Numerical Ranges and Dilations

Dedicated to Professor Yik-Hoi Au-Yeung on the occasion of his retirement

Man-Duen Choi
Department of Mathematics, University of Toronto, Toronto, Ontario, Canada M5S 3G3. E-mail: choi@math.toronto.edu

and

Chi-Kwong Li
Department of Mathematics, The College of William and Mary, P.O. Box 8795, Williamsburg, Virginia 23187-8795, USA. E-mail: ckli@math.wm.edu

Abstract

Denote by $W(A)$ the numerical range *of a bounded linear operator A. For two operators A and B (which may act on different Hilbert spaces), we study the relation between the inclusion relation $W(A) \subseteq W(B)$ and the condition that A can be dilated to an operator of the form $B \otimes I$. We also investigate the possibilities of dilating an operator A to operators with simple structure under the assumption that $W(A)$ is included in a special region.*

Keywords: Numerical range, dilation, Hilbert space, operator.
AMS Classifications: 15A60, 47A12, 47A30.

1. Introduction

Given a Hilbert space $\mathcal{H}$, let $\mathcal{B}(\mathcal{H})$ the algebra of all bounded linear operators on $\mathcal{H}$. If $\mathcal{H}$ has dimension $n < \infty$, $\mathcal{B}(\mathcal{H})$ will be identified with the algebra of $n \times n$ complex matrices, denoted by M_n. The *numerical range* of an operator $A \in \mathcal{B}(\mathcal{H})$ is defined by

$$W(A) = \{(Ax, x) : x \in \mathcal{H},\ (x, x) = 1\}.$$

A result of Ando [An], which also follows from a more general theorem of Arveson [Ar, Theorem 3.1.1], asserts the following.

Theorem 1.1 *Let $A \in \mathcal{B}(\mathcal{H})$. The following conditions are equivalent.*

(a) *$W(A)$ is included in the closed unit disk $\mathcal{D} = \{z \in \mathbf{C} : |z| \le 1\}$.*

(b) $A = V^* \begin{pmatrix} 0 & 2I_{\mathcal{H}} \\ 0 & 0 \end{pmatrix} V$ *for some V satisfying $V^*V = I_{\mathcal{H}}$.*

Notice that $\begin{pmatrix} 0 & 2I \\ 0 & 0 \end{pmatrix} = B \otimes I$ with $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$, and $W(B) = \mathcal{D}$. Thus condition Theorem 1.1 (a) is equivalent to:

(c) $W(A) \subseteq W(B)$ *with* $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$.

Given two operators X and Y acting on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, we say that Y is a *dilation* of X if there exists an operator $V : \mathcal{H}_1 \to \mathcal{H}_2$ satisfying $V^*V = I$ and $X = V^*YV$. It is easy to see that $W(X) \subseteq W(Y)$ if Y is a dilation of X, and $W(Y \otimes I) = W(Y)$. Hence, if A can be dilated to an operator of the form $B \otimes I$, it follows readily that $W(A) \subseteq W(B)$. It is remarkable that the converse also holds if $B = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$.

Another interesting result along the same line is the following theorem due to Mirman [M] (also see [Na]).

Theorem 1.2 *Let $A \in \mathcal{B}(\mathcal{H})$ and let $\gamma_1, \gamma_2, \gamma_3 \in \mathbf{C}$. The following conditions are equivalent.*

(a) *$W(A)$ is included in the triangle with vertices $\gamma_1, \gamma_2, \gamma_3$.*

(b) *$A = V^*(B \otimes I)V$, where $B = \operatorname{diag}(\gamma_1, \gamma_2, \gamma_3)$, I is the identity operator on a certain Hilbert space $\mathcal{K}$, and $V : \mathcal{H} \to \mathbf{C}^3 \otimes \mathcal{K}$ satisfies $V^*V = I_{\mathcal{H}}$.*

(c) *$W(A)$ is included in $W(B)$ with $B = \operatorname{diag}(\gamma_1, \gamma_2, \gamma_3)$.*

The results of Ando and Mirman show that there are interesting relations between the conditions that $W(A) \subseteq K$ for some special region $K \subseteq \mathbf{C}$, or simply $W(A) \subseteq W(B)$ for some operator B, and the dilation property that

$$A = V^*(B \otimes I)V \quad \text{for some } \; V \; \text{ satisfying } V^*V = I_{\mathcal{H}}. \tag{1}$$

If (1) holds, we say that A can be dilated to an operator of the form $B \otimes I$. The purpose of this paper is to further study the numerical range inclusion relation $W(A) \subseteq K$ or $W(A) \subseteq W(B)$ and the dilation property (1).

