Received: 25-01-2022
Accepted: 05-03-2022

International Journal of Advanced Multidisciplinary Research and Studies

## ISSN: 2583-049X

# Characterization of Numerical ranges of Matrix Operators in Hilbert Spaces 

${ }^{1}$ Benard Okelo, ${ }^{2}$ James Mbugua Ngigi<br>${ }^{1,2}$ School of Biological, Physical, Mathematics and Actuarial Sciences, Jaramogi Oginga Odinga University of Science and Technology, Bondo, Kenya<br>Corresponding Author: Benard Okelo


#### Abstract

Let H be a Hilbert space. Let ${ }^{A}$ be a $3 \times 3$ complex matrix whose numerical range, $W(A)$, is an ellipse. Also let $W(A) \subset \Omega$ where $\Omega$ is a convex subset of the complex


plane. We determine the value of $\psi_{D}(A)$ by reducing it to the case of $2 \times 2$ matrix with elliptical numerical range.

Keywords: Matricial Operator, Hilbert Space, Holomorphic

## 1. Introduction

The notion of numerical range was first introduced by O. Toeplitz in 1918 for matrices ${ }^{[1-5]}$. Toeplitz introduced the numerical range for finite dimensional spaces ${ }^{[6-9]}$. The concept was independently extended by G. Lumer and F. Bauer in sixties to a bounded linear operator on arbitrary Banach space. Lumer used numerical range techniques to characterize the isometries of certain reflexive Orlicz spaces ${ }^{[10-17]}$. In ${ }^{1975}$, lightbourne and Martin extended this concept by employing a class of semi norms generated by a family of supplementary projection. Bonsall and Duncan (1971) studied extensively numerical ranges of operators on normed spaces and of elements of normed algebras ${ }^{[18]}$. In 1981 Miroslav Fielder studied geometry of numerical ranges of matrices. He developed some techniques for the study of the algebraic curve $C(A)$ which generates the numerical range $W(A)$ of an $n \times n$ matrix $A$ as its convex hull is developed. These enable one to give an explicit point equation of $C(A)$ and a formula for the curvature of $C(A)$ at a boundary point of $W(A)$. Dennis S. Keeler (1997) gave a series of tests, allowing one to determine the shape of $W(A)$ for $3 \times 3$ matrices ${ }^{[19]}$. He showed that for a matrix $A$ itself or its canonical unitarily equivalent forms it is possible to determine when numerical range of a $3 \times 3$ matrix is an ellipse, a set with a flat portion on its boundary, or an ovular set. Michele Benzi (1999) in his article on bounds for the entries of matrix functions with applications to preconditioning showed that when a matrix $A$ is banded, the entries of $f(A)$ are bounded in an exponentially decaying manner away from the main diagonal. In 2004 Michel Crouzeix studied the bounds of functions of $2^{\times 2}$ matrices ${ }^{[20]}$. He showed that the bounds are attained and gave an explicit formula for $2 \times 2$ matrices ${ }^{[21]}$.

## Preliminaries

In this section, we give some basic definitions, results and theorems used in this study.
Definition 1.2.1: A set $S \subset \mathbb{R}^{n}$ or $\mathbb{C}^{n}$ is said to be convex if the line segment connecting $x$ and $y$ is contained in S i.e $\forall x, y \in S\{t x+(1-t) y: t \in[0,1]\} \subset S$.

Definition 1.2.2: Let $A \in \mathbb{C}^{n \times n}$. if $A x=\lambda x$ for $\lambda \in \mathbb{C}, x \neq 0$ then ${ }^{\lambda}$ is the eigenvalue of $A$ and ${ }^{x}$ is the eigenvector of ${ }^{A}$. The set of all eigenvalues of $A$ is the spectrum of ${ }^{A}$ denoted as $\sigma(A)$.

Definition 1.2.3: The convex hull of a set ${ }^{S}$, denoted $\operatorname{co}^{(S)}$, is the minimal convex set containing $S$.
Definition 1.2.4: For any analytic function $f$ defined on a set $D \subset \mathbb{C}$, it holds

$$
f(x)=\frac{1}{2 \pi i} \oint_{C} \frac{f(z)}{z-x} d z
$$

where ${ }^{C}$ is a closed curve inside the domain ${ }^{D}$ enclosing ${ }^{x}$. This is called the Cauchy integral formula.
Definition 1.2.5: Let $T: H_{1} \rightarrow H_{2}$ be a bounded linear operator, where $H_{1}$ and $H_{2}$ are Hilbert spaces. Then the Hilbert adjoint operator $T^{*}$ of $T$ is the operator $T^{*}: H_{2} \rightarrow H_{1}$ such that for all $x \in H_{1}$ and $y \in H_{2},(T x, y\rangle=\left\langle x, T^{*} y\right\rangle$.

Definition 1.2.6: Two vectors ${ }^{x}$ and $y$ in an inner product space $E$ are said to be orthogonal if and only if $\langle x, y\rangle=0, x \neq y$, denoted by $x^{\perp} y$.

Definition 1.2.7: Let ${ }^{S}$ be a subset of a Hilbert space $H$. The set of all vectors orthogonal to $S$ is called the orthogonal complement of S denoted by $\mathrm{S}^{\perp}$.i.e. $S^{\perp}=\left\{x \in H: x^{\perp} s, \forall s \in S\right\}$.

Definition 1.2.8: Let $A \in M_{n}$. Write $A=H+i K$ with $H, K$ Hermitiian, and let $L_{A}(u, v, w)=\operatorname{det}(u H+v K+w l)$. The equation $L_{A}(u, v, w)=0$, with $u, v, w$ viewed as homogeneous line coordinates defines an algebraic curve of class ${ }^{n}$ called Kippenhahn polynomial.

Definition 1.2.9: The real part of the algebraic curve $L_{A}(u, v, w)=0$ is called the associated curve denoted by $C(A)$.
Definition 1.2.10: We say that a matrix $A$ is reducible if there exist a unitary matrix $U$ such that

$$
U^{*} A U=\operatorname{diag}\left[A_{1}, A_{2}\right], \text { where both diagonal blocks are of non-zero size. }
$$

Definition 1.2.11: Let H be a Hilbert space. An operator $T: H \rightarrow H$ is called Hermitian or self-adjoint if $\quad T=T^{*}$.
Definition 1.2.12: Let H be a Hilbert space. An operator $T: H \rightarrow H$ is called unitary if $T T^{*}=T^{*} T=I$.
Definition 1.2.13: A complex valued function ${ }^{h}$ of a complex variable ${ }^{\lambda}$ is said to be holomorphic (or analytic) on a domain $G$ of the complex ${ }^{\lambda-}$ plane if $h$ is defined and differentiable on ${ }^{G}$. That is, the derivative $h^{\prime}$ of $h$ defined by

$$
h^{\prime}(\lambda)=\lim _{\Delta \lambda} \frac{h(\lambda+\Delta \lambda)-h(\lambda)}{\Delta \lambda} \text { exists for every } \lambda \in G .
$$

Definition 1.2.14: Let $a \in \bar{D}(0,1)$, the unit disk centred at origin. Then the Blaschke factor is defined by

$$
B_{a}(z)=\frac{z-a}{1-\bar{a} z}
$$

Definition 1.2.15: A Blaschke product is an expression of the form

$$
B(z)=z^{m} \prod_{j=1}^{\infty} \frac{-\overline{a_{j}}}{\left\|a_{j}\right\|} B a_{j}(z)
$$

Where ${ }^{m}$ is non-negative integer.
Definition 1.2.16: (Young's inequality) If $a$ and $b$ are non-negative real numbers and $p$ and $q$ are positive real numbers such

$$
\begin{aligned}
& \text { that } \frac{1}{p}+\frac{1}{q}=1 \text {, } \\
& \text { then } a b \leq \frac{a^{p}}{p}+\frac{b^{q}}{q} \text { When } p=q=2 \text {, we have } a b \leq \frac{a^{2}}{2}+\frac{b^{2}}{2} \text {. Equality occurs if and only if } b=a^{p-1} \text {. }
\end{aligned}
$$

