Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 429 (2008) 716–726
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Numerical ranges of nilpotent operators

Hwa-Long Gau [a,*,1], Pei Yuan Wu [b,2]

[a] Department of Mathematics, National Central University, Chungli 32001, Taiwan
[b] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 300, Taiwan

Received 1 February 2008; accepted 31 March 2008
Available online 16 May 2008
Submitted by C.-K. Li

Abstract

For any operator A on a Hilbert space, let $W(A)$, $w(A)$ and $w_0(A)$ denote its numerical range, numerical radius and the distance from the origin to the boundary of its numerical range, respectively. We prove that if $A^n = 0$, then $w(A) \leqslant (n-1)w_0(A)$, and, moreover, if A attains its numerical radius, then the following are equivalent: (1) $w(A) = (n-1)w_0(A)$, (2) A is unitarily equivalent to an operator of the form $aA_n \oplus A'$, where a is a scalar satisfying $|a| = 2w_0(A)$, A_n is the n-by-n matrix

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ & 0 & \ddots & \vdots \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

and A' is some other operator, and (3) $W(A) = bW(A_n)$ for some scalar b.
© 2008 Elsevier Inc. All rights reserved.

AMS classification: 47A12; 15A60

Keywords: Numerical range; Numerical radius; Nilpotent operator

* Corresponding author.
E-mail addresses: hlgau@math.ncu.edu.tw (H.-L. Gau), pywu@math.nctu.edu.tw (P.Y. Wu).
[1] Research supported by the National Science Council of the Republic of China under NSC 96-2115-M-008-006.
[2] Research supported by the National Science Council of the Republic of China under NSC 96-2115-M-009-013-MY3 and by the MOE-ATU.

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.03.029

0. Introduction

Let A be a bounded linear operator on a complex Hilbert space H. The *numerical range* and *numerical radius* 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 associated norm in H. It is known that $W(A)$ is a bounded convex subset of the plane. When H is finite dimensional, it is even compact. Its closure $\overline{W(A)}$ contains $\sigma(A)$, the spectrum of A. For its other properties, the reader may consult [7, Chapter 22] or [5].

Assume that A is *nilpotent*, say, $A^n = 0$ for some $n \geqslant 1$. Then 0 is in $\sigma(A)$ and hence in $\overline{W(A)}$. In the following, we will show that if A is nonzero, then 0 is actually in the interior of $W(A)$ (cf. Corollary 1.2). To estimate the extent to which 0 belongs to this interior, we define the quantity

$$w_0(A) \equiv \operatorname{dist}(0, \partial W(A)) = \inf\{|z| : z \in \partial W(A)\}.$$

Thus $w_0(A)$ is the radius of the largest closed circular disc centered at the origin which is contained in $\overline{W(A)}$. We show in Theorem 1.1 that the inequality $w(A) \leqslant (n-1)w_0(A)$ holds and, moreover, if A *attains its numerical radius*, that is, if there is a unit vector x in H with $|\langle Ax, x\rangle| = w(A)$, then the following conditions are equivalent: (1) $w(A) = (n-1)w_0(A)$, (2) A is unitarily equivalent to an operator of the form $aA_n \oplus A'$, where a is a scalar satisfying $|a| = 2w_0(A)$, A_n is the n-by-n matrix

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ & 0 & \ddots & \vdots \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}$$

and A' is some other operator, and (3) $W(A) = bW(A_n)$ for some scalar b. This shows the prominent role the matrix A_n plays in this regard. We investigate its numerical range detailedly in Lemma 1.3. For example, it is shown that $w(A_n) = (n-1)/2$, $w_0(A_n) = 1/2$ (for $n \geqslant 2$), and the boundary of $W(A_n)$ contains one line segment (for $n \geqslant 3$). In Section 2, we consider an analogous inequality by Caston et al. [1, Lemma 2.6] in the opposite situation when 0 is not in the interior of the numerical range of an n-by-n matrix A. In this case, the inequality $w(A) \leqslant nr(A)$ holds, where $r(A)$ is the spectral radius of A. This, together with a characterization of those A's for which $w(A) = nr(A)$ is true, is given in Theorem 2.1.

Results in this paper are analogous to the ones in [6,10] relating the numerical radius and norm of a nilpotent operator. Indeed, it is shown therein that if A is nilpotent with $A^n = 0$ ($n \geqslant 1$), then $w(A) \leqslant \|A\| \cos(\pi/(n+1))$ and, moreover, if A attains its numerical radius and $w(A) = \|A\| \cos(\pi/(n+1))$, then A is unitarily equivalent to $\|A\| J_n \oplus A'$ for some operator A', where J_n denotes the n-by-n Jordan block

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

In our situation, the role of J_n is played by A_n and $2A_n + I_n$ in Theorems 1.1 and 2.1, respectively.

1. Nilpotent operators

We start with the following main result of this paper.

