is defined as

$$\Lambda_\epsilon(A) = \{\zeta \in \mathbf{C} : \ \|(\zeta I - A)^{-1}\| \geq \epsilon^{-1}\}.$$

These sets are especially useful because they indicate how the eigenvalues of A can change when the matrix is perturbed by a matrix of given norm; that is, an equivalent definition of the ϵ-pseudospectrum is [3]:

$$\Lambda_\epsilon(A) = \{\zeta \in \mathbf{C} : \quad \zeta \text{ is an eigenvalue of } A + E \text{ for some } E \text{ with } \|E\| \leq \epsilon \ \}.$$

In [17] bounds were derived on the pseudospectra of Toeplitz matrices, and these bounds are of course applicable to the simplest Toeplitz matrix, a Jordan block. It was shown that the ϵ-pseudospectrum of an n by n Jordan block contains the disk about the eigenvalue of radius $\epsilon^{1/n}$ and is contained in the disk about the eigenvalue of radius $1 + \epsilon$.

Since the resolvent can be written as a polynomial in the matrix:

$$(\zeta I - J(\lambda))^{-1} = \sum_{j=0}^{n-1} (\zeta - \lambda)^{-(j+1)} (J(\lambda) - \lambda I)^j, \quad \zeta \neq \lambda,$$

it follows from Corollary 2.3 that

$$\|(\zeta I - J(\lambda))^{-1}\| \geq \max_{z \in \mathcal{H}_{n-1}(J(\lambda))} \left| \sum_{j=0}^{n-1} (\zeta - \lambda)^{-(j+1)} (z - \lambda)^j \right| =$$

$$\max_{z \in \mathcal{H}_{n-1}(J(\lambda))} \left| \frac{1}{\zeta - \lambda} \sum_{j=0}^{n-1} \left(\frac{z - \lambda}{\zeta - \lambda} \right)^j \right| = \max_{z \in \mathcal{H}_{n-1}(J(\lambda))} \left| \frac{((z - \lambda)/(\zeta - \lambda))^n - 1}{z - \zeta} \right|. \tag{15}$$

The quantity on the right in (15) is maximized by taking $z-\lambda = r_{n-1,n}(\zeta-\lambda)/|\zeta-\lambda|$. To simplify the notation, we can take $\lambda = 0$, since adding a scalar times the identity to $J(0)$ just corresponds to shifting ζ by that scalar. Then expression (15) can be written as

$$\|(\zeta I - J(0))^{-1}\| \geq \frac{(r_{n-1,n}/|\zeta|)^n - 1}{r_{n-1,n} - |\zeta|}. \tag{16}$$

Thus, on the disk of radius $|\zeta|$, the resolvent norm is greater than or equal to the expression in (16), or, put another way, the $\epsilon \equiv \epsilon(|\zeta|) = (r_{n-1,n} - |\zeta|)/[(r_{n-1,n}/|\zeta|)^n - 1]$ pseudospectrum of $J(0)$ contains the disk of radius $|\zeta|$ about 0. For $|\zeta|$ small, this result is slightly weaker than the one in [17], since $\epsilon > |\zeta|^n$; specifically,

$$\lim_{|\zeta|\to 0} \frac{\epsilon(|\zeta|)}{|\zeta|^n} = \lim_{|\zeta|\to 0} \frac{r_{n-1,n} - |\zeta|}{r_{n-1,n}^n - |\zeta|^n} = \frac{1}{r_{n-1,n}^{n-1}} > 1.$$

For larger values of $|\zeta|$, however, such as $|\zeta| \to r_{n-1,n}$ or $|\zeta| > r_{n-1,n}$, the expression for ϵ is less than $|\zeta|^n$; for example,

$$\lim_{|\zeta|\to r_{n-1,n}} \epsilon(|\zeta|) = \lim_{|\zeta|\to r_{n-1,n}} \frac{-1}{-n(r_{n-1,n}/|\zeta|)^{n-1} r_{n-1,n}|\zeta|^{-2}} = \frac{r_{n-1,n}}{n} < r_{n-1,n}^n,$$

and

$$\lim_{|\zeta|\to\infty} \frac{\epsilon(|\zeta|)}{1+|\zeta|} = \lim_{|\zeta|\to\infty} (r - |\zeta|)/(1+|\zeta|)(r/|\zeta|)^n - 1 = 1.$$

Thus for large values of $|\zeta|$ this inner bound on the ϵ-pseudospectrum approaches the outer bound of [17].

