Linear Algebra and its Applications 438 (2013) 3054–3061

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Higher-rank numerical ranges and Kippenhahn polynomials

Hwa-Long Gau,[a,*,1], Pei Yuan Wu[b,2]

[a] Department of Mathematics, National Central University, Chung-Li 320, Taiwan
[b] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

ARTICLE INFO

Article history:
Received 7 August 2012
Accepted 18 November 2012
Available online 8 January 2013

Submitted by P. Šemrl

AMS classification:
15A60

Keywords:
Higher-rank numerical range
Kippenhahn polynomial

ABSTRACT

We prove that two n-by-n matrices A and B have their rank-k numerical ranges $\Lambda_k(A)$ and $\Lambda_k(B)$ equal to each other for all k, $1 \leq k \leq \lfloor n/2 \rfloor + 1$, if and only if their Kippenhahn polynomials $p_A(x, y, z) \equiv \det(x\mathrm{Re}\, A + y\mathrm{Im}\, A + zI_n)$ and $p_B(x, y, z) \equiv \det(x\mathrm{Re}\, B + y\mathrm{Im}\, B + zI_n)$ coincide. The main tools for the proof are the Li-Sze characterization of higher-rank numerical ranges, Weyl's perturbation theorem for eigenvalues of Hermitian matrices and Bézout's theorem for the number of common zeros for two homogeneous polynomials.

For an n-by-n complex matrix A, its *rank-k numerical range* $(1 \leqslant k \leqslant n)$ is, by definition,

$$\Lambda_k(A) = \{\lambda \in \mathbb{C} : X^*AX = \lambda I_k \text{ for some } n\text{-by-}k \text{ matrix } X \text{ with } X^*X = I_k\}.$$

Motivated by investigations in connection with the quantum error correction, researchers started to study the higher-rank numerical ranges in [2]. The research was then pursued in a flurry of papers [4,3,15,10,12,11,6,5]. It is now known that $\Lambda_k(A)$, $1 \leqslant k \leqslant n$, is always convex [15], and, moreover, it consists of those λ's in $\mathbb{C}$ for which $\mathrm{Re}\,(e^{-i\theta}\lambda) \leqslant \lambda_k(\mathrm{Re}\,(e^{-i\theta}A))$ for all real θ [10, Theorem 2.2]. Here, and for our later discussions, we use $\mathrm{Re}\, X = (X + X^*)/2$ and $\mathrm{Im}\, X = (X - X^*)/(2i)$ to denote the *real* and *imaginary parts* of a finite matrix X, and, for an n-by-n Hermitian matrix Y, $\lambda_1(Y) \geqslant \cdots \geqslant \lambda_n(Y)$ denote its (ordered) eigenvalues. Note that the rank-one numerical range $\Lambda_1(A)$ coincides with the

* Corresponding author.
E-mail addresses: hlgau@math.ncu.edu.tw (H.-L. Gau), pywu@math.nctu.edu.tw (P.Y. Wu).
1 The research of Hwa-Long Gau is supported in part by the National Science Council of the Republic of China under project NSC 101-2115-M-008-006.
2 The research of Pei Yuan Wu is supported in part by the National Science Council of the Republic of China under project NSC 101-2115-M-009-004 and by the MOE-ATU.

0024-3795/$ - see front matter
http://dx.doi.org/10.1016/j.laa.2012.11.017

classical numerical range $W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ of A, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ are the standard inner product and its associated norm in $\mathbb{C}^n$.

The purpose of this paper is to determine when two matrices of the same size have all their higher-rank numerical ranges equal to each other. The following is the main theorem, which provides the answer.

Theorem 1. *The following conditions are equivalent for n-by-n matrices A and B:*

(a) $\Lambda_k(A) = \Lambda_k(B)$ *for all k*, $1 \leqslant k \leqslant \lfloor n/2\rfloor + 1$,
(b) $\det(x\mathrm{Re}\,A + y\mathrm{Im}\,A + zI_n) = \det(x\mathrm{Re}\,B + y\mathrm{Im}\,B + zI_n)$ *for all complex x, y and z, and*
(c) the eigenvalues of $\mathrm{Re}\,(e^{-i\theta}A)$ *and* $\mathrm{Re}\,(e^{-i\theta}B)$ *coincide (with the same multiplicities) for all real* θ.

Here $\lfloor n/2\rfloor$ denotes the largest integer which is less than or equal to $n/2$. For an n-by-n matrix X, we call $p_X(x, y, z) = \det(x\mathrm{Re}\,X + y\mathrm{Im}\,X + zI_n)$ the *Kippenhahn polynomial* of X. It is a degree-n homogeneous polynomial in x, y and z with real coefficients.

Note that when the n-by-n matrices A and B are such that p_A or p_B is irreducible, the equality of $\Lambda_1(A)$ and $\Lambda_1(B)$ already guarantees that p_A and p_B coincide (cf. [7, Corollary 2.4]). On the other hand, the number $\lfloor n/2\rfloor + 1$ in Theorem 1(a) cannot be further reduced as the 3-by-3 matrices $A = \mathrm{diag}\,(0, 1, 1)$ and $B = \mathrm{diag}\,(0, 0, 1)$ with $\Lambda_1(A) = \Lambda_1(B) = [0, 1]$, $p_A(x, y, z) = z(x+z)^2$ and $p_B(x, y, z) = z^2(x+z)$ show. Also, the conditions in Theorem 1 cannot be strengthened to the unitary equivalence of A and B. For example, if

$$A = \begin{bmatrix} 0 & 1 & \\ & 0 & 2 \\ & & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 2 & \\ & 0 & 1 \\ & & 0 \end{bmatrix},$$

then $\Lambda_1(A) = \Lambda_1(B) = \{z \in \mathbb{C} : |z| \leqslant \sqrt{5}/2\}$, $\Lambda_2(A) = \Lambda_2(B) = \{0\}$, $\Lambda_3(A) = \Lambda_3(B) = \emptyset$ and $p_A(x, y, z) = p_B(x, y, z) = z^3 - (5/4)(x^2 + y^2)z$, but A and B are not unitarily equivalent (cf. [16, Example 4]). However, for certain special-type matrices, we do have the unitary equivalence.

