Numerical Ranges of the Powers of an Operator

Man-Duen Choi[1]
Department of Mathematics, University of Toronto,
Toronto, Ontario, Canada M5S 2E4
E-mail: choi@math.toronto.edu

Chi-Kwong Li[2]
Department of Mathematics, The College of William and Mary,
Williamsburg, VA 23187, USA
E-mail: ckli@math.wm.edu

Abstract

The numerical range $W(A)$ of a bounded linear operator A on a Hilbert space is the collection of complex numbers of the form (Av, v) with v ranging over the unit vectors in the Hilbert space. In terms of the location of $W(A)$, inclusion regions are obtained for $W(A^k)$ for positive integers k, and also for negative integers k if A^{-1} exists. Related inequalities on the numerical radius $w(A) = \sup\{|\mu| : \mu \in W(A)\}$ and the Crawford number $c(A) = \inf\{|\mu| : \mu \in W(A)\}$ are deduced.

AMS Subject Classification 47A12, 47A50, 15A60.

Keywords Numerical range, numerical radius, Crawford number, bounded linear operator, Hilbert space.

1 Introduction

Let $\mathcal{H}$ be a Hilbert space with the inner product (u, v), and let $\mathcal{B}(\mathcal{H})$ be the set of bounded linear operators on $\mathcal{H}$. The *numerical range* of $A \in \mathcal{B}(\mathcal{H})$ is defined by

$$W(A) = \{(Av, v) : v \in \mathcal{S},\ (v, v) = 1\}.$$

Furthermore, define the *numerical radius* and the *Crawford number* of A by

$$w(A) = \sup\{|z| : z \in W(A)\} \qquad \text{and} \qquad c(A) = \inf\{|z| : z \in W(A)\},$$

respectively. These concepts are useful in studying linear operators and have attracted the attention of many authors in the last few decades (e.g., see [2, 7, 12, 13], and their references). In applications of these concepts to other areas such as perturbation theory, generalized eigenvalue problems, numerical analysis, system theory, and dilation theory (e.g., see [1, 4, 5, 9, 10, 12, 15, 18]), it is useful to know the location of $W(A^k)$ for positive integers k and also negative integers k if A is invertible. The following facts are well-known; see [12, 13] for example.

[1]Research partially supported by a Canada NSERC grant.
[2]Research partially support by a USA NSF grant and the W&M Plumeri award. Li is an honorary professor of the University of Hong Kong and an honorary professor of the Taiyuan University of Technology.

Facts 1.1 *Let $A \in \mathcal{B}(\mathcal{H})$ and k be a positive integer.*

(a) *If $W(A) \subseteq \mathbf{D}$ with $\mathbf{D} = \{z \in \mathbb{C} : |z| \le 1\}$ then $W(A^k) \subseteq \mathbf{D}$.*

(b) *If $W(A) \subseteq S = \{re^{it} : r \ge 0,\ t \in [-t_0, t_0]\}$ for some $t_0 \in [0, \pi/2)$ and A is invertible, then $W(A^{-1}) \subseteq S$ as well.*

(c) *Suppose A is normal. Then $\overline{W(A^k)}$ is the convex hull of the spectrum of A^k, which is a subset of the convex hull of the set*
$$\overline{W(A)}^k = \{\eta^k : \eta \in \overline{W(A)}\}.$$
If, in addition, $0 \notin \overline{W(A)}$, then the above conclusion holds for negative integers k as well.

It is desirable to find a good inclusion region for $W(A^k)$ in terms of $W(A)$. In view of (a)—(c) above, one may wonder whether it is always the case that $\overline{W(A^k)}$ is a subset of the convex hull of $\overline{W(A)}^k = \{z^k : z \in \overline{W(A)}\}$ for a general operator. It is not true as shown in the following example.

Example 1.2 *Let $k > 1$ be a positive integer. If $A = \begin{pmatrix} 1 & 2s \\ 0 & 1 \end{pmatrix}$ with $s = \sin(\pi/(2k))$, then $1 < ks$ and $W(A) = \{z \in \mathbb{C} : |z - 1| \le s\}$ and*
$$W(A) \subseteq \left\{ re^{it} \in \mathbb{C} : r \ge 0,\ \frac{-\pi}{2k} \le t \le \frac{\pi}{2k} \right\}.$$
So, $\{\eta_1 \cdots \eta_k : \eta_i \in W(A), i = 1, \ldots, k\}$ is a subset of the right half plane $\{z \in \mathbb{C} : \operatorname{Re} z \ge 0\}$. Since $A^k = \begin{pmatrix} 1 & 2ks \\ 0 & 1 \end{pmatrix}$, it follows that $W(A^k)$ is the circular disk with radius ks centered at 1. In particular, 0 is an interior point of $W(A^k)$. So,
$$W(A^k) \not\subseteq \operatorname{conv}\{\eta_1 \cdots \eta_k : \eta_i \in W(A), i = 1, \ldots, k\}.$$

In this paper, we will study the location of $W(A^k)$ for integers k. Related inequalities on the numerical radius and the Crawford number of A^k will be obtained. We will focus on positive powers of operators A in Section 2, and turn to negative powers in Section 3.

In our discussion, we will identify $\mathcal{H}$ with $\mathbb{C}^n$ if $\mathcal{H}$ has dimension n. In such a case, we will identify $\mathcal{B}(\mathcal{H})$ with M_n. The following basic facts (see [13, 12]) will be used frequently in our discussion.

