Linear Algebra and its Applications 433 (2010) 2243–2254

Contents lists available at ScienceDirect

Linear Algebra and its Applications

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

Crawford numbers of powers of a matrix

Kuo-Zhong Wang [a,1], Pei Yuan Wu [a,*,2], Hwa-Long Gau [b,3]

[a] Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan
[b] Department of Mathematics, National Central University, Chungli 32001, Taiwan

ARTICLE INFO

Article history:
Received 31 March 2010
Accepted 29 July 2010
Available online 30 August 2010

Submitted by C.K. Li

AMS classification:
15A60
47A12

Keywords:
Numerical range
Crawford number
Generalized Crawford number

ABSTRACT

For an n-by-n matrix A, its Crawford number $c(A)$ (resp., generalized Crawford number $C(A)$) is, by definition, the distance from the origin to its numerical range $W(A)$ (resp., the boundary of its numerical range $\partial W(A)$). It is shown that if A has eigenvalues $\lambda_1, \ldots, \lambda_n$ arranged so that $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$, then $\overline{\lim}_k c(A^k)^{1/k}$ (resp., $\overline{\lim}_k C(A^k)^{1/k}$) equals 0 or $|\lambda_n|$ (resp., $|\lambda_j|$ for some j, $1 \leqslant j \leqslant n$). For a normal A, more can be said, namely, $\overline{\lim}_k c(A^k)^{1/k} = |\lambda_n|$ (resp., $\overline{\lim}_k C(A^k)^{1/k} = |\lambda_j|$ for some j, $3 \leqslant j \leqslant n$). In these cases, the above possible values can all be assumed by some A.

1. Introduction

Let A be a bounded linear operator on a complex Hilbert space H. The *numerical range* of A is, by definition, $W(A) = \{\langle Ax, x\rangle : x \in H, \|x\| = 1\}$, where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ are the inner product and its associated norm in H. It is known that $W(A)$ is a bounded convex subset of the complex plane with its closure $\overline{W(A)}$ containing the spectrum $\sigma(A)$ of A. For its other properties, the reader may consult [4, Chapter 22], [3] or [5, Chapter 1].

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

[1] Research supported by a post-doctor fellowship of the National Science Council of the Republic of China.
[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 project.
[3] Research supported by the National Science Council of the Republic of China under NSC 98-2628-M-008-007.

0024-3795/$ - see front matter
doi:10.1016/j.laa.2010.08.004

To measure the location and relative size of $W(A)$, one frequently used quantity is the *numerical radius* of A: $w(A) = \sup\{|z| : z \in W(A)\}$. In this paper, we consider two other quantities, which are not so well known but still quite useful, namely, the *Crawford number* $c(A)$ and *generalized Crawford number* $C(A)$ of A. These are defined as $c(A) = \inf\{|z| : z \in W(A)\}$ and $C(A) = \inf\{|z| : z \in \partial W(A)\}$, respectively. The former was first considered (for finite matrices) in [1] while the latter appeared in [2] in the study of the numerical ranges of nilpotent operators. We are concerned here with the asymptotic behavior of the homogenized (generalized) Crawford numbers of powers of a matrix A. Note that such a sequence $c(A^k)^{1/k}$ (resp., $C(A^k)^{1/k}$), $k \geqslant 1$, may not converge in general $\left(\text{as witness the matrix } A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right)$. Hence we will consider instead its limit supremum. In the literature, there are several results of this nature for different operator parameters. For example, it is well known that, for any operator A, the limit of $\|A^k\|^{1/k}$ as k approaches infinity exists and is equal to the *spectral radius* $r(A)$ $(\equiv \sup\{|z| : z \in \sigma(A)\})$ of A (cf. [4, Problem 88]). Since the numerical radius and the norm of A are related by $\|A\|/2 \leqslant w(A) \leqslant \|A\|$, it can be easily seen that $\lim_k w(A^k)^{1/k}$ also equals $r(A)$. For the *minimum modulus* $m(A)$ $(\equiv \inf\{\|Ax\| : \|x\| = 1\})$, it has been proven that $\lim_k m(A^k)^{1/k} = \text{dist}\,(0, \sigma_l(A))$, the distance from the origin to the *left spectrum* $\sigma_l(A)$ $(\equiv \{z \in \mathbb{C} : A - zI \text{ not left invertible}\})$ of A (cf. [10, Theorem 3]). If A acts on an n-dimensional space, the results on $\|\cdot\|$ and $m(\cdot)$ can be extrapolated to singular numbers of A: $\lim_k s_j(A^k)^{1/k} = |\lambda_j|$ for each j, $1 \leqslant j \leqslant n$, where $s_j(A^k)$ is the jth largest singular number of A^k and λ_j is the jth largest, in modulus, eigenvalue of A (cf. [9] and also [7, Theorem 2]). The contrast between our results for the Crawford numbers versus the ones for the numerical radius is analogous to that for the minimum modulus versus the norm. More precisely, we show that if A is an n-by-n matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$ arranged so that $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$, then $\overline{\lim}_k\, c(A^k)^{1/k}$ equals either 0 or $|\lambda_n|$ while $\overline{\lim}_k\, C(A^k)^{1/k}$ equals some $|\lambda_j|$, $1 \leqslant j \leqslant n$. For A a normal matrix or of size two, more detailed information can be obtained. For example, if A is normal, then $\overline{\lim}_k\, c(A^k)^{1/k}$ is always equal to $|\lambda_n|$ and $\overline{\lim}_k\, C(A^k)^{1/k}$ equals some $|\lambda_j|$ with $3 \leqslant j \leqslant n$. For a 2-by-2 matrix A, we give precise conditions on A under which these limit suprema equal each of the asserted values. Such results for the Crawford number (resp., generalized Crawford number) will be presented in Section 2 (resp., Section 3) below.

We end this section by giving some basic properties of the (generalized) Crawford numbers of general operators.

Proposition 1.1. *Let A and A_n, $n \geqslant 1$, be operators on H. Then the following hold:*

(1) $c(A) = c(A^*)$ *and* $C(A) = C(A^*)$.
(2) $c(\lambda A) = |\lambda| c(A)$ *and* $C(\lambda A) = |\lambda| C(A)$ *for any scalar* λ.
(3) $0 \leqslant c(A) \leqslant C(A) \leqslant w(A)$.
(4) $c(A) > 0$ (*resp.*, $C(A) > 0$) *if and only if* $0 \notin \overline{W(A)}$ (*resp.*, $0 \notin \partial W(A)$).
(5) *If* $c(A) > 0$, *then* $c(A) = C(A) \leqslant \text{dist}\,(0, \sigma(A))$.
(6) *If A is invertible, then* $c(A) > 0$ (*resp.*, $C(A) > 0$) *if and only if* $c(A^{-1}) > 0$ (*resp.*, $C(A^{-1}) > 0$).
(7) *If* $A_n \to A$ *in norm, then* $c(A_n) \to c(A)$ *and* $C(A_n) \to C(A)$.
(8) $\overline{\lim}_k\, c(A^k)^{1/k} \leqslant \text{dist}\,(0, \sigma(A))$ *and* $\overline{\lim}_k\, C(A^k)^{1/k} \leqslant r(A)$.

