ISSN: 1017-060X (Print)

ISSN: 1735-8515 (Online)

Special Issue of the

Bulletin of the Iranian Mathematical Society

in Honor of Professor Heydar Radjavi's 80th Birthday

Vol. 41 (2015), No. 7, pp. 15–27

Title:

Upper and lower bounds for numerical radii of block shifts

Author(s):

H.-L. Gau and P. Y. Wu

Published by Iranian Mathematical Society
http://bims.ims.ir

Bull. Iranian Math. Soc.
Vol. 41 (2015), No. 7, pp. 15–27
Online ISSN: 1735-8515

UPPER AND LOWER BOUNDS FOR NUMERICAL RADII OF BLOCK SHIFTS

H.-L. GAU AND P. Y. WU*

(Communicated by Peter Šemrl)

Dedicated to Professor Heydar Radjavi on his 80th birthday

ABSTRACT. For an n-by-n complex matrix A in a block form with the (possibly) nonzero blocks only on the diagonal above the main one, we consider two other matrices whose nonzero entries are along the diagonal above the main one and consist of the norms or minimum moduli of the diagonal blocks of A. In this paper, we obtain two inequalities relating the numeical radii of these matrices and also determine when either of them becomes an equality.
Keywords: Numerical radius, block shift, minimum modulus.
MSC(2010): Primary: 15A60; Secondary: 47A12.

1. Introduction

An n-by-n complex matrix A is called a *block shift* if it is of the form

$$\begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix},$$

where A_j's are in general rectangular matrices. In this paper, we obtain sharp upper and lower bounds for the numerical radius of such a matrix. Recall that the *numerical radius* $w(X)$ of an n-by-n matrix X is the quantity

$$\max\{|\langle Xx, x\rangle| : x \in \mathbb{C}^n, \|x\| = 1\},$$

where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the standard inner product and norm of vectors in $\mathbb{C}^n$, respectively. Note that $w(X)$ is the radius of the smallest circular disc

Article electronically published on December 31, 2015.
Received: 4 September 2014, Accepted: 2 December 2014.
*Corresponding author.

centered at the origin which contains the *numerical range*

$$W(X) = \{\langle Xx, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$$

of X. For properties of the numerical range and numerical radius, the reader is referred to [2, Chapter 22] or [3, Chapter 1].

Note that if A is a block shift of the above form, then it is unitarily similar to $e^{i\theta}A$ for all real θ. Hence its numerical range is a closed circular disc centered at the origin with radius equal to its numerical radius. To estimate the latter, we consider the scalar matrices

$$B = \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \quad \text{and} \quad B' = \begin{bmatrix} 0 & m(A_1) & & \\ & 0 & \ddots & \\ & & 0 & m(A_{k-1}) \\ & & & 0 \end{bmatrix},$$

where $m(\cdot)$ denotes the minimum modulus of a matrix. Recall that the *minimum modulus* $m(X)$ of an m-by-n matrix X is, by definition, $\min\{\|Xx\| : x \in \mathbb{C}^n, \|x\| = 1\}$. In Sections 2 and 3 below, we show that $w(B') \le w(A) \le w(B)$ always hold, and that, under the extra condition that A_j's are all nonzero (resp., $m(A_j)$'s are nonzero), $w(A) = w(B)$ (resp., $w(A) = w(B')$) implies that B (resp., B') is a direct summand of A (cf. Theorem 2.1 and Corollary 3.3). Examples are given showing that the nonzero conditions on A_j's are essential.

2. **Upper bound**

The main result of this section is the following theorem.

Theorem 2.1. *Let*

$$A = \begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix} \quad \text{on} \quad \mathbb{C}^n = \mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k} \tag{2.1}$$

be an n-by-n block shift, where A_j is an n_j-by-n_{j+1} matrix for $1 \le j \le k-1$, and let

$$B = \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \quad \text{on} \quad \mathbb{C}^k.$$

Express A and B as $\sum_{j=1}^{m} \oplus A'_j$ and $\sum_{j=1}^{m} \oplus B'_j$, respectively, where A'_j (resp., B'_j) is either a zero matrix or of the form

$$\begin{bmatrix} 0 & A_s & & \\ & 0 & \ddots & \\ & & \ddots & A_t \\ & & & 0 \end{bmatrix} \quad (resp. \begin{bmatrix} 0 & \|A_s\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_t\| \\ & & & 0 \end{bmatrix})$$

with $1 \le s \le t \le k-1$ and the A_j's in such expressions all nonzero. Then

(a) *$w(A) \le w(B)$,*

(b) *$w(A) = w(B)$ if and only if A is unitarily similar to $B'_{j_0} \oplus C$, where j_0 ($1 \le j_0 \le m$) is such that $w(A'_{j_0}) = \max_j w(A'_j)$ $(= w(A))$, and C is a block shift with $w(C) \le w(B'_{j_0})$, and*

(c) *under the assumption that $A_j \ne 0$ for all j in (2.1), we have $w(A) = w(B)$ if and only if A is unitarily similar to $B \oplus C$, where C is a block shift with $w(C) \le w(B)$.*