(1) the spectrum $\sigma(A)$ of A is always a subset of the closure of $W(A)$,

(2) $W(aA + bI) = aW(A) + b$ for any $a, b \in \mathbb{C}$,

(3) $W(B \oplus C) = \operatorname{conv}\{W(B) \cup W(C)\}$, where $\operatorname{conv} S$ denotes the convex hull of $S \subseteq \mathbb{C}$.

(4) If $A \in M_2$ has eigenvalues λ_1, λ_2, then $W(A)$ is an elliptical disk with foci λ_1, λ_2 and the length of minor axis equals $\{\operatorname{tr}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2\}^{1/2}$.

2 Positive powers of operators

2.1 Inclusion regions of $W(A^k)$ in terms of $W(A)$

Let k be a positive integer, and $A \in \mathcal{B}(\mathcal{H})$. As shown in Section 1, the inclusion $W(A^k) \subseteq W(A)^k$ is not true in general. But there are situations under which the inclusion is valid. Here is a scheme to amend the situation so that $W(A)$ will be useful in estimating $W(A^k)$.

Let $\mu = e^{i2\pi/k}$ and $\widetilde{A} = A \oplus \mu A \oplus \cdots \oplus \mu^{k-1}A$. Then $W(\widetilde{A}^k) = W(A^k)$. It is natural to use $W(\widetilde{A})$ to determine an inclusion region for $W(A^k)$. Denote by $\mathcal{C}^k = \{\eta^k : \eta \in \mathcal{C}\}$ for any $\mathcal{C} \subseteq \mathbb{C}$. We have the following result.

Theorem 2.1 *Suppose $A \in \mathcal{B}(\mathcal{H})$ and $\mu = e^{i2\pi/k}$ for a positive integer $k > 1$. Let $\widetilde{A} = A \oplus \mu A \oplus \cdots \oplus \mu^{k-1}A$. Then*

$$W(\widetilde{A})^k = \{\mathrm{conv}\,[\cup_{j=1}^k \mu^j W(A)]\}^k$$

is a convex set satisfying the following inclusion

$$W(A^k) = W(\widetilde{A}^k) \subseteq W(\widetilde{A})^k.$$

To show that $W(\widetilde{A})^k$ is convex, we prove the following general result on convex subsets in $\mathbb{C}$, which is of independent interest.

Proposition 2.2 *Suppose $\mathcal{C} \subseteq \mathbb{C}$ is convex and satisfies $\mathcal{C} = \mu\mathcal{C}$ with $\mu = e^{i2\pi/k}$. Then the set $\mathcal{C}^k = \{\nu^k : \nu \in \mathcal{C}\}$ is convex.*

Proof. Suppose $\alpha, \beta \in \mathcal{C}$. We need only to show that the line segment joining α^k and β^k lies in the set $\mathcal{C}^k$. The conclusion is clear if α or β is 0. Suppose it is not the case. We may assume that the argument of $\alpha^{-1}\beta$ lies in $[0, \pi/k]$. Otherwise, replace β by $e^{i2j\pi/k}\beta$ for a suitable $j \in \{1, \ldots, k-1\}$. Let $\Delta = \mathrm{conv}\,\{0, \alpha, \beta\}$. Then $\{z^k : z \in \Delta\}$ is a convex set with boundary:

$$\{t\alpha^k : 0 \le t \le 1\} \cup \{t\beta^k : 0 \le t \le 1\} \cup \{[t\alpha + (1-t)\beta]^k : 0 \le t \le 1\}.$$

□

The following identity will be useful in our discussion.

Lemma 2.3 *Let $T \in \mathcal{B}(\mathcal{H})$, $a \in \mathbb{C}$, and $\mu = e^{i2\pi/k}$ for a positive integer $k > 1$. If $g_j(z) = (a^k - z^k)/(a - \mu^j z)$ for $j = 1, \ldots, k$, then*

$$k\bar{a}^{k-1}(a^k I - T^k) = \sum_{j=1}^k g_j(T)^*(aI - \mu^j T)g_j(T).$$

Proof. Let $g_j(z) = (a^k - z^k)/(a - \mu^j z)$ for $j = 1, \ldots, k$. Then

$$\sum_{j=1}^k g_j(z) = \frac{d}{da}(a^k - z^k) = ka^{k-1}$$

is independent of z. Thus, $k\bar{a}^{k-1} = \sum_{j=1}^{k} \overline{g_j(z)}$ and

$$k\overline{a}^{k-1}(a^k I - T^k) = \sum_{j=1}^{k} g_j(T)^*(a^k I - T^k) = \sum_{j=1}^{k} g_j(T)^*(aI - \mu^j T)g_j(T). \qquad \square$$

We are now ready to present the following.

Proof of Theorem 2.1. Since $W(B \oplus C) = \operatorname{conv}\{W(B) \cup W(C)\}$, if $\widetilde{A}$ is defined as in the theorem then $W(A^k) = W(\widetilde{A}^k)$ and $W(\widetilde{A})^k = \{\operatorname{conv}[\cup_{j=1}^{k} \mu^j W(A)]\}^k$. Applying Proposition 2.2 to $\mathcal{C} = W(\widetilde{A})$, we see that $W(\widetilde{A})^k$ is convex.

