國立交通大學應用數學系

博 士 論 文

算子和矩陣的數值域研究

A Study of Numerical Ranges of Operators and Matrices

研究生：蔡明誠

指導教授：吳培元教授

中華民國一百年五月

算子和矩陣的數值域研究

A Study of Numerical Ranges of Operators and Matrices

研究生：蔡明誠　　　　Student：Ming-Cheng Tsai

指導教授：吳培元　　　　Advisor：Pei Yuan Wu

國 立 交 通 大 學

應 用 數 學 系

博 士 論 文

A Thesis

Submitted to Department of Applied Mathematics

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Applied Mathematics

May 2011

Hsinchu, Taiwan, Republic of China

中華民國一百年五月

算子和矩陣的數值域研究

學 生 : 蔡 明 誠　　　　　　　　指 導 教 授 : 吳 培 元 教 授

國 立 交 通 大 學 應 用 數 學 系

摘　　　　　要

在本論文中, 我們對於某些算子和矩陣的數值域做一個研究。

首先, 我們考慮 C_0 收縮算子和二次算子。我們證明如果 A 是一個 C_0 收縮算子有著最小函數 ϕ 使得 $w(A)=w(S(\phi))$ 而且 B 和A 可交換, 其中 $w(\cdot)$ 表示算子的數値半徑, 則 $w(AB)\leq w(A)\|B\|$。因此, 對於所有二次算子 A 和任意和 A 可以交換的算子 B, 我們也得到 $w(AB)\leq w(A)\|B\|$。

接著, 令 A 代表 n 維 $(n\geq 2)$ 的帶權移動矩陣 $[t_{ij}]_{i,j=1}^{n}$, 其中 $t_{i,i+1}=a_i$, $i=1,2,\ldots,n-1$, $t_{n,1}=a_n$ 而且其餘 $t_{i,j}=0$。我們證明它的數値域邊界具有一個線段的充分必要條件是這些 a_i 不爲零而且 A 的 $n-1$ 維主要子矩陣的數値域都相同。由此我們得到如果一個 n 維的帶權移動矩陣 A, 其中這些 a_i 是非零而且 $|a_i|$ 是週期的, 則它的數値域邊界具有一個線段。我們也證明了它的數値域邊界含有一個非圓的橢圓弧若且唯若這些 a_i 不爲零, n 是偶數, $|a_1|=|a_3|=\cdots=|a_{n-1}|$, $|a_2|=|a_4|=\cdots=|a_n|$ 而且 $|a_1|\neq|a_2|$。最後, 我們刻劃 A 是可約的情形而且完整描述它的數値域。

再來, 我們證明一個四維的實冪零矩陣 A 的數値域邊界最多具有兩個線段。我們也給了一個四維的冪零矩陣 A 的數値域邊界具有兩個平行線段的一個充分必要條件。

最後, 我們將證明一個有限維矩陣 $A=(\sum_{i=1}^{k_1}\oplus A_i)\oplus\operatorname{diag}(w_1,\ldots,w_{k_2})$, 其中

$$A_i=\begin{bmatrix} x_i & z_i \\ 0 & y_i \end{bmatrix}, i=1,\ldots,k_1,$$ 是兩個非負收縮矩陣的乘積若且唯若 $0\leq x_i,y_i,w_j\leq 1$ 而且, 對於所有 i,j, $|z_i|\leq|\sqrt{x_i}-\sqrt{y_i}|\sqrt{(1-x_i)(1-y_i)}$。藉此我們可以在 n 維的二次算子上得到一個類似的結果。

A Study of Numerical Ranges of Operators and Matrices

Student : Ming Cheng Tsai Advisor : Pei Yuan Wu

Department of Applied Mathematics
National Chiao Tung University

Abstract

In this thesis, we study properties of the numerical ranges of some operators and matrices.

First, we consider C_0 contractions and quadratic operators. We show that if A is a C_0 contraction with minimal function ϕ such that $w(A) = w(S(\phi))$ and if B commutes with A, where $w(\cdot)$ denotes the numerical radius of an operator, then $w(AB) \leq w(A)\|B\|$. As a consequence, we also obtain $w(AB) \leq w(A)\|B\|$ for any quadratic operator A and any B commuting with A.

Second, let A be the n-by-n ($n \geq 2$) weighted shift matrix $[t_{ij}]_{i,j=1}^{n}$, where $t_{i,i+1} = a_i$ for $i = 1, 2, \cdots, n-1$, $t_{n,1} = a_n$ and $t_{i,j} = 0$ otherwise. We show that the boundary of its numerical range contains a line segment if and only if the a_i's are nonzero and the numerical ranges of the $(n-1)$-by-$(n-1)$ principal submatrices of A are all equal. Using this, we obtain that the boundary of the numerical range of an n-by-n weighted shift matrix A has a line segment if the a_i's are nonzero and their moduli are periodic. We also prove that $\partial W(A)$ contains a noncircular elliptic arc if and only if the a_i's are nonzero, n is even, $|a_1| = |a_3| = \cdots = |a_{n-1}|$, $|a_2| = |a_4| = \cdots = |a_n|$ and $|a_1| \neq |a_2|$. Finally, we give a criterion for A to be reducible and completely characterize the numerical ranges of such matrices.

Next, we show that if A is a 4-by-4 nilpotent real matrix, then the boundary of its numerical range has at most two line segments. We also give a necessary and sufficient condition for the boundary of $W(A)$ to have a pair of parallel line segments.

Finally, we give a necessary and sufficient condition for a finite matrix $A = (\sum_{i=1}^{k_1} \oplus A_i) \oplus \operatorname{diag}(w_1, \ldots, w_{k_2})$, where $A_i = \begin{bmatrix} x_i & z_i \\ 0 & y_i \end{bmatrix}$ for all i, to be a product of two nonnegative contractions: $0 \leq x_i, y_i, w_j \leq 1$ and $|z_i| \leq |\sqrt{x_i} - \sqrt{y_i}|\sqrt{(1-x_i)(1-y_i)}$ for all i, j. Applying this, we obtain an analogous characterization for an n-by-n quadratic operator.

誌　謝

首先感謝我的指導教授 吳培元老師。由於他的耐心教導及在學習和研究上的協助，我才能完成這篇論文。另外老師對做研究積極認眞的精神和嚴謹治學的態度更是我學習追求的目標。這幾年來老師辛苦帶我做研究的過程，在此致上我最誠摯的敬意和謝枕。此外要感謝前年通往天國的碩士指導教授 蕭勝彥老師在前幾年對我持續的關懷勉勵。我會永遠記得老師對於做研究所給我的建議。其次對於 高華隆老師，我也非常感謝他在做研究及生活上所給予我的幫助，同時對於系上老師學業上的教導在此也一併致謝。

另外，感謝同研究室的昌源學長、同學金龍、偉國、光暉和建綸學弟在這幾年來課業及生活上所給予的幫助，還有室友其儒一起準備資格考的日子。當然也感謝 David 郭學長及同學忠澤、育慈、文貴、惠蘭、小巴、恭儉、信儀、偉強、偉碩、宛諭等學弟妹在學業上或生活上一起討論相處的日子，使我有個愉快且充實的交大生涯。

最後，我要將此論文獻給我親愛的父母和關心我、愛護我的姊姊以及姊夫。沒有他們在後面的支持和幫助就沒有這一切。

蔡明誠　2011 \ 5 \ 19 於交大

Contents

Chapter 1 Introduction

Let A be a bounded linear operator on a complex Hilbert space H. The *numerical range* $W(A)$ and *numerical radius* $w(A)$ of A are, by definition,

$$W(A) = \{\langle Ax, x\rangle : x \in H, \|x\| = 1\}$$

and

$$w(A) = \sup\{|z| : z \in W(A)\},$$

respectively, where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the inner product and its corresponding norm in H. In addition, we let $r(A) = \max\{|z| : z \in \sigma(A)\}$ denote the *spectral radius* of A. For an operator A, let A^T denote its transpose, A^* its adjoint, Re A its real part $(A + A^*)/2$ and Im A its imaginary part $(A - A^*)/2i$. For any subset $\triangle$ of $\mathbb{C}$, $\triangle^\wedge$ denotes its convex hull, that is, $\triangle^\wedge$ is the smallest convex set containing $\triangle$. We list several important properties of the numerical range.

(1) $W(U^*AU) = W(A)$ for any unitary operator U.

(2) $W(A)$ is a bounded subset of $\mathbb{C}$ and compact if H is finite dimensional.

(3) $W(aA + bI) = aW(A) + b$ for any scalars a and b.

(4) $W(\text{Re } A) = \text{Re } W(A)$ and $W(\text{Im } A) = \text{Im } W(A)$.

(5) If $A = \begin{bmatrix} B & * \\ * & * \end{bmatrix}$, then $W(B) \subseteq W(A)$.

(6) By the celebrated Hausdorff–Toeplitz theorem, $W(A)$ is always a convex subset of $\mathbb{C}$.

(7) $\sigma(A) \subseteq \overline{W(A)}$.

(8) If A is normal, then $\overline{W(A)}$ is equal to $\sigma(A)^\wedge$.

(9) $W(\sum_n \oplus A_n) = (\cup_n W(A_n))^\wedge$ and $w(\sum_n \oplus A_n) = \sup_n w(A_n)$.

(10) $r(A) \leq w(A) \leq \|A\| \leq 2w(A)$.

For other properties, the reader may consult [20, Chapter 22] or [19].

In Chapter 2, we consider C_0 contractions. Recall that an operator A (on a separable Hilbert space) is of *class* C_0 if it is a contraction ($\|A\| \leq 1$), it is completely nonunitary (i.e., it has no unitary direct summand) and it satisfies $\phi(A) = 0$ for some ϕ in the Hardy space H^∞ of bounded analytic functions on the open unit disc. The *minimal function* of a C_0 contraction A is the smallest function ϕ in H^∞ with $\phi(A) = 0$ (i.e., it divides all other annihilating functions of A). More important for us here is the *compression of the shift* $S(\phi)$ defined as follows. Let ϕ be an inner function ($\phi \in H^\infty$ with $|\phi| = 1$ a.e. on the unit circle) and let $S(\phi)$ be defined on $H = H^2 \ominus \phi H^2$ by

$$S(\phi)f = P_H(zf(z))|H \quad \text{for } f \in H.$$

In Theorem 2.1, we show that if A is a C_0 contraction with minimal function ϕ such that $w(A) = w(S(\phi))$ and if B commutes with A, then $w(AB) \leq w(A)\|B\|$. This is in contrast to the known fact that if $A = S(\phi)$ (even on a finite-dimensional space) and B commutes with A, then $w(AB) \leq \|A\|w(B)$ is not necessarily true. As a consequence, we have $w(AB) \leq w(A)\|B\|$ for any quadratic operator A and any B commuting with A. Here, an operator A is said to be *quadratic* if it is annihilated by a quadratic polynomial, that is, if it satisfies $A^2 + aA + bI = 0$ for some scalars a and b.

In Chapter 3, we consider an n-by-n ($n \geq 3$) matrix A of the form

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

In Section 3.2, we show that 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 $W(A)$ is the equilateral triangular region with vertices the three cubic roots of $a_1a_2a_3$. 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. In Section 3.4, we show that if the a_j's are nonzero and their moduli are periodic, then the boundary of its numerical range contains a line segment. We also prove that $\partial W(A)$ contains a noncircular elliptic arc if and only if the a_j's are nonzero, n is even, $|a_1| = |a_3| = \cdots = |a_{n-1}|$, $|a_2| = |a_4| = \cdots = |a_n|$ and $|a_1| \neq |a_2|$. Finally, we give a criterion for A to be reducible (i.e., it is unitarily equivalent to the direct sun of two other matrices) and completely characterize the numerical ranges of such matrices.

In Chapter 4, let A be a 4-by-4 nilpotent matrix ($A^k = 0$ for some $k \geq 1$). We show that if A is a real matrix, then the boundary of its numerical range has at most two line segments. We also prove that the boundary of $W(A)$ has one pair of parallel line segments if and only if A is unitarily equivalent to a matrix of the form

$$\alpha \begin{bmatrix} 0 & r_1 & ir_2 & -r_3 \\ & 0 & r_3 & -ir_2 \\ & & 0 & r_1 \\ & & & 0 \end{bmatrix},$$

where $\alpha \in \mathbb{C} \setminus \{0\}$, r_1, $r_3 > 0$ and $r_2 \in \mathbb{R}$. Moreover, in this case, $\partial W(A)$ has no other line segment. Finally, we give a special form for A to have a line segment on the boundary of its numerical range.

In Chapter 5, a bounded linear operator A on a complex Hilbert space H is *nonnegative* if $\langle Ax, x\rangle \geq 0$ for any x in H. Theorem 5.18 says that if a bounded linear operator A of the form $\begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix}$ is a product of two nonnegative contractions, then so are A_1 and A_2. In addition, in Corollary 5.22, a necessary and sufficient condition for a finite matrix $A = (\sum_{i=1}^{k_1} \oplus A_i) \oplus \mathrm{diag}(w_1, \ldots, w_{k_2})$, where $A_i = \begin{bmatrix} x_i & z_i \\ 0 & y_i \end{bmatrix}$ for all i, to be a product of two nonnegative contractions is

$$0 \le x_i, y_i, w_j \le 1 \text{ and } |z_i| \le |\sqrt{x_i} - \sqrt{y_i}|\sqrt{(1-x_i)(1-y_i)} \text{ for all } i, j.$$

Here, a diagonal matrix with diagonals $a_1, \ldots, a_n$ is denoted by $\mathrm{diag}(a_1, \ldots, a_n)$. It follows in Corollary 5.23 that an n-by-n quadratic operator is a product of two nonnegative contractions if and only if it is unitarily equivalent to a matrix of the form

$$aI_1 \oplus bI_2 \oplus (\sum_{i=1}^{k} \oplus \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix}),$$

where $0 \le a$, $b \le 1$ and $|c_i| \le |\sqrt{a} - \sqrt{b}|\sqrt{(1-a)(1-b)}$ for all i. Finally, we show that $W(A)$ is not a circular disc if A is a product of two n-by-n nonnegative contractions (Corollary 5.26).

Chapter 2 Numerical radius inequality for C_0 contractions

2.1 Introduction

Let A be a bounded linear operator on a complex Hilbert space H. The numerical radius inequalities we discuss here have their genesis from the power inequality. The latter asserts that $w(A^n) \leq w(A)^n$ for all $n \geq 1$ or, equivalently, that $w(A) \leq 1$ implies $w(A^n) \leq 1$ for all n. The first proof of it is given by Berger in his Ph.D. thesis [2] by way of his structure theorem for numerical contractions: $w(A) \leq 1$ if and only if there is a unitary operator U on a space K containing H such that $A^n = 2P_H U^n|H$ for all $n \geq 1$, where P_H denotes the (orthogonal) projection from K onto H (Lemma 2.6(b)). A totally elementary proof of the power inequality is later provided by Pearcy [38] (cf. also [20, Problem 221]).

In 1969, Holbrook [21] asked whether, for commuting operators A and B, the inequalities $w(AB) \leq w(A)\|B\|$ and $w(AB) \leq \|A\|w(B)$ hold. It is known that this is indeed the case when A and B doubly commute (i.e., $AB = BA$ and $AB^* = B^*A$) (cf. [21, Theorem 3.4]). Another known case is when A is an isometry (cf. [4, Lemma 2]). On the other hand, Crabb showed that for commuting A and B the inequality $w(AB) \leq (\sqrt{2+2\sqrt{3}}/2)w(A)\|B\|$ is true (cf. [37]). More recently, Holbrook [22] proved $w(AB) \leq w(A)w(B)$ for commuting 2-by-2 matrices A and B. So much for the partial positive confirmations. It came as a surprise when in 1988 Müller [34] gave an example of two 12-by-12 commuting matrices A and B with $w(AB) > \|A\|w(B)$. The example involves pure computations with no revealing reason why this should be the case. The day is saved by Davidson and Holbrook [9] that $w(AB) > \|A\|w(B)$ is already true for $A = J_9$ and $B = J_9^3 + J_9^7$. Here J_n, $n \geq 1$, denotes the n-by-n *Jordan* *block*

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

whose numerical range is known to be $\{z \in \mathbb{C} : |z| \leq \cos(\pi/(n+1))\}$. In [23], Holbrook and Schoch showed that $w(AB) > \|A\|w(B)$ can occur even for 3-by-3 commuting A and B.

Let ϕ be an inner function and let $S(\phi)$ be defined on $H = H^2 \ominus \phi H^2$ by

$$S(\phi)f = P_H(zf(z))|H \quad \text{for } f \in H.$$

Then $S(\phi)$ is a C_0 contraction with minimal function ϕ and satisfies

$$\text{rank}\,(I - S(\phi)^* S(\phi)) = 1.$$

Such operators were first studied by Sarason [41] and later developed by Sz.-Nagy and Foiaş in the 1960s and '70s; they form the building blocks for the "Jordan model" for general C_0 contractions (cf. [44] and [1]). Among other things, $S(\phi)$ is known to have the commutant lifting property: every operator B commuting with $S(\phi)$ is of the form $f(S(\phi))$ for some f in H^∞ with $\|f\|_\infty = \|B\|$ (Lemma 2.3). If $\phi(z) = z^n$, $n \geq 1$,

$$(\text{resp.}, \phi(z) = \frac{z-a}{1-\bar{a}z} \cdot \frac{z-b}{1-\bar{b}z}, \quad |a|, |b| < 1),$$

then $S(\phi)$ is unitarily equivalent to J_n

$$(\text{resp.}, \begin{bmatrix} a & (1-|a|^2)^{1/2}(1-|b|^2)^{1/2} \\ 0 & b \end{bmatrix}).$$

2.2 Numerical radius inequality for C_0 contractions

The main theorem of this section is

Theorem 2.1. *If A is a C_0 contraction with minimal function ϕ such that $w(A) = w(S(\phi))$ and if B commutes with A, then $w(AB) \leq w(A)\|B\|$.*

We start the proof with the following lemma.

Lemma 2.2. *If $A = S(\phi)$ and B commutes with A, then $w(AB) \leq w(A)\|B\|$.*

For its proof, we need the following lemmas.

Lemma 2.3. (Sarason's Commutant Lifting Theorem). *If B is an operator on $H^2 \ominus \phi H^2$ that commutes with $S(\phi)$, then there is a function f in H^∞ such that $\|f\|_\infty = \|B\|$ and $B = f(S(\phi))$.*

Let $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

Lemma 2.4. *Assume $w(A) \leq 1$. Let $f : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ be analytic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$. Then $w(f(A)) \leq 1 + 2|f(0)|$.*

The proof of Lemma 2.3 is in [41, Theorem 3] and Lemma 2.4 is obtained in Berger and Stampfli [3, Corollary 2] or Kato [27, Theorem 5].

Using the preceding lemmas, we now prove Lemma 2.2.

Proof of Lemma 2.2. We may assume that $A \neq 0$ and $\|B\| = 1$. By Lemma 2.3, $B = f(A)$ for some f in H^∞ with $\|f\|_\infty = \|B\| = 1$. Letting $A_1 = A/w(A)$ and $g(z) = zf(w(A)z)$, we have $A_1 B = g(A_1)$ with g in H^∞, $g(0) = 0$ and $\|g\|_\infty \leq \|f\|_\infty = 1$. By Lemma 2.4, we obtain

$$w(A_1 B) = w(g(A_1)) \leq \|g\|_\infty \leq 1.$$

Thus $w(AB) \leq w(A)$ follows as required. ■

As was remarked before, the Davidson–Holbrook example of $A = J_9$ and $B = J_9^3 + J_9^7$ shows the falsity of $w(AB) \leq \|A\| w(B)$ for $A = S(\phi)$ and B commuting with A.

Using the extension of a C_0 contraction to the direct sum of the compressions of the shift and a "completely bounded" version of the result of Kato or Berger and Stampfli, one can generalize Lemma 2.2 from $S(\phi)$ to the more general C_0 contractions.

We start with the following lemmas. For any operator X on H and any integer $d \geq 1$, let $X^{(d)}$ denote the direct sum of d copies of X on $H^{(d)}$, the direct sum of d copies of H.

Lemma 2.5. *Let A be a C_0 contraction with minimal function ϕ. Then*

(a) *A can be extended to an operator A_1 on a larger space which is unitarily equivalent to $S(\phi)^{(d_A)}$, where $d_A = \dim \overline{\operatorname{ran}(I - A^*A)^{1/2}} \leq \infty$, and*

(b) *every operator B commuting with A can be extended to an operator B_1 commuting with A_1 with $\|B_1\| = \|B\|$.*

The proof of this lemma is based on the Sz.-Nagy–Foiaş functional model for C_0 contractions (cf. [44, Section VI.3]).

Proof of Lemma 2.5. (a) We represent the C_0 contraction A^* on $H = H^2(K) \ominus \Theta H^2(K)$ by $A^* f = P_H(zf(z))$ for f in H, where K is a space of dimension d_A, $H^2(K)$ is the Hardy space of K-valued analytic square-integrable functions on the unit disc, and Θ is the characteristic function of A^*. Since the minimal function of A^* is $\widetilde{\phi}$ given by $\widetilde{\phi}(z) = \overline{\phi(\overline{z})}$ for $|z| < 1$, we have $\widetilde{\phi}(A^*) = 0$ and hence $\widetilde{\phi} H^2(K) \subseteq \Theta H^2(K)$. Let A_1^* be the operator defined on $H_0 = H^2(K) \ominus \widetilde{\phi} H^2(K)$ by $A_1^* g = P_{H_0}(zg(z))$ for g in H_0. Then A_1 is unitarily equivalent to $S(\phi)^{(d_A)}$. Since $H \subseteq H_0$ and A and A_1 are given by

$$Af = \frac{1}{z}(f(z) - f(0)) \quad \text{for} \quad f \in H$$

and

$$A_1 g = \frac{1}{z}(g(z) - g(0)) \quad \text{for} \quad g \in H_0,$$

respectively, we obtain $A = A_1|H$ as required.

(b) Since B^* commutes with A^*, it can be represented as $B^* f = P_H(\Phi f)$ for f in H, where Φ is a K-valued bounded analytic function on the unit disc with $\Phi\Theta H^2(K) \subseteq \Theta H^2(K)$ and $\|\Phi\|_\infty = \|B^*\|$. Let B_1^* be defined on H_0 by $B_1^* g = P_{H_0}(\Phi g)$ for g in H_0. Then B_1 commutes with A_1 and

$$\|B_1\| = \|B_1^*\| \leq \|\Phi\|_\infty = \|B^*\| = \|B\|.$$

On the other hand, if C denotes the adjoint of the operator $f \mapsto \Phi f$ on $H^2(K)$, then $B = C|H$ and $B_1 = C|H_0$. It follows that $B = B_1|H$ and hence $\|B\| \leq \|B_1\|$. Therefore, $\|B\| = \|B_1\|$ as required. ■

Part (a) of the preceding lemma is due to Nakazi and Takahashi [35, Lemma 4].

Lemma 2.6. *Let A be an operator on H. Then*

(a) (Sz-Nagy's Power Dilation Theorem). *A is a contraction if and only if there is a unitary operator U on a space K containing H such that $A^n = P_H U^n|H$ for all $n \geq 1$.*

(b) (Berger's Dilation Theorem). *$w(A) \leq 1$ if and only if there is a unitary operator U on a space K containing H such that $A^n = 2P_H U^n|H$ for all $n \geq 1$. Here P_H denotes the (orthogonal) projection from K onto H.*

The proof of the former can be deduced from [43, Theorem] and the latter is proved in [2].

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. As before, we may assume that $A \neq 0$ and $\|B\| = 1$. By Lemma 2.5, A extends to (an operator unitarily equivalent to) $A_1 = S(\phi)^{(d)}$ on L, where $d = d_A$, and B extends to B_1 on L with $B_1 A_1 = A_1 B_1$ and $\|B_1\| = \|B\|$. A matrix version of Lemma 2.2 [41, Theorem 3] implies that B_1 can be represented as $[f_{ij}(S(\phi))]_{i,j=1}^d$, where the f_{ij}'s are in H^∞ for all i and j and $\|[f_{ij}]\|_\infty = \|B_1\|$. Let $g_{ij}(z) = f_{ij}(w(A)z)$ for $|z| < 1$. We have

$$\begin{aligned} \|[g_{ij}]\|_\infty &= \sup\{\|[g_{ij}(z)]\| : |z| < 1\} \\ &\leq \sup\{\|[f_{ij}(w)]\| : |w| < 1\} \\ &= \|[f_{ij}]\|_\infty = \|B_1\| = \|B\| = 1. \end{aligned}$$

If $C = S(\phi)/w(S(\phi))$, then $w(C) = 1$ and $B_1 = [g_{ij}(C)]$. By Lemma 2.6 (b), there is a unitary operator U on a space K containing $H \equiv H^2 \ominus \phi H^2$ such that $C^n = 2P_H U^n|H$

for all $n \geq 1$. Since $Cg_{ij}(C) = 2P_H(Ug_{ij}(U))|H$ for all i and j, we obtain

$$(A_1B_1/w(A))^n = [Cg_{ij}(C)]^n = 2P_L[Ug_{ij}(U)]^n|L$$

for all $n \geq 1$. On the other hand, because

$$\|[Ug_{ij}(U)]\| \leq \|[zg_{ij}(z)]\|_\infty \leq \|[g_{ij}]\|_\infty \leq 1,$$

Lemma 2.6 (a) yields a unitary operator W on a space containing K such that

$$[Ug_{ij}(U)]^n = P_{K^{(d)}}W^n|K^{(d)}$$

for all n. Combining these two dilations, we obtain $(A_1B_1/w(A))^n = 2P_LW^n|L$ for all n. This implies, by Lemma 2.6 (b) again, that $w(A_1B_1/w(A)) \leq 1$ or $w(AB) \leq w(A_1B_1) \leq w(A)$ as required. ■

Note that, for any C_0 contraction A with minimal function ϕ, Lemma 2.5 (a) implies that $w(A) \leq w(S(\phi))$. In Theorem 2.1, the extra condition on their equality is essential for otherwise the example of

$$B = \frac{J_9 + \frac{1}{4}J_9^5}{\|J_9 + \frac{1}{4}J_9^5\|}$$

and $A = B^3$ attests to the falsity of the assertion there (cf. [8]). When A acts on a finite-dimensional space, the next proposition gives some equivalent conditions for the equality $w(A) = w(S(\phi))$.

Proposition 2.7. *For a C_0 contraction A with minimal function ϕ on a finite-dimensional space, the following are equivalent*:

(a) $w(A) = w(S(\phi))$;

(b) $\partial W(A) \cap \partial W(S(\phi)) \neq \emptyset$;

(c) $W(A) = W(S(\phi))$.

Proof. (a)$\Rightarrow$(b). We infer from Lemma 2.5 (a) that $W(A) \subseteq W(S(\phi))$. If $\partial W(A) \cap \partial W(S(\phi)) = \emptyset$, then obviously $w(A) < w(S(\phi))$. This proves (a)$\Rightarrow$(b).

(b)$\Rightarrow$(c). By [17, Lemma 3.3], (b) implies that $S(\phi)$ is a direct summand of A. Thus $W(S(\phi)) \subseteq W(A)$. Together with $W(A) \subseteq W(S(\phi))$ from Lemma 2.5 (a), this yields (c).

(c)$\Rightarrow$(a). This is trivial. ∎

We conclude this section by asking the following remaining question concerning this topic.

Is it true that if $A = S(\phi)$, then $w(A^{n+1}) \leq w(A^n)$ for all $n \geq 1$? More generally, if A is a C_0 contraction with minimal function ϕ such that $w(A) = w(S(\phi))$, then is $w(A^{n+1}) \leq w(A^n)$ true for all $n \geq 1$?

Recall that a general inner function ϕ has a canonical factorization as $c\phi_1\phi_2$, where c is a complex number with $|c| = 1$, ϕ_1 is a Blaschke product

$$\phi_1(z) = \prod_{n=1}^{\infty} \frac{\bar{a_n}}{|a_n|} \frac{z - a_n}{1 - \bar{a_n} z}$$

with zeros a_n in $\mathbb{D}$ satisfying $\sum_n (1 - |a_n|) < \infty$ and ϕ_2 is a singular function

$$\phi_2(z) = \exp(-\int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(\theta)),$$

where μ is a positive measure on $\partial\mathbb{D}$ which is singular with respective to the Lebesgue measure on $\partial\mathbb{D}$. From [1, Theorem 2.4.11], we know that $\sigma(S(\phi))$, the spectrum of

$S(\phi)$, consists of those points $\lambda \in \mathbb{D}$ such that $\phi(\lambda) = 0$ and those λ in $\partial\mathbb{D}$ such that ϕ cannot be continued analytically across λ.

Lemma 2.8.

(a) *If H is a finite-dimensional space, then* $w(S(\phi)^n) < 1$ *for all* n.

(b) *If H is an infinite-dimensional space, then* $w(S(\phi)^n) = 1$ *for all* n.

Proof. (a) This follows from the fact that $\sigma(T) \subseteq \mathbb{D}$ if and only if $W(T) \subseteq \mathbb{D}$ for any finite-dimensional contraction T (cf. [20, Solution 212]).