Corollary 2. *Let A and B be n-by-n matrices. If n = 2 or A and B are both normal or both companion matrices, then the conditions in Theorem 1 are equivalent to the unitary equivalence of A and B.*

Proof. It is well-known that if $X = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$, then $\Lambda_1(X)$ equals the elliptic disc with foci a and b and with minor axis of length $|c|$. Hence, for 2-by-2 matrices, $\Lambda_1(A) = \Lambda_1(B)$ implies the unitary equivalence of A and B. For general matrices, if $p_A(x, y, z) = p_B(x, y, z)$ for all x, y and z, then plugging in $x = 1$ and $y = i$ yields $\det(A + zI_n) = \det(B + zI_n)$ for all z, which implies that the eigenvalues of A and B coincide (with the same algebraic multiplicities). In particular, if A and B are normal or companion matrices, then obviously they are unitarily equivalent (in fact, equal in the latter case). $\square$

To prepare for the proof of Theorem 1, we review some basic properties of the Kippenhahn polynomials and numerical ranges. Recall that the *complex projective plane* $\mathbb{CP}^2$ consists of the equivalence classes of ordered triple $[x, y, z]$ of complex numbers x, y and z which are not all equal to zero under the equivalence relation: $[x, y, z] \sim [x', y', z']$ if $[x, y, z] = \lambda[x', y', z']$ for some nonzero scalar λ. The point $[x, y, z]$ in $\mathbb{CP}^2$ with $z \neq 0$ corresponds to the point $(x/z, y/z)$ in $\mathbb{C}^2$ and, conversely, (u, v) in $\mathbb{C}^2$ corresponds to $[u, v, 1]$ in $\mathbb{CP}^2$. For a homogeneous polynomial p in x, y and z, the *dual* of the algebraic curve $p(x, y, z) = 0$ in $\mathbb{CP}^2$ is the curve

$$\{[u, v, w] \in \mathbb{CP}^2 : ux + vy + wz = 0 \text{ is a tangent line of } p(x, y, z) = 0\}.$$

It is known that the dual of the dual curve is the original one (cf. [14, Theorem 1.5.3]). They have bearings on the numerical range because of Kippenhahn's result [8]: *the numerical range* $W(A)$ *of an* *n-by-n matrix A equals the convex hull of the real points* $(u/w, v/w)$ *of the dual curve of* $p(x, y, z) = 0$. In the following, we also need Bézout's theorem [9, Theorem 3.9], which counts the number of intersection points of two algebraic curves: *if two homogeneous polynomials p and q in x, y and z of degrees m and n, respectively, have no common factor, then the number of common zeros of p and q is at most mn.*

We now proceed to prove Theorem 1. The main part is to show the implication (a)⇒(b). This is done via a series of lemmas. Note that the Kippenhahn polynomial p_A of an n-by-n matrix A can be factored as the product of irreducible (real homogeneous) polynomials: $p_A = q_1^{n_1} \cdots q_m^{n_m}$, where the q_j's are distinct and $n_j \geqslant 1$ for all j. Under the condition in Theorem 1(a), we will show that each $q_j^{n_j}$ is also a factor of p_B and thus p_A divides p_B. Condition (b) then follows by symmetry. We start with the following lemma dealing with an irreducible q having degree at least two.

Lemma 3. *Let q be an irreducible real homogeneous polynomial in x, y and z with degree at least two. If C is the curve in the plane consisting of the real points of the dual of* $q(x, y, z) = 0$*, then the convex hull of C has no corner.*

Recall that, for a nonempty compact convex subset $\triangle$ of the plane, a point λ on the boundary of $\triangle$ is a *corner* of $\triangle$ if $\triangle$ has more than one supporting lines passing through it; otherwise, λ is a *differentiable point* of $\partial\triangle$.

Proof of Lemma 3. Let λ be a corner of the convex hull $\triangle$ of C. Then there are some θ_1 and θ_2, $\theta_1 < \theta_2$, such that $x\cos\theta + y\sin\theta = \mathrm{Re}\,(e^{-i\theta}\lambda)$ is a supporting line of $\triangle$ for all θ in (θ_1, θ_2). By duality, this implies that $q(\cos\theta, \sin\theta, -\mathrm{Re}\,(e^{-i\theta}\lambda)) = 0$ for all such θ's. On the other hand, $[\cos\theta, \sin\theta, -\mathrm{Re}\,(e^{-i\theta}\lambda)]$ is also a zero of the linear polynomial $\lambda_1 x + \lambda_2 y + z$, where $\lambda_1 = (\lambda + \overline{\lambda})/2$ and $\lambda_1 = (\lambda - \overline{\lambda})/(2i)$. Bézout's theorem then implies that $\lambda_1 x + \lambda_2 y + z$ is a factor of q, which contradicts the irreducibility of q. Hence $\triangle$ cannot have any corner. □

Essentially the same arguments as above were used in [8] to prove that every corner of $W(A)$ for a finite matrix A is an eigenvalue of A.

Another observation which we need is the following lemma, whose proof we omit.

