The Polynomial Numerical Hulls of Jordan Blocks and Related Matrices

Vance Faber* Anne Greenbaum† Donald E. Marshall‡

June 2, 2003

Abstract

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

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

While these sets have been computed numerically for a number of matrices, the computations have not been verified analytically in most cases.

In this paper we show analytically that the 2-norm polynomial numerical hulls of degrees 1 through $n-1$ for an n by n Jordan block are disks about the eigenvalue with radii approaching 1 as $n \to \infty$, and we prove a theorem characterizing these radii $r_{k,n}$. In the special case where $k = n-1$, this theorem leads to a known result in complex approximation theory: For n even, $r_{n-1,n}$ is the positive root of $2r^n + r - 1 = 0$, and for n odd, it satisfies a similar formula. For large n, this means that $r_{n-1,n} \approx 1 - \log(2n)/n + \log(\log(2n))/n$. These results are used to obtain bounds on the polynomial numerical hulls of certain degrees for banded triangular Toeplitz matrices and for block diagonal matrices with triangular Toeplitz blocks.

*Cytoprint, Inc., 3900 Paseo Del Sol Suite 351, Santa Fe, NM 87507; Mapping Science, Inc., 2525 221st Ave. NE, Sammamish, WA 98074.

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

‡University of Washington, Mathematics Dept., Box 354350, Seattle, WA 98195. This work was supported in part by NSF grant DMS-0201435.

1 Introduction

The *polynomial numerical hull of degree* k for an n by n matrix A was introduced by Nevanlinna in [12, 13] and further studied by Greenbaum in [8, 9]. 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. It is defined as

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

where k is a positive integer and $\mathcal{P}_k$ denotes the set of polynomials of degree k or less.

A few simple properties of these sets are easily observed. The polynomial numerical hull of any degree contains the spectrum of A since if λ is an eigenvalue and v a corresponding normalized eigenvector of A, then $p(A)v = p(\lambda)v$ implies $\|p(A)\| \geq |p(\lambda)|$ for any polynomial p and any matrix norm compatible with the given vector norm. In this paper, the matrix norm will always be the one induced by a given vector norm ($\|B\| = \sup_{\|w\|=1} \|Bw\|$), and the vector norm, unless otherwise stated, will be the 2-norm. For k greater than or equal to the degree of the minimal polynomial of A, the polynomial numerical hull of degree k consists precisely of the spectrum since the minimal polynomial of A has roots only at the eigenvalues but satisfies $\|p(A)\| = 0$. The polynomial numerical hull of any degree is a closed bounded set; it is a subset, for instance, of $\{z \in \mathbf{C} : \ |z| \leq \|A\|\}$.

The following theorem lists some elementary properties of the 2-norm polynomial numerical hull of degree k. Some of these can be found in [8] or [13]:

Theorem 1.

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

(ii) If Q is an n by j matrix with orthonormal columns and if $AQ = QB$ for some j by j matrix B, then $\mathcal{G}_k(B) \subset \mathcal{G}_k(A)$.

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

(iv) If $\mathcal{F}(A)$ denotes the *field of values* ($\mathcal{F}(A) = \{q^*Aq : \ q^*q = 1\}$) and $\sigma(A)$ the spectrum, then $\mathcal{F}(A) = \mathcal{G}_1(A) \supset \mathcal{G}_2(A) \supset \ldots \supset \mathcal{G}_m(A) =$

$\mathcal{G}_{m+1}(A) = \ldots = \sigma(A)$, where m is the degree of the minimal polynomial of A.

(v) Let $\mathrm{pco}_k(S)$ denote the *polynomially convex hull of degree* k for a compact set $S \subset \mathbf{C}$:

$$\mathrm{pco}_k(S) \equiv \{\zeta \in \mathbf{C} : \ |p(\zeta)| \leq \max_{z \in S} |p(z)| \ \ \forall p \in \mathcal{P}_k\}. \tag{2}$$

If $\mathcal{G}_k(A) \supset S$, then $\mathcal{G}_k(A) \supset \mathrm{pco}_k(S)$. If, for all $p \in \mathcal{P}_k$, $\|p(A)\| = \max_{z \in S} |p(z)|$, then $\mathcal{G}_k(A) = \mathrm{pco}_k(S)$.

(vi) If A is a block diagonal matrix:

$$A = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_\ell \end{pmatrix},$$

then

$$\mathcal{G}_1(A) = \mathrm{co}\left(\bigcup_{i=1}^{\ell} \mathcal{G}_1(B_i)\right) = \mathrm{pco}_1\left(\bigcup_{i=1}^{\ell} \mathcal{G}_1(B_i)\right),$$

where $\mathrm{co}(\cdot)$ denotes the convex hull. For $k > 1$,

$$\mathcal{G}_k(A) \supset \mathrm{pco}_k\left(\bigcup_{i=1}^{\ell} \mathcal{G}_k(B_i)\right).$$

Proof:

(i) This follows from the invariance of the 2-norm under unitary similarity transformations: If Q is a unitary matrix and p any polynomial, then $p(Q^*AQ) = Q^*p(A)Q \Rightarrow \|p(Q^*AQ)\| = \|p(A)\| \Rightarrow \mathcal{G}_k(Q^*AQ) = \mathcal{G}_k(A)$.

(ii) If the columns of Q form an orthonormal basis for a j dimensional invariant subspace of A; i.e., if $AQ = QB$ for some j by j matrix B, then for any polynomial p we have $p(A)Q = Qp(B)$, or, $p(B) = Q^*p(A)Q$. It follows that $\|p(B)\| \leq \|p(A)\|$ and hence that $\mathcal{G}_k(B) \subset \mathcal{G}_k(A)$ for any $k \geq 1$.