Proof. (a) Let $x = [x_1 \ \dots \ x_k]^T$ be a unit vector in $\mathbb{C}^n$ such that $|\langle Ax, x\rangle| = w(A)$. Hence

$$\begin{aligned} w(A) &= \left| \left\langle \begin{bmatrix} 0 & A_1 & & \\ & 0 & \ddots & \\ & & \ddots & A_{k-1} \\ & & & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix}, \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} \right\rangle \right| \\ &= \Big| \sum_{j=1}^{k-1} \langle A_j x_{j+1}, x_j \rangle \Big| \\ &\le \sum_{j=1}^{k-1} |\langle A_j x_{j+1}, x_j \rangle| \\ &\le \sum_{j=1}^{k-1} \|A_j\| \|x_{j+1}\| \|x_j\| \qquad (2.2) \\ &= \left\langle \begin{bmatrix} 0 & \|A_1\| & & \\ & 0 & \ddots & \\ & & \ddots & \|A_{k-1}\| \\ & & & 0 \end{bmatrix} \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_k\| \end{bmatrix}, \begin{bmatrix} \|x_1\| \\ \vdots \\ \|x_k\| \end{bmatrix} \right\rangle \\ &\le w(B), \qquad (2.3) \end{aligned}$$

where the last inequality follows from the fact that $[\|x_1\| \ \dots \ \|x_k\|]^T$ is a unit vector in $\mathbb{C}^k$.

(b) Assume that $w(A) = w(B)$. Let A'_{j_0} $(1 \le j_0 \le m)$ be such that $w(A'_{j_0}) = \max_j w(A'_j) = w(A)$. Then

$$w(A) = w(A'_{j_0}) \le w(B'_{j_0}) \le w(B) = w(A),$$

where the first inequality follows from (a). This yields equalities throughout. Hence considering A'_{j_0} and B'_{j_0} instead of A and B, we may assume that A_j's in (2.1) are all nonzero. The assumption $w(A) = w(B)$ also yields equalities throughout the chain of inequalities in the proof of (a). Since B is an (entrywise) nonnegative matrix with irreducible real part, the equality in (2.3) yields, by [4, Proposition 3.3], that $x_j \neq 0$ for all j. Let $\widehat{x}_j = [0 \ \dots \ 0 \ \underset{j\text{th}}{x_j} \ 0 \ \dots \ 0]^T$ for $1 \le j \le k$, and let K be the subspace of $\mathbb{C}^n$ spanned by $\widehat{x}_j$'s. The equality in (2.2) implies that

$$|\langle A_j x_{j+1}, x_j\rangle| = \|A_j x_{j+1}\|\|x_j\| = \|A_j\|\|x_{j+1}\|\|x_j\|. \tag{2.4}$$

Hence $A_j x_{j+1} = a_j x_j$ for some scalar a_j. Therefore, $A\widehat{x}_1 = 0$ and

$$A\widehat{x}_j = [0 \ \dots \ 0 \ \underset{(j-1)\text{st}}{A_{j-1}x_j} \ 0 \ \dots \ 0]^T = [0 \ \dots \ 0 \ \underset{(j-1)\text{st}}{a_{j-1}x_{j-1}} \ 0 \ \dots \ 0]^T = a_{j-1}\widehat{x}_{j-1}$$

is in K for all j, $2 \le j \le k$. This shows that $AK \subseteq K$.

We next prove that $A^*K \subseteq K$. Indeed, we have $A^*\widehat{x}_j = [0 \dots 0 \ \underset{(j+1)\text{st}}{A_j^* x_j} \ 0 \dots 0]^T$ for $1 \le j \le k-1$. Since

$$|a_j|\|x_j\|^2 = \|a_j x_j\|\|x_j\| = \|A_j x_{j+1}\|\|x_j\| = \|A_j\|\|x_{j+1}\|\|x_j\|$$

by (2.4), the nonzeroness of A_j's and x_j's yields the same for a_j's. Letting $B_j = A_j/\|A_j\|$ and $y_j = (\|A_j\|/a_j)x_{j+1}$, we have $B_j y_j = (1/a_j)A_j x_{j+1} = x_j$ with $\|B_j\| = 1$ and

$$\|y_j\| = \frac{\|A_j\|}{|a_j|}\|x_{j+1}\| = \frac{\|A_j x_{j+1}\|}{|a_j|} = \|x_j\|$$

by (2.4). It follows from an extension of a lemma of Riesz and Sz.-Nagy that $B_j^* x_j = y_j$ (cf. [6, p. 215]). Therefore, we have $A_j^* x_j = (\|A_j\|^2/a_j)x_{j+1}$, which shows that $A_j^*\widehat{x}_j = (\|A_j\|^2/a_j)\widehat{x}_{j+1}$ is in K for $1 \le j \le k-1$. Moreover, we also have $A^*\widehat{x}_k = 0$. Thus $A^*K \subseteq K$ as asserted.

