



Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

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

Numerical ranges and Geršgorin discs

Chi-Tung Chang^{a,*}, Hwa-Long Gau^{b,2}, Kuo-Zhong Wang^{a,2}, Pei Yuan Wu^{a,2}^a Department of Applied Mathematics, National Chiao Tung University, Hsin chu 300, Taiwan^b Department of Mathematics, National Central University, Chung-Li 320, Taiwan

ARTICLE INFO

Article history:

Received 3 October 2011

Accepted 9 September 2012

Available online 11 October 2012

Submitted by R.A. Horn

AMS classification:

15A60

Keywords:

Numerical range

Geršgorin disc

Unitarily irreducible matrix

Permutationally irreducible matrix

ABSTRACT

For a complex matrix $A = [a_{ij}]_{i,j=1}^n$, let $W(A)$ be its numerical range, and let $G(A)$ be the convex hull of $\bigcup_{i=1}^n \{z \in \mathbb{C} : |z - a_{ii}| \leq (\sum_{i \neq j} (|a_{ij}| + |a_{ji}|)) / 2\}$ and $G'(A) = \bigcap \{G(U^*AU) : U \text{ } n\text{-by-}n \text{ unitary}\}$. It is known that $W(A)$ is always contained in $G(A)$ and hence in $G'(A)$. In this paper, we consider conditions for $W(A)$ to be equal to $G(A)$ or $G'(A)$. We show that if $W(A) = G'(A)$, then the boundary of $W(A)$ consists only of circular arcs and line segments. If, moreover, A is unitarily irreducible, then $W(A)$ is a circular disc. (Almost) complete characterizations of 2-by-2 and 3-by-3 matrices A for which $W(A) = G'(A)$ are obtained. We also give criteria for the equality of $W(A)$ and $G(A)$. In particular, such A 's among the permutationally irreducible ones must have even sizes. We also characterize those A 's with size 2 or 4 which satisfy $W(A) = G(A)$.

Crown Copyright © 2012 Published by Elsevier Inc. All rights reserved.

1. Introduction

For an n -by- n matrix A , let $W(A)$ be its *numerical range*

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

where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the standard inner product and its associated norm in \mathbb{C}^n , respectively. If $A = [a_{ij}]_{i,j=1}^n$, then, for each i , $1 \leq i \leq n$, let

* Corresponding author.

E-mail addresses: d937208@oz.nthu.edu.tw (C.-T. Chang), hlgau@math.ncu.edu.tw (H.-L. Gau), kzwang@math.nctu.edu.tw (K.-Z. Wang), pywu@math.nctu.edu.tw (P.Y. Wu).

¹ Research supported by the National Science Council of the Republic of China under the post-doctoral fellowship NSC 100-2811-M-009-054.

² Research supported by the National Science Council of the Republic of China under the projects NSC 100-2115-M-008-004, NSC 99-2115-M-009-013-MY2 and NSC 99-2115-M-009-002-MY2, respectively.

$$g_i(A) = \frac{1}{2} \sum_{\substack{1 \leq j \leq n \\ j \neq i}} (|a_{ij}| + |a_{ji}|)$$

and

$$C_i(A) = \{z \in \mathbb{C} : |z - a_{ii}| \leq g_i(A)\},$$

and let

$$G(A) = \left(\bigcup_{i=1}^n C_i(A) \right)^\wedge,$$

the convex hull of $\bigcup_{i=1}^n C_i(A)$, and

$$G'(A) = \bigcap \{G(U^*AU) : U \text{ } n\text{-by-}n \text{ unitary matrix}\}.$$

The $C_i(A)$'s, $G(A)$ and $G'(A)$ are called the *Geršgorin discs*, *Geršgorin region* and *unitarily reduced Geršgorin region* of A , respectively. The purpose of this paper is to discuss when $W(A)$ and $G(A)$ (respectively, $G'(A)$) are equal.

The set $\bigcup_i C_i(A)$ was first proposed by S. Geršgorin [4] in 1931 to serve as an inclusion region for the eigenvalues of A . Its relation to $W(A)$ was considered by C. R. Johnson [9]; he proved that $W(A) \subseteq G(A)$ is always true. Note that both $W(A)$ and $G'(A)$ are invariant under the unitary similarity of A while $G(A)$, depending on the entries of A , is not. From these, we easily obtain $W(A) \subseteq G'(A)$. The main concern now is when the extremum cases $W(A) = G(A)$ and $W(A) = G'(A)$ hold.

In Section 2 below, we first prove a decomposition theorem (Theorem 2.4) for a matrix A with $G(A)$ contained in the closed half-plane $H \equiv \{z \in \mathbb{C} : \operatorname{Re} z \geq 0\}$ and with $W(A) \cap \partial H$ nonempty. It says that in this case A is permutationally similar to a direct sum $A_1 \oplus \cdots \oplus A_\ell \oplus B$, where each A_k , $1 \leq k \leq \ell$, is such that $\operatorname{Re} A_k$ is permutationally irreducible, $W(A_k) \cap \partial H$ is a singleton together with many other nice properties, and B satisfies $W(B) \cap \partial H = \emptyset$. This will be the main tool in proving conditions for $W(A) = G'(A)$ and $W(A) = G(A)$ in Sections 3 and 4, respectively. Using this, we derive in Proposition 2.6 that if $\operatorname{Re} A$ is permutationally irreducible, then $W(A) \cap \partial G(A)$ is contained in the circle with center $(1/n) \operatorname{tr} A$ and radius $(\sum_{i=1}^n g_i(A))/n$. In Section 3, we consider the equality of $W(A)$ and $G'(A)$. Among other things, we show that (1) if $W(A) = G'(A)$, then $\partial W(A)$ consists of circular arcs and line segments and, moreover, if A is unitarily irreducible, then $W(A)$ must be a circular disc (Theorem 3.4), (2) a 2-by-2 matrix A satisfies $W(A) = G'(A)$ if and only if it is unitarily similar to

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \text{ or } \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \text{ for some scalars } a \text{ and } b \text{ (Proposition 3.5), and (3) if a 3-by-3 matrix } A \text{ satisfies}$$

$W(A) = G'(A)$, then it is unitarily similar to

$$(i) \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix}, \quad (ii) \begin{bmatrix} a & b & \\ 0 & a & \\ & & c \end{bmatrix}, \quad \text{or} \quad (iii) \begin{bmatrix} a & b & \\ & a & c \\ & & a \end{bmatrix},$$

for some a, b and c , and, conversely, if A is of one of these forms with $|b| = |c|$ in (iii), then it satisfies $W(A) = G'(A)$ (Proposition 3.6). Finally, in Section 4, we study the property $W(A) = G(A)$. The main results here (Theorems 4.2 and 4.4) say that an n -by- n matrix A satisfies $W(A) = G(A)$ if and only if it is permutationally similar to a direct sum $D \oplus A_1 \oplus \cdots \oplus A_\ell \oplus B$, where D is a diagonal matrix, each A_k , $1 \leq k \leq \ell$, is such that $\operatorname{Re} A_k$ is permutationally irreducible and $W(A_k)$ is a circular disc which coincides with all the Geršgorin discs $C_i(A_k)$, and B satisfies $G(B) \subseteq W(D \oplus A_1 \oplus \cdots \oplus A_\ell)$. In this case,

if $\text{Re } A$ is permutationally irreducible, then n must be even. Complete characterizations for 2-by-2 and 4-by-4 matrices A with $W(A) = G(A)$ are obtained (Propositions 4.5 and 4.6).

For any nonzero complex number z , its *argument*, $\arg z$, is the unique number θ in $[0, 2\pi)$ such that $z = |z|e^{i\theta}$. The *trace* of a matrix A is denoted by $\text{tr}A$, and its *real part* $(A + A^*)/2$ and *imaginary part* $(A - A^*)/(2i)$ by $\text{Re } A$ and $\text{Im } A$, respectively. A^T is the *transpose* of A . We use $\text{diag}(a_1, \dots, a_n)$ to denote the n -by- n diagonal matrix with eigenvalues a_1, \dots, a_n . A *permutation matrix* is one each of whose rows and columns contains exactly one 1 and whose all other entries are 0. Two n -by- n matrices A and B are said to be *permutationally similar* if there is a permutation matrix P such that $P^TAP = B$; they are *unitarily similar* if $U^*AU = B$ for some unitary matrix U . In this paper, two notions of irreducibility will be used. An n -by- n matrix A is *permutationally reducible* if either $n = 1$ and $A = [0]$

or $n \geq 2$ and there is an n -by- n permutation matrix P such that P^TAP is of the form $\begin{bmatrix} B & C \\ O & D \end{bmatrix}$, where B

and D are square matrices; otherwise, A is *permutationally irreducible*. It is obvious that $A = [a_{ij}]_{i,j=1}^n$ is permutationally irreducible if and only if for any i and j with $1 \leq i \neq j \leq n$, there are distinct indices $r_0 = i, r_1, \dots, r_{\ell-1}, r_\ell = j$ such that for all $s, 1 \leq s \leq \ell, a_{r_{s-1}r_s}$ is nonzero. A is *unitarily reducible* if it is unitarily similar to a direct sum of other matrices; otherwise, it is *unitarily irreducible*.

Note that these two notions are different. For example, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is permutationally irreducible but

unitarily reducible, while $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is unitarily irreducible but permutationally reducible. It is known

that every matrix is permutationally similar (respectively, unitarily similar) to a direct sum of matrices with permutationally irreducible real parts (respectively, a direct sum of unitarily irreducible matrices), and the summands are unique up to permutations and permutational similarities (respectively, unitary similarities). In particular, the real part of a unitarily irreducible matrix must be permutationally irreducible. The above permutationally irreducible assertion can be proven by an easy graph-theoretic argument while the unitarily irreducible assertion was proven in [2, Corollary 3.2]. The former notion will be mostly referred to in Sections 2 and 4 while the latter in Section 3.

The general references for this paper are the two monographs [7,8] by Horn and Johnson. In particular, [7, Sections 6.1 and 6.2] contains some discussions on Geršgorin discs, [8, Chapter 1] on numerical ranges, and [8, p. 39, Problem 4] specifically asks whether $W(A) = G(A)$ is true. Other references for the numerical range are [5] and [6, Chapter 22]. [9] contains results on numerical ranges of 3-by-3 matrices, which will be used in Section 3. In the literature, there are papers discussing the containment relations between (generalized) Geršgorin regions and (generalized) numerical ranges. It seems that, other than [8, p. 39, Problem 4] and [17, Question 2], none has touched on their equality.

2. Nonemptiness of $W(A) \cap \partial G(A)$

We start by a result relating the attaining vector for a point of $W(A)$ to the entries of A . If $A = [a_{ij}]_{i,j=1}^n$ and $K = \{k_1, \dots, k_p\} \subseteq \{1, 2, \dots, n\}$ with $1 \leq k_1 < \dots < k_p \leq n$, we let $A^{[K]}$ denote the p -by- p matrix $[a_{k_i k_j}]_{i,j=1}^p$.

