S_n-矩陣的數值域和壓縮

中文摘要：

設 A 係一S_n-矩陣，即 A 是一$n \times n$收縮矩陣，其特徵值的絕對值均小於 1，且滿足$I_n - A^* A$的秩爲 1。設$W(A)$表示 A 的數值域。在本篇論文中，我們證明：

(一) 設 B 是一個$k \times k$ $(1 \le k < n)$大小的 A 的壓縮，則$W(B)$包含於且不等於$W(A)$，

(二) 設 A 表示成上三角的標準形式，且 B 是一個$k \times k$ $(1 \le k < n)$的 A 的主要子矩陣，則$W(B)$包含在$W(A)$的內部，

(三) 設$k(A)$表示最大的k的值使得存在一個$k \times k$大小的 A 的壓縮，其對角元素都在$W(A)$的邊界上，則$k(A) = 2$ 如果$n = 2$，且$k(A) = \lceil n/2 \rceil$ 如果$n \ge 3$。

關鍵字：數值域，壓縮，S_n-矩陣

Operators and Matrices

[0648] First Galley Proofs

NUMERICAL RANGES AND COMPRESSIONS OF S_n–MATRICES

HWA-LONG GAU AND PEI YUAN WU

(*Communicated by H. Bercovici*)

Abstract. Let A be an n-by-n $(n \geqslant 2)$ S_n-matrix, that is, A is a contraction with eigenvalues in the open unit disc and with $\operatorname{rank}(I_n - A^*A) = 1$, and let $W(A)$ denote its numerical range. We show that (1) if B is a k-by-k $(1 \leqslant k < n)$ compression of A, then $W(B) \subsetneqq W(A)$, (2) if A is in the standard upper-triangular form and B is a k-by-k $(1 \leqslant k < n)$ principal submatrix of A, then $\partial W(B) \cap \partial W(A) = \emptyset$, and (3) the maximum value of k for which there is a k-by-k compression of A with all its diagonal entries in $\partial W(A)$ is equal to 2 if $n = 2$, and $\lceil n/2 \rceil$ if $n \geqslant 3$.

1. Introduction

Let A be an n-by-n complex matrix. Its *numerical range* $W(A)$ is, by definition, the set $\{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$, where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the standard inner product and its associated norm of vectors in $\mathbb{C}^n$, respectively. It is well known that $W(A)$ is a nonempty compact convex subset of the complex plane. For other properties of the numerical range, we refer the reader to [14, Chapter 1]. A k-by-k matrix B is a *compression* of A if $B = V^*AV$ for some n-by-k matrix V with $V^*V = I_k$. In this case, we also say that A is a *dilation* of B or B *dilates* to A.

An n-by-n matrix A is said to be of *class* S_n if it is a contraction, that is, $\|A\| \equiv \max_{\|x\|=1} \|Ax\| \leqslant 1$, has all its eigenvalues in the open unit disc $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$, and satisfies $\operatorname{rank}(I_n - A^*A) = 1$. Such matrices are finite-dimensional versions of the so-called $S(\phi)$-contractions, first studied by Sarason [20] back in 1967. Their many nice properties have been explored since then. For the past 15 years or so, the focus of investigations was shifted to their numerical ranges. These were mainly done by three groups of researchers, namely, the present authors [6, 7, 8, 9, 10, 11, 12, 21], Mirman *et al.* [15, 16, 17, 18, 19] and Gorkin *et al.* [2, 3, 4].

In this paper, we obtain further properties of the numerical range $W(A)$ of an S_n-matrix A by relating it to the compression of A. We start, in Section 2 below, by reviewing some results on such numerical ranges which will be needed in later discussions. The main result in Section 3, Theorem 3.3, says that (a) if B is a k-by-k $(1 \leqslant k < n)$

Mathematics subject classification (2010): 15A60.

Keywords and phrases: Numerical range, compression, S_n-matrix, unitary dilation.

Research supported in part by the National Science Council of the Republic of China under project NSC 100-2115-M-008-004.

Research supported in part by the National Science Council of the Republic of China under project NSC 99-2115-M-009-002-MY2 and by the MOE-ATU.

compression of the S_n-matrix A, then $W(B)$ is properly contained in $W(A)$, and (b) if A is represented as the standard upper-triangular form as specified in Theorem 2.1 and B is any k-by-k ($1 \leqslant k < n$) principal submatrix of A, then $W(B)$ is even contained in the interior of $W(A)$. As a consequence, in the case of (b), a unit vector x in $\mathbb{C}^n$ for which $\langle Ax, x\rangle$ belongs to the boundary $\partial W(A)$ of $W(A)$ can have only nonzero components (Corollary 3.4). Then in Section 4, we consider, for an S_n-matrix A, its compressions B with all the diagonal entries in the boundary of $W(A)$. We show that the maximum size of B is 2 if $n = 2$, and is $\lceil n/2 \rceil$ if $n \geqslant 3$ (Proposition 4.3 and Theorem 4.4). This is the same as the maximum number of orthonormal vectors x for which $\langle Ax, x\rangle$ belongs to $\partial W(A)$. This is proven via the inscribing-circumscribing property, as described in Section 2, between the unit circle $\partial\mathbb{D}$ and the boundary $\partial W(A)$ of the $(n+1)$-gons formed by the eigenvalues of $(n+1)$-by-$(n+1)$ unitary dilations of A.

2. Numerical range of S_n-matrix

In this section, we briefly review the basic ingredients which we would need for the proofs of our results in Sections 3 and 4. The first one is from [8, Corollary 1.3].

THEOREM 2.1. *An n-by-n matrix A is of class S_n if and only if it is unitarily equivalent to the upper-triangular matrix $[a_{ij}]_{i,j=1}^n$ with $|a_{ii}| < 1$ for all i and*

$$a_{ij} = (1 - |a_{ii}|^2)^{1/2}(1 - |a_{jj}|^2)^{1/2} \prod_{k=i+1}^{j-1} (-\overline{a}_{kk})$$

for $j > i$.

