Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 419 (2006) 37–47
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Characterizations of the polynomial numerical hull of degree k

James V. Burke [1], Anne Greenbaum [*,2]

Department of Mathematics, University of Washington, Box 354350, Seattle, WA 98195, United States

Received 28 December 2005; accepted 6 April 2006
Available online 14 July 2006
Submitted by R.A. Brualdi

Abstract

Six characterizations of the polynomial numerical hull of degree k are established for bounded linear operators on a Hilbert space. It is shown how these characterizations provide a natural distinction between interior and boundary points. One of the characterizations is used to prove that the polynomial numerical hull of any fixed degree k for a Toeplitz matrix whose symbol is piecewise continuous approaches all or most of that of the infinite-dimensional Toeplitz operator, as the matrix size goes to infinity.

AMS classification: 15A60; 65F15; 65F35

Keywords: Polynomial numerical hull; Nonnormal matrix; Toeplitz matrix

1. Introduction

The polynomial numerical hull of degree k for a bounded linear operator A on a Hilbert space is defined as

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

where $\mathscr{P}_k$ denotes the set of polynomials of degree k or less [10,11,5]. The sets $\mathscr{H}_k(A)$ are nonempty and compact, and $\mathscr{H}_1(A)$ is equal to the closure of the field of values of A [10].

* Corresponding author.
E-mail addresses: burke@math.washington.edu (J.V. Burke), greenbau@math.washington.edu (A. Greenbaum).
[1] Research supported in part by the National Science Foundation Grant DMS-0505712.
[2] Research supported in part by the National Science Foundation Grant DMS-0208353.

0024-3795/$ - see front matter
doi:10.1016/j.laa.2006.04.023

Polynomial numerical hulls were introduced as a tool for understanding and estimating $\|f(A)\|$ for various classes of functions f. When the operator A is normal and $\|\cdot\|$ is the 2-norm, the spectrum of A suffices for understanding $\|f(A)\|$, but when A is nonnormal the spectrum often does not supply the kind of information required in applications [14]. Polynomial numerical hulls have been shown to be a useful tool in a number of applications where the underlying operator A is nonnormal [5–7]. Nonetheless, very little is known about these sets.

The purpose of this paper is to present several characterizations of polynomial numerical hulls and their boundary. It is hoped that our characterizations of the boundary will eventually permit us to apply methods similar to those used for pseudospectra to compute the polynomial numerical hull. Notwithstanding this numerical interest, an understanding of the geometry of the boundary of the polynomial numerical hull is fundamental to an understanding of the geometry of the entire set. We conclude this note by showing that the polynomial numerical hull of any fixed degree k for a Toeplitz matrix whose symbol is piecewise continuous approaches all or most of that of the infinite-dimensional Toeplitz operator, as the matrix size goes to infinity. More precisely, we show that the polynomial numerical hull of degree k for the Toeplitz operator contains the *uniform* and *partial* limits [12] (or, equivalently, the inner and outer limits [13], respectively) as $N \to \infty$ of the polynomial numerical hulls of degree k for the $N \times N$ Toeplitz matrix, which in turn contain the closure of the interior of that of the Toeplitz operator.

The notation we use is standard. However, for the convenience of the reader, we list some of this notation here. Let $\mathbb{H}$ be a complex Hilbert space with inner product $\langle\cdot,\cdot\rangle$. We also use this notation to denote the usual inner product on $\mathbb{C}^k$: $\langle z, w\rangle := \sum_{j=1}^{k} z_j \bar{w}_j \ \forall z, w \in \mathbb{C}^k$, where $\bar{w}_j$ is the complex conjugate of w_j. The space of linear transformations from $\mathbb{H}$ to itself is denoted $\mathscr{L}(\mathbb{H})$, and the subset of these that are bounded is denoted $\mathscr{B}(\mathbb{H})$. We use the notation $\|\cdot\|$ for the Hilbert space norm for vectors and the corresponding operator norm for linear operators and matrices. The interior, boundary, and convex hull of a subset S of $\mathbb{C}^k$ are denoted int S, bdry S, and co S, respectively. If $K \subset \mathbb{C}^k$ is convex and $x \in K$, we denote the normal cone to K at x by

$$N_K(x) := \{z \in \mathbb{C}^k : \operatorname{Re}\langle z, y - x\rangle \leqslant 0 \ \forall y \in K\}.$$

It is straightforward to show that $x \in \operatorname{int} K$ if and only if $N_K(x) = \{0\}$.

2. The field of values

The polynomial numerical hull is intimately related to the notion of the field of values, or numerical range, of an operator. Indeed, the polynomial numerical hull of degree 1 of an operator coincides with the closure of the field of values. But there are other connections as well. For this reason, we begin our discussion with the field of values.

Let B be a linear operator on the Hilbert space $\mathbb{H}$. The field of values, or numerical range, of B is defined as

$$\mathscr{F}(B) := \{\langle Bq, q\rangle : \|q\| = 1\}. \tag{2}$$

The field of values is always a convex set, but it is not necessarily closed in the infinite dimensional case. For this reason, define $\overline{\mathscr{F}}(B) := \operatorname{cl} \mathscr{F}(B)$ to be the *closure* of the field of values of B. The results of the following lemma are known, at least in the case of finite dimensional spaces (see, for example [8,5]), but we include the proofs here for completeness.

