We show that if $A$ is an $n$-by-$n$ ($n \geq 3$) matrix of the form
\[
\begin{bmatrix}
0 & a_1 & 0 & \cdots & 0 \\
0 & 0 & a_2 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & a_{n-1} & 0 \\
\end{bmatrix},
\]
then the boundary of its numerical range contains a line segment if and only if the $a_j$'s are nonzero and the numerical ranges of the $(n-1)$-by-$(n-1)$ principal submatrices of $A$ are all equal. For $n = 3$, this is the case if and only if $|a_1| = |a_2| = |a_3| \neq 0$, in which case $W(A)$, the numerical range of $A$, is the equilateral triangular region with vertices the three cubic roots of $a_1 a_2 a_3$. For $n = 4$, the condition becomes $|a_1| = |a_3| \neq 0$ and $|a_2| = |a_4| \neq 0$, in which case $W(A)$ is the convex hull of two (degenerate or otherwise) ellipses.

© 2011 Elsevier Inc. All rights reserved.
Theorem 1. Let $A$ be an $n$-by-$n$ $(n \geq 2)$ weighted shift matrix $A$ is one of the form

\[
\begin{bmatrix}
0 & a_1 \\
0 & \ddots \\
a_n & 0 \\
\end{bmatrix},
\]

where the $a_i$'s, called the weights of $A$, are complex numbers. The purpose of this paper is to study the numerical ranges of such matrices.

Recall that for any $n$-by-$n$ complex matrix $A$, its numerical range $W(A)$ is by definition the subset \{ $\langle Ax, x \rangle : x \in \mathbb{C}^n$, $\|x\| = 1$ \} of the plane, where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ denote the standard inner product and its associated norm in $\mathbb{C}^n$. It is known that $W(A)$ is a nonempty compact convex subset of $\mathbb{C}$. For its other properties, the reader may consult [7, Chapter 1] or [5].

The numerical ranges of certain weighted shift matrices are easier to determine. For example, if any of the weights of an $n$-by-$n$ weighted shift matrix $A$ is zero, then its numerical range is a circular disc centered at the origin. On the other hand, if all the weights of $A$ have equal (nonzero) moduli, then $W(A)$ is a polygonal region with its boundary a regular $n$-gon. The main theorem below gives necessary and sufficient conditions for the boundary of $W(A)$ to have a line segment. More specifically, it is shown that this is the case if and only if the weights are all nonzero and all its $(n-1)$-by-$(n-1)$ principal submatrices have identical numerical ranges (which are necessarily circular discs centered at the origin). This is then used to give a characterization of such a matrix $A$ of size 4 with line segments on $\partial W(A)$ purely in terms of its weights, namely, for

\[
A = \begin{bmatrix}
0 & a_1 \\
0 & a_2 \\
0 & a_3 \\
a_4 & 0 \\
\end{bmatrix},
\]

the boundary of $W(A)$ has a line segment if and only if $|a_1| = |a_3| \neq 0$ and $|a_2| = |a_4| \neq 0$. Along the way, we also prove various properties of the numerical ranges of such matrices.

In the literature, there are works on the numerical ranges and numerical radii of weighted shift matrices and operators. For example, [12, Lemma 2] gives a method to compute the numerical radius of a weighted shift matrix with at least one zero weight. The authors in Refs. [11, 12, 1] discuss properties of the numerical ranges and numerical radii of weighted shifts on $l^2$ with periodic or geometric weights.

For an $n$-by-$n$ matrix $A$, let $A^T$ denote its transpose, $A^*$ its adjoint and Re $A$ its real part $(A + A^*)/2$. For $1 \leq i_1 < \cdots < i_m \leq n$, let $A[i_1, \ldots, i_m]$ denote the $(n-m)$-by-$(n-m)$ principal submatrix of $A$ obtained by deleting its rows and columns indexed by $i_1, \ldots, i_m$. The numerical radius $w(A)$ and generalized Crawford number $w_0(A)$ of $A$ are, by definition, max $\{|z| : z \in W(A)\}$ and min $\{|z| : z \in \partial W(A)\}$, respectively. A diagonal matrix with diagonals $a_1, \ldots, a_n$ is denoted by $\text{diag}(a_1, \ldots, a_n)$. Our basic reference for properties of matrices is [6].

For any nonzero complex number $z = x + iy$ ($x$ and $y$ real), $\arg z$ is the angle $\theta$, $0 \leq \theta < 2\pi$, from the positive $x$-axis to the vector $(x, y)$. If $z = 0$, then $\arg z$ can be an arbitrary real number. In the following, let $B(r) = \{z \in \mathbb{C} : |z| \leq r\}$ for $r > 0$ and $\omega_n = e^{2\pi i/n}$ for $n \geq 1$. For any subset $\Delta$ of $\mathbb{C}$, $\Delta^\triangle$ denotes its convex hull.

The main result of this paper is the following.

**Theorem 1.** Let $A$ be an $n$-by-$n$ $(n \geq 3)$ weighted shift matrix with weights $a_1, \ldots, a_n$. Then $\partial W(A)$ has a line segment if and only if the $a_i$’s are nonzero and $W(A[1]) = \cdots = W(A[n])$. In this case, $W(A[j])$ is the circular disc centered at the origin with radius $w_0(\theta)$, the line segment lies on one of the lines $x \cos \theta_k + y \sin \theta_k = w_0(A)$, where $\theta_k = (2k + 1)\pi + \sum_{j=1}^n \arg a_j)/n$, $0 \leq k < n$, and there are exactly $n$ line segments on $\partial W(A)$. 

The necessity of this theorem is easier to establish. It follows from the next two results.

**Lemma 2.** Let $A$ and $B$ be $n$-by-$n$ ($n \geq 2$) weighted shift matrices with weights $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, respectively.

1. If, for some fixed $k$, $1 \leq k \leq n$, $b_j = a_{k+j}$ ($a_{n+j} \equiv a_j$) for all $j$, then $A$ is unitarily equivalent to $B$.
2. If $|a_j| = |b_j|$ for all $j$, then $A$ is unitarily equivalent to $e^{i\psi k}B$, where $\psi_k = (2k\pi + \sum_{j=1}^n (\arg a_j - \arg b_j))/n$ for $0 \leq k < n$.

