Polynomial Numerical Hulls of Matrices

Chandler Davis,[*] Chi-Kwong Li[†] and Abbas Salemi[‡]

August 21, 2007

Abstract. For any n-by-n complex matrix A, we use the joint numerical range $W(A, A^2, \ldots, A^k)$ to study the polynomial numerical hull of order k of A, denoted by $V^k(A)$. We give an analytic description of $V^2(A)$ when A is normal. The result is then used to characterize those normal matrices A satisfying $V^2(A) = \sigma(A)$, and to show that a unitary matrix A satisfies $V^2(A) = \sigma(A)$ if and only if its eigenvalues lie in a semicircle, where $\sigma(A)$ denotes the spectrum of A. When $A = \operatorname{diag}(1, w, \ldots, w^{n-1})$ with $w = e^{i2\pi/n}$, we determine $V^k(A)$ for $k \in \{2\} \cup \{j \in \mathbb{N} : j \geq n/2\}$. We also consider matrices $A \in M_n$ such that A^2 is Hermitian. For such matrices we show that $V^4(A)$ is the spectrum of A, and give a description of the set $V^2(A)$.

Key words Polynomial numerical hull, joint numerical range, normal matrix.
AMS Subject Classification 15A60, 15A18.

1 Introduction

Let M_n be the set of $n \times n$ complex matrices. Motivated by the study of convergence of iterative methods in solving linear systems (e.g., see [2, 3, 4]), researchers studied the *polynomial numerical hull of order* k of a matrix $A \in M_n$, which is defined and denoted by

$$V^k(A) = \{\xi \in \mathbb{C} : |p(\xi)| \leq \|p(A)\| \text{ for all } p(z) \in \mathbf{P}_k[\mathbb{C}]\},$$

where $\mathbf{P}_k[\mathbb{C}]$ is the set of complex polynomials with degree at most k. The *joint numerical range* of $(A_1, A_2, \ldots, A_m) \in M_n \times \cdots \times M_n$ is denoted by

$$W(A_1, A_2, \ldots, A_m) = \{(x^*A_1x, x^*A_2x, \ldots, x^*A_mx) : x \in \mathbb{C}^n, x^*x = 1\}.$$

By the result in [2] (see also [3])

$$V^k(A) = \{\zeta \in \mathbb{C} : (0, \ldots, 0) \in \operatorname{conv} W((A - \zeta I), (A - \zeta I)^2, \ldots, (A - \zeta I)^k)\},$$

where $\operatorname{conv} X$ denotes the convex hull of $X \subseteq \mathbb{C}^k$.

In this paper, we use the joint numerical range $W(A, A^2, \ldots, A^k)$ to study $V^k(A)$ for $A \in M_n$. Denote by $\sigma(A)$ the spectrum of $A \in M_n$. In Section 2, we give an analytic description of $V^2(A)$ when $A \in M_n$ is normal. The result is then used to characterize those normal matrices A satisfying

[*]Department of Mathematics, University of Toronto, Ontario, Canada M5S 2E4 (davis@math.toronto.edu.)
[†]Department of Mathematics, College of William & Mary, Williamsburg, VA 23185, USA (ckli@math.wm.edu.) Li is an honorary professor of the University of Hong Kong.
[‡]Department of Mathematics, Shahid Bahonar University of Kerman, Kerman, Iran (salemi@mail.uk.ac.ir.)

$V^2(A) = \sigma(A)$, and to show that a unitary matrix A satisfies $V^2(A) = \sigma(A)$ if and only if its eigenvalues lie in a semicircle. When $A = \text{diag}(1, w, \ldots, w^{n-1})$ with $w = e^{i2\pi/n}$, we determine $V^k(A)$ for $k \in \{2\} \cup \{j \in \mathbb{N} : j \geq n/2\}$ in Section 3. Section 4 concerns those matrices $A \in M_n$ such that A^2 is Hermitian. For such matrices we show that $V^4(A) = \sigma(A)$, and give a description of the set $V^2(A)$. Additional results and remarks are given in Section 5.

Below we state some properties of the polynomial numerical hull of $A \in M_n$; one may see [1, 3] for details.

1) $\sigma(A) \subseteq V^{k+1}(A) \subseteq V^k(A) \subseteq V^1(A) = W(A)$, for all $k \geq 1$.

2) If m is the degree of the minimal polynomial of A, then $V^k(A) = \sigma(A)$ for all $k \geq m$.

3) $V^k(\alpha A + \beta I) = \alpha V^k(A) + \beta$ for all α and β in the complex plane $\mathbb{C}$.

4) Let $A = A^*$. Then $V^2(A) = \sigma(A)$.

5) Let A be a normal matrix. Then $\partial(W(A)) \cap V^2(A) \subseteq \sigma(A)$, where $\partial(D)$ means the boundary of D.

6) If A is normal or an upper triangular Toeplitz matrix, then $W(A, \ldots, A^k)$ is convex, and hence

$$\begin{aligned} V^k(A) &= \{\zeta \in \mathbb{C} : (\zeta, \ldots, \zeta^k) \in W(A, \ldots, A^k)\} \\ &= \{x^*Ax : x \in \mathbb{C}^n, x^*x = 1, \text{ and } (x^*Ax)^j = x^*A^jx, j = 1, 2, \ldots, k\}. \end{aligned}$$

In fact, it can be shown that $(0, \ldots, 0) \in \text{conv}\, W((A - \zeta I), \ldots, (A - \zeta I)^k)$ if and only if $(\zeta, \ldots, \zeta^k) \in \text{conv}\, W(A, \ldots, A^k)$; see Theorem 5.3.

2 Polynomial numerical hull of order two for normal matrices

In the following, we will develop a scheme to give an analytic description of $V^2(A)$ for a normal matrix $A = H + iG = \text{diag}(a_1, \ldots, a_n)$, where H and G are Hermitian. By (6) in Section 1, $\mu = x + iy \in V^2(A)$ if and only if $(\mu, \mu^2) \in W(A, A^2)$, equivalently,

$$(x, y, x^2 - y^2, 2xy) \in W(H, G, H^2 - G^2, HG + GH) \subset \mathbb{R}^4.$$

By [1, Theorem 3.2], if $a_1, \ldots, a_n$ lie in a rectangular hyperbola, then so does $V^2(A)$. However, exactly which part of the hyperbola belongs to $V^2(A)$ was not determined. The following result addresses this problem.

Theorem 2.1 *Let $A = H + iG$ with $H^* = H = \text{diag}(h_1, \ldots, h_n)$ and $G^* = G = \text{diag}(g_1, \ldots, g_n)$ be such that*

$$\{(h_j, g_j) : 1 \leq j \leq n\} \subseteq R = \{(x, y) : r_1(x^2 - y^2) + r_2 xy = r_3 x + r_4 y + r_5\},$$

where $r_1 r_2 \neq 0$. Then $(x, y) \in V^2(A)$ if and only if $(x, y) \in R$ and one or both of the following holds:

(a) $(x, y, x^2 - y^2) \in W(H, G, H^2 - G^2)$ *if* $r_1 \neq 0$.

(b) $(x, y, xy) \in W(H, G, HG)$ *if* $r_2 \neq 0$.

Proof. The necessity follows from Theorem 3.2 in [1] and (6) in section 1. For the converse, by the convexity of $W(H, G, H^2 - G^2)$, there exist $t_1, \ldots, t_n \geq 0$ with $t_1 + \cdots + t_n = 1$ such that

$$(x, y, x^2 - y^2) = \sum_{j=1}^{n} t_j (h_j, g_j, h_j^2 - g_j^2) \in W(H, G, H^2 - G^2).$$

Then $(x, y) \in R$ implies that

$$r_2 xy = r_3 x + r_4 y + r_5 + r_1(y^2 - x^2) = \sum_{j=1}^{n} t_j (r_3 h_j + r_4 g_j + r_5 + r_1(g_j^2 - h_j^2)) = \sum_{j=1}^{n} r_2 t_j h_j g_j.$$

Thus, $(x, y, x^2 - y^2, xy) \in W(H, G, H^2 - G^2, GH)$. The result follows.