Proof. (1), (2), (3) and (4) are trivial. (5) follows from the fact that $\sigma(A) \subseteq \overline{W(A)}$. To prove (6), we need check that $0 \in \overline{W(A)}$ (resp., $0 \in \partial W(A)$) if and only if $0 \in \overline{W(A^{-1})}$ (resp., $0 \in \partial W(A^{-1})$). Assume first that $0 \in \overline{W(A)}$. Then there are unit vectors x_n, $n \geqslant 1$, in H such that $\langle Ax_n, x_n \rangle \to 0$ as $n \to \infty$. Hence

$$\left\langle A^{-1}\left(\frac{Ax_n}{\|Ax_n\|}\right), \frac{Ax_n}{\|Ax_n\|} \right\rangle = \frac{1}{\|Ax_n\|^2} \langle x_n, Ax_n \rangle = \frac{1}{\|Ax_n\|^2} \overline{\langle Ax_n, x_n \rangle} \to 0$$

since $1/\|Ax_n\|^2 \leqslant \|A^{-1}\|^2$ for all n. This shows that 0 is in $\overline{W(A^{-1})}$. That $0 \in \overline{W(A^{-1})}$ implies $0 \in \overline{W(A)}$ follows by symmetry. Next assume that $0 \in \partial W(A) \backslash \partial W(A^{-1})$. By what we have just proven, 0 is in the interior of $W(A^{-1})$. Let $r > 0$ be such that $re^{i\theta}$ is in $W(A^{-1})$ for all real θ. Then $re^{i\theta} = \langle A^{-1}x_\theta, x_\theta \rangle$ for some unit vector x_θ in H. Hence

$$\frac{re^{-i\theta}}{\|A^{-1}x_\theta\|^2} = \overline{\left\langle \frac{A^{-1}x_\theta}{\|A^{-1}x_\theta\|}, \frac{x_\theta}{\|A^{-1}x_\theta\|} \right\rangle} = \left\langle A\left(\frac{A^{-1}x_\theta}{\|A^{-1}x_\theta\|}\right), \frac{A^{-1}x_\theta}{\|A^{-1}x_\theta\|} \right\rangle,$$

which shows that $re^{-i\theta}/\|A^{-1}x_\theta\|^2$ is in $W(A)$ for all real θ. Since $|re^{-i\theta}|/\|A^{-1}x_\theta\|^2 \geqslant r/\|A^{-1}\|^2$ and 0 is in $\overline{W(A)}$, we infer from the convexity of $W(A)$ that $re^{-i\theta}/\|A^{-1}\|^2$ is in $W(A)$ for all real θ and hence 0 is in the interior of $W(A)$. This contradicts our assumption. Thus $0 \in \partial W(A)$ implies $0 \in \partial W(A^{-1})$. The converse follows by symmetry.

(7) follows from the fact that $A_n \to A$ in norm implies $\overline{W(A_n)} \to \overline{W(A)}$ and also $\partial W(A_n) \to \partial W(A)$ in the Hausdorff metric (cf. [4, Problem 220]). (We remark that the assertion for the Crawford numbers also follows from the easily verified inequality $|c(A_n) - c(A)| \leqslant \|A_n - A\|$ for all n.)

For the proof of (8), note that $c(A^k)^{1/k} \leqslant \text{dist}\,(0, \sigma(A))$ for all $k \geqslant 1$ by (5). Hence $\overline{\lim}_k\, c(A^k)^{1/k} \leqslant \text{dist}\,(0, \sigma(A))$ holds. For the generalized Crawford number, we make use of $C(A) \leqslant w(A)$ from (3). Hence

$$\overline{\lim_k}\, C(A^k)^{1/k} \leqslant \overline{\lim_k}\, w(A^k)^{1/k} = r(A)$$

as was noted before. □

In the remaining part of this paper, we consider only finite matrices unless otherwise stated.

2. Crawford number

The main result of this section is the following theorem.

Theorem 2.1. *If A is an n-by-n matrix, then* $\overline{\lim}_k\, c(A^k)^{1/k}$ *equals either* 0 *or* dist $(0, \sigma(A))$.

We start the proof with the following result of Kronecker.

Lemma 2.2. *If* $|\lambda_j| = 1$ *for* $1 \leqslant j \leqslant n$, *then there are positive integers* n_k, $k \geqslant 1$, *such that* $\lim_k \lambda_j^{n_k} = 1$ *for all* j.

Proof. Let $\lambda_j = \exp(2\pi i\theta_j)$, where $0 \leqslant \theta_j < 1$, for each j. We may assume that $\{\theta_1, \ldots, \theta_m, 1\}$ $(0 \leqslant m \leqslant n)$ is a maximal independent set over the field of rational numbers. Then the same is true for $\{2\pi\theta_1, \ldots, 2\pi\theta_m, \pi\}$. Kronecker's theorem [6, Theorem VI.9.1] says that there are positive integers m_k, $k \geqslant 1$, such that $\lim_k \lambda_j{}^{m_k} = 1$ for all j, $1 \leqslant j \leqslant m$. Let

$$\theta_j = \frac{p_1^{(j)}}{q_1^{(j)}}\theta_1 + \cdots + \frac{p_m^{(j)}}{q_m^{(j)}}\theta_m + \frac{p_{m+1}^{(j)}}{q_{m+1}^{(j)}}, \quad m+1 \leqslant j \leqslant n,$$

where $p_u^{(j)}$ and $q_u^{(j)}$ are relatively prime integers for each u, $1 \leqslant u \leqslant m+1$. If $n_k = m_k \prod_{u,v} q_u^{(v)}$ for $k \geqslant 1$, then it is easily seen that $\lim_k \lambda_j^{n_k} = 1$ for all j, $1 \leqslant j \leqslant n$. □

The normal case can now be easily treated.

Proposition 2.3. *If A is an n-by-n normal matrix, then* $\overline{\lim}_k\, c(A^k)^{1/k} = \sup_k c(A^k)^{1/k} = \text{dist}\,(0, \sigma(A))$. *In this case,* $\overline{\lim}_k\, c(A^k)^{1/k} = 0$ *if and only if* 0 *is in* $W(A^k)$ *for all* $k \geqslant 1$.

Proof. If 0 is in $\sigma(A)$, then it is in $\sigma(A^k)$ and hence in $W(A^k)$ for all k. Thus $\overline{\lim}_k\, c(A^k)^{1/k} = \sup_k c(A^k)^{1/k} = \text{dist}\,(0, \sigma(A)) = 0$. Hence we may assume that 0 is not in $\sigma(A)$. In view of Proposition 1.1 (2), we may further assume that $A = \text{diag}\,(\lambda_1, \ldots, \lambda_{n-1}, 1)$, where $|\lambda_j| \geqslant 1$ for all j, $1 \leqslant j \leqslant n-1$. By Lemma 2.2, there are positive integers n_k, $k \geqslant 1$, such that $\lim_k (\lambda_j/|\lambda_j|)^{n_k} = 1$ for all j. Since $W(A^{n_k})$ is the convex hull of $\{\lambda_1^{n_k}, \ldots, \lambda_{n-1}^{n_k}, 1\}$, we infer that $\lim_k c(A^{n_k}) = 1$. Hence

$$1 = \lim_k c(A^{n_k})^{1/n_k} \leqslant \overline{\lim_k}\, c(A^k)^{1/k} \leqslant \sup_k c(A^k)^{1/k} \leqslant \text{dist}\,(0, \sigma(A)) = 1,$$

where the last inequality is by Proposition 1.1 (5). Our assertions follow. □

We next prove for matrices of size two. The proof engenders all the essential ingredients of the one for general matrices. Recall that if $A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$, then $W(A)$ equals the elliptic disc with foci a and b and with minor axis of length $|c|$ (cf. [4, p. 113]).

