EXCURSIONS IN NUMERICAL RANGES

HWA-LONG GAU [1,a] AND PEI YUAN WU[2,b]

Dedicated to Neil Trudinger on his 70th birthday

[1]Department of Mathematics, National Central University, Chungli 32001, Taiwan.

[a]E-mail: hlgau@math.ncu.edu.tw

[2]Department of Applied Mathematics, National Chiao Tung University, Hsinchu 30010, Taiwan.

[b]E-mail: pywu@math.nctu.edu.tw

Abstract

We survey some of our recent results on three topics in the study of numerical ranges, namely, (1) Anderson's condition for the numerical range of a finite matrix to equal a circular disc, (2) Holbrook's conjecture on the numerical radius inequality concerning the product of two commuting operators, and (3) Williams and Crimmins's structure theorem on an operator when its numerical radius equals half of its norm.

1. Introduction

The *numerical range* of a bounded linear operator A on a complex Hilbert space H is, by definition, the subset

$$W(A) = \{\langle Ax, x \rangle : x \in H, \|x\| = 1\}$$

of the complex plane $\mathbb{C}$, where $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$ are the inner product in H and its associated norm, respectively. The *numerical radius* of A is $w(A) = \sup\{|z| : z \in W(A)\}$. Thus $W(A)$ is the image of the unit circle in H under the quadratic form $f(x) \equiv \langle Ax, x \rangle$ from H to $\mathbb{C}$, and $w(A)$ is the smallest radius of a circular disc centered at the origin which contains $W(A)$.

The study of the numerical range has a history of almost one hundred years now. It started with the amazing result of Toeplitz–Hausdorff [38, 21]

that the numerical range is always a convex set in the plane. This means that the quadratic form f maps the unit circle $\|x\| = 1$ of H (no interior) to a subset of $\mathbb{C}$ with all its interior filled up. The subject was picked up slowly in the beginning. There are various generalizations of the numerical range to different contexts over the years. A group of British functional analysts considered it for elements of a normed algebra, which culminates, in the early 1970s, in the two monographs [3] and [4]. For the past decade or so, the topic has gone through a renaissance, due in a large part to the biennial "Workshop on Numerical Ranges and Numerical Radii" (WONRA) organized by the active and efficient Chi-Kwong Li. In recent years, a prominent development out of this is the applications of the higher-rank numerical ranges to the quantum information theory.

In this article, we will concentrate on the investigations concerning the classical numerical ranges of operators and finite matrices. The general problems to be considered in this area are the following:

(1) Given an operator A, what can be said about its numerical range $W(A)$?
(2) Conversely, if we know certain properties of $W(A)$, what can we deduce about the structure of A?
    and
(3) Which nonempty bounded convex subset of $\mathbb{C}$ is the numerical range $W(A)$ of some operator or matrix A (in some special classes)?

We start, in Section 2 below, by presenting the basic properties and examples of numerical ranges of operators. The ensuing three sections give the developments in recent years on three topics from the 1960s and early '70s. Section 3 is concerned with Anderson's result that if A is an n-by-n matrix whose numerical range $W(A)$ is contained in a circular disc D with their boundaries $\partial W(A)$ and ∂D intersect at most than n points, then $W(A)$ and D coincide with each other. Several analogues, generalizations and applications will be presented. If A is nilpotent, then we may decrease the number of intersection points from n to $n-2$ (cf. Theorem 3.3). On the other hand, we may also deduce a weaker conclusion when the circular disc is assumed to be contained in $W(A)$ (cf. Theorems 3.4, 3.5 and 3.6). The analogue for compact operators on an infinite-dimensional space, in which case we need infinitely many intersection points, is also given (cf. Theorem

3.10). In Section 4, we consider the "conjecture" of Holbrook on the inequality $w(AB) \leq \|A\|w(B)$ for commuting operators A and B. Although this is refuted by Müller and, subsequently, by Davidson and Holbrook in 1988, we are able to show, in contrast to the counterexample of the latter, that if the roles of A and B are switched and A is of class S_n, then $w(AB) \leq w(A)\|B\|$ still holds for commuting A and B (cf. Theorem 4.6). As a consequence, the same is true for A a quadratic operator. However, it is unknown whether $w(AB) \leq \|A\|w(B)$ if A is quadratic and $AB = BA$. In this respect, our contribution is that it is indeed the case if A is square-zero or idempotent. Finally, Section 5 deals with generalizations of Williams and Crimmins's result on an operator A with $w(A) = \|A\|/2$. The latter says that if A is such that $w(A) = |z|$ for some z in $W(A)$ and $w(A) = 1/2$ and $\|A\| = 1$, then it is unitarily equivalent to an operator of the form $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus B$. Since $w(A) \geq \|A\|/2$ for any operator A, this says something about the structure of A in the extremal case. Two generalizations are to be presented. One concerns a finite Blaschke product $f(z) = z \prod_{i=2}^{n} (z - a_i)/(1 - \overline{a}_i z)$, $|a_i| < 1$ for all i, of A with $w(A) = 1$ and $\|f(A)\| = 2$ (cf. Corollary 5.4). Another yields the inequality $w(A) \geq \cos(\pi/(k+2))$ for any contraction A with $\|A^k x\| = 1$ for some $k \geq 1$ and some unit vector x, and a corresponding Jordan block summand J_{k+1} for A when the inequality becomes an equality (cf. Theorem 5.5).

2. Preliminaries

In this section, we give some basic properties and examples of numerical ranges of operators which are to be used in later discussions.