Lemma 2.1. *Let $B \in \mathscr{B}(\mathbb{H})$. The following statements are equivalent*:

(i) $0 \notin \overline{\mathscr{F}}(B)$.

(ii) *There exists* $c \in \mathbb{C}$ *such that* $\overline{\mathscr{F}}(cB)$ *lies in the open right half plane.*
(iii) $\min\{\|I - cB\| : c \in \mathbb{C}\} < 1$.

Moreover, $0 \in \operatorname{bdry}(\overline{\mathscr{F}}(B))$ *if and only if there is a* $c \in \mathbb{C}$ *such that* $\overline{\mathscr{F}}(cB)$ *lies in the closed right half plane with zero on its boundary.*

Proof. [(i) $\Longleftrightarrow$ (ii)] Use the convexity of $\overline{\mathscr{F}}(B)$ and the fact that $\overline{\mathscr{F}}(cB) = c\overline{\mathscr{F}}(B)$ to rotate $\overline{\mathscr{F}}(B)$ to the right half plane.

[(ii) $\Longleftrightarrow$ (iii)] Observe the following equivalences:

$$\min_{c\in\mathbb{C}} \|I - cB\| < 1 \tag{3}$$
$$\Longleftrightarrow$$
$$\exists \epsilon > 0,\ c \in \mathbb{C}\setminus\{0\} \text{ such that } \|(I - cB)x\|^2 < 1 - \epsilon\ \forall \|x\| = 1$$
$$\Longleftrightarrow$$
$$\exists \epsilon > 0,\ c \in \mathbb{C}\setminus\{0\} \text{ such that } -2\mathrm{Re}\langle cBx, x\rangle + \|cBx\|^2 < -\epsilon\ \forall \|x\| = 1. \tag{4}$$

The final equivalence (4) implies that there exists $\epsilon > 0$ and $c \in \mathbb{C}\setminus\{0\}$ such that

$$\epsilon < \mathrm{Re}\langle cBx, x\rangle\ \forall \|x\| = 1 \tag{5}$$

or equivalently, $\overline{\mathscr{F}}(cB)$ lies in the open right half plane.

Next suppose that $\overline{\mathscr{F}}(cB)$ lies in the open right half plane, or equivalently, (5) holds with $\epsilon > 0$. Let $\lambda > 0$ be such that $0 < \lambda < \epsilon/\|cB\|^2$, or equivalently,

$$-2\lambda\epsilon + \|\lambda cB\|^2 < -\lambda\epsilon \quad \text{with } \lambda > 0. \tag{6}$$

Set $\hat{c} = \lambda c \in \mathbb{C}\setminus\{0\}$ and $\hat{\epsilon} = \lambda\epsilon > 0$ so that (6) gives $-2\hat{\epsilon} + \|\hat{c}B\|^2 < -\hat{\epsilon}$. Then $-2\mathrm{Re}\langle \hat{c}Bx, x\rangle + \|\hat{c}Bx\|^2 \leqslant -2\hat{\epsilon} + \|\hat{c}B\|^2 < -\hat{\epsilon}$ for all $\|x\| = 1$ which, by (4), is equivalent to (3).

Let us now suppose that $0 \in \operatorname{bdry}(\overline{\mathscr{F}}(B))$. The result follows again by rotation and using the convexity of $\overline{\mathscr{F}}(B)$. $\square$

We also make use of a generalization of the field of values known as the k-dimensional field of values of k transformations $\{B_j\}_{j=1}^k \subset \mathscr{B}(\mathbb{H})$ [8]. This set is given by

$$\{(\langle B_1 q, q\rangle, \ldots, \langle B_k q, q\rangle)^{\mathrm{T}} : \|q\| = 1\}. \tag{7}$$

Our results refer to the convex hull of this set which we denote by $\mathscr{F}(\{B_j\}_{j=1}^k)$. The set $\mathscr{F}(\{B_j\}_{j=1}^k)$ is convex by definition, but may not be closed in infinite dimensions, and so we define $\overline{\mathscr{F}}(\{B_j\}_{j=1}^k) := \operatorname{cl}\mathscr{F}(\{B_j\}_{j=1}^k)$. We now connect this notion to that of the classical field of values thereby allowing us to apply the results of the previous lemma to the k-dimensional field of values.

Lemma 2.2. *Let* $B_j \in \mathscr{B}(\mathbb{H})$ *for* $j = 1, \ldots, k$. *Then we have:*

1. $0 \in \overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ *if and only if* $0 \in \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$ *for all* $c \in \mathbb{C}^k$.
2. $0 \in \operatorname{bdry}(\overline{\mathscr{F}}(\{B_j\}_{j=1}^k))$ *if and only if* $0 \in \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$ *for all* $c \in \mathbb{C}^k$ *and there exists* $\hat{c} \in \mathbb{C}^k \setminus \{0\}$ *such that*

$$0 \in \operatorname{bdry}\left(\overline{\mathscr{F}}\left(\sum_{j=1}^k \hat{c}_j B_j\right)\right).$$