Theorem 1.1. *Let A be an operator on H with $A^n = 0$ ($n \geqslant 1$).*

(a) *The inequality $w(A) \leqslant (n-1)w_0(A)$ holds.*
(b) *Assume that A attains its numerical radius. Then the following conditions are equivalent*:

(1) $w(A) = (n-1)w_0(A)$,
(2) *A is unitarily equivalent to an operator of the form $aA_n \oplus A'$, where a is a scalar satisfying $|a| = 2w_0(A)$ and A' is some other operator, and*
(3) *$W(A) = bW(A_n)$ for some scalar b.*

An easy consequence of the preceding theorem is the following:

Corollary 1.2. *If A is a nonzero nilpotent operator, then 0 belongs to the interior of $W(A)$ and $\partial W(A)$ is a differentiable curve.*

Proof. For the nonzero nilpotent A, we have $w_0(A) > 0$ by Theorem 1.1(a) and hence 0 is in the interior of $W(A)$. On the other hand, since any nondifferentiable point λ on $\partial W(A)$ must be in the approximate point spectrum of A by [11, Theorem 2] and the latter is the singleton $\{0\}$ for the nilpotent A, we have $\lambda = 0$, which contradicts $w_0(A) > 0$. Hence $\partial W(A)$ must be differentiable. □

We remark that the preceding corollary is not necessarily true for a nonzero quasinilpotent operator A ($\sigma(A) = \{0\}$). Indeed, if A is the Volterra operator

$$(Af)(x) = \int_0^x f(t)dt \quad \text{for } f \text{ in } L^2(0,1),$$

then A is quasinilpotent and 0 is in the boundary of $W(A)$ (cf. [7, pp. 98–99]).

To prove Theorem 1.1, we need a fuller understanding of the numerical range of A_n. This is provided by the next lemma.

Lemma 1.3. (1) *$W(A_n)$ is symmetric with respect to the x-axis.*
(2) $w(A_n) = \langle A_n x_0, x_0\rangle = (n-1)/2$, *where* $x_0 = [1/\sqrt{n} \ \cdots \ 1/\sqrt{n}]^{\mathrm{T}}$.
(3) $(-1/2)I_n \leqslant \operatorname{Re} A_n \leqslant ((n-1)/2)I_n$.
(4) *If $n \geqslant 3$, then $\partial W(A_n)$ has exactly one line segment, which is on the line $x = -1/2$.*
(5) *For any point λ of $\partial W(A_n)$ which is not on its line segment, the set $\{x \in \mathbb{C}^n : \langle A_n x, x\rangle = \lambda\|x\|^2\}$ is a subspace of dimension one.*
(6) *$w_0(A_n) = 1/2$ for $n \geqslant 2$.*
(7) *A point λ with $|\lambda| = 1/2$ is in $\partial W(A_n)$ ($n \geqslant 3$) if and only if $\lambda = -1/2$.*
(8) *The polynomial $p_{A_n}(x, y, z) = \det(x\operatorname{Re} A_n + y\operatorname{Im} A_n + zI_n)$ is irreducible.*

Here $\text{Re}\, X = (X + X^*)/2$ and $\text{Im} X = (X - X^*)/(2\text{i})$ denote the *real* and *imaginary parts* of a matrix X, respectively.

Note that p_{A_n} and $W(A_n)$ are related by a result of Kippenhahn [8], the numerical range of an n-by-n matrix A equals the convex hull of the real points $(u/w, v/w)$ of the dual $\{[u, v, w] \in \mathbb{CP}^2 : ux + vy + wz = 0$ is a tangent line of $p_A(x, y, z) = 0\}$ of the curve $p_A = 0$ in the complex projective plane $\mathbb{CP}^2$.

To prove (4) of the preceding lemma, we need another result. For any n-by-n matrix A and real θ, let L be the supporting line of the convex set $W(A)$ perpendicular to the ray which emanates from the origin and forms angle θ from the positive x-axis, and let d be the (signed) distance from the origin to L. It is easily seen that d is the largest eigenvalue of $\text{Re}\,(\text{e}^{-\text{i}\theta}A)$, and a unit vector in $\mathbb{C}^n$ is such that $\langle Ax, x\rangle$ belongs to $\partial W(A) \cap L$ if and only if $\text{Re}\,(\text{e}^{-\text{i}\theta}A)x = dx$. From this, we obtain the following lemma immediately.

Lemma 1.4. *Let A be an n-by-n matrix. Then $\partial W(A)$ has a line segment on $x\cos\theta + y\sin\theta = d$ if and only if d is the largest eigenvalue of* $\text{Re}\,(\text{e}^{-\text{i}\theta}A)$ *which has unit eigenvectors x_1 and x_2 such that* $\text{Im}\langle \text{e}^{-\text{i}\theta}Ax_1, x_1\rangle \neq \text{Im}\langle \text{e}^{-\text{i}\theta}Ax_2, x_2\rangle$.

Proof of Lemma 1.3. (1) This follows easily from the fact that A_n is a real matrix.

