Operators and Matrices
Volume 11, Number 4 (2017), 999–1014 doi:10.7153/oam-2017-11-69

LOWER BOUNDS FOR THE NUMERICAL RADIUS

HWA-LONG GAU AND PEI YUAN WU

(*Communicated by C.-K. Li*)

Abstract. We show that if $A=[a_{ij}]_{i,j=1}^n$ is an n-by-n complex matrix and $A'=[a'_{ij}]_{i,j=1}^n$, where

$$a'_{ij}=\begin{cases} a_{ij} & \text{if } (i,j)=(1,2),\ldots,(n-1,n) \text{ or } (n,1),\\ 0 & \text{otherwise,}\end{cases}$$

then $w(A)\geqslant w(A')$, where $w(\cdot)$ denotes the numerical radius of a matrix. Moreover, if n is odd and $a_{12},\ldots,a_{n-1,n},a_{n1}$ are all nonzero, then $w(A)=w(A')$ if and only if $A=A'$. For an even n, under the same nonzero assumption, we have $W(A)=W(A')$ if and only if $A=A'$, where $W(\cdot)$ is the numerical range of a matrix.

1. Introduction

Let $A=[a_{ij}]_{i,j=1}^n$ be an n-by-n complex matrix. The *numerical range* and *numerical radius* of A are $W(A)=\{\langle Ax,x\rangle : x\in\mathbb{C}^n, \|x\|=1\}$ and $w(A)=\max\{|z| : z\in W(A)\}$, respectively, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ denote the standard inner product and its associated norm of vectors in $\mathbb{C}^n$. In this paper, we obtain various lower bounds for the numerical radius of A. The primary one is $w(A)\geqslant w(A')$, where A' is the n-by-n matrix obtained from A by replacing all its entries other than $a_{12},\ldots,a_{n-1,n}$ and a_{n1} by zeros (cf. Proposition 3.1). We also consider when the equality $w(A)=w(A')$ holds. Under the assumptions of odd n and nonzero $a_{12},\ldots,a_{n-1,n}$ and a_{n1}, this is the case only when $A=A'$ (cf. Theorem 3.2). On the other hand, if n is even, then, under the same nonzero assumption, we need the stronger condition $W(A)=W(A')$ to guarantee the equality of A and A' (cf. Theorem 3.5). Another lower bound for $w(A)$ is $w(A'')$, where A'' is the matrix obtained from A' by replacing its $(n,1)$-entry ($=a_{n1}$) by zero. Again, we obtain conditions for the equality $w(A)=w(A'')$ (cf. Theorem 4.2).

In the following, we start in Section 2 with 2-by-2 matrices. The results therein motivate the later developments. Section 3 gives the lower bound $w(A')$ for $w(A)$ and discusses when this can be attained. In Section 4, we consider some special cases, generalizations and related consequences of the main results in Section 3, one of which is Theorem 4.2 that we mentioned above.

We use 0_n and I_n to denote the n-by-n zero matrix and identity matrix, respectively. For a square matrix A, we use $\mathrm{Re}A$ for its *real part* $(A+A^*)/2$. The column

Mathematics subject classification (2010): 15A60.
Keywords and phrases: Numerical range, numerical radius.
The first author was research supported in part by the Ministry of Science and Technology of the Republic of China under project MOST 105-2115-M-008-005.

vector with components $x_1,\dots,x_n$ is denoted by $[x_1 \ \dots \ x_n]^T$ and the diagonal matrix with diagonals $a_1,\dots,a_n$ by $\operatorname{diag}(a_1,\dots,a_n)$. If $A=[a_{ij}]_{i,j=1}^n$ and $B=[b_{ij}]_{i,j=1}^n$, their *Hadamard product* $A\circ B$ is $[a_{ij}b_{ij}]_{i,j=1}^n$. When A and B are real matrices, $A\preccurlyeq B$ means that $a_{ij}\leqslant b_{ij}$ for all i and j. We refer to [7] for other matrix notations and properties. Our reference for the numerical range and numerical radius is [6, Chapter 1].

2. 2-by-2 matrices

We start with the following preliminary result for 2-by-2 matrices. It lights the way forward to the general n-by-n case.

PROPOSITION 2.1. *Let* $A=\begin{bmatrix} a & b \\ c & d\end{bmatrix}$ *and* $A'=\begin{bmatrix} 0 & b \\ c & 0\end{bmatrix}$. *Then*

(a) $w(A)\geqslant w(A')$,

(b) $w(A)=w(A')$ *if and only if* $a+d=0$ *and* ad *is in* $[0,bc]$, *the line segment connecting* 0 *and* bc,

(c) $w(A)=w(A')$ *implies that* $W(A)\supseteq W(A')$, *and*

(d) $W(A)=W(A')$ *if and only if* $A=A'$.

The example of $A=\begin{bmatrix} i & 1 \\ 1 & -i\end{bmatrix}$ shows that (c) and (d) above cannot be further strengthened. Indeed, since A is unitarily similar to $\begin{bmatrix} 0 & 2 \\ 0 & 0\end{bmatrix}$, we have $W(A)=\{z\in\mathbb{C}: |z|\leqslant 1\}$. On the other hand, we can easily derive that $W(A')=W(\begin{bmatrix} 0 & 1 \\ 1 & 0\end{bmatrix})=[-1,1]$. Thus $A\neq A'$, $w(A)=w(A')=1$, and $W(A)\supsetneqq W(A')$. This shows that, in (c) (resp., (d)) of the preceding proposition, the conclusion $W(A)\supseteq W(A')$ (resp., the condition $W(A)=W(A')$) cannot be replaced by $W(A)=W(A')$ (resp., $w(A)=w(A')$).

Proof of Proposition 2.1. (a) Since $W(A)$ is an elliptic disc with center $(a+d)/2$, it is easily seen that $w(A)\geqslant w(B)$, where

$$B=A-\frac{1}{2}(a+d)I_2=\begin{bmatrix} \frac{1}{2}(a-d) & b \\ c & -\frac{1}{2}(a-d)\end{bmatrix}$$

(cf. [1, Lemma 5 (a)]). Note that a matrix C of the form $\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & -c_{11}\end{bmatrix}$ is unitarily similar to

$$\begin{bmatrix} (c_{11}^2+c_{12}c_{21})^{1/2} & (2|c_{11}|^2+|c_{12}|^2+|c_{21}|^2-2|c_{11}^2+c_{12}c_{21}|)^{1/2} \\ 0 & -(c_{11}^2+c_{12}c_{21})^{1/2}\end{bmatrix}$$

whence $w(C)$ can be computed to be

$$\frac{1}{2}(2|c_{11}|^2+|c_{12}|^2+|c_{21}|^2+2|c_{11}^2+c_{12}c_{21}|)^{1/2}. \tag{1}$$

Thus $w(B) = (1/2)\big((1/2)|a-d|^2 + |b|^2 + |c|^2 + (1/2)|(a-d)^2 + 4bc|\big)^{1/2}$ and $w(A') = (|b| + |c|)/2$. Therefore, $w(A) \geqslant w(B) \geqslant w(A')$ as asserted.

(b) From (a), we have $w(A) = w(A')$ if and only if $w(A) = w(B)$ and $w(B) = w(A')$. Since $(a+d)/2$ is the scalar approximant to A under $w(\cdot)$ by [1, Lemma 5 (a)], its uniqueness follows from the Loewner–Behrend theorem (cf. [2, Theorem 11.8.10.7]). Hence the equality $w(A) = w(B)$ is equivalent to $a + d = 0$. On the other hand, $w(B) = w(A')$ is equivalent to $|a-d|^2 + |(a-d)^2 + 4bc| = 4|bc|$, which is the same as $(a-d)^2 + 4bc = -t(a-d)^2$ for some $t \geqslant 0$ or $-(a-d)^2/4$ being in $[0, bc]$. The assertion in (b) then follows by combining these two conditions together.