Proof. 1. Let us first suppose that $0 \in \overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$. By Caratheodory's theorem [13, p. 55], there exist sequences $\{\lambda_\nu\} \subset \mathbb{R}_+^{2k+1}$, $\{q_{\nu j}\} \subset \mathbb{H}$, $j = 1, \ldots, 2k+1$, satisfying $\|q_{\nu j}\| = 1$ for all $j = 1, \ldots, 2k+1$ and $\nu = 1, 2, \ldots, \sum_{j=1}^{2k+1} \lambda_{\nu j} = 1$ for $\nu = 1, 2, \ldots$, and

$$\sum_{j=1}^{2k+1} \lambda_{\nu j}(\langle B_1 q_{\nu j}, q_{\nu j}\rangle, \ldots, \langle B_k q_{\nu j}, q_{\nu j}\rangle)^{\mathrm{T}} \longrightarrow 0.$$

Therefore, given any $c \in \mathbb{C}^k$, we have

$$\left\langle c, \sum_{j=1}^{2k+1} \lambda_{\nu j} \begin{pmatrix} \langle B_1 q_{\nu j}, q_{\nu j}\rangle \\ \vdots \\ \langle B_k q_{\nu j}, q_{\nu j}\rangle \end{pmatrix} \right\rangle = \sum_{j=1}^{2k+1} \lambda_{\nu j} \left\langle \left[\sum_{\ell=1}^{k} c_\ell B_\ell\right] q_{\nu j}, q_{\nu j} \right\rangle \longrightarrow 0$$

and so $0 \in \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$ since this set is convex and closed.

Next we suppose that $0 \notin \overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ and show that there exists $c \in \mathbb{C}^k$ such that $0 \notin \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$. Since $0 \notin \overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ where $\overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ is a nonempty closed convex set, the Hahn–Banach theorem yields the existence of a $c \in \mathbb{C}^k \setminus \{0\}$ such that

$$0 > \mathrm{Re}\langle c, (\langle B_1 q, q\rangle, \ldots, \langle B_k q, q\rangle)^{\mathrm{T}}\rangle = \mathrm{Re}\left\langle \left[\sum_{j=1}^{k} c_j B_j\right] q, q \right\rangle$$

for all $\|q\| = 1$, which implies that $0 \notin \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$.

2. If $0 \in \mathrm{bdry}\,\overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$, then $0 \in \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$ for all $c \in \mathbb{C}^k$, by part 1, and the normal cone to $\overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ at 0 must contain a nonzero element $\hat{c} \in \mathbb{C}^k$. But then 1 is an element of the normal cone to $\overline{\mathscr{F}}(\sum_{j=1}^k \hat{c}_j B_j)$ at 0 since for every $\|q\| = 1$,

$$\mathrm{Re}\left\langle \left[\sum_{j=1}^{k} \hat{c}_j B_j\right] q, q \right\rangle = \mathrm{Re}\langle \hat{c}, (\langle B_1 q, q\rangle, \ldots, \langle B_k q, q\rangle)^{\mathrm{T}}\rangle \leqslant 0.$$

Hence $0 \in \mathrm{bdry}\overline{\mathscr{F}}(\sum_{j=1}^k \hat{c}_j B_j)$ since its normal cone at 0 contains a nonzero element.

Next assume that $0 \in \overline{\mathscr{F}}(\sum_{j=1}^k c_j B_j)$ for all $c \in \mathbb{C}^k$ and there exists $\hat{c} \in \mathbb{C}^k \setminus \{0\}$ such that $0 \in \mathrm{bdry}\overline{\mathscr{F}}(\sum_{j=1}^k \hat{c}_j B_j)$. Again, by part 1, $0 \in \overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$. Since $0 \in \mathrm{bdry}\,\overline{\mathscr{F}}(\sum_{j=1}^k \hat{c}_j B_j)$, the normal cone to $\overline{\mathscr{F}}(\sum_{j=1}^k \hat{c}_j B_j)$ at 0 must contain a nonzero element $\zeta \in \mathbb{C}$. But then $\zeta\hat{c}$ is a nonzero element of the normal cone to $\overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ at 0 since for every $\|q\| = 1$.

$$\mathrm{Re}\left\langle \zeta\hat{c}, \begin{pmatrix} \langle B_1 q, q\rangle \\ \vdots \\ \langle B_k q, q\rangle \end{pmatrix} \right\rangle = \mathrm{Re}\,\zeta \left\langle \left[\sum_{j=1}^{k} \hat{c}_j B_j\right] q, q \right\rangle \leqslant 0.$$

Since the normal cone to $\overline{\mathscr{F}}(\{B_j\}_{j=1}^k)$ at 0 contains a nonzero element, 0 must be on the boundary of this set. □

The relationship between the polynomial numerical hull and the k-dimensional field of values is obtained by setting the operators B_j in definition (7) equal to powers of a single operator $A - \zeta I$ with $A \in \mathscr{B}(\mathbb{H})$ yielding the following observation about the field of values of polynomials in A.