(b) Since H is finite dimensional if and only if $S(\phi) = c\phi_1$ where c is a complex number with $|c| = 1$ and ϕ_1 is a finite Blaschke product [10, Lemma 7.35]. Hence, $\sum_n (1 - |a_n|) < \infty$ implies $\lim_{n\to\infty} |a_n| = 1$. There exists γ in $\partial\mathbb{D}$ such that $\gamma \in \sigma(S(\phi))$. Thus, $r(S(\phi)^n) = r(S(\phi))^n = 1$ by spectral mapping theorem. Hence, $w(S(\phi)^n) = 1$. ■

If A is a Jordan block or a quadratic contraction (to be considered in Section 2.3), then we do have $w(A^{n+1}) \leq w(A^n)$ for all n. The latter is a consequence of Proposition 2.11 in Section 2.3. For the negative side, although $w(A^2) \leq w(A)$ for any contraction A by the power inequality, $w(A^4) \leq w(A^3)$ is in general false by the example in [8].

2.3 Numerical radius inequality for quadratic operators

An operator A is said to be *quadratic* if it is annihilated by a quadratic polynomial, that is, if it satisfies $A^2 + aA + bI = 0$ for some scalars a and b. The structure of quadratic operators is well-understood (cf. [46]).

Lemma 2.9. *Let A be a quadratic operator on H. Then*

(a) *the spectrum of A consists of the two zeros α and β of the polynomial z^2+az+b,*

(b) *A is unitarily equivalent to an operator of the form*

$$\begin{bmatrix} \alpha I_1 & D \\ 0 & \beta I_2 \end{bmatrix},$$

(c)

$$\|A\| = \| \begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix} \|,$$

and

(d) *the numerical range of A is the (open or closed) elliptic disc with foci at α and β, major axis of length $(|\alpha-\beta|^2 + \|D\|^2)^{1/2}$ and minor axis of length $\|D\|$.*

The proof of this Lemma can be found in [46, Theorems 1.1 and 2.1] and in the proof of [46, Lemma 2.2].

Using Theorem 2.1, we can now prove a numerical radius inequality for quadratic operators.

Theorem 2.10. *If A is a quadratic operator and B commutes with A, then $w(AB) \leq w(A)\|B\|$.*

Proof. If A is normal, then A is unitarily equivalent to $\alpha I_1 \oplus \beta I_2$, in which case the asserted inequality can be easily verified. Hence we may assume that A is nonnormal and has norm one. Then A is unitarily equivalent to an operator of the form $\begin{bmatrix} \alpha I_1 & D \\ 0 & \beta I_2 \end{bmatrix}$ with $D \neq 0$ by Lemma 2.9 (b). Since $1 = \|A\| \geq (|\alpha|^2 + \|D\|^2)^{1/2}$ and $D \neq 0$, we have $|\alpha| < 1$ and, similarly, $|\beta| < 1$. Hence A is a C_0 contraction with $\phi(A) = 0$, where ϕ is the inner function

$$\phi(z) = \frac{z-\alpha}{1-\overline{\alpha}z} \cdot \frac{z-\beta}{1-\overline{\beta}z}.$$

On the other hand, since

$$1 = \|A\| = \| \begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix} \|$$

by Lemma 2.9 (c), we obtain that $S(\phi)$ is unitarily equivalent to $\begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix}$. Therefore, $w(A) = w(S(\phi))$ by Lemma 2.9 (d). The asserted inequality then follows from Theorem 2.1. ■

Before we move on, two remarks are in order. (1) The inequality in the preceding theorem is not necessarily true if A is annihilated by a cubic polynomial. In [9, Corollary 4], it was shown that if

$$A = \begin{bmatrix} 0 & I_3 & J_3 \\ & 0 & I_3 \\ & & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} J_3 & & \\ & J_3 & \\ & & J_3 \end{bmatrix},$$

then $A^3 = B^3 = 0$, $AB = BA$, $w(A) = \cos(\pi/10)$, $\|B\| = 1$ and $w(AB) = 1$ and thus $w(AB) > w(A)\|B\|$. (2) The stronger inequality $w(AB) \leq w(A)w(B)$ is not necessarily true for a quadratic A and commuting B. An example was given in [21, p. 168]. Here is another one: $A = J_2 \oplus J_2$ and $B = J_4^2$. In this case, $A^2 = 0$, $AB = BA$ and $w(A) = w(B) = w(AB) = 1/2$.

In contrast to (2) above, if B is a polynomial of the quadratic A, then we do have $w(AB) \leq w(A)w(B)$.

Proposition 2.11. *If A is a quadratic operator and B is a polynomial of A, then* $w(AB) \leq w(A)w(B)$.

Proof. We may assume that $A = \begin{bmatrix} \alpha I_1 & D \\ 0 & \beta I_2 \end{bmatrix}$ and $B = A + \lambda I$. By Lemma 2.9 (d), we have $w(A) = w(\begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix})$. Similarly,

$$w(B) = w(\begin{bmatrix} (\alpha+\lambda) I_1 & D \\ 0 & (\beta+\lambda) I_2 \end{bmatrix}) = w(\begin{bmatrix} \alpha+\lambda & \|D\| \\ 0 & \beta+\lambda \end{bmatrix})$$

and

$$\begin{aligned} w(AB) &= w(\begin{bmatrix} \alpha(\alpha+\lambda) I_1 & (\alpha+\beta+\lambda) D \\ 0 & \beta(\beta+\lambda) I_2 \end{bmatrix}) \\ &= w(\begin{bmatrix} \alpha(\alpha+\lambda) & |\alpha+\beta+\lambda|\|D\| \\ 0 & \beta(\beta+\lambda) \end{bmatrix}) \\ &= w(\begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix} \begin{bmatrix} \alpha+\lambda & \|D\| \\ 0 & \beta+\lambda \end{bmatrix}). \end{aligned}$$

The latter is less than or equal to the product of $w(\begin{bmatrix} \alpha & \|D\| \\ 0 & \beta \end{bmatrix}) = w(A)$ and $w(\begin{bmatrix} \alpha+\lambda & \|D\| \\ 0 & \beta+\lambda \end{bmatrix}) = w(B)$ by [22]. This completes the proof. ■

Although Proposition 2.11 is proved via [22], it also generalizes the latter. Indeed, if A and B are commuting 2-by-2 matrices, then, assuming that $A = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}$ with $\gamma \neq 0$, B must be a (linear) polynomial of A and thus $w(AB) \leq w(A)w(B)$ by Proposition 2.11.

If A is square-zero ($A^2 = 0$) or idempotent ($A^2 = A$), the inequality $w(AB) \leq \min\{w(A)\|B\|, \|A\|w(B)\}$ (for commuting B) was proved in [13] using a completely different approach. The same can be said for $w(AB) \leq \|A\|w(B)$ when A satisfies $A^2 = aI$ for some scalar a and B commutes with A (cf. [39]).

We conclude this section by stating the following remaining question concerning this topic:

Is it true that $w(AB) \leq \|A\|w(B)$ for A quadratic and B commuting with A?

Note that this is false if A is annihilated by a cubic polynomial: the example of [9, Corollary 4] with

$$A = \begin{bmatrix} J_3 & & \\ & J_3 & \\ & & J_3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & I_3 & J_3 \\ & 0 & I_3 \\ & & 0 \end{bmatrix}$$

attests to this.

Chapter 3 Numerical ranges of weighted shift matrices

3.1 Introduction

An n-by-n ($n \geq 2$) *weighted shift matrix* A is one of the form

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

where the a_j's, called the *weights* of A, are complex numbers. The purpose of this chapter 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. It is known that $W(A)$ is a nonempty compact convex subset of $\mathbb{C}$. For any subset $\triangle$ of $\mathbb{C}$, $\triangle^\wedge$ denotes its convex hull, that is, $\triangle^\wedge$ is the smallest convex set containing $\triangle$. $W(A)$ contains the convex hull of the spectrum of A and, when A is normal, they are equal ([19, Theorem 1.4-4]). For other properties, we may consult [25, Chapter 1] or [19].

In Section 3.2, we know that 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 of Section 2 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 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 $\theta_k = (\sum_{j=1}^{n}\arg a_j + (2k+1)\pi)/n$, $0 \le k < n$, and there are exactly n line segments on $\partial W(A)$ (Theorem 3.1). 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, [42, Lemma 2] gives a method to compute the numerical radius of a weighted shift matrix with at least one zero weight. [40, 42, 5] discuss properties of the numerical ranges and numerical radii of weighted shifts on l^2 with periodic or geometric weights.

In Section 3.3, we state Theorem 3.15 on the numerical ranges of matrices which has an analogous structure as the one in Theorem 3.1, namely, the nilpotent matrices of the form

(i) $$A = \begin{bmatrix} 0 & a_1 & 0 & \cdots & 0 & a_n \\ & 0 & a_2 & \ddots & & 0 \\ & & \ddots & \ddots & \ddots & \vdots \\ & & & \ddots & \ddots & 0 \\ & & & & 0 & a_{n-1} \\ & & & & & 0 \end{bmatrix}$$

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 Section 2 for the latter have their analogues for the former. The main 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. A study of the matrix of the form (i) with $a_1 = \cdots = a_n = 1$ was made in [18, Proposition 3.2].

In Section 3.4, we study the numerical ranges of the n-by-n weighted shift matrices with periodic weights. In our discussions, we may assume that the weights are all nonnegative. Then in Theorem 3.21 below, we show that if an n-by-n weighted shift matrix A has periodic nonzero weights, then all its $(n-1)$-by-$(n-1)$ principal submatrices have identical numerical ranges. Using Theorem 3.1, we obtain that the boundary of its numerical range has a line segment. In Theorem 3.27, we give a necessary and sufficient condition for the boundary of $W(A)$ to have a noncircular elliptic arc. More specifically, it is shown that this is the case if and only if the a_j's are nonzero, n is even, $|a_1| = |a_3| = \cdots = |a_{n-1}|$, $|a_2| = |a_4| = \cdots = |a_n|$ and $|a_1| \neq |a_2|$. For $n = 4$, this essentially generalizes Proposition 3.13. Finally, we give a criterion for A to be reducible and characterize their numerical ranges in Theorem 3.28. In particular, it says that, for $n = 4$, A is reducible if and only if either (1) $a_i = a_j = 0$ for some i and j, $1 \leq i < j \leq n$, or (2) $|a_1| = |a_3| \neq 0$ and $|a_2| = |a_4| \neq 0$.

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 $i_1, \ldots, i_m$. Recall that 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. Let $l(A)$ denote the number of line segments on $\partial W(A)$. A diagonal matrix with diagonals $a_1, \ldots, a_n$ is denoted by $\mathrm{diag}(a_1, \ldots, a_n)$. Our basic reference for properties of matrices is [24].

For an n-by-n matrix A, consider the degree-n homogeneous polynomial $p_A(x, y, z) = \det(x\text{Re } A + y\text{Im } A + zI_n)$. The *Kippenhahn curve* $C(A)$ of A is the algebraic curve dual to the one determined by $p_A(x, y, z) = 0$ in the complex projective plane $\mathbb{CP}^2$, that is, $C(A)$ consists of all points $[u, v, w]$ in $\mathbb{CP}^2$ such that $ux+vy+wz = 0$ is a tangent line to $p_A(x, y, z) = 0$. As usual, we identify the point (x, y) in $\mathbb{C}^2$ with $[x, y, 1]$ in $\mathbb{CP}^2$ and identify any point $[x, y, z]$ in $\mathbb{CP}^2$ such that $z \neq 0$ with $(x/z, y/z)$ in $\mathbb{C}^2$. The real part of the curve $C(A)$, namely, the set $\{a + bi \in \mathbb{C} : a, b \in \mathbb{R}$ and $ax + by + z = 0$ is tangent to $p_A(x, y, z) = 0\}$, will be denoted by $C_{\mathbb{R}}(A)$. A result of Kippenhahn [28, p. 199] says that the numerical range $W(A)$ is the convex hull of the real points of the curve $p_A(x, y, z) = 0$, that is, $W(A) = C_{\mathbb{R}}(A)^\wedge$. The point $[x_0, y_0, z_0]$ is said to be a *focus* of the curve C if it is not equal to $[1, \pm i, 0]$ and the lines through $[x_0, y_0, z_0]$ and $[1, \pm i, 0]$ are tangent to C at points other than $[1, \pm i, 0]$.

For any nonzero complex number $z = x + iy$ (x and y real), arg z is the angle θ, $0 \leq \theta < 2\pi$, 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 $B(0; r) = \{z \in \mathbb{C} : |z| \leq r\}$ for $r > 0$ and $\omega_n = e^{2\pi i/n}$ for $n \geq 1$.

3.2 Numerical ranges of weighted shift matrices

The main result of this section is the following.

Theorem 3.1. *Let A be an n-by-n ($n \geq 2$) weighted shift matrix with weights $a_1, \ldots, a_n$. 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 $\theta_k = (\sum_{j=1}^{n} \arg\, a_j + (2k+1)\pi)/n$, $0 \leq k < n$, and there are exactly n line segments on $\partial W(A)$.*

For the proof of Theorem 3.1, we need the following lemmas. We start with the necessity of the proof with a lemma from [15, Lemma 5]. It relates the line segments on $\partial W(A)$ to the numerical ranges of submatrices of A.

Lemma 3.2. *If A is an n-by-n matrix and B is an $(n-1)$-by-$(n-1)$ submatrix of A, then every line segment on $\partial W(A)$ intersects $\partial W(B)$.*

Lemma 3.3. *Let A and B be n-by-n ($n \geq 3$) 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 $e^{i\alpha_k}B$, where $\alpha_k = (2k\pi + \sum_{j=1}^{n}(\arg\, a_j - \arg\, b_j))/n$ for $0 \leq k < n$. In particular, A is unitarily equivalent to $w_n A$ and hence $W(A)$ has n symmetry.*

(3) (a) *Either the intersection number of $\partial W(A)$ and $\partial B(0; w(A))$ is n or $W(A) =$*

$B(0; w(A))$.

(b) *Either the intersection number of* $\partial W(A)$ *and* $\partial B(0; w_0(A))$ *is* n *or* $W(A) = B(0; w_0(A))$.

(4) *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* θ, *and*

(c) $W(A)$ *is a circular disc centered at the origin.*

(5) *If* $\partial W(A)$ *has a line segment* L, *then* $\text{dist}(0, L) = w_0(A)$ *and there are exactly* n *line segments on* $\partial W(A)$.

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 $\phi_j = \phi_{j-1} + (\arg b_{j-1} - \arg a_{j-1}) + \alpha_k$ for $2 \leq j \leq n$, then U is unitary and $AU = U(e^{i\alpha_k}B)$. In particular, A is unitarily equivalent to w_nA by letting $B = A$ and $k = 1$.

(3) (a) follows from J. Anderson's result (cf. [45, Lemma 6] or [33, Theorem 4.12]) and (2). (b) follows from [18, Theorem 2.5 (a), (b)] and (2).

(4) If (a) holds, then the α_k's in (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. [33, Theorem 4.2]). Hence $\det A = (-1)^{n+1}a_1 \cdots a_n = 0$, which shows that $a_j = 0$ for some j.

(5) If L is a line segment on $\partial W(A)$, then, by Lemma 3.2, L intersects $\partial W(A[j])$ for every j, $1 \leq j \leq n$. Since $W(A[j]) \subseteq W(A)$ and $W(A[j])$ is a circular disc centered at the origin, we obtain $\operatorname{dist}(0, L) = w_0(A[j]) \leq w_0(A)$ for every j. But $\operatorname{dist}(0, L) \geq w_0(A)$ is obviously true. This shows that $\operatorname{dist}(0, L) = w_0(A[1]) = \cdots = w_0(A[n]) = w_0(A)$. Additionally, there are exactly n line segments on $\partial W(A)$ by (2) and (3b). ■

Therefore, the necessity of Theorem 3.1 follows easily from Lemmas 3.2 and 3.3 (4), (5). We now proceed to prepare ourselves for the proof of the sufficiency of Theorem 3.1. This will be done in a series of lemmas and propositions. Note that some of them have already been obtained in [26], the Ph.D. dissertation of Issos on irreducible nonnegative matrices. For example, (2) follows from [26, Theorem 7]. We start with the following.

Lemma 3.4. *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) *$W(A)$ is symmetric with respect to the lines $y = x\tan((k\pi/n)+\theta)$ for $0 \leq k < n$.*

(2) *We have $\{\arg\lambda : \lambda \in \partial W(A), |\lambda| = w(A)\} = \{(2k\pi/n) + \theta : 0 \leq k < n\}$ and $\{\arg\lambda : \lambda \in \partial W(A), |\lambda| = w_0(A)\} = \{((2k+1)\pi/n) + \theta : 0 \leq k < n\}$.*

(3) *$w(A) \leq w_0(A)\sec(\pi/n)$ and*

$$B(0; w_0(A)) \subseteq W(A) \subseteq w_0(A)(\sec\frac{\pi}{n})e^{i\theta}\{1, \omega_n, \ldots, \omega_n^{n-1}\}^\wedge.$$

Proof. (1) We need only consider $a_j \neq 0$ for all j by Lemma 3.3 (4). Lemma 3.3 (2) implies that A is unitarily equivalent to $e^{i((2k\pi/n)+\theta)}B$ and $e^{i(((2k+1)\pi/n)+\theta)}C$, where B and C are the n-by-n weighted shift matrices with weights $|a_1|, \ldots, |a_n|$ and $|a_1|, \ldots, |a_{n-1}|, -|a_n|$, respectively, where $0 \leq k < n$. Hence our assertion follows from the fact that the numerical range of a finite square matrix with real entries is symmetric with respect to the x-axis.

(2) Assume that $a_j \neq 0$ for all j. Then A is unitarily equivalent to $e^{i\theta}B$, where B is the n-by-n matrix with weights $|a_1|, \ldots, |a_n|$ by Lemma 3.3 (2). By the Perron-Frobenius Theory [31, Theorem 15.5.1], we know that $w(A) = w(B) \in W(B)$. Therefore our assertion follows from Lemma 3.3 (3a), (3b) and (1).

(3) Since the points $w(A)e^{i((2k\pi/n)+\theta)}$, $0 \leq k < n$, are in $W(A)$ by (2), the regular n-polygonal region R whose vertices are these points is contained in $W(A)$. Hence

$$w_0(A) \geq \text{dist}(0, R) = w(A)\frac{1}{2}|e^{i\theta} + e^{i((2\pi/n)+\theta)}| = w(A)\cos\frac{\pi}{n}.$$

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

By (2), we have the containment $B(0; w_0(A)) \subseteq W(A)$. 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 $w_0(A)\sec(\pi/n)e^{i\theta}$ and $w_0(A)\sec(\pi/n)e^{i((2\pi/n)+\theta)}$, from the origin, then, by (1), its symmetric point u' with respect to the line connecting 0 and $w_0(A)e^{i((\pi/n)+\theta)}$ is also in $W(A)$. Thus $(u+u')/2$ is in $W(A)$, which would contradict the fact that $w_0(A)e^{i((\pi/n)+\theta)}$ 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

$$\begin{aligned} W(A) &\subseteq w_0(A)(\sec\frac{\pi}{n})\{e^{i((2k\pi/n)+\theta)} : 0 \leq k < n\}^{\wedge} \\ &= w_0(A)(\sec\frac{\pi}{n})e^{i\theta}\{\omega_n^k : 0 \leq k < n\}^{\wedge}. \end{aligned}$$

■

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 [26, Theorem 13].

Proposition 3.5. *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* $\mathrm{diag}(\lambda, \lambda\omega_n, \ldots, \lambda\omega_n^{n-1})$, *where* $\lambda = (a_1 \cdots a_n)^{1/n}$,

(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* $|a_1 \cdots a_n|^{1/n}$,

(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 $|a_1|e^{i(\sum_{j=1}^n \arg\, a_j)/n}B$, where B is the n-by-n weighted shift matrix with weights $1, \ldots, 1$ by Lemma 3.3 (2). It is easily seen that B is unitarily equivalent to $\mathrm{diag}(1, \omega_n, \ldots, \omega_n^{n-1})$ and $|a_1|e^{i(\sum_{j=1}^n \arg\, a_j)/n} = (a_1 \cdots a_n)^{1/n} = \lambda$. Hence (3) follows. For (5) $\Rightarrow$ (1), if λ is a nondifferentiable point of $\partial W(A)$, then so are $\lambda\omega_n^k$, $0 \leq k < n$, by Lemma 3.3 (2). Since each of such points is a reducing eigenvalue of A, we obtain that A is unitarily equivalent to $\mathrm{diag}(\lambda, \lambda\omega_n, \ldots, \lambda\omega_n^{n-1})$. In particular, A is normal, that is, (1) holds. Finally, if (4) holds, then (2) is true and hence

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

that is, (6) holds. Conversely, if (6) is true, then Lemma 3.4 (3) says that $W(A) \subseteq w(A)e^{i\theta}\{1, \omega_n, \ldots, \omega_n^{n-1}\}^\wedge$, where $\theta = (\sum_{j=1}^{n} \arg a_j)/n$. 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 Lemma 3.4 (2). Hence we must have $W(A) = w(A)e^{i\theta}\{1, \omega_n, \ldots, \omega_n^{n-1}\}^\wedge$. Hence $\partial W(A)$ has nondifferentiable points, that is, (5) holds. This completes the proof. ■

For the sufficiency of Theorem 3.1, we also need the following lemma.

Lemma 3.6. *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 $W(A) \subsetneqq W(B)$.*

(3) *If the a_j's are nonzero, then $W(A[n]) \subsetneqq W(A)$.*

Proof. In view of Lemma 3.3 (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 [32, Corollary 3.6]. To prove (3), let $C = A[n] \oplus [0]$. Then $W(A[n]) = W(C) \subsetneqq W(A)$ by (2). This completes the proof. ■

The next lemma is needed for the proof of Proposition 3.8.

Lemma 3.7. *If A and B are n-by-n ($n \geq 2$) weighted shift matrices with weights $a_1, \ldots, a_{n-1}, 0$ and $a_{n-1}, \ldots, a_1, 0$, respectively, then $W(A) = W(B)$.*

Proof. Since $W(A)$ and $W(B)$ are circular discs centered at the origin by Lemma 3.3 (4), we need only check that $w(A) = w(B)$. By Lemma 3.3 (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

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

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

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:
J(Re A)=(Re B)J, where $J = [J_{ij}]_{i,j=1}^n$ is the n-by-n skew identity matrix with

$$J_{ij} = \begin{cases} 1 & \text{if } i + j = n + 1, \\ 0 & \text{otherwise,} \end{cases}$$

and hence $w(A) = \|\text{Re } A\| = \|\text{Re } B\| = w(B)$.

Proposition 3.8. *Let A be an n-by-n ($n \geq 3$) 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 3.3 (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 Lemma 3.3 (4). 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 3.7. Since $a_1 = \cdots = a_{n-3}$, by Lemma 3.3 (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 3.1. We conclude that $W(A_1) = W(A_2)$. Therefore, $a_{n-2} = a_n$ by Lemma 3.6 (2). This completes the proof. ■

From Proposition 3.8, 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 3.8 and Lemma 3.3 (1). The sufficiency in (1) has already been proven in Proposition 3.5. Note that the condition in (2) is actually also sufficient, but its proof has to wait until the proving of Theorem 3.1 (cf. Proposition 3.13 later).

The next proposition is the major step in proving the sufficiency of Theorem 3.1.

Proposition 3.9. *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* $\text{Re}\,(e^{-i\theta}A)$.

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

For the proof, we need the following lemma.

Lemma 3.10. *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 $B = \mathrm{Re}\, A[n]$ 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 λ 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 3.6 (3). Hence λ, being the radius of the circular disc $W(A[n])$, does not belong to $W(A[j+1, \ldots, n])$. In particular, λ 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

$$\begin{aligned}
0 &= \det(\lambda I_{n-1} - B) \\
&= \det(\lambda I_j - A_j) \det((\lambda I_{n-j-1} - D_j) - (-C_j)(\lambda I_j - A_j)^{-1}(-B_j)) \\
&= \det(\lambda I_j - A_j) \det((\lambda I_{n-j-1} - D_j) - \frac{1}{4} a_j^2 \frac{\det(\lambda I_{j-1} - A_{j-1})}{\det(\lambda I_j - A_j)} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}) \\
&= \det(\lambda I_j - A_j)(\det(\lambda I_{n-j-1} - D_j) - \frac{1}{4} a_j^2 \frac{\det(\lambda I_{j-1} - A_{j-1})}{\det(\lambda I_j - A_j)} \det(\lambda I_{n-j-2} - D_{j+1})),
\end{aligned}$$

from which we obtain

$$a_j^2 = 4 \frac{\det(\lambda I_j - A_j) \det(\lambda I_{n-j-1} - D_j)}{\det(\lambda I_{j-1} - A_{j-1}) \det(\lambda I_{n-j-2} - D_{j+1})}$$

for $2 \leq j \leq n-3$. Taking the product of the a_j^2's yields

$$\begin{aligned}
a_2^2 \ldots a_{n-3}^2 \lambda^2 &= 4^{n-4} \frac{\det(\lambda I_{n-3} - A_{n-3}) \det(\lambda I_{n-3} - D_2)}{\det(\lambda I_1 - A_1) \det(\lambda I_1 - D_{n-2})} \lambda^2 \\
&= 4^{n-4} \det(\lambda I_{n-3} - A_{n-3}) \det(\lambda I_{n-3} - D_2)
\end{aligned}$$

since A_1 and D_{n-2} are both the 1-by-1 zero matrix. This completes the proof. ■

Proof of Proposition 3.9. (1) We may assume, by Lemma 3.3 (1), that $W(A[n-1]) = W(A[n]) = W(A[1]) = B(0;r)$. Also, by Lemma 3.3 (2), A is unitarily equivalent to $e^{i\theta}C$, 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 $\partial W(C)$ by Lemma 3.4 (2) 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 nth row to obtain

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

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

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

(resp.,

$$d_{n,n-1} = \frac{1}{2^{n-1}}|a_1 \cdots a_{n-2}a_n| - \frac{1}{2}|a_{n-1}| \det C_2),$$

where

$$C_1 = \begin{bmatrix} r & -|a_2|/2 & & \\ -|a_2|/2 & r & \ddots & \\ & \ddots & \ddots & -|a_{n-2}|/2 \\ & & -|a_{n-2}|/2 & r \end{bmatrix}$$

(resp.,

$$C_2 = \begin{bmatrix} r & -|a_1|/2 & & \\ -|a_1|/2 & r & \ddots & \\ & \ddots & \ddots & -|a_{n-3}|/2 \\ & & -|a_{n-3}|/2 & r \end{bmatrix}).$$

Hence

$$\begin{aligned} \det(rI_n - \text{Re } C) &= \frac{1}{2^n}|a_1 \cdots a_n| - \frac{1}{4}|a_n|^2 \det C_1 + \frac{1}{2^n}|a_1 \cdots a_n| - \frac{1}{4}|a_{n-1}|^2 \det C_2 \\ &= \frac{1}{2^{n-1}}|a_1 \cdots a_n| - \frac{1}{4}|a_n|^2 \det C_1 - \frac{1}{4}|a_{n-1}|^2 \det C_2. \end{aligned} \tag{ii}$$

On the other hand, let

$$D_j = \begin{bmatrix} r & -|a_j|/2 & & \\ -|a_j|/2 & r & \ddots & \\ & \ddots & \ddots & -|a_{n+j-5}|/2 \\ & & -|a_{n+j-5}|/2 & r \end{bmatrix}$$

for $j = 1$, 2 and 3. Since $\det(rI_{n-1} - \text{Re } C[n]) = 0$, expanding this determinant along its first row (resp., its last row) yields $r \det C_1 = (|a_1|^2/4) \det D_3$ (resp., $r \det C_2 = (|a_{n-2}|^2/4) \det D_1$). Similarly, from $\det(rI_{n-1} - \text{Re } C[1]) = 0$ (resp., $\det(rI_{n-1} - \text{Re } C[n-1]) = 0$), we obtain $r \det C_1 = (|a_{n-1}|^2/4) \det D_2$ (resp., $r \det C_2 = (|a_n|^2/4) \det D_2$). Since r is the maximum eigenvalue of $\text{Re } C[j]$ for $j = 1, n-1$ and n, we have $\det C_j \geq 0$ for $j = 1$ and 2, and $\det D_j \geq 0$ for $j = 1$, 2 and 3. Thus (ii) becomes

$$\det(rI_n - \text{Re } C) = \frac{1}{2^{n-1}}|a_1 \cdots a_n| - \frac{1}{4}|a_n|^2 \frac{|a_{n-1}|^2}{4r} \det D_2 - \frac{1}{4}|a_{n-1}|^2 \frac{|a_n|^2}{4r} \det D_2$$

$$\begin{aligned} &= \frac{1}{2^{n-1}}|a_1 \cdots a_n| - \frac{1}{8r}|a_{n-1}a_n|^2 \frac{|a_{n-2}|}{|a_n|}(\det D_1)^{1/2} \frac{|a_1|}{|a_{n-1}|}(\det D_3)^{1/2} \\ &= \frac{1}{2^{n-1}}|a_1 \cdots a_n| - \frac{1}{8r}|a_1 a_{n-2} a_{n-1} a_n|(\det D_1 \cdot \det D_3)^{1/2} = 0 \end{aligned}$$

by Lemma 3.10. Hence $\det(rI_n - \text{Re }(e^{-i\theta}A)) = 0$. Since r is the maximum eigenvalue of Re $(e^{-i\theta}A)[1]$, this shows that it is either the largest or the second largest eigenvalue of Re $(e^{-i\theta}A)$.