Definition 1.2.17: (Nevanlinna-Pick interpolation problem). In complex analysis, given initial data consisting of $n$ points $\lambda_{1}, \lambda_{2}, \ldots, \lambda_{n}$ in the complex unit disc $D$ and target data consisting of $n$ points $z_{1}, z_{2}, \ldots, z_{n}$ in $D$, the Nevanlinna -Pick interpolation problem seeks to find a holomorphic function $\varphi$ that interpolates the datarthat is for all ${ }^{i}$,

$$
\varphi(i)=z_{i}
$$

Subject to the constrain

$$
|\varphi(\lambda)| \leq 1
$$

For all $\lambda \in D$.
Definition 2.0.1: The numerical range of $A \in M_{n}$, is the subset $W(A) \subset \mathbb{C}$, given by

$$
W(A)=\left\{\langle A x, x): x \in \mathbb{C}^{n} ;\|x\|=1\right\}
$$

where $\|\cdot\|$ denotes the 2 -norm. Note that $W(A)$ is the continuous image of a compact set, and is thus itself a compact set in $\mathbb{C}$. As we will show, the numerical range of a linear operator is a convex set. This is a consequence of the Toeplitz-Hausdorff Theorem. We first review some basic properties of the numerical range.

## 2. Main results

In this section, we give the results of our study. We begin with the following proposition.
Proposition 2.1.1: Let $A \in M_{n}$, then the following properties hold.

1. For any $\alpha, \beta \in \mathbb{C}_{,}$we have that $W(\alpha A+\beta I)=\alpha W(A)+\beta$.
2. $W\left(U^{*} A U\right)=W(A)$ for any unitary $U \in M_{n}$.
3. If $k \in\{1, \ldots, n-1\}$ and $X \in \mathbb{C}^{n \times k}$ satisfies $X^{*} X=I_{k}$, where $I_{k}$ denotes the $k \times k$ identity matrix, then $W\left(X^{*} A X\right) \subset W(A)$.

Proof: To show (i), we calculate

$$
\begin{aligned}
W(\alpha A+ & \beta I)=\left\{x^{*}(\alpha A+\beta I) x: x \in \mathbb{C}^{n},\|x\|=1\right\} \\
& =\left\{\alpha x^{*} A x+\beta x^{*} x: x \in \mathbb{C}^{n},\|x\|=1\right\} \\
& =\left\{\alpha x^{*} A x+\beta: x \in \mathbb{C}^{n},\|x\|=1\right\} \\
& =\alpha\left\{x^{*} A x: x \in \mathbb{C}^{n},\|x\|=1\right\}+\beta \\
& =\alpha W(A)+\beta
\end{aligned}
$$

To show (ii), let $\lambda \in W\left(U^{*} A U\right)$. Then there exists a unit vector $x \in \mathbb{C}^{n}$ such that $\left\langle<U^{*} A U x, x\right\rangle=\lambda$. Since U is self-adjoint, we can write $\langle A U x, U x\rangle=\lambda$. Now let $y=U x$. since multiplying by a unitary matrix preserves norm, that is,

$$
\|y\|^{2}=y^{*} y=x^{*} U^{*} U x=x^{*} x=\|x\|^{2}
$$

we have that $\mathrm{y} \in \mathbb{C}^{n}$ is also a unit vector. Thus $\langle A y, y\rangle=\lambda_{\text {so }} \lambda \in W(A)$.
To show the reverse inclusion, note that $W(A)=W\left(U U^{*} A U U^{*}\right) \subset W\left(U^{*} A U\right)$ by what was just shown. Thus

$$
W(A)=W\left(U^{*} A U\right)
$$

For (iii), let $\lambda \in W\left(X^{*} A X\right)$. Then there exists a unit vector $\mathrm{y} \in \mathbb{C}^{n}$ such that $\left\langle X^{*} A X y, y\right)=\lambda$. Note that

$$
\begin{aligned}
& \|X y\|^{2}=y^{*} X^{*} X y=y^{*} y=1 \text {. Thus setting } v=X y \text { yields } \\
& v^{*} A v=\langle A v, v\rangle=\lambda . \text { Hence } W\left(X^{*} A X\right) \subset W(A) .
\end{aligned}
$$

Proposition 2.1.2: Let $A, B \in M_{n}$. Then

1. $W\left(A^{*}\right)=\{\bar{\lambda}: \lambda \in W(A)\}=\overline{W(A)}$.
2. (Subadditivity) $W(B+A) \subset W(B)+W(A)$.

Proof: For (i), we have the following:

$$
\begin{aligned}
& W\left(A^{*}\right)=\left\{\left\langle A^{*} x, x\right\rangle:\|x\|=1\right\} \\
& =\{\langle x, A x\rangle:\|x\|=1\} \\
& =\{\overline{\langle A x, x)}:\|x\|=1\} \\
& =\overline{W(A)}
\end{aligned}
$$

For (ii), let $\xi \in W(B+A)$ with unit vector $x$ satisfying $x^{*}(B+A) x=\xi \cdot$ Let $\xi_{B}=x^{*} B x \quad$ and $\xi_{A}=x^{*} A x$. Then $\xi_{B} \in W(B)$ and $\xi_{A} \in W(A)$ and

$$
\begin{aligned}
& \xi=x^{*}(B+A) x \\
& =x^{*} B x+x^{*} A x \\
& =\xi_{B}+\xi_{A} .
\end{aligned}
$$

So $\xi$ is the sum of an element in $W(B)$ and an element in $W(A)$, thus

$$
\xi \in W(B)+W(A)
$$

The next two results will be very useful in showing how to sketch the numerical range as well as proving its convexity, but first we need the following lemma:

Lemma 2.1.3: Let $A \in M_{n}$. If $\langle A x, x\rangle=0$ for all $x \in \mathbb{C}^{n}$, then $A=0$.
Proof: First suppose that ${ }^{A}$ is Hermitian. Then for any $x, y \in \mathbb{C}^{n}$, we have that

$$
\langle x, A y\rangle=\langle A x, y\rangle .
$$

By the hypothesis, we also have that

$$
\langle A(x+y), x+y\rangle=0 .
$$

Therefore,

$$
\begin{aligned}
& 0=\langle A x, x\rangle+\langle A y, x\rangle+\langle A x, y\rangle+\langle A y, y\rangle \\
& =\langle A x, y\rangle+\langle y, A x\rangle .
\end{aligned}
$$

Letting $y=A x$ then yields ${ }^{0}=2\|A x\|^{2}$, which implies $A x=0$ for all $x \in \mathbb{C}^{n}$. Thus

$$
A=0
$$

Now let $A \in M_{n}$ be arbitrary. If we let

$$
H=\frac{A+A^{*}}{2} \quad \text { and } \quad K=\frac{A-A^{*}}{2 i}
$$

then $A=H+i K$ with $H^{H}$ and ${ }^{K}$ both Hermitian (we call ${ }^{H}$ the Hermitian part of ${ }^{A}$ and ${ }^{i K}$ the skew-Hermitian part of ${ }^{A}$ ). Thus, if $\langle A x, x\rangle=x^{*} A x=0$ for all $x \in \mathbb{C}^{n}$, then since $\left(x^{*} A x\right)^{*}=x^{*} A^{*} x=0$ also, we have

$$
\begin{aligned}
& x^{*} A x=x^{*} H x+i x^{*} K x=0, \text { and } \\
& x^{*} A^{*} x=x^{*} H x-i x^{*} K x=0,
\end{aligned}
$$

for all $x \in \mathbb{C}^{n}$. Adding these two equations gives us that $2 x^{*} H x=0$ for all $x \in \mathbb{C}^{n}$ and so by the result above, $\mathrm{H}=0$.

$$
\text { Similarly, } K=0 \text { and so } A=0
$$

Proposition 2.1.4: The numerical range of a Hermitian matrix $A \in \mathbb{C}^{n}$ is an interval $\left[\lambda_{1}, \lambda_{n}\right] \subset \mathbb{R}$ where $\lambda_{1}$ is the smallest eigenvalue of $A$ and $\lambda_{n}$ is the largest eigenvalue of $A$. Moreover, the set $L_{A}\left(\lambda_{n}\right)=\left\{x \in \mathbb{C}^{n}:\|x\|=1, x^{*} A x=\lambda_{n}\right\}$ is the set of all unit eigenvectors of $A$ corresponding to $\lambda_{n}$ and similarly for $\lambda_{1}$. We also have that if $W(A) \subset \mathbb{R}$, then $A$ is Hermitian.