Since $\{\widehat{x}_j/\|x_j\|\}_{j=1}^k$ is an orthonormal basis of K, $A(\widehat{x}_1/\|x_1\|) = 0$, and

$$\begin{aligned} A\left(\frac{\widehat{x}_j}{\|x_j\|}\right) &= \frac{a_{j-1}\|x_{j-1}\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} = \frac{a_{j-1}}{|a_{j-1}|}\frac{\|a_{j-1}x_{j-1}\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} \\ &= \frac{a_{j-1}}{|a_{j-1}|}\frac{\|A_{j-1}x_j\|}{\|x_j\|}\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} = \frac{a_{j-1}}{|a_{j-1}|}\|A_{j-1}\|\frac{\widehat{x}_{j-1}}{\|x_{j-1}\|} \end{aligned}$$

for $2 \le j \le k$ by (2.4), we derive that the restriction $A|_K$ is unitarily similar to B. Thus A is unitarily similar to $B \oplus (A|_{K^\perp})$. We now show that $A|_{K^\perp}$ is unitarily similar to a block shift. Indeed, let $\widehat{H}_j = 0 \oplus \cdots \oplus 0 \oplus \underset{j\text{th}}{\mathbb{C}^{n_j}} \oplus 0 \oplus \cdots \oplus 0$, $K_j = \mathbb{C}^{n_j} \ominus \bigvee\{x_j\}$, and $\widehat{K}_j = 0 \oplus \cdots \oplus 0 \oplus \underset{j\text{th}}{K_j} \oplus 0 \oplus \cdots \oplus 0$ for $1 \leq j \leq k$. Then $K^{\perp} = K_1 \oplus \cdots \oplus K_k$. Since $A\widehat{H}_{j+1} \subseteq \widehat{H}_j$ and $A^*\widehat{x}_j \in \bigvee\{\widehat{x}_{j+1}\}$ from before, we have $A\widehat{K}_{j+1} \subseteq \widehat{K}_j$ for $1 \leq j \leq k-1$. Moreover, $A\widehat{H}_k = \{0\}$ implies that $A\widehat{K}_k = \{0\}$. We conclude that $C \equiv A|_{K^{\perp}}$ is unitarily similar to a block shift with $w(C) \leq w(A) = w(B)$. This proves one direction of (b). The converse is trivial.

(c) is an easy consequence of (b). □

Corollary 2.2. *Let A be an n-by-n block shift as in (2.1). Then*

(a) $w(A) \leq \|A\| \cos(\pi/(k+1))$, *and*
(b) $w(A) = \|A\| \cos(\pi/(k+1))$ *if and only if A is unitarily similar to $(\|A\| J_k) \oplus C$, where C is a block shift with $w(C) \leq \|A\| \cos(\pi/(k+1))$.*

Here J_k denotes the k-by-k *Jordan block*

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

whose numerical range is known to be $\{z \in \mathbb{C} : |z| \leq \cos(\pi/(k+1))\}$ (cf. [5]).

The assertions in Corollary 2.2 are easy consequences of Theorem 2.1 and [4, Corollary 3.6].

We remark that the assertion in Theorem 2.1(c) still holds for $n \leq 5$ even without the nonzero assumption on A_j's. This can be proven via a case-by-case verification by invoking, in most cases, the known result on the numerical ranges of square-zero matrices (cf. [8, Theorem 2.1]), which we omit. This is no longer the case for $n \geq 6$. Here we give a counterexample for $n = 6$.

Example 2.3. Let

$$A = \begin{bmatrix} 0 & \sqrt{2} & & & & \\ & 0 & 0 & & & \\ & & 0 & 1 & 0 & \\ & & & 0 & 0 & 0 \\ & & & 0 & 0 & 1 \\ & & & & & 0 \end{bmatrix}$$

with $A_1 = [\sqrt{2}]$, $A_2 = [0]$, $A_3 = [1\ 0]$ and $A_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then

$$B = \begin{bmatrix} 0 & \sqrt{2} & & & \\ & 0 & 0 & & \\ & & 0 & 1 & \\ & & & 0 & 1 \\ & & & & 0 \end{bmatrix},$$

and thus

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

Hence $w(A) = w(B) = \sqrt{2}/2$, but B is not a direct summand of A. To see the latter, note that $\ker A \cap \ker A^* = \{0\}$. Hence A cannot have the 1-by-1 zero matrix $[0]$ as a direct summand, and thus A cannot be unitarily similar to $B \oplus [0]$, or B is not a direct summand of A. However, A has the direct summand $\begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$ as dictated by Theorem 2.1(b).