(iii) If $\beta = 0$, the result $\mathcal{G}_k(\alpha I) = \{\alpha\}$ is obvious, 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 $\zeta \in \mathcal{G}_k(A)$ if and only if $\|p(A)\| \geq |p(\zeta)|\ \forall p \in \mathcal{P}_k$ if and only if $\|q(\alpha I + \beta A)\| \geq |q(\alpha + \beta\zeta)|\ \forall q \in \mathcal{P}_k$ if and only if $\alpha + \beta\zeta \in \mathcal{G}_k(\alpha I + \beta A)$.

(iv) For a proof that $\mathcal{G}_1(A) = \mathcal{F}(A)$, see [8] or [12]. The inclusions $\mathcal{G}_j(A) \supset \mathcal{G}_{j+1}(A)$ are clear from definition (1), and it has already been noted that for $j \geq m$, $\mathcal{G}_j(A) = \sigma(A)$.

(v) If $\mathcal{G}_k(A) \supset S$, then it is clear from definition (1) that $\mathcal{G}_k(A) \supset \mathrm{pco}_k(S)$. If for all $p \in \mathcal{P}_k$, $\|p(A)\| = \max_{z \in S} |p(z)|$, then $\mathcal{G}_k(A)$ contains S and hence $\mathrm{pco}_k(S)$. On the other hand, if $\zeta \notin \mathrm{pco}_k(S)$, then there is a polynomial $p \in \mathcal{P}_k$ such that $|p(\zeta)| > \max_{z \in S} |p(z)| = \|p(A)\|$, so $\zeta \notin \mathcal{G}_k(A)$. Therefore $\mathcal{G}_k(A) = \mathrm{pco}_k(S)$.

(vi) Since $\mathcal{G}_1(A) = \mathcal{F}(A)$, the first statement here is a known property of the field of values [10, p. 12], and it follows from the convexity of that set. In general, we have $\|p(A)\| = \max_{i=1,\ldots,\ell} \|p(B_i)\|$ and so $\mathcal{G}_k(A) \supset \cup_{i=1}^{\ell} \mathcal{G}_k(B_i)$ and by (v) it contains also the polynomially convex hull of degree k for this union of sets. $\square$

In [11] and [8] a number of numerical examples were presented in which the 2-norm polynomial numerical hulls of various degrees for different matrices were computed. In particular, it was demonstrated numerically in [8] that the 2-norm polynomial numerical hulls of degrees 1 through $n-1$ for an n by n Jordan block resemble disks about the eigenvalue of radius slightly less than 1. In this paper we establish this result analytically. The problem of determining the radius of the polynomial numerical hull of degree $n-1$ for an n by n Jordan block turns out to be equivalent to a classical problem in complex approximation theory, closely related to the Carathéodory-Fejér interpolation problem [2, 5] and explicitly solved by Schur and Szegö [15] and then rediscovered with a different proof by Goluzin [6], [7, Theorem 6, pp. 522–523]. Specifically, the result states that if n is even then this radius is the positive root of

$$2r^n + r - 1 = 0,$$

while if n is odd, the radius is the largest value of r satisfying

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

This latter condition is most stringent when $d/(n-1) \approx 0$ and $(\pi - d)/n \approx 0$, and then it is almost the same as the first. In either case, for large n, $r \approx 1 - \log(2n)/n + \log(\log(2n))/n$. In this paper, we prove a somewhat more general result that characterizes the radii of the polynomial numerical hulls of all degrees $k < n$ and leads immediately to the result for $k = n - 1$.

In section 3, the results for Jordan blocks are used to obtain bounds on the polynomial numerical hulls of certain degrees for banded triangular Toeplitz matrices and for block diagonal matrices whose blocks are triangular Toeplitz matrices. In section 4, some comparisons are made with ϵ-pseudospectra.

2 Main Theorems

Let A be an n by n Jordan block with eigenvalue 0:

$$A = \begin{pmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}. \tag{3}$$

This is the only Jordan block that we need consider since by Theorem 1 (iii) the polynomial numerical hulls of a Jordan block with eigenvalue λ are just those of this matrix, shifted by λ. More generally, knowing $\mathcal{G}_k(A)$, we can use (iii) to determine the polynomial numerical hull of degree k for any bidiagonal Toeplitz matrix:

$$\begin{pmatrix} \lambda & \beta & & \\ & \ddots & \ddots & \\ & & \ddots & \beta \\ & & & \lambda \end{pmatrix}.$$

It was noted in Theorem 1 (iv) that the polynomial numerical hull of degree 1 is just the field of values. The field of values of a Jordan block is known [10, p. 25,45]. For A of the form (3), the field of values is a disk about the origin of radius $\rho(A+A^T)/2$, where $\rho(\cdot)$ denotes the spectral radius. The eigenvalues of $(A+A^T)/2$ are $\cos(j\pi/(n+1))$, $j = 1,\ldots,n$, so the spectral radius is $\cos(\pi/(n+1))$, which, for large n, is approximately $1-\frac{1}{2}(\pi/(n+1))^2$.

Lemma 2. Let A be an n by n Jordan block with eigenvalue 0. For any $k < n$, the polynomial numerical hull of degree k for A is a disk about the origin of radius $r_{k,n} < 1$.

Proof: First note that since $\mathcal{G}_k(A) \subset \mathcal{G}_1(A)$, it is contained in the disk about the origin of radius $\cos(\pi/(n+1)) < 1$.