The case for $(x, y, xy) \in W(H, G, GH)$ can be proved in a similar way. □

If $n = 2$ then $V^2(A) = \sigma(A)$. If $n = 3$ then $V^2(A) = \sigma(A)$ or $V^2(A) = \sigma(A) \cup \{\mu\}$ if the orthocenter μ of the triangle with vertices eigenvalues a_1, a_2, a_3 of A lies in $W(A)$; see for example [1, Theorem 2.4].

Suppose $A \in M_4$ is normal. If there are $\mu, \nu \in \mathbb{C}$ with $\mu \neq 0$ such that the eigenvalues of $\mu A + \nu I$ lie in $\mathbb{R} \cup i\mathbb{R}$, then one can apply the results in [1, Section 2] (see also Theorem 4.4) to determine $V^2(A)$. If it is not the case, then Theorem 2.2 below gives a complete description of $V^2(A)$. In particular, the result shows that one can reduce the problem to the special case where $A = \operatorname{diag}(-1, 1, \mu, \nu)$, so that the intersection of the open intervals $(-1, 1)$ and $(\operatorname{Re}(\mu), \operatorname{Re}(\nu))$ will determine the set $V^2(A)$ readily. As we will see, Theorem 2.2 is the key result allowing us to give an analytic description for $V^2(N)$ for a normal matrix $N \in M_n$ for any $n \in \mathbb{N}$.

Theorem 2.2 *Let $A = \operatorname{diag}(a_1, \ldots, a_4)$ be such that $a_1, \ldots, a_4$ are not contained in two perpendicular lines. Suppose $R \subseteq \mathbb{C} \equiv \mathbb{R}^2$ is the rectangular hyperbola uniquely determined by a_1, a_2, a_3, a_4 and is the union of the two branches R_1 and R_2. Then $V^2(A) \subseteq R$, and $V^2(A)$ can be determined as follows.*

(a) *Suppose each branch of R contains two of the eigenvalues, say, $a_1, a_2 \in R_1$ and $a_3, a_4 \in R_2$. Let $(u_1, v_1) = (2, a_1 + a_2)/(a_1 - a_2)$ and $u_1 A - v_1 I = \operatorname{diag}(1, -1, x_3 + iy_3, x_4 + iy_4)$. Then $z \in R_1$ belongs to $V^2(A)$ if and only if $z = a_1, a_2$ or $u_1 z - v_1 = x + iy$ with $x \in (-1, 1)$ such that x lies between x_3 and x_4. Let $(u_2, v_2) = (2, a_3 + a_4)/(a_3 - a_4)$ and $u_2 A - v_2 I = \operatorname{diag}(x_1 + iy_1, x_2 + iy_2, 1, -1)$. Then $z \in R_2$ belongs to $V^2(A)$ if and only if $z = a_3, a_4$ or $u_2 z - v_2 = x + iy$ with $x \in (-1, 1)$ such that x lies between x_1 and x_2.*

(b) *Suppose one of the branches of R contains three of the eigenvalues, say, $a_1, a_2, a_3 \in R_1$ with a_3 lying between a_1 and a_2. Let $(u, v) = (2, a_1 + a_2)/(a_1 - a_2)$ and $uA - vI = \operatorname{diag}(1, -1, x_3 + iy_3, x_4 + iy_4)$. Then $z \in R_1$ belongs to $V^2(A)$ if and only if $z = a_1, a_2, a_3$ or $uz - v = x + iy$ with $x \in (-1, 1)$ such that x lies between x_3 and x_4.*

(c) *Suppose one of the branches of R contains four eigenvalues a_1, a_2, a_3, a_4. Then $V^2(A) = \sigma(A)$.*

Proof. By Theorem 3.1 in [1], $V^2(A) \subseteq R \cap W(A)$.

(a) Since $V^2(uA - vI) = uV^2(A) - v$, we may replace A by $uA - vI$ and assume that $A = \mathrm{diag}\,(1, -1, x_3 + iy_3, x_4 + iy_4)$. Since $\overline{V^2(A)} = V^2(A^*)$, we may replace A by A^* if necessary, and assume that $y_3, y_4 > 0$. Furthermore, we may assume that $x_3 < x_4$. Otherwise, replace A by $-A$ and relabel the third and fourth eigenvalues. Then the rectangular hyperbola passing through $(1, 0), (-1, 0), (x_3, y_3), (x_4, y_4)$ satisfies a formula of the form

$$y^2 + (ax + c)y + (1 - x^2) = 0. \tag{1}$$

Let $A = H + iG$ and consider the joint numerical range $W(H, G, H^2 - G^2)$, which is the convex hull of the points

$$(-1, 0, 1), (1, 0, 1), (x_3, y_3, x_3^2 - y_3^2), (x_4, y_4, x_4^2 - y_4^2).$$

Since $(x_3, y_3), (x_4, y_4)$ satisfy (1), we see that

$$x_j^2 - y_j^2 = (ax_j + c)y_j + 1 \qquad \text{for } j = 3, 4.$$

So, the plane passing through the point $(-1, 0, 1), (1, 0, 1), (x_j, y_j, x_j^2 - y_j^2)$, make an angle $\theta_j \in [-\pi/2, \pi/2]$ with the plane $\mathbf{P} = \{(x, y, 1) : x, y \in \mathbb{R}\}$, where

$$\tan\theta_j = (x_j^2 - y_j^2 - 1)/(y_j - 0) = [(ax_j + c)y_j + 1 - 1]y_j = ax_j + c, \qquad j = 3, 4.$$

Now, for any point $(x, y) \in R_1$, the line joining $(x, y, x^2 - y^2) \in R_1$ and the point $(x, 0, 1)$ will make an angle θ with the plane $\mathbf{P}$ such that

$$\tan\theta = ((x^2 - y^2) - 1)/(y - 0) = [(ax + c)y + 1 - 1]/y = ax + c.$$

Thus, the values $(x, y, x^2 - y^2)$ lie between the two triangular laminas

$$T_j = \mathrm{conv}\,\{(-1, 0, 1), (1, 0, 1), (x_j, y_j, x_j^2 - y_j^2)\}, \qquad j = 3, 4,$$

if and only if $\tan\theta$ lies between $\tan\theta_3$ and $\tan\theta_4$, equivalently, $x \in [x_3, x_4]$. Note that for $(x, y) \in R_1$, the point $(x, y, x^2 - y^2)$ lies between the two triangular laminas T_3 and T_4 if and only if $(x, y, x^2 - y^2) \in W(H, G, H^2 - G^2)$. By Theorem 2.1, we see that $(x, y) \in V^2(A) \cap R_1$ if and only if $x \in [x_3, x_4]$.

The proof for the other branch in (a) and the proof for (b) can be done similarly.

(c) If a_1, a_2, a_3, a_4 belongs to a branch of R, then $W(A) = \mathrm{conv}\,\{a_1, a_2, a_3, a_4\}$ only intersect R at a_1, a_2, a_3, a_4. So $\sigma(A) \subseteq V^2(A) \subseteq R \cap W(A) = \sigma(A)$. □

By the above theorem and the results in [1, Section 2] (see also Theorem 4.4), we have the following.

Corollary 2.3 *Let $B = \mathrm{diag}\,(a_1, a_2, a_3, a_4)$ be such that $a_1, \ldots, a_4 \in \mathbb{C}$ are distinct. Then the following conditions are equivalent.*

(a) $V^2(B) = \sigma(B)$

(b) $V^2(B)$ *is finite.*

(c) *One of the following holds.*

(c.1) $a_1, \ldots, a_4$ *are contained in a straight line.*

(c.2) *One of the points* $a_1, \ldots, a_4$ *is the orthocenter of the triangle with the other three points as vertices.*

(c.3) *There are* $\mu, \nu \in \mathbb{C}$ *with* $\mu \neq 0$ *such that* $\{\mu a_j + \nu : 1 \leq j \leq 4\} = \{b_1, b_2, b_3, i\}$ *satisfying* $\{b_1, b_2, b_3\} \subseteq [0, \infty)$, $\{b_1, b_2, b_3\} \subseteq (-\infty, 0]$, *or* $\{b_1, b_2, b_3\} \subseteq \mathbb{R} \setminus \{0\}$ *with the property that* $\operatorname{conv}\{b_p, b_q, i\}$ *is not an acute angle triangle for any* $p, q \in \{1, 2, 3\}$.