(2) From our assumption and the proof of (1), we have $\det(rI_{n-1} - \text{Re } C[j]) = \det(rI_n - \text{Re } C) = 0$ for all j, $1 \leq j \leq n$. Thus if $p(z) = \det(zI_n - \text{Re } C)$, then $p'(r) = \sum_{j=1}^{n} \det(rI_{n-1} - \text{Re } C[j]) = 0$ (cf. [24, p. 43, Problem 4]). This shows that the eigenvalue r of 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 $r \leq w_0(C)$. If $r < w_0(C)$, then we deduce from the facts that $w_0(C)$ is the maximum eigenvalue of 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 3.6 (3) since the a_j's are nonzero. Hence we must have $r = w_0(C) = w_0(A)$, which is the largest eigenvalue of Re $(e^{-i\theta}A)$ with multiplicity at least two. ■

Another result which we need is the following condition for the line segment on the boundary of a numerical range. It is from [16, Lemma 1.4].

Lemma 3.11. *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* Re $(e^{-i\theta}A)$, *which has unit eigenvectors x_1 and x_2 such that* Im $\langle e^{-i\theta}Ax_1, x_1\rangle \neq$ Im $\langle e^{-i\theta}Ax_2, x_2\rangle$.

Lemma 3.12. *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* Re A *with multiplicity at least two.*

Proof. In view of Lemma 3.11, 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 $a_1 b_2 + a_n b_n = 0$, $2db_2 = a_2 b_3$, $2db_j = a_{j-1}b_{j-1} + a_j b_{j+1}$ 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

$$\begin{aligned}
& a_n b_n c_1 = -a_1 b_2 c_1 = a_2 b_2 c_3 = (2db_3 - a_3 b_4)c_3 \\
= \ & a_3(b_3 c_4 - b_4 c_3) = a_3 b_3 c_4 - (2dc_4 - a_4 c_5) b_4 \\
= \ & a_4(b_4 c_5 - b_5 c_4) \\
= \ & \cdots \\
= \ & a_{n-1}(b_{n-1}c_n - b_n c_{n-1}).
\end{aligned}$$

Letting $x_1 = (b + ic)/\|b + ic\|$ and $x_2 = b/\|b\|$, we have

$$\begin{aligned}
\text{Im}\langle Ax_1, x_1\rangle &= \frac{1}{\|b+ic\|^2}\Big(-a_1 b_2 c_1 + a_2 b_2 c_3 + \sum_{j=3}^{n-1} a_j(b_j c_{j+1} - b_{j+1}c_j) + a_n b_n c_1\Big) \\
= \ & -\frac{n a_1 b_2 c_1}{\|b+ic\|^2} \neq 0 = \text{Im}\langle Ax_2, x_2\rangle.
\end{aligned}$$

Our assertion follows from Lemma 3.11. ■

We are now ready to prove the sufficiency of Theorem 3.1.

Proof of Theorem 3.1. Assume that $a_j \neq 0$ and $W(A[j]) = B(0; r)$ for all j, $1 \leq j \leq n$. By Lemma 3.3 (2), A is unitarily equivalent to $e^{i\theta}C$, where C is the n-by-n weighted shift matrix with weights $|a_1|, \ldots, |a_{n-1}|, -|a_n|$ and $\theta = (\pi + \sum_{j=1}^{n} \arg\ a_j)/n$. By Proposition 3.9 (2), $r = w_0(C) = w_0(A)$ is the largest eigenvalue of Re C with multiplicity at least two. Lemma 3.12 then implies that $\partial W(C)$ has a line segment on the line $x = r$. Thus $\partial W(A)$ has a line segment on $x\cos\theta + y\sin\theta = r = w_0(A)$. This completes the proof. ■

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 H.-L. Gau and P. Y. Wu some years ago.

Proposition 3.13. *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) $\Leftrightarrow$ (2). Since the characteristic polynomial of Re $A[1]$ is

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

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 3.1 and Lemma 3.3 (4).

(2) $\Rightarrow$ (3). Since $\det(zI_4 - A) = z^4 - a_1a_2a_3a_4$, the eigenvalues of A are $\alpha_j \equiv (a_1a_2a_3a_4)^{1/4}\omega_4^j$, $0 \leq j < 4$. Their respective eigenvectors can be computed to be (multiples of) $x_j \equiv [1 \ \ \alpha_j/a_1 \ \ {\alpha_j}^2/(a_1a_2) \ \ {\alpha_j}^3/(a_1a_2a_3)]^T$, $0 \leq j < 4$. Note that

$$\langle x_j, x_k\rangle = 1 + \frac{1}{|a_1|^2}\alpha_j\overline{\alpha_k} + \frac{1}{|a_1a_2|^2}(\alpha_j\overline{\alpha_k})^2 + \frac{1}{|a_1a_2a_3|^2}(\alpha_j\overline{\alpha_k})^3$$

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

$$\langle x_1, x_2\rangle = \langle x_1, x_4\rangle = \langle x_3, x_2\rangle = \langle x_3, x_4\rangle = 0.$$

Let $y_1 = x_1 - x_3 = [0 \ \ 2\alpha_0/a_1 \ \ 0 \ \ 2{\alpha_0}^3/(a_1a_2a_3)]^T$, $y_2 = x_1 + x_3 = [2 \ \ 0 \ \ 2{\alpha_0}^2/(a_1a_2) \ \ 0]^T$, $y_3 = x_2 - x_4 = [0 \ \ 2i\alpha_0/a_1 \ \ 0 \ \ -2i{\alpha_0}^3/(a_1a_2a_3)]^T$ and $y_4 = x_2 + x_4 = [2 \ \ 0 \ \ -2{\alpha_0}^2/(a_1a_2) \ \ 0]^T$, and let M be the subspace of $\mathbb{C}^4$ spanned by y_1 and y_2. Since $Ay_1 = Ax_1 - Ax_3 = \alpha_1x_1 - \alpha_3x_3$ and $Ay_2 = Ax_1 + Ax_3 = \alpha_1x_1 + \alpha_3x_3$, and M is also spanned by x_1 and x_3, we have $AM \subseteq M$. A simple computation shows that $A^*y_1 = (\overline{a_4}{\alpha_0}^3/(a_1a_2a_3))y_2$ and $A^*y_2 = (|a_1|^2/\alpha_0)y_1$, 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 $\langle y_1, y_2\rangle = \langle y_3, y_4\rangle = \langle Ay_j, y_j\rangle = 0$ for all j. Therefore, A is unitarily equivalent to a matrix of the form $\begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & c_1 \\ c_2 & 0 \end{bmatrix} \equiv B \oplus C$ on $M \oplus M^\perp$. Since x_1 and x_3 are in M, α_1 and α_3 are eigenvalues of B. Hence

$$-b_1b_2 = \det B = \alpha_1\alpha_3 = \alpha_0^{1/2}\omega_4^2 = -\alpha_0^{1/2}.$$

A similar argument with C yields $-c_1c_2 = \alpha_0^{1/2}$. It follows that $b_1b_2 = -c_1c_2$.

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

$$b_1 = \langle A\frac{y_2}{\|y_2\|}, \frac{y_1}{\|y_1\|}\rangle = \alpha_0\frac{\|y_1\|}{\|y_2\|},$$

$$b_2 = \langle A\frac{y_1}{\|y_1\|}, \frac{y_2}{\|y_2\|}\rangle = \alpha_0\frac{\|y_2\|}{\|y_1\|},$$

$$c_1 = \langle A\frac{y_4}{\|y_4\|}, \frac{y_3}{\|y_3\|}\rangle = i\alpha_0\frac{\|y_3\|}{\|y_4\|},$$

and

$$c_2 = \langle A\frac{y_3}{\|y_3\|}, \frac{y_4}{\|y_4\|}\rangle = i\alpha_0\frac{\|y_4\|}{\|y_3\|},$$

and $\|y_1\| = \|y_3\|$ and $\|y_2\| = \|y_4\|$. Thus

$$\begin{aligned} |b_1|^2 + |b_2|^2 &= |\alpha_0|^2(\frac{\|y_1\|^2}{\|y_2\|^2} + \frac{\|y_2\|^2}{\|y_1\|^2}) \\ = \quad & |\alpha_0|^2(\frac{\|y_3\|^2}{\|y_4\|^2} + \frac{\|y_4\|^2}{\|y_3\|^2}) = |c_1|^2 + |c_2|^2 \end{aligned}$$

as asserted.

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

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

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

Recall that a set S is *nowhere dense* in $\mathbb{C}^n$ if its closure contains no nonempty open subset of $\mathbb{C}^n$. By Theorem 3.1, we can obtain the following result.

Proposition 3.14. *Let* $S = \{[a_1, \ldots, a_n]^T \in \mathbb{C}^n : \partial W(\begin{bmatrix} 0 & a_1 & & \\ & 0 & \ddots & \\ & & \ddots & a_{n-1} \\ a_n & & & 0 \end{bmatrix})$ *has a line segment*$\}$ $(n \geq 2)$. *Then* S *is closed and nowhere dense in* $\mathbb{C}^n$.

Proof. (1) For any $a = [a_1, \ldots, a_n]^T \in \mathbb{C}^n$, let A be the n-by-n $(n \geq 3)$ weighted shift matrix with weights $a_1, \ldots, a_n$. Then Theorem 3.1 implies

$$\begin{aligned} S &= \{[a_1, \ldots, a_n]^T \in \mathbb{C}^n : \partial W(A) \text{ has a line segment}\} \\ &= \{[a_1, \ldots, a_n]^T \in \mathbb{C}^n : W(A[1]) = \cdots = W(A[n]),\ a_i \neq 0 \text{ for all } i\} \\ &= \{[a_1, \ldots, a_n]^T \in \mathbb{C}^n : \|\text{Re } (A[1])\| = \cdots = \|\text{Re } (A[n])\|,\ a_i \neq 0 \text{ for all } i\} \end{aligned}$$

Let $\{r_k\} = \{[a_{1k}, \ldots, a_{nk}]^T\} \subseteq S$ be a convergent sequence, $a = [a_1, \ldots, a_n]^T = \lim r_k$. Since $|a_{ik} - a_i| \leq \|r_k - a\|$ for all i and k, we have $a_{ik} \to a_i$ as $k \to \infty$ for all i. Let A_k (resp., A) be the n-by-n weighted shift matrices with weights $a_{1k}, \ldots, a_{nk}$ (resp., $a_1, \ldots, a_n$), then we obtain $\|\text{Re } ((A_k)[i])\| \to \|\text{Re } (A[i])\|$ as $k \to \infty$ for all i. Since $\{r_k\} \subseteq S$, we have $\|\text{Re } ((A_k)[i])\| = \|\text{Re } ((A_k)[j])\|$ for all i, j, k, $1 \leq i < j \leq n$ and the a_{ik}'s are nonzero. Therefore, we show that $\|\text{Re } (A[1])\| = \cdots = \|\text{Re } (A[n])\|$ and the a_j's are nonzero. This implies $a \in S$ and hence S is closed in $\mathbb{C}^n$.

(2) We need only show that S has no interior point. Suppose there exists $a = [a_1, \ldots, a_n]^T$ in the interior of S. Then we have $B(a; r) \subseteq S$ for some r, $r > 0$. Let $b = [(|a_1| + r/2)e^{i\arg a_1}, a_2, \ldots, a_n]^T \in \mathbb{C}^n$ and A (resp., B) be the n-by-n weighted shift

matrix with weights $a_1, \ldots, a_n$ (resp., $(|a_1|+r/2)e^{i\arg a_1}, a_2, \ldots, a_n$). Hence both a and b are in S by Theorem 3.1. These imply $W(A[n]) = W(A[1])$, $W(B[n]) = W(B[1])$ and the a_j's are nonzero. Since $A[1] = B[1]$, we obtain that $a_1 = (|a_1| + r/2)e^{i\arg a_1}$ by Lemma 3.6 (2). It is impossible. This completes the proof. ■

3.3 Nilpotent matrices with n weights

Consider the n-by-n nilpotent matrix of the form (i) with weights $a_1, \ldots, a_n$. We start with the main result of this section, which is analogous to Theorem 3.1.

Theorem 3.15. *Let A be an n-by-n ($n \geq 3$) nilpotent matrix of the form* (i) *with weights $a_1, \ldots, a_n$. 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 $\theta_k = ((\sum_{j=1}^{n-1}\text{arg } a_j) - \text{arg } a_n + (2k+1)\pi)/(n-2)$, $0 \leq k < n-3$, and there are exactly $n-2$ line segments on $\partial W(A)$.*

For the proof of Theorem 3.15, we also need a fuller understanding of the numerical range of the n-by-n nilpotent matrix of the form (i) with weights $a_1, \ldots, a_n$. This is provided by the following lemmas.

Lemma 3.16. *Let A and B be n-by-n ($n \geq 3$) nilpotent matrix of the form* (i) *with weights $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$, respectively.*

(1) *If $|a_j| = |b_j|$ for all j, then A is unitarily equivalent to $e^{i\beta_k}B$, where $\beta_k = (\sum_{j=1}^{n-1}(\arg a_j - \arg b_j) + (\arg b_n - \arg a_n) + 2k\pi)/(n-2)$ for $0 \leq k < n-2$. In particular, A is unitarily equivalent to $w_{n-2}A$ and hence $W(A)$ has $n-2$ symmetry.*

(2) (a) *Either the intersection number of $\partial W(A)$ and $\partial B(0; w(A))$ is $n-2$ or $W(A) = B(0; w(A))$.*

(b) *Either the intersection number of $\partial W(A)$ and $\partial B(0; w_0(A))$ is $n-2$ or $W(A) = B(0; w_0(A))$.*

(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* θ,

(c) $W(A)$ *is a circular disc centered at the origin, and*

(d) $\partial W(A)$ *contains an elliptic arc.*

(4) *If* $\partial W(A)$ *has a line segment* L, *then* $\mathrm{dist}(0, L) = w_0(A)$ *and there are exactly* $n-2$ *line segments on* $\partial W(A)$.

Proof. (1) If $U = \mathrm{diag}(e^{i\phi_1}, \ldots, e^{i\phi_n})$, where $\phi_1 = 0$ and $\phi_j = \phi_{j-1} + (\arg b_{j-1} - \arg a_{j-1}) + \beta_k$ for $2 \le j \le n$, then U is unitary and $AU = U(e^{i\beta_k}B)$. In particular, A is unitarily equivalent to $w_{n-2}A$ by letting $B = A$ and $k = 1$.

(2) (a) follows from [18, Proposition 3.1] and (1). (b) follows from [18, Theorems 2.5 (b), 3.5 (a) and (b)] and (1).

(3) If (a) holds, then the β_k's in (1) can be arbitrary. Letting $B = A$ in there, we obtain (b). The implications (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (d) are trivial. To prove (c) $\Rightarrow$ (a), let λ be the maximum eigenvalue of $\mathrm{Re}(\omega A)$ for all $|\omega| = 1$. Then we have $\det(\lambda I - \mathrm{Re}(\omega A)) = 0$ for all $|\omega| = 1$. Since

$$\det(\lambda I - \mathrm{Re}(\omega A)) = f(\lambda) - \frac{a_1 \cdots a_{n-1}\overline{a_n}}{2^n}\omega^{n-2} - \frac{\overline{a_1 \cdots a_{n-1}}a_n}{2^n}\overline{\omega}^{n-2}$$

for some polynomial $f(\lambda)$ which is independent of ω with degree $f(\lambda) \le n$, it can be considered as a trigonometric polynomial in ω which has infinitely many zeros. Hence the coefficients of ω^{n-2} and $\overline{\omega}^{n-2}$ are both zero and we obtain $a_j = 0$ for some j. We finally show that (d) $\Rightarrow$ (c). If E is an elliptic disc and $\partial W(A)$ contains an arc of ∂E, then we have $E \subseteq W(A)$ and the two foci of ∂E are the eigenvalues of A by [14].

Hence $E = B(0; w_0(A))$ is a closed circular disc centered at the origin. Therefore we have $W(A) = E$ by (2b) and this proves our assertion.

(4) For $w_0(A) = \text{dist}(0, L)$, the proof is the same as the one in Lemma 3.3 (5). Additionally, there are exactly $n-2$ line segments on $\partial W(A)$ by (1) and (2b). ■

Therefore, the necessity of Theorem 3.15 follows easily from Lemmas 3.2 and 3.16 (3), (4). To prove the sufficiency of Theorem 3.15, we need several lemmas and propositions. We start with the following.

Lemma 3.17. *Let A be an n-by-n ($n \geq 3$) nilpotent matrix of the form* (i) *with weights $a_1, \ldots, a_n$, and let $\theta = ((\sum_{j=1}^{n-1} \arg a_j) - \arg a_n)/(n-2)$.*

(1) *$\partial W(A)$ is a differentiable curve.*

(2) *$W(A)$ is symmetric with respect to the lines $y = x \tan((k\pi/(n-2)) + \theta)$ for $0 \leq k < n-2$.*

(3) *We have $\{\arg \lambda : \lambda \in \partial W(A), |\lambda| = w(A)\} = \{(2k\pi/(n-2)) + \theta : 0 \leq k < n-2\}$ and $\{\arg \lambda : \lambda \in \partial W(A), |\lambda| = w_0(A)\} = \{((2k+1)\pi/(n-2)) + \theta : 0 \leq k < n-2\}$.*

(4) *For $n \geq 5$, $w(A) \leq w_0(A) \sec(\pi/(n-2))$ and*

$$B(0; w_0(A)) \subseteq W(A) \subseteq w_0(A)(\sec \frac{\pi}{n-2}) e^{i\theta} \{1, \omega_{n-2}, \ldots, \omega_{n-2}^{n-3}\}^{\wedge}.$$

In addition, $w(A) < w_0(A) \sec(\pi/(n-2))$ if $a_j \neq 0$ for all j.

Proof. (1) If there exists $\lambda \in \partial W(A)$ such that λ is not differentiable, then λ is a reducing eigenvalue of A. However $\sigma(A) = \{0\}$ implies $\lambda = 0$. This contradicts the fact that 0 does not belong to $\partial W(A)$. Hence λ is differentiable for all $\lambda \in W(A)$.

The proofs of (2), (3) and (4) are essentially similar to those of Lemma 3.4 (1), (2) and (3) by replacing $\theta = (\sum_{j=1}^{n} \arg a_j)/n$ in $\theta = ((\sum_{j=1}^{n-1} \arg a_j) - \arg a_n)/(n-2)$. We only need show that $w(A) \neq w_0(A)\sec(\pi/(n-2))$ if $a_j \neq 0$ for all j in (4). If $w(A) = w_0(A)\sec(\pi/(n-2))$, then (4) says that $W(A) \subseteq w(A)e^{i\theta}\{1, \omega_{n-2}, \ldots, \omega_{n-2}^{n-3}\}^{\wedge}$. But the vertices of this latter regular $n-2$-polygonal region, namely, $w(A)e^{i\theta}\omega_{n-2}^{k}$, $0 \leq k < n-2$, are in $W(A)$ by (3). Hence we must have $W(A) = w(A)e^{i\theta}\{1, \omega_{n-2}, \ldots, \omega_{n-2}^{n-3}\}^{\wedge}$. Hence $\partial W(A)$ has nondifferentiable points. It contradicts (1). Consequently our assertion follows. ■

The next proposition is the major step in proving the sufficiency of Theorem 3.15.

Proposition 3.18. *Let A be an n-by-n $(n \geq 3)$ nilpotent matrix of the form* (i) *with nonzero weights $a_1, \ldots, a_n$, and let $\theta = (\pi + (\sum_{j=1}^{n-1}\text{arg } a_j) - \text{arg } a_n)/(n-2)$.*

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.*

Proof. (1) By Lemma 3.16 (1), A is unitarily equivalent to $e^{i\theta}B$, where B is the n-by-n nilpotent matrix of the form (i) with nonzero weights $|a_1|, \ldots, |a_{n-1}|, -|a_n|$. Let C be the n-by-n weighted shift matrix with weights $|a_1|, \ldots, |a_{n-1}|, -|a_n|$. Then we have $W(A_k) = W(B_k) = W(C_k) = B(0; r)$ for $k = j-1$, j and $j+1$. Since $\det(rI_n - \text{Re}$

$B) = \det(rI_n - \text{Re } C)$, Proposition 3.9 (1) implies that r is either the largest or the second largest eigenvalue of Re $(e^{-i\theta}A)$.

(2) In the proof of (1), by our assumption and Proposition 3.9 (2), we obtain that $r = w_0(C)$ is the largest eigenvalue of Re C with multiplicity at least two. Since $\det(rI_n - \text{Re } B) = \det(rI_n - \text{Re } C)$, Lemmas 3.4 (2) and 3.17 (3) implies that $r = w_0(C) = w_0(B) = w_0(A)$, which is the largest eigenvalue of Re $(e^{-i\theta}A)$ *with multiplicity at least two.* ■

To prove the sufficiency of Theorem 3.15, we also need the following lemma:

Lemma 3.19. *Let A be an n-by-n ($n \geq 3$) nilpotent matrix of the form* (i) *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* Re A *with multiplicity at least two.*

Proof. In view of Lemma 3.11, we need only prove the sufficiency part. Let B and C be n-by-n weighted shift matrices with nonzero real weights $a_1, \ldots, a_n$ and $0, \ldots, 0, -2a_n$, respectively. Then Re A = Re B and Im A = Im B + Im C. Hence by our assumption and the proof of Lemma 3.12, there exist real unit vectors x_1 and x_2 such that $\langle \text{Im } Bx_1, x_1 \rangle = \text{Im } \langle Bx_1, x_1 \rangle \neq \text{Im } \langle Bx_2, x_2 \rangle = \langle \text{Im } Bx_2, x_2 \rangle$. In addition, we also have $\langle \text{Im } Cx_1, x_1 \rangle = \langle \text{Im } Cx_2, x_2 \rangle = 0$. Therefore, $\text{Im } \langle Ax_1, x_1 \rangle = \langle \text{Im } Bx_1, x_1 \rangle \neq \langle \text{Im } Bx_2, x_2 \rangle = \text{Im } \langle Ax_2, x_2 \rangle$. Finally, Lemma 3.11 implies our assertion. ■

We are now ready to prove the sufficiency of Theorem 3.15.

Proof of Theorem 3.15. Assume that $a_j \neq 0$ and $W(A[j]) = B(0; r)$ for all j,

$1 \leq j \leq n$. By Lemma 3.16 (1), A is unitarily equivalent to $e^{i\theta}B$, where B is the n-by-n nilpotent matrix of the form (i) with weights $|a_1|, \ldots, |a_{n-1}|, -|a_n|$ and $\theta = (\pi + (\sum_{j=1}^{n-1} \arg a_j) - \arg a_n)/(n-2)$. By Proposition 3.18 (2), $r = w_0(B) = w_0(A)$ is the largest eigenvalue of Re B with multiplicity at least two. Lemma 3.19 then implies that $\partial W(B)$ has a line segment on the line $x = r$. Thus $\partial W(A)$ has a line segment on $x\cos\theta + y\sin\theta = r = w_0(A)$. This completes the proof. ■

The following is an easy corollary.

Corollary 3.20. *Let A (resp., B) be the n-by-n ($n \geq 3$) weighted shift matrix (resp., nilpotent matrix of the form* (i)*) 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.*

Proof. (1) If the a_i's are real, then Re $A =$ Re B. Hence we can assume that the a_i's are real by Lemmas 3.3 (2) and 3.16 (1) and the result can be obtained by Lemmas 3.4 (2) and 3.17 (3).

(2) The proof is similar to the one of (1).

(3) This follows directly from Theorems 3.1 and 3.15 and that $W(A_i) = W(B_i)$ for all i. ■

3.4 Weighted shift matrix with periodic weights

The purpose of this section is to study the numerical ranges of the n-by-n weighted shift matrices with periodic weights. The main result of this section is the following.

Theorem 3.21. *Let A be an n-by-n $(n \geq 3)$ weighted shift matrix with nonzero weights $a_1, \ldots, a_n$. Assume that $|a_j| = |a_{k+j}| = \cdots = |a_{(m-1)k+j}|$ for all $1 \leq j \leq k$, where $n = km$ for some k* and *m, $k, m \geq 2$. Then*

(a) *p_A is reducible and $W(A) = W(B)$, where $B = C \oplus (e^{i\theta}C) \oplus \cdots \oplus (e^{i(m-1)\theta}C)$ and C is the k-by-k weighted shift matrix with weights $a_1, \ldots, a_{k-1}, \alpha a_k$, $\alpha = (a_1 \cdots a_n)^{1/m}/(a_1 \cdots a_k)$ and $\omega_n = e^{2\pi i/n}$.*

(b) *$\partial W(A)$ has a line segment L and $\operatorname{dist}(0, L) = w_0(A) = w(A[i]) =$ maximum zero of $\det(\lambda I_{n-1} - \operatorname{Re} A[i])$ for every i, $1 \leq i \leq n$.*

Note that $\partial W(A)$ has a line segment for $k = 1$ by Proposition 3.5.

An easy consequence of the preceding theorem is the following:

Corollary 3.22. *Let A be an n-by-n $(n \geq 3)$ weighted shift matrix with weights $a_1, \ldots, a_n$. Suppose that $n-2$ of the a_j's have equal absolute value and the remaining two terms are a_k and a_l. Then $\partial W(A)$ has a line segment if and only if all the a_i's are nonzero and either*

(a) *n is even, $|k - l| = n/2$, $|a_k| = |a_l|$, or*

(b) *all the a_i's have the same absolute values.*

Proof. The sufficiency follows easily from Theorem 3.21. Now we prove the necessity. By Theorem 3.1, we have that the a_i's are nonzero and $W(A[1]) = \cdots = W(A[n])$. If n is even and $|k-l| = n/2$, then we may assume that $k = n/2$, $l = n$, $a_i = a_j > 0$ for $1 \le i < j \le n-1$, i, $j \ne n/2$ and $a_{n/2}$, $a_n > 0$ by Lemma 3.3 (1) and (2). Also, $W(A[n/2]) = W(A[n])$ implies that $|a_{n/2}| = |a_n|$ by Lemma 3.6. Otherwise, we may assume that $1 \le k < n/2$, $l = n$, $a_i = a_j > 0$ for $1 \le i < j \le n-1$, i, $j \ne k$ and a_k, $a_n > 0$ by Lemma 3.3 (1) and (2). Let $a \equiv a_i$, where $i \ne k$, n. Note that we have $W(A[k]) = W(A[2k]) = W(A[n-k]) = W(A[n])$ and $A[n]$ is the $(n-1)$-by-$(n-1)$ matrix $[s_{ij}]_{i,j=1}^{n-1}$, where $s_{i,i+1} = a$ for $1 \le i \le n-2, i \ne k$, $s_{k,k+1} = a_k$ and $s_{i,j} = 0$ otherwise. By Lemma 3.3 (1), we may assume that $A[k]$ is the $(n-1)$-by-$(n-1)$ matrix $[t_{ij}]_{i,j=1}^{n-1}$, where $t_{i,i+1} = a$ for $1 \le i \le n-2, i \ne n-k$, $t_{n-k,n-k+1} = a_n$ and $t_{i,j} = 0$ otherwise. For the orders of $\{a_n, a_k, a\}$, consider the following three cases:

(1) $a_n \ge a \ge a_k$ or $a_n \le a \le a_k$. Since $W(A[n]) = W(A[k])$, by Lemma 3.6 (2), we infer that $a_n = a = a_k$.

(2) $a_n \ge a_k \ge a$ or $a_n \le a_k \le a$. By Lemma 3.3 (1), we may assume that $A[n-k]$ is the $(n-1)$-by-$(n-1)$ matrix $[u_{ij}]_{i,j=1}^{n-1}$, where $u_{i,i+1} = a$ for $1 \le i \le n-2, i \ne k, 2k$, $u_{k,k+1} = a_n$, $u_{2k,2k+1} = a_k$ and $u_{i,j} = 0$ otherwise. Since $W(A[n]) = W(A[n-k])$, by Lemma 3.6 (2), we also infer that $a_n = a = a_k$.

(3) $a_k \ge a_n \ge a$ or $a_k \le a_n \le a$. By Lemma 3.3 (1), we may assume that $A[2k]$ is the $(n-1)$-by-$(n-1)$ matrix $[v_{ij}]_{i,j=1}^{n-1}$, where $v_{i,i+1} = a$ for $1 \le i \le n-2, i \ne n-k, n-2k$, $v_{n-2k,n-2k+1} = a_n$, $v_{n-k,n-k+1} = a_k$ and $v_{i,j} = 0$ otherwise. Since $W(A[k]) = W(A[2k])$, by Lemma 3.6 (2), we obtain that $a_n = a = a_k$ and complete the proof. ∎

We now proceed to prepare ourselves for the proof of Theorem 3.21. This will be done in a series of lemmas. We start our work with the following lemma.