Lemma 4. *Let* $\triangle$ *be a nonempty nonsingleton compact convex subset of the plane. Let* $\triangle = \cap_{\theta\in[0,2\pi)} H_\theta$*, where* $H_\theta = \{x + iy \in \mathbb{C} : x\cos\theta + y\sin\theta \leq d(\theta)\}$ *with* $\theta \mapsto d(\theta)$ *continuous, and let* λ *be a point in the boundary of* $\triangle$*.*

(a) If $\triangle$ *is not a line segment and* λ *is a differentiable point of* $\partial\triangle$*, then some* ∂H_θ *is the unique supporting line of* $\triangle$ *which passes through* λ*.*
(b) If $\triangle$ *is not a line segment and* λ *is a corner of* $\triangle$*, then there are* θ_1 *and* θ_2 *in* $[0, 2\pi)$ *with* $\theta_1 < \theta_2$ *such that* ∂H_{θ_1} *and* ∂H_{θ_2} *are supporting lines of* $\triangle$ *and* $\partial H_{\theta_1} \cap \partial H_{\theta_2} = \{\lambda\}$*. If we further require that* $(\theta_2 - \theta_1) \pmod{\pi}$ *be maximal, then* θ_1 *and* θ_2 *are unique.*
(c) If $\triangle$ *is a line segment, then there are unique* θ_1 *and* θ_2 *in* $[0, 2\pi)$ *with* $\theta_2 - \theta_1 = \pi$ *such that* $\triangle \subseteq \partial H_{\theta_1} \cap \partial H_{\theta_2}$*.*

In the following, this will be applied for an n-by-n matrix A with $\triangle = \Lambda_k(A)$ and $H_\theta = \{x + iy \in \mathbb{C} : x\cos\theta + y\sin\theta \leq \lambda_k(\mathrm{Re}\,(e^{-i\theta}A))\}$, $1 \leq k \leq n$. Note that in this case ∂H_θ is in general not a supporting line of $\triangle$ nor the converse. One example is

$$A = \begin{bmatrix} -1/\sqrt{2} & -1/2 & 1/(2\sqrt{2}) & 1/4 \\ & -1/\sqrt{2} & -1/2 & -1/(2\sqrt{2}) \\ & & 1/\sqrt{2} & -1/2 \\ & & & 1/\sqrt{2} \end{bmatrix}$$

as it is known that $\Lambda_2(A)$ has exactly two corners, around which the ∂H_θ's and the supporting lines are completely different (cf. [13, Example 7]).

We now proceed to obtain a characterization of p_A with the power of an irreducible factor of degree at least two in terms of the relative positions of the $\Lambda_k(A)$'s.

Lemma 5. *Let A be an n-by-n matrix, q be an irreducible real homogeneous polynomial in x, y and z with degree at least two, and C be the real part of the dual curve of* $q(x, y, z) = 0$. *Then* q^m *divides* p_A $(m \geq 1)$ *if and only if* $\partial\Lambda_{k_0}(A) \cap \partial\Lambda_{k_0-1}(A) \cap \cdots \cap \partial\Lambda_{k_0-m+1}(A)$ *contains an arc of C for some* k_0, $1 \leq k_0 \leq \lfloor n/2 \rfloor$.

A result we need in the proof is Weyl's perturbation theorem for (ordered) eigenvalues of Hermitian matrices (cf. [1, Theorem VI.2.1]): *if X and Y are n-by-n Hermitian matrices with eigenvalues* $\lambda_1(X) \geq \cdots \geq \lambda_n(X)$ *and* $\lambda_1(Y) \geq \cdots \geq \lambda_n(Y)$, *respectively, then* $|\lambda_j(X) - \lambda_j(Y)| \leq \|X - Y\|$ *for all* j, $1 \leq j \leq n$.

Proof of Lemma 5. Let $\triangle$ be the convex hull of C. Assume first that q^m divides p_A. Let k_0 be the largest integer for which $\triangle$ is contained in $\Lambda_{k_0}(A)$. Since $\triangle \subseteq \Lambda_1(A)$ by Kippenhahn's result, we have $k_0 \geq 1$. On the other hand, since $\Lambda_{\lfloor n/2 \rfloor+1}(A)$ is either a singleton or an empty set [2, Proposition 2.2], if $k_0 > \lfloor n/2 \rfloor$, then $\triangle \subseteq \Lambda_{\lfloor n/2 \rfloor+1}(A)$ and hence $\triangle$ is a singleton. By duality, this says that q is of degree one, contradicting our assumption that q has degree at least two. Thus $1 \leq k_0 \leq \lfloor n/2 \rfloor$. Note that, by Lemma 3, $\triangle$ has no corner. Hence, for each real θ, $\triangle$ has a unique supporting line $x\cos\theta + y\sin\theta = d(\theta)$ with $x\cos\theta + y\sin\theta \leq d(\theta)$ for all $x + iy$ in $\triangle$. Since $\triangle$ is not contained in $\Lambda_{k_0+1}(A)$, there is some λ_0 in $\triangle \setminus \Lambda_{k_0+1}(A)$. By the Li-Sze characterization [10, Theorem 2.2] of $\Lambda_{k_0+1}(A)$, we have $\mathrm{Re}\,(e^{-i\theta_0}\lambda_0) > \alpha_{k_0+1}(\theta_0)$ for some θ_0. Here $\alpha_k(\theta)$ denotes $\lambda_k(\mathrm{Re}\,(e^{-i\theta}A))$ for $1 \leq k \leq n$ and θ in $\mathbb{R}$. Weyl's perturbation theorem then implies that there is some $\delta > 0$ such that $\mathrm{Re}\,(e^{-i\theta}\lambda_0) > \alpha_{k_0+1}(\theta)$ for all θ in $(\theta_0 - \delta, \theta_0 + \delta)$. On the other hand, since λ_0 is in $\triangle$ and $\triangle$ is contained in $\Lambda_{k_0}(A)$, we also have $\mathrm{Re}\,(e^{-i\theta}\lambda_0) \leq d(\theta) \leq \alpha_{k_0}(\theta)$ for all θ. Thus $\alpha_{k_0+1}(\theta) < d(\theta) \leq \alpha_{k_0}(\theta)$ for all θ in $(\theta_0 - \delta, \theta_0 + \delta)$. Since $[\cos\theta, \sin\theta, -d(\theta)]$ is a zero of $q(x, y, z)$ by duality, the fact that q^m divides p_A implies that $[\cos\theta, \sin\theta, -d(\theta)]$ is a zero of $p_A(x, y, z)$ with multiplicity at least m. We infer from above that $d(\theta) = \alpha_k(\theta)$ for all k, $k_0 - m + 1 \leq k \leq k_0$, and all θ in $(\theta_0 - \delta, \theta_0 + \delta)$. We then obtain from $\triangle \subseteq \Lambda_k(A)$ that $x\cos\theta + y\sin\theta = d(\theta)$ is the unique supporting line of $\Lambda_k(A)$ for all such k's and θ's. Hence $\partial\Lambda_{k_0}(A) \cap \partial\Lambda_{k_0-1}(A) \cap \cdots \cap \partial\Lambda_{k_0-m+1}(A)$ contains an arc of C.