It remains to show that $W(\widetilde{A}^k) \subseteq W(\widetilde{A})^k$. By Lemma 2.3, if $g_j(z) = (a^k - z^k)/(a - \mu^j z)$ for $j = 1, \dots, k$, then

$$k\overline{a}^{k-1}(a^k I - \widetilde{A}^k) = \sum_{j=1}^{k} g_j(\widetilde{A})^*(aI - \mu^j \widetilde{A})g_j(\widetilde{A}). \tag{2.1}$$

Now, suppose $b = a^k \in W(A^k) = W(\widetilde{A}^k)$. Then there is a unit vector $v \in \mathcal{H}$ such that

$$0 = ((a^k I - \widetilde{A}^k)v, v).$$

We consider two cases.

Case 1. Suppose $g_j(\widetilde{A})v$ is nonzero for each $j = 1, \dots, k$ in (2.1). Let

$$K = W(aI - \mu\widetilde{A}) = \cdots = W(aI - \mu^{k-1}\widetilde{A}) = W(aI - \widetilde{A}).$$

By (2.1), if

$$\eta_j = ((aI - \mu^j \widetilde{A})g_j(\widetilde{A})v, g_j(\widetilde{A})v)/\|g_j(\widetilde{A})v\|^2 \in W(aI - \mu^j \widetilde{A})$$

for $j = 1, \dots, k$, then for $\gamma = \sum_{j=1}^{k} \|g_j(\tilde{A})v\|^2$,

$$0 = \gamma^{-1}((a^k I - \widetilde{A}^k)v, v) = \gamma^{-1} \sum_{j=1}^{k} \|g_j(\widetilde{A})v\|^2 \eta_j$$

is a convex combination of k elements in the convex set $K = W(aI - \widetilde{A})$. So, $a \in W(\widetilde{A})$ and $a^k \in W(\widetilde{A})^k$.

Case 2. Suppose there is $j \in \{1, \dots, k\}$ such that

$$0 = g_j(\widetilde{A})v = \prod_{1 \le q \le k, q \ne j} (aI - \mu^q \widetilde{A})v.$$

Then there is $q \in \{1, \dots, k\}$ and a nonzero vector y such that $(a - \mu^q \widetilde{A})y = 0$. It follows that $a \in W(\mu^q \widetilde{A}) = W(\widetilde{A})$. Thus, $a^k \in W(\widetilde{A})^k$. $\square$

Note that one can also deduce Theorem 2.1 from [14, Theorem 1] once Proposition 2.2 is verified. Also, one can use Proposition 2.2 and arguments similar to those in the proof of Theorem 2.1 to prove the following general version of Theorem 2.1.

Theorem 2.4 *Let $k > 1$ be a positive integer. Suppose $A \in \mathcal{B}(\mathcal{H})$ is such that $W(A) \subseteq \Gamma$, where Γ is a convex subset of $\mathbb{C}$ satisfying $e^{i2\pi/k}\Gamma = \Gamma$. Then $W(A^k) \subseteq \operatorname{conv}\{z^k : z \in \Gamma\}$.*

Now, if one considers Γ to be the set of k-sided polygons if $k \geq 3$ or parallel strips when $k = 2$ containing $W(A)$, then the intersection of all such Γ equals $W(\widetilde{A})$ in Theorem 2.1, and we have

$$W(A^k) \subseteq \{\mu^k : \mu \in W(\widetilde{A})\} \subseteq \bigcap_{\Gamma}\{\mu^k : \mu \in \Gamma\}.$$

2.2 Inclusion regions for $W(A^k)$ in terms of $w(A)$ and $c(A)$

In this subsection, we study inclusion regions for $W(A^k)$ in terms of the numerical radius and the Crawford number of A when $W(A)$ does not contain the origin. In particular, it provides estimates for $c(A^k)$ in terms of $c(A)$.

Theorem 2.5 *Let $A \in \mathcal{B}(\mathcal{H})$ be such that $W(A)$ is a subset of the segment*

$$\{z \in \mathbb{C} : |z| \leq 1,\ \operatorname{Re} z \geq \cos\phi\}$$

for some $\phi \in [0, \pi/2]$. For any positive integer $m = 2k$ or $2k-1$ such that $2k\phi \leq \pi$, we have $W(A^m)$ is a subset of the segment

$$\{z \in \mathbb{C} : |z| \leq 1,\ \operatorname{Re} z \geq \cos(m\phi)\}.$$

Proof. By the power inequality, $w(A^m) \leq w(A)^m$ for any positive integer m. Thus, $W(A^m)$ lies inside the unit disk. Denote by $A^m = H_m + iG_m$ with $H_m = H_m^*$ and $G_m = G_m^*$. It remains to show that:

$$\textit{For } m = 2k-1 \textit{ or } k,\quad H_m \geq (\cos(m\phi))I.$$

Here $X \geq Y$ means that $X - Y$ is positive semi-definite.

