Linear Algebra and its Applications 438 (2013) 514–532

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

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

Diagonals and numerical ranges of weighted shift matrices

Kuo-Zhong Wang [*,1], Pei Yuan Wu [2]

Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan, ROC

ARTICLE INFO

Article history:
Received 17 May 2012
Accepted 13 August 2012
Available online 17 September 2012

Submitted by R.A. Brualdi

AMS classification:
15A60

Keywords:
Numerical ranges
Weighted shift matrix
Compression

ABSTRACT

For any n-by-n matrix A, we consider the maximum number $k = k(A)$ for which there is a k-by-k compression of A with all its diagonal entries in the boundary $\partial W(A)$ of the numerical range $W(A)$ of A. For any such compression, we give a standard model under unitary equivalence for A. This is then applied to determine the value of $k(A)$ for A of size 3 in terms of the shape of $W(A)$. When A is a matrix of the form

$$\begin{pmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ w_n & & & 0 \end{pmatrix},$$

we show that $k(A) = n$ if and only if either $|w_1| = \cdots = |w_n|$ or n is even and $|w_1| = |w_3| = \cdots = |w_{n-1}|$ and $|w_2| = |w_4| = \cdots = |w_n|$. For such matrices A with exactly one of the w_j's zero, we show that any k, $2 \le k \le n-1$, can be realized as the value of $k(A)$, and determine exactly when the equality $k(A) = n - 1$ holds.

1. Introduction

For an n-by-n complex matrix A, let $W(A) = \{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\}$ denote its *numerical range*, where $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ are the standard inner product and its associated norm in $\mathbb{C}^n$, respectively,

* Corresponding author.
E-mail addresses: kzwang@math.nctu.edu.tw (K.-Z. Wang), pywu@math.nctu.edu.tw (P.Y. Wu).

1 Research supported by the National Science Council of the Republic of China under NSC 99-2115-M-009-013-MY2.
2 Research supported by the National Science Council of the Republic of China under NSC 99-2115-M-009-002-MY2 and by the MOE-ATU project.

0024-3795/$ - see front matter
http://dx.doi.org/10.1016/j.laa.2012.08.007

and let $k(A)$ be the maximum number k of orthonormal vectors $x_1, \ldots, x_k$ in $\mathbb{C}^n$ with $\langle Ax_j, x_j\rangle$ in the boundary $\partial W(A)$ of $W(A)$ for all j. Note that $k(A)$ is also the maximum size of a compression of A with all its diagonal entries in $\partial W(A)$. Recall that a k-by-k matrix B is a *compression* of A if $B = V^*AV$ for some n-by-k matrix V with $V^*V = I_k$. The number $k(A)$ was first introduced in [5]. It relates properties of the numerical range and the compressions of A. In particular, it was shown in [5, Lemma 4.1 and Theorem 4.4] that $2 \le k(A) \le n$ for any n-by-n ($n \ge 2$) matrix A, and $k(A) = \lceil n/2 \rceil$ for any S_n-matrix A ($n \ge 3$). Recall that an n-by-n matrix A is of *class* S_n if it is a *contraction*, that is, $\|A\| \equiv \max_{\|x\|=1} \|Ax\| \le 1$, its eigenvalues are all in the open unit disc $\mathbb{D} \equiv \{z \in \mathbb{C} : |z| < 1\}$, and the rank of $I_n - A^*A$ equals one. One particular example is the n-by-n *Jordan block*

$$J_n = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}.$$

In this paper, we proceed to study $k(A)$ for other classes of A. In particular, we are interested in knowing when $k(A)$ equals the size of A. In Section 2 below, we first give a structure theorem (Theorem 2.7) of A when it has a compression with all its diagonal entries in $\partial W(A)$. This is then used to determine the value of $k(A)$ for A of size 3 in terms of the shape of its numerical range $W(A)$ (Proposition 2.11). Then, in Section 3, we consider the n-by-n ($n \ge 2$) *weighted shift matrix*

$$A = \begin{pmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ w_n & & & 0 \end{pmatrix}. \tag{1}$$

For such an A, we determine in Theorem 3.1 exactly when its $k(A)$ equals n. We show that this is the case if and only if either $|w_1| = \cdots = |w_n|$ or n is even and $|w_1| = |w_3| = \cdots = |w_{n-1}|$ and $|w_2| = |w_4| = \cdots = |w_n|$. In particular, this implies that, for A of the form (1) with $n \ge 3$ and with exactly one zero weight, $k(A)$ is never equal to n. We then concentrate on those A's in this latter class, and show that in this case its $k(A)$ can be any integer between 2 and $n-1$ (Theorem 3.5). We also completely characterize among such A's those with $k(A) = n-1$ (Theorem 3.10).

Our reference for properties of the numerical range is [6, Chapter 1].

We end this section by fixing some notations. For any finite square matrix A, we use Re $A = (A + A^*)/2$ and Im $A = (A - A^*)/(2i)$ to denote its *real* and *imaginary parts*, respectively, and ker A and ran A to denote its *kernel* and *range*, respectively. A is said to be *reducible* if it is unitarily equivalent to the direct sum of two other matrices; otherwise, A is *irreducible*. The set of eigenvalues of A is denoted by $\sigma(A)$. 0_n and I_n are the n-by-n zero and identity matrices, respectively. The n-by-n diagonal matrix with diagonals $a_1, \ldots, a_n$ is diag$(a_1, \ldots, a_n)$. The *argument*, arg z, of a nonzero complex number z is the unique number θ in $[0, 2\pi)$ such that $z = |z|e^{i\theta}$; arg 0 can be any number in $[0, 2\pi)$. Finally, for any $n \ge 1$, the *nth primitive root of unity* $e^{2\pi i/n}$ is denoted by ω_n.

2. Generalities

In this section, we prove some general results on the number $k(A)$ of a finite matrix A, and start by reviewing a few basic facts concerning the boundary points of $W(A)$.

For an n-by-n matrix A, a point a in $\partial W(A)$ and a supporting line L of $W(A)$ which passes through a, there is a θ in $[0, 2\pi)$ such that the ray R_θ from the origin which forms angle θ from the positive x-axis is perpendicular to $L = L_\theta$ (cf. Fig. 2.1). In this case, Re $(e^{-i\theta}a)$ is the maximum eigenvalue of

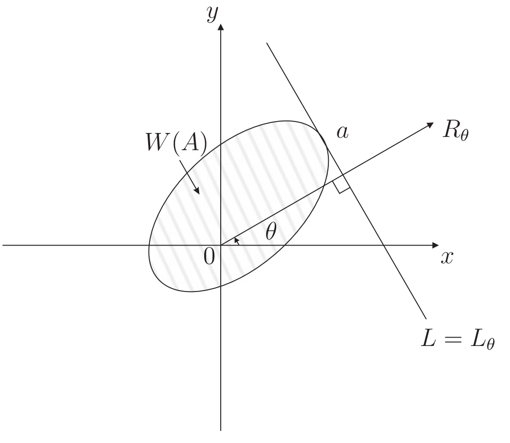

Fig. 2.1. Supporting line of $W(A)$.

$\mathrm{Re}\ (e^{-i\theta}A)$ with the corresponding eigenspace $E_{a,L}(A) \equiv \ker\ \mathrm{Re}\ (e^{-i\theta}(A - aI_n))$. Let $K_a(A)$ denote the set $\{x \in \mathbb{C}^n : \langle Ax, x\rangle = a\|x\|^2\}$ and $H_a(A)$ the subspace generated by $K_a(A)$. If the matrix A is clear from the context, we will abbreviate these to $E_{a,L}$, K_a and H_a, respectively. Note that these three sets are in general not equal. For example, if $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $a = 0$ or 1, then $W(A) = [0, 1]$ has infinitely many supporting lines L at a. It is easily seen that $E_{a,L} = \mathbb{C}^2$ if L is the x-axis, and $\mathbb{C} \oplus \{0\}$ if otherwise, and $K_a = H_a = \{0\} \oplus \mathbb{C}$ or $\mathbb{C} \oplus \{0\}$. On the other hand, if $0 < a < 1$, then L must be the x-axis, $E_{a,L} = H_a = \mathbb{C}^2$, and $K_a = \{(\sqrt{a}e^{i\theta_1}) \oplus (\sqrt{1-a}e^{i\theta_2}) : \theta_1,\ \theta_2 \in \mathbb{R}\}$. The next proposition gives precise information on their relationship.

Proposition 2.2. *Let A be an n-by-n matrix, a be a point in $\partial W(A)$, and L be a supporting line of $W(A)$ which passes through a. Then the following hold:*

(a) *H_a is contained in $E_{a,L}$.*
(b) *K_a is a subspace of $\mathbb{C}^n$, that is, $K_a = H_a$ if and only if a is an extreme point of $W(A)$.*
(c) *If a is not extreme for $W(A)$, then L is unique and $H_a = \cup\{K_b : b \in L \cap \partial W(A)\}$.*
(d) *$H_a = E_{a,L}$ if and only if either a is an extreme point of $W(A)$ and $L \cap \partial W(A) = \{a\}$ or a is not extreme for $W(A)$.*
(e) *If $L \cap \partial W(A)$ is a (nondegenerate) line segment of $\partial W(A)$, then $\dim E_{a,L} \geq 2$. The converse is in general false.*
(f) *If A is irreducible and $\dim E_{a,L} > n/2$, then $L \cap \partial W(A)$ is a line segment.*

Proof. (a) is trivial, (b) and (c) were proven in [2, Theorem 1], and (d) follows easily from (b) and (c). The assertion in (e) is trivial. For the converse, let

$$A = \begin{pmatrix} 0 & 0 & -2\sqrt{3+2\sqrt{2}} & 0 \\ 0 & 0 & 0 & -2\sqrt{3-2\sqrt{2}} \\ 2\sqrt{3+2\sqrt{2}} & 0 & 4 & -4 \\ 0 & 2\sqrt{3-2\sqrt{2}} & 4 & 4 \end{pmatrix}.$$

Then $W(A)$ is as in Fig. 2.3 with the y-axis as its supporting line L, which satisfies $L \cap \partial W(A) = \{0\}$ and $\dim E_{0,L} = 2$ (cf. [11, Example 4, Fig. 8]). It can be verified that the only (orthogonal) projection which commutes with A is 0_4 or I_4, and thus A is irreducible.