3. **Lower bound**

Let A be an n-by-n block shift as in (2.1). For each j, $1 \le j \le k-1$, let $B_j = [0]$ if $A_j = 0$, and, when $A_j \ne 0$, let

(3.1)
$$B_j = \begin{bmatrix} 0 & m(A_{s_j}) & & & & & & & \\ & 0 & \ddots & & & & & & \\ & & \ddots & m(A_{j-1}) & & & & & \\ & & & 0 & \|A_j\| & & & & \\ & & & & 0 & m(A_{j+1}^*) & & & \\ & & & & & 0 & \ddots & & \\ & & & & & & \ddots & m(A_{t_j}^*) \\ & & & & & & & 0 \end{bmatrix} \quad \text{on } \mathbb{C}^{t_j - s_j + 2},$$

where $s_j = \min\{\ell : 1 \le \ell \le j-1, m(A_\ell)\cdots m(A_{j-1}) \ne 0\}$ and $t_j = \max\{\ell : j+1 \le \ell \le k-1, m(A_{j+1}^*)\cdots m(A_\ell^*) \ne 0\}$. Note that

$$B_j = \begin{bmatrix} 0 & \|A_j\| & & & \\ & 0 & m(A_{j+1}^*) & & \\ & & 0 & \ddots & \\ & & & \ddots & m(A_{t_j}^*) \\ & & & & 0 \end{bmatrix} \text{(resp., } \begin{bmatrix} 0 & m(A_{s_j}) & & & \\ & 0 & \ddots & & \\ & & \ddots & m(A_{j-1}) & \\ & & & 0 & \|A_j\| \\ & & & & 0 \end{bmatrix} \text{)}$$

if $j = 1$ or $m(A_{j-1}) = 0$ (resp., $j = k-1$ or $m(A_{j+1}^*) = 0$).

The following is the main result of this section.

Theorem 3.1. *Let A and B_j, $1 \le j \le k-1$, be as above. Then*

(a) $w(A) \ge \max_j w(B_j)$,
(b) $w(A) = w(B_j)$ *for some* j *if and only if* A *is unitarily similar to* $B_j \oplus C$, *where* C *is a block shift with* $w(C) \le w(B_j)$.

The next lemma gives some basic properties of the minimum modulus of a rectangular matrix. For a square matrix (or, for that matter, an operator on a possibly infinite-dimensional Hilbert space), these appeared in [1, Theorem 1].

Lemma 3.2. *Let A be an m-by-n matrix. Then*

(a) $m(A) > 0$ *if and only if* A *is left invertible,*
(b) $m(A)$ *equals the minimum singular value of* A, *and*
(c) *if* $m < n$, *then* $m(A) = 0$.

Proof. (a) Note that $m(A) > 0$ means that there is a $c > 0$ such that $\|Ax\| \ge c\|x\|$ for all x in $\mathbb{C}^n$, which is equivalent to the well-definedness of the linear transformation $Ax \mapsto x$ from the range of A to $\mathbb{C}^n$, or to the left-invertibility of A.

(b) Consider the polar decomposition of A: $A = V(A^*A)^{1/2}$, where V is an m-by-n partial isometry with $\ker V = \ker A$ (cf. [2, Problem 134]). Then

$$\begin{aligned} m(A) &= \min\{\|Ax\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \min\{\|V(A^*A)^{1/2}x\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \min\{\|(A^*A)^{1/2}x\| : x \in \mathbb{C}^n, \|x\| = 1\} \\ &= \text{minimum eigenvalue of } (A^*A)^{1/2} \\ &= \text{minimum singular value of } A. \end{aligned}$$

(c) This is an easy consequence of (a) or (b). □

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. (a) We need to show that $w(A) \ge w(B_j)$ for all j. If $A_j = 0$, then obviously $w(A) \ge 0 = w(B_j)$. Now we assume that $A_j \ne 0$. Let x_{j+1} be a unit vector in $\mathbb{C}^{n_{j+1}}$ such that $\|A_j x_{j+1}\| = \|A_j\|$. For any i, $s_j \le i \le j-1$, we have $m(A_i) > 0$ and hence $\ker A_i = \{0\}$ by Lemma 3.2(a). Thus we may let $x_i = A_i x_{i+1}/\|A_i x_{i+1}\|$ for $i = j, j-1, \ldots, s_j$ successively. Similarly, since $m(A_i^*) > 0$ for $j+1 \le i \le t_j$, we may let $x_i = A_{i-1}^* x_{i-1}/\|A_{i-1}^* x_{i-1}\|$ for each i, $j+2 \le i \le t_j+1$. Such x_i's are unit vectors in $\mathbb{C}^{n_i}$'s. On the other hand, since B_j is an (entrywise) nonnegative matrix with irreducible real part, there is a unit vector $u = [r_{s_j} \ \ldots \ r_{t_j+1}]^T$ in $\mathbb{C}^{t_j-s_j+2}$ with $r_j > 0$ for all j such that $\langle B_j u, u\rangle = w(B_j)$ (cf. [4, Proposition 3.3]). Let $\widehat{u} =$