Proposition 2.4. *Let* $A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$. *Then*

$$\overline{\lim_k}\, c(A^k)^{1/k} = \begin{cases} \min\{|a|, |b|\} & \text{if either } c = 0 \text{ or } c \neq 0, a \neq b \text{ and } |a| = |b|, \\ 0 & \text{otherwise}. \end{cases}$$

In this case, $\overline{\lim}_k\, c(A^k)^{1/k} = 0$ *if and only if* 0 *is in* $W(A^k)$ *for all large* k.

Proof. If $a = 0$ or $b = 0$, then 0 is in $\sigma(A^k)$ and hence in $W(A^k)$ for all k. Thus $c(A^k) = 0$ for all k and therefore $\lim_k c(A^k)^{1/k} = 0$. For the remaining part of the proof, we assume that $a, b \neq 0$. Four cases are considered separately:

(1) $c \neq 0$ and $a = b$. In this case,

$$A^k = \begin{bmatrix} a^k & cka^{k-1} \\ 0 & a^k \end{bmatrix} = a^k \begin{bmatrix} 1 & ck/a \\ 0 & 1 \end{bmatrix}.$$

Since 0 is in $W\left(\begin{bmatrix} 1 & ck/a \\ 0 & 1 \end{bmatrix}\right)$ and hence in $W(A^k)$ for all large k, we have $c(A^k) = 0$ for such k's. Thus $\lim_k c(A^k)^{1/k} = 0$.

(2) $c \neq 0$ and $|a| \neq |b|$. In view of Proposition 1.1 (2), we may assume that $a = 1 > |b|$. Since

$$A^k = \begin{bmatrix} 1 & c(1 - b^k)/(1 - b) \\ 0 & b^k \end{bmatrix} \to \begin{bmatrix} 1 & c/(1 - b) \\ 0 & 0 \end{bmatrix}$$

as $k \to \infty$, we obtain $0 \in W(A^k)$ for all large k. Hence $\lim_k c(A^k)^{1/k} = 0$.

(3) $c = 0$. In this case, A is normal. Hence our assertion follows form Proposition 2.3.

(4) $c \neq 0, a \neq b$ and $|a| = |b|$. We may assume that $|a| = |b| = 1$. Let $n_k, k \geqslant 1$, be positive integers such that $\lim_k a^{n_k} = \lim_k b^{n_k} = 1$ by Lemma 2.2. Then

$$A^{n_k} = \begin{bmatrix} a^{n_k} & c(a^{n_k} - b^{n_k})/(a - b) \\ 0 & b^{n_k} \end{bmatrix} \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

as $k \to \infty$. Hence $\overline{\lim}_k\, c(A^k)^{1/k} \geqslant \lim_k c(A^{n_k})^{1/n_k} = 1$. The converse inequality follows by Proposition 1.1 (8). □

Note that in general $\overline{\lim}_k\, c(A^k)^{1/k}$ is smaller than $\sup_k c(A^k)^{1/k}$ even for a 2-by-2 matrix A. One example is $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ for which the former equals 0 while the latter 1/2. For general matrices, the condition for their equality will be given in Proposition 2.8.

For the ease of exposition, we define three types of matrices, which correspond roughly to the four cases in the proof of Proposition 2.4.

Definition 2.5. A matrix A is said to be of *type I* if its eigenvalues have equal moduli and it is unitarily equivalent to a matrix of the form

$$\begin{bmatrix} * & * & * \\ 0 & A' & * \\ 0 & 0 & * \end{bmatrix},$$

where A' is not a scalar matrix and has equal eigenvalues, it is of *type II* if it is irreducible and has two eigenvalues with unequal moduli, and it is of *type III* if it is unitarily equivalent to a matrix of the form

$$\begin{bmatrix} \lambda_1 I_1 & & * \\ & \ddots & \\ 0 & & \lambda_m I_m \end{bmatrix},$$

where $|\lambda_1| = \cdots = |\lambda_m|$ and the λ_j's are distinct.

Recall that a matrix is *irreducible* if it is not unitarily equivalent to the direct sum of any two other matrices.

Note that a matrix of type I or II has size at least two while a size-one matrix is of type III. A general (resp., normal) matrix is unitarily equivalent to a direct sum of irreducible matrices of type I, II or III (resp., of type III).

The next lemma deals with matrices of types I and II.

Lemma 2.6. *If A is an n-by-n matrix of type I (even without the requirement of equal-moduli eigenvalues) or type II, then* 0 *is in* $W(A^k)$ *for all large k and hence* $\lim_k c(A^k)^{1/k} = 0$.

Proof.

(1) A is of type I. Let

$$A = \begin{bmatrix} * & * & * \\ 0 & A' & * \\ 0 & 0 & * \end{bmatrix},$$

where $A' = [a_{ij}]_{i,j=1}^m$ $(2 \leqslant m \leqslant n)$ is upper triangular ($a_{ij} = 0$ for all $i > j$), invertible ($\lambda_i \equiv a_{ii} \neq 0$ for all i) with $\lambda_1 = \cdots = \lambda_m$, and is not a scalar matrix. Then there are i_0 and j_0 with $i_0 < j_0$ such that $a_{i_0 j_0} \neq 0$ and $a_{ij} = 0$ for all i and j satisfying either $i < i_0$ or $i = i_0$ and $j < j_0$. If

$$B = \begin{bmatrix} \lambda_1 & a_{i_0 j_0} \\ 0 & \lambda_1 \end{bmatrix},$$

then B^k is a submatrix of A^k for all $k \geqslant 1$ and 0 is in $W(B^k)$ for all large k by Proposition 2.4 (1). Hence 0 is in $W(A^k)$ for all large k.

(2) A is of type II. We may assume that $A = [a_{ij}]_{i,j=1}^n$ with $a_{ij} = 0$ for $i > j$, $\lambda_i \equiv a_{ii} \neq 0$ for all i, and $|\lambda_1| = \cdots = |\lambda_m| > |\lambda_{m+1}|, \ldots, |\lambda_n|$ for some m, $1 \leqslant m < n$. Since A is irreducible, there are i_0 and j_0 with $1 \leqslant i_0 \leqslant m$ and $m < j_0 \leqslant n$ such that $a_{i_0 j_0} \neq 0$ and $a_{ij} = 0$ for all i and j satisfying either $i = i_0$ and $m < j < j_0$ or $i_0 < i \leqslant m$ and $m < j \leqslant n$. Let

$$B = \begin{bmatrix} \lambda_{i_0} & a_{i_0 j_0} \\ 0 & \lambda_{j_0} \end{bmatrix}.$$

By Proposition 2.4 (2), we have $0 \in W(B^k)$ for all large k. Since B^k is a submatrix of A^k for all k, we obtain $0 \in W(A^k)$ for all large k. □

The following lemma is useful in proving for type-III matrices.