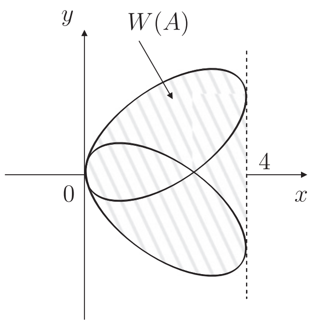

Fig. 2.3. Numerical range of A.

(f) After an affine transformation of A, we may assume that L is the y-axis and $a = 0$ is an eigenvalue of Re A with multiplicity bigger than $n/2$, that is, $m \equiv \dim M > n/2$, where $M = \ker \operatorname{Re} A$. Consider Re A as $0 \oplus B$ on $\mathbb{C}^n = M \oplus M^\perp$. For any unit vector x in M, we have

$$\langle Ax, x\rangle = \langle (\operatorname{Re} A)x, x\rangle + i\langle (\operatorname{Im} A)x, x\rangle = i\langle (\operatorname{Im} A)x, x\rangle.$$

Assume that $L \cap \partial W(A) = \{0\}$. This implies that $\langle (\operatorname{Im} A)x, x\rangle = 0$ for all x in M. Thus

$$A = \begin{pmatrix} 0 & 0 \\ 0 & B \end{pmatrix} + i \begin{pmatrix} 0 & C \\ C^* & D \end{pmatrix} = \begin{pmatrix} 0 & iC \\ iC^* & B + iD \end{pmatrix}$$

for some matrices C and D. Let $iC = USV$ be the singular value decomposition of iC, where U and V are unitary and S is of the form

$$\begin{pmatrix} s_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_{n-m} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

with $s_1 \geq \cdots \geq s_{n-m} \geq 0$. Hence A is unitarily equivalent to a matrix A' of the form $\begin{pmatrix} 0 & S \\ -S^* & E \end{pmatrix}$. If $s_{n-m} = 0$, then A' is reducible, contradicting our assumption on the irreducibility of A. Thus we have $s_{n-m} > 0$. Therefore, we have

$$\begin{aligned} \ker \operatorname{Im} A' &= \ker \begin{pmatrix} 0 & -iS \\ iS^* & \operatorname{Im} E \end{pmatrix} = \{(\underbrace{0, \ldots, 0}_{n-m}, x_{n-m+1}, \ldots, x_m, \underbrace{0, \ldots, 0}_{n-m}) : x_{n-m+1}, \ldots, x_m \in \mathbb{C}\} \\ &\subseteq \ker \begin{pmatrix} 0 & 0 \\ 0 & \operatorname{Re} E \end{pmatrix} = \ker \operatorname{Re} A'. \end{aligned}$$

Hence

$$\ker A' \cap \ker A'^* = \ker \operatorname{Re} A' \cap \ker \operatorname{Im} A' = \ker \operatorname{Im} A',$$

which is of dimension $2m-n > 0$. This shows that A' is reducible, again a contradiction. Thus $L\cap\partial W(A)$ is a line segment. □

We remark that Proposition 2.2(f) is a consequence of [1, Lemmas 2.1 and 2.2]. The proof here is more direct and matrix theoretic in nature. The case $n = 3$ was in [7, Proposition 3.2].

Using Proposition 2.2, we can give a lower bound for $k(A)$.

Proposition 2.4. *Let A be an n-by-n matrix, a be a point in $\partial W(A)$, and $k = \dim H_a$. If $W(A)$ is either the singleton $\{a\}$ or a line segment $[b, c]$ with a in (b, c), then $k(A) = k = n$; otherwise, $k(A) \geq k + 1$.*

Proof. If $W(A) = \{a\}$, then $A = aI_n$ and our assertion is obvious. On the other hand, if $W(A) = [b, c]$ with $a \in (b, c)$, then A is normal with eigenvalues in $[b, c]$. Hence we may diagonalize A to obtain $k(A) = n$. Since $H_a = \cup\{K_\lambda : \lambda \in [b, c]\} = \mathbb{C}^n$ by [2, Theorem 1] or Proposition 2.2(c), we also have $k = \dim H_a = n$. For the remaining case, consider a supporting line L_θ of $W(A)$ at a with the associated angle θ in $[0, 2\pi)$ such that $H_a = E_{a,L_\theta}$ (cf. Proposition 2.2(d)). Let $L_{\theta+\pi}$ be the supporting line of $W(A)$ which is parallel to L_θ, and let b be any point in $L_{\theta+\pi} \cap \partial W(A)$. Then E_{a,L_θ} (resp., $E_{b,L_{\theta+\pi}}$) is the eigenspace of $\operatorname{Re}(e^{-i\theta}A)$ for its maximum (resp., minimum) eigenvalue $\operatorname{Re}(e^{-i\theta}a)$ (resp., $\operatorname{Re}(e^{-i\theta}b)$). Since $W(\operatorname{Re}(e^{-i\theta}A))$ is not a singleton by our assumption, these two eigenvalues are distinct. Thus E_{a,L_θ} and $E_{b,L_{\theta+\pi}}$ are orthogonal to each other and hence the same is true for H_a and H_b. Therefore, we can find at least $m \equiv \dim H_a + \dim H_b$ many orthonormal vectors $x_1, \ldots, x_m$ in $\mathbb{C}^n$ with $\langle Ax_j, x_j\rangle$ in $\partial W(A)$ for all j. This shows that $k(A) \geq m = \dim H_a + \dim H_b \geq k + 1$ as asserted. □

Similar arguments as above together with Proposition 2.2(e) yield the following.

Corollary 2.5. *Let A be an n-by-n ($n \geq 3$) matrix*

(a) *If $\partial W(A)$ contains a line segment, then $k(A) \geq 3$.*
(b) *If $\partial W(A)$ has two parallel line segments, then $k(A) \geq 4$.*

Another easy corollary is the following necessary condition for $k(A) = 2$.

Corollary 2.6. *If A is an n-by-n nonscalar matrix with $k(A) = 2$, then $\dim H_a = 1$ for all a in $\partial W(A)$.*

The converse of the above is false. For example, if $A = J_5$, the 5-by-5 Jordan block, then it is known that $\dim H_a = 1$ for all a in $\partial W(A) = \{z \in \mathbb{C} : |z| = \cos(\pi/6)\}$, but $k(A) = 3$ (cf. [5, Theorem 4.4]). There are even 4-by-4 counterexamples to the converse as, for example, the matrix

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

(cf. Theorem 3.10 below). For 3-by-3 matrices, such a phenomenon cannot occur as will be seen in our discussions later in this section.

The main result of this section is the following structure theorem for matrix A which has a compression with diagonal entries all in $\partial W(A)$.

Theorem 2.7. *An n-by-n ($n \geq 2$) matrix A has a k-by-k compression with all its diagonal entries in $\partial W(A)$ if and only if it is unitarily equivalent to a matrix of the form*

$$\begin{pmatrix} B_1 & \cdots & 0 & e^{i\theta_1}C_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & B_m & e^{i\theta_m}C_m \\ -e^{i\theta_1}C_1^* & \cdots & -e^{i\theta_m}C_m^* & C \end{pmatrix}, \qquad (2)$$

where $\theta_1, \ldots, \theta_m$ *are distinct numbers in* $[0, \pi)$ *and* B_j, $1 \le j \le m$, *is of the form*

$$\left(\begin{array}{ccc|ccc} \alpha_1^{(j)} & \cdots & 0 & & & \\ \vdots & \ddots & \vdots & & e^{i\theta_j}D_j & \\ 0 & \cdots & \alpha_{s_j}^{(j)} & & & \\ \hline & & & \beta_1^{(j)} & \cdots & 0 \\ & -e^{i\theta_j}D_j^* & & \vdots & \ddots & \vdots \\ & & & 0 & \cdots & \beta_{t_j}^{(j)} \end{array}\right) \qquad (3)$$

with $s_j + t_j \ge 1$ *for all* j, $\sum_{j=1}^m (s_j + t_j) = k$, $\mathrm{Re}\ (e^{-i\theta_j}\alpha_1^{(j)}) = \cdots = \mathrm{Re}\ (e^{-i\theta_j}\alpha_{s_j}^{(j)}) = \max\ \sigma(\mathrm{Re}\ (e^{-i\theta_j}A))$ *and* $\mathrm{Re}\ (e^{-i\theta_j}\beta_1^{(j)}) = \cdots = \mathrm{Re}\ (e^{-i\theta_j}\beta_{t_j}^{(j)}) = \min\ \sigma(\mathrm{Re}\ (e^{-i\theta_j}A))$.

Geometrically, the conditions on the matrix B_j simply say that its diagonal entries $\alpha_1^{(j)}, \ldots, \alpha_{s_j}^{(j)}$ (resp., $\beta_1^{(j)}, \ldots, \beta_{t_j}^{(j)}$) are on the supporting line L_{θ_j} (resp., the parallel supporting line $L_{\theta_j+\pi}$) of $W(A)$ (cf. Fig. 2.1).

The proof of Theorem 2.7 depends on the corresponding result for 2-by-2 matrices (cf. [15, Corollary 4] or [5, Proposition 4.3]). This we state below for easy reference.