$[0\ \dots\ 0\ r_{s_j}x_{s_j}\ \dots\ r_{t_j+1}x_{t_j+1}\ 0\ \dots\ 0]^T$ in $\mathbb{C}^n$. Then $\widehat{u}$ is a unit vector and

$$
\begin{aligned}
\langle A\widehat{u},\widehat{u}\rangle &= \sum_{i=s_j}^{t_j}\langle A_i(r_{i+1}x_{i+1}), r_ix_i\rangle \\
&= \sum_{i=s_j}^{j} r_{i+1}r_i\Big\langle A_ix_{i+1}, \frac{A_ix_{i+1}}{\|A_ix_{i+1}\|}\Big\rangle + \sum_{i=j+1}^{t_j} r_{i+1}r_i\Big\langle \frac{A_i^*x_i}{\|A_i^*x_i\|}, A_i^*x_i\Big\rangle \\
&= \sum_{i=s_j}^{j} r_{i+1}r_i\|A_ix_{i+1}\| + \sum_{i=j+1}^{t_j} r_{i+1}r_i\|A_i^*x_i\| \\
&\geq \Big(\sum_{i=s_j}^{j-1} r_{i+1}r_im(A_i)\Big) + r_{j+1}r_j\|A_j\| + \Big(\sum_{i=j+1}^{t_j} r_{i+1}r_im(A_i^*)\Big) \qquad (3.2) \\
&= \langle B_ju,u\rangle = w(B_j).
\end{aligned}
$$

Hence $w(A)\geq\langle A\widehat{u},\widehat{u}\rangle\geq w(B_j)$ as asserted.

(b) Assume that $w(A)=w(B_j)$ for some j. From above, we have $w(A)=\langle A\widehat{u},\widehat{u}\rangle = w(B_j)$ and an equality in (3.2). We derive from the latter that $\|A_ix_{i+1}\| = m(A_i)$ for $s_j\leq i\leq j-1$, and $\|A_i^*x_i\| = m(A_i^*)$ for $j+1\leq i\leq t_j$. We now check that $A_{s_j-1}x_{s_j}=0$ and $A_{t_j+1}^*x_{t_j+1}=0$. To prove the former, assume otherwise that $A_{s_j-1}x_{s_j}\neq 0$ and $s_j\geq 2$. Then let $x_{s_j-1} = A_{s_j-1}x_{s_j}/\|A_{s_j-1}x_{s_j}\|$ and

$$
D = \left[\begin{array}{c|c} 0 & \|A_{s_j-1}x_{s_j}\|\ 0\ \dots\ 0 \\ \hline & B_j \end{array}\right]
$$

$$
= \begin{bmatrix}
0 & \|A_{s_j-1}x_{s_j}\| & & & & & & & & \\
 & 0 & m(A_{s_j}) & & & & & & & \\
 & & 0 & \ddots & & & & & & \\
 & & & \ddots & m(A_{j-1}) & & & & & \\
 & & & & 0 & \|A_j\| & & & & \\
 & & & & & 0 & m(A_{j+1}^*) & & & \\
 & & & & & & 0 & \ddots & & \\
 & & & & & & & \ddots & m(A_{t_j}^*) \\
 & & & & & & & & 0
\end{bmatrix}.
$$

Since $\|A_{s_j-1}x_{s_j}\|, m(A_{s_j}),\dots,m(A_{j-1}),\|A_j\|,m(A_{j+1}^*),\dots,m(A_{t_j}^*)>0$, we infer from [7, Lemma 5 (3)] that $w(D)>w(B_j)$ and from [4, Proposition 3.3] that there is a unit vector $v=[p_{s_j-1}\ \dots\ p_{t_j+1}]^T$ in $\mathbb{C}^{t_j-s_j+3}$ with $p_i>0$ for all i such that $\langle Dv,v\rangle = w(D)$. Let $\widehat{v}=[0\ \dots\ 0\ p_{s_j-1}x_{s_j-1}\ \dots\ p_{t_j+1}x_{t_j+1}\ 0\ \dots\ 0]^T$

in $\mathbb{C}^n$. Then $\widehat{v}$ is a unit vector and

$$\begin{aligned}
\langle A\widehat{v},\widehat{v}\rangle &= \sum_{i=s_j-1}^{t_j}\langle A_i(p_{i+1}x_{i+1}),p_ix_i\rangle \\
&= \sum_{i=s_j-1}^{t_j} p_{i+1}p_i\langle A_ix_{i+1},x_i\rangle \\
&= \sum_{i=s_j-1}^{j} p_{i+1}p_i\|A_ix_{i+1}\| + \sum_{i=j+1}^{t_j} p_{i+1}p_i\|A_i^*x_i\| \\
&= p_{s_j}p_{s_j-1}\|A_{s_j-1}x_{s_j}\| + \Big(\sum_{i=s_j}^{j-1} p_{i+1}p_im(A_i)\Big) + p_{j+1}p_j\|A_j\| + \Big(\sum_{i=j+1}^{t_j} p_{i+1}p_im(A_i^*)\Big) \\
&= \langle Dv,v\rangle = w(D) > w(B_j).
\end{aligned}$$

This yields $w(A)\geq\langle A\widehat{v},\widehat{v}\rangle > w(B_j)$, which contradicts our assumption. Thus we must have $A_{s_j-1}x_{s_j}=0$.

