Eötvös Loránd University
Faculty of Science

Kristóf Havasréti

Numerical Ranges of Operators

MSc Thesis

Advisor: Zsigmond Tarcsay
Department of Applied Analysis and Computational Mathematics

Budapest, 2021

Acknowledgement

I would like to express my gratitude to my advisor, Zsigmond Tarcsay for all his help.

Contents

Introduction

The theory of numerical ranges grew out of the study of quadratic forms on Hilbert spaces. The classical definition in finite-dimensional spaces was introduced by Otto Toeplitz in 1918. By this definition the *numerical range* (also called the *field of values*) of a bounded linear operator is the image of the associated quadratic form when restricted to the unit sphere. In the same year Felix Hausdorff showed that this is a convex set. W, the most common symbol for the numerical range comes from the German expression *Wertvorrat* used by Toeplitz and Hausdorff. The finite-dimensional definition easily extends into infinite dimensions.

A definition for numerical range in general normed spaces was created by Günter Lumer in 1961, and by Heinz Bauer in 1962. Lumer's approach was based on semi-inner products, while Bauer's definition relied on the dual space. These two independent approaches ended up being closely related. In the following years the theory of numerical ranges was expanded by mathematicians such as P. Halmos, I. Vidav, F.F. Bonsall, J. Duncan, A.M. Sinclair, and T.W. Palmer. This thesis will present mostly their work. Even though it will not be covered here at all, there are numerous applications of numerical ranges for example in numerical analysis, differential equations, operator semigroups, quantum physics, and quantum computing.

The aim of this thesis is to present and contrast the properties of the numerical range in three different settings: in Hilbert spaces, in Banach algebras (including C^*-algebras), and in general normed spaces. The first chapter is a brief overview of the definitions and theorems used later on. Chapter 2 is about the numerical range of Hilbert space operators. Chapter 3 covers a more abstract object, the algebraic numerical range. Chapter 4 will discuss the most general definition, the spatial numerical range, and it will tie together these three notions.

Notation

$\overline{z}$	complex conjugate of $z \in \mathbb{C}$
$\mathrm{Re}(z)$	real part of $z \in \mathbb{C}$
$\overline{A}$	closure of a set A
$\mathrm{conv}(A)$	convex hull of a set A
$\overline{\mathrm{conv}}(A)$	closed convex hull of a set A
$\mathrm{span}(x_1, x_2, ...)$	linear subspace generated by the vectors $x_1, x_2, ...$
X	normed space, Banach space
X'	topological dual
$\\|x\\|$	norm
$S(X)$	unit sphere
$\mathscr{A}$	normed algebra, Banach algebra
$\mathbb{1}$	unit (multiplicative identity)
$G(\mathscr{A})$	set of invertible elements
a^*	involution of a
$\mathrm{Sp}(a), \mathrm{Sp}(T)$	spectrum
$r(a), r(T)$	spectral radius
$\mathscr{S}(\mathscr{A}, a)$	support functionals of a
$\mathscr{S}(\mathscr{A})$	state space of $\mathscr{A}$
$\mathscr{B}(X)$	the space of bounded linear operators on a normed space
T	a bounded linear operator
$\\|T\\|$	operator norm
$\ker T$	kernel of an operator
$\mathrm{ran}\, T$	range/image of an operator
T^{-1}	inverse operator
T^*	adjoint operator
I	identity operator
U	unitary operator
Tr	matrix trace
$\mathcal{H}$	Hilbert space
$\langle ., . \rangle$	inner/scalar product
$[., .]$	L-semi-inner product

$W(T)$	numerical range
$w(T)$	numerical radius
$V(a)$	algebraic numerical range
$v(a)$	algebraic/spatial numerical radius
$\boldsymbol{V}(T)$	spatial numerical range
$n(X)$	numerical index

Chapter 1

Preliminaries

This chapter will give an overview of the spaces that will provide the setting for studying the numerical range. Thorough introduction and proofs can be found for example in [17] or [5]. In the definitions $\mathbb{F}$ means either the real or the complex number field.

Let X denote a *normed space* over $\mathbb{F}$ with the norm $\|.\| : X \to \mathbb{R}_+$. If X is complete with the given norm, it is called a *Banach space*. If X, Y are two normed spaces, then $\mathscr{B}(X, Y)$ is the set of all bounded (or equivalently, continuous) linear functions from X to Y. The elements of $\mathscr{B}(X, Y)$ are called (linear) operators. With the operator norm

$$\|T\| = \sup\{\|Tx\| : x \in X, \|x\| \leq 1\} \quad (T \in \mathscr{B}(X, Y))$$

$\mathscr{B}(X, Y)$ is a normed space as well. If Y is a Banach space, then $\mathscr{B}(X, Y)$ is a Banach space too. $\mathscr{B}(X, X)$ is simply written as $\mathscr{B}(X)$. $X' = \mathscr{B}(X, \mathbb{F})$ is the *topological dual* (or dual space) of X. The elements of X' are called the continuous linear functionals on X. The operator norm on X' is called the functional norm or the dual norm. A $T \in \mathscr{B}(X)$ operator is *invertible* if it is bijective and $A^{-1} \in \mathscr{B}(X)$. The *Banach adjoint* of an operator $T \in \mathscr{B}(X)$ is an operator $T^* \in \mathscr{B}(X')$ such that $T^*(f) = f \circ T$ for $f \in X'$. The Banach-adjoint is also called the adjoint or the transpose, but it is different from the Hilbert space adjoint.

An associative algebra $\mathscr{A}$ over $\mathbb{F}$ is a *normed algebra* with the norm $\|.\|$, if $\|xy\| \leq \|x\|\|y\|$ for every $x, y \in \mathscr{A}$. An algebra is *unital* if there is an element $\mathbb{1} \in \mathscr{A}$ such that $\mathbb{1} \cdot x = x \cdot \mathbb{1} = x$ for every $x \in \mathscr{A}$. We will suppose that for the unit $\|\mathbb{1}\| = 1$. The set of invertible elements in $\mathscr{A}$ is denoted by $G(\mathscr{A})$. A *Banach algebra* is a complete normed algebra. If X is a normed (or Banach) space, then $\mathscr{B}(X)$ is a normed (or Banach) algebra where the multiplication is the composition of operators, and the algebra norm is the operator norm. If $\mathscr{A}$ is a unital Banach algebra, then the *spectrum* of $a \in \mathscr{A}$ is defined as

$$\mathrm{Sp}(a) := \{\lambda \in \mathbb{F} : a - \lambda\mathbb{1} \notin G(\mathscr{A})\}.$$

The spectrum of a complex algebra is always a nonempty, compact set. The *spectral radius* is defined as

$$r(a) := \inf_{n\in\mathbb{N}, n\geq 1} \|a^n\|^{\frac{1}{n}}.$$

By the Gelfand formula $r(a) = \lim_{n\to\infty} \|a^n\|^{\frac{1}{n}}$. More intuitively the spectral radius is the maximum of the absolute value of the spectrum, i.e. $r(a) = \max\{|\lambda| : \lambda \in \mathrm{Sp}(a)\}$. If $a, b \in \mathscr{A}$ are commuting elements, then

$$\begin{aligned} r(a+b) &\leq r(a) + r(b), \\ r(ab) &\leq r(a)r(b). \end{aligned}$$

A complex algebra $(\mathscr{A}, *)$ is a **-algebra* (or involutive algebra) if $* : \mathscr{A} \to \mathscr{A}$ is an involution: for every $x, y \in \mathscr{A}$ and $\lambda \in \mathbb{C}$

(1) $(x^*)^* = x$,

(2) $(x+y)^* = x^* + y^*$ and $(\lambda x)^* = \overline{\lambda} x^*$,

(3) $(xy)^* = y^* x^*$.

A *-algebra $\mathscr{A}$ is a *Banach *-algebra* with the algebra norm $\|.\|$, if $\|x^*\| = \|x\|$ for every $x \in \mathscr{A}$. A Banach *-algebra is a *C*-algebra* if it also has the C^*-property, that for every $x \in \mathscr{A}$ $\|x^*x\| = \|x\|^2$. In Banach *-algebras there are a few classes of special elements:

(1) an element a is *normal*, if $aa^* = a^*a$,

(2) a normal element u is *unitary*, if $uu^* = u^*u = \mathbb{1}$,

(3) a normal element a is *self-adjoint*, if $a = a^*$,

(4) a self-adjoint element a is an *orthogonal projection* if it is idempotent as well.

A pre-Hilbert space $\mathcal{H}$ is a vector space equipped with a scalar product $\langle .,. \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{F}$ which induces the norm $\|x\| = \sqrt{\langle x, x\rangle}$ $(x \in \mathcal{H})$. The scalar product is also called the inner product. A pre-Hilbert space is a *Hilbert space* if it is a Banach space with the induced norm. Two vectors $x, y \in \mathcal{H}$ are *orthogonal* if $\langle x, y\rangle = 0$. Let T^* denote the adjoint of an operator $T \in \mathscr{B}(\mathcal{H})$. $\mathscr{B}(\mathcal{H})$ is a C^*-algebra with adjoining as the involution. A self-adjoint $T \in \mathscr{B}(\mathcal{H})$ is *positive* if $\langle Tx, x\rangle \geq 0$ for every $x \in \mathcal{H}$.
If $\mathscr{A}$ is a unital complex Banach algebra, and $\varphi : \mathbb{C} \to \mathbb{C}$ is an entire function, then the theory of entire (or holomorphic) functional calculus gives meaning to $\varphi(a)$. The value of φ at a is the limit of the absolute convergent series

$$\varphi(a) := \sum_{n=0}^{\infty} \frac{\varphi^{(n)}(0)}{n!} a^n.$$

A consequence of the functional calculus is the spectral mapping theorem:

Theorem 1.1. *If $\mathscr{A}$ is a complex unital Banach algebra, and φ is an entire complex function, then for $a \in \mathscr{A}$ the spectrum of $\varphi(a)$ is*

$$\mathrm{Sp}(\varphi(a)) = \varphi(\mathrm{Sp}(a)).$$

For normal elements of C^*-algebras the functional calculus can be extended into a continuous functional calculus. This makes possible to define the square root of a positive operator.

Theorem 1.2. *If $\mathcal{H}$ is a complex Hilbert space, $T \in \mathscr{B}(\mathcal{H})$ is a positive operator, then there is a unique positive operator $T^{\frac{1}{2}} \in \mathscr{B}(\mathcal{H})$ such that $(T^{\frac{1}{2}})^2 = T$.*

In finite dimensional Hilbert spaces normal matrices can be diagonalized. The infinite dimensional analogue is the spectral decomposition of normal operators [10].

Theorem 1.3 (Spectral theorem)**.** *If $N \in \mathscr{B}(\mathcal{H})$ is a normal operator, then there is a unique regular spectral measure P on the Borel sets of $\mathrm{Sp}(N)$ such that*

$$N = \int_{\mathrm{Sp}(N)} \lambda \, dP(\lambda).$$

For any nonempty open set $A \subseteq \mathrm{Sp}(N)$ we have $P(A) \neq 0$. A $T \in \mathscr{B}(\mathcal{H})$ operator commutes with N if and only if it commutes with $P(B)$ for every Borel set $B \subseteq \mathrm{Sp}(N)$.

Two operators S, T are said to *double commute* if $ST = TS$ and $S^*T = TS^*$ (this implies that $ST^* = T^*S$ too). One theorem that can be proved with the aid of the spectral theorem is Fuglede's theorem, which states if an operator commutes with a normal operator then they double commute.

Theorem 1.4 (Fuglede)**.** *If $T, N \in \mathscr{B}(\mathcal{H})$, N is normal, and $TN = NT$, then $TN^* = N^*T$ as well.*

Chapter 2

Numerical Range in Hilbert Spaces

2.1 Convexity and Spectral Inclusion

The definition of the numerical range of a Hilbert space operator is not very complicated. It is the image of the unit sphere under the quadratic form associated with the operator. Two strong properties can be proven from this definition: this is a convex set, and its closure contains the spectrum. (Unless noted otherwise the following results in this chapter are based on Gustafson's and Rao's book on numerical ranges [7].)

Definition 2.1. Let $T \in \mathscr{B}(\mathcal{H})$ be a bounded operator on a real or complex Hilbert space. Then its *numerical range* is

$$W(T) := \{\langle Tx, x\rangle : x \in \mathcal{H}, \|x\| = 1\}.$$

Remark. For any $z, w \in \mathbb{C}$ complex number, arbitrary $S, T \in \mathscr{B}(\mathcal{H})$ and unitary $U \in \mathscr{B}(\mathcal{H})$ operator the numerical range has the following properties:

$$\begin{gathered} W(zT + wI) = zW(T) + \{w\}, \\ W(T^*) = \{\bar{z} \mid z \in W(T)\}, \\ W(U^*TU) = W(T), \\ W(S+T) \subseteq W(S) + W(T), \end{gathered}$$

where $I \in \mathscr{B}(\mathcal{H})$ is the identity operator.

While the definition allows both real and complex numbers as the underlying field of the Hilbert space, the real case lacks features of the complex case. For example it is possible in the real setting that $W(T) = \{0\}$ for a nonzero operator. To illustrate this let $\mathcal{H} = \mathbb{R}^2$, and let T be the rotation by 90 degrees:

$$T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

For $x = (x_1, x_2) \in \mathbb{R}^2$ we have $\langle Tx, x\rangle = \langle(-x_2, x_1), (x_1, x_2)\rangle = 0$, so $W(T) = \{0\}$. In complex Hilbert spaces as a consequence of Theorem 2.7 $W(T) = \{0\}$ if and only if T is the zero operator. From now on except when noted otherwise it will be assumed that the Hilbert space is complex.

In two dimensions the numerical range has a simple characterisation: it is just an ellipse (allowing the case that the ellipse is just a line segment). This fact will also help to prove the convexity of the numerical range.