Proposition 2.8. *The following conditions are equivalent for a* 2-*by*-2 *matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

(a) $a \in \partial W(A)$,
(b) $be^{-i\theta} + \bar{c}e^{i\theta} = 0$ *for some* θ *in* $[0, 2\pi)$,
(c) $|b| = |c|$,
(d) $d \in \partial W(A)$.

Under these conditions, if A *is normal and* $W(A)$ *equals the line segment* $[a, d]$, *then* $b = c = 0$; *otherwise, the tangent lines to the (nondegenerate) ellipse* $\partial W(A)$ *at* a *and* d *are parallel to each other with the common slope* $-\cot\theta$.

Proof of Theorem 2.7. We need only prove the necessity. Let B be a k-by-k compression of A with the asserted property. We may assume, after a unitary equivalence, that $A = [a_{ij}]_{i,j=1}^n$ and $B = [a_{ij}]_{i,j=1}^k$. Consider all those diagonal entries of B which are on the same supporting line L_{θ_1} (resp., the parallel supporting line $L_{\theta_1+\pi}$) of $W(A)$ for some θ_1 in $[0, \pi)$. Call them $\alpha_1^{(1)}, \ldots, \alpha_{s_1}^{(1)}$ (resp., $\beta_1^{(1)}, \ldots, \beta_{t_1}^{(1)}$). Then $\mathrm{Re}\ (e^{-i\theta_1}\alpha_j^{(1)}) = \max\ \sigma(\mathrm{Re}\ (e^{-i\theta_1}A))$ for $1 \le j \le s_1$ (resp., $\mathrm{Re}\ (e^{-i\theta_1}\beta_j^{(1)}) = \min\ \sigma(\mathrm{Re}\ (e^{-i\theta_1}A))$ for $1 \le j \le t_1$). After a suitable permutation of rows and columns, we may further assume that

$$a_{jj} = \begin{cases} \alpha_j^{(1)} & \text{for } 1 \leq j \leq s_1, \\ \beta_{j-s_1}^{(1)} & \text{for } s_1 + 1 \leq j \leq s_1 + t_1. \end{cases}$$

Applying Proposition 2.8 repeatedly to the 2-by-2 principal submatrices $\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix}$, $1 \leq i, j \leq n$, of A, yields that A is of the form

$$\begin{pmatrix} B_1' & e^{i\theta_1} D_1 & e^{i\theta_1} C_1' \\ -e^{i\theta_1} D_1^* & B_1'' & e^{i\theta_1} C_1'' \\ -e^{i\theta_1} C_1'^* & -e^{i\theta_1} C_1''^* & E \end{pmatrix},$$

where $B_1' = \operatorname{diag}(\alpha_1^{(1)}, \ldots, \alpha_{s_1}^{(1)})$ and $B_1'' = \operatorname{diag}(\beta_1^{(1)}, \ldots, \beta_{t_1}^{(1)})$. We next apply the above arguments to E to obtain

$$E = \begin{pmatrix} B_2' & e^{i\theta_2} D_2 & e^{i\theta_2} C_2' \\ -e^{i\theta_2} D_2^* & B_2'' & e^{i\theta_2} C_2'' \\ -e^{i\theta_2} C_2'^* & -e^{i\theta_2} C_2''^* & E' \end{pmatrix},$$

where $\theta_2 \in [0, \pi)$ and $\theta_2 + \pi$ are distinct from θ_1 and $\theta_1 + \pi$, $B_2' = \operatorname{diag}(\alpha_1^{(2)}, \ldots, \alpha_{s_2}^{(2)})$ and $B_2'' = \operatorname{diag}(\beta_1^{(2)}, \ldots, \beta_{t_2}^{(2)})$. For any 2-by-2 submatrix

$$\begin{pmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{pmatrix} \tag{4}$$

with $1 \leq i \leq s_1 + t_1$ and $s_1 + t_1 + 1 \leq j \leq s_1 + t_1 + s_2 + t_2$, the diagonal entries a_{ii} (equal to either $\alpha_i^{(1)}$ or $\beta_{i-s_1}^{(1)}$) and a_{jj} (to $\alpha_{j-s_1-t_1}^{(2)}$ or $\beta_{j-s_1-t_1-s_2}^{(2)}$) are on distinct and nonparallel supporting lines of $W(A)$. Hence the submatrix (4) is normal with numerical range equal to $[a_{ii}, a_{jj}]$. We infer from Proposition 2.8 that $a_{ij} = a_{ji} = 0$. Repeating the above to E' and so forth, we thus obtain the asserted form for A.

The following lemma is useful on some occasions.

Lemma 2.9. *If $A = A_1 \oplus A_2$ with $W(A_2)$ contained in the interior of $W(A_1)$, then $k(A) = k(A_1)$.*

Proof. We obviously have $k(A) \geq k(A_1)$. To prove the converse inequality, assume that A, A_1 and A_2 are of sizes n, n_1 and n_2, respectively. Let $k = k(A)$ and let $u_1 = x_1 \oplus y_1, \ldots, u_k = x_k \oplus y_k$ be orthonormal vectors in $\mathbb{C}^n = \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2}$ such that $a_j \equiv \langle Au_j, u_j \rangle$ is in $\partial W(A) = \partial W(A_1)$ for all j. We claim that y_j must all be 0. Indeed, if $y_j \neq 0$ for some j, then

$$\begin{aligned} a_j &= \langle A_1 x_j, x_j \rangle + \langle A_2 y_j, y_j \rangle \\ &= \|x_j\|^2 \left\langle A_1 \frac{x_j}{\|x_j\|}, \frac{x_j}{\|x_j\|} \right\rangle + \|y_j\|^2 \left\langle A_2 \frac{y_j}{\|y_j\|}, \frac{y_j}{\|y_j\|} \right\rangle \\ &\equiv \|x_j\|^2 b_j + \|y_j\|^2 c_j \end{aligned}$$

if $x_j \neq 0$, and $a_j = c_j$ if otherwise. This shows that a_j is a convex combination of b_j and c_j. Since a_j, b_j and c_j are in $\partial W(A)$, $W(A_1)$ and $W(A_2)$, respectively, and $W(A_2)$ is contained in the interior of $W(A_1)$, we infer that $x_j \neq 0$ and a_j must be equal to b_j. It follows that $y_j = 0$, which is a contradiction. Thus $y_j = 0$ for all j and $x_1, \ldots, x_k$ are orthonormal vectors in $\mathbb{C}^{n_1}$ with $\langle A_1 x_j, x_j \rangle = a_j$ in $\partial W(A_1)$ for all j. This shows that $k(A_1) \geq k = k(A)$ and hence $k(A) = k(A_1)$. □

An easy consequence of Theorem 2.7 and Lemma 2.9 is the following upper bound for $k(A)$.

Proposition 2.10. *If A is an n-by-n ($n \geq 3$) matrix with* $\dim H_a = 1$ *for all a in $\partial W(A)$, then $k(A) \leq n-1$.*

Proof. If $k(A) = n$, then, by Theorem 2.7, A is unitarily equivalent to a direct sum $\sum_{j=1}^{m} \oplus B_j$, where each B_j is of the form (3). Our assumption on H_a implies that $\partial W(A)$ has no line segment and $H_a = E_{a,L}$ for any supporting line L of $W(A)$ (cf. Proposition 2.2(e) and (d)). As $W(A)$ equals the convex hull of $\cup_{j=1}^{m} W(B_j)$, these force the existence of some j_0, $1 \leq j_0 \leq m$, such that $W(B_j)$ is contained in the interior of $W(B_{j_0})$ for all $j \neq j_0$. Lemma 2.9 then yields that $k(B_{j_0}) = k(A) = n$. If $m > 1$, then, obviously, $k(B_{j_0}) \leq s_{j_0} + t_{j_0} < n$, which is a contradiction. Hence we must have $m = 1$ or A is unitarily equivalent to B_1. Then the fact that $\dim E_{a,L} = 1$ for any a in $\partial W(B_1)$ and any supporting line L of $W(B_1)$ implies that $s_1, t_1 \leq 1$. Therefore, B_1, together with A, is of size at most 2, which contradicts our assumption that $n \geq 3$. Thus $k(A) \leq n-1$ as asserted. □

We now combine Proposition 2.4 and Proposition 2.10 to determine $k(A)$ for a 3-by-3 matrix A. Recall that, in this case, $W(A)$ is of one of the following shapes (cf. [7]):

(a) a triangular region (or, in the degenerate case, a line segment or a singleton) if A is normal,
(b) an elliptic disc,
(c) an elliptic disc with a cone attached to it if A is unitarily equivalent to, say, $A' \oplus [a]$, where A' is a 2-by-2 nonnormal matrix and a is not in the elliptic disc $W(A')$,
(d) the convex hull of a heart-shaped region, in which case $\partial W(A)$ contains a line segment, and
(e) an oval region.

In cases (d) and (e) above, A is irreducible. The next proposition gives the value of $k(A)$ in terms of the shape of $W(A)$.

Proposition 2.11. *Let A be a 3-by-3 matrix. Then $k(A) = 2$ if $W(A)$ is either an elliptic disc, except when A has an eigenvalue on $\partial W(A)$, or an oval region. In all other cases, $k(A) = 3$.*

Proof. If $\partial W(A)$ contains a line segment, then $k(A) = 3$ by Corollary 2.5(a). This covers cases (a), (c) and (d) above. For the remaining part of the proof, we assume that $\partial W(A)$ contains no line segment. If A is irreducible, then $\dim H_a = 1$ for all a in $\partial W(A)$ by Proposition 2.2(a) and (f), and hence $k(A) \leq 2$ by Proposition 2.10. Therefore, in this case we have $k(A) = 2$ by Proposition 2.4 or [5, Lemma 4.1]. In particular, $k(A) = 2$ in case (e) above. For the remaining case (b), if A is irreducible, then $k(A) = 2$ as proven above. Now assume that A is reducible. Let A be unitarily equivalent to $A' \oplus [a]$, where A' is a 2-by-2 nonnormal matrix and a is in $W(A')$. If a is in $\partial W(A')$, then $\dim H_a = 2$ and hence $k(A) = 3$ by Proposition 2.4. On the other hand, if a is in the interior of $W(A')$, then $k(A) = k(A') = 2$ by Lemma 2.9 and [5, Lemma 4.1]. This completes the proof. □

The next corollary is already proven in the above.