Lemma 2.7. *Let* $A = [A_{ij}]_{i,j=1}^n$ $(n \geqslant 2)$ *on* $H = \sum_{j=1}^n \oplus H_j$ *be an n-by-n upper-triangular operator matrix with* $A_{ij} = 0$ *for all* $i > j$ *and* $A_{ii} = \lambda_i I_i$ *for all i, where the* λ_i*'s are distinct with equal moduli and* I_i *denotes the identity operator on* H_i. *If* $A^k = [(A^k)_{ij}]_{i,j=1}^n$ *for* $k \geqslant 1$, *then, for each pair* (i,j) *with* $i < j$, *there are operators* B_l, $i+1 \leqslant l \leqslant j$, *from* H_j *to* H_i *such that*

$$(A^k)_{ij} = \sum_{l=i+1}^{j} \frac{\lambda_i^k - \lambda_l^k}{\lambda_i - \lambda_l} B_l \quad \textit{for all } k \geqslant 1.$$

Proof. We prove this by induction on n. For $n = 2$, we have

$$A^k = \begin{bmatrix} \lambda_1^k I_1 & [(\lambda_1^k - \lambda_2^k)/(\lambda_1 - \lambda_2)]A_{12} \\ 0 & \lambda_2^k I_2 \end{bmatrix}.$$

Hence $B_2 = A_{12}$ meets our requirement. Next assuming that our assertion is true for all operator matrices of size less than n, we prove it for n. If $(i,j) \neq (1,n)$, then this follows by the induction hypothesis for the $(n-1)$-by-$(n-1)$ submatrix $[A_{ij}]_{i,j=1}^{n-1}$ or $[A_{ij}]_{i,j=2}^{n}$ of A. Hence we need only consider for $(i,j) = (1,n)$. In this case, we have

$$(A^k)_{1n} = (AA^{k-1})_{1n} = \lambda_1 (A^{k-1})_{1n} + \sum_{j=2}^{n} A_{1j}(A^{k-1})_{jn}$$

and similarly

$$(A^{k-1})_{1n} = \lambda_1 (A^{k-2})_{1n} + \sum_{j=2}^{n} A_{1j}(A^{k-2})_{jn}.$$

Substituting the latter into the former yields

$$(A^k)_{1n} = \lambda_1^2 (A^{k-2})_{1n} + \sum_{j=2}^{n} A_{1j}\left((A^{k-1})_{jn} + \lambda_1 (A^{k-2})_{jn}\right).$$

Continuing in this fashion, we obtain

$$\begin{aligned}(A^k)_{1n} &= \lambda_1^{k-1} A_{1n} + \sum_{j=2}^{n} A_{1j}\left(\sum_{p=0}^{k-2} \lambda_1^p (A^{k-p-1})_{jn}\right) \\ &= \left(\sum_{p=0}^{k-1} \lambda_1^p \lambda_n^{k-p-1}\right) A_{1n} + \sum_{j=2}^{n-1} A_{1j}\left(\sum_{p=0}^{k-2} \lambda_1^p (A^{k-p-1})_{jn}\right). \end{aligned} \tag{1}$$

The induction hypothesis on the submatrix $[A_{ij}]_{i,j=m}^{n}$, $2 \leqslant m \leqslant n$, says that for any j, $2 \leqslant j \leqslant n$, there are operators B_{jl}, $j+1 \leqslant l \leqslant n$, from H_n to H_j such that

$$(A^k)_{jn} = \sum_{l=j+1}^{n} \frac{\lambda_j^k - \lambda_l^k}{\lambda_j - \lambda_l} B_{jl} \tag{2}$$

for all $k \geqslant 1$. Substituting (2) into (1) yields

$$(A^k)_{1n} = \frac{\lambda_1^k - \lambda_n^k}{\lambda_1 - \lambda_n} A_{1n} + \sum_{j=2}^{n-1} A_{1j}\left(\sum_{p=0}^{k-2} \lambda_1^p \left(\sum_{l=j+1}^{n} \frac{\lambda_j^{k-p-1} - \lambda_l^{k-p-1}}{\lambda_j - \lambda_l} B_{jl}\right)\right). \tag{3}$$

Since

$$\sum_{p=0}^{k-2} \lambda_1^p \frac{\lambda_j^{k-p-1} - \lambda_l^{k-p-1}}{\lambda_j - \lambda_l} = \frac{1}{\lambda_j - \lambda_l}\left(\frac{\lambda_1^k - \lambda_j^k}{\lambda_1 - \lambda_j} - \frac{\lambda_1^k - \lambda_l^k}{\lambda_1 - \lambda_l}\right),$$

from (3) we obtain

$$(A^k)_{1n} = \frac{\lambda_1^k - \lambda_n^k}{\lambda_1 - \lambda_n} A_{1n} + \sum_{j=2}^{n-1} \sum_{l=j+1}^{n} \left(\frac{\lambda_1^k - \lambda_j^k}{\lambda_1 - \lambda_j} - \frac{\lambda_1^k - \lambda_l^k}{\lambda_1 - \lambda_l}\right)\left(\frac{1}{\lambda_j - \lambda_l} A_{1j} B_{jl}\right),$$

which, when expanded, gives the asserted form for $(A^k)_{1n}$. □

We are now ready to prove Theorem 2.1.

Proof of Theorem 2.1. We may assume that $A = A_1 \oplus \cdots \oplus A_q$, where the A_j's are all upper triangular and irreducible. Assume that $\overline{\lim}_k\, c(A^k)^{1/k} \neq 0$. Then, in particular, A is invertible and $\overline{\lim}_k\, c(A_j^k)^{1/k} \neq 0$ for all j. By Lemma 2.6, each A_j is of type III, say, of the form

$$A_j = \begin{bmatrix} \lambda_{j1} I_{j1} & & * \\ & \ddots & \\ 0 & & \lambda_{jm_j} I_{jm_j} \end{bmatrix}, \quad 1 \leqslant j \leqslant q,$$

where $|\lambda_{j1}| = \cdots = |\lambda_{jm_j}|$ for all j. Assume that dist $(0, \sigma(A)) = |\lambda_{11}|$. Then replacing A by A/λ_{11}, we may assume that $\lambda_{11} = 1$ and $|\lambda_{jl}| \geqslant 1$ for all j and l. By Lemma 2.2, there are positive integers n_k, $k \geqslant 1$, such that $\lim_k (\lambda_{jl}/|\lambda_{jl}|)^{n_k} = 1$ for all j and l. Then A^{n_k} behaves asymptotically like $\sum_{j,l} \oplus (\lambda_{jl}^{n_k} I_{jl})$ as $k \to \infty$ by Lemma 2.7. Since the numerical range of the latter matrix is the convex hull of $\{\lambda_{jl}^{n_k} : 1 \leqslant j \leqslant q, 1 \leqslant l \leqslant m_j\}$, we infer that $\lim_k c(A^{n_k}) = 1$. Thus

$$\text{dist}\,(0, \sigma(A)) = 1 = \lim_k c(A^{n_k})^{1/n_k} \leqslant \overline{\lim_k}\, c(A^k)^{1/k} \leqslant \sup_k c(A^k)^{1/k} \leqslant \text{dist}\,(0, \sigma(A)),$$

where the last inequality is by Proposition 1.1 (5). Therefore, $\overline{\lim}_k\, c(A^k)^{1/k} = \sup_k c(A^k)^{1/k} =$ dist $(0, \sigma(A))$ and our assertion follows. □

The above proof also yields the following proposition.