Our paper is organized as follows. We first obtain extensions of the results of Ando and Mirman in section 2, and consider the set Γ of operators B such that condition (1) holds whenever $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq W(B)$. In section 3, we shift our focus to the set Λ of operators A such that if B is an operator satisfying $W(A) \subseteq W(B)$ then condition (1) holds. In section 4, we prove that if $A \in M_n$ has numerical range lying in a trapezoidal region K, then A can be dilated to an operator B, which is a direct sum of no more than

n matrices in M_2 such that $W(B) \subseteq K$. We conclude the paper with some remarks in section 5. Open questions are mentioned throughout the paper.

Suppose $f(x+iy) = (ax+by+r) + i(cx+dy+s)$ with $a,b,c,d,r,s \in \mathbf{R}$ is an affine transform on $\mathbf{C}$. If $A = A_1 + iA_2 \in \mathcal{B}(\mathcal{H})$ such that A_1 and A_2 are self-adjoint operators, define

$$f(A) = (aA_1 + bA_2 + rI) + i(cA_1 + dA_2 + sI).$$

It is easy to check that f is invertible if and only if $ad - bc \neq 0$. We shall use the following fact in our discussion.

Proposition 1.3 *Let A and B be bounded operators that may act on different Hilbert spaces, and let f be an invertible affine transform on $\mathbf{C}$. Then A can be dilated to an operator of the form $B \otimes I$ if and only if $f(A)$ can be dilated to an operator of the form $f(B) \otimes I$.*

2. Extensions of the results of Ando and Mirman

We begin with an extension of the result of Ando.

Theorem 2.1 *Let $\mathcal{E}$ be the closed elliptical disk in $\mathbf{C}$ with foci μ_1 and μ_2, and let d be the length of the minor axis. Then $\mathcal{E} = W(B)$ with $B = \begin{pmatrix} \mu_1 & d \\ 0 & \mu_2 \end{pmatrix}$. Moreover, let $A \in \mathcal{B}(\mathcal{H})$. Then A satisfies $W(A) \subseteq W(B) = \mathcal{E}$ if and only if A can be dilated to an operator of the form $B \otimes I$.*

Proof. The fact that $\mathcal{E} = W(B)$ is the well known elliptical range theorem for the numerical range of a 2×2 matrix (see e.g., [HoJ, 1.3.3]).

If $A \in \mathcal{B}(\mathcal{H})$ can be dilated to an operator of the form $B \otimes I$, then $W(A) \subseteq W(B \otimes I) = W(B)$. Conversely, suppose $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq W(B)$. We prove that A can be dilated to an operator of the form $B \otimes I$ by considering three cases.

If $W(B)$ is a singleton $\{\mu\}$, then so is $W(A)$. It follows that $B = \mu I_2$ and $A = \mu I_{\mathcal{H}}$, and hence the conclusion holds.

If $W(B)$ is a non-degenerate line segment, then $d = 0$ and $\mu_1 \neq \mu_2$. Replace A and B by $\tilde{A} = (A - \mu_2 I_{\mathcal{H}})/(\mu_1 - \mu_2)$ and $\tilde{B} = (B - \mu_2 I_2)/(\mu_1 - \mu_2) = \mathrm{diag}(1, 0)$. Then $W(A) \subseteq W(B)$ if and only if $W(\tilde{A}) \subseteq W(\tilde{B})$; and condition (1) is equivalent to the fact that $\tilde{A}$ can be dilated to an operator of the form $\tilde{B} \otimes I$. It is clear that $\tilde{B} \otimes I = I \oplus 0$ and $\tilde{A} = V^*(I \oplus 0)V$, where $V = \begin{pmatrix} \sqrt{\tilde{A}} & \sqrt{I - \tilde{A}} \end{pmatrix}^*$ satisfying $V^*V = I_{\mathcal{H}}$. Thus the conclusion holds.