Proposition 2.1. *For operators A and B on spaces H and K, respectively, the following hold:*

(a) *$W(A)$ is a nonempty bounded convex subset of $\mathbb{C}$. If H is finite dimensional, then $W(A)$ is even compact.*

(b) *$W(aA + bI) = aW(A) + b$ for any complex a and b.*

(c) *$W(U^*AU) = W(A)$ for any unitary operator U on H.*

(d) *If A is unitarily equivalent to an operator of the form $\begin{bmatrix} B & * \\ * & * \end{bmatrix}$ (that is, A is a dilation of B or B is a compression of A), then $W(B) \subseteq W(A)$.*

(e) *A complex number z is in $W(A)$ if and only if A is unitarily equivalent to an operator of the form $\begin{bmatrix} z & * \\ * & * \end{bmatrix}$.*

(f) *The spectrum $\sigma(A)$ of A is always contained in $\overline{W(A)}$.*

(g) *$W(A \oplus B) = (W(A) \cup W(B))^\wedge$, the convex hull of the union $W(A) \cup W(B)$.*

(h) *If A is normal, then $\overline{W(A)} = \sigma(A)^\wedge$.*

(i) *$\|A\|/2 \le w(A) \le \|A\|$.*

We next give some commonly seen examples of numerical ranges.

Example 2.2. If A is an operator on a two-dimensional space represented as a 2-by-2 upper-triangular matrix $\begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$, then $W(A)$ is the elliptic disc with foci the eigenvalues a and c of A and with its minor axis of length $|b|$.

Example 2.3. If A is a normal operator on an n-dimensional space with eigenvalues $a_1, \ldots, a_n$, then $W(A)$ is a polygonal region with some of the a_j's as vertices.

Example 2.4. If A is the n-by-n *Jordan block*

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

then $W(A) = \{z \in \mathbb{C} : |z| \le \cos(\pi/(n+1))\}$.

Example 2.5. If A is the *unilateral shift*

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

on ℓ^2, then $W(A)$ equals the open unit disc $\mathbb{D} \equiv \{z \in \mathbb{C} : |z| < 1\}$.

In studying the numerical ranges of finite matrices, we have an additional tool which is not available for general operators, namely, the Kippenhahn polynomial. For an n-by-n matrix A, let $p_A(x, y, z) = \det(x\operatorname{Re} A +$

$y\mathrm{Im}\,A + zI_n)$ denote its *Kippenhahn polynomial*, where $\mathrm{Re}\,A = (A + A^*)/2$ and $\mathrm{Im}\,A = (A - A^*)/(2i)$ denote the *real* and *imaginary parts* of A, respectively. Note that $p_A(x, y, z)$ is a degree-n homogeneous polynomial in the three variables x, y and z. It encodes plenty of the spectral and numerical range information of A. This is because $p_A(-1, -i, z) = \det(zI_n - A)$ is the characteristic polynomial of A, whose zeros are exactly the eigenvalues of A. On the other hand, $p_A(-\cos\theta, -\sin\theta, z) = \det(zI_n - \mathrm{Re}\,(e^{-i\theta}A))$ is the characteristic polynomial of $\mathrm{Re}\,(e^{-i\theta}A)$ for any real θ, whose maximum zero is the (signed) distance from the origin to the supporting line L_θ of $W(A)$, which is perpendicular to the ray R_θ from the origin with inclination θ from the positive x-axis (cf. Figure 1). As θ runs over all numbers in $[0, 2\pi)$, the supporting lines L_θ form an envelope of $W(A)$, which clearly yields the numerical range of A.
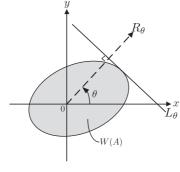

Figure 1

The technical statement of the above is the following theorem of Kippenhahn (cf. [26, 43]).

Theorem 2.6. *If A is an n-by-n matrix, then $W(A)$ equals the convex hull of the real points (u, v) for which $ux + vy + z = 0$ is a tangent line to the curve $p_A(x, y, z) = 0$ in the complex projective plane $\mathbb{CP}^2$.*

Our main references for the numerical ranges of operators are [20, Chapter 22] and [19], and for matrices [25, Chapter 1].

3. Anderson's Theorem

In the early 1970s, Anderson obtained a condition for the numerical range of a finite matrix to be equal to a circular (elliptic) disc.

Theorem 3.1. *If A is an n-by-n matrix such that $W(A)$ is contained in a closed elliptic disc E and their boundaries $\partial W(A)$ and ∂E intersect at more than n points, then $W(A) = E$.*

Anderson's original proof, which is never published, is based on Kippenhahn's result and Bézout's theorem. Recall that the latter says that if $p(x, y, z)$ and $q(x, y, z)$ are homogeneous complex polynomials of degrees m and n, respectively, which have no common factor, then they have exactly mn common zeros counting multiplicities (cf. [27, Theorem 3.1]). Theorem 3.1 can be seen by assuming that $E = \overline{\mathbb{D}}$ (since the involved properties are all invariant under affine transforms) and noting that $p_A(x, y, z) = 0$ and the irreducible quadratic curve $q(x, y, z) = x^2 + y^2 - z^2 = 0$, which corresponds to the unit circle $\partial \mathbb{D}$, have more than n intersection points, each with a common tangent line. Hence Bézout's theorem yields that q is a factor of p_A and thus $\overline{\mathbb{D}} \subseteq W(A)$. The assertion of Theorem 3.1 then follows.