The result follows from the fact that for any angle θ, the Jordan block A is similar to $e^{\imath\theta}A$, via the unitary similarity transformation $A = D(e^{\imath\theta}A)D^{-1}$, where $D = \operatorname{diag}(1, e^{\imath\theta}, \ldots, e^{\imath(n-1)\theta})$. It follows from Theorem 1 (i,iii) that $\mathcal{G}_k(A) = \mathcal{G}_k(e^{\imath\theta}A) = e^{\imath\theta}\mathcal{G}_k(A)$. This means that if $\mathcal{G}_k(A)$ contains any point ζ with $|\zeta| > 0$, then it contains the circle $\{|\zeta|e^{\imath\theta} : \ 0 \le \theta < 2\pi\}$, and then by the maximum principle it must contain the disk $\{z : \ |z| \le |\zeta|\}$. $\square$

Let $p(z) = \sum_{j=0}^{k} c_j z^j$ be an arbitrary polynomial of degree k or less. If A is the n by n matrix in (3), then $p(A)$ is the upper triangular Toeplitz matrix:

$$T_c = \begin{pmatrix} c_0 & \ldots & c_k & 0 & \ldots & 0 \\ & \ddots & & \ddots & \ddots & \vdots \\ & & \ddots & & \ddots & 0 \\ & & & \ddots & & c_k \\ & & & & \ddots & \vdots \\ & & & & & c_0 \end{pmatrix}. \tag{4}$$

Theorem 3. (Carathéodory-Fejér) [2, 5] The matrix T_c in (4) is a contraction (i.e., $\|T_c\| \le 1$) if and only if there is a function $h(z) \in H^\infty$ such that

$$\|\sum_{j=0}^{k} c_j z^j + z^n h(z)\|_\infty \le 1.$$

Here H^∞ denotes the Hardy space of bounded analytic functions in the open unit disk $\mathbf{D}$, with $\|f\|_\infty = \sup_{|z|<1} |f(z)|$.

Since the matrix $T_c/\|T_c\|$ is always a contraction, the Carathéodory-Fejér theorem tells us that there is always a function analytic in the unit disk and bounded in absolute value by 1 there for which $\sum_{j=0}^{k}(c_j/\|T_c\|)z^j$ are the first k terms in the power series expansion; that is, there is a function $h \in H^\infty$ such that $\|\sum_{j=0}^{k} c_j z^j + z^n h(z)\|_\infty \leq \|T_c\|$. On the other hand, if, for some $h \in H^\infty$, $\|\sum_{j=0}^{k} c_j z^j + z^n h(z)\|_\infty$ were strictly less than $\|T_c\|$, it would imply that T_c divided by this smaller number was a contraction, which is not the case. Therefore an equivalent statement of the Carathéodory-Fejér theorem is:

$$\|T_c\| = \min_{h \in H^\infty} \|\sum_{j=0}^{k} c_j z^j + z^n h(z)\|_\infty. \tag{5}$$

Our problem is to find the radius $r_{k,n}$ of the disk in which the kth partial sum $\sum_{j=0}^{k} c_j z^j$ is bounded in absolute value by $\|T_c\|$. For $k = n-1$, this problem was solved by Schur and Szegö [15] and then rediscovered with a different proof by Goluzin [6], [7, Theorem 6, pp. 522–523]. Here we present a somewhat more general result that characterizes the radii $r_{k,n}$, $k \leq n-1$, and leads to the result of Schur and Szegö for $k = n-1$. We could not find a way to derive the more general result with the methods of [6], [7], or [15].

Theorem 4. For given $r \in (0,1)$ and $k \leq n-1$, let

$$K(z) = \sum_{j=0}^{k} r^j z^j, \quad z \in \partial\mathbf{D}. \tag{6}$$

Then $r \leq r_{k,n}$ if and only if there exists a polynomial q of degree less than $n-k-1$ such that

$$\Re\left(K(z) - \frac{1}{2} + z^{k+1}q(z)\right) \geq 0, \quad \forall z \in \partial\mathbf{D}. \tag{7}$$

Proof: Let $I = \{j \in \mathbf{Z} : \ k+1 \leq j \leq n-1\}$, let C denote the space of continuous functions on the unit circle, and let

$$X = \{f \in C : \ \hat{f}(j) = 0 \quad \text{if either} \quad j < 0 \quad \text{or} \quad j \in I\}.$$

Here $\hat{f}(j)$ denotes the Fourier coefficient, so that X consists of all functions of the form $f(z) = p(z) + z^n h(z)$, where $p \in \mathcal{P}_k$ and $h \in H^\infty$. By the F. and M. Riesz theorem

$$X^\perp = \{g \in L^1 : \ \hat{g}(\ell) = 0 \quad \text{if both} \quad \ell \leq 0 \quad \text{and} \quad -\ell \notin I\}.$$

Define a linear functional ϕ on X by

$$\phi(f) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{\imath\theta}) \overline{K(e^{\imath\theta})} \, d\theta. \tag{8}$$

Note that if $f(z) = p(z) + z^n h(z)$, where $p \in \mathcal{P}_k$ and $h \in H^\infty$, then $\phi(f) = p(r)$. Since by (5) we have, for any $p \in \mathcal{P}_k$, $\|p(A)\| = \min_{h \in H^\infty} \|p(z) + z^n h(z)\|$, and by the definition of ϕ we have $|p(r)| = |\phi(p(z) + z^n h(z))|$ for any $h \in H^\infty$, we will have $|\phi(f)| \leq \|f\|_\infty \;\; \forall f \in X$ (i.e., $\|\phi\| \leq 1$) if and only if $|p(r)| \leq \|p(A)\|$ for all $p \in \mathcal{P}_k$. This is the condition that r be in $\mathcal{G}_k(A)$; i.e., that $r \leq r_{k,n}$.