Lemma 2.2. *If* $\dim(\mathcal{H}) = 2$, $W(T)$ *is a filled ellipse.*

Proof. Let $D = \{z \in \mathbb{C} : |z| \leq 1\} \subseteq \mathbb{C}$. In the special case when

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

$W(T) = \frac{1}{2}D$. If $x = (z_1, z_2)$ and $\|x\| = |z_1|^2 + |z_2|^2 = 1$, then $\langle Tx, x\rangle = z_2\bar{z}_1$. By the inequality of arithmetic and geometric means $|\langle Tx, x\rangle| = |z_1||z_2| \leq \frac{1}{2}(|z_1|^2 + |z_2|^2) = \frac{1}{2}$. So $W(T) \subseteq \frac{1}{2}D$. And for every $z \in \frac{1}{2}D$ there is an $x \in \mathcal{H}$ such that $|\langle Tx, x\rangle| = z$. If $z = re^{i\theta}$, $0 \leq r \leq \frac{1}{2}$, let $x = (\cos\alpha, e^{i\theta}\sin\alpha)$, where $\sin 2\alpha = 2r$ and $0 \leq \alpha \leq \frac{\pi}{4}$. This way $\langle Tx, x\rangle = e^{i\theta}\sin\alpha\cos\alpha = re^{i\theta}$.

In the general case let us take the Schur decomposition: $T = USU^*$ where U is a unitary matrix and S is an upper triangular matrix with the eigenvalues of T on its diagonal:

$$S = \begin{bmatrix} \lambda_1 & z \\ 0 & \lambda_2 \end{bmatrix}.$$

Because of unitary similarity $W(S) = W(T)$.

If $\lambda_1 = \lambda_2 = \lambda$,

$$S - \lambda I = \begin{bmatrix} 0 & z \\ 0 & 0 \end{bmatrix} = z\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

So $W(S - \lambda I) = \frac{1}{2}|z|D$ and $W(S) = \frac{1}{2}|z|D + \{\lambda\}$.

If $\lambda_1 \neq \lambda_2$ and $z = 0$,

$$S = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Denoting $x = (z_1, z_2)$, then $\langle Tx, x\rangle = \lambda_1|z_1|^2 + \lambda_2|z_2|^2$. Because $|z_1|^2 + |z_2|^2 = 1$, $W(S)$ is the line segment joining λ_1 and λ_2.

If $\lambda_1 \neq \lambda_2$ and $z \neq 0$, $S_1 := S - \frac{\operatorname{Tr} S}{2} I$. If the eigenvalues of S_1 are μ and $-\mu$, $S_2 := \frac{1}{\mu} S_1$. If S_2 is not normal, it is unitarily similar to

$$\begin{bmatrix} 1 & 2c \\ 0 & -1 \end{bmatrix} =: S_3,$$

where $c > 0$. $S_4 := \frac{1}{2}((S_3 + S_3^*) + \frac{\sqrt{1+c^2}}{c}(S_3 - S_3^*))$. Both eigenvalues of S_4 are zero, and it is unitarily similar to

$$\begin{bmatrix} 0 & 2\sqrt{1+c^2} \\ 0 & 0 \end{bmatrix},$$

so $W(S_4) = \sqrt{1+c^2} D$, and

$$u + vi \in W(S_3) \Leftrightarrow u + \frac{\sqrt{1+c^2}}{c} vi \in W(S_4).$$

Because $\partial W(S_4) = \{\sqrt{1+c^2} e^{it} \mid t \in \mathbb{R}\}$, $\partial W(S_3) = \{\sqrt{1+c^2}\cos t + ic\sin t \mid t \in \mathbb{R}\}$. So $W(S_3)$ is an elliptical disk, and we get $W(S)$ by scaling and translating that ellipse. □

Theorem 2.3 (Toeplitz-Hausdorff)**.** *The numerical range is a convex set in the complex plane.*

Proof. We have $z_1, z_2 \in W(T)$, $z_1 = \langle Tx_1, x_1 \rangle$, $z_2 = \langle Tx_2, x_2 \rangle$, where $\|x_1\| = \|x_2\| = 1$. We need that the line segment joining z_1 and z_2, denoted by $[z_1, z_2]$, is within $W(T)$. Let P be the orthogonal projection on the subspace $\operatorname{span}(x_1, x_2)$. Because $\langle PTPx_i, x_i \rangle = \langle Tx_i, x_i \rangle$ $(i = 1, 2)$, $z_1, z_2 \in W(PTP)$. By the previous lemma $W(PTP)$ is an ellipse, therefore $[z_1, z_2] \subseteq W(PTP)$. If $z = \langle PTPy, y \rangle \in W(PTP)$, then $z = \langle TPy, Py \rangle \in W(T)$ as well. We got that $W(PTP) \subseteq W(T)$ and $W(T)$ contains the line segment. □

The following lemma about invertible operators, and the proof of the spectral inclusion theorem is from [17] (Lemma 7.29 and Theorem 7.30).

Lemma 2.4. *$T \in \mathscr{B}(\mathcal{H})$ is invertible if and only if* $\ker(T^*) = \{0\}$ *and there is such a $C > 0$ constant that $\|Tx\|^2 \geq C\|x\|^2$ for every $x \in \mathcal{H}$.*

Theorem 2.5 (Spectral inclusion)**.** *The spectrum of an operator $T \in \mathscr{B}(\mathcal{H})$ is contained in the closure of the numerical range, i.e.*

$$\operatorname{Sp}(T) \subseteq \overline{W(T)}.$$

Proof. If $\lambda \in \operatorname{Sp}(T)$, then by the previous lemma either $\ker(T^* - \bar{\lambda} I) \neq \{0\}$ or there is a $(x_n)_{n \in \mathbb{N}}$ sequence of unit-norm vectors such that $(T - \lambda I)x_n \to 0$. In the first case let us pick x from $\ker(T^* - \bar{\lambda} I)$ with $\|x\| = 1$. For this vector

$$0 = \left\langle x, (T^* - \bar{\lambda} I)x \right\rangle = \langle (T - \lambda I)x, x \rangle = \langle Tx, x \rangle - \lambda.$$

So $\lambda = \langle Tx, x\rangle \in W(T)$. In the second case let $\lambda_n := \langle Tx_n, x_n\rangle$. Because $\|x_n\| = 1$ for every n, $\lambda_n \in W(T)$, and

$$|\lambda_n - \lambda| = |\,\langle Tx_n - \lambda x_n, x_n\rangle\,| \leq \|(T - \lambda I)x_n\| \to 0.$$

Thus $\lambda_n \to \lambda$, which means that $\lambda \in \overline{W(T)}$. □

For an arbitrary operator the spectrum can be much smaller than the numerical range. For example we have seen that $W(T) = \{z \in \mathbb{C} : |z| \leq \frac{1}{2}\}$ for the two-dimensional shift operator $T = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, but $\operatorname{Sp}(T) = \{0\}$. For normal operators the spectrum and the numerical range is more closely linked. In Theorem 2.27 we will see that for normal operators the closure of the numerical range is the same as the convex hull of the spectrum.

2.2 Numerical Radius

Analogously to the spectrum and the spectral radius, we can define the numerical radius based on the numerical range. This way we have three similar numerical quantities that characterize an operator: the spectral radius, the numerical radius, and the operator norm. This section is about the relationship between these three numbers and some other properties of the numerical radius.

Definition 2.6. The *numerical radius* of $T \in \mathscr{B}(\mathcal{H})$ is

$$w(T) := \sup\{|z| : z \in W(T)\}.$$

Theorem 2.7. *The numerical radius is a norm on $\mathscr{B}(\mathcal{H})$. It is equivalent to the operator norm, and $w(T) \leq \|T\| \leq 2w(T)$.*

Proof. (Following [17], Theorems 3.62 and 3.63). Being defined by absolute values, $w(T) \geq 0$. Because for every $S, T \in \mathscr{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$

$$w(\lambda T) = \sup_{\|x\|=1} |\,\langle \lambda Tx, x\rangle\,| = |\lambda| \sup_{\|x\|=1} |\,\langle Tx, x\rangle\,| = |\lambda| w(T),$$
$$w(S+T) = \sup_{\|x\|=1} |\,\langle (S+T)x, x\rangle\,| \leq \sup_{\|x\|=1} |\,\langle Sx, x\rangle\,| + \sup_{\|x\|=1} |\,\langle Tx, x\rangle\,| = w(S) + w(T),$$

so we have that $w : \mathscr{B}(\mathcal{H}) \to \mathbb{R}_+$ is a seminorm. If $w(T) = 0$, then for every $x \in \mathcal{H}$ $\langle Tx, x\rangle = 0$. This means that $T = 0$ (again because $\mathcal{H}$ is complex) and w is a norm. If $x \in \mathcal{H}$, $\|x\| = 1$, then by the Cauchy-Schwart inequality

$$|\,\langle Tx, x\rangle\,| \leq \|Tx\|\|x\| \leq \|T\|.$$

Using the fact that $|\langle Tx, x\rangle| \leq w(T)\|x\|^2$, the polarization principle, and the parallelogram law we get that for $x, y \in \mathcal{H}$

$$
\begin{aligned}
4 \cdot |\langle Tx, y\rangle| &\leq \sum_{k=0}^{3} |\langle T(x+i^k y), x+i^k y\rangle| \\
&\leq \sum_{k=0}^{3} w(T)\|x+i^k x\|^2 \\
&= w(T)(\|x+y\|^2 + \|x-y\|^2) + w(T)(\|x+iy\|^2 + \|x-iy\|^2) \\
&= 2w(T)(\|x\|^2 + \|y\|^2) + 2w(T)(\|x\|^2 + \|iy\|^2) \\
&= 4w(T)(\|x\|^2 + \|y\|^2).
\end{aligned}
$$

Substituting normalized vectors $\frac{x}{\|x\|}, \frac{y}{\|y\|}$,

$$|\langle Tx, y\rangle| \leq 2w(T)\|x\|\|y\|.$$

By choosing $y = Tx$ we get

$$\|Tx\|^2 \leq 2w(T)\|x\|\|Tx\|.$$

Putting the two bounds together $\frac{1}{2}\|T\| \leq w(T) \leq \|T\|$. □

Theorem 2.5 and Theorem 2.7 together produce the following chain of inequalities.

Corollary 2.8. *For every $T \in \mathscr{B}(\mathcal{H})$ operator $r(T) \leq w(T) \leq \|T\|$.*

Both inequalities in Theorem 2.7 are sharp. In the next section we will see that for normal operators $r(T) = w(T) = \|T\|$. Propositions 2.9 to 2.11 are about some of the other edge cases.

Proposition 2.9. *If $w(T) = \|T\|$, then $r(T) = \|T\|$.*

Proof. We will show the non-invertibility of $T - \lambda I$ by applying Lemma 2.4. According to the assumption there must be a $\lambda \in \mathbb{C}$ number and a $(x_n)_{n\in\mathbb{N}} \subseteq \mathcal{H}$ sequence of unit-norm vectors such that $|\lambda| = \|T\|$ and $\langle Tx_n, x_n\rangle \to \lambda$. Because $|\langle Tx_n, x_n\rangle| \leq \|Tx_n\| \leq \|T\|$, by squeezing we get $\|Tx_n\| \to \|T\|$. Therefore

$$
\begin{aligned}
\|(T-\lambda I)x_n\|^2 &= \langle Tx_n - \lambda x_n, Tx_n - \lambda x_n\rangle \\
&= \|Tx_n\|^2 - \langle Tx_n, \lambda x_n\rangle - \langle \lambda x_n, Tx_n\rangle + |\lambda|^2\|x_n\|^2 \\
&\leq \|T\|^2 - \overline{\lambda}\langle Tx_n, x_n\rangle - \lambda\langle x_n, Tx_n\rangle + \|T\|^2 \\
&\to 2\|T\|^2 - 2\overline{\lambda}\lambda = 0.
\end{aligned}
$$

So $\lambda \in \mathrm{Sp}(T)$ and $r(T) \geq \|T\|$. □

Proposition 2.10. *If $\lambda \in W(T), |\lambda| = \|T\|$, then λ is an eigenvalue.*

Proof. If $\lambda = \langle Tx, x\rangle, \|x\| = 1$, then $\|T\| = |\lambda| = |\langle Tx, x\rangle| \leq \|Tx\|\|x\| \leq \|T\|\|x\|^2$. Applying the Cauchy-Schwarz inequality again to $|\langle Tx, x\rangle| = \|Tx\|\|x\|$ we get that $Tx = \mu x$ for some $\mu \in \mathbb{C}$. But $\mu = \langle \mu x, x\rangle = \langle Tx, x\rangle = \lambda$. □

Proposition 2.11. *If the range of T is orthogonal to the range of T^*, then $w(T) = \frac{1}{2}\|T\|$.*

Proof. Let us take any $x \in \mathcal{H}$ unit vector and decompose it into $x = x_1 + x_2$ where $x_1 \in \ker T, x_2 \in (\ker T)^\perp = \overline{\operatorname{ran} T^*}$. Since $Tx_2 \in \operatorname{ran} T, x_2 \in \overline{\operatorname{ran} T^*}$, $\langle Tx_2, x_2\rangle = 0$ due to the orthogonality assumption. Using this, and that $Tx_1 = 0$ we get

$$\langle Tx, x\rangle = \langle Tx_1 + Tx_2, x_1 + x_2\rangle = \langle Tx_2, x_1\rangle .$$

By the Cauchy-Schwarz inequality and the inequality of means

$$|\langle Tx, x\rangle| \leq \|T\|\|x_1\|\|x_2\| \leq \|T\|\frac{\|x_1\|^2 + \|x_2\|^2}{2} = \frac{\|T\|}{2}\|x\|.$$

Because this stands for any unit vector, $w(T) \leq \frac{1}{2}\|T\| \leq w(T)$. □

Example 2.12. Let T be the left shift operator on $\mathbb{C}^2$. Then $\operatorname{ran} T = \{(z, 0) : z \in \mathbb{C}\}$, $\operatorname{ran} T^* = \{(0, z) : z \in \mathbb{C}\}$, $r(T) = 0$, $w(T) = \frac{1}{2}$ and $\|T\| = 1$. So $r(T) < w(T) < \|T\|$ and $w(T) = \frac{1}{2}\|T\|$. This is a special case of a more general construction. If $e, f \in \mathcal{H}$ then let us define $e \otimes f \in \mathscr{B}(\mathcal{H})$ as $(e \otimes f)x = \langle x, e\rangle f$ for $x \in \mathcal{H}$. The adjoint is $(e \otimes f)^*x = (f \otimes e)x$. If we choose e and f in such a way that they are orthonormal, then

$$\langle (e \otimes f)x, (e \otimes f)^*x\rangle = \langle x, e\rangle \overline{\langle x, f\rangle} \langle f, e\rangle = 0,$$

and if $\|x\| = 1$, by the Cauchy-Schwarz inequality

$$\|(e \otimes f)x\| = |\langle x, e\rangle|^2 \langle f, f\rangle \leq \langle x, x\rangle \langle e, e\rangle = 1.$$

We have $\|(e \otimes f)e\| = 1$, so $\|e \otimes f\| = 1$. And we can use the calculations in the previous proof to show that $w(e \otimes f) = \frac{1}{2}$.