(2) Since

$$\begin{aligned}\langle A_n x_0, x_0\rangle &= (n-1)\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} + (n-2)\frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} + \cdots + \frac{1}{\sqrt{n}}\frac{1}{\sqrt{n}} \\ &= \frac{1}{n}[(n-1) + (n-2) + \cdots + 1] \\ &= \frac{1}{2}(n-1),\end{aligned}$$

we have $w(A_n) \geqslant (n-1)/2$. On the other hand, for any unit vector $x = [x_1 \ \cdots \ x_n]^{\text{T}}$, we deduce from

$$|\langle A_n x, x\rangle| = \left|\sum_{i<j} x_j \overline{x_i}\right| \leqslant \sum_{i<j} |x_i||x_j|$$

and

$$\begin{aligned}0 &\leqslant \sum_{i<j}(|x_i| - |x_j|)^2 = (n-1)\sum_i |x_i|^2 - 2\sum_{i<j}|x_i||x_j| \\ &= n - 1 - 2\sum_{i<j}|x_i||x_j|\end{aligned}$$

that $\sum_{i<j}|x_i||x_j| \leqslant (n-1)/2$. Hence $w(A_n) \leqslant (n-1)/2$ and thus $w(A_n) = (n-1)/2$ as asserted.

(3) We first show that $-1/2$ is the smallest eigenvalue of $\text{Re}\, A_n$. Indeed, since

$$\text{Re}\, A_n + \frac{1}{2}I_n = \begin{bmatrix} 1/2 & \cdots & 1/2 \\ \vdots & & \vdots \\ 1/2 & \cdots & 1/2 \end{bmatrix}$$

is of rank one, $-1/2$ is an eigenvalue of Re A_n of multiplicity $n-1$. To prove that it is the smallest one, let $t < -1/2$ and let $x = [x_1 \ \cdots \ x_n]^{\mathrm{T}}$ be a unit vector in $\mathbb{C}^n$ such that (Re $A_n)x = tx$. A simple computation shows that $(1/2)\sum_{j=1}^n x_j = (t+(1/2))x_k$ for $1 \leqslant k \leqslant n$. Since $t+(1/2) < 0$, we obtain $x_1 = \cdots = x_n$. These equalities then yield $nx_k/2 = (t+(1/2))x_k$ or $(t+(1/2)-(n/2))x_k = 0$. Hence $x_k = 0$ for all k. This shows that no such t can be an eigenvalue of Re A_n. Thus $(-1/2)I_n \leqslant$ Re A_n. On the other hand, since $w(\text{Re}\, A_n) \leqslant w(A_n) = (n-2)/2$ by (2), the inequality Re $A_n \leqslant ((n-2)/2)I_n$ follows.

(4) Note that $x_\pm = [1/2 \ \pm \mathrm{i}/2 \ -(1\pm \mathrm{i})/2 \ 0 \ \cdots \ 0]^{\mathrm{T}}$ are unit eigenvectors of Re A_n for the smallest eigenvalue $-1/2$ with $\langle A_n x_\pm, x_\pm\rangle = (-1/2) \pm (1/4)\mathrm{i}$. Lemma 1.4 implies that $\partial W(A_n)$ contains a line segment on $x = -1/2$.

We next show that, for every real θ with $\mathrm{e}^{-\mathrm{i}\theta} \neq -1$, the largest eigenvalue d of Re $(\mathrm{e}^{-\mathrm{i}\theta} A_n)$ has multiplicity one. Indeed, if $\dim \ker(dI_n - \text{Re}(\mathrm{e}^{-\mathrm{i}\theta}A_n)) \geqslant 2$, then there is a nonzero vector x of the form $[x_1 \ \cdots \ x_{n-1} \ 0]^{\mathrm{T}}$ in $\ker(dI_n - \text{Re}(\mathrm{e}^{-\mathrm{i}\theta}A_n))$. A simple computation shows that

$$dx_k - \frac{1}{2}\mathrm{e}^{-\mathrm{i}\theta}\sum_{j>k} x_j - \frac{1}{2}\mathrm{e}^{\mathrm{i}\theta}\sum_{j<k} x_j = 0 \quad \text{for } 1 \leqslant k \leqslant n-1 \tag{i}$$

and

$$\frac{1}{2}\mathrm{e}^{\mathrm{i}\theta}\sum_{j=1}^{n-1} x_j = 0.$$

The latter yields $\sum_j x_j = 0$, which, together with the former gives, for $k = 1$,

$$dx_1 - \frac{1}{2}\mathrm{e}^{-\mathrm{i}\theta}\sum_{j=2}^{n-1} x_j = \left(d + \frac{1}{2}\mathrm{e}^{-\mathrm{i}\theta}\right)x_1 = 0.$$