**Proof.**

1. If $U$ is the $n$-by-$n$ weighted shift matrix with weights $1, \ldots, 1$, then $U$ is unitary and $AU^{n-k} = U^{n-k}B$. This proves the unitary equivalence of $A$ and $B$.
2. If $U = \text{diag}(e^{i\phi_1}, \ldots, e^{i\phi_n})$, where $\phi_1 = 0$ and $\phi_j = \phi_{j-1} + (\arg b_{j-1} - \arg a_{j-1}) + \psi_k$ for $2 \leq j \leq n$, then $U$ is unitary and $AU = U(e^{i\psi_k}B)$. □

Using this lemma, we can deduce many properties of the numerical range of a weighted shift matrix. These are gathered together in the following. Note that some of them have already been obtained in [8], the Ph.D. dissertation of Issos on irreducible nonnegative matrices. For example, (1) below is a special case of [8, Theorem 6] and (4) follows from [8, Theorem 7].

**Proposition 3.** Let $A$ be an $n$-by-$n$ ($n \geq 3$) weighted shift matrix with weights $a_1, \ldots, a_n$, and let $\theta = (\sum_{j=1}^n \arg a_j)/n$.

1. $W(A)$ satisfies $W(A) = \omega_n W(A)$.
2. $W(A)$ is symmetric with respect to the lines $y = \pm x \tan \gamma_k$, where $\gamma_k = (k\pi/n) + \theta$ for $0 \leq k < n$.
3. The following conditions are equivalent:
   a. $a_j = 0$ for some $j$,
   b. $A$ is unitarily equivalent to $e^{i\theta}A$ for all real $\theta$, and
   c. $W(A)$ is a circular disc centered at the origin.
4. If $a_j \neq 0$ for all $j$, then $\partial W(A)$ intersects $\partial B(0; w(A))$ (resp., $\partial B(0; w_0(A))$) at exactly the $n$ points $w(A)e^{i\alpha_k}$ (resp., $w_0(A)e^{i\beta_k}$), where $\alpha_k = (2k\pi/n) + \theta$ (resp., $\beta_k = ((2k-1)\pi/n) + \theta$) for $0 \leq k < n$.
5. $w(A) \leq w_0(A) \sec(\pi/n)$ and $B(0; w_0(A)) \subseteq W(A) \subseteq w_0(A) \left(\sec\frac{\pi}{n}\right) e^{i\theta}\{1, \omega_n, \ldots, \omega_n^{n-1}\}^\wedge$.
6. If $\partial W(A)$ has a line segment $L$, then $\text{dist}(0, L) = w_0(A)$, $L$ lies on one of the lines $x \cos \beta_k + y \sin \beta_k = w_0(A)$, where the $\beta_k$’s are as in (4), and there are exactly $n$ line segments on $\partial W(A)$.

**Proof.** (1) Letting $B = A$ and $k = 1$ in Lemma 2 (2) yields the unitary equivalence of $A$ and $\omega_n A$. The assertion $W(A) = \omega_n W(A)$ then follows. 

(2) By Lemma 2 (2), $A$ is unitarily equivalent to $e^{i\alpha_k}B$ (resp., $e^{i\beta_k}C$), where $B$ (resp., $C$) is the $n$-by-$n$ weighted shift matrix with weights $|a_1|, \ldots, |a_n|$ (resp., $|a_1|, \ldots, |a_{n-1}|$ and $-|a_n|$) and $\alpha_k = (2k\pi/n) + \theta$ (resp., $\beta_k = ((2k-1)\pi/n) + \theta$) for $0 \leq k < n$. Since the numerical ranges of the real matrices $B$ and $C$ are symmetric with respect to the $x$-axis, the unitary equivalences above yield the symmetry of $W(A)$ with respect to the lines $l_k : y = x \tan \alpha_k$ (resp., $l_k : y = x \tan \beta_k$), $0 \leq k < n$. Note that for odd $n$, $l_k$ and $L_{(k+(n+1)/2)}(\text{mod } n)$ coincide while for even $n$, $l_k$ and $L_{(k+(n/2))}(\text{mod } n)$ (resp., $L_{k} \text{ and } L_{(k+(n/2))}(\text{mod } n)$) coincide. Thus $W(A)$ is symmetric with respect to the $n$ distinct lines among them, namely, the lines $y = x \tan \gamma_k$, $0 \leq k < n$.

(3) If (a) holds, then the $\psi_k$’s in Lemma 2 (2) can be arbitrary. Letting $B = A$ in there, we obtain (b). The implication (b) $\Rightarrow$ (c) is trivial. To prove (c) $\Rightarrow$ (a), note that $0$, the center of the circular disc $W(A)$,
is an eigenvalue of $A$ (cf. [10, Theorem 4.2]). Hence $\det A = (-1)^{n+1}a_1, \ldots, a_n = 0$, which shows that $a_j = 0$ for some $j$.

(4) We prove the assertion for $w(A)$. As shown in (2) above, $A$ is unitarily equivalent to $e^{i\alpha_k}B$ for $0 \leq k < n$. Since $B$ is a nonnegative matrix, $w(B)$ belongs to $W(B)$ (cf. [9, Proposition 3.3]). Hence $w(A)$ is in $e^{-i\alpha_k}W(A)$ and thus $w(A)e^{i\alpha_k}$ is in $W(A)$ for $0 \leq k < n$. If there are more than $n$ points in $\partial W(A) \cap \partial B(0; w(A))$, then Anderson’s theorem (cf. [13, Lemma 6] or [10, Theorem 4.12]) implies that $W(A)$ and $B(0; w(A))$ coincide. This would contradict the assertions in (3) since $a_j \neq 0$ for all $j$. Hence $\partial W(A) \cap \partial B(0; w(A))$ consists of exactly the $n$ points $w(A)e^{i\alpha_k}, 0 \leq k < n$.