It is known that an idempotent operator is an orthogonal projection if and only if their operator norm is less than or equal to 1. This statement can be made a little sharper by using numerical radius in the condition instead.

Proposition 2.13. *If $T^2 = T$ and $w(T) \leq 1$, then T is an orthogonal projection.*

Proof. We will show that $T = 0$ if T is restricted to $(\operatorname{ran} T)^\perp$. Let us pick an $x \in \operatorname{ran}(T)^\perp$, and $y := Tx$ (so $\langle x, y\rangle = 0$). Because of idempotence $Ty = y$, so for any $t \geq 0$ $\langle T(x + ty), x + ty\rangle = \langle y + ty, x + ty\rangle = \langle (1 + t)y, x + ty\rangle = \langle (1 + t)y, ty\rangle = (1 + t)t\|y\|^2$.

Because $w(T) \leq 1$, $\langle T(x+ty), x+ty\rangle \leq \|x+ty\|^2$. Putting the two together we get $(1+t)t\|y\|^2 \leq \|x+ty\|^2 = \|x\|^2 + t^2\|y\|^2$ or $t\|y\| \leq \|x\|$. This is true for any t, so $y = 0$. □

The numerical radius is not submultiplicative so it is not an algebra norm. It is not submultiplicative for commuting elements either. It is possible to show a case where submultiplicativity fails even for T and T^2 ([8], section 221). In the construction of the example we will use the following proposition about the numerical range in the direct sum of Hilbert spaces.

Proposition 2.14. *If $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$ is a direct sum of Hilbert spaces, then*

$$w(T_1 \oplus T_2) = \max\{w(T_1), w(T_2)\},$$

where $T_1 \in \mathscr{B}(\mathcal{H}_1), T_2 \in \mathscr{B}(\mathcal{H}_2)$.

Proof. If $T = T_1 \oplus T_2$, then $W(T_1), W(T_2) \subseteq W(T)$ and $\mathrm{conv}(W(T_1), W(T_2)) \subseteq W(T)$. If $x \in \mathcal{H}$ is a unit vector, then $x = x_1 + x_2$ where $x_1 \in \mathcal{H}_1, x_2 \in \mathcal{H}_2$ and $\|x_1\|^2 + \|x_2\|^2 = 1$. With unit vectors $u_1 = \frac{x_1}{\|x_1\|} \in \mathcal{H}_1, u_2 = \frac{u_2}{\|u_2\|} \in \mathcal{H}_2$

$$\begin{aligned}\langle Tx, x\rangle &= \langle T(x_1+x_2), x_1+x_2\rangle \\ &= \langle Tx_1, x_1\rangle + \langle Tx_2, x_2\rangle \\ &= \|x_1\|^2 \langle T_1u_1, u_1\rangle + \|x_2\|^2 \langle T_2u_2, u_2\rangle \\ &\in \mathrm{conv}(W(T_1), W(T_2)).\end{aligned}$$

Thus $W(T_1 \oplus T_2) = W(T) = \mathrm{conv}(W(T_1), W(T_2))$, which implies the statement. □

Now we are ready to analyze a relatively simple operator T for which the numerical radius does not behave very well. We will show that $w(T)w(T^2) \nleq w(T^3)$, furthermore $w(T^*T) \neq w(T)^2$.

Example 2.15. Let us look at this $T \in \mathscr{B}(\mathbb{C}^4)$ shift operator:

$$T = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

If $x = (z_1, z_2, z_3, z_4) \in \mathbb{C}^4, |z_1|^2 + |z_2|^2 + |z_3|^2 + |z_4|^2 = 1$, then

$$|\langle Tx, x\rangle| = |z_1||z_2| + |z_2||z_3| + |z_3||z_4|.$$

By the Cauchy-Schwarz inequality

$$|\langle Tx, x\rangle| \leq (|z_2|^2 + |z_3|^2 + |z_4|^2) \leq 1.$$

But since

$$|z_1||z_2| + |z_2||z_3| + |z_3||z_4| = |z_1|^2 + |z_2|^2 + |z_3|^2 + |z_4|^2$$

is not possible, $w(T) < 1$. By exchanging the proper rows and columns we get that T^2 is unitarily similar to

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

and T^3 is unitarily similar to

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Therefore $w(T^2) = w(T^3) = \frac{1}{2}$. But this also means that $w(TT^2) > w(T)w(T^2)$. The C^*-property does not hold either: $T^*T = \operatorname{diag}(0, 1, 1, 1)$ so $w(T^*T) = 1$, on the other hand $w(T)^2 < 1$.

Submultiplicativity can hold with additional assumptions, like in Theorem 2.31. In general a weaker property can be proved, the power inequality. The following proof is by Carl Pearcy [14].

Lemma 2.16. *Let $x \in \mathcal{H}$ be a unit vector, $T \in \mathscr{B}(\mathcal{H})$, $n \in \mathbb{N}_+$, $\omega_j := e^{\frac{2\pi i j}{n}}$, and*

$$x_j := \left(\prod_{k=1, k\neq j}^{n} (1 - \omega_k T) \right) x.$$

In this case

$$1 - \langle T^n x, x\rangle = \frac{1}{n} \sum_{j=1}^{n} \|x_j\|^2 \left(1 - \omega_j \left\langle \frac{Tx_j}{\|x_j\|}, \frac{x_j}{\|x_j\|} \right\rangle \right).$$

Proof. Using

$$1 - T^n = \prod_{k=1}^{n} (1 - \omega_k T)$$

and

$$1 = \frac{1}{n} \sum_{j=1}^{n} \prod_{\substack{k=1 \\ k\neq j}}^{n} (1 - \omega_k T)$$

we get that

$$\frac{1}{n}\sum_{j=1}^{n}\|x_j\|^2\left(1-\omega_j\left\langle\frac{Tx_j}{\|x_j\|},\frac{x_j}{\|x_j\|}\right\rangle\right)=$$
$$\frac{1}{n}\sum_{j=1}^{n}\langle(1-\omega_jT)x_j,x_j\rangle=\frac{1}{n}\sum_{j=1}^{n}\langle(1-T^n)x,x_j\rangle=$$
$$\left\langle(1-T^n)x,\frac{1}{n}\sum_{j=1}^{n}\left(\prod_{k=1,k\neq j}^{n}(1-\omega_kT)\right)x\right\rangle=$$
$$\langle(1-T^n)x,x\rangle=1-\langle T^nx,x\rangle\,.$$

□

Theorem 2.17 (Power inequality)**.** *If $T\in\mathscr{B}(\mathcal{H})$ and $n\in\mathbb{N}$,*

$$w(T^n)\leq w(T)^n.$$

Proof. First let us suppose that $w(T)\leq 1$. We will show that $w(T^n)\leq 1$ $(n\in\mathbb{N})$. If $x\in\mathcal{H}$ is a unit vector, $\theta\in\mathbb{R}$, and we use the previous lemma with $e^{i\theta}T$, then

$$1-e^{in\theta}\langle T^nx,x\rangle=\frac{1}{n}\sum_{j=1}^{n}\|x_j\|^2\left(1-\omega_je^{i\theta}\left\langle\frac{Tx_j}{\|x_j\|},\frac{x_j}{\|x_j\|}\right\rangle\right).$$

Because $w(T)\leq 1$, $\operatorname{Re}\left(1-\omega_je^{i\theta}\left\langle\frac{Tx_j}{\|x_j\|},\frac{x_j}{\|x_j\|}\right\rangle\right)\geq 0$, so $\operatorname{Re}\left(1-e^{in\theta}\langle T^nx,x\rangle\right)\geq 0$ as well. This true for every $\theta\in\mathbb{R}$, therefore $\|\langle T^nx,x\rangle\|\leq 1$. In general $w(\frac{T}{w(T)})=1$, so $\frac{1}{w(T)^n}w(T^n)=w(\frac{T^n}{w(T)^n})\leq 1$. □

2.3 Self-adjoint and Normal Operators

It is possible to make stronger statements about the numerical range and numerical radius if we make the additional assumption that the operator is normal or that it is self-adjoint. Conversely, normality and self-adjointness can be characterized in terms of the numerical range.

Theorem 2.18. *$T\in\mathscr{B}(\mathcal{H})$ is self-adjoint if and only if $W(T)\subseteq\mathbb{R}$.*

Proof. If $T=T^*$, for every $x\in\mathcal{H}$ $\langle Tx,x\rangle=\langle x,Tx\rangle=\overline{\langle Tx,x\rangle}$, so $W(T)$ is real. If $\langle Tx,x\rangle\in\mathbb{R}$ for every $x\in\mathcal{H}$, $0=\langle Tx,x\rangle-\langle x,Tx\rangle=\langle(T-T^*)x,x\rangle$. In a complex Hilbert space if $A,B\in\mathscr{B}(\mathcal{H})$ and for every $x\in\mathcal{H}$ $\langle Ax,x\rangle=\langle Bx,x\rangle$, then $A=B$. So $T-T^*$ must be the null operator and $T=T^*$. □

Corollary 2.19. *$T \in \mathscr{B}(\mathcal{H})$ is a positive operator if and only if $W(T) \subseteq \mathbb{R}_+$.*

From Theorem 2.18 and Theorem 2.5 we can derive the following property of self-adjoint and positive operators. By the end of this section we will be able to prove the converse of this as well.

Corollary 2.20. $\mathrm{Sp}(T) \subseteq \mathbb{R}$ *if T is self-adjoint,* $\mathrm{Sp}(T) \subseteq \mathbb{R}_+$ *if T is a positive operator.*

Theorem 2.21. *In both real and complex Hilbert spaces, $\|T\| = w(T)$ for self-adjoint operators.*

Proof. For any $x, y \in \mathcal{H}$

$$2(\langle Tx, y\rangle + \langle Ty, x\rangle) = \langle T(x+y), x+y\rangle - \langle T(x-y), x-y\rangle\,,$$

so

$$\begin{aligned} 2|\,\langle Tx, y\rangle + \langle Ty, x\rangle\,| &\le |\,\langle T(x+y), x+y\rangle\,| + |\,\langle T(x-y), x-y\rangle\,| \\ &\le w(T)(\|x+y\|^2 + \|x-y\|^2) \\ &\le 2w(t)(\|x\|^2 + \|x\|^2). \end{aligned}$$

If we apply this to $\frac{x}{\|x\|}$ and $\frac{Tx}{\|Tx\|}$ and a self-adjoint T,

$$2\frac{\|Tx\|}{\|x\|} = \frac{2}{\|Tx\|\|x\|}\langle Tx, Tx\rangle = \left\langle T\frac{x}{\|x\|}, \frac{Tx}{\|Tx\|}\right\rangle + \left\langle T\frac{Tx}{\|Tx\|}, \frac{x}{\|x\|}\right\rangle \le 2w(T).$$

So $\|T\| \le w(T)$ for self-adjoint T. And by the Cauchy-Schwarz inequality $w(T) \le \|T\|$ for any operator. □

Proposition 2.22. *In case $W(T)$ is a line segment, T is a normal operator.*

Proof. If $z \in W(T)$, and the line segment has inclination θ then $W(e^{-i\theta}(T - zI)) \subseteq \mathbb{R}$, so because of Theorem 2.18 $e^{-i\theta}(T - zI)$ is self-adjoint. This means that $e^{-i\theta}(T - zI) = e^{i\theta}(T^* - \bar{z}I)$, so $T^* = e^{-2i\theta}Tz + (\bar{z} - e^{-2i\theta}z)I$. Becaue T^* is a polynomial of T, they commute and T is normal. □

Theorem 2.23. *If T is normal, then $r(T) = w(T) = \|T\|$.*

Proof. For normal operators $\|Tx\| = \|T^*x\|$ for every $x \in \mathcal{H}$, so $\|T^2x\| = \|T(Tx)\| = \|T^*(Tx)\|$. Because of this and the C^*-property $\|T^2\| = \|T^*T\| = \|T\|^2$. By induction we get the property of normal operators that $\|T^n\| = \|T\|^n$. Using this and the Gelfand limit formula for the spectral radius

$$r(T) = \lim_{n\to\infty} \|T^n\|^{1/n} = \|T\|.$$

By Corollary 2.8 we always have $r(T) \le w(T) \le \|T\|$, so we are done. □

For an operator $T \in \mathscr{B}(\mathcal{H})$ we have $w(T) = \|T\|$ if T is self-adjoint and $\mathcal{H}$ is real, or if T is normal and $\mathcal{H}$ is complex . This not true if T is normal and $\mathcal{H}$ is real. Let us suppose that $\mathcal{H} = \mathbb{R}^2$ and

$$T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Then $w(T) = 0$ but $\|T\| = 1$.