For $f \in X$,

$$\begin{aligned} |\phi(f)| &= \frac{1}{2\pi} \left| \int_0^{2\pi} f(e^{\imath\theta}) \overline{K(e^{\imath\theta})} \, d\theta \right| \\ &= \frac{1}{2\pi} \left| \int_0^{2\pi} f(e^{\imath\theta}) (\overline{K(e^{\imath\theta})} + g(e^{\imath\theta})) \, d\theta \right|, \quad \forall g \in X^\perp \\ &\leq \|f\|_\infty \inf_{g \in X^\perp} \|\overline{K} + g\|_{L^1}. \end{aligned}$$

Since $\phi(1) = 1$, it follows that $\|\phi\| \geq 1$. If $\|\phi\| = 1$, then by compactness, there exists $g \in X^\perp$ such that

$$1 = \|\phi\| = \|\overline{K} + g\|_{L^1} = \frac{1}{2\pi} \int_0^{2\pi} 1 \cdot \left(\overline{K(e^{\imath\theta})} + g(e^{\imath\theta}) \right) \, d\theta.$$

This implies that $\overline{K(z)} + g(z) \geq 0$ for all $z \in \partial\mathbf{D}$. Looking at the Fourier coefficents of $\overline{K(z)} + g(z)$, we conclude that $g(z) = K(z) - 1 + z^{k+1} q(z) + \overline{z^{k+1} q(z)}$, where q is a polynomial of degree less than $n-k-1$. The condition $\overline{K(z)} + g(z) \geq 0$ is then equivalent to that in (7). Conversely, if (7) holds, then there is a function $g(z)$ of the form $g(z) = K(z) - 1 + z^{k+1} q(z) + \overline{z^{k+1} q(z)}$, where $\deg(q) < n - k - 1$, such that $\overline{K(z)} + g(z) \geq 0$ for all $z \in \partial\mathbf{D}$. In this case, for any $f \in X$,

$$|\phi(f)| \leq \|f\|_\infty \|\overline{K} + g\|_{L^1} = \|f\|_\infty \frac{1}{2\pi} \int_0^{2\pi} (\overline{K(e^{\imath\theta})} + g(e^{\imath\theta})) \, d\theta = \|f\|_\infty,$$

and $\|\phi\| = 1$. $\square$

Taking $k = n - 1$ and replacing z by $e^{\imath\theta}$ in (7), one can find an explicit characterization of the largest r for which the left-hand side of (7) is greater than or equal to 0 for all θ. Summing the geometric series and using some algebra we can write:

$$\begin{aligned}\Re\left(\sum_{j=0}^{n-1} r^j e^{\imath j\theta} - \frac{1}{2}\right) &= \Re\left(\frac{r^n e^{\imath n\theta} - 1}{r e^{\imath\theta} - 1}\right) - \frac{1}{2} \\ &= \frac{1}{2}\left[\frac{r^n e^{\imath n\theta} - 1}{r e^{\imath\theta} - 1} + \frac{r^n e^{-\imath n\theta} - 1}{r e^{-\imath\theta} - 1}\right] - \frac{1}{2} \\ &= \frac{2r^{n+1}\cos((n-1)\theta) - 2r^n\cos(n\theta) + 1 - r^2}{2|re^{\imath\theta} - 1|^2},\end{aligned}$$

and so the condition becomes

$$2r^{n+1}\cos((n-1)\theta) - 2r^n\cos(n\theta) - r^2 + 1 \geq 0 \quad \forall\theta. \tag{9}$$

When n is even, the left-hand side of (9) is minimal when $\theta = \pi$, and so we need:

$$1 - r^2 - 2r^n - 2r^{n+1} = (1+r)(1 - r - 2r^n) \geq 0.$$

The factor $(1+r)$ is always positive and the factor $(1 - r - 2r^n)$ is positive for $r = 0$ and is a decreasing function of r, so the solution to our problem is the positive root of

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

For n odd, equation (10) is a sufficient condition on r to make the left-hand side of (9) nonnegative, but one may be able to find a slightly larger value of r for which this holds. Writing θ in the form $\theta = (n+1)\pi/n + d/(n(n-1))$ for some d, inequality (9) becomes

$$2r^{n+1}\cos\left(n\pi - \frac{\pi - d}{n}\right) - 2r^n\cos\left((n+1)\pi + \frac{d}{n-1}\right) - r^2 + 1 =$$

$$-2r^{n+1}\cos\left(\frac{\pi - d}{n}\right) - 2r^n\cos\left(\frac{d}{n-1}\right) - r^2 + 1 \geq 0.$$

Dividing by $(1+r)$, as was done in the case of even n, this becomes

$$1 - r - 2r^n + 2r^n\frac{[1 - \cos(d/(n-1))] + r[1 - \cos((\pi - d)/n)]}{1+r} \geq 0 \quad \forall d. \tag{11}$$

This condition is most stringent when $d/(n-1) \approx 0$ and $(\pi - d)/n \approx 0$, and then it is almost the same as (10). For example, taking $d = (n-1)\pi/(2n-1)$, so that $d/(n-1) = (\pi - d)/n$, we find the necessary condition:

$$1 - r - 2r^n \cos(\pi/(2n-1)) \geq 0. \tag{12}$$

This establishes the theorem of Schur and Szegö[1]:

Corollary 5. The polynomial numerical hull of degree $n-1$ for the n by n Jordan block (3) is a disk about the origin of radius $r_{n-1,n}$ where: for n even, $r_{n-1,n}$ is the positive root of

$$2r^n + r - 1 = 0,$$

and for n odd, $r_{n-1,n}$ is greater than or equal to the positive root of this equation and is the largest value of r that satisfies (11).

It remains to say something about how the roots of equation (10) behave for n large. This result also can be found in [15].