For the ease of exposition, we introduce some notations which are used in Lemmas 3.23, 3.25 and 3.26 below. Fix $k \in \mathbb{N}$. For every $n \geq 2$, there is a unique $m \in \mathbb{N}$ and a unique $l \in \mathbb{N}$, $0 \leq l \leq k-1$, such that $n = km - l$. Moreover, for every $i \in \mathbb{Z}$, let $(a_i, a_{i+1}, \ldots, a_{i+n-2})$ be any sequence of nonzero real numbers with $a_{i+j} = a_{i+k+j} = \cdots = a_{i+(m-1)k+j}$ if $0 \leq j \leq k-2$ and $a_{i+(k-1)} = a_{i+2k-1} = \cdots = a_{i+(m-1)k-1}$. Finally, let $A_n(a_i)$ be the n-by-n tridiagonal self-adjoint matrix with zeros on its diagonal and $(a_i, a_{i+1}, \ldots, a_{i+n-2})$ the sequence of entries on its subdiagonal.

Lemma 3.23. *Fix $k \geq 3$. For $m \in \mathbb{N}$ and $1 \leq i \leq k$, let the a_i's be real. Then*

(1) $\det(\lambda I_{k-1} - A_{k-1}(a_i)) | \det(\lambda I_{km-1} - A_{km-1}(a_i))$, *and*

(2)

$$\begin{aligned} &\frac{\det(\lambda I_{km-1} - A_{km-1}(a_i))}{\det(\lambda I_{k-1} - A_{k-1}(a_i))} \\ = \ &\det(\lambda I_{k(m-1)} - A_{k(m-1)}(a_{i-1})) \\ &-a_{i+k-2}^2 \det(\lambda I_{k-2} - A_{k-2}(a_i)) \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_i))}{\det(\lambda I_{k-1} - A_{k-1}(a_i))} \qquad (a_0 \equiv a_k). \end{aligned}$$

Proof. (1) For $\lambda \in \sigma(A_{k-1}(a_i))$, there is a nonzero vector $u = [x_1, \ldots, x_{k-1}]^T$ in $\ker(\lambda I_{k-1} - A_{k-1}(a_i))$ with $x_1, x_{k-1} \neq 0$. Letting

$$v = [x_1, \ldots, x_{k-1}, 0, rx_1, \ldots, rx_{k-1}, 0, \ldots, rx_1, \ldots, rx_{k-1}]^T$$

with $r = -a_{i+k-2}x_{k-1}/a_{i+k-1}x_1$, we have $a_{i+k-2}x_{k-1} + a_{i+k-1}rx_1 = 0$. Thus $\lambda \in \sigma(A_{km-1}(a_i))$.

(2) For simplicity, we may assume that $i = 1$. Let $A_{km-1}(a_1)$ 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-$(km-1-j)$, $(km-1-j)$-by-j and $(km-1-j)$-by-$(km-1-j)$, respectively. Thus we have

$$\begin{aligned}
&\frac{\det(\lambda I_{km-1} - A_{km-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} \\
= \ & \det((\lambda I_{k(m-1)} - D_{k-1}) - (-C_{k-1})(\lambda I_{k-1} - A_{k-1})^{-1}(-B_{k-1})) \\
= \ & \det((\lambda I_{k(m-1)} - D_{k-1}) - a_{k-1}^2 \frac{\det(\lambda I_{k-2} - A_{k-2})}{\det(\lambda I_{k-1} - A_{k-1})} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 \end{bmatrix}) \\
= \ & \det(\lambda I_{k(m-1)} - D_{k-1}) - a_{k-1}^2 \det(\lambda I_{k-2} - A_{k-2}) \frac{\det(\lambda I_{k(m-1)-1} - D_k)}{\det(\lambda I_{k-1} - A_{k-1})} \\
= \ & \det(\lambda I_{k(m-1)} - A_{k(m-1)}(a_k)) \\
& -a_{k-1}^2 \det(\lambda I_{k-2} - A_{k-2}(a_1)) \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))}.
\end{aligned}$$

■

Now we consider the circular symmetric functions $S_r(a_1, \ldots, a_n)$, where n and r are nonnegative integers, defined first in [42, p. 496]. S_0 is defined to be 1, while for $r \geq 1$, $S_r(a_1, \ldots, a_n) = \sum\{\prod_{k=1}^{r} a_{\pi(k)} \mid \pi : (1, \ldots, r) \to (1, \ldots, n)$, where $\pi(k) + 1 \leq \pi(k+1)$ for $1 \leq k < r$, and if $\pi(1) = 1$ then $\pi(r) \neq n\}$. These have a nice description: imagine a regular n-gonal with vertices labeled a_1 through a_n. Draw a convex r-gonal in it, with vertices among the a_i's with the restriction that it cannot use an edge of the original polygon. Each term in $S_r(a_1, \ldots, a_n)$ is the product of the vertices of such an r-gon. These functions satisfy many identities, but we need only the following:

(1) $S_r(a_1, \ldots, a_n, 0) = S_r(a_n, \ldots, a_1, 0)$,

(2) $S_{r+1}(a_1,\ldots,a_{n+1},0)=S_{r+1}(a_1,\ldots,a_n,0)+a_{n+1}S_r(a_1,\ldots,a_{n-1},0)$.

Another result from [42, Lemma 1] which we need is the following:

Lemma 3.24. *Let A be an $(n+1)$-by-$(n+1)$ tridiagonal self-adjoint matrix with zeros on its diagonal, and let $(a_1,\ldots,a_n)$ be the sequence of entries on its subdiagonal. Then*

$$\det(I_{n+1}-\mu A)=\sum_{l=0}^{\lfloor (n+1)/2\rfloor} S_l(|a_1|^2,\ldots,|a_n|^2,0)(-1)^l\mu^{2l}$$

for any scalar μ.

Hence we have the following lemma which is derived from the above identities and Lemma 3.24.

Lemma 3.25. *For $n\geq 4$, $1\leq i\leq n$ ($a_0\equiv a_n$), let the a_i's be real and $g_i(\lambda)=\det(\lambda I_n-A_n(a_i))-a_{i-1}^2\det(\lambda I_{n-2}-A_{n-2}(a_{i+1}))$. Then $g_i(\lambda)=g_j(\lambda)$ for $1\leq i\neq j\leq n$.*

Proof. For simplicity, we may assume that $i=1$ and $j=2$. By [42, Lemma 1], we have

$$\begin{aligned}
g_1(\lambda) &= \sum_{l=0}^{\lfloor n/2\rfloor} S_l(a_1^2,\ldots,a_{n-1}^2,0)(-1)^l\lambda^{n-2l}\\
&\quad -a_n^2\sum_{l=0}^{\lfloor (n-2)/2\rfloor} S_l(a_2^2,\ldots,a_{n-2}^2,0)(-1)^l\lambda^{n-2l-2}\\
&= \lambda^n+\sum_{l=1}^{\lfloor n/2\rfloor} S_l(a_1^2,\ldots,a_{n-1}^2,0)(-1)^l\lambda^{n-2l}\\
&\quad -a_n^2\sum_{l_1=1}^{\lfloor n/2\rfloor} S_{l_1-1}(a_2^2,\ldots,a_{n-2}^2,0)(-1)^{l_1-1}\lambda^{n-2l_1}
\end{aligned}$$

$$= \lambda^n + \sum_{l=1}^{\lfloor n/2 \rfloor} (-1)^l [S_l(a_1^2, \ldots, a_{n-1}^2, 0) + a_n^2 S_{l-1}(a_2^2, \ldots, a_{n-2}^2, 0)] \lambda^{n-2l}.$$

$$= \lambda^n + \sum_{l=1}^{\lfloor n/2 \rfloor} (-1)^l [S_l(a_2^2, \ldots, a_{n-1}^2, 0) + a_1^2 S_{l-1}(a_3^2, \ldots, a_{n-1}^2, 0)$$
$$+ S_l(a_2^2, \ldots, a_n^2, 0) - S_l(a_2^2, \ldots, a_{n-1}^2, 0)] \lambda^{n-2l}$$

$$= \lambda^n + \sum_{l=1}^{\lfloor n/2 \rfloor} (-1)^l [S_l(a_2^2, \ldots, a_n^2, 0) + a_1^2 S_{l-1}(a_3^2, \ldots, a_{n-1}^2, 0)] \lambda^{n-2l},$$

where the fourth equality follows from the above identities (1) and (2). Similarly, we also have $g_2(\lambda) = \lambda^n + \sum_{l=1}^{\lfloor n/2 \rfloor} (-1)^l [S_l(a_2^2, \ldots, a_n^2, 0) + a_1^2 S_{l-1}(a_3^2, \ldots, a_{n-1}^2, 0)] \lambda^{n-2l}$ from the above third equality. Thus $g_1(\lambda) = g_2(\lambda)$. ■

The next lemma is the major step in proving Theorem 3.21.

Lemma 3.26. *Fix $k \geq 3$. For $m \geq 3$, $1 \leq i \neq j \leq k$, let the a_i's be real. Then*

(1) $\det(\lambda I_{k-2} - A_{k-2}(a_i)) \det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_i))$
$- \det(\lambda I_{k-1} - A_{k-1}(a_i)) \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_i))$
$= - \prod_{l=1, l \neq i-2}^{k} a_l^2 \det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_i))$
$(a_{-1} \equiv a_{k-1}$, $a_0 \equiv a_k$ *and* $A_1(a_i) \equiv 0)$,

(2) $\frac{\det(\lambda I_{km-1} - A_{km-1}(a_i))}{\det(\lambda I_{k-1} - A_{k-1}(a_i))} - \frac{\det(\lambda I_{km-1} - A_{km-1}(a_j))}{\det(\lambda I_{k-1} - A_{k-1}(a_j))}$
$= \prod_{l=1}^{k} a_l^2 [\frac{\det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_i))}{\det(\lambda I_{k-1} - A_{k-1}(a_i))} - \frac{\det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_j))}{\det(\lambda I_{k-1} - A_{k-1}(a_j))}]$, *and*

(3) $\frac{\det(\lambda I_{km-1} - A_{km-1}(a_i))}{\det(\lambda I_{k-1} - A_{k-1}(a_i))} = \frac{\det(\lambda I_{km-1} - A_{km-1}(a_j))}{\det(\lambda I_{k-1} - A_{k-1}(a_j))}$ *for $m \geq 2$, $1 \leq i \neq j \leq k$.*

Proof. For simplicity, we may assume that $i = 1$ and $j = 2$.

(1) $\quad \det(\lambda I_{k-2} - A_{k-2}(a_1)) \det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_1))$
$- \det(\lambda I_{k-1} - A_{k-1}(a_1)) \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_1))$

$$
\begin{aligned}
&= \det(\lambda I_{k-2} - A_{k-2}(a_1))[\lambda \det(\lambda I - A_{k(m-1)-2}(a_1)) \\
&\quad -a_{k-2}^2 \det(\lambda I_{k(m-1)-3} - A_{k(m-1)-3}(a_1))] \\
&\quad -[\det(\lambda I_{k-2} - A_{k-2}(a_1)) - a_{k-2}^2 \det(\lambda I_{k-3} - A_{k-3}(a_1))] \\
&\quad \cdot \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_1)) \\
&= a_{k-2}^2[\det(\lambda I_{k-3} - A_{k-3}(a_1)) \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_1)) \\
&\quad - \det(\lambda I_{k-2} - A_{k-2}(a_1)) \det(\lambda I_{k(m-1)-3} - A_{k(m-1)-3}(a_1))] \\
&= \cdots \qquad \text{(by induction)} \\
&= a_{k-2}^2 \dots a_2^2 a_1^2[\det(\lambda I_{k(m-1)-(k-1)} - A_{k(m-1)-(k-1)}(a_1)) \\
&\quad -\lambda \det(\lambda I_{k(m-1)-k} - A_{k(m-1)-k}(a_1))] \\
&= -a_1^2 \dots a_{k-2}^2 a_k^2 \det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_1)).
\end{aligned}
$$

(2) By Lemma 3.23 (2), we have

$$
\begin{aligned}
& \frac{\det(\lambda I_{km-1} - A_{km-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} - \frac{\det(\lambda I_{km-1} - A_{km-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))} \\
= \; & [\det(\lambda I_{k(m-1)} - A_{k(m-1)}(a_k)) - a_{k-1}^2 \det(\lambda I_{k-2} - A_{k-2}(a_1)) \\
& \cdot \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))}] - [\det(\lambda I_{k(m-1)} - A_{k(m-1)}(a_1)) \\
& -a_k^2 \det(\lambda I_{k-2} - A_{k-2}(a_2)) \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))}] \\
= \; & a_k^2 \det(\lambda I_{k-2} - A_{k-2}(a_2)) \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))} \\
& -a_{k-1}^2 \det(\lambda I_{k-2} - A_{k-2}(a_1)) \frac{\det(\lambda I_{k(m-1)-1} - A_{k(m-1)-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} \\
& +a_{k-1}^2 \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_1)) - a_k^2 \det(\lambda I_{k(m-1)-2} - A_{k(m-1)-2}(a_2)) \\
= \; & \frac{a_k^2}{\det(\lambda I_{k-1} - A_{k-1}(a_2))}(-a_1^2 \dots a_{k-1}^2) \det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_2)) \\
& -\frac{a_{k-1}^2}{\det(\lambda I_{k-1} - A_{k-1}(a_1))}(-a_1^2 \dots a_{k-2}^2 a_k^2) \det(\lambda I_{k(m-2)-1} - A_{k(m-2)-1}(a_1)) \\
= \; & \prod_{l=1}^{k} a_l^2 [\frac{\det(\lambda I_{km-1} - A_{km-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} - \frac{\det(\lambda I_{km-1} - A_{km-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))}],
\end{aligned}
$$

where the second and third equalities follow from (1) and Lemma 3.25, respectively.

(3) By the second equality in the proof of (2), the assertion for $m = 2$ is easily seen to be true. For $m = 3$, we have

$$\begin{aligned} & \frac{\det(\lambda I_{3k-1} - A_{3k-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} - \frac{\det(\lambda I_{3k-1} - A_{3k-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))} \\ = \ & (a_1a_2a_3)^2[\frac{\det(\lambda I_{k-1} - A_{k-1}(a_1))}{\det(\lambda I_{k-1} - A_{k-1}(a_1))} - \frac{\det(\lambda I_{k-1} - A_{k-1}(a_2))}{\det(\lambda I_{k-1} - A_{k-1}(a_2))}] = 0. \end{aligned}$$

Therefore our assertion follows from the above results, (2) and induction. ■

We are now ready to prove Theorem 3.21.

Proof of Theorem 3.21. (a) Let $B = C \oplus (e^{i\theta}C) \oplus \cdots \oplus (e^{i(m-1)\theta}C)$, where C is the k-by-k weighted shift matrix with weights $a_1, \ldots, a_{k-1}, \alpha a_k$, $\alpha = (a_1 \cdots a_n)^{1/m}/(a_1 \cdots a_k)$ and $\theta = 2\pi/n$. Since $|a_j| = |a_{k+j}| = \cdots = |a_{(m-1)k+j}| \neq 0$ for $1 \leq j \leq k$ and arg $(a_1 \cdots a_n)/((a_1 \cdots a_k)^m \alpha^m) = 0$, we may assume that A is the n-by-n weighted shift matrix with periodic weights $a_1, \ldots, a_{k-1}, \alpha a_k, \ldots, a_1, \ldots, a_{k-1}, \alpha a_k$ by Lemma 3.3 (2). Let the matrix $x\text{Re } A + y\text{Im } A + zI_n$ be partitioned as $\begin{bmatrix} C_{11} & \cdots & C_{1m} \\ \vdots & & \vdots \\ C_{m1} & \cdots & C_{mm} \end{bmatrix}$ with C_{ij} of sizes k-by-k for all i, j, $1 \leq i, j \leq n$. Since $C_{1j} + \cdots + C_{mj} = x\text{Re } C + y\text{Im } C + zI_k$, for all j, $1 \leq j \leq m$, we have

$$\begin{aligned} p_A(x, y, z) &= \det(x\text{Re } A + y\text{Im } A + zI_n) \\ &= \det \begin{bmatrix} C_{11} + \cdots + C_{m1} & C_{12} + \cdots + C_{m2} & \cdots & C_{1m} + \cdots + C_{mm} \\ C_{21} & C_{22} & \cdots & C_{2m} \\ \vdots & \vdots & & \vdots \\ C_{m1} & C_{m2} & \cdots & C_{mm} \end{bmatrix} \end{aligned}$$

$$= \det \begin{bmatrix} x\text{Re } C + y\text{Im } C + zI_k & 0 & \cdots & 0 \\ * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}.$$

Hence $p_C | p_A$. Since A and $\omega_n^j A$ are unitarily equivalent for all integer j, then $p_C | p_{\omega_n^j A}$ for all integer j, consequently, $p_{\omega_n^j C} | p_A$ for all $j = 0, 1, \ldots, m-1$. Note that the real foci of the curve $p_{\omega_n^j C} = 0$ are eigenvalues of $\omega_n^j C$ for each j. Since $\sigma(C) = \{\lambda, \omega_k \lambda, \ldots, \omega_k^{k-1}\lambda\}$, where $\lambda = (a_1 \cdots a_k)^{1/k}$, it follows that $\sigma(\omega_n^i C) \cap \sigma(\omega_n^j C) = \emptyset$ for any $0 \le i < j \le m-1$, thus the homogeneous polynomials $p_{\omega_n^i C}$ and $p_{\omega_n^j C}$ have no common factor for any $0 \le i < j \le m-1$. Therefore, we deduce that $p_A = \prod_{j=0}^{m-1} p_{\omega_n^j C}$ or $W(A) = W(B)$. This completes the proof.

(b) By Lemma 3.3 (5) and its proof, we need only prove that $\partial W(A)$ has a line segment. By assumption and Lemma 3.3 (2), we may assume that $a_j = a_{k+j} = \cdots = a_{(m-1)k+j} > 0$ for every j, $1 \le j \le k$. For $k = 2$, since

$$\begin{aligned}
\det(\lambda I_{2\cdot 2-1} - A_{2\cdot 2-1}(\frac{a_2}{2})) &= \lambda^3 - (a_1^2 + a_2^2)\frac{\lambda}{4} = \det(\lambda I_{2\cdot 2-1} - A_{2\cdot 2-1}(\frac{a_1}{2})) \ , \\
\det(\lambda I_{n-1} - \text{Re } A[1]) &= \det(\lambda I_{2m-1} - A_{2m-1}(\frac{a_2}{2})) \\
&= \lambda \det(\lambda I_{2(m-1)} - A_{2(m-1)}(\frac{a_2}{2})) \\
&\quad - \det(\lambda I_{2(m-1)-1} - A_{2(m-1)-1}(\frac{a_2}{2}))\frac{a_1^2}{4} \ , \text{and} \\
\det(\lambda I_{n-1} - \text{Re } A[2]) &= \det(\lambda I_{2m-1} - A_{2m-1}(\frac{a_1}{2})) \\
&= \lambda \det(\lambda I_{2(m-1)} - A_{2(m-1)}(\frac{a_2}{2})) \\
&\quad - \det(\lambda I_{2(m-1)-1} - A_{2(m-1)-1}(\frac{a_1}{2}))\frac{a_1^2}{4} \ ,
\end{aligned}$$

we obtain $\det(\lambda I_{n-1} - \text{Re } A[1]) = \det(\lambda I_{n-1} - \text{Re } A[2])$ by induction. Similarly, we have $\det(\lambda I_{n-1} - \text{Re } A[i]) = \det(\lambda I_{n-1} - \text{Re } A[j])$ for $1 \le i, j \le n$. Hence $W(A[i]) = W(A[j])$ for every i, j, $1 \le i, j \le n$. For $k \ge 3$, we have $\|A_{km-1}(a_1/2)\| >$

$\|A_{k-1}(a_1/2)\|$ by Lemma 3.6. Thus $w(A[n]) = \|\text{Re } A[n]\| = \|A_{km-1}(a_1/2)\|$ = maximum zero of $\frac{\det(\lambda I - \text{Re } A[n])}{\det(\lambda I - \text{Re } A[k,\ldots,n])}$ by Lemma 3.23. Similarly, we also have $w(A[i])$ = maximum zero of $\frac{\det(\lambda I - \text{Re } A[i])}{\det(\lambda I - \text{Re } A[1,\ldots,i,i+k,\ldots,n])}$ for every i, $1 \leq i \leq n-1$. Hence $W(A[i]) = W(A[j])$ for every i, j, $1 \leq i, j \leq n$ by Lemma 3.26 (3). Therefore, by Theorem 3.1, $\partial W(A)$ has a line segment L with $\text{dist}(0, L) = w(A[i])$ for all i. ■

The next theorem gives a necessary and sufficient condition for an n-by-n weighted shift matrix A to have a noncircular elliptic arc in $\partial W(A)$. Moreover, in this case, $\partial W(A)$ also has a line segment.

Theorem 3.27. *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 noncircular elliptic arc if and only if the a_j's are nonzero, n is even, $|a_1| = |a_3| = \cdots = |a_{n-1}|$, $|a_2| = |a_4| = \cdots = |a_n|$ and $|a_1| \neq |a_2|$. In this case, $W(A) = W(B)$, where $B = C \oplus (e^{i\theta}C) \oplus \cdots \oplus (e^{((n/2)-1)\theta}C)$, $C = \begin{bmatrix} 0 & a_1 \\ \alpha a_2 & 0 \end{bmatrix}$, $\alpha = (a_1 \cdots a_n)^{2/n}/(a_1 a_2)$ and $\theta = 2\pi/n$, and $\partial W(A)$ has a line segment.*

Proof. The sufficiency follows easily from Theorem 3.21 (a) and the fact that $W(\begin{bmatrix} 0 & a_1 \\ \alpha a_2 & 0 \end{bmatrix})$ is a noncircular elliptic disc as $|a_1| \neq |\alpha a_2|$ and both are nonzero, where $\alpha = (a_1 \cdots a_n)^{2/n}/(a_1 a_2)$.

To prove the necessity, by Lemma 3.3 (2), we have that A is unitarily equivalent to $e^{i\phi}A'$, where $\phi = (\sum_{j=1}^{n} \arg a_j)/n$ and A' is the n-by-n weighted shift matrix with weights $|a_1|, \ldots, |a_n|$. Then $\sigma(A) = \{|a_1 \cdots a_n|^{1/n}\omega_n^j : j = 0, 1, \ldots, n-1\}$. Since $\partial W(A)$ has a noncircular elliptic arc, by [14], there is a 2-by-2 matrix C_1 such that $p_{C_1} | p_{A'}$ and $\sigma(C_1) \subseteq \sigma(A')$, say, $\sigma(C_1) = \{\beta, \gamma\}$. From Lemma 3.3 (2), we infer that

$p_{\omega_n^j C_1} | p_{A'}$ and $\sigma(\omega_n^j C_1) \subseteq \sigma(A')$ for all $j = 0, 1, \ldots, n-1$. Therefore, $\sigma(A') \supseteq \{\omega_n^j \beta : j = 0, \ldots, n-1\} \cup \{\omega_n^j \gamma : j = 0, \ldots, n-1\}$. Since these sets $\sigma(A')$, $\{\omega_n^j \beta : j = 0, \ldots, n-1\}$ and $\{\omega_n^j \gamma : j = 0, \ldots, n-1\}$ consist of n distinct elements, we deduce that $\sigma(A') = \{\omega_n^j \beta : j = 0, \ldots, n-1\} = \{\omega_n^j \gamma : j = 0, \ldots, n-1\}$. Therefore, we may assume that $\beta = |a_1 \cdots a_n|^{1/n}$ and $\gamma = \omega_n^{j_0} \beta$ for some j_0. Now, if $\omega_n^{j_0} \neq -1$ or n is odd, then these irreducible homogeneous polynomials $p_{C_1}, p_{\omega_n C_1}, \ldots, p_{\omega_n^{\lfloor n/2 \rfloor} C_1}$ are distinct, it follows that $p_{A'}$ can be divided by the homogeneous polynomial $\prod_{j=0}^{\lfloor n/2 \rfloor} p_{\omega_n^j C_1}$ of degree $2(\lfloor n/2 \rfloor + 1) > n$, this contradicts to the fact that $p_{A'}$ is of degree n. Therefore, we deduce that $\omega_n^{j_0} = -1$ and n is even. Moreover, $p_{A'} = \prod_{j=0}^{(n/2)-1} p_{\omega_n^j C_1}$. On the other hand, since C_1 is 2-by-2 with eigenvalues $\pm |a_1 \cdots a_n|^{1/n}$, by unitarily equivalence, we may assume that $C_1 = \begin{bmatrix} 0 & b_1 \\ b_2 & 0 \end{bmatrix}$, where $b_1, b_2 > 0$, $b_1 \neq b_2$ and $b_1 b_2 = |a_1 \cdots a_n|^{2/n}$. Let $B' = C_1 \oplus \omega_n C_1 \oplus \cdots \oplus \omega_n^{(n/2)-1} C_1$ and B_1 be the n-by-n weighted shift matrix with periodic weights $b_1, b_2, b_1, b_2, \ldots, b_1, b_2$. By Theorem 3.21 (a), we have $p_{B_1} = p_{B'} = p_{A'}$. Compute now the coefficients of $(x^2 + y^2) z^{n-2}$ and y^n of $p_{A'}$ and p_{B_1}. Since $p_{A'} = p_{B_1}$, we have $\sum_{j=1}^n |a_j|^2 = (b_1^2 + b_2^2) n/2$ and $(\prod_{j=1}^{n/2} |a_{2j-1}| - \prod_{j=1}^{n/2} |a_{2j}|)^2 = (b_1^{n/2} - b_2^{n/2})^2$. Hence we may assume that $b_1^{n/2} - b_2^{n/2} = \prod_{j=1}^{n/2} |a_{2j-1}| - \prod_{j=1}^{n/2} |a_{2j}|$. In addition, $b_1 b_2 = |a_1 \cdots a_n|^{2/n}$ implies that $b_1^{n/2} = \prod_{j=1}^{n/2} |a_{2j-1}|$ and $b_2^{n/2} = \prod_{j=1}^{n/2} |a_{2j}|$. We also have

$$\begin{aligned} \sum_{j=1}^{n} |a_j|^2 &= \sum_{j=1}^{\frac{n}{2}} |a_{2j-1}|^2 + \sum_{j=1}^{\frac{n}{2}} |a_{2j}|^2 \\ &\geq \frac{n}{2} (\prod_{j=1}^{\frac{n}{2}} |a_{2j-1}|^2)^{\frac{2}{n}} + \frac{n}{2} (\prod_{j=1}^{\frac{n}{2}} |a_{2j}|^2)^{\frac{2}{n}} = \frac{n}{2} (b_1^2 + b_2^2). \end{aligned}$$

Therefore, the equality holds if and only if $b_1 = |a_{2j-1}| \neq 0$, $b_2 = |a_{2j}| \neq 0$ for all j, $1 \leq j \leq n/2$ and $b_1 \neq b_2$. Let $C = e^{i\phi} C_1$ and $B = e^{i\phi} B_1$. Then C is unitarily equivalent to $\begin{bmatrix} 0 & a_1 \\ \alpha a_2 & 0 \end{bmatrix}$, where $\alpha = e^{i(2\phi - \arg a_1 - \arg a_2)} = (a_1 \cdots a_n)^{2/n} / (a_1 a_2)$ and $W(A) = e^{i\phi} W(A') = e^{i\phi} W(B_1) = W(B)$. This proves our assertion. In particular, it follows from Theorem 3.21 (b) that $\partial W(A)$ has a line segment. ∎

Note that the weighted shift matrix A in the above theorem is a special case of the ones considered in Theorem 3.21. The next theorem is another special case. Recall that a matrix A is said to be *reducible* if it is unitarily equivalent to the direct sun of two other matrices; otherwise, A is *irreducible*. We characterize those n-by-n weighted shift matrices A which are reducible in the following theorem.

Theorem 3.28. *Let A be an n-by-n ($n \geq 2$) weighted shift matrix with weights $a_1, \ldots, a_n$. Then A is reducible if and only if one of the following cases holds*:

(1) $a_i = a_j = 0$ *for some* $1 \leq i < j \leq n$,

(2) n *is odd*, $|a_i| = |a_j| \neq 0$ *for all* $1 \leq i < j \leq n$,

(3) n *is even*, $|a_i| = |a_{i+(n/2)}| \neq 0$ *for all* $1 \leq i \leq n/2$.