Corollary 2.12. *A 3-by-3 matrix A is such that $k(A) = 2$ if and only if* $\dim H_a = 1$ *for all a in $\partial W(A)$.*

Corollary 2.13. *The following conditions are equivalent for the matrix*

$$A = \begin{pmatrix} 0 & a & \\ & 0 & b \\ c & & 0 \end{pmatrix}:$$

(a) $k(A) = 3$,
(b) $|a| = |b| = |c|$,

(c) *A is normal*, *and*
(d) *either* $A = 0_3$ *or* $\partial W(A)$ *contains a line segment.*

Proof. The equivalence of (b) and (c) was proven in [13, Proposition 4]; that of (b) and (d) was noted in [13, p. 248]. The implication $(d) \Rightarrow (a)$ is by Corollary 2.5. Finally, assume that (a) is true and $A \neq 0_3$. According to Proposition 2.11, either A is unitarily equivalent to $A' \oplus [a]$, where A' is a 2-by-2 nonnormal matrix and a is in $\partial W(A')$, or $\partial W(A)$ has a line segment. The former cannot happen since A is unitarily equivalent to $\omega_3 A$ (cf. [13, Proposition 3 (1)]). Thus (a) implies (d), completing the proof. □

In the next section, we consider the n-by-n weighted shift matrix (1) and determine when its $k(A)$ is equal to n, thus generalizing the preceding corollary.

3. Weighted shift matrices

An n-by-n *weighted shift matrix* A is one of the form (1), where the w_j's are called the *weights* of A. Properties of such matrices, especially those concerning their numerical ranges, were studied recently in [13,12]. Using the results there, we are able to give, among such matrices A, a characterization of the ones with $k(A) = n$.

Theorem 3.1. *Let A be an n-by-n* $(n \geq 2)$ *weighted shift matrix with weights* $w_1, \ldots, w_n$. *Then* $k(A) = n$ *if and only if either* $|w_1| = \cdots = |w_n|$ *or n is even and* $|w_1| = |w_3| = \cdots = |w_{n-1}|$ *and* $|w_2| = |w_4| = \cdots = |w_n|$.

The proof of this theorem depends on Theorem 2.7 and a corrected version of [12, Theorem 4] on the reducibility of weighted shift matrices, which appears in ([4], Theorem 3.1 and Corollary 3.3).

Theorem 3.2. *Let A be an n-by-n* $(n \geq 2)$ *weighted shift matrix with weights* $w_1, \ldots, w_n$. *Then A is reducible if and only if either at least two of the* w_j*'s are zero or the moduli of the weights* $|w_j|$ *are periodic. Moreover, if A is reducible and* $w_j \neq 0$ *for all j, then A is unitarily equivalent to* $e^{i\phi} \sum_{k=0}^{m-1} \oplus \omega_n^k B$, *where* $\phi = (\sum_{j=1}^{n} \arg w_j)/n$, *p is the period of the* $|w_j|$*'s*, $m = n/p$, *and B is the p-by-p irreducible weighted shift matrix with weights* $|w_1|, \ldots, |w_p|$.

Recall that the *period* of $\{|w_j|\}_{j=1}^{n}$ is the smallest integer p, $1 \leq p \leq n$, such that $|w_j| = |w_{p+j}|$ for all j ($w_m \equiv w_{m \,(\mathrm{mod}\, n)}$ for $m > n$). $\{|w_j|\}_j$ is *periodic* if the above p is such that $1 \leq p < n$, in which case we necessarily have $p|n$.

The next two lemmas facilitate the proof of Theorem 3.1.

Lemma 3.3. *Let A be an n-by-n* $(n \geq 2)$ *weighted shift matrix with nonzero weights* $w_1, \ldots, w_n$, *a be a point in* $\partial W(A)$, *and L be a supporting line of* $W(A)$ *which passes through a. Then* $\dim E_{a,L} \leq 2$. *Furthermore*, $\dim E_{a,L} = 2$ *if and only if* $L \cap \partial W(A)$ *is a (nondegenerate) line segment.*

Proof. Let θ in $[0, 2\pi)$ be such that the ray R_θ from the origin which forms angle θ from the positive x-axis is perpendicular to L (cf. Fig. 2.1), and let $x = [x_1, \ldots, x_n]^T$ be any vector in $E_{a,L} = \ker(\mathrm{Re}\,(e^{-i\theta}(A - aI_n)))$. Then $\mathrm{Re}\,(e^{-i\theta}A)x = \mathrm{Re}\,(e^{-i\theta}a)x \equiv \lambda x$, which is the same as

$$\frac{1}{2}(e^{-i\theta}w_1x_2 + e^{i\theta}\bar{w}_nx_n) = \lambda x_1,$$

$$\frac{1}{2}(e^{i\theta}\bar{w}_{j-1}x_{j-1} + e^{-i\theta}w_jx_{j+1}) = \lambda x_j, \quad 2 \leq j \leq n-1,$$

and

$$\frac{1}{2}(e^{-i\theta}w_nx_1 + e^{i\theta}\bar{w}_{n-1}x_{n-1}) = \lambda x_n.$$

Hence

$$x_2 = \frac{2\lambda e^{i\theta}}{w_1}x_1 - \frac{\bar{w}_n e^{2i\theta}}{w_1}x_n \equiv \alpha_2 x_1 + \beta_2 x_n, \tag{5}$$

$$x_{j+1} = \frac{2\lambda e^{i\theta}}{w_j}x_j - \frac{\bar{w}_{j-1} e^{2i\theta}}{w_j}x_{j-1}, \quad 2 \le j \le n-1 \tag{6}$$

and

$$x_{n-1} = -\frac{w_n e^{-2i\theta}}{\bar{w}_{n-1}}x_1 + \frac{2\lambda e^{-i\theta}}{\bar{w}_{n-1}}x_n \equiv \alpha_{n-1}x_1 + \beta_{n-1}x_n.$$

Iterating (6) and then applying (5), we may express each $x_{j+1}, 2 \le j \le n-3$, as $x_{j+1} = \alpha_{j+1}x_1 + \beta_{j+1}x_n$ for some scalars α_{j+1} and β_{j+1}. Let $u = [1, \alpha_2, \ldots, \alpha_{n-1}, 0]^T$ and $v = [0, \beta_2, \ldots, \beta_{n-1}, 1]^T$. Then x is a linear combination of u and v. Since u and v depend only on λ, θ and the w_j's, we obtain $\dim E_{a,L} \le 2$ as asserted. The second assertion was proven before in [13, Lemma 11]. □

Lemma 3.4. *Let A be an n-by-n ($n \ge 2$) irreducible weighted shift matrix with nonzero weights. Then $k(A) = n$ if and only if $n = 2$.*

Proof. Assume that $k(A) = n$. The irreducibility of A implies, by Theorem 2.7 and Lemma 3.3, that A is unitarily equivalent to a matrix of one of the following forms:

$$\begin{pmatrix} \alpha_1 & * \\ * & \beta_1 \end{pmatrix}, \quad \left(\begin{array}{cc|c} \alpha_1 & 0 & * \\ 0 & \alpha_2 & \\ \hline * & & \beta_1 \end{array}\right) \text{ and } \left(\begin{array}{cc|cc} \alpha_1 & 0 & * & \\ 0 & \alpha_2 & & \\ \hline * & & \beta_1 & 0 \\ & & 0 & \beta_2 \end{array}\right),$$

where α_1 and α_2 (resp., β_1 and β_2) are on a line segment of $\partial W(A)$. In particular, n can only be 2, 3 or 4. If $n = 3$ (resp., 4), then the existence of a line segment on $\partial W(A)$ yields that A is normal by Corollary 2.13 (resp., A is unitarily equivalent to the direct sum of two 2-by-2 matrices by [13, Proposition 12]), which contradicts the irreducibility of A. Thus we must have $n = 2$. Conversely, if $n = 2$, then $k(A) = 2$ by [5, Lemma 4.1], completing the proof. □

We are now ready to prove Theorem 3.1.