Proposition 2.1. *Let $A = [a_{ij}]_{i,j=1}^n$ and let $z = \langle Ax, x \rangle$ be a point in $W(A)$, where $x = [x_1 \dots x_n]^T$ is a unit vector in \mathbb{C}^n . If $K = \{i : x_i \neq 0\} = \{k_1, \dots, k_p\}$, where $1 \leq k_1 < \dots < k_p \leq n$, then*

$$\left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| \leq \sum_{i=1}^p |x_{k_i}|^2 g_i(A^{[K]}). \tag{1}$$

Moreover, (1) is an equality if and only if the following two conditions hold:

- (a) If $z = \sum_i |x_{k_i}|^2 a_{k_i k_i}$, then $A^{[K]} = \text{diag}(a_{k_1 k_1}, \dots, a_{k_p k_p})$; otherwise, $\arg(a_{k_s k_t} x_{k_t} \overline{x_{k_s}}) = \arg(z - \sum_i |x_{k_i}|^2 a_{k_i k_i}) \equiv \theta$ for all $k_s \neq k_t$ with $a_{k_s k_t} \neq 0$.
- (b) $|x_{k_s}| = |x_{k_t}|$ for all $k_s \neq k_t$ with $a_{k_s k_t}$ or $a_{k_t k_s}$ nonzero.

In particular, if (1) is an equality and $a_{k_s k_t} \neq 0$ for some $k_s \neq k_t$, then $x_{k_s} = x_{k_t} e^{i(\theta_{st} - \theta)}$, where $\theta_{st} = \arg a_{k_s k_t}$.

Proof. Since

$$z = \sum_{i,j=1}^n a_{ij} x_j \overline{x_i} = \sum_{i=1}^p a_{k_i k_i} |x_{k_i}|^2 + \sum_{1 \leq i \neq j \leq p} a_{k_i k_j} x_{k_j} \overline{x_{k_i}},$$

we have

$$\left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| = \left| \sum_{1 \leq i \neq j \leq p} a_{k_i k_j} x_{k_j} \overline{x_{k_i}} \right| \leq \sum_{i \neq j} |a_{k_i k_j}| |x_{k_j}| |x_{k_i}| = \frac{1}{2} \sum_{i \neq j} |x_{k_j}| |x_{k_i}| (|a_{k_i k_j}| + |a_{k_j k_i}|) \tag{2}$$

$$\leq \frac{1}{4} \sum_{i \neq j} (|x_{k_i}|^2 + |x_{k_j}|^2) (|a_{k_i k_j}| + |a_{k_j k_i}|) \tag{3}$$

$$= \frac{1}{2} \sum_{i \neq j} |x_{k_i}|^2 (|a_{k_i k_j}| + |a_{k_j k_i}|) = \sum_{i=1}^p |x_{k_i}|^2 g_i (A^{[K]}).$$

This proves (1). Moreover, (1) becomes an equality if and only if the inequalities in (2) and (3) are equalities, which are equivalent to (a) and (b).

Finally, if (1) is an equality and $a_{k_s k_t} \neq 0$ for some $k_s \neq k_t$, then

$$|a_{k_s k_t}| e^{i\theta_{st}} x_{k_t} \overline{x_{k_s}} = a_{k_s k_t} x_{k_t} \overline{x_{k_s}} = |a_{k_s k_t}| |x_{k_t}| |x_{k_s}| e^{i\theta} = |a_{k_s k_t}| |x_{k_t}| e^{-i(\arg x_{k_t})} \overline{x_{k_s}} e^{i(\arg x_{k_s})} e^{i\theta},$$

where the second equality is ensured by (a). Since $a_{k_s k_t}$, x_{k_t} and x_{k_s} are all nonzero, we obtain $e^{i(\theta_{st} - \theta)} = e^{i(\arg x_{k_s} - \arg x_{k_t})}$. Together with (b), this yields $x_{k_s} = x_{k_t} e^{i(\theta_{st} - \theta)}$ as asserted. \square

Note that if $\text{Re } A$ is permutationally irreducible, then Proposition 2.1(b) is equivalent to

$$(b') \quad |x_{k_i}| = 1/\sqrt{p} \quad \text{for all } i, 1 \leq i \leq p.$$

Indeed, under the irreducibility of $\text{Re } A$, for any $1 \leq i \neq j \leq p$, there are distinct indices $r_0 = i, r_1, \dots, r_{\ell-1}, r_\ell = j$ such that for all $s, 1 \leq s \leq \ell$, either $a_{k_{r_{s-1}} k_{r_s}}$ or $a_{k_{r_s} k_{r_{s-1}}}$ is nonzero. Hence Proposition 2.1(b) implies that

$$|x_{k_i}| = |x_{k_{r_1}}| = \dots = |x_{k_{r_{\ell-1}}}| = |x_{k_j}|.$$

This obviously yields (b'). That (b') implies (b) is trivial.

In the following, we will frequently use some easily derived properties of $G(A)$, which we gather together in the next lemma.

Lemma 2.2. *If A is an n -by- n matrix, then*

- (a) $G(aA + bI_n) = aG(A) + b$ for any scalars a and b ,
- (b) $G(A) = G(U^*AU)$ for any diagonal unitary matrix U , and
- (c) $G(A) \supseteq G(B)$ for any principal submatrix B of A .

We now apply Proposition 2.1 to derive some necessary conditions for the nonemptiness of $W(A) \cap \partial G(A)$.

Proposition 2.3. *Let $A = [a_{ij}]_{i,j=1}^n$ be such that $G(A) \subseteq H \equiv \{z \in \mathbb{C} : \operatorname{Re} z \geq 0\}$ and $W(A) \cap \partial H \neq \emptyset$. If $z = \langle Ax, x \rangle$ is in $W(A) \cap \partial H$, where $x = [x_1 \dots x_n]^T$ is a unit vector in \mathbb{C}^n , $K = \{i : x_i \neq 0\} = \{k_1, \dots, k_p\}$ with $1 \leq k_1 < \dots < k_p \leq n$, and $J = \{j : C_j(A) \cap \partial H \neq \emptyset\}$, then the following hold:*

- (a) $K \subseteq J$.
- (b) $\operatorname{Re} a_{k_i k_i} = g_{k_i}(A)$ for all i , $1 \leq i \leq p$.
- (c) $\left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| = \sum_{i=1}^p |x_{k_i}|^2 g_{k_i}(A)$.
- (d) $\operatorname{Im} z = \sum_{i=1}^p |x_{k_i}|^2 \operatorname{Im} a_{k_i k_i}$.
- (e) *If $\operatorname{Re} a_{k_i k_i} = 0$ for all i , then $A^{[K]} = \operatorname{diag}(a_{k_1 k_1}, \dots, a_{k_p k_p})$. Otherwise, if $\operatorname{Re} a_{k_i k_i} > 0$ for some i , then for all $k_s \neq k_t$ with $a_{k_s k_t} \neq 0$, we have $x_{k_s} = -x_{k_t} e^{i(\arg a_{k_s k_t})}$, and, in particular, $\arg a_{k_s k_t} = \arg \overline{a_{k_t k_s}}$ for $a_{k_s k_t}, a_{k_t k_s} \neq 0$.*
- (f) $g_i(A^{[K]}) = g_i(\operatorname{Re} A^{[K]}) = g_{k_i}(A) = g_{k_i}(\operatorname{Re} A)$ for all i , $1 \leq i \leq p$.
- (g) A is permutationally similar to $A^{[K]} \oplus B$ for some $(n - p)$ -by- $(n - p)$ matrix B .

Proof. Since $G(A) \subseteq H$, we have $\operatorname{Re} a_{jj} \geq g_j(A)$ for all j , $1 \leq j \leq n$, and $\operatorname{Re} a_{jj} = g_j(A)$ if and only if j is in J .

- (a) To prove $K \subseteq J$, assume to the contrary that there is a k_0 in $K \setminus J$. The inequality (1) in Proposition 2.1 says that

$$\left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| \leq \sum_{i=1}^p |x_{k_i}|^2 g_{k_i}(A).$$

Hence

$$\begin{aligned} 0 = \operatorname{Re} z &\geq \operatorname{Re} \left(\sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right) - \sum_{i=1}^p |x_{k_i}|^2 g_{k_i}(A) \\ &= \sum_{i=1}^p |x_{k_i}|^2 (\operatorname{Re} a_{k_i k_i} - g_{k_i}(A)) \geq |x_{k_0}|^2 (\operatorname{Re} a_{k_0 k_0} - g_{k_0}(A)) > 0 \end{aligned}$$

since $k_0 \in J \setminus K$ implies that $x_{k_0} \neq 0$ and $\operatorname{Re} a_{k_0 k_0} > g_{k_0}(A)$. This contradiction yields that $K \subseteq J$.

- (b) Since each k_i , $1 \leq i \leq p$, in K is in J by (a), we have $C_{k_i}(A) \cap \partial H \neq \emptyset$ and hence $\operatorname{Re} a_{k_i k_i} = g_{k_i}(A) \geq 0$.
- (c) From (b), we derive that

$$\begin{aligned} \left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| &\geq \left| \operatorname{Re} \left(z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right) \right| \\ &= \sum_{i=1}^p |x_{k_i}|^2 \operatorname{Re} a_{k_i k_i} = \sum_{i=1}^p |x_{k_i}|^2 g_{k_i}(A) \geq \sum_{i=1}^p |x_{k_i}|^2 g_i(A^{[K]}). \end{aligned} \tag{4}$$

Combined with inequality (1), this yields equalities throughout. In particular, we have

$$\left| z - \sum_{i=1}^p |x_{k_i}|^2 a_{k_i k_i} \right| = \sum_{i=1}^p |x_{k_i}|^2 g_{k_i}(A).$$

- (d) As all the inequalities in (4) are equalities, we deduce from its first one that $\text{Im } z = \sum_{i=1}^p |x_{k_i}|^2 \text{Im } a_{k_i k_i}$.
- (e) If $\text{Re } a_{k_i k_i} = 0$ for all i , then $z = \sum_i |x_{k_i}|^2 a_{k_i k_i}$ by (d) and hence $A^{[K]} = \text{diag}(a_{k_1 k_1}, \dots, a_{k_p k_p})$ by Proposition 2.1(a). On the other hand, if $\text{Re } a_{k_i k_i} > 0$ for some i , then $z \neq \sum_i |x_{k_i}|^2 a_{k_i k_i}$ by (4). Note that $z - \sum_i |x_{k_i}|^2 a_{k_i k_i}$ is real and equals $-\sum_i |x_{k_i}|^2 \text{Re } a_{k_i k_i}$ by (d). Thus $\arg(z - \sum_i |x_{k_i}|^2 a_{k_i k_i}) = \pi$. Therefore, Proposition 2.1(a) ensures that $x_{k_s} = -x_{k_t} e^{i(\arg a_{k_s k_t})}$ for any $k_s \neq k_t$ with $a_{k_s k_t} \neq 0$.
- (f) Since $g_{k_i}(A) \geq g_i(A^{[K]})$ for all $i, 1 \leq i \leq p$, the equalities in (4) yield that they are actually equal to each other. To prove $g_i(A^{[K]}) = g_i(\text{Re } A^{[K]})$, we need show that $|a_{k_s k_t}| + |a_{k_t k_s}| = |a_{k_s k_t} + \overline{a_{k_t k_s}}|$ for all $k_s \neq k_t$. If one of $a_{k_s k_t}$ and $a_{k_t k_s}$ is zero, then this is obvious; otherwise, this follows from $\arg a_{k_s k_t} = \arg \overline{a_{k_t k_s}}$ in (e). Finally, from

$$g_{k_t}(\text{Re } A) \leq g_{k_t}(A) = g_i(A^{[K]}) = g_i(\text{Re } A^{[K]}) \leq g_{k_t}(\text{Re } A),$$

where the two equalities have just been proven, we infer that $g_{k_i}(A) = g_{k_i}(\text{Re } A)$ as asserted.

