Generalizations of the Field of Values Useful in the Study of Polynomial Functions of a Matrix

Anne Greenbaum*

October 14, 2001

Abstract

For a given square matrix A and positive integer k, we consider sets Ω in the complex plane satisfying

$$\|p(A)\| \geq \max_{z \in \Omega} |p(z)|,$$

for all polynomials p of degree k or less. The largest such set, referred to as the *polynomial numerical hull of degree k*, was introduced by Nevanlinna [*Convergence of Iterations for Linear Equations*, Birkhäuser, 1993] and a number of properties of this set were derived for both matrices and linear operators. We give several equivalent characterizations of the polynomial numerical hull of degree k and we actually compute these sets for several matrices. For $k = 1$, this set is just the field of values of A, and for $k \geq m$, where m is the degree of the minimal polynomial of A, it is the spectrum of A. For $1 < k < m$, these sets are intermediate between the field of values and the spectrum and sometimes resemble pseudospectra.

1 Introduction

Let A be a given n by n matrix. In order to estimate $\|f(A)\|$ for various functions f, it is helpful if one can associate A with some set in the complex

*University of Washington, Mathematics Dept., Box 354350, Seattle, WA 98195. This work was supported by NSF contract DMS-9802919.

plane and relate $\|f(A)\|$ to the size of f on this set. For normal matrices, an appropriate set is the spectrum of A, since if $A = Q\Lambda Q^*$, where Q is unitary and Λ is a diagonal matrix of eigenvalues, then $f(A) = Qf(\Lambda)Q^*$ and $\|f(A)\| = \|f(\Lambda)\|$, for any norm invariant under unitary similarity transformations. In this paper we will be concerned mainly with the spectral norm and, unless otherwise stated, $\|\cdot\|$ will denote the Euclidean norm for vectors and the induced spectral norm for matrices. Some of the results will be shown to hold more generally, however, for any norm that is greater than or equal to the numerical radius: $\|B\| \geq \nu(B) \equiv \max\{|q^*Bq| : \; q^*q = 1\}$.

For nonnormal matrices, it is less clear what (if any) set(s) in the complex plane should be associated with A. One would like to identify sets $\hat{\Omega}$ providing upper bounds: $\|f(A)\| \leq \max_{z\in\hat{\Omega}} |f(z)|$, as well as sets Ω providing lower bounds: $\|f(A)\| \geq \max_{z\in\Omega} |f(z)|$, for all functions f in some class of interest. In this paper we consider polynomials of a fixed degree k or less and look for sets Ω satisfying

$$\|p(A)\| \geq \max_{z\in\Omega} |p(z)| \tag{1}$$

for all such polynomials p.

Specifically, we define

$$\mathcal{F}_k(A) = \{q^*Aq : \; q^*q = 1 \text{ and } q^*A^jq = (q^*Aq)^j, \; j = 1, \ldots, k\}, \tag{2}$$

and show that this set has the desired property (1). It is shown that this definition is equivalent to

$$\mathcal{F}_k(A) = \{\zeta \in \mathbf{C} : \; 0 \in F(\{(A - \zeta I)^j\}_{j=1}^k)\}, \tag{3}$$

where $F(\{B_j\}_{j=1}^k)$ denotes the k-dimensional generalized field of values [8]:

$$F(\{B_j\}_{j=1}^k) = \left\{ \begin{pmatrix} q^*B_1q \\ \vdots \\ q^*B_kq \end{pmatrix} : \; q^*q = 1 \right\}.$$

Using results from Faber, et al [4], it is shown that the largest set with property (1) is:

$$\mathcal{G}_k(A) = \{\zeta \in \mathbf{C} : \; 0 \in \text{co}[F(\{(A - \zeta I)^j\}_{j=1}^k)]\}, \tag{4}$$

where co[$\cdot$] denotes the convex hull.

The largest set satisfying (1) was earlier considered by Nevanlinna [14] and was referred to as the *polynomial numerical hull of degree k*. A number of properties of this set were derived for both matrices and linear operators [14, 15], but it was never characterized as in (4) or computed for specific matrices. For nonnormal matrices, these sets are, in some ways, the analog of eigenvalues for a normal matrix; for any kth degree polynomial p, we have $\|p(A)\| \geq \|p(B)\|$ for any normal matrix B with eigenvalues spread densely throughout the polynomial numerical hull of degree k, and this does not hold if B has eigenvalues outside this set.

For $k = 1$, each of these sets is the field of values of A, and for $k \geq m$, where m is the degree of the minimal polynomial of A, they are shown to be the spectrum of A. For $1 < k < m$, the sets $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$ are intermediate between the field of values and the spectrum: $\mathcal{F}(A) \supset \mathcal{G}_k(A) \supset \mathcal{F}_k(A) \supset \sigma(A)$, where $\mathcal{F}(\cdot)$ denotes the field of values and $\sigma(\cdot)$ the spectrum.

One application of this analysis is in describing the convergence rate of the GMRES algorithm for solving linear systems $Ax = b$. An upper bound on the residual norm at iteration k of the algorithm is:

$$\min_{p \in \mathcal{P}_k(0)} \|p(A)\|, \tag{5}$$

where $\mathcal{P}_k(0)$ denotes the set of polynomials of degree k or less with value 1 at the origin. This quantity is said to be the residual norm for the *ideal* GMRES algorithm [7]. It provides an upper bound on the actual residual norm in the GMRES algorithm for the worst possible initial residual, but it is not always sharp [4, 17]. The worst-case behavior of the actual GMRES algorithm is governed by

$$\max_{\|b\|=1} \min_{p \in \mathcal{P}_k(0)} \|p(A)b\|. \tag{6}$$

It follows from property (1) that the quantities

$$\min_{p \in \mathcal{P}_k(0)} \max_{z \in \Omega} |p(z)|, \tag{7}$$

where $\Omega = \mathcal{F}_k(A)$ or $\mathcal{G}_k(A)$ provide lower bounds on the ideal GMRES residual norm (5). We present numerical examples to illustrate that they are often good *estimates* of (5) as well.