Proof of Theorem 3.1. Assume that $k(A) = n$. If A is irreducible, then $n = 2$ by Lemma 3.4 and we are done. Hence we may assume that A is reducible and also $A \ne 0_n$. Then Theorem 3.2 says that either at least two of the w_j's are zero or $\{|w_j|\}_{j=1}^n$ is periodic. In the former case, we may express A as $A_1 \oplus \cdots \oplus A_m$, where each A_k is either the 1-by-1 zero matrix 0_1 or a n_k-by-n_k ($n_k \ge 2$) weighted shift matrix with exactly one zero weight. Since the numerical ranges of the A_k's are either the singleton $\{0\}$ or a circular disc centered at the origin (cf. [13, Proposition 3 (3)]), we may assume that there is some l, $1 \le l \le m$, such that $W(A_1) = \cdots = W(A_l)$ and $W(A_{l+1}), \ldots, W(A_m)$ are all contained in the interior of $W(A_1)$. From Lemma 2.9, we deduce that $k(A) = k(A_1 \oplus \cdots \oplus A_l)$. Since $k(A) = n$ and $k(A_1 \oplus \cdots \oplus A_l) \le \sum_{k=1}^{l} n_k \le n$, we have $n = \sum_{k=1}^{l} n_k$ or $A = A_1 \oplus \cdots \oplus A_l$, where the A_k's are each of size at least 2 and have equal numerical ranges. Note also that the A_k's are irreducible. This is because if some

$$A_k = \begin{pmatrix} 0 & u_1 & & \\ & 0 & \ddots & \\ & & \ddots & u_{n_k-1} \\ & & & 0 \end{pmatrix},$$

where $u_j \neq 0$ for all j, is reducible, say, it is unitarily equivalent to $A' \oplus A''$, where A' and A'' are of sizes n' and n'' ($1 \leq n', n'' \leq n_k - 1$), respectively, then, since A_k, A' and A'' are all nilpotent, we have $A_k^p = A'^p \oplus A''^p = 0_{n_k}$, where $p = \max\{n', n''\} \leq n_k - 1$. This yields that

$$A_k^{n_k-1} = \begin{pmatrix} 0 & \cdots & 0 & \prod_{j=1}^{n_k-1} u_j \\ & 0 & & 0 \\ & & \ddots & \vdots \\ & & & 0 \end{pmatrix} = 0_{n_k},$$

and hence some u_j is equal to 0, which contradicts our assumption. (This fact appeared in [4, Proposition 3.2] with a different proof. The above was communicated to the second author by H.-L. Gau. Compare also [8, Lemma 2.4].) Next we claim that the A_k's are all of size exactly 2. To prove this, note that, by Theorem 2.7, A is unitarily equivalent to $\sum_{j=1}^{m} \oplus B_j$, where B_j, $1 \leq j \leq m$, is of the form (3). Since each B_j can be further decomposed as the direct sum of irreducible matrices and the irreducible summands of any matrix are unique up to ordering and unitary equivalence (cf. [3, Theorem 3.1]), we infer that B_j is unitarily equivalent to the direct sum of some of the A_k's, say, $A_{j_1} \oplus \cdots \oplus A_{j_q}$. Note that for any point α in $\partial W(B_j) = \partial W(A_{j_i})$, $1 \leq i \leq q$, we have $\dim E_{\alpha,L_\alpha}(B_j) = \sum_{i=1}^{q} \dim E_{\alpha,L_\alpha}(A_{j_i}) = q$, where L_α is the unique supporting line of the circular disc $W(B_j) = W(A_{j_i})$ at α (cf. Lemma 3.6 below). Since the diagonal entries $\alpha_1^{(j)}, \ldots, \alpha_{s_j}^{(j)}$ (resp., $\beta_1^{(j)}, \ldots, \beta_{t_j}^{(j)}$) of (3) for B_j are all in $E_{\alpha_1^{(j)},L_{\alpha_1}}(B_j)$ (resp., $E_{\beta_1^{(j)},L_{\beta_1}}(B_j)$), we infer that the size $s_j + t_j$ of B_j is at most $2q$. Thus the same is true for $A_{j_1} \oplus \cdots \oplus A_{j_q}$. Since each A_{j_i} is of size at least 2, we conclude that it has size exactly equal to 2. The same holds for all the A_k's. If $A_k = \begin{pmatrix} 0 & v_k \\ 0 & 0 \end{pmatrix}$ for $1 \leq k \leq l$, the equality of their numerical ranges $\{z \in \mathbb{C} : |z| \leq |v_k|/2\}$ yields that $|v_1| = \cdots = |v_l| \equiv v$. Hence n is even and $|w_1| = |w_3| = \cdots = |w_{n-1}| = v$ and $|w_2| = |w_4| = \cdots = |w_n| = 0$.

We next consider the case of periodic weights. Assume that $w_j > 0$ for all j and $\{w_j\}_{j=1}^n$ is periodic with period $p \geq 3$. Theorem 3.2 says that A is unitarily equivalent to $\sum_{k=0}^{m-1} \oplus \omega_n^k B$, where $m = n/p$ and B is the p-by-p irreducible weighted shift matrix with weights $w_1, \ldots, w_p$. On the other hand, since $k(A) = n$, Theorem 2.7 implies that A is also unitarily equivalent to $\sum_{i=1}^{q} \oplus B_i$, where each B_i is of the form (3). Note that, by Lemma 3.3, B_i, $1 \leq i \leq q$, is of size at most 4. As before, by the uniqueness of the irreducible summands of A [3, Theorem 3.1], we infer that each B_i is unitarily equivalent to the direct sum of some of the $\omega_n^k B$'s. Since B is of size at least 3, each B_i can be of size 3 or 4 only. Hence B_i is unitarily equivalent to one single $\omega_n^k B$ and $\partial W(B_i) = \partial W(B)$ has a line segment by Lemma 3.3. These facts combined together yield, via Corollary 2.13 or [13, Proposition 12], that B is reducible, which leads to a contradiction. Thus p must be equal to 1 or 2. This yields our assertion on the weights of A.

Conversely, if $|w_1| = \cdots = |w_n|$, then A is normal and is unitarily equivalent to $e^{i\phi}|w_1|B$, where $\phi = (\sum_{j=1}^{n} \arg\ w_j)/n$ and $B = \mathrm{diag}(1,\ \omega_n, \ldots, \omega_n^{n-1})$. Thus $k(A) = n$ obviously. On the other hand, if n is even and $|w_1| = |w_3| = \cdots = |w_{n-1}|$ and $|w_2| = |w_4| = \cdots = |w_n|$, then A is unitarily equivalent to $e^{i\phi} \sum_{k=0}^{(n/2)-1} \oplus \omega_n^k C$, where $C = \begin{pmatrix} 0 & |w_1| \\ |w_2| & 0 \end{pmatrix}$, by Theorem 3.2. We easily obtain $k(A) = n$ from Proposition 2.8.

Note that under the conditions of Theorem 3.1, the numerical range of A is either a regular n-polygonal region with vertices $e^{i(2k\pi + \sum_j \arg\ w_j)/n}|w_1|$, $0 \leq k \leq n-1$, or the convex hull of the union of $n/2$ elliptic discs with foci $\pm e^{i(2k\pi + \sum_j \arg\ w_j)/n}|w_1 w_2|$, $0 \leq k \leq (n/2) - 1$, and minor axis of length $||w_1| - |w_2||$.

A consequence of Theorem 3.1 is that if A is an n-by-n ($n \geq 3$) weighted shift matrix with exactly one zero weight, then $k(A)$ is never equal to n. In the remaining part of this section, we restrict ourselves to such matrices A. It turns out that in this case $k(A)$ can be any integer from 2 to $n-1$.

Theorem 3.5. *For any $n \geq 3$ and any k, $2 \leq k \leq n-1$, there is a matrix A of the form*

$$\begin{pmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ & & & 0 \end{pmatrix} \tag{7}$$

with $w_j \neq 0$ for all j such that $k(A) = k$.

This will be proven after a series of lemmas, the first of which gives conditions for two unit vectors x and y with $\langle Ax, x\rangle$ and $\langle Ay, y\rangle$ in $\partial W(A)$ to be orthogonal to each other.

Recall that the *numerical radius* $w(A)$ of a matrix A is the quantity $\max\{|z| : z \in W(A)\}$.

Lemma 3.6. *Let A be an n-by-n ($n \geq 2$) matrix of the form (7) with $w_j > 0$ for all j. Then the following hold:*

(a) $W(A) = \{z \in \mathbb{C} : |z| \leq w(A)\}$.
(b) *There is a unique unit vector $x = [x_1, \ldots, x_n]^T$ in $\mathbb{C}^n$ with $x_j > 0$ for all j such that $\langle Ax, x\rangle = w(A)$.*
(c) *For any $a = w(A)e^{i\theta}$, $\theta \in [0, 2\pi)$, in $\partial W(A)$, let $x_\theta = [x_1,\ e^{i\theta}x_2, \ldots, e^{(n-1)i\theta}x_n]^T$. Then $a = \langle Ax_\theta, x_\theta\rangle$ and H_a is generated by x_θ.*
(d) *Let $a_j = w(A)e^{i\theta_j}$ ($\theta_j \in [0, 2\pi)$), $j = 1,\ 2$, be two points in $\partial W(A)$. Then x_{θ_1} and x_{θ_2} are orthogonal to each other if and only if $e^{i(\theta_1-\theta_2)}$ is a zero of the polynomial $x_1^2 + x_2^2 z + \cdots + x_n^2 z^{n-1}$.*

Proof. Since $U_\theta^* A U_\theta = e^{i\theta}A$ for any real θ, where $U_\theta = \mathrm{diag}(1,\ e^{i\theta},\ e^{2i\theta}, \ldots, e^{(n-1)i\theta})$ is unitary, (a) follows immediately. (b) is a consequence of [9, Proposition 3.3] since A is a nonnegative matrix with Re A (permutationally) irreducible. To prove (c), note that

$$\begin{aligned} a = w(A)e^{i\theta} &= \langle e^{i\theta}Ax, x\rangle = \langle U_\theta^* A U_\theta x, x\rangle \\ &= \langle A(U_\theta x), U_\theta x\rangle = \langle Ax_\theta, x_\theta\rangle, \end{aligned}$$

which shows that x_θ is in H_a. That $\dim H_a = 1$ is by [9, Corollary 3.10]. Thus H_a is generated by x_θ. (d) follows from the fact that $\langle x_{\theta_1}, x_{\theta_2}\rangle = \sum_{k=1}^n e^{(k-1)i(\theta_1-\theta_2)}x_k^2$. This completes the proof. □

Thus, for a matrix A of the form (7) with $w_j > 0$ for all j, $k(A)$ equals the maximum number of $\theta_1, \ldots, \theta_k$ in $[0, 2\pi)$ for which $e^{i(\theta_j-\theta_l)}$ is a zero of $x_1^2 + x_2^2 z + \cdots + x_n^2 z^{n-1}$ for all j and l, $1 \leq j \neq l \leq k$. To actually construct the matrix A in Theorem 3.5, we need another tool, namely, a parametric representation of the weights w_j from [14, Theorem 3.1(b)].