Since Re $\begin{bmatrix} 0 & \mathrm{e}^{-\mathrm{i}\theta} \\ 0 & 0 \end{bmatrix}$ is a submatrix of Re $(\mathrm{e}^{-\mathrm{i}\theta}A_n)$, its maximum eigenvalue $1/2$ is less than or equal to the maximum eigenvalue d of Re $(\mathrm{e}^{-\mathrm{i}\theta}A_n)$. Thus, in particular, $d + (\mathrm{e}^{-\mathrm{i}\theta}/2) \neq 0$. We obtain from above that $x_1 = 0$. The equality in (i) for $k = 2$ then gives

$$dx_2 - \frac{1}{2}\mathrm{e}^{-\mathrm{i}\theta}\sum_{j=3}^{n-1} x_j = \left(d + \frac{1}{2}\mathrm{e}^{-\mathrm{i}\theta}\right)x_2 = 0$$

and hence $x_2 = 0$. Proceeding successively with the remaining equalities in (i), we obtain $x_k = 0$ for all k, which contradicts our assumption on x. This proves our assertion. In particular, it follows from Lemma 1.4 that $\partial W(A_n)$ has no line segment other than the one on $x = -1/2$.

(5) We may assume that $n \geqslant 3$. By the remarks preceding Lemma 1.4, for the given λ, there is a real θ for which Re $(\mathrm{e}^{-\mathrm{i}\theta}\lambda)$ is the largest eigenvalue of Re $(\mathrm{e}^{-\mathrm{i}\theta}A_n)$ whose eigenvectors x are exactly those satisfying $\langle A_n x, x\rangle = \lambda\|x\|^2$. Hence $\{x \in \mathbb{C}^n : \langle A_n x, x\rangle = \lambda\|x\|^2\}$ is a subspace of dimension one from the proof of (4).

(6) We need only consider $n \geqslant 3$. In this case, the boundary of $W(A_n)$ contains a line segment on $x = -1/2$ by (4). Hence, obviously, $w_0(A_n) \leqslant 1/2$. On the other hand, since $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ is a submatrix of A_n, we have

$$W\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = \left\{z \in \mathbb{C} : |z| \leqslant \frac{1}{2}\right\} \subseteq W(A_n).$$

Therefore, $w_0(A_n) \geqslant 1/2$. Our assertion follows.

(7) Let λ in $\partial W(A_n)$ be such that $|\lambda| = 1/2$. Since

$$\left\{z \in \mathbb{C} : |z| \leqslant \frac{1}{2}\right\} = W(A_2) \subseteq W(A_3) \subseteq W(A_n),$$

the point λ is also in $\partial W(A_3)$. Hence if $\theta = \arg \lambda$, then $\operatorname{Re}(\mathrm{e}^{-\mathrm{i}\theta}\lambda) = 1/2$ is the largest eigenvalue of $\operatorname{Re}(\mathrm{e}^{-\mathrm{i}\theta}A_3)$. Therefore,

$$\begin{aligned} 0 &= \det \begin{bmatrix} 1/2 & -\mathrm{e}^{-\mathrm{i}\theta}/2 & -\mathrm{e}^{-\mathrm{i}\theta}/2 \\ -\mathrm{e}^{\mathrm{i}\theta}/2 & 1/2 & -\mathrm{e}^{-\mathrm{i}\theta}/2 \\ -\mathrm{e}^{\mathrm{i}\theta}/2 & -\mathrm{e}^{\mathrm{i}\theta}/2 & 1/2 \end{bmatrix} \\ &= -(1/8)(2 + \mathrm{e}^{\mathrm{i}\theta} + \mathrm{e}^{-\mathrm{i}\theta}) \\ &= -(1 + \operatorname{Re} \mathrm{e}^{\mathrm{i}\theta})/4, \end{aligned}$$

which yields $\mathrm{e}^{\mathrm{i}\theta} = -1$ or $\lambda = \mathrm{e}^{\mathrm{i}\theta}/2 = -1/2$.

(8) Assume that $p_{A_n}(x, y, z) = \prod_{j=1}^{m} p_j(x, y, z)$, where the p_j's are irreducible homogeneous polynomials in x, y and z with real coefficients. As shown in (1), the eigenvalues of $\operatorname{Re} A_n$ are $(n-1)/2$ and $-1/2$, the latter with multiplicity $n-1$. Hence the roots of $p_{A_n}(1, 0, z) = 0$ are $-(n-1)/2$ and $1/2$ (of multiplicity $n-1$). We may assume, for convenience, that $p_1(1, 0, -(n-1)/2) = 0$ and let $q = \prod_{j=2}^{m} p_j$. Then the only root of $q(1, 0, z) = 0$ is 1/2. Hence if Δ is the convex hull of the real points of the dual of the curve $q(x, y, z) = 0$, then by duality the only vertical supporting line of Δ is $x = -1/2$. This implies that Δ can only be contained in the line $x = -1/2$. Dually, this says that $q(x, y, z) = \prod_j (-(1/2)x + a_j y + z)$ for some real a_j's. From

$$\begin{aligned} z^n = p_{A_n}(1, i, z) &= p_1(1, i, z) q(1, i, z) \\ &= p_1(1, i, z) \prod_{j=1}^{m} (-(1/2) + a_j i + z), \end{aligned}$$

we reach a contradiction. This proves the irreducibility of p_{A_n}. □

Some of the assertions in Lemma 1.3 can be proved from results in [2, Section 3,3, Section 3] on properties of p_A and $W(A)$ for more general nilpotent Toeplitz matrices A. In particular, the irreducibility of p_{A_n} in (8) follows from [2, Theorem 3.2] and the existence of a line segment on $\partial W(A_n)$ in (4) from [3, Theorem 3]. Our proofs are more direct.