- (g) Since $g_i(A^{[K]}) = g_{k_i}(A)$ for $1 \leq i \leq p$ from (f), we obtain $a_{k_j k_i} = a_{j k_i} = 0$ for all $j \notin K$. This yields the permutational similarity of A and $A^{[K]} \oplus B$ for some B . \square

Geometrically, condition (b) of the preceding proposition says that all the Geršgorin discs $C_{k_i}(A)$, $1 \leq i \leq p$, are tangent to the y -axis ∂H . Moreover, if $\text{Re } A$ is permutationally irreducible, then $|x_{k_i}| = 1/\sqrt{p}$ for all i (by condition (b') after Proposition 2.1), and hence (d) says that the “height” of z on ∂H is an average of those of the centers of the $C_{k_i}(A)$'s.

We are now ready for the main result of this section, a decomposition theorem for A with $G(A) \subseteq H$ and $W(A) \cap \partial H \neq \emptyset$. It refines Proposition 2.3.

Theorem 2.4. *Let $A = [a_{ij}]_{i,j=1}^n$ be such that $G(A) \subseteq H \equiv \{z \in \mathbb{C} : \text{Re } z \geq 0\}$ and $W(A) \cap \partial H \neq \emptyset$. Then A is permutationally similar to a matrix of the form $A_1 \oplus \dots \oplus A_\ell \oplus B$, where $\ell = \dim \ker(\text{Re } A)$, $W(B) \cap \partial H = \emptyset$, and for each $k, 1 \leq k \leq \ell$, $A_k = [a_{ij}^{(k)}]_{i,j=1}^{n_k}$ is a matrix of size n_k , which satisfies the following conditions:*

- (a) $\text{Re } a_{ii}^{(k)} = g_i(A_k) = g_i(\text{Re } A_k)$ for all $i, 1 \leq k \leq \ell$.
- (b) $\text{Re } A_k$ is permutationally irreducible.
- (c) If $z = \langle A_k x, x \rangle$ is in $W(A_k) \cap \partial H$, where $x = [x_1 \dots x_{n_k}]^T$ is a unit vector in \mathbb{C}^{n_k} , then $\text{Im } z = (\sum_{i=1}^{n_k} \text{Im } a_{ii}^{(k)})/n_k$ and $|x_i| = 1/\sqrt{n_k}$ for all i . In particular, $W(A_k) \cap \partial H$ is a singleton.
- (d) If $n_k > 1$, then $\text{Re } a_{ii}^{(k)} > 0$ for all i and $x_s = -x_t e^{i(\arg a_{st}^{(k)})}$ for all $s \neq t$ with $a_{st}^{(k)} \neq 0$.
- (e) $\dim \ker(\text{Re } A_k) = 1$.
- (f) If $U = \text{diag}(e^{i(\arg x_{k_1})}, \dots, e^{i(\arg x_{k_p})})$ and $U^* A_k U = [b_{ij}^{(k)}]_{i,j=1}^{n_k}$, then $\text{Re } b_{ii}^{(k)} \geq 0$ for all i and $b_{ij}^{(k)} \leq 0$ for all $i \neq j$.

Proof. Let $z' = \langle Ax', x' \rangle$ be a point in $W(A) \cap \partial H$, where $x' = [x'_1 \dots x'_n]^T$ is a unit vector in \mathbb{C}^n , and let $K = \{i : x'_i \neq 0\} = \{k_1, \dots, k_p\}$ with $1 \leq k_1 < \dots < k_p \leq n$. We may choose z' and x' to be such that p is the smallest. Then, by Proposition 2.3(g), A is permutationally similar to $A^{[K]} \oplus B_1$ for some

$(n - p)$ -by- $(n - p)$ matrix B_1 . Let $A_1 = A^{[K]} = [a_{ij}^{(1)}]_{i,j=1}^p$. We now show that A_1 satisfies the asserted properties (a) \sim (f) with n_1 there replaced by p .

- (a) This follows from Proposition 2.3(b) and (f).
- (b) Assume that $\text{Re } A_1$ is permutationally reducible. Using Proposition 2.3(e), we deduce that A_1 is permutationally similar to, say, $A^{[K_1]} \oplus A^{[K_2]}$, where $\{K_1, K_2\}$ is a partition of K . Since $z' = \langle A_1 x'', x'' \rangle$, where $x'' = [x''_1 \dots x''_p]^T$ is a unit vector in \mathbb{C}^p given by $x''_i = x''_{k_i}$ for $1 \leq i \leq p$, we have $z' \in W(A_1) \cap \partial H$. Furthermore, since $W(A_1) = (W(A^{[K_1]}) \cup W(A^{[K_2]}))^\wedge$, we may assume that $W(A^{[K_1]}) \cap \partial H \neq \emptyset$. If $K_1 = \{r_1, \dots, r_q\}$ with $1 \leq r_1 < \dots < r_q \leq p$, then there is a unit vector $u = [u_1 \dots u_q]^T$ in \mathbb{C}^q such that $\langle A^{[K_1]} u, u \rangle$ is in $W(A^{[K_1]}) \cap \partial H$. Let $v = [v_1 \dots v_n]^T$ be given by

$$v_j = \begin{cases} u_i & \text{if } j = k_{r_i} \text{ for some } i, 1 \leq i \leq q, \\ 0 & \text{otherwise.} \end{cases}$$

Then v is a unit vector in \mathbb{C}^n and $\langle Av, v \rangle = \langle A^{[K_1]} u, u \rangle$ is in $W(A) \cap \partial H$. Since the cardinality of $\{j : v_j \neq 0\}$ is at most q , which is in turn strictly less than p , this contradicts our choice of p in the first place. Hence $\text{Re } A_1$ is indeed permutationally irreducible.

- (c) If $z = \langle A_1 x, x \rangle$ is any point in $W(A_1) \cap \partial H$, where $x = [x_1 \dots x_p]^T$ is a unit vector in \mathbb{C}^p , then the x_j 's must all be nonzero. This is because if some of the x_j 's are zero, then Proposition 2.3(g) applied to A_1 gives the permutational reducibility of $\text{Re } A_1$, which contradicts (b). By Propositions 2.3(c) and 2.1(b'), we have $|x_j| = 1/\sqrt{p}$ for all j . Hence Proposition 2.3(d) yields $\text{Im } z = (\sum_{i=1}^p \text{Im } a_{ii}^{(1)})/p$ as required. In particular, this shows that z is unique and thus $z = z'$ and $W(A_1) \cap \partial H$ is a singleton.
- (d) If $p > 1$ and $\text{Re } a_{ii}^{(1)} = 0$ for some i , then (a) implies that $g_i(\text{Re } A_1) = 0$. This would result in the permutational reducibility of $\text{Re } A_1$, contradicting (b). Hence we must have $\text{Re } a_{ii}^{(1)} > 0$ for all i . Proposition 2.3(e) then implies that $x_s = -x_t e^{i(\arg a_{st}^{(1)})}$ for all $s \neq t$ with $a_{st}^{(1)} \neq 0$.
- (e) Since $\text{Re } A_1$ is permutationally irreducible by (b), for each $j, 1 < j \leq p$, there are distinct indices $r_0 = j, r_1, \dots, r_{q-1}, r_q = 1$ such that either $a_{r_{s-1}r_s}^{(1)}$ or $a_{r_s r_{s-1}}^{(1)}$ is nonzero for all $s, 1 \leq s \leq q$. Let

$$\theta_j^s = \begin{cases} \arg \left(-a_{r_{s-1}r_s}^{(1)} \right) & \text{if } a_{r_{s-1}r_s}^{(1)} \neq 0, \\ \arg \left(-\overline{a_{r_s r_{s-1}}^{(1)}} \right) & \text{if } a_{r_s r_{s-1}}^{(1)} \neq 0 \end{cases}$$

for $1 \leq s \leq q$. Note that if both $a_{r_{s-1}r_s}^{(1)}$ and $a_{r_s r_{s-1}}^{(1)}$ are nonzero, then $\arg \left(-a_{r_{s-1}r_s}^{(1)} \right) = \arg \left(-\overline{a_{r_s r_{s-1}}^{(1)}} \right)$ by Proposition 2.3(e), and thus the θ_j^s above is well-defined. Moreover, when $a_{r_{s-1}r_s}^{(1)} \neq 0$, we have $x_{r_{s-1}} = -x_{r_s} e^{i(\arg a_{r_{s-1}r_s}^{(1)})}$ by Proposition 2.3(e) again. Thus

$$x_{r_{s-1}} = x_{r_s} e^{i(\pi + \arg a_{r_{s-1}r_s}^{(1)})} = x_{r_s} e^{i\theta_j^s}.$$

Similarly, the same is true for $a_{r_s r_{s-1}}^{(1)} \neq 0$. For each $j, 1 < j \leq p$, let $\theta_j = \sum_{s=1}^q \theta_j^s$. Then

$$x_j = x_{r_0} = x_{r_1} e^{i\theta_j^1} = x_{r_2} e^{i(\theta_j^1 + \theta_j^2)} = \dots = x_{r_q} e^{i(\sum_{s=1}^q \theta_j^s)} = x_1 e^{i\theta_j}.$$

Thus $x = x_1 [1 e^{i\theta_2} \dots e^{i\theta_p}]^T$ and hence $\dim \ker(\text{Re } A_1) = 1$.

(f) If $A'_1 = U^*A_1U = [b_{ij}^{(1)}]_{i,j=1}^p$ and $y = \left[\underbrace{1/\sqrt{p} \dots 1/\sqrt{p}}_p \right]^T$, then $Uy = x$ and

$$\langle A'_1y, y \rangle = \langle U^*A_1Uy, y \rangle = \langle A_1x, x \rangle = z = z'.$$

Note that $G(A'_1) \subseteq H$ and z' is in $W(A'_1) \cap \partial H$. Thus Proposition 2.3(e) applied to y yields that $1/\sqrt{p} = -\left(1/\sqrt{p}\right) e^{i(\arg b_{st}^{(1)})}$ for any $s \neq t$ with $b_{st}^{(1)} \neq 0$. Hence $b_{st}^{(1)} \leq 0$ for $s \neq t$. On the other hand, since $b_{ss}^{(1)} = a_{ss}^{(1)}$, we have $\text{Re } b_{ss}^{(1)} = \text{Re } a_{ss}^{(1)} = g_s(A_1) \geq 0$ by (a).

Note that if $W(B_1) \cap \partial H = \emptyset$, then we are done; otherwise, apply the above arguments to B_1 and proceed repeatedly. \square

In the remaining part of this section, we use the preceding theorem to derive properties of $W(A) \cap \partial G(A)$ and $G(A)$. For convenience, we restrict ourselves to matrices with $\text{Re } A$ permutationally irreducible.