Related work in the literature has been aimed at finding sets $\hat{\Omega} \subset \mathbf{C}$ such that $\|f(A)\| \leq \gamma \max_{z \in \hat{\Omega}} |f(z)|$, for some moderate size number γ, and either for all analytic functions f or specifically for the polynomial P_k that achieves the minimum in (5). Trefethen has suggested taking $\hat{\Omega}$ to be the ϵ-pseudospectrum of A [21]:

$$\Lambda_\epsilon(A) = \{z \in \mathbf{C} : \ \|(zI - A)^{-1}\| \geq \epsilon^{-1}\}. \tag{8}$$

Expressing $f(A)$ as a Cauchy integral about Γ_ϵ, the boundary of Λ_ϵ, and replacing the norm of the Cauchy integral by the integral of the norm of the integrand, one obtains the following upper bound:

$$\|f(A)\| \leq \frac{\mathcal{L}(\Gamma_\epsilon)}{2\pi\epsilon} \max_{z \in \Lambda_\epsilon} |f(z)|, \tag{9}$$

where $\mathcal{L}(\Gamma_\epsilon)$ denotes the length of Γ_ϵ [21]. To obtain a good estimate of $\|f(A)\|$, one must try to choose ϵ large enough so that the first factor is of moderate size, but small enough so that the set Λ_ϵ is not too large.

More generally, if Γ is the boundary of any set $\hat{\Omega}$ containing the spectrum of A, then we have from the Cauchy integral formula

$$f(A) = \frac{1}{2\pi\imath} \int_\Gamma (zI - A)^{-1} f(z)\, dz,$$

and taking norms on each side,

$$\begin{aligned} \|f(A)\| &\leq \frac{1}{2\pi} \int_\Gamma \|(zI - A)^{-1}\| \ |f(z)| \ |dz| \\ &\leq \left(\frac{1}{2\pi} \int_\Gamma \|(zI - A)^{-1}\| \ |dz|\right) \max_{z \in \hat{\Omega}} |f(z)|. \end{aligned} \tag{10}$$

Numerical computations suggest that the polynomial numerical hull of degree $k < m$ is sometimes very similar to the ϵ-pseudospectrum for a moderate size value of ϵ. Using either (9), or (10) with Γ taken to be the boundary of the polynomial numerical hull of degree k, one can thus obtain realistic upper bounds on $\|p(A)\|$, when p is a polynomial of degree much less than the minimal polynomial of A.

Another approach to obtaining upper bounds on the size of the ideal GMRES polynomial P_k uses the field of values of A. Eiermann has shown that if $\hat{\Omega}$ is a compact convex set containing $\mathcal{F}(A)$ but not containing the origin, then

$$\min_{p \in \mathcal{P}_k(0)} \|p(A)\| \leq c_k \min_{p \in \mathcal{P}_k(0)} \max_{z \in \hat{\Omega}} |p(z)|, \tag{11}$$

where the constant c_k depends on $\hat{\Omega}$ but not on A [2, 3]. For this bound to provide useful information, it is necessary that 0 be outside both $\mathcal{F}(A)$ and $\hat{\Omega}$, but that $\hat{\Omega}$ be chosen large enough so that c_k is of moderate size.

Other sets Ω satisfying $\|P_k(A)\| \geq \min_{p \in \mathcal{P}_k(0)} \max_{z \in \Omega} |p(z)|$ have been identified by Huhtanen and Nevanlinna [9]. They showed that if A differs from a normal matrix B by a matrix of small rank, then (5) is bounded below by the minimum over $p \in \mathcal{P}_k(0)$ of the maximum absolute value of p at certain eigenvalues of B [9]. Huhtanen then extended this analysis to low-rank modifications of the matrix A that yield matrices B with well-conditioned eigenvectors and showed that the quantity in (5) is greater than or equal to one over the condition number of the eigenvectors of B times the maximum absolute value of p on certain eigenvalues of B [10].

In this paper we study the sets $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$ defined in (2) and (4). In section 2, we establish basic properties of these sets and relate $\|p(A)\|$ to the maximum absolute value of p on these sets, for polynomials p of degree k or less. Section 3 describes a method for computing $\mathcal{G}_k(A)$, although it is hoped that better methods can be developed. It also contains numerical examples showing the polynomial numerical hull of degree k for several matrices and various values of k. Section 4 states some conclusions and possible applications.

2 Basic Properties of $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$

Theorem 1. Let $\mathcal{F}_k(A)$ be defined by (2). For any polynomial p of degree k or less we have

$$\|p(A)\| \geq \max_{z \in \mathcal{F}_k(A)} |p(z)|, \tag{12}$$

where $\|\cdot\|$ can be any norm that satisfies $\|B\| \geq \max\{|q^*Bq| : \; q^*q = 1\}$, for all n by n matrices B; e.g., the spectral norm, the Frobenius norm, etc.

Proof: Let $z = q^*Aq$ be an element of $\mathcal{F}_k(A)$. Then for any polynomial p of degree k or less, we have $q^*p(A)q = p(q^*Aq)$, so $\|p(A)\| \geq |q^*p(A)q| = |p(z)|$. □

Theorem 2. The definitions (2) and (3) of $\mathcal{F}_k(A)$ are equivalent.

Proof: Let q be any vector with $q^*q = 1$. The condition $q^*(A - \zeta I)q = 0$

is equivalent to $q^*Aq = \zeta$. If this condition is satisfied, then the condition $q^*(A-\zeta I)^2 q = 0$ is equivalent to $q^*A^2q = \zeta^2 = (q^*Aq)^2$, etc. $\square$

The set $\mathcal{F}_k(A)$ is not quite the largest set for which inequality (1) holds, but the following theorem shows that if $\zeta \notin \mathcal{F}_k(A)$, then for any vector b with $\|b\| = 1$, one can construct a polynomial p for which $\|p(A)b\| < |p(\zeta)|$. This is the more relevant question in analyzing the behavior of the true GMRES algorithm for the worst possible right-hand side vector [4, 17].