For the assertion on $w_0(A)$, as in (2) above, $A$ is unitarily equivalent to $e^{i\beta_k}C$ for $0 \leq k < n$. Let $w_0(C)e^{i\beta}$ ($\beta$ real) be any point in $\partial W(C) \cap \partial B(0; w_0(C))$. Then $w_0(A)e^{i(\beta+\beta_k)}$ is in $\partial W(A) \cap \partial B(0; w_0(A))$ for $0 \leq k < n$. In particular, $W(A)$ contains the regular $n$-polygonal region with these points as vertices. From the convexity of $W(A)$ and the minimality of $w_0(A)$, we infer that $B(0; w_0(A)) \subseteq W(A)$. If there are more than $n$ points in $\partial W(A) \cap \partial B(0; w_0(A))$, then [3, Theorem 2.5 (b)] implies that $\partial W(A)$ contains at least one arc of $\partial B(0; w_0(A))$. The $n$-symmetry of $W(A)$ from (1) then yields that either $\partial W(A) \neq \partial B(0; w_0(A))$ or $\partial W(A)$ contains $n$ arcs of $\partial B(0; w_0(A))$. The former implies that $W(A) = B(0; w_0(A))$, which contradicts (3), while the latter would contradict [3, Theorem 2.5 (a)]. Thus $\partial W(A) \cap \partial B(0; w_0(A))$ consists of exactly $n$ points. These points must be on the lines $y = x \tan \gamma_k$ in (2) for otherwise by symmetry there will be at least $2n$ points in $\partial W(A) \cap \partial B(0; w_0(A))$ contradicting what we have proved above. We conclude that the points in $\partial W(A) \cap \partial B(0; w_0(A))$ are exactly $w_0(A)e^{i\beta_k}, 0 \leq k < n$.

(5) Since the points $w(A)e^{i\alpha_k}, 0 \leq k < n$, are in $W(A)$ by (4), the regular $n$-polygonal region $R$ whose vertices are these points is contained in $W(A)$. Hence

$$w_0(A) \geq \text{dist}(0, R) = w(A)\left\{\frac{1}{2}e^{i\alpha_0} + e^{i\alpha_1}\right\} = w(A)\cos\frac{\pi}{n}.$$ 

This proves that $w(A) \leq w_0(A)\sec(\pi/n)$.

The containment $B(0; w_0(A)) \subseteq W(A)$ was already noted in the proof of (4). For the other direction, note that if $u$ is any point of $W(A)$ which is in a different half-plane, determined by the line $L$ connecting $w_0(A)\sec(\pi/n)e^{i\alpha_0}$ and $w_0(A)\sec(\pi/n)e^{i\alpha_1}$, from the origin, then, by (2), its symmetric point $u'$ with respect to the line connecting $0$ and $w_0(A)e^{i\beta_1}$ is also in $W(A)$. Thus $(u + u')/2$ is in $W(A)$, which would contradict the fact that $w_0(A)e^{i\beta_1}$ is on the boundary of $W(A)$. This shows that $W(A)$ is contained in the same half-plane of $L$ as the origin. The $n$-symmetry of $W(A)$ from (1) then yields that

$$W(A) \subseteq w_0(A)\left\{\sec\frac{\pi}{n}e^{i\alpha_k} : 0 \leq k < n\right\}^\wedge = w_0(A)\left\{\sec\frac{\pi}{n}e^{i\theta} : \omega^k_n : 0 \leq k < n\right\}^\wedge.$$ 

(6) If $L$ is a line segment on $\partial W(A)$, then $L$ intersects $\partial W(A[j])$ for every $j$, $1 \leq j \leq n$ (cf. [2, Lemma 5]). Since $W(A[j]) \subseteq W(A)$ and $W(A[j])$ is a circular disc centered at the origin, we obtain $\text{dist}(0, L) = w_0(A[j]) \leq w_0(A)$ for every $j$. But $\text{dist}(0, L) \geq w_0(A)$ is obviously true. This shows that $\text{dist}(0, L) = w_0(A[1]) = \cdots = w_0(A[n]) = w_0(A)$. It follows from (4) that $L$ must lie on one of the lines $x \cos \beta_k + y \sin \beta_k = w_0(A), 0 \leq k < n$. Together with (1), this implies that there are exactly $n$ line segments on $\partial W(A)$. □

The necessity of Theorem 1 then follows easily from Proposition 3 (3) and (6).

As a side result, the next proposition gives conditions for a weighted shift matrix to have a regular polygonal region as its numerical range. The equivalence of some conditions below can also be derived from [8, Theorem 13].

**Proposition 4.** Let $A$ be a nonzero $n$-by-$n$ ($n \geq 3$) weighted shift matrix with weights $a_1, \ldots, a_n$. Then the following conditions are equivalent:

1. $A$ is normal,
2. $|a_1| = \cdots = |a_n|.$
(3) $A$ is unitarily equivalent to diag($\lambda$, $\lambda \omega_n$, ..., $\lambda \omega_n^{n-1}$), where $\lambda = (a_1, \ldots, a_n)^{1/n}$.

(4) $W(A)$ is a regular $n$-polygonal region with center at the origin and the distance from the center to its vertices equal to $|a_1, \ldots, a_n|^{1/n}$.

(5) $\partial W(A)$ has a nondifferentiable point, and

(6) $w(A) = w_0(A) \sec(\pi/n)$.