Lemma 2.3. *Given $A \in \mathscr{B}(\mathbb{H})$ and $\zeta \in \mathbb{C}$, we have*

$$0 \in \overline{\mathscr{F}}\left(\sum_{j=1}^{k} c_j (A - \zeta I)^j\right) \ \forall c \in \mathbb{C}^k \iff p(\zeta) \in \overline{\mathscr{F}}(p(A))\ \forall p \in \mathscr{P}_k.$$

Proof. Every polynomial $p \in \mathscr{P}_k$ can be written in the form $p(A) = p(\zeta)I + \sum_{j=1}^{k} c_j (A - \zeta I)^j$, for some $c \in \mathbb{C}^k$, and conversely every polynomial of this form is in $\mathscr{P}_k$. Hence the lemma is an immediate consequence of the identity

$$\{\omega\} + \overline{\mathscr{F}}(B) = \overline{\mathscr{F}}(\omega I + B)$$

for every $\omega \in \mathbb{C}$ and $B \in \mathscr{B}(\mathbb{H})$. $\square$

3. Characterizations of $\mathscr{H}_k(A)$

Let $p \in \mathscr{P}_k$ and $z \in \mathbb{C}$ be such that $p(z) \neq 0$. Expanding p at z we have

$$p(\zeta) = c_0 + c_1(\zeta - z) + c_2(\zeta - z)^2 + \cdots + c_k(\zeta - z)^k,$$

where $p(z) = c_0 \neq 0$. Hence, given $A \in \mathscr{B}(\mathbb{H})$, the inequality $\|p(A)\| \geqslant |p(z)|$ is equivalent to

$$\|c_0 I + c_1(A - zI) + c_2(A - zI)^2 + \cdots + c_k(A - zI)^k\| \geqslant |c_0|.$$

Divide this inequality through by $|c_0| \neq 0$ to obtain the equivalent inequality

$$\left\| I - \sum_{j=1}^{k} \hat{c}_j (A - zI)^j \right\| \geqslant 1,$$

where $\hat{c}_j = -c_j/|c_0|$, $j = 1, \ldots, k$. Since the inequality $\|p(A)\| \geqslant |p(z)|$ holds trivially when $p(z) = 0$, we obtain the equivalence

$$\|p(A)\| \geqslant |p(z)|\ \forall p \in \mathscr{P}_k \iff \min_{c \in \mathbb{C}^k} \left\| I - \sum_{j=1}^{k} c_j (A - zI)^j \right\| = 1, \tag{8}$$

where equality is attained in the minimization at $c = 0$. Combining this equivalence with the results of the previous section gives a variety of characterizations of the polynomial numerical hull.

Theorem 3.1. *Let $A \in \mathscr{B}(\mathbb{H})$ and $k \in \mathbb{N}$. Then the following statements are equivalent*:

1. $z \in \mathscr{H}_k(A)$.
2. $\min_{c \in \mathbb{C}^k} \|I - \sum_{j=1}^{k} c_j (A - zI)^j\| = 1$.
3. $0 \in \overline{\mathscr{F}}(\sum_{j=1}^{k} c_j (A - zI)^j)$ *for all* $c \in \mathbb{C}^k$.
4. $\inf_{\|q\|=1} \mathrm{Re}\langle [\sum_{j=1}^{k} c_j (A - zI)^j] q, q\rangle \leqslant 0$ *for all* $c \in \mathbb{C}^k$.
5. $0 \in \overline{\mathscr{F}}(\{(A - zI)^j\}_{j=1}^{k})$.
6. $p(z) \in \overline{\mathscr{F}}(p(A))$ *for all* $p \in \mathscr{P}_k$.

Remarks

(1) The equivalence of 1 and 2 was first established in [5, Theorem 4]. The equivalence of 1 and 5 is given in [5, Corollary 5].
(2) In [2], Davies introduces the set

$\mathrm{Num}_k(A) := \{z \in \mathbb{C} : p(z) \in \overline{\mathscr{F}}(p(A))\ \forall p \in \mathscr{P}_k\}.$

The equivalence of 1 and 6 above implies that $\mathrm{Num}_k(A) = \mathscr{H}_k(A)$. Note that the definition of the set $\mathrm{Num}_k(A)$ does not require the use of the norm. Therefore, this definition may prove useful in extending the notion of the polynomial numerical hull to unbounded operators as well as to linear operators on more general spaces.

Proof. The equivalence of 1 and 2 is simply a restatement of (8). The equivalence of 2 and 3 is a consequence of the equivalence of (i) and (iii) in Lemma 2.1. The equivalence of 3 and 4 follows from the equivalence of (i) and (ii) in Lemma 2.1. The equivalence of 3 and 5 follows from part 1 of Lemma 2.2. Finally, the equivalence of 3 and 6 is the content of Lemma 2.3. □

In the next result we use Theorem 3.1 to characterize the boundary points of the polynomial numerical hull.