(c) Under the assumption $w(A) = w(A')$, we obtain from the proof of (b) that $A = B$ and $a^2 + bc = sbc$ for some s, $0 \leqslant s \leqslant 1$. The latter condition yields that the foci $\pm(a^2+bc)^{1/2}$ of the elliptic disc $W(A)$ and the foci $\pm(bc)^{1/2}$ of $W(A')$ are on the same line passing through the origin. Moreover, their major axes are both of length $2w(A)$ and minor axes of lengths $2(|b|^2 + |c|^2 - 2|a^2 + bc|)^{1/2}$ and $2||b| - |c||$, respectively. Since

$$|b|^2 + |c|^2 - 2|a^2 + bc| = |b|^2 + |c|^2 - 2s|bc| \geqslant |b|^2 + |c|^2 - 2|bc| = ||b| - |c||^2,$$

we conclude that $W(A) \supseteq W(A')$.

(d) If $W(A) = W(A')$, then the coincidence of their foci yields from (c) that a is zero. Moreover, d is also zero from $a + d = 0$ (by (b)). Thus $A = A'$ as required. □

We remark that Proposition 2.1 (a) is also a special case of [3, Theorem 2.1].

We now seek along the line of the preceding proposition the largest lower bound for the numerical radius of a given 2-by-2 matrix. This is done via the next proposition.

PROPOSITION 2.2. *Let* $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. *Then*

(a) *A is unitarily similar to*

$$B \equiv \frac{1}{2}\begin{bmatrix} a+c & \big((|a-c|^2+|b|^2)^{1/2}+|b|\big)e^{i\theta} \\ \big((|a-c|^2+|b|^2)^{1/2}-|b|\big)e^{i\theta} & a+c \end{bmatrix},$$

where the real θ satisfies $a - c = |a-c|e^{i\theta}$,

(b) *the maximum value of $|x|$ for which a matrix of the form $\begin{bmatrix} * & x \\ * & * \end{bmatrix}$ is unitarily similar to A is $\big(|b| + (|a-c|^2+|b|^2)^{1/2}\big)/2$, which occurs when $\begin{bmatrix} * & x \\ * & * \end{bmatrix}$ equals B,*

(c) *the maximum value of $|x| + |y|$ for which a matrix of the form $\begin{bmatrix} * & x \\ y & * \end{bmatrix}$ is unitarily similar to A is $(|a-c|^2+|b|^2)^{1/2}$, which occurs when $\begin{bmatrix} * & x \\ y & * \end{bmatrix}$ equals B.*

Proof. (a) This follows by showing, via a simple computation, that A and B have equal traces, determinants and Frobenius norms.

(b) If $\begin{bmatrix} z & x \\ y & w \end{bmatrix}$ is unitarily similar to A, then $\begin{bmatrix} z-\lambda & x \\ y & w-\lambda \end{bmatrix}$ is unitarily similar to $A-\lambda I_2$ for any λ in $\mathbb{C}$. Hence

$$|x| \leqslant \min_{\lambda \in \mathbb{C}} \|A-\lambda I_2\| = \frac{1}{2}\left(|b| + (|a-c|^2+|b|^2)^{1/2}\right)$$

by [1, Lemma 5 (b)]. From (a), the inequality becomes an equality when $\begin{bmatrix} z & x \\ y & w \end{bmatrix}$ equals B. Our assertion follows.

(c) If $\begin{bmatrix} z & x \\ y & w \end{bmatrix}$ is unitarily similar to A, then $\begin{bmatrix} z-a & x \\ y & w-a \end{bmatrix}$ is unitarily similar to $\begin{bmatrix} 0 & b \\ 0 & c-a \end{bmatrix}$. From $|x|^2+|y|^2+|z-a|^2+|w-a|^2 = |b|^2+|c-a|^2$ and $\det \begin{bmatrix} z-a & x \\ y & w-a \end{bmatrix} = (z-a)(w-a)-xy = 0$, we obtain

$$\begin{aligned} (|x|+|y|)^2 &= |x|^2+|y|^2+2|xy| \\ &= (|a-c|^2+|b|^2-|z-a|^2-|w-a|^2)+2|(z-a)(w-a)| \\ &\leqslant |a-c|^2+|b|^2. \end{aligned}$$

Hence $|x|+|y| \leqslant (|a-c|^2+|b|^2)^{1/2}$. The equality is attained for $\begin{bmatrix} z & x \\ y & w \end{bmatrix} = B$ from (a). This proves (c). □

COROLLARY 2.3. *If* $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, *then* $\{x \in \mathbb{C} : \left[\begin{smallmatrix} * & x \\ * & * \end{smallmatrix}\right]$ *is unitarily similar to* $A\} = \{z \in \mathbb{C} : |z| \leqslant \left(|b| + (|a-c|^2+|b|^2)^{1/2}\right)/2\}$.

Proof. It was known that the set of x's for which $\left[\begin{smallmatrix} * & x \\ * & * \end{smallmatrix}\right]$ is unitarily similar to A is a closed circular disc centered at the origin (cf. [9, Theorem 4] or [6, p. 84, Exercise]). Our assertion follows from Proposition 2.2 (b). □

PROPOSITION 2.4. *If* $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, *then* $w(A) \geqslant (|a-c|^2+|b|^2)^{1/2}/2$. *Moreover, the equality holds if and only if* $a+c=0$.

Proof. Since A is unitarily similar to the matrix B in Proposition 2.2 (a), we have $w(A) = w(B) \geqslant w(B')$ by Proposition 2.1 (a), where

$$B' = \frac{1}{2}\begin{bmatrix} 0 & \left((|a-c|^2+|b|^2)^{1/2}+|b|\right)e^{i\theta} \\ \left((|a-c|^2+|b|^2)^{1/2}-|b|\right)e^{i\theta} & 0 \end{bmatrix}.$$

As $w(B') = (|a-c|^2+|b|^2)^{1/2}/2$ from (1), our first assertion follows.

From Proposition 2.1 (b), we have that $w(B) = w(B')$ if and only if $a+c=0$ and $(a+c)^2$ is in $[0, -|a-c|^2 e^{2i\theta}]$. Since the latter condition follows obviously from the former, we obtain the second assertion. □

There is a simpler geometric proof for Proposition 2.4. This is seen by noting that the elliptic disc $W(A)$ is contained in the circular disc $\{z \in \mathbb{C} : |z| \leqslant w(A)\}$. Hence the length $(|a-c|^2+|b|^2)^{1/2}$ of the major axis of the former is less than or equal to the diameter $2w(A)$ of the latter. Moreover, their equality is equivalent to the coincidence of their centers, that is, $(a+c)/2 = 0$ or $a+c=0$.

3. Lower bound and its attainment

We start by generalizing the inequality in Proposition 2.1 (a) to matrices of size n.

PROPOSITION 3.1. *If* $A = [a_{ij}]_{i,j=1}^{n}$ $(n \geqslant 2)$ *and*

$$A' = \begin{bmatrix} 0 & a_{12} & & & \\ & 0 & a_{23} & & \\ & & 0 & \ddots & \\ & & & \ddots & a_{n-1,n} \\ a_{n1} & & & & 0 \end{bmatrix},$$

then $w(A) \geqslant w(A')$.