Proposition 2.8. *Let A be an n-by-n matrix. Then $\overline{\lim}_k\, c(A^k)^{1/k} = \sup_k c(A^k)^{1/k}$ if and only if either $\overline{\lim}_k\, c(A^k)^{1/k} \neq 0$ or 0 is in $W(A^k)$ for all $k \geqslant 1$.*

Our last result of this section gives conditions on A for which $\overline{\lim}_k\, c(A^k)^{1/k}$ equals 0.

Proposition 2.9. *Let A be an n-by-n matrix. Then the following conditions are equivalent:*

(1) $\overline{\lim}_k\, c(A^k)^{1/k} = 0$;
(2) *one of the following holds:*

  (a) *A is noninvertible;*
  (b) *A is unitarily equivalent to a matrix of the form $\begin{bmatrix} * & * & * \\ 0 & A' & * \\ 0 & 0 & * \end{bmatrix}$, where A' is not a scalar matrix and has equal eigenvalues;*
  (c) *A is unitarily equivalent to $\begin{bmatrix} * & * & * \\ 0 & A' & * \\ 0 & 0 & * \end{bmatrix}$, where A' is irreducible and has two eigenvalues with unequal moduli;*

(3) 0 *is in $W(A^k)$ for all large k.*

Proof. To prove (1) ⇒ (2), assume that (1) holds and $A = A_1 \oplus \cdots \oplus A_q$, where the A_j's are upper triangular and irreducible. If (a), (b) and (c) are all false, then all the A_j's are invertible and of type III. Arguing as in the proof of Theorem 2.1, we obtain $\overline{\lim}_k\, c(A^k)^{1/k} =$ dist $(0, \sigma(A)) \neq 0$, which contradicts (1). Hence (1) ⇒ (2) is true. Finally, (2) ⇒ (3) is a consequence of Lemma 2.6 and (3) ⇒ (1) is trivial. □

3. Generalized Crawford number

In this section, we consider the limit supremum of $C(A^k)^{1/k}$, $k \geqslant 1$, for a finite matrix A. The next theorem is our main result.

Theorem 3.1. *For any n-by-n matrix A with eigenvalues* $\lambda_1, \ldots, \lambda_n$, *the quantity* $\overline{\lim}_k C(A^k)^{1/k}$ *equals some* $|\lambda_j|$, $1 \leqslant j \leqslant n$.

We start with the 2-by-2 matrices. Its proof is the harbinger of the one for the general case.

Proposition 3.2. *Let* $A = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$. *Then*

$$\overline{\lim_k} C(A^k)^{1/k} = \begin{cases} \min\{|a|, |b|\} & \text{if } c = 0, \\ \max\{|a|, |b|\} & \text{if } c \neq 0. \end{cases}$$

Proof. We consider four cases separately:

(1) $c \neq 0$ and $a = b$. If $a = b = 0$, then $A^k = 0$ and $C(A^k) = 0$ for all $k \geqslant 2$. Thus $\overline{\lim}_k C(A^k)^{1/k} = 0 = \max\{|a|, |b|\}$ in this case. Assume next that $a = b \neq 0$. Then

$$A^k = \begin{bmatrix} a^k & cka^{k-1} \\ 0 & a^k \end{bmatrix} = a^k \begin{bmatrix} 1 & ck/a \\ 0 & 1 \end{bmatrix}$$

and

$$C(A^k) = |a|^k \left(\frac{|c|k}{2|a|} - 1 \right)$$

for all large k. Hence

$$C(A^k)^{1/k} = |a| \left(\frac{|c|k}{2|a|} - 1 \right)^{1/k} \to |a|$$

as $k \to \infty$. This shows that $\overline{\lim}_k C(A^k)^{1/k} = |a| = \max\{|a|, |b|\}$.

(2) $c \neq 0$ and $|a| \neq |b|$. We may assume that $a = 1 > |b|$. Then

$$A^k = \begin{bmatrix} 1 & c(1-b^k)/(1-b) \\ 0 & b^k \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & c/(1-b) \\ 0 & 0 \end{bmatrix} \equiv B$$

as $k \to \infty$. Hence $C(A^k) \to C(B) > 0$ by Proposition 1.1 (7) and thus $C(A^k)^{1/k} \to 1 = \max\{|a|, |b|\}$ as $k \to \infty$.

(3) $c = 0$. On the one hand, we have

$$\overline{\lim_k} C(A^k)^{1/k} \geqslant \overline{\lim_k} c(A^k)^{1/k} = \min\{|a|, |b|\}$$

by Proposition 2.4. On the other, since dist $(0, \partial W(A)) \leqslant$ dist $(0, \sigma(A))$ for any 2-by-2 normal matrix A, we have $C(A^k)^{1/k} \leqslant \min\{|a|, |b|\}$ for all k. Thus $\overline{\lim}_k C(A^k)^{1/k} \leqslant \min\{|a|, |b|\}$ and our asserted equality holds.

(4) $c \neq 0, a \neq b$ and $|a| = |b|$. In this case, $\overline{\lim}_k C(A^k)^{1/k} = |a|$ can be proved as in case (4) of the proof of Proposition 2.4. $\square$

The next two lemmas are useful in the proof of Theorem 3.1.

Lemma 3.3. *Let A be an n-by-n matrix. If* $\overline{\lim}_k C(A^k)^{1/k} = 0$, *then A is not invertible*.

Proof. Our assumption implies that $\overline{\lim}_k\, c(A^k)^{1/k} = 0$. If A is invertible, then, by Proposition 2.9, A is unitarily equivalent to a matrix of the form $\begin{bmatrix} * & * & * \\ 0 & A' & * \\ 0 & 0 & * \end{bmatrix}$, where either A' is not a scalar matrix and has equal eigenvalues or it is irreducible and has two eigenvalues with unequal moduli. In either case, the proof of Lemma 2.6 yields a 2-by-2 nonnormal matrix B such that $r(B) = r(A')$, B^k is a submatrix of A^k for all $k \geqslant 1$, and 0 is in $W(B^k)$ for all large k. Then

$$0 = \overline{\lim_k}\, C(A^k)^{1/k} \geqslant \overline{\lim_k}\, C(B^k)^{1/k} = r(B) = r(A') > 0$$

by Proposition 3.2, which is absurd. Hence A is not invertible. $\square$

A complete characterization of matrices A with $\overline{\lim}_k\, C(A^k)^{1/k} = 0$ is given in Theorem 3.8.

In the following, the convex hull of a subset $\triangle$ of the plane is denoted by $\triangle^\wedge$.