The proof for $A^*_{t_j+1}x_{t_j+1}=0$ is analogous to the above. Indeed, assume that $A^*_{t_j+1}x_{t_j+1}\neq 0$ and $t_j\leq k-2$. Let $x_{t_j+2}=A^*_{t_j+1}x_{t_j+1}/\|A^*_{t_j+1}x_{t_j+1}\|$ and

$$D=\left[\begin{array}{c|c} B_j & \begin{matrix}0\\ \vdots\\ 0\\ \|A^*_{t_j+1}x_{t_j+1}\|\end{matrix} \\ \hline & 0\end{array}\right] = \begin{bmatrix} 0 & m(A_{s_j}) & & & & & & & & \\ & 0 & \ddots & & & & & & & \\ & & \ddots & m(A_{j-1}) & & & & & & \\ & & & 0 & \|A_j\| & & & & & \\ & & & & 0 & m(A^*_{j+1}) & & & & \\ & & & & & 0 & \ddots & & & \\ & & & & & & \ddots & m(A^*_{t_j}) & & \\ & & & & & & & 0 & \|A^*_{t_j+1}x_{t_j+1}\| \\ & & & & & & & & 0 \end{bmatrix}.$$

As before, we have $w(D)>w(B_j)$ and there is a unit vector $w=[q_{s_j}\ \dots\ q_{t_j+2}]^T$ in $\mathbb{C}^{t_j-s_j+3}$ with $q_i>0$ for all i such that $\langle Dw,w\rangle=w(D)$. Let $\widehat{w}=[0\ \dots\ 0\ q_{s_j}x_{s_j}\ \dots\ q_{t_j+2}x_{t_j+2}\ 0\ \dots\ 0]^T$ in $\mathbb{C}^n$. Then $\widehat{w}$ is a unit vector and

$$\langle A\widehat{w}, \widehat{w}\rangle = \sum_{i=s_j}^{t_j+1} \langle A_i(q_{i+1}x_{i+1}), q_i x_i\rangle$$

$$= \sum_{i=s_j}^{t_j+1} q_{i+1}q_i \langle A_i x_{i+1}, x_i\rangle$$

$$= \Big(\sum_{i=s_j}^{j} q_{i+1}q_i \|A_i x_{i+1}\|\Big) + \Big(\sum_{i=j+1}^{t_j+1} q_{i+1}q_i \langle x_{i+1}, A_i^* x_i\rangle\Big)$$

$$= \Big(\sum_{i=s_j}^{j-1} q_{i+1}q_i m(A_i)\Big) + q_{j+1}q_j\|A_j\| + \Big(\sum_{i=j+1}^{t_j} q_{i+1}q_i m(A_i^*)\Big) + q_{t_j+2}q_{t_j+1}\|A_{t_j+1}^* x_{t_j+1}\|$$

$$= \langle Dw, w\rangle = w(D) > w(B_j).$$

We thus obtain $w(A) \geq \langle A\widehat{w}, \widehat{w}\rangle > w(B_j)$, a contradiction. Hence $A_{t_j+1}^* x_{t_j+1} = 0$ holds.

Let $\widehat{x}_i = [0\ \dots\ 0\ x_i\ 0\ \dots\ 0]^T$ for $s_j \leq i \leq t_j + 1$, and let K be the subspace of $\mathbb{C}^n$ spanned by the $\widehat{x}_i$'s. Since $A_{s_j-1}x_{s_j} = 0$ as proven above, we have $A\widehat{x}_{s_j} = 0$. Since $A_{i-1}x_i = \|A_{i-1}x_i\| x_{i-1}$ for $s_j + 1 \leq i \leq j + 1$, we also have

$$A\widehat{x}_i = [0\ \dots\ 0\ A_{i-1}x_i\ 0\ \dots\ 0]^T = \|A_{i-1}x_i\|\widehat{x}_{i-1}$$

for such i's. We now check that $A\widehat{x}_i = m(A_{i-1}^*)\widehat{x}_{i-1}$ for $j + 2 \leq i \leq t_j + 1$. Indeed, since $\|A_{i-1}^* x_{i-1}\| = m(A_{i-1}^*)$ from before, we have

$$\Big\langle \big(A_{i-1}A_{i-1}^* - m(A_{i-1}^*)^2 I_{n_{i-1}}\big)x_{i-1}, x_{i-1}\Big\rangle = \|A_{i-1}^* x_{i-1}\|^2 - m(A_{i-1}^*)^2 = 0.$$

The positive semidefiniteness of $A_{i-1}A_{i-1}^* - m(A_{i-1}^*)^2 I_{n_{i-1}}$ yields that $A_{i-1}A_{i-1}^* x_{i-1} = m(A_{i-1}^*)^2 x_{i-1}$. Hence

$$A_{i-1}x_i = A_{i-1}\frac{A_{i-1}^* x_{i-1}}{\|A_{i-1}^* x_{i-1}\|} = \frac{m(A_{i-1}^*)^2 x_{i-1}}{m(A_{i-1}^*)} = m(A_{i-1}^*)x_{i-1},$$

and therefore $A\widehat{x}_i = m(A_{i-1}^*)\widehat{x}_{i-1}$ as asserted. These show that $AK \subseteq K$.