Proposition 2.5. *Let $A = [a_{ij}]_{i,j=1}^n$ be such that $\text{Re } A$ is permutationally irreducible.*

- (a) *If $W(A) \cap \partial G(A)$ consists of two points, then the a_{ii} 's are on a line and $G(A)$ is the convex hull of two circular discs.*
- (b) *If $W(A) \cap \partial G(A)$ consists of at least three points, then the Geršgorin discs $C_i(A)$'s all coincide with each other and thus $a_{11} = a_{22} = \dots = a_{nn}$ and $G(A)$ is a circular disc.*

Proof. (a) Assume that $W(A) \cap \partial G(A) = \{z_1, z_2\}$ ($z_1 \neq z_2$). Let $L_j, j = 1, 2$, be a supporting line of $G(A)$ at z_j . After a translation and rotation, we may assume that L_1 is the y -axis and $G(A) \subseteq \{z \in \mathbb{C} : \text{Re } z \geq 0\}$. Thus Theorem 2.4 is applicable. Condition (a) there says that each Geršgorin disc $C_i(A), 1 \leq i \leq n$, is tangent to both L_j 's. Thus their centers a_{ii} are all on the bisecting line of L_1 and L_2 , and $G(A)$ is the convex hull of the two Geršgorin discs which are farthest apart.

(b) Assume that z_1, z_2 and z_3 are three distinct points in $W(A) \cap \partial G(A)$. Let $L_j, 1 \leq j \leq 3$, be a supporting line of $G(A)$ at z_j . As above, each $C_i(A), 1 \leq i \leq n$, is tangent to all the L_j 's. Note that the L_j 's are distinct. Indeed, if $L_1 = L_2 \equiv L$, then L contains two distinct points z_1 and z_2 of $W(A) \cap \partial G(A)$. This contradicts the assertion in Theorem 2.4(c) that $W(A) \cap L$ is a singleton. On the other hand, since the L_j 's are all supporting lines of the convex set $G(A)$, they cannot be all parallel to one another nor can they intersect at one single point. Thus the centers a_{ii} of the $C_i(A)$'s, $1 \leq i \leq n$, are all on all three bisecting lines of the L_j 's. It follows that the $C_i(A)$'s, and, in particular, their centers a_{ii} all coincide and hence $G(A) = C_i(A)$ is a circular disc. \square

The next two propositions give more information on the points in $W(A) \cap \partial G(A)$.

Proposition 2.6. *Let $A = [a_{ij}]_{i,j=1}^n$ ($n \geq 2$) be such that $\text{Re } A$ is permutationally irreducible. Then $W(A) \cap \partial G(A)$ is contained in the circle with center $(1/n) \text{tr}A$ and radius $(\sum_{i=1}^n g_i(A))/n$.*

Proof. Let $z = \langle Ax, x \rangle$ be any point in $W(A) \cap \partial G(A)$, where $x = [x_1 \dots x_n]^T$ is a unit vector in \mathbb{C}^n , and let L be a supporting line of $G(A)$ at z . After a translation and rotation, we may assume that L is the y -axis and $G(A) \subseteq \{z \in \mathbb{C} : \text{Re } z \geq 0\}$. Applying Theorem 2.4(c) and (f), we may further assume that $x_j = 1/\sqrt{n}$ for all j and $a_{ij} \leq 0$ for all $i \neq j$. Hence

$$z = \sum_{i,j=1}^n a_{ij}x_j\bar{x}_i = \frac{1}{n} \text{tr}A + \frac{1}{n} \sum_{1 \leq i \neq j \leq n} a_{ij}.$$

It follows that

$$\left| z - \frac{1}{n} \operatorname{tr} A \right| = -\frac{1}{n} \sum_{i \neq j} a_{ij} = \frac{1}{n} \sum_i g_i(A).$$

Thus z is indeed on the asserted circle. \square

Note that, in the above proof, the asserted circle is also tangent to L . This is because

$$\operatorname{Re} \left(\frac{1}{n} \operatorname{tr} A \right) = \frac{1}{n} \sum_i \operatorname{Re} a_{ii} = \frac{1}{n} \sum_i g_i(A)$$

by Theorem 2.4(a).

Proposition 2.7. *Let $A = [a_{ij}]_{i,j=1}^n$ be such that $\operatorname{Re} A$ is permutationally irreducible. If there are distinct indices $k_j, 1 \leq j \leq p$, with $1 \leq k_j \leq n$ for all j so that some permutation $k'_j, 1 \leq j \leq p$, of them satisfies $k'_j \neq k_j$ for all j and $a_{k_1 k'_1}, \dots, a_{k_p k'_p} \neq 0$, then $W(A) \cap \partial G(A)$ consists of at most p points, and these points are some of the vertices of a regular p -gon inscribed on the circle with center $(1/n) \operatorname{tr} A$ and radius $(\sum_{i=1}^n g_i(A)) / n$.*

Proof. Let $z = \langle Ax, x \rangle$ and $z' = \langle Ay, y \rangle$ be points in $W(A) \cap \partial G(A)$, where $x = [x_1 \dots x_n]^T$ and $y = [y_1 \dots y_n]^T$ are unit vectors in \mathbb{C}^n , and let L be a supporting line of $G(A)$ at z' . As before, we may assume that L is the y -axis, $G(A) \subseteq \{z \in \mathbb{C} : \operatorname{Re} z \geq 0\}$, $y_j = 1/\sqrt{n}$ for all j , and $a_{ij} \leq 0$ for all $i \neq j$. On the other hand, after a translation and a rotation by θ ($0 \leq \theta < 2\pi$), we also have, by Theorem 2.4(c) and (d), $|x_t| = 1/\sqrt{n}$ for all t and

$$x_s = -x_t e^{i(\theta + \arg a_{st})} = x_t e^{i\theta}$$

for all $s \neq t$ with $a_{st} \neq 0$. In particular, this is true for all x_{k_j} 's and $x_{k'_j}$'s. Hence

$$\prod_{j=1}^p x_{k_j} = \left(\prod_{j=1}^p x_{k'_j} \right) e^{ip\theta} = \left(\prod_{j=1}^p x_{k_j} \right) e^{ip\theta},$$

from which we deduce that $e^{ip\theta} = 1$. Thus $a_{ij} x_j \bar{x}_i = (1/n) a_{ij} e^{i\theta}$ for all $i \neq j$, and therefore

$$z - \frac{1}{n} \operatorname{tr} A = \sum_{i,j=1}^n a_{ij} x_j \bar{x}_i - \frac{1}{n} \sum_{i=1}^n a_{ii} = \frac{e^{i\theta}}{n} \sum_{1 \leq i \neq j \leq n} a_{ij}$$

and, similarly,

$$z' - \frac{1}{n} \operatorname{tr} A = \sum_{i,j=1}^n a_{ij} y_j \bar{y}_i - \frac{1}{n} \sum_{i=1}^n a_{ii} = \frac{1}{n} \sum_{1 \leq i \neq j \leq n} a_{ij}.$$

It follows that $z - (1/n) \operatorname{tr} A = (z' - (1/n) \operatorname{tr} A) e^{i\theta}$ and

$$\left| z - \frac{1}{n} \operatorname{tr} A \right| = -\frac{1}{n} \sum_{i \neq j} a_{ij} = \frac{1}{n} \sum_i g_i(A). \quad \square$$

3. Equality of $W(A)$ and $G'(A)$

In this section, we consider matrices A for which $W(A)$ and $G'(A)$ are equal, obtain some necessary/sufficient conditions, and give complete characterizations when A is of size 2. In contrast, we remark that if the unitary similarity of A in the definition of $G'(A)$ is replaced by similarity, then the intersection is much easier to characterize: $\bigcap \{G(X^{-1}AX) : X \text{ invertible}\}$ is equal to the convex hull of the spectrum of A (see [7, p. 351, Problem 2] or [8, p. 60, Problem 30]). This can be proven by invoking the Jordan canonical form of A , which is in turn similar to a direct sum of matrices of the form

$$\begin{bmatrix} a & \epsilon & & & \\ & a & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & \epsilon \\ & & & & a \end{bmatrix}$$

with arbitrarily small positive ϵ . Note also that the containment $W(A) \subseteq G'(A)$ is always true (cf. [9]). We start with the following lemma.

Lemma 3.1. *If A is an n -by- n matrix with $G'(A)$ contained in the closed right half-plane $H = \{z \in \mathbb{C} : \operatorname{Re} z \geq 0\}$, then there is an n -by- n unitary matrix U such that $G(U^*AU)$ is contained in H .*

Note that, in general, it may happen that a sequence of compact convex sets has its intersection contained in H , but none of these sets is in H . One example is given by the closed discs $\{z \in \mathbb{C} : |z - 1| \leq 1 + (1/n)\}$, $n \geq 1$. Lemma 3.1 says that this is not the case for $G'(A)$.

Proof of Lemma 3.1. For each $t_k = -1/k$, $k \geq 1$, our assumption of $G'(A) \subseteq H$ implies that there is an n -by- n unitary matrix U_k such that t_k is not in $G(U_k^*AU_k)$. Let L_k and M_k be two supporting lines of $G(U_k^*AU_k)$ which pass through t_k , and let H_k denote the closed sector in the open left half-plane $\{z \in \mathbb{C} : \operatorname{Re} z < 0\}$ formed by L_k and M_k with vertex t_k . The convexity of $G(U_k^*AU_k)$ implies that $G(U_k^*AU_k) \cap H_k = \emptyset$ (cf. Figure 1). Note that there is a subsequence $\{U_{k_j}\}_{j=1}^\infty$ which converges to a unitary matrix U . Then $G(U_{k_j}^*AU_{k_j})$ converges to $G(U^*AU)$ in the Hausdorff metric as j approaches infinity. We infer that $G(U_k^*AU_k) \cap H_k = \emptyset$ for all $k \geq 1$. Thus

$$G(U^*AU) \cap M = G(U^*AU) \cap \left(\bigcup_{k=1}^\infty H_k \right) = \emptyset,$$

that is, $G(U^*AU) \subseteq H$ as asserted. \square

Recall that the Hausdorff metric h is defined, for nonempty compact subsets Δ_1 and Δ_2 of the plane, by

$$h(\Delta_1, \Delta_2) = \max\left\{ \max_{z_1 \in \Delta_1} \min_{z_2 \in \Delta_2} |z_1 - z_2|, \max_{z_2 \in \Delta_2} \min_{z_1 \in \Delta_1} |z_1 - z_2| \right\}.$$

It can be proven that if $\{A_k\}_{k=1}^\infty$ is a sequence of n -by- n matrices which converges to A in norm, then $G(A_k)$ converges to $G(A)$ in the Hausdorff metric.

To prepare for the next theorem, we state some general facts concerning the support function of a compact convex subset of the complex plane. Recall that if Δ is a nonempty compact convex subset of \mathbb{C} , then its support function $d(\theta)$ is, for each real θ , the signed distance from the origin to the supporting line L_θ of Δ which is perpendicular to the ray R_θ from the origin forming angle θ from the positive x -axis (cf. Figure 2). Thus L_θ is given by the equation $x \cos \theta + y \sin \theta = d(\theta)$,

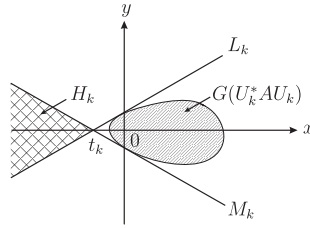


Fig. 1. Disjointness of $G(U_k^*AU_k)$ and H_k .

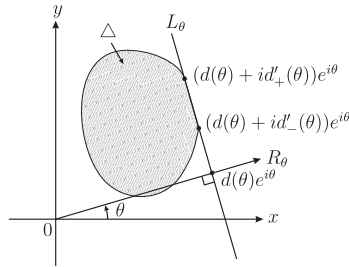


Fig. 2. Support function $d(\theta)$ of Δ .