Another proof, due to the second author [37, Lemma 6], is via the classical theorem of Riesz–Fejér (cf. [31, p. 77, Problem 40]). Indeed, the assumption $W(A) \subseteq \overline{\mathbb{D}}$ is equivalent to $\operatorname{Re}(e^{-i\theta} A) \leq I_n$ for all real θ. If $p(e^{i\theta}) = \det(I_n - \operatorname{Re}(e^{-i\theta} A))$, then $p(e^{i\theta}) = \sum_{j=-n}^{n} a_j e^{ij\theta}$ is a trigonometric polynomial which assumes only nonnegative values. Thus the Riesz–Fejér theorem yields that $p(e^{i\theta}) = |q(e^{i\theta})|^2$ for some polynomial of degree at most n. Since a point $e^{i\theta_0}$ is in $\partial W(A)$ if and only if $p(e^{i\theta_0}) = 0$, the assumption on the intersection points of $\partial W(A)$ and $\partial \mathbb{D}$ implies that $p(e^{i\theta}) = 0$ for more than n many θ's. Thus the same is true for q, which yields $q \equiv 0$ and hence $p \equiv 0$. The latter is equivalent to $W(A) = \overline{\mathbb{D}}$.

A generalization of Theorem 3.1 using only the fact that a degree-n polynomial can have at most n zeros is given in [13].

Anderson's theorem rules out the possibility for the half disc $\{z \in \overline{\mathbb{D}} : \operatorname{Re} z \geq 0\}$ to be the numerical range of any finite matrix.

Another easy consequence of Theorem 3.1 is the following.

Proposition 3.2. *If $A = \sum_{j=1}^{k} \oplus A_j$ is an n-by-n matrix whose numerical range is an elliptic disc E, then $W(A_j) = E$ for some j.*

Since $W(A) = (\cup_{j=1}^{k} W(A_j))^{\wedge} = E$, the pigeonhole principle yields that, for some j, $\partial W(A_j) \cap \partial E$ contains infinitely many points, and thus Theorem 3.1 is applicable to A_j to give $W(A_j) = E$. This is a generalization of [37, Theorem 3].

If A is an n-by-n nilpotent matrix ($A^n = 0$), then the critical number of intersection points n in Theorem 3.1 can be reduced to $n - 2$ to achieve the same conclusion.

Theorem 3.3. *If A is an n-by-n nilpotent matrix with $W(A)$ contained in $\overline{\mathbb{D}}$ and $\partial W(A) \cap \partial \mathbb{D}$ containing more than $n - 2$ points, then $W(A) = \overline{\mathbb{D}}$. In this case, the number "$n - 2$" is sharp.*

This is in [17, Propositions 3.1 and 3.2]. Again, it can be proven via the Riesz–Fejér theorem. The magic number $n - 2$ is explained by the fact that, in this case, the trigonometric polynomial $p(e^{i\theta}) = \det(I_n - \mathrm{Re}\,(e^{-i\theta}A))$ is of degree at most $n - 2$. Its sharpness can be seen by the matrix

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

which is such that $W(A_n) \subsetneq \overline{\mathbb{D}}$ and $\partial W(A_n) \cap \partial \mathbb{D}$ contains exactly $n - 2$ points.

It is interesting to know what conclusion can be drawn about $W(A)$ when it is assumed to contain an elliptic disc E. The next theorem [17, Theorem 2.5] gives the answer. It says that though $W(A) = E$ may not be true in general, the boundary $\partial W(A)$ does contain an arc of ∂E.

Theorem 3.4. *Let A be an n-by-n ($n \geq 3$) matrix. Then*

(a) *$\partial W(A)$ can contain at most $n - 2$ arcs of any ellipse, and*

(b) *if $W(A)$ contains an elliptic disc E and $\partial W(A)$ and ∂E intersect at more than n points, then $\partial W(A)$ contains an arc of ∂E.*

In this case, both numbers "$n-2$" and "n" are sharp.

The proof of (a) depends on some analytic arguments on the polynomial $p(z, e^{i\theta}) \equiv \det(zI_n - \text{Re}\,(e^{-i\theta}A))$ for z in $\mathbb{C}$ and real θ. (b) is then proven by enlarging A to an $(n+2)$-by-$(n+2)$ matrix and making use of (a). The sharpness of "$n-2$" and "n" is seen by the matrices $A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus \text{diag}\,(r, r\omega_{n-2}, r\omega_{n-2}^2, \ldots, r\omega_{n-2}^{n-3})$, where $1 < r < \sec(\pi/(n-2))$ and $\omega_{n-2} = e^{2\pi i/(n-2)}$, and $A = \text{diag}\,(1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1})$, where $\omega_n = e^{2\pi i/n}$, respectively.

For a nilpotent A, Theorem 3.4 has the following analogue (cf. [17, Theorem 3.5]).

Theorem 3.5. *Let A be an n-by-n nilpotent matrix.*

(a) *If $n \geq 5$, then $\partial W(A)$ contains at most $n-4$ arcs of any circle centered at the origin.*

(b) *If $W(A)$ contains a closed circular disc D centered at the origin and $\partial W(A)$ and ∂D intersect at more than $n-2$ points, then $W(A) = D$ if $2 \leq n \leq 4$, and $\partial W(A)$ contains at least one arc of ∂D if $n \geq 5$.*

In this case, both "$n-4$" and "$n-2$" are sharp.

The proof, which makes full use of the nilpotency of A, is a modification of that for Theorem 3.4.