We next show that $A^*K \subseteq K$. Indeed, for $s_j \leq i \leq j-1$, we have $\|A_i x_{i+1}\| = m(A_i)$. Hence

$$\Big\langle \big(A_i^* A_i - m(A_i)^2 I_{n_{i+1}}\big)x_{i+1}, x_{i+1}\Big\rangle = \|A_i x_{i+1}\|^2 - m(A_i)^2 = 0.$$

Since $A_i^* A_i \geq m(A_i)^2 I_{n_{i+1}}$, we infer that $A_i^* A_i x_{i+1} = m(A_i)^2 x_{i+1}$ and hence $A_i^* x_i = \big(m(A_i)^2/\|A_i x_{i+1}\|\big)x_{i+1} = m(A_i)x_{i+1}$. It follows that $A^*\widehat{x}_i = m(A_i)\widehat{x}_{i+1}$ is in K for $s_j \leq i \leq j - 1$. For $i = j$, we have

$$\|A_j^* x_j\| \leq \|A_j^*\| = \|A_j\| = \|A_j x_{j+1}\| = \langle A_j x_{j+1}, x_j\rangle = \langle x_{j+1}, A_j^* x_j\rangle \leq \|A_j^* x_j\|.$$

Thus the equalities hold throughout. In particular, this implies that $A_j^* x_j$ is a multiple of x_{j+1}. Again, $A^*\widehat{x}_j$ is in K. For $j + 1 \leq i \leq t_j$, we have

$A_i^* x_i = \|A_i^* x_i\| x_{i+1}$, which implies that $A^*\widehat{x}_i = \|A_i^* x_i\|\widehat{x}_{i+1}$ is in K. Finally, for $i = t_j + 1$, since $A_{t_j+1}^* x_{t_j+1} = 0$, we have $A^*\widehat{x}_{t_j+1} = 0$. Thus $A^*K \subseteq K$ as asserted.

From above, we conclude that A is unitarily similar to $(A|_K) \oplus (A|_{K^\perp})$. Since $\{\widehat{x}_{s_j}, \ldots, \widehat{x}_{t_j+1}\}$ is an orthonormal basis of K and

$$A\widehat{x}_i = \begin{cases} 0 & \text{if } i = s_j, \\ m(A_{i-1})\widehat{x}_{i-1} & \text{if } s_j + 1 \le i \le j, \\ \|A_j\|\widehat{x}_j & \text{if } i = j+1, \\ m(A_{i-1}^*)\widehat{x}_{i-1} & \text{if } j+2 \le i \le t_j + 1, \end{cases}$$

we infer that $A|_K$ is unitarily similar to B_j. The unitary similarity of $A|_{K^\perp}$ to a block shift follows as in the last part of the proof of Theorem 2.1(b). This proves one direction of (b). The converse is trivial. □

Corollary 3.3. *Let A be an n-by-n block shift as in (2.1), and let*

$$B' = \begin{bmatrix} 0 & m(A_1) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{k-1}) \\ & & & 0 \end{bmatrix}$$

and

$$B'' = \begin{bmatrix} 0 & m(A_1^*) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{k-1}^*) \\ & & & 0 \end{bmatrix} \quad \text{on } \mathbb{C}^k.$$

Then

(a) *$w(A) \ge w(B'), w(B'')$, and*

(b) *under the assumption of $m(A_j) > 0$ for all j (resp. $m(A_j^*) > 0$ for all j), we have $w(A) = w(B')$ (resp. $w(A) = w(B'')$) if and only if A is unitarily similar to $B' \oplus C$ (resp. $B'' \oplus C$), where C is a block shift with $w(C) \le w(B')$ (resp. $w(C) \le w(B'')$). In this case, $m(A_{k-1}) = \|A_{k-1}\|$ (resp. $m(A_1^*) = \|A_1\|$).*

Proof. We only prove for B'. The case involving B'' can be dealt with analogously.

(a) Note that B' is unitarily similar to a matrix of the form $(\sum_{i=1}^{r} \oplus B_i') \oplus 0_m$, where, for each i,

$$B_i' = \begin{bmatrix} 0 & m(A_{p_i}) & & \\ & 0 & \ddots & \\ & & \ddots & m(A_{q_i}) \\ & & & 0 \end{bmatrix}$$

with $m(A_\ell) > 0$ for $p_i \leq \ell \leq q_i$, $1 \leq p_i \leq q_i < p_{i+1} \leq q_{i+1} \leq k-1$, and 0_m denotes the m-by-m zero matrix. Since $m(A_j) \leq \|A_j\|$ for all j, we have $w(B_i') \leq w(C_i)$, where

$$C_i = \begin{bmatrix} 0 & m(A_{p_i}) & & & \\ & 0 & \ddots & & \\ & & \ddots & m(A_{q_i-1}) & \\ & & & 0 & \|A_{q_i}\| \\ & & & & 0 \end{bmatrix}$$

by [4, Corollary 3.6]. But, obviously, $w(C_i) \leq w(B_{q_i})$, where B_{q_i} is given by (3.1). Thus we obtain

$$w(B') = \max_{1\leq i\leq r} w(B_i') \leq \max_{1\leq i\leq r} w(C_i) \leq \max_{1\leq i\leq r} w(B_{q_i}) \leq \max_{1\leq j\leq k-1} w(B_j) \leq w(A)$$

by Theorem 3.1(a).