Theorem 3.2. *Let $A \in \mathscr{B}(\mathbb{H})$, $\zeta \in \mathbb{C}$, and $k \in \mathbb{N}$. Then the following statements are equivalent*:

1. $\zeta \in \mathrm{bdry}(\mathscr{H}_k(A))$.
2. *For all $c \in \mathbb{C}^k$*

$$0 \in \overline{\mathscr{F}}\left(\sum_{j=1}^{k} c_j (A - \zeta I)^j\right) \tag{9}$$

and there exists $\hat{c} \in \mathbb{C}^k \setminus \{0\}$ such that

$$0 \in \mathrm{bdry}\,\overline{\mathscr{F}}\left(\sum_{j=1}^{k} \hat{c}_j (A - \zeta I)^j\right). \tag{10}$$

3. *For all $c \in \mathbb{C}^k$*

$$\inf_{\|q\|=1} \mathrm{Re}\left\langle \left[\sum_{j=1}^{k} c_j (A - \zeta I)^j\right] q, q \right\rangle \leqslant 0$$

and there exists $\hat{c} \in \mathbb{C}^k \setminus \{0\}$ such that this infimum is zero.
4. $0 \in \mathrm{bdry}\overline{\mathscr{F}}(\{(A - \zeta I)^j\}_{j=1}^{k})$.
5. *For all $p \in \mathscr{P}_k$, $p(\zeta) \in \overline{\mathscr{F}}(p(A))$ and there exists a nonconstant $\hat{p} \in \mathscr{P}_k$ such that $\hat{p}(\zeta) \in \mathrm{bdry}\,\overline{\mathscr{F}}(\hat{p}(A))$.*

Proof. The equivalence of 2 and 3 follows from Lemma 2.1; the equivalence of 3 and 4 is a consequence of Lemma 2.2; and the equivalence of 2 and 5 follows from Lemmas 2.2 and 2.3. We conclude by establishing the equivalence of 1 and 2.

Condition (9) was already shown to be equivalent to the statement $\zeta \in \mathscr{H}_k(A)$; hence it holds for $\zeta \in \text{bdry}\mathscr{H}_k(A)$ since the set is closed.

To show that $\zeta \in \text{bdry}\mathscr{H}_k(A)$ implies (10), let ζ be a point on the boundary of $\mathscr{H}_k(A)$, and let $\{\zeta_m\}_{m=0}^{\infty}$ be a sequence of points outside $\mathscr{H}_k(A)$ and converging to ζ. For each point ζ_m, it follows from 3 in Theorem 3.1 that there exist coefficients $c_{1,m}, \ldots, c_{k,m}$ such that $0 \notin \overline{\mathscr{F}}(\sum_{j=1}^{k} c_{j,m}(A - \zeta_m I)^j)$. We can take these coefficients to satisfy $\sum_{j=1}^{k} |c_{j,m}|^2 = 1$, since multiplying an operator by a nonzero scalar just multiplies its field of values by that scalar and does not affect whether or not the closure of the field of values contains the origin. Since the vectors $(c_{1,m}, \ldots, c_{k,m}, \zeta_m)$ are bounded, there is a convergent subsequence, $\{(c_{1,m_\ell}, \ldots, c_{k,m_\ell}, \zeta_{m_\ell})\}_{\ell=1}^{\infty}$, converging to, say, $(\hat{c}_1, \ldots, \hat{c}_k, \zeta)$. Since the field of values is a continuous function of the linear operator [8] and since, by 3 in Theorem 3.1,

$$0 \notin \overline{\mathscr{F}}\left(\sum_{j=1}^{k} c_{j,m_\ell}(A - \zeta_{m_\ell} I)^j\right) \to \overline{\mathscr{F}}\left(\sum_{j=1}^{k} \hat{c}_j (A - \zeta I)^j\right) \ni 0,$$

it follows that $0 \in \text{bdry}\overline{\mathscr{F}}(\sum_{j=1}^{k} \hat{c}_j (A - \zeta I)^j)$.

Conversely, suppose ζ is a point in the interior of $\mathscr{H}_k(A)$ and suppose that $0 \in \text{bdry}\,\overline{\mathscr{F}}(\sum_{j=1}^{k} \times c_j(A - \zeta I)^j)$, for certain scalars $c_1, \ldots, c_k$. We will show that each c_j must be 0.

Since $\mathscr{H}_k(A)$ contains a disk of radius $r > 0$ about ζ, it follows from 3 in Theorem 3.1 that for any scalars $d_1, \ldots, d_k$, any $\delta \in [0, r)$, and any $\theta \in [0, 2\pi)$,

$$0 \in \overline{\mathscr{F}}\left(\sum_{j=1}^{k} d_j (A - (\zeta - \delta e^{i\theta}) I)^j\right). \tag{11}$$

Expanding using the binomial formula, we find

$$\sum_{j=1}^{k} d_j (A - (\zeta - \delta e^{i\theta}) I)^j = \sum_{j=1}^{k} d_j \sum_{\ell=0}^{j} \binom{j}{\ell} \delta^\ell e^{i\ell\theta} (A - \zeta I)^{j-\ell}.$$

Separating out the $\ell = 0$ and $\ell = j$ terms we get

$$\begin{aligned}
&\sum_{j=1}^{k} d_j (A - \zeta I)^j + \sum_{j=1}^{k} d_j \sum_{\ell=1}^{j-1} \binom{j}{\ell} \delta^\ell e^{i\ell\theta} (A - \zeta I)^{j-\ell} + \left(\sum_{j=1}^{k} d_j \delta^j e^{ij\theta}\right) I \\
&\quad = \sum_{j=1}^{k} \left(d_j + \sum_{\ell=1}^{k-j} d_{j+\ell} \binom{j+\ell}{\ell} \delta^\ell e^{i\ell\theta}\right) (A - \zeta I)^j + \left(\sum_{j=1}^{k} d_j \delta^j e^{ij\theta}\right) I.
\end{aligned}$$

