Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Linear Algebra and its Applications 383 (2004) 151–161

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

On polynomial numerical hulls of normal matrices

Ch. Davis [a,*], A. Salemi [b]

[a] Department of Mathematics, University of Toronto, Toronto, Ont., Canada M5S 3G3

[b] Department of Mathematics, Shahid Bahonar University of Kerman, Kerman 7616914111, Iran

Received 26 September 2003; accepted 7 December 2003

Submitted by R. Bhatia

Dedicated to Heydar Radjavi

Abstract

The notion of polynomial numerical hull was introduced by O. Nevanlinna [Convergence of Iteration for linear equations, Birkhäuser, 1993]. In this note we determine the polynomial numerical hulls of matrices of the form $A = A_1 \oplus \mathrm{i}A_2$, where A_1, A_2 are hermitian matrices. Also we study the relationship between rectangular hyperbolas and polynomial numerical hulls of order two for normal matrices. The polynomial numerical hulls of order two for some special matrices is studied.

AMS classification: 15A60; 15A18

Keywords: Polynomial numerical hull; Numerical order; Rectangular hyperbola

1. Introduction

Let $A \in M_n(\mathbb{C})$, where $M_n(\mathbb{C})$ denotes the set of all $n \times n$ complex matrices. The numerical range of A is denoted by

$$W(A) := \{x^*Ax : \|x\| = 1\}. \tag{1}$$

* Corresponding author.

E-mail addresses: davis@math.toronto.edu (Ch. Davis), salemi@mail.uk.ac.ir (A. Salemi).

0024-3795/$ - see front matter © 2004 Published by Elsevier Inc.

doi:10.1016/j.laa.2003.12.025

Let $p(\lambda)$ be any complex polynomial. Define

$$V_p(A) := \{\lambda : |p(\lambda)| \leqslant \|p(A)\|\}. \tag{2}$$

If p is not constant, $V_p(A)$ is a compact set containing $\sigma(A)$. (For more details see [2].) The *polynomial numerical hull of A of order k* is denoted by

$$V^k(A) := \bigcap V_p(A), \tag{3}$$

where the intersection is taken over all polynomials p of degree at most k. Greenbaum [4] proved that if A is a normal matrix, then

$$V^k(A) = \{x^*Ax : x^*x = 1 \text{ and } x^*A^ix = (x^*Ax)^i,\ i = 1, \ldots, k\}. \tag{4}$$

The intersection over all polynomials is called the *polynomial numerical hull of A* and is denoted by

$$V(A) := \bigcap_{k=1}^{\infty} V^k(A). \tag{5}$$

The following proposition due to Nevanlinna states the relationship between the polynomial numerical hull of order one and the numerical range of a bounded operator.

Proposition 1.1. *Let L be a bounded operator in a Hilbert space H. Then $V^1(L) = \overline{W(L)}$. (See* [2], *Theorem* 2.10.5.)

In the above proposition, if $\dim H < \infty$, then $V^1(L) = W(L)$.

If $V^m(A) = V(A)$ and $V^i(A) \neq V(A)$, $\forall i < m$, then the integer m is called the *numerical order* of A and is denoted by $\mathrm{num}(A)$. It is obvious that if the degree of the minimal polynomial of $A \in M_n(\mathbb{C})$ is k, then $V^k(A) = V^{k+1}(A) = \cdots = V(A) = \sigma(A)$ and hence the numerical order of A is less than or equal to the degree of the minimal polynomial of A.

Throughout the paper all direct sums are assumed to be orthogonal.

Lemma 1.2. *Let $A \in M_n(\mathbb{C})$ and $\alpha \in V^k(A)$. Let $A' = A \oplus [\alpha]$, where $[\alpha]$ denotes a 1×1 matrix with entry α. Then $V^k(A') = V^k(A)$.*

Proof. Let p be an arbitrary polynomial of degree at most k. Since $|p(\alpha)| \leqslant \|p(A)\|$ and $\|p(A')\| = \max\{\|p(A)\|, |p(\alpha)|\}$, it follows that $\|p(A)\| = \|p(A')\|$ and hence $V^k(A') = V^k(A)$. □

The following proposition is due to Nevanlinna ([3], Proposition 3.10); for a shorter proof see [4].

Proposition 1.3. *Let L be a bounded self-adjoint operator in a Hilbert space H. Then* $\mathrm{num}(L) \leqslant 2$ *and hence $V^2(L) = \sigma(L)$.*

In the following, we show that if A is a normal matrix and a boundary point λ of $W(A)$ belongs to $V^2(A)$, then λ is an eigenvalue of A.