Lemma 3.7. *Let A be an n-by-n ($n \geq 2$) matrix of the form (7) with $w_j > 0$ for all j. Then there is a unique sequence $\{a_j\}_{j=1}^n$ with $a_1 = -1$, $-1 < a_j < 1$ for $2 \leq j \leq n-1$, and $a_n = 1$ such that $w_j/w(A) = \sqrt{(1-a_j)(1+a_{j+1})}$ for all j. In this case, if $y_1 = 1$, $y_j = \prod_{k=1}^{j-1}\sqrt{(1-a_k)/(1+a_{k+1})}$ for $2 \leq j \leq n$, and $y = [y_1, \ldots, y_n]^T$, then $\langle A(y/\|y\|), y/\|y\|\rangle = w(A)$.*

Proof. We need only prove the second assertion. Note that $(\mathrm{Re}\, A)y = (1/2)[w_1y_2,\ w_1y_1 + w_2y_3, \ldots, w_{n-2}y_{n-2} + w_{n-1}y_n,\ w_{n-1}y_{n-1}]^T$. Here

$$w_1y_2 = w(A)\sqrt{(1-a_1)(1+a_2)}\sqrt{\frac{1-a_1}{1+a_2}} = w(A)(1-a_1) = 2w(A) = 2w(A)y_1,$$

$$\begin{aligned} &w_jy_j + w_{j+1}y_{j+2} \\ &= w(A)\left[\sqrt{(1-a_j)(1+a_{j+1})}\prod_{k=1}^{j-1}\sqrt{\frac{1-a_k}{1+a_{k+1}}} + \sqrt{(1-a_{j+1})(1+a_{j+2})}\prod_{k=1}^{j+1}\sqrt{\frac{1-a_k}{1+a_{k+1}}}\right] \\ &= w(A)\left(\prod_{k=1}^{j}\sqrt{\frac{1-a_k}{1+a_{k+1}}}\right)\left[\sqrt{(1-a_j)(1+a_{j+1})}\sqrt{\frac{1+a_{j+1}}{1-a_j}}\right. \\ &\quad \left.+\sqrt{(1-a_{j+1})(1+a_{j+2})}\sqrt{\frac{1-a_{j+1}}{1+a_{j+2}}}\right] \\ &= w(A)y_{j+1}[(1+a_{j+1})+(1-a_{j+1})] \\ &= 2w(A)y_{j+1}, \qquad 1 \le j \le n-2, \end{aligned}$$

and

$$\begin{aligned} w_{n-1}y_{n-1} &= w(A)\sqrt{(1-a_{n-1})(1+a_n)}\prod_{k=1}^{n-2}\sqrt{\frac{1-a_k}{1+a_{k+1}}} \\ &= w(A)\left(\prod_{k=1}^{n-1}\sqrt{\frac{1-a_k}{1+a_{k+1}}}\right)\sqrt{(1-a_{n-1})(1+a_n)}\sqrt{\frac{1+a_n}{1-a_{n-1}}} \\ &= w(A)y_n(1+a_n) \\ &= 2w(A)y_n. \end{aligned}$$

This shows that $(\mathrm{Re}\ A)y = w(A)y$. Since

$$\langle A\frac{y}{\|y\|}, \frac{y}{\|y\|}\rangle = \langle(\mathrm{Re}\ A)y, y\rangle\frac{1}{\|y\|^2} + i\ \mathrm{Im}\ \langle A\frac{y}{\|y\|}, \frac{y}{\|y\|}\rangle = w(A) + i\ \mathrm{Im}\ \langle A\frac{y}{\|y\|}, \frac{y}{\|y\|}\rangle$$

is in the circular disc $W(A) = \{z \in \mathbb{C} : |z| \le w(A)\}$, we infer that $\langle A(y/\|y\|), y/\|y\|\rangle = w(A)$ as asserted. □

Our construction of the matrix A in Theorem 3.5 is based on the following lemma. Here, for any real t, let $\lfloor t \rfloor$ denote the largest integer which is less than or equal to t.

Lemma 3.8. *For any positive integers l and m, let*

$$p(z) = (1+z)^l(1+z+\cdots+z^m) \equiv 1 + \alpha_1 z + \cdots + \alpha_{l+m-1}z^{l+m-1} + z^{l+m}.$$

Then the following hold:

(a) *$p(z)$ is a self-inversive polynomial, that is, its coefficients satisfy $\alpha_j = \alpha_{l+m-j}$ for all j, $1 \le j \le l+m-1$.*
(b) $1 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_{\lfloor (l+m)/2 \rfloor}$.
(c) *There is a sequence $\{a_j\}_{j=1}^{l+m+1}$ with $a_1 = -1$, $-1 < a_j < 1$ for $2 \le j \le l+m$, and $a_{l+m+1} = 1$ such that $\alpha_j = \prod_{k=1}^{j}(1-a_k)/(1+a_{k+1})$ for $1 \le j \le l+m-1$ and $\prod_{k=1}^{l+m}(1-a_k)/(1+a_{k+1}) = 1$.*

Proof of (a) and (b). It is easily seen that

$$\alpha_j = \begin{cases} \sum_{k=0}^{j} \binom{l}{k} & \text{if } 1 \le j \le l, \\ \sum_{k=0}^{l} \binom{l}{k} & \text{if } l+1 \le j \le m, \\ \sum_{k=j-m}^{l} \binom{l}{k} & \text{if } m+1 \le j \le l+m-1, \end{cases} \quad \text{or} \quad \begin{cases} \sum_{k=0}^{j} \binom{l}{k} & \text{if } 1 \le j \le m, \\ \sum_{k=j-m}^{j} \binom{l}{k} & \text{if } m+1 \le j \le l, \\ \sum_{k=j-m}^{l} \binom{l}{k} & \text{if } l+1 \le j \le l+m-1 \end{cases} \tag{8}$$

depending on whether $l \le m$ or $l > m$. (a) and (b) follow immediately.

Note that (a) can also be proved by comparing the coefficients of $p(z)$ and $z^{l+m}\overline{p(1/\overline{z})}$. The equality of these two polynomials follows from the fact that the leading coefficient, the constant term and the moduli of the zeros of $p(z)$ are all equal to 1 (cf. [10, p. 17, Theorem 2.1.2]).

To prove (c), we need another lemma.

Lemma 3.9. *(a) If $1 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_{n-1}$, then there is a sequence $\{a_j\}_{j=1}^n$ with $a_1 = -1$ and $-1 < a_j < 1$ for $2 \le j \le n$ such that $\alpha_j = \prod_{k=1}^{j}(1-a_k)/(1+a_{k+1})$ for $1 \le j \le n-1$.*

(b) Let l and m be positive integers such that $l \ge m$ and $l+m$ is even. If

$$\alpha_j = \begin{cases} \sum_{k=0}^{j} \binom{l}{k} & \text{for } 1 \le j \le m, \\ \sum_{k=j-m}^{j} \binom{l}{k} & \text{for } m+1 \le j \le \frac{1}{2}(l+m), \end{cases}$$

then, letting a_j, $1 \le j \le (l+m)/2$, be as in (a) and $a_{((l+m)/2)+1} = 0$, we have $\alpha_j = \prod_{k=1}^{j}(1-a_k)/(1+a_{k+1})$ for $1 \le j \le (l+m)/2$.

Proof. (a) Let $a_1 = -1$, and define a_j, $2 \le j \le n$, inductively by

$$a_j = \left(\frac{1-a_{j-1}}{\alpha_{j-1}} \prod_{k=1}^{j-2} \frac{1-a_k}{1+a_{k+1}}\right) - 1.$$

Then $\alpha_j = \prod_{k=1}^{j}(1-a_k)/(1+a_{k+1})$ for all j. Now we show that $-1 < a_j < 1$ for $2 \le j \le n$ by induction. Since $1 < \alpha_1 = 2/(1+a_2)$, we have $-1 < a_2 < 1$. In general, if $-1 < a_{j_0} < 1$ for some j_0, $2 \le j_0 < n$, then

$$1 + a_{j_0+1} = \frac{1-a_{j_0}}{\alpha_{j_0}} \prod_{k=1}^{j_0-1} \frac{1-a_k}{1+a_{k+1}} = \frac{1-a_{j_0}}{\alpha_{j_0}} \alpha_{j_0-1} \le 1 - \alpha_{j_0} < 2,$$

from which we obtain $-1 < a_{j_0+1} < 1$. Thus $-1 < a_j < 1$ for all j as asserted.