Corollary 2.24. *For a unitary operator U, $w(U) = 1$.*

Lemma 2.25. *If T is a normal operator and $z \notin \mathrm{Sp}(T)$, then*

$$\|(T - zI)^{-1}\| = \frac{1}{d(z, \mathrm{Sp}(T))}.$$

Proof. Because $(T - zI)^{-1}$ is normal as well, we can use Theorem 2.23, and

$$\|(T - zI)^{-1}\| = r((T - zI)^{-1}) = \sup\{|\lambda| : \lambda \in \mathrm{Sp}((T - zI)^{-1})\}.$$

Since we have $\mathrm{Sp}((T - zI)^{-1}) = \{1/(\lambda - z) : \lambda \in \mathrm{Sp}(T)\}$, therefore

$$\|(T - zI)^{-1}\| = \sup\left\{\frac{1}{|\lambda - z|} : \lambda \in \mathrm{Sp}(T)\right\} = \frac{1}{\inf\{|z - \lambda| : \lambda \in \mathrm{Sp}(T)\}} = \frac{1}{d(z, \mathrm{Sp}(T))}.$$

□

Corollary 2.26. *If T is a normal operator and $z \notin \mathrm{Sp}(T)$, then $\|(T - zI)x\| \geq d(z, \mathrm{Sp}(T))$ for every unit vector $x \in \mathcal{H}$.*

Theorem 2.27. *If T is normal, then $\overline{W(T)} = \mathrm{conv}(\mathrm{Sp}(T))$.*

Proof. It is enough to show that any closed half plane of the complex plane which contains the spectrum also contains the numerical range. Furthermore it is enough to check the special case when $\mathrm{Sp}(T) \subset \{\mathrm{Re}\,\lambda \leq 0\}$ and the imaginary axis is a supporting line of the convex hull.
Let us suppose that the numerical range extends into the right half plane, so there is a point $a + bi \in W(T)$ where $a > 0$. If x is the unit vector for which $\langle Tx, x\rangle = a + bi$, then $Tx = (a + bi)x + y$ where $\langle x, y\rangle = 0$. Let us pick any point c on the positive real axis. This point is not in $\mathrm{Sp}(T)$ so by the previous corollary $d(c, \mathrm{Sp}(T)) \leq \|(T - cI)x\|$, that is

$$c^2 \leq \|(a + bi)x + y - cx\|^2 = |(a - c) + bi|^2\|x\|^2 + \|y\|^2 = (a - c)^2 + b^2 + \|y\|^2.$$

So $2ac \leq a^2 + b^2$ where a and c is positive, but this leads to a contradiction because c can be arbitrarily chosen. □

Putting together Theorem 2.27 with Theorem 2.18 produces the spectral characterization of self-adjoint and positive operators.

Corollary 2.28. *An operator T is self-adjoint if and only if it is normal and* $\mathrm{Sp}(T) \subseteq \mathbb{R}$. *$T$ is positive if and only if it is normal and* $\mathrm{Sp}(T) \subseteq \mathbb{R}_+$.

2.4 Commuting Operators

While the numerical radius is not submultiplicative, it is possible to bound $w(ST)$ with some constant multiple of $w(S)$ and $w(T)$. In the general case this is a simple consequence of the equivalence of the numerical radius to the norm. For commuting operators the bound can be made tighter.

Proposition 2.29. *For any two operators S, T we have $w(ST) \leq 4w(S)w(T)$. Moreover if $ST = TS$, then $w(ST) \leq 2w(S)w(T)$.*

Proof. Because of norm equivalence it is true in general that

$$w(ST) \leq \|ST\| \leq \|S\|\|T\| \leq 4w(S)w(T).$$

To prove the commutative case first let us assume that $w(S) = w(T) = 1$:

$$\begin{aligned}
w(ST) &= w\left(\frac{1}{4}((S+T)^2 - (S-T)^2)\right) && (S \text{ and } T \text{ commute}) \\
&\leq \frac{1}{4}(w((S+T)^2) + w((S-T)^2)) && (w \text{ is subadditive}) \\
&\leq \frac{1}{4}((w(S+T))^2) + (w(S-T))^2) && (\text{power inequality}) \\
&\leq \frac{1}{4}((w(S)+w(T))^2 + (w(S)+w(T))^2) && (\text{subadditivity again}) \\
&= 2.
\end{aligned}$$

If $w(S), w(T) \neq 1$, we can replace S and T with $\frac{1}{w(S)}S$ and $\frac{1}{w(T)}T$. □

Both of the above inequalities are sharp. First let us have

$$T_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } T_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

We have seen in the proof of 2.2 that $w(T_1) = \frac{1}{2}$, and $w(T_2) = w(T_1)$ because $T_2 = T_1^*$. If $x = (z_1, z_2) \in \mathbb{C}^2$, then $|\langle T_1T_2x, x\rangle| = |z_1|^2 \leq 1$. So $w(T_1T_2) = 4w(T_1)w(T_2)$. For the commutative case let us look at the matrices

$$T_3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad T_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad T_3T_4 = T_4T_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

With unitary similarities $T_3 \sim T_1 \oplus T_1$, $T_4 \sim T_1 \oplus T_1$, and $T_3T_4 \sim \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \oplus T_1$. So $w(T_3) = w(T_4) = w(T_3T_4) = \frac{1}{2}$ and $w(T_3T_4) = 2w(T_3)w(T_4)$.

We can say more about the numerical range of the product of two operators if we make other assumptions (that one of the operators is self-adjoint or normal, or that the operators double commute) besides the operators commuting.

Proposition 2.30. *If T is a positive operator and $ST = TS$, then $W(ST) \subseteq W(S)W(T)$.*

Proof. Because T is positive by Theorem 1.2 it has a unique square root $T^{1/2}$ for which $ST^{1/2} = T^{1/2}S$ as well. If $\|x\| = 1$, $T^{1/2}x \neq 0$ and $y := \frac{T^{1/2}x}{\|T^{1/2}x\|}$, then $\langle STx, x\rangle = \left\langle ST^{1/2}x, T^{1/2}x\right\rangle = \left\langle Sy\|T^{1/2}x\|, y\|T^{1/2}x\|\right\rangle = \langle Sy, y\rangle \langle Tx, x\rangle \in W(S)W(T)$. □

Theorem 2.31. *If T is a normal operator and $ST = TS$, then $w(ST) \leq w(S)w(T)$.*

Proof. We will use the spectral decomposition of normal operators (Theorem 1.3). Because T is normal it can be approximated in the operator norm by a sum $T_n = \sum_{i=1}^{n} \lambda_i P_i$ where $\lambda_i \in \mathrm{Sp}(T)$, every $P_i \in \mathscr{B}(\mathcal{H})$ is an orthogonal projection, for every $i \neq j$ $P_i \perp P_j$, $\sum_{i=1}^{n} P_i = I$, and $P_iS = SP_i$. We will show that for such a T_n it is true that $w(ST_n) \leq w(S)w(T_n)$. If $x \in \mathcal{H}$, $\|x\| = 1$, then

$$\begin{aligned}
|\langle ST_nx, x\rangle| &= |\langle S(\sum_{i=1}^{n} \lambda_i P_i)x\rangle, \langle(\sum_{j=1}^{n} P_j)x\rangle| \\
&= |\sum_{i,j=1}^{n} \lambda_i \langle SP_ix, P_jx\rangle| \\
&= |\sum_{i,j=1}^{n} \lambda_i \langle P_jSP_ix, P_jx\rangle| && (P_j^* = P_j = P_j^2) \\
&= |\sum_{i,j=1}^{n} \lambda_i \langle S(P_iP_j)x, P_jx\rangle| && (SP_i = P_iS) \\
&= |\sum_{i=1}^{n} \lambda_i \langle SP_ix, P_ix\rangle| && \text{(pairwise orthogonality)}
\end{aligned}$$

$$
\begin{aligned}
|\sum_{i=1}^{n} \lambda_i \langle SP_i x, P_i x\rangle| &\leq \sum_{i=1}^{n} |\lambda_i| |\langle SP_i x, P_i x\rangle| \\
&\leq \sum_{i=1}^{n} w(T)|\langle S(P_i x), (P_i x)\rangle|. \qquad (|\lambda_i| \leq r(T) = w(T)) \\
&\leq \sum_{i=1}^{n} w(T)w(S)\|P_i x\|^2 \\
&= w(T)w(S)\|\sum_{i=1}^{n} P_i x\|^2 \qquad \text{(pairwise orthogonality)} \\
&= w(T)w(S)\|Ix\|^2 \\
&= w(T)w(S).
\end{aligned}
$$

Because $w(ST_n) \leq w(S)w(T_n)$ for every approximating operator T_n, we can conclude that $w(ST) \leq w(S)w(T)$ as well. □

Corollary 2.32. *If U is unitary and $TU = UT$, then $w(UT) \leq w(T)$.*

Proof. $w(UT) \leq w(U)w(T) = \|U\|w(T) = w(T)$. □

Corollary 2.33. *If U is an isometry and $TU = UT$, then $w(UT) \leq w(T)$.*

Proof. That U is an isometry means that $U^*U = I$, so $\langle UTx, x\rangle = \langle U^*UUTx, x\rangle = \langle (UT)Ux, Ux\rangle$. Therefore it is enough to look at only $\mathrm{ran}(U)$. For $y \in \mathrm{ran}(U)$ there is x such that $y = Ux$, and $UU^*y = UU^*Ux = UIx = y$. So U acts as a unitary operator on its range and we can apply the previous result to show that $w(UT) \leq w(T)$ on $\mathrm{ran}(U)$. □

Theorem 2.34. *If $ST = TS$ and $ST^* = T^*S$ as well, then $w(ST) \leq w(S)\|T\|$.*

Proof. We will show that if $\|T\| \leq 1$ then $w(ST) \leq w(S)$. If T is a contraction, $I - T^*T$ is non-negative. So we can define the defect operator $D_T = (I - T^*T)^{1/2}$ and the Julia operator [4] acting on $\mathcal{K} = \mathcal{H} \oplus \mathcal{H}$ with the matrix representation

$$
V = \begin{bmatrix} T & D_{T^*} \\ D_T & -T^* \end{bmatrix}.
$$

From $T(I - T^*T) = (I - T^*T)T$ it follows that $TD_T = D_{T^*}T$ and that V is unitary. In other words V is a unitary dilation of T. Let $S_{\mathcal{K}} := S \oplus S$. For $x \in \mathcal{H}$ $\langle S_{\mathcal{K}}V(x, 0), (x, 0)\rangle_{\mathcal{K}} = \langle STx, x\rangle$ so $w(ST) \leq w_{\mathcal{K}}(S_{\mathcal{K}}V)$. Because S and T double commute $S_{\mathcal{K}}$ commutes with V and we can use Corollary 2.32 to show that $w(ST) \leq w_{\mathcal{K}}(S_{\mathcal{K}}V) \leq w_{\mathcal{K}}(S_{\mathcal{K}}) = w(S)$. □

This result on double commuting operators can be proved directly too, as described for example in [9]. Then Theorem 2.31 is true as a consequence of Fuglede's theorem (Theorem 1.4).

Double commuting is necessary for the proven result, it is not true that if only $ST = TS$, then $w(ST) \leq w(T)\|S\|$. On the other hand it can be shown that there is a $c < 2$ constant such that for every pair of commuting operators $w(ST) \leq c \cdot w(S)\|T\|$.

Chapter 3

Numerical Range in Banach Algebras

3.1 Algebraic Numerical Range

In this section a different numerical range is introduced, the so-called algebraic numerical range. This construct shares most of the important properties of the Hilbert space numerical range: convexity, spectral inclusion, and the equivalence of the numerical radius to the norm. Unless stated otherwise the exposition in this chapter and the next follows the monograph by Bonsall and Duncan [2].

Definition 3.1. Let $\mathscr{A}$ be a unital normed algebra over $\mathbb{C}$, and $\mathscr{A}'$ the set of every continuous linear functional over $\mathscr{A}$ (the topological dual). Let $S(\mathscr{A})$ be the unit sphere in $\mathscr{A}$,

$$S(\mathscr{A}) := \{x \in \mathscr{A} : \|x\| = 1\}.$$

For an element $x \in S(\mathscr{A})$ let us define its set of *support functionals* as

$$\mathscr{S}(\mathscr{A}, x) := \{f \in \mathscr{A}' : f(x) = \|f\| = 1\},$$

and the *states* (or normalized states) of $\mathscr{A}$ as the support functionals of the unit,

$$\mathscr{S}(\mathscr{A}) := \mathscr{S}(\mathscr{A}, \mathbb{1}) = \{f \in \mathscr{A}' : f(\mathbb{1}) = \|f\| = 1\}.$$

These are nonempty sets due to the Hahn-Banach theorem.

Definition 3.2. If $a \in \mathscr{A}$ and $x \in S(\mathscr{A})$ then

$$V(\mathscr{A}, a, x) := \{f(ax) : f \in \mathscr{S}(\mathscr{A}, x)\},$$

and let us define the *algebraic numerical range* of a as

$$V(\mathscr{A}, a) := \bigcup_{x \in S(\mathscr{A})} V(\mathscr{A}, a, x) = \{f(ax) : x \in S(\mathscr{A}), f \in \mathscr{S}(\mathscr{A}, x)\},$$

and the *algebraic numerical radius* of a as

$$v(a) := \sup\{|z| : z \in V(\mathscr{A}, a)\}.$$

The algebraic numerical range is also called the *algebra numerical range* or the *intrinsic numerical range*. In this chapter the algebraic numerical range (radius) will be simply called numerical range (radius). Also in this chapter the notation for algebraic numerical range will be simplified to $V(a)$ when it does not cause confusion.

Remark. Similarly to the Hilbert space case, for $a, b, \mathbb{1} \in \mathscr{A}$ and $z \in \mathbb{C}$

$$V(a+b) \subseteq V(a) + V(b),$$
$$V(za + w\mathbb{1}) = zV(a) + \{w\}.$$

Lemma 3.3. $V(a) = V(\mathscr{A}, a, \mathbb{1}) = \{f(a) : f \in \mathscr{S}(\mathscr{A})\}$.