Theorem 1.4. *Let $A \in M_n(\mathbb{C})$ be a normal matrix. Then $\partial(W(A)) \cap V^2(A) \subseteq \sigma(A)$, where $\partial(D)$ means the boundary points of D.*

Proof. We know that $W(A)$ is a polygon. Let I be an arbitrary side of the polygon $W(A)$. As we are free to translate and rotate, we assume without loss of generality that $I \subseteq \mathbb{R}$ and $\Im(W(A)) \geqslant 0$. And by unitary conjugation we can assume our operator is in the form $A = A_1 \oplus A_2$, where $W(A_1) = I$ and $\Im(W(A_2)) > 0$. It is enough to show that $V^2(A) \cap I \subseteq \sigma(A)$. Let $x \in \mathbb{C}^n$ be any unit vector and write $x = x_1 \oplus x_2$. If $x^*Ax \in I$ then since $x^*Ax \in \mathbb{R}$ and $x_1^*A_1x_1 \in \mathbb{R}$ it follows that $x_2^*A_2x_2 \in \mathbb{R}$ and hence $\Im(x_2^*A_2x_2) = 0$. Since $\Im(x_2^*A_2x_2) > 0$, $x_2 = 0$ and $\|x_1\| = \|x\| = 1$. If also $x^*Ax \in V^2(A)$, then $x^*A^2x = (x^*Ax)^2$, thus $x_1^*A_1^2x_1 = (x_1^*A_1x_1)^2$; that is, $x_1^*A_1x_1 \in V^2(A_1)$. But A_1 is Hermitian, so by Proposition 1.3, $x^*Ax = x_1^*A_1x_1 \in \sigma(A_1) \subseteq \sigma(A)$. $\square$

Throughout the paper we fix some notations. Define $\mathrm{i}[a, b] = \{\mathrm{i}t : a \leqslant t \leqslant b\}$ and $\mathrm{i}(a, b) = \{\mathrm{i}t : a < t < b\}$, where a and b are real numbers. Also define $\overline{\{z_1, z_2\}}$ to be the line passing through the points z_1 and z_2, where z_1 and z_2 are complex numbers.

2. Polynomial numerical hull of order 2

In this section we consider matrices $A \in M_n(\mathbb{C})$ of the form

$$A = A_1 \oplus \mathrm{i}A_2, \quad A_1^* = A_1, A_2^* = A_2. \tag{6}$$

Theorem 2.1. *Let A be of the form* (6). *Let A_1 and A_2 be semidefinite matrices. Then* $\mathrm{num}(A) \leqslant 2$.

Proof. Let $x = x_1 \oplus x_2$ be a unit vector such that $x^*Ax = x_1^*A_1x_1 + \mathrm{i}x_2^*A_2x_2 \in V^2(A)$; we are to prove that $x^*Ax \in \sigma(A)$. We know

$$\begin{aligned}(x^*Ax)^2 &= (x_1^*A_1x_1)^2 - (x_2^*A_2x_2)^2 + 2\mathrm{i}(x_1^*A_1x_1)(x_2^*A_2x_2),\\ (x^*A^2x) &= (x_1^*A_1^2x_1) - (x_2^*A_2^2x_2).\end{aligned} \tag{7}$$

Using (4), $(x_1^*A_1x_1)(x_2^*A_2x_2) = 0$. Assume without loss of generality that $x_2^*A_2x_2 = 0$ and $x_1^*A_1x_1 \neq 0$ (otherwise $x^*Ax = 0 \in \sigma(A)$). Since A_2 is semidefinite, $A_2x_2 = 0$ and hence $(x_1^*A_1x_1)^2 = x_1^*A_1^2x_1 = \|A_1x_1\|^2$. By the Cauchy–Schwarz inequality, $(x_1^*A_1x_1)^2 \leqslant \|x_1\|^2\|A_1x_1\|^2$, and hence $\|x_1\| = 1$ and equality holds. Then x_1 is parallel to A_1x_1, so that $x_1^*A_1x_1 \in \sigma(A_1) \subseteq \sigma(A)$. $\square$

Theorem 2.2. *Let A be of the form* (6) *and A_2 be a positive semidefinite matrix. Then*

$$V^2(A) \subseteq \sigma(A_1) \cup \mathrm{i}[0, r(A_2)],$$

where $r(A_2)$ is the spectral radius of A_2.

Proof. Let $x = x_1 \oplus x_2$ be a unit vector such that $x^*Ax = x_1^*A_1x_1 + x_2^*A_2x_2 \in V^2(A)$. Since $x^*A^2x = (x^*Ax)^2$, by (7), $(x_1^*A_1x_1)(x_2^*A_2x_2) = 0$. If $x_2^*A_2x_2 = 0$, and $x_1^*A_1x_1 \neq 0$, then in the same manner as in the proof of Theorem 2.1, $x^*Ax = x_1^*A_1x_1 \in \sigma(A_1)$. If $x_1^*A_1x_1 = 0$, then $x^*Ax = \mathrm{i}x_2^*A_2x_2 \subseteq \{\mathrm{i}\alpha : 0 \leqslant \alpha \leqslant r(A_2)\}$. Thus $V^2(A) \subseteq \sigma(A_1) \cup \mathrm{i}[0, r(A_2)]$. $\square$

The following example shows that the inclusion asserted in Theorem 2.2 for $V^2(A)$ can become equality.

Example 1. Let $A = \mathrm{diag}(-1, 1, 0) \oplus \{\mathrm{i}\}$. Then $V^2(A) = \{-1, 1, 0\} \cup \{\mathrm{i}\alpha : 0 \leqslant \alpha \leqslant 1\}$.

The following key lemma plays an important role throughout the paper.