In case (1), *A is unitarily equivalent to $B_1 \oplus B_2$, where B_1 and B_2 are the weighted shift matrices with weights $a_{j+1}, \ldots, a_{i-1}, 0$ and $a_{i+1}, \ldots, a_{j-1}, 0$, respectively ($a_r \equiv a_{n+r}$ for $1 \leq r \leq n$, $B_1 \equiv [0]$ if $i = 1$, $j = n$ and $B_2 \equiv [0]$ if $i = j - 1$). Hence $W(A)$ is a circular disc centered at the origin. In case* (2), *A is unitarily equivalent to* $\operatorname{diag}(\alpha, \alpha\omega_n, \ldots, \alpha\omega_n^{n-1})$, *where $\omega_n = e^{2\pi i/n}$ and $\alpha = (a_1 \cdots a_n)^{1/n}$. Hence $W(A)$ is a closed regular n-gonal region centered at the origin and the distance from the origin to its vertices equals $|a_1 \cdots a_n|^{1/n}$. In case* (3), *A is unitarily equivalent to $A_1 \oplus e^{i\theta} A_1$, where $\theta = 2\pi/n$ and A_1 is an $(n/2)$-by-$(n/2)$ weighted shift matrix with weights $a_1, \ldots, a_{(n/2)-1}, \alpha a_{n/2}$, $\alpha = (a_1 \cdots a_n)^{1/2}/(a_1 \cdots a_{n/2})$. In particular, $\partial W(A)$ has a line segment.*

Proof. (1) Let $a_i = a_j = 0$ for some i, j, $1 \leq i < j \leq n$. Also, by Lemma 3.3 (1), we may assume that $j = n$. Then $A = B_1 \oplus B_2$, where B_1 and B_2 are the weighted shift matrices with weights $a_{j+1}, \ldots, a_n, a_1, \ldots, a_{i-1}, 0$ and $a_{i+1}, \ldots, a_{j-1}, 0$, respectively ($a_r \equiv a_{n+r}$ for $1 \leq r \leq n$, $B_1 \equiv [0]$ if $i = 1$, $j = n$ and $B_2 \equiv [0]$ if $i = j - 1$). Hence $W(A)$ is a circular disc centered at the origin. Let $a_i = 0$ for some i, $1 \leq i \leq n$, and $a_j \neq 0$ for all $j \neq i$. Again, by Lemma 3.3 (1), we may assume that $i = n$. Then for any orthogonal projection $P = [p_{ij}]_{i,j=1}^n$ such that $AP = PA$, we have $a_i(p_{i,i} - p_{i+1,i+1}) = a_{i+1}(p_{i+1,i+1} - p_{i+2,i+2}) = 0$ for $1 \leq i \leq n-1$. Thus $p_{1,1} = p_{2,2} = \cdots = p_{n,n}$. In addition, $AP = PA$ also implies that $a_i p_{i+1,1} = 0$ for $1 \leq i \leq n-1$. We substitute $p_{i+1,1} = 0$ in these equalities for $AP = PA$. Then $a_i p_{i+1,2} = a_1 p_{i,1} = 0$ for $2 \leq i \leq n-1$. Proceeding successively with the remaining equalities for $AP = PA$, we have $p_{i,j} = 0$ for $i > j$. Hence the assumption $P = P^* = P^2$ implies that $P = 0$ or $P = I_n$. Therefore, A is irreducible.

(2) If n is odd and $a_i \neq 0$ for all $1 \leq i \leq n$, then we may assume that $a_i > 0$ by Lemma 3.3 (2). For any orthogonal projection $P = [p_{ij}]_{i,j=1}^n$ such that $AP = PA$, we have $a_1(p_{1,1} - p_{2,2}) = a_2(p_{2,2} - p_{3,3}) = \cdots = a_n(p_{n,n} - p_{1,1}) = 0$. Thus $p_{1,1} = p_{2,2} = \cdots = p_{n,n}$. In addition, $AP = PA$ also implies that $a_i p_{i+1,i+2} = a_{i+1} p_{i,i+1}$ and $a_{i+1} p_{i+2,i+1} = a_i p_{i+1,i}$ for $1 \leq i \leq n$ ($p_{n,n+1} \equiv p_{n,1}$, $p_{n+1,n+2} \equiv p_{1,2}$, $p_{n+1,n} \equiv p_{1,n}$, $p_{n+2,n+1} \equiv p_{2,1}$, $a_{n+1} \equiv a_1$). Since $P = P^*$, we have $a_{i+1} p_{i+1,i+2} = a_i p_{i,i+1}$ for $1 \leq i \leq n$. Thus $p_{i,i+1} = 0$ for some i or $a_1 = \cdots = a_n$. Hence $p_{i,i+1} = 0$ for every i, $1 \leq i \leq n$ or $a_1 = \cdots = a_n$. Since $n - 1$ is even, by the same process, we have $p_{i,j} = 0$ for all $i < j$ or $a_1 = \cdots = a_n$. Thus $P = P^* = P^2$ implies that P equals 0 or I_n, or $a_1 = \cdots = a_n$. That is, A is reducible if and only if $|a_1| = \cdots = |a_n| \neq 0$. Hence the assertion on $W(A)$ follows from Proposition 3.5.

(3) If n is even and $a_i \neq 0$ for all $1 \leq i \leq n$, then we may assume that

$a_i > 0$ by Lemma 3.3 (2). For any orthogonal projection $P = [p_{ij}]_{i,j=1}^n$ such that $AP = PA$, following a similar argument as in the proof of (2), we obtain $p_{1,1} = p_{2,2} = \cdots = p_{n,n}$ and $p_{i,j} = 0$ for all $i \neq j$, $|i - j| \neq n/2$. In addition, we also have $a_i p_{i+1,(n/2)+i+1} = a_{(n/2)+i} p_{i,(n/2)+i}$ and $a_{(n/2)+i} p_{(n/2)+i+1,i+1} = a_i p_{(n/2)+i,i}$ for every i, $1 \leq i \leq n/2$ ($p_{(n/2)+1,n+1} \equiv p_{(n/2)+1,1}$, $p_{n+1,(n/2)+1} \equiv p_{1,(n/2)+1}$). Hence $P = P^* = P^2$ implies that P equals 0 or I_n, or $a_1 = a_{(n/2)+1}, \ldots, a_{n/2} = a_n$. Therefore, A is reducible if and only if $|a_i| = |a_{i+(n/2)}|$ for all i, $1 \leq i \leq n/2$. Hence $\partial W(A)$ has a line segment by Theorem 3.21 (b). Moreover, by Lemma 3.3 (2), A is unitarily equivalent to $e^{i\psi}B$, where $\psi = (\sum_{j=1}^n \arg\ a_j)/n$ and B is the n-by-n weighted shift matrix with weights $|a_1|, \ldots, |a_{n/2}|, |a_1|, \ldots, |a_{n/2}|$. Let $U = (1/\sqrt{2}) \begin{bmatrix} I_{n/2} & I_{n/2} \\ I_{n/2} & -I_{n/2} \end{bmatrix}$. Then $U^*BU = B_1 \oplus e^{i\theta}B_1$, where $\theta = 2\pi/n$ and B_1 is the $(n/2)$-by-$(n/2)$ weighted shift matrix with weights $|a_1|, \ldots, |a_{n/2}|$. Hence A is unitarily equivalent to $(e^{i\psi}B_1) \oplus e^{i\theta}(e^{i\psi}B_1)$. Let $A_1 = e^{i\psi}B_1$. Then A_1 is the $(n/2)$-by-$(n/2)$ weighted shift matrix with weights $a_1, \ldots, a_{(n/2)-1}, \alpha a_{n/2}$, where $\alpha = e^{i\phi}$ and $\phi = (n/2)\theta - (\sum_{j=1}^{n/2} \arg\ a_j) = (n/2)(\sum_{j=1}^{n} \arg\ a_j)/n - (\sum_{j=1}^{n/2} \arg\ a_j) = (\sum_{j=1}^{n/2} \arg\ a_{(n/2)+j} - \sum_{j=1}^{n/2} \arg\ a_j)/2$. This proves our assertion. ■

An immediate corollary of Theorem 3.28 and [14] is the following:

Corollary 3.29. *Let A be an n-by-n ($n \geq 3$) weighted shift matrix with weights $a_1, \ldots, a_n$ and $a_i = 0$ for some i, $1 \leq i \leq n$. Then*

(1) *p_A is reducible.*

(2) *A is reducible if and only if $a_j = 0$ for some $j \neq i$, $1 \leq j \leq n$.*

Recall that the reducibility of an n-by-n matrix A implies the reducibility of p_A but the converse is in general not true. We give two examples of weighted shift matrices A for which p_A is reducible but A is irreducible.

Example 3.30.

(1) *If* $A = J_n$ $(n \geq 3)$, *then* A *is irreducible*, p_A *is reducible and* $\partial W(A)$ *has no line segment.*

(2) *If* A *is a* 6-by-6 *weighted shift matrix with weights* $1, 2, 1, 2, 1, 2$, *then* A *is irreducible*, p_A *is reducible but* $\partial W(A)$ *has a line segment.*

Proof. (1) From Lemma 3.3 (4), we obtain that $W(A)$ is a circular disc centered at the origin. Hence the assertion follows directly from [14] and Theorem 3.28.

(2) Follow directly from Theorems 3.21 and 3.28. ■

Chapter 4 Numerical ranges of 4-by-4 nilpotent real matrices

4.1 Introduction

A matrix A is *nilpotent* if $A^k = 0$ for some $k \geq 1$. In this section, we are concerned with the number of line segments on the boundary of the numerical range of a 4-by-4 nilpotent matrix. The study of this problem was started in [18]. It was proven in [18, Theorem 3.4] that the number of line segments on the boundary of the numerical range of a 4-by-4 nilpotent matrix A is less than or equal to 2 under the additional condition that A has a 3-by-3 principal submatrix B with $W(B)$ a circular disc centered at the origin. But it is unknown whether the numbers are less than or equal to 2 for all 4-by-4 nilpotent matrices (cf. Corollary 4.3 below). In Theorem 4.1 below, we show that if A is a 4-by-4 nilpotent real matrix, then $\partial W(A)$ has at most two line segments. The proof is based on Lemmas 3.2 and 3.11 in Chapter 3. The former says that if A is an n-by-n matrix and B is any $(n-1)$-by-$(n-1)$ principal submatix of A, then every line segment of $\partial W(A)$ intersects $\partial W(B)$. The latter gives a necessary and sufficient condition for the existence of line segments on $\partial W(A)$. More precisely, it says that for an n-by-n matrix A, $\partial W(A)$ has a line segment on the line $x\cos\theta + y\sin\theta = d$ if and only if d is the maximum eigenvalue of Re $(e^{-i\theta}A)$ with unit eigenvectors x_1 and x_2 such that Im $\langle e^{-i\theta}Ax_1, x_1\rangle \neq$ Im $\langle e^{-i\theta}Ax_2, x_2\rangle$. In Theorem 4.4, we give a necessary and sufficient condition for the boundary of $W(A)$ to have a pair of parallel line segments. More specifically, it is shown that this is the case if and only if A is unitarily equivalent to a matrix of the form

$$\alpha \begin{bmatrix} 0 & r_1 & ir_2 & -r_3 \\ & 0 & r_3 & -ir_2 \\ & & 0 & r_1 \\ & & & 0 \end{bmatrix},$$

where $\alpha \in \mathbb{C} \setminus \{0\}$, r_1, $r_3 > 0$ and $r_2 \in \mathbb{R}$. Moreover, in this case, $\partial W(A)$ has no other line segment. in Theorem 4.6, we show that a 4-by-4 nilpotent matrix A which has a line segment on the boundary of its numerical range must be unitarily equivalent to a matrix of a special form.

4.2 Numerical ranges of 4-by-4 nilpotent real matrice

The main result of this section is the following.

Theorem 4.1. *Let A be a 4-by-4 nilpotent real matrix. Then $\partial W(A)$ has at most two line segments.*

The proof of Theorem 4.1 depends on some lemmas and propositions. The first one is Lemma 3.2 in Chapter 3. It relates the line segments on $\partial W(A)$ to the numerical ranges of submatrices of A.

Lemma 3.2. *Let A be an n-by-n ($n \geq 2$) matrix and let B be any $(n-1)$-by-$(n-1)$ principal submatrix of A. Then every line segment of $\partial W(A)$ intersects $\partial W(B)$.*

Proposition 4.2. *Let A be the 4-by-4 nilpotent matrix*

$$\begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_4 & a_5 \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_j \in \mathbb{C}$ for $1 \leq j \leq 6$. Then

(1) *$\partial W(A)$ has at most three line segments if p_A is irreducible and at most one line segment if p_A is reducible.*

(2) *$\partial W(A)$ has at most two line segments if at least one of the a_j's is zero.*

Proof. (1) If p_A is irreducible, then our assertion on $\partial W(A)$ follows from [18, Lemma 2.1]. Otherwise, we may consider the following three cases. First, assume that p_A

is the product of two factors $p_A = p_1 p_2$ with p_1 irreducible cubic and p_2 linear. Let C_j be the dual curve of $p_j = 0$, $j = 1, 2$. If the real point (a, b) is a focus of C_1, then the line $x + iy - (a + ib)z = 0$ which passes $[a, b, 1]$ and $[1, i, 0]$ is tangent to C_1. Hence $p_1(1, i, -(a + ib)) = 0$. Since $\det(A - (a + ib)I_4) = p_A(1, i, -(a + ib)) = p_1(1, i, -(a + ib))p_2(1, i, -(a + ib))$, we have that $a + ib$ is an eigenvalue of A. Then $\sigma(A) = \{0\}$ implies $a = b = 0$. This shows that 0 is in C_1. Similarly, by the same argument, we also have 0 in C_2. Since p_2 is linear, we obtain that C_2 is a single point. Thus $C_2 = \{0\}$ and hence $W(A) = C_1^{\wedge}$. Note that if p is a degree-n homogeneous irreducible polynomial, then the number of singular points of the curve $p = 0$ is at most $(n - 1)(n - 2)/2$ (cf. [36, p. 59, Exercises 5]), and every line segment on the boundary of the convex hull of the dual curve of $p = 0$ corresponds to a singular point of the curve $p = 0$ through duality. Applying these to the cubic p_1, we deduce that there can be at most one line segment on $\partial W(A)$. Next, if p_A is the product of an irreducible quadratic factor and a (possibly reducible) quadratic factor, then [14] and $\sigma(A) = \{0\}$ imply that $\partial W(A)$ has no line segment. Finally, if p_A is the product of four linear factors, then similarly, $W(A) = \{0\}$. Hence $\partial W(A)$ has also no line segment.

(2) Since $a_i = 0$ for some i, by [7, Theorem 1], we obtain that $W(A[k])$ is a circular disc centered at the origin for some k, $1 \leq k \leq 4$. Thus our assertion follows from [18, Theorem 3.4]. ∎

An easy consequence of Theorem 4.1 and the preceding two results is the following:

Corollary 4.3. *Let A be a 4-by-4 nilpotent matrix. Suppose that $\partial W(A)$ has two line segments L_1, L_2 with* $\text{dist}(0, L_1) = \text{dist}(0, L_2)$. *Then $\partial W(A)$ has exactly two line* *segments.*

Proof. After applying a suitable affine transformation, we may assume, without loss of generality, that for some θ, $0 < \theta < 2\pi$, both $\partial W(A)$ and $\partial W(e^{-i\theta}A)$ have a line segment on $x = 1/2$. Since A is nilpotent, by Proposition 4.2 (2), we may assume that

$$A = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_4 & a_5 \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_j \neq 0$ for $1 \leq j \leq 6$, and $a_j > 0$ for $j = 1, 4, 6$. We deduce from Lemma 3.2 that $x = 1/2$ is tangent to $\partial W(A[k])$ and $\partial W((e^{-i\theta}A)[k])$ for every k, $1 \leq k \leq 4$. Thus $\det((1/2)I_3 - \text{Re }(A[k])) = 0$ and $\det((1/2)I_3 - \text{Re }((e^{-i\theta}A)[k])) = 0$ for $1 \leq k \leq 4$. A simple computation shows that $\text{Re }(a_1\overline{a_2}a_4) = \text{Re }(e^{-i\theta}a_1\overline{a_2}a_4)$, $\text{Re }(a_1\overline{a_3}a_5) = \text{Re }(e^{-i\theta}a_1\overline{a_3}a_5)$ and $\text{Re }(a_4\overline{a_5}a_6) = \text{Re }(e^{-i\theta}a_4\overline{a_5}a_6)$. Hence, by the first and third equalities, we have $\overline{a_j}(1 - e^{-i\theta}) = -a_j(1 - e^{i\theta})$ for $j = 2, 5$. This and the second equality yield $\overline{a_2}a_5 = a_2\overline{a_5}$ and $\overline{a_3}a_5^2 = a_3\overline{a_5}^2$. Therefore, if $a_2 = |a_2|e^{i\phi}$ for some real ϕ, then $a_5 = \pm|a_5|e^{i\phi}$ and $a_3 = \pm|a_3|e^{2i\phi}$. Finally, let $U = \text{diag}(e^{4i\phi}, e^{3i\phi}, e^{2i\phi}, e^{i\phi})$. Then U is unitary and

$$U^*AU = e^{-i\phi} \begin{bmatrix} 0 & a_1 & |a_2| & \pm|a_3| \\ 0 & 0 & a_4 & \pm|a_5| \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Consequently, our assertion follows from Theorem 4.1. ■

We conclude from Theorem 4.1, Proposition 4.2 and Corollary 4.3 that $l(A) \leq 3$ for a 4-by-4 nilpotent complex matrix A and $l(A) \leq 2$ for a 4-by-4 nilpotent complex matrix A satisfying any one of the following four conditions: (1) A is a real matrix; (2) p_A is reducible; (3) A is unitarily equivalent to the form of Proposition 4.2 and at least one of the a_j's is zero and (4) $\partial W(A)$ has two line segments L_1, L_2 with $\text{dist}(0, L_1) = \text{dist}(0, L_2)$. However, it is still unknown that $l(A) \leq 2$ for all 4-by-4 nilpotent complex matrices A.

For our main theorem, another result which we need is Theorem 3.11 in Chapter 3:

Lemma 3.11. *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* $\text{Re}\,(e^{-i\theta}A)$ *with unit eigenvectors x_1 and x_2 such that* $\text{Im}\,\langle e^{-i\theta}Ax_1, x_1\rangle \neq \text{Im}\,\langle e^{-i\theta}Ax_2, x_2\rangle$.

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. By Proposition 4.2 (1), we may assume that p_A is irreducible and $l(A) \leq 3$. Suppose that $l(A) = 3$. Since the numerical range of the real matrix A is symmetric with respect to the x-axis, we may assume that for some $r > 0$ and θ, $0 < \theta < \pi$, $\partial W(A)$ (resp., $\partial W(e^{\pm i\theta}A)$) has a line segment on $x = r$ (resp., $x = 1/2$). Since A is nilpotent, by Proposition 4.2 (2), we may assume that

$$A = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_4 & a_5 \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_i \in \mathbb{R}$, $a_i \neq 0$ for $1 \leq i \leq 6$ and $a_j > 0$ for $j = 1, 4, 6$. We deduce from Lemma 3.2 that $x = r$ (resp., $x = 1/2$) is tangent to $\partial W(A[k])$ (resp., $\partial W((e^{\pm i\theta}A)[k])$ for every

k, $1 \le k \le 4$. Thus $\det(rI_3 - \text{Re }(A[k])) = 0$ and $\det((1/2)I_3 - \text{Re }((e^{\pm i\theta}A)[k])) = 0$ for $1 \le k \le 4$. A simple computation shows that

$$\begin{aligned} 4r^2 - \frac{a_1a_2a_4}{r} &= 1 - 2a_1a_2a_4\cos\theta, \\ 4r^2 - \frac{a_1a_3a_5}{r} &= 1 - 2a_1a_3a_5\cos\theta, \\ 4r^2 - \frac{a_2a_3a_6}{r} &= 1 - 2a_2a_3a_6\cos\theta, \\ 4r^2 - \frac{a_4a_5a_6}{r} &= 1 - 2a_4a_5a_6\cos\theta. \end{aligned}$$

If $1/r = 2\cos\theta$, then from the first equality above we obtain $\cos^2\theta = 1$. This is impossible since $0 < \theta < \pi$. Hence from the above four equalities, we have $a_1a_2 = a_5a_6$, $a_1a_3 = a_4a_6$, $a_1a_4 = a_3a_6$, $a_1a_5 = a_2a_6$, $a_2a_3 = a_4a_5$ and $a_2a_4 = a_3a_5$. This shows that either $a_4 = a_3$, $a_5 = a_2$, and $a_6 = a_1$, or $a_4 = -a_3$, $a_5 = -a_2$, and $a_6 = -a_1$. Because a_1, $a_6 > 0$, we only need to consider the former case. Since $\partial W(A)$ has a line segment on $x = r$, Lemma 3.11 implies that $\dim\ker(rI_4 - \text{Re }A) \ge 2$. Moreover, for every $[x_1 \ldots x_4]^T \in \ker(rI_4 - \text{Re }A)$, a simple computation shows that either

(a) $\ker(rI_4 - \text{Re }A) \subseteq \{[x_1 \ldots x_4]^T : x_1 + x_2 + x_3 + x_4 = 0\}$, or

(b) $r = (a_1 + a_2 + a_3)/2$.

In case (a), substituting $x_4 = -x_1 - x_2 - x_3$ in the component equalities for $(rI_4 - \text{Re }A)[x_1 \ldots x_4]^T = 0$, we obtain the two equalities $(2r - a_1 + a_2 + a_3)(x_1 + x_2) = 0$ and $(2r + a_1 + a_2 - a_3)(x_2 + x_3) = 0$. Therefore, we need to consider the following four cases:

(1) $2r - a_1 + a_2 + a_3 = 0$ and $x_2 + x_3 = 0$. Substituting them and $x_4 = -x_1 - x_2 - x_3$ into the component equalities for $(rI_4 - \text{Re }A)[x_1 \ldots x_4]^T = 0$ and using $\dim\ker(rI_4 - \text{Re }A) \ge 2$, we obtain $a_1 = a_2$ and $r = -a_3/2$. Since $a_3 = a_4 > 0$, we have $r < 0$, which contradicts our assumption on r.

(2) $2r + a_1 + a_2 - a_3 = 0$ and $x_1 + x_2 = 0$. Similarly, we obtain $a_2 = a_3$ and $r = -a_1/2$. Since $a_1 > 0$, we also have $r < 0$. This is impossible.

(3) $x_1 + x_2 = 0$ and $x_2 + x_3 = 0$. Since $x_4 = -x_1 - x_2 - x_3$ and $\dim \ker(rI_4 - \text{Re } A) \geq 2$, we are led to a contradiction.

(4) $2r - a_1 + a_2 + a_3 = 0$ and $2r + a_1 + a_2 - a_3 = 0$. From above we obtain $a_1 = a_3$ and $a_2 = -2r < 0$. If $U = \text{diag}(1, -1, 1, -1)$, then U is unitary and

$$U^*AU = \begin{bmatrix} 0 & -a_1 & a_2 & -a_1 \\ 0 & 0 & -a_1 & a_2 \\ 0 & 0 & 0 & -a_1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = -\begin{bmatrix} 0 & a_1 & -a_2 & a_1 \\ 0 & 0 & a_1 & -a_2 \\ 0 & 0 & 0 & a_1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Hence we may assume, without loss of generality, that

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

where $a > 0$. Since A is a 4-by-4 upper triangular nilpotent matrix, by [6, p. 147, (4)], the eigenvalues of Re $(e^{-i\theta}A)$ are given by $\mu_{-1}(\theta) = -(a/2) - \sin(\theta/2)$, $\mu_0(\theta) = (a/2) + \cos(\theta/2)$, $\mu_1(\theta) = -(a/2) + \sin(\theta/2)$ and $\mu_2(\theta) = (a/2) - \cos(\theta/2)$. If $0 < a < 1$, then we have $\mu_0(0) = 1 + (a/2) > \mu_k(0)$ for $k = \pm 1$, 2 and $\mu_1(\pi) = 1 - (a/2) > \mu_j(\pi)$ for $j = -1$, 0, 2. Let

$$\mu(\theta) = \max\{\mu_k(\theta) : k = -1, 0, 1, 2\}.$$

Since $\mu_k(0)$ and $\mu_k(\pi)$ are differentiable functions for all k, we have $(\mu)'_+(0) = (\mu_0)'_+(0) = (\mu_0)'_-(0) = (\mu)'_-(0)$ and $(\mu)'_+(\pi) = (\mu_1)'_+(\pi) = (\mu_1)'_-(\pi) = (\mu)'_-(\pi)$.

Hence [6, Theorem 2] implies that $\partial W(A)$ has no vertical line segment. In addition, the fact that A is a real matrix implies that $\partial W(A)$ has at most two line segments. If $a = 1$, then $\partial W(A)$ has exactly one line segment by [6, Theorem 2]. If $a > 1$, then we obtain $\mu_0(\theta) = a/2 + \cos(\theta/2) > \mu_k(\theta)$ for $k = \pm 1, 2$. Similarly, for every θ, $0 \le \theta < \pi$, we have $(\mu)'_+(\theta) = (\mu_0)'_+(\theta) = (\mu_0)'_-(\theta) = (\mu)'_-(\theta)$. Again, [6, Theorem 2] and the fact that A is a real matrix imply that $\partial W(A)$ has at most one line segment. These all contradict our assertion.

In case (b), we substitute $r = (a_1 + a_2 + a_3)/2$ in the component equalities for $(rI_4 - \text{Re } A)[x_1 \dots x_4]^T = 0$. A simple computation shows that

$$(a_2 + a_3)[(a_1 + a_2)x_1 - (a_1 + a_3)x_2 - (a_2 - a_3)x_3] = 0$$

and

$$(a_1 + a_2)[(a_1 - a_2)x_1 - (a_1 + a_3)x_2 + (a_2 + a_3)x_3] = 0.$$

If $a_1 + a_2 \neq 0$ and $a_2 + a_3 \neq 0$, then from above we have $x_1 = x_3$. Substituting $x_1 = x_3$ into the second equality above, we obtain $x_1 = x_2$ since $a_1 > 0$ and $a_3 = a_4 > 0$. Thus substituting $r = (a_1 + a_2 + a_3)/2$ and $x_1 = x_2 = x_3$ into the component equalities for $(rI_4 - \text{Re } A)[x_1 \dots x_4]^T = 0$ would imply that $x_1 = x_2 = x_3 = x_4$, which contradicts $\dim \ker(rI_4 - \text{Re } A) \ge 2$. Therefore, we only need to consider the following two cases:

(5) $a_1 + a_2 = 0$. Then $a_2 = -a_1$ and hence

$$A = \begin{bmatrix} 0 & a_1 & -a_1 & a_3 \\ 0 & 0 & a_3 & -a_1 \\ 0 & 0 & 0 & a_1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_1,\ a_3 > 0$. Since $\partial W(e^{-i\theta}A)$ has a line segment on $x = 1/2$, Lemma 3.11 implies that $\dim \ker((1/2)I_4 - \text{Re } (e^{-i\theta}A)) \ge 2$, where $0 < \theta < \pi$. Thus there is a nonzero vector x of the form $[0\ x_2\ x_3\ x_4]^T$ in $\ker((1/2)I_4 - \text{Re}\ e^{-i\theta}A)$. A simple computation shows that $a_1(x_2 - x_3) = -a_3x_4$, $x_2 = e^{-i\theta}(a_3x_3 - a_1x_4) = 0$, $x_3 = a_3e^{i\theta}x_2 + a_1e^{-i\theta}x_4$ and $a_1e^{i\theta}(x_2 - x_3) = -x_4$. By the first and last equalities, we have $(a_3e^{i\theta} - 1)x_4 = 0$. Since $a_3e^{i\theta} - 1 \neq 0$, we obtain $x_4 = 0$. Hence we have $x_2 = x_3$ from the first equality. Substituting them into the second equality yields $(1 - a_3e^{-i\theta})x_2 = 0$. Thus we also have $x_2 = 0$, which contradicts our assumption on the nonzeroness of x.

(6) $a_2 + a_3 = 0$. Then $a_2 = -a_3$ and hence

$$A = \begin{bmatrix} 0 & a_1 & -a_3 & a_3 \\ 0 & 0 & a_3 & -a_3 \\ 0 & 0 & 0 & a_1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where a_1, $a_3 > 0$. Using the same argument as in the proof of case (5), we are also led to a contradiction. This completes the proof. ■

We next give a necessary and sufficient condition for the boundary of the numerical range of a 4-by-4 nilpotent matrix A to have a pair of parallel line segments.

Theorem 4.4. *Let A be a 4-by-4 nilpotent matrix. Then $\partial W(A)$ has a pair of parallel line segments if and only if A is unitarily equivalent to a matrix of the form*

$$\alpha \begin{bmatrix} 0 & r_1 & ir_2 & -r_3 \\ & 0 & r_3 & -ir_2 \\ & & 0 & r_1 \\ & & & 0 \end{bmatrix},$$

where $\alpha \in \mathbb{C} \setminus \{0\}$, r_1, $r_3 > 0$ and $r_2 \in \mathbb{R}$. In this case, $\partial W(A)$ has no other line segment.