Theorem 3. Let $\|\cdot\|$ denote the Euclidean norm for vectors and the induced spectral norm for matrices. If $\zeta \notin \mathcal{F}_k(A)$, then for any vector b with $\|b\| = 1$, there is a polynomial p of degree k or less such that $\|p(A)b\| < |p(\zeta)|$.

Proof: Since $\zeta \notin \mathcal{F}_k(A)$, there is no vector b with norm one satisfying $b^*(A-\zeta I)^j b = 0$, for all $j = 1, \ldots k$; that is, any vector b with norm one has a nonzero projection onto the Krylov space

$$\text{span}\{(A-\zeta I)b, (A-\zeta I)^2 b, \ldots, (A-\zeta I)^k b\}.$$

It follows that there is a polynomial p_{k-1} of degree $k-1$ or less in $(A-\zeta I)$ such that

$$\|b - (A-\zeta I)p_{k-1}(A-\zeta I)b\| < 1.$$

Define $p(z) = 1-(z-\zeta)p_{k-1}(z-\zeta)$, so that $p(\zeta) = 1$ but $\|p(A)b\| < 1$. $\square$

The next corollary shows that the largest set satisfying inequality (1) is the set $\mathcal{G}_k(A)$ defined in (4). It follows easily from the following result in [4], whose proof we include here for completeness (and because it is so pretty!):

Theorem 4 (Faber, et al [4]). For any n by n matrix B and any positive integer k, we have $\min_{p\in\mathcal{P}_k(0)} \|p(B)\| = 1$ if and only if $0 \in \text{co}[F(\{B^j\}_{j=1}^k)]$.

Proof: By the Hahn-Banach theorem, if 0 is not in the convex set $S = \text{co}[F(\{B^j\}_{j=1}^k)]$, then there exists a separating hyperplane for 0 and S; that is, there is a k-vector $c = (c_1, \ldots, c_k)^T$ such that $\text{Re}(c^*w) > 0$ for all $w \in S$. This implies that for any q with $q^*q = 1$, we have

$$\text{Re}\left(\sum_{j=1}^k \bar{c}_j q^* B^j q\right) = \text{Re}\left(q^*(\sum_{j=1}^k \bar{c}_j B^j)q\right) > 0. \tag{13}$$

Define $p_{k-1}(B) = \sum_{j=1}^k \bar{c}_j B^{j-1}$. Then (13) implies that the field of values of $Bp_{k-1}(B)$ lies in the right half-plane, and it follows that there is a positive number α such that $\|p(B)\| \equiv \|I - \alpha B p_{k-1}(B)\| < 1$. To see this, note that if $C = Bp_{k-1}(B)$ then $\|I - \alpha C\|$ is the square root of the largest eigenvalue of $I - \alpha(C + C^*) + \alpha^2 C^* C$, where $C + C^*$ is positive definite. For α sufficiently small (e.g., $\alpha < \lambda_{min}(C + C^*)/\lambda_{max}(C^* C)$), this matrix will have all its eigenvalues less than 1.

Conversely, if $0 \in S$, then there is no separating hyperplane for 0 and S, or even for 0 and $F(\{B^j\}_{j=1}^k)$. That is, there is no vector c for which the quantity in (13) has the same sign for all vectors q with $q^* q = 1$. Hence the field of values of any polynomial of the form $Bp_{k-1}(B)$, where p_{k-1} is any polynomial of degree $k - 1$ or less, must contain the origin. It follows that for any $p \in \mathcal{P}_k(0)$, since $p(z)$ must be of the form $p(z) = 1 + zp_{k-1}(z)$, there is a vector q with $q^* q = 1$ such that $q^* p(B) q = 1$, and this implies that $\|p(B)\| \geq 1$. □

Definition.[14] For a given n by n matrix A and positive integer k, the *polynomial numerical hull of A of degree k* is the largest set $\Omega \subset \mathbf{C}$ for which inequality (1) holds for all polynomials p of degree k or less.

Corollary 5. The set $\mathcal{G}_k(A)$ defined in (4) is the polynomial numerical hull of A of degree k.

Proof: By the previous theorem, there is a polynomial $\hat{p} \in \mathcal{P}_k(0)$ with $\|\hat{p}(A - \zeta I)\| < 1$ if and only if $\zeta \notin \mathcal{G}_k(A)$; that is, there is a polynomial p of the form

$$p(z) = 1 + c_1(z - \zeta) + \ldots + c_k(z - \zeta)^k \tag{14}$$

with $\|p(A)\| < 1 = |p(\zeta)|$ if and only if $\zeta \notin \mathcal{G}_k(A)$. Therefore $\mathcal{G}_k(A)$ contains the polynomial numerical hull of degree k.

It also follows that if $\zeta \in \mathcal{G}_k(A)$, then, since any polynomial q of degree k or less that is nonzero at ζ can be written as a multiple of one of the form (14), $q(z) = q(\zeta)p(z)$, there is no such polynomial with $\|q(A)\| < |q(\zeta)|$, or, equivalently, every such polynomial satisfies $\|q(A)\| \geq |q(\zeta)|$. This inequality obviously holds also if $q(\zeta) = 0$, so $\mathcal{G}_k(A)$ is contained in the polynomial numerical hull of degree k. □

Having established the relationship between $\|p(A)\|$ and the magnitude of p on $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$, we now list some simple properties of these sets. Some of these properties of $\mathcal{G}_k(A)$ also can be found in [13], as can a different approach to computing these sets.

Theorem 6.

(i) $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$ are invariant under unitary similarity transformations of A.

(ii) For $\gamma \in \mathbf{C}$, $\mathcal{F}_k(\gamma I + A) = \gamma + \mathcal{F}_k(A)$ and $\mathcal{G}_k(\gamma I + A) = \gamma + \mathcal{G}_k(A)$.

(iii) $\mathcal{F}_1(A) = \mathcal{G}_1(A) = \mathcal{F}(A)$.