For the converse, if $\partial\Lambda_{k_0}(A) \cap \cdots \cap \partial\Lambda_{k_0-m+1}(A)$ contains an arc of C, then the supporting line $x\cos\theta + y\sin\theta = d(\theta)$ of $\triangle$ is also a supporting line of $\Lambda_k(A)$ for all k, $k_0 - m + 1 \leq k \leq k_0$, and all θ in some (θ_1, θ_2). This implies by the Li-Sze characterization [10, Theorem 2.2] of $\Lambda_k(A)$ and Lemma 4(a) that $d(\theta) = \alpha_k(\theta)$. Hence $[\cos\theta, \sin\theta, -d(\theta)]$ is a zero of $p_A(x, y, z)$ with multiplicity at least m for all such θ's. Since $[\cos\theta, \sin\theta, -d(\theta)]$ is a zero of $q(x, y, z)$ for all θ by duality, we obtain that $p_A(x, y, z)$ and $q(x, y, z)$ have infinitely many common zeros of the form $[\cos\theta, \sin\theta, -d(\theta)]$. Bézout's theorem yields that the irreducible q divides p_A. Next we claim that the number of θ's in $[0, 2\pi)$ for which $-d(\theta)$ is a zero of $q(\cos\theta, \sin\theta, z)$ with multiplicity at least two is finite. Indeed, if otherwise, then $p(x, y, z) \equiv \partial q(x, y, z)/\partial z$ and $q(x, y, z)$ have infinitely many common zeros of the form $[\cos\theta, \sin\theta, -d(\theta)]$. Since q is irreducible, Bézout's theorem implies that q divides p, which is impossible for the degree of $q(\cos\theta, \sin\theta, z)$ is one bigger than that of $p(\cos\theta, \sin\theta, z)$. Hence again we can apply Bézout's theorem to p_A/q and q to obtain that q divides p_A/q. Repeating these arguments yield that q divides p_A/q^j for all j, $0 \leq j \leq m - 1$. Thus q^m divides p_A, completing the proof. □

Corollary 6. *An n-by-n matrix A is normal if and only if* $\Lambda_k(A)$ *is a (closed) polygonal region for all k,* $1 \leq k \leq n$.

Here a polygonal region is one whose boundary is a polygon. In the degenerate case, this may be an empty set, a singleton or a line segment.

Proof of Corollary 6. The necessity follows from [10, Corollary 2.4]. The sufficiency is an easy consequence of Lemma 5 since the latter implies that p_A has only linear factors. □

We remark that, in relation to the preceding corollary, some geometric properties of the higher-rank numerical ranges of normal matrices have been studied in [5].

The next corollary is another consequence of Lemma 5.

Corollary 7. *Let A and B be n-by-n matrices with $\Lambda_k(A) = \Lambda_k(B)$ for all k, $1 \leqslant k \leqslant \lfloor n/2 \rfloor + 1$. Then p_A and p_B contain the same powers of irreducible factors with degrees at least two.*

To show that p_A and p_B contain the same powers of linear factors under the above conditions, we need a characterization, analogous to the one in Lemma 5, for powers of linear factors. Unfortunately, a complete analog of Lemma 5 is not true. We have only had the following necessary condition.

Lemma 8. *Let A be an n-by-n matrix. If $(ax + by + z)^m$ divides $p_A(x, y, z)$, where a and b are real and $m \geqslant 1$, then there is a k_0, $k_0 \geqslant m$, such that $a + bi$ is a corner of $\Lambda_k(A)$ for all k, $k_0 - m + 1 \leqslant k \leqslant k_0$, and is not in $\Lambda_{k_0+1}(A)$.*

Proof. Let $k_0 = \max\{k \geqslant 1 : a + bi \text{ is in } \Lambda_k(A)\}$. Then $a + bi$ is in $\Lambda_k(A)$ for all k, $1 \leq k \leq k_0$, and is not in $\Lambda_{k_0+1}(A)$. Because $a + bi$ is not in $\Lambda_{k_0+1}(A)$, there is some θ_0 such that $a\cos\theta_0 + b\sin\theta_0 > \alpha_{k_0+1}(\theta_0) \equiv \lambda_{k_0+1}(\mathrm{Re}\,(e^{-i\theta_0}A))$. By Weyl's perturbation theorem, we obtain $a\cos\theta + b\sin\theta > \alpha_{k_0+1}(\theta)$ on $(\theta_0 - \delta, \theta_0 + \delta)$ for some $\delta > 0$, where, for $1 \leqslant k \leqslant n$ and θ in $\mathbb{R}$, $\alpha_k(\theta)$ denotes $\lambda_k(\mathrm{Re}\,(e^{-i\theta}A))$. On the other hand, $a + bi$ being in $\Lambda_{k_0}(A)$ implies that $a\cos\theta + b\sin\theta \leqslant \alpha_{k_0}(\theta)$ for all real θ. Since $[\cos\theta, \sin\theta, -(a\cos\theta + b\sin\theta)]$ is a zero of $(ax + by + z)^m$ and hence a zero of $p_A(x, y, z)$ with multiplicity at least m, $a\cos\theta + b\sin\theta$ is an eigenvalue of $\mathrm{Re}\,(e^{-i\theta}A)$ with multiplicity at least m. Hence $k_0 \geq m$ and $a\cos\theta + b\sin\theta = \alpha_k(\theta)$ for all k, $k_0 - m + 1 \leqslant k \leqslant k_0$, on $(\theta_0 - \delta, \theta_0 + \delta)$. This means that $a + bi$ is a corner of $\Lambda_k(A)$ for all such k's. □

The next example shows that the converse of the assertion in Lemma 8 is not necessarily true. Recall that, for any subset $\triangle$ of the plane, $\triangle^\wedge$ denotes its convex hull.