Proof. Let U be the n-by-n unitary matrix

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

Since A' is equal to the Hadamard product $A \circ U$ of A and U, we have $w(A') = w(A \circ U) \leqslant w(A)\|U\| = w(A)$ (cf. [4, p. 293]). $\square$

Next we consider when the inequality $w(A) \geqslant w(A')$ in Proposition 3.1 becomes an equality. If the size of A is odd, then there is a satisfactory answer.

THEOREM 3.2. *Let A and A' be as in Proposition 3.1. If n is odd and $a_{12}, \ldots, a_{n-1,n}$ and a_{n1} are all nonzero, then the following conditions are equivalent:*

(a) $W(A) = W(A')$,

(b) $w(A) = w(A')$, *and*

(c) $A = A'$.

Note that the implication (b) $\Rightarrow$ (c) here is not valid for even n as the 2-by-2 matrix $A = \begin{bmatrix} i & 1 \\ 1 & -i \end{bmatrix}$ shows (cf. the paragraph after Proposition 2.1). If exactly one of the entries $a_{12}, \ldots, a_{n-1,n}$ and a_{n1} of A is zero, then the same conclusion holds irrespective of the parity of n. This will be proven in Theorem 4.2. However, it cannot be further relaxed to two zeros as seen by the n-by-n matrix $(n \geqslant 3)$ $A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus \operatorname{diag}(\underbrace{a, 0, \ldots, 0}_{n-2})$ with $0 < |a| \leqslant 1$, in which case $A' = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus 0_{n-2}$ and hence $w(A) = w(A') = 1$ but $A \neq A'$.

For the proof of Theorem 3.2, we need the following lemmas. The first one is a standard result from the nonnegative matrix theory.

LEMMA 3.3. *Let*

$$A=\begin{bmatrix} 0 & a_1 & & & \\ & 0 & a_2 & & \\ & & 0 & \ddots & \\ & & & \ddots & a_{n-1} \\ a_n & & & & 0 \end{bmatrix}$$

with $a_k \geqslant 0$ *for all* k.

(a) *There is a unit vector* $x=[x_1 \ \ldots \ x_n]^T$ *in* $\mathbb{C}^n$ *with* $x_k \geqslant 0$ *for all* k *such that* $\langle Ax,x\rangle = w(A)$. *Moreover, if* $a_k>0$ *for all* k, *then such an* x *is unique and* $x_k>0$ *for all* k.

(b) *If* $\omega_j=\exp(2\pi ij/n)$ *and* $x_{\omega_j}=[x_1 \ x_2\omega_j \ x_3\omega_j^2 \ \ldots \ x_n\omega_j^{n-1}]^T$ *for* $0\leqslant j\leqslant n-1$, *then* $\langle Ax_{\omega_j},x_{\omega_j}\rangle=\omega_j w(A)$ *for all* j.

Proof. (a) is a consequence of [8, Proposition 3.3]. For the proof of (b), we have

$$\begin{aligned}\langle Ax_{\omega_j},x_{\omega_j}\rangle &= \Big(\sum_{k=1}^{n-1} a_k(x_{k+1}\omega_j^k)(x_k\overline{\omega}_j^{k-1})\Big)+a_nx_1(x_n\overline{\omega}_j^{n-1})\\ &=\omega_j\Big(\big(\sum_{k=1}^{n-1} a_kx_{k+1}x_k\big)+a_nx_1x_n\Big)=\omega_j\langle Ax,x\rangle=\omega_j w(A). \quad \square\end{aligned}$$

LEMMA 3.4. *Let* A *and* A' *be as in Proposition* 3.1. *Then* $w(A)=w(A')$ *if and only if* $\langle Ax_{\omega_j},x_{\omega_j}\rangle = e^{i\psi}\omega_j w(A)$ (*equivalently,* $\big(\mathrm{Re}\,(e^{-i\psi}\overline{\omega}_jA)\big)x_{\omega_j}=w(A)x_{\omega_j}$) *for all* j, $0\leqslant j\leqslant n-1$, *where* x,ω_j *and* x_{ω_j} *are as in Lemma* 3.3 (*with* A' *replacing* A *there*) *and* $\psi=\big((\sum_{k=1}^{n-1}\arg a_{k,k+1})+\arg a_{n1}\big)/n$.

Proof. Let $\theta_k=\arg a_{k,k+1}$ for $1\leqslant k\leqslant n-1$ and $\theta_n=\arg a_{n1}$.

If $U=\mathrm{diag}\Big(\exp(i\psi),\exp\big(i(2\psi-\theta_1)\big),\ldots,\exp\big(i(n\psi-\sum_{k=1}^{n-1}\theta_k)\big)\Big)$, then U is unitary and U^*AU is of the form

$$e^{i\psi}\begin{bmatrix} * & |a_{12}| & * & \cdots & * \\ \cdot & * & |a_{23}| & \ddots & \vdots \\ \cdot & & \ddots & \ddots & * \\ \cdot & & & \ddots & |a_{n-1,n}| \\ |a_{n1}| & \cdot & \cdot & \cdot & * \end{bmatrix}.$$

Hence we may assume without loss of generality that the $a_{k,k+1}$'s and a_{n1} are all nonnegative and ψ is zero. Let $B=A-A'$.

Assume first that $w(A) = w(A')$. Since $\operatorname{Re}(\overline{\omega}_j\langle Ax_{\omega_j}, x_{\omega_j}\rangle) \leqslant |\langle Ax_{\omega_j}, x_{\omega_j}\rangle| \leqslant w(A)$ and $\operatorname{Re}(\overline{\omega}_j\langle A'x_{\omega_j}, x_{\omega_j}\rangle) = w(A')$ by Lemma 3.3 (b), we have $\operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) \leqslant 0$ for all j. Let B_k, $-(n-1) \leqslant k \leqslant n-1$, denote the matrix $[b_{ij}^{(k)}]_{i,j=1}^n$, where

$$b_{ij}^{(k)} = \begin{cases} a_{ij} & \text{if } j-i=k, \\ 0 & \text{otherwise,} \end{cases}$$

and let $b_k = \langle B_k x, x\rangle$. We have

$$\begin{aligned} \operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) &= \sum_{\substack{k=-(n-2) \\ k\neq 1}}^{n-1} \operatorname{Re}(\overline{\omega}_j\langle B_k x_{\omega_j}, x_{\omega_j}\rangle) \\ &= \sum_{\substack{k=-(n-2) \\ k\neq 1}}^{n-1} \operatorname{Re}(\overline{\omega}_j\langle B_k x, x\rangle \omega_j^k) = \sum_{\substack{k=-(n-2) \\ k\neq 1}}^{n-1} \operatorname{Re}(b_k \omega_j^{k-1}) \end{aligned}$$

for all j. Adding these together yields

$$\sum_{j=0}^{n-1} \operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) = \sum_{\substack{k=-(n-2) \\ k\neq 1}}^{n-1} \operatorname{Re}(b_k \sum_{j=0}^{n-1} \omega_j^{k-1}) = 0. \tag{2}$$

Since $\operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) \leqslant 0$ for all j, their sum being zero implies that they are all zero. Hence

$$\operatorname{Re}(\overline{\omega}_j\langle Ax_{\omega_j}, x_{\omega_j}\rangle) = \operatorname{Re}(\overline{\omega}_j\langle A'x_{\omega_j}, x_{\omega_j}\rangle) = w(A') = w(A), \quad 0 \leqslant j \leqslant n-1.$$

This together with $|\overline{\omega}_j\langle Ax_{\omega_j}, x_{\omega_j}\rangle| \leqslant w(A)$ yields that $\langle Ax_{\omega_j}, x_{\omega_j}\rangle = \omega_j w(A)$. Hence $w(\operatorname{Re}(\overline{\omega}_j A)) \geqslant \langle(\operatorname{Re}(\overline{\omega}_j A))x_{\omega_j}, x_{\omega_j}\rangle = w(A) \geqslant w(\operatorname{Re}(\overline{\omega}_j A))$ and thus the resulted equalities yield $\operatorname{Re}(\overline{\omega}_j A)x_{\omega_j} = w(A)x_{\omega_j}$. The implication from $\operatorname{Re}(\overline{\omega}_j A)x_{\omega_j} = w(A)x_{\omega_j}$ to $\langle Ax_{\omega_j}, x_{\omega_j}\rangle = \omega_j w(A)$ also follows from above.