For $1 \leq j \leq k$, $\mathcal{F}(A) \supset \mathcal{F}_j(A) \supset \mathcal{F}_k(A) \supset \sigma(A)$ and $\mathcal{F}(A) \supset \mathcal{G}_j(A) \supset \mathcal{G}_k(A) \supset \sigma(A)$.

For k greater than or equal to the degree of the minimal polynomial of A, $\mathcal{F}_k(A) = \mathcal{G}_k(A) = \sigma(A)$.

(iv) $\mathcal{F}_k(A) \subset \mathcal{G}_k(A) \subset \bigcap_{j=1}^{k}[\mathcal{F}(A^j)]^{1/j}$.

(v) If A is a normal matrix or an upper triangular Toeplitz matrix, then $\mathcal{F}_k(A) = \mathcal{G}_k(A)$ for all k.

(vi) If A is Hermitian, then

$$\mathcal{F}_k(A) = \mathcal{G}_k(A) = \begin{cases} \mathrm{co}(\sigma(A)), & \text{for } k = 1, \\ \sigma(A), & \text{for } k > 1. \end{cases} \tag{15}$$

Proof:

(i) This follows from the invariance of the k-dimensional generalized field of values under unitary similarity transformations.

(ii) We have $\xi \in \mathcal{G}_k(\gamma I + A)$ if and only if there exist unit vectors q_1 and q_2 and a number $t \in [0,1]$ such that

$$t \begin{pmatrix} q_1^*(A + (\gamma - \xi)I)q_1 \\ \vdots \\ q_1^*(A + (\gamma - \xi)I)^k q_1 \end{pmatrix} + (1-t) \begin{pmatrix} q_2^*(A + (\gamma - \xi)I)q_2 \\ \vdots \\ q_2^*(A + (\gamma - \xi)I)^k q_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix},$$

and we have $\xi \in \mathcal{F}_k(\gamma I + A)$ if and only if this holds for $t = 1$. Each of these is easily seen to be equivalent to the condition that $\zeta = \xi - \gamma$ be in $\mathcal{G}_k(A)$ or $\mathcal{F}_k(A)$, respectively.

(iii) The identity $\mathcal{F}_1(A) = \mathcal{F}(A)$ is immediate from definition (2), as is the inclusion $\mathcal{F}(A) \supset \mathcal{F}_j(A) \supset \mathcal{F}_k(A) \supset \sigma(A)$, for $1 \leq j \leq k$. It follows from (3) and (4), since $\mathcal{F}(A - \zeta I)$ is a convex set, that $\mathcal{G}_1(A) = \mathcal{F}_1(A)$. It is also clear from (4) that for $j \leq k$, $\mathcal{G}_j(A) \supset \mathcal{G}_k(A) \supset \mathcal{F}_k(A)$. This establishes the first two parts of (iii). The third part then follows from Theorem 1 and Corollary 5, taking p to be the minimal polynomial of A.

(iv) It follows from definition (2) that if $q^*Aq \in \mathcal{F}_k(A)$, then $(q^*Aq)^j \in \mathcal{F}(A^j)$. Thus $\mathcal{F}_k(A) \subset \bigcap_{j=1}^{k}[\mathcal{F}(A^j)]^{1/j}$. To see that $\mathcal{G}_k(A) \subset \bigcap_{j=1}^{k}[\mathcal{F}(A^j)]^{1/j}$, note that $\zeta \in \mathcal{G}_k(A)$ if and only if there exist unit vectors q_1 and q_2 and a number $t \in [0,1]$ such that $tq_1^*(A - \zeta I)^j q_1 + (1-t)q_2^*(A - \zeta I)^j q_2 = 0$, for all $j = 1, \ldots, k$. For $j = 1$, this says $\zeta = tq_1^*Aq_1 + (1-t)q_2^*Aq_2$, which implies $\zeta \in \text{co}[\mathcal{F}(A)] = \mathcal{F}(A)$. Taking $j = 2$ and substituting this expression for ζ, we find that $\zeta^2 = tq_1^*A^2q_1 + (1-t)q_2^*A^2q_2$, so $\zeta^2 \in \text{co}[\mathcal{F}(A^2)] = \mathcal{F}(A^2)$. Continuing in this way, we see that each $\zeta^j \in \text{co}[\mathcal{F}(A^j)] = \mathcal{F}(A^j)$, $j = 1, \ldots, k$.

(v) This follows from results in Faber et al [4] showing that for normal matrices and for upper triangular Toeplitz matrices the k-dimensional generalized field of values is convex. (See also [6, 11] for analysis leading to the result for normal matrices.)

(vi) This result is given in [15, Proposition 3.10] for bounded linear operators. In the finite dimensional case, for $k = 1$, it is just the statement that the field of values of a Hermitian matrix is the convex hull of its eigenvalues. For $k > 1$ and for q a unit vector, q^*Aq in $\mathcal{F}_k(A)$ with A Hermitian implies $(q^*Aq)^2 = q^*A^2q = \|Aq\|^2$. By the Cauchy-Schwarz inequality, this can happen only if q and Aq are parallel; i.e., $Aq = \lambda q$, and then $q^*Aq = \lambda \in \sigma(A)$. $\square$

When A is normal but nonhermitian, one might expect a result like (15) to hold, but in this case, depending on the eigenvalue distribution, $\mathcal{F}_k(A)$, $k > 1$, may contain more than just the spectrum of A. Although $\|p(A)\| = \max_{\lambda \in \sigma(A)} |p(\lambda)|$, there may be other points $z \in \mathbf{C}$ for which $|p(z)| \leq \max_{\lambda \in \sigma(A)} |p(\lambda)|$, for all polynomials of degree k or less. For example, if the eigenvalues of A are densely distributed around the unit circle, then for $k << n$, $\mathcal{F}_k(A)$ will consist of almost the entire unit disk, since, by the maximum principal, a polynomial must obtain its maximum absolute value on the boundary of the disk.