**Proof.** That $(1) \Rightarrow (2)$, $(3) \Rightarrow (4)$ and $(4) \Rightarrow (5)$ are trivial. To prove $(2) \Rightarrow (3)$, note that, under $(2)$, $A$ is unitarily equivalent to $|a_1| e^{i\theta} B$, where $B$ is the $n$-by-$n$ weighted shift matrix with weights $1, \ldots, 1$ by Lemma 2 (2). It is easily seen that $B$ is unitarily equivalent to diag(1, $\omega_n$, ..., $\omega_n^{n-1}$) and $|a_1| e^{i\theta} = |a_1| e^{i(\sum_{j=1}^n \arg a_j)/n} = (a_1, \ldots, a_n)^{1/n} = \lambda$. Hence $(3)$ follows. For $(5) \Rightarrow (1)$, if $\lambda$ is a nondifferentiable point of $\partial W(A)$, then so are $\lambda \omega_n^k, 0 \leq k < n$, by Proposition 3 (1). Since each of such points is a reducing eigenvalue of $A$, we obtain that $A$ is unitarily equivalent to diag($\lambda$, $\lambda \omega_n$, ..., $\lambda \omega_n^{n-1}$). In particular, $A$ is normal, that is, $(1)$ holds. Finally, if $(4)$ holds, then $(2)$ is true and hence

$$w(A) = |a_1, \ldots, a_n|^{1/n} = |a_1| = w_0(A) \sec(\pi/n),$$

that is, $(6)$ holds. Conversely, if $(6)$ is true, then Proposition 3 (5) says that $W(A) \subseteq w(A) e^{i\theta} \{1, \omega_n, \ldots, \omega_n^{n-1}\}$. But the vertices of this latter regular $n$-polygonal region, namely, $w(A) e^{i\theta} \omega_n^k, 0 \leq k < n$, are in $W(A)$ by Proposition 3 (4). Hence we must have $W(A) = w(A) e^{i\theta} \{1, \omega_n, \ldots, \omega_n^{n-1}\}$. Hence $\partial W(A)$ has nondifferentiable points, that is, $(5)$ holds. \[\square\]

We now proceed to prepare ourselves for the proof of the sufficiency of Theorem 1. This will be done in a series of lemmas and propositions. We start with the following.

**Lemma 5.** Let $A$ and $B$ be the $n$-by-$n$ ($n \geq 2$) weighted shift matrices with weights $a_1, \ldots, a_{n-1}, 0$ and $b_1, \ldots, b_{n-1}, 0$, respectively.

1. If $|a_j| \leq |b_j|$ for all $j$, then $W(A) \subseteq W(B)$.

2. If the $b_j$’s are nonzero, $|a_j| \leq |b_j|$ for all $j$ and $|a_k| < |b_k|$ for some $k$, then $W(A) \nsubseteq W(B)$.

3. If the $a_j$’s are nonzero, then $W(A[n]) \subsetneq W(A)$.

**Proof.** In view of Lemma 2 (2), we may assume that the $a_j$’s and $b_j$’s are all nonnegative. Since $W(A)$ and $W(B)$ are circular discs centered at the origin by Proposition 3 (3), the assertions in (1) and (2) are equivalent to $w(A) \leq w(B)$ and $w(A) < w(B)$, respectively. These in turn follow from [9, Corollary 3.6]. To prove (3), let $C = A[n] \oplus [0]$. Then $W(A[n]) = W(C) \subsetneq W(A)$ by (2). \[\square\]

The next lemma is needed for the proof of Proposition 7.

**Lemma 6.** If $A$ and $B$ are $n$-by-$n$ ($n \geq 2$) weighted shift matrices with weights $a_1, \ldots, a_{n-1}, 0$ and $a_n$, respectively, then $W(A) = W(B)$.

**Proof.** Since $W(A)$ and $W(B)$ are circular discs centered at the origin by Proposition 3 (3), we need only check that $w(A) = w(B)$. By Lemma 2 (2), we may assume that $a_j \geq 0$ for all $j$. Let $x = [x_1, \ldots, x_n]^T$ be a unit vector with nonnegative components such that $w(A) = \langle Ax, x \rangle$. Then

$$w(A) = \sum_{j=1}^{n-1} a_j x_{j+1} x_j = \langle B y, y \rangle \leq w(B),$$

where $y = [x_n, \ldots, x_1]^T$. Similarly, we have $w(B) \leq w(A)$. Thus $w(A) = w(B)$ as asserted. \[\square\]

As indicated by the referee, the preceding lemma can also be proven by noting, under $a_j \geq 0$ for all $j$, that Re $A$ and Re $B$ are unitarily equivalent:
Proposition 7. Let $A$ be an $n$-by-$n$ weighted shift matrix with weights $a_1, \ldots, a_n$. If $|a_1| = \cdots = |a_{n-3}|$ and $\partial W(A)$ has a line segment, then $|a_{n-2}| = |a_n| \neq 0$.

Proof. By Lemma 2 (2), we may assume that $a_j \geq 0$ for all $j$. Since $\partial W(A)$ has a line segment, we even have $a_j > 0$ by Proposition 3 (3). Let $A_1$ and $A_2$ be the $(n-1)$-by-$(n-1)$ weighted shift matrices with weights $a_1, \ldots, a_{n-3}, a_{n-2}, 0$ and $a_1, \ldots, a_{n-3}, a_n, 0$, respectively. Then $A_1 = A[n]$ and $W(A_2) = W(A_3)$, where $A_3$ is the $(n-1)$-by-$(n-1)$ weighted shift matrix with weights $a_n, a_{n-3}, \ldots, a_1, 0$. By Lemma 6. Since $a_1 = \cdots = a_{n-3}$, by Lemma 2 (1), $A_3$ is unitarily equivalent to $A[n-1]$. Thus $W(A_3) = W(A[n-1])$. Note that the existence of a line segment on $\partial W(A)$ guarantees that $W(A[n]) = W(A[n-1])$ by the necessity part of Theorem 1. We conclude that $W(A_1) = W(A_2)$. Therefore, $a_{n-2} = a_n$ by Lemma 5 (2). □

From Proposition 7, we can derive the following for weighted shift matrices of size 3 or 4: (1) a 3-by-3 weighted shift matrix $A$ with weights $a_1, a_2, a_3$ is such that $\partial W(A)$ contains a line segment if and only if $|a_1| = |a_2| = |a_3| \neq 0$, and (2) if the 4-by-4 weighted shift matrix $A$ with weights $a_1, a_2, a_3, a_4$ is such that $\partial W(A)$ contains a line segment, then $|a_1| = |a_3| \neq 0$ and $|a_2| = |a_4| \neq 0$. The necessity in (1) and (2) is a consequence of Proposition 7 and Lemma 2 (1). The sufficiency in (1) has already been proven in Proposition 4. Note that the condition in (2) is actually also sufficient, but its proof has to wait until the proving of Theorem 1 (cf. Proposition 12).