(c.4) $\mathbf{Q} = \operatorname{conv}\{a_1, a_2, a_3, a_4\}$ *is a quadrangle such that* $\operatorname{conv}\{\mu, a_p, a_q\}$ *is not an acute angle triangle for any* $p, q \in \{1, 2, 3, 4\}$, *where* μ *is the intersection of the diagonals of* $\mathbf{Q}$.

Proof. We consider three cases.

Case 1. Suppose $\sigma(B)$ is a subset of the straight line. Then $V^2(B) = \sigma(B)$.

Case 2. Suppose $\sigma(B)$ is a subset of two perpendicular lines. By the results in [1, Section 2] (see also Theorem 4.4 and the two examples following it), either

(i) $V^2(B) \neq \sigma(B)$ and $V^2(B)$ contains a nontrivial line segment, or

(ii) $V^2(B) = \sigma(B)$ so that (c.2) or (c.3) holds.

Case 3. Suppose $\sigma(B)$ is not a subset of two perpendicular lines. Then a_1, a_2, a_3, a_4 determine a unique rectangular hyperbola R not equal to a pair of perpendicular lines, and one of the conditions (a) – (c) of Theorem 2.2 holds.

Suppose Theorem 2.2 (a) holds. We can assume that a_1, a_2 lie in one branch of R and a_3, a_4 lie in anther branch. Following the arguments in the proof of Theorem 2.2, we see that $V^2(B) \neq \sigma(B)$ if and only if $(-1, 1) \cap (x_3, x_4) \neq \emptyset$ and $(-1, 1) \cap (x_1, x_2) \neq \emptyset$. One can check that these conditions are equivalent to the existence of a non-degenerate acute angle triangle of the form $\operatorname{conv}\{\mu, a_p, a_q\}$, where μ is the intersection of the diagonals of $\mathbf{Q}$ and $p, q \in \{1, 2, 3, 4\}$. Thus, either

(i) $V^2(B) \neq \sigma(B)$ and $V^2(B)$ contains a nontrivial segment of R, or

(ii) $V^2(B) = \sigma(B)$ and condition (c.4) holds.

Suppose Theorem 2.2 (b) holds, say, a_1, a_2, a_3 lie in one branch of R so that a_1 and a_2 are the end points of the segment of the curve. Following the arguments in the proof of Theorem 2.2, we see that $V^2(B) \neq \sigma(B)$ if and only if $(-1, 1) \cap (x_3, x_4) \neq \emptyset$. Thus,

(i) $V^2(B)$ contains a nontrivial segment of R, unless

(ii) a_3 is the orthocenter of the triangle $\operatorname{conv}\{a_1, a_2, a_4\}$.

However, if (ii) holds, then $\sigma(B)$ will lie in the union of two perpendicular line, which is a contradiction. So, (i) must hold in this case.

If Theorem 2.2 (c) holds, then $V^2(B) = \sigma(B)$.

Combining the analysis in Cases 1–3, we see that $V^2(B) \neq \sigma(B)$ if and only if $V^2(B)$ is infinite. Moreover, $V^2(B) = \sigma(B)$ if and only if one of the conditions (c.1)–(c.4) holds. □

Remark 2.4 Consider $A = H + iG \in M_n$ with $n \geq 5$. Note that $W(H, G, H^2 - G^2, HG)$ is a polyhedron in $\mathbb{R}^4$. By elementary convex analysis, we have the following observations.

(a) Every point in $W(H, G, H^2 - G^2, HG)$ is a convex combination of at most 5 vertices.

(b) Every boundary point of $W(H, G, H^2 - G^2, HG)$ is a convex combination of at most 4 vertices.

(c) Suppose $(\mu, \mu^2) \in W(A, A^2)$ is an interior point, i.e., $(\mu+\varepsilon_1, \mu^2+\varepsilon_2) \in W(A, A^2)$ for $\varepsilon_1, \varepsilon_2 \in \mathbb{C}$ with $|\varepsilon_1|^2 + |\varepsilon_2|^2 < d$ for some $d > 0$. Then clearly μ lies in the interior of $V^2(A)$. So, if μ is a boundary point of $V^2(A)$, then (μ, μ^2) is a boundary point of $W(A, A^2)$ and is determined by 4 vertices of $W(A, A^2)$.

By observation (c) above, we have the following result giving an analytic description of $V^2(A)$ for a normal matrix $A \in M_n$ with more than four distinct eigenvalues.

Theorem 2.5 *Suppose $A \in M_n$ is a normal matrix with distinct eigenvalues $a_1, \ldots, a_m$ such that $m > 4$. Then the boundary of $V^2(A)$ is a subset of*

$$\mathbf{S} = \cup\{V^2(\mathrm{diag}\,(a_{j_1}, a_{j_2}, a_{j_3}, a_{j_4})) : 1 \leq j_1 < j_2 < j_3 < j_4 \leq m\}.$$

Consequently, $V^2(A)$ is equal to the union of the set $\mathbf{S}$ and the set of complex numbers enclosed by the closed curves in the set $\mathbf{S}$.

Proof. The first statement follows from Remark 2.4 (c). Since $V^2(A)$ is polynomially convex; see [4] and [1, Lemma 3.5], the set includes all the points inside the bounded closed regions enclosed by the boundary curves as well. □

We illustrate this theorem with the following example.

Example 2.6 Suppose $A = \mathrm{diag}\,(1+i/2, 1-i/2, -1+i/2, -1-i/2, 0)$. Then $V^2(A) = R_1 \cup R_2 \cup \{0\}$, where $R_1 \subseteq \mathbb{C} \equiv \mathbb{R}^2$ is the closed region bounded by the following:

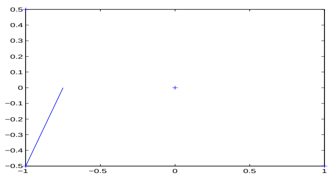
$L_1 = \{t(1, 1/2) + (1 - t)(3/4, 0) : t \in [0, 1]\}$,
$L_2 = \{t(1, -1/2) + (1 - t)(3/4, 0) : t \in [0, 1]\}$, and
$C_1 = \{(x, y) : x^2 - y^2 = 3/4, x \in [\sqrt{3}/2, 1]\}$,

and R_2 is the closed region bounded by the following:

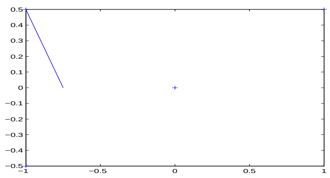
$L_3 = \{t(-1, 1/2) + (1 - t)(-3/4, 0) : t \in [0, 1]\}$,
$L_4 = \{t(-1, -1/2) + (1 - t)(-3/4, 0) : t \in [0, 1]\}$, and
$C_2 = \{(x, y) : x^2 - y^2 = 3/4, x \in [-1, -\sqrt{3}/2]\}$.

Proof. Using the four points $\{1 + i/2, 1 - i/2, -1 + i/2, -1 - i/2\}$, we get the set $C_1 \cup C_2$. The four line segments L_1, L_2, L_3, L_4 are obtained using 0 and three other nonzero points. The union of these sets cover the boundary points of $V^2(A)$. Taking the interior of those regions enclosed by closed curves, we get the set $V^2(A)$.

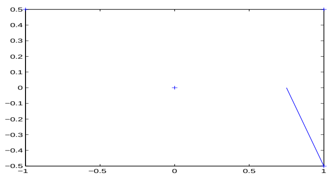
Here we depict the sets $V^2(A_1), V^2(A_2), V^2(A_3), V^2(A_4), V^2(A_5)$, where A_j is obtained from A by removing the jth row and jth column.



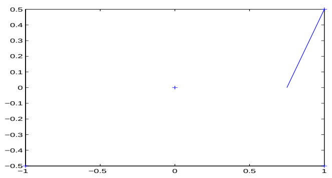
$V^2(A_1) = L_4 \cup \{0\}$



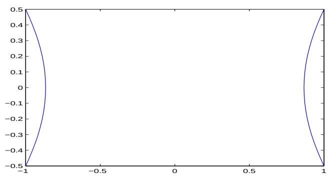
$V^2(A_2) = L_3 \cup \{0\}$



$V^2(A_3) = L_2 \cup \{0\}$

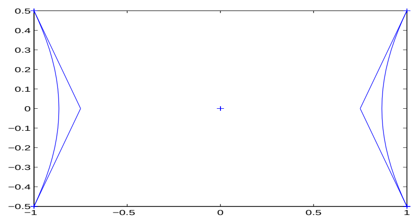


$V^2(A_4) = L_1 \cup \{0\}$

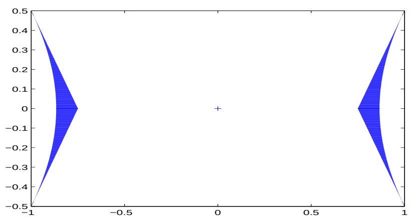


$V^2(A_5) = C_1 \cup C_2$

Taking the union of these curves, we get the boundary of $V^2(A)$. We can then fill in all the points enclosed by closed curves.