Figure 2 shows a contour plot of the logarithm base 10 of the resolvent norm $\|(\zeta I - J(0))^{-1}\|$ and the lower bound (16) (with $r_{n-1,n}$ replaced by its lower bound $1 - \log(2n)/n$), for a 50 by 50 Jordan block with eigenvalue 0. (The picture is just shifted by λ for a nonzero eigenvalue.) As can be seen from the figure, the inner bounds on pseudospectra derived from (16) are fairly close to the actual pseudospectra.

For one more example, we consider the triangular Toeplitz matrix whose symbol is:

$$s(z) = 10\left(-\frac{1}{16} - \frac{1}{4}z + z^2 - \frac{1}{4}z^3 - \frac{1}{16}z^4\right). \tag{17}$$

This example was considered in [1], where it was noted that the powers of the matrix show many peaks and valleys. Since a triangular Toeplitz matrix is just a polynomial in the Jordan block with eigenvalue zero, any function of a triangular Toeplitz matrix is a function of $J = J(0)$, so that for this example $A^k = [s(J)]^k$. Figure 3 shows $\|A^k\|$ and the lower bound (8) applied to J with $f(z) = s(z)^k$. A matrix of size $n = 18$ was used. It is interesting that the peaks and valleys in the actual norm have corresponding peaks and valleys in the lower bound. To compute the lower bound we differentiated $f(z)$ symbolically and then used a global search followed by bisection to find the maximum value of $\left|\sum_{i=0}^{n-1} \frac{f^{(i)}(0)}{i!} z^i\right|$ on the circle about the origin of radius $1 - \log(2n)/n$.

Finally, consider the GMRES algorithm applied to a Jordan block $J(\lambda)$ (or to any matrix unitarily similar to $J(\lambda)$). It follows from (3) and Theorem 3.1 that ideal

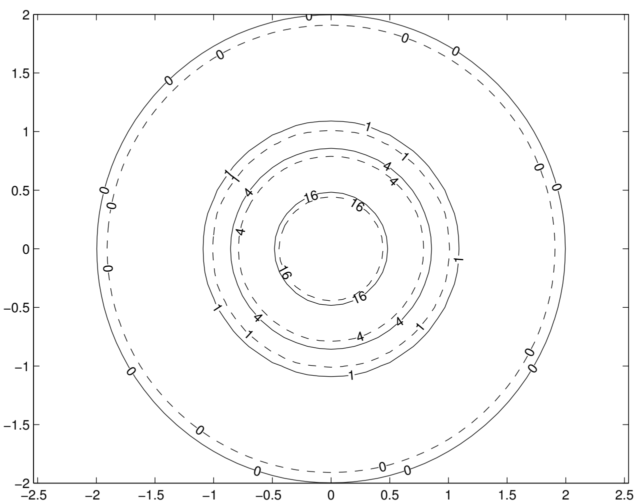

FIG. 2. *Contour plot of the logarithm base 10 of the resolvent norm (solid) and the lower bound (16) (dashed) for a* 50 *by* 50 *Jordan block with eigenvalue* 0.