Proof. We first prove the sufficiency. In the assumption, we may assume that $\alpha = 1$ and $r_1^2 + r_2^2 + r_3^2 = 1$. From a simple computation, we obtain, for every μ with $|\mu| = 1$ and for every $z \in \mathbb{C}$,

$$\begin{aligned} p(z,\mu) &\equiv \det(zI_4 - \mathrm{Re}(\mu A)) \\ &= z^4 - (\frac{2r_1^2 + 2r_2^2 + 2r_3^2}{4})z^2 + (\frac{r_1^4 + r_2^4 + r_3^4 + 2r_1^2r_2^2 + 2r_2^2r_3^2 + (\mu^2 + \overline{\mu}^2)r_1^2r_3^2}{16}) \\ &= z^4 - (\frac{1}{2})z^2 + (\frac{1 + (\mu^2 + \overline{\mu}^2 - 2)r_1^2r_3^2}{16}), \end{aligned}$$

and hence

$$p(z,1) = z^4 - \frac{1}{2}z^2 + \frac{1}{16} = (z - \frac{1}{2})^2(z + \frac{1}{2})^2.$$

Thus both $1/2$ and $-1/2$ are the eigenvalues of Re A with multiplicity two. Since r_1, $r_3 > 0$ and $r_1^2 + r_2^2 + r_3^2 = 1$, we have $r_1 < 1$ and $r_3 < 1$. Some calculations on $\ker((1/2)I_4 - \mathrm{Re}\ A)$ show that

$$\ker(\frac{1}{2}I_4 - \mathrm{Re}A) = \{\lambda_1[\frac{t_1}{1-r_1^2}\ \frac{t_2}{1-r_1^2}\ 1\ 0]^T + \lambda_2[-\frac{t_2}{1-r_1^2}\ -\frac{t_1}{1-r_1^2}\ 0\ 1]^T : \lambda_1, \lambda_2 \in \mathbb{C}\},$$

where $t_1 = r_1r_3 + ir_2$ and $t_2 = r_3 + ir_1r_2$. Note that if B is a 4-by-4 nilpotent matrix of the form

$$\begin{bmatrix} 0 & b_1 & b_2 & b_3 \\ 0 & 0 & b_4 & b_5 \\ 0 & 0 & 0 & b_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $b_j \in \mathbb{C}$ for $1 \le j \le 6$, then for every nonzero $r \in \mathbb{R}$ and $x = [x_1\ x_2\ x_3\ x_4]^T \in \ker(rI_4 - \mathrm{Re}\ B)$, we have $2rx_1 = b_1x_2 + b_2x_3 + b_3x_4$ and $-\overline{b_1}x_1 + 2rx_2 = b_4x_3 + b_5x_4$. Thus

$$\begin{aligned} \mathrm{Im}\langle Bx, x\rangle &= \mathrm{Im}(\overline{x_1}(b_1x_2 + b_2x_3 + b_3x_4) + \overline{x_2}(b_4x_3 + b_5x_4) + b_6\overline{x_3}x_4) \\ &= \mathrm{Im}(2r|x_1|^2 - \overline{b_1}x_1\overline{x_2} + 2r|x_2|^2 + b_6\overline{x_3}x_4) \\ &= \mathrm{Im}(b_6\overline{x_3}x_4 - \overline{b_1}x_1\overline{x_2}). \qquad \text{(iii)} \end{aligned}$$

We derive from (iii) that Im $\langle Ax, x\rangle = r_1$ Im $(\overline{x_3}x_4 - x_1\overline{x_2})$. Letting $y = [t_1 - it_2 \;\; t_2 - it_1 \;\; 1 - r_1^2 \;\; (1 - r_1^2)i]^T$ and $z = [t_1 + it_2 \;\; t_2 + it_1 \;\; 1 - r_1^2 \;\; -(1 - r_1^2)i]^T$, we obtain

$$\begin{aligned}
\mathrm{Im}\langle Ay, y\rangle &= r_1\mathrm{Im}(i(1 - r_1^2)^2 - (t_1 - it_2)(\overline{t_2} + i\overline{t_1})) \\
&= r_1\mathrm{Im}(i(1 - r_1^2)^2 - (t_1\overline{t_2} + \overline{t_1}t_2 + i|t_1|^2 - i|t_2|^2)) \\
&= r_1((1 - r_1^2)^2 + |t_2|^2 - |t_1|^2) \\
&= r_1((1 - r_1^2)^2 + r_3^2 + r_1^2r_2^2 - r_1^2r_3^2 - r_2^2) \\
&= r_1((1 - r_1^2)^2 + (1 - r_1^2)(r_3^2 - r_2^2)) \\
&= r_1(1 - r_1^2)(1 - r_1^2 - r_2^2 + r_3^2) \\
&= r_1(1 - r_1^2)(1 + r_3^2 - (1 - r_3^2)) = 2r_1r_3^2(1 - r_1^2) > 0
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Im}\langle Az, z\rangle &= r_1\mathrm{Im}(-i(1 - r_1^2)^2 - (t_1 + it_2)(\overline{t_2} - i\overline{t_1})) \\
&= r_1\mathrm{Im}(-i(1 - r_1^2)^2 - (t_1\overline{t_2} + \overline{t_1}t_2 - i|t_1|^2 + i|t_2|^2)) \\
&= r_1(-(1 - r_1^2)^2 - |t_2|^2 + |t_1|^2) \\
&= -\mathrm{Im}\langle Ay, y\rangle < 0.
\end{aligned}$$

Finally, letting $y_1 = y/\|y\|$ and $z_1 = z/\|z\|$, we have $\mathrm{Im}\langle Ay_1, y_1\rangle \neq \mathrm{Im}\langle Az_1, z_1\rangle$. Therefore, by Lemma 3.11, $\partial W(A)$ has a line segment on $x = 1/2$. In a similar fashion, we also derive that $\partial W(A)$ has a line segment on $x = -1/2$. This proves one direction.

Now we prove the necessity. After applying a suitable affine transformation, we may assume, without loss of generality, that for some $r \in \mathbb{R}$, $\partial W(A)$ has two line segments on $x = 1/2$ and $x = r$. Lemma 3.11 implies that dim $\ker((1/2)I_4 - \mathrm{Re}\ A) = \dim \ker(rI_4 - \mathrm{Re}\ A) = 2$. Thus both $1/2$ and r are the eigenvalues of Re A with multiplicity two. Since Re A has zeros on its diagonal, we have $(1/2) + r =$ tr

(Re A)= 0, where tr (Re A) denotes the sum of the diagonal entries of Re A. Thus $r = -1/2$. Since A is nilpotent, we may assume that

$$A = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_4 & a_5 \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_i \in \mathbb{C}$ for $1 \le i \le 6$ and $a_j \ge 0$ for $j = 1, 4, 6$. Fix $|\mu| = 1$. For every $z \in \mathbb{C}$, some simple computations show that

$$\begin{aligned} p(z,\mu) &\equiv \det(zI_4 - \mathrm{Re}\mu A) \\ &= z^4 - \left(\frac{a_1^2 + |a_2|^2 + |a_3|^2 + a_4^2 + |a_5|^2 + a_6^2}{4}\right)z^2 \\ &\quad - \frac{\mathrm{Re}(\mu(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4))}{4} z \\ &\quad + \frac{a_1^2 a_6^2 + |a_2|^2|a_5|^2 + |a_3|^2 a_4^2 - \beta + (1-\mu^2) a_1 \overline{a_3} a_4 a_6 + (1-\overline{\mu}^2) a_1 a_3 a_4 a_6}{16}, \end{aligned}$$

where $\alpha_1 \equiv a_1\overline{a_2}a_4$, $\alpha_2 \equiv a_1\overline{a_3}a_5$, $\alpha_3 \equiv a_2\overline{a_3}a_6$, $\alpha_4 \equiv a_4\overline{a_5}a_6$ and $\beta = 2\ \mathrm{Re}(a_1\overline{a_2}a_5a_6 + a_2\overline{a_3}a_4a_5 + a_1\overline{a_3}a_4a_6)$. Hence we obtain

$$\begin{aligned} \frac{\partial p}{\partial z}(z,\mu) &= 4z^3 - \left(\frac{a_1^2 + |a_2|^2 + |a_3|^2 + a_4^2 + |a_5|^2 + a_6^2}{2}\right)z \\ &\quad - \frac{\mathrm{Re}(\mu(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4))}{4}. \end{aligned}$$

Thus $p(\pm 1/2, 1) = \frac{\partial p}{\partial z}(\pm 1/2, 1) = 0$. By our assumption, we deduce from Lemma 3.2 that both $x = 1/2$ and $x = -1/2$ are tangent to $\partial W(A[k])$ for every k, $1 \le k \le 4$. Then $\det(\pm 1/2 I_3 - \mathrm{Re}\ A[k]) = 0$ for $1 \le k \le 4$. A simple computation shows that $a_1^2 + |a_2|^2 + a_4^2 = 1 \pm 2\ \mathrm{Re}\ \alpha_1$, $a_1^2 + |a_3|^2 + |a_5|^2 = 1 \pm 2\ \mathrm{Re}\ \alpha_2$, $|a_2|^2 + |a_3|^2 + a_6^2 = 1 \pm 2$ Re α_3 and $a_4^2 + |a_5|^2 + a_6^2 = 1 \pm 2\ \mathrm{Re}\alpha_4$. Hence Re $\alpha_j = 0$ for $1 \le j \le 4$. Moreover, taking the sum of the first three equalities above and subtracting the last equality from it yields $a_1^2 + |a_2|^2 + |a_3|^2 = 1$. From this and the first equality, we have $|a_3| = a_4$.

Therefore, the first three equalities imply that $|a_2| = |a_5|$ and $a_1 = a_6$. Substituting them into the equality $p(1/2, 1) = 0$, we have $1 = a_1^4 + |a_2|^4 + |a_3|^4 - \beta$. Thus

$$\begin{aligned} 1 &= (a_1^2 + |a_2|^2 + |a_3|^2)^2 - 2a_1^2|a_2|^2 - 2|a_2|^2|a_3|^2 - 2a_1^2|a_3|^2 - \beta \\ &= 1 - (|a_1\overline{a_2} + \overline{a_5}a_6|^2 + |a_2\overline{a_3} + a_4\overline{a_5}|^2 + |a_1\overline{a_3} + a_4a_6|^2). \end{aligned}$$

Consequently, we obtain $a_1\overline{a_2} = -\overline{a_5}a_6$, $a_2\overline{a_3} = -a_4\overline{a_5}$ and $a_1\overline{a_3} = -a_4a_6$. Consider the following cases:

(1) $a_1a_4 = 0$. Suppose that $a_1 = 0$ (resp., $a_4 = 0$). Then $a_6 = 0$ (resp., $a_3 = 0$) and hence $W(A)$ is a circular disc centered at the origin by [7, Theorem 1]. This is a contradiction.

(2) $a_1a_4 \neq 0$ and $a_2 = 0$. Hence $a_5 = 0$. From $a_1\overline{a_3} = -a_4a_6$ and $a_1 = a_6$ and $|a_3| = a_4$, we have $a_3 = -a_4$.

(3) $a_1a_2a_4 \neq 0$. Hence $a_j \neq 0$ for all j, $1 \leq j \leq 6$. We also have $a_3 = -a_4 < 0$ by the same argument as in case (2). In addition, $a_1\overline{a_2} = -\overline{a_5}a_6$ (resp., $a_2\overline{a_3} = -a_4\overline{a_5}$) implies that $a_2 = -a_5$ (resp., $a_2 = \overline{a_5}$). Thus $a_2 = -a_5 = -\overline{a_2}$. That is, $a_2 = ir$ for some $r \in \mathbb{R}$ and $r \neq 0$. This proves our assertion.

Finally, that $\partial W(A)$ has no line segment other than those on $x = \pm 1/2$ follows from Corollary 4.3. ■

An immediate consequence of the preceding theorem is the following:

Corollary 4.5. *Let A be a 4-by-4 nilpotent real matrix. Then $\partial W(A)$ has two* *parallel line segments if and only if A is unitarily equivalent to a matrix of the form*

$$\alpha \begin{bmatrix} 0 & r_1 & 0 & -r_3 \\ & 0 & r_3 & 0 \\ & & 0 & r_1 \\ & & & 0 \end{bmatrix},$$

where $\alpha \in \mathbb{R} \setminus \{0\}$, and r_1, $r_3 > 0$.

Proof. The sufficiency follows easily from Theorem 4.4. For the necessity, we may assume, by Theorem 4.4, that

$$A = \alpha \begin{bmatrix} 0 & r_1 & ir_2 & -r_3 \\ & 0 & r_3 & -ir_2 \\ & & 0 & r_1 \\ & & & 0 \end{bmatrix},$$

where $\alpha \in \mathbb{C} \setminus \{0\}$, r_1, $r_3 > 0$ and $r_2 \in \mathbb{R}$. Since A is a real matrix, we obtain $\alpha \in \mathbb{R}$. Thus $i\alpha r_2 = -i\alpha r_2$ and hence $r_2 = 0$. Our assertion follows. ■

Finally, it is shown in the next theorem that a 4-by-4 nilpotent matrix A which has a line segment on $\partial W(A)$ must be unitarily equivalent to a matrix of a special form.

Theorem 4.6. *Let A be a 4-by-4 nilpotent matrix. If $\partial W(A)$ has a line segment, then A is unitarily equivalent to a matrix of the form*

(iv)
$$\alpha \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ & 0 & a_4 & a_5 \\ & & 0 & a_6 \\ & & & 0 \end{bmatrix},$$

where

$$a_4 = \overline{a_1}a_2 + ((1-|a_1|^2)(1-|a_2|^2))^{1/2}e^{i\theta_1},$$
$$a_5 = \overline{a_1}a_3 + ((1-|a_1|^2)(1-|a_3|^2))^{1/2}e^{i\theta_2},$$
$$a_6 = \overline{a_2}a_3 + ((1-|a_2|^2)(1-|a_3|^2))^{1/2}e^{i(\theta_2-\theta_1)},$$

$|a_i| \leq 1$ *for* $1 \leq i \leq 3$, *and* $\alpha \in \mathbb{C} \setminus \{0\}$, θ_1, $\theta_2 \in \mathbb{R}$. (*If* $|a_1| = 1$ *or* $|a_2| = 1$, *then* θ_1 *can be an arbitrary real number. Similarly, if* $|a_1| = 1$ *or* $|a_3| = 1$, *then* θ_2 *can be an arbitrary real number.*)

Proof. Since A is nilpotent, we may assume that

$$A = \begin{bmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & 0 & a_4 & a_5 \\ 0 & 0 & 0 & a_6 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $a_j \in \mathbb{C}$ for all j. After applying a suitable affine transformation, we may assume, without loss of generality, that $\partial W(A)$ has a line segment on $x = -1/2$. Lemma 3.11 implies that $-1/2$ is the minimum eigenvalue of Re A and dim ker$((1/2)I_4+$Re $A) \geq 2$. Hence the fact that Re $\begin{bmatrix} 0 & a_j \\ 0 & 0 \end{bmatrix}$ is a submatrix of Re A yields $W(\begin{bmatrix} 0 & a_j \\ 0 & 0 \end{bmatrix}) \subseteq W(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix})$ for every j. Thus $|a_j| \leq 1$ for $1 \leq j \leq 6$. We deduce from Lemma 3.2 that $x = -1/2$ is tangent to $\partial W(A[k])$ and $\partial W((e^{-i\theta}A)[k])$ for every k, $1 \leq k \leq 4$. Then we have $\det((1/2)I_3 + \text{Re }(A[k])) = 0$ and $\det((1/2)I_3 + \text{Re }((e^{-i\theta}A)[k])) = 0$

for $1 \leq k \leq 4$. A simple computation shows that

$$\begin{aligned}
|a_1|^2 + |a_2|^2 + |a_4|^2 &= 1 + 2\text{Re } a_1\overline{a_2}a_4, \\
|a_1|^2 + |a_3|^2 + |a_5|^2 &= 1 + 2\text{Re } a_1\overline{a_3}a_5, \\
|a_2|^2 + |a_3|^2 + |a_6|^2 &= 1 + 2\text{Re } a_2\overline{a_3}a_6, \\
|a_4|^2 + |a_5|^2 + |a_6|^2 &= 1 + 2\text{Re } a_4\overline{a_5}a_6.
\end{aligned}$$

Let $r_1 = \sqrt{1-|a_1|^2}$, $r_2 = \sqrt{1-|a_2|^2}$ and $r_3 = \sqrt{1-|a_3|^2}$. From the first three equalities above, we obtain

$$\begin{aligned}
|a_4 - \overline{a_1}a_2|^2 &= (1-|a_1|^2)(1-|a_2|^2) = r_1^2 r_2^2, \\
|a_5 - \overline{a_1}a_3|^2 &= (1-|a_1|^2)(1-|a_3|^2) = r_1^2 r_3^2, \\
|a_6 - \overline{a_2}a_3|^2 &= (1-|a_2|^2)(1-|a_3|^2) = r_2^2 r_3^2.
\end{aligned}$$

Thus we may assume that $a_4 = \overline{a_1}a_2 + r_1 r_2 e^{i\theta_1}$, $a_5 = \overline{a_1}a_3 + r_1 r_3 e^{i\theta_2}$ and $a_6 = \overline{a_2}a_3 + r_2 r_3 e^{i\theta_3}$, where $\theta_j \in \mathbb{R}$ for $1 \leq j \leq 3$. By means of a succession of elementary row operations, we can transform $(1/2)I_4$+Re A into the matrix

$$D = \begin{bmatrix} 1 & a_1 & a_2 & a_3 \\ 0 & 1-|a_1|^2 & a_4 - \overline{a_1}a_2 & a_5 - \overline{a_1}a_3 \\ 0 & \overline{a_4} - a_1\overline{a_2} & 1-|a_2|^2 & a_6 - \overline{a_2}a_3 \\ 0 & \overline{a_5} - a_1\overline{a_3} & \overline{a_6} - a_2\overline{a_3} & 1-|a_3|^2 \end{bmatrix} = \begin{bmatrix} 1 & a_1 & a_2 & a_3 \\ 0 & r_1^2 & r_1 r_2 e^{i\theta_1} & r_1 r_3 e^{i\theta_2} \\ 0 & r_1 r_2 e^{-i\theta_1} & r_2^2 & r_2 r_3 e^{i\theta_3} \\ 0 & r_1 r_3 e^{-i\theta_2} & r_2 r_3 e^{-i\theta_3} & r_3^2 \end{bmatrix}.$$

Since dim ker($(1/2)I_4$+Re A) ≥ 2, we have dim ran($(1/2)I_4$+Re A) $= 1$ or 2. If dim ran($(1/2)I_4$+Re A) $= 1$, then dim ran$D = 1$. Hence we have $r_1 = r_2 = r_3 = 0$. This implies that $a_4 = \overline{a_1}a_2$, $a_5 = \overline{a_1}a_3$ and $a_6 = \overline{a_2}a_3$. On the other hand, if dim ran($(1/2)I_4$+Re A) $= 2$, then dim ran$D = 2$. Thus we obtain $r_2 r_3 e^{i\theta_3} = r_2 r_3 e^{i(\theta_2 - \theta_1)}$. This completes the proof. ∎

For the converse of Theorem 4.6, the following example shows that this is not true in general.

Example 4.7. *Let*

$$A = \begin{bmatrix} 0 & (\sqrt{2+\sqrt{3}})/2 & 1/2 & \sqrt{2}/2 \\ 0 & 0 & (\sqrt{2+\sqrt{3}}+\sqrt{6-3\sqrt{3}})/4 & (\sqrt{4+2\sqrt{3}}+\sqrt{4-2\sqrt{3}})/4 \\ 0 & 0 & 0 & (\sqrt{2}+\sqrt{6})/4 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then A is of the form (iv) *and $\partial W(A)$ has no line segment.*

Proof. It is easily seen that A is of the form (iv) by letting $\alpha = 1$ and $\theta_1 = \theta_2 = 0$. Now suppose that for some $r > 0$ and θ, $0 \leq \theta < 2\pi$, $\partial W(e^{-i\theta}A)$ has line segment on $x = r$. We deduce from Lemma 3.2 that $x = 1/2$ is tangent to $\partial W((e^{-i\theta}(A/(2r)))[k])$ for every k, $1 \leq k \leq 4$. Thus $\det((1/2)I_3 - \mathrm{Re}\,((e^{-i\theta}(A/(2r)))[k]) = 0$ for $1 \leq k \leq 4$. A simple computation shows that

$$\begin{aligned} 1 - (\frac{1}{2r})^2(\frac{5+\sqrt{3}}{4}) &= 2(\frac{1}{2r})^3(\frac{1+\sqrt{3}}{8})\cos\theta, \\ 1 - (\frac{1}{2r})^2(\frac{7+\sqrt{3}}{4}) &= 2(\frac{1}{2r})^3(\frac{3+\sqrt{3}}{8})\cos\theta. \end{aligned}$$

From the above two equalities, we obtain $\cos\theta \neq 0$ and hence

$$\frac{1 - (\frac{1}{2r})^2(\frac{5+\sqrt{3}}{4})}{1 - (\frac{1}{2r})^2(\frac{7+\sqrt{3}}{4})} = \frac{1+\sqrt{3}}{3+\sqrt{3}}.$$

This shows that $r = 1/2$. Substituting $r = 1/2$ in the first equality, we have $\cos\theta = -1$. We next show that $\partial W(A)$ has no line segment on $x = -1/2$. By means of a succession of elementary row operations, we can transform $(-1/2)I_4 - \mathrm{Re}\,A$ into the matrix

$$E = \begin{bmatrix} 1 & (\sqrt{2+\sqrt{3}})/2 & 1/2 & \sqrt{2}/2 \\ 0 & \sqrt{2-\sqrt{3}} & \sqrt{3} & \sqrt{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus some calculations on $\ker((-1/2)I_4 - \text{Re } A)$ show that

$$\ker(-\frac{1}{2}I_4 - \text{Re}A) = \{\lambda_1[t_1 \ \ -\sqrt{3}t_2 \ 1 \ 0]^T + \lambda_2[\frac{\sqrt{2}}{2}t_1 \ \ -\sqrt{2}t_2 \ 0 \ 1]^T : \lambda_1, \lambda_2 \in \mathbb{C}\},$$

where $t_1 = 1 + \sqrt{3}$ and $t_2 = \sqrt{2+\sqrt{3}}$. Hence by (iii) in the proof of Theorem 4.4, we derive that for any $x = [x_1 \ x_2 \ x_3 \ x_4]^T \in \ker((-1/2)I_4 - \text{Re } A)$,

$$\begin{aligned}
\text{Im}\langle Ax, x\rangle &= \text{Im}(\frac{\sqrt{2}+\sqrt{6}}{4}\overline{x_3}x_4 - \frac{\sqrt{2+\sqrt{3}}}{2}x_1\overline{x_2}) \\
&= \text{Im}(\frac{\sqrt{2}}{4}t_1\overline{x_3}x_4 - \frac{1}{2}t_2(t_1x_3 + \frac{\sqrt{2}}{2}t_1x_4)(-\sqrt{3}t_2\overline{x_3} - \sqrt{2}t_2\overline{x_4})) \\
&= \text{Im}((\frac{\sqrt{2}}{4}t_1 + \frac{\sqrt{6}}{4}t_1t_2^2)\overline{x_3}x_4 + \frac{\sqrt{2}}{2}t_1t_2^2x_3\overline{x_4}) \\
&= \text{Im}(\frac{\sqrt{2}}{4}(1+\sqrt{3})(1+\sqrt{3}(2+\sqrt{3}))\overline{x_3}x_4 + \frac{\sqrt{2}}{2}(1+\sqrt{3})(2+\sqrt{3})x_3\overline{x_4}) \\
&= \frac{\sqrt{2}}{2}(1+\sqrt{3})(2+\sqrt{3})\text{Im}(\overline{x_3}x_4 + x_3\overline{x_4}) = 0.
\end{aligned}$$

Thus by Lemma 3.11, $\partial W(A)$ has no line segment on $x = -1/2$. Therefore, it is impossible for $\partial W(A)$ to have a line segment.

Chapter 5 Products of two 2-by-2 nonnegative contractions

5.1 Introduction

Recall that a bounded linear operator A on a complex Hilbert space H is nonnegative (resp., positive) if $\langle Ah, h\rangle \geq 0$ for any h in H (resp., $\langle Ah, h\rangle > 0$ for any $h \neq 0$ in H). For convenience, we denote it by $A \geq 0$ (resp., $A > 0$). We know that $A \geq 0$ (resp., $A > 0$) if and only if $A = A^*$ and $\sigma(A) \geq 0$ (resp., $\sigma(A) > 0$). It is well-known that if A and B satisfy $0 \leq A$, $B \leq I$, then Re $AB \geq -1/8$ and $-1/4 \leq$ Im $AB \leq 1/4$ (cf. [12, Theorem 1.1 and Corollary 4.3]). The purpose of this chapter is to study matrices which are products of 2-by-2 nonnegative contractions. We also give several properties of a matrix which is a product of n-by-n ones and discuss those of its numerical range.

In Section 5.2, we consider a matrix which is a product of two 2-by-2 nonnegative contractions. Let A be a 2-by-2 matrix in upper-triangular form

$$\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}.$$

In Theorem 5.1 below, we give a necessary and sufficient condition for A to be a product of two nonnegative contractions. More specifically, it is shown that this is the case if and only if

$$0 \leq x, y \leq 1 \text{ and } |z| \leq |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}.$$

Next, we list several propositions and corollaries of Theorem 5.1. For example, Corollary 5.6 says that $W(A)$ is not a circular disc if a 2-by-2 matrix A is a product of two nonnegative contractions. It was proven in Corollary 5.12 that if A is a 2-by-2 matrix which is a product of finitely many nonnegative contractions, then

$$\|A\| = 1 \text{ if and only if } A \cong \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}, \text{ where } 0 \le x \le 1.$$

In Corollary 5.13 (resp., Corollary 5.14), we give another proof of the fact that if B and C satisfy $0 \le B, C \le I_2$, then $(-1/4)I_2 \le \text{Im } BC \le (1/4)I_2$ (resp., Re $BC \ge (-1/8)I_2$) and

$$\{w \in \mathbb{C} : \text{Im } w = \frac{i}{4}\} \cap \partial W(BC) \neq \varnothing \text{ if and only if } BC \cong \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$$

(resp.,

$$-\frac{1}{8} \in \partial W(BC) \text{ if and only if } BC \cong \begin{bmatrix} 0 & \frac{\sqrt{3}}{4} \\ 0 & \frac{1}{4} \end{bmatrix}).$$

Finally, in Corollary 5.15 (resp., Corollary 5.17), we give some equivalent conditions for a 2-by-2 matrix to be the product of two nonnegative contractions at least one of which is noninvertible (resp., invertible).

In Section 5.3, Theorem 5.18 says that if a bounded linear operator A of the form $\begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix}$ is a product of two nonnegative contractions, then so are A_1 and A_2. We also have several immediate corollaries (cf. Corollaries 5.21, 5.22, 5.23 and 5.24 below). For example, in Corollary 5.22, it is shown that a finite matrix $A = (\sum_{i=1}^{k_1} \oplus A_i) \oplus \text{diag}(w_1, \ldots, w_{k_2})$, where $A_i = \begin{bmatrix} x_i & z_i \\ 0 & y_i \end{bmatrix}$ for all i, is a product of two nonnegative contractions if and only if

$$0 \le x_i, y_i, w_j \le 1 \text{ and } |z_i| \le |\sqrt{x_i} - \sqrt{y_i}|\sqrt{(1-x_i)(1-y_i)} \text{ for all } i, j.$$

In Corollary 5.23, it is shown that an n-by-n quadratic operator is a product of two nonnegative contractions if and only if it is unitarily equivalent to a matrix of the form

$$aI_1 \oplus bI_2 \oplus (\sum_{i=1}^{k} \oplus \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix}),$$

where $0 \leq a,\ b \leq 1$ and $|c_i| \leq |\sqrt{a} - \sqrt{b}|\sqrt{(1-a)(1-b)}$ for all i. In addition, we generalize Corollary 5.6 in Corollary 5.26. More specifically, in Theorem 5.25, it is shown that if an n-by-n matrix A whose numerical range is a circular disc centered at a, then a is an eigenvalue of A with its geometric multiplicity less than its algebraic multiplicity. Hence from Theorem 5.25 we obtain Corollary 5.26, which says that $W(A)$ is not a circular disc if A is a product of two n-by-n nonnegative contractions. Finally, from Kuo and Wu [29, Theorem 3.1], we know that an n-by-n matrix A is a product of finitely many orthogonal projections if and only if it is unitarily equivalent to $I_k \oplus A_1$, where $0 \leq k \leq n$ and A_1 is singular with $\|A_1\| < 1$. Therefore, an n-by-n matrix A is a product of finitely many orthogonal projections if and only if it is unitarily equivalent to $I_k \oplus A_1$, where $0 \leq k \leq n$ and A_1 is a product of finitely many orthogonal projections with $\|A_1\| < 1$. In Proposition 5.30, we also show that A is a product of infinitely (resp., finitely) many nonnegative contractions if and only if $A \cong I_k \oplus B$, where $0 \leq k \leq n$ and B is a product of infinitely (resp., finitely) many nonnegative contractions with $\|B\| < 1$.

5.2 Products of two 2-by-2 nonnegative contractions

The following theorem is the main result of this section.

Theorem 5.1. *Let A be a 2-by-2 matrix of the form* $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$. *Then A is a product of two nonnegative contractions B and C if and only if $0 \leq x,\, y \leq 1$ and $|z| \leq |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}$. In this case, we can choose B and C such that $\|B\| = \|C\| = 1$.*

In order to prove this theorem, we need the following lemmas. The proof of Lemma 5.2 can be obtained by [4, Theorem 3] and Lemma 5.3 is a common result in linear algebra.