The boundary of $V^2(A)$



The set $V^2(A)$.

$\square$

It is well-known that a normal matrix $A \in M_n$ with three distinct eigenvalues a_1, a_2, a_3 satisfies $V^2(A) = \sigma(A)$ if and only if conv $\{a_1, a_2, a_3\}$ is not an acute triangle. Using Theorems 2.2 and 2.5, we can characterize those normal matrices A such that $V^2(A) = \sigma(A)$ in general. Again, the key is checking the 4-by-4 case.

Theorem 2.7 *Let $A \in M_n$ be a normal matrix with at least four distinct eigenvalues. The following conditions are equivalent.*

(a) $V^2(A) = \sigma(A)$.

(b) *The set $V^2(A)$ is finite.*

(c) *For any four distinct eigenvalues a_1, a_2, a_3, a_4 of A, one of the conditions* (c.1)–(c.4) *in Corollary* 2.3 *holds.*

Proof. The implication (a) $\Rightarrow$ (b) is clear.

To prove (b) $\Rightarrow$ (c), suppose (c) is not valid. Let $B = \text{diag}\,(a_1, a_2, a_3, a_4)$ be such that $V^2(B) \neq \sigma(B)$. By Corollary 2.3, $V^2(B)$ is an infinite. Since $V^2(B) \subseteq V^2(A)$, $V^2(A)$ is infinite as well.

Finally, we consider the implication (c) $\Rightarrow$ (a). Suppose (c) holds. For any four eigenvalues $a_1, \dots, a_4$ of A, we can assume that they are distinct. Otherwise, we can add other eigenvalues to the collection. Let $B = \operatorname{diag}(a_1, \dots, a_4)$. Then $V^2(B) = \sigma(B) \subseteq \sigma(A)$ by Corollary 2.3. By Remark 2.4 (c), the boundary of $V^2(A)$ is a subset of $\sigma(A)$. Thus, $V^2(A) = \sigma(A)$. $\square$

The next theorem characterizes those unitary matrices A satisfying $V^2(A) = \sigma(A)$.

Theorem 2.8 *Suppose $A \in M_n$ is a unitary matrix. Then $V^2(A) = \sigma(A)$ if and only if $\sigma(A)$ lies in a semi-circle (including end points).*

Proof. The result is clear if A has less than four distinct eigenvalues. So, assume it is not the case. Suppose the eigenvalues of A do not belong to a semicircle. Then there are three points in $\sigma(A)$ such that the triangle generated by them is an acute angle triangle, and its orthocenter does not belong to $\sigma(A)$. So, $V^2(A) \neq \sigma(A)$.

Conversely, suppose all the eigenvalues of A lie in a semicircle. Then for any four distinct eigenvalues a_1, a_2, a_3, a_4 of A, $\operatorname{conv}\{a_1, a_2, a_3, a_4\}$ is a quadrangle satisfying Corollary 2.3 (c.4). By Theorem 2.7, $V^2(A) = \sigma(A)$. $\square$

3 Polynomial numerical hulls of the basic circulant matrix

Let $P_n = E_{12} + \cdots + E_{n-1,n} + E_{n1}$ be the basic circulant matrix, whose powers span the algebra of circulant matrices. Then P_n is unitarily similar to

$$D_n = \operatorname{diag}(1, w, \dots, w^{n-1}), \tag{2}$$

where $w = e^{i2\pi/n}$. Then $V^k(P_n) = V^k(D_n), k = 1, \dots, n$. We begin with a characterization of $V^k(D_n)$ when $k \geq n/2$.

Theorem 3.1 *Let $D_n = \operatorname{diag}(1, w, \dots, w^{n-1})$ with $w = e^{i2\pi/n}$. If k is a positive integer such that $n/2 < k < n$, then*

$$V^k(D_n) = \sigma(D_n) \cup \{0\}.$$

If $n = 2k$ is even, then

$$V^k(D_n) = \bigcup_{j=0}^{n-1} w^j[0, 1].$$

Proof. It is easy to check that $(D_n^k)^* = D_n^{n-k}$ for all $k = 1, \dots, n$. Suppose $n/2 < k < n$ and $z \in V^k(D_n)$. Then there is a unit vector v such that $v^* D_n^j v = z^j$ for $j = 1, \dots, k$. Hence, $z^k = \bar{z}^{n-k}$. Applying absolute value on both sides, we see that $|z| = 1$ or $z = 0$. In the former case, $z^n = 1$ and hence $z \in \sigma(D_n)$. Thus, $\sigma(D_n) \subseteq V^k(D_n) \subseteq \sigma(D_n) \cup \{0\}$. Let $v = [1, \dots, 1]^t/\sqrt{n}$ be a unit vector, then $v^* D_n^j v = 0$ for all $j = 1, \dots, n-1$. Hence $V^k(D_n) = \sigma(D_n) \cup \{0\}$.

Now, suppose $n = 2k$ and $z \in V^k(D_n)$. We continue to assume that v is a unit vector such that $v^* D_n^j v = z^j$ for $j = 1, \dots, k$. Since $D_n^k = \operatorname{diag}(1, -1, 1, -1, \dots, 1, -1)$, we see that $z^k \in [-1, 1]$.

Thus, $z = re^{i\theta}$ for some $r \in [0, 1]$ and θ satisfying $e^{ik\theta} \in \mathbb{R}$. Hence, $z \in \cup_{j=0}^{n-1} w^j[0, 1]$. So, we have $V^k(D_n) \subseteq \cup_{j=0}^{n-1} w^j[0, 1]$.

We claim that $[0, 1] \subseteq V^k(D_n)$. Once this is proved, we can use the fact that D_n is unitarily similar to $w^j D_n$, $j = 1, \ldots, n-1$, to conclude that $\bigcup_{j=1}^{n-1} w^j[0, 1] \subseteq V^k(D_n)$.

To prove our claim, let $r \in [0, 1]$. The result is clear if $r = 1$. So, assume that $r < 1$. We will show that there exists a unit vector $v = [\sqrt{t_1}, \ldots, \sqrt{t_n}]^t$ with $t_1, \ldots, t_n \geq 0$ such that

$$v^* D_n^j v = r^j \qquad \text{for } j = 1, \ldots, k. \tag{3}$$

Let $F \in M_n$ be such that the (p, q) entry of F is $w^{(p-1)(q-1)}$, and let $T = [t_1, \ldots, t_n]$. Then (3) holds if and only if

$$TF = [1, r, \cdots, r^k, r_{k+2}, \ldots, r_n]$$

for some numbers $r_{k+2}, \ldots, r_n$. Denote by F_j the jth column of F. Then for $j > 1$, F_j is the conjugate of F_{n-j+2}. As a result, for $j \geq k+2$,

$$r_j = TF_j = \overline{TF}_{n-j+2} = \bar{r}^{n-j+1} = r^{n-j+1}.$$

Note that $F^{-1} = F^*/n$. To finish our proof, we need only to show that for any $r \in [0, 1)$, the vector

$$nT = [1, r, \ldots, r^{k-1}, r^k, r^{k-1}, \ldots, r]F^*$$

has nonnegative entries. Now, for $j \in \{1, \ldots, n\}$, let $\nu = \bar{w}^{j-1}$. Then

$$\begin{aligned} nt_j &= [1 \ r \ \cdots \ r^{k-1} \ r^k \ r^{k-1} \ \cdots \ r]\overline{F}_j \\ &= 1 + r\nu + \cdots + (r\nu)^{k-1} + (r\nu)^k + (r\bar{\nu})^{k-1} + (r\bar{\nu})^{k-2} + \cdots + r\bar{\nu} \\ &= \xi + \bar{\xi}, \end{aligned}$$

where

$$\begin{aligned} \xi &= [1 + r\nu + \cdots + (r\nu)^{k-1}] - (1 - (r\nu)^k)/2 \\ &= (1 - (r\nu)^k)(1 - r\nu)^{-1} - (1 - (r\nu)^k)/2 \\ &= (1 - (r\nu)^k)[(1 - r\nu)^{-1} - 1/2] \\ &= (1 - (r\nu)^k)[(1 - r\nu)^{-1} - 1/2]. \end{aligned}$$

Note that $\nu^k \in \{-1, 1\}$. Since $r \in [0, 1)$, we have $1 - (r\nu)^k > 0$, and the real part of $[(1 - r\nu)^{-1} - 1/2]$ is

$$\frac{2 - r\nu - r\bar{\nu} - |1 - r\nu|^2}{2|1 - r\nu|^2} = \frac{1 - |r\nu|^2}{2|1 - r\nu|^2} > 0.$$

So, our claim is proved. □

Next, we give an analytic description of the set $V^2(D_n)$ using the idea in the proof of [1, Theorem 2.6], which dealt with $V^2(D_5)$.