For $m = 1$, we have $H_1 \geq (\cos\phi)I$ by the given assumption. For $m = 2$, we have $H_1^2 \geq (\cos^2\phi)I$. Since $W(A)$ lies inside the unit disk and $H_1 \geq (\cos\phi)I$, we have $-(\sin\phi)I \leq G_1 \leq (\sin\phi)I$ and hence $G_1^2 \leq (\sin^2\phi)I$. By the fact that

$$A^2 = (H_1 + iG_1)^2 = (H_1^2 - G_1^2) + i(H_1G_1 + G_1H_1),$$

we have

$$H_2 = H_1^2 - G_1^2 \geq (\cos^2\phi)I - (\sin^2\phi)I = (\cos(2\phi))I.$$

Next, for $m = 2k-1$ or $2k$ with $k > 1$, we assume by induction hypothesis that

$$H_{2\tilde{k}-1} \geq \cos((2\tilde{k}-1)\phi)I \quad \text{and} \quad H_{2\tilde{k}} \geq (\cos(2\tilde{k})\phi)I$$

for $\tilde{k} < k$. In particular, $H_{k-1} \geq (\cos((k-1)\phi))I$ and $H_k \geq (\cos(k\phi))I$. Since $W(A^k)$ and $W(A^{k-1})$ lies inside the closed unit disk, it follows that

$$-(\sin((k-1)\phi))I \leq G_{k-1} \leq (\sin((k-1)\phi))I \quad \text{and} \quad -(\sin(k\phi))I \leq G_k \leq (\sin(k\phi))I \tag{2.2}$$

For $m = 2k$, we have $H_k \geq (\cos(k\phi))I$ by induction assumption, and $G_k^2 \leq (\sin^2(k\phi))I$ by (2.2). Thus,

$$H_{2k} = H_k^2 - G_k^2 \geq (\cos^2(k\phi))I - (\sin^2(k\phi))I = (\cos(2k\phi))I.$$

Suppose $m = 2k-1$. Note that

$$\begin{aligned} G_k &= (-i/2)[A^k - (A^*)^k] \\ &= (-i/2)[AA^{k-1} - A^*(A^*)^{k-1}] \\ &= (-i/2)[(H_1 + iG_1)(H_{k-1} + iG_{k-1}) - (H_1 - iG_1)(H_{k-1} - iG_{k-1})] \\ &= H_1 G_{k-1} + G_1 H_{k-1}. \end{aligned}$$

Also, since $G_k = G_k^*$, we have

$$G_k = H_1 G_{k-1} + G_1 H_{k-1} = G_{k-1} H_1 + H_{k-1} G_1. \tag{2.3}$$

Writing $A^m = A^{k-1}AA^{k-1} = (H_{k-1} + iG_{k-1})(H_1 + iG_1)(H_{k-1} + iG_{k-1})$, we have

$$\begin{aligned} 2H_m &= 2H_{k-1}H_1H_{k-1} - G_{k-1}G_1H_{k-1} - H_{k-1}G_1G_{k-1} \\ &\qquad - G_{k-1}(H_1G_{k-1} + G_1H_{k-1}) - (G_{k-1}H_1 + H_{k-1}G_1)G_{k-1}. \end{aligned}$$

By (2.3),

$$2H_m = H_{k-1}[2H_1 - H_{k-1}^{-1}G_{k-1}G_1 - G_1G_{k-1}H_{k-1}^{-1}]H_{k-1} - (G_kG_{k-1} + G_{k-1}G_k). \tag{2.4}$$

Since $H_{k-1} \geq (\cos((k-1)\phi))I$ by induction assumption, and $\|G_{k-1}\| \leq \sin((k-1)\phi)$ by (2.2),

$$\|H_{k-1}^{-1}G_{k-1}G_1 + G_1G_{k-1}H_{k-1}^{-1}\| \leq 2\|G_1\|\|G_{k-1}\|\|H_{k-1}^{-1}\| \leq \frac{2\sin\phi\sin((k-1)\phi)}{\cos((k-1)\phi)}.$$

It follows that

$$2H_1 - H_{k-1}^{-1}G_{k-1}G_1 - G_1G_{k-1}H_{k-1}^{-1} \geq 2(\cos\phi)I - \frac{2\sin\phi\sin((k-1)\phi)I}{\cos((k-1)\phi)} = \frac{2(\cos(k\phi))I}{\cos((k-1)\phi)}. \tag{2.5}$$

By (2.2) again, we have

$$\|G_kG_{k-1} + G_{k-1}G_k\| \leq 2\|G_k\|\|G_{k-1}\| \leq 2\sin(k\phi)\sin((k-1)\phi). \tag{2.6}$$

Putting (2.5) and (2.6) in (2.4) and using the assumption that $H_{k-1} \geq (\cos((k-1)\phi))I$, we have

$$2H_m \geq \frac{2\cos(k\phi)}{\cos((k-1)\phi)}H_{k-1}H_{k-1} - 2\sin((k-1)\phi)\sin(k\phi)I$$

$$
\begin{aligned}
&\geq 2\cos(k\phi)\cos((k-1)\phi)I - 2\sin((k-1)\phi)\sin(k\phi)I \\
&= 2(\cos((2k-1)\phi))I,
\end{aligned}
$$

which is the desired inequality. □

Clearly, if $W(A)$ lies in a segment of a circle, we have bounds for $w(A)$ and $c(A)$. By the power inequality, we know that $w(A^k) \leq w(A)^k$. By these facts and Theorem 2.5, one can use $w(A)$ and $c(A)$ to obtain bounds for $c(A^m)$ as follows.

Theorem 2.6 *Let m be a positive integer and $A \in \mathcal{B}(\mathcal{H})$ be nonzero such that $c(A) = w(A)\cos\phi$ with $m\phi \in [0, \pi/2]$. Then*

$$
c(A^m)/w(A^m) \geq c(A^m)/w(A)^m \geq \cos(m\phi).
$$

Proof. The result is trivial if $m = 1$. Assume that $m > 1$. Then $m\phi \in [0, \pi/2]$ implies that $c(A) \geq w(A)\cos(\phi) > 0$. Hence, A is invertible, and so is A^m. Thus, $0 < w(A^m)$ and $w(A^m) \leq w(A)^m$ by the power inequality. The first inequality in the theorem follows.