Another specialization of the preceding results is to the companion matrices. Recall that a *companion matrix* is one of the form

$$
\begin{bmatrix}
0 & 1 & & & \\
 & 0 & \ddots & & \\
 & & \ddots & 1 & \\
 & & & 0 & 1 \\
-a_n & -a_{n-1} & \cdots & -a_2 & -a_1
\end{bmatrix}.
$$

It is well known that the characteristic and minimal polynomials of such a matrix are both $z^n + \sum_{j=1}^{n} a_j z^{n-j}$. For companion matrices, we have the following stronger conclusion.

Theorem 3.6. *If A is an n-by-n companion matrix with $W(A)$ containing D, a closed circular disc centered at the origin, and $\partial W(A)$ and ∂D intersecting at more than n points, then A equals the Jordan block J_n and, in particular, $W(A) = D$.*

This is proven first in [14, Theorem 2.9] under the assumption $W(A) = D$, and in [16, Theorem 1] for the general case $W(A) \supseteq D$. Its proof makes use of the fact that J_{n-1} is a principal submatrix of A and also the Riesz–Fejér theorem.

The next generalization of Anderson's theorem [41, Theorem 1] is useful in deducing, when $W(A)$ is a circular disc, properties of its center.

Theorem 3.7. *If A is an n-by-n matrix of the form*

$$\begin{bmatrix} aI_m & B \\ 0 & C \end{bmatrix},$$

where $0 \leq m < n$, such that $W(A)$ is contained in a closed circular disc D centered at a and $\partial W(A) \cap \partial D$ has more than $n-m$ points, then $W(A) = D$.

The case $m = 0$ corresponds to Anderson's theorem. Again, the Riesz–Fejér theorem plays a major role in its proof.

Note that if A is an n-by-n matrix such that $W(A)$ is an elliptic disc E, then the two foci of ∂E are eigenvalues of A. This can be seen via the fact that the irreducible quadratic polynomial $q(x, y, z)$ which corresponds to the dual ellipse of ∂E is a factor of $p_A(x, y, z)$ by Bézout's theorem. Thus, in particular, if $W(A)$ is a circular disc, then its center is an eigenvalue of A with algebraic multiplicity at least 2. Theorem 3.7 can be used to deduce more precise spectral information of the center.

Theorem 3.8. *If A is an n-by-n matrix such that $\partial W(A)$ contains more than $2n$ points of a circle C centered at a, then $W(A) \supseteq C^{\wedge}$ and a is an eigenvalue of A with its geometric multiplicity strictly less than its algebraic multiplicity. In this case, the number "$2n$" is sharp.*

Note that, in the above situation, $W(A)$ may not be equal to $C^{\wedge}$ as the example $A = J_{n-1} \oplus [1]$ shows. The sharpness of "$2n$" is seen by the n-by-n diagonal matrix $A = (1+\varepsilon)\text{diag}\,(1, \omega_n, \omega_n^2, \ldots, \omega_n^{n-1})$, where $\varepsilon > 0$ is sufficiently small.

The preceding theorem rules out the circular discs as the numerical ranges of certain special types of matrices.

Corollary 3.9. *Let A be an n-by-n matrix. If* (a) *A is similar to a normal matrix,* (b) *A is nonnegative and permutationally irreducible, or* (c) *A is row (resp., column) stochastic, then $W(A)$ cannot be a circular disc.*

In case (a), every eigenvalue of A has equal algebraic and geometric multiplicities. Thus $W(A)$ is not a circular disc by Theorem 3.8. (b) is proven in [28, Theorem 4.5]. It can be seen by noting that if $W(A)$ is a circular disc, then its center coincides with the strictly positive maximum eigenvalue of A by [28, Theorem 4.8], the two multiplicities of which are both equal to 1, thus contradicting Theorem 3.8. Recall that a matrix $A = [a_{ij}]_{i,j=1}^n$ is *row* (resp., *column*) *stochastic* if $a_{ij} \geq 0$ for all i and j, and $\sum_{j=1}^n a_{ij} = 1$ for all i (resp., $\sum_{i=1}^n a_{ij} = 1$ for all j). Every row (resp., column) stochastic A is permutationally similar to a direct sum $\sum_{j=1}^k \oplus A_j$ with A_j row (resp., column) stochastic and $\text{Re}\,A_j$ permutationally irreducible for all j. Assuming that $W(A)$ is a circular disc D, we infer from Proposition 3.2 that $W(A_j) = D$ for some j. Since the center of D equals the maximum eigenvalue 1 of A_j by [28, Theorem 4.8], and the algebraic and geometric multiplicities of 1 are equal to each other by [34, Lemma 6.3 (vi)], we have a contradiction as before.

Theorems 3.7, 3.8 and Corollary 3.9 (a) were proven later in [6] by elementary arguments using only properties of polynomials and continuous functions.

For compact operators on an infinite-dimensional space, there is also an analogue of Anderson's theorem.

Theorem 3.10. *If A is a compact operator such that $W(A)$ is contained in a closed circular disc D centered at the origin and $\partial W(A) \cap \partial D$ has infinitely many points, then $W(A) = D$.*

This is in [15, Theorem 1]. If the compact A acts on an infinite-dimensional space, then, instead of the trigonometric polynomial $p(e^{i\theta}) = \det(I_n - \mathrm{Re}\,(e^{-i\theta}A))$ defined in terms of the determinant, we need consider an analytic branch of the function $d_A(e^{i\theta}) \equiv \max \overline{W(\mathrm{Re}\,(e^{-i\theta}A))}$ for real θ. Our assumptions yield that $d_A(e^{i\theta}) \leq r$ for all θ and $d_A(e^{i\theta}) = r$ for infinitely many θ's, where r is the radius of D. Thus $d_A(e^{i\theta})$ is identically equal to r, which then gives $W(A) = D$. As before, this theorem implies that no half-disc centered at the origin can be the numerical range of a compact operator.