Another result which we need is the following:

Proposition 1.5. *Assume that* $A^n = 0$ $(n \geqslant 1)$. *Then* $W(A) \subseteq W(A_n)$ *if and only if* $\operatorname{Re} A \geqslant (-1/2)I$.

The proof of this is based on the one for the corresponding numerical range containment for a nilpotent contraction and the Jordan block.

Lemma 1.6. *The operator A on H is a contraction with* $A^n = 0$ $(n \geqslant 1)$ *if and only if A is power dilated to* $J_n \oplus \cdots \oplus J_n$, *that is, there is an isometry V from H to* $\mathbb{C}^n \oplus \cdots \oplus \mathbb{C}^n$ *such that* $A^k = V^*(J_n \oplus \cdots \oplus J_n)^k V$ *for all* $k \geqslant 1$, *where the number of summands is* $\dim \overline{\operatorname{ran}(I - A^*A)}$.

This can be proved by modifying the arguments for [7, Solution 152]. It was essentially reproduced in [10].

Proof of Proposition 1.5. Since Re $A_n \geqslant (-1/2)I_n$ as proved in Lemma 1.3(3), the necessity of our assertion is obvious. For the converse, assume that Re $A \geqslant (-1/2)I$ and let $B = A(I+A)^{-1}$. For any vector x, if $y = (I+A)^{-1}x$, then we have

$$\begin{aligned}\|x\|^2 - \|Bx\|^2 &= \langle x, x\rangle - \langle A(I+A)^{-1}x, A(I+A)^{-1}x\rangle \\ &= \langle (I+A)y, (I+A)y\rangle - \langle Ay, Ay\rangle \\ &= \langle y, y\rangle + \langle Ay, y\rangle + \langle y, Ay\rangle \\ &= \langle (I + 2\mathrm{Re}\, A)y, y\rangle \geqslant 0,\end{aligned}$$

showing that B is a contraction. We also have $B^n = 0$. Lemma 1.6 then yields the power dilation of B to $J_n \oplus \cdots \oplus J_n$. Since

$$A = B(I-B)^{-1} = B + B^2 + \cdots + B^{n-1}$$

and

$$A_n = J_n + J_n^2 + \cdots + J_n^{n-1},$$

we obtain the dilation of A to $A_n \oplus \cdots \oplus A_n$. Thus $W(A) \subseteq W(A_n)$ as asserted. □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. (a) To show that $w(A) \leqslant (n-1)w_0(A)$, we represent A as

$$\begin{bmatrix} 0 & A_{12} & \cdots & A_{1n} \\ & 0 & \ddots & \vdots \\ & & \ddots & A_{n-1,n} \\ & & & 0 \end{bmatrix}$$

on $H = \ker A \oplus (\ker A^2 \ominus \ker A) \oplus \cdots \oplus (H \ominus \ker A^{n-1})$. Let

$$A' = \begin{bmatrix} 0 & \|A_{12}\| & \cdots & \|A_{1n}\| \\ & 0 & \ddots & \vdots \\ & & \ddots & \|A_{n-1,n}\| \\ & & & 0 \end{bmatrix}$$

and, for any unit vector $x = [x_1 \ \cdots \ x_n]^{\mathrm{T}}$ in H, let $x' = [\|x_1\| \ \cdots \ \|x_n\|]^{\mathrm{T}}$. We have

$$\begin{aligned}|\langle Ax, x\rangle| &\leqslant \langle A'x', x'\rangle && \text{(ii)} \\ &\leqslant \max\{\|A_{ij}\| : 1 \leqslant i < j \leqslant n\}\langle A_n x', x'\rangle && \text{(iii)} \\ &\leqslant 2w_0(A)w(A_n) && \text{(iv)} \\ &= (n-1)w_0(A), && \text{(v)}\end{aligned}$$

where the inequality in (iv) follows from the facts that $\begin{bmatrix} 0 & A_{ij} \\ 0 & 0 \end{bmatrix}$ dilates to A and thus

$$\left\{z \in \mathbb{C} : |z| \leqslant \frac{1}{2}\|A_{ij}\|\right\} = \overline{W\left(\begin{bmatrix} 0 & A_{ij} \\ 0 & 0 \end{bmatrix}\right)} \subseteq \overline{W(A)}$$

(cf. [13, Theorem 2.1]), and the equality in (v) is by Lemma 1.3 (2). The inequality $w(A) \leqslant (n-1)w_0(A)$ follows.