Theorem 6. For large n,

$$r_{n-1,n} = 1 - \frac{\log(2n)}{n} + \frac{\log(\log(2n))}{n} - \frac{\epsilon_n}{n}, \tag{13}$$

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

Proof: First assume that n is even. We will show that if $r = 1 - \log(2n)/n + \log(\log(2n))/n$ then $2r^n + r - 1$ is positive for sufficiently large n, while if $r = 1 - \log(2n)/n + \log(\log(2n))/n - c/n$, for some constant $c > 0$, then $2r^n + r - 1$ is negative for sufficiently large n. It will follow that the root $r_{n-1,n}$ lies between these two values when n is sufficiently large and so satisfies (13).

First let $r = 1 - \log(2n)/n + \log(\log(2n))/n$. Then

$$2r^n + r - 1 > 0 \iff n \log r > \log\left(\frac{1-r}{2}\right) \iff n \log\left(\frac{1}{r}\right) < \log\left(\frac{2}{1-r}\right)$$

[1] Schur and Szegö [15] and later Goluzin [6], [7, Theorem 6, pp. 522–523] proved that if r is as described in Corollary 5, then for any $f \in H^\infty$ with $\|f\|_\infty \leq 1$, if $f(z) = p(z) + z^n h(z)$, where $p \in \mathcal{P}_{n-1}$, then $|p(z)| \leq 1$ in the disk of radius r. They did not explicitly combine this with Theorem 3 to obtain information about the norm of the Toeplitz matrix.

$$\iff n\log\left(\frac{1}{1-\frac{\log(2n)}{n}+\frac{\log(\log(2n))}{n}}\right) < \log\left(\frac{2n}{\log(2n)-\log(\log(2n))}\right). \quad (14)$$

We will use the expansion

$$\log\left(\frac{1}{1-x}\right) = x + \frac{x^2}{2} + \frac{x^3}{3} + \dots, \quad |x| < 1,$$

and the resulting inequalities

$$x < \log\left(\frac{1}{1-x}\right) < x + x^2, \quad \text{for } x > 0 \text{ and sufficiently small.} \quad (15)$$

Let $y = \log(2n) - \log(\log(2n))$. It follows from (15) that, for sufficiently large n, the left-hand side of (14) is less than

$$y + \frac{1}{n}y^2,$$

while the right-hand side satisfies

$$\log\left(\frac{2n/\log(2n)}{1-\log(\log(2n))/\log(2n)}\right) = y + \log\left(\frac{1}{1-\log(\log(2n))/\log(2n)}\right) >$$

$$y + \frac{\log(\log(2n))}{\log(2n)}.$$

Hence inequality (14) will hold provided

$$\frac{1}{n}y^2 = \frac{1}{n}\left(\log(2n) - \log(\log(2n))\right)^2 < \frac{\log(\log(2n))}{\log(2n)}.$$

Since the $1/n$ term on the left dominates all of the log terms, this inequality will hold for n sufficiently large.

Now suppose $r = 1 - \log(2n)/n + \log(\log(2n))/n - c/n$ for some constant $c > 0$. Using the inequality

$$\left(1 - \frac{x}{n}\right)^n \leq e^{-x},$$

which holds for all x and $n \geq 1$, we find that

$$2r^n + r - 1 \leq \frac{\log(2n)}{n}(e^{-c} - 1) + \frac{\log(\log(2n)) - c}{n}.$$

For any $c > 0$, the right-hand side is negative for n sufficiently large. This completes the proof of (13) when n is even.

For n odd, the radius $r_{n-1,n}$ is greater than or equal to the positive root of (10) so, by the above arguments, it is greater than or equal to $1-\log(2n)/n+\log(\log(2n))/n - c/n$ for any constant $c > 0$. It must satisfy inequality (12), however, which is equivalent to the inequality

$$n \log\left(\frac{1}{r}\right) \geq \log\left(\frac{2}{1-r}\right) - \log\left(\frac{1}{\cos\left(\frac{\pi}{2n-1}\right)}\right) = \log\left(\frac{2}{1-r}\right) - O(n^{-2}). \quad (16)$$

It was shown above that if $r = 1 - \log(2n)/n + \log(\log(2n))/n$, then for n sufficiently large

$$n \log\left(\frac{1}{r}\right) < \log\left(\frac{2}{1-r}\right),$$

and the difference is at least

$$\log\left(\frac{2}{1-r}\right) - n\log\left(\frac{1}{r}\right) > \frac{\log(\log(2n))}{\log(2n)} - \frac{1}{n}\left(\log(2n) - \log(\log(2n))\right)^2 >> O(n^{-2}).$$

Hence inequality (16) is not satisfied, so $r_{n-1,n}$ must be less than $1-\log(2n)/n+\log(\log(2n))/n$. $\square$

Theorem 4 can be used to derive additional information about the radii $r_{k,n}$, $k < n-1$, as well. For example, we can show that for any k and n, $r_{k,n} \leq \sqrt{r_{[k/2],[(n+1)/2]}}$, where $[\cdot]$ denotes the integer part. To see this, let $H(z) = K(z) - 1/2 + z^{k+1}q(z)$ be the function in (7). If the real part of $H(z)$ is nonnegative for all $z \in \partial\mathbf{D}$, then the same holds for the real part of $(H(z) + H(-z))/2$, and

$$\frac{H(z) + H(-z)}{2} = \frac{1}{2} + r^2 z^2 + \ldots + r^{2[k/2]} z^{2[k/2]} + z^{2[k/2]+2} \hat{q}(z^2),$$

where $\hat{q}$ is a polynomial of degree less than $(n - 2[k/2] - 2)/2$. The condition that this function have nonnegative real part for all $z^2 \in \partial\mathbf{D}$ or, equivalently, for all $z \in \partial\mathbf{D}$, is precisely the condition from Theorem 4 that r^2 be less than or equal to $r_{[k/2],[(n+1)/2]}$:

$$r^2 \leq r_{[k/2],[(n+1)/2]} \quad \text{if and only if}$$

$$\Re\left(\frac{1}{2}+(r^2)z+(r^2)^2z^2+\ldots+(r^2)^{[k/2]}z^{[k/2]}+z^{[k/2]+1}\hat{q}(z)\right)\geq 0 \quad \forall z\in\partial\mathbf{D},$$

for some $\hat{q}$ of degree less than $[(n+1)/2]-[k/2]-1$.