Such a matrix is said to be in a *standard upper-triangular form*.

The next two theorems are concerned with $(n+1)$-by-$(n+1)$ unitary dilations of an S_n-matrix, the first of which is from [6, Theorem 2.1 and Lemma 2.2] while the second is essentially proven in [6, Theorems 3.1 and 3.2].

THEOREM 2.2. *If A is an S_n-matrix, then for any point λ in $\partial\mathbb{D}$, there is a unique $(n+1)$-gon P such that λ is a vertex of P, and P is inscribed on $\partial\mathbb{D}$ and circumscribed about $W(A)$ with each edge tangent to $\partial W(A)$ at exactly one point. Moreover, such $(n+1)$-gons are in one-to-one correspondence with $(n+1)$-by-$(n+1)$ unitary dilations U (up to unitary equivalence) of A under which the vertices of P are exactly the eigenvalues of U. In particular, every point a on $\partial W(A)$ has a unique supporting line of $W(A)$ which passes it and is an extreme point of $W(A)$ with its associated set $\{x \in \mathbb{C}^n : \langle Ax, x\rangle = a\|x\|^2\}$ a one-dimensional subspace of $\mathbb{C}^n$.*

THEOREM 2.3. *Let A be an S_n-matrix with the $(n+1)$-by-$(n+1)$ unitary dilation $U = \mathrm{diag}(\lambda_1, \ldots, \lambda_{n+1})$, where the λ_j's are arranged counterclockwise around $\partial\mathbb{D}$. Let $a_j = t_j\lambda_j + (1-t_j)\lambda_{j+1}$, where $0 < t_j < 1$, be the tangent point of the edge $[\lambda_j, \lambda_{j+1}]$ of the $(n+1)$-gon $\lambda_1 \ldots \lambda_{n+1}$ with the boundary $\partial W(A)$ of $W(A)$, $1 \leqslant j \leqslant n+1$ ($\lambda_{n+2} \equiv \lambda_1$), and let*

$$x_j = [0 \ \ldots \ 0 \ \underset{jth}{\sqrt{t_j}} \ \sqrt{1-t_j} \ 0 \ \ldots \ 0]^T$$

for $1 \leqslant j \leqslant n$, *and* $x_{n+1} = [\sqrt{1-t_{n+1}}\ 0\ \ldots\ 0\ \sqrt{t_{n+1}}]^T$ *in* $\mathbb{C}^{n+1}$. *If* V *is the* $(n+1)$*-by-*n *matrix* $[x_1\ \ldots\ x_n]$ *and* $A' = V^*UV$, *then* A' *is of class* S_n *and is unitarily equivalent to* A.

In the next two sections, we use these results to prove certain properties relating the numerical ranges and compressions of S_n-matrices.

3. Compression of S_n-matrix

We start with a general criterion for a compression B of a matrix A to have $W(B)$ properly contained in (resp., contained in the interior of) $W(A)$. Note that if B is a compression of A, then $W(B)$ is always contained in $W(A)$.

For an n-by-n matrix A and a point a in $\partial W(A)$, let K_a denote the set $\{x \in \mathbb{C}^n : \langle Ax, x\rangle = a\|x\|^2\}$ and let H_a be the subspace generated by K_a. It is known from [5, Theorem 1] that (1) a is an extreme point of the convex set $W(A)$ if and only if K_a is a subspace of $\mathbb{C}^n$, and (2) if a is not an extreme point of $W(A)$, then $H_a = \cup\{K_b : b \in L\}$, where L is the supporting line of $W(A)$ at a.

LEMMA 3.1. *Let A be an n-by-n matrix.*

(a) If $\dim H_a = 1$ *for all* a *in* $\partial W(A)$ *and* $\mathbb{C}^n$ *is generated by* $\cup\{K_a : a \in \partial W(A)\}$, *then* $W(B) \subsetneqq W(A)$ *for any* k*-by-*k ($1 \leqslant k < n$) *compression* B *of* A.

(b) If, for all a *in* $\partial W(A)$, *the vectors in* H_a *have only nonzero components, then* $\partial W(B) \cap \partial W(A) = \emptyset$ *for any* k*-by-*k ($1 \leqslant k < n$) *principal submatrix* B *of* A.

For any n-by-n matrix A, let A_j, $1 \leqslant j \leqslant n$, denote the $(n-1)$-by-$(n-1)$ principal submatrix of A obtained by deleting the jth row and jth column of A.

Proof of Lemma 3.1. (a) Let B be a k-by-k ($1 \leqslant k < n$) compression of A with $W(B) = W(A)$, and let V be an n-by-k matrix with $V^*V = I_k$ and $B = V^*AV$. For any a in $\partial W(B)$, let x be a unit vector in $\mathbb{C}^k$ such that $\langle Bx, x\rangle = a$. Then

$$a = \langle Bx, x\rangle = \langle V^*AVx, x\rangle = \langle A(Vx), Vx\rangle.$$

This shows that Vx is a unit vector in H_a. Since the latter is of dimension one, it consists of scalar multiples of Vx. It then follows from our assumption that $\mathbb{C}^n$ is generated by all the Vx's with $\langle Bx, x\rangle \in \partial W(B)$. This is absurd since the latter space is of dimension at most k. Our assertion that $W(B) \subsetneqq W(A)$ follows.

(b) Let $B = A_j$, $1 \leqslant j \leqslant n$, be such that $\partial W(B) \cap \partial W(A) \neq \emptyset$, and let $a = \langle Bx, x\rangle$ be in $\partial W(B) \cap \partial W(A)$, where $x = [x_1\ \ldots\ x_{n-1}]^T$ is a unit vector in $\mathbb{C}^{n-1}$. Since $a = \langle Ax', x'\rangle$, where $x' = [x'_1\ \ldots\ x'_n]^T$ is such that

$$x'_i = \begin{cases} x_i & \text{if } 1 \leqslant i < j, \\ 0 & \text{if } i = j, \\ x_{i-1} & \text{if } j < i \leqslant n, \end{cases}$$

x' is a unit vector in H_a. This contradicts our assumption on the nonzeroness of the components of x'. Hence $\partial W(B) \cap \partial W(A) = \emptyset$ as asserted. □

We also need the following lemma.