Conversely, assume that our asserted condition holds. Since $\operatorname{Re}(\overline{\omega}_j\langle Ax_{\omega_j}, x_{\omega_j}\rangle) = w(A)$ and $\operatorname{Re}(\overline{\omega}_j\langle A'x_{\omega_j}, x_{\omega_j}\rangle) \leqslant w(A')$, we have $\operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) \geqslant 0$ for all j, $0 \leqslant j \leqslant n-1$, by Proposition 3.1. Inferring from the identity in (2), we obtain $\operatorname{Re}(\overline{\omega}_j\langle Bx_{\omega_j}, x_{\omega_j}\rangle) = 0$ for all j. In particular, the case $j=0$ yields $w(A) = \operatorname{Re}\langle Ax, x\rangle = \operatorname{Re}\langle A'x, x\rangle \leqslant w(A')$. By Proposition 3.1 again, we have $w(A) = w(A')$. □

We remark that the preceding lemma can also be proven by using the condition for the equality case of the Hadamard product in [4, Theorem 3.2].

We are now ready to prove Theorem 3.2.

Proof of Theorem 3.2. We need only prove (b) $\Rightarrow$ (c). As before, we may assume that $a_{12}, \ldots, a_{n-1,n}$ and a_{n1} are all strictly positive. Let x, ω_j and x_{ω_j} be as in Lemma 3.3 (with A' replacing A there) and let $B = A - A'$. Under the assumption $w(A) = w(A')$, we have $(\operatorname{Re}(\overline{\omega}_j B))x_{\omega_j} = 0$ as in the proof of Lemma 3.4 or $Bx_{\omega_j} = -\omega_j^2 B^* x_{\omega_j}$ for all j, $0 \leqslant j \leqslant n-1$. If $0 \leqslant j \neq k \leqslant n-1$, then

$$\begin{aligned} \langle Bx_{\omega_j}, x_{\omega_k}\rangle &= -\omega_j^2\langle B^* x_{\omega_j}, x_{\omega_k}\rangle = -\omega_j^2\langle x_{\omega_j}, Bx_{\omega_k}\rangle \\ &= \left(\frac{\omega_j}{\omega_k}\right)^2 \langle x_{\omega_j}, B^* x_{\omega_k}\rangle = \left(\frac{\omega_j}{\omega_k}\right)^2 \langle Bx_{\omega_j}, x_{\omega_k}\rangle. \end{aligned}$$

Since n is odd, we have $\omega_j^2 \neq \omega_k^2$. Hence $\langle Bx_{\omega_j}, x_{\omega_k}\rangle = 0$ for all $j \neq k$. On the other hand, we also have

$$\langle Bx_{\omega_j}, x_{\omega_j}\rangle = \langle Ax_{\omega_j}, x_{\omega_j}\rangle - \langle A'x_{\omega_j}, x_{\omega_j}\rangle = \omega_j w(A) - \omega_j w(A') = 0$$

for all j. Since the vectors x_{ω_k}, $0 \leqslant k \leqslant n-1$, form a basis of $\mathbb{C}^n$, we infer from above that $Bx_{\omega_j} = 0$ for all j and hence $B = 0_n$. This proves (c). □

We next consider the case of even n. The following theorem generalizes Proposition 2.1 (d) for $n = 2$.

THEOREM 3.5. *Let A and A' be as in Proposition 3.1 with n even and $a_{12}, \ldots, a_{n-1,n}$ and a_{n1} nonzero. Then $W(A) = W(A')$ if and only if $A = A'$.*

The proof is an elaboration of the one for odd n, for which we need the following lemma.

LEMMA 3.6. *Let A be as in Lemma 3.3 with $a_k > 0$ for all k. Let θ be in $[0, 2\pi) \setminus \{(2j+1)\pi/n : 0 \leqslant j \leqslant n-1\}$, r be the maximum eigenvalue of $\mathrm{Re}\,(e^{-i\theta}A)$, and $y = [y_1 \ \ldots \ y_n]^T$ be a unit vector satisfying $\big(\mathrm{Re}\,(e^{-i\theta}A)\big)y = ry$.*

(a) *We have $y_k \neq 0$ for all k. In particular, the eigenspace $\ker\big(rI_n - \mathrm{Re}\,(e^{-i\theta}A)\big)$ is one dimensional.*

(b) *If $\omega_j = \exp(2\pi i j/n)$ and $y_{\omega_j} = [y_1 \ y_2\omega_j \ y_3\omega_j^2 \ \ldots \ y_n\omega_j^{n-1}]^T$ for $0 \leqslant j \leqslant n-1$, then $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A)\big)y_{\omega_j} = ry_{\omega_j}$.*

Proof. (a) Assume otherwise that $y_k = 0$ for some k, $1 \leqslant k \leqslant n$. Let $\widehat{A}$ be the $(n-1)$-by-$(n-1)$ principal submatrix of A obtained by deleting the kth row and kth column of A, and let $\widehat{y}$ be the unit vector $[y_1 \ \ldots \ y_{k-1} \ y_{k+1} \ \ldots \ y_n]^T$ in $\mathbb{C}^{n-1}$. From our assumptions on r and y, we infer that $\lambda \equiv \langle Ay, y\rangle$ is in the boundary $\partial W(A)$ of $W(A)$. On the other hand, $W(\widehat{A})$ is contained in $W(A)$ and is a circular disc centered at the origin (cf. [11, Proposition 3 (3)]). Hence $\langle \widehat{A}\widehat{y}, \widehat{y}\rangle = \langle Ay, y\rangle = \lambda$ is also in $\partial W(\widehat{A})$ and therefore $\lambda = re^{i\theta}$ is in $\partial W(A) \cap \partial W(\widehat{A})$. However, by [11, Proposition 3 (4)], $\partial W(A)$ intersects $\partial W(\widehat{A})$ at exactly the n points $r\exp\big((2j+1)\pi i/n\big)$, $0 \leqslant j \leqslant n-1$, which contradicts our assumption on θ. Hence $y_k \neq 0$ for all k as asserted. Moreover, if $\ker\big(rI_n - \mathrm{Re}\,(e^{-i\theta}A)\big)$ has dimension bigger than one, then a suitable linear combination of two linearly independent vectors in it would result in a nonzero vector with, say, its first component equal to zero, which contradicts to what has just been proved. Thus $\ker\big(rI_n - \mathrm{Re}\,(e^{-i\theta}A)\big)$ must be of dimension one.