If $W(B)$ is neither a point nor a line segment, then the boundary of $W(B)$ is a non-degenerate ellipse. Since $W(aX + bI) = aW(X) + b$ for any $a, b \in \mathbf{C}$, there exist

$\alpha, \beta \in \mathbf{C}$ such that the boundary of the numerical range of the matrix $\tilde{B} = \alpha(B - \beta I_2)$ is an ellipse centered at the origin with -1 and 1 as the endpoints of its major axis. Let $\tilde{A} = \alpha(A - \beta I_{\mathcal{H}}) = A_1 + iA_2$ and $\tilde{B} = B_1 + iB_2$, where A_1, A_2, B_1 and B_2 are self-adjoint. Suppose B_2 has eigenvalues $\pm\gamma$. Then 2γ is the length of the minor axis of $W(\tilde{B})$. Consider $\hat{A} = A_1 + iA_2/\gamma$ and $\hat{B} = B_1 + iB_2/\gamma$. One easily checks that $W(\hat{B})$ is the closed unit disk, and $W(\hat{A}) \subseteq W(\hat{B})$. Since $\hat{B}$ is a 2×2 matrix with $W(\hat{B}) = \mathcal{D}$, by the elliptical range theorem there is a unitary matrix U such that $U^*\hat{B}U = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$. Now since $W(\hat{A}) \subset \mathcal{D}$, by Ando's theorem [An] we have

$$\hat{A} = V^* \begin{pmatrix} 0 & 2I_{\mathcal{H}} \\ 0 & 0 \end{pmatrix} V$$

for some V satisfying $V^*V = I_{\mathcal{H}}$. Let $\tilde{V} = (U \otimes I_{\mathcal{H}})V$. Then $\tilde{V}^*\tilde{V} = I_{\mathcal{H}}$ and $\hat{A} = \tilde{V}^*(\hat{B} \otimes I_{\mathcal{H}})\tilde{V}$. By Proposition 1.3, A has a dilation of the form $B \otimes I$. $\square$

We remark that our proof of Theorem 2.1 depends on the result of Ando. Moreover, Theorem 2.1 does not hold if $\mathcal{E} = W(B)$ for a general 3×3 matrix B even when $W(B)$ is a circular disk as shown in the following example.

Example 2.2 Let $B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ and $A = \begin{pmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}$. Then $W(A) \subseteq W(B)$, but A cannot be dilated to an operator of the form $B \otimes I$ as $\|A\| > \|B\|$.

Next, we turn to a generalization of the result of Mirman. Our proof covers the theorem of Mirman and is different from those in [M] and [Na].

Proposition 2.3 *Let B be an operator such that $W(B)$ is the convex hull of k complex numbers with $k \leq 3$. Suppose $A \in \mathcal{B}(\mathcal{H})$. Then A satisfies $W(A) \subseteq W(B)$ if and only if A has a dilation of the form $B \otimes I$.*

Proof. The ($\Leftarrow$) part is clear. To prove the converse, we consider three cases, namely, $W(B)$ is a singleton, a line segment, and a triangular region.

The proof of the first case is similar to that in Theorem 2.1. For the second case, we use Proposition 1.3 and apply a suitable affine transform on $\mathbf{C}$ so that $W(A) \subseteq W(B) = [0, 1]$. By the proof of Theorem 2.1, A has a dilation of the form $B_1 = \operatorname{diag}(0, 1) \otimes I$. Since $W(B) = [0, 1]$, it follows (see e.g., [Do]) that 0 and 1 are eigenvalues of B. Hence B_1 has a dilation of the form $B \otimes I$, and so does A.

For the third case, we again use Proposition 1.3 and apply a suitable affine transform on $\mathbf{C}$ so that $W(B)$ is the convex hull of $\{0, 1, i\}$. Thus $W(B) = W(B_1)$ with $B_1 =$

$\mathrm{diag}\,(0,1,i)$ and $0,1,i$ are eigenvalues of B (see [Do]). The result will follow if we can show that A has a dilation of the form $B_1 \otimes I$. To this end, let $A = H + iG \in \mathcal{B}(\mathcal{H})$, where H and G are self-adjoint. Then $W(A) \subseteq W(B_1)$ if and only if $H \geq 0$, $G \geq 0$ and $H + G \leq I$. Set $V = (\sqrt{I - H - G} \quad \sqrt{H} \quad \sqrt{G}\,)^*$. Then $V^*V = I$ and $A = V^*(B_1 \otimes I)V$. $\square$

The next example shows that the conclusion of Proposition 2.3 may not hold if $W(B)$ is a convex polygonal region with more than 3 vertices.

Example 2.4 Let $A = \begin{pmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix}$, and let $B = \mathrm{diag}(1,-1,i,-i)$. Then $W(A) \subseteq W(B)$ and $W(B)$ is the convex hull of the numbers $1,-1,i,-i$. However, A cannot be dilated to an operator of the form $B \otimes I$ as $\|A\| > \|B\|$.