LEMMA 3.2. *If λ_1 and λ_2 are two scalars such that $|\lambda_1| \leqslant |\lambda_2|$, then*

$$W\left(\begin{bmatrix} A & \lambda_1 B \\ 0 & C \end{bmatrix}\right) \subseteq W\left(\begin{bmatrix} A & \lambda_2 B \\ 0 & C \end{bmatrix}\right).$$

Proof. We may assume that $\lambda_2 \neq 0$. If $\lambda = \lambda_1/\lambda_2$, then $|\lambda| \leqslant 1$ and we need only check that

$$W\left(\begin{bmatrix} A & \lambda B \\ 0 & C \end{bmatrix}\right) \subseteq W\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right).$$

Since

$$\begin{bmatrix} A & \lambda B \\ 0 & C \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & |\lambda| B \\ 0 & C \end{bmatrix}$$

are unitarily equivalent, we may further assume that $0 \leqslant \lambda \leqslant 1$. Thus

$$\begin{aligned} W\left(\begin{bmatrix} A & \lambda B \\ 0 & C \end{bmatrix}\right) &= W\left(\lambda \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} + (1-\lambda)\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}\right) \\ &\subseteq \lambda W\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) + (1-\lambda) W\left(\begin{bmatrix} A & 0 \\ 0 & C \end{bmatrix}\right) \\ &\subseteq \lambda W\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) + (1-\lambda) W\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right) \\ &\subseteq W\left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix}\right), \end{aligned}$$

completing the proof. □

Using these, we are able to prove the main result of this section.

THEOREM 3.3. *Let A be an S_n-matrix ($n \geqslant 2$).*
(a) If B is a k-by-k ($1 \leqslant k < n$) compression of A, then $W(B) \subsetneqq W(A)$.
(b) If $A = [a_{ij}]_{i,j=1}^n$ is of the standard upper-triangular form in Theorem 2.1, and B is a k-by-k ($1 \leqslant k < n$) principal submatrix of A, then $\partial W(B) \cap \partial W(A) = \emptyset$.

Note that assertions (a) and (b) above are comparable to the result that if A is an S_n-matrix, K is a proper invariant subspace of A, and $B = A|K$, then $\partial W(B) \cap \partial W(A) = \emptyset$ (cf. [7, Corollary 3.4] or [1, Proposition 2 (1)]).

Proof of Theorem 3.3. (a) Let U, λ_j's, a_j's, t_j's, x_j's, V and A' be as in Theorem 2.3, and let $x_j' = V^* x_j$ for $1 \leqslant j \leqslant n$. Then

$$a_j = \langle U x_j, x_j \rangle = \langle A'(V^* x_j), V^* x_j \rangle = \langle A' x_j', x_j' \rangle$$

for each j. This shows that x_j' is a unit vector in $H_{a_j} = \{x \in \mathbb{C}^n : \langle A' x, x \rangle = a_j \|x\|^2\}$. Since the x_j''s are linearly independent, we obtain that $\mathbb{C}^n$ is generated by $\cup\{H_a : a \in \partial W(A)\}$. This, together with the fact from Theorem 2.2 that $\dim H_a = 1$ for all a in $\partial W(A)$, yields, via Lemma 3.1 (a), that $W(B) \subsetneqq W(A)$.

(b) We may assume that $k = n-1$ and $B = A_j$ $(1 \leqslant j \leqslant n)$. If $j = n$, then B is the restriction of A to its invariant subspace $\mathbb{C}^{n-1} \oplus \{0\}$ and thus $\partial W(B) \cap \partial W(A) = \emptyset$ follows from [7, Corollary 3.4]. Applying this to the S_n-matrix A^* and its restriction to the invariant subspace $\{0\} \oplus \mathbb{C}^{n-1}$ yields our assertion for $B = A_1$. Thus we need only consider for $B = A_j$, $2 \leqslant j \leqslant n-1$. For this, we express B as

$$\begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$

where B_{11} and B_{22} are of sizes $j-1$ and $n-j$, respectively. If $a_{jj} = 0$, then $B_{12} = 0$ and hence $B = B_{11} \oplus B_{22}$. By [7, Corollary 3.4] again, we have $\partial W(B_{ii}) \cap \partial W(A) = \emptyset$ for $i = 1,2$. Since $W(B) = (W(B_{11}) \cup W(B_{22}))^\wedge$, the convex hull of $W(B_{11})$ and $W(B_{22})$, it follows that $\partial W(B) \cap \partial W(A) = \emptyset$ as asserted. On the other hand, for $a_{jj} \neq 0$, let $B'_{12} = (-1/\overline{a}_{jj})B_{12}$ and

$$B' = \begin{bmatrix} B_{11} & B'_{12} \\ 0 & B_{22} \end{bmatrix}.$$

Since $|-1/\overline{a}_{jj}| > 1$, we have $W(B) \subseteqq W(B')$ by Lemma 3.2. Note that if A' denotes the S_n-matrix $[a'_{ij}]_{i,j=1}^n$, where

$$a'_{ii} = \begin{cases} a_{ii} & \text{if } 1 \leqslant i \leqslant j-1, \\ a_{i+1\,i+1} & \text{if } j \leqslant i \leqslant n-1, \\ a_{jj} & \text{if } i = n, \end{cases}$$

and

$$a'_{ij} = \begin{cases} (1-|a'_{ii}|^2)^{1/2}(1-|a'_{jj}|^2)^{1/2}\prod_{k=i+1}^{j-1}(-\overline{a}'_{kk}) & \text{if } j > i, \\ 0 & \text{if } j < i \end{cases}$$