Proof: Let $A \in M_{n}$ be Hermitian. Then there exists a set of $n$ orthonormal eigenvectors of $A$, denoted $\left\{x_{1}, \ldots, x_{n}\right\}$, with corresponding eigenvalues $\lambda_{1}, \ldots, \lambda_{n}$, which are arranged so that $\lambda_{1} \leq \lambda_{2} \leq \ldots \leq \lambda_{n}$. Let $x=c_{1} x_{1}+\ldots+c_{n} x_{n}$ be a unit vector in $\mathbb{C}^{n}$ (so $\left|c_{1}\right|^{2}+\ldots+\left|c_{n}\right|^{2}=1$ ). Using the fact that eigenvalues of $A$ are all real, we have that

$$
\begin{aligned}
& x^{*} A x=\lambda_{1}\left|c_{1}\right|^{2}+\lambda_{2}\left|c_{2}\right|^{2}+\ldots+\lambda_{n}\left|c_{n}\right|^{2} \\
& \leq \lambda_{n}\left(\left|c_{1}\right|^{2}+\left|c_{2}\right|^{2}+\ldots+\left|c_{n}\right|^{2}\right)=\lambda_{n}
\end{aligned}
$$

The first line of this equation implies that $x^{*} A x \in \mathbb{R}$ for all unit vectors $x \in \mathbb{C}^{n}$ and hence $W(A) \subset \mathbb{R}$. Furthermore, it also shows that for all $z \in W(A), z \leq \lambda_{n}$. Similarly, we can show that $x^{*} A x \geq \lambda_{1}$ for all unit vectors $x \in \mathbb{C}^{n}$, and thus $W(A) \subset\left[\lambda_{1}, \lambda_{n}\right]$. We still need to show that $W(A)=\left[\lambda_{1}, \lambda_{n}\right]$, i.e., for all $c \in\left[\lambda_{1}, \lambda_{n}\right]$, there exists a unit vector $x \in \mathbb{C}^{n}$ such that $x^{*} A x=c$. To do this let $x_{s}=\sqrt{s x_{1}}+\sqrt{1-s x_{n}}$ for $0 \leq{ }_{\mathrm{s}} \leq 1$. Then $\left\|x_{s}\right\|=1$ and $x_{s}^{*} A x_{s}=s \lambda_{1}+(1-s) \lambda_{n}$. So given any $c \in\left[\lambda_{1}, \lambda_{n}\right]$, we can find a unit vector $x_{s}$ such that $x_{s}^{*} A x_{s}=c$ by choosing an appropriate ${ }^{s}$. Thus $W(A)=\left[\lambda_{1}, \lambda_{n}\right]$.
For the second assertion, we claim that $x^{*} A x=\lambda_{n}$ if and only if $x$ is a unit eigenvector of $A$ corresponding to $\lambda_{n}$. The reverse implication is clear. For the forward direction, we prove the contrapositive. Suppose ${ }^{x}$ is not an eigenvector of ${ }^{A}$ corresponding to $\lambda_{n}$. Then ${ }^{x}$ cannot be a linear combination of eigenvectors corresponding to $\lambda_{n}$ either (for such a vector is, in fact, an eigenvector corresponding to $\lambda_{n}$ ), so in the representation $x=c_{1} x_{1}+\ldots+c_{n} x_{n}$,
we must have that $c_{j} \neq 0$ for some $j$ where $\lambda_{j} \neq \lambda_{n}$. Since $\lambda_{n}$ is the maximum eigenvalue, this means that $\lambda_{j}<\lambda_{n}$ so in this case, the inequality above is strict, i.e., $x^{*} A x<\lambda_{n}$. This proves the second assertion for $\lambda_{n}$. The proof for $\lambda_{1}$ is similar.
Finally, to show the last statement, let $A \in M_{n}$, not necessarily Hermitian, and suppose $x^{*} A x \in \mathbb{R}$ for all unit vectors $x \in \mathbb{C}^{n}$. Then $x^{*} A x=\overline{x^{*} A x}=x^{*} A^{*} x$ for all unit vectors $x \in \mathbb{C}^{n}$. Rearranging the terms, we get $x^{*} A x-x^{*} A^{*} x=0 \Rightarrow x^{*}\left(A-A^{*}\right) x=0$ for all unit vectors $x \in \mathbb{C}^{n}$. By Lemma 1.1.3, $A-A^{*}=0$ and hence $A=A^{*}$.

Proposition 2.1.5: For all $A \in M_{n}$, let $H(A)=\left(A+A^{*}\right) / 2$ and $K(A)=\left(A-A^{*}\right) / 2 i$ denote the Hermitian and skewHermitian parts of ${ }^{A}$, respectively. Then

$$
\operatorname{Re}(W(A))=W(H(A)) \quad \text { and } \quad \operatorname{Im}(W(A))=W(K(A)),
$$

Where,

$$
\operatorname{Re}(W(A))=\{\operatorname{Re} z: z \in W(A)\} \text { and } \operatorname{Im}(W(A))=\{\operatorname{Im} z: z \in W(A)\} .
$$

Proof: For all unit vectors $x \in \mathbb{C}^{n}$, we have that

$$
\begin{aligned}
& x^{*} H(A) x=x^{*} \frac{1}{2}\left(A+A^{*}\right) x \\
& =\frac{1}{2}\left(x^{*} A x+x^{*} A^{*} x\right) \\
& =\frac{1}{2}\left(x^{*} A x+\left(x^{*} A x\right)^{*}\right) \\
& =\frac{1}{2}\left(x^{*} A x+\overline{x^{*} A x}\right) \\
& =\operatorname{Re}\left(x^{*} A x\right) .
\end{aligned}
$$

So, every point of $W(H(A))$ is of the form $R e z$ for some $z \in W(A)$ and conversely.
Similarly, for $K(A)$, we have that for all unit vectors $x \in \mathbb{C}^{n}$,

$$
\begin{aligned}
& x^{*} K(A) x=x^{*} \frac{1}{2 i}\left(A-A^{*}\right) x \\
& =\frac{1}{2 i}\left(x^{*} A x-x^{*} A^{*} x\right) \\
& =\frac{1}{2 i}\left(x^{*} A x-\left(x^{*} A x\right)^{*}\right) \\
& =\frac{1}{2 i}\left(x^{*} A x-\overline{x^{*} A x}\right) \\
& =\frac{1}{2 i}\left(2 i \operatorname{Im}\left(x^{*} A x\right)\right) \\
& =\operatorname{Im}\left(x^{*} A x\right)
\end{aligned}
$$

At this point we consider convexity. One of the most significant properties of the numerical range is the fact that for any $A \in M_{n}$, the numerical range of $A$ is convex. This fact was proved by Toeplitz and Hausdorff. Toeplitz showed that the boundary of the numerical range is a convex curve and later, Hausdorff showed that the numerical range is itself convex. Thus this theorem has been named the Toeplitz-Hausdorff Theorem. There are various different proofs of this theorem. We present one of the more common ones below.