In the special case where $k=n-2$, and $n=2m$, where m is even, we obtain an equality: $r_{n-2,n}=\sqrt{r_{m-1,m}}$. To see this, note that Theorem 4 states that $r\leq r_{n-2,n}$ if and only if there is a constant c such that

$$\Re\left(\frac{1}{2}+rz+r^2z^2+\ldots+r^{n-2}z^{n-2}+cz^{n-1}\right)\geq 0 \quad \forall z\in\partial\mathbf{D}. \qquad (17)$$

Summing the geometric series in (17) and taking its real part gives the equivalent conditions:

$$c\,\Re(z^{2m-1})+\Re\left(\frac{1-r^{2m-1}z^{2m-1}}{1-rz}\right)-\frac{1}{2}\geq 0 \ \ \forall z\in\partial\mathbf{D},$$

or,

$$g(\theta)\equiv c\cos((2m-1)\theta)(1-2r\cos\theta+r^2)-\left[-\frac{1}{2}+\frac{1}{2}r^2+r^{2m}\right]-$$
$$r^{2m-1}\left(\cos((2m-1)\theta)-r(1+\cos((2m-2)\theta))\right)\geq 0 \ \ \forall\theta. \qquad (18)$$

For $r=\sqrt{r_{m-1,m}}$, and m even, the term in brackets is 0, and then $g(\pi/2)=0$. In order that $g(\theta)$ be nonnegative it must therefore be the case that $g'(\pi/2)=0$, and this leads to a formula for c:

$$c=\frac{r^{2m-1}}{1+r^2}. \qquad (19)$$

Now we must show that inequality (18) holds for all θ, when $r=\sqrt{r_{m-1,m}}$ and c is given by (19). The needed inequality can be written as:

$$\frac{1}{1+r^2}\cos((2m-1)\theta)(1-2r\cos\theta+r^2)-\cos((2m-1)\theta)+r+r\cos((2m-2)\theta)\geq 0,$$

and making the substitution $\cos((2m-2)\theta)=\cos((2m-1)\theta)\cos(\theta)+\sin((2m-1)\theta)\sin\theta$ and doing some algebra, this reduces to:

$$(1+r^2)+(1+r^2)\sin((2m-1)\theta)\sin\theta-(1-r^2)\cos((2m-1)\theta)\cos\theta\geq 0,$$

or,

$$1+r^2-\cos(2m\theta)+r^2\cos((2m-2)\theta)\geq 0,$$

which clearly holds for all θ.

3 Generalizations

3.1 Banded Triangular Toeplitz Matrices

Let T be the n by n banded triangular Toeplitz matrix:

$$T = \begin{pmatrix} t_0 & \dots & t_b & & \\ & \ddots & & \ddots & \\ & & \ddots & & t_b \\ & & & \ddots & \vdots \\ & & & & t_0 \end{pmatrix},$$

whose *symbol* is $q(z) = \sum_{j=0}^{b} t_j z^j$. Note that $T = q(A)$ where A is the Jordan block in (3). Consider any polynomial p of degree k or less, where $kb < n$. We have $p(T) = p(q(A))$ and so

$$\|p(T)\| = \|p(q(A))\| \geq |p(q(z))| \quad \forall z \in \mathcal{G}_{kb}(A).$$

It follows that the polynomial numerical hull of degree k for T contains the image under q of the polynomial numerical hull of degree kb for A; that is, it contains the image under q of the closed disk about the origin of radius $r_{kb,n}$ which, for n large and $kb < n$, is most of the unit disk. By Theorem 1 (v), it also contains $\mathrm{pco}_k(q(\mathcal{G}_{kb}(A)))$.

The norm of an infinite triangular Toeplitz matrix is the maximum value of its symbol in the closed unit disk $\bar{\mathbf{D}}$. See, for example, [1]. Let T_∞ denote the infinite triangular Toeplitz matrix whose upper left n by n block is T and whose other diagonals are 0. For any polynomial p, the matrix $p(T_\infty)$ is again Toeplitz and its symbol is just $p(q(z))$. Hence $\|p(T_\infty)\| = \max_{z \in \bar{\mathbf{D}}} |p(q(z))| = \max_{z \in q(\bar{\mathbf{D}})} |p(z)|$. It follows from Theorem 1 (v) that $\mathcal{G}_k(T_\infty) = \mathrm{pco}_k(q(\bar{\mathbf{D}}))$. Now $p(T)$ can be written as $I^*_{\infty \times n} p(T_\infty) I_{\infty \times n}$, where the columns of $I_{\infty \times n}$ are the first n unit vectors in ℓ^2. It follows that $\|p(T)\| \leq \|p(T_\infty)\| \;\; \forall p \in \mathcal{P}_k$ and so $\mathcal{G}_k(T) \subset \mathcal{G}_k(T_\infty)$.