Theorem 3.2 *Let* $n > 3$ *and* $D_n = \text{diag}\,(1, w, \ldots, w^{n-1})$ *with* $w = e^{i2\pi/n}$.

(a) *Suppose* $n = 2k$ *is even. For* $j = 1, \ldots, k$, *let* $A_j = \text{diag}\,(w^j, w^{j+1}, w^{j+k}, w^{j+k+1})$. *Then* $\text{conv}\,\sigma(A_j)$ *is a rectangle, and* $V^2(A_j)$ *consists of two segments of the rectangular hyperbola passing through* $\sigma(A_j)$ *such that one of them joins* w^j *and* w^{j+1} *and the other one joins* w^{j+k} *and* w^{j+k+1}. *Moreover,* $V^2(A)$ *is the bounded region enclosed by the closed curve* $\bigcup_{j=1}^k V^2(A_j)$.

(b) *Suppose* $n = 2k+1$ *is odd. For* $j = 0, 1, \ldots, n-1$, *let* μ_j *be the orthocenter of the triangle* $\text{conv}\,\{w^j, w^{j+k}, w^{j+k+1}\}$, *and let* $B_j = \text{diag}\,(w^j, w^{j+1}, w^{j+k}, w^{j+k+1})$. *Then* $\text{conv}\,\sigma(B_j)$ *is a trapezoid, and* $V^2(B_j)$ *consists of two segments of the rectangular hyperbola passing through* $\sigma(B_j)$ *such that one of them joins* μ_{j+k+1} *and* w^{j+1} *and the other one joins* w^{j+k} *and* μ_j. *(Note that* $\text{conv}\,\{\mu_{j+k+1}, w^{j+1}, w^{j+k}, \mu_j\}$ *is a rectangle.) Moreover,* $V^2(A)$ *is the bounded region enclosed by the closed curve* $\bigcup_{j=0}^{n-1} V^2(B_j)$.

Proof. Suppose $n = 2k$. Consider the submatrices A_j for $j = 1, 2, \ldots, k$ defined as in (a). Then $\sigma(A_j)$ determine uniquely a rectangular hyperbola R_j, and $\sigma(A)$ lies in the closed region between the two branches of R_j, and so is $V^2(A)$ by [1, Lemma 3.3]. Consequently, $V^2(A)$ lies in the intersection of these regions, which is the closed bounded region with boundary $\bigcup_{j=1}^k V^2(A_j)$. By Theorem 2.5, we get the reverse inclusion, namely, the closed bounded region enclosed by the curve $\bigcup_{j=1}^k V^2(A_j)$ is a subset of $V^2(A)$.

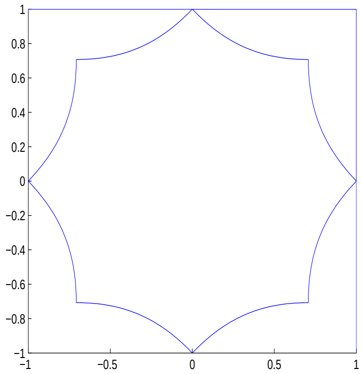
Suppose $n = 2k+1$. For $j = 0, \ldots, n-1$, consider B_j defined as in (b). By Theorem 2.2, $V^2(B_j)$ has the asserted form, and one can check that $\bigcup_{j=0}^{n-1} V^2(B_j)$ is a closed curve. Similar to the proof in case (a), one can show that for each $j = 0, \ldots, n-1$, $\sigma(B_j)$ determine uniquely a rectangular hyperbola $\hat{R}_j$, and that $V^2(A)$ lies in the closed region between the two branches of $\hat{R}_j$. Thus, $V^2(A)$ lies in the intersection of these regions, which is the closed bounded region with boundary $\bigcup_{j=0}^{n-1} V^2(B_j)$. Evidently, the closed bounded region enclosed by the curve $\bigcup_{j=1}^k V^2(B_j)$ is a subset of $V^2(A)$. The conclusion follows. □

For $3 \le k < n/2$, we do not have a complete description for $V^k(D_n)$. Nevertheless, we have the following result.

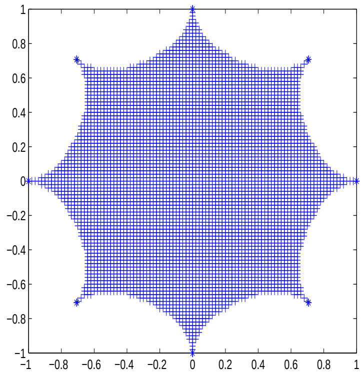
Theorem 3.3 *Let* $D_n = \text{diag}\,(1, w, \ldots, w^{n-1})$ *and* $F = (w^{(p-1)(q-1)}) \in M_n$ *with* $w = e^{i2\pi/n}$. *Suppose* $3 < k < n/2$. *Then* $\mu \in V^k(D_n)$ *if and only if there exist complex numbers* $z_{k+2}, \ldots, z_{n-k}$ *such that* $z_j = \bar{z}_{n-j+2}$ *and* $F^{-1}[1, \mu, \ldots, \mu^k, z_{k+2}, \ldots, z_{n-k}, \bar{\mu}^k, \ldots, \bar{\mu}]^t$ *is a nonnegative vector.*

Proof. Note that for any vector $v \in \mathbb{R}^n$, if $Fv = [z_1, \ldots, z_n]^t$ then $z_j = \bar{z}_{n-j+2}$ for $j = 2, \ldots, n$. Consequently, $(\mu, \ldots, \mu^k) \in W(D_n, D_n^2, \ldots, D_n^k)$ if and only if there is nonnegative vector v and complex numbers $z_{k+2}, \ldots, z_{n-k}$ such that $Fv = [1, \mu, \ldots, \mu^k, z_{k+2}, \ldots, z_{n-k}, \bar{\mu}^k, \ldots, \bar{\mu}]^t$, which is the desired conclusion. □

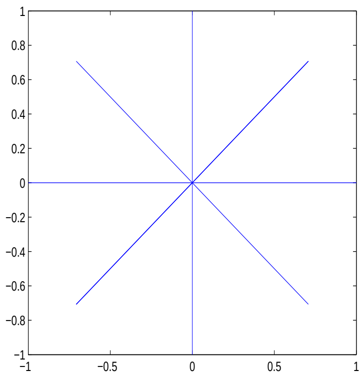
Let us depict the boundary of $V^2(D_8)$, and the sets $V^3(D_8)$ and $V^4(D_8)$. For comparison purpose, we also put the $V^2(D_8)$ and $V^3(D_8)$ in the same frame so as to illustrate that $V^3(D_8)$ is a proper subset of $V^2(D_8)$.



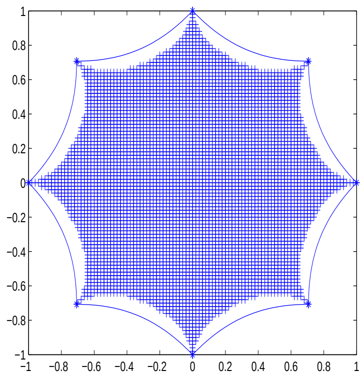
The boundary of $V^2(D_8)$.



The set $V^3(D_8)$.