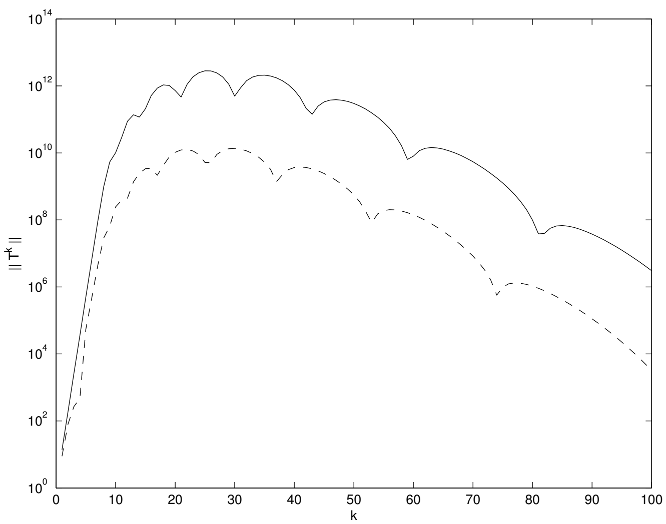

FIG. 3. *Norms of powers of an* 18 *by* 18 *triangular Toeplitz matrix with symbol (17). Solid line is* $\|A^k\|$*; dashed line is lower bound from (8) applied to* $J(0)$ *with* $f(z) = [s(z)]^k$.

GMRES(k) converges if and only if $|\lambda| > r_{k,n}$; that is,

$$\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J(\lambda))\| < 1 \quad \Longleftrightarrow \quad 0 \notin \mathcal{H}_k(J(\lambda)) \quad \Longleftrightarrow \quad |\lambda| > r_{k,n}.$$

Since $J(\lambda)$ is a triangular Toeplitz matrix, it follows from [5] that

$$\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J(\lambda))\| < 1 \quad \Longleftrightarrow \quad \max_{\|w\|=1} \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J(\lambda))w\| < 1.$$

so one obtains the same criterion for the convergence of ordinary GMRES(k), with the worst possible initial vector.

A lower bound on the rate of convergence of ideal GMRES is obtained by noting that for any polynomial $p \in \mathcal{P}_k$ with $p(0) = 1$,

$$\|p(J(\lambda))\| \geq \max_{z \in \mathcal{H}_k(J(\lambda))} |p(z)| = \max_{|z-\lambda| \leq r_{k,n}} |p(z)|.$$

When $|\lambda| > r_{k,n}$, the right-hand side of this inequality is minimized by taking $p(z) = (z-\lambda)^k/(-\lambda)^k$. See, for instance, [19, Lemma 6.26, p. 201]. It follows that for $|\lambda| > r_{k,n}$,

$$\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J(\lambda))\| \geq \left|\frac{r_{k,n}}{\lambda}\right|^k \geq \left|\frac{1-\log(2n)/n}{\lambda}\right|^k.$$

An upper bound on $\min_{p \in \mathcal{P}_k,\ p(0)=1} \|p(J(\lambda))\|$ is obtained by considering an infinite Jordan block J_∞. For $|\lambda| > 1$, we have

$$\min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J(\lambda))\| \leq \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \|p(J_\infty(\lambda))\| = \min_{\substack{p \in \mathcal{P}_k \\ p(0)=1}} \max_{|z-\lambda| \leq 1} |p(z)| = \left|\frac{1}{\lambda}\right|^k.$$

For related results involving specific initial vectors, see [12] or [14].

5. The Field of Values of a Toeplitz Matrix. While we have not yet been able to derive expressions for all of the polynomial numerical hulls of an arbitrary Toeplitz matrix, we can use the result about Jordan blocks to derive bounds on the hull of degree 1 (i.e., the field of values) of an arbitrary Toeplitz matrix. It is shown in [18] that the field of values of an n by n Toeplitz matrix approaches the closure of the field of values of the infinite Toeplitz operator as $n \to \infty$. It is further shown in [2] that for banded Toeplitz matrices with a fixed bandwidth b, the rate of convergence is $O(n^{-2})$. Here we give an explicit inner bound on the field of values of an arbitrary Toeplitz matrix that approaches that of the infinite operator at the rate that $r_{b,n}$ approaches 1. Related work can also be found in [13].