Example 9. Let $A = \mathrm{diag}\,(1, i, -1, -i, 1/2, i/2, -1/2, -i/2, (1+i)/3)$. Then $\Lambda_1(A) = \{\pm 1, \pm i\}^\wedge$, $\Lambda_2(A) = \{\pm 1/2, \pm i/2, (\pm 1 \pm i)/3\}^\wedge$, $\Lambda_3(A) = \{0, 1/4, i/4, (1+i)/3\}^\wedge$, $\Lambda_4(A) = \{0\}$ and $\Lambda_k(A) = \emptyset$ for $5 \leqslant k \leqslant 9$ (cf. Fig. 1). Hence $(1+i)/3$ is a corner of $\Lambda_2(A)$ and $\Lambda_3(A)$, but $((1/3)x + (1/3)y + z)^2$ does not divide $p_A(x, y, z) = (x^2 - z^2)(y^2 - z^2)((1/4)x^2 - z^2)((1/4)y^2 - z^2)((1/3)x + (1/3)y + z)$.

For an n-by-n matrix A and $1 \leq \ell \leq \lfloor n/2 \rfloor$, let

$$V_\ell(A) = \{a + bi : ax + by + z \text{ is a real linear factor of } p_A(x, y, z) \text{ with multiplicity } m,\ 1 \leq m \leq \lfloor n/2 \rfloor - \ell + 1, \text{ and } a + bi \in \Lambda_{\ell+m-1}(A) \setminus \Lambda_{\ell+m}(A)\}.$$

We remark that if $ax+by+z$ is a real linear factor of $p_A(x, y, z)$ with multiplicity m, then Lemma 8 yields that there is a k_0, $k_0 \geqslant m$, such that $a+bi$ is a corner of $\Lambda_k(A)$ for all k, $k_0 - m + 1 \leqslant k \leqslant k_0$, and is not in $\Lambda_{k_0+1}(A)$. If, moreover, $k_0 \leq \lfloor n/2 \rfloor$ or, equivalently, $a+bi \notin \Lambda_{\lfloor n/2 \rfloor+1}(A)$, then $a+bi$ is in $V_{k_0-m+1}(A)$. Conversely, if $a + bi$ is in $V_{\ell_0}(A)$ and $a + bi \in \Lambda_{k_0}(A) \setminus \Lambda_{k_0+1}(A)$, then the definition of $V_{\ell_0}(A)$ yields that $a + bi \notin \Lambda_{\lfloor n/2 \rfloor+1}(A)$, $\ell_0 \leq k_0 \leq \lfloor n/2 \rfloor$ and $ax + by + z$ is a real linear factor of $p_A(x, y, z)$ with multiplicity $k_0 - \ell_0 + 1$. Obviously, the $V_\ell(A)$'s and $\Lambda_{\lfloor n/2 \rfloor+1}(A)$ are mutually disjoint and $V_1(A) \cup V_2(A) \cup \cdots \cup V_{\lfloor n/2 \rfloor}(A) \cup \Lambda_{\lfloor n/2 \rfloor+1}(A) = \{a+bi : ax+by+z \text{ is a real linear factor of } p_A(x, y, z)\}$. For the proof of the latter, note that $\Lambda_{\lfloor n/2 \rfloor+1}(A)$ is either empty or a singleton (cf. [2, Proposition 2.2]). Hence we need only show that if $\Lambda_{\lfloor n/2 \rfloor+1}(A) = \{a+bi\}$, then $ax+by+z$ is a factor of $p_A(x, y, z)$. Indeed, in this case, $[\cos\theta, \sin\theta, -\lambda_{\lfloor n/2 \rfloor+1}(\mathrm{Re}\,(e^{-i\theta}A))]$ is a zero of both $p_A(x, y, z)$ and $ax+by+z$ for all real θ by the Li-Sze characterization of $\Lambda_{\lfloor n/2 \rfloor+1}(A)$. The assertion then follows from Bézout's theorem. As

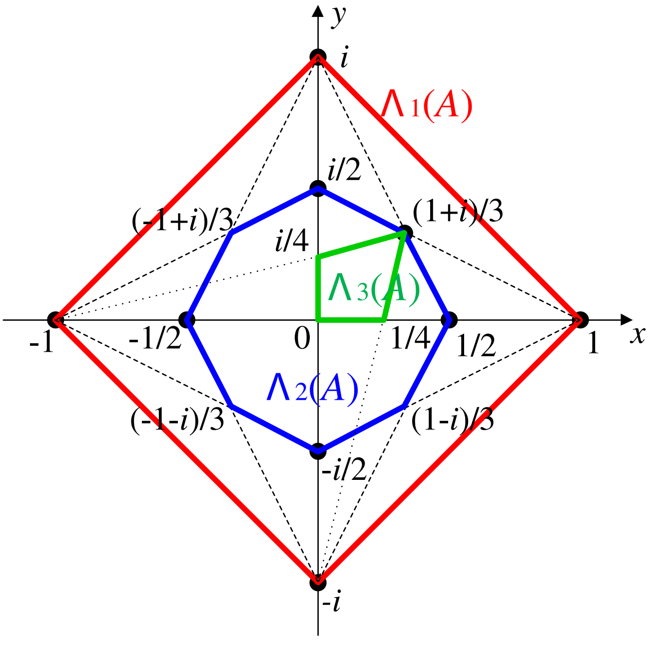

Fig. 1. Higher-rank numerical ranges of A.

an example, in Example 9, we have $V_1(A) = \{\pm 1, \pm i\}$, $V_2(A) = \{\pm 1/2, \pm i/2\}$, $V_3(A) = \{(1+i)/3\}$ and $V_4(A) = \emptyset$.

The next lemma would help us to conclude the proof of Theorem 1.