In section 4, we shall obtain more dilation results concerning those $A \in \mathcal{B}(\mathcal{H})$ with $W(A)$ lying in a square. In connection to the above example, we have the next theorem characterizing those $A \in \mathcal{B}(\mathcal{H})$ which has a dilation of the form $\mathrm{diag}(1,-1,i,-i) \otimes I$. It is interesting to note that the condition also involves the numerical range of a certain operator generated by A.

Theorem 2.5 *Let $A \in \mathcal{B}(\mathcal{H})$. The following conditions are equivalent.*

(a) *A has a dilation of the form $B = \mathrm{diag}(1,-1,i,-i) \otimes I$.*

(b) *There exist positive semi-definite operators A_1, A_2, A_3, A_4 such that $A_1 + A_2 + A_3 + A_4 = I$ and $A = (A_1 - A_2) + i(A_3 - A_4)$.*

(c) *The numerical range of the operator $\tilde{A} = \begin{pmatrix} 0 & A + A^* \\ i(A^* - A) & 0 \end{pmatrix}$ is a subset of the closed unit disk $\mathcal{D}$.*

Proof. The equivalence of (a) and (b) are clear. We prove the equivalence of (b) and (c) in the following.

Suppose (b) holds. Let

$$V = \begin{pmatrix} \sqrt{A_1} & \sqrt{A_2} & 0 & 0 & 0 & 0 & \sqrt{A_3} & -\sqrt{A_4} \\ 0 & 0 & \sqrt{A_3} & \sqrt{A_4} & \sqrt{A_1} & -\sqrt{A_2} & 0 & 0 \end{pmatrix}^*.$$

Then $V^*V = I$ and $V^* \begin{pmatrix} 0 & 2I \\ 0 & 0 \end{pmatrix} V = \tilde{A}$. By the result of Ando [An], $W(\tilde{A}) \subseteq \mathcal{D}$.

Conversely, suppose (c) holds. By the result of Ando [An], there exists

$$V = \begin{pmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{pmatrix}$$

such that $V^*V = I$ and

$$V^* \begin{pmatrix} 0 & 2I \\ 0 & 0 \end{pmatrix} V = \tilde{A}.$$

Hence,

$$X_1^*X_1 + Y_1^*Y_1 + X_2^*X_2 + Y_2^*Y_2 = 2I,$$

$$2X_1^*Y_2 = A + A^* = 2Y_2^*X_1, \quad \text{and} \quad 2X_2^*Y_1 = i(A^* - A) = 2Y_1^*X_2.$$

Set

$$A_1 = (X_1 + Y_2)^*(X_1 + Y_2)/4, \qquad A_2 = (X_1 - Y_2)^*(X_1 - Y_2)/4,$$

$$A_3 = (X_2 + Y_1)^*(X_2 + Y_1)/4 \quad \text{and} \quad A_4 = (X_2 - Y_1)^*(X_2 - Y_1)/4.$$

One easily checks that condition (b) holds. $\square$

Let Γ be the set of operators B such that: if $A \in \mathcal{B}(\mathcal{H})$ satisfies $W(A) \subseteq W(B)$ then A admits a dilation of the form $B \otimes I$. We have shown that all 2×2 matrices and all matrices B such that $W(B)$ is a polygonal disk with no more than 3 vertices belong to Γ. It would be nice to solve

Problem 2.6 *Characterize those operators lying in* Γ.

Notice that in both Example 2.2 and Example 2.4, B is unitarily similar to its transpose B^t. Thus the examples actually show that $W(A) \subseteq W(B)$ does not even imply that A can be dilated to an operator of the form $(B \oplus B^t) \otimes I$, i.e.,

$$A = V^*((B \oplus B^t) \otimes I)V \quad \text{for some } V \text{ satisfying } V^*V = I_{\mathcal{H}}. \tag{2}$$

One may also ask the following problem.

Problem 2.7 *Characterize those operators* B *such that condition* (2) *holds whenever* $A \in \mathcal{B}(\mathcal{H})$ *satisfies* $W(A) \subseteq W(B)$.

3. Other dilation results

In this section, we shift our focus to the set Λ of linear operators A such that A admits a dilation of the form $B \otimes I$ whenever B is an operator satisfying $W(A) \subseteq W(B)$.

Recall that an operator A is a *convexoid* (see e.g., [HoJ] and [H2]) if the closure of $W(A)$ equals the convex hull of the spectrum of A. We shall show if $A \in \Lambda$, then A must be a convexoid.