Lemma 5.2. *If $B \geq 0$ and A is any bounded linear operator, then $\sigma(AB)^\wedge \subseteq \overline{W(A)} \cdot \overline{W(B)}$.*

Lemma 5.3. *Let A and B be two 2-by-2 matrices. Then A and B are unitarily equivalent if and only if* $\mathrm{tr}\ A = \mathrm{tr}\ B$, $\det A = \det B$ and $\mathrm{tr}\ (A^*A) = \mathrm{tr}\ (B^*B)$. *In particular, normal A and B are unitarily equivalent if and only if* $\mathrm{tr}\ A = \mathrm{tr}\ B$ *and* $\det A = \det B$.

The next lemma is the major step in proving Theorem 5.1.

Lemma 5.4. *Let A be a 2-by-2 matrix of the form* $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$. *Then*

(1) *$A = BC$, where B (resp., C) is unitarily equivalent to* $\mathrm{diag}(1, \lambda_1)$ (*resp.*, $\mathrm{diag}(1, \lambda_2)$), *where $0 \leq \lambda_1 \leq 1$ (resp., $0 \leq \lambda_2 \leq 1$), if and only if $xy = \lambda_1\lambda_2$, where $0 \leq x,\, y \leq$*

1, $\min\{x,y\} \le \lambda_1, \lambda_2 \le \max\{x,y\}$ *and* $|z|^2 = (1-x)(1-y)[(x+y)-(\lambda_1+\lambda_2)]$.

(2) *A is a product of two nonnegative contractions which are both unitarily equivalent to* $\mathrm{diag}(1,\lambda)$ *for some* λ, $0 \le \lambda \le 1$, *if and only if* $0 \le x$, $y \le 1$ *and* $|z| = |\sqrt{x}-\sqrt{y}|\sqrt{(1-x)(1-y)}$. *In this case,* $\lambda = \sqrt{xy}$.

(3) *A is a product of two nonnegative contractions which are both unitarily equivalent to* $\mathrm{diag}(\lambda_1,\lambda_2)$ *for some* λ_1 *and* λ_2, $0 \le \lambda_1$, $\lambda_2 \le 1$, *if and only if* $0 \le x$, $y \le 1$ *and* $|z| \le |\sqrt{x}-\sqrt{y}|\sqrt{(1-x)(1-y)}$.

Proof. (1) We first prove the sufficiency. By assumption, if $y=1$, then $x=\lambda_1\lambda_2$ and $z=0$. It is easy to see that our assertion holds in this case. Hence we may assume, without loss of generality, that $xy=\lambda_1\lambda_2$, where $0 \le x \le \lambda_1 \le \lambda_2 \le y < 1$ and $|z|^2 = (1-x)(1-y)[(x+y)-(\lambda_1+\lambda_2)]$. Let $B = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_1 \end{bmatrix}$, $C = \begin{bmatrix} a & b \\ \bar{b} & d \end{bmatrix}$ and $A_1 = BC = \begin{bmatrix} a & b \\ \bar{b}\lambda_1 & d\lambda_1 \end{bmatrix}$, where $a = 1-c$, $d = \lambda_2 + c$, $c = (1-x)(1-y)/(1-\lambda_1)$ and $|b| = \sqrt{ad-\lambda_2}$. Then we obtain that $a+d = 1+\lambda_2$ and $ad-|b|^2 = \lambda_2$. Then Lemma 5.3 implies that $C \cong \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. Hence we only need to prove that $A \cong A_1$. Some simple computations show that

$$\begin{aligned} \mathrm{tr}\ A_1 &= a + d\lambda_1 = 1 - c + \lambda_1(\lambda_2 + c) = 1 + \lambda_1\lambda_2 - c(1-\lambda_1) \\ &= 1 + xy - (1-x)(1-y) = x + y = \mathrm{tr}\ A \end{aligned}$$

and

$$\det A_1 = ad\lambda_1 - |b|^2\lambda_1 = \lambda_1(ad - |b|^2) = \lambda_1\lambda_2 = xy = \det A.$$

In addition,

$$\begin{aligned}
|z|^2 &= (1-x)(1-y)[(x+y)-(\lambda_1+\lambda_2)] \\
&= (1-x)(1-y)[(1-\lambda_1)(1-\lambda_2)-(1-x)(1-y)] \\
&= c(1-\lambda_1)[(1-\lambda_1)(1-\lambda_2)-c(1-\lambda_1)] \\
&= (1-\lambda_1)^2[c-c\lambda_2-c^2] \\
&= (1-\lambda_1)^2[(\lambda_2+c)(1-c)-\lambda_2] \\
&= (1-\lambda_1)^2(ad-\lambda_2) = |b|^2(1-\lambda_1)^2.
\end{aligned}$$

This implies that

$$\begin{aligned}
\mathrm{tr}(A_1^*A_1) &= a^2+|b|^2+|b|^2\lambda_1^2+d^2\lambda_1^2 \\
&= (a+d\lambda_1)^2+|b|^2(1-\lambda_1)^2-2\lambda_1(ad-|b|^2) \\
&= 1+\lambda_1\lambda_2-c(1-\lambda_1) \\
&= (x+y)^2-2xy+|z|^2 = \mathrm{tr}(A^*A).
\end{aligned}$$

From these equalities and Lemma 5.3, we obtain $A_1 \cong A$ as asserted. This proves the sufficiency.

For the converse, Lemma 5.2 implies that $\sigma(A) \subseteq W(B)W(C) \subseteq [0,\ 1]$ and hence $0 \le x,\ y \le 1$. If $\lambda_1 = 1$, then $A \cong C \cong \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$. This implies that $z = 0$ and $\max\{x, y\} = 1$. It is easy to see that our assertion holds for the case $\max\{x, y\} = 1$. Hence we may assume that $0 \le x \le y < 1$ and $\lambda_1 < 1$. By our assumptions, we may further assume that $B = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_1 \end{bmatrix}$, $C = \begin{bmatrix} a & b \\ \bar{b} & d \end{bmatrix}$ with $C \cong \begin{bmatrix} 1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$ and $A \cong BC = \begin{bmatrix} a & b \\ \bar{b}\lambda_1 & d\lambda_1 \end{bmatrix}$. Hence Lemma 5.3 implies that $a+d = 1+\lambda_2$, $ad-|b|^2 = \lambda_2$, and $x+y = a+d\lambda_1$, $xy = (ad-|b|^2)\lambda_1$, $x^2+y^2+|z|^2 = a^2+|b|^2+(|b|^2+d^2)\lambda_1^2$.

Thus we have

$$|z|^2 = [a^2 + |b|^2 + (|b|^2 + d^2)\lambda_1^2] - (a + d\lambda_1)^2 + 2(ad - |b|^2)\lambda_1 = |b|^2(1 - \lambda_1)^2.$$

In addition, we also obtain that

$$x + y = a + d\lambda_1 = (1 + \lambda_2 - d) + d\lambda_1 = 1 + \lambda_2 - (1 - \lambda_1)d$$

and hence

$$\begin{aligned} d &= \frac{1}{1-\lambda_1}(1 + \lambda_2 - x - y) \\ &= \frac{1}{1-\lambda_1}[(1-x)(1-y) - xy + \lambda_2] \\ &= \frac{1}{1-\lambda_1}[\lambda_2(1-\lambda_1) + (1-x)(1-y)]. \end{aligned}$$

Let $c = (1-x)(1-y)/(1-\lambda_1)$. Then $d = \lambda_2 + c$ and $a = 1 - c$. This implies that

$$\begin{aligned} |z|^2 &= |b|^2(1-\lambda_1)^2 = (1-\lambda_1)^2(ad - \lambda_2) \\ &= (1-\lambda_1)^2[(1-c)(\lambda_2 + c) - \lambda_2] \\ &= c(1-\lambda_1)^2(1 - \lambda_2 - c) \\ &= c(1-\lambda_2)[(1-\lambda_1)(1-\lambda_2) - c(1-\lambda_1)] \\ &= (1-x)(1-y)[(x+y) - (\lambda_1 + \lambda_2)]. \end{aligned}$$

Hence $x + y \geq \lambda_1 + \lambda_2$. By the equality $xy = (ad - |b|^2)\lambda_1 = \lambda_1\lambda_2$, we derive that $x^2 + \lambda_1\lambda_2 = x^2 + xy \geq x\lambda_1 + x\lambda_2$. That is, $(\lambda_1 - x)(\lambda_2 - x) \geq 0$. Therefore, $x \leq \lambda_1, \lambda_2 \leq y$ follows. This proves the necessity.

(2) By (1), the sufficiency follows easily from letting $\lambda_1 = \lambda_2 = \sqrt{xy}$. For the necessity, let $\lambda_1 = \lambda_2 = \lambda$. Then (1) implies that $\lambda = \sqrt{xy}$ and hence $|z|^2 = (1-x)(1-y)[(x+y) - 2\lambda] = (1-x)(1-y)(\sqrt{x} - \sqrt{y})^2$. Our assertion follows.

(3) Assume that $A = BC$, where B, C are both unitarily equivalent to $\text{diag}(\lambda_1, \lambda_2)$ for some λ_1 and nonzero λ_2, $0 \le \lambda_1 \le \lambda_2 \le 1$. Considering $A/(\lambda_2{}^2) = (B/\lambda_2)(C/\lambda_2)$, we deduce from (2) that $0 \le x/\lambda_2,\ y/\lambda_2 \le 1$ and

$$|\frac{z}{\lambda_2}| = |\sqrt{\frac{x}{\lambda_2}} - \sqrt{\frac{y}{\lambda_2}}|\sqrt{(1-\frac{x}{\lambda_2})(1-\frac{y}{\lambda_2})}.$$

Hence $0 \le x,\ y \le 1$ and

$$|z| = |\sqrt{x} - \sqrt{y}|\sqrt{(\lambda_2 - x)(1-\frac{y}{\lambda_2})} \le |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}.$$

This proves the sufficiency. For the necessity, we may assume that $0 \le x \le y \le 1$. Consider the function $f(t) = |\sqrt{x} - \sqrt{y}|\sqrt{(t-x)(1-(y/t))}$ on $[y,\ 1]$. We have $f(y) = 0$ and $f(1) = |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}$. Hence the real-valued continuous function f must assume the value z at some point α, $y \le \alpha \le 1$. This says that $0 \le x/\alpha \le y/\alpha \le 1$ and $|z/\alpha| = |\sqrt{x/\alpha} - \sqrt{y/\alpha}|\sqrt{(1-(x/\alpha))(1-(y/\alpha))}$. Therefore, by (2), $A = \alpha(A/\alpha) = \alpha BC = (\sqrt{\alpha}B)(\sqrt{\alpha}C)$ for some nonnegative contractions $\sqrt{\alpha}B$ and $\sqrt{\alpha}C$ which are both unitarily equivalent to $\text{diag}(\sqrt{\alpha}, \lambda\sqrt{\alpha})$ for some λ, $0 \le \lambda \le 1$. This completes the proof. ■

Now, we are ready to prove Theorem 5.1 by using the preceding lemmas.

Proof of Theorem 5.1. The sufficiency follows easily from Lemma 5.4 (3). To prove the necessity, Lemma 5.2 implies that $0 \le x,\ y \le 1$. First, we consider $\|B\| = \|C\| = 1$. Then B (resp., C) is unitarily equivalent to $\text{diag}(1, \lambda_1)$ (resp., $\text{diag}(1, \lambda_2)$) for some λ_1 (resp., λ_2), $0 \le \lambda_1 \le 1$ (resp., $0 \le \lambda_2 \le 1$). By Lemma 5.4 (1), we obtain that $xy = \lambda_1\lambda_2$, where $0 \le x, y \le 1$, $\min\{x, y\} \le \lambda_1, \lambda_2 \le \max\{x, y\}$ and $|z|^2 = (1-x)(1-y)[(x+y) - (\lambda_1 + \lambda_2)]$. Thus we have the inequality $\lambda_1 + \lambda_2 \ge 2\sqrt{\lambda_1\lambda_2} = 2\sqrt{xy}$. This implies that $|z| \le \sqrt{(1-x)(1-y)}\sqrt{(x+y) - 2\sqrt{xy}} = \sqrt{(1-x)(1-y)}|\sqrt{x} - \sqrt{y}|$. In general, since $A = \alpha \begin{bmatrix} x/\alpha & z/\alpha \\ 0 & y/\alpha \end{bmatrix} = \alpha(B/\|B\|)(C/\|C\|)$, where $0 < \alpha =$

$\|B\|\|C\| \leq 1$, the scalars x, y, z in the above can be replaced by $x/\alpha, y/\alpha, z/\alpha$, respectively, to get $0 \leq x/\alpha, y/\alpha \leq 1$ and $|z/\alpha| \leq \sqrt{(1-(x/\alpha))(1-(y/\alpha))}|\sqrt{x/\alpha}-\sqrt{y/\alpha}|$. This shows that $0 \leq x,\ y \leq \alpha \leq 1$ and $|z| \leq |\sqrt{x}-\sqrt{y}|\sqrt{(\alpha-x)(1-(y/\alpha))} \leq |\sqrt{x}-\sqrt{y}|\sqrt{(1-x)(1-y)}$, which proves our assertion. ■

Next, we list below some corollaries of Theorem 5.1.

Corollary 5.5. *Let a 2-by-2 matrix A of the form* $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$ *be a product of two nonnegative contractions. Then any* $A' = \begin{bmatrix} x & z' \\ 0 & y \end{bmatrix}$ *with* $|z'| \leq |z|$ *is also a product of two nonnegative contractions.*

Proof. This follows directly from Theorem 5.1. ■

Corollary 5.6. *If a 2-by-2 matrix A is a product of two nonnegative contractions, then $W(A)$ is not a circular disc.*

Proof. We may assume that A is of the form $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$ and $W(A)$ is a circular disc. It implies that $x = y$, and hence $z = 0$ by Theorem 5.1. This yields that $W(A) = \{x\}$, contradicting our assumption. Our assertion follows. ■

Corollary 5.7. *A 2-by-2 matrix A is a product of two nonnegative contractions if and only if A is a product of two nonnegative contractions which are both unitarily equivalent to* $\mathrm{diag}(\lambda_1, \lambda_2)$ *for some* λ_1 *and* λ_2, $0 \leq \lambda_1,\ \lambda_2 \leq 1$.

Proof. This follows directly from Theorem 5.1 and Lemma 5.4 (3). ■

Corollary 5.8. *Let A be a 2-by-2 matrix. Then $A = BC$, where B, C are both orthogonal projections ($B = B^* = B^2$, $C = C^* = C^2$) if and only if either $A = I_2$ or A is unitarily equivalent to a matrix of the form* $\begin{bmatrix} 0 & \sqrt{y(1-y)} \\ 0 & y \end{bmatrix}$ *for some y,* $0 \leq y \leq 1$.

Proof. We may assume that A is a 2-by-2 matrix of the form $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$. Since a 2-by-2 nonzero, nonidentity orthogonal projection P is unitarily equivalent to a matrix of the form $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, by Lemma 5.4 (2) we know that A is a product of two orthogonal projections if and only if either $A = I_2$ or $x = 0$, $0 \leq y \leq 1$ and $|z| = \sqrt{y(1-y)}$. ■

The next corollary which is analogous to Lemma 5.4 (2) and (3) is needed for the proof of Corollary 5.10.

Corollary 5.9. *Let A be a 2-by-2 matrix of the form* $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$. *Then*

(1) *A is a product of two nonnegative matrices which are both unitarily equivalent to* $\mathrm{diag}(1, \lambda)$ *for some λ, $\lambda \geq 1$, if and only if x, $y \geq 1$ and $|z| = |\sqrt{x} - \sqrt{y}|\sqrt{(x-1)(y-1)}$. In this case, $\lambda = \sqrt{xy}$.*

(2) *A is a product of two nonnegative matrices which are both unitarily equivalent to* $\mathrm{diag}(\lambda_1, \lambda_2)$ *for some λ_1 and λ_2, λ_1, $\lambda_2 \geq 1$, if and only if x, $y \geq 1$ and $|z| \leq |\sqrt{x} - \sqrt{y}|\sqrt{(x-1)(y-1)}$.*

Proof. Note that $A = BC$, where B, C are both unitarily equivalent to $\mathrm{diag}(1, \lambda)$ for some λ, $\lambda \geq 1$, if and only if $A^{-1} = \begin{bmatrix} 1/x & -z/xy \\ 0 & 1/y \end{bmatrix} = C^{-1}B^{-1}$, where B, C are both unitarily equivalent to $\mathrm{diag}(1, 1/\lambda)$ for some λ, $0 < 1/\lambda \leq 1$. In addition, x, $y \geq 1$ and $|z| = |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}$ (resp., $|z| \leq |\sqrt{x} - \sqrt{y}|\sqrt{(1-x)(1-y)}$) if and only if $0 < 1/x$, $1/y \leq 1$ and $|-z/xy| = |\sqrt{1/x} - \sqrt{1/y}|\sqrt{(1-(1/x))(1-(1/y))}$ (resp., $|-z/xy| \leq |\sqrt{1/x} - \sqrt{1/y}|\sqrt{(1-(1/x))(1-(1/y))}$). Hence our assertions in (1) and (2) follow from Lemma 5.4 (2) and (3), respectively. ■

Corollary 5.10. *Let A be a 2-by-2 matrix. Then A is a product of two positive contractions if and only if A^{-1} is unitarily equivalent to a matrix of the form* $\begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$, *where a, $b \geq 1$ and $|c| \leq |\sqrt{a} - \sqrt{b}|\sqrt{(a-1)(b-1)}$.*

Proof. This follows directly from Theorem 5.1 and the proof of Corollary 5.9. ■

Corollary 5.11. *Let A be a 2-by-2 matrix, which is a product of two nonnegative contractions. Then*

$$\|A\| = 1 \;\; if \; and \; only \; if \;\; A \cong \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}, \; where \; 0 \leq x \leq 1.$$

Proof. We need only prove the necessity. Let A be a 2-by-2 matrix of the form $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$. Since the assumption and the fact that $\|A\| = 1$ imply $|z| = \sqrt{(1-|x|^2)(1-|y|^2)}$, by Theorem 5.1 we obtain $0 \leq x$, $y \leq 1$ and

$$(1-x^2)(1-y^2) \leq (1-x)(1-y)(\sqrt{x} - \sqrt{y})^2.$$

A simple computation shows that either one of x and y is 1, or $(1 + \sqrt{xy})^2 \leq 0$. Note that the latter is a contradiction. The former implies $z = 0$ and our assertion follows.

■

Corollary 5.12. *Let the* 2*-by-*2 *matrix* A *be a product of finitely many nonnegative contractions. Then*

$$\|A\| = 1 \ \textit{ if and only if } \ A \cong \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}, \ \textit{where } 0 \leq x \leq 1.$$

Proof. We need only prove the necessity. Assume that $A = A_1 \cdots A_n$, where $0 \leq A_i \leq I_2$ for all i, $1 \leq i \leq n$. Since $\|A\| = 1$, we have that $\|A_1 A_2\| = 1$. Corollary 5.11 implies that $A_1 A_2 \cong \begin{bmatrix} x_1 & 0 \\ 0 & 1 \end{bmatrix}$, where $0 \leq x_1 \leq 1$ and hence $0 \leq A_1 A_2 \leq I_2$. Repeating the above arguments with A_1 replacing $A_1 A_2$ and so forth, we obtain $A \cong \begin{bmatrix} x & 0 \\ 0 & 1 \end{bmatrix}$, where $0 \leq x \leq 1$. This proves our assertion. ■

Recall that if A and B satisfy $0 \leq A, B \leq I_n$, then it is known that Re $AB \geq (-1/8)I_n$ and $(-1/4)I_n \leq \text{Im } AB \leq (1/4)I_n$ (cf. [12]). Now, we give another proof for the case $n = 2$ in the following.

Corollary 5.13. *Let the* 2*-by-*2 *matrix* A *be a product of two nonnegative contractions. Then* $(-1/4)I_2 \leq \text{Im } A \leq (1/4)I_2$, *and*

(v) $$\{w \in \mathbb{C} : \text{Im } w = \frac{i}{4}\} \cap \partial W(A) \neq \varnothing \ \textit{ if and only if } \ A \cong \begin{bmatrix} 0 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}.$$

In case (v), $\{w \in \mathbb{C} : \text{Im } w = -i/4\} \cap \partial W(A) \neq \varnothing$.

Proof. By Theorem 5.1, we may assume that A is a 2-by-2 matrix of the form $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$, where $0 \leq x \leq y \leq 1$ and $0 \leq z \leq (\sqrt{y} - \sqrt{x})\sqrt{(1-x)(1-y)}$. This implies that

$$z \leq \sqrt{y(1-y)} = \sqrt{-(y-\frac{1}{2})^2+\frac{1}{4}} \leq \frac{1}{2}$$

and hence $z = 1/2$ if and only if $y = 1/2$ and $x = 0$. Since $W(A)$ is the elliptic disc with foci x, y and minor axis of length z, we obtain $(-1/4)I_2 \leq \text{Im } A \leq (1/4)I_2$ and thus the sufficiency of (v). For the necessity of (v), if $A = \begin{bmatrix} 0 & 1/2 \\ 0 & 1/2 \end{bmatrix}$, then $\text{Im } A = \begin{bmatrix} 0 & -i/4 \\ i/4 & 0 \end{bmatrix}$. Hence $\sigma(\text{Im}A) = \{-1/4, 1/4\}$. This implies that Im $W(A) = W(\text{Im}A) = [-1/4, 1/4]$, completing the proof. ■

Corollary 5.14. *Let the 2-by-2 matrix A be a product of two nonnegative contractions. Then* Re $A \geq (-1/8)I_2$, *and*

(vi) $$-\frac{1}{8} \in \partial W(A) \ \ if\ and\ only\ if \ \ A \cong \begin{bmatrix} 0 & \frac{\sqrt{3}}{4} \\ 0 & \frac{1}{4} \end{bmatrix}.$$

Proof. By Theorem 5.1, we may assume that A is a 2-by-2 matrix of the form $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$, where $0 \leq x \leq y \leq 1$ and $0 \leq z \leq (\sqrt{y}-\sqrt{x})\sqrt{(1-x)(1-y)}$. This implies that

$$\begin{aligned} z^2+(y-x)^2 &\leq (1-x)(1-y)(\sqrt{y}-\sqrt{x})^2+(\sqrt{y}+\sqrt{x})^2(\sqrt{y}-\sqrt{x})^2 \\ &= (1+\sqrt{xy})^2(\sqrt{y}-\sqrt{x})^2. \end{aligned}$$

Note that $W(A)$ is the elliptic disc with foci x, y and major axis of length $\sqrt{z^2+(y-x)^2}$. Therefore, by our assumption we have, for any $\lambda \in \text{Re } W(A)$,

$$\begin{aligned} \lambda &\geq \frac{(x+y)-\sqrt{z^2+(y-x)^2}}{2} \\ &\geq \frac{1}{2}(x+y+\sqrt{x}-\sqrt{y}+x\sqrt{y}-y\sqrt{x}). \end{aligned} \quad \text{(vii)}$$

Consider the function $f(x,y) = 1/2(x + y + \sqrt{x} - \sqrt{y} + x\sqrt{y} - y\sqrt{x})$ on R, where $R = \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 1, x \leq y \leq 1\}$. We observe that f is a continuous function on R. Consider the following cases:

(1) $x = 0$, $0 \leq y \leq 1$. Hence $f(x,y) \geq 1/2(y - \sqrt{y}) = 1/2((\sqrt{y} - 1/2)^2 - 1/4) \geq -1/8$ and f assumes a local minimum $-1/8$ if and only if $x = 0$, $y = 1/4$.

(2) $0 \leq x \leq 1$, $y = 1$. Hence $f(x,y) \geq 1/2(x + 1 + \sqrt{x} - 1 + x - \sqrt{x}) = x \geq 0$.

(3) $0 \leq x \leq 1$, $y = x$. Hence $f(x,y) \geq 1/2(x + x + \sqrt{x} - \sqrt{x} + x\sqrt{x} - x\sqrt{x}) = x \geq 0$.

(4) $0 < x < y < 1$. Some simple computations involving $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$ show that $x = y = 0$. This is a contradiction.

Hence f assumes an absolute minimum $-1/8$ if and only if $x = 0$ and $y = 1/4$. Therefore, by (vii), for any $\lambda \in$ Re $W(A)$, we have $\lambda \geq -1/8$ and thus obtain the sufficiency of (vi). For the necessity of (vi), if $A = \begin{bmatrix} 0 & \sqrt{3}/4 \\ 0 & 1/4 \end{bmatrix}$, then Re $A = \begin{bmatrix} 0 & \sqrt{3}/8 \\ \sqrt{3}/8 & 1/4 \end{bmatrix}$. Hence $\sigma(\text{Re}A) = \{-1/8, 3/8\}$. This implies that Re $W(A) = W(\text{Re}A) = [-1/8, 3/8]$, completing the proof. ■

Finally, we will give some equivalent conditions for a 2-by-2 matrix to be the product of two nonnegative contractions at least one of which is noninvertible (resp., invertible) in Corollary 5.15 (resp., Corollary 5.17).

Corollary 5.15. *The following conditions are equivalent for a* 2*-by-*2 *matrix* A:

(1) *A is noninvertible and is a product of two nonnegative contractions.*

(2) *A is a product of two nonnegative contractions which are both noninvertible.*

(3) *A is a product of two nonnegative contractions at least one of which is noninvertible.*

Proof. We need only prove (1) $\Rightarrow$ (2). This follows easily from Corollary 5.7 and our assumption that A is noninvertible. ■

To prove the next corollary, we need the following lemma.

Lemma 5.16. *Let* $A = \begin{bmatrix} 0 & \sqrt{y(1-y)} \\ 0 & y \end{bmatrix}$ *be a 2-by-2 matrix, where* $0 < y < 1$, *and* $A = BC$ *for some* B *and* C, $0 \le B, C \le I_2$. *Then both* B *and* C *are nontrivial orthogonal projections.*

Proof. We may assume that $B \cong \begin{bmatrix} b_1 & 0 \\ 0 & b_2 \end{bmatrix}$ (resp., $C \cong \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$), where $b_2 \neq 0$, $c_2 \neq 0$ and $0 \le b_1 \le b_2 \le 1$, $0 \le c_1 \le c_2 \le 1$. Considering $A/(b_2c_2) = (B/b_2)(C/c_2)$, we obtain, by Lemma 5.4 (1), that $b_1c_1 = 0$ and

$$\text{(viii)} \qquad \frac{y(1-y)}{b_2^2c_2^2} = (1 - \frac{y}{b_2c_2})(\frac{y}{b_2c_2} - \frac{b_1}{b_2} - \frac{c_1}{c_2}).$$

Hence we may assume, without loss of generality, that $b_1 = 0$. Plugging this into (viii) yields $(b_2(c_1 + c_2) - 1)y = b_2^2c_1c_2$. If $b_2(c_1 + c_2) \neq 1$, then $0 < y < 1$ implies that $(1 - b_2c_1)(1 - b_2c_2) < 0$. This contradicts our assumption. It follows that $b_2(c_1 + c_2) = 1$. Thus $c_1 = 0$ and hence $b_2 = c_2 = 1$. This proves our assertion. ■

Now, we prove the following corollary.

Corollary 5.17. *A 2-by-2 matrix A is a product of two nonnegative contractions at least one of which is invertible if and only if either* (1) *A is a product of two positive contractions, or* (2) *A is unitarily equivalent to a matrix of the form* $\begin{bmatrix} x & z \\ 0 & y \end{bmatrix}$, *where* $x = 0$, $0 \le y \le 1$ *and* $|z| < \sqrt{y(1-y)}$ (*if* $y = 0$ *or* 1, *then* $z = 0$).

Proof. Note that the necessity follows easily from Theorem 5.1 and Lemma 5.16. For the converse, we may assume that $A = \begin{bmatrix} 0 & z \\ 0 & y \end{bmatrix}$, where $0 < y < 1$ and $0 < |z| < \sqrt{y(1-y)}$. Let $\lambda_1 = 0$ and $\lambda_2 = y - (|z|^2/(1-y))$. Then $0 < \lambda_2 \le y < 1$ and $|z|^2 = (1-y)(y - \lambda_2)$. Therefore, our assertion follows from Lemma 5.4 (1). ■

5.3 Products of two n-by-n nonnegative contractions

On a finite-dimensional Hilbert space, Wu [47, Corollary 2.3] has shown that if $A = \begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix}$ is a product of two nonnegative matrices, then so are A_1 and A_2. Here we give another proof which holds for both finite- and infinite-dimensional Hilbert spaces. In fact, it is even true that nonnegative matrices are replaced by nonnegative contractions. The main result of this section is the following:

Theorem 5.18. *Let A be a bounded linear operator of the form*

$$\begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix} \text{ on } H \bigoplus K,$$

where H and K are both Hilbert spaces. If A is a product of two nonnegative contractions, then so are A_1 and A_2.

In order to prove this theorem, we need the following lemmas. The proof of the lemma can be found in [11, p. 547].