(b) To prove (1) $\Rightarrow$ (2), assume that $w(A) = (n-1)w_0(A)$. Represent A as in (a) and let $x = [x_1 \ \cdots \ x_n]^{\mathrm{T}}$ be a unit vector in H such that $\langle Ax, x\rangle = w(A)\mathrm{e}^{\mathrm{i}\theta}$ for some real θ. Replacing A by $\mathrm{e}^{-\mathrm{i}\theta}A$, we may assume that $\langle Ax, x\rangle = w(A)$. Then $\langle Ax, x\rangle = (n-1)w_0(A)$ implies that the inequalities in (ii), (iii) and (iv) are actually equalities. The one from (ii) yields

$$\sum_{i<j}\langle A_{ij}x_j, x_i\rangle = \sum_{i<j}\|A_{ij}\|\|x_j\|\|x_i\|,$$

which in turn implies that $\langle A_{ij}x_j, x_i\rangle = \|A_{ij}x_j\|\|x_i\| = \|A_{ij}\|\|x_j\|\|x_i\|$ for all $i<j$. Thus $A_{ij}x_j = a_{ij}x_i$ for some scalar a_{ij}. On the other hand, the equalities from (iii) and (iv) yield $\|A_{ij}\| = 2w_0(A)$ for all $i<j$ and $\langle A_n x', x'\rangle = w(A_n)$. By Lemma 1.3(2) and (5), the latter implies that $\|x_j\| = 1/\sqrt{n}$ for all j. Hence

$$\begin{aligned}\frac{1}{n}a_{ij} &= a_{ij}\|x_i\|^2 = \langle A_{ij}x_j, x_i\rangle = \|A_{ij}\|\|x_j\|\|x_i\| \\ &= \frac{1}{n}\|A_{ij}\| = \frac{2}{n}w_0(A)\end{aligned}$$

and therefore $A_{ij}x_j = 2w_0(A)x_i$ for all $i<j$. If $e_j = [0 \ \cdots \ 0 \ \underset{j\text{th}}{\sqrt{n}x_j} \ 0 \ \cdots \ 0]^{\mathrm{T}}$ in H for $1 \leqslant j \leqslant n$, then $\{e_1, \ldots, e_n\}$ forms an orthonormal set in H. Let M be the subspace generated by the e_j's. We check that M is a reducing subspace of A with $A|M$ unitarily equivalent to $2w_0(A)A_n$. Indeed, we have

$$\begin{aligned}Ae_j &= \sqrt{n}[A_{1j}x_j \ \cdots \ A_{j-1,j}x_j \ 0 \ \cdots \ 0]^{\mathrm{T}} \\ &= 2\sqrt{n}w_0(A)[x_1 \ \cdots \ x_{j-1} \ 0 \ \cdots \ 0]^{\mathrm{T}} \\ &= 2w_0(A)(e_1 + \cdots + e_{j-1})\end{aligned}$$

for all j, which shows that M is invariant under A. On the other hand, we also have

$$A^*e_i = \sqrt{n}[0 \ \cdots \ 0 \ A^*_{i,i+1}x_i \ \cdots \ A^*_{in}x_i]^{\mathrm{T}}.$$

Note that, for any contraction T ($\|T\| \leqslant 1$) and vectors x and y with $\|x\| = \|y\|$, $Tx = y$ and $T^*y = x$ are equivalent. (This is essentially due to B. Sz.-Nagy and can be proved as in [12, Proposition I.3.1]) Applying it to $(A_{ij}/\|A_{ij}\|)x_j = x_i$ yields $(A^*_{ij}/\|A_{ij}\|)x_i = x_j$ or $A^*_{ij}x_i = 2w_0(A)x_j$ for $i<j$. Thus

$$\begin{aligned}A^*e_i &= 2\sqrt{n}w_0(A)[0 \ \cdots \ 0 \ x_{i+1} \ \cdots \ x_n]^{\mathrm{T}} \\ &= 2w_0(A)(e_{i+1} + \cdots + e_n)\end{aligned}$$

for all i and hence M is invariant under A^*. This proves that M reduces A. Since

$$\begin{aligned}\langle Ae_j, e_i\rangle &= 2w_0(A)(\langle e_1, e_i\rangle + \cdots + \langle e_{j-1}, e_i\rangle) \\ &= \begin{cases}0 & \text{if } j \leqslant i, \\ 2w_0(A) & \text{if } i<j,\end{cases}\end{aligned}$$

we have the unitary equivalence of $A|M$ and $[\langle Ae_j, e_i\rangle]_{i,j=1}^n = 2w_0(A)A_n$. Hence A is unitarily equivalent to $2w_0(A)A_n \oplus A'$ for some A'. This proves (2).