Given an operator X, let $\sigma(X)$ denote its spectrum, $r(X)$ denote its *spectral radius*, and $w(X) = \sup\{|z| : z \in W(X)\}$ denote its *numerical radius*. We have the following result.

Proposition 3.1 *Suppose $A \in \Lambda$. Then the following two equivalent conditions hold.*

(a) $\|A - \mu I\| = w(A - \mu I)$ *for all* $\mu \in \mathbf{C}$.

(b) $\|A - \mu I\| = r(A - \mu I)$ *for all* $\mu \in \mathbf{C}$.

Furthermore, these conditions imply the following equivalent conditions.

(c) $w(A - \mu I) = r(A - \mu I)$ *for all* $\mu \in \mathbf{C}$.

(d) *A is a convexoid.*

Proof. The equivalence of (a) and (b) is due to Wintner [W]. We establish (a) under the hypothesis that $A \in \Lambda$ in the following. Suppose there exists $\mu \in \mathbf{C}$ such that $\|A - \mu I\| > w(A - \mu I)$. Let $\tilde{A} = (A - \mu I)/w(A - \mu I)$. Then $\|\tilde{A}\| > w(\tilde{A}) = 1$. Let m be a positive integer such that $\|\tilde{A}\| > \gamma$ with $\gamma = 1/\cos(\pi/(m+1))$, and let $\tilde{B} = (b_{ij})$ be an $m \times m$ matrix such that

$$b_{ij} = \begin{cases} \gamma & \text{if } j = i + 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then $W(\tilde{B})$ is the closed unit disk $\mathcal{D}$ so that $W(\tilde{A}) \subseteq W(\tilde{B})$. Since $\|\tilde{A}\| > \gamma = \|\tilde{B}\|$, it is impossible for $\tilde{A}$ to have a dilation of the form $\tilde{B} \otimes I$. Now, let $B = w(A - \mu I)\tilde{B} + \mu I$. Then $W(A) \subseteq W(B)$, but A does not have a dilation of the form $B \otimes I$.

It is known (see e.g., [FN]) that an operator A is a convexoid if and only if condition (c) holds. Since $r(X) \leq w(X) \leq \|X\|$ for any operator X, condition (c) follows from condition (b). $\square$

The following theorem gives a different description for those $A \in \Lambda$. Together with Proposition 3.1, it helps to identify some more operators in Λ.

Proposition 3.2 *Let $A \in \mathcal{B}(\mathcal{H})$. The following conditions are equivalent.*

(a) $A \in \Lambda$.

(b) *There is a family of positive semi-definite operators A_μ, $\mu \in W(A)$, such that* $\sum_{\mu \in W(A)} A_\mu = I$ *and* $A = \sum_{\mu \in W(A)} \mu A_\mu$.

(c) *A has a dilation of the form $D \otimes I$, where D is a diagonal operator with $W(A) = \sigma(D)$.*

In case $W(A)$ is closed, then conditions (a) – (c) are equivalent to:

(d) *A has a dilation of the form $\tilde{D} \otimes I$, where $\tilde{D}$ is a diagonal operator with $\sigma(\tilde{D})$ equal to the set of extreme points of $W(A)$.*

Proof. The equivalence of (b) and (c) is clear. Suppose (a) holds. If D is a diagonal operator with $W(A) = \sigma(D)$, then $W(A) \subseteq W(D)$. Hence A admits a dilation of the form $D \otimes I$, i.e., condition (c) holds.

Conversely, if (c) holds, and if $W(D) = W(A) \subseteq W(B)$, then D admits a dilation of the form $B \otimes I$ and so does A. Hence condition (a) holds.

If $W(A)$ is closed, then it is a compact convex set, and is equal to the convex hull of its extreme points. Consequently, if D and $\tilde{D}$ are the diagonal operators defined in (c) and (d), then D has a dilation of the form $\tilde{D} \otimes I$. Hence, condition (c) implies condition (d). The reverse implication is clear. Thus the last assertion of the theorem follows. $\square$

Let $A \in \mathcal{B}(\mathcal{H})$. Then $W(A)$ is a convex polygon with interior having vertices $\mu_1, \ldots, \mu_k$ if and only if A is unitarily similar to $D \oplus A_1$ with $D = \operatorname{diag}(\mu_1, \ldots, \mu_k)$ (see [Do] and [HoJ]). If this condition holds, then clearly A is a convexoid. The converse is valid in the finite dimensional case. By Proposition 3.2, we have the following corollary, which, in particular, characterizes those $A \in M_n$ lying in Λ.