Lemma 5.19. *Let A be a bounded linear operator of the form*

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{bmatrix} \text{ on } H \bigoplus K,$$

where H and K are Hilbert spaces. Then A is nonnegative if and only if A_{11} and A_{22} are both nonnegative and there exists a contraction D mapping K into H satisfying $A_{12} = A_{11}^{1/2} D A_{22}^{1/2}$.

Lemma 5.20. *If A is a nonnegative bounded linear operator on the Hilbert space H, then there exists a (possibly unbounded) linear operator B on* ran A *such that*

$BA = P_{\overline{\text{ran}\, A}}$.

Proof. If A_1 is the restriction of A to $(\ker A)^\perp$, then A_1 is injective. So we can consider the (possibly unbounded) inverse A_1^{-1}: $\text{ran}\ A \to (\ker\ A)^\perp$. Since A is nonnegative, it is easy to check that $(\ker\ A)^\perp = \overline{\text{ran}\, A}$ and $A_1^{-1}A = P_{\overline{\text{ran}\, A}}$. Hence our assertion follows. ■

Proof of Theorem 5.18. By our assumption and Lemma 5.19, we may assume that $A = BC$, where B (resp., C) is of the form

$$\begin{bmatrix} B_1 & B_1^{1/2}D_1B_2^{1/2} \\ B_2^{1/2}D_1^*B_1^{1/2} & B_2 \end{bmatrix} \left(\text{resp.,} \begin{bmatrix} C_1 & C_1^{1/2}D_2C_2^{1/2} \\ C_2^{1/2}D_2^*C_1^{1/2} & C_2 \end{bmatrix}\right) \text{ on } H \bigoplus K,$$

$0 \le B_1 \le I_H$ (resp., $0 \le C_1 \le I_H$), $0 \le B_2 \le I_K$ (resp., $0 \le C_2 \le I_K$) and D_1 (resp., D_2) is a contraction from K into H. From $A = BC$, we obtain that

$$\text{(ix)} \qquad A_1 = B_1C_1 + B_1^{1/2}D_1(B_2^{1/2}C_2^{1/2}D_2^*C_1^{1/2}),$$

$$\text{(x)} \qquad B_2^{1/2}(D_1^*B_1^{1/2}C_1^{1/2})C_1^{1/2} = -B_2^{1/2}(B_2^{1/2}C_2^{1/2}D_2^*)C_1^{1/2},$$

$$A_2 = (B_2^{1/2}D_1^*B_1^{1/2}C_1^{1/2})D_2C_2^{1/2} + B_2C_2.$$

Since $0 \le B_2^{1/2} \le I_K$, we deduce from Lemma 5.20 that there exists a (possibly unbounded) linear operator E on $\text{ran}\, B_2^{1/2}$ such that $EB_2^{1/2} = P_{\overline{\text{ran}\, B_2^{1/2}}}$. Hence by (x), we derive that

$$B_2^{1/2}C_2^{1/2}D_2^*C_1^{1/2} = P_{\overline{\text{ran}\, B_2^{1/2}}}(B_2^{1/2}C_2^{1/2}D_2^*C_1^{1/2}) = -P_{\overline{\text{ran}\, B_2^{1/2}}}(D_1^*B_1^{1/2}C_1).$$

Moreover, substitute this into (ix) to get

$$\begin{aligned} A_1 &= B_1C_1 - B_1^{1/2}D_1(P_{\overline{\text{ran}\, B_2^{1/2}}}(D_1^*B_1^{1/2}C_1)) \\ &= [B_1^{1/2}(I_H - D_1P_{\overline{\text{ran}\, B_2^{1/2}}}D_1^*)B_1^{1/2}]C_1 \\ &= [B_1^{1/2}(I_H - (P_{\overline{\text{ran}\, B_2^{1/2}}}D_1^*)^*(P_{\overline{\text{ran}\, B_2^{1/2}}}D_1^*))B_1^{1/2}]C_1. \end{aligned}$$

Note that $\|P_{\overline{\operatorname{ran} B_2^{1/2}}} D_1^*\| \leq 1$ implies that $0 \leq (I_H - (P_{\overline{\operatorname{ran} B_2^{1/2}}} D_1^*)^*(P_{\overline{\operatorname{ran} B_2^{1/2}}} D_1^*)) \leq I_H$. Therefore, $A_1 = [(B_1^{1/2} P_1^*) P_1 B_1^{1/2}] C_1$, where $P_1^* P_1 = I_H - (P_{\overline{\operatorname{ran} B_2^{1/2}}} D_1^*)^*(P_{\overline{\operatorname{ran} B_2^{1/2}}} D_1^*)$ for some P_1, $0 \leq P_1 \leq I_H$. This shows that A_1 is a product of two nonnegative contractions. In a similar fashion, we also derive that A_2^* is a product of two nonnegative contractions, and hence so is A_2. This completes our proof. ∎

The following corollaries are easy consequences of the preceding theorem. Recall that the first one has been proven in Wu [47, Corollary 2.3].

Corollary 5.21. *If an n-by-n* ($n \geq 2$) *matrix* $A = \begin{bmatrix} A_1 & A_3 \\ 0 & A_2 \end{bmatrix}$ *is a product of two n-by-n nonnegative matrices, then so are* A_1 *and* A_2.

Proof. We assume that $A = BC$, where $B, C \geq 0$. Since $A/(\|B\|\|C\|) = (B/\|B\|)(C/\|C\|)$, Theorem 5.18 implies that each of $A_1/(\|B\|\|C\|)$ and $A_2/(\|B\|\|C\|)$ is a product of two nonnegative contractions. Hence each of A_1 and A_2 is a product of two nonnegative matrices. Our assertion follows. ∎

Corollary 5.22. *An n-by-n matrix* $A = (\sum_{i=1}^{k_1} \oplus A_i) \oplus \operatorname{diag}(w_1, \ldots, w_{k_2})$, *where* $A_i = \begin{bmatrix} x_i & z_i \\ 0 & y_i \end{bmatrix}$ *for all i, is a product of two nonnegative contractions if and only if* $0 \leq x_i, y_i, w_j \leq 1$ *and* $|z_i| \leq |\sqrt{x_i} - \sqrt{y_i}|\sqrt{(1-x_i)(1-y_i)}$ *for all i, j.*

Proof. Note that Theorem 5.18 implies that A is a product of two nonnegative contractions if and only if A_i, $i = 1, \ldots, k_1$, and diag $(w_1, \ldots, w_{k_2})$ are products of two nonnegative contractions. Thus our assertion follows from Theorem 5.1. ∎

Corollary 5.23. *Let A be a finite quadratic matrix. Then A is a product of two* *nonnegative contractions if and only if it is unitarily equivalent to a matrix of the form*

$$aI_1 \oplus bI_2 \oplus (\sum_{i=1}^{k} \oplus \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix}),$$

where $0 \leq a, b \leq 1$ *and* $0 < c_i \leq |\sqrt{a} - \sqrt{b}|\sqrt{(1-a)(1-b)}$ *for all* i.

Proof. Since A is quadratic, [46, Theorem 1.1] says that it is unitarily equivalent to a matrix of the form

$$aI_1 \oplus bI_2 \oplus \begin{bmatrix} aI_3 & D \\ 0 & bI_4 \end{bmatrix}, \text{ where } D > 0.$$

Hence we may assume that it is unitarily equivalent to a matrix of the form

$$A = aI_1 \oplus bI_2 \oplus (\sum_{i=1}^{k} \oplus \begin{bmatrix} a & c_i \\ 0 & b \end{bmatrix}), \text{ where } c_i > 0.$$

Therefore, our assertion follows from Corollary 5.22. ■

Corollary 5.24. *Let* A *be a bounded linear operator of the form*

$$\begin{bmatrix} 0 & \sqrt{y(1-y)} \\ 0 & y \end{bmatrix} \oplus A_1 \text{ on } \mathbb{C}^2 \bigoplus K,$$

where $0 < y < 1$ *and* K *is a Hilbert space. If* A *is a product of two nonnegative contractions* B *and* C, *then we have* $\{0,1\} \in \sigma(B)$ *and* $\{0,1\} \in \sigma(C)$.

Proof. This follows easily from Lemma 5.16 and the proof of Theorem 5.18. ■

In order to prove that the numerical range of a product of nonnegative contractions on a finite-dimensional space cannot be a circular disc, we give a more general result in the following. These are from [48]

Theorem 5.25. *Let A be an n-by-n matrix. If $W(A)$ is a circular disc centered at a, then a is an eigenvalue of A with its geometric multiplicity less than its algebraic multiplicity.*

Proof. Note that a is an eigenvalue of A with algebraic multiplicity at least two (cf. [33, Corollary 4.4]). Hence we may assume that A is unitarily equivalent to a matrix of the form $\begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}$, where $A_1 = \begin{bmatrix} a & * & * \\ 0 & \ddots & * \\ 0 & 0 & a \end{bmatrix}_{k\times k}$ and $A_3 = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_{n-k} \end{bmatrix}$ with $\lambda_j \neq a$ for all j. Suppose that the geometric multiplicity of a is equal to its algebraic multiplicity. Then we may assume that A is similar to B, where $B = aI_k \oplus B_1$ and $a \notin \sigma(B_1)$. This implies that $\text{rank}\,(A - aI_n) = \text{rank}\,(B - aI_n) = n - k$ and hence $A_1 = aI_k$. In addition, let λ be the maximum eigenvalue of Re $(\omega(A - aI_n))$, $|\omega| = 1$. Then our assumption that $W(A)$ is a circular disc centered at a implies $\det(\lambda I_n - \text{Re}(\omega(A - aI_n)) = 0$. Since $\det(\lambda I_n - \text{Re}(\omega(A - aI_n)))$ can be considered as a trigonometric polynomial in ω with infinitely many zeros, the coefficients of ω^j for $j = 0, \pm 1, \ldots, \pm(n-k)$ are all zero. Since the coefficient of ω^{n-k} can be computed to be $(-1/2)^{n-k}\lambda^k(\lambda_1 - a)\cdots(\lambda_{n-k} - a)$, it follows that $\lambda_i = a$ for some i. This contradicts our assumption. Our proof is then completed. ■

Corollary 5.26. *Let the n-by-n matrix A be a product of two nonnegative matrices. Then $W(A)$ is not a circular disc. In particular, $W(A)$ is not a circular disc if A is a product of two nonnegative contractions.*

Proof. Suppose that $W(A)$ is a circular disc centered at a. Then a is an eigenvalue of A with multiplicity at least two. Since A is a product of two nonnegative matrices, by [47, Theorem 2.2], A is similar to a nonnegative matrix D. Hence the geometric multiplicity of the eigenvalue a of A is equal to its algebraic multiplicity. Therefore, Theorem 5.25 leads to a contradiction. This proves our assertion. ∎

Note that the result of Corollary 5.26 may be false if the assumption on nonnegativity is replaced by self-adjointness. We give an easy example here. Suppose that $A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $A = A_1A_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Then A_1 and A_2 are both self-adjoint but $W(A)$ is a circular disc.

We know that if a bounded linear operator A is a product of two nonnegative contractions, then $\sigma(A) = \sigma(A^*)$. However, $W(A)$ may not be equal to $W(A^*)$. We give an example in the following. Note that this also means that $W(A_1A_2) \neq W(A_2A_1)$ for some A_1, A_2, $0 \leq A_1$, $A_2 \leq I_n$.

Example 5.27. *Let*

$$A = \begin{bmatrix} 0 & i & i \\ 0 & \frac{1}{2} & i \\ 0 & 0 & 1 \end{bmatrix}.$$

Then there exists $\epsilon > 0$ such that ϵA is a product of two nonnegative contractions and $W(\epsilon A) \neq W(\epsilon A^*)$.

Proof. It is easy to see that both A and A^* are similar to $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Hence by [47, Theorem 2.2], we know that $A = BC$ for some nonnegative matrices B and C.

Thus ϵA is a product of two nonnegative contractions for some $\epsilon > 0$. In addition,

$$\text{Im } A = \frac{1}{2i}\begin{bmatrix} 0 & i & i \\ i & 0 & i \\ i & i & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

and $\text{Im } A^* = -\text{ Im } A$. It is easy to verify that $\sigma(\text{Im } A) = \{-1/2, 1\}$ and $\sigma(\text{Im } A^*) = \{-1, 1/2\}$. Hence

$$\text{Im } W(A) = W(\text{Im } A) = \sigma(\text{Im } A)^\wedge = [-1/2, 1]$$

and

$$\text{Im } W(A^*) = W(\text{Im } A^*) = -W(\text{Im } A) = [1/2, -1].$$

This shows that $W(A) \neq W(A^*)$ and hence $W(\epsilon A) \neq W(\epsilon A^*)$. ■

Now, we will give several propositions about a product of finitely many nonnegative contractions. The first one is analogous to [47, Proposition 3.5].

Proposition 5.28. *Let A be a bounded linear operator of the form*

$$\begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \text{ on } H \bigoplus K,$$

where H and K are Hilbert spaces. If $A = BCD$ for some nonnegative contractions B, C, D and either B or D is invertible, then A_1 is also a product of three nonnegative contractions.

Proof. We may assume, without loss of generality, that D is invertible. The proof is analogous to the proof of [47, Proposition 3.5]. Let

$$D^{-1} = \begin{bmatrix} D_1 & D_2 \\ D_3 & D_4 \end{bmatrix}.$$

Then $D^{-1} \geq I_{H \bigoplus K}$ and

$$AD^{-1} = \begin{bmatrix} A_1 D_1 & A_1 D_2 \\ 0 & 0 \end{bmatrix} = BC$$

is the product of two nonnegative contractions. It follows from Theorem 5.18 that the same is true for $A_1 D_1$. Since $D^{-1} \geq I_{H \bigoplus K}$, we know by Lemma 5.19 that $D_1 \geq I_H$. Hence $0 < D_1^{-1} \leq I_H$. Therefore, A_1 is the product of three nonnegative contractions. ■

Proposition 5.29. *For* $n \geq 1$, αI_n *is a product of four nonnegative contractions if and only if* $0 \leq \alpha \leq 1$.

Proof. The sufficiency is trivial. For the converse, we only need to consider $\alpha > 0$. We may assume that $\alpha I_n = A_1 A_2 A_3 A_4$, where $0 < A_i \leq I_n$, $1 \leq i \leq 4$. Hence $|\alpha| \leq 1$ and $\alpha A_4^{-1} A_3^{-1} = A_1 A_2$. This implies that $\alpha\sigma(A_4^{-1} A_3^{-1}) = \sigma(A_1 A_2)$ and hence $0 < \alpha \leq 1$ follows from Lemma 5.2. ■

For the ease of exposition, we introduce some notations. Let

$S_1 = \{$ products of finitely many n-by-n orthogonal projections$\}$,

$S_2 = \{$ products of finitely many n-by-n nonnegative contractions$\}$

and

$S_3 = \{$ products of infinitely many n-by-n nonnegative contractions$\}$.

It is clear that $S_1 \subseteq S_2 \subseteq S_3$. From Kuo and Wu [29, Theorem 3.1], we know that an n-by-n matrix A is in S_1 if and only if A is unitarily equivalent to $I_k \oplus A_1$, where $0 \leq k \leq n$ and A_1 is singular with $\|A_1\| < 1$. They also derived that if an n-by-n matrix A is in S_2, then there exists $k > 0$ such that $\|(I - A)x\|^2 \leq k(\|x\|^2 - \|Ax\|^2)$ for all x (cf. [30, Proposition 2.1]). Moreover, if an n-by-n matrix A is in S_2, then A

is unitarily equivalent to $I_k \oplus A_2$, where $0 \le k \le n$, A_2 is a completely nonunitary contraction and $A_2 \in S_2$ (cf. [30, Proposition 2.3]). Note that $\|A_2\| < 1$ follows from the proof of [30, Theorem 2.2]. Here we give several related propositions and corollaries.

Proposition 5.30. *Let A be an n-by-n matrix. Then for every i, $1 \le i \le 3$, A is in S_i if and only if $A \cong I_k \oplus B_i$, where $0 \le k \le n$ and $B_i \in S_i$ with $\|B_i\| < 1$.*

Proof. From the above results, we only need consider the case $i = 3$. The sufficiency is trivial. For the converse, we assume that $A_m = T_1 \cdots T_m$ and $A = \prod_{j=1}^{\infty} T_j = \lim_{m\to\infty} A_m$ with $\|A\| = 1$. Since $1 \ge \|A_m\| \searrow \|A\|$, we have $\|A_m\| = 1$ and hence $1 \in \sigma(A_m)$ by [30, Proposition 2.3]. In addition, we know that $\limsup_m \sigma(A_m) \subseteq \sigma(A)$ (cf. [20, Problem 103]). This implies that $1 \in \sigma(A)$. Consider $H_m = \ker(I_n - A_m)$ and $H = \cap_{m=1}^{\infty} H_m$. We want to prove that $H_m \searrow H$. Fix $m_0 \in \mathbb{N}$, let $m \ge m_0$ and let v be a nonzero vector in H_m. Then $v = A_m v = T_1 \cdots T_m v$ and hence $\|v\| = \|T_m v\|$. [30, Proposition 2.1] shows that $T_m v = v$. By the same process, we obtain that $T_j v = v$ for every j, $1 \le j \le m$. It follows that $A_{m_0} v = v$, that is, $v \in H_{m_0}$. Thus we obtain $H_m \searrow H$. This implies that there exists an $N \in \mathbb{N}$ such that $H_m = H$ for all m, $m \ge N$. Hence by the above arguments, for every m, $m \ge N$, and $v \in H = H_m$, we have $A_m v = v$ and $T_j v = v$ for every j, $1 \le j \le m$. This also implies that $A_m^* v = (T_m \cdots T_1) v = v$. Therefore, for every j, $1 \le j \le m$, there exists some unitary matrix U such that $T_j = U^*(I_k \oplus T_j')U$ on $H \bigoplus H^\perp$ for some k, $0 \le k \le n$ and $A_m = U^*(I_k \oplus \prod_{j=1}^{m} T_j')U$ for every m. Note that $\|\prod_{j=1}^{m} T_j'\| < 1$ for every j, $j \ge N$ by the results preceding this proposition. Let $A = U^* \begin{bmatrix} B & C \\ D & E \end{bmatrix} U$ on $H \bigoplus H^\perp$.

Then $\lim_{m\to\infty} A_m = A$ implies that

$$\| \begin{bmatrix} I_k - B & -C \\ -D & \prod_{j=1}^m T_j' - E \end{bmatrix} \| = \|U(A_m - A)U^*\| = \|A_m - A\| \to 0 \text{ as } m \to \infty.$$

Hence

$$\|I_k - B\| \le \|A_m - A\| \to 0,$$

$$\| - C\| = \|P_H(A_m - A)|_{H^\perp}\| \le \|A_m - A\| \to 0,$$

$$\| - D\| = \|P_{H^\perp}(A_m - A)|_H\| \le \|A_m - A\| \to 0$$

and

$$\| \prod_{j=1}^{m} T_j' - E\| \le \|A_m - A\| \to 0 \text{ as } m \to \infty.$$

Therefore, $B = I_k$, $C = D = 0$ and $E = \lim_m \prod_{j=1}^m T_j'$. Finally, $\| \prod_{j=1}^m T_j'\| \le \| \prod_{j=1}^N T_j'\|$ for every m, $m \ge N$, implies that

$$\|E\| = \lim_m \| \prod_{j=1}^{m} T_j'\| \le \| \prod_{j=1}^{N} T_j'\| < 1.$$

This completes our proof. ∎

Proposition 5.31. *Let A be an n-by-n noninvertible matrix.*

(1) *The following conditions are equivalent :*

(a) $A \cong I_k \oplus A_1$, *where $0 \le k < n$ and A_1 is singular with $\|A_1\| < 1$,*

(b) $A \in S_1$,

(c) $A \in S_2$, *and*

(d) $A \in S_3$.

(2) $A \in \overline{S_2} \;\Leftrightarrow\; \|A\| \le 1$.

Proof. (1) (a) $\Leftrightarrow$ (b) follows easily from the noninvertibility of A and [29, Theorem 3.1]. It is clear that (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (d). (d) $\Rightarrow$ (a) follows easily from the fact that A is noninvertible and Proposition 5.30.

(2) The necessity is trivial. For the sufficiency, consider $A_n = (1-(1/n))A$, for all n. Then $\|A_n\| = (1-(1/n))\|A\| \leq (1-(1/n)) < 1$. Since A is noninvertible, we have A_n is noninvertible. Hence, by (1), $A_n \in S_2$. Therefore, the fact that A_n converges to A leads to our assertion. ■

Corollary 5.32. $S_2 \subseteq S_3 \subsetneqq \overline{S_2}$.

Proof. It is trivial that $S_2 \subseteq S_3 \subseteq \overline{S_2}$. In order to prove $\overline{S_2} \backslash S_3 \neq \emptyset$, we consider $A = [-1] \oplus [0]_{n-1}$. Since A is noninvertible, Proposition 5.31 (1) and (2) imply that $A \in \overline{S_2} \backslash S_3$. This proves our assertion. ■

In Section 5.2, we have mentioned that if A and B satisfy $0 \leq A$, $B \leq I_n$, then

$$\inf\{\lambda | \lambda \in \sigma(\mathrm{Re}AB), 0 \leq A, B \leq I_n\} = -1/8,$$

$$\sup\{\lambda | \lambda \in \sigma(\mathrm{Re}AB), 0 \leq A, B \leq I_n\} = 1,$$

$$\inf\{\lambda | \lambda \in \sigma(\mathrm{Im}AB), 0 \leq A, B \leq I_n\} = -1/4$$

and

$$\sup\{\lambda | \lambda \in \sigma(\mathrm{Im}AB), 0 \leq A, B \leq I_n\} = 1/4.$$

Moreover, we also have the following proposition.

Proposition 5.33. *For* $i = 1$, 2, *and* 3, *we have*

(1) $\inf\{\lambda | \lambda \in \sigma(\mathrm{Re}A), A \in S_i\} = \inf\{\lambda | \lambda \in \sigma(\mathrm{Im}A), A \in S_i\} = -1,$

(2) $\sup\{\lambda | \lambda \in \sigma(\mathrm{Re}A), A \in S_i\} = \sup\{\lambda | \lambda \in \sigma(\mathrm{Im}A), A \in S_i\} = 1.$

Proof. It is trivial that if A_i satisfy $0 \le A_i \le I_n$ for every i, $1 \le i \le m$, $m \in \mathbb{N}$, then

$$-I_n \le \ \mathrm{Re} \prod_{i=1}^{m} A_i \le I_n \text{ and } -I_n \le \ \mathrm{Im} \prod_{i=1}^{m} A_i \le I_n.$$

Hence our assertion follows from Proposition 5.31 (1). This completes the proof. ■

References

[1] H. Bercovici, Operator Theory and Arithmetic in H^∞, Amer. Math. Soc., Providence, 1988.

[2] C. A. Berger, Normal Dilations, Ph.D. dissertation, Cornell Univ., 1963.

[3] C. A. Berger, J. G. Stampfli, Mapping theorems for the numerical range, Amer. J. Math. 89 (1967) 1047–1055.

[4] R. Bouldin, The numerical range of a product, II, J. Math. Anal. Appl. 33 (1971) 212–219.

[5] M.-T. Chien, H. Nakazato, The numerical radius of a weighted shift operator with geometric weights, Electron. J. Linear Algebra 18 (2009) 58–63.

[6] M.-T. Chien, H. Nakazato, Flat portions on the boundary of the numerical ranges of certain Toeplitz matrices, Linear Multilinear Algebra 56 (2008) 143–162.

[7] M.-T. Chien, B.-S. Tam, Circularity of the numerical range, Linear Algebra Appl. 201 (1994) 113–133.

[8] V. Chkliar, Numerical radii of simple powers, Linear Algebra Appl. 265 (1997) 119–121.

[9] K. R. Davidson, J. A. R. Holbrook, Numerical radii of zero-one matrices, Michigan Math. J. 35 (1988) 261–267.

[10] R. G. Douglas, Banach Algebra Techniques in Operator Theory, Academic Press, New York, 1972.

[11] C. Foias, A. E. Frazho, The Commutant Lifting Approach to Interpolation Problems, Birkhäuser-Verlag, Basel, 1990.

[12] J. I. Fujii, M. Fujii, S. Izumino, F. Kubo, R. Nakamoto, Strang's inequality, Math. Japon. 37 (1992) 479–486.

[13] H.-L. Gau, C.-Y. Huang, P. Y. Wu, Numerical radius inequalities for square-zero and idempotent operators, Oper. Matrices 2 (2008) 137–141.

[14] H.-L. Gau, P. Y. Wu, Condition for the numerical range to contain an elliptic disc, Linear Algebra Appl. 364 (2003) 213–222.

[15] H.-L. Gau, P. Y. Wu, Numerical ranges of companion matrices, Linear Algebra Appl. 421 (2007) 202–218.

[16] H.-L. Gau, P. Y. Wu, Numerical ranges of nilpotent operators, Linear Algebra Appl. 429 (2008) 716–726.

[17] H.-L. Gau, P. Y. Wu, Dilation to inflations of $S(\phi)$, Linear Multilinear Algebra 45 (1998) 109–123.

[18] H.-L. Gau, P. Y. Wu, Line segments and elliptic arcs on the boundary of a numerical range, Linear Multilinear Algebra 56 (2008) 131–142.

[19] K. Gustafson, D. K. M. Rao, Numerical Range. The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.

[20] P. R. Halmos, A Hilbert Space Problem Book, second ed., Springer, New York, 1982.

[21] J. A. R. Holbrook, Multiplicative properties of the numerical radius in operator theory, J. Reine Angew. Math. 237 (1969) 166–174.

[22] J. A. R. Holbrook, Inequalities of von Neumann type for small matrices, in: Function Spaces, K. Jarosz, Ed., Marcel Dekker, New York, 1992, pp. 189–193.

[23] J. A. R. Holbrook, J.-P. Schoch, Theory vs. experiment: multiplicative inequalities for the numerical radius of commuting matrices, Oper. Theory Adv. Appl. 202 (2010) 273–284.

[24] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge Univ. Press, Cambridge, 1985.

[25] R. A. Horn, C. R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, Cambridge, 1991.

[26] J. N. Issos, The Field of Values of Non-negative Irreducible Matrices, Ph.D. dissertation, Auburn University, 1966.

[27] T. Kato, Some mapping theorems for the numerical range, Proc. Japan Acad. Ser. A Math. Sci. 41 (1965) 652–655.

[28] R. Kippenhahn, Über den Wertevorrat einer Matrix, Math. Nachr. 6 (1951) 193–228.

[29] K.-H. Kuo, P. Y. Wu, Factorization of matrices into partial isometries, Proc. Amer. Math. Soc. 105 (1989) 263–272.

[30] K.-H. Kuo, P. Y. Wu, Products of orthogonal projections, in: Functional Analysis and Related Topics, S. Koshi, ed., World Scientific, Singapore, 1991, pp. 127–137.

[31] P. Lancaster and M. Tismentsky, The Theory of Matrices with Applications, second ed., Academic Press, New York, 1985.

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

[33] J. Maroulas, P. J. Psarrakos, M. J. Tsatsomeros, Perron–Frobenius type results on the numerical range, Linear Algebra Appl. 348 (2002) 49–62.

[34] V. Müller, The numerical radius of a commuting product, Michigan Math. J. 35 (1988) 255–260.

[35] T. Nakazi, K. Takahashi, Two-dimensional representations of uniform algebras, Proc. Amer. Math. Soc. 123 (1995) 2777–2784.

[36] M. Namba, Geometry of Projective Algebraic Curves, Marcel Dekker, New York, 1984.

[37] K. Okubo, T. Ando, Operator radii of commuting products, Proc. Amer. Math. Soc. 56 (1976) 203–210.

[38] C. Pearcy, An elementary proof of the power inequality for the numerical radius, Michigan Math. J. 13 (1966) 289–291.

[39] D. K. M. Rao, Rango numerico de operadores commutativos, Rev. Colombiana Mat. 27 (1994) 231–233.

[40] W. C. Ridge, Numerical range of a weighted shift with periodic weights, Proc. Amer. Math. Soc. 55 (1976) 107–110.

[41] D. Sarason, Generalized interpolation in H^∞, Trans. Amer. Math. Soc. 127 (1967) 179–203.

[42] Q. F. Stout, The numerical range of a weighted shift, Proc. Amer. Math. Soc. 88 (1983) 495–502.

[43] B. Sz.-Nagy, Sur les contractions de l'espace de Hilbert, Acta Sci. Math. (Szeged) 15 (1953) 87–92.

[44] B. Sz.-Nagy, C. Foiaş, Harmonic Analysis of Operators on Hilbert space, North-Holland, Amsterdam, 1970.

[45] B.-S. Tam, S. Yang, On matrices whose numerical ranges have circular or weak circular symmetry, Linear Algebra Appl. 302/303 (1999) 193–221.

[46] S.-H. Tso, P. Y. Wu, Matricial ranges of quadratic operators, Rocky Mountain J. Math. 29 (1999) 1139–1152.

[47] P. Y. Wu, Products of positive semidefinite matrices, Linear Algebra Appl. 111 (1988) 53–61.

[48] P. Y. Wu, Numerical ranges as circular discs, preprint.