To prove the second inequality, we may replace A by $A/w(A)$ and assume that $w(A) = 1$ and $c(A) = \cos\phi$ with $m\phi \in [0, \pi/2]$. Then there is $\theta \in [0, 2\pi)$ such that

$$
W(A) \subseteq e^{i\theta}\{z \in \mathbb{C} : |z| \leq 1,\ \mathrm{Re}\, z \geq \cos\phi\}.
$$

By Theorem 2.5, for each $m = 2k$ or $2k - 1$ with $2k\phi \leq 2m\phi \leq \pi$, we have

$$
W(A^m) \subseteq e^{im\theta}\{z \in \mathbb{C} : |z| \leq 1,\ \mathrm{Re}\, z \geq \cos(m\phi)\}.
$$

Thus, $c(A^m) \geq \cos(m\phi)$. □

Remark 2.7 *Note that the conclusion on $W(A^m)$ in Theorem 2.5 can be deduced from Theorem 2.6 if $m\phi \in [0, \pi/2]$.*

To see this, suppose Theorem 2.6 holds. Assume

$$
W(A) \subseteq \{z \in \mathbb{C} : |z| \leq 1,\ \mathrm{Re}\, z \geq \cos\phi\},
$$

where $m\phi \in [0, \pi/2]$. Then $B = \mathrm{diag}\,(e^{i\phi}, e^{-i\phi}) \oplus A$ satisfies $w(B) = 1$ and $c(B) = \cos\phi$. By Theorem 2.6, $w(B^m) \leq 1$ and $c(B^m) \geq \cos(m\phi)$. Thus, there is $\theta \in [0, 2\pi)$ such that

$$
W(A^m) \subseteq W(B^m) \subseteq e^{i\theta}\{z \in \mathbb{C} : |z| \leq 1,\ \mathrm{Re}\, z \geq \cos(m\phi)\}.
$$

Since $e^{im\phi}, e^{-im\phi} \in W(B^m)$, we see that $\theta = 0$. Hence, we have

$$
W(A^m) \subseteq \{z \in \mathbb{C} : |z| \leq 1,\ \mathrm{Re}\, z \geq \cos(m\phi)\}.
$$

□

Next, we prove a result using the information $W(A^k)$ to deduce information of $W(A)$. In a certain sense, it can be viewed as the converse of Theorem 2.6.

Theorem 2.8 *Suppose k is a positive integer and $A \in \mathcal{B}(\mathcal{H})$ satisfies*

$$W(A) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge \cos(\pi/k)\}. \tag{2.7}$$

If $d \in \mathbb{R}$ is such that

$$W(A^k) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge d^k\},$$

then

$$W(A) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge d\}.$$

Proof. If $d < \cos(\pi/k)$, the result trivially holds. Assume that $d \ge \cos(\pi/k)$. By Lemma 2.3 and the assumption on $W(A^k)$, we have

$$0 \ge \mathrm{Re}\left(kd^{k-1}(d^k I - A^k)\right) = \mathrm{Re}\left(\sum_{j=1}^{k} g_j(A)^*(dI - \mu^j A)g_j(A)\right), \tag{2.8}$$

where $\mu, g_1, \ldots, g_k$ are defined as in Lemma 2.3. Since $\mathrm{Re}\,(\mu^j A) \le \cos(\pi/k) < d$, we see that $d \notin \overline{W(\mu^j A)}$ for $j \in \{1, \ldots, k-1\}$. Hence $g_k(A) = \prod_{j=1}^{k-1}(dI - \mu^j A)$ is invertible. Thus, the negativity of $g_k(A)^*\mathrm{Re}\,(dI - A)g_k(A)$ implies that of $\mathrm{Re}\,(dI - A)$. □

Note that if $c(A)/w(A) \le \cos(\pi/(2k))$, then 0 may lie in $W(A^k)$. So, we always assume that $c(A)/w(A) \ge \cos(\pi/(2k))$ when we study bounds of $c(A^k)$. Using the contra-positive of Theorem 2.6, we can get an upper bound for $c(A)$ in terms of that of $c(A^k)$, namely, if $w(A) = 1$ and $c(A^k) = \cos(k\phi)$ with $\phi \in [0, \pi/(2k))$, then $c(A) \le \cos\phi$. Using Theorem 2.8, we can prove the following.

Theorem 2.9 *Let k be a positive integer. Suppose $A \in \mathcal{B}(\mathcal{H})$ is nonzero and satisfies $c(A)/w(A) \ge \cos(\pi/(2k))$. Then $c(A)^k \ge c(A^k)$.*