(cf. Theorem 2.1), then $B' = A'_n$. From what was proven before, we have $\partial W(B') \cap \partial W(A') = \emptyset$. Since A' and A are both of class S_n and have the same eigenvalues, they are unitarily equivalent. Thus $W(A') = W(A)$. Combining these together, we obtain $\partial W(B) \cap \partial W(A) = \emptyset$ as asserted. □

COROLLARY 3.4. *If $A = [a_{ij}]_{i,j=1}^n$ is an S_n-matrix of the standard upper-triangular form in Theorem 2.1, and $x = [x_1 \ \dots \ x_n]^T$ is a unit vector in $\mathbb{C}^n$ such that $\langle Ax, x\rangle \in \partial W(A)$, then $x_j \neq 0$ for all j.*

Proof. Assume that $x_j = 0$ for some j, $1 \leqslant j \leqslant n$. Let A_j be the jth $(n-1)$-by-$(n-1)$ principal submatrix of A. Then $x' \equiv [x_1 \ \dots \ x_{j-1}\ x_{j+1} \ \dots \ x_n]^T$ is a unit vector in $\mathbb{C}^{n-1}$ such that $\langle A_j x', x'\rangle = \langle Ax, x\rangle$ is in $\partial W(A)$. This shows that $\partial W(A_j) \cap \partial W(A) \neq \emptyset$, contradicting Theorem 3.3 (b). Hence we must have $x_j \neq 0$ for all j. □

In case $A = J_n$, the n-by-n Jordan block

$$\begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix},$$

it is known that $W(A) = \{z \in \mathbb{C} : |z| \leqslant \cos(\pi/(n+1))\}$ and a unit vector x in $\mathbb{C}^n$ is such that $\langle Ax, x\rangle$ belongs to $\partial W(A)$ if and only if

$$x = \lambda\sqrt{\frac{2}{n+1}}\left[e^{i\theta}\sin\left(\frac{\pi}{n+1}\right) \;\; e^{2i\theta}\sin\left(\frac{2\pi}{n+1}\right) \;\; \ldots \;\; e^{ni\theta}\sin\left(\frac{n\pi}{n+1}\right)\right]^T$$

for some scalar λ with $|\lambda| = 1$ and some real θ (cf. [13, Proposition 1 (3)] or [22, p. 134, Lemma 7]). This confirms the assertion in Corollary 3.4 for this case.

4. Diagonal of S_n-matrix

For any n-by-n matrix A, let $k(A)$ denote the maximum size k of a compression B of A for which the diagonal eatries of B are all in $\partial W(A)$. Equivalently, $k(A)$ equals the maximum number of orthonormal vectors $x_1, \ldots, x_k$ in $\mathbb{C}^n$ for which $\langle Ax_j, x_j\rangle$ is in $\partial W(A)$ for all j. It is obvious that $1 \leqslant k(A) \leqslant n$. The next lemma says that we even have $k(A) \geqslant 2$ for $n \geqslant 2$.

For any matrix A or scalar a, $\mathrm{Re}A$ (resp., $\mathrm{Re}\,a$) denotes its *real part* $(A+A^*)/2$ (resp., $(a+\overline{a})/2$).

LEMMA 4.1. *If A is an n-by-n matrix with $n \geqslant 2$, then $2 \leqslant k(A) \leqslant n$.*

Proof. For any point a in $\partial W(A)$, let L_a be a supporting line of $W(A)$ which passes a, and R_a be the ray from the origin which is perpendicular to L_a. If $\theta \in [0, 2\pi)$ is the angle from the positive x-axis to R_a and x is a unit vector in $\mathbb{C}^n$ with $\langle Ax, x\rangle = a$, then it is easily seen that $\mathrm{Re}\,(e^{-i\theta}a)$ is the maximum eigenvalue of the Hermitian matrix $\mathrm{Re}\,(e^{-i\theta}A)$ with the eigenvector x. Let L_b be the supporting line of $W(A)$ which is parallel to L_a and which passes a point, say, b in $\partial W(A)$ (cf. Figure 4.2). If y is a unit vector in $\mathbb{C}^n$ such that $\langle Ay, y\rangle = b$, then, similarly, $\mathrm{Re}\,(e^{-i\theta}b)$ is the minimum eigenvalue of $\mathrm{Re}\,(e^{-i\theta}A)$ with the eigenvector y. Thus, in case $\mathrm{Re}\,(e^{-i\theta}a) \neq \mathrm{Re}\,(e^{-i\theta}b)$, x and y are orthogonal to each other and, therefore, $k(A) \geqslant 2$. On the other hand, if $\mathrm{Re}\,(e^{-i\theta}a) = \mathrm{Re}\,(e^{-i\theta}b)$, then $W(e^{-i\theta}A)$ is a line segment. In this case, A is a normal matrix and thus $k(A) \geqslant 2$ obviously. This completes the proof. □

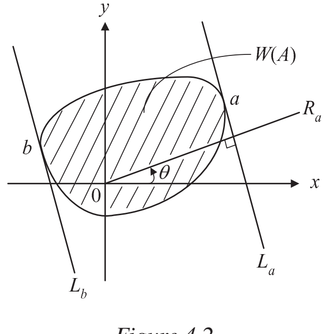

Figure 4.2

An easy corollary of the above is that $k(A) = 2$ for any 2-by-2 matrix A. The next proposition gives more precise information on a 2-by-2 A with diagonal entries in $\partial W(A)$.

PROPOSITION 4.3. *The following conditions are equivalent for* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:
(a) $a \in \partial W(A)$,
(b) $be^{-i\theta} + \overline{c}e^{i\theta} = 0$ *for some real* θ,
(c) $|b| = |c|$, *and*
(d) $d \in \partial W(A)$.
Under these conditions, if A is not normal, then the tangent lines to the (nondegenerate) ellipse $\partial W(A)$ at a and d are parallel to each other while if A is normal and $W(A)$ equals the line segment $[a,d]$, then $b = c = 0$.