Proof. $V(\mathscr{A}, a, \mathbb{1}) \subseteq V(a)$ because $\mathbb{1} \in S(\mathscr{A})$. The other way around, if $z \in V(a)$ then $z = f_1(ax_1)$ for some $x_1 \in S(\mathscr{A})$ and $f_1 \in \mathscr{S}(\mathscr{A}, x_1)$. Let us define $f_2 \in \mathscr{A}'$ as $f_2(x) = f_1(xx_1)$. Defined this way $f_2 \in \mathscr{S}(\mathscr{A})$, and $z = f_2(a) = f_2(a\mathbb{1}) \in V(\mathscr{A}, a, \mathbb{1})$. □

Corollary 3.4. *For any $a \in \mathscr{A}$ it is true that $v(a) \leq \|a\|$.*

Theorem 3.5. *$V(a) \subseteq \mathbb{C}$ is a compact convex set for every $a \in \mathscr{A}$.*

Proof. Let $B = \{f \in \mathscr{A}' : \|f\| \leq 1\}$ be the closed unit ball in the dual space. According to the Banach-Alaoglu theorem this is compact in the weak-* topology. Because $\|\mathbb{1}\| = 1$, $\mathscr{S}(\mathscr{A}) = \{f \in \mathscr{A}' : \|f\| \leq 1 \text{ and } f(\mathbb{1}) = 1\}$. $\mathscr{S}(\mathscr{A}) \subseteq B$ is convex, and a closed subset, so it is compact as well. The evaluating functional $F_a : \mathscr{A}' \to \mathbb{C}$, $F_a(f) = f(a)$ is a weak-* continuous linear mapping. Because $V(a) = F_a(\mathscr{S}(\mathscr{A}))$, it is compact and convex too. □

It is an important property of the algebraic numerical range that it is invariant under restriction to a subalgebra.

Theorem 3.6. *If $\mathscr{B} \subseteq \mathscr{A}$ is a unital subalgebra, then $V(\mathscr{B}, b) = V(\mathscr{A}, b)$ for every $b \in \mathscr{B}$.*

Proof. By the Hahn-Banach theorem every $f \in \mathscr{S}(\mathscr{B})$ must be a restriction of some $g \in \mathscr{S}(\mathscr{A})$. So $V(\mathscr{A}, b, \mathbb{1}) = V(\mathscr{B}, b, \mathbb{1})$ and by Lemma 3.3 $V(\mathscr{B}, b) = V(\mathscr{A}, b)$. □

Corollary 3.7. *$V(\mathscr{A}, a) = V(\hat{\mathscr{A}}, a)$ where $\hat{\mathscr{A}}$ is the completion of $\mathscr{A}$. Furthermore $V(\mathscr{A}, a) = V(\mathscr{B}, a)$ where $\mathscr{B} \subseteq \mathscr{A}$ is the subalgebra generated by a and $\mathbb{1}$.*

Because of this property we can assume that $\mathscr{A}$ is complete, that it is a (complex, unital) Banach algebra.

Theorem 3.8 (Lumer). *If $a \in \mathscr{A}$, then*

$$\max\{\operatorname{Re} z : z \in V(a)\} = \inf_{t>0} \frac{1}{t}(\|\mathbb{1} + ta\| - 1) = \lim_{t\to 0+} \frac{1}{t}(\|\mathbb{1} + ta)\| - 1).$$

Proof. $\mu := \max\{\operatorname{Re} z : z \in V(a)\}$. If $f \in \mathscr{S}(\mathscr{A})$ and $t > 0$ we can write $f(a)$ as

$$f(a) = \frac{1}{t}(f(\mathbb{1} + ta) - 1).$$

Because $\operatorname{Re} f(a) \le |f(a)| \le \|a\|$

$$\operatorname{Re} f(a) \le \frac{1}{t}(\|\mathbb{1} + ta\| - 1)$$

and

$$\mu \le \inf_{t>0} \frac{1}{t}(\|\mathbb{1} + ta\| - 1).$$

Let us suppose $a \neq 0$ and $0 < t < \|a\|^{-1}$. If $x \in S(\mathscr{A})$ and $f \in \mathscr{S}(\mathscr{A}, x)$ then

$$\|(\mathbb{1} - ta)x\| \ge \operatorname{Re} f((\mathbb{1} - ta)x) \ge 1 - t\mu \ge 1 - t\|a\| > 0.$$

So for any $x \in \mathscr{A}$ we have $\|(\mathbb{1} - ta)x\| \ge (1 - t\mu)\|x\|$. Let us choose $x = \mathbb{1} + ta$ so that

$$\|(\mathbb{1} - ta)\| \le \frac{\|\mathbb{1} - t^2a^2\|}{1 - t\mu} \le \frac{1 + t^2\|a^2\|}{1 - t\mu}.$$

Therefore if $0 < t < \|a\|^{-1}$,

$$\frac{\|\mathbb{1} + ta\|}{t} \le \frac{1 + t^2\|a^2\|}{t(1 - t\mu)} = \frac{1 + t^2\|a^2\| + t\mu - t\mu}{t(1 - t\mu)} = \frac{\mu + t\|a^2\|}{1 - t\mu} + \frac{1}{t}.$$

Putting it all together

$$\mu \le \inf_{t>0} \frac{1}{t}(\|\mathbb{1} + ta\| - 1) \le \lim_{t\to 0+} \frac{1}{t}(\|\mathbb{1} + ta)\| - 1) \le \lim_{t\to 0+} \frac{\mu + t\|a^2\|}{1 - t\mu} = \mu.$$

□

Theorem 3.9. *If $\mathscr{A}$ is a Banach algebra, then* $\operatorname{Sp}(a) \subseteq V(a)$ *for every $a \in \mathscr{A}$.*

Proof. Let us pick any $\lambda \in \operatorname{Sp}(a)$. Then $\lambda\mathbb{1} - a \notin G(\mathscr{A})$ and let us suppose that the left inverse does not exist. In this case $\mathscr{I} = \mathscr{A}(\lambda\mathbb{1} - a)$ is a proper left ideal. By the Hahn-Banach theorem for the vector $\mathbb{1}$ there exists a functional $f \in \mathscr{A}'$ with the properties $f(\mathbb{1}) = 1$, $\|f\| = 1$, and $f(\mathscr{I}) = \{0\}$. This means that $f \in \mathscr{S}(\mathscr{A})$, $f(\lambda\mathbb{1} - a) = 0$, and $\lambda = f(a) \in V(a)$. □

In Banach algebras the spectral radius has a minimality property: if $\mathscr{P}$ is the set of algebra norms on $\mathscr{A}$ which are equivalent to $\|.\|$ and for which $p(\mathbb{1}) = 1$ if $p \in \mathscr{P}$, then

$$r(a) = \inf_{p \in \mathscr{P}} p(a).$$

The convex hull of the spectrum can be described in terms of this set $\mathscr{P}$ and the numerical range. We will use the statement of the following lemma (its proof can be found in [2] as Lemma 2.7).

Lemma 3.10. *For every bounded semigroup $S \subseteq \mathscr{A}$ (with respect to the algebra multiplication) there is a $p \in \mathscr{P}$ norm such that $p(s) \leq 1$ for every $s \in S$.*

Lemma 3.11. *If $a_1, ..., a_n \in \mathscr{A}$ are mutually commuting elements and $\varepsilon > 0$, then there is such a $p \in \mathscr{P}$ that $p(a_k) < r(a_k) + \varepsilon$ $(k = 1, ..., n)$.*

Proof. Let S be the multiplicative semigroup generated by $b_1, ..., b_n$ where

$$b_k = \frac{1}{r(a_k) + \varepsilon} a_k.$$

These $b_1, ..., b_n$ are mutually commuting too and $r(b_k) < 1$, so $S \subseteq \mathscr{A}$ is a bounded semigroup and we can use the previous lemma. There is a $p \in \mathscr{P}$ such that $p(b_k) \leq 1$ and $p(a_k) \leq r(a_k) + \varepsilon$. □

Theorem 3.12. *Let V_p be the numerical range when the Banach algebra is equipped with the algebra norm $p \in \mathscr{P}$. Then for $a \in \mathscr{A}$*

$$\operatorname{conv} \operatorname{Sp}(a) = \bigcap_{p \in \mathscr{P}} V_p(a).$$

Proof. It follows from Theorems 3.5 and 3.9 that $\operatorname{conv} \operatorname{Sp}(a) \subseteq \bigcap\{V_p(a) : p \in \mathscr{P}\}$. $\operatorname{Sp}(a)$ is compact, so $\operatorname{conv} \operatorname{Sp}(a)$ is a compact convex set and it is equal to the intersection of all the open discs that contain $\operatorname{Sp}(a)$. Let us suppose that one of these discs is $B_t(z_0) = \{z \in \mathbb{C} : |z - z_0| < t\}$. Then $\operatorname{Sp}(a) \subseteq B_t(z_0)$ and $\operatorname{Sp}(a - z_0\mathbb{1}) \subseteq B_t(0)$. So $r(a - z_0) < t$ and by Lemma 3.11 there is a $p \in \mathscr{P}$ for which $p(a - z_0\mathbb{1}) < t$. If $z \in V_p(a)$, there is a $f \in \mathscr{A}'$ such that $z = f(a)$ and

$$|z - z_0| = |f(a) - z_0| = |f(a - z_0\mathbb{1})| \leq p(a - z_0\mathbb{1}) < t.$$

So we have that $V_p(a) \subseteq B_t(z_0)$. This is true for any open disc that contains $\operatorname{Sp}(a)$, therefore $\bigcap\{V_p(a) : p \in \mathscr{P}\} \subseteq \operatorname{conv} \operatorname{Sp}(a)$. □

3.2 The Exponential Function

The holomorphic functional calculus makes it possible to define the exponential on a Banach algebra. This exponential function will provide a useful characterization of the numerical range.

Definition 3.13. If $a \in \mathscr{A}$, then the *exponential* of a is defined as

$$\exp(a) := \mathbb{1} + \sum_{n=1}^{\infty} \frac{1}{n!} a^n.$$

Remark. If $a, b \in \mathscr{A}$ are commuting elements, then the $\exp : \mathscr{A} \to \mathscr{A}$ function has the following properties:

(1) $\exp(a + b) = \exp(a)\exp(b)$,

(2) $\exp(a) \in G(\mathscr{A})$,

(3) $\exp(a)\exp(-a) = \mathbb{1}$,

(4) $\exp(a) = \lim\limits_{n\to\infty} \left(\mathbb{1} + \frac{1}{n} a\right)^n$.

The next theorem is an analogue of Theorem 3.8 phrased in terms of exponentials.

Theorem 3.14. *If $a \in \mathscr{A}$, then*

$$\max\{\operatorname{Re} z : z \in V(a)\} = \sup_{t>0} \frac{1}{t} \log \|\exp(ta)\| = \lim_{t\to 0+} \frac{1}{t} \log \|\exp(ta)\|.$$

Proof. Let us define μ as $\max\{\operatorname{Re} V(a)\}$, and suppose that $t > 0$. The same way as in the proof of Theorem 3.8 for $f \in \mathscr{S}(\mathscr{A}, x)$ and $x \in \mathscr{A}$

$$\|(\mathbb{1} - ta)x\| \geq (1 - t\mu)\|x\|.$$

If $1 - t\mu \geq 0$ then we can use induction to get

$$\|(\mathbb{1} - ta)^n x\| \geq (1 - t\mu)^n \|x\| \quad (n = 1, 2, ...).$$

For sufficiently large n we have $t\mu \leq n$, so

$$\|(\mathbb{1} + \frac{1}{n}(-ta))^n x\| \geq (1 + \frac{1}{n}(-t\mu))^n \|x\|.$$

If $n \to \infty$ the two sides of the above expression converge to

$$\|\exp(-ta)x\| \geq \exp(-t\mu)\|x\|.$$

With the choice of $x = \exp(ta)$ we get $\|\exp(ta)\| \leq \exp(t\mu)$ and

$$\sup\{\frac{1}{t}\log\|\exp(ta)\| : t > q\} \leq \mu.$$

Using the power series expansion of the exponential function $\|\exp(ta)\| = \|\mathbb{1} + ta\| + f(t)$ where $|f(t)| \leq ct^2$ for some $c > 0$ if $0 \leq t \leq 1$. In general $\log t \geq \frac{y-1}{y}$ if $y > 0$, so if we choose $y = \|\exp(ta)\|$

$$\log\|\exp(ta)\| \geq \frac{\|\exp(ta)\| - 1}{\|\exp(ta)\|} = \frac{\|\mathbb{1} + ta\| + f(t) - 1}{\|\mathbb{1} + ta\| + f(t)}.$$

If we divide by t, the right hand side of the inequality will converge to μ

$$\frac{1}{t}\log\|\exp(ta)\| \geq \frac{\frac{1}{t}(\|\mathbb{1} + ta\| - 1) + \frac{f(t)}{t}}{\|\mathbb{1} + ta\| + f(t)} \to \mu \quad (t \to 0+),$$

because by the proof of Theorem 3.8 $\frac{1}{t}(\|\mathbb{1} + ta\| - 1)$ converges to μ as t goes to $0+$. □

Theorem 3.16 is a version of Theorem 3.14 for the spectrum of an element. Its proof will use the following auxiliary lemma.