(b) Deriving as in the proof of Lemma 3.3 (b), we have $\langle Ay_{\omega_j}, y_{\omega_j}\rangle = \omega_j\langle Ay, y\rangle$. Hence

$$\begin{aligned}\big\langle \big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A)\big)y_{\omega_j}, y_{\omega_j}\big\rangle &= \mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle Ay_{\omega_j}, y_{\omega_j}\rangle) = \mathrm{Re}\,(e^{-i\theta}\langle Ay, y\rangle) \\ &= \big\langle \big(\mathrm{Re}\,(e^{-i\theta}A)\big)y, y\big\rangle = \langle ry, y\rangle = r.\end{aligned}$$

Since A is unitarily similar to $\overline{\omega}_j A$ by [11, Lemma 2 (2)], r is also the maximum eigenvalue of $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A)$. Thus $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A)\big)y_{\omega_j} = ry_{\omega_j}$ follows. □

Proof of Theorem 3.5. As before, we may assume that $a_{12},\ldots,a_{n-1,n}$ and a_{n1} are all strictly positive. Assume that $W(A) = W(A')$. Let θ be in $(0,\pi/n)$, r be the maximum eigenvalue of $\mathrm{Re}\,(e^{-i\theta}A')$, and $y = [y_1 \ldots y_n]^T$ be a unit vector such that $\big(\mathrm{Re}\,(e^{-i\theta}A')\big)y = ry$. Then, by Lemma 3.6, we have $y_k \neq 0$ for all k, $1 \leqslant k \leqslant n$, and $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A')\big)y_{\omega_j} = ry_{\omega_j}$ for all j, $0 \leqslant j \leqslant n-1$, where ω_j and y_{ω_j} are as before. On the other hand, since r is also the maximum eigenvalue of $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A')$ (cf. the proof of Lemma 3.6 (b)), it is equal to $\max \mathrm{Re}\,\big(e^{-i\theta}\overline{\omega}_j W(A')\big)$ $(= \max \mathrm{Re}\,\big(e^{-i\theta}\overline{\omega}_j W(A)\big))$. As $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle Ay_{\omega_j}, y_{\omega_j}\rangle)$ is in $\mathrm{Re}\,\big(e^{-i\theta}\overline{\omega}_j W(A)\big)$, we obtain that $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle Ay_{\omega_j}, y_{\omega_j}\rangle) \leqslant r = \mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle A'y_{\omega_j}, y_{\omega_j}\rangle)$. Thus if $B = A - A'$, then $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle By_{\omega_j}, y_{\omega_j}\rangle) \leqslant 0$ for all j. As the identity in (2) shows, their sum is equal to zero. It follows that $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle By_{\omega_j}, y_{\omega_j}\rangle) = 0$ or $\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j\langle Ay_{\omega_j}, y_{\omega_j}\rangle) = r$ for all j. As $r = \max \mathrm{Re}\,\big(e^{-i\theta}\overline{\omega}_j W(A)\big)$, we conclude that $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A)\big)y_{\omega_j} = ry_{\omega_j}$. Together with $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j A')\big)y_{\omega_j} = ry_{\omega_j}$, this yields $\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j B)\big)y_{\omega_j} = 0$. Hence y_{ω_j} is in $\ker\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j B)\big)$ for each j.

We now check that $\ker\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j B)\big)$ is equal to $\ker\big(\mathrm{Re}\,(e^{-i\theta}B)\big)$ for all j. Indeed, let $x = [x_1 \ \ldots \ x_n]^T$ be a unit vector in $\mathbb{C}^n$ with $x_k > 0$ for all k such that $\langle A'x, x\rangle = w(A')$, and let x_{ω_j} be as before. Then, as in the proof of Theorem 3.2, we have $Bx_{\omega_j} = -\omega_j^2 B^* x_{\omega_j}$ for all j, and $\langle Bx_{\omega_j}, x_{\omega_k}\rangle = 0$ for $0 \leqslant j,k \leqslant n-1$ with $|j-k| \neq n/2$. Let

$$b_j = \begin{cases} \langle Bx_{\omega_j}, x_{\omega_{j+(n/2)}}\rangle & \text{if } 0 \leqslant j \leqslant \frac{n}{2}-1, \\ \langle Bx_{\omega_j}, x_{\omega_{j-(n/2)}}\rangle & \text{if } \frac{n}{2} \leqslant j \leqslant n-1, \end{cases}$$

and let X be the n-by-n matrix $[x_{\omega_0} \ x_{\omega_1} \ \ldots \ x_{\omega_{n-1}}]$. Then

$$X^*BX = \begin{bmatrix} 0_{n/2} & B_2 \\ B_1 & 0_{n/2} \end{bmatrix},$$

where $B_1 = \mathrm{diag}\,(b_0, b_1, \ldots, b_{(n/2)-1})$ and $B_2 = \mathrm{diag}\,(b_{n/2}, b_{(n/2)+1}, \ldots, b_{n-1})$. Note that the b_j's are related in the following way:

$$\begin{aligned} b_{j+(n/2)} &= \langle Bx_{\omega_{j+(n/2)}}, x_{\omega_j}\rangle = -\omega_{j+(n/2)}^2\langle B^* x_{\omega_{j+(n/2)}}, x_{\omega_j}\rangle \\ &= -\omega_j^2\langle x_{\omega_{j+(n/2)}}, Bx_{\omega_j}\rangle = -\omega_j^2\overline{\langle Bx_{\omega_j}, x_{\omega_{j+(n/2)}}\rangle} = -\omega_j^2\overline{b}_j, \quad 0 \leqslant j \leqslant \frac{n}{2}-1. \end{aligned} \tag{3}$$

Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{C}^n$. We check that $\ker\big(\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j X^*BX)\big)$ is spanned by those vectors e_k for which $b_k = 0$. Indeed, since

$$\mathrm{Re}\,(e^{-i\theta}\overline{\omega}_j X^*BX) = \frac{1}{2}\begin{bmatrix} 0_{n/2} & C_j^* \\ C_j & 0_{n/2} \end{bmatrix},$$

where $C_j = e^{-i\theta}\overline{\omega}_j B_1 + e^{i\theta}\omega_j B_2^* = \operatorname{diag}(e^{-i\theta}\overline{\omega}_j b_0 + e^{i\theta}\omega_j \overline{b}_{n/2}, \ldots, e^{-i\theta}\overline{\omega}_j b_{(n/2)-1} + e^{i\theta}\omega_j \overline{b}_{n-1})$, we obtain from (3) that

$$\begin{aligned} e^{-i\theta}\overline{\omega}_j b_k + e^{i\theta}\omega_j \overline{b}_{k+(n/2)} &= e^{-i\theta}\overline{\omega}_j b_k + e^{i\theta}\omega_j(-\omega_k^2 b_k) \\ &= e^{-i\theta}\overline{\omega}_j b_k\big(1-(e^{i\theta}\omega_j\omega_k)^2\big) \end{aligned}$$

for $0 \leqslant j \leqslant n-1$ and $0 \leqslant k \leqslant (n/2)-1$. As $0 < \theta < \pi/n$, $(e^{i\theta}\omega_j\omega_k)^2$ is never equal to 1. Hence $e^{-i\theta}\overline{\omega}_j b_k + e^{i\theta}\omega_j \overline{b}_{k+(n/2)} = 0$ if and only if $b_k = 0$. Our assertion on the kernel of $\operatorname{Re}(e^{-i\theta}\overline{\omega}_j X^*BX)$ follows. In particular, we have $\ker\big(\operatorname{Re}(e^{-i\theta}\overline{\omega}_j X^*BX)\big) = \ker\big(\operatorname{Re}(e^{-i\theta}X^*BX)\big)$ or $\ker\big(X^*\big(\operatorname{Re}(e^{-i\theta}\overline{\omega}_j B)\big)X\big) = \ker\big(X^*\big(\operatorname{Re}(e^{-i\theta}B)\big)X\big)$ for all j. Since X is invertible, the latter equality yields that $\ker\big(\operatorname{Re}(e^{-i\theta}\overline{\omega}_j B)\big) = \ker\big(\operatorname{Re}(e^{-i\theta}B)\big)$ for all j as asserted.