$d(\theta) = \max\{\text{Re}(e^{-i\theta}z) : z \in \Delta\}$, and $\Delta = \bigcap_{\theta \in \mathbb{R}} \{z \in \mathbb{C} : \text{Re}(e^{-i\theta}z) \leq d(\theta)\}$. If the origin is in Δ (respectively, in the interior of Δ), then $d(\theta) \geq 0$ (respectively, $d(\theta) > 0$) for all θ . One reference for the support function is [16, Part V, Section A].

Lemma 3.2. Let Δ , $d(\theta)$ and L_θ be given as above. Then

- (a) $d(\theta)$ is both left and right differentiable for all θ ,
- (b) $d(\theta)$ is continuously differentiable for all but a countable number of values of θ ,
- (c) d is differentiable at θ if and only if $L_\theta \cap \partial\Delta$ is a singleton, in which case $L_\theta \cap \partial\Delta$ consists of the point $(d(\theta) + id'(\theta))e^{i\theta}$, and
- (d) d is not differentiable at θ if and only if $\partial\Delta$ contains a line segment on L_θ , in which case the endpoints of the line segment are $(d(\theta) + id'_\pm(\theta))e^{i\theta}$.

Proof. (a) follows from [16, Theorem 10.5] and (b) from [16, Corollary 10.3] or [1, Section I.4, Corollary 3.2]. We now prove that $L_\theta \cap \partial\Delta$ is the line segment with endpoints $(d(\theta) + id'_\pm(\theta))e^{i\theta}$. (c) and (d) will then follow from this easily. Indeed, after a rotation of R_θ by the angle $-\theta$, we may assume that Δ has a vertical supporting line $L = L_\theta$ given by the equation $x = d(0)$ and the corresponding perpendicular ray R_0 (cf. Figure 3). We check that the lowest point of $L \cap \partial\Delta$ is $d(0) + id'_-(0)$. Indeed, from Figure 3 we have

$$\begin{aligned} d'_-(0) &= \lim_{\alpha \rightarrow 0^-} \frac{d(\alpha) - d(0)}{\alpha} = \lim_{\alpha \rightarrow 0^-} \frac{a_\alpha \cos \alpha - (a_\alpha + b_\alpha \tan(-\alpha))}{\alpha} \\ &= \lim_{\alpha \rightarrow 0^-} \left(a_\alpha \cdot \frac{\cos \alpha - 1}{\alpha} + b_\alpha \cdot \frac{\tan \alpha}{\alpha} \right) = 0 + \lim_{\alpha \rightarrow 0^-} b_\alpha = c. \end{aligned}$$

Similarly, the highest point of $L \cap \partial\Delta$ is $d(0) + id'_+(0)$. This proves our assertion and hence also (c) and (d). \square

The next lemma gives a condition for a part of the boundary of a compact convex set to be a circular arc.

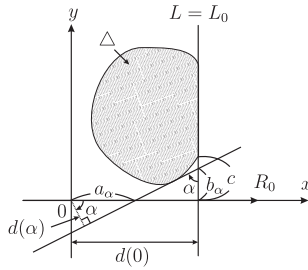


Fig. 3. Proof for lowest point of $L \cap \partial \Delta$.

Lemma 3.3.

- (a) Let Δ be a nonempty compact convex subset of the plane, and let α be an arcwise connected subset of $\partial \Delta$. Then α is a circular arc if and only if there is a point p in \mathbb{C} such that, for any q in α , the supporting line L_q of Δ at q is perpendicular to the line L_{pq} connecting p and q .
- (b) Let A be an n -by- n matrix and let $\partial W(A)$ be composed of the algebraic arcs $\alpha_1, \dots, \alpha_m$. Then, for any $\alpha_k, 1 \leq k \leq m$, and any point p in \mathbb{C} , either α_k is a circular arc with center p or there are only finitely many points q in α_k with supporting line L_q of $W(A)$ at q perpendicular to L_{pq} .

Proof.

- (a) If α is a circular arc, then with p the center of α we have the asserted perpendicular property. For the converse, we may assume that p is the origin. Let $d(\theta)$ be the support function of Δ and let (θ_1, θ_2) correspond to the arc α . Our assumption on α implies that it cannot contain any line segment. Hence $d(\theta)$ is differentiable for all θ in (θ_1, θ_2) by Lemma 3.2(c) or (d). Thus each point q of α is given by both $(d(\theta) + id'(\theta))e^{i\theta}$ and $d(\theta)e^{i\theta}$. Their equality then yields that $d'(\theta) = 0$ or $d(\theta)$ is a constant for all θ in (θ_1, θ_2) . Thus α is a circular arc.
- (b) Note that, by [11, Theorem 10], the boundary of the numerical range of a matrix is always composed of finitely many algebraic curves. Let $\alpha_k, 1 \leq k \leq m$, be given by

$$f_k(x, y) \equiv \sum_{0 \leq i+j \leq \ell_k} a_{ij}^{(k)} x^i y^j = 0,$$

where f_k is an irreducible polynomial of degree ℓ_k . An implicit differentiation of f_k with respect to x yields

$$\sum_{i,j} a_{ij}^{(k)} \left(ix^{i-1}y^j + jx^i y^{j-1} \frac{dy}{dx} \right) = 0.$$

Let $p = (x_0, y_0)$ and $q = (x_1, y_1)$. The perpendicular condition says that

$$\frac{dy}{dx}(x_1, y_1) = \frac{s_k(x_1, y_1)}{r_k(x_1, y_1)} = -\frac{1}{\frac{y_1 - y_0}{x_1 - x_0}},$$

where $s_k(x, y) = -\sum_{i,j} a_{ij}^{(k)} ix^{i-1}y^j$ and $r_k(x, y) = \sum_{i,j} a_{ij}^{(k)} jx^i y^{j-1}$, that is, (x_1, y_1) is a zero of the polynomial

$$g_k(x, y) \equiv s_k(x, y)(y - y_0) + r_k(x, y)(x - x_0)$$

of degree ℓ_k . If f_k and g_k have a common factor, then the irreducibility of f_k implies that f_k is a factor of g_k . Thus all the points of α_k have the perpendicular property. (a) then implies that α_k is

a circular arc with center p . On the other hand, if f_k and g_k have no common factor, then Bézout’s theorem [12, Theorem 3.9] implies that the number of their common zeros is at most ℓ_k^2 . In this case, there are at most ℓ_k^2 many points q in α_k with L_q perpendicular to L_{pq} . \square

We are now ready to prove the main result of this section. It gives the possible shape of the numerical range $W(A)$ of A when it is equal to $G'(A)$. Note that though the boundary of $G(A)$ consists of circular arcs and line segments, that of $G'(A)$ can in general be quite arbitrary as witness the fact that any nonempty convex compact subset of the plane is the intersection of closed polygonal regions (respectively, closed circular discs) containing it.

Theorem 3.4. *Let A be an n -by- n matrix with $W(A) = G'(A)$.*

- (a) *If A is nonscalar, then $\partial W(A)$ is composed of finitely many circular arcs and line segments.*
- (b) *If A is unitarily irreducible, then $W(A)$ is a circular disc centered at $(1/n)\text{tr}A$ and A is unitarily similar to a matrix $B = [b_{ij}]_{i,j=1}^n$ of the form*

$$\begin{bmatrix} \left(\frac{1}{n}\text{tr}A\right)I_1 & B_1 & & & \\ & \left(\frac{1}{n}\text{tr}A\right)I_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & B_{\ell-1} \\ & & & & \left(\frac{1}{n}\text{tr}A\right)I_\ell \end{bmatrix}$$

with $b_{ij} \geq 0$ for all $i \neq j$.

Proof.

- (a) Assume that $\partial W(A)$ is composed of the algebraic curves $\alpha_1, \dots, \alpha_m$. Let A be unitarily similar to $A_1 \oplus \dots \oplus A_r$, where each A_j , $1 \leq j \leq r$, is unitarily irreducible (cf. [2, Corollary 3.2]). For each ordered sequence of indices $1 \leq i_1 < \dots < i_s \leq r$, let $B_i = A_{i_1} \oplus \dots \oplus A_{i_s}$ and let $p_i = (1/m_i)\text{tr}B_i$, where m_i is the size of B_i . Assume that some α_j is not a circular arc nor a line segment. Then, by Lemma 3.3 (b), for each p_i there are only finitely many points q in α_j such that the supporting line L_q of $W(A)$ at q is perpendicular to L_{p_iq} , the line connecting p_i and q . Since α_j contains infinitely many points, by the pigeonhole principle, there is a point q_0 in α_j which is not equal to any of such q ’s, that is, the supporting line L_{q_0} of $W(A)$ at q_0 is not perpendicular to $L_{p_iq_0}$ for all the p_i ’s. Let H be the closed half-plane with boundary L_{q_0} which contains $W(A) = G'(A)$. By Lemma 3.1, there is a unitary matrix U such that $G(U^*AU)$ is contained in H . Now apply Theorem 2.4 to infer that U^*AU is permutationally similar to a matrix of the form $A'_1 \oplus \dots \oplus A'_t \oplus B$, where A'_k , $1 \leq k \leq t$, is an n_k -by- n_k matrix with $\text{Re } A'_k$ permutationally irreducible, $W(A'_k) \cap L_{q_0}$ is the singleton $\{q_0\}$ (because α_j is not a line segment) and the line perpendicular to L_{q_0} at q_0 passes through the point $(1/n_k)\text{tr}A'_k$. Since each A'_k (and also B) is unitarily similar to a direct sum of unitarily irreducible matrices and the unitarily irreducible decomposition is unique (up to the permutations and unitary similarities of the summands) by [2, Theorem 3.1], we obtain the unitary similarity of each A'_k to some B_i . Hence the line perpendicular to L_{q_0} at q_0 passes through $p_i = (1/m_i)\text{tr}B_i$, that is, L_{q_0} is perpendicular to $L_{p_iq_0}$, contradicting our choice of q_0 . It follows that each α_j is either a circular arc or a line segment.
- (b) In this case, for any point q_0 on $\partial W(A)$, let L be a supporting line of $W(A)$ at q_0 . Apply Lemma 3.1 to obtain a unitary matrix U such that $G(U^*AU)$ is contained in the closed half-plane which contains $W(A) = G'(A)$ and has boundary L . Theorem 2.4 then implies that $W(A) \cap L = \{q_0\}$ and the line perpendicular to L at q_0 passes through $(1/n)\text{tr}A$. Hence $W(A)$ is a circular disc centered at $(1/n)\text{tr}A$ by Lemma 3.3(a).

To prove the remaining part, we may assume, by considering $A - ((1/n)\text{tr } A)I_n$ instead of A , that $\text{tr } A = 0$ and hence $W(A) = G'(A) = \{z \in \mathbb{C} : |z| \leq r\}$ ($r > 0$). Since $G'(rI_n - A) \subseteq H \equiv \{z \in \mathbb{C} : \text{Re } z \geq 0\}$, Lemma 3.1 implies the existence of a unitary matrix U such that $G(rI_n - U^*AU) \subseteq H$. Applying Theorem 2.4(b) and (f) to $rI_n - U^*AU$, we obtain that $\text{Re } U^*AU$ is permutationally irreducible and there is a diagonal matrix V such that all nondiagonal entries of $(UV)^*A(UV)$ are nonnegative. [13, Theorem 1] then says that $(UV)^*A(UV)$ is permutationally similar to a matrix of the form

$$\begin{bmatrix} 0 & B_1 & & & \\ & 0 & \ddots & & \\ & & \ddots & B_{\ell-1} & \\ & & & & 0 \end{bmatrix}. \quad \square$$

We can now characterize 2-by-2 matrices A for which $W(A) = G'(A)$.