(b) If $m(A_j) > 0$ for all j, then $r = 1$, $B' = B_1'$ and $C_1 = B_{k-1}$ in (a). Hence if $w(A) = w(B')$, then $w(B') = w(B_{k-1}) = w(A)$. The first equality yields $m(A_{k-1}) = \|A_{k-1}\|$ by [4, Corollary 3.6] and thus $B' = B_{k-1}$ while the second equality implies, by Theorem 3.1(b), that A is unitarily similar to $B_{k-1} \oplus C$ for some block shift C with $w(C) \leq w(B_{k-1})$. Our assertion follows. The converse is trivial. □

Corollary 3.4. *Let A be an n-by-n block shift as in (2.1), and let $m = \min_j m(A_j)$. Then*

(a) *$w(A) \geq m \cdot \cos(\pi/(k+1))$, and*
(b) *$w(A) = m \cdot \cos(\pi/(k+1))$ if and only if A is unitarily similar to $(mJ_k) \oplus B$, where B is a block shift with $w(B) \leq m \cdot \cos(\pi/(k+1))$.*

This can be proven as Corollary 2.2 by using Corollary 3.3 and [4, Corollary 3.6].

Analogous to the situation in Section 2, the assertions in Corollary 3.3(b) remain true for $n \leq 3$ without the strict positivity assumptions on $m(A_j)$'s or $m(A_j^*)$'s. This is no longer the case for $n \geq 4$. A counterexample for $n = 4$ is given below.

Example 3.5. Let

$$A = \begin{bmatrix} 0 & 1 & 1 & \\ & 0 & 0 & 1 \\ & 0 & 0 & -1 \\ & & & 0 \end{bmatrix}$$

with $A_1 = [1\ 1]$ and $A_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. In this case,

$$B' = \begin{bmatrix} 0 & 0 & \\ & 0 & \sqrt{2} \\ & & 0 \end{bmatrix} \quad \text{and} \quad B'' = \begin{bmatrix} 0 & \sqrt{2} & \\ & 0 & 0 \\ & & 0 \end{bmatrix}.$$

Since $A^2 = 0$, we have $w(A) = \|A\|/2 = \sqrt{2}/2$ (cf. [8, Theorem 2.1]). On the other hand, we also have $w(B') = w(B'') = \sqrt{2}/2$. But neither B' nor B'' is a direct summand of A. This is because if it is, then A would be unitarily similar to $B' \oplus [0]$, which is impossible since $\ker A \cap \ker A^* = \{0\}$. However, A has the direct summand $\begin{bmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}$ as dictated by Theorem 3.1(b) (cf. also [8, Theorem 1.1]).

Acknowledgments

The two authors acknowledge supports from the Ministry of Science and Technology of the Republic of China under projects MOST 103-2115-M-008-006 and NSC 102-2115-M-009-007, respectively. They also thank the (anonymous) referee for his incisive comments, which lead to major improvements of the results in Section 3.

References

[1] R. Bouldin, The essential minimum modulus, *Indiana Univ. Math. J.* **30** (1981), no. 4, 513–517.

[2] P. R. Halmos, A Hilbert Space Problem Book, 2nd Edition, Springer-Verlag, New York-Berlin, 1982.

[3] R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[4] C. K. Li, B. S. Tam and P. Y. Wu, The numerical range of a nonnegative matrix, *Linear Algebra Appl.* **350** (2002) 1–23.

[5] M. Marcus and B. N. Shure, The numerical range of certain $0,1$-matrices, *Linear Multilinear Algebra* **7** (1979), no. 2, 111–120.

[6] D. Petz, Application of a lemma of Riesz and Sz.-Nagy, *Acta Sci. Math. (Szeged)* **57** (1993), no. 1-4, 215–222.

[7] M. C. Tsai and P. Y. Wu, Numerical ranges of weighted shift matrices, *Linear Algebra Appl.* **435** (2011), no. 9, 243–254.

[8] S. H. Tso and P. Y. Wu, Matricial ranges of quadratic operators, *Rocky Mountain J. Math.* **29** (1999), no. 3, 1139–1152.

(Hwa-Long Gau) Department of Mathematics, National Central University, Chungli 32001, Taiwan.

E-mail address: `hlgau@math.ncu.edu.tw`

(Pei Yuan Wu) Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan.

E-mail address: `pywu@math.nctu.edu.tw`