4. Holbrook's Conjecture

In 1969, Holbrook [22] considered a numerical radius inequality of the commuting product of two operators: If A and B are two commuting operators, then $w(AB) \leq \min\{\|A\|w(B), w(A)\|B\|\}$. Note that if no commutativity assumption is made on A and B, then, in comparing the three numerical radii $w(AB)$, $w(A)$ and $w(B)$, the best we can have is that $w(AB) \leq 4w(A)w(B)$. For example, if $A = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$ and $B = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$, then $AB = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, in which case $w(A) = w(B) = 1/2$ and $w(AB) = 1$, and thus the equality $w(AB) = 4w(A)w(B)$ holds. On the other hand, if A and B commute, then the constant "4" on the right-hand side of the inequality can be reduced to "2", that is, $w(AB) \leq 2w(A)w(B)$ is true. Again, in this case, "2" is sharp. This is seen by $A = \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right] \oplus \left[\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right]$ and

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

in which case we have $w(AB) = w(A) = w(B) = 1/2$.

Back then, Holbrook had already shown the validity of his conjecture for *doubly commuting* A and B, that is, for A and B satisfying $AB = BA$ and $AB^* = B^*A$. This is also obtained independently by Sz.-Nagy [35].

Theorem 4.1. *If A and B doubly commute, then $w(AB) \leq \min\{\|A\|w(B), w(A)\|B\|\}$. In particular, this is the case if A or B is normal and $AB = BA$.*

The normal case is proven by first checking for diagonal operators and then by approximating them to the normal ones. For the doubly commuting case, this is proven by simultaneously dilating such a pair to two commuting normal operators.

Another partial generalization of the normal case is the next result from [5, Lemma 2] concerning isometries.

Theorem 4.2. *If A and B commute and A is an isometry, then $w(AB) \leq \min\{\|A\|w(B), w(A)\|B\|\}$.*

Since A can be extended to a unitary operator U, the commuting B can also be extended to an operator C which commutes with U. The inequality $w(AB) \leq \|A\|w(B)$ then follows by showing that $\overline{W(B)} = \overline{W(C)}$. The other inequality $w(AB) \leq w(A)\|B\|$ is trivial.

In [22, Theorem 4.4], Holbrook showed that the smallest constant c for which $w(AB) \leq c\|A\|w(B)$ holds for all commuting A and B is strictly less than 2. One specific such constant was obtained by Crabb [30, pp. 209–210].

Theorem 4.3. *If A and B commute, then $w(AB) \leq (\sqrt{2 + 2\sqrt{3}}/2)\|A\|w(B)$.*

Note that the constant $\sqrt{2 + 2\sqrt{3}}/2$ equals $1.1687\ldots$, which is bigger than 1. The proof of Theorem 4.3 involves some ingenious arguments, which can be further refined to yield even smaller estimates of c.

In light of all such partial results, it came as a surprise when in 1988 Müller [29] gave an example showing that the best constant c for which $w(AB) \leq c\|A\|w(B)$ holds for all commuting operators A and B is greater than 1.01 instead of the long-suspected 1. His example involves operators on a 12-dimensional space and relies on a computer check for its validity. The day is saved by Davidson and Holbrook: they gave in [8] simpler examples involving only zero-one matrices with the additional advantages that the computations can be carried out directly and the lower bound on the constant c can be improved. Here is one of their examples.

Example 4.4. If $A = J_9$ and $B = J_9^3 + J_9^7$, then A and B commute, $w(AB) = \|A\| = 1$ and $w(B) = \cos(\pi/10)$. Thus $w(AB) > \|A\|w(B)$.

This also shows that the constant c for which $w(AB) \leq c\|A\|w(B)$ holds for all A and B with $AB = BA$ is at least $1/\cos(\pi/10) = 1.0514\ldots$ In recent years, Holbrook and Schoch [24, p. 278] gave examples to show that $w(AB) > \|A\|w(B)$ may even occur for commuting 3-by-3 matrices. However, no such examples can exist for matrices of size 2. This was proven again by Holbrook [23].

Theorem 4.5. *If A and B are commuting 2-by-2 matrices, then $w(AB) \leq w(A)w(B)$.*

The proof of this theorem is based on the following lemma, which has some independent interest.

Lemma 4.1. *If A and B are commuting 2-by-2 matrices with $w(A), w(B) \leq 1$, then there is another 2-by-2 matric C with $w(C) \leq 1$ and there are analytic functions f and g from $\overline{\mathbb{D}}$ to $\overline{\mathbb{D}}$ such that $A = f(C)$ and $B = g(C)$.*

Another proof of Theorem 4.5, based on the Pick interpolation condition, is given in [24, Corollary 3.2].

Note that, in Example 4.4, the inequality $w(AB) \leq w(A)\|B\|$ is still true, revealing its asymmetric nature. Our contributions to this subject is that the latter inequality holds for a much larger class of operators.