THEOREM 5.1. *Suppose $B = s_1(A) + s_2(A^*)$, where s_1 and s_2 are polynomials of degree b or less and A^* denotes the complex conjugate transpose. Then*

$$\mathcal{H}_1(B) = \mathcal{F}(B) \supset co\left(\{s_1(\zeta) + s_2(\bar{\zeta}) : \ \zeta \in \mathcal{F}_b(A)\}\right).$$

Here $\mathcal{F}_b$ is the set defined in (5), and co denotes the convex hull.

Proof. Let ζ be a point in $\mathcal{F}_b(A)$. Then ζ can be written as q^*Aq for a certain vector q satisfying $q^*q = 1$ and $q^*A^jq = (q^*Aq)^j$, $j = 1, \ldots, b$. For this vector we have

$$q^*Bq = q^*s_1(A)q + q^*s_2(A^*)q = q^*s_1(A)q + \overline{q^*\bar{s}_2(A)q} = s_1(q^*Aq) + s_2(\overline{q^*Aq}).$$

Hence $s_1(\zeta) + s_2(\bar{\zeta})$ lies in the field of values of B. Since this set is convex, it also contains the convex hull of $\{s_1(\zeta) + s_2(\bar{\zeta}) : \ \zeta \in \mathcal{F}_b(A)\}$. $\square$

COROLLARY 5.2. *If A in Theorem 5.1 is a normal matrix or a triangular Toeplitz matrix, then*

$$\mathcal{H}_1(B) = \mathcal{F}(B) \supset co\left(\{s_1(\zeta) + s_2(\bar{\zeta}) : \ \zeta \in \mathcal{H}_b(A)\}\right).$$

Proof. It is shown in [5] that $\mathcal{F}_b(A) = \mathcal{H}_b(A)$ when A is a normal matrix or a triangular Toeplitz matrix. □

COROLLARY 5.3. *If T is an n by n Toeplitz matrix with symbol $s(z) = s_1(z) + s_2(\bar{z})$, where $s_1(z) = \sum_{j=0}^{b} a_j z^j$ and $s_2(z) = \sum_{j=1}^{b} a_{-j} z^j$, then*

$$\mathcal{H}_1(T) = \mathcal{F}(T) \supset co\left(\{s_1(\zeta) + s_2(\bar{\zeta}) : \ |\zeta| \leq r_{b,n}\}\right).$$

Proof. $T = s_1(J) + s_2(J^*)$. □

Since the closure of the field of values of the Toeplitz operator with symbol $s(z) = s_1(z) + s_2(\bar{z}) = \sum_{j=0}^{b} a_j z^j + \sum_{j=1}^{b} a_{-j} \bar{z}^j$ is [10, 13]

$$\mathrm{clos}\left(\mathcal{F}(T_\infty)\right) = \mathrm{co}\left(\{s_1(z) + s_2(\bar{z}) : \ |z| = 1\}\right),$$

Corollary 5.3 implies that the field of values of an n by n Toeplitz matrix with a fixed bandwidth b approaches that of the Toeplitz operator at least at the rate that $r_{b,n}$ approaches 1.

Solutions to the equation $\mathbf{y}'(t) = T\mathbf{y}(t)$ initially grow in norm for certain initial vectors $\mathbf{y}(0) = \mathbf{y}_0$ if and only if the field of values of T extends into the right half plane, since $\frac{d}{dt}\left(\|\mathbf{y}(t)\|^2\right) = 2\Re\langle \mathbf{y}(t), T\mathbf{y}(t)\rangle$. It follows from Corollary 5.3 that there is (at least transient) growth in $\|\mathbf{y}(t)\| = \|e^{tT}\mathbf{y}_0\|$, for a certain $\mathbf{y}_0$, if $\{s_1(\zeta) + s_2(\bar{\zeta}) : \ |\zeta| \leq r_{b,n}\}$ extends into the right half plane. Similar criteria are derived by different means in [1].