(b) Letting $n = (l+m)/2$, we need only show that $\alpha_n = \alpha_{n-1}(1-a_n)$. This is done by first expressing $1-a_j$, $2 \le j \le n$, in terms of α_0 ($\equiv 1$), $\alpha_1, \ldots, \alpha_{j-1}$, namely,

$$1 - a_j = \frac{2}{\alpha_{j-1}}(\alpha_{j-1} - \alpha_{j-2} + \cdots + (-1)^{j-1}\alpha_0). \tag{9}$$

Indeed, for $j = 2$, we have

$$1 - a_2 = 2 - (1 + a_2) = 2 - \frac{2}{\alpha_1} = \frac{2(\alpha_1 - \alpha_0)}{\alpha_1}.$$

Assume next that (9) holds for all $j < j_0$, $2 \le j_0 \le n$. Then

$$\begin{aligned}1 - a_{j_0} &= 2 - (1 + a_{j_0}) = 2 - \frac{(1 - a_{j_0-1})\alpha_{j_0-2}}{\alpha_{j_0-1}} \\ &= 2 - \frac{2}{\alpha_{j_0-1}}(\alpha_{j_0-2} - \alpha_{j_0-3} + \cdots + (-1)^{j_0-2}\alpha_0) \\ &= \frac{2}{\alpha_{j_0-1}}(\alpha_{j_0-1} - \alpha_{j_0-2} + \cdots + (-1)^{j_0-1}\alpha_0).\end{aligned}$$

Hence (9) holds by induction. To complete the proof, we need the identity

$$\alpha_n = 2(\alpha_{n-1} - \alpha_{n-2} + \cdots + (-1)^{n-1}\alpha_0). \tag{10}$$

Indeed, we have

$$\begin{aligned}&(-1)^{n-m-1}\alpha_m + (-1)^{n-m}\alpha_{m-1} + \cdots + (-1)^{n-1}\alpha_0 \\ &\quad = (-1)^{n-m-1}\sum_{k=0}^{m}\binom{l}{k} + (-1)^{n-m}\sum_{k=0}^{m-1}\binom{l}{k} + \cdots + (-1)^{n-1}\alpha_0\end{aligned}$$

$$= \begin{cases} (-1)^{n-m-1}[\binom{l}{0} + \binom{l}{2} + \cdots + \binom{l}{m}] & \text{if } m \text{ is even,} \\ (-1)^{n-m-1}[\binom{l}{1} + \binom{l}{3} + \cdots + \binom{l}{m}] & \text{if } m \text{ is odd,} \end{cases} \tag{11}$$

and

$$\begin{aligned}&\alpha_{n-1} - \alpha_{n-2} + \cdots + (-1)^{n-m-2}\alpha_{m+1} \\ &\quad = \sum_{k=n-m-1}^{n-1}\binom{l}{k} - \sum_{k=n-m-2}^{n-2}\binom{l}{k} + \cdots + (-1)^{n-m-2}\sum_{k=1}^{m+1}\binom{l}{k} \\ &\quad = \begin{cases} [\binom{l}{m+2} + \binom{l}{m+4} + \cdots + \binom{l}{n-1}] - [\binom{l}{1} + \binom{l}{3} + \cdots + \binom{l}{n-m-2}] & \text{if } n-m-1 \text{ is even,} \\ [\binom{l}{m+3} + \binom{l}{m+5} + \cdots + \binom{l}{n-1}] - [\binom{l}{2} + \binom{l}{4} + \cdots + \binom{l}{n-m-2}] & \\ \quad + [\binom{l}{1} + \binom{l}{2} + \cdots + \binom{l}{m+1}] & \text{if } n-m-1 \text{ is odd.} \end{cases}\end{aligned} \tag{12}$$

For m and $n - m - 1$ both even, adding (11) and (12) yields

$$\begin{aligned}&2(\alpha_{n-1} - \alpha_{n-2} + \cdots + (-1)^{n-1}\alpha_0) \\ &\quad = 2\left[\binom{l}{n-1} + \binom{l}{n-3} + \cdots + \binom{l}{0}\right] - 2\left[\binom{l}{n-m-2} + \binom{l}{n-m-4} + \cdots + \binom{l}{1}\right] \\ &\quad = 2\left[\sum_{k=0}^{n-1}\binom{l-1}{k} - \sum_{k=0}^{n-m-2}\binom{l-1}{k}\right] \\ &\quad = 2\left[\sum_{k=n-m-1}^{n-1}\binom{l-1}{k}\right],\end{aligned}$$

where the second equality follows by using the identity

$$\binom{l}{k} = \binom{l-1}{k} + \binom{l-1}{k-1}.$$

On the other hand, we also have

$$\begin{aligned}\alpha_n &= \sum_{k=n-m}^{n} \binom{l}{k} = \sum_{k=n-m}^{n} \left[\binom{l-1}{k} + \binom{l-1}{k-1} \right] \\ &= 2\left[\sum_{k=n-m-1}^{n-1} \binom{l-1}{k} \right] - \binom{l-1}{n-m-1} + \binom{l-1}{n} \\ &= 2\left[\sum_{k=n-m-1}^{n-1} \binom{l-1}{k} \right].\end{aligned}$$

This shows that (10) is indeed true in this case. For other parities of m and $n-m-1$, analogous arguments as above show that (10) also holds. This completes the proof. □

Proof of Lemma 3.8(c). We need only check that there is a sequence $\{a_j\}_{j=1}^{\lfloor (l+m)/2 \rfloor + 1}$ with $a_1 = -1$, $-1 < a_j < 1$ for $2 \le j \le \lfloor (l+m)/2 \rfloor + 1$, and, in addition, $a_{\lfloor (l+m)/2 \rfloor + 1} = 0$ if $l+m$ is even such that $\alpha_j = \prod_{k=1}^{j} (1-a_k)/(1+a_{k+1})$ for $1 \le j \le \lfloor (l+m)/2 \rfloor$. Indeed, if this is the case, then, letting $a_j = -a_{l+m+2-j}$ for $\lfloor (l+m)/2 \rfloor + 2 \le j \le l+m+1$, we obtain, by Lemma 3.8(a), that

$$\alpha_j = \alpha_{l+m-j} = \prod_{k=1}^{l+m-j} \frac{1-a_k}{1+a_{k+1}} = \prod_{k=1}^{j} \frac{1-a_k}{1+a_{k+1}}$$

for $\lfloor (l+m)/2 \rfloor + 1 \le j \le l+m-1$ and $\prod_{k=1}^{l+m} (1-a_k)/(1+a_{k+1}) = 1$ as required.

Now consider the case of odd $l+m$. Since $1 < \alpha_1 \le \alpha_2 \le \cdots \le \alpha_{\lfloor (l+m)/2 \rfloor}$ by Lemma 3.8(b), Lemma 3.9(a) yields a sequence $\{a_j\}_{j=1}^{\lfloor (l+m)/2 \rfloor + 1}$ with the properties in the preceding paragraph.

Next assume that $l+m$ is even. If $l \le m$, then (8) yields that $\alpha_j = \sum_{k=0}^{j} \binom{l}{k}$ for $1 \le j \le l$. Hence Lemma 3.9(b) (with $l = m$) guarantees the existence of $\{a_j\}_{j=1}^{l+1}$ with $a_1 = -1$, $-1 < a_j < 1$ for $2 \le j \le l$, and $a_{l+1} = 0$ such that $\alpha_j = \prod_{k=1}^{j} (1-a_k)/(1+a_{k+1})$ for $1 \le j \le l$. If $l \le m-1$, then, letting $a_j = 0$ for $l+2 \le j \le ((l+m)/2) + 1$, we have, by Lemma 3.8(a) and (8),

$$\alpha_j = \alpha_{l+m-j} = \prod_{k=1}^{l+m-j} \frac{1-a_k}{1+a_{k+1}} = \prod_{k=1}^{l} \frac{1-a_k}{1+a_{k+1}} = \alpha_l$$

for $l+1 \le j \le (l+m)/2$. This shows that the required properties for $\{\alpha_j\}_{j=1}^{((l+m)/2)+1}$ in the first paragraph are satisfied, and thus we are done.

Finally, consider $l > m$. In this case, we have

$$\alpha_j = \begin{cases} \sum_{k=0}^{j} \binom{l}{k} & \text{if } 1 \le j \le m, \\ \sum_{k=j-m}^{j} \binom{l}{k} & \text{if } m+1 \le j \le (l+m)/2 \end{cases}$$

by (8). Lemma 3.9(b) yields a sequence $\{a_j\}_{j=1}^{((l+m)/2)+1}$, which satisfies the required properties in the first paragraph. This completes the proof.

Now we are ready to prove Theorem 3.5.

Proof of Theorem 3.5. For the given n and k, let

$$p(z) = (1+z)^{n-k}(1+z+\cdots+z^{k-1}) \equiv 1+\alpha_1 z+\cdots+\alpha_{n-2}z^{n-2}+z^{n-1}.$$

Let $\{a_j\}_{j=1}^n$ be the sequence given in Lemma 3.8(c) with $a_1=-1$, $-1<a_j<1$ for $2\le j\le n-1$, $a_n=1$ such that $\alpha_j=\prod_{k=1}^j(1-a_k)/(1+a_{k+1})$ for $1\le j\le n-2$ and $\prod_{k=1}^{n-1}(1-a_k)/(1+a_{k+1})=1$, let $w_j=\sqrt{(1-a_j)(1+a_{j+1})}$ for $1\le j\le n-1$, and let

$$A=\begin{pmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ 0 & & & 0 \end{pmatrix}.$$

If $y=[1,\ \sqrt{\alpha_1},\ldots,\sqrt{\alpha_{n-2}},\ \ 1]^T$, then $\langle Ay,y\rangle=\|y\|^2w(A)$ by Lemma 3.7. Hence, according to Lemma 3.6(d), $k(A)$ equals the maximum number of $\theta_1,\ldots,\theta_k$ in $[0,2\pi)$ for which $e^{i(\theta_j-\theta_l)}$ is a zero of $p(z)$ for all j and l, $1\le j\ne l\le k$. Since the zeros of $p(z)$ are $-1,\ \omega_k,\ \omega_k^2,\ldots,\omega_k^{k-1}$, one maximal choice of the θ_j's is $2\pi j/k$, $0\le j\le k-1$, and thus $k(A)=k$ as required.

Our final result is a characterization of the n-by-n weighted shift matrices A with exactly one zero weight for which $k(A)=n-1$.

Theorem 3.10. *For $n\ge 3$, let*

$$A=\begin{pmatrix} 0 & w_1 & & \\ & 0 & \ddots & \\ & & \ddots & w_{n-1} \\ & & & 0 \end{pmatrix}$$

with $w_j\ne 0$ for all j and $w(A)=1$. Then $k(A)=n-1$ if and only if either

(a) *n is even, $|w_1|=|w_{n-1}|=\sqrt{2}$ and $|w_2|=\cdots=|w_{n-2}|=1$, or*
(b) *n is odd, and $|w_1|=2/\sqrt{1+\alpha}$, $|w_{2j}|=2\alpha/(1+\alpha)$, $|w_{2j+1}|=2/(1+\alpha)$ for $1\le j\le (n-3)/2$, and $|w_{n-1}|=2\sqrt{\alpha/(1+\alpha)}$ for some $\alpha>0$.*

Proof. We may assume that $w_j>0$ for all j.