Proposition 3.5. *A 2-by-2 matrix A is such that $W(A) = G'(A)$ if and only if it is unitarily similar to $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ or $\begin{bmatrix} a & c \\ 0 & a \end{bmatrix}$ for some scalars a, b and c .*

Since the unitary similarity of two 2-by-2 matrices is equivalent to the equality of their numerical ranges, the preceding proposition says that, for a 2-by-2 matrix A , a necessary and sufficient condition for $W(A) = G'(A)$ is that $W(A)$ equals a singleton, a line segment or a circular disc.

Proof of Proposition 3.5. Let A be unitarily similar to $B = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$ for some a, b and c . Assume first that $W(A) = G'(A)$. If A is unitarily reducible, then $c = 0$ and we are done. Otherwise, $W(A)$ is a circular disc by Theorem 3.4(b). Since a and b are the foci of the elliptic disc $W(A) = W(B)$, we must have $a = b$ and hence A is unitarily similar to $\begin{bmatrix} a & c \\ 0 & a \end{bmatrix}$ as asserted. Conversely, if $B = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$ (respectively, $B = \begin{bmatrix} a & c \\ 0 & a \end{bmatrix}$), then $W(B)$ and $G(B)$ both equal the line segment with endpoints a and b (respectively, the circular disc with center a and radius $|c|/2$). In either case, we obviously have $W(A) = W(B) = G'(B) = G'(A)$. \square

The next proposition gives necessary/sufficient conditions for a 3-by-3 matrix A to satisfy $W(A) = G'(A)$.

Proposition 3.6. *If a 3-by-3 matrix A is such that $W(A) = G'(A)$, then it is unitarily similar to a matrix of the form*

$$(i) \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix}, \quad (ii) \begin{bmatrix} a & b & \\ & a & \\ & & c \end{bmatrix} \quad \text{or} \quad (iii) \begin{bmatrix} a & b & \\ & a & c \\ & & a \end{bmatrix}$$

for some scalars a, b and c . Conversely, if A is of one of the above forms with $|b| = |c|$ in (iii), then it satisfies $W(A) = G'(A)$.

Unfortunately, at the present time we are not yet able to verify whether a matrix of type (iii) above with nonzero b and c satisfying $|b| \neq |c|$ has the property $W(A) \subsetneq G'(A)$.

Proof of Proposition 3.6. Let A be a 3-by-3 matrix with $W(A) = G'(A)$. Obviously, we need only consider the case of unitarily irreducible A . Let a_1, a_2 and a_3 be its eigenvalues. Under our assumptions, $W(A)$ is a circular disc with center $(a_1 + a_2 + a_3)/3$ by Theorem 3.4(b). It is known that, in this case, two of the eigenvalues of A , say, a_1 and a_2 are equal to the center of $W(A)$ (cf. [10, Corollary 2.5]). Hence $a_1 = a_2 = (a_1 + a_2 + a_3)/3$, from which we obtain $a_1 = a_2 = a_3 \equiv a$. Thus $A - aI_3$ is nilpotent and hence is unitarily similar to a matrix A' of the form

$$\begin{bmatrix} 0 & b & d \\ & 0 & c \\ & & 0 \end{bmatrix}.$$

Since $W(A') = W(A) - a$ is a circular disc, [10, Theorem 4.1] implies that $bcd = 0$. If b or c equals 0, then $\text{rank } A' = 1$. In this case, A' has the matrix representation $A_1 \oplus A_2$ on the decomposition $\mathbb{C}^3 = K \oplus K^\perp$, where K is the span of the ranges of A' and A'^* with $1 \leq \dim K \leq 2$. This contradicts our assumption on the unitary irreducibility of A . We conclude that b and c are both nonzero. Thus $d = 0$ and hence A is of the form (iii) as required.

For the converse, if A is unitarily reducible, then it is unitarily similar to either

$$B = \begin{bmatrix} a & & \\ & b & \\ & & c \end{bmatrix} \text{ or } C = \begin{bmatrix} a & b & \\ & d & \\ & & c \end{bmatrix}.$$

In the former case, since $W(A) = W(B) = G(B)$, we have $W(A) = G'(A)$. For the latter, if $a \neq d$ and $b \neq 0$, then

$$W(A) = W(C) = \left(W \left(\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \right) \cup \{c\} \right)^\wedge.$$

In this case, the boundary of $W(A)$ contains a (noncircular) elliptic arc, and hence $W(A) \subsetneq G'(A)$ by Theorem 3.4(a). Thus we must have $a = d$. Then

$$W(A) = W(C) = G(C) \supseteq G'(C) = G'(A).$$

Hence $W(A) = G'(A)$.

It remains to show that if

$$A = \begin{bmatrix} a & b & \\ & a & c \\ & & a \end{bmatrix},$$

where $|b| = |c| \neq 0$, then $W(A) = G'(A)$. We already know that $W(A) \subseteq G'(A)$. To prove the converse containment, we may assume, by considering $U^*((A - aI_3)/|b|)U$, where $U = \text{diag}(1, e^{-i \arg b}, e^{-i(\arg b + \arg c)})$, instead of A , that

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

In this case, $W(A)$ is the circular disc with center the origin and radius $1/\sqrt{2}$. Let V be the unitary matrix

$$\begin{bmatrix} \frac{1}{2\sqrt{3}} + \frac{1}{\sqrt{6}}e^{-\frac{7\pi}{12}i} & -\frac{1}{2\sqrt{3}} + \frac{1}{2}e^{\frac{\pi}{4}i} + \frac{1}{2\sqrt{6}}e^{-\frac{7\pi}{12}i} & \frac{1}{2\sqrt{3}} + \frac{1}{2}e^{\frac{\pi}{4}i} - \frac{1}{2\sqrt{6}}e^{-\frac{7\pi}{12}i} \\ -\frac{1}{\sqrt{6}} + \frac{1}{\sqrt{3}}e^{-\frac{7\pi}{12}i} & \frac{1}{\sqrt{6}} + \frac{1}{2\sqrt{3}}e^{-\frac{7\pi}{12}i} & -\frac{1}{\sqrt{6}} - \frac{1}{2\sqrt{3}}e^{-\frac{7\pi}{12}i} \\ \frac{1}{2\sqrt{3}} + \frac{1}{\sqrt{6}}e^{-\frac{7\pi}{12}i} & -\frac{1}{2\sqrt{3}} - \frac{1}{2}e^{\frac{\pi}{4}i} + \frac{1}{2\sqrt{6}}e^{-\frac{7\pi}{12}i} & \frac{1}{2\sqrt{3}} - \frac{1}{2}e^{\frac{\pi}{4}i} - \frac{1}{2\sqrt{6}}e^{-\frac{7\pi}{12}i} \end{bmatrix}.$$

Via some computations, we obtain

$$V^*AV = \begin{bmatrix} \frac{1}{3\sqrt{2}} & \frac{4\sqrt{2}+\sqrt{3}+3}{12} & \frac{-4\sqrt{2}+\sqrt{3}+3}{12} \\ \frac{4\sqrt{2}-\sqrt{3}-3}{12} & -\frac{1}{6\sqrt{2}} - \frac{\sqrt{3}}{4}i & \frac{\sqrt{2}-2\sqrt{3}+3}{12} \\ \frac{-4\sqrt{2}-\sqrt{3}-3}{12} & \frac{\sqrt{2}+2\sqrt{3}-3}{12} & -\frac{1}{6\sqrt{2}} + \frac{\sqrt{3}}{4}i \end{bmatrix}.$$

Its three Geršgorin discs are

$$C_1(V^*AV) = \left\{ z \in \mathbb{C} : \left| z - \frac{1}{3\sqrt{2}} \right| \leq \frac{2\sqrt{2}}{3} \right\},$$

$$C_2(V^*AV) = \left\{ z \in \mathbb{C} : \left| z - \left(-\frac{1}{6\sqrt{2}} - \frac{\sqrt{3}}{4}i \right) \right| \leq \frac{5\sqrt{2}}{12} \right\},$$

and

$$C_3(V^*AV) = \left\{ z \in \mathbb{C} : \left| z - \left(-\frac{1}{6\sqrt{2}} + \frac{\sqrt{3}}{4}i \right) \right| \leq \frac{5\sqrt{2}}{12} \right\}.$$

Therefore, $W(A) \cap \partial G(V^*AV)$ is the singleton $\{-1/\sqrt{2}\}$. If $U_\theta = \text{diag}(1, e^{i\theta}, e^{2i\theta})$ for any real θ , then $U_\theta^*AU_\theta = e^{i\theta}A$. Hence we obtain $W(A) \cap \partial G((U_\theta V)^*A(U_\theta V)) = \{- (1/\sqrt{2}) e^{i\theta}\}$. This shows that

$$G'(A) \subseteq \bigcap_{\theta \in \mathbb{R}} G((U_\theta V)^*A(U_\theta V)) = W(A),$$

and thus $W(A) = G'(A)$ as asserted. \square

We end this section by some remarks. It can be proven that if each A_j , $1 \leq j \leq m$, satisfies $W(A_j) = G'(A_j)$, then $A \equiv A_1 \oplus \dots \oplus A_m$ also satisfies $W(A) = G'(A)$. However, the converse is false. A counterexample is given by

$$A_1 = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & 1 & 1 \\ & 0 & 1 \\ & & & 0 \end{bmatrix}.$$

Here $W(A_1) = G'(A_1) = \{z \in \mathbb{C} : |z| \leq 1\}$. Since A_2 is unitarily irreducible and $\partial W(A_2)$ contains one line segment (cf. [10, Corollary 3.3]), we have $W(A_2) \subsetneq G'(A_2)$ by Theorem 3.4(b). On the other hand, since $G(A_2) = \{z \in \mathbb{C} : |z| \leq 1\}$, we can derive that

$$\begin{aligned} W(A_1 \oplus A_2) &= (W(A_1) \cup W(A_2))^\wedge = (G(A_2) \cup W(A_2))^\wedge = G(A_2) \\ &= (G(A_1) \cup G(A_2))^\wedge = G(A_1 \oplus A_2), \end{aligned}$$

from which it follows that $W(A_1 \oplus A_2) = G'(A_1 \oplus A_2)$. More generally, using Lemma 3.1, we can prove that if matrices A_1 and A_2 are such that $W(A_1) = G'(A_1)$ and $G(A_2) \subseteq W(A_1)$, then $A \equiv A_1 \oplus A_2$ satisfies $W(A) = G'(A)$.

4. Equality of $W(A)$ and $G(A)$

In this section, we consider matrices A with the property $W(A) = G(A)$. Complete characterizations for such matrices are obtained. We also give more specific forms for such matrices of size 2 or 4.