This appeared in [23] as a consequence of a more general result [23, Theorem 2]. Here we give a simple computational proof.

Proof of Proposition 4.3. (a) $\Rightarrow$ (b). Considering $A - ((a+d)/2)I_2$ instead of A, we may assume that $A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$. As in the proof of Lemma 4.1, $a \in \partial W(A)$ implies that, for some real θ, $\operatorname{Re}(e^{-i\theta}a)$ is the maximum eigenvalue of $\operatorname{Re}(e^{-i\theta}A)$. Since a simple computation yields that the eigenvalues of $\operatorname{Re}(e^{-i\theta}A)$ are

$$\pm\frac{1}{2}(4(\operatorname{Re}(e^{-i\theta}a))^2 + |be^{-i\theta} + \overline{c}e^{i\theta}|^2)^{1/2},$$

we have

$$\operatorname{Re}(e^{-i\theta}a) = \frac{1}{2}(4(\operatorname{Re}(e^{-i\theta}a))^2 + |be^{-i\theta} + \overline{c}e^{i\theta}|^2)^{1/2},$$

from which we obtain $be^{-i\theta} + \overline{c}e^{i\theta} = 0$ as required.

(b) $\Rightarrow$ (c). Trivial.

(c) $\Rightarrow$ (a). As in the proof above, we may assume that $A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ with $|b| = |c|$. Since A is unitarily equivalent to

$$\begin{bmatrix} \sqrt{a^2+bc} & \alpha \\ 0 & -\sqrt{a^2+bc} \end{bmatrix},$$

where $|\alpha| = \sqrt{2}(|a|^2 + |b|^2 - |a^2+bc|)^{1/2}$, and $\partial W(A)$ is the ellipse with foci $\pm\sqrt{a^2+bc}$ and minor axis of length $|\alpha|$, to prove $a \in \partial W(A)$ we need only check that

$$|a - \sqrt{a^2+bc}| + |a + \sqrt{a^2+bc}| = 2(|a^2+bc| + \frac{1}{4}|\alpha|^2)^{1/2}.$$

Since the square of the left-hand side (resp., right-hand side) of the above equality equals

$$\begin{aligned} &|a - \sqrt{a^2+bc}|^2 + |a + \sqrt{a^2+bc}|^2 + 2|a^2 - (a^2+bc)| \\ =\ &2(|a|^2 + |a^2+bc|) + 2|bc| \end{aligned}$$

(resp.,

$$\begin{aligned}&4(|a^2+bc|+\frac{1}{4}|\alpha|^2)\\&=4|a^2+bc|+2(|a|^2+|b|^2-|a^2+bc|)\\&=2(|a^2+bc|+|a|^2+|b|^2)),\end{aligned}$$

they are indeed equal.

(d) $\Leftrightarrow$ (c). This follows by symmetry of the equivalence of (a) and (c).

Assume that these conditions hold. If A is not normal. then a and d form a chord of the (nondegenerate) ellipse $\partial W(A)$ which passes its center and hence their tangent lines are parallel. On the other hand, if A is normal and $W(A)=[a,d]$, then a and d are eigenvalues of A. Thus $a,d=(a+d\pm((a+d)^2-4(ad-bc))^{1/2})/2$. A simple computation then yields $bc=0$ and hence $b=c=0$. □

We remark that the implication (c) $\Rightarrow$ (b) in the preceding proposition can be interpreted geometrically as saying that if $b=|b|e^{i\theta_1}$ and $c=|c|e^{i\theta_2}$ (θ_1 and θ_2 real) are such that $|b|=|c|$, then the rotation around the origin by the angle $-(\theta_1+\theta_2+\pi)/2$ transforms b and c to the symmetric, relative to the y-axis, points $|b|e^{i(\theta_1-\theta_2-\pi)/2}$ and $-|c|e^{-i(\theta_1-\theta_2-\pi)/2}$, respectively.

The main result of this section is the following theorem, which determines $k(A)$ for any S_n-matrix A.

THEOREM 4.4. *If A is a matrix of class S_n ($n\geqslant 3$), then $k(A)=\lceil n/2\rceil$, the ceiling of $n/2$. Moreover, if $k=k(A)$, then there are $a_1,\ldots,a_k$ in $\partial W(A)$ such that* $\operatorname{diag}(a_1,\ldots,a_k)$ *dilates to A.*

To prove this theorem, we need the following two lemmas, the first of which is an improvement of [12, Corollary 4.2].

LEMMA 4.5. *Let A be an S_n-matrix ($n\geqslant 2$), $U=\operatorname{diag}(\lambda_1,\ldots,\lambda_{n+1})$ be an $(n+1)$-by-$(n+1)$ unitary dilation of A with the λ_j's arranged counterclockwise around $\partial\mathbb{D}$, and $a_j=t_j\lambda_j+(1-t_j)\lambda_{j+1}$, $0<t_j<1$, be the tangent point of the edge $[\lambda_j,\lambda_{j+1}]$ of the $(n+1)$-gon $\lambda_1\ldots\lambda_{n+1}$ with $\partial W(A)$, $1\leqslant j\leqslant n+1$ ($\lambda_{n+2}\equiv\lambda_1$). If*

$$x_j=[0\ \ldots\ 0\ \underset{jth}{\sqrt{t_j}}\ \sqrt{1-t_j}\ 0\ \ldots\ 0]^T$$

for $2\leqslant j\leqslant n$, V is the $(n+1)$-by-$(n-1)$ matrix $[x_2\ \ldots\ x_n]$, and $B=V^*UV$, then B is an S_{n-1}-matrix and $\partial W(B)\cap\partial W(A)=\{a_2,\ldots,a_n\}$.