Corollary 3.3 *Let $A \in \mathcal{B}(\mathcal{H})$ be such that $W(A)$ is a convex polygon with interior having vertices $\mu_1, \ldots, \mu_k$. Then $A \in \Lambda$ if and only if A can be dilated to an operator of the form $D \otimes I$ with $D = \operatorname{diag}(\mu_1, \ldots, \mu_k)$.*

The situation in the infinite dimensional case is more complicated if A is a convexoid such that $W(A)$ is not closed as shown in the following example.

Example 3.4 Let A be the bilateral shift operator such that $Ae_j = e_{j+1}$ for all integers j. Suppose B is the diagonal operator (acting on an inseparable Hilbert space) with diagonal elements e^{it} with $t \in [0, 2\pi)$. Then A is unitary and $W(A) \subseteq W(B)$. However, A cannot be dilated to a unitary operator of the form $B \otimes I$ because of the following reason. Since both A and B are unitary, if there exists unitary operator U such that $U^*(B \otimes I)U = \begin{pmatrix} A & R \\ S & T \end{pmatrix}$ then R and S must be zeros. But then A is a direct summand of an operator with eigenvalues, which is a contradiction.

In general, it would be interesting to solve the following problem.

Problem 3.5 *Characterize those infinite dimensional A that lie in Λ.*

Note that the question is open even if we assume that A is normal. In view of Proposition 3.2 and Example 3.4, the solution of Problem 3.5 should involve conditions on the boundary of $W(A)$.

Also, in view of Proposition 3.1, one may consider imposing norm inequalities on the operators A and B in addition to inclusion relation that $W(A) \subseteq W(B)$ to ensure that A has a dilation of the form $B \otimes I$. For instance, if A has a dilation of the form $B \otimes I$, then $W(A) \subseteq W(B)$, and $\|A - \mu I\| \le \|B - \mu I\|$ for all $\mu \in \mathbf{C}$. One may wonder whether the converse holds. Unfortunately, the answer is negative as shown by the following example.

Example 3.6 Let $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$ and $B = A^t$. Then $W(A) = W(B)$ and $\|\alpha A + \beta A^* + \gamma I\| = \|\alpha B + \beta B^* + \gamma I\|$ for all $\alpha, \beta, \gamma \in \mathbf{C}$. However,

$$\|2AA^* + A^*A\| = 9 > 8 = \|2BB^* + B^*B\|,$$

showing that A cannot be dilated to an operator of the form $B \otimes I$.

4. Trapezoidal region

In this section, we consider an extension of Proposition 2.3 in a different direction, namely, we show that $A \in M_n$ can be dilated to operators with simple structure if $W(A)$ is included in a trapezoidal region.

Theorem 4.1 *Let K be a trapezoidal region in $\mathbf{C}$. Then for any $A \in M_n$ satisfying $W(A) \subseteq K$, there exist $k \le n$ and $B_1, \dots, B_k \in M_2$ with $W(B_j) \subseteq K$ such that $B_1 \oplus \cdots \oplus B_k$ is a dilation of A. Furthermore, if K is a square with vertices $0, 1, i$ and $1+i$, then $\operatorname{Re} B_j$ and $\operatorname{Im} B_j$ can be chosen as rank one projections for all j.*

Proof. We start with the special case that K is the square with vertices $0, 1, i$ and $1+i$. Suppose $A = A_1 + iA_2$, where A_1 and A_2 are self-adjoint. Then $W(A) \subseteq K$ if and only if A_1 and A_2 are positive semi-definite contractions. Thus $C_j = X_j^* X_j$ with $X_j = \begin{pmatrix} {A_j}^{1/2} & (I - A_j)^{1/2} \end{pmatrix}$ is a rank n projection for $j = 1, 2$. It is well known (see e.g., [D], [H1]) that $C_1 + iC_2$ is unitarily similar to a direct sum of $B_1, \dots, B_n \in M_2$ such that $\operatorname{Re} B_s$ and $\operatorname{Im} B_s$ are rank one projections for all $s = 1, \dots, n$.