(a) If n is even and $w_1=w_{n-1}=\sqrt{2}$ and $w_2=\cdots=w_{n-2}=1$, then $x\equiv(1/\sqrt{n-1})[1/\sqrt{2},\ 1,\ldots,1,\ 1/\sqrt{2}]^T$ is the unique unit vector in $\mathbb{C}^n$ with positive components such that $\langle Ax,x\rangle=1=w(A)$. Hence, by Lemma 3.6(d), $k(A)$ equals the maximum number of $\theta_1,\ldots,\theta_k$ in $[0,2\pi)$ for which $e^{i(\theta_j-\theta_l)}$ is a zero of the polynomial $p(z)=(1/(2(n-1)))(1+2z+\cdots+2z^{n-2}+z^{n-1})$ for all $j\ne l$. Since the zeros of $p(z)$ are -1 and ω_{n-1}^j, $1\le j\le n-2$, we infer that a maximal choice of the θ_j's are 0 and $2\pi j/(n-1)$, $1\le j\le n-2$, and hence $k(A)=n-1$.

Conversely, if n is even and $k(A)=n-1$, then let $x=[x_1,\ldots,x_n]^T$ be the unit vector in $\mathbb{C}^n$ with $x_j>0$ for all j such that $\langle Ax,x\rangle=w(A)=1$, and let $p(z)=x_1^2+x_2^2z+\cdots+x_n^2z^{n-1}$. Since $x'\equiv[x_1,\ e^{i\pi}x_2,\ldots,e^{(n-1)i\pi}x_n]^T$ is a unit vector satisfying $\langle Ax',x'\rangle=-1$ and is orthogonal to x (because they are eigenvectors of Re A corresponding to the eigenvalues -1 and 1, respectively), by Lemma 3.6(d), -1 is a zero of $p(z)$. On the other hand, since $k(A)=n-1$ and $p(1)=1\ne 0$, there are $\theta_1,\ldots,\theta_{n-1}$ with $0\le\theta_1<\theta_2<\cdots<\theta_{n-1}<2\pi$ such that $e^{i(\theta_j-\theta_l)}$, $1\le j\ne l\le n-1$, are all zeros of $p(z)$. Note that, except -1, the zeros of the real polynomial $p(z)$ appear in conjugate pairs. Thus for each fixed l, $1 \le l \le n-1$, the zeros of $p(z)$ are exactly -1 together with $e^{i(\theta_j-\theta_l)}$, $1 \le j \ne l \le n-1$. Since the latter are counterclockwise around the unit circle, we have, in particular, that $e^{i(\theta_2-\theta_1)} = e^{i(\theta_3-\theta_2)}$ and $e^{i(\theta_3-\theta_1)} = e^{i(\theta_4-\theta_2)}$. Hence

$$(\theta_3-\theta_1)-(\theta_2-\theta_1)=(\theta_4-\theta_2)-(\theta_3-\theta_2)=\theta_4-\theta_3=(\theta_4-\theta_1)-(\theta_3-\theta_1).$$

In a similar fashion, we may show that $(\theta_{j+1}-\theta_1)-(\theta_j-\theta_1)=(\theta_{j+2}-\theta_1)-(\theta_{j+1}-\theta_1)$ for all j, $1 \le j \le n-2$ ($\theta_n \equiv 2\pi+\theta_1$). Therefore, $e^{i(\theta_{j+1}-\theta_1)}$, $1 \le j \le n-2$, are equally distributed over the unit circle, that is, $e^{i(\theta_{j+1}-\theta_1)} = \omega_{n-1}^j$ for all j. It follows that

$$p(z)=\frac{1}{2(n-1)}(1+z)(1+z+\cdots+z^{n-2})=\frac{1}{2(n-1)}(1+2z+\cdots+2z^{n-2}+z^{n-1}),$$

that is, $x_1 = x_n = 1/\sqrt{2(n-1)}$ and $x_2=\cdots=x_{n-1}=1/\sqrt{n-1}$. Let $\{a_j\}_{j=1}^n$ be the sequence given by Lemma 3.7 such that $a_1=-1$, $-1<a_j<1$ for all j, $2 \le j \le n-1$, $a_n=1$ and $w_j=\sqrt{(1-a_j)(1+a_{j+1})}$ for all j. Moreover, if $y_1=1$ and $y_j=\prod_{k=1}^{j-1}\sqrt{(1-a_k)/(1+a_{k+1})}$ for $2 \le j \le n$, then x is the normalized vector of $y \equiv [y_1,\ldots,y_n]^T$. Thus $\prod_{k=1}^{j-1}(1-a_k)/(1+a_{k+1})=2$ for all j, $2 \le j \le n-1$, from which we obtain $a_j=0$ for $2 \le j \le n-1$. Hence $w_1=w_{n-1}=\sqrt{2}$ and $w_2=\cdots=w_{n-2}=1$ as asserted.

(b) Assume that n is odd. We may argue analogously as before except that this time -1 and $\omega_{n-1}^{(n-1)/2}$ coincide. So, for one direction, if the w_j's are of the asserted form for some $\alpha>0$, then $x=(1/\sqrt{(n-1)(1+\alpha)})[1,\ \sqrt{1+\alpha},\ldots,\sqrt{1+\alpha},\ \sqrt{\alpha}]^T$ is the unique unit vector in $\mathbb{C}^n$ with $\langle Ax,x\rangle = 1 = w(A)$, and $p(z)=(1/((n-1)(1+\alpha)))(1+(1+\alpha)z+\cdots+(1+\alpha)z^{n-2}+\alpha z^{n-1})$ has zeros $-1/\alpha$ and ω_{n-1}^j, $1 \le j \le n-2$. From this, we infer as before that $k(A)=n-1$.

Conversely, if $k(A)=n-1$, let $x=[x_1,\ldots,x_n]^T$, $p(z)=x_1^2+x_2^2z+\cdots+x_n^2z^{n-1}$, and $\theta_1,\ldots,\theta_{n-1}$ with $0 \le \theta_1<\theta_2<\cdots<\theta_{n-1}<2\pi$ be as in (a). Then for each fixed l, $1 \le l \le n-1$, the $n-2$ distinct numbers $e^{i(\theta_j-\theta_l)}$, $1 \le j \ne l \le n-1$, are zeros of $p(z)$. Let the real β be the remaining zero of $p(z)$. As before, we have $e^{i(\theta_{j+1}-\theta_1)}=\omega_{n-1}^j$ for $1 \le j \le n-2$ and hence

$$\begin{aligned}p(z)&=\frac{1}{(n-1)(1-\beta)}(-\beta+z)(1+z+\cdots+z^{n-2})\\&=\frac{1}{(n-1)(1-\beta)}(-\beta+(1-\beta)z+\cdots+(1-\beta)z^{n-2}+z^{n-1}).\end{aligned}$$

In particular, this implies that $\beta<0$ and $x_1=\sqrt{-\beta/((n-1)(1-\beta))}$, $x_j=1/\sqrt{n-1}$ for $2 \le j \le n-1$, and $x_n=1/\sqrt{(n-1)(1-\beta)}$. Let $\alpha=-1/\beta$ and let $\{a_j\}_{j=1}^n$ be as given in Lemma 3.7. We can derive as before that the w_j's are of the form as asserted. □

Acknowledgement

We thank Hsin-Yi Lee for his suggestions, which lead to the present improved Proposition 2.10.

References

[1] J. Eldred, L. Rodman, I.M. Spitkovsky, Numerical ranges of companion matrices: flat portions on the boundary, Linear Multilinear Algebra 60 (2012), in press.

[2] M.R. Embry, The numerical range of an operator, Pacific J. Math. 32 (1970) 647–650.

[3] J.S. Fang, C.-L. Jiang, P.Y. Wu, Direct sums of irreducible operators, Studia Math. 155 (2003) 37–49.

[4] H.-L. Gau, M.C. Tsai, H.-C. Wang, Weighted shift matrices: unitary equivalence, reducibility and numerical ranges, preprint, arXiv: 1206.1975.

[5] H.-L. Gau, P.Y. Wu, Numerical ranges and compressions of S_n-matrices, Oper. Matrices, in press.
[6] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, Cambridge, 1991.
[7] D.S. Keeler, L. Rodman, I.M. Spitkovsky, The numerical range of 3×3 matrices, Linear Algebra Appl. 252 (1997) 115–139.
[8] D.P. Kimsey, H.J. Woerdeman, Minimal normal and commuting completions, Int. J. Inf. Syst. Sci. 4 (2008) 50–59.
[9] C.-K. Li, B.-S. Tam, P.Y. Wu, The numerical range of a nonnegative matrix, Linear Algebra Appl. 350 (2002) 1–23.
[10] G.V. Milovanović, D.S. Mitrinović, Th. M. Rassias, Topics in Polynomials: Extremal Problems, Inequalities, Zeros, World Scientific, Singapore, 1994.
[11] H. Nakazato, Quartic curves associated to 4 by 4 matrices, Sci. Rep. Hirosaki Univ. 43 (1986) 209–221.
[12] M.C. Tsai, Numerical ranges of weighted shift matrices with periodic weights, Linear Algebra Appl. 435 (2011) 2296–2302.
[13] M.C. Tsai, P.Y. Wu, Numerical ranges of weighted shift matrices, Linear Algebra Appl. 435 (2011) 243–254.
[14] K.-Z. Wang, P.Y. Wu, Numerical ranges of weighted shifts, J. Math. Anal. Appl. 381 (2011) 897–909.
[15] H.-Y. Zhang, Y.-N. Dou, M.-F. Wang, H.-K. Du, On the boundary of numerical ranges of operators, Appl. Math. Lett. 24 (2011) 620–622.