The set $V^4(D_8)$.



The boundary of $V^2(D_8)$ and the set $V^3(D_8)$.

It would be nice to give an analytic description of $V^k(D_n)$ for $3 \leq k < n/2$.

In the proof of Theorem 3.2, we use at most n 4-by-4 submatrices instead of $\binom{n}{4}$ 4-by-4 submatrices to determine $V^2(D_n)$. In general, it is natural to ask the following:

Question Can we use a (small) sub-collection of 4-by-4 submatrices to determine $V^2(A)$ for diagonal matrices A, instead of all the $\binom{n}{4}$ of them?

4 Matrices whose squares are Hermitian

Suppose $A \in M_n$ is such that $e^{it}A^2$ is Hermitian for some $t \in [0, 2\pi)$. Then $B = e^{it/2}A$ satisfies B^2 is Hermitian. The joint numerical range $W(B, B^2)$ lies in $\mathbb{C} \times \mathbb{R} \equiv \mathbb{R}^3$, and it is convex if $n \geq 3$. We can use the theory of joint numerical ranges to characterize $V^2(B)$. We first obtain a canonical form for those matrices $A \in M_n$ such that A^2 is Hermitian.

Theorem 4.1 *Let $A \in M_n$. Then A^2 is Hermitian if and only if A is unitarily similar to a direct sum of Hermitian matrix, a skew-Hermitian matrix, and 2-by-2 matrices of the form:*

$$\begin{pmatrix} \mu & i\nu \\ i\nu & -\mu \end{pmatrix} \qquad \text{with } \mu, \nu > 0. \tag{4}$$

Proof. Let $A = H + iG$ so that $H = (h_{ij})$ and $G = (g_{ij})$ are Hermitian. Applying a unitary similarity to A, we may assume that $H = \text{diag}\,(h_1, \ldots, h_n)$. Then A^2 is Hermitian if and only if $HG+GH = 0$, which means $(g_{ij}(h_i + h_j)) = 0$, for all $1 \le i, j \le n$. Consequently, $g_{ij} = 0$ whenever $h_i + h_j \ne 0$. In particular, $g_{jj} = 0$ whenever $h_j \ne 0$. Assume $H = H_0 \oplus H_1 \oplus \cdots \oplus H_s \oplus 0_k$ such that $H_0, \ldots, H_s$ are nonsingular diagonal matrices with disjoint spectra, for $j = 0, 1, \ldots, s$, the spectrum of H_j equals $\{\mu_j, -\mu_j\}$ with $\mu_j > 0$, and for each eigenvalue λ of $H_0 \in M_l$ the value $-\lambda$ is not an eigenvalue of H. Then $G = 0_l \oplus G_1 \oplus \cdots G_s \oplus G_0$ such that $G_0 \in M_k$ and for each $j = 1, \ldots, s$, the matrix $H_j + iG_j$ has the form

$$\begin{pmatrix} \mu_j I & iR_j \\ iR_j^* & -\mu_j I \end{pmatrix}.$$

Let $W_j = U_j \oplus V_j$ be such that $U_j R_j V_j^*$ has it singular values lying on the diagonal positions. Then $W_j(H_j + iG_j)W_j^*$ is permutationally similar to a direct sum of 2-by-2 matrices of the form

$$\begin{pmatrix} \mu_j & i\nu \\ i\nu & -\mu_j \end{pmatrix} \qquad \text{with } \nu > 0$$

and a real scalar matrix if R_j is not a square matrix. Since this is true for every $j = 1, \ldots, s$, the conclusion follows. $\square$

Using Theorem 4.1, we can prove the following.

Theorem 4.2 *Suppose $A \in M_n$ is such that A^2 is Hermitian. Then $V^4(A) = \sigma(A)$.*

Proof. By Theorem 4.1, we may assume that $A = R \oplus S \oplus A_1 \oplus \cdots \oplus A_r$, where $R = R^*, S = -S^*, A_1, A_2, \ldots, A_r$ be as in (4). Suppose $\mu \in V^4(A)$. Then $\mu^2 \in \mathbb{R}$ and there is a unit vector $x \in \mathbb{C}^n$ such that $x^*A^jx = \mu^j$ for $j = 1, \ldots, 4$. Thus, $|x^*A^2x|^2 = |x^*A^4x| = \|A^2x\|^2$, and hence x is an eigenvector of A^2 such that $A^2x = \mu^2 x$. We need to prove that $\mu \in \sigma(A)$. If μ^2 is an eigenvalue of A_j^2 with $j \ge 1$, then $\pm\mu \in \sigma(A_j)$, and hence $\mu \in \sigma(A)$. Suppose it is not the case. Then μ^2 is an eigenvalue of R^2 or S^2 depending on $\mu^2 > 0$ or $\mu^2 < 0$. Assume $\mu^2 > 0$. If R has both eigenvalues $\pm\mu$, then again $\mu \in \sigma(A)$. Otherwise, the eigenspace of μ^2 of A^2 must be the eigenspace of an eigenvalue of A. Thus, x is a unit eigenvector of A, and hence $\mu = x^*Ax$ is an eigenvalue of A. If $\mu^2 < 0$, we can show that $\mu \in \sigma(S) \subseteq \sigma(A)$ by a similar argument. $\square$

To determine $V^2(A)$, we need the following result.

Theorem 4.3 *Suppose $A \in M_n$ and A^2 is Hermitian. Assume that A is unitarily similar to a direct sum of $R = \text{diag}\,(h_1, \ldots, h_p)$, $S = i\text{diag}\,(g_1, \ldots, g_q)$, and $A_j = \begin{pmatrix} \mu_j & i\nu_j \\ i\nu_j & -\mu_j \end{pmatrix}$ for $j = 1, \ldots, r$, such that $h_1 \ge \cdots \ge h_p$, $g_1 \ge \cdots \ge g_q$. Then the joint numerical range $W(A, \ldots, A^m)$ is convex for any positive integer m. Moreover, let*

$$\mathcal{E}_j = \{(x, y, \mu_j^2 - \nu_j^2) : x + iy \in W(A_j)\} = W(\text{Re}A_j, \text{Im}A_j, A_j^2).$$

Then the joint numerical range $W(A, A^2)$ is the convex hull of the set

$$\{(h_j, 0, h_j^2) : 1 \le j \le p\} \cup \{(0, g_j, -g_j^2) : 1 \le j \le q\} \cup \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_r.$$

Proof. Note that every point in $W(A, A^2, \ldots, A^m)$ is a convex combination of elements in $W(R, R^2, \ldots, R^m)$, $W(S, S^2, \ldots, S^m)$, $W(A_1, A_1^2, \ldots, A_1^m)$, …, and $W(A_r, A_r^2, \ldots, A_r^m)$. Since R and S are normal matrices, $W(R, R^2, \ldots, R^m)$ and $W(S, S^2, \ldots, S^m)$ are convex. Also, for each $j = 1, \ldots r$, A_j^k is a multiple of A_j or I. For example, $A_j^2 = \gamma_j I$, $A_j^3 = \gamma_j A_j$ and $A_j^4 = \gamma_j^2 I$, where $\gamma_j = \mu_j^2 - \nu_j^2$. So $W(A_j, A_j^2, \ldots, A_j^m)$ is a convex. Therefore, $W(A, A^2, \ldots, A^m)$ is a convex sum of points from $(r+2)$ convex set, and is convex. The second assertion can be easily verified. □

Now, we can characterize $V^2(A)$ for those $A \in M_n$ such that A^2 is Hermitian.

Theorem 4.4 *Suppose $A \in M_n$ satisfies the hypotheses of Theorem 4.3. Let K_1 be the convex hull of the union of the sets:*

(a.1) $\{(h_j, h_j^2) : 1 \le j \le p\}$,

(a.2) $\{(\pm\mu_j, \mu_j^2 - \nu_j^2) : 1 \le j \le r\}$,

(a.3) $\{(0, g_1 g_q), (0, \tilde{g})\}$ *if* $g_1 g_q \le 0$, *where* $\tilde{g} = \max\{g_u g_v : g_u g_v \le 0,\ 1 \le u < v \le q\}$.