Lemma 2.3. *Let $A \in M_n(\mathbb{C})$ be a normal matrix. Assume the* (*complex*) *plane is divided into four* (*closed*) *quadrants Q_1, Q_2, Q_3 and Q_4* (*ordered counter clockwise*) *made by two perpendicular lines. Let $\sigma(A)$ be contained in the union $Q_1 \cup Q_3$ of the opposite quadrants Q_1 and Q_3. Then $V^2(A) \subseteq Q_1 \cup Q_3$.*

Proof. Assume without loss of generality that, up to a translation and a rotation, the perpendicular lines are the coordinate axes and each Q_i is the usual ith quadrant of the xoy-plane ($i = 1, 2, 3, 4$). We know that the set $\{z^2 : z \in V^2(A)\} \subseteq W(A^2)$. Since $W(A^2)$ is the convex hull of $\sigma(A^2)$ and $\sigma(A^2) \subseteq Q_1 \cup Q_2$, it follows that $W(A^2) \subset Q_1 \cup Q_2$. Let z be a complex number in the interior of Q_2 or Q_4. Thus z^2 belongs to the interior of $Q_3 \cup Q_4$. Therefore, z does not belong to $V^2(A)$. We have proved that $V^2(A) \subseteq Q_1 \cup Q_3$. $\square$

The following theorem characterizes the polynomial numerical hulls of all 3×3 normal matrices. Note that if the spectrum of a normal matrix A lies on a line, then it follows from Proposition 1.3 that $V^2(A) = \sigma(A)$.

Theorem 2.4. *Let $A \in M_n(\mathbb{C})$ be a normal matrix whose spectrum consists of three non-colinear points. Then $V^2(A) = \sigma(A) \cup (\{\lambda_0\} \cap W(A))$, where λ_0 is the orthocenter of the triangle $\sigma(A)$.*

Proof. In view of Lemma 1.2, we can assume without loss of generality that $n = 3$. Making a translation followed by a rotation, putting the longest side on the real axis, we can write $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma)$, where α, β, and γ are positive numbers. Applying the key lemma to the four quadrants made by the altitude from the vertex α of the triangle $\{\alpha, -\beta, \mathrm{i}\gamma\}$ to the opposite side $\overline{\{-\beta, \mathrm{i}\gamma\}}$, it follows that $V^2(A)$ is a subset of the union of these two lines. Another application of the key lemma implies that $V^2(A)$ is a subset of the union of the altitude from the vertex $-\beta$ and the opposite line $\overline{\{\alpha, \mathrm{i}\gamma\}}$. Therefore $V^2(A) \subseteq \sigma(A) \cup \{\lambda_0\}$. Since $\sigma(A) \subseteq V^2(A) \subseteq W(A)$, it remains only to show that if $\lambda_0 \in W(A)$, then $\lambda_0 \in V^2(A)$. Since $W(A)$ is the convex hull of $\sigma(A)$, it follows that $W(A)$ is a triangle. If the triangle has a right angle (at $\mathrm{i}\gamma$), then $\lambda_0 = \mathrm{i}\gamma \in \sigma(A) \subseteq V^2(A)$. If the triangle has an obtuse angle (at $\mathrm{i}\gamma$), then λ_0 does not belong to $W(A)$. It remains to consider the case that the triangle $\{\alpha, -\beta, \mathrm{i}\gamma\}$ is an acute-angle triangle. In this case $\alpha\beta < \gamma^2$ and $\lambda_0 = \mathrm{i}(\alpha\beta/\gamma)$. Define a unit vector $X = (x_1, x_2, x_3)^{\mathrm{T}}$, where $|x_1|^2 = \beta(\gamma^2 - \alpha\beta)/[\gamma^2(\alpha+\beta)]$, $|x_2|^2 = \alpha(\gamma^2 - \alpha\beta)/[\gamma^2(\alpha+\beta)]$, and $|x_3|^2 = \alpha\beta/\gamma^2$. Hence $X^*AX = \mathrm{i}(\alpha\beta/\gamma)$ and $X^*A^2X = -(\alpha\beta/\gamma)^2 = (X^*AX)^2$. Thus $\lambda_0 = \mathrm{i}(\alpha\beta/\gamma) \in V^2(A)$. $\square$

We now study normal matrices $A \in M_n(\mathbb{C})$ of the form (6) whose spectra have four points. As in the case of Theorem 2.4, it is enough to consider matrices whose spectra are non-colinear. We take up different configurations in turn.

Theorem 2.5. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, 0)$, *where* α, β, *and* γ *are positive numbers. Then* $V^2(A) = \sigma(A) \cup \{\mathrm{i}s : 0 \leqslant s \leqslant \min\{(\alpha\beta)/\gamma, \gamma\}\}$ *and so* $\mathrm{num}(A) > 2$.