While the sets $\mathcal{F}_k(A)$ and $\mathcal{G}_k(A)$ are neither convex nor necessarily connected, they are compact sets. They are bounded since they are subsets of $\mathcal{F}(A)$, and they are closed since the k-dimensional generalized field of values, being the continuous image of the unit sphere, is closed. To see that $\mathcal{F}_k(A)$ is closed, note that if ζ_ℓ, $\ell = 1, 2, \ldots$, is a sequence of points in $\mathcal{F}_k(A)$ converging to some value ζ, then $(\zeta_\ell, \ldots, \zeta_\ell^k)^T$ is a sequence of vectors in the k-dimensional generalized field of values $F(\{A^j\}_{j=1}^k)$, converging to the vector $(\zeta, \ldots, \zeta^k)^T$, which also must be in this generalized field of values; i.e., $\zeta \in \mathcal{F}_k(A)$.

3 Computation of the Polynomial Numerical Hull of a Given Degree and Numerical Examples

Using the fact that the polynomial numerical hull of degree k is a subset of the field of values of A, or, more specifically, that it is a subset of $\bigcap_{j=1}^k \mathcal{F}(A^j)^{1/j}$, one can proceed to test points ζ in this region to determine if $0 \in \mathrm{co}[F(\{(A-\zeta I)^j\}_{j=1}^k)]$. According to Theorem 4, this is equivalent to determining if $\min_{p \in \mathcal{P}_k(0)} \|p(A-\zeta I)\|$ is equal to one. The problem of finding the polynomial $p \in \mathcal{P}_k(0)$ that minimizes $\|p(A-\zeta I)\|$, for a given value of ζ, can be cast as a semidefinite programming problem, and an algorithm and software for its solution have been developed by Toh, et al [19, 18].

Here we use this software to approximate polynomial numerical hulls of various degrees for a number of example problems. It is hoped that methods can be developed for computing these sets without simply testing all possible points. Of course, once a closed boundary curve of the polynomial numerical hull of degree k has been identified, one can argue by the maximum principle that the region inside this curve also must be contained in the set. Hence some of our computations involve testing of only enough points to determine a good approximation to the boundary curve.

Example 1. Jordan Block. As a first example, we consider a Jordan block of size $n = 24$ with eigenvalue 0 and compute the polynomial numerical hull of degree $k = 23$. Figure 1 shows a plot of the field of values (the region inside

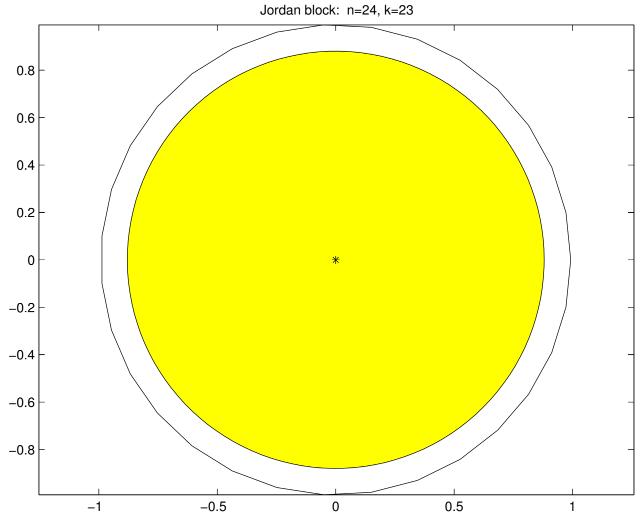

Figure 1: Field of values and polynomial numerical hull of degree k for a Jordan block. $n = 24,\ k = 23$.

the outer curve) and an approximation to the polynomial numerical hull of degree $n - 1 = 23$ (the shaded region). The single eigenvalue zero is marked by an asterisk. This indicates that for polynomials of degree less than n, a Jordan block behaves in many ways like a normal matrix with eigenvalues spread throughout most of its field of values; that is, $\|p(A)\| \geq \|p(B)\|$, when B is a normal matrix with eigenvalues throughout the shaded region.

Using inequality (10) with Γ taken to be the boundary of this set and noting that the resolvent norm on Γ is roughly 100 (i.e., Γ closely resembles the 10^{-2}-pseudospectrum of A), we obtain fairly tight upper and lower bounds on the norms of polynomials applied to this Jordan block:

$$\max_{z \in \mathcal{G}_{n-1}(A)} |p(z)| \leq \|p(A)\| \leq 90 \max_{z \in \mathcal{G}_{n-1}(A)} |p(z)|,$$

for all polynomials p of degree $n - 1 = 23$ or less.

Example 2. Gauss-Seidel Matrix. It is well-known that the convergence rate of the Gauss-Seidel method may depend on whether an upwind or downwind direction is chosen for the sweeps; that is, it is sometimes better to use the upper triangular part and sometimes better to use the lower triangular part of the matrix as a preconditioner, even when the spectral radii of the two iteration matrices are the same. Trefethen has used this phenomenon

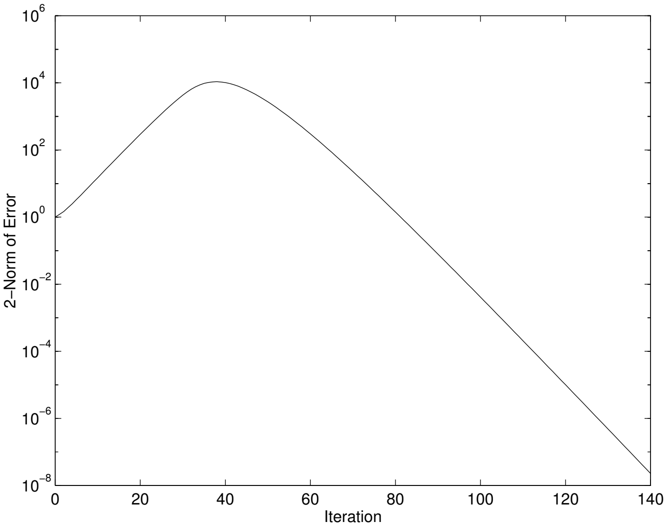

Figure 2: Convergence of the Gauss-Seidel iteration.

to illustrate the relation between pseudospectra and the performance of the Gauss-Seidel iteration [22].