Recall that a contraction A ($\|A\| \leq 1$) is *of class C_0* if it is *completely nonunitary* (meaning that A has no unitary direct summand) and satisfies $\phi(A) = 0$ for some ϕ in the Hardy space H^∞ of bounded analytic functions on $\mathbb{D}$. The *minimal function* of a C_0 contraction A is the annihilating ϕ in H^∞ which divides all other such annihilating functions of A. In this case, ϕ must be *inner*, namely, it satisfies $|\phi| = 1$ a.e. on the unit circle $\partial\mathbb{D}$. One example of C_0 contractions is the *compression of the shift* $S(\phi)$ for any inner function ϕ, defined on the space $H \equiv H^2 \ominus \phi H^2$ by $S(\phi)f = P_H(zf(z))|H$ for f in H, where P_H denotes the orthogonal projection from H^2 onto H. Note that the minimal function of $S(\phi)$ is ϕ and rank $(I - S(\phi)^*S(\phi)) = 1$ holds. If $\phi(z) = z^n$, then $S(\phi)$ is unitarily equivalent to the n-by-n Jordan block J_n. Such operators were first studied by Sarason [33] and later featured prominently in the Sz.-Nagy–Foiaş contraction theory [36] as the building blocks of the "Jordan model" for C_0 contractions [36, 2].

Theorem 4.6. *If A is a C_0 contraction with minimal function ϕ such that $W(A) = W(S(\phi))$ and if B commutes with A, then $w(AB) \le w(A)\|B\|$.*

This is proven in [42] first for A a compression of the shift $S(\phi)$ and, for the general C_0 contraction A with minimal function ϕ, by extending it to the direct sum A_1 of copies of $S(\phi)$ and by extending the commuting B to an operator B_1 which commutes with A_1 and satisfies $\|B_1\| = \|B\|$.

Note that the extra condition that $W(A) = W(S(\phi))$ in Theorem 4.6 is essential for otherwise the example of $B = (J_9 + (1/4)J_9^5)/\|J_9 + (1/4)J_9^5\|$ and $A = B^3$ attests the falsity of $w(AB) \le w(A)\|B\|$.

An operator A is said to be *quadratic* if it satisfies $A^2 + aA + bI = 0$ for some complex numbers a and b. Using Theorem 4.6, we can prove a numerical radius inequality for quadratic operators (cf. [42, Theorem 5]). It is a partial generalization of Theorem 4.5.

Theorem 4.7. *If A is a quadratic operator and B commutes with A, then $w(AB) \le w(A)\|B\|$.*

Note that the inequality above is not necessarily true if A is annihilated by a cubic polynomial. For example, it was shown in [8, Corollary 4] that if

$$
A = \begin{bmatrix} 0 & I_3 & J_3 \\ & 0 & I_3 \\ & & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} J_3 & & \\ & J_3 & \\ & & J_3 \end{bmatrix},
$$

then $A^3 = B^3 = 0$, $AB = BA$, $w(A) = \cos(\pi/10)$, $\|B\| = 1$ and $w(AB) = 1$. Thus, in this case, $w(AB) > w(A)\|B\|$. The remaining open question is whether $w(AB) \le \|A\|w(B)$ holds for A quadratic and B commuting with A. Some special cases of it are known to be true. For example, this is the case if A satisfies $A^2 = aI$ for some scalar a (cf. [32]).

Theorem 4.8. *If A is an operator satisfying $A^2 = aI$ for some scalar a and B commutes with A, then $w(AB) \le \|A\|w(B)$.*

Another known case is the following.

Theorem 4.9. *If A is an idempotent operator ($A^2 = A$) and B commutes with A, then $w(AB) \le \|A\|w(B)$.*

This is from [10, Theorem 3]. Its proof makes use of Theorem 4.1, the doubly commuting case.

5. Williams–Crimmins's Result

As we have seen in Proposition 2.1 (i) that $w(A) \geq \|A\|/2$ holds for any operator A. The result of Williams and Crimmins [39] gives the structure of A when this becomes an equality.

Theorem 5.1. *Let A be an operator on H such that $\|Ax\| = \|A\|$ for some unit vector x in H. If $w(A) = 1$ and $\|A\| = 2$, then A is unitarily equivalent to an operator of the form $\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus A'$ and $W(A) = \overline{\mathbb{D}}$.*

The proof is quite straightforward. Indeed, if $y = (1/2)Ax$, then y and x are orthonormal vectors. Let K be the subspace of H generated by y and x. Then A is unitarily equivalent to $\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \oplus A'$ on the decomposition $H = K \oplus K^\perp$ of H.

A generalization of the preceding theorem was obtained by Crabb [7].

Theorem 5.2. *If A is an operator on H with $w(A) \leq 1$ and $\|A^n x\| = 2$ for some $n \geq 1$ and some unit vector x in H, then A is unitarily equivalent to an operator of the form $B \oplus C$, where B is the $(n+1)$-by-$(n+1)$ matrix*

$$\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \quad or \quad \begin{bmatrix} 0 & \sqrt{2} & & & & \\ & 0 & 1 & & & \\ & & 0 & \ddots & & \\ & & & \ddots & 1 & \\ & & & & 0 & \sqrt{2} \\ & & & & & 0 \end{bmatrix}$$

depending on whether $n = 1$ or $n \geq 2$. In this case, $W(A) = \overline{\mathbb{D}}$.

As for Theorem 5.1, the subspace K of H on which B acts is generated by the mutually orthogonal vectors $x, Ax, \ldots, A^n x$. The proof itself is quite involved, not easily to be further generalized.

In 2009, the authors were able to give in [18] a generalization of the preceding theorem.

Theorem 5.3. *Let f be a function in H^∞ with $\|f\|_\infty \le 1$. Then there is an operator A on H with $w(A) \le 1$ which has no unitary summand and has a unit vector x in H such that $\|f(A)x\| = 2$ if and only if f is inner and $f(0) = 0$. In this case, A has a direct summand similar to $S(\phi)$, where $\phi(z) = zf(z)$ for z in $\mathbb{D}$.*

A finite-dimensional version of this confirms Drury's Conjecture 6 in [9].