Proof. That B is of class S_{n-1} follows from [12, Corollary 4.2]. For the proof of $\partial W(B)\cap\partial W(A)=\{a_2,\ldots,a_n\}$, note that one containment is trivial. To prove the other, assume that a is in $\partial W(B)\cap\partial W(A)$ is such that $a\neq a_j$ for all j, $2\leqslant j\leqslant n$. Let L be the common supporting line of $W(A)$ and $W(B)$ at the point a, and let z_1 and z_2 be the intersection points of L and $\partial\mathbb{D}$. Since $a\neq a_j$ for $2\leqslant j\leqslant n$ by our assumption, z_1 and z_2 are distinct from all the λ_j's. Let $\alpha_1,\ldots,\alpha_n$ (resp.,

$\beta_1,\ldots,\beta_{n-1}$) be the eigenvalues of A (resp., B), and let $\phi(z)=\prod_{j=1}^{n}(z-\alpha_j)/(1-\overline{\alpha}_j z)$ (resp., $\psi(z)=\prod_{j=1}^{n-1}(z-\beta_j)/(1-\overline{\beta}_j z)$) be the corresponding finite Blaschke product. Applying condition (6) of [10, Theorem 2.1] to A (resp., B) yields that

$$\frac{\phi(z)}{z\phi(z)-(-1)^n\prod_{j=1}^{n}\lambda_j}=\sum_{j=1}^{n+1}\frac{m_j}{z-\lambda_j},\quad z\neq\lambda_1,\ldots,\lambda_{n+1}, \tag{1}$$

$$\left(\text{resp.,}\ \frac{\psi(z)}{z\psi(z)-(-1)^{n-1}\prod_{j=2}^{n}\lambda_j}=\sum_{j=2}^{n+1}\frac{m_j'}{z-\lambda_j},\quad z\neq\lambda_2,\ldots,\lambda_{n+1}\right), \tag{2}$$

where

$$m_j=\frac{t_1\cdots t_{j-1}(1-t_j)\cdots(1-t_n)}{\sum_{k=1}^{n+1}t_1\cdots t_{k-1}(1-t_k)\cdots(1-t_n)},\quad 1\leqslant j\leqslant n+1$$

$$\left(\text{resp.,}\ m_j'=\frac{t_2\cdots t_{j-1}(1-t_j)\cdots(1-t_n)}{\sum_{k=2}^{n+1}t_2\cdots t_{k-1}(1-t_k)\cdots(1-t_n)},\quad 2\leqslant j\leqslant n+1\right).$$

Since $m_j'=m_j/(1-m_1)$ for $2\leqslant j\leqslant n+1$, we may combine (1) and (2) together to obtain

$$\frac{\phi(z)}{z\phi(z)-(-1)^n\prod_{j=1}^{n}\lambda_j}=\frac{m_1}{z-\lambda_1}+(1-m_1)\frac{\psi(z)}{z\psi(z)-(-1)^{n-1}\prod_{j=2}^{n}\lambda_j} \tag{3}$$

for $z\neq\lambda_1,\ldots,\lambda_{n+1}$. Plugging z_1 and z_2 into (3), dividing the resulting equalities, and noting that $z_1\phi(z_1)=z_2\phi(z_2)$ (by [10, Theorem 2.1 (8)]) yield that

$$\frac{z_2}{z_1}=\frac{\phi(z_1)}{\phi(z_2)}=\frac{\frac{m_1}{z_1-\lambda_1}+(1-m_1)\frac{\psi(z_1)}{z_1\psi(z_1)-(-1)^{n-1}\prod_j\lambda_j}}{\frac{m_1}{z_2-\lambda_1}+(1-m_1)\frac{\psi(z_2)}{z_2\psi(z_2)-(-1)^{n-1}\prod_j\lambda_j}}.$$

Since $z_1\psi(z_1)=z_2\psi(z_2)$ by [10, Theorem 2.1 (8)] and $m_1>0$, we cross-multiply the above fractions to obtain, after cancellation, $z_1(z_2-\lambda_1)=z_2(z_1-\lambda_1)$ or $z_1=z_2$, which contradicts our assumption. Thus a must be equal to some a_j, $2\leqslant j\leqslant n$, completing the proof. □

We remark that the key equality (3) above was first proven in [3, Theorem 6 (a)].

LEMMA 4.6. *Let A be an S_n-matrix ($n\geqslant 2$), a be a point in $\partial W(A)$ and x be a unit vector in $\mathbb{C}^n$ such that $\langle Ax,x\rangle=a$. Let $U=\operatorname{diag}(\lambda_1,\ldots,\lambda_{n+1})$ be the $(n+1)$-by-$(n+1)$ unitary dilation of A with the λ_j's arranged counterclockwise around $\partial\mathbb{D}$ such that the edge $[\lambda_1,\lambda_2]$ contains a, and let a_j, $1\leqslant j\leqslant n+1$, be the tangent point of $[\lambda_j,\lambda_{j+1}]$ with $\partial W(A)$ ($\lambda_{n+2}\equiv\lambda_1$ and $a_1=a$). Then a point b in $\partial W(A)$ is such that the unit vector y with $\langle Ay,y\rangle=b$ is orthogonal to x if and only if b is either*

(a) the tangent point a' of the supporting line, parallel to $[\lambda_1,\lambda_2]$, of $W(A)$ with $\partial W(A)$, or

(b) one of the tangent points $a_3,\ldots,a_n$.

Note that in the case of $A=J_n$, for any point a in $\partial W(J_n)$, the number of points b with the asserted property is $n-1$ or $n-2$ depending on whether n is even or odd.

This is because, in this case, the eigenvalues $\lambda_1,\ldots,\lambda_{n+1}$ of the unitary dilation U form a regular $(n+1)$-gon, and thus a' is distinct from $a_3,\ldots,a_n$ for n even while $a' = a_{(n+3)/2}$ for n odd.

Proof of Lemma 4.6. Let t_j's, x_j's, V and A' be as in Theorem 2.3. Since A' and A are unitarily equivalent, we may assume that $A = A'$ and $x = V^*x_1$. To prove one direction, if $b = a'$, then the orthogonality of y and x follows as in the proof of Lemma 4.1. On the other hand, if $b = a_j$ for some j, $3 \leqslant j \leqslant n$, then, since x_1 and x_j are obviously orthogonal, the same is true for V^*x_1 and V^*x_j or x and y.