Our first proposition relates the numerical radius $w(A) = \max\{|z| : z \in W(A)\}$ and Geršgorin radius $g(A) = \max\{|z| : z \in G(A)\}$ of a matrix A . Note that if $A = [a_{ij}]_{i,j=1}^n$, then $g(A) = \max_{1 \leq i \leq n} \sum_{j=1}^n (|a_{ij}| + |a_{ji}|)/2$.

Proposition 4.1. *If $A = [a_{ij}]_{i,j=1}^n$ is an n -by- n matrix, then $w(A) \leq g(A)$. Moreover, if $\operatorname{Re} A$ is permutationally irreducible, then $w(A) = g(A)$ if and only if $W(A)$ and all the Geršgorin discs $C_i(A)$ have a common supporting line L at a common point and the a_{ii} 's and the origin are all on a line and on one side of L .*

Proof. Since $W(A) \subseteq G(A)$, we obviously have $w(A) \leq g(A)$. Now assume that $\operatorname{Re} A$ is permutationally irreducible and $w(A) = g(A)$. Let z_0 in $W(A)$ be such that $|z_0| = w(A)$, and let L be a supporting line of $W(A)$ at z_0 which is tangent to the circle $|z| = w(A)$. Since $|z_0| = g(A)$ by our assumption and $W(A) \subseteq G(A) \subseteq D \equiv \{z \in \mathbb{C} : |z| \leq g(A)\}$, we infer that L is a common supporting line of $W(A)$ and $G(A)$ at z_0 . After an affine transform of A , we may assume that $z_0 = 0$, L is the y -axis and $G(A) \subseteq \{z \in \mathbb{C} : \operatorname{Re} z \geq 0\}$. Since $\operatorname{Re} A$ is permutationally irreducible, in the decomposition of A in Theorem 2.4 we have $A = A_1$. Hence (a) there, together with $C_i(A) \subseteq D$ for all i , yields that the $C_i(A)$'s are all tangent to L at the common point z_0 , and the centers a_{ii} of the $C_i(A)$'s and z_0 are on a common line and on the same side of L (cf. Figure 4). This proves the necessity. The sufficiency is trivial. \square

The main results of this section are the following two theorems.

Theorem 4.2. *Let $A = [a_{ij}]_{i,j=1}^n$ be an n -by- n ($n \geq 2$) matrix with $\operatorname{Re} A$ permutationally irreducible. Then the following conditions are equivalent:*

- (a) $W(A) = G(A)$;
- (b) $W(A)$ is a circular disc, which coincides with all the Geršgorin discs $C_i(A)$;

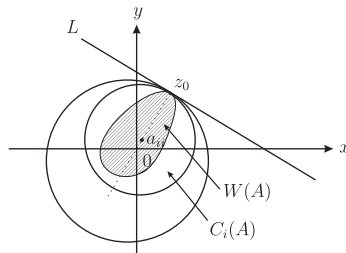


Fig. 4. $C_i(A)$'s with a common supporting line L .

(c) there are a permutation matrix P and a diagonal unitary U such that $B \equiv (PU)^*A(PU) = [b_{ij}]_{i,j=1}^n$ is of the form

$$\begin{bmatrix} \left(\frac{1}{n}\text{tr}A\right) I_1 & B_1 & & & \\ & \left(\frac{1}{n}\text{tr}A\right) I_2 & B_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & B_{m-1} \\ & & & & \left(\frac{1}{n}\text{tr}A\right) I_m \end{bmatrix} \quad (m \geq 2) \tag{5}$$

with $b_{ij} \geq 0$ for all $i \neq j$ and $g_i(B)$'s constant.

In this case, a_{ii} and $g_i(A)$, $1 \leq i \leq n$, are both constants, and n is even.

Proof. By considering $A - ((1/n)\text{tr}A)I_n$, we may assume that $\text{tr}A = 0$. The implication (a) \Rightarrow (b) follows from Proposition 2.5(b). To prove (b) \Rightarrow (c), note that (b) implies that $a_{ii} = 0$ for all i and the $g_i(A)$'s are equal to one another. Now apply Theorem 2.4 (f) to $g_1(A)I_n - A$ to obtain a permutation matrix P_1 and a diagonal unitary V such that $C \equiv (P_1V)^*A(P_1V) = [c_{ij}]_{i,j=1}^n$ satisfies $c_{ii} = 0$ for all i and $c_{ij} \geq 0$ for all $i \neq j$. Since $W(C) = W(A)$ is a circular disc centered at the origin and $\text{Re } C$, as $\text{Re } A$, is permutationally irreducible, [13, Theorem 1] says that there is another permutation matrix P_2 such that $B \equiv P_2^T C P_2$ is of the form (5). Letting $P = P_1 P_2$ and $U = P_2^T V P_2$, we obtain $B = (PU)^*A(PU)$ as asserted.

We now prove (c) \Rightarrow (a). Under condition (c), we have $g_i(B) = g_i(A) = g(A)$ for all i and hence $G(A) = G(B) = \{z \in \mathbb{C} : |z| \leq g(A)\}$. Therefore, $W(A) = W(B) \subseteq \{z \in \mathbb{C} : |z| \leq g(A)\}$. On the other hand, since B is unitarily similar to $e^{i\theta}B$ for all real θ , its numerical range $W(B)$ is also a circular disc centered at the origin. Letting 1_n denote the n -vector whose components are all equal to 1, we have

$$w(A) = w(B) = \|\text{Re } B\| \geq \frac{\|(\text{Re } B)1_n\|}{\|1_n\|} = \frac{\|g(B)1_n\|}{\|1_n\|} = g(B) = g(A).$$

It follows that $w(A) = g(A)$. Hence $W(A) = G(A) = \{z \in \mathbb{C} : |z| \leq g(A)\}$, that is, (a) holds.

Finally, we show that n is even. Assume that B_k , $1 \leq k \leq m - 1$, is an n_k -by- n_{k+1} matrix, and let

$$f_{ik}(B) = \begin{cases} \frac{1}{2} \sum_{j=1}^{n_2} (B_1)_{ij} & \text{if } k = 1 \text{ and } 1 \leq i \leq n_1, \\ \frac{1}{2} \left(\sum_{j=1}^{n_{k-1}} (B_{k-1})_{ji} + \sum_{j=1}^{n_{k+1}} (B_k)_{ij} \right) & \text{if } 2 \leq k \leq m - 1 \text{ and } 1 \leq i \leq n_k, \\ \frac{1}{2} \sum_{j=1}^{n_{m-1}} (B_{m-1})_{ji} & \text{if } k = m \text{ and } 1 \leq i \leq n_m, \end{cases}$$

where, for any matrix D , $(D)_{ij}$ denotes its (i, j) -entry. We have $f_{ik}(B) = g(B)$ for all i and k . Hence $\sum_{i=1}^{n_k} f_{ik}(B) = n_k g(B)$ for all k . Since

$$\begin{aligned} \sum_{k=1}^m (-1)^{k-1} n_k g(B) &= \sum_{k=1}^m (-1)^{k-1} \sum_{i=1}^{n_k} f_{ik}(B) \\ &= \frac{1}{2} \left[\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (B_1)_{ij} - \sum_{i=1}^{n_2} \left(\sum_{j=1}^{n_1} (B_1)_{ji} + \sum_{j=1}^{n_3} (B_2)_{ij} \right) + \dots \right] \end{aligned}$$

$$\begin{aligned}
 &+ (-1)^{m-2} \sum_{i=1}^{n_{m-1}} \left(\sum_{j=1}^{n_{m-2}} (B_{m-2})_{ji} + \sum_{j=1}^{n_m} (B_{m-1})_{ij} \right) + (-1)^{m-1} \sum_{i=1}^{n_m} \sum_{j=1}^{n_{m-1}} (B_{m-1})_{ji} \Big] \\
 &= 0
 \end{aligned}$$

and $g(B) \neq 0$ by the permutational irreducibility of $\text{Re } A$ and $\text{Re } B$, we obtain $\sum_{k=1}^m (-1)^{k-1} n_k = 0$. On the other hand, we also have $\sum_{k=1}^m n_k = n$. Adding these two together yields $n = 2 \sum_{k=1, k \text{ odd}}^m n_k$, which shows that n is even as asserted. \square

Note that if a matrix A is such that $W(A) = G(U^*AU)$ for some unitary U , then, obviously, $W(A) = G'(A)$. The converse is false as the following corollary shows. It is an easy consequence of Proposition 3.6 and Theorem 4.2.

Corollary 4.3. *If*

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

then $W(A) = G'(A)$, but $W(A) \not\subseteq G(U^*AU)$ for any unitary U .

The next theorem deals with general matrices with the property $W(A) = G(A)$.

Theorem 4.4. *Let A be an n -by- n matrix. Then $W(A) = G(A)$ if and only if A is permutationally similar to a matrix of the form $D \oplus A_1 \oplus \dots \oplus A_\ell \oplus B$, where D is a diagonal matrix, each A_k , $1 \leq k \leq \ell$, is such that $\text{Re } A_k$ is permutationally irreducible and satisfies $W(A_k) = G(A_k)$, and B is such that $G(B) \subseteq (W(D) \cup (\cup_{k=1}^\ell W(A_k)))^\wedge$. In this case, $W(A) = W(D \oplus A_1 \oplus \dots \oplus A_\ell)$.*

Here it goes without saying that some of the summands in the above decomposition may be absent. The proof of this theorem is analogous to the one for Theorem 3.4(a).

Proof of Theorem 4.4. Assume first that $W(A) = G(A)$. Then $W(A) = G'(A)$ holds. Hence, by Theorem 3.4, the boundary of $W(A)$ consists of circular arcs and line segments. Let α be one of such arcs. Note that A is permutationally similar to a direct sum $A'_1 \oplus \dots \oplus A'_m$ with each $\text{Re } A'_j$ permutationally irreducible. For any A'_j of size $n_j \geq 2$, let $p_j = (1/n) \text{tr} A'_j$. We now show that one of the p_j 's is the center of the circular arc α . Indeed, assume to the contrary that none of the p_j 's is the center of α . Then Lemma 3.3(b) implies that for any p_j there are only finitely many points q in α with supporting line L_q of $W(A)$ at q perpendicular to the line L_{p_jq} connecting p_j and q . The pigeonhole principle guarantees the existence of a point q_0 in α such that its supporting line L_{q_0} is not perpendicular to L_{p_jq} for any j . Let H be the closed half-plane with boundary L_{q_0} which contains $W(A) = G(A)$. Since $W(A) \cap \partial H \neq \emptyset$, an application of Theorem 2.4 yields some A'_j with all its Geršgorin discs tangent to L_{q_0} at q_0 , and, in particular, with L_{q_0} perpendicular to $L_{p_jq_0}$. This contradicts our choice of q_0 . Hence α is indeed a circular arc with center, say, p_{j_0} . Also, as the above arguments show, α and all the Geršgorin discs $C_i(A'_{j_0})$ of A'_{j_0} are tangent to each other at a common point. We infer from this and the fact that p_{j_0} , the center of α , is the average of the centers of the $C_i(A'_{j_0})$'s that the $C_i(A'_{j_0})$'s must all coincide and α is part of their common boundary. On the other hand, since $W(A'_{j_0}) \cap L_q = \{q\}$ for all points q in α by Theorem 2.4(c), we also have $\alpha \subseteq \partial W(A'_{j_0})$. Hence $C_i(A'_{j_0}) \subseteq W(A'_{j_0})$ by [14, Lemma] or [3, Theorem]. Since $W(A'_{j_0}) \subseteq G(A'_{j_0}) = C_i(A'_{j_0})$, they are equal to each other, that is, $W(A'_{j_0}) = C_i(A'_{j_0})$. Note that the

above goes for every circular arc α of $\partial W(A)$. Hence we may rename all those A_j' 's thus obtained as A_1, \dots, A_ℓ , let D be the direct sum of those A_j' 's which have size one, and let B be the direct sum of the remaining A_j' 's. As

$$G(B) \subseteq G(A) = W(A) = \left(W(D) \cup \left(\bigcup_{k=1}^{\ell} W(A_k) \right) \right)^\wedge,$$

$D \oplus A_1 \oplus \dots \oplus A_\ell \oplus B$ is the asserted decomposition of A .