Proof. We may replace A by ξA for some complex number ξ with $|\xi| = 1/w(A)$ and assume that $c(A) = \cos\theta \in \overline{W(A)}$ with $\theta \in [0, \pi/(2k)]$ and

$$W(A) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge \cos\theta\}.$$

Suppose $\gamma \in \overline{W(A^k)}$ satisfies $|\gamma| = c(A^k)$. By Theorem 2.6,

$$\gamma \in \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge \cos(k\theta)\}$$

so that $\gamma = d^k e^{ikt}$ for some $d > 0$ and $t \in [-\pi/(2k), \pi/(2k)]$. Now, replace A by $e^{-it}A$ so that

$$W(A) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge \cos(|\theta| + |t|)\} \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge \cos(\pi/k)\}$$

as $\theta, t \in [-\pi/(2k), \pi/(2k)]$. Since $d^k = |\gamma|$ is a point in $\overline{W(A^k)}$ nearest to the origin, we see that

$$W(A^k) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge d^k\}.$$

By Theorem 2.8, we see that

$$W(A) \subseteq \{z \in \mathbb{C} : |z| \le 1,\ \mathrm{Re}\, z \ge d\}.$$

Thus, $c(A) \ge d$, and hence $c(A)^k \ge d^k = c(A^k)$ as asserted. □

3 Negative powers of operators

First, we study inclusion region for $W(A^{-1})$. To ensure that A^{-1} exists, we often assume that $0 \notin \overline{W(A)}$. Note that this condition is stronger than the assumption that A is invertible. We have the following.

Proposition 3.1 *Let $A \in \mathcal{B}(\mathcal{H})$. If H has dimension 2 and A is invertible, then $W(A^{-1}) = W(A)/\det(A)$. If H has dimension at least 3 and if $0 \notin \overline{W(A)}$, then*

$$W(A^{-1}) \subseteq \bigcup \{W(B)/\det(B) : B = X^*AX, X^*X = I_2\}.$$

Proof. Assume that $A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in M_2$. Then $A^{-1} = \frac{1}{ac}\begin{pmatrix} c & -b \\ 0 & a \end{pmatrix}$, which is unitarily similar to $A/\det(A)$. So, $W(A^{-1}) = W(A)/\det(A)$.

Suppose $A \in \mathcal{B}(\mathcal{H})$ is invertible, and $v \in \mathcal{H}$ is a unit vector. Let B be the compression of A on a two dimensional subspace of $\mathcal{H}$ containing v and $A^{-1}v$. Since $0 \notin W(B) \subseteq \overline{W(A)}$, B is invertible. Then $(A^{-1}v, v) = (B^{-1}v, v) \in W(B^{-1})$. □

If A is a nonzero multiple of a positive definite matrix, then $W(A^{-1}) = \operatorname{conv}\{z^{-1} : z \in W(A)\}$. But the equality may not hold even for a normal operator A of the form $\alpha I + \beta H$ for some self-adjoint H.

Example 3.2 *Let $A = \operatorname{diag}(1+i, 1-i)$. Then $W(A)$ is the line segment joining the points $1+i$ and $1-i$, but $W(A^{-1})$ is the line segment joining the points $(1-i)/2$ and $(1+i)/2$, and is a proper subset of* $\operatorname{conv}\{z^{-1} : z \in W(A)\}$.

For non-normal $A \in \mathcal{B}(\mathcal{H})$, the inclusion relation $W(A^{-1}) \subseteq \{z^{-1} : z \in W(A)\}$ may not hold as shown in the following example.

Example 3.3 *Let $A = \begin{pmatrix} 1 & 2c \\ 0 & 1 \end{pmatrix}$ with $0 < c < 1$. Then $A^{-1} = \begin{pmatrix} 1 & -2c \\ 0 & 1 \end{pmatrix}$, and $W(A^{-1}) = W(A)$ is a circular disk with $[1-c, 1+c]$ as diameter. However, $\mathcal{K} = \{z^{-1} : z \in W(A)\}$ is the disk with $[1/(1+c), 1/(1-c)]$ as diameter. So $W(A^{-1}) \not\subseteq \mathcal{K}$ as $1-c < 1/(1+c)$.*

Using Theorem 6 in [14], we have the following result.

Proposition 3.4 *Suppose $A \in \mathcal{B}(\mathcal{H})$ is invertible. If*

$$W(A) \subseteq \{z \in \mathbb{C} : \ |z-1| \leq 1\},$$

then

$$W(A^{-1}) \subseteq \{z : \operatorname{Re} z \geq -1/2\}.$$

By Proposition 3.4, one can obtain a left supporting line for $W(A^{-1})$ if $W(A)$ is included in a certain circle. In the following, we show that one can obtain a circular inclusion region for $W(A^{-1})$ in terms of a supporting line of $W(A)$ that separates $W(A)$ and the origin.

Proposition 3.5 *Let $A \in \mathcal{B}(\mathcal{H})$ be such that $\operatorname{Re} A \geq sI$ for some $s > 0$. Then*

$$\|A^{-1} - I/(2s)\| \leq 1/(2s),$$

and

$$W(A^{-1}) \subseteq \{z \in \mathbb{C} : |z - 1/(2s)| \leq 1/(2s)\}.$$

Proof. If $A + A^* \geq 2sI > 0$, then $2sA^{-1}(A + A^* - 2sI)A^{*-1} \geq 0$ and hence

$$I \geq I - 2sA^{-1}(A + A^* - 2sI)A^{*-1} = (I - 2sA^{-1})(I - 2sA^{-1})^*,$$

which is equivalent to $\|I - 2sA^{-1}\| \leq 1$. Hence $w(I - 2sA^{-1}) \leq 1$ and the result follows. □

Note that none of Proposition 3.4 or Proposition 3.5 gives us any information about $c(A^{-1})$. In connection with this, we have the following result.