The next proposition is the major step in proving the sufficiency of Theorem 1.

Proposition 8. Let $A$ be an $n$-by-$n$ $(n \geq 3)$ weighted shift matrix with nonzero weights $a_1, \ldots, a_n$, and let $\theta = (\pi + \sum_{j=1}^{n} \arg a_j)/n$.

(1) If $W(A[j-1]) = W(A[j]) = W(A[j+1]) = B(0; r)$ for some $j$, $1 \leq j \leq n$ ($A[0] \equiv A[n]$ and $A[n+1] \equiv A[1]$) and some $r > 0$, then $r$ is either the largest or the second largest eigenvalue of $\Re(e^{-i\theta}A)$.

(2) If $W(A[1]) = \cdots = W(A[n]) = B(0; r)(r > 0)$, then $r = w_0(A)$ is the largest eigenvalue of $\Re(e^{-i\theta}A)$ with multiplicity at least two.

For the proof, we need the following lemma.

Lemma 9. Let $A$ be an $n$-by-$n$ $(n \geq 5)$ weighted shift matrix with nonzero real weights $a_1, \ldots, a_n$. For $1 \leq j \leq n - 2$, let $B = \Re A[n]$ be partitioned as

\[
\begin{bmatrix}
A_j & B_j \\
C_j & D_j
\end{bmatrix}
\]

with $A_j, B_j, C_j$ and $D_j$ of sizes $j$-by-$j$, $j$-by-$(n-j-1)$, $(n-j-1)$-by-$j$ and $(n-j-1)$-by-$(n-j-1)$, respectively. If $\lambda$ is the maximum eigenvalue of $B$, then $a_2^2, \ldots, a_{n-3}^2\lambda^2 = 4^{n-4} \det(\lambda I_{n-3} - A_{n-3}) \det(\lambda I_{n-3} - D_2)$.

Proof. Since the $a_j$'s are nonzero, $W(A[n])$ properly contains $W(A[j+1], \ldots, n])$ for any $j$, $1 \leq j \leq n - 2$, by Lemma 5 (3). Hence $\lambda$, being the radius of the circular disc $W(A[n])$, does not belong to $W(A[j+1], \ldots, n])$. In particular, $\lambda$ is not an eigenvalue of $A_j = \Re A[j+1, \ldots, n]$ and therefore $\lambda I_j - A_j$ is invertible for all $j$, $1 \leq j \leq n - 2$. Similarly, the same is true for $\lambda I_{n-j-1} - D_j$. Thus

\[
0 = \det(\lambda I_{n-1} - B)
= \det(\lambda I_j - A_j) \det((\lambda I_{n-j-1} - D_j) - (-C_j)(\lambda I_j - A_j)^{-1}(-B_j))
\]
Proof of Proposition 8. (1) We may assume, by Lemma 2 (1), that $W(A[n − 1]) = W(A[n]) = W(A[1]) = B(0; r)$. Also, by Lemma 2 (2), $A$ is unitarily equivalent to $e^{θD}C$, where $C$ is the $n$-by-$n$ weighted shift matrix with weights $|a_1|, \ldots, |a_{n−1}|, −|a_n|$. Then $w_0(A) = w_0(C)$ is in $∂W(C)$ by Proposition 3 (4) and $W(A[j]) = W(C[j])$ for all $j$. Thus $w_0(C)$ is the maximum eigenvalue of $Re C$ and $r$ is the maximum eigenvalue of $Re C[j]$ for $j = 1, n − 1$ and $n$. We now expand the determinant of $rI_n − Re C$ by minors along its $n$th row to obtain

$$\det(rI_n − Re C) = \frac{1}{2} |a_n|(-1)^{n+1}d_{n1} − \frac{1}{2} |a_{n−1}|(-1)^{2n−1}d_{n,n−1} + r \det(rI_{n−1} − Re C[n])$$

$$= \frac{1}{2} |a_n|(-1)^{n+1}d_{n1} − \frac{1}{2} |a_{n−1}|(-1)^{2n−1}d_{n,n−1},$$

where $(-1)^{n+j}d_{nj}$ denotes the cofactor of the $(n,j)$-entry of $Re C$ in $Re C, j = 1, n − 1$. The expansion of the determinant $d_{n1}$ (resp., $d_{n,n−1}$) along its first row (resp., its last row) yields

$$d_{n1} = \frac{(-1)^{n−1}}{2^{n−1}} |a_1, \ldots, a_{n−1}| + \frac{(-1)^n}{2} |a_n| \det C_1$$

$$\left(\text{resp., } d_{n,n−1} = \frac{1}{2^{n−1}} |a_1, \ldots, a_{n−2}a_n| − \frac{1}{2} |a_{n−1}| \det C_2\right),$$

where

$$C_1 = \begin{bmatrix} r & −|a_2|/2 \\ −|a_2|/2 & r & \ddots \\ \vdots & \ddots & \ddots & −|a_{n−2}|/2 \\ −|a_{n−2}|/2 & \cdots & −|a_{n−2}|/2 & r \end{bmatrix}$$
\[
\begin{pmatrix}
  r & -|a_1|/2 \\
  -|a_1|/2 & r \\
  \vdots & \ddots & \ddots \\
  -|a_{n-3}|/2 & \ddots & -|a_{n-3}|/2 \\
  -|a_{n-3}|/2 & \ddots & \ddots & \ddots \\
  -|a_{n-3}|/2 & \ddots & \ddots & \ddots & r \\
  \end{pmatrix}
\]

Hence
\[
\det(rI_n - \text{Re } C) = \frac{1}{2^n}|a_1, \ldots, a_n| - \frac{1}{4}|a_n|^2 \det C_1 + \frac{1}{2^n}|a_1, \ldots, a_n| - \frac{1}{4}|a_{n-1}|^2 \det C_2
\]
\[
= \frac{1}{2^{n-1}}|a_1, \ldots, a_n| - \frac{1}{4}|a_n|^2 \det C_1 - \frac{1}{4}|a_{n-1}|^2 \det C_2.
\]