Consider the matrix

$$C = \begin{pmatrix} 1 & -1.16 & & \\ .16 & \ddots & \ddots & \\ & \ddots & \ddots & -1.16 \\ & & .16 & 1 \end{pmatrix}.$$

The spectral radius of $I - M^{-1}C$ is about .73 when M is taken to be either the lower or the upper triangle of C. Hence, asymptotically the iteration $x_{k+1} = x_k + M^{-1}(b - Cx_k)$ converges to the solution of the linear system $Cx = b$, reducing the error by approximately the factor .73 at each step in the later stages. If M is taken to be the lower triangle of C, however, as in the Gauss-Seidel method, then the iteration matrix $A = I - M^{-1}C$ is highly nonnormal and the error grows by several orders of magnitude before reaching this asymptotically convergent regime. This is shown in Figure 2.

The explanation is that the polynomial numerical hull of any degree $k < n$ for the iteration matrix is much larger than suggested by the eigenvalues. Figure 3 shows the field of values (inside the outer curve) and the polynomial numerical hull of degree $k = 29$ (shaded) for the Gauss-Seidel iteration matrix

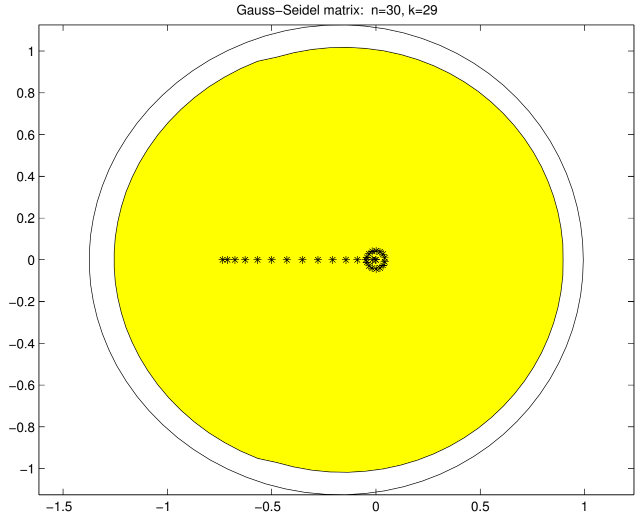

Figure 3: Field of values and polynomial numerical hull of degree k for the Gauss-Seidel iteration matrix. $n = 30,\ k = 29$.

of size $n = 30$. The eigenvalues of the matrix are marked with asterisks. Again it is seen that the polynomial numerical hull of degree $n-1$ fills most of the field of values. The largest absolute value of a point in the polynomial numerical hull of degree 29 is about 1.256, indicating that over the first 29 steps the error in the Gauss-Seidel iteration (more precisely, $\|A^k\|$) must grow to at least $1.256^{29} \approx 700$. In fact, it grows to about 14 times this value before beginning to decrease, as shown in Figure 2.

Again using inequality (10) with Γ taken to be the boundary of the polynomial numerical hull of degree $k = n-1$ (which closely resembles the 10^{-2}-pseudospectrum of A), we obtain fairly tight upper and lower bounds:

$$\max_{z \in \mathcal{G}_{n-1}(A)} |p(z)| \leq \|p(A)\| \leq 100 \max_{z \in \mathcal{G}_{n-1}(A)} |p(z)|,$$

for all polynomials p of degree $n-1 = 29$ or less.

Since this example involves simple powers of the matrix A, one also can obtain lower and upper bounds on $\sup_{j \geq 0} \|A^j\|$ using the Kreiss matrix theorem [12, 16, 23]. According to that theorem (in its current sharp form) [16],

$$r(A) \leq \sup_{j \geq 0} \|A^j\| \leq e\ n\ r(A), \quad \text{where} \quad r(A) = \sup_{|z|>1} (|z|-1) \|(zI - A)^{-1}\|.$$

A numerical computation shows that $r(A) \approx 4324$, and the factor en is approximately 82. The polynomial numerical hull of degree k provides lower and upper bounds on $\|A^j\|$ for $j \leq k$, while the Kreiss matrix theorem gives bounds on $\sup_{j \geq 0} \|A^j\|$.

Example 3. The Ehrenfests' Urn. This example is similar to the previous one in that it involves the norms of powers of a matrix. The setup represents a simple model of diffusion and displays a cutoff phenomenon described by Diaconis [1]. Consider two urns and d balls. Initially all of the balls are in urn 2. At each step, the probability of moving a ball from one urn to the other is proportional to the number of balls in the urn. Letting the state space be the number of balls $\{0, 1, \ldots, d\}$ in urn 1, the transition probabilities are: $P(i, i-1) = i/(d+1)$, $P(i,i) = 1/(d+1)$, and $P(i, i+1) = (d-i)/(d+1)$. Applying the probability transition matrix P^T to the initial vector $(1, 0, \ldots, 0)^T$ many times, the system approaches its stationary state, which is a binomial distribution: $\pi(j) = \begin{pmatrix} d \\ j \end{pmatrix} / 2^d$, $0 \leq j \leq d$.

Asymptotically, the difference between the current state of the system and the stationary state converges to zero at a rate determined by the second largest eigenvalue of the matrix, but a number of steps are required before any convergence towards the stationary state is seen. The number of steps depends on the norm in which this difference is measured. The most appropriate norm is the total variation distance described in [1], and this is closely related to the 1-norm [20]. The cutoff is less pronounced, but can still be seen, when the difference is measured in the 2-norm.

Figure 4 shows a plot of the 1-norm (solid) and the 2-norm (dashed) of the powers of the matrix $A = P^T - v_1 w_1^T$, where v_1 is a right eigenvector and w_1 a left eigenvector associated with the largest eigenvalue 1; that is, if $P^T = V \Lambda V^{-1}$, where Λ is a diagonal matrix of eigenvalues with $\Lambda_{11} = 1$, then v_1 is the first column of V and w_1^T is the first row of V^{-1}. Powers of this matrix A, applied to the initial state vector, give the differences between the current and stationary states. The plot is for a problem with $d = 50$ balls, so the matrix is of order $n = 51$. The second largest eigenvalue of the probability transition matrix is about .9608, and the 2-norm condition number of the matrix V of eigenvectors is about 10^7. By step $k = 100$, the ratio of norms of successive powers $\|A^k\|/\|A^{k-1}\|$ is very close to its asymptotic value of .9608. Early in the process, however, we see that $\|A^k\|/\|A^{k-1}\|$ is very close to 1, especially for the 1-norm.