Let K_2 be the convex hull of the union of the sets:

(b.1) $\{(g_j, -g_j^2) : 1 \le j \le q\}$,

(b.2) $\{(\pm\nu_j, \mu_j^2 - \nu_j^2) : 1 \le j \le r\}$,

(b.3) $\{(0, -h_1 h_p), (0, -\tilde{h})\}$ *if* $h_1 h_p \le 0$, *where* $\tilde{h} = \max\{h_u h_v : h_u h_v \le 0,\ 1 \le u < v \le p\}$.

Then

$$V^2(A) = \{\mu \in \mathbb{R} : (\mu, \mu^2) \in K_1\} \cup \{i\mu \in i\mathbb{R} : (\mu, -\mu^2) \in K_2\} \subseteq \mathbb{R} \cup i\mathbb{R}.$$

Proof. Use the fact that $\xi \in V^2(A)$ if and only if $(\xi, \xi^2) \in W(A, A^2)$. Since A^2 is Hermitian, $\xi^2 \in \mathbb{R}$. Thus, $\xi \in \mathbb{R} \cup i\mathbb{R}$.

By Theorem 4.3, $W(\mathrm{Re}A, \mathrm{Im}A, A^2)$ is the convex hull of the set

$$\{(h_j, 0, h_j^2) : 1 \le j \le p\} \cup \{(0, g_j, -g_j^2) : 1 \le j \le q\} \cup \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_r.$$

Let $P_1 = \{(\mu, 0, \mu^2) : \mu \in \mathbb{R}\}$ and $P_2 = \{(0, \mu, -\mu^2) : \mu \in \mathbb{R}\}$. Then $\{(h_j, 0, h_j^2) : 1 \le j \le p\} \subseteq P_1$, $\{(0, g_j, -g_j^2) : 1 \le j \le q\} \subseteq P_2$, and $\mathcal{E}_1, \ldots, \mathcal{E}_r$ are symmetric about the (x, z)-plane and the (y, z)-plane.

Let S_1 be the intersection of $W(\mathrm{Re}A, \mathrm{Im}A, A^2)$ and the (x, z)-plane. Then S_1 is the convex hull of the union of the sets:

(a.1)' $\{(h_j, 0, h_j^2) : 1 \le j \le p\}$,

(a.2)' $\{(\pm\mu_j, 0, \mu_j^2 - \nu_j^2) : 1 \le j \le r\}$,

(a.3)' $\{(0, 0, g_1 g_q), (0, \tilde{g})\}$ if $g_1 g_q \le 0$, where $\tilde{g} = \max\{g_u g_v : g_u g_v \le 0,\ 1 \le u < v \le q\}$.

As a result, $\mu \in V^2(A) \cap \mathbb{R}$ if and only if $(\mu, 0, \mu^2) \in P_1 \cap W(\mathrm{Re}A, \mathrm{Im}A, A^2) = P_1 \cap S_1$, equivalently, $(\mu, \mu^2) \in K_1$.

Similarly, we can show that $i\mu \in V^2(A) \cap i\mathbb{R}$ if and only if $(\mu, -\mu^2) \in K_2$. □

Using Theorem 4.4, one can recover many known results and obtain new ones. In particular, the next two examples cover Theorems 2.5 - 2.11 in [1].

Example 4.5 Let $A = \operatorname{diag}(h_1, h_2, h_3, i)$ with $h_1, h_2, h_3 \in \mathbb{R}$ such that $h_1 < h_2 < h_3$.

(a) If $h_1 \geq 0$ or $h_3 \leq 0$ then $V^2(A) = \sigma(A)$.

(b) If $h_1 h_3 < 0$ and for $\tilde{h} = \max\{h_r h_s : h_r h_s \leq 0,\ 1 \leq r < s \leq 3\}$, then

$$V^2(A) = \sigma(A) \cup \left\{ i\gamma : |\tilde{h}| \leq \gamma \leq \min\{|h_1 h_3|, 1\} \right\}.$$

Consequently, $V^2(A) = \sigma(A)$ if and only if $h_2 \neq 0$ and none of the triangles $\operatorname{conv}\{i, h_r, h_s\}$ be an acute angle triangle for any $r, s \in \{1, 2, 3\}$; otherwise, $V^2(A)$ contains a non-trivial line segment in $i\mathbb{R}$.

Example 4.6 Let $A = \operatorname{diag}(h_1, h_2, i, ig)$ with $h_1, h_2, g \in \mathbb{R} \setminus \{0\}$ with $h_2 < h_1$ and $g < 1$.

(a) If $\{h_1 h_2, g\} \subseteq (0, \infty)$, then $V^2(A) = \sigma(A)$.

(b) If $h_1 h_2 < 0$ and $g < 0$, then $V^2(A) \neq \sigma(A)$ and

$$V^2(A) = \sigma(A) \cup \left\{ \gamma : \gamma \in [h_2, h_1] \cap \left[\frac{-g}{h_2}, \frac{-g}{h_1} \right] \right\} \cup \left\{ i\gamma : \gamma \in [g, 1] \cap \left[\frac{-h_1 h_2}{g}, -h_1 h_2 \right] \right\},$$

which contains non-trivial line segments in $\mathbb{R} \cup i\mathbb{R}$.

(c) If $h_1 h_2 < 0 < g$, then

$$V^2(A) = \sigma(A) \cup \left\{ i\gamma : \gamma \in [g, 1] \cap \left[-h_1 h_2, \frac{-h_1 h_2}{g} \right] \right\}.$$

Consequently, $V^2(A) = \sigma(A)$ if and only if ig is the orthocenter of $\operatorname{conv}\{h_1, h_2, i\}$. Otherwise, $V^2(A)$ contains a non-trivial line segment in $i\mathbb{R}$.

Example 4.7 Let $A = \operatorname{diag}(3, -3, i, -i) \oplus \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Then

$$K_1 = \operatorname{conv}\{(3, 9), (-3, 9), (0, -1), (1, 0), (-1, 0)\},$$

$$K_2 = \operatorname{conv}\{(1, 1), (-1, 1), (1, 0), (-1, 0)\},$$

and hence

$$V^2(A) = \{-3, 3\} \cup [-3/2, 3/2] \cup \{i\gamma : \gamma \in [-1, 1]\}.$$

Question Note that it is possible that $V^3(A) \setminus \sigma(A)$ can be empty or non-empty. It would be nice to determine $V^3(A)$ if A^2 is Hermitian.

5 Additional results and remarks

In Section 3, we show that $V^{n-1}(D_n) = \sigma(D_n) \cup \{0\}$. Here we show that for any normal matrix $A \in M_n$, the set $V^{n-1}(A)$ is the union of the spectrum and at most one extra point. This conjecture was introduced to us by Anne Greenbaum via private communication.

Theorem 5.1 *Let $A \in M_n$ be a normal matrix with $n \geq 3$. Then $V^{n-1}(A)$ has at most $n+1$ points.*

Proof. Let $A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$, where $a_1, a_2, \ldots, a_n$ are n complex numbers. If $a_i = a_j$ for some $1 \leq i, j \leq n$. Then the degree of the minimal polynomial of A is less than n, and hence $V^{n-1}(A) = \sigma(A)$. Also, if $a_1, a_2, \ldots, a_n$ are collinear, then $V^{n-1}(A) = V^2(A) = \sigma(A)$. Now, we assume A has n distinct non-collinear eigenvalues. Let $\mu \in V^{n-1}(A) \setminus \sigma(A)$. We will show that $V^{n-1}(A) = \sigma(A) \cup \{\mu\}$. Assume if possible $\mu \neq \nu \in V^{n-1}(A) \setminus \sigma(A)$. Without loss of generality (by rotation and translation), we assume that $\mu = 0$ and $\nu = 1$. Let

$$W := \begin{pmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{pmatrix}, \quad \hat{\mu} := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad \hat{\nu} := \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