From what were proven in the preceding two paragraphs, we obtain that y_{ω_j} is in $\ker\big(\operatorname{Re}(e^{-i\theta}B)\big)$ for all j. Since the components of y are all nonzero, the y_{ω_j}'s form a basis of $\mathbb{C}^n$. Hence $\ker\big(\operatorname{Re}(e^{-i\theta}B)\big) = \mathbb{C}^n$. As this kernel is spanned by those e_k's for which $b_k = 0$, we infer that $b_k = 0$ for all k. Hence $B_1 = B_2 = 0_{n/2}$, $X^*BX = 0_n$, or $B = 0_n$. This proves $A = A'$ as required. □

4. Ramifications

In this section, we discuss some results which are related to the main theorems in Section 3. We start with one class of matrices A for which $w(A) = w(A')$ implies $A = A'$ irrespective of the parity of its size. It is also a strengthening of [5, Theorem 3.11].

THEOREM 4.1. *Let A be the n-by-n companion matrix*

$$\begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix},$$

and let

$$A' = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Then (a) $w(A) \geqslant w(A')$, *and* (b) $w(A) = w(A')$ *if and only if* $a_1 = a_2 = \cdots = a_{n-1} = 0$.

Proof. (a) is by Proposition 3.1. To prove (b), assume that $w(A) = w(A')$. If $a_n = 0$, then $w(A) = w(A') = \cos\big(\pi/(n+1)\big)$ and hence $A = A'$ by [5, Theorem 3.11]. Therefore, we may assume that $a_n \neq 0$. Let

$$B = A - A' = \begin{bmatrix} 0 & & & & 0 \\ & 0 & & & \\ & & \ddots & & \\ & & & 0 & \\ 0 & -a_{n-1} & \cdots & -a_2 & -a_1 \end{bmatrix}.$$

We may assume that $a_n < 0$. Let x, ω_j and x_{ω_j} be as in Lemma 3.3 (with A' replacing A there). By Lemmas 3.3 (b) and 3.4, we have

$$\langle A' x_{\omega_j}, x_{\omega_j} \rangle = \omega_j w(A') = \omega_j w(A) = \langle A x_{\omega_j}, x_{\omega_j} \rangle$$

for all j, $0 \leqslant j \leqslant n-1$. Thus

$$\begin{aligned} 0 = \langle B x_{\omega_j}, x_{\omega_j} \rangle &= -x_n \overline{\omega}_j^{n-1} (a_{n-1} x_2 \omega_j + a_{n-2} x_3 \omega_j^2 + \cdots + a_1 x_n \omega_j^{n-1}) \\ &= -(a_{n-1} x_2 x_n \overline{\omega}_j^{n-2} + a_{n-2} x_3 x_n \overline{\omega}_j^{n-3} + \cdots + a_2 x_{n-1} x_n \overline{\omega}_j + a_1 x_n^2). \end{aligned}$$

This shows that the degree-$(n-2)$ polynomial $p(z) \equiv \sum_{k=2}^{n} a_{n-k+1} x_k x_n z^{n-k}$ has $\overline{\omega}_j$, $0 \leqslant j \leqslant n-1$, as zeros. Hence p must be the zero polynomial. Since $x_k > 0$ for all k, we obtain $a_1 = a_2 = \cdots = a_{n-1} = 0$ as asserted. $\square$

Another example of the equality of the numerical radii implying the equality of the matrices is given in the next theorem.

THEOREM 4.2. *Let* $A = [a_{ij}]_{i,j=1}^{n}$ $(n \geqslant 2)$ *and*

$$A'' = \begin{bmatrix} 0 & a_{12} & & & \\ & 0 & a_{23} & & \\ & & 0 & \ddots & \\ & & & \ddots & a_{n-1,n} \\ & & & & 0 \end{bmatrix}.$$

Then (a) $w(A) \geqslant w(A'')$, *and* (b) *if* $a_{12}, \ldots, a_{n-1,n}$ *are all nonzero, then the following conditions are equivalent:*

(i) $W(A) = W(A'')$,

(ii) $w(A) = w(A'')$, *and*

(iii) $A = A''$.

The following corollary of it is another generalization of [5, Theorem 3.11].

COROLLARY 4.3. *Let $A = [a_{ij}]_{i,j=1}^n$ with $a_{i,i+1} = 1$ for all i, $1 \leqslant i \leqslant n-1$. Then* (a) $w(A) \geqslant \cos\big(\pi/(n+1)\big)$ *and* (b) $w(A) = \cos\big(\pi/(n+1)\big)$ *if and only if $A = J_n$, the n-by-n Jordan block*

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

Proof. This is an easy consequence of the preceding theorem and the fact that $w(J_n) = \cos\big(\pi/(n+1)\big)$. □

For the proof of Theorem 4.2, we need the following lemma, which may have some independent interest.

LEMMA 4.4. *Let A and B be n-by-n matrices. Then $A = B$ if and only if there is a vector $x = [x_1 \ \dots \ x_n]^T$ in $\mathbb{C}^n$ with $x_k \neq 0$ for all k such that $\big(\mathrm{Re}\,(e^{-i\theta}A)\big)x_\theta = \big(\mathrm{Re}\,(e^{-i\theta}B)\big)x_\theta$, where $x_\theta = [x_1 \ \ x_2e^{i\theta} \ \ x_3e^{2i\theta} \ \ \dots \ \ x_ne^{(n-1)i\theta}]^T$, for at least $n+2$ distinct values of θ in $[0, 2\pi)$.*

Proof. To prove the sufficiency, we may assume that $B = 0_n$ and $A = [a_{ij}]_{i,j=1}^n$. Our assumption on A yields that $Ax_\theta + e^{2i\theta}A^*x_\theta = 0$ for $n+2$ values of θ in $[0, 2\pi)$ and hence, for z equal to such $e^{i\theta}$'s and $1 \leqslant j \leqslant n$,

$$\begin{aligned} p_j(z) &\equiv \sum_{k=1}^n a_{jk}x_kz^{k-1} + \sum_{k=1}^n \overline{a}_{kj}x_kz^{k+1} \\ &= (a_{j1}x_1 + a_{j2}x_2z) + \left(\sum_{k=2}^{n-1}(a_{j,k+1}x_{k+1} + \overline{a}_{k-1,j}x_{k-1})z^k\right) + (\overline{a}_{n-1,j}x_{n-1}z^n + \overline{a}_{nj}x_nz^{n+1}) \\ &= 0. \end{aligned}$$

It follows that $a_{j1} = a_{j2} = a_{n-1,j} = a_{nj} = 0$ and

$$a_{j,k+1}x_{k+1} + \overline{a}_{k-1,j}x_{k-1} = 0 \tag{4}$$

for all j, $1 \leqslant j \leqslant n$, and all k, $2 \leqslant k \leqslant n-1$. In particular, for $j = 1$ and $2 \leqslant k \leqslant n-1$, we have $a_{1,k+1}x_{k+1} + \overline{a}_{k-1,1}x_{k-1} = a_{1,k+1}x_{k+1} = 0$ and thus $a_{1,k+1} = 0$. Similarly, for $j = 2$ and $2 \leqslant k \leqslant n-1$, we obtain $a_{2,k+1} = 0$. In a similar fashion, we infer from (4) that $a_{k-1,n-1} = 0$ (resp., $a_{k-1,n} = 0$) for $j = n-1$ (resp., $j = n$) and $2 \leqslant k \leqslant n-1$. This shows that the jth row and kth column of A are all zeros for $j,k = 1,2,n-1$ and n. Continuing this process, we deduce successively from (4) that the remaining rows and columns of A are also zeros. Hence $A = 0_n$ as required. □

Note that, in the preceding lemma, the number "$n+2$" is sharp as seen by the following example.