Theorem 2.2.1: (Toeplitz-Hausdorff). Let $A \in M_{n}$. Then $W(A) \subset \mathbb{C}$ is convex.
For the proof, we need the preliminary result stating that for a $2 \times 2$ matrix, the numerical range is an elliptical disk whose foci are the eigenvalues of the matrix. There are several different ways of proving this fact. We present here one of the proofs. We will need the following lemma:

Lemma 2.2.2: Given any $A \in M_{2}$, there exists a unitary matrix $U \in M_{2}$ such that the two main diagonal entries of $U^{*} A U$ are equal.

Proof: Without loss of generality, we can suppose that $\operatorname{tr} A=0$. To see why this is so, simply replace $A$
 $U^{*}(A-\alpha I) U$ are equal. Then if the $(1,1)$ entry and the $(2,2)$ entry of $U^{*} A U$ are $a_{11}^{\prime}$ and $a_{22}^{\prime}$, respectively, we would have that $a_{11}^{\prime}-\alpha=a_{22}^{\prime}-\alpha$ and so $a_{11}^{\prime}=a_{22}^{\prime}$. Thus, we can suppose that $\operatorname{tr} A=0$, and our task is reduced to finding a unitary matrix $U \in M_{2}$ such that the two main diagonal entries of $U * A U$ are zero.
In order to do this, it suffices to show that there exists a nonzero $w \in \mathbb{C}^{n}$ such that $w^{*} A w=0$. This is because if we normalize $w$ and set it as the first column of a unitary matrix ${ }^{W}$, we will have

$$
W^{*} A W=\left(\begin{array}{ll}
0 & x \\
x & \times
\end{array}\right)
$$

and since $\operatorname{tr}\left(W^{*} A W\right)=\operatorname{tr} A=0$, it must follow that the $(2,2)$ entry is zero also. To construct the vector $w$, first note that since $\operatorname{tr} A=0$, it is easily verified that the eigenvalues of $A$ are $\pm \lambda$, for some $\lambda \in \mathbb{C}$. Let $x$ and $y$ be the normalized eigenvectors for ${ }^{-\lambda}$ and $\lambda$, respectively. If $\lambda=0$, note that we can simply take $w=x$. Otherwise, let $w=e^{i \theta} x+y$. Since ${ }^{x}$ and $y$ are independent, ${ }^{w}$ is nonzero for all $\theta \in \mathbb{R}$, and

$$
\begin{aligned}
w^{*} A w & =\left(e^{-i \theta} x^{*}+y^{*}\right) A\left(e^{i \theta} x+y\right) \\
& =\left(e^{-i \theta} x^{*}+y^{*}\right)\left(-e^{i \theta} \lambda x+\lambda y\right) \\
& =\lambda\left(e^{-i \theta} x^{*} y-e^{i \theta} y^{*} x\right) \\
& =2 i \lambda I m\left(e^{-i \theta} x^{*} y\right) .
\end{aligned}
$$

The result then follows by picking $\theta$ so that $e^{-i \theta} x^{*} y$ is real.
Continuing along the same vein, let $A \in M_{2}$ and set $\alpha=(-1 / 2) \operatorname{tr} A$. By Proposition 2.1.1, it suffices to consider $W(A+\alpha I)$. Further, $\operatorname{tr}(A+\alpha I)=0$ and by the preceding lemma, we can suppose that the two main diagonal entries are both zero.

Now, if $c=|c| e^{i \theta_{1}}$ and $d=|d| e^{i \theta 2}$, let $\theta=(1 / 2)\left(\theta_{1}-\theta_{2}\right)$. Then the above matrix product equals

$$
e^{i \phi}\left(\begin{array}{cc}
0 & |c| \\
|d| & 0
\end{array}\right), \quad \phi=\frac{1}{2}\left(\theta_{1}+\theta_{2}\right)
$$

So, by unitary and scalar invariance we are done.
Lemma 2.2.3: Let $A \in M_{2}$. Then $W(A)$ is an elliptical disk whose foci are the eigenvalues of $A$.
Proof: By the above results, we can assume that $A$ is of the form (2.2.1). Without loss of generality, suppose $a \geq b \geq 0$. Let $z \in \mathbb{C}^{n}$ be an arbitrary unit vector. The goal is to show that all numbers of the form $z^{*} A z$ form an elliptical disk with the desired properties. Note that $\left(e^{i \theta} z\right)^{*} A\left(e^{i \theta} z\right)=z^{*} A z$ for all $\theta \in \mathbb{R}$ and so given any unit vector $z \in \mathbb{C}^{n}$, we can suppose that the first component of $z$ is real and nonnegative. Since ${ }^{z}$ is a unit vector, this means that ${ }^{z}$, with the first component real and nonnegative, has the form $z=\left(t, e^{i \theta}\left(1-t^{2}\right)^{1 / 2}\right)^{T}$, where $t \in[0,1]$ and $\theta \in[0,2 \pi]$. Therefore,

$$
\begin{aligned}
& z^{*} A z=\left(t, e^{-i \theta} \sqrt{1-t^{2}}\right)\left(\begin{array}{ll}
0 & a \\
b & 0
\end{array}\right)\binom{t}{e^{i \theta} \sqrt{1-t^{2}}} \\
& =\left(t, e^{-i \theta} \sqrt{1-t^{2}}\right)\binom{a e^{i \theta} \sqrt{1-t^{2}}}{b t} \\
& =t a e^{i \theta} \sqrt{1-t^{2}}+t b e^{-i \theta} \sqrt{1-t^{2}} \\
& =t \sqrt{1-t^{2}}((a+b) \cos (\theta)+i(a-b) \sin (\theta)) .
\end{aligned}
$$

Letting $\theta$ vary from ${ }^{0}$ to $2 \pi$, the point $(a+b) \cos (\theta)+i(a-b) \sin (\theta)$ traces out an ellipse $E^{E}$ with center $(0,0)$. (Note that the ellipse could be degenerate, as would be the case if $A$ were Hermitian.) As ${ }^{t}$ varies from ${ }^{0}$ to ${ }^{1}$, the term $t \sqrt{1-t^{2}}$ varies from 0 to $1 / 2$ and back to 0 . This shows that every point in the interior of $(1 / 2) E$ is attained for some z . Lastly, by considering the angles $\theta=0$ and ${ }^{\theta}=\pi / 2$, we see that the major axis of the ellipse extends from $-(a+b) / 2$ to $(a+$ b) $/ 2$ along the real axis and the minor axis extends from $i(b-a) / 2$ to $i(a-b) / 2$ along the imaginary axis. So, the distance from the center to the foci is

$$
\left(\frac{1}{4}(a+b)^{2}-\frac{1}{4}(a-b)^{2}\right)^{1 / 2}=\sqrt{a b}
$$

which means the foci are given by $\pm \sqrt{ } a b$, which are precisely the eigenvalues of $A$.
Proof of Theorem 2.2.1: Since we know the result is true in the $2 \times 2$ case, suppose $A \in M_{n}$, where $n>2$. Let $\gamma, \mu \in W(A)$. We need to show that the line segment connecting $\gamma$ and ${ }^{\mu}$ denoted $[\gamma, \mu]$, is contained in $W(A)$. Let $x, y \in \mathbb{C}^{n}$ be unit vectors such that $\gamma=x^{*} A x$ and ${ }^{\mu}=y^{*} A y$. Let $X \in \mathbb{C}^{n \times 2}$ be such that the column space of $X$ contains ${ }^{x}$ and $y$ and $X^{*} X=I_{2}$. Then there exists unit vectors $v, w \in \mathbb{C}^{2}$ such that $X v=x$ and $X w=y$. Thus $\gamma=x^{*} A x=v^{*} X^{*} A X v$ and $\mu=y^{*} A y=w^{*} X A X w$ which means that $\gamma, \mu \in W\left(X^{*} A X\right)$.
Note that $X^{*} A X$ is a 2 by 2 matrix, thus by Lemma 2.2.3, $W\left(X^{*} A X\right)$ is an ellipse. Since an ellipse is convex, we have that $\left[{ }^{\gamma}, \mu\right]$ $\subset W\left(X^{*} A X\right)$. But $W\left(X^{*} A X\right) \subset W(A)$ by Proposition 2.1.1. Thus $W(A)$ contains $[\gamma, \mu]$ which shows that $W(A)$ is convex.
The next proposition shows one of the applications of the Toeplitz-Hausdorff theorem.
Proposition 2.2.4: For any matrix $A \in M_{n}, W(A)$ contains the convex hull of the eigenvalues of $A$, denoted $\operatorname{co}(\sigma(A))$. Moreover, if $A$ is normal, then $W(A)=c o(\sigma(A))$.