Suppose $d_k = c_k$ and d_j, $j = k-1, \ldots, 1$ are determined by

$$d_j = c_j - \sum_{\ell=1}^{k-j} d_{j+\ell} \binom{j+\ell}{\ell} \delta^\ell e^{i\ell\theta}. \tag{12}$$

Then

$$\sum_{j=1}^{k} d_j (A - (\zeta - \delta e^{i\theta}) I)^j = \sum_{j=1}^{k} c_j (A - \zeta I)^j + \left(\sum_{j=1}^{k} d_j \delta^j e^{ij\theta}\right) I \tag{13}$$

and the field of values of this matrix is that of $\sum_{j=1}^{k} c_j(A-\zeta I)^j$ shifted by $\sum_{j=1}^{k} d_j\delta^j \mathrm{e}^{\mathrm{i}j\theta}$. To express $\sum_{j=1}^{k} d_j\delta^j \mathrm{e}^{\mathrm{i}j\theta}$ in terms of the c_j's, note that if $\sum_{j=1}^{k} c_j(A-\zeta I)^j = \sum_{j=1}^{k} c_j(A-(\zeta-\delta\mathrm{e}^{\mathrm{i}\theta})I-\delta\mathrm{e}^{\mathrm{i}\theta}I)^j$ is expanded using the binomial formula, then the expansion involves powers from 1 to k of $A-(\zeta-\delta\mathrm{e}^{\mathrm{i}\theta})I$ plus the term $(\sum_{j=1}^{k} c_j(-1)^j\delta^j\mathrm{e}^{\mathrm{i}j\theta})I$. Comparing this with (13), it follows that

$$\sum_{j=1}^{k} d_j\delta^j\mathrm{e}^{\mathrm{i}j\theta} = \sum_{j=1}^{k} c_j(-1)^{j+1}\delta^j\mathrm{e}^{\mathrm{i}j\theta}. \tag{14}$$

If the origin is on the boundary of $\overline{\mathscr{F}}(\sum_{j=1}^{k} c_j(A-\zeta I)^j)$, then there is a line through the origin that separates the field of values from a half-plane. Suppose $c_1 = \cdots = c_{\ell-1} = 0$ but $c_\ell \neq 0$. Choose θ so that $\mathrm{e}^{\mathrm{i}\ell\theta}(-1)^{\ell+1}c_\ell$ lies in the half-plane containing $\overline{\mathscr{F}}(\sum_{j=1}^{k} c_j(A-\zeta I)^j)$ and is orthogonal to the separating line. Then if the field of values is shifted in the direction of $\mathrm{e}^{\mathrm{i}\ell\theta}(-1)^{\ell+1}c_\ell$ its closure will exclude the origin. By choosing $\delta > 0$ sufficiently small, one can make the shift term in (14) arbitrarily close to $c_\ell(-1)^{\ell+1}\delta^\ell\mathrm{e}^{\mathrm{i}\ell\theta}$, and so one can exclude the origin from the closure of the field of values of $\sum_{j=1}^{k} d_j(A-(\zeta-\delta\mathrm{e}^{\mathrm{i}\theta})I)^j$, but this contradicts (11). Therefore each coefficient c_j must be 0. □

4. Polynomial numerical hulls of Toeplitz matrices and operators

In this section T denotes an infinite-dimensional Toeplitz operator:

$$T = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots \\ a_1 & a_0 & a_{-1} & \dots \\ a_2 & a_1 & a_0 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{15}$$

The *symbol* of T evaluated at a point z is defined as

$$\mathsf{a}(z) := a_0 + \sum_{\ell=1}^{\infty} a_\ell z^\ell + \sum_{\ell=1}^{\infty} a_{-\ell} z^{-\ell}. \tag{16}$$

We will assume throughout that $\mathsf{a} \in L^\infty(\mathscr{U})$, where $\mathscr{U} := \{z \in \mathbb{C} : |z| = 1\}$ denotes the unit circle, so that the matrix in (15) corresponds to a bounded linear operator on ℓ^2 (see, e.g. [1]).

The polynomial numerical hull of degree 1 for T, i.e., the closure of the field of values, is known [9]:

$$\mathscr{H}_1(T) = \mathrm{co}(\mathsf{a}(\mathscr{U})), \tag{17}$$

where $\mathsf{a}(\mathscr{U}) := \{\mathsf{a}(z) : z \in \mathscr{U}\}$. An extension to $\mathscr{H}_k(T)$ is obtained using the *polynomially convex hull* of degree k [4], defined for any compact set $S \subset \mathbb{C}$ as

$$\mathrm{pco}_k(S) := \{z \in \mathbb{C} : |p(z)| \leqslant \max_{\zeta\in S} |p(\zeta)|\ \forall p \in \mathscr{P}_k\}.$$

Theorem 4.1. *The polynomial numerical hull of degree $k \geqslant 1$ for T satisfies*

$$\mathrm{co}(\mathsf{a}(\mathscr{U})) \supset \mathscr{H}_k(T) \supset \mathrm{pco}_k(\mathsf{a}(\mathscr{U})). \tag{18}$$