For (2) $\Rightarrow$ (3), we may assume that $A = A_n \oplus A'$ with $w_0(A) = 1/2$. Then, obviously, $W(A_n) \subseteq W(A)$. We now check that $\operatorname{Re} A' \geqslant (-1/2)I$ or, equivalently, $W(A') \subseteq \{z \in \mathbb{C} : \operatorname{Re} z \geqslant -1/2\}$. Assume otherwise that there is some z_0 in $W(A')$ with $\operatorname{Re} z_0 < -1/2$. If $[z_1, \overline{z_1}]$ is the vertical line segment on $\partial W(A_n)$, then the triangular region $\Delta z_0 z_1 \overline{z_1}$ is contained in $W(A)$. Let z_2 be a point in $\Delta z_0 z_1 \overline{z_1}$ which is on the x-axis. Then $\Delta z_2 z_1 \overline{z_1}$ is also contained in $W(A)$. If Δ is the convex hull of the set $\Delta z_2 z_1 \overline{z_1} \cup W(A_n)$, then $\Delta \subseteq W(A)$ and $\mathrm{dist}(0, \partial\Delta) > 1/2$ by Lemma 1.3(7). Thus $w_0(A) = \mathrm{dist}(0, \partial W(A)) > 1/2$, contradicting our assumption. Hence $\mathrm{Re}\, A' \geqslant (-1/2)I$ as asserted. Therefore, $\mathrm{Re}\, A = \mathrm{Re}\, A_n \oplus \mathrm{Re}\, A' \geqslant (-1/2)I$ by Lemma 1.3 (3). Proposition 1.5 then implies that $W(A) \subseteq W(A_n)$. Thus $W(A) = W(A_n)$, which proves (3).

The implication (3) $\Rightarrow$ (1) is an easy consequence of Lemma 1.3 (2) and (6). □

2. Numerical and spectral radii

In this section, we consider a situation which is opposite to the nilpotent case, namely, we consider an n-by-n matrix A with 0 not in the interior of its numerical range. Under this condition, [1, Lemma 2.6] says that $w(A) \leqslant nr(A)$ holds, where $r(A) = \max\{|z| : z \in \sigma(A)\}$ is the *spectral radius* of A. This inequality resembles in appearance the one we have had in Section 1, namely, $w(A) \leqslant (n-1)w_0(A)$ for A satisfying $A^n = 0$. As the following theorem shows, the resemblance extends even to the equality case.

Theorem 2.1. *Let A be an n-by-n matrix with $0 \notin \mathrm{Int}\,W(A)$.*

(a) *The inequality $w(A) \leqslant nr(A)$ holds.*
(b) *The following conditions are equivalent*:

(1) $w(A) = nr(A)$,
(2) *A is unitarily equivalent to aB_n for some a in $\mathbb{C}$, where B_n is the n-by-n matrix*

$$\begin{bmatrix} 1 & 2 & \cdots & 2 \\ & 1 & \ddots & \vdots \\ & & \ddots & 2 \\ & & & 1 \end{bmatrix}$$

and
(3) *$W(A) = aW(B_n)$ for some a in $\mathbb{C}$.*

Here $B_n = 2A_n + I_n$ plays the role of A_n in Theorem 1.1. (a) is from [1, Lemma 2.6]. We include its proof below for completeness. (b) depends on the numerical range analogue of the Perron–Frobenius theorem for nonnegative matrices. Our reference is [9].

For a complex matrix $A = [a_{ij}]_{i,j=1}^n$, we use $|A|$ to denote the matrix $[|a_{ij}|]_{i,j=1}^n$. If $A = [a_{ij}]_{i,j=1}^n$ and $B = [b_{ij}]_{i,j=1}^n$ are real matrices, then $A \preccurlyeq B$ (or $B \succcurlyeq A$) means that $a_{ij} \leqslant b_{ij}$ for all i and j. Similar notations extend to vectors.

Proof of Theorem 2.1. (a) Since 0 is not in $\mathrm{Int}\,W(A)$, we may assume, after a suitable rotation, that $\mathrm{Re}\, A \geqslant 0$ and also

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{bmatrix}.$$

Since

$$\operatorname{Re}\begin{bmatrix} a_{ii} & a_{ij} \\ 0 & a_{jj} \end{bmatrix} = \begin{bmatrix} \operatorname{Re} a_{ii} & a_{ij}/2 \\ \overline{a_{ij}}/2 & \operatorname{Re} a_{jj} \end{bmatrix} \geqslant 0$$

for all $i < j$, we have $|a_{ij}|^2/4 \leqslant (\operatorname{Re} a_{ii})(\operatorname{Re} a_{jj})$ or $|a_{ij}| \leqslant 2|a_{ii}a_{jj}|^{1/2}$. It follows that $|A| \preccurlyeq r(A)B_n$ and hence $w(A) \leqslant r(A)w(B_n)$ by [9, Corollary 3.6]. Since

$$w(B_n) = w(2A_n + I_n) = 2w(A_n) + 1 = 2 \cdot \frac{1}{2}(n-1) + 1 = n$$

by Lemma 1.3 (2), the asserted inequality follows.