Proof: Assume $A x=\lambda x$ with $\lambda_{\in} \sigma(A)$ and $\|x\|=1$. Thus $x^{*} A x=\lambda x^{*} x=\lambda$. So $\sigma(A) \subset W(A)$. The fact that the convex hull of $\sigma(A)$ is contained in $W(A)$ then follows from Theorem 2.2.1.
To show the second assertion, suppose ${ }^{A}$ is normal. Then ${ }^{A}$ is unitarily diagonalizable, i.e., $A=U^{*} D U$ for some unitary matrix $U_{\text {and }} D=\operatorname{diag}\left(\lambda_{1}, \ldots, \lambda_{n}\right)$ where $(\lambda)_{i=1}^{n}$ are the eigenvalues of $A$. By the unitary invariance property shown in Proposition 2.1.1, we have that $W(A)=W(D)$. Now let $x \in \mathbb{C}^{n}$ be a unit vector. Then

$$
x^{*} D x=\sum_{i-1}^{n} \lambda_{i}\left|x_{i}\right|^{2}
$$

Since ${ }^{x}$ is a unit vector, $\sum_{i-1}^{n}\left|x_{i}\right|^{2}=1$. So, we see that $W(D)$ is the set of all convex combinations of the eigenvalues of $A$. Thus $W(A)=W(D)=c o(\sigma(A))$.

Note that by Proposition 2.1.3 and Proposition 2.2.4 we have for any two $A, S \in M_{n}$, that $\sigma(A+S) \subset W(A+S) \subset W(A)+W(S)$. So while in general, $\sigma(A+S)$ is unrelated to $\sigma(A)$ and $\sigma(S)$, we can use the numerical range to say something about where the eigenvalues of $A+S$ are located in the complex plane.
One of the consequences of knowing the numerical range is convex is the advantage it provides in sketching it. Since we know it is convex, we only need to determine the boundary of $W(A)$ and then just shade in the interior. This idea provides a nice segue into the next section, which deals with the boundary of the numerical range.
For any matrix $A \in M_{n}, W(A)$ is a compact subset of $\mathbb{C}$. Thus, it is natural to want to know what can be said about the boundary of $W(A)$. We will denote the boundary of the numerical range by $\partial W(A)$. The following is a result dealing with the case where $W(A)$ has empty interior.

Proposition 2.3.1: Let $A \in M_{n}$. Then

1. $W(A)=\{\mu\}$ for some $\mu \in \mathbb{C}$ if and only if $A=\mu I$.
2. $W(A)$ has empty interior (meaning $W(A)=\partial W(A)$ is a line segment) if and only if $\frac{1}{n} \operatorname{tr} A \in \partial W(A)$.

Proof: (i) Suppose that $W(A)=\{\mu\}$ for some $\in \mathbb{C}$. Thus, for all unit vectors $x \in \mathbb{C}^{n}$ we have that $x^{*} A x=\mu$. Therefore,

$$
x^{*} A x=\mu x^{*} x \Rightarrow x^{*} A x-\mu x^{*} x=0 \Rightarrow x^{*}(A-\mu I) x=0 .
$$

Since this holds for all unit vectors $x \in \mathbb{C}^{n}$, by Lemma 2.1.4 we must have that $A-\mu I=0$ and so $A=\mu I$.
Conversely, if $A=\mu I$ for some $\mu \in \mathbb{C}$, then for all unit vectors $x \in \mathbb{C}^{n}$, we have

$$
x^{*} A x=x^{*} \mu I x=\mu x^{*} x=\mu
$$

Thus $W(A)=\{\mu\}$.
(ii) For the second assertion, suppose that $W(A)$ has empty interior. Then, since $W(A)$ is convex, $W(A)=\partial W(A)$ is a line segment. So $W(A)=r(t)=c_{1} t+(1-t) c_{2}$ for some $c_{1}, c_{2} \in \mathbb{C}, t \in[0,1]$. By Proposition 2.1.1, we can without loss of generality assume that $c_{1}=1$ and $c_{2}=0$ so $W(A)=[0,1]$. By Proposition 2.1.5, we can conclude that $A$ is Hermitian. Then $A$ is unitarily similar to a diagonal matrix $D=\operatorname{diag}\left(\lambda_{1}, \ldots, \lambda_{n}\right)$ where $\lambda_{i}$,
$i=1, \ldots, n$ are the eigenvalues of $A$. So again by Proposition 2.1.1, we have that $W(A)=W(D)$.
Let $x \in \mathbb{C}^{n}$ and $\mathrm{x}=\left(\frac{1}{\sqrt{n}} \ldots, \frac{1}{\sqrt{n}}\right)$.
Then $\|x\|=1$ and $x^{*} D x=\sum_{i=0}^{n} \frac{1}{n} \lambda_{i}=\frac{1}{n} \operatorname{tr} D=\frac{1}{n} \operatorname{tr} A$
Therefore,

$$
\frac{1}{n} \operatorname{tr} A \in W(A)=\partial W(A)
$$

For the reverse direction, let $\zeta=\frac{1}{n} \operatorname{tr} A$ and suppose $\frac{1}{n} \operatorname{tr} A \in \partial W(A)$. Let $B=A-\zeta I$. Then a direct calculation shows that $\frac{1}{n} \operatorname{tr} B=0$ and so $0 \in W(B)$. By the hypothesis, and Proposition 2.1.1 we can further say that $0 \in \partial W(B)$. Since $W(B)$ is convex, there exists $\theta \in \mathbb{R}$ such that $W\left(e^{i \theta} B\right)$ is contained in the closed upper half plane,
$U H P=\{z \in \mathbb{C}: \operatorname{Im} z \geq 0\}$.
Let $C=e^{i \theta} B$. At this point, we have the following:

$$
W(C) \subset U H P, \quad 0 \in \partial W(C), \quad \text { and } \quad \frac{1}{n} \operatorname{tr} C=0
$$

Note that $c_{i i} \in W(C)$ and $\operatorname{Im} c_{i i} \geq 0$ for $1 \leq i \leq n$. But we must have that $\frac{1}{n} \Sigma \operatorname{Im} c_{i i}=0$ and so each $c_{i i} \in \mathbb{R}$. Let $i, j \in\{1, \ldots, n\}$ be arbitrary indices and let $\Gamma_{i j}$ denote the $2 \times 2$ principal submatrix lying in the rows and columns of $C$ indexed by $i, j$. Let $\lambda_{1}, \lambda_{2}$ denote the eigenvalues of $\Gamma_{i j}$. By Proposition 2.1.1, $W\left(\Gamma_{i j}\right) \subset W(C) \subset U H P$.Therefore, $\lambda_{1}, \lambda_{2} \in$ UHP. But $\operatorname{Im}\left(\lambda_{1}\right)+\operatorname{Im}\left(\lambda_{2}\right)=\operatorname{Im}\left(\lambda_{1}+\lambda_{2}\right)=\operatorname{Im}\left(\operatorname{tr} \Gamma_{i j}\right)=0$. Therefore, $\lambda_{1}, \lambda_{2}$ are real. Since $W\left(\Gamma_{i j}\right)_{\text {is }}$ an ellipse with foci $\lambda_{1}, \lambda_{2}$ and has the real axis as its major axis, we must have that $W\left(\Gamma_{i j}\right) \subset \mathbb{R}$. By Proposition 2.1.5, we see that $\Gamma_{i j}$ is Hermitian. Since ${ }^{i, j}$ were arbitrary, it follows that $C$ is also Hermitian and so $W(C)$ is an interval in the real line and thus has empty interior. By the invariance properties of the numerical range, this implies that $W(A)$ has empty interior.
We now turn our attention to the corners of the numerical range. In order to discuss about the corners, we need the following definition.A point $\mu \in W(A)$ is called a corner of $W(A)$ if there exist angles $\theta_{1}$ and $\theta_{2}$ satisfying $0 \leq \theta_{1}<\theta_{2}<2 \pi$ such that Re $e^{i \theta} \mu=\max \left\{\operatorname{Re} \beta: \beta \in W\left(e^{i \theta} A\right)\right\} \quad$ for all $\theta \in\left(\theta_{1}, \theta_{2}\right)$. We have the following result concerning the corners of the numerical range.