Proof. The left inclusion follows from (17) and the fact that $\mathscr{H}_k(T) \subset \mathscr{H}_1(T)$. To establish the right inclusion, let $p(T)$ be any polynomial in T. By the spectral mapping theorem, $\sigma(p(T)) =$

$p(\sigma(T))$, where $\sigma(\cdot)$ denotes the spectrum. By the Hartman–Wintner theorem (see, e.g. [1, Theorem 1.25, p. 27]), the essential spectrum of T contains the range of its symbol $\mathsf{a}(\mathscr{U})$. Hence

$$\|p(T)\| \geqslant \sup_{\zeta\in\sigma(p(T))} |\zeta| \geqslant \sup_{|z|=1} |p(\mathsf{a}(z))|.$$

The result then follows from definition (1) of $\mathscr{H}_k(T)$. □

In words, the polynomial numerical hull of degree k for an infinite-dimensional Toeplitz operator lies somewhere between the polynomially convex hull of degree k and the ordinary convex hull of the image, under the symbol, of the unit circle $\mathscr{U}$.

Let T_N be the N by N Toeplitz matrix consisting of the upper left block of the operator in (15):

$$T_N = \begin{pmatrix} a_0 & a_{-1} & a_{-2} & \dots & a_{-(N-1)} \\ a_1 & a_0 & a_{-1} & \dots & a_{-(N-2)} \\ a_2 & a_1 & a_0 & \dots & a_{-(N-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N-1} & a_{N-2} & a_{N-3} & \dots & a_0 \end{pmatrix}. \tag{19}$$

It is known that if T is banded or, more generally, if a is *piecewise continuous* on the unit circle, then the polynomial numerical hull of degree 1 of T_N approaches that of T as $N \to \infty$ [3,6,12]. Moreover, if $p(T_N)$ is any polynomial in T_N and if a is piecewise continuous on the unit circle, then it follows from results in [12, Theorem 2] that

$$\text{p-lim}_{N\to\infty}\overline{\mathscr{F}}(p(T_N)) = \text{u-lim}_{N\to\infty}\overline{\mathscr{F}}(p(T_N)) = \overline{\mathscr{F}}(p(T)). \tag{20}$$

Here $\text{u-lim}_{N\to\infty}S_N$, for a sequence of sets S_N, denotes the set of all limits of sequences of points $\{s_N \in S_N\}_{N=1}^{\infty}$, while $\text{p-lim}_{N\to\infty}S_N$ denotes all limits of subsequences $\{s_{N_\ell} \in S_{N_\ell}\}_{\ell=1}^{\infty}$, where $(N_1, N_2, \dots)$ is a subsequence of $(1, 2, \dots)$. Roch [12] refers to these set-valued limits as the *uniform* and *partial* limits of the sequence $\{S_N\}$, respectively. An extensive literature on these notions of limit dates from the original work of Painlevé in 1902. The uniform limit is more commonly known as the limit infimum, or inner limit, and the partial limit is more commonly known as the limit supremum, or outer limit [13, Chapter 4]. Clearly $\text{u-lim}_{N\to\infty}S_N \subset \text{p-lim}_{N\to\infty}S_N$. Note that if $\zeta \in \text{int}(\overline{\mathscr{F}}(p(T)))$, then (20) implies that for N large enough, $\zeta \in \text{int}(\overline{\mathscr{F}}(p(T_N)))$, since these sets are convex and they contain sequences of points approaching points on a circle about ζ in $\text{int}(\overline{\mathscr{F}}(p(T)))$ (see, for example [13, Proposition 4.15]).

Theorem 4.2. *Assume* a *in* (16) *is piecewise continuous on the unit circle. For any fixed degree* $k \geqslant 1$,

$$\mathscr{H}_k(T) \supset \text{p-lim}_{N\to\infty}\mathscr{H}_k(T_N) \supset \text{u-lim}_{N\to\infty}\mathscr{H}_k(T_N) \supset \text{cl}(\text{int}(\mathscr{H}_k(T))), \tag{21}$$

where $\text{cl}(\cdot)$ *denotes the closure.*

Proof. First suppose $\zeta_{N_\ell} \in \mathscr{H}_k(T_{N_\ell})$ is a convergent subsequence with $\zeta_{N_\ell} \to \zeta$ as $\ell \to \infty$. It follows from 3 Theorem 3.1 that for all $c_1, \dots, c_k$, $0 \in \overline{\mathscr{F}}(\sum_{j=1}^{k} c_j(T_{N_\ell} - \zeta_{N_\ell} I)^j)$. Since the field of values is a continuous multi-valued function of the operator, it follows that for each $c_1, \dots, c_k$, there is a sequence of points $s_{N_\ell} \in \overline{\mathscr{F}}(\sum_{j=1}^{k} c_j(T_{N_\ell} - \zeta I)^j)$ such that $s_{N_\ell} \to 0$ as $\ell \to \infty$. By (20), this implies that $0 \in \overline{\mathscr{F}}(\sum_{j=1}^{k} c_j(T - \zeta I)^j)$, and hence, again by 3 Theorem

3.1, $\zeta \in \mathscr{H}_k(T)$. This establishes the first inclusion in (21). The second inclusion is clear from the definitions of u-lim and p-lim.