Proof. By Theorem 2.2, $V^2(A) \subseteq \{\alpha, -\beta\} \cup \{\mathrm{i}s : 0 \leqslant s \leqslant \gamma\}$. Now, consider the four quadrants Q_1, Q_2, Q_3 and Q_4 (ordered counter-clockwise) of the plane made by the altitude from the vertex $-\beta$ of the triangle $\{\alpha, -\beta, \mathrm{i}\gamma\}$ and the opposite line $\overline{\{\alpha, \mathrm{i}\gamma\}}$, such that Q_1 is the one containing 0. Applying the key lemma to A implies that $V^2(A) \subseteq Q_1 \cup Q_3$, and hence $V^2(A) \subseteq \{\alpha, -\beta\} \cup \{\mathrm{i}s : 0 \leqslant s \leqslant \min\{(\alpha\beta)/\gamma, \gamma\}\}$. For the converse, let $0 \leqslant \eta \leqslant \min\{(\alpha\beta)/\gamma, \gamma\}$; we must show that $\mathrm{i}\eta \in V^2(A)$. Define a unit vector $X = (x, y, z, t)^{\mathrm{T}}$ such that $|x|^2 = (\gamma\eta - \eta^2)/(\alpha\beta + \alpha^2)$, $|y|^2 = (\gamma\eta - \eta^2)/(\alpha\beta + \beta^2)$, $|z|^2 = \eta/\gamma$, and $|t|^2 = (\gamma - \eta)(\alpha\beta - \eta\gamma)/\alpha\beta\gamma$. It is easy to see that $\mathrm{i}\eta = X^*AX$ and $X^*A^2X = (X^*AX)^2$. Therefore, $V^2(A) = \sigma(A) \cup \{\mathrm{i}s : 0 \leqslant s \leqslant \min\{(\alpha\beta)/\gamma, \gamma\}\}$, and hence $\mathrm{num}(A) > 2$. $\square$

Let $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, -\mathrm{i}\theta)$, where α, β, γ and θ are positive numbers. Define $\hat{\alpha} := \min\{\gamma\theta/\alpha, \alpha\}$, $\hat{\beta} := \min\{\gamma\theta/\beta, \beta\}$, $\hat{\gamma} := \min\{\alpha\beta/\gamma, \gamma\}$, $\hat{\theta} := \min\{\alpha\beta/\theta, \theta\}$. We conclude this section by studying other matrices whose spectra consists of four points lying on the union of two perpendicular lines. We assume without loss of generality that the lines are coordinate axes.

Theorem 2.6. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, -\mathrm{i}\theta)$, *where* α, β, γ *and* θ *are positive numbers. Let* $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ *and* $\hat{\theta}$ *be as above. Then* $V^2(A) = \sigma(A) \cup [-\hat{\beta}, \hat{\alpha}] \cup \mathrm{i}[-\hat{\theta}, \hat{\gamma}]$ *and so* $\mathrm{num}(A) > 2$.

Proof. We know that $V^2(A) \subseteq W(A)$. Applying the key lemma to the matrix A with respect to the coordinate axes, it follows that $V^2(A) \subseteq [-\beta, \alpha] \cup \mathrm{i}[-\theta, \gamma]$. Consider a unit vector $X = (x, y, z, t)^{\mathrm{T}}$ such that $|t|^2 = \alpha\beta\gamma/[(\gamma\theta + \alpha\beta)(\theta + \gamma)]$, $|x|^2 = \theta(\gamma + \theta)|t|^2/[\alpha(\alpha + \beta)]$, $|y|^2 = \theta(\gamma + \theta)|t|^2/[\beta(\alpha + \beta)]$, and $|z|^2 = \theta|t|^2/\gamma$. It is easy to show that $X^*AX = 0$ and $X^*A^2X = (X^*AX)^2 = 0$, so $0 \in V^2(A)$. Define $\widetilde{A} = \mathrm{diag}(\alpha, -\beta, 0, \mathrm{i}\gamma, -\mathrm{i}\theta)$. By Lemma 1.2, $V^2(A) = V^2(\widetilde{A})$. Now apply Theorem 2.5 successively to direct summands of $\widetilde{A}$, $\mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, 0)$, $\mathrm{diag}(\alpha, -\beta, 0, -\mathrm{i}\theta)$, $\mathrm{diag}(\alpha, 0, \mathrm{i}\gamma, -\mathrm{i}\theta)$ and $\mathrm{diag}(-\beta, 0, \mathrm{i}\gamma, -\mathrm{i}\theta)$; it follows that $\sigma(A) \cup [-\hat{\beta}, \hat{\alpha}] \cup \mathrm{i}[-\hat{\theta}, \hat{\gamma}] \subseteq V^2(A)$. It remains to show that if any one of the orthocenters $\mathrm{i}\alpha\beta/\gamma$, $-\mathrm{i}\alpha\beta/\theta$, $\gamma\theta/\alpha$, $-\gamma\theta/\beta$ is inside the corresponding triangle $\{\alpha, -\beta, \mathrm{i}\gamma\}$, $\{\alpha, -\beta, -\mathrm{i}\theta\}$, $\{\mathrm{i}\gamma, -\mathrm{i}\theta, \alpha\}$, or $\{\mathrm{i}\gamma, -\mathrm{i}\theta, -\beta\}$, respectively, then the open line segment joining that orthocenter to the corresponding vertex is not in $V^2(A)$. For instance, assuming $\tilde{\gamma} = \alpha\beta/\gamma < \gamma$, we must show that $\mathrm{i}(\tilde{\gamma}, \gamma)$ is not included in $V^2(A)$. For this case apply the key lemma to the matrix $\widetilde{A}$ with respect to the four quadrants made by the perpendicular lines $\overline{\{\alpha, \mathrm{i}\gamma\}}$ and $\overline{\{-\beta, \mathrm{i}\hat{\gamma}\}}$; it implies that the segment $\mathrm{i}(\hat{\gamma}, \gamma)$ is excluded from $V^2(A)$. Thus $V^2(A) = \sigma(A) \cup [-\hat{\beta}, \hat{\alpha}] \cup \mathrm{i}[-\hat{\theta}, \hat{\gamma}]$ and so $\mathrm{num}(A) > 2$. □