Lemma 10. *Let A and B be n-by-n matrices with $\Lambda_k(A) = \Lambda_k(B)$ for all k, $1 \leqslant k \leqslant \lfloor n/2 \rfloor + 1$. Then $V_\ell(A) = V_\ell(B)$ and $\lambda_\ell(\text{Re}\,(e^{-i\theta}A)) = \lambda_\ell(\text{Re}\,(e^{-i\theta}B))$ for all ℓ, $1 \le \ell \le \lfloor n/2 \rfloor$, and θ in $\mathbb{R}$.*

Proof. For $1 \le k \le n$ and real θ, let $\alpha_k(\theta) = \lambda_k(\text{Re}\,(e^{-i\theta}A))$ and $\beta_k(\theta) = \lambda_k(\text{Re}\,(e^{-i\theta}B))$. We now prove our assertion by induction on ℓ.

If $\ell = 1$, then, since A and B have the same numerical range, we have $\alpha_1(\theta) = \beta_1(\theta)$ for all real θ. We now check that $V_1(A) \subseteq V_1(B)$. Indeed, if $a + bi \in V_1(A)$ and $ax + by + z$ is a real linear factor of $p_A(x, y, z)$ with multiplicity m, $1 \le m \le \lfloor n/2 \rfloor$, then Lemma 8 says that $a + bi$ is a corner of $\Lambda_k(A)$, $1 \le k \le m$, and is not in $\Lambda_{m+1}(A)$. Since $\Lambda_k(A) = \Lambda_k(B)$ for all k, $1 \leqslant k \leqslant \lfloor n/2 \rfloor + 1$, and $m \le \lfloor n/2 \rfloor$, we obtain that $a + bi$ is also a corner of $\Lambda_k(B)$ for $1 \leqslant k \leqslant m$ and is not in $\Lambda_{m+1}(B)$. As $a + bi$ is a corner of $\Lambda_1(B)$, it is an eigenvalue of B and $ax + by + z$ divides $p_B(x, y, z)$. Moreover, there are θ_1 and θ_2, $\theta_1 < \theta_2$, such that $a\cos\theta + b\sin\theta = \beta_1(\theta)$ for all θ in (θ_1, θ_2). On the other hand, $a + bi$ being in $\Lambda_m(B)$ implies that $a\cos\theta + b\sin\theta \leqslant \beta_m(\theta)$ for all real θ by the Li-Sze characterization of $\Lambda_m(B)$. We thus have $a\cos\theta + b\sin\theta = \beta_1(\theta) = \cdots = \beta_m(\theta)$ and therefore $(a\cos\theta + b\sin\theta + z)^m$ divides $p_B(\cos\theta, \sin\theta, z)$ for θ in (θ_1, θ_2). Using Bézout's theorem repeatedly, we conclude that $(ax+by+z)^m$ divides $p_B(x, y, z)$. This means that $ax + by + z$ is a real linear factor of $p_B(x, y, z)$ with multiplicity at least m. But since $a + bi \in \Lambda_m(B) \setminus \Lambda_{m+1}(B)$, $1 \le m \le \lfloor n/2 \rfloor$, by Lemma 8 and the definition of $V_1(B)$, we deduce that $ax + by + z$ is a real linear factor of $p_B(x, y, z)$ with multiplicity m and $a + bi$ is in $V_1(B)$. Therefore, we have $V_1(A) \subseteq V_1(B)$. By symmetry, we also obtain $V_1(B) \subseteq V_1(A)$. Hence we conclude that $V_1(A) = V_1(B)$.

Next assume that our assertion is true for all ℓ, $1 \le \ell < \ell_0$, and $\ell_0 \le \lfloor n/2 \rfloor$. We prove its validity for ℓ_0. Firstly, we check that $\alpha_{\ell_0}(\theta) = \beta_{\ell_0}(\theta)$ for all real θ. Indeed, if otherwise, then we have $\alpha_{\ell_0}(\theta_0) \neq \beta_{\ell_0}(\theta_0)$ for some θ_0. Without loss of generality, we may assume that $\alpha_{\ell_0}(\theta_0) < \beta_{\ell_0}(\theta_0)$. Weyl's perturbation theorem then yields that $\alpha_{\ell_0}(\theta) < \beta_{\ell_0}(\theta)$ for all θ in some neighborhood $I \equiv (\theta_0 - \delta, \theta_0 + \delta)$

of θ_0 ($\delta > 0$). Since $\beta_{\ell_0}(\theta)$ is an eigenvalue of Re $(e^{-i\theta}B)$, we have $p_B(\cos\theta, \sin\theta, -\beta_{\ell_0}(\theta)) = 0$ for all real θ. Note that p_B has only finitely many irreducible factors. Thus there is some irreducible factor q of p_B and an infinite subset I_1 of I such that $q(\cos\theta, \sin\theta, -\beta_{\ell_0}(\theta)) = 0$ for all θ in I_1.

If q is of degree at least two, then, by Corollary 7, q is also a factor of p_A. Hence $p_A(\cos\theta, \sin\theta, -\beta_{\ell_0}(\theta)) = 0$ and thus $\beta_{\ell_0}(\theta)$ is an eigenvalue of Re $(e^{-i\theta}A)$ for all $\theta \in I_1$. Moreover, by the induction hypothesis, we have $\alpha_{\ell_0-1}(\theta) = \beta_{\ell_0-1}(\theta) \geq \beta_{\ell_0}(\theta) > \alpha_{\ell_0}(\theta)$ for all $\theta \in I$, and thus $\beta_{\ell_0}(\theta) = \alpha_{\ell_0-1}(\theta) = \beta_{\ell_0-1}(\theta) > \alpha_{\ell_0}(\theta)$ and $q(\cos\theta, \sin\theta, -\alpha_{\ell_0-1}(\theta)) = 0$ for all θ in I_1. Let m be the multiplicity of $q(x, y, z)$ in $p_A(x, y, z)$. Then $\alpha_{\ell_0}(\theta) < \alpha_{\ell_0-1}(\theta) = \cdots = \alpha_{\ell_0-m}(\theta)$ for all $\theta \in I_1$. The induction hypothesis says that $\alpha_\ell(\theta) = \beta_\ell(\theta)$ for all ℓ, $1 \leq \ell < \ell_0$, and all real θ, and thus $\beta_{\ell_0-m}(\theta) = \cdots = \beta_{\ell_0-1}(\theta) = \beta_{\ell_0}(\theta)$ for all $\theta \in I_1$. Since the set I_1 is infinite, using Bézout's theorem repeatedly (as in the proof of Lemma 5), we obtain that q^{m+1} divides p_B. This means that q is an irreducible factor of p_B with multiplicity at least $m+1$, which contradicts the assertion of Corollary 7. Hence q must be linear.