Lemma 3.15. *If $f : \mathbb{R} \to \mathbb{R}_+$ is continuous and subadditive, then*

$$\lim_{t\to\infty} \frac{f(t)}{t} = \inf_{t>0} \frac{f(t)}{t}.$$

Proof. For every α that is greater than $\inf\{\frac{1}{t}f(t) : t > 0\}$ we can pick a number $s > 0$ for which $\frac{1}{s}f(s) < \alpha$. Because f is continuous the set below is bounded:

$$\sup\{f(t) : s \leq t \leq 2s\} = m < \infty.$$

If it is true for a positive integer n and a positive real t that $(n+1)s \leq t \leq (n+2)s$, then

$$f(t) \leq f(ns) + f(t - ns) \leq nf(s) + m,$$

and

$$\frac{f(t)}{t} < \frac{ns}{t}\alpha + \frac{m}{t}.$$

If $t \to \infty$ and $n \to \infty$ while $(n+1)s \leq t \leq (n+2)s$ is still true, we get that

$$\inf_{t\to\infty} \frac{f(t)}{t} \leq \liminf_{t\to\infty} \frac{f(t)}{t} \leq \limsup_{t\to\infty} \frac{f(t)}{t} \leq \alpha.$$

This is true for arbitrary α greater than the infimum, so the limit and the infimum is the same. □

Theorem 3.16. *For every $a \in \mathscr{A}$*

$$\max\{\operatorname{Re}\lambda : \lambda \in \operatorname{Sp}(a)\} = \inf_{t>0} \frac{1}{t} \log \| \exp(ta)\| = \lim_{t\to\infty} \frac{1}{t} \log \| \exp(ta)\|.$$

Proof. Because $\exp(a)$ is always invertible, $r(\exp(a)) > 0$, and by using Lemma 3.15

$$\log r(\exp(a)) = \log \lim_{n\to\infty} \| \exp(a)^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \frac{1}{n} \log \| \exp(na)\| = \inf_{t>0} \frac{1}{t} \log \| \exp(ta)\|.$$

At the same time we have

$$r(\exp(a)) = \max_{\lambda\in\operatorname{Sp}(a)} |\exp(\lambda)| = \max_{\lambda\in\operatorname{Sp}(a)} \{\exp(\operatorname{Re}\lambda)\} = \exp(\max\{\operatorname{Re}\operatorname{Sp}(a)\}).$$

because according to Theorem 1.1 $\operatorname{Sp}(\exp(a)) = \exp(\operatorname{Sp}(a))$. □

Corollary 3.17. *For $a \in \mathscr{A}$ the spectral radius can be expressed as*

$$r(a) = \max_{|z|=1} \lim_{t\to\infty} \frac{1}{t} \log \| \exp(tza)\|$$

where $z \in \mathbb{C}$ and $t > 0$.

We can use the algebra exponential to prove the following version of the Hirschfeld-Zelazko theorem. (This is somewhat weaker than the standard version, which can be found for example in [13] as Proposition 3.1.8).

Theorem 3.18. *If there is $k > 0$ such that $r(a) \geq k\|a\|$ for every $a \in \mathscr{A}$, then $\mathscr{A}$ is a commutative Banach algebra.*

Proof. For $a, b \in \mathscr{A}$ let us define the function $\Phi : \mathbb{C} \to \mathscr{A}$ as

$$\Phi(z) = \exp(-za) b \exp(za).$$

Because $\operatorname{Sp}(\Phi(z)) = \operatorname{Sp}(b)$

$$k\|\Phi(z)\| \leq r(\Phi(z)) \leq r(b) \quad (z \in \mathbb{C}).$$

Let us choose an $f \in \mathscr{A}'$ such that $\|f\| \leq 1$. Then $f \circ \Phi : \mathbb{C} \to \mathbb{C}$ is a bounded entire function. According to Liouville's theorem this must be constant. Moreover by the Hahn-Banach theorem Φ is constant and $\Phi(1) = \Phi(0)$. So we have

$$\exp(a)b = b \exp(a).$$

If we replace a with λa ($\lambda \in \mathbb{C}$), expand the exponentials into their power series, and equate the coefficients of λ, we get that $ab = ba$. □

3.3 Numerical Radius

The numerical radius in Banach algebras retains the property that it is equivalent to the norm. The lower bounding constant $\frac{1}{2}$ of the Hilbert space case will be replaced with $\frac{1}{e}$. The range of possible bounding constants will be discussed in the next chapter, in the section about the numerical index. As a further similarity to Hilbert spaces, the numerical radius equals the norm for self-adjoint elements. This will be one of the topics of the next section about Hermitian elements. On the other hand the power inequality does not apply to the algebraic numerical radius.

Theorem 3.19. *The numerical radius is equivalent to the algebra norm, in particular for every $a \in \mathscr{A}$*

$$\frac{1}{e}\|a\| \leq v(a) \leq \|a\|.$$

Proof. It is a consequence of Lemma 3.3 that $v(a) \leq \|a\|$. We need to show that $\|a\| \leq e \cdot v(a)$. Let us choose a $b \in \mathscr{A}$ in such a way that for some $\mu < 1$ bound $v(b) \leq \mu$. For some $x \in S(\mathscr{A})$, $f \in \mathscr{S}(\mathscr{A}, x)$ and $z \in \mathbb{C}$ with $|z| \leq 1$

$$\|(\mathbb{1} - zb)x\| \geq |f((\mathbb{1} - zb)x)| = |1 - zf(bx)| \geq |1 - |z||f(bx)|| \geq 1 - \mu.$$

We used that $f(bx) \in V(b)$ so $|f(bx)| \leq v(a) \leq \mu$. For any $x \in \mathscr{A}$ and $|z| \leq 1$

$$\|(\mathbb{1} - zb)x\| \geq (1 - \mu)\|x\|.$$

By Theorem 3.9 $r(b) \leq v(b) \leq \mu < 1$ and if $|z| \leq 1$ then $|1/z| \geq r(b)$. Therefore $\frac{1}{z}\mathbb{1} - b$ is invertible, so $\mathbb{1} - zb$ is invertible too. If we choose $x = (\mathbb{1} - zb)^{-1}$ then

$$\|(\mathbb{1} - zb)^{-1}\| \leq \frac{1}{1 - \mu} \quad (|z| \leq 1).$$

Let $\omega_1, ..., \omega_n$ be the n-th roots of unity. They have the property that for any j integer

$$\sum_{i=1}^{n} \omega_i^j = \begin{cases} 0 & \text{if } j \not\equiv 0 \pmod{n} \\ n & \text{if } j \equiv 0 \pmod{n}. \end{cases}$$

Let us define for $k = 1, 2, ...$ the sum

$$S(k, n) = \frac{1}{n} \sum_{i=1}^{n} \omega_i^{-1} (\mathbb{1} - \omega_i b)^{-k}.$$

Expanding $(\mathbb{1} - \omega_i b)^{-k}$ into a power series

$$\omega_i^{-1}(\mathbb{1} - \omega_i b)^{-k} = \omega_i^{-1} + kb + \frac{k(k+1)}{2!}\omega_i b^2 + \dots$$

so $S(k,n)$ can be rewritten as

$$S(k,n) = kb + \frac{k(k+1)...(k+n)}{(n+1)!} b^{n+1} + \dots$$

We have that $r(b)^n \to 0$ because $r(b) < 1$, $r(b^n) \to 0$ because $r(b^n) \le r(b)^n$, so $b^n \to 0$ as $n \to \infty$. Therefore in the limit

$$\lim_{n\to\infty} S(k,n) = kb.$$

Using that $\|(\mathbb{1} - \omega_i b)^{-1}\| \le 1/(1-\mu)$ we get

$$\|S(k,n)\| \le \frac{1}{n} \sum_{i=1}^{n} |\omega_i| \|(\mathbb{1} - \omega_i b)^{-1}\|^k \le \frac{1}{(1-\mu)^k}.$$

Putting these results together $k\|b\| \le (1-\mu)^{-k}$. To finish the proof, let us pick an arbitrary element $a \in \mathscr{A}$. If $k \ge 2$ let us choose

$$\mu = \frac{1}{k} \text{ and } b = \frac{\mu}{v(a)} a.$$

This way $v(b) = \mu < 1$ and for every $k \ge 2$

$$\frac{1}{v(a)} \|a\| \le \left(1 - \frac{1}{k}\right)^{-k}.$$

Therefore

$$v(a) \ge \|a\| \lim_{k\to\infty} \left(1 - \frac{1}{k}\right)^{k} = \|a\| \frac{1}{e}.$$

□

Now we have the Banach algebra analogues of spectral inclusion (Theorems 2.5 and 3.9) and the equivalence of the numerical radius to the norm (Theorems 2.7 and 3.19). Therefore we can re-state Corollary 2.8 for Banach algebras.

Corollary 3.20. *For every $a \in \mathscr{A}$ element $r(a) \le v(a) \le \|a\|$.*

Corollary 3.21. *If either $r(a) = \|a\|$ or $r(a) = v(a)$ for every $a \in \mathscr{A}$, then $\mathscr{A}$ is a commutative algebra, and the spectral radius is an algebra norm equivalent to the given norm.*

Proof. In the first case $r(a) \ge 1 \cdot \|a\|$, in the second case by Theorem 3.19 $r(a) \ge \frac{1}{e}\|a\|$, so by Theorem 3.18 $\mathscr{A}$ is commutative. Because every element commutes, $r(a+b) \le r(a) + r(b)$ and $r(ab) \le r(a)r(b)$ for every $a, b \in \mathscr{A}$, so r is an algebra norm. □

The algebraic numerical radius is not submultiplicative. In a general normed algebra even the power inequality is false. It is possible to show a complex unital Banach algebra $\mathscr{A}$ and $0 \neq a \in \mathscr{A}$ such that

$$\|a^n\| = n! \left(\frac{e}{n}\right)^n v(a)^n \quad (n = 1, 2, ...).$$

On the other hand the power inequality is true in C^*-algebras ([1] Corollaries 23.10 and 24.10).

3.4 Hermitian Elements

Even if no involution is defined on a Banach algebra, it is possible to define a property similar to self-adjointness in terms of the numerical range. An element $h \in \mathscr{A}$ will be called Hermitian if its numerical range is real. We will see that in C^*-algebras the two notions actually coincide. Furthermore positive and normal elements can be defined in a similar way, without making reference to an involution.

Definition 3.22. An element $h \in \mathscr{A}$ is *Hermitian* if $V(h) \subseteq \mathbb{R}$. $H(\mathscr{A})$ denotes the set of Hermitian elements in $\mathscr{A}$.

Lemma 3.23. *For $h \in \mathscr{A}$ the following characterizations are equivalent:*

(i) $V(h) \subseteq \mathbb{R}$,

(ii) $\lim_{t\to 0} \frac{1}{t}(\|1 + ith\| - 1) = 0 \quad (t \in \mathbb{R})$,

(iii) $\|\exp(ith)\| = 1 \quad (t \in \mathbb{R})$.

Proof. Phrased in terms of maxima $h \in \mathscr{A}$ is Hermitian if and only if

$$\max\{\operatorname{Re} V(ih)\} = \max\{\operatorname{Re} V(-ih)\} = 0.$$

We get the second property by applying Theorem 3.8 and the third by Theorem 3.14. □

Proposition 3.24. *If $\mathscr{A}$ is a C^*-algebra, then an element is Hermitian if and only if it is self-adjoint.*

Proof. First let us suppose that $h = h^*$ for $h \in \mathscr{A}$. Then for $t \in \mathbb{R}$

$$\|\mathbb{1} + t^2h^2\| = \|(\mathbb{1} + ith)(\mathbb{1} - ith)\| = \|(\mathbb{1} + ith)(\mathbb{1} + ith)^*\| = \|\mathbb{1} + ith\|^2.$$

We can get that $h \in H(\mathscr{A})$ by the limit characterization:

$$\lim_{t\to 0} \frac{1}{t}(\|\mathbb{1} + ith\| - 1) = \lim_{t\to 0} \frac{1}{t}(\sqrt{\|\mathbb{1} + t^2h^2\|} - 1) = 0,$$

because

$$\left|\frac{\sqrt{\|\mathbb{1}+t^2h^2\|}-1}{t}\right| = \left|\frac{\|\mathbb{1}+t^2h^2\|-1}{t(\sqrt{\|\mathbb{1}+t^2h^2\|}+1)}\right| \leq \left|\frac{\|t^2h^2\|+\|\mathbb{1}\|-1}{t(\sqrt{\|\mathbb{1}+t^2h^2\|}+1)}\right| = \frac{|t|\|h^2\|}{\sqrt{\|\mathbb{1}+t^2h^2\|}+1}.$$

Conversely, let us suppose that $h \in H(\mathscr{A})$. We can decompose h as $a + ib$ where $a = \frac{1}{2}(h+h^*)$, $b = \frac{1}{2i}(h-h^*)$, so $a = a^*$, $b = b^*$. Then the numerical range can be written as

$$V(h) = \{f(a) + if(b) : f \in \mathscr{S}(\mathscr{A})\}.$$

For every $f \in \mathscr{S}(\mathscr{A})$ $f(b) = 0$ because a, b, and h are all Hermitian. This means that $V(b) = 0$ and $b = 0$. □

Lemma 3.25. *$H(\mathscr{A})$ is a real Banach space, and $i(hk-kh) \in H(\mathscr{A})$ if $h,k \in H(\mathscr{A})$.*

Proof. As a consequence of part (iii) of Lemma 3.23 $\|a\exp(ith)\| \leq \|a\|$. This implies that for $s \in \mathbb{R}$

$$\|\exp(isk)\exp(ish)\exp(-isk)\exp(-ish)\| = 1.$$

By expanding the power series of the exponentials we get that there is constant C such that

$$|\|\mathbb{1} - s^2(hk-kh)\| - 1| \leq Cs^3 \quad (0 \leq s \leq 1).$$

Substituting t for s^2

$$\lim_{t\to 0}\frac{1}{t}(\|\mathbb{1} + it(i(hk-kh))\| - 1) = 0.$$

By part (ii) of 3.23 $i(hk-kh)$ is Hermitian. □

In certain ways the elements of $H(\mathscr{A})$ are analogous to the real numbers in the complex field. This motivates the idea to use Hermitian elements in a similar role as the real and imaginary parts of complex numbers. Let us define the set

$$J(\mathscr{A}) := \{h + ik : h,k \in H(\mathscr{A})\}.$$

Even if $\mathscr{A}$ is not a *-algebra, it is possible to introduce an involution on $J(\mathscr{A})$. The following lemma is necessary to make sure that this involution will be a well-defined operation.