Let $\zeta \in \operatorname{int}(\mathscr{H}_k(T))$. Then, by 2 Theorem 3.2, we know that for all $c_1, \ldots, c_k$ with $\sum_{j=1}^k |c_j|^2 = 1$, $0 \in \operatorname{int}(\overline{\mathscr{F}}(\sum_{j=1}^k c_j(T - \zeta I)^j))$. It follows from (20) that there exists $n \equiv n(c_1, \ldots, c_k)$ such that for all $N \geqslant n$, $0 \in \operatorname{int}(\overline{\mathscr{F}}(\sum_{j=1}^k c_j(T_N - \zeta I)^j))$. Since the coefficients come from a compact set and since the field of values is a continuous multi-valued function of $c_1, \ldots, c_k$, there exist coefficients $\hat{c}_1, \ldots, \hat{c}_k$ for which the required minimum n-value, $\hat{n} \equiv n(\hat{c}_1, \ldots, \hat{c}_k)$, is maximal and finite. Therefore, for $N \geqslant \hat{n}$ and for all $c_1, \ldots, c_k$ with $\sum_{j=1}^k |c_j|^2 = 1$, $0 \in \operatorname{int}(\overline{\mathscr{F}}(\sum_{j=1}^k c_j(T_N - \zeta I)^j))$. Hence, by 2 Theorem 3.2, $\zeta \in \operatorname{int}(\mathscr{H}_k(T_N))$. This shows that every interior point of $\mathscr{H}_k(T)$ eventually lies in $\mathscr{H}_k(T_N)$ and hence that every limit of interior points of $\mathscr{H}_k(T)$ is a limit of points in $\mathscr{H}_k(T_N)$. This proves the third inclusion in (21). □

The preceding theorems leave open the possibility that $\mathscr{H}_k(T)$ contains isolated points or curves between the region enclosed by $\mathsf{a}(\mathscr{U})$ and $\operatorname{co}(\mathsf{a}(\mathscr{U}))$, and, in this case, it is not known if there is a sequence or subsequence of points in $\mathscr{H}_k(T_N)$ converging to these points. In the simplest case, where the region enclosed by $\mathsf{a}(\mathscr{U})$ is a convex set, Theorems 4.1 and 4.2 imply that $\text{p-lim}_{N\to\infty}\mathscr{H}_k(T_N) = \text{u-lim}_{N\to\infty}\mathscr{H}_k(T_N) = \mathscr{H}_k(T) = \operatorname{co}(\mathsf{a}(\mathscr{U}))$.

Finally, note that Theorem 4.2 makes no direct use of Toeplitz properties; it uses only relation (20) and properties of the polynomial numerical hull. Hence for any sequence of matrices or operators A_N satisfying

$$\text{p-lim}_{N\to\infty}\overline{\mathscr{F}}(p(A_N)) = \text{u-lim}_{N\to\infty}\overline{\mathscr{F}}(p(A_N)) = \overline{\mathscr{F}}(p(A))$$

for all polynomials p, the analogue of relation (21) will hold, namely,

$$\mathscr{H}_k(A) \supset \text{p-lim}_{N\to\infty}\mathscr{H}_k(A_N) \supset \text{u-lim}_{N\to\infty}\mathscr{H}_k(A_N) \supset \operatorname{cl}(\operatorname{int}(\mathscr{H}_k(A)))$$

for any fixed $k \geqslant 1$.

Acknowledgment

The authors thank Nick Trefethen for helpful comments on a draft of this paper.

References

[1] A. Böttcher, B. Silbermann, Introduction to Large Truncated Toeplitz Matrices, Springer, 1999.
[2] E.B. Davies, Spectral bounds using higher order numerical ranges, London Math. Soc. J. Comput. Math. 8 (2005) 17–45.
[3] M. Eiermann, Fields of values and iterative methods, Linear Algebra Appl. 180 (1993) 167–197.
[4] V. Faber, A. Greenbaum, D.E. Marshall, The polynomial numerical hulls of Jordan blocks and related matrices, Linear Algebra Appl. 374 (2003) 231–246.
[5] A. Greenbaum, Generalizations of the field of values useful in the study of polynomial functions of a matrix, Linear Algebra Appl. 347 (2002) 233–249.
[6] A. Greenbaum, Some theoretical results derived from polynomial numerical hulls of Jordan blocks, ETNA 18 (2004) 81–90.
[7] A. Greenbaum, Card shuffling and the polynomial numerical hull of degree k, SIAM J. Sci. Comput. 25 (2004) 408–416.
[8] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
[9] E.M. Klein, The numerical range of a Toeplitz operator, Proc. Amer. Math. Soc. 35 (1972) 101–103.
[10] O. Nevanlinna, Convergence of Iterations for Linear Equations, Birkhäuser, Basel, 1993.

[11] O. Nevanlinna, Hessenberg matrices in Krylov subspaces and the computation of the spectrum, Numer. Funct. Anal. Optimiz. 16 (1995) 443–473.
[12] S. Roch, Numerical ranges of large Toeplitz matrices, Linear Algebra Appl. 282 (1998) 185–198.
[13] R.T. Rockafellar, J.B. Wets, Variational Analysis, Springer, 1998.
[14] L.N. Trefethen, M. Embree, Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators, Princeton University Press, 2005.