Remark 1. This solves the problem of finding $V^*(A)$ when $A \in M_n(\mathbb{C})$ is a normal matrix whose spectrum consists of the vertices of a convex quadrilateral whose diagonals are orthogonal. (Without loss of generality, we took the diagonals along the coordinate axes.) We now study the case in which the quadrilateral is non-convex. This is covered by the following theorems.

Theorem 2.7. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, \mathrm{i}\theta)$, *such that* α, β *are positive numbers and* $0 \leqslant \theta < \gamma \leqslant (\alpha\beta)^{1/2}$. *Then* $V^2(A) = \sigma(A) \cup \mathrm{i}[\theta, \gamma]$, *and so* $\mathrm{num}(A) > 2$.

Proof. Assume without loss of generality that $\gamma = 1$. By Theorem 2.2, $V^2(A) \subseteq \{\alpha, -\beta\} \cup \{\mathrm{i}t : 0 \leqslant t \leqslant 1\}$. In order that $\mathrm{i}\eta \in V^2(A)$, $0 < \eta < 1$. there must exist a unit vector $X = (x, y, z, t)^{\mathrm{T}}$ such that $X^*AX = \mathrm{i}(\theta|z|^2 + |t|^2) = \mathrm{i}\eta$. Since $X^*A^2X = (X^*AX)^2$, it follows that $\alpha|x|^2 = \beta|y|^2$ and $\alpha^2|x|^2 + \beta^2|y|^2 - \theta^2|z|^2 - |t|^2 = -(\theta|z|^2 + |t|^2)^2$. Thus $(\theta|z|^2 + |t|^2)^2 + \alpha\beta(1 - |z|^2 - |t|^2) - \theta^2|z|^2 - |t|^2 = 0$. Since $\eta = \theta|z|^2 + |t|^2$, it follows that $\eta^2 - (\alpha\beta + 1)\eta - (1 - \theta)(\alpha\beta - \theta)|z|^2 + \alpha\beta = 0$. Now if this equation is regarded as expressing $|z|^2$ as a function of η, one easily shows that as η runs from θ to 1, $|z|^2$ is strictly decreasing from 1 to 0. So all these values of η are possible, and $\mathrm{i}[\theta, 1] \subseteq V^2(A)$. For the converse, we consider the line $\overline{\{-\beta, \mathrm{i}\theta\}}$ and its perpendicular line passing through α. Applying the key lemma to these two lines, $V^2(A) = \sigma(A) \cup \mathrm{i}[\theta, 1]$. □

Theorem 2.8. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, \mathrm{i}\theta)$, *such that* α, β *are positive numbers and* $0 < (\alpha\beta)^{1/2} \leqslant \theta < \gamma$. *Then* $V^2(A) = \sigma(A) \cup \mathrm{i}[\alpha\beta/\gamma, \alpha\beta/\theta]$.

Proof. Since $0 < (\alpha\beta)^{1/2} \leqslant \theta < \gamma$, it follows from Theorem 2.4 that $\alpha\beta/\gamma$ and $\alpha\beta/\theta$ belong to $V^2(A)$. Define $\widetilde{A} = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\alpha\beta/\gamma, \mathrm{i}\alpha\beta/\theta)$. By Lemma 1.2, $V^2(\widetilde{A}) \subseteq V^2(A)$. Applying Theorem 2.7 to the matrix $\widetilde{A}$ yields $\sigma(\widetilde{A}) \cup \mathrm{i}[\alpha\beta/\gamma, \alpha\beta/\theta] = V^2(\widetilde{A}) \subseteq V^2(A)$. (Of course the other points of $\sigma(A)$ are also in $V^2(A)$.) Applying the key lemma to the matrix A with respect to the perpendicular lines made by the altitude from the vertex $-\beta$ of the triangle $\{\alpha, -\beta, \mathrm{i}\theta\}$ and the opposite line $\overline{\{\alpha, \mathrm{i}\theta\}}$, implies that the segment $\mathrm{i}(\alpha\beta/\theta, \theta)$ is excluded from $V^2(A)$. Finally, applying the key lemma to the matrix A with respect to the four quadrants made by the line $\overline{\{-\beta, \mathrm{i}\theta\}}$ and a line perpendicular to it passing through $\mathrm{i}\gamma$, implies that the segment $\mathrm{i}(\theta, \gamma)$ is excluded from $V^2(A)$. Thus $V^2(A) = \sigma(A) \cup \mathrm{i}[\alpha\beta/\gamma, \alpha\beta/\theta]$. $\square$

In the same manner as in the proof of Theorems 2.7 and 2.8, one proves the following.

Theorem 2.9. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, \mathrm{i}\theta)$, *such that* α, β *are positive numbers and* $0 \leqslant \theta \leqslant (\alpha\beta)^{1/2} \leqslant \gamma$. *Then* $V^2(A) = \sigma(A) \cup \mathrm{i}[\alpha\beta/\gamma, \theta]$.