We have shown that when $kb < n$ the polynomial numerical hull of degree k for T lies between the polynomially convex hull of degree k for: the image under q of the closed disk about the origin of radius $r_{kb,n}$ and the image under q of the closed unit disk.

An example is shown in Figure 1. Here T is a Toeplitz matrix of order $n = 50$ with -1's on its first superdiagonal and 1's on its fourth superdiagonal

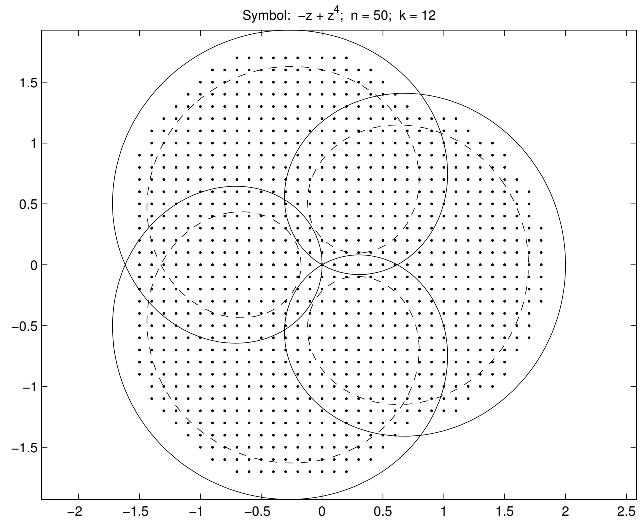

Figure 1: Image of the unit circle (solid) and of the circle of radius .934 (dashed) under the mapping $q(z) = -z + z^4$. Region with dots is the numerically computed polynomial numerical hull of degree 12 for the 50 by 50 Toeplitz matrix with symbol q.

so that $q(z) = -z+z^4$. The outer solid curve in the figure is the image under q of the unit circle and the inner dashed curve is the image under q of the circle of radius $r_{n-1,n} \approx .934$. The polynomial numerical hulls of degrees 1 through 12 for this matrix must contain pco_k of the trefoil-shaped region enclosed by the dashed curve, and they must be contained in pco_k of the trefoil-shaped region enclosed by the solid curve. The dots in the figure represent points in the numerically computed polynomial numerical hull of degree 12 and can be seen to satisfy these inclusions.

3.2 Block Diagonal Matrices with Triangular Toeplitz Blocks

Let A be a matrix in Jordan form or, more generally, a block diagonal matrix with triangular Toeplitz blocks:

$$A = \begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_\ell \end{pmatrix},$$

where each block T_i is of dimension n_i and has symbol $q_i(z)$. Theorem 1 (vi) gives an exact expression for $\mathcal{G}_1(A)$ as the convex hull of $\cup_{i=1}^{\ell} \mathcal{G}_1(T_i)$, but for $k > 1$ it states only that

$$\mathcal{G}_k(A) \supset \operatorname{pco}_k \left(\bigcup_{i=1}^{\ell} \mathcal{G}_k(T_i) \right). \tag{20}$$

In general, this is not an equality.

It was shown in the previous subsection that $\mathcal{G}_k(T_i)$ contains the image under q_i of a disk about the origin of radius r_{kb_i,n_i}, where b_i is the bandwidth of T_i. If $kb_i \geq n_i$, then we interpret this radius to be 0 and the image under q_i of the origin to be the eigenvalue of T_i. Then expression (20) can be replaced by:

$$\mathcal{G}_k(A) \supset \operatorname{pco}_k \left(\bigcup_{i=1}^{\ell} q_i(\{z \in \mathbf{C} : \ |z| \leq r_{kb_i,n_i}\}) \right). \tag{21}$$

We can obtain an outer bound on $\mathcal{G}_k(A)$ by using the fact that for any polynomial p we have $\|p(T_i)\| \leq \|p(T_{i,\infty})\|$, where $T_{i,\infty}$ is the infinite triangular Toeplitz matrix whose upper left n_i by n_i block is T_i and whose other diagonals are 0. The norm of a polynomial function of this infinite Toeplitz matrix is $\|p(T_{i,\infty})\| = \max_{z \in q_i(\bar{\mathbf{D}})} |p(z)|$, as was shown in the previous subsection. Since $\|p(A)\|$ satisfies:

$$\|p(A)\| = \max_{i=1,\ldots,\ell} \|p(T_i)\| \leq \max_{i=1,\ldots,\ell} \|p(T_{i,\infty})\| = \max_{z \in \cup_{i=1}^{\ell} q_i(\bar{\mathbf{D}})} |p(z)|,$$

it follows that

$$\mathcal{G}_k(A) \subset \operatorname{pco}_k \left(\bigcup_{i=1}^{\ell} q_i(\bar{\mathbf{D}}) \right), \tag{22}$$

because if $\zeta \notin \mathrm{pco}_k(\cup_{i=1}^{\ell} q_i(\bar{\mathbf{D}}))$, then there is a polynomial $p \in \mathcal{P}_k$ such that

$$|p(\zeta)| > \max_{z \in \cup_{i=1}^{\ell} q_i(\bar{\mathbf{D}})} |p(z)| \geq \|p(A)\|,$$

so $\zeta \notin \mathcal{G}_k(A)$.

The inner and outer bounds (21) and (22) are equal when the blocks are 1 by 1 (i.e., when A is a diagonal matrix), and in that case expression (20) is an equality. They are close when the block sizes n_i are large and each $kb_i < n_i$, since in that case r_{kb_i,n_i} is close to 1. They may differ significantly if some $kb_i \geq n_i$ and $n_i > 1$.