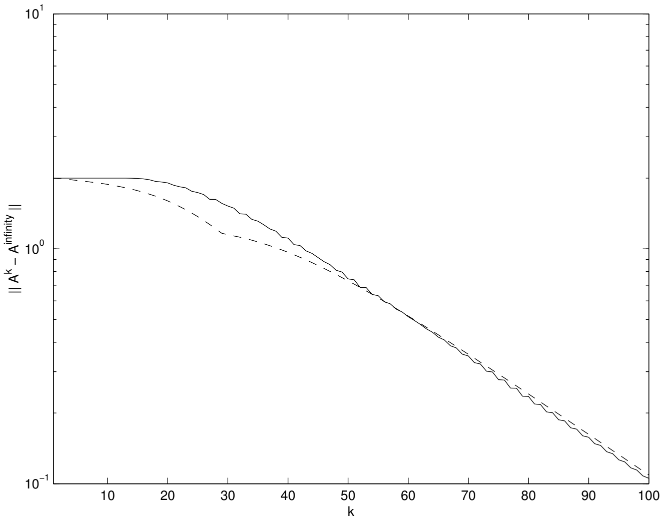

Figure 4: Norms of powers of the Ehrenfests' urn matrix. 1-norm (solid), 2-norm (dashed).

The behavior of the matrix powers in the 2-norm is *partially* explained by the polynomial numerical hull of degree $k = 11$ pictured in Figure 5. Again the outer curve denotes the boundary of the field of values, the polynomial numerical hull of degree $k = 11$ is shaded, and the eigenvalues, which lie on the real axis, are marked with asterisks. The field of values and the polynomial numerical hulls of degree $k \leq 10$ for A all contain points with absolute value greater than 1, indicating that $\|A^k\|$ cannot be less than 1 for $k \leq 10$. The point $\zeta = 1$ is right on the border of the polynomial numerical hull of degree $k = 11$ and is outside the set for $k = 12$. In actuality, it can be seen from Figure 4 that $\|A^k\|$ does not drop below 1 until step $k = 39$, so the polynomial numerical hulls of various degrees only partially explain the behavior of the matrix powers in this case. Moreover, the polynomial numerical hulls plotted here are associated with the 2-norm. Different sets such as $\{z \in \mathbf{C} : \ \|p(A)\|_1 \geq |p(z)| \text{ for all } p \text{ of degree } k \text{ or less}\}$ must be used to study the 1-norm behavior of polynomial functions of a matrix.

Example 4. Grcar matrix. The following matrix was introduced by Grcar [5] and has been studied in connection with pseudospectra by Trefethen [22]

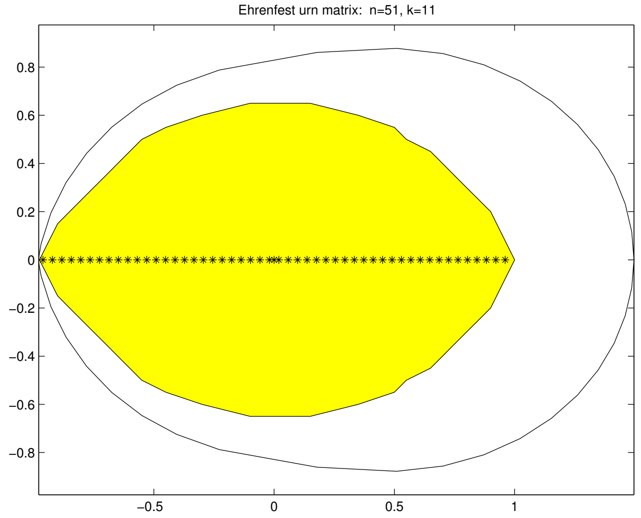

Figure 5: Field of values and polynomial numerical hull of degree k for the Ehrenfests' urn matrix. $n = 51,\ k = 11$.

and by Toh and Trefethen [19]:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & & \\ -1 & 1 & 1 & 1 & \ddots & \\ & \ddots & \ddots & \ddots & \ddots & 1 \\ & & -1 & 1 & 1 & 1 \\ & & & -1 & 1 & 1 \\ & & & & -1 & 1 \end{pmatrix},$$

It is a matrix whose polynomial numerical hulls of various degrees are significantly different from either its eigenvalues or its field of values. Using a matrix of size $n = 48$, the field of values and eigenvalues are plotted in Figure 6, while the polynomial numerical hulls of degree $k = 4$, 8, 16, and 32 are plotted in Figure 7. As can be seen from the figures, while the field of values contains the origin, the polynomial numerical hull of degree 4 does not, showing that the GMRES algorithm for solving a linear system with coefficient matrix A will make some (small amount of) progress within the first 4 steps: $\min_{p \in \mathcal{P}_4(0)} \|p(A)\| < 1$.

While the computation of polynomial numerical hulls of various degrees,

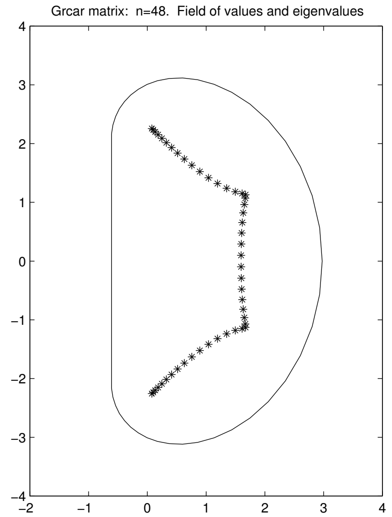

Figure 6: Field of values and eigenvalues of the Grcar matrix. $n = 48$.