Let $q(x, y, z) = ax+by+z$. Then $\alpha_{\ell_0}(\theta) < \beta_{\ell_0}(\theta) = a\cos\theta+b\sin\theta$ for all $\theta \in I_1$, and hence $a+bi$ is not in $\Lambda_{\ell_0}(A) = \Lambda_{\ell_0}(B)$ by the Li-Sze characterization of $\Lambda_{\ell_0}(A)$. This implies that $a+bi$ is in $V_\ell(B)$ for some $\ell < \ell_0$. The induction hypothesis yields that $a+bi$ is also in $V_\ell(A)$ $(= V_\ell(B))$. In particular, $ax+by+z$ is a real linear factor of $p_A(x, y, z)$. Hence $p_A(\cos\theta, \sin\theta, -\beta_{\ell_0}(\theta)) = 0$ and thus $\beta_{\ell_0}(\theta)$ is an eigenvalue of Re $(e^{-i\theta}A)$ for all $\theta \in I_1$. Moreover, by the induction hypothesis, we have $\alpha_{\ell_0-1}(\theta) = \beta_{\ell_0-1}(\theta) \geq \beta_{\ell_0}(\theta) > \alpha_{\ell_0}(\theta)$ for all $\theta \in I$. This forces that $\alpha_{\ell_0-1}(\theta) = \beta_{\ell_0}(\theta) > \alpha_{\ell_0}(\theta)$ for all $\theta \in I_1$. Let m be the multiplicity of $ax+by+z$ in $p_A(x, y, z)$. Then $\alpha_{\ell_0}(\theta) < \alpha_{\ell_0-1}(\theta) = \cdots = \alpha_{\ell_0-m}(\theta)$ for all $\theta \in I_1$. The induction hypothesis implies that $\beta_{\ell_0-m}(\theta) = \cdots = \beta_{\ell_0-1}(\theta) = \beta_{\ell_0}(\theta)$ for all $\theta \in I_1$. Using Bézout's theorem repeatedly, we obtain that $(ax+by+z)^{m+1}$ divides $p_B(x, y, z)$. This means that $ax+by+z$ is a linear factor of $p_B(x, y, z)$ with multiplicity at least $m+1$. On the other hand, since $a+bi \in V_\ell(A) = V_\ell(B)$, under Lemma 8 and our assumption, $a+bi$ is a corner of $\Lambda_k(A) = \Lambda_k(B)$ for $\ell \leqslant k \leqslant \ell+m-1$ and is not in $\Lambda_{\ell+m}(A) = \Lambda_{\ell+m}(B)$. Hence the definition of $V_\ell(B)$ yields that the multiplicity of $ax+by+z$ in $p_B(x, y, z)$ is m, a contradiction. Thus $\alpha_{\ell_0}(\theta) = \beta_{\ell_0}(\theta)$ for all real θ as asserted.

We now show that $V_{\ell_0}(A) \subseteq V_{\ell_0}(B)$. Suppose that $a+bi \in V_{\ell_0}(A)$ and $ax+by+z$ is a real linear factor of $p_A(x, y, z)$ with multiplicity m. Lemma 8 yields that $a+bi$ is a corner of $\Lambda_k(A)$ for $\ell_0 \leqslant k \leqslant k_0$ and is not in $\Lambda_{k_0+1}(A)$, where $k_0 = \ell_0+m-1$. Note that $a+bi \notin \Lambda_{\lfloor n/2 \rfloor+1}(A)$ implies $k_0 \leq \lfloor n/2 \rfloor$. Since $\Lambda_k(A) = \Lambda_k(B)$ for all k, $1 \leq k \leq \lfloor n/2 \rfloor + 1$, $a+bi$ is also a corner of $\Lambda_k(B) = \Lambda_k(A)$ for $\ell_0 \leqslant k \leqslant k_0$ and is not in $\Lambda_{k_0+1}(B) = \Lambda_{k_0+1}(A)$. The former implies that $a\cos\theta + b\sin\theta \leqslant \alpha_{k_0}(\theta), \beta_{k_0}(\theta)$ for all real θ while the latter, by Lemma 4(b) and (c), that $a\cos\theta_0 + b\sin\theta_0 > \alpha_{k_0+1}(\theta_0), \beta_{k_0+1}(\theta_0)$ for some common θ_0, both by the Li-Sze characterization of the higher-rank numerical ranges. Weyl's perturbation theorem then yields that $a\cos\theta+b\sin\theta > \alpha_{k_0+1}(\theta), \beta_{k_0+1}(\theta)$ for all θ in some neighborhood $I \equiv (\theta_0-\delta, \theta_0+\delta)$ of θ_0 ($\delta > 0$). Since $(ax+by+z)^m$ divides $p_A(x, y, z)$, $a\cos\theta+b\sin\theta$ appears as m $(= k_0-\ell_0+1)$ values of the $\alpha_k(\theta)$'s. Thus $a\cos\theta+b\sin\theta = \alpha_{k_0}(\theta) = \cdots = \alpha_{\ell_0}(\theta)$ for θ in I. Since we have proved that $\alpha_{\ell_0}(\theta) = \beta_{\ell_0}(\theta)$ for all real θ, thus $a\cos\theta+b\sin\theta = \beta_{\ell_0}(\theta)$ for all θ in I. It follows that $a\cos\theta+b\sin\theta = \beta_{k_0}(\theta) = \cdots = \beta_{\ell_0}(\theta)$ for θ in I. Using Bézout's theorem repeatedly, we obtain that $(ax+by+z)^m$ divides p_B. This means that $ax+by+z$ is a real linear factor of $p_B(x, y, z)$ with multiplicity at least m. Moreover, from the definition of the $V_\ell(B)$'s, $a+bi \in \Lambda_{\ell_0+m-1}(B) \setminus \Lambda_{\ell_0+m}(B)$ implies that $a+bi \in V_{\ell_1}(B)$ for some $\ell_1 \leq \ell_0$. If $\ell_1 < \ell_0$, the induction hypothesis yields that $a+bi$ is also in $V_{\ell_1}(A)$, which contradicts the mutual disjointness of the $V_\ell(A)$'s. Hence we conclude that $\ell_1 = \ell_0$ or $a+bi \in V_{\ell_0}(B)$ as desired. For the converse, interchanging A with B in the above arguments, we also obtain $V_{\ell_0}(B) \subseteq V_{\ell_0}(A)$. Thus $V_{\ell_0}(A) = V_{\ell_0}(B)$, completing the proof. □

We are now ready to prove Theorem 1.