The matrix W is an invertible Vandermonde matrix. Since $\mu = 0$ and $\nu = 1$ are in $V^{n-1}(A)$, the equations $WX = \hat{\mu}$ and $WX = \hat{\nu}$ have solutions $X_1 = [t_1, \ldots, t_n]^t$ and $X_2 = [s_1, \ldots, s_n]^t$, respectively such that t_i and s_i are positive numbers, $i = 1, \ldots, n$. By Cramer's rule, we know that $t_k = \prod_{i=1, i\neq k}^{n} \frac{(a_i - 0)}{(a_i - a_k)}$ and $s_k = \prod_{i=1, i\neq k}^{n} \frac{(a_i - 1)}{(a_i - a_k)}$. Define $f(x) = \prod_{i=1}^{n}(a_i - x)$, Then $0 < \frac{t_k}{s_k} = \frac{(a_k - 1) f(0)}{(a_k - 0) f(1)}$. Thus, the argument $\arg(\frac{a_k - 1}{a_k}) = \arg(\frac{f(1)}{f(0)}) = \gamma$, for all $k = 1, \ldots, n$. Hence, $\arg(1 - \frac{1}{a_i}) = \arg(1 - \frac{1}{a_j}), \forall i, j = 1, \ldots, n$. Since $\mu = 0$ is an interior point of $W(A)$, there exist $1 \leq i, j \leq n$ such that $0 < \arg(a_i) < \pi$ and $-\pi < \arg(a_j) < 0$. Let $b_l := -\frac{1}{a_l}, l = i, j$. It is easy to see that $\arg(b_i) = \pi - \arg(a_i) > 0$ and $\arg(b_j) = \pi - \arg(a_j) < 0$. Therefore, $\arg(1 + b_i) > 0$ and $\arg(1 + b_j) < 0$, which is a contradiction. □

Remark 5.2 Let $A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$, where $a_1, a_2, \ldots, a_n$ are n complex numbers and $\mu \in V^{n-1}(A) \setminus \sigma(A)$. We may assume that $\mu = 0$ by replacing A by $A - \mu I$. Then $0 \in V^{n-1}(A) \setminus \sigma(A)$ if and only if $t_k = \prod_{i=1, i\neq k}^{n} \frac{a_i}{(a_i - a_k)}, k = 1, \ldots, n$ are positive numbers. In the case $n = 3$, we know that 0 is an orthocenter of the triangle generated by $\{a_1, a_2, a_3\}$. It would be nice to find some geometric interpretation for $\mu = 0$, if $n > 3$. Also it would be interesting to characterize those normal matrices $A \in M_n$ with n distinct eigenvalues such that $V^{n-1}(A) = \sigma(A)$.

Theorem 5.3 *Let $A \in M_n$, and $k \in \{1, \ldots, n\}$. Then $\mu \in V^k(A)$ if and only if $(\mu, \ldots, \mu^k) \in \operatorname{conv} W(A, \ldots, A^k)$. Moreover, every point $(\mu, \ldots, \mu^k)$ is a convex combination of no more than m elements in $W(A, \ldots, A^k)$ with $m \leq \min\{n, \sqrt{2k}\}$.*

Proof. There are nonnegative real numbers $t_1, \ldots, t_m$ summing to 1, and unit vectors $x_1, \ldots, x_m$ such that

$$(0, \ldots, 0) = \sum_{j=1}^{m} t_j \left(x_j^*(A - \mu I)x_j, \ldots, x_j^*(A - \mu I)^k x_j \right)$$

if and only if

$$\begin{aligned}
0 &= \sum_{j=1}^{m} t_j x_j^*(A - \mu I)x_j = \sum_{j=1}^{m} t_j(x_j^* A x_j) - \mu, \\
0 &= \sum_{j=1}^{m} t_j x_j^*(A - \mu I)^2 x_j = \sum_{j=1}^{m} t_j(x_j^* A^2 x_j - 2\mu x^* A x + \mu^2) = \sum_{j=1}^{m} t_j x_j^* A^2 x_j - \mu^2, \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
0 &= \sum_{j=1}^{m} t_j x_j^*(A - \mu I)^k x_j = \qquad \cdots\cdots\cdots \qquad = \sum_{j=1}^{m} t_j(x_j^* A^k x_j) - \mu^k.
\end{aligned}$$

Thus, the first assertion follows.

Let $A^j = H_j + iG_j$ with $H_j = (A + A_j^*)/2$ for $j = 1, \ldots, k$. By the result in [5], each point in $\operatorname{conv} W(H_1, G_1, \ldots, H_k, G_k)$ is a combination of no more than m points in $W(H_1, G_1, \ldots, H_k, G_k)$ with $m = \min\{n, \sqrt{2k} + \delta_{n^2, 2k+1}\}$. The result follows. $\square$

Remark 5.4 By Theorem 5.3, one may consider using the following approaches to study $V^k(A)$.

(a) Determine $\mu \in W(A)$ such that $(\mu, \mu^2, \ldots, \mu^k)$ is a convex combination of at most m elements in $W(A, \ldots, A^k)$. [Note that the m elements may not be extreme points of $\operatorname{conv} W(A, \ldots, A^k)$.] In particular, every $(\mu, \mu^2) \in \operatorname{conv} W(A, A^2)$ is a convex combination of at most two points in $W(A, A^2)$. Thus, one may study $V^2(A)$ by considering those lines joining (x^*Ax, x^*A^2x) and (y^*Ay, y^*A^2y) in $W(A, A^2)$.

(b) Use the fact that $\mu \notin V^k(A)$ if and only if $(\mu, \mu^2, \ldots, \mu^k) \notin \operatorname{conv} W(A, A^2, \ldots, A^k)$. For $j = 1, \ldots, k$, let $A^j = H_j + iG_j$, where H_j, G_j are Hermitian. Then the condition is equivalent to the fact the linear span of $\{H_1 - \operatorname{Re}(\mu)I, G_1 - \operatorname{Im}(\mu)I, \ldots, H_k - \operatorname{Re}(\mu^k)I, G_k - \operatorname{Im}(\mu^k)I\}$ contains a positive definite matrix. This condition can be readily checked by positive semidefinite programming. Alternatively, one can check whether the largest eigenvalue of a linear combination of $H_1 - \operatorname{Re}(\mu)I, G_1 - \operatorname{Im}(\mu)I, \ldots, H_k - \operatorname{Re}(\mu^k)I, G_k - \operatorname{Im}(\mu^k)I$ is negative.

(c) If A is normal with eigenvalues $a_1, \ldots, a_n$, we need to check whether

$$(0, \ldots, 0) \in \operatorname{conv} \left\{ \left((a_j - \mu), \ldots, (a_j - \mu)^k \right) : 1 \le j \le n \right\},$$

equivalently,

$$(\mu, \mu^2, \ldots, \mu^k) \in \operatorname{conv} \{(a_j, \ldots, a_j^k) : 1 \le j \le n\}.$$

This condition can be checked by standard linear programming package.

[In fact, this is how we generate $V^3(D_8)$ in Section 3.]

Question Can we determine $V^k(A)$ analytically for special classes of matrices A?

Some techniques in the previous sections can be further exploited. Here are two observations, which can be easily verified.

Theorem 5.5 *Let $A \in M_n$ and $k \in \{2, \ldots, n\}$.*

1) *If A^k is Hermitian, then $V^k(A) \subseteq \{\mu \in \mathbb{C} : \mu^k \in \mathbb{R}\}$.*

2) *Let $k \geq n/2$ and $A \in M_n$ such that $W(A, A^2, \ldots, A^k)$ is convex and $A^k = \alpha(A^*)^{n-k}$, where $\alpha \in \mathbb{C}$. Then $V^k(A) \subseteq \{re^{i\theta} : r \geq 0, r^{2k-n}e^{in\theta} = \alpha\}$.*

Acknowledgment

Research of the second author was supported by a USA NSF grant and a HK RCG grant. Research of the third author was supported by Shahid Bahonar University of Kerman and INSF Grant No. 84041. This research was done while he visiting the College of William and Mary in the academic year 2006-2007. He would like to thank the staff of the College of William and Mary. Also the authors would like to express their thanks to Professors Michael Lewis and David Phillips for their help in writing a Matlab program to generate $V^k(A)$.

References

[1] Ch. Davis and A. Salemi, On polynomial numerical hulls of normal matrices. Linear Algebra Appl. 383 (2004), 151–161.

[2] V. Faber, W. Joubert, M. Knill, and T. Manteuffel, Minimal residual method stronger than polynomial preconditioning, SIAM J. Matrix Anal. Appl., 17 (1996), 707-729.

[3] 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.

[4] O. Nevalinna, Convergence of iterations for linear equation, Birkhäuser, Basel, 1993.

[5] Y.T. Poon, On the convex hull of the multiform numerical range, Linear and Multilinear Algebra 37 (1994), 221-224.