(b) We first prove (1) $\Rightarrow$ (2). Let $w \equiv w(A) = nr(A)$ and $B_n' = r(A)B_n$. From the proof in (a), we have $w = w(B_n')$. Let x be a unit vector in $\mathbb{C}^n$ such that $|\langle Ax, x\rangle| = w$. Since

$$w = |\langle Ax, x\rangle| \leqslant \langle |A||x|, |x|\rangle \leqslant \langle B_n'|x|, |x|\rangle \leqslant w,$$

the inequalities here are actually equalities. In particular, we have $\langle B_n'|x|, |x|\rangle = w$. Hence $|x| = [1/\sqrt{n} \ \cdots \ 1/\sqrt{n}]^{\mathrm{T}}$ by Lemma 1.3 (5) and (2). Since $\langle (B_n' - |A|)|x|, |x|\rangle = 0$ and $B_n' - |A| \succcurlyeq 0$ (from (a)), we infer that $|A| = B_n'$. Let $x = [x_1 \ \cdots \ x_n]^{\mathrm{T}}$ and $D = \operatorname{diag}(|x_1|/x_1, \ldots, |x_n|/x_n)$. If $\lambda = \langle Ax, x\rangle/w$, then

$$\left\langle \frac{1}{\lambda} DAD^{-1}|x|, |x| \right\rangle = \frac{1}{\lambda}\langle Ax, x\rangle = w = \langle B_n'|x|, |x|\rangle.$$

Since $|(1/\lambda)DAD^{-1}| = |A| = B_n'$, we infer, by taking into account the fact that $|x| = [1/\sqrt{n} \ \cdots \ 1/\sqrt{n}]^{\mathrm{T}}$, that $(1/\lambda)DAD^{-1} = B_n'$ or

$$A = \lambda D^{-1}B_n'D = \lambda r(A)D^{-1}B_nD,$$

which proves (2).

Since (2) $\Rightarrow$ (3) is trivial, to complete the proof we need only show (3) $\Rightarrow$ (1). If (3) holds, then

$$w(A) = |a|w(B_n) = n|a|$$

as proved in (a). Note that $p_{A_n}(x, y, z) = p_{B_n}(x/2, y/2, z - (x/2))$ for all x, y and z. The irreducibility of p_{A_n} (by Lemma 1.3 (8)) implies that of p_{B_n} and also of $p_{B_n'}$, where $B_n' = aB_n$. Since $W(A) = W(B_n')$, we infer from [4, Corollary 2.4] that $p_A = p_{B_n'}$. Hence $\sigma(A) = \sigma(B_n') = \{a\}$. In particular, this gives $r(A) = |a|$. Thus $w(A) = nr(A)$ as asserted. □

We remark that the above proof for (1) $\Rightarrow$ (2) is modeled after that of [9, Lemma 3.8].

Acknowledgement

We thank the (anonymous) referee for prompting us to a deeper probing of the conditions in Theorem 1.1, which results in a vast improvement, and also for reminding us of the Refs. [2,6].

References

[1] L. Caston, M. Savova, I. Spitkovsky, N. Zobin, On eigenvalues and boundary curvature of the numerical range, Linear Algebra Appl. 322 (2001) 129–140.

[2] M.-T. Chien, H. Nakazato, Boundary generating curves of the c-numerical range, Linear Algebra Appl. 294 (1999) 67–84.

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

[4] H.-L. Gau, P.Y. Wu, Companion matrices: reducibility, numerical ranges and similarity to contractions, Linear Algebra Appl. 383 (2004) 127–142.
[5] K. Gustafson, D.K.M. Rao, Numerical Range. The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.
[6] U. Haagerup, P. de la Harpe, The numerical radius of a nilpotent operator on a Hilbert space, Proc. Amer. Math. Soc. 115 (1992) 371–379.
[7] P.R. Halmos, A Hilbert Space Problem Book, second ed., Springer, New York, 1982.
[8] R. Kippenhahn, Über den Wertevorrat einer Matrix, Math. Nachr. 6 (1951) 193–228., English translation: P.F. Zachlin, M.E. Hochstenbach, On the numerical range of a matrix, Linear and Multilinear Algebra 56 (2008) 185–225.
[9] C.-K. Li, B.-S. Tam, P.Y. Wu, The numerical range of a nonnegative matrix, Linear Algebra Appl. 350 (2002) 1–23.
[10] C. Pop, On a result of Haagerup and de la Harpe, Rev. Roumaine Math. Pures Appl. 43 (1998) 869–871.
[11] B. Sims, On a connection between the numerical range and spectrum of an operator on a Hilbert space, J. London Math. Soc. 8 (1974) 57–59.
[12] B. Sz-Nagy, C. Foiaş, Harmonic Analysis of Operators on Hilbert Space, North Holland, Amsterdam, 1970.
[13] S.-H. Tso, P.Y. Wu, Matricial ranges of quadratic operators, Rocky Mountain J. Math. 29 (1999) 1139–1152.