Conversely, if A is permutationally similar to $D \oplus A_1 \oplus \dots \oplus A_\ell \oplus B$ with the asserted properties, then

$$\begin{aligned} G(A) &= G(D \oplus A_1 \oplus \dots \oplus A_\ell \oplus B) = \left(G(D) \cup \left(\bigcup_{k=1}^{\ell} G(A_k) \right) \cup G(B) \right)^\wedge \\ &\subseteq \left(W(D) \cup \left(\bigcup_{k=1}^{\ell} W(A_k) \right) \right)^\wedge \subseteq W(A). \end{aligned}$$

Since $W(A) \subseteq G(A)$ always holds, we obtain $W(A) = G(A)$. \square

We now characterize matrices A of small sizes for which $W(A)$ and $G(A)$ are equal. Since, by Theorem 4.2, A can only have an even size if this holds, in the following we only consider 2-by-2 and 4-by-4 matrices.

Proposition 4.5. *A 2-by-2 matrix A satisfies $W(A) = G(A)$ if and only if $A = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \begin{bmatrix} a & b \\ 0 & a \end{bmatrix}$ or $\begin{bmatrix} a & 0 \\ b & a \end{bmatrix}$*

for some a and b .

This follows easily from Theorems 4.2 and 4.4.

The next proposition deals with 4-by-4 matrices.

Proposition 4.6. *A 4-by-4 matrix A satisfies $W(A) = G(A)$ if and only if it is permutationally similar to one of the following forms:*

(a) $\text{diag}(a, b, c, d)$,

(b) $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \oplus \begin{bmatrix} c & 0 \\ 0 & d \end{bmatrix}$,

(c) $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \oplus \begin{bmatrix} c & d \\ 0 & c \end{bmatrix}$,

(d) $\begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \oplus B$, where B is a 2-by-2 matrix with $G(B) \subseteq \{z \in \mathbb{C} : |z - a| \leq |b|/2\}$,

(e) $\begin{bmatrix} \lambda & 0 & a & b \\ & \lambda & c & d \\ & & \lambda & 0 \\ & & & \lambda \end{bmatrix}$ with $|a| = |d| \neq 0, |b| = |c| \neq 0$, and $\arg(ad) = \arg(bc)$, and

$$(f) \begin{bmatrix} \lambda & a & b & 0 \\ & \lambda & 0 & c \\ & & \lambda & d \\ & & & \lambda \end{bmatrix} \text{ with } |a| = |d| \neq 0, |b| = |c| \neq 0, \text{ and } \arg(ac) = \arg(bd).$$

Proof. In view of Theorem 4.4 and Proposition 4.5, we need only prove, for a 4-by-4 matrix A with $\text{Re } A$ permutationally irreducible, that $W(A) = G(A)$ if and only if A is permutationally similar to a matrix of the form (e) or (f). Assume that $W(A) = G(A)$, $\text{Re } A$ is permutationally irreducible, and $\text{tr} A = 0$. Then, by Theorem 4.2(c), there are a permutation matrix P and a diagonal unitary matrix U such that $B \equiv (PU)^* A (PU) = [b_{ij}]_{i,j=1}^4$ is of the form (5) with $b_{ii} = 0$ for all i , $b_{ij} \geq 0$ for $i \neq j$ and $g_i(B)$'s constant. It is easily seen that if

$$B = \begin{bmatrix} 0 & a & & \\ & 0 & b & \\ & & 0 & c \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & a & & \\ & 0 & 0 & b & c \\ & & 0 & 0 \\ & & & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & a & \\ & 0 & 0 & b \\ & & 0 & c \\ & & & 0 \end{bmatrix}, \begin{bmatrix} 0 & a & b & c \\ & 0 & 0 & 0 \\ & & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \end{bmatrix}, \text{ or } \begin{bmatrix} 0 & 0 & 0 & a \\ & 0 & 0 & 0 & b \\ & & 0 & 0 & 0 & c \\ & & & 0 & 0 & 0 \end{bmatrix},$$

then the constancy of the $g_i(B)$'s implies $B = 0$, contradicting the permutational irreducibility of $\text{Re } B$. This leaves only the cases

$$B = \begin{bmatrix} 0 & 0 & a & b \\ & 0 & 0 & c & d \\ & & 0 & 0 \\ & & & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & a & b & 0 \\ & 0 & 0 & c \\ & & 0 & 0 & d \\ & & & 0 \end{bmatrix}.$$

In either one, the constancy of the $g_i(B)$'s yields that $a = d \geq 0$ and $b = c \geq 0$. If any of them is zero, then $\text{Re } B$, and hence $\text{Re } A$, is permutationally reducible, again a contradiction. Let $U = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}, e^{i\theta_4})$ with real θ_j 's. Then

$$P^T A P = U B U^* = \begin{bmatrix} 0 & 0 & a e^{i(\theta_1 - \theta_3)} & b e^{i(\theta_1 - \theta_4)} \\ 0 & 0 & b e^{i(\theta_2 - \theta_3)} & a e^{i(\theta_2 - \theta_4)} \\ & & 0 & 0 \\ & & & 0 & 0 \end{bmatrix} \text{ or } \begin{bmatrix} 0 & a e^{i(\theta_1 - \theta_2)} & b e^{i(\theta_1 - \theta_3)} & 0 \\ & 0 & 0 & b e^{i(\theta_2 - \theta_4)} \\ & & 0 & a e^{i(\theta_3 - \theta_4)} \\ & & & 0 \end{bmatrix},$$

which shows that A is permutationally similar to a matrix of the form (e) or (f).

Conversely, if A is of the form (e) with $\lambda = 0$, then, letting $U = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}, e^{i\theta_4})$, where $\theta_1 = 0, \theta_2 = \arg a - \arg c = \arg b - \arg d, \theta_3 = \arg a$ and $\theta_4 = \arg b$, we have

$$U A U^* = \begin{bmatrix} 0 & 0 & |a| & |b| \\ 0 & 0 & |c| & |d| \\ & & 0 & 0 \\ & & & 0 & 0 \end{bmatrix}.$$

Thus $W(A) = G(A)$ by Theorem 4.2. Similarly, if A is of the form (f) with $\lambda = 0$, then $U = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3}, e^{i\theta_4})$, where $\theta_1 = 0, \theta_2 = \arg a, \theta_3 = \arg b$ and $\theta_4 = \arg a + \arg c = \arg b + \arg d$, is such that

$$UAU^* = \begin{bmatrix} 0 & |a| & |b| & 0 \\ & 0 & |c| & \\ & & 0 & |d| \\ & & & 0 \end{bmatrix}.$$

Again, we have $W(A) = G(A)$ by Theorem 4.2. \square

Our final example shows that, in the situation of Theorem 4.2, condition (c) there cannot be relaxed to requiring that A be permutationally similar to a matrix B of the form (5) with $g_i(B)$'s constant.

Example 4.7. Let

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

Then $G(A) = \{z \in \mathbb{C} : |z| \leq 1\}$. Since $A^2 = 0$ and $\left\| \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right\| = \sqrt{2}$, we have $W(A) = \{z \in \mathbb{C} : |z| \leq \sqrt{2}/2\}$ (see [15, Theorem 2.1]). Thus $W(A) \subsetneq G(A)$. A similar example of the form in Proposition 4.6(f) can also be constructed to this effect.

We conclude this paper by asking the following questions:

- (a) Is it true that $W(A) \subsetneq G'(A)$ for

$$A = \begin{bmatrix} 1 & b \\ & 1 & c \\ & & 1 \end{bmatrix}$$

with $b, c > 0$ and $b \neq c$? If this is the case, then it yields, together with Proposition 3.6, a complete characterization of 3-by-3 matrices A satisfying $W(A) = G'(A)$.

- (b) Does the n -by- n Jordan block

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

satisfy $W(J_n) = G'(J_n)$ for $n \geq 4$? Note that this is indeed the case for $n = 2$ (Proposition 3.5) and $n = 3$ (Proposition 3.6).

- (c) Is there some general criterion of A for which $W(A) = G'(A)$ holds, just as Theorems 4.2 and 4.4 for $W(A) = G(A)$? In particular, is there a complete characterization for 4-by-4 matrices satisfying $W(A) = G'(A)$?

Acknowledgements

The authors thank the referee of this paper for his detailed comments, which lead to improvements in both the contents and the presentation.

References

- [1] N. Bourbaki, *Fonctions d'une Variable Réelle*, Hermann, Paris, 1958.
- [2] J.S. Fang, C.-L. Jiang, P.Y. Wu, Direct sums of irreducible operators, *Studia Math.* 155 (2003) 37–49.
- [3] H.-L. Gau, P.Y. Wu, Condition for the numerical range to contain an elliptic disc, *Linear Algebra Appl.* 364 (2003) 213–222.
- [4] S. Geršgorin, Über die Abgrenzung der Eigenwerte einer Matrix, *Izv. Akad. Nauk. S.S.S.R.* 7 (1931) 749–754.
- [5] K. Gustafson, D.K.M. Rao, Numerical Range, The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.
- [6] P.R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Springer, New York, 1982.
- [7] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1985.
- [8] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1991.
- [9] C.R. Johnson, A Geršgorin inclusion set for the field of values of a finite matrix, *Proc. Amer. Math. Soc.* 41 (1973) 57–60.
- [10] D.S. Keeler, L. Rodman, I.M. Spitkovsky, The numerical range of 3×3 matrices, *Linear Algebra Appl.* 252 (1997) 115–139.
- [11] R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.* 6 (1951) 193–228. English translation: P.F. Zschlin, M.E. Hochstenbach, On the numerical range of a matrix, *Linear Multilinear Algebra* 56 (2008) 185–225.
- [12] F. Kirwan, *Complex Algebraic Curves*, Cambridge Univ. Press, Cambridge, 1992.
- [13] B.-S. Tam, S. Yang, On matrices whose numerical ranges have circular or weak circular symmetry, *Linear Algebra Appl.* 302/303 (1999) 193–221.
- [14] R.C. Thompson, The upper numerical range of a quaternionic matrix is not a complex numerical range, *Linear Algebra Appl.* 254 (1997) 19–28.
- [15] S.-H. Tso, P.Y. Wu, Matricial ranges of quadratic operators, *Rocky Mountain J. Math.* 29 (1999) 1139–1152.
- [16] F.A. Valentine, *Convex Sets*, McGraw-Hill, New York, 1964.
- [17] A. Zalewska-Mitura, J. Zemánek, The Gerschgorin discs under unitary similarity, in: *Linear Operators (Warsaw, 1994)*, vol. 38, Banach Center Pub., Warsaw, Poland, 1997, pp. 427–441.