A very simple example is a 3 by 3 matrix with one Jordan block of size 2 and another of size 1:

$$A = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}, \quad T_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad T_2 = (1)\,. \tag{23}$$

The polynomial numerical hull of degree 1 for this matrix is $\mathrm{co}(\mathcal{G}_1(T_1) \cup \mathcal{G}_1(T_2)) = \mathrm{co}(\{z \in \mathbf{C} : \ |z| \leq \frac{1}{2}\} \cup \{1\})$. The polynomial numerical hull of degree 3 is just the eigenvalues $\{0, 1\}$. By Theorem 1 (vi) or expression (21), the polynomial numerical hull of degree 2 contains $\mathrm{pco}_2(\mathcal{G}_2(T_1) \cup \mathcal{G}_2(T_2)) = \{0, 1\}$, but, in fact, it contains more than this. Expression (22) says that it is contained in $\mathrm{pco}_2(\bar{\mathbf{D}} \cup \{1\}) = \bar{\mathbf{D}}$. It can be shown with some algebra that the polynomial numerical hull of degree 2 actually consists of the segment of the real axis $[-1/3, 1]$.

4 Comparison with Pseudospectra

The *ϵ-pseudospectrum* is another type of set that can be associated with a matrix to give more information than the spectrum alone can provide [4]. It is defined as

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} : \ \|(zI - A)^{-1}\| \geq \epsilon^{-1}\},$$

and an equivalent definition is

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

The pseudospectra of a matrix describe how the eigenvalues change under perturbations to the matrix of various sizes.

It is natural to think that there would be some relation between these sets and the polynomial numerical hulls of various degrees. If a small change in the matrix results in much larger changes in the eigenvalues, then probably eigenvalues do not determine the behavior of polynomial functions of the matrix, and the polynomial numerical hulls of various degrees will contain far more than just the polynomially convex hull of the given degree for the eigenvalues. Numerical computations in [8] show a striking resemblance for certain matrices between the polynomial numerical hull of degree k and pco_k of the ϵ-pseudospectrum, for properly matched values of k and ϵ. The reasons for this close resemblance are still not fully understood.

For Jordan blocks it turns out that for each k there is an ϵ such that the polynomial numerical hull of degree k is identical to the ϵ-pseudospectrum; that is, they are both disks about the eigenvalue and, for a certain ϵ, their radii are equal. Bounds on the pseudospectra of Jordan blocks and, more generally, of Toeplitz matrices were derived analytically in [14]. There 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 $1+\epsilon$. Numerical experiments there suggest that the lower bound is typically much sharper than the upper bound. To determine the approximate value of ϵ that corresponds to the polynomial numerical hull of degree $n-1$ for an n by n Jordan block, then, we can set $\epsilon^{1/n}$ equal to $r_{n-1,n}$. Substituting $\epsilon^{1/n}$ for r in the first term of (10) and $r_{n-1,n} \approx 1 - \log(2n)/n + \log(\log(2n))/n$ in the second term, we obtain the approximation

$$\epsilon \approx \frac{\log(2n)}{2n} - \frac{\log(\log(2n))}{2n}.$$

The polynomial numerical hulls of lower degrees correspond to ϵ-pseudospectra with larger values of ϵ.

There is a similar correspondence for banded triangular Toeplitz matrices. It was shown in [14] that the ϵ-pseudospectrum of such a matrix contains the image, under the symbol, of the disk of radius $(\epsilon/c)^{1/n}$ for a certain constant c that can be taken to be the sum of absolute values of the entries in the first row of the matrix. It is contained in the image, under the symbol, of the unit disk, plus a disk of radius ϵ. Compare this to the bounds of section 3.1 on the polynomial numerical hulls of certain degrees for banded triangular Toeplitz matrices. It remains to be seen whether these results, like those for pseudospectra, can be extended to more general Toeplitz matrices and to polynomial numerical hulls of higher degree.

A *difference* between pseudospectra and polynomial numerical hulls can be seen in the case of block diagonal matrices. The ϵ-pseudospectrum of a block diagonal matrix is just the union of the ϵ-pseudospectra of the blocks, but it was seen in the example of section 3.2 that this is not necessarily so for the polynomial numerical hull of degree k. The polynomial numerical hull of degree k contains pco_k of the union of the hulls of degree k for the blocks, but it may contain more. The polynomial numerical hull of degree 2 for the 3 by 3 matrix in (23) bears little resemblance to any ϵ-pseudospectrum of the matrix. The precise implications of these similarities and differences remain to be explored.

References

[1] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, New York, 1999.

[2] C. Carathéodory and L. Fejér, *Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz*, Rend. Circ. Mat. Palermo, 32 (1911), pp. 218–239.

[3] P. L. Duren, *Theory of H^p Spaces*, Academic Press, New York, 1970.

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

[5] C. Foias and A. E. Frazho, *The Commutant Lifting Approach to Interpolation Problems*, Birkhäuser, Basel, 1990.

[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, *Card shuffling and the polynomial numerical hull of degree k*, submitted to SIAM J. Sci. Comput.

[10] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[11] X. Lü, *On the polynomial numerical hull of a matrix*, Licentiate thesis, Helsinki University of Technology, HUT-95, 1995.

[12] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

[13] O. Nevanlinna, *Hessenberg matrices in Krylov subspaces and the computation of the spectrum*, Numer. Funct. Anal. and Optimiz., 16 (1995), pp. 443–473.

[14] L. Reichel and L. N. Trefethen, *Eigenvalues and pseudo-eigenvalues of Toeplitz matrices*, Lin. Alg. Appl. 162-164 (1992), pp. 153–185.

[15] I. Schur and G. Szegö, *Über die Abschnitte einer im Einheitskreise beschränkten Potenzreihe*, Sitzungsberichte Preuss. Akad. Wiss. Berlin (1925), pp. 545–560.