Lemma 3.4. *Let $A = B \oplus C_1 \oplus \cdots \oplus C_q$ be an n-by-n matrix, where the C_j's are all of type III. If 0 is in the interior of $W(A)$ and $W(A) = (W(C_1) \cup \cdots \cup W(C_q))^\wedge$, then $\overline{\lim}_k\, C(A^k)^{1/k} \geqslant \min\{r(C_1), \ldots, r(C_q)\}$.*

Proof. We may assume that $c \equiv \min\{r(C_1), \ldots, r(C_q)\} > 0$ and even $c = 1$ (by considering A/c instead of A). Since 0 is in the interior of $W(A)$, there is an $\varepsilon > 0$ such that $D_\varepsilon \equiv \{z \in \mathbb{C} : |z| < \varepsilon\}$ is contained in $W(A)$. Assume that

$$C_j = \begin{bmatrix} \lambda_{j1} I_{j1} & & * \\ & \ddots & \\ 0 & & \lambda_{jm_j} I_{jm_j} \end{bmatrix}, \quad 1 \leqslant j \leqslant q,$$

where $|\lambda_{j1}| = \cdots = |\lambda_{jm_j}| = r(C_j)$ for all j. By Lemma 2.2, there are positive integers n_k, $k \geqslant 1$, such that $\lim_k (\lambda_{jl}/r(C_j))^{n_k} = 1$ for all j and l. We deduce, using Lemma 2.7, that, for each j, $(C_j/r(C_j))^{n_k} \to I$ as $k \to \infty$ and hence $W(C_j^{n_k+1})/r(C_j)^{n_k}$ is asymptotically close to $W(C_j)$ (in the Hausdorff metric) for large k. Since $D_\varepsilon \subseteq W(A) = (W(C_1) \cup \cdots \cup W(C_q))^\wedge$, we obtain $D_{\varepsilon/2} \subseteq ((W(C_1^{n_k+1})/r(C_1)^{n_k}) \cup \cdots \cup (W(C_q^{n_k+1})/r(C_q)^{n_k}))^\wedge$ for all large k. It follows that

$$D_{\varepsilon/2} \subseteq (W(C_1^{n_k+1}) \cup \cdots \cup W(C_q^{n_k+1}))^\wedge \subseteq W(A^{n_k+1})$$

for large k because $r(C_j) \geqslant 1$ for all j. Hence $C(A^{n_k+1}) \geqslant \varepsilon/2$ for large k and therefore

$$\overline{\lim_k}\, C(A^k)^{1/k} \geqslant \overline{\lim_k}\, C(A^{n_k+1})^{1/(n_k+1)} \geqslant 1 = \min\{r(C_1), \ldots, r(C_q)\}. \quad \square$$

We are now ready for the proof of Theorem 3.1.

Proof of Theorem 3.1. Assume that $A = B_1 \oplus \cdots \oplus B_p \oplus C_1 \oplus \cdots \oplus C_q$, where the B_i's are of type I or II and the C_j's are of type III. Let $c = \overline{\lim}_k\, C(A^k)^{1/k}$ and let the eigenvalues λ_j of A be arranged as $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$. If $c = 0$, then $\lambda_n = 0$ by Lemma 3.3, in which case $c = |\lambda_n|$as required. For the remaining part of the proof, we assume that $c > 0$. Let n_k, $k \geqslant 1$, be positive integers such that $\lim_k C(A^{n_k})^{1/n_k} = c$. We may assume that 0 is in the interior of $W(A^{n_k})$ for all k. Indeed, if there are infinitely many k's, say, k' for which 0 is not in $\operatorname{Int} W(A^{n_{k'}})$, then $C(A^{n_{k'}}) = c(A^{n_{k'}})$ for all such k''s. Hence we have

$$c = \lim_{k'} C(A^{n_{k'}})^{1/n_{k'}} = \lim_{k'} c(A^{n_{k'}})^{1/n_{k'}} \leqslant \overline{\lim_k}\, c(A^k)^{1/k} \leqslant \overline{\lim_k}\, C(A^k)^{1/k} = c.$$

This shows that $c = \overline{\lim}_k\, c(A^k)^{1/k}$. Since $c > 0$, we obtain $c = |\lambda_n|$ by Theorem 2.1. Hence in the following we may assume that $0 \in \operatorname{Int} W(A^{n_k})$ for all k. Since there are only finitely many subsets of the set of summands $\{B_i, C_j : 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant q\}$ of A, by the pigeonhole principle there is a subsequence $\{m_l\}_{l=1}^{\infty}$ of $\{n_k\}_{k=1}^{\infty}$ and there are $B_{i_1},\dots,B_{i_s},C_{j_1},\dots,C_{j_t}$ $(1\leqslant i_1<\cdots<i_s\leqslant p$ and $1\leqslant j_1<\cdots<j_t\leqslant q)$ such that, for all $l\geqslant 1$,

$$W(A^{m_l})=\left(\left(\bigcup_{1\leqslant u\leqslant s}W\left(B_{i_u}^{m_l}\right)\right)\cup\left(\bigcup_{1\leqslant v\leqslant t}W\left(C_{j_v}^{m_l}\right)\right)\right)^{\wedge}, \tag{4}$$

$$\partial W(A^{m_l})\cap\partial W\left(B_{i_u}^{m_l}\right)\neq\emptyset \quad \text{for all } u,\ 1\leqslant u\leqslant s,$$

and

$$\partial W(A^{m_l})\cap\partial W(C_{j_v}^{m_l})\neq\emptyset \quad \text{for all } v,\ 1\leqslant v\leqslant t. \tag{5}$$

Let $d=\min\{r(B_{i_u}),r(C_{j_v}):1\leqslant u\leqslant s,1\leqslant v\leqslant t\}$.We claim that $c=d$.

To prove $c\leqslant d$, let z be any point in $\partial W(A^{m_l})\cap\partial W(B_{i_u}^{m_l})$. We have $C(A^{m_l})\leqslant|z|\leqslant\|B_{i_u}^{m_l}\|$. Hence

$$c=\lim_l C(A^{m_l})^{1/m_l}\leqslant\lim_l\|B_{i_u}^{m_l}\|^{1/m_l}=r(B_{i_u})$$

for all u. Similarly, we have $c\leqslant r(C_{j_v})$ for all v. Thus $c\leqslant d$ as asserted.

For the proof of $c\geqslant d$, we assume that $d>0$. Two cases are considered separately: (a) $s\geqslant 1$. In this case, the proof is analogous to the one for Lemma 3.3. Indeed, assume that $\min\{r(B_{i_u}):1\leqslant u\leqslant s\}=r(B_{i_1})$. Since B_{i_1} is of type I or type II, by the proof of Lemma 2.6, there is a 2-by-2 nonnormal matrix B such that $r(B)=r(B_{i_1})$, B^k is a submatrix of $B_{i_1}^k$ for all $k\geqslant 1$ and 0 is in $W(B^k)$ for all large k. Then

$$c=\overline{\lim_k}\,C(A^k)^{1/k}\geqslant\overline{\lim_k}\,C(B^k)^{1/k}=r(B)=r(B_{i_1})\geqslant d,$$

where the equality of $\overline{\lim}_k\,C(B^k)^{1/k}$ and $r(B)$ is by Proposition 3.2. This gives $c\geqslant d$ in case (a). (b) $s=0$. In this case, since 0 is in the interior of $W(A^{m_1})$ and $W(A^{m_1})=(\cup_{1\leqslant v\leqslant t}W(C_{j_v}^{m_1}))^{\wedge}$ from (4), we have $\overline{\lim}_k\,C(A^{m_1k})^{1/k}\geqslant\min\{r(C_{j_v}^{m_1}):1\leqslant v\leqslant t\}=d^{m_1}$ by Lemma 3.4 and hence

$$c\geqslant\overline{\lim_k}\,C(A^{m_1k})^{1/(m_1k)}\geqslant d$$

as desired. This shows that $c=d$. $\square$

For normal matrices, more can be said.