Corollary 5.4. *Let $f : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ be a function analytic on $\mathbb{D}$ and continuous on $\overline{\mathbb{D}}$. Then there is an operator A on H with $w(A) \le 1$ and $\|f(A)x\| = 2$ for some unit vector x in H if and only if $f(z) = z \prod_{i=2}^{n}(z - a_i)/(1 - \overline{a}_i z)$ for some $n \ge 1$ and $a_2, \ldots, a_n$ in $\mathbb{D}$. In this case, A is unitarily equivalent to an operator of the form $B \oplus C$, where $B = [b_{ij}]_{i,j=1}^{n+1}$ is such that $b_{11} = b_{n+1,n+1} = 0$, $b_{ii} = a_i$ for $2 \le i \le n$,*

$$
b_{ij} = \begin{cases}
\sqrt{2}a_{ij} & \text{if } 1 = i < j \le n \text{ or } 2 \le i < j = n+1, \\
2a_{ij} & \text{if } i = 1 \text{ and } j = n+1, \\
a_{ij} & \text{if } 2 \le i < j \le n, \\
0 & \text{if } i > j,
\end{cases}
$$

and

$$
a_{ij} = (-1)^{j-i-1}\overline{a}_{i+1}\cdots\overline{a}_{j-1}[(1 - |a_i|^2)(1 - |a_j|^2)]^{1/2} \quad \text{for } i < j.
$$

Moreover, in this case, $W(A) = \overline{\mathbb{D}}$.

Note that the case $f(z) = z$ corresponds exactly to Theorem 5.1, the Williams–Crimmins result, while $f(z) = z^n$ ($n \ge 1$) corresponds to Theorem 5.2.

The proof of Theorem 5.3 is quite intricate. It depends on a factorization theorem of Ando [1, Theorem 2] for operators A with $w(A) \le 1$. The matrix form of B in Corollary 5.4 is a consequence of the necessity proof of Theorem 5.3 and the upper-triangular matrix representation of the finite-dimensional compression of the shift $S(\phi)$ [12, Corollary 1.3].

Another finite-dimensional generalization of Theorem 5.1 was obtained more recently in [11, Theorem 2.10]. Later on, the first author found out that it is even true for operators.

Theorem 5.5. *Let A be a contraction on H. If $\|A^k x\| = 1$ for some $k \geq 1$ and some unit vector x in H, then $w(A) \geq \cos(\pi/(k+2))$. Moreover, in this case, the equality $w(A) = \cos(\pi/(k+2))$ holds if and only if A is unitarily equivalent to an operator of the form $J_{k+1} \oplus B$ with $w(B) \leq \cos(\pi/(k+2))$. Also, if this is the case, then $W(A) = \{z \in \mathbb{C} : |z| \leq \cos(\pi/(k+2))\}$.*

In [11, Theorem 2.5], the inequality $w(A) \geq \cos(\pi/(k+2))$ (for A a finite matrix) was proven by showing the equivalence of $\|A^k\| = 1$ and $w(A \otimes J_{k+1}) = w(J_{k+1})$. When the equality is true, we show that the subspace K of $\mathbb{C}^n$ generated by $x, Ax, \ldots, A^k x$, where x is a unit vector in $\mathbb{C}^n$ with $\|A^k x\| = 1$, is reducing for A and the restriction of A to K is unitarily equivalent to J_{k+1}. For an operator A, a more direct proof has been discovered by the first author recently.

In the preceding theorem, the case $k = 1$ corresponds to the Williams–Crimmins result. Another extremal case is for k to be equal to $n-1$ as the following corollary shows.

Corollary 5.6. *For an n-by-n matrix A with $\|A\| = 1$, the following conditions are equivalent:*

(a) $\|A^{n-1}\| = 1$ *and* $w(A) = \cos(\pi/(n+1))$,
(b) *A is unitarily equivalent to J_n, and*
(c) $\|A^{n-1}\| = 1$ *and* $A^n = 0$.

We conclude this section with one more equivalent condition, besides the ones in Corollary 5.6.

Theorem 5.7. *An n-by-n matrix A is unitarily equivalent to J_n if and only if $\|A\| = 1$ and $W(A) = \{z \in \mathbb{C} : |z| \leq \cos(\pi/(n+1))\}$.*

This is in [40, Theorem 3]. The proof is of a matricial nature, which is independent of the ones for Theorem 5.5 and Corollary 5.6.

Acknowledgments

Versions of this paper have been reported by the second author in various occasions in recent years, the most recent of which is "The International Workshop on Matrix Theory and Ring Theory" (July 19–22, 2013 at Nanjing, China). He likes to thank the organizers Jianlong Chen and Yimin Wei for their hospitality during the conference.

References

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

2. H. Bercovici, *Operator Theory and Arithmetic in H^∞*, Amer. Math. Soc., Providence, 1988.

3. F. F. Bonsall and J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, Cambridge University Press, Cambridge, 1971.

4. F. F. Bonsall and J. Duncan, *Numerical Ranges II*, Cambridge University Press, Cambridge, 1973.

5. R. Bouldin, The numerical range of a product, II, *J. Math. Anal. Appl.*, **33** (1971), 212-219.

6. W.-S. Cheung and C.-K. Li, Elementary proofs for some results on the circular symmetry of the numerical range, *Linear Multilinear Algebra*, 61 (2013), 596-602.