Lemma 3.26. *Every $a \in J(\mathscr{A})$ is uniquely represented by a pair of elements $h,k \in H(\mathscr{A})$ such that $a = h + ik$.*

Proof. Because $H(\mathscr{A})$ is a real vector space, it is sufficient that if $h,k \in H(\mathscr{A})$ and $h+ik = 0$ then $h = k = 0$. If $h = -ik$, then $V(h) = -iV(k)$. $V(h) \subseteq \mathbb{R}$ and $-iV(k) \subseteq i\mathbb{R}$, so $V(h) = -iV(k) \subseteq \mathbb{R} \cap i\mathbb{R} = \{0\}$. Because of Theorem 3.19 the numerical range of an element is the set containing only zero if and only if the element is zero, therefore $h = k = 0$. □

Proposition 3.27. *$J(A)$ is complex Banach space with norm of $\mathscr{A}$, and the natural involution $(h+ik)^* = h - ik$ is continuous ($(h, k \in H(\mathscr{A}))$).*

Proof. Suppose $h, k \in \mathscr{A}$. Then the numerical range of h and k is real, therefore

$$v(h) \leq v(h+ik) \leq \|h+ik\| \text{ and } v(k) \leq v(h+ik) \leq \|h+ik\|.$$

According to the norm equivalence in Theorem 3.19

$$\|h\| \leq e\|h+ik\| \text{ and } \|k\| \leq e\|h+ik\|.$$

The involution is continuous because

$$\|(h+ik)^*\| \leq \|h\| + \|k\| \leq 2e\|h+ik\|.$$

To show the closedness of $J(\mathscr{A})$, let us suppose that $(h_n) + i(k_n)$ is a Cauchy sequence in $J(\mathscr{A})$. Since $\|h\|, \|k\| \leq e\|(h+ik)\|$, (h_n) and (k_n) are Cauchy sequences too. By Lemma 3.25 $H(\mathscr{A})$ is closed, so (h_n) and (k_n) converges. So $(h_n) + i(k_n)$ is convergent as well. □

As a consequence of Proposition 3.24, for every C^*-algebra $\mathscr{A} = J(\mathscr{A})$. More surprisingly the converse is true as well, so $J(\mathscr{A})$ provides a way to characterize C^*-algebras among Banach algebras. A relatively short proof is described in [16].

Theorem 3.28 (Vidav-Palmer)**.** *If every element of complex unital Banach algebra $\mathscr{A}$ can be written in the form $h + ik$ where $h, k \in \mathscr{A}$ are Hermitian, then $\mathscr{A}$ is a C^*-algebra with the natural involution.*

We can prove an algebra version of Theorem 2.27 (for normal Hilbert space operators the closure of the numerical range is the convex hull of the spectrum). The proof uses Vidav's lemma, which states that for Hermitian elements the algebraic numerical range equals the spectral radius. The Vidav lemma could be diretly proved, but A. M. Sinclair provided a stronger result, that for Hermitian elements the spectral radius equals the algebra norm [15]. The Vidav lemma is a consequence of Sinclair's theorem by Corollary 3.20. The following simplified proof of Sinclair's theorem is from [1] (Lemma 26.1 and Theorem 26.2).

Lemma 3.29. *Let us suppose that the sequence $(c_k) \subseteq \mathbb{R}_+$ provides the coefficients of a power series expansion of the inverse sine function,*

$$\arcsin(t) = \sum_{k=1}^{\infty} c_k t^k \quad (|t| \leq 1),$$

and let us define the function series

$$F_n(z) = \sum_{k=1}^{n} c_k (\sin z)^k \quad (z \in \mathbb{C}).$$

If $K \subseteq (-\frac{\pi}{2}, \frac{\pi}{2})$ is compact, then there is an open subset $U \subseteq \mathbb{C}$ such that $K \subseteq U$ and $F_n(z) \to z$ uniformly on U.

Proof. It is possible to find an open set $U \subseteq \mathbb{C}$ which contains K, and for which $|\sin z| \leq 1$ if $z \in U$. Hence we can apply the Weierstrass M-test to show that $F_n(z)$ converges uniformly on U, and because of this the limit function is analytic on U. For real $t \in K$ we have $F_n(t) \to t$, therefore $F_n(z) \to z$ if $z \in U$. □

Theorem 3.30 (Sinclair)**.** *If $h \in \mathscr{A}$ is Hermitian, then $r(h) = \|h\|$.*

Proof. Because the norm and the spectral radius are both absolute homogeneous, we can suppose that $r(h) < \frac{\pi}{2}$. In this case $\mathrm{Sp}(h) \subseteq (-\frac{\pi}{2}, \frac{\pi}{2})$, and by Lemma 3.29 and the functional calculus of entire functions

$$h = \sum_{k=1}^{\infty} c_k (\sin h)^k,$$

where (c_k) is defined as in Lemma 3.29. According to Lemma 3.23

$$\|\sin h\| \leq \frac{1}{2}\|\exp(ih)\| + \frac{1}{2}\|\exp(-ih)\| = 1.$$

Therefore the norm of h can bounded with

$$\|h\| \leq \sum_{k=1}^{\infty} c_k = \frac{\pi}{2}.$$

This is true for any $r(h) < \frac{\pi}{2}$, so $r(h)$ equals $\|h\|$. □

Corollary 3.31 (Vidav lemma)**.** *If $h \in \mathscr{A}$ is Hermitian, then $v(a) = r(a)$, so*

$$\max V(h) = \max \mathrm{Sp}(h).$$

Corollary 3.32. *If $h \in \mathscr{A}$ is Hermitian, then $V(h) = \mathrm{conv}\,\mathrm{Sp}(h)$.*

Proof. By the Vidav lemma

$$\min V(h) = -\max V(-h) = -\max \mathrm{Sp}(-h) = \min \mathrm{Sp}(h).$$

This is enough, since $\mathrm{Sp}(h) \subseteq V(h) \subseteq \mathbb{R}$, and $V(h)$ is compact and convex. □

There is an alternate definition of normality in terms of Hermitian elements. In Banach *-algebras this coincides with the usual definition.

Definition 3.33. An element $a \in \mathscr{A}$ is *normal* if $a = h + ik$ for some $h, k \in H(\mathscr{A})$ and $hk = kh$.

For normal elements we can derive the algebra counterpart of Theorem 2.27.

Theorem 3.34. *If $a \in \mathscr{A}$ is normal, then $V(a) = \operatorname{conv} \operatorname{Sp}(a)$.*

Proof. We know by Theorems 3.5 and 3.9 that $\operatorname{conv} \operatorname{Sp}(a) \subseteq V(a)$. Because the convex hull is the intersection of those closed half-planes that contain the spectrum, for the reverse inclusion we will show that if a half-plane contains $\operatorname{Sp}(a)$, then it contains $V(a)$ as well. Since a is normal, $a = h + ik$ with $h, k \in H(\mathscr{A})$ and $hk = kh$. Let $\mathscr{B}$ denote the maximal commutative subalgebra of $\mathscr{A}$ which contains both h and k, furthermore

$$\operatorname{Sp}_{\mathscr{B}}(a) := \{\lambda \in \mathbb{C} : a - \lambda \mathbb{1} \notin G(\mathscr{B})\}.$$

Because $\mathscr{B}$ is a maximal commutative subalgebra, $\operatorname{Sp}(a) = \operatorname{Sp}_{\mathscr{B}}(a)$. Let $X(\mathscr{B}) \subseteq \mathscr{B}'$ be the characters of $\mathscr{B}$, that is the set of all nonzero multiplicative functionals on $\mathscr{B}$. It is a property of commutative Banach-algebras that

$$\operatorname{Sp}_{\mathscr{B}}(a) = \{\chi(a) : \chi \in X(\mathscr{B})\}.$$

Therefore $\operatorname{Sp}(a) = \{\chi(h) + i\chi(k) : \chi \in X(\mathscr{B})\}$ and $\operatorname{Sp}(h) = \{\chi(h) : \chi \in X(\mathscr{B})\}$. As h is Hermitian, $\operatorname{Sp}(h) \subseteq \mathbb{R}$. Let us suppose that $\operatorname{Sp}(a)$ is in the right half-plane, that $\operatorname{Re} \operatorname{Sp}(a) \subseteq \mathbb{R}_+$. In this case $\chi(h) \in \mathbb{R}_+$ for every $\chi \in X(\mathscr{B})$, so $\operatorname{Sp}(h) \subseteq \mathbb{R}_+$. By Corollary 3.31 $V(h) \subseteq \mathbb{R}_+$, therefore $\operatorname{Re} V(A) \subseteq \mathbb{R}_+$. When a is normal, $\lambda a + \mu \mathbb{1}$ is normal as well, so the containment is true for every other half-plane too. □

Corollary 3.35. *If $a \in \mathscr{A}$ is normal, then $r(a) = v(a)$.*

Chapter 4

Numerical Range in Normed Spaces

4.1 Spatial Numerical Range

Let us suppose there is a normed vector space X, and a bounded linear operator T acting on X. One way to define the numerical range of T is to treat it as an element of the unital normed algebra $\mathscr{B}(X)$. In this case we have the algebraic numerical range $V(\mathscr{B}(X), T)$ with all the properties discussed in the previous chapter. A different definition is possible, the spatial numerical range. This definition does not use the dual of $\mathscr{B}(X)$, only X itself and its dual space X'.
Let $S(X)$ be the unit sphere in X, and $S(X')$ the unit sphere in X' by the dual norm

$$\|f\| = \sup\{|f(x)| : \|x\| \leq 1, x \in X\} \quad (f \in X').$$

Let us denote by $\Pi(X)$ the subset of the Cartesian product $X \times X'$ defined as

$$\Pi(X) := \{(x, f) : x \in S(X), f \in S(X'), f(x) = 1\}.$$

Definition 4.1. If $T \in \mathscr{B}(X)$, then its *spatial numerical range* is defined as

$$\boldsymbol{V}(T) := \{f(Tx) : (x, f) \in \Pi(X)\}.$$

If $\mathscr{S}(X, x) := \{f \in X' : \|f\| = f(x) = 1\}$ is the set of support functionals of x, then

$$\boldsymbol{V}(T) := \{f(Tx) : x \in S(X), f \in \mathscr{S}(X, x)\}.$$

If X is a Hilbert space $\mathcal{H}$, then $(x, f) \in \Pi(\mathcal{H})$ exactly when $x \in S(\mathcal{H})$ and f is in the form $f(y) = \langle y, x\rangle$ where $y \in \mathcal{H}$. So for Hilbert space operators we get back Definition 2.1, and $\boldsymbol{V}(T) = W(T)$ if $T \in \mathscr{B}(\mathcal{H})$.

Proposition 4.2. *For an operator $T \in \mathscr{B}(X)$ the spatial numerical range is contained in the algebraic numerical range.*

Proof. For every $(x, f) \in \Pi(X)$ we can define a functional $F \in \mathscr{B}(X)'$ as $F(A) = f(Ax)$ where $A \in \mathscr{B}(X)$. $F(I) = f(x) = 1$, so F is in the state space $\mathscr{B}(X)$. Therefore $f(Tx) = F(T) \in V(\mathscr{B}(X), T)$ and $\boldsymbol{V}(T) \subseteq V(\mathscr{B}(X), T)$. □

Let π_1 denote the natural projection $\pi_1(\Gamma) = \{x : (x, f) \in \Gamma\}$ where $\Gamma \subseteq \Pi(X)$.

Lemma 4.3. *Let us suppose $\Gamma \subseteq \Pi(X)$ is such that $\pi_1(\Gamma)$ is a dense subset in $S(X)$. Then for every operator $T \in \mathscr{B}(X)$*

$$\sup\{\operatorname{Re} f(Tx) : (x, f) \in \Gamma\} = \inf_{t>0} \frac{1}{t}(\|I + tT\| - 1).$$

Proof. Let μ denote the supremum, $\mu := \sup\{\operatorname{Re} f(Tx) : (x, f) \in \Gamma\}$. By the previous proposition and Theorem 3.8

$$\mu \leq \inf_{t>0} \frac{1}{t}(\|I + tT\| - 1).$$

Let us suppose that $T \neq 0$, and let us pick $0 < t < \frac{1}{\|T\|}$, $\varepsilon > 0$, $x \in S(X)$. Because of the density of $\pi_1(\Gamma)$ in $S(X)$, there must be a $(y, g) \in \Gamma$ such that $\|x - y\| < \varepsilon$. Because

$$\operatorname{Re} g(Ty) \leq \mu \leq \|T\|,$$

we have

$$\|(I - tT)y\| \geq \operatorname{Re} g((I - tT)y) = 1 - t \operatorname{Re} g(Ty) \geq 1 - t\mu > 0.$$

If we look at x instead of y, and suppose that ε is small enough that $\|y\| > \|x - y\|$, then

$$\begin{aligned}\|(I - tT)x\| &= \|(I - tT)y - (I - tT)(y - x)\| \\ &\geq \|(I - tT)y\| - \|(I - tT)(y - x)\| \\ &\geq (1 - t\mu) - \|I - tT\|\varepsilon.\end{aligned}$$

Because ε can be arbitrarily small, $\|(I - tT)x\| \geq 1 - t\mu$. Then for arbitrary $x \in X$

$$\|(I - tT)x\| \geq (1 - t\mu)\|x\|,$$

and if x is taken to be $(I + tT)x$, then

$$\|(I + tT)x\| \leq \frac{1}{1 - t\mu}\|(I - t^2T^2)x\|.$$

Therefore

$$\|I + tT\| \leq \frac{1 + t^2\|T^2\|}{1 - t\mu},$$

And the proof completes the same way as the proof of Theorem 3.8. □

Lemma 4.4. *Let us suppose that $\Gamma \subseteq \Pi(X)$ is such that $\pi_1(\Gamma)$ is a dense subset in $S(X)$. Then for every operator $T \in \mathscr{B}(X)$*

$$\overline{\text{conv}}\{f(Tx) : (x, f) \in \Gamma\} = V(\mathscr{B}(X), T).$$

Proof. According to Lemma 3.8 and Lemma 4.3

$$\sup\{\operatorname{Re} f(Tx) : (x, f) \in \Gamma\} = \sup\{\operatorname{Re} \lambda : \lambda \in V(\mathscr{B}(X), T)\}.$$

This remains true if we replace T with any zT where $z \in \mathbb{C}$. Because the algebraic numerical range is always closed and convex, it has to be the closed convex hull of the spatial range. □

Replacing Γ with $\Pi(X)$ in the previous statement gives a description of the relationship between the algebraic and the spatial numerical range: The algebraic numerical range is the closed convex hull of the spatial numerical range. Therefore there is no need to distinguish between algebraic and spatial numerical radius, the respective suprema coincide.