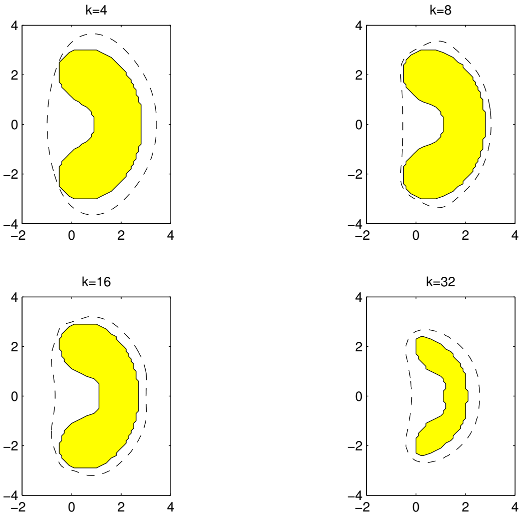

Figure 7: Polynomial numerical hulls of degree $k = 4, 8, 16, 32$ for the Grcar matrix. $n = 48$. Dashed curves are the lemniscates $\|\rho_k(A)\| = |\rho_k(z)|$, where ρ_k is the kth degree Arnoldi polynomial for a random starting vector.

as in Figure 7, is quite costly, these sets might be approximated by

$$\{z \in \mathbf{C} : \ \|\rho_k(A)\| \geq |\rho_k(z)|\}$$

for some specific polynomial ρ_k such as that produced in the GMRES or Arnoldi algorithm. This set necessarily contains the polynomial numerical hull of degree k. The dashed curves in Figure 7 are the lemniscates $\{z \in \mathbf{C} : \ \|\rho_k(A)\| = |\rho_k(z)|\}$, where ρ_k is the polynomial produced at step k of the Arnoldi algorithm, using a random initial vector. Trefethen has suggested looking at the regions enclosed by such lemniscates as approximations to certain pseudospectra and we recommend them also as approximations to the polynomial numerical hull of the given degree. In this case, the approximation is rough, however, and all of the lemniscates up to $k = 32$ enclose the origin.

4 Summary and Applications

The goal of this paper is to convince readers that to study the behavior of polynomial functions of a matrix, when the polynomials are of moderate degree compared to the minimal polynomial, one should consider polynomial numerical hulls of various degrees rather than eigenvalues. The applications include every field in which eigenvalue analysis is used to predict anything other than asymptotic behavior — stability of difference schemes, convergence of iterative methods, cutoff phenomena in random processes, etc., etc. In order to understand the growth or stationarity of the norms of powers of a matrix as illustrated in Figures 2 and 4, one must have a *lower* bound for $\|p(A)\|$. Using the Cauchy integral formula, one can easily derive upper bounds, and if the upper and lower bounds turn out to be close, as they were for a number of the examples in Section 3, then this tells us that the set $\mathcal{G}_k(A)$ really does determine the behavior of A to a close approximation under the action of polynomials of degree k or less. One might be able to derive similar sets in the complex plane that determine the behavior of A under the action of other classes of functions. Another interesting class might be functions of the form $f(A) = e^{tA}$, $0 \leq t \leq T$.

The difficulty in computing polynomial numerical hulls of various degrees remains an obstacle to their use. Since even a rough idea of what these sets look like can be useful, however, their approximation via the Arnoldi algorithm or some other means may prove sufficient in practice to deduce important information about the behavior of the matrix.

Acknowledgments: The author thanks Nick Trefethen and Mark Embree for helpful comments on a draft of this paper. She thanks Marko Huhtanen and an anonymous referee for the references [14, 15] to Nevanlinna's work on the sets studied here.

References

[1] P. Diaconis, *The cutoff phenomenon in finite Markov chains*, Proc. Natl. Acad. Sci., 93 (1996), pp. 1659–1664.

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

[3] M. Eiermann, *Fields of values and iterative methods*, talk presented at Oberwolfach meeting on Iterative Methods and Scientific Computing, Oberwolfach, Germany, April 1997.

[4] 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.

[5] J. Grcar, *Operator coefficient methods for linear equations*, Sandia National Lab. Rep. SAND89-8691, Nov., 1989.

[6] A. Greenbaum and L. Gurvits, *Max-min properties of matrix factor norms*, SIAM J. Sci. Comput., 15 (1994), pp. 427–439.

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

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

[9] M. Huhtanen and O. Nevanlinna, *Minimal decompositions and iterative methods*, Numer. Math., 86 (2000), pp. 257–281.

[10] M. Huhtanen, *Ideal GMRES can be bounded from below by three factors*, Helsinki University of Technology, Math Res. Rep. A412, 1999.

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

[12] H. O. Kreiss, *Über die stabilitätsdefinition für differenzengleichungen die partielle differenzialgleichungen approximieren* BIT, 2 (1962), pp. 153–181.

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

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

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

[16] M. N. Spijker, *On a conjecture by LeVeque and Trefethen related to the Kreiss matrix theorem*, BIT, 31 (1991), pp. 551–555.

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

[18] K. C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3–a Matlab software package for semidefinite programming*, Optim. Meth. Softw. 11-12 (1999), pp. 545–581.

[19] K. C. Toh and L. N. Trefethen, *The Chebyshev polynomials of a matrix*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 400–419.

[20] L. N. Trefethen and L. M. Trefethen, *How many shuffles to randomize a deck of cards?*, Proc. Roy. Soc. Lond. A 456 (2000), pp. 2561–2568.

[21] L. N. Trefethen, *Approximation theory and numerical linear algebra*, in Algorithms for Approximation II, J. Mason and M. Cox, eds., Chapman and Hall, London, 1990.

[22] L. N. Trefethen, *Pseudospectra of matrices*, in D. F. Griffiths and G. A. Watson, eds., Numerical Analysis 1991 (Dundee, 1991), Harlow, Essex, UK: Longman Sci. Tech, 1992, pp. 234–266.

[23] E. Wegert and L. N. Trefethen, *From the Buffon needle problem to the Kreiss matrix theorem*, Amer. Math. Monthly, 101 (1994), pp. 132–139.