EXAMPLE 4.5. Let $\theta_1,\dots,\theta_{n-1}$ be any $n-1$ distinct numbers in $[0,2\pi)$ which are also distinct from $\pi/2$ and $3\pi/2$, and let $x_{\theta_j} = [1 \quad e^{i\theta_j} \quad e^{2i\theta_j} \quad \dots \quad e^{(n-1)i\theta_j}]^T$ for $1 \leqslant j \leqslant n-1$. Since the x_{θ_j}'s are linearly independent, there is a nonzero vector y in $\mathbb{C}^n$ such that $\langle y, x_{\theta_j}\rangle = 0$ for all j. Let $A = yy^*$, $B = 0_n$ and $x = [1 \ \dots \ 1]^T$. Then

$$\big(\mathrm{Re}\,(e^{-i\theta_j}A)\big)x_{\theta_j} = \frac{1}{2}(e^{-i\theta_j}A + e^{i\theta_j}A^*)x_{\theta_j} = (\cos\theta_j)yy^*x_{\theta_j} = 0, \ \ 1 \leqslant j \leqslant n-1.$$

These together with $\mathrm{Re}\,(\pm iA) = 0_n$ show that $\big(\mathrm{Re}\,(e^{-i\theta}A)\big)x_\theta = 0$ for $n+1$ distinct values of θ in $[0,2\pi)$ does not guarantee that $A = 0_n$.

We remark that Lemma 4.4 is not applicable in the proofs of Theorems 3.2 and 3.5 since in each case we have only had n values of θ to satisfy the required condition.

Proof of Theorem 4.2. (a) As before, we may assume that $a_{12},\dots,a_{n-1,n}$ and a_{n1} are all nonnegative. If

$$A' = \begin{bmatrix} 0 & a_{12} & & & \\ & 0 & a_{23} & & \\ & & 0 & \ddots & \\ & & & \ddots & a_{n-1,n} \\ a_{n1} & & & & 0 \end{bmatrix},$$

then $0_n \preccurlyeq A'' \preccurlyeq A'$, which implies that $w(A'') \leqslant w(A')$ by [8, Corollary 3.6]. Together with $w(A') \leqslant w(A)$ from Proposition 3.1, this yields $w(A'') \leqslant w(A)$.

(b) Assume that $a_{12},\dots,a_{n-1,n}$ are all strictly positive and that $w(A) = w(A'')$. Let $x = [x_1 \ \dots \ x_n]^T$ be a unit vector in $\mathbb{C}^n$ with $x_k > 0$ for all k such that $\langle A''x, x\rangle = w(A'')$, and, for any real θ, let $x_\theta = [x_1 \quad x_2e^{i\theta} \quad x_3e^{2i\theta} \quad \dots \quad x_ne^{(n-1)i\theta}]^T$. As in the proof of Lemma 3.3 (b), we can easily verify that $\langle A''x_\theta, x_\theta\rangle = e^{i\theta}w(A'')$ and hence $\mathrm{Re}\,(e^{-i\theta}\langle A''x_\theta, x_\theta\rangle) = w(A'')$. On the other hand, we also have

$$\mathrm{Re}\,(e^{-i\theta}\langle Ax_\theta, x_\theta\rangle) \leqslant |\langle Ax_\theta, x_\theta\rangle| \leqslant w(A) = w(A'').$$

Thus if $B = A - A''$, then $\mathrm{Re}\,(e^{-i\theta}\langle Bx_\theta, x_\theta\rangle) \leqslant 0$ for all θ. For each k, $-(n-1) \leqslant k \leqslant n-1$ and $k \neq 1$, let A_k be the n-by-n matrix whose (i,j)-entry is a_{ij} if $j - i = k$, and 0 if otherwise. Then

$$\langle Bx_\theta, x_\theta\rangle = \sum_{\substack{k=-(n-1)\\ k\neq 1}}^{n-1} \langle A_kx_\theta, x_\theta\rangle = \sum_{\substack{k=-(n-1)\\ k\neq 1}}^{n-1} e^{ki\theta}\langle A_kx, x\rangle.$$

This shows that

$$-\mathrm{Re}\,(e^{-i\theta}\langle Bx_\theta, x_\theta\rangle) = -\mathrm{Re}\Big(\sum_{\substack{k=-(n-1)\\ k\neq 1}}^{n-1} e^{(k-1)i\theta}\langle A_kx, x\rangle\Big)$$

is a trigonometric polynomial of degree at most n which has no constant term and assumes only nonnegative values for all θ. By the Riesz–Fejér theorem [10, p. 77, Problem 40], there is a polynomial $p(z)=\sum_{j=0}^{n} b_j z^j$ of degree at most n such that $-\mathrm{Re}\,(e^{-i\theta}\langle Bx_\theta,x_\theta\rangle)=|p(e^{i\theta})|^2$ for all θ. Since the constant term of the latter is given by $\sum_{j=0}^{n}|b_j|^2$, we obtain $b_j=0$ for all j, $0\leqslant j\leqslant n$. Thus $\mathrm{Re}\,(e^{-i\theta}\langle Bx_\theta,x_\theta\rangle)=0$ for all θ. It follows that

$$\mathrm{Re}\,(e^{-i\theta}\langle Ax_\theta,x_\theta\rangle)=\mathrm{Re}\,(e^{-i\theta}\langle A''x_\theta,x_\theta\rangle)=w(A'')=w(A).$$

Thus

$$\big(\mathrm{Re}\,(e^{-i\theta}A)\big)x_\theta=w(A)x_\theta=w(A'')x_\theta=\big(\mathrm{Re}\,(e^{-i\theta}A'')\big)x_\theta$$

for all θ. It then follows from Lemma 4.4 that $A=A''$. $\square$

The inequalities in Proposition 3.1 and Theorem 4.2 (a) can be further generalized as follows.

Let $A=[a_{ij}]_{i,j=1}^{n}$. For any permutation σ on the integers $1,2,\ldots,n$ given by $\sigma(\ell)=k_\ell$ for $1\leqslant\ell\leqslant n$, and for any m, $1\leqslant m\leqslant n$, let A_{σ_m} be the n-by-n matrix $[a'_{ij}]_{i,j=1}^{n}$, where

$$a'_{ij}=\begin{cases} a_{ij} & \text{if } (i,j)=(k_\ell,k_{\ell+1}) \text{ for } 1\leqslant\ell\leqslant m\ (k_{n+1}\equiv k_1),\\ 0 & \text{otherwise.}\end{cases}$$

COROLLARY 4.6. *If A and A_{σ_m} are as above, then $w(A)\geqslant w(A_{\sigma_m})$.*

Proof. For any permutation σ as above, it is easily seen that A is unitarily similar to a matrix $B=[b_{ij}]_{i,j=1}^{n}$ of the form

$$\begin{bmatrix} * & a_{k_1,k_2} & * & \cdots & * \\ \vdots & * & a_{k_2,k_3} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & * \\ * & & & \ddots & a_{k_{n-1},k_n} \\ a_{k_n,k_1} & * & \cdots & \cdots & * \end{bmatrix}.$$

Let $B_{\sigma_m}=[b'_{ij}]_{i,j=1}^{n}$ be given by

$$b'_{ij}=\begin{cases} a_{k_i,k_{i+1}} & \text{if } (i,j)=(\ell,\ell+1) \text{ for } 1\leqslant\ell\leqslant m\ (k_{n+1}\equiv k_1),\\ 0 & \text{otherwise.}\end{cases}$$

Then B_{σ_m} is also unitarily similar to A_{σ_m}. In particular, we have $w(A)=w(B)\geqslant w(B_{\sigma_n})=w(A_{\sigma_n})$ by Proposition 3.1. To prove our assertion for general A_{σ_m}, we may assume that $a_{k_\ell,k_{\ell+1}}\geqslant 0$ for all ℓ. Then $B_{\sigma_n}\succcurlyeq B_{\sigma_m}$ yields that $w(B_{\sigma_n})\geqslant w(B_{\sigma_m})$ by [8, Corollary 3.6]. Therefore, $w(A)\geqslant w(B_{\sigma_n})\geqslant w(B_{\sigma_m})=w(A_{\sigma_m})$ as asserted. $\square$

We conclude this paper with an example showing that for a matrix A of even size the equality of $w(A)$ and $w(A')$ does not imply that $W(A)$ contains $W(A')$. This is in contrast to the case of A of size two (cf. Proposition 2.1 (c)).