For the other direction, let $b = \langle Ay,y\rangle$ in $\partial W(A)$ with $\|y\| = 1$ and $\langle x,y\rangle = 0$, and let L be the supporting line of $W(A)$ which passes b. Assume that $b \neq a'$. Then L is not parallel to the edge $[\lambda_1,\lambda_2]$ of the $(n+1)$-gon $\lambda_1\ldots\lambda_{n+1}$. Let M be the n-by-2 matrix $[x\ y]$ and let $B = M^*AM$. Then B is of the form $\begin{bmatrix} a & * \\ * & b \end{bmatrix}$. If $W(B)$ is a (nondegenerate) elliptic disc, then $[\lambda_1,\lambda_2]$ and L, being tangent lines of the ellipse $\partial W(B)$ at a and b, respectively, are parallel by Proposition 4.3, contradicting our assumption. Thus $B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ by Proposition 4.3 again. This implies that

$$\langle BM^*x, M^*y\rangle = \left\langle B\begin{bmatrix} x^* \\ y^* \end{bmatrix} x, \begin{bmatrix} x^* \\ y^* \end{bmatrix} y \right\rangle = \left\langle B\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\rangle = 0.$$

Since

$$MM^*x = [x\ y]\begin{bmatrix} x^* \\ y^* \end{bmatrix} x = [x\ y]\begin{bmatrix} 1 \\ 0 \end{bmatrix} = x$$

and, similarly, $MM^*y = y$, we obtain

$$\langle Ax,y\rangle = \langle AMM^*x, MM^*y\rangle = \langle MM^*AMM^*x, y\rangle = \langle BM^*x, M^*y\rangle = 0.$$

In a similar fashion, $x' \equiv Vx$ and $y' \equiv Vy$ are orthonormal vectors in $\mathbb{C}^{n+1}$ satisfying $\langle Ux',x'\rangle = a = a_1$, $\langle Uy',y'\rangle = b$ and $\langle Ux',y'\rangle = 0$. Since $\langle Ax,x\rangle = a$, we may assume that $x' = x_1 = [\sqrt{t_1}\ \sqrt{1-t_1}\ 0\ \ldots\ 0]^T$. If $y' = [y_1\ y_2\ \ldots\ y_{n+1}]^T$, then we obtain from $\langle x',y'\rangle = \langle Ux',y'\rangle = 0$ that

$$y_1\sqrt{t_1} + y_2\sqrt{1-t_1} = 0 \tag{4}$$

and

$$y_1\overline{\lambda}_1\sqrt{t_1} + y_2\overline{\lambda}_2\sqrt{1-t_1} = 0. \tag{5}$$

Multiplying (4) by $\overline{\lambda}_1$ and then subtracting (5) from it result in $(\overline{\lambda}_1 - \overline{\lambda}_2)y_2\sqrt{1-t_1} = 0$ or $y_2 = 0$ (since $\lambda_1 \neq \lambda_2$ and $\sqrt{1-t_1} > 0$). From (4), we then also have $y_1 = 0$ (since $\sqrt{t_1} > 0$). Thus $y' = [0\ 0\ y_3\ \ldots\ y_{n+1}]^T$. Let N_1 be the $(n+1)$-by-$(n-1)$ matrix $[x_2\ \ldots\ x_n]$, and let $C_1 = N_1^*UN_1$. Then C_1 is an S_{n-1}-compression of A by Theorem 2.3. Note that $b = \langle Uy',y'\rangle = \langle C_1N_1^*y', N_1^*y'\rangle$, showing that b is in $\partial W(C_1) \cap \partial W(A)$. Hence, by Lemma 4.5, b is equal to some a_{j_1}, $2 \leqslant j_1 \leqslant n$. Analogously, considering $N_2 = [x_3\ \ldots\ x_{n+1}]$ and $C_2 = N_2^*UN_2$, we also obtain $b = a_{j_2}$ for some j_2, $3 \leqslant j_2 \leqslant n+1$. Thus b equals one of the points $a_3,\ldots,a_n$, completing the proof. □

Now we are finally ready to prove Theorem 4.4.

Proof of Theorem 4.4. Let $k = k(A)$ and let B be a k-by-k compression of A with diagonal entries $b_1, \ldots, b_k$ in $\partial W(A)$. According to Lemma 4.6, we may assume that $b_1 = a_1 = a$ and $\{b_2, \ldots, b_k\}$ is a subset of $\{a', a_3, \ldots, a_n\}$. In the following, we show that the latter b_j's can all be chosen from $a_3, \ldots, a_n$.

Indeed, following the proof of Lemma 4.6, let $x'_j = V^* x_j$ for $3 \leqslant j \leqslant n$. Then x'_j is a unit vector in $\mathbb{C}^n$ satisfying $\langle Ax'_j, x'_j\rangle = a_j$ for each j. Also, let $a' = \langle Au', u'\rangle$ for some unit vector in $\mathbb{C}^n$ and let $u = Vu'$. Since the supporting lines of $W(A)$ at b_1 $(= a_1)$ and a' are parallel, a' is inbetween, say, a_j and a_{j+1} for some j, $2 \leqslant j \leqslant n$, around $\partial W(A)$. If Q is the $(n+1)$-by-3 matrix $[x_j\ x_{j+1}\ x_{j+2}]$ $(x_{n+2} \equiv x_1)$ and $D = Q^* U Q$, then a' is in the numerical range of D and hence u is of the form $[0\ \ldots\ 0\ u_j\ u_{j+1}\ u_{j+2}\ 0\ \ldots\ 0]^T$. Assume first that j is odd. If u is orthogonal to x_{j+2}, then $u_{j+2} = 0$. Hence

$$a' = \langle Au', u'\rangle = \langle Uu, u\rangle = \lambda_j |u_j|^2 + \lambda_{j+1} |u_{j+1}|^2$$