Proposition 3.5. *If A is an n-by-n $(n\geqslant 3)$ normal matrix with eigenvalues $\lambda_1,\dots,\lambda_n$ satisfying $|\lambda_1|\geqslant\cdots\geqslant|\lambda_n|$, then $\overline{\lim}_k\,C(A^k)^{1/k}=\sup_k C(A^k)^{1/k}=|\lambda_j|$ for some j, $3\leqslant j\leqslant n$.*

Proof. Let $c=\overline{\lim}_k\,C(A^k)^{1/k}$ and assume that $A=\operatorname{diag}(\lambda_1,\dots,\lambda_n)$. Then A can be considered as the direct sum $[\lambda_1]\oplus\cdots\oplus[\lambda_n]$ of the 1-by-1 type-III matrices $[\lambda_j]$. Arguing as in the proof of Theorem 3.1, we see that if either $c=0$ or 0 is not in $\operatorname{Int}W(A^{n_k})$ for infinitely many indices n_k with $\lim_k C(A^{n_k})^{1/n_k}=c$, then $c=|\lambda_n|$. For the remaining case, it was shown in the proof of Theorem 3.1 that $c=\min\{|\lambda_{j_v}|:1\leqslant v\leqslant t\}=|\lambda_{j_t}|$ for some indices j_v $(1\leqslant j_1<\cdots<j_t\leqslant n)$. Since 0 is in $\operatorname{Int}W(A^{m_1})=\operatorname{Int}\{\lambda_{j_v}^{m_1}:1\leqslant v\leqslant t\}^{\wedge}$ by (4) and $\lambda_{j_v}^{m_1}$ is in $\partial W(A^{m_1})$ for all v by (5), we infer that $t\geqslant 3$ and hence $j_t\geqslant 3$.

To prove that $c=\sup_k C(A^k)^{1/k}$, we need only check $C(A^{k_0})^{1/k_0}\leqslant c$ for all $k_0\geqslant 1$. If 0 is not in $W(A^{k_0})$, then

$$C(A^{k_0})^{1/k_0}=c(A^{k_0})^{1/k_0}\leqslant|\lambda_n|\leqslant c$$

by Proposition 1.1 (5) and Theorem 3.1. If 0 is in $\partial W(A^{k_0})$, then $C(A^{k_0})^{1/k_0}=0\leqslant c$. Hence we may assume that 0 is in $\operatorname{Int}W(A^{k_0})$. Assume further that $W(A^{k_0})$ is the polygonal region with vertices $\lambda_{j_1}^{k_0},\dots,\lambda_{j_t}^{k_0}$ $(3\leqslant t\leqslant n$ and $1\leqslant j_1<\cdots<j_t\leqslant n)$ and $\lambda_{j_t}=1$ (by considering A/λ_{j_t} instead of A). By Lemma 2.2, there are positive integers n_k, $k\geqslant 1$, such that $\lim_k(\lambda_j^{k_0}/|\lambda_j^{k_0}|)^{n_k}=1$ for all nonzero λ_j,

$1 \leqslant j \leqslant n$. Letting $m_k = k_0(n_k + 1)$ for $k \geqslant 1$, we infer that $W(A^{m_k})$ is asymptotically close to $W(A^{k_0})$ (in the Hausdorff metric) and hence $\{z \in \mathbb{C} : |z| < C(A^{k_0})/2\} \subseteq W(A^{m_k})$ for all large k. Thus $C(A^{m_k}) \geqslant C(A^{k_0})/2$ for all large k. It follows that

$$c \geqslant \overline{\lim_k} C(A^{m_k})^{1/m_k} \geqslant 1 = \lambda_{j_t} \geqslant C(A^{k_0})^{1/k_0}$$

as required. □

Note that, for an n-by-n normal matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$ satisfying $|\lambda_1| \geqslant \cdots \geqslant |\lambda_n|$, if $n \leqslant 3$, then $\overline{\lim}_k C(A^k)^{1/k} = |\lambda_n|$ by Propositions 3.2 and 3.5 while if $n \geqslant 4$, then $\overline{\lim}_k C(A^k)^{1/k}$ can be any of $|\lambda_j|$, $3 \leqslant j \leqslant n$. This is seen by the following example.

Example 3.6. For any $n \geqslant 4$ and any j_0, $3 \leqslant j_0 \leqslant n$, let

$$\lambda_j = \begin{cases} (j_0 - j + 1)\exp(2\pi i/3) & \text{if } 1 \leqslant j \leqslant j_0 - 2, \\ 2\exp(4\pi i/3) & \text{if } j = j_0 - 1, \\ 1 & \text{if } j = j_0, \\ 1/j & \text{if } j_0 < j \leqslant n, \end{cases}$$

and let $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Then $C(A^k) \leqslant 1$ for all $k \geqslant 1$ and $\lim_k C(A^{3k+1}) = \sqrt{3}/2$. Hence $\overline{\lim}_k C(A^k)^{1/k} = 1 = |\lambda_{j_0}|$.

Similarly, for a general matrix A, $\overline{\lim}_k C(A^k)^{1/k}$ can be any of the $|\lambda_j|$'s as the preceding example (for $n \geqslant 4$ and $j \geqslant 3$) and the next (for $n \geqslant 3$ and $j = 1$ or 2) show.

Example 3.7. For any $n \geqslant 3$, let

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1/2 \end{bmatrix} \oplus \operatorname{diag}\left(\frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}\right)$$

and

$$A_2 = [1] \oplus \begin{bmatrix} 1/2 & 1 \\ 0 & 1/3 \end{bmatrix} \oplus \operatorname{diag}\left(\frac{1}{4}, \frac{1}{5}, \ldots, \frac{1}{n}\right).$$

Using Proposition 3.2, we can easily show that $\overline{\lim}_k C(A_1^k)^{1/k} = 1$ and $\overline{\lim}_k C(A_2^k)^{1/k} = 1/2$.

We conclude this section with the following theorem characterizing those matrices A with $\overline{\lim}_k C(A^k)^{1/k} = 0$.

Theorem 3.8. *For an n-by-n matrix A, $\overline{\lim}_k C(A^k)^{1/k} = 0$ if and only if there is an $m \geqslant 1$ such that $0 \notin \partial W(A^{m-1})$ and $0 \in \partial W(A^k)$ for all $k \geqslant m$. In this case, A is unitarily equivalent to a matrix of the form $B \oplus C$, where B must be present satisfying $B^{m-1} \neq 0$ and $B^m = 0$ for some $m \geqslant 1$, and C is invertible* (*but may be absent*).