A set of four points, one of which is the orthocenter of the other three, is called an *orthocentric system*. In such a set, *each* of the points is the orthocenter of the other three.

Theorem 2.10. *Let* $A = \mathrm{diag}(\alpha, -\beta, \mathrm{i}\gamma, \mathrm{i}\theta)$, *such that* α, β, γ *and* θ *are distinct positive numbers. Then* $\mathrm{num}(A) = 2$ *if and only if* $\sigma(A)$ *is an orthocentric system.*

Proof. If $\sigma(A)$ is an orthocentric system, then by Theorem 2.4, $\sigma(A) = V^2(A)$ and hence $\mathrm{num}(A) = 2$. For the converse, assume without loss of generality that $0 < \theta < \gamma$. If the angle at $\mathrm{i}\gamma$ in the triangle $\{\alpha, -\beta, \mathrm{i}\gamma\}$ is an obtuse or right angle, then by Theorem 2.7, $\mathrm{i}[\theta, \gamma] \subseteq V^2(A)$, contradicting $\mathrm{num}(A) = 2$. Thus, the angle $\mathrm{i}\gamma$ is an acute angle and, by Theorem 2.4, $\hat{\gamma} = \alpha\beta/\theta \in V^2(A)$. Now $\hat{\gamma} < \gamma$, so it must be $\hat{\gamma} = \theta$. Therefore, $\sigma(A)$ is an orthocentric system. $\square$

In the following we consider matrices of the form (6) whose spectra have more than four points.

Theorem 2.11. *Let* H *be a non-singular Hermitian* $n \times n$ *matrix. Define* $A = H \oplus [\mathrm{i}\theta]$, *where* θ *is a non-zero real number. Then* $\mathrm{num}(A) = 2$ *if and only if none of the triangles generated by* $\sigma(A)$ *is an acute-angle triangle.*

Proof. Let $\mathrm{num}(A) = 2$. If one of the triangles generated by $\sigma(A)$ is an acute-angle triangle, then its orthocenter belongs to $V^2(A)$, and hence $\mathrm{num}(A) > 2$; a contradiction. For the converse, assume without loss of generality that $A = \mathrm{diag}(\alpha_1, \ldots, \alpha_n, -\beta_1, \ldots, -\beta_m, \mathrm{i}\theta)$, where $0 < \alpha_1 \leqslant \alpha_2 \leqslant \cdots \leqslant \alpha_n$, and $0 < \beta_1 \leqslant \beta_2 \leqslant \cdots \leqslant \beta_m$

and $\theta > 0$. By Theorem 2.2, $V^2(A) \subseteq \sigma(A) \cup \mathrm{i}[0, \theta]$. Since $\alpha_1\beta_1 \geqslant \theta^2$, apply the key lemma to the two lines $\overline{\{\alpha_1, \mathrm{i}\theta\}}$ and the perpendicular to it passing through $-\beta_1$; it follows that $V^2(A) = \sigma(A)$. $\square$

We know that $\mathrm{num}(A) = 1$ if and only if A is a scalar matrix. In the following corollary, we summarize some of the results in this section.

Corollary 2.12. *Let A be a non-scalar matrix of the form* (6). *Then* $\mathrm{num}(A) = 2$ *if and only if either $\sigma(A)$ is an orthocentric system or no triangle of points of $\sigma(A)$ is an acute-angle triangle.*

3. Rectangular hyperbolas

In this section we study the relationship between rectangular hyperbolas and polynomial numerical hulls of order 2 for normal matrices. A hyperbola for which the asymptotes are perpendicular is called a *rectangular hyperbola*. Let us state some properties of a rectangular hyperbola needed in this section. If the three vertices of a triangle lie on a rectangular hyperbola, then so does the orthocenter. If four non-colinear points do not form an orthocentric system, then there is a unique rectangular hyperbola, perhaps degenerate, passing through them [1,5].

In the previous sections, we considered the degenerate case of a rectangular hyperbola; i.e., two perpendicular lines. In this section we study the case that $\sigma(A)$ lies on a general rectangular hyperbola.

Theorem 3.1. *Let $A \in M_n(\mathbb{C})$ be a normal matrix such that $\sigma(A) \subseteq \mathscr{R}$, where $\mathscr{R}$ is a non-degenerate rectangular hyperbola. Then $V^2(A) \subseteq \mathscr{R}$.*

Proof. Without loss of generality we assume that $A = \mathrm{diag}(a_1, a_2, \ldots, a_n)$ and $\mathscr{R} = \{x + \mathrm{i}y : x^2 - y^2 = 1\}$. Let $z = x + \mathrm{i}y \in V^2(A)$. Since $z \in V^2(A) \subseteq W(A)$, there exists a unit vector $X = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ such that $z = X^*AX$ and $z^2 = X^*A^2X = \sum_{i=1}^n |x_i|^2 a_i^2$. We know that $\sigma(A) \subseteq \mathscr{R}$, so $\Re(a_i^2) = 1$ for $(i = 1, 2, \ldots, n)$. Now, $\Re(z^2) = \Re(\sum_{i=1}^n |x_i|^2 a_i^2) = \sum_{i=1}^n |x_i|^2 \Re(a_i^2) = \sum_{i=1}^n |x_i|^2 = 1$, which implies that $V^2(A) \subseteq \mathscr{R}$. $\square$

Theorem 3.2. *Let A be a normal matrix such that $\sigma(A) \subseteq \mathscr{R}$, where $\mathscr{R}$ is any rectangular hyperbola. Then* $\mathrm{num}(A) \leqslant 4$.