If K is a parallelogram, then there exist real numbers $a_1, a_2, b_1, b_2, c_1, c_2$ so that the affine transformation $f(x+iy) = (a_1x + b_1y + c_1) + i(a_2x + b_2y + c_2)$ on $\mathbf{C}$ is invertible and f transforms K to the square with vertices $0, 1, i$ and $1+i$. Suppose $W(A) \subseteq K$. Then $f(A)$, defined as in Proposition 1.3, satisfies $W(f(A)) = f(W(A)) \subseteq f(K)$. By the result in the preceding paragraph, $f(A)$ can be dilated to a matrix of the form $B_1 \oplus \cdots \oplus B_n$ with $B_j \in M_2$ for all j. One easily checks that A can be dilated to $f^{-1}(B_1) \oplus \cdots \oplus f^{-1}(B_n)$.

Now, if K is a trapezoidal region, then there exist real constants $a_1, a_2, b_1, b_2, c_1, c_2$ so that the affine transformation $f(x+iy) = (a_1x + b_1y + c_1) + i(a_2x + b_2y + c_2)$ on $\mathbf{C}$ is invertible and f transforms K to the trapezoidal region with $0, 1, i$, and $1 + (1-a)i$ for some $a < 1$. Suppose $A \in M_n$ satisfies $W(A) \subseteq K$. Then $W(f(A)) = f(W(A)) \subseteq f(K)$. Similar to the previous case, we need only to prove that $f(A)$ can be dilated to a direct sum of no more than n matrices in M_2. The result will then follow. To this end, write $f(A) = A_1 + iA_2$. Then $f(A)$ lies in the trapezoidal region implies that

$$0 \le A_1 \le I \quad \text{and} \quad 0 \le A_2 \le I - aA_1.$$

We first dilated A_1 to a projection

$$\tilde{A}_1 = \begin{pmatrix} A_1 & (A_1 - A_1^2)^{1/2} \\ (A_1 - A_1^2)^{1/2} & I - A_1 \end{pmatrix}.$$

As $T = I - aA_1 \geq A_2$, there is a positive semi-definite contraction J such that $A_2 = T^{1/2}JT^{1/2}$. Since $T = I - aA_1$ is the left top corner of the matrix $\tilde{T} = I_{2n} - a\tilde{A}_1 \geq 0$, if we let $G = -aT^{-1/2}(A_1 - A_1^2)^{1/2}$ and $H = (1-a)^{1/2}T^{-1/2}$, then $\tilde{T}$ can be written as X^*X, where

$$X = \begin{pmatrix} T^{1/2} & G \\ 0_n & H \end{pmatrix}$$

Consequently, $\tilde{A}_2 = X^*PX$ is a dilation of $A_2 = T^{1/2}JT^{1/2}$ satisfying $0 \leq \tilde{A}_2 \leq I_{2n} - a\tilde{A}_1$, where $P = (\, J^{1/2} \quad (I_n - J)^{1/2}\,)^* (\, J^{1/2} \quad (I_n - J)^{1/2}\,)$ is a projection. By the polar decomposition, $X = U\tilde{T}^{1/2}$ where U is unitary matrix. Thus

$$\tilde{A}_2 = X^*PX = \tilde{T}^{1/2}U^*PU\tilde{T}^{1/2} = (I_{2n} - a\tilde{A}_1)^{1/2}Q(I_{2n} - a\tilde{A}_1)^{1/2}$$

with $Q = U^*PU$. Since $\tilde{A}_1$ and Q are projections, they are simultaneously unitarily similar to direct sums of 2×2 projections (see e.g. [D] and [H1]). It follows that $\tilde{A}_1 + i\tilde{A}_2$ is unitarily similar to a direct sum of 2×2 matrices. $\square$

It is not difficult to see from the proof that the choice of matrices $B_1, \ldots, B_k \in M_2$ in the statement of Theorem 4.1 depends heavily on the given operator A and is not unique.

In view of the above remark, one may ask whether there is a finite collection of $B_1, \ldots, B_m \in M_2$ with $W(B_j) \in K$ for all j so that every $A \in \mathcal{B}(\mathcal{H})$ satisfying $W(A) \subseteq K$ can be dilated to a matrix of the form $(B_1 \oplus \cdots \oplus B_m) \otimes I$. Evidently, $B = B_1 \oplus \cdots \oplus B_m$ is a finite matrix. Thus, more generally, one may ask whether there is a finite matrix B with $W(B) = K$ so that every operator A with $W(A) \subseteq K$ has a dilation of the form $B \otimes I$. The following result shows that the answer of this question is negative.