(1) becomes
\[
\det(rI_n - \text{Re } C) = \frac{1}{2^n-1}|a_1, \ldots, a_n| - \frac{1}{4}|a_n|^2 |a_{n-1}|^2 \det D_2 - \frac{1}{4}|a_{n-1}|^2 |a_n|^2 \det D_2
\]
\[
= \frac{1}{2^n-1}|a_1, \ldots, a_n| - \frac{1}{8r}|a_{n-1}a_n|^2 |a_{n-2}|^2 |a_n| |(\det D_1)^{1/2} |a_1| |(\det D_2)^{1/2} |a_{n-1}|
\]
\[
= \frac{1}{2^n-1}|a_1, \ldots, a_n| - \frac{1}{8r}|a_{n-2}a_{n-1}a_n| |(\det D_1 \cdot \det D_2)^{1/2} = 0
\]
by Lemma 9. Hence \(\det(rI_n - \text{Re } (e^{-i\theta} A)) = 0\). Since \(r\) is the maximum eigenvalue of \(\text{Re } (e^{-i\theta} A)[1]\), this shows that it is either the largest or the second largest eigenvalue of \(\text{Re } (e^{-i\theta} A)\).

(2) From our assumption and the proof of (1), we have \(\det(rI_n - \text{Re } C[j]) = \det(rI_n - \text{Re } C) = 0\) for all \(j\), \(1 \leq j \leq n\). Thus if \(p(z) = \det(zI_n - \text{Re } C)\), then \(p'(r) = \sum_{j=1}^{n} \text{det}(rI_n - \text{Re } C[j]) = 0\) (cf. [6, p. 43, Problem 4]). This shows that the eigenvalue \(r\) of \(\text{Re } C\) has (algebraic) multiplicity at least two or, equivalently, \(\dim \ker(rI_n - \text{Re } C) \geq 2\). Since \(B(0; r) = W(C[n]) \subseteq W(C)\), we have \(r \leq \omega_0(C)\). If \(r < \omega_0(C)\), then we deduce from the facts that \(\omega_0(C)\) is the maximum eigenvalue of \(\text{Re } C\) and \(\dim \ker(rI_n - \text{Re } C) \geq 2\) that \(B(0; r) = W(C[n]) = W(C[n - 1, n])\). This contradicts Lemma 5 (3) since the \(a_j\)'s are nonzero. Hence we must have \(r = \omega_0(C) = \omega_0(A)\), which is the largest eigenvalue of \(\text{Re } (e^{-i\theta} A)\) with multiplicity at least two.

To prove the sufficiency of Theorem 1, we need the following condition for the line segment on the boundary of a numerical range. It is from [4, Lemma 1.4].

**Lemma 10.** Let \(A\) be an \(n\)-by-\(n\) (\(n \geq 2\)) matrix. Then \(\partial W(A)\) has a line segment on the line \(x \cos \theta + y \sin \theta = d\) if and only if \(d\) is the maximum eigenvalue of \(\text{Re } (e^{-i\theta} A)\), which has unit eigenvectors \(x_1\) and \(x_2\) such that \(\text{Im } \langle e^{-i\theta} A x_1, x_1 \rangle \neq \text{Im } \langle e^{-i\theta} A x_2, x_2 \rangle\).
Lemma 11. Let $A$ be an $n$-by-$n$ ($n \geq 2$) weighted shift matrix with nonzero real weights $a_1, \ldots, a_n$. Then $\partial W(A)$ has a line segment on the line $x = d$ if and only if $d$ is the maximum eigenvalue of $\text{Re} A$ with multiplicity at least two.

Proof. In view of Lemma 10, we need only prove the sufficiency part. Since $\dim \ker(dI_n - \text{Re} A) \geq 2$, there are real vectors $b = [0\ b_2\ \ldots\ b_n]^T$ and $c = [c_1\ 0\ c_3\ \ldots\ c_n]^T$ in $\ker(dI_n - \text{Re} A)$ with $b_2, c_1 \neq 0$. Then we obtain $a_1b_2 + a_n b_n = 0, 2db_2 = a_2b_3, 2db_j = a_{j-1}b_{j-1} + a_j b_{j+1}$ for $3 \leq j \leq n-1$, and $a_{n-1}b_{n-1} = 2db_n$ (resp., $2dc_1 = a_n c_n, a_1 c_1 + a_2 c_3 = 0, 2dc_3 = a_3 c_4, 2dc_j = a_{j-1} c_{j-1} + a_j c_{j+1}$ for $4 \leq j \leq n-1$, and $a_n c_1 + a_{n-1} c_{n-1} = 2dc_n$). Simple computations show that
\[
\begin{align*}
\frac{a_n b_n c_1}{\|b + ic\|^2} &= -a_1 b_2 c_1 + a_2 b_2 c_3 = (2db_3 - a_3 b_4)c_3 \\
&= a_3(b_3 c_4 - b_4 c_3) = a_3 b_3 c_4 - (2dc_4 - a_4 c_5)b_4 \\
&= a_4(b_4 c_5 - b_5 c_4) \\
&= \cdots \\
&= a_{n-1}(b_{n-1} c_n - b_n c_{n-1}).
\end{align*}
\]
Letting $x_1 = (b + ic)/\|b + ic\|$ and $x_2 = b/\|b\|$, we have
\[
\text{Im}(Ax_1, x_1) = \frac{1}{\|b + ic\|^2} \left(-a_1 b_2 c_1 + a_2 b_2 c_3 + \sum_{j=3}^{n-1} a_j (b_j c_{j+1} - b_{j+1} c_j) + a_n b_n c_1 \right)
\]
\[
= -\frac{n a_1 b_2 c_1}{\|b + ic\|^2} \neq 0 = \text{Im}(Ax_2, x_2).
\]
Our assertion follows from Lemma 10. \qed

We are now ready to prove the sufficiency part of Theorem 1.