Proposition 2.3.3: For any $A \in M_{n,}$ if $\mu$ is a corner of $W(A)$, then $\mu$ is an eigenvalue of $A$.
Proof: Let ${ }^{\mu}$ be a corner of $W(A)$, then there exists $\theta_{1}$ and $\theta_{2}$ as in the definition such that

$$
\operatorname{Re} e^{i \theta} \mu=\max \left\{\operatorname{Re} \beta: \beta \in W\left(e^{i \theta} A\right)\right\} \quad \text { for all } \theta \in\left(\theta_{1}, \theta_{2}\right) .
$$

By Proposition 2.1.5 and Proposition 2.1.6, we have that this quantity is the same as the largest eigenvalue of the Hermitian part of $e^{i \theta} A, \theta \in\left(\theta_{1}, \theta_{2}\right)$. Thus, for all such $\theta$ there exists a unit vector $x_{\theta}$, such that $x_{\theta} * A x_{\theta}=\mu$ and

$$
x_{\theta} * H\left(e^{i \theta} A\right) x_{\theta}=\lambda_{n}\left(H\left(e^{i \theta} A\right)\right)=\operatorname{Re} e^{i \theta} \mu .
$$

In fact, we have that the same vector $x_{\theta}$ works for all $\theta^{\theta}$ in the interval $\left(\theta_{1}, \theta_{2}\right)$. To see why this is true, suppose $\theta^{\prime} \in\left(\theta_{1}, \theta_{2}\right)$ is different from ${ }^{\theta}$. Note that

$$
\begin{aligned}
& x_{\theta}^{*} H\left(e^{i \theta} A\right) x_{\theta}=\frac{x_{\theta}^{*} e^{i \theta^{\theta}} A x_{\theta}+x_{\theta}^{*} e^{-i \theta^{\theta}} A x_{\theta}}{2} \\
& =\frac{1}{2}\left(e^{i \theta^{\prime}} x_{\theta}^{*} A x_{\theta}+\overline{e^{i \theta^{\prime}} x_{\theta}^{*} A x_{\theta}}\right) \\
& =\operatorname{Re} x_{\theta}^{*} e^{i \theta^{\prime}} A x_{\theta} \\
& =\operatorname{Re} e^{i \theta^{\prime}} \mu \\
& =\lambda_{n}\left(H\left(e^{i \theta^{\prime}} A\right)\right)
\end{aligned}
$$

where the last equality follows from the fact that $\theta^{\prime} \in\left(\theta_{1}, \theta_{2}\right)$. From this, we can conclude by Proposition 2.1.5 that $H\left(e^{i \theta^{y}} A\right)^{x_{\theta}}=\lambda_{n}\left(H\left(e^{i \theta^{y}} A\right)\right)^{x_{\theta}}=e^{i \theta^{y}} \mu x_{\theta}$. So, from now on, we will simply write ${ }^{x}$ instead of $x^{x_{\theta}}$.
Now let $\lambda_{\theta}=\operatorname{Re} e^{i \theta} \mu$. Since ${ }^{x}$ is independent of $\theta$, we can take the derivative of the eigenvector equation $H\left(e^{i \theta} A\right) x=\lambda_{\theta} x$ with respect to ${ }^{\theta}$ to obtain

$$
H\left(i e^{i \theta} A\right)^{x}=\lambda_{\theta}^{\prime} x
$$

It is easily verified that this is the same as

$$
i K\left(e^{i \theta} A\right) x=-i \lambda_{\theta}^{\prime} x
$$

Now if we add this last equation to the eigenvector equation $H\left(e^{i \theta} A\right) x=\lambda_{\theta} x$, we get

$$
e^{i \theta} A x=\left(\lambda_{\theta-}-i \lambda_{\theta}^{\prime}\right)^{x} \quad \text { or } \quad A x=e^{-i \theta}\left(\lambda_{\theta}-i \lambda_{\theta}^{\prime}\right)^{x} .
$$

Thus $e^{-i \theta}\left(\lambda_{\theta}-i \lambda_{\theta}^{\prime}\right)$ is an eigenvalue of $A$. But

$$
x^{*} A x=e^{-i \theta}\left(\lambda_{\theta}-i \lambda_{\theta}^{\prime}\right)
$$

which must equal $\mu$ since these last two equations hold for all $\theta \in\left(\theta_{1}, \theta_{2}\right)$. Thus $\mu$ is an eigenvalue of $A$.

The numerical radius of an operator $A \in M_{n}$, is given by $w(A)=\sup \{|\lambda|: \lambda \in W(A)\}$.
Remark 2.4.2: The definition of numerical radius immediately implies that $w(A) \geq 0$, where equality holds if and only if $W(A)=\{0\}$ which, by Proposition 2.3.1, is true if and only if $A=0$. So, the numerical radius satisfies one of the requirements for a norm on $M_{n}$. Next, we also have that

$$
\begin{aligned}
w(z A) & =\sup \left\{|\langle z A x, x\rangle|: x \in \mathbb{C}^{n},\|x\|=1\right\} \\
& =|z|\left\{|\langle A x, x\rangle|: x \in \mathbb{C}^{n},\|x\|=1\right\} \\
& =|z| w(A),
\end{aligned}
$$

and Proposition 2.1.3 shows that the numerical radius also satisfies the triangle inequality. Thus the numerical radius is a norm on $M_{n \text {. By Proposition 2.2.5, we also have that }}$

$$
r(A) \leq w(A) \quad \text { for all } A \in M_{n} .
$$

Next, we introduce some basic results on the numerical radius. Since all norms on finite dimensional vector spaces are equivalent, we have that the numerical radius is equivalent to the matrix 2 -norm of $A$. This next result states this more precisely.


$$
w(A) \leq\|A\| \leq 2 w(A) .
$$

Proof: Let $\mu=\langle A x, x\rangle$ where $\|x\|=1$. Then by the Cauchy-Schwarz inequality, we have

$$
|\mu|=|\langle A x, x\rangle| \leq\|A x\| \leq\|A\| .
$$

Since this is true for all $\mu \in W(A)$, we have that $w(A) \leq\|A\|$.
For the other inequality, first note that for any nonzero $x \in \mathbb{C}^{n}$, we have that

$$
\begin{equation*}
\langle A x, x\rangle=\left(\frac{A x}{\|x\|}, \frac{x}{\|x\|}\right)\|x\|^{2} \leq w(A)\|x\|^{2} \tag{2.4.1}
\end{equation*}
$$

We will also make use of the following polarization identity:

$$
4\langle A x, y\rangle=\langle A(x+y), x+y\rangle-\langle A(x-y), x-y\rangle+i\langle A(x+i y), x+i y\rangle-i\langle A(x-i y), x-i y) .
$$