with $|u_j|^2 + |u_{j+1}|^2 = \|u\|^2 = 1$, which shows that a' lies on the edge $[\lambda_j, \lambda_{j+1}]$. Since a' is also on $\partial W(A)$, it thus coincides with a_j. Similarly, if j is even, then we can also deduce from the orthogonality of u and x_{j-1} that $a' = a_{j+1}$. This means that a' and either of its neighboring a_j and a_{j+1} cannot both belong to the set $\{b_1, \ldots, b_k\}$ unless they coincide. Hence to achieve the maximum value of k, we need only choose the b_j's to be $a_1, a_3, a_5, \ldots, a_n$ (resp., $a_1, a_3, a_5, \ldots, a_{n-1}$) if n is odd (resp., even). Therefore, $k(A) = \lceil n/2 \rceil$ as asserted. In this case, $\{x'_j : j = 1, 3, 5, \ldots, n$ (resp., $1, 3, 5, \ldots, n-1$)$\}$ is an orthonormal set of vectors in $\mathbb{C}^n$ with

$$\langle Ax'_j, x'_i\rangle = \langle Ux_j, x_i\rangle = \begin{cases} a_j & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases}$$

which shows that $\operatorname{diag}(a_1, a_3, \ldots, a_n)$ (resp., $\operatorname{diag}(a_1, a_3, \ldots, a_{n-1})$) dilates to A, completing the proof. $\square$

We end this paper by remarking that, in Theorem 4.4, the maximum value of k can be achieved by different choices of the diagonal entries $b_1, \ldots, b_k$ of B even when one of them, say, b_1 is fixed. For example, if $A = J_4$ and $b_1 = \cos(\pi/5)$, then $k(A) = 2$ and b_2 can be any one of $-\cos(\pi/5)$, $e^{i(4\pi/5)}\cos(\pi/5)$ and $e^{i(6\pi/5)}\cos(\pi/5)$.

Acknowledgement

The second author likes to thank his colleague Kuo-Zhong Wang for his inspiring weekly chats which lead to Lemma 3.1.

REFERENCES

[1] H. BERCOVICI, *Numerical ranges of operators of class C_0*, Linear Multilinear Algebra **50** (2002), 219–222.

[2] I. CHALENDAR, P. GORKIN AND J. R. PARTINGTON, *Numerical ranges of restricted shifts and unitary dilations*, Oper. Matrices **3** (2009), 271–281.

[3] U. DAEPP, P. GORKIN AND R. MORTINI, *Ellipses and finite Blaschke products*, Amer. Math. Monthly **109** (2002), 785–795.

[4] U. DAEPP, P. GORKIN AND K. VOSS, *Poncelet's theorem, Sendov's conjecture, and Blaschke products*, J. Math. Anal. Appl. **365** (2010), 93–102.
[5] M. R. EMBRY, *The numerical range of an operator*, Pacific J. Math. **32** (1970), 647–650.
[6] H.-L. GAU AND P. Y. WU, *Numerical range of $S(\phi)$*, Linear Multilinear Algebra **45** (1998), 49–73.
[7] H.-L. GAU AND P. Y. WU, *Dilation to inflations of $S(\phi)$*, Linear Multilinear Algebra **45** (1998), 109–123.
[8] H.-L. GAU AND P. Y. WU, *Lucas' theorem refined*, Linear Multilinear Algebra **45** (1999), 359–373.
[9] H.-L. GAU AND P. Y. WU, *Numerical range and Poncelet property*, Taiwanese J. Math. **7** (2003), 173–193.
[10] H.-L. GAU AND P. Y. WU, *Numerical range circumscribed by two polygons*, Linear Algebra Appl. **382** (2004), 155–170.
[11] H.-L. GAU AND P. Y. WU, *Numerical range of a normal compression*, Linear Multilinear Algebra **52** (2004), 195–201.
[12] H.-L. GAU AND P. Y. WU, *Numerical range of a normal compression II*, Linear Algebra Appl. **390** (2004), 121–136.
[13] U. HAAGERUP AND P. DE LA HARPE, *The numerical radius of a nilpotent operator on a Hilbert space*, Proc. Amer. Math. Soc. **115** (1992), 371–379.
[14] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
[15] B. MIRMAN, *Numerical ranges and Poncelet curves*, Linear Algebra Appl. **281** (1998), 59–85.
[16] B. MIRMAN, *UB-matrices and conditions for Poncelet polygon to be closed*, Linear Algebra Appl. **360** (2003), 123–150.
[17] B. MIRMAN, *Poncelet's porism in the finite real plane*, Linear Multilinear Algebra **57** (2009), 439–458.
[18] B. MIRMAN, V. BOROVIKOV, L. LADYZHENSKY AND R. VINOGRAD, *Numerical ranges, Poncelet curves, invariant measures*, Linear Algebra Appl. **329** (2001), 61–75.
[19] B. MIRMAN AND P. SHUKLA, *A characterization of complex plane Poncelet curves*, Linear Algebra Appl. **408** (2005), 86–119.
[20] D. SARASON, *Generalized interpolation in H^∞*, Trans. Amer. Math. Soc. **127** (1967), 179–203.
[21] P. Y. WU, *Polygons and numerical ranges*, Amer. Math. Monthly **107** (2000), 528–540.
[22] T. YOSHINO, *Introduction to Operator Theory*, Longman, Harlow, Essex, 1993.
[23] H.-Y. ZHANG, Y.-N. DOU, M.-F. WANG AND H.-K. DU, *On the boundary of numerical ranges of operators*, Appl. Math. Lett. **24** (2011), 620–622.

(Received March 1, 2012)

Hwa-Long Gau
Department of Mathematics
National Central University
Chungli 32001
Taiwan
e-mail: hlgau@math.ncu.edu.tw

Pei Yuan Wu
Department of Applied Mathematics
National Chiao Tung University
Hsinchu 30010
Taiwan
e-mail: pywu@math.nctu.edu.tw