Proof of Theorem 1. The implication (b)⇒(c) is trivial and the implication (c)⇒(a) follows from the Li-Sze characterization of the higher-rank numerical ranges. We need only prove (a)⇒(b). Suppose that $q(x, y, z)$ is an irreducible factor of $p_A(x, y, z)$ with multiplicity m. If q is of degree at least two, then Corollary 7 implies that $q(x, y, z)$ is also an irreducible factor of $p_B(x, y, z)$ with multiplicity m. Assume next that q is of degree one, say, $q(x, y, z) = ax + by + z$. If $a + bi$ is not in $\Lambda_{\lfloor n/2\rfloor+1}(A)$, then $a + bi$ is in $V_{\ell_0}(A)$ and $a + bi \in \Lambda_{\ell_0+m-1}(A) \setminus \Lambda_{\ell_0+m}(A)$ for some ℓ_0 with $\ell_0 + m \leq \lfloor n/2\rfloor + 1$. Lemma 10 and the condition in (a) implies that $a + bi$ is also in $V_{\ell_0}(B)$ and $a + bi \in \Lambda_{\ell_0+m-1}(B) \setminus \Lambda_{\ell_0+m}(B)$. Hence $ax + by + z$ is also a factor of $p_B(x, y, z)$ with multiplicity m. Therefore, if $\Lambda_{\lfloor n/2\rfloor+1}(A)$ is empty, then, since p_A and p_B have the same degree, we have $p_A = p_B$.

On the other hand, if $\Lambda_{\lfloor n/2\rfloor+1}(A)$ is nonempty, then it must be a singleton, say, $\Lambda_{\lfloor n/2\rfloor+1}(A) = \{c + di\}$. We have $\Lambda_{\lfloor n/2\rfloor+1}(B) = \Lambda_{\lfloor n/2\rfloor+1}(A) = \{c + di\}$, and $cx + dy + z$ is a real linear factor of both $p_A(x, y, z)$ and $p_B(x, y, z)$ (cf. the paragraph after Example 9). Since the degrees of p_A and p_B coincide, we infer from what were proved before that $p_A(x, y, z) = p_B(x, y, z)$, completing the proof. □

The study of the higher-rank numerical ranges is useful in understanding the structure of matrices in general. How special features of the former and properties of the latter are related is worthy of further explorations. For example, what can be inferred about a matrix A if some of its $\Lambda_k(A)$'s have a common corner or a common line segment on their boundaries? and, conversely, which matrix A has such common corners or common line segments? The interplay between them should make interesting research topics.

References

[1] R. Bhatia, Matrix Analysis, Springer, New York, 1997.
[2] M.-D. Choi, D.W. Kribs, K. Życzkowski, Higher-rank numerical ranges and compression problems, Linear Algebra Appl. 418 (2006) 828–839.
[3] M.-D. Choi, M. Giesinger, J.A. Holbrook, D.W. Kribs, Geometry of higher-rank numerical ranges, Linear and Multilinear Algebra 56 (2008) 53–64.
[4] M.-D. Choi, J.A. Holbrook, D.W. Kribs, K. Życzkowski, Higher-rank numerical ranges of unitary and normal matrices, Oper. Matrices 1 (2007) 409–426.
[5] H.-L. Gau, C.-K. Li, Y.-T. Poon, N.-S. Sze, Higher rank numerical ranges of normal matrices, SIAM J. Matrix Anal. Appl. 32 (2011) 23–43.
[6] H.-L. Gau, C.-K. Li, P.Y. Wu, Higher-rank numerical ranges and dilations, J. Operator Theory 63 (2010) 181–189.
[7] H.-L. Gau, P.Y. Wu, Companion matrices: reducibility, numerical ranges and similarity to contractions, Linear Algebra Appl. 383 (2004) 127–142.
[8] R. Kippenhahn, Über den Wertevorrat einer Matrix, Math. Nachr. 6 (1951) 193–228., English translation: P.F. Zachlin, M.E. Hochstenbach, On the numerical range of a matrix, Linear and Multilinear Algebra 56 (2008) 185–225.
[9] F. Kirwan, Complex Algebraic Curves, Cambridge University Press, Cambridge, 1992.
[10] C.-K. Li, N.-S. Sze, Canonical forms, higher rank numerical ranges, totally isotropic subspaces, and matrix equations, Proc. Amer. Math. Soc. 136 (2008) 3013–3023.
[11] C.-K. Li, Y.-T. Poon, N.-S. Sze, Condition for the higher rank numerical range to be non-empty, Linear and Multilinear Algebra 57 (2009) 365–368.
[12] R.A. Martínez-Avendaño, Higher-rank numerical range in infinite-dimensional Hilbert space, Oper. Matrices 2 (2008) 249–264.
[13] B. Mirman, Numerical ranges and Poncelet curves, Linear Algebra Appl. 281 (1998) 59–85.
[14] M. Namba, Geometry of Projective Algebraic Curve, Marcel Dekker, New York, 1984.
[15] H. Woerdeman, The higher rank numerical range is convex, Linear and Multilinear Algebra 56 (2008) 65–67.
[16] P.Y. Wu, A numerical range characterization of Jordan blocks, Linear and Multilinear Algebra 43 (1998) 351–361.