Now applying (2.4.1) to (2.4.2), we get that

$$
\begin{aligned}
& 4|\langle A x, y\rangle| \leq w(A)\left[\|x+y\|^{2}+\|x-y\|^{2}+\|x+i y\|^{2}+\|x-i y\|^{2}\right] \\
& =\left[\|x\|^{2}+\langle y, x\rangle+\langle x, y\rangle+\|y\|^{2}+\|x\|^{2}-\langle x, y\rangle-\langle y, x\rangle+\|y\|^{2}+\|x\|^{2}+i\langle y, x\rangle-i\langle x, y\rangle+\|y\|^{2}+\|x\|^{2}-\right. \\
& \left.i\langle y, x\rangle+i\langle x, y\rangle+\|y\|^{2}\right] \\
& =4 w(A)\left[\|x\|^{2}+\|y\|^{2}\right]
\end{aligned}
$$

Since ${ }^{x}$ and $y$ were arbitrary, we can pick $\|x\|=\|y\|=1$ so that

$$
|\langle A x, y\rangle| \leq 2 w(A)
$$

Now let $y=A x /\|A x\|$. Then

$$
\frac{\| A x, A x\rangle \mid}{\|A x\|} \leq 2 w(A)
$$

Hence $\|A x\| \leq 2 w(A)$. Taking the supremum over all $x \in \mathbb{C}^{n}$, with $\|x\|=1$ implies $\|A\| \leq 2 w(A)$.

The following result deals with one of the extreme cases of Theorem 2.4.3, namely when the numerical radius equals the norm of A.

Theorem 2.4.4: $\operatorname{\text {If}} w(A)=\|A\|$, then $r(A)=\|A\|$
Proof: Since ${ }^{w(A)}=\|A\|$, we can write
$\operatorname{Sup}\{|\langle A x, x\rangle|:\|x\|=1\}=\|A\|$.
Since $W(A)$ is compact, there exists a unit vector $x \in \mathbb{C}^{n}$ such that this supremum is attained, that is $\mid\langle A x, x\rangle\|=\| A \|$ for this particular ${ }^{x}$. But, by the Cauchy-Schwarz inequality,

$$
\|A\|=\mid\langle A x, x\rangle\|\leq\| A x\|\leq\| A \| .
$$

So, we must have equalities throughout which implies $A x=\lambda x$ for some $\lambda \in \mathbb{C}$. Thus $\lambda \in \sigma(A)$ and so $r(A) \geq|\lambda|=|\langle A x, x\rangle|=\|A\|$. But since in general, $r(A) \leq w(A) \leq\|A\|$, this implies that $r(A)=\|A\|$.

The next theorem deals with the other extreme case of Theorem 2.4.3, which is $w(A)=\frac{1}{2}\|A\|$. First, let $R(A)=\left\{A x: x \in \mathbb{C}^{n}\right\}$ denote the range of a matrix ${ }^{A}$ and $N(A)=\left\{x \in \mathbb{C}^{n}: A x=0\right\}$ denote the nullspace of ${ }^{A}$. Then we have the following:

Theorem 2.4.5: If $R(A) \perp R\left(A^{*}\right)$, then $w(A)=\frac{1}{2}\|A\|$.
Proof: Let ${ }^{x}$ be a unit vector in $\mathbb{C}^{n}$. We can write ${ }^{x}$ as $x_{1}+x_{2}$ where $x_{1} \in N(A)$ and $x_{2} \in R\left(A^{*}\right)$, since $R\left(A^{*}\right)=N(A)^{\perp}$ by the fundamental theorem of linear algebra. So $A x_{1}=0$ and since $R(A) \perp R\left(A^{*}\right)$, we also have that $\left\langle A x_{2}, x_{2}\right\rangle=0$. Therefore,

$$
\begin{aligned}
& \langle A x, x\rangle=\left\langle A\left(x_{1}+x_{2}\right), x_{1}+x_{2}\right\rangle \\
& =\left\langle A x_{1}, x_{1}\right\rangle+\left\langle A x_{1}, x_{2}\right\rangle+\left\langle A x_{2}, x_{1}\right\rangle+\left\langle A x_{2}, x_{2}\right\rangle \\
& =\left\langle A x_{2}, x_{1}\right\rangle
\end{aligned}
$$

This implies that

$$
|\langle A x, x\rangle| \leq\|A\|\left\|x_{2}\right\|\left\|x_{1}\right\| \leq \frac{\|A\|}{2}\left(\left\|x_{1}\right\|^{2}+\left\|x_{2}\right\|^{2}\right)=\frac{\|A\|}{2},
$$

where the second inequality follows from Young's inequality, or the fact that $(a-b)^{2} \geq 0$ for any $a, b \in \mathbb{R}$. Now since ${ }^{x}$ is arbitrary, we can take the supremum on the left-hand side, which yields

$$
w(A) \leq \frac{\|A\|}{2},
$$

and since $\frac{\|A\|}{2} \leq w(A)$ by Theorem 2.4.3, we get that $w(A)=\frac{1}{2}\|A\|$.
The last result we show here is the well-known power inequality.
Theorem 2.4.6: Let $A \in M_{n}$. Then for any positive integer $m$, we have that $w\left(A^{m}\right) \leq w(A)^{m}$.
Before proving this theorem, we prove the following lemma.
Lemma 2.4.7: If $w(A) \leq 1$ implies $w\left(A^{m}\right) \leq 1$ for all $m \in N$, then $w\left(A^{m}\right) \leq w(A)_{\mathrm{m}}$ for all $A \in M_{n}$.
Proof: Let $w(A)=c$, for some $c \geq 0$. If $c=0$, then $A=0$ and the result holds trivially, so suppose $c>0$.
Then $\frac{1}{c} w(A)=w\left(\frac{1}{c} A\right)$ by Remark 2.4.2. Letting $B=\frac{1}{c} A$, we have by hypothesis that $w(B) \leq 1$ and so ${ }^{w}\left(B^{m}\right) \leq 1$.
Therefore, $w\left(\frac{1}{c^{m}} A^{m}\right)=\frac{1}{c^{m}} w\left(A^{m}\right) \leq_{1}$
which implies $w\left(A^{m}\right) \leq c^{m}=w(A)_{\mathrm{m}}$.

Lemma 2.4.8: Let $A \in M_{n}$ and $z \in \mathbb{C}_{\text {with }}|z|<1$. Then the following are equivalent:

1. $w(A) \leq 1$.
2. $\operatorname{Re}\langle(I-z A) x, x\rangle \geq 0$ for all $x \in \mathbb{C}^{n}$.
3. $\operatorname{Re}\left\langle(I-z A)^{-1} y, y\right\rangle \geq 0$ for all $y \in \mathbb{C}^{n}$, provided $I-z A$ is invertible.

Proof: (i) $\Leftrightarrow$ (ii) Suppose that $w(A) \leq 1$. Recall that $|\langle A x, x\rangle| \leq w(A)\|x\|^{2}$ for all $x \in \mathbb{C}^{n}$ (see (2.4.1)) and $\operatorname{Re} z \leq|z|$ for all $z \in \mathbb{C}$. Now let ${ }^{z}$ be any number in $\mathbb{C}$ with $|z|<1$.