Theorem 3.6 *Suppose $A \in \mathcal{B}(\mathcal{H})$ is invertible. Then*

$$c(A^{-1}) \geq c(A)/w(A)^2.$$

Proof. The result is trivial if $c(A) = 0$. So, we assume that $c(A) > 0$. If $A \in M_2$, then $W(A^{-1}) = W(A)/\det(A)$ by Proposition 3.1. Since $|\det(A)| \leq w(A)^2$, we have $c(A^{-1}) \geq c(A)/(w(A))^2$.

This result can be extended to general operators $A \in \mathcal{B}(\mathcal{H})$ with $0 \notin \overline{W(A)}$. In fact, for such an operator A, let $v \in \mathcal{H}$ be any unit vector, and let B be the compression of A on a two dimensional subspace of $\mathcal{H}$ containing the vectors v and $A^{-1}v$. Since $0 \notin W(B)$, B is invertible. By the result on 2×2 matrices, we have

$$\begin{aligned} |(A^{-1}v, v)| &= |(B^{-1}v, v)| \\ &\geq \inf\{|z| : z \in W(B^{-1})\} \\ &\geq \inf\{|z| : z \in W(B)/w(B)^2\} \\ &\geq \inf\{|z| : z \in W(A)/w(A)^2\} \end{aligned}$$

Thus $c(A^{-1}) \geq c(A)/w(A)^2$. □

Theorem 3.7 *Let $S = \{z \in \mathbb{C} : |z| \leq r, \operatorname{Re} z \geq s\}$ with $r > s > 0$. Suppose $A \in \mathcal{B}(\mathcal{H})$ is such that $W(A) \subseteq S$. Then*

$$W(A^{-1}) \subseteq \operatorname{conv}\{z^{-1} : z \in S\} = \{z \in \mathbb{C} : |z - 1/(2s)| \leq 1/(2s),\ \operatorname{Re} z \geq s/r^2\}.$$

Proof. We may replace A by $A/w(A)$ and assume that $r = 1$ and $s = \cos\theta$ with $\theta \in (0, \pi/2)$. Let $B = A \oplus \text{diag}\,(e^{i\theta}, e^{-i\theta})$. Then $W(B) \subseteq S$ and $W(A^{-1}) \subseteq W(B^{-1})$. By Theorem 3.6, $c(B^{-1}) \geq c(B)/w(B)^2 = \cos\theta$. Since $e^{i\theta}, e^{-i\theta} \in W(B^{-1})$, it follows that $s = \cos\theta \in W(B^{-1})$ is the point in $\overline{W(B^{-1})}$ nearest to the origin, and the support line of the convex set $\overline{W(B^{-1})}$ at $\cos\theta$ must pass through $e^{i\theta}$ and $e^{-i\theta}$. Thus, $W(A^{-1}) \subseteq W(B^{-1}) \subseteq \{z : \text{Re}\, z \geq s\}$. By Proposition 3.5, we see that $W(A^{-1}) \subseteq \{z : |z - 1/(2s)| \leq 1/(2s)\}$. Hence, $W(A^{-1}) \subseteq \{z : |z-1/(2s)| \leq 1/(2s),\ \text{Re}\, z \geq s\}$, which equals the set $\text{conv}\,\{z^{-1} : z \in S\}$. □

By Theorem 2.5 and Theorem 3.7, we have the following.

Corollary 3.8 *Let $\theta \in [0, \pi/(2p)]$ for a positive integer p. Suppose $A \in \mathcal{B}(\mathcal{H})$ is such that $W(A) \subseteq \{z : |z| \leq 1, \text{Re}\, z \geq \cos\theta\}$. Then for $\gamma = 2\cos(p\theta)$,*

$$W(A^{-p}) \subseteq \{z \in \mathbb{C} : |z - 1/(2\gamma)| \leq 1/(2\gamma),\ \text{Re}\, z \geq \gamma\}.$$

By the power inequality, we have $w(A^k) \leq w(A)^k$ for any $A \in \mathcal{B}(\mathcal{H})$ and any positive integer k. In the following, we obtain a similar inequality for the negative powers of A for an invertible $A \in \mathcal{B}(\mathcal{H})$, and characterize those A which attained the equality.

Theorem 3.9 *Suppose $A \in \mathcal{B}(\mathcal{H})$ is invertible and p is a positive integer. Then*

$$w(A^{-p}) \geq w(A)^{-p}.$$

The equality holds if and only if A is a nonzero multiple of a unitary operator.

Proof. Using the fact that $\rho(B) \leq w(B)$ for any $B \in \mathcal{B}(\mathcal{H})$, one easily sees that

$$w(A^{-1}) \geq \rho(A^{-1}) = 1/\inf\{|z| : z \in \sigma(A)\} \geq \rho(A)^{-1} \geq w(A)^{-1}.$$

Replacing A by A^p, we have $w(A^{-p}) \geq w(A^p)^{-1}$. Using the fact that $w(A^p) \leq w(A)^p$, we have $w(A^{-p}) \geq w(A)^{-p}$.