6. Further Discussion. The polynomial numerical hull of degree $m-1$ for a matrix A with minimal polynomial of degree m plays a special role: the norm of any function of A can be bounded below by the maximum absolute value of a certain polynomial on this set. For Jordan blocks and triangular Toeplitz matrices of at least moderate size, this lower bound turns out to be a fairly good estimate of $\|f(A)\|$ for a variety of functions f. This might not be the case for other matrices, and better estimates might be obtained by approximating $f(A)$ by a polynomial of some lower degree k and relating $\|f(A)\|$ to the size of this polynomial on the hull of degree k.

It appears to be difficult to determine (theoretically) the polynomial numerical hulls of most matrices. However, as this paper illustrates, once these hulls are known, a great deal of information can be derived easily about the behavior of functions of the matrix.

Acknowledgments: The author thanks Albrecht Böttcher and Nick Trefethen for helpful comments on a draft of this paper. She thanks the anonymous referee for pointing out results in [2] about the rate of convergence of the field of values of a Toeplitz matrix to that of the Toeplitz operator.

REFERENCES

[1] A. Böttcher, *Transient behavior of powers and exponentials of large Toeplitz matrices*, this issue of ETNA.

[2] M. Eiermann, *Fields of values and iterative methods*, Lin. Alg. Appl. 180 (1993), pp. 167–197.

[3] M. Embree and L. N. Trefethen, *Pseudospectra Gateway*, www.comlab.ox.ac.uk/pseudospectra.

[4] V. Faber, A. Greenbaum, and D. E. Marshall, *The polynomial numerical hulls of Jordan blocks and related matrices*, Lin. Alg. Appl., 374 (2003), pp. 231–246.

[5] V. Faber, W. Joubert, M. Knill, and T. Manteuffel, *Minimal residual method stronger than polynomial preconditioning*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 707–729.
[6] G. M. Goluzin, *Some estimates for bounded functions* (Russian), Mat. Sbornik N.S. 26(68), (1950), pp. 7–18.
[7] G. M. Goluzin, *Geometric Theory of Functions of a Complex Variable*, v. 26, Translations of Mathematical Monographs, American Mathematical Society, 1969.
[8] A. Greenbaum, *Generalizations of the field of values useful in the study of polynomial functions of a matrix*, Lin. Alg. Appl. 347 (2002), pp. 233–249.
[9] A. Greenbaum and L. N. Trefethen, *GMRES/CR and Arnoldi/Lanczos as matrix approximation problems*, SIAM J. Sci. Comput., 15 (1994), pp. 359–368.
[10] P. R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, 1982.
[11] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
[12] I. C. F. Ipsen, Expressions and bounds for the GMRES residual, BIT, 40 (2000), pp. 524–535.
[13] E. M. Klein, *The numerical range of a Toeplitz operator*, Proc. Amer. Math. Soc. 35 (1972), pp. 101–103.
[14] J. Liesen and Z. Strakos, *Convergence of GMRES for tridiagonal Toeplitz matrices*, to appear in SIAM J. Matrix Anal.
[15] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.
[16] O. Nevanlinna, *Hessenberg matrices in Krylov subspaces and the computation of the spectrum*, Numer. Funct. Anal. and Optimiz., 16 (1995), pp. 443–473.
[17] L. Reichel and L. N. Trefethen, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices*, Lin. Alg. Appl. 162-164 (1992), pp. 153–185.
[18] S. Roch, *Numerical ranges of large Toeplitz matrices*, Lin. Alg. Appl. 282 (1998), pp. 185–198.
[19] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2003.
[20] I. Schur and G. Szegö, *Über die Abschnitte einer im Einheitskreise beschränkten Potenzreihe*, Sitzungsberichte Preuss. Akad. Wiss. Berlin (1925), pp. 545–560.