Then, $\operatorname{Re}\langle(I-z A) x, x\rangle=\|x\|^{2}-\operatorname{Re}\langle z A x, x\rangle \geq\|x\|^{2}-|z \||\langle A x, x\rangle|$

$$
\geq\|x\|^{2}-|z| w(A)\|x\|^{2} \geq\|x\|^{2}(1-|z|) \geq 0
$$

where the penultimate inequality follows from $w(A) \leq 1$. Conversely, we suppose that
$\operatorname{Re}\langle(I-z A) x, x\rangle \geq 0$ for all $|z|<1$. Simplifying, we have that $\|x\|^{2} \geq \operatorname{Re}\langle z A x, x\rangle$.
Writing $z=t e^{i \theta}$ and letting $\mathrm{t} \rightarrow 1$, we get

$$
\operatorname{Re}\left\langle e^{i \theta} A x, x\right\rangle \leq\|x\|^{2}
$$

This implies that $\mathrm{w}(\mathrm{A}) \leq 1$. To see why this is true, suppose that $\mu \in W(A)$. Then $\mu=r e^{-i \theta}$, where $r>0$ and $0 \leq \theta<2 \pi$. Also, $\mu=\langle A y, y\rangle$ for some unit vector $y \in \mathbb{C}^{n}$. So, we have that $|\mu|=r$ and $r e^{-i \theta}=\langle A y, y\rangle$. Thus $r=e^{i \theta}\langle A y, y\rangle \in \mathbb{R}$ so $r=\operatorname{Re} e^{i \theta}\langle A y, y\rangle \leq\|y\|^{2} \leq 1$, by the above equation. Since $r|\mu|$, and $\mu \in W(A)$ was arbitrary, we have that $w(A) \leq 1$.
(ii) $\Leftrightarrow$ (iii) Suppose $z \in \mathbb{C},|z|<1$ is such that $I-z A$ is invertible. Then

$$
x=(I-z A)^{-1} y
$$

for some $\mathrm{y} \in \mathbb{C}^{n}$. Therefore, by plugging in $(I-z A)^{-1} y$ for x , we get

$$
\begin{align*}
& \operatorname{Re}\langle(I-z A) x, x\rangle \geq 0 \quad \text { for all } x \in \mathbb{C}^{n} \\
& \Leftrightarrow \operatorname{Re}\left\langle\left(y,(I-z A)^{-1} y\right\rangle \geq 0 \text { for all } y \in \mathbb{C}^{n}\right. \\
& \Leftrightarrow \operatorname{Re}\left\langle(I-z A)^{-1} y, y\right) \geq 0 \quad \text { for all } \mathrm{y} \in \mathbb{C}^{n} \tag{2.4.3}
\end{align*}
$$

This completes the proof.
Proof of Theorem 2.4.6: Assume $w(A) \leq 1$. By Lemma 2.4.7, it suffices to show that this implies $w^{w}\left(A^{m}\right) \leq 1$ for all $m \in \mathbb{N}$. To do this, we use Lemma 2.4 .8 (iii). The invertibility of $I-z A$ for $|z|<1$ follows from the fact that $r(A) \leq w(A) \leq 1$. Furthermore, $r(A) \leq 1$ implies $r\left(A^{m}\right) \leq 1$ for all $m \in \mathbb{N}$, so by similar reasoning, $I-z^{m} A^{m}$ is invertible.
So, to prove the theorem, it is sufficient to show that for all $x \in \mathbb{C}^{n}$

$$
\operatorname{Re}\left\langle\left(I-z^{m} A^{m}\right)^{-1} x, x\right\rangle \geq_{0} \quad \text { where } z \in \mathbb{C},|z|<1
$$

since this condition will imply that $\mathrm{w}\left(\mathrm{A}^{\mathrm{m}}\right) \leq 1$. To do this, we use the following identity:

$$
\begin{equation*}
\left(\mathrm{I}-z^{m} A^{m}\right)^{-1}=\frac{1}{m} \sum_{k=0}^{m-1}\left(I-\omega^{k} z A\right)^{-1} \tag{2.4.4}
\end{equation*}
$$

where $\omega$ is a primitive $m^{\text {th }}$ root of unity. Note that for $0 \leq k \leq m-1,\left|\omega^{k}\right|=1$ so $\left|\omega^{k} z\right|<1$ and since $w(A) \leq 1$, we have, by Lemma 2.4.8, that

$$
\operatorname{Re}\left\langle\left(I-\omega^{k} z A\right)^{-1} x, x\right) \geq 0 \quad \text { for all } x \in \mathbb{C}^{n},|z|<1, k=0, \ldots, m-1
$$

Therefore,

$$
\begin{aligned}
\operatorname{Re}\left(\left(I-z^{m} A^{m}\right)^{-1} x, x\right) & =\operatorname{Re}\left\langle\frac{1}{m} \sum_{k=0}^{m-1}\left(I-\omega^{k} z A\right)^{-1} x, x\right) \\
& =\frac{1}{m} \sum_{k=0}^{m-1} \operatorname{Re}\left(\left(I-\omega^{k} z A\right) x \quad, x\right\rangle \geq_{0} \quad \text { for all } \mathrm{x} \in \mathbb{C}^{n},|z|<1
\end{aligned}
$$

By Lemma 2.4.8, this implies that $w\left(A^{m}\right) \leq 1$. Thus $w\left(A^{m}\right) \leq w(A)^{m}$ by Lemma 2.4.7.

## 3. Conclusion

In this paper, for a Hilbert space $H$, we have determined the numerical range of $2 \times 2$ matrix with elliptical properties.

## 4. References

1. Bonsall FF, Duncan J. Numerical Ranges of Operator on Normed Spaces and of Elements of Normed Algebras, Cambridge university press, 1971. ISBN 978-0-521-07988-4
2. Bernard D, Francis D. Generalization of von Neumann's spectral sets and integral representation of operators, Bull. Soc. Math. France. 1999; 127(1):25-41.
3. Badea C, Crouzeix M, Delyon B. Convex Domains and k-spectral Sets. Math. Z. 2006; 252(2):345-365.
4. Cowen C, Harel E. An effective algorithm for computing the numerical range, 1995.
5. Keeler DS, Rodman L, Spitkovsky IM. The numerical range of $3 \times 3$ matrices, Linear algebra and its applications. 1997; 252:115-139.
6. Hwa LG, Wu PY. Higher Rank numerical Ranges and Kippenhahn Polynomials, linear algebra and its applications. 2013; 438:3054-3061.
7. Shapiro J. Notes on numerical range, 2003. http:/www.math.msu.edu/~shapiro/pubvit/downloads.numrangenotes/numrange-notes-pdf.
8. Akaashueek M, Ball JA, Gohberg I. Nevanlinna-Pick interpolation for time-varying input-output maps. The continuous time case, sc, Birkhausser, 1992.
9. Gustafson K, Duggirala K, Rao M. Numerical range, Universitext. Springer-Verlag. New York, 1997.
10. Luo SM. Crouzeix conjecture and the GMRES algorithm. All Thesis and Dissertations, 2011, article-2819.
11. Crouzeix M. Bounds for Analytic Functions of Matrices. Integral equations operator theory. 2004; 48(4):461-477.
12. Choi MD, Kribs DW, Zyczkowsi K. Higher Rank Numerical Ranges and Compression Problems. Linear algebra and its applications. 2006; 418:828-839.
13. Miguel M, Javier M, Rafael P. On the intrinsic and the spatial numerical range. Mathematical analysis and applications. 2006; 318:175-189.
14. Crouzeix M. Numerical Range and Functional Calculus in Hilbert Space. Journal of Functional Analysis. 2007; 244(2):668-690.
15. Fieldler M. Geometry of the Numerical Range of Matrices. Linear algebra and its applications. 1981; 37:81-96.
16. Crouzeix M, Delyon B. Some estimates for analytic functions of strip or sectorial operators. Arch. Math. (Basel). 2003; 81(5):559-566.
17. Halmos PR. A Hilbert space problem book. D. Van Nostrand Co., Inc., Princeton, N.J-Toronto, Ont.-London, 1967.
18. Panayiotis JP, Tsatsomeros MJ. Numerical range :( in) a matrix nutshell, May 2002.
19. Ando T. Structure of operators with numerical radius one. Acta Sci. Math. (Szeged). 1973; 34:11-15.
20. Tosio K. Some mapping theorems for the numerical range. Proc. Japan Acad. 1965; 41:652-655.
21. Young WH. On classes of summable functions and their Fourier series. Proceedings of the royal society. 1912; A87(594):225-229.