Proof. Assume that $c \equiv \overline{\lim}_k C(A^k)^{1/k} = 0$ and $A = B_1 \oplus \cdots \oplus B_p \oplus C_1 \oplus \cdots \oplus C_q$, where the B_i's are of type I or II and the C_j's are of type III. We first check that the B_i's are all nilpotent. Indeed, if $r(B_{i_0}) > 0$ for some i_0, then, by the proof of Lemma 2.6, there is a 2-by-2 nonnormal matrix B' such that $r(B') = r(B_{i_0})$, B'^k is a submatrix of $B_{i_0}^k$ for all $k \geqslant 1$ and 0 is in $W(B'^k)$ for all large k. Then

$$c \geqslant \overline{\lim_k} C(B'^k)^{1/k} = r(B') = r(B_{i_0}) > 0,$$

where the equality of $\overline{\lim}_k C(B'^k)^{1/k}$ and $r(B')$ is by Proposition 3.2. This leads to a contradiction. Hence the B_i's must all be nilpotent. Let B be the direct sum of the B_i's together with those nilpotent C_j's. Say, $A = B \oplus C_1 \oplus \cdots \oplus C_{q'}$, where the C_j's, $1 \leqslant j \leqslant q'$, are all invertible type-III matrices. Assume that $0 \notin \partial W(A^k)$ for infinitely many k's. Since $c = 0$, Theorem 3.1 implies that $0 \in \sigma(A)$ and hence $0 \in W(A^k)$

for all $k \geqslant 1$. Thus 0 is in Int $W(A^k)$ for infinitely many k's. Let $k_0 \geqslant n$ be such that $0 \in \text{Int}\, W(A^{k_0})$. Then $A^{k_0} = 0 \oplus C_1^{k_0} \oplus \cdots \oplus C_{q'}^{k_0}$, and hence

$$W(A^{k_0}) = \left(\{0\} \cup W\left(C_1^{k_0}\right) \cup \cdots \cup W\left(C_{q'}^{k_0}\right)\right)^{\wedge} = \left(\{0\} \cup \left(W\left(C_1^{k_0}\right) \cup \cdots \cup W\left(C_{q'}^{k_0}\right)\right)^{\wedge}\right)^{\wedge}.$$

If 0 is not in $(W(C_1^{k_0}) \cup \cdots \cup W(C_{q'}^{k_0}))^{\wedge}$, then it will be in $\partial W(A^{k_0})$, a contradiction. Hence we have $W(A^{k_0}) = (W(C_1^{k_0}) \cup \cdots \cup W(C_{q'}^{k_0}))^{\wedge}$. From Lemma 3.4, we infer that

$$c \geqslant \overline{\lim_k}\, C\left(A^{k_0 k}\right)^{1/(k_0 k)} \geqslant \left(\min\left\{r\left(C_1^{k_0}\right), \ldots, r\left(C_{q'}^{k_0}\right)\right\}\right)^{1/k_0} = \min\left\{r(C_1), \ldots, r(C_{q'})\right\} > 0,$$

which leads to a contradiction. Hence there is an $m \geqslant 1$ such that $0 \notin \partial W(A^{m-1})$ and $0 \in \partial W(A^k)$ for all $k \geqslant m$. Conversely, under this condition, we obviously have $C(A^k) = 0$ for all $k \geqslant m$ and hence $c = 0$.

In this case, the arguments in the preceding paragraph yield $A = B \oplus C_1 \oplus \cdots \oplus C_{q'}$ as above. We claim that $0 \in \partial W(A^m)$ implies $B^m = 0$. Indeed, if otherwise, then B^m, being a nonzero nilpotent matrix, is unitarily equivalent to an upper-triangular matrix $[b_{ij}]$ with $b_{ij} = 0$ for all $i \geqslant j$ and $b_{i_0 j_0} \neq 0$ for some $i_0 < j_0$. Therefore, 0 is in Int $W\left(\begin{bmatrix} 0 & b_{i_0 j_0} \\ 0 & 0 \end{bmatrix}\right)$ and hence in Int $W(A^m)$, which contradicts our assumption of 0 in $\partial W(A^m)$. Thus we have $B^m = 0$.

Note that B must be present in the decomposition $A = B \oplus C_1 \oplus \cdots \oplus C_{q'}$ of A. This is because if otherwise then $0 = c \geqslant \text{dist}\,(0, \sigma(A))$ (by Theorem 3.1) implies that $A = C_1 \oplus \cdots \oplus C_{q'}$ is noninvertible, a contradiction.

Finally, assume that $B^{m-1} = 0$. Since A is noninvertible, 0 is in $\sigma(A^{m-1})$ and hence in $W(A^{m-1})$. Our assumption $0 \notin \partial W(A^{m-1})$ implies that $0 \in \text{Int}\, W(A^{m-1})$. On the other hand, since $A^{m-1} = 0 \oplus C_1^{m-1} \oplus \cdots \oplus C_{q'}^{m-1}$, we have

$$\begin{aligned} W(A^{m-1}) &= \left(\{0\} \cup W\left(C_1^{m-1}\right) \cup \cdots \cup W\left(C_{q'}^{m-1}\right)\right)^{\wedge} \\ &= \left(\{0\} \cup \left(W\left(C_1^{m-1}\right) \cup \cdots \cup W\left(C_{q'}^{m-1}\right)\right)^{\wedge}\right)^{\wedge}. \end{aligned}$$

If 0 is not in $(W(C_1^{m-1}) \cup \cdots \cup W(C_{q'}^{m-1}))^{\wedge}$, then it must be in $\partial W(A^{m-1})$, a contradiction. Thus $W(A^{m-1}) = (W(C_1^{m-1}) \cup \cdots \cup W(C_{q'}^{m-1}))^{\wedge}$. Lemma 3.4 then implies that

$$\begin{aligned} c \geqslant \overline{\lim_k}\, C(A^{(m-1)k})^{1/((m-1)k)} &\geqslant \left(\min\left\{r\left(C_1^{m-1}\right), \ldots, r\left(C_{q'}^{m-1}\right)\right\}\right)^{1/(m-1)} \\ &= \min\{r(C_1), \ldots, r(C_{q'})\} > 0, \end{aligned}$$

again a contradiction. This shows that $B^{m-1} \neq 0$ as asserted. □

As pointed out by the referee, the decomposition $B \oplus C$ of A in the preceding theorem is a special type of the Fitting decomposition of A (cf. [8, pp. 151–152]).

References

[1] C.R. Crawford, A stable generalized eigenvalue problem, SIAM J. Numer. Anal. 13 (1976) 854–860.
[2] H.-L. Gau, P.Y. Wu, Numerical ranges of nilpotent operators, Linear Algebra Appl. 429 (2008) 716–726.
[3] K. Gustafson, D.K.M. Rao, Numerical Range. The Field of Values of Linear Operators and Matrices, Springer, New York, 1997.
[4] P.R. Halmos, A Hilbert Space Problem Book, second ed., Springer, New York, 1982.
[5] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge Univ. Press, Cambridge, 1991.
[6] Y. Katznelson, An Introduction to Harmonic Analysis, second ed., Dover, New York, 1976.
[7] R. Mathias, Two theorems on singulars values and eigenvalues, Amer. Math. Monthly 97 (1990) 47–50.
[8] V.S. Varadarajan, Lie Groups, Lie Algebras, and their Representations, Springer, New York, 1984.
[9] T. Yamamoto, On the extreme values of the roots of matrices, J. Math. Soc. Japan 19 (1967) 173–178.
[10] E. Makai Jr., J. Zemánek, The surjectivity radius, packing numbers and boundedness below of linear operators, Integral Equations Operator Theory 6 (1983) 372–384.