Proof. Without loss of generality we assume that $\mathscr{R} = \{x + \mathrm{i}y : x^2 - y^2 = 1\}$. Since $\sigma(A) \subseteq \mathscr{R}$, it follows that if $z \in \sigma(A)$, then $\Re(z^2) = 1$. Thus $\Re(\sigma(A^2)) = 1$. Therefore, $W(A^2)$ is a line segment and hence $\mathrm{num}(A^2) \leqslant 2$. Let $X^*AX \in V^4(A)$ for unit vector X. Then $(X^*A^2X)^2 = (X^*AX)^4 = X^*A^4X$. Since $\mathrm{num}(A^2) \leqslant 2$, $X^*A^2X \in \sigma(A^2)$. But $X^*A^2X = (X^*AX)^2 \in \sigma(A^2)$. It follows that either X^*AX

or its negative is in $\sigma(A)$ (by symmetry, of course, this means both are). Thus $V^4(A) = \sigma(A)$. $\square$

Now we need some definitions: An *inner cross* is a rectangular hyperbola together with the points between its two branches. An *outer cross* is a rectangular hyperbola together with the points not between its two branches. A *cross* is a degenerate rectangular hyperbola together with two opposite quadrants bounded by it. Note that a cross is a degenerate case of an inner cross and of an outer cross. The following is, we will see, a counterpart to the key lemma of the last section.

Lemma 3.3. *Let A be a normal matrix and let Δ be an inner or outer cross. If $\sigma(A) \subseteq \Delta$, then $V^2(A) \subseteq \Delta$.*

Proof. First, we assume that Δ is an inner cross. Assume without loss of generality that $A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$. and $\Delta = \{z : \Re(z^2) \leqslant 1\}$. Let $z = x + \mathrm{i}y \in V^2(A)$. Since $z \in V^2(A) \subseteq W(A)$, there exists a unit vector $X = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ such that $z = X^*AX$ and $z^2 = X^*A^2X = \sum_{i=1}^n |x_i|^2 a_i^2$. We know that $\sigma(A) \subseteq \{z : \Re(z^2) \leqslant 1\}$. Then $\Re(a_i^2) \leqslant 1$ for $(i = 1, 2, \ldots, n)$. Since $\Re(z^2) = \Re(\sum_{i=1}^n |x_i|^2 a_i^2) = \sum_{i=1}^n |x_i|^2 \Re(a_i^2) \leqslant \sum_{i=1}^n |x_i|^2 = 1$, it follows that $V^2(A) \subseteq \Delta$. The proof for an outer cross follows in a similar manner. $\square$

Remark 2. The key lemma is just the degenerate case of Lemma 3.3, whose statement remains true for all degenerate or non-degenerate crosses Δ.

In the following we determine $V^2(A)$, where A is a normal matrix whose spectrum consists of the four vertices of a rectangle.

Theorem 3.4. *Let* $A = \operatorname{diag}(1 + \mathrm{i}\alpha, 1 - \mathrm{i}\alpha, -1 + \mathrm{i}\alpha, -1 - \mathrm{i}\alpha)$, *where* $0 < \alpha \leqslant 1$. *Then* $V^2(A) = \{z \in W(A) : \Re(z^2) = 1 - \alpha^2\}$.

Proof. Let $\mathscr{R}$ be the rectangular hyperbola $\{z = x + \mathrm{i}y : x^2 - y^2 = 1 - \alpha^2\}$. Since $\sigma(A) \subseteq \mathscr{R}$, it follows from Theorem 3.1 that $V^2(A) \subseteq \mathscr{R}$. Then $V^2(A) \subseteq \mathscr{R} \cap W(A)$. For the converse, let $z = x + \mathrm{i}y \in \mathscr{R} \cap W(A)$, then $\Re(z^2) = x^2 - y^2 = 1 - \alpha^2$. Let us define a unit vector $X = (x_1, x_2, x_3, x_4)^{\mathrm{T}}$ such that $|x_1|^2 = (1 + x)(1 + s)/4$, $|x_2|^2 = (1 + x)(1 - s)/4$, $|x_3|^2 = (1 - x)(1 + s)/4$ and $|x_4|^2 = (1 - x)(1 - s)/4$, where $s = y/\alpha$. It is easy to check that $z = X^*AX$ and $(X^*AX)^2 = X^*A^2X$, hence $z \in V^2(A)$. We have proved that $\mathscr{R} \cap W(A) \subseteq V^2(A)$. $\square$

Let D_n be an $n \times n$ diagonal matrix of the form

$$D_n = \operatorname{diag}(1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1}), \tag{8}$$

where ω_n is a primitive nth root of unity (i.e., $\omega_n^n = 1$ and $\omega_n^k \neq 1$ for $k = 1, 2, \ldots, n - 1$). The nth roots of unity form the vertices of a regular n-sided polygon inscribed in the unit circle of radius 1. We have seen that $V^2(D_1) = \{1\}$, $V^2(D_2) = \{1, -1\}$, $V^2(D_3) = \{0, 1, e^{i2\pi/3}, e^{i4\pi}/3\}$ and $V^2(D_4) = [-1, 1] \cup i[-1, 1]$. Continuing the methods of this section, we have results for higher n. The next theorem will give a full description of $V^2(D_5)$.