If A is a multiple of a unitary operator, then we have $w(A^{-p}) = w(A)^{-p}$. For the converse, suppose the equality holds. We may replace A by γA for a suitable nonzero γ and assume that $w(A^{-p}) = w(A)^{-p} = 1$. Thus,

$$1 = w(A^{-p}) \geq w(A^p)^{-1} \geq w(A)^{-p} = 1.$$

So, $1 = w(A^{-p}) = w(A^p)$. By [17, Corollary 1] (see also [16]), A^p is unitary. Since $w(A) = 1$, by the result of Ando [1], there exist a self-adjoint $Z \in \mathcal{B}(\mathcal{H})$ and a contraction $C \in \mathcal{B}(\mathcal{H})$ such that $-I \leq Z \leq I$ and $A = (I+Z)^{1/2}C(I-Z)^{1/2}$. Now,

$$V = A^p = (I+Z)^{1/2}\widetilde{C}(I-Z)^{1/2} \tag{3.1}$$

is unitary, where $\widetilde{C} = C(I-Z^2)^{1/2}C \cdots C(I-Z^2)^{1/2}C$ is a contraction. So, $(I+Z)$ and $(I-Z)$ are invertible, and

$$\begin{pmatrix} I+Z & V \\ V^* & I-Z \end{pmatrix}$$

is positive semi-definite. Applying Schur complement, we see that

$$I+Z \geq V(I-Z)^{-1}V^*. \tag{3.2}$$

Suppose $\operatorname{conv}\sigma(Z) = [a,b]$. Since $(I+Z)$ and $(I-Z)$ are invertible, we have $[a,b] \subseteq (-1,1)$ and $\operatorname{conv}(\sigma((I-Z)^{-1})) = [(1-a)^{-1},(1-b)^{-1}]$. Comparing the spectrum on both sides of (3.2), we see that $1+a \geq (1-a)^{-1}$ and $1+b \geq (1-b)^{-1}$. Hence, $a=b=0$, i.e., $Z=0$. By (3.1) and the fact that $Z=0$, we have $A^p = \widetilde{C} = C^p$ is unitary. Because C is a contraction, we conclude that C is unitary. □

Corollary 1 in [17] asserts that $A \in \mathcal{B}(\mathcal{H})$ is unitary if $w(A) = w(A^{-1}) = 1$. The above theorem can be viewed as a generalization of this result. Clearly, if $w(A) = w(A^{-1}) = 1$, then $w(A) \leq 1$ and $\sigma(A)$ lies on the unit circle. In the finite-dimensional case, the converse is also valid. In the infinite-dimensional case, there exists non-unitary $A \in \mathcal{B}(\mathcal{H})$ such that $w(A) \leq 1$ with $\sigma(A)$ lying on the unit circle. For instance, if V is the Volterra operator, then $A = (I+V)^{-1}$ is such an example (e.g., see [13, Problem 190]).

Remark 3.10 *From the proof of Theorem 3.9, we see that*

$$w(A^{-1})^p \geq w(A^{-p}) \geq w(A^p)^{-1} \geq w(A)^{-p}.$$

We have shown that $w(A^{-p}) = w(A)^{-p}$ if and only if A is a multiple of a unitary operator. Replacing A by A^{-1}, we see that $w(A^{-1})^p = w(A^p)^{-1}$ if and only if A is a multiple of a unitary operator.

Very recently, Ando and Li [3] have extended Theorem 3.9 by replacing the numerical radius with any operator radius.

Acknowledgment

The authors would like to thank the referee for some helpful comments.

References

[1] T. Ando, Structure of operators with numerical radius one, Acta Sci. Math. (Szeged) 34 (1973), 11-15.

[2] T. Ando and C.K. Li (Editors), Special Issue: The Numerical Range and Numerical Radius, Linear and Multilinear Algebra 37, nos. 1–3, 1994.

[3] T. Ando and C.K. Li, Operator radii and unitary operators, preprint.

[4] O. Axelsson, H. Lu and B. Polman, On the numerical radius of matrices and its application to iterative solution methods, Linear and Multilinear Algebra 37 (1994), 225-238.

[5] C.A. Berger, A strange dilation theorem, Amer. Math. Soc. Not. 12 (1965), 590.

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

[7] F.F. Bonsall and Duncan, Studies in Functional Analysis – Numerical ranges, Studies in Mathematics Vol. 21, The Mathematical Association of America, 1980.

[8] M.D. Choi, Adjunction and inversion of invertible Hilbert space operators, Indiana U. Math. J. 23 (1973), 413-419.

[9] M.D. Choi and C.K. Li, Numerical ranges and dilations, Linear and Multilinear Algebra 47 (2000), 35-48.

[10] M.D. Choi and C.K. Li, Constrained Unitary Dilations and Numerical Ranges, J. Operator Theory, 46 (2001), 435-447.

[11] M.J. Crabb, J. Duncan, and C.M. McGregor, Mapping theorems and the numerical radius, Proc. London Math. Soc. 25 (1972), 486-502.

[12] K.E. Gustafson and D.K.M. Rao, Numerical range: The field of values of linear operators and matrices, Springer, New York, 1997.

[13] P. Halmos, A Hilbert Space Problem Book, Second Edition, Springer-Verlag, New York, 1974.

[14] T. Kato, Some mapping theorems for the numerical range, Proc. Japan. Acad. 41 (1965), 652-655.

[15] T. Kato, Perturbation theory for linear operators, 2nd ed., Springer-Verlag, New York, 1980.

[16] T. Sano and A. Uchiyama, Numerical radius and unitarity, Acta Sci. Math. (Szeged), to appear.

[17] J.G. Stampfli, Minimal range theorems for operators with thin spectra, Pac. J. Math. 23 (1967), 601-612.

[18] G.W. Stewart and J.-g. Sun, Matrix Perturbation Theory, Academic Press, New York, 1990.