EXAMPLE 4.7. If

$$A = \begin{bmatrix} 1 & 2 & -1/2 & 0 \\ 0 & -2/5 & 4 & 2/5 \\ -1/2 & 0 & 1/4 & 4 \\ 2 & 2/5 & 0 & -2/5 \end{bmatrix} \quad \text{and} \quad A' = \begin{bmatrix} 0\,2\,0\,0 \\ 0\,0\,4\,0 \\ 0\,0\,0\,4 \\ 2\,0\,0\,0 \end{bmatrix},$$

then $w(A) = w(A') = \sqrt{10}$, but $W(A) \not\supseteq W(A')$. Indeed, the characteristic polynomials of $\operatorname{Re}(e^{i\theta}A)$ and $\operatorname{Re}(e^{i\theta}A')$ for θ in $[0,2\pi)$ can be computed to be

$$\begin{aligned} p_\theta(z) &\equiv \det\left(zI_4 - \operatorname{Re}(e^{i\theta}A)\right) \\ &= z^4 - \left(\frac{9}{20}\cos\theta\right)z^3 - (\cos^2\theta + 10)z^2 + \left(\frac{9}{2}\cos\theta\right)z \\ &\quad + \left(5\cos^2\theta + 5\cos^2\theta \cdot \cos(2\theta) - 8\cos(4\theta) + 8\right) \end{aligned}$$

and

$$q_\theta(z) \equiv \det\left(zI_4 - \operatorname{Re}(e^{i\theta}A')\right) = z^4 - 10z^2 + \left(8 - 8\cos(4\theta)\right),$$

respectively. Since $q_0(z) = z^4 - 10z^2$ has zeros 0 and $\pm\sqrt{10}$, $p_0(\sqrt{10}) = 0$,

$$\begin{aligned} p_\theta(\sqrt{10}) &= 100 - \frac{9}{2}\sqrt{10}\cos\theta - (\cos^2\theta + 10)10 + \frac{9}{2}\sqrt{10}\cos\theta \\ &\quad + \left(5\cos^2\theta + 5\cos^2\theta \cdot \cos(2\theta) - 8\cos(4\theta) + 8\right) \\ &= -5\cos^2\theta + 5\cos^2\theta \cdot \cos(2\theta) - 8\cos(4\theta) + 8 \\ &= \frac{27}{2}\sin^2(2\theta) \geqslant 0, \end{aligned}$$

and $p_\theta(z)$ is strictly increasing on $[\sqrt{10},\infty)$ for any θ in $[0,2\pi)$ (because

$$\begin{aligned} p'_\theta(z) &= 4z^3 - \left(\frac{27}{20}\cos\theta\right)z^2 - 2(\cos^2\theta + 10)z + \frac{9}{2}\cos\theta \\ &\geqslant 4z\left(z^2 - \frac{27}{80}z - \frac{11}{2}\right) - \frac{9}{2} \\ &= 4z\left(z - \frac{27 + \sqrt{141529}}{160}\right)\left(z - \frac{27 - \sqrt{141529}}{160}\right) - \frac{9}{2} > 0 \end{aligned}$$

for $z \geqslant \sqrt{10}$), we conclude that the maximum eigenvalue of $\operatorname{Re}(e^{i\theta}A)$ is at most $\sqrt{10}$ for any θ and hence $w(A) = \sqrt{10}$. On the other hand, we also have $q_0(\sqrt{10}) = 0$,

$$q_\theta(\sqrt{10}) = 16\sin^2(2\theta) \geqslant 0$$

and $q_\theta(z)$ is strictly increasing on $[\sqrt{10},\infty)$ for θ in $[0,2\pi)$ (because

$$q'_\theta(z) = 4z^3 - 20z = 4z(z^2 - 5) > 0$$

for $z \geqslant \sqrt{10}$). Hence $w(A') = \sqrt{10}$.

Finally, we check that $W(A) \not\supseteq W(A')$. Since the zeros of $q_{\pi/4}(z) = z^4 - 10z^2 + 16$ are $\pm\sqrt{2}$ and $\pm 2\sqrt{2}$, the maximum eigenvalue of $\operatorname{Re}(e^{(\pi/4)i}A')$ is $2\sqrt{2}$. On the other hand, we have

$$p_{\pi/4}(z) = z^4 - \frac{9\sqrt{2}}{40}z^3 - \frac{21}{2}z^2 + \frac{9\sqrt{2}}{4}z + \frac{37}{2}.$$

Hence $p_{\pi/4}(2\sqrt{2}) = 3/10 > 0$ and $p_{\pi/4}(z)$ is strictly increasing on $[2\sqrt{2}, \infty)$ as above. Thus the maximum eigenvalue of $\operatorname{Re}(e^{(\pi/4)i}A)$ is less than $2\sqrt{2}$. This shows that $W(A) \not\supseteq W(A')$ as asserted.

Acknowledgements. This research was motivated by trying to answer an inquiry of Chi-Kwong Li to the first author concerning Corollary 4.3 here. We thank him for such an inspirational question. The second author also thanks his colleague Kuo-Zhong Wang for discussions on this topic, who also gave an independent proof of Theorem 4.2.

REFERENCES

[1] A. ABU-OMAR AND P. Y. WU, *Scalar approximants of quadratic operators with applications*, preprint.
[2] M. BERGER, *Geometry I*, Springer, Berlin, 1987.
[3] R. BHATIA, M. D. CHOI AND C. DAVIS, *Comparing a matrix to its off-diagonal part*, Oper. Theory Adv. Appl. **40** (1989), 151–164.
[4] H.-L. GAU AND P. Y. WU, *Numerical radius of Hadamard product of matrices*, Linear Algebra Appl. **504** (2016), 292–308.
[5] H.-L. GAU AND P. Y. WU, *Crawford numbers of companion matrices*, Oper. Matrices **10** (2016), 863–880.
[6] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[7] R. A. HORN AND C. R. JOHNSON, *Matrix Analysis*, 2nd ed., Cambridge University Press, Cambridge, 2013.
[8] C.-K. LI, B.-S. TAM AND P. Y. WU, *The numerical range of a nonnegative matrix*, Linear Algebra Appl. **350** (2002), 1–23.
[9] W. V. PARKER, *Sets of complex numbers associated with a matrix*, Duck Math. J. **15** (1948), 711–715.
[10] G. POLYA AND G. SZEGÖ, *Problems and Theorems in Analysis*, Vol. II, Springer, Berlin, 1976.
[11] M.-C. TSAI AND P. Y. WU, *Numerical ranges of weighted shift matrices*, Linear Algebra Appl. **435** (2011), 243–254.

(Received January 31, 2017)

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