Remark 3. Let A be any normal matrix such that $\sigma(A)$ has at most four points. Then $\sigma(A)$, and hence by Theorem 3.1 also $V^2(A)$, are subsets of a rectangular hyperbola. Thus $m_2(V^2(A)) = 0$, where m_2 is 2-dimensional Lebesgue measure. Therefore, we need at least five points in the spectrum of A such that $m_2(V^2(A)) > 0$. The next theorem will confirm that five points suffice.

First we state a lemma of wide applicability.

Lemma 3.5. *Let A be an arbitrary matrix. Let D be a bounded domain in $\mathbb{C}$ and ∂D be its boundary. If $\partial D \subseteq V^k(A)$, then $D \subseteq V^k(A)$.*

Proof. Let $z \in D$ and p be any polynomial of degree less than or equal k. Because p is analytic, by the Maximum Modulus Principle there exists $w \in \partial D$ such that $|p(z)| \leqslant |p(w)|$. Since $\partial D \subseteq V^k(A)$, then $|p(w)| \leqslant \| p(A) \|$. Thus $z \in V^k(A)$. □

Theorem 3.6. *Let D_5 be as in* (8) *and let m_2 be* 2*-dimensional Lebesgue measure on $\mathbb{C}$ (identified as $\mathbb{R}^2$). Then $m_2(V^2(D_5)) > 0$. Moreover $V^2(D_5)$ is the intersection of five inner crosses.*

Proof. Let $a_j = \omega_5^j$, $j = 0, \ldots, 4$. Define $b_0 := \mathrm{orth}(a_0a_2a_3)$, $b_1 := \mathrm{orth}(a_1a_3a_4)$, $b_2 := \mathrm{orth}(a_2a_4a_0)$, $b_3 := \mathrm{orth}(a_3a_0a_1)$ and $b_4 := \mathrm{orth}(a_4a_1a_2)$, where $\mathrm{orth}(abc)$ means the orthocenter of the triangle with vertices a, b and c (in $\mathbb{C}$). By Theorem 2.4, $\{b_0, \ldots, b_4\} \subseteq V^2(D_5)$. Now consider the five rectangles $q_0 = \{a_0a_1b_4b_2\}$, $q_1 = \{a_1a_2b_0b_3\}$, $q_2 = \{a_2a_3b_1b_4\}$, $q_3 = \{a_3a_4b_2b_0\}$ and $q_4 = \{a_4a_0b_3b_1\}$. Let Q_i be a 4×4 diagonal matrix whose diagonal entries are the vertices of q_i and let $\mathscr{R}_i$ be the (unique) rectangular hyperbola passing through $\sigma(Q_i)$, $(i = 0, \ldots, 4)$. Applying Theorem 3.4, it follows that $\mathscr{R}_i \cap W(Q_i) = V^2(Q_i) \subseteq V^2(D_5)$. Also let $\varDelta_i$ be the inner cross of the rectangular hyperbola $\mathscr{R}_i$, and let $\varDelta = \cap_{i=0}^4 \varDelta_i$. Since $\sigma(D_5) \subseteq \varDelta_i$, it follows from Lemma 3.3 that $V^2(D_5) \subseteq \varDelta_i$ $(i = 0, \ldots, 4)$. Thus $V^2(D_5) \subseteq \varDelta$. Since $\partial\varDelta \subseteq \cup_{i=0}^4 (\mathscr{R}_i \cap W(Q_i)) \subseteq V^2(D_5)$ and $\varDelta$ is a closed domain in the complex plane, it follows by Lemma 3.5 that $\varDelta \subseteq V^2(D_5)$. Thus, $V^2(D_5) = \varDelta$. It is easy to check that $\varDelta$ contains the pentagon $\{b_0b_1b_2b_3b_4\}$. Surely, then, $m_2(V^2(D_5)) > 0$. □

Acknowledgement

We would like to thank Professor Mehdi Radjabalipour for helpful discussions and the referee for useful comments.

References

[1] H.S.M. Coxeter, S.L. Greitzer, The Nine-point Circle; §1.8 in Geometry Revisited, Random House, New York, 1967, pp. 20–22.
[2] O. Nevalinna, Convergence of Iterations for Linear Equation, Birkhäuser, Basel, 1993.
[3] O. Nevalinna, Hessenberg matrices in Krylov subspaces and the computation of the spectrum, Numer. Funct. Anal. Optimiz. 16 (3,4) (1995) 443–447.
[4] A. Greenbaum, Generalizations of the field of values useful in the study of polynomial functions of matrix, Linear Algebra Appl. 347 (2002) 233–249.
[5] D. Wells, The Penguin Dictionary of Curious and Interesting Geometry, Penguin, London, 1991.