Proposition 4.2 *Suppose K is the square region with vertices $0, 1, i$ and $1+i$. Let $\mathcal{C}$ be the collection of matrices*

$$C_t = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + i \begin{pmatrix} \cos^2 t & \sin t \cos t \\ \sin t \cos t & \sin^2 t \end{pmatrix}, \quad t \in [0, \pi).$$

If $A \in M_n$ satisfies $W(A) \subseteq K$, then A has a dilation of the form B, where B is a direct sum of no more than n matrices in $\mathcal{C}$. Moreover, if C_t has a dilation C with $W(C) \subseteq K$, *then C is unitarily similar to $C_t \oplus \tilde{C}$ for some $\tilde{C}$. Consequently, there is no finite matrix B with $W(B) = K$ so that every operator A with $W(A) \subseteq K$ has a dilation of the form $B \otimes I$.*

Proof. If $B \in M_2$ is such that both $\operatorname{Re} B$ and $\operatorname{Im} B$ are rank one projections, then B is unitarily similar to a matrix of the form

$$C_t = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + i \begin{pmatrix} \cos^2 t & \sin t \cos t \\ \sin t \cos t & \sin^2 t \end{pmatrix}$$

for some $t \in [0, \pi)$. By Theorem 4.1, we get the first assertion of the theorem.

Now suppose C is a dilation of C_t, and $W(C) \subseteq K$. Then $0 \le \operatorname{Re} C, \operatorname{Im} C \le I$. Moreover, $\operatorname{Re} C$ contains the rank one projection $\operatorname{Re} C_t$, and $\operatorname{Im} C$ contains the rank one projection $\operatorname{Im} C_t$. The second assertion of the theorem follows.

To prove the last assertion, assume there exists $B \in M_n$ $W(B) = K$ so that every operator A with $W(A) \subseteq K$ has a dilation of the form $B \otimes I$. Then for each $t \in [0, \pi)$, B is unitarily similar to $C_t \oplus \tilde{B}_t$ for some $\tilde{B}_t$, which is impossible. $\square$

Note that every parallelogram can be converted to a square with prescribed vertices by an affine transform on $\mathbf{C}$. By Proposition 1.3, one can extend Proposition 4.2 to the case when K is a parallelogram. With some more effort, one may further extend the results to trapezoidal regions. In general, it would be interesting to study

Problem 4.3 *Is it possible to extend the results of Theorem 4.1 to arbitrary quadrangles or convex polygons?*

5. Remarks

Our study can be viewed as the starting point of using the numerical range to study the theory of dilation. As shown in the previous sections, there are many interesting questions that deserve further research. Moreover, one may consider our problem in some more general contexts involving other concepts. In a forthcoming project, we will study relations among numerical ranges, dilations, C^*-convex sets and positive linear maps.

Acknowledgement

Research of the first author was supported by an NSERC grant. Research of the second author was supported by an NSF grant. This paper was done while he was visiting the University of Toronto in the academic year 1998-1999 supported by a Faculty Research Grant of the College of William and Mary. He would like to thank the staff of the University of Toronto for their warm hospitality. Also, the authors would like to thank Dr. Nam-Kiu

Tsing for some helpful discussion. In particular, Theorem 2.1 was first obtained by him using a different proof.

References

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

[Ar] W.B. Arveson, Subalgebras of C^*-algebras II, *Acta Math.* 123 (1972), 271-308.

[D] C. Davis, Separation of two linear subspaces, *Acta Sci. Math.* (*Szeged*) 19 (1958), 172-187.

[Do] W.F. Donoghue, Jr., On the numerical range of a bounded operator, *Michigan J. Math.* 4 (1957), 261-263.

[FN] T. Furuta and R. Nakamota, On the numerical range of an operator, *Proc. Japan Acad.* 47 (1971), 279-284.

[H1] P.A. Halmos, Two subspaces, *Trans. of Amer. Math. Soc.* 144 (1969), 381-389.

[H2] P.A. Halmos, *A Hilbert Space Problem Book,* 2nd ed., Springer-Verlag, New York, 1982.

[HoJ] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis,* Cambridge University Press, New York, 1991.

[M] B.A. Mirman, Numerical range and norm of a linear operator, *Voronež. Gos. Univ. Trudy Sem. Funkcional Anal.,* 10 (1968), 51-55.

[Na] Y. Nakamura, Numerical range and norm, *Math. Japonica* 27 (1982), 149-150.

[W] A. Wintner, *Spectral theorie der unendlichen Matrizen: Einführung in den analytischen Apparat der Quantenmechanik,* S. Hirzel, Leipzig, 34-37, 1929.