7. M. J. Crabb, The powers of an operator of numerical radius one, *Michigan Math. J.*, **18** (1971), 252-256.

8. K. R. Davidson and J. A. R. Holbrook, Numerical radii of zero-one matrices, *Michigan Math. J.*, **35** (1988), 261-267.

9. S. W. Drury, Symbolic calculus of operators with unit numerical radius, *Linear Algebra Appl.*, **428** (2008), 2061-2069.

10. H.-L. Gau, C.-Y. Huang and P. Y. Wu, Numerical radius inequalities for square-zero and idempotent operators, *Oper. Matrices*, **2** (2008), 137-141.

11. H.-L. Gau, K.-Z. Wang and P. Y. Wu, Numerical radii for tensor products of matrices, *Linear Multilinear Algebra*, to appear, arXiv: 1306.2423.

12. H.-L. Gau and P. Y. Wu, Lucas' theorem refined, *Linear Multilinear Algebra*, **45** (1999), 359-373.

13. H.-L. Gau and P. Y. Wu, Condition for the numerical range to contain an elliptic disc, *Linear Algebra Appl.*, **364** (2003), 213-222.

14. H.-L. Gau and P. Y. Wu, Companion matrices: reducibility, numerical ranges and similarity to contractions, *Linear Agebra Appl.*, **383** (2004), 127-142.

15. H.-L. Gau and P. Y. Wu, Anderson's theorem for compact operators, *Proc. Amer. Math. Soc.*, **134** (2006), 3159-3162.

16. H.-L. Gau and P. Y. Wu, Numerical ranges of companion matrices, *Linear Algebra Appl.*, **421** (2007), 202-218.

17. H.-L. Gau and P. Y. Wu, Line segments and elliptic arcs on the boundary of a numerical range, *Linear Multilinear Algebra*, **56** (2008), 131-142.

18. H.-L. Gau and P. Y. Wu, Inner functions of numerical contractions, *Linear Algebra Appl.*, **430** (2009), 2182-2191.

19. K. E. Gustafson and D. K. M. Rao, *Numerical Range. The Field of Values of Linear Operators and Matrices*, Springer, New York, 1997.

20. P. R. Halmos, *A Hilbert Space Problem Book*, 2nd ed., Springer, New York, 1982.

21. F. Hausdorff, Der Wertevorrat einer Bilinearform, *Math. Z.*, **3** (1919), 314-316.

22. J. A. R. Holbrook, Multiplicative properties of the numerical radius in operator theory, *J. Reine Angew. Math.*, **237** (1969), 166-174.

23. J. A. R. Holbrook, Inequalities of von Neumann type for small matrices, in: *Function Spaces*, K. Jarosz, ed., Marcel Dekker, New York, 1992, pp.189-193.

24. J. A. R. Holbrook and J.-P. Schoch, Theory vs. experiment: multiplicative inequalities for the numerical radius of commuting matrices, *Oper. Theory Adv. Appl.*, **202** (2010), 273-284.

25. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

26. R. Kippenhahn, Über den Wertevorrat einer Matrix, *Math. Nachr.*, **6** (1951), 193-228.

27. F. Kirwan, *Complex Algebraic Curves*, Cambridge University Press, Cambridge, 1992.

28. J. Maroulas, P. J. Psarrakos and M. J. Tsatsomeros, Perron–Frobenius type results on the numerical range, *Linear Algebra Appl.*, **348** (2002), 49-62.

29. V. Müller, The numerical radius of a commuting product, *Michigan Math. J.*, **35** (1988), 255-260.

30. K. Okubo and T. Ando, Operator radii of commuting products, *Proc. Amer. Math. Soc.*, **56** (1976), 203-210.

31. G. Polya and G. Szegö, *Problems and Theorems in Analysis*, vol. II, Springer, Berlin, 1976.

32. D. K. M. Rao, Rango numérico de productos commutativos, *Rev. Colombiana Mat.*, **27** (1993), 231-233.

33. D. Sarason, Generalized interpolation in H^∞, *Trans. Amer. Math. Soc.*, **127** (1967), 179-203.

34. S. Serra-Capizzno, Google PageRanking problem: the model and the analysis, *J. Comput. Applied Math.*, **234** (2010), 3140-3169.

35. B. Sz.-Nagy, Products of operators of classes C_ρ, *Rev. Roumaine Math. Pures Appl.*, **13** (1968), 897-899.

36. B. Sz.-Nagy, C. Foiaş, H. Bercovici and L. Kérchy, *Harmonic Analysis of Operators on Hilbert Space*, 2nd ed., Springer, New York, 2010.

37. B.-S. Tam and S. Yang, On matrices whose numerical ranges have circular or weak circular symmetry, *Linear Algebra Appl.*, **302/303** (1998), 193-221.

38. O. Toeplitz, Das algebraische Analogon zu einem Satze von Fejér, *Math. Z.*, **2** (1918), 187-197.

39. J. P. Williams and T. Crimmins, On the numerical radius of a linear operator, *Amer. Math. Monthly*, **74** (1967), 832-833.

40. P. Y. Wu, A numerical range characterization of Jordan blocks, *Linear Multilinear Algebra*, **43** (1998), 351-361.

41. P. Y. Wu, Numerical ranges as circular discs, *Applied Math. Lett.*, **24** (2011), 2115-2117.

42. P. Y. Wu, H.-L. Gau and M. C. Tsai, Numerical radius inequality for C_0 contractions, *Linear Algebra Appl.*, **430** (2009), 1509-1516.

43. P. E. Zachlin and M. E. Hochstenbach, On the numerical range of a matrix, *Linear Multilinear Algebra*, **56** (2008), 185-225.