Proof of Theorem 1. Assume that $a_j \neq 0$ and $W(A[j]) = B(0; r)$ for all $j, 1 \leq j \leq n$. By Lemma 2 (2), $A$ is unitarily equivalent to $e^{i\psi} C$, where $C$ is the $n$-by-$n$ weighted shift matrix with weights $|a_1|, \ldots, |a_{n-1}|, -|a_n|$ and $\psi = (\pi + \sum_{j=1}^n \arg a_j)/n$. By Proposition 8 (2), $r = w_0(C) = w_0(A)$ is the largest eigenvalue of $\text{Re} C$ with multiplicity at least two. Lemma 11 then implies that $\partial W(C)$ has a line segment on the line $x = r$. Thus $\partial W(A)$ has a line segment on $x\cos\psi + y\sin\psi = r = w_0(A)$. \qed

The next proposition characterizes those 4-by-4 weighted shift matrices $A$ with $\partial W(A)$ containing a line segment in terms of the weights of $A$. It was worked out by Gau and the second author some years ago.

Proposition 12. Let $A$ be a 4-by-4 weighted shift matrix with weights $a_1, \ldots, a_4$. Then the following conditions are equivalent:

1. $\partial W(A)$ has a line segment,
2. $|a_1| = |a_3| \neq 0$ and $|a_2| = |a_4| \neq 0$,
3. $A$ is unitarily equivalent to $\begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & c_1 \\ c_2 & 0 \end{bmatrix}$, where $b_1 b_2 = -c_1 c_2 \neq 0$ and $|b_1|^2 + |b_2|^2 = |c_1|^2 + |c_2|^2$, and
4. $A$ is unitarily equivalent to $\begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix} \oplus i \begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix}$ with $b_1, b_2 \neq 0$. 
In this case, \( W(A) \) is the convex hull of the two (orthogonal) ellipses \( E_1 \) and \( E_2 \) (may degenerate to line segments if \( |b_1| = |b_2| \)) with \( E_1 \) having foci \( \pm (b_1 b_2)^{1/2} \) and minor axis of length \(|b_1| - |b_2|\) and \( E_2 = iE_1 \). In particular, \( \partial W(A) \) has four line segments.

**Proof.** (1) \( \iff \) (2). Since the characteristic polynomial of Re \( A[1] \) is

\[
\det(zI_3 - \text{Re} A[1]) = z^3 - \frac{1}{4} \left( |a_2|^2 + |a_3|^2 \right) z,
\]

we have \( w(A[1]) = \|\text{Re} A[1]\| = (|a_2|^2 + |a_3|^2)^{1/2} / 2 \). Similarly, we obtain values of \( w(A[j]) \) for \( 2 \leq j \leq 4 \). Thus the equivalence of (1) and (2) follows from Theorem 1 and Proposition 3 (3).

(2) \( \Rightarrow \) (3). Since \( \det(zI_4 - A) = z^4 - a_1 a_2 a_3 a_4 \), the eigenvalues of \( A \) are \( \alpha_j = (a_1 a_2 a_3 a_4)^{1/4} \alpha_j^j, 0 \leq j < 4 \). Their respective eigenvectors can be computed to be (multiples of) \( x_j \equiv [1 \alpha_j / a_1 \alpha_j^2 / (a_1 a_2) \alpha_j^3 / (a_1 a_2 a_3)]^T, 0 \leq j < 4 \). Note that

\[
\langle x_j, x_k \rangle = 1 + \frac{1}{|a_1|^2} \alpha_j \alpha_k + \frac{1}{|a_1 a_2|^2} (\alpha_j \alpha_k)^2 + \frac{1}{|a_1 a_2 a_3|^2} (\alpha_j \alpha_k)^3
\]

for any \( j \) and \( k \). From this, it is easy to verify that

\[
\langle x_1, x_2 \rangle = \langle x_3, x_4 \rangle = \langle x_3, x_4 \rangle = 0.
\]

Let \( y_1 = x_1 - x_3 = [0 2a_0 / a_1 0 2a_0^2 / (a_1 a_2 a_3)]^T, y_2 = x_1 + x_3 = [2 0 2a_0^2 / (a_1 a_2) 0]^T, y_3 = x_2 - x_4 = [0 2ia_0 / a_1 0 -2ia_0^2 / (a_1 a_2 a_3)]^T \) and \( y_4 = x_2 + x_4 = [2 0 -2a_0^2 / (a_1 a_2) 0]^T \), and let \( M \) be the subspace of \( \mathbb{C}^4 \) spanned by \( y_1 \) and \( y_2 \). Since \( A y_1 = A x_1 - A x_3 = \alpha_1 x_1 - \alpha_3 x_3 \) and \( A y_2 = A x_1 + A x_3 = \alpha_1 x_1 + \alpha_3 x_3 \), and \( M \) is also spanned by \( x_1 \) and \( x_3 \), we have \( A M \subseteq M \). A simple computation shows that \( A^* y_1 = (a_0 a_0^0 / (a_1 a_2 a_3)) y_2 \) and \( A^* y_2 = (|a_1|^2 / a_0) y_1 \), where the assumptions that \( |a_1| = |a_3| \) and \( |a_2| = |a_4| \) are used. This shows that \( A^* M \subseteq M \). Thus \( M \) is a reducing subspace of \( A \). Moreover, it is easily seen that \( M^\perp \) is spanned by \( y_3 \) and \( y_4 \), and \( \langle y_1, y_2 \rangle = \langle y_3, y_4 \rangle = \langle A y_j, y_j \rangle = 0 \) for all \( j \).

Therefore, \( A \) is unitarily equivalent to a matrix of the form

\[
\begin{bmatrix}
0 & b_1 \\ b_2 & 0
\end{bmatrix} \oplus \begin{bmatrix}
0 & c_1 \\ c_2 & 0
\end{bmatrix} \equiv B \oplus C \text{ on } M \oplus M^\perp.
\]

Since \( x_1 \) and \( x_3 \) are in \( M, \alpha_1 \) and \( \alpha_3 \) are eigenvalues of \( B \). Hence

\[
-b_1 b_2 = \det B = \alpha_1 \alpha_3 = a_0^{1/2} a_0^{2} = -a_0^{1/2}.
\]

A similar argument with \( C \) yields \( -c_1 c_2 = a_0^{1/2} \). It follows that \( b_1 b_2 = -c_1 c_2 \).

To prove \( |b_1|^2 + |b_2|^2 = |c_1|^2 + |c_2|^2 \), note that simple computations give

\[
b_1 = \left\langle A \frac{y_2}{\|y_2\|}, \frac{y_1}{\|y_1\|} \right\rangle = a_0 \frac{\|y_1\|}{\|y_2\|},
\]

\[
b_2 = \left\langle A \frac{y_1}{\|y_1\|}, \frac{y_2}{\|y_2\|} \right\rangle = a_0 \frac{\|y_2\|}{\|y_1\|},
\]

\[
c_1 = \left\langle A \frac{y_4}{\|y_4\|}, \frac{y_3}{\|y_3\|} \right\rangle = i a_0 \frac{\|y_3\|}{\|y_4\|},
\]

and

\[
c_2 = \left\langle A \frac{y_3}{\|y_3\|}, \frac{y_4}{\|y_4\|} \right\rangle = i a_0 \frac{\|y_4\|}{\|y_3\|},
\]

and \( \|y_1\| = \|y_3\| \) and \( \|y_2\| = \|y_4\| \). Thus
\[ |b_1|^2 + |b_2|^2 = |\alpha_0|^2 \left( \frac{||y_1||^2}{||y_2||^2} + \frac{||y_2||^2}{||y_1||^2} \right) \]
\[ = |\alpha_0|^2 \left( \frac{||y_3||^2}{||y_4||^2} + \frac{||y_4||^2}{||y_3||^2} \right) = |c_1|^2 + |c_2|^2 \]
as asserted.

(3) \implies (4). Note that \[
\begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix} \quad \text{(resp., } \begin{pmatrix} 0 & c_1 \\ c_2 & 0 \end{pmatrix} \text{)}
\]
is unitarily equivalent to \[
\begin{pmatrix} (b_1b_2)^{1/2} & |b_1| - |b_2| \\ 0 & -(b_1b_2)^{1/2} \end{pmatrix} \quad \text{(resp., } \begin{pmatrix} (c_1c_2)^{1/2} & |c_1| - |c_2| \\ 0 & -(c_1c_2)^{1/2} \end{pmatrix} \).
\]
From the assumption in (3), we have \[(b_1b_2)^{1/2} = \pm i(c_1c_2)^{1/2} \text{ and } ||b_1| - |b_2|| = ||c_1| - |c_2||.\] Thus \[
\begin{pmatrix} 0 & c_1 \\ c_2 & 0 \end{pmatrix}
\]
is unitarily equivalent to \[
i \begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix}, \text{ and (4) follows.}
\]

(4) \implies (1). Since \(W \left( \begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix} \right)\) is the elliptic disc with foci \(\pm (b_1b_2)^{1/2}\) and minor axis of length \(||b_1| - |b_2||, that is, \(W \left( \begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix} \right) = E_1^\wedge \) and \(W \left( i \begin{pmatrix} 0 & b_1 \\ b_2 & 0 \end{pmatrix} \right) = (iE_1)^\wedge\), it is obvious that \(\partial W(A)\) contains four line segments. This also proves our assertion on \(W(A)\), completing the proof. \(\Box\)

For \(n > 4\), we can use the same arguments as in the proof of (1) \(\iff\) (2) above to obtain conditions in terms of the weights. They turn out to be too complicated to be useful.

In a forthcoming paper \([14]\) by the first author, more specific information is obtained for the numerical ranges of weighted shift matrices with periodic weights.

We conclude this paper by stating a theorem on the numerical ranges of matrices with an analogous structure, namely, the nilpotent matrices of the form

\[
A = \begin{pmatrix}
0 & a_1 & 0 & \cdots & 0 & a_n \\
0 & a_2 & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & 0 \\
0 & a_{n-1} \\
0 & 0 & \cdots & \cdots & \cdots & 0
\end{pmatrix}
\]

with weights \(a_1, \ldots, a_n\). Note that \(A\) and the weighted shift matrix with weights \(a_1, \ldots, a_n\) (\(a_n\) real) have the same real parts, which explains why (almost) all results in this paper for the latter have their analogs for the former. The only difference is that in the present case \(A\) is unitarily equivalent to \(\omega_{n-2}A\) and hence \(W(A)\) has the \((n-2)\)-symmetry property. The full details can be found in the first author’s Ph.D. dissertation \([15]\).

**Theorem 13.** Let \(A\) be an \(n\)-by-\(n\) \((n \geq 3)\) matrix of the form (ii). Then \(\partial W(A)\) has a line segment if and only if the \(a_j\)'s are nonzero and \(W(A[1]) = \cdots = W(A[n])\). In this case, \(W(A[j])\) is the circular disc centered at the origin with radius \(w_0(A)\), the line segment lies on one of the lines \(x \cos \theta_k + y \sin \theta_k = w_0(A)\), where
\[
\theta_k = \left( (2k + 1)\pi - \arg a_n + \sum_{j=1}^{n-1} \arg a_j \right) / (n - 2),
\]
0 \leq k < n - 2, and there are exactly \( n - 2 \) line segments on \( \partial W(A) \).

The following is an easy corollary.

**Corollary 14.** Let \( A \) (resp., \( B \)) be the \( n \)-by-\( n \) (\( n \geq 3 \)) weighted shift matrix (resp., nilpotent matrix of the form (ii)) with weights \( a_1, \ldots, a_n \). Then

1. \( w(A) = w(B) \),
2. \( w_0(A) = w_0(B) \), and
3. \( \partial W(A) \) has a line segment if and only if \( \partial W(B) \) has.

A study of the matrix of the form (ii) with \( a_1 = \cdots = a_n = 1 \) was made in [3, Proposition 3.2].

**Acknowledgments**

The authors thank the (anonymous) referees for their incisive comments. One of them pointed out [8], the Ph.D. dissertation of Issos, to us, which contains some relevant information.

**References**