Theorem 4.5. *For every $T \in \mathscr{B}(X)$ operator*

(1) $\overline{\text{conv}}\, \mathbf{V}(T) = V(\mathscr{B}(X), T)$,

(2) $\sup\{|\lambda| : \lambda \in \mathbf{V}(T)\} = \sup\{|\lambda| : \lambda \in V(\mathscr{B}(X), T)\}$.

Spectral inclusion is true for the spatial numerical range too.

Theorem 4.6. *If X is a Banach space, $T \in \mathscr{B}(X)$, then $\operatorname{Sp}(T) \subseteq \overline{\mathbf{V}(T)}$.*

Proof. Let us pick a $z \in \mathbb{C}$ such that $z \notin \overline{\mathbf{V}(T)}$. Let $\varepsilon := \inf\{|z - w| : w \in V(T)\}$. Then $\varepsilon > 0$, and for every $(x, f) \in \Pi(X)$

$$\|(zI - T)x\| \geq |f((zI - T)x)| = |z - f(Tx)| \geq \varepsilon.$$

Looking at the adjoint of $zI - T$ we get $\|(zI - T)^* f\| \geq \varepsilon$. According to the Bishop-Phelps theorem if $g \in S(X')$ and $\delta > 0$, then there is a $(x, f) \in \Pi(X)$ such that $\|f - g\| < \delta$. If δ is small enough that $\|f\| > \|f - g\|$, then

$$\begin{aligned}\|(zI - T)^* g\| &= \|(zI - T)^* f - (zI - T)^*(f - g)\| \\ &\geq \|(zI - T)^* f\| - \|(zI - T)^*(f - g)\| \\ &\geq \varepsilon - \|zI - T\|\delta.\end{aligned}$$

Because δ can be arbitrarily small, for any $g \in X'$ we have $\|(zI - T)^* g\| \geq \varepsilon \|g\|$. The same way for $x \in X$ we have $\|(zI - T)x\| \geq \varepsilon \|x\|$. So $zI - T$ is bounded below, therefore it is a bijective mapping from X onto a closed subspace $Y \subseteq X$. Moreover $(zI - T)^*$ is bounded below too, so it is injective, and Y must be dense in X. Consequently $X = Y$, and by the Banach linear homeomorphism theorem $zI - T$ is invertible, so z cannot be in the spectrum. □

The following theorem will help describe the behavior of spatial numerical range with respect to adjoint operators.

Theorem 4.7. *For every $U \in \mathscr{B}(X')$ dual operator*

$$\overline{\operatorname{conv}}\{(Uf)(x) : (x, f) \in \Pi(X)\} = \overline{\operatorname{conv}}\, \mathbf{V}(U) = V(\mathscr{B}(X'), U).$$

Proof. If $x \in X, f \in X'$ then let $\hat{x} \in X''$ be defined as $\hat{x}(f) = f(x)$. By the Bishop-Phelps theorem the support functionals are dense in X', so if $\Gamma = \{(f, \hat{x}) : (x, f) \in \Pi(X)\} \subseteq \Pi(X')$, then $\pi_1(\Gamma)$ is dense in $S(X')$. Consequently

$$\overline{\operatorname{conv}}\{(Uf)(x) : (x, f) \in \Pi(X)\} = \overline{\operatorname{conv}}\{\hat{x}(Uf) : (f, \hat{x}) \in \Gamma\} = V(\mathscr{B}(X'), U)$$

because we can apply Lemma 4.4. □

Corollary 4.8. *If T^* is the adjoint of an operator $T \in \mathscr{B}(X)$, then*

(1) $\mathbf{V}(T) \subseteq \mathbf{V}(T^*)$,

(2) $\overline{\operatorname{conv}}\, \mathbf{V}(T) = \overline{\operatorname{conv}}\, \mathbf{V}(T^*)$,

(3) $v(T) = v(T^*)$.

4.2 Semi-Inner Products

Usually the term semi-inner product means a scalar product which relaxes the requirement that it should be strictly positive. In this section semi-inner product will mean the L-semi-inner product, the semi-inner product in the sense of Lumer [11]. This differs from the usual scalar product in that it is not necessarily linear in the second variable.

Definition 4.9. For a vector space X an *L-semi-inner-product* is a $[.,.] : X \times X \to \mathbb{C}$ function such that for every $x, y, z \in X$ and $\lambda \in \mathbb{C}$

(1) $[x + y, z] = [x, z] + [y, z]$,

(2) $[\lambda x, y] = \lambda\, [x, y]$,

(3) $[x,x] > 0$ if $x \neq 0$,

(4) $|\,[x,y]\,|^2 \leq [x,x]\,[y,y]$.

Proposition 4.10. *If* $[.,.]$ *is a semi-inner product on the vector space* X*, then the function* $\|x\| = \sqrt{[x,x]} \quad (x \in X)$ *is a norm on* X*.*

Proof. The triangle inequality follows from

$$\|x+y\|^2 = [x+y, x+y] = [x, x+y] + [y, x+y] \leq (\|x\| + \|y\|)\|x+y\|.$$

For homogeneity we have $\|\lambda x\|^2 = \lambda\,[x, \lambda x] \leq |\lambda|\|x\|\|\lambda x\|$. Therefore $\|\lambda x\| \leq |\lambda|\|x\|$. On the other hand, if $\lambda \neq 0$, then $\|x\| = \|\frac{1}{\lambda}\lambda x\| \leq \frac{1}{|\lambda|}\|\lambda x\|$. So $|\lambda|\|x\| \leq \|\lambda x\|$ as well. □

Definition 4.11. The semi-inner product $[.,.]$ *determines the norm* of a normed space X, if $\|x\| = \sqrt{[x,x]}$ for $x \in X$.

Proposition 4.12. *For every normed space* X *there is at least one semi-inner product such that it determines the norm.*

Proof. By the Hahn-Banach theorem, for every $y \in S(X)$ the set $\mathscr{S}(X,y)$ is nonempty. Therefore there is a $F : S(X) \to S(X')$ mapping such that $F(y) \in \mathscr{S}(X,y)$ for every $y \in S(X)$. Then the function

$$[x,y] = \begin{cases} 0 & \text{if } y = 0 \\ \|y\|F(\frac{1}{\|y\|}y)(x) & \text{if } y \neq 0 \end{cases}$$

is a semi-inner product, and it determines the norm of X. □

Definition 4.13. If $[.,.]$ is a semi-inner product on X, $T \in \mathscr{B}(X)$, then the *corresponding numerical range* is defined as

$$W(T) = \{[Tx,x] : [x,x] = 1, x \in X\}.$$

It follows from the proof of Proposition 4.12 that

$$\boldsymbol{V}(T) = \bigcup_{\alpha} W_\alpha(T)$$

where α indexes all the possible numerical ranges corresponding to a semi-inner product that determines norm. In certain terminology, the $W(T)$ corresponding to a semi-inner product is called the spatial numerical range, and $\boldsymbol{V}(T)$ is called the *total spatial numerical range.*

Theorem 4.14. *Let us suppose that $[.,.]$ is a semi-inner product that determines the norm of X. Then for $T \in \mathscr{B}(X)$ the corresponding numerical range $W(T)$ has the following properties:*

(1) $W(T) \subseteq \mathbf{V}(T)$,

(2) $\overline{\operatorname{conv}}\, W(T) = V(\mathscr{B}(X), T)$,

(3) $\sup\{\operatorname{Re}\lambda : \lambda \in W(T)\} = \inf_{t>0} \frac{1}{t}(\|I + tT\| - 1)$,

(4) $\sup\{|\lambda| : \lambda \in W(T)\} = v(T)$.

Proof. For $x \in X, y \in S(X)$ let us define the function $f_y(x) = [x, y]$. Then $(y, f_y) \in \Pi(X)$ because $f_y \in X'$, $\|f_y\| \leq 1$, and $f_y(y) = [y, y] = \|y\|^2 = 1$. If $\Gamma = \{(y, f_y) : y \in S(X)\}$, then $\pi_1(\Gamma) = S(X)$, and we may apply Lemma 4.4. □

4.3 Numerical Index

The numerical index is a constant which helps to describe the relationship between the norm and the numerical radius in a normed space. A survey of the numerical index of different normed spaces can be found in [12]. This section presents a few of these results. Let us suppose in this section that the dimension of the space X is always greater than 1.

Definition 4.15. The *numerical index* of a (real or complex) normed space X is

$$n(X) := \inf\{v(T) : T \in \mathscr{B}(X), \|T\| = 1\}.$$

Equivalently $n(X) = \sup\{k : k\|T\| \leq v(T) \text{ for every } T \in \mathscr{B}(X), 0 \leq k\}$. The numerical index is the same for isometrically isomorphic spaces. In general $0 \leq n(X) \leq 1$, and $n(X) > 0$ exactly when v and $\|.\|$ are equivalent. Theorem 3.19 says that in a complex space $\frac{1}{e} \leq n(X) \leq 1$. If $\mathcal{H}$ is a complex Hilbert space then $n(\mathcal{H}) = \frac{1}{2}$ by Theorem 2.7. In a real Hilbert space it is possible to construct an operator T such that $\|T\| = 1$ and $\langle Tx, x\rangle = 0$ for every $x \in S(\mathcal{H})$, therefore $n(\mathcal{H}) = 0$ if $\mathcal{H}$ is real.
Duncan, McGregor, Pryce, and White showed that for every $t \in [0, 1]$ there is a real normed space x such that $n(X) = t$, and for every $t \in [\frac{1}{e}, 0]$ there is a complex normed space X such that $n(X) = t$ [6]. In both cases X can be just a 2-dimensional space. They also showed that the Banach space of continuous scalar-valued functions on a compact Hausdorff space has numerical index 1. Since then the disk algebra has become another classical space whose numerical index has been found (its index is also 1). The numerical index of various other classical Banach spaces is still an open question.
Looking at the dual of X, $n(X') \leq n(X)$, and $n(X') = n(X)$ if X is reflexive. In [3] it was shown that it is possible that $n(X')$ is strictly smaller than $n(X)$. Furthermore for

every $\alpha, \beta \in [0,1]$ where $\alpha \geq \beta$ there is a Banach space $X_{\alpha,\beta}$ such that $n(X_{\alpha,\beta}) = \alpha$ and $n(X'_{\alpha,\beta}) = \beta$.

4.4 Convexity and Connectedness

While we know that $\boldsymbol{V}(T)$ is bounded, it is not necessarily closed or convex. Because for Hilbert spaces $\boldsymbol{V}(T) = W(T)$, the spatial numerical range can be open, closed, or neither. In the following example we are going to construct an operator whose spatial numerical range is not convex. On the other hand it can be proven that $\boldsymbol{V}(T)$ is always connected.

Example 4.16. Suppose $X_p = (\mathbb{C}^2, \|.\|_p)$, where $1 < p < \infty$, and for $(z,w) \in X_p$

$$\|(z,w)\|_p = (|z|^p + |w|^p)^{\frac{1}{p}}.$$

For our example let us pick the operator

$$T(z,w) = (iz + w, -z - iw).$$

If $f \in S(X'_p)$, then $f(z,w) = \lambda z + \mu w$ for some $(\lambda,\mu) \in S(X_q)$ where $\frac{1}{p} + \frac{1}{q} = 1$. By the equality case of the Hölder inequality there is a unique $(\lambda,\mu) \in S(X_q)$ such that $\lambda z + \mu w = 1$, concretely

$$(\lambda,\mu) = (\overline{z}|z|^{p-2}, \overline{w}|w|^{p-2}).$$

Therefore the spatial numerical range is

$$\boldsymbol{V}(T) = \{i|z|^p - i|w|^p + w\overline{z}|z|^{p-2} - \overline{w}z|w|^{p-2} : |z|^p + |w|^p = 1\}.$$

Let us convert this into exponential form. With $z = |z|e^{i\varphi}, w = |w|e^{i\psi}, \varphi, \psi \in \mathbb{R}$ and $\theta = \varphi - \psi$ we can write

$$\begin{aligned}\boldsymbol{V}(T) &= \{i|z|^p - i|w|^p + |z||w|e^{i(\psi-\varphi)}|z|^{p-2} - |z||w|e^{-i(\psi-\varphi)}|w|^{p-2} : |z|^p + |w|^p = 1\} \\ &= \{\cos\theta|z||w|(|z|^{p-2} - |w|^{p-2}) + i(|z|^p - |w|^p + \sin\theta|z||w|(|z|^{p-2} - |w|^{p-2})) : \\ &\quad |z|^p + |w|^p = 1\}.\end{aligned}$$

Let us look at the suprema $s_1 = \sup\{\operatorname{Re} \boldsymbol{V}(T)\}$ and $s_2 = \sup\{\boldsymbol{V}(T) \cap \mathbb{R}\}$. We have

$$\begin{aligned}s_1 &= \sup\{|z||w|(|z|^{p-2} - |w|^{p-2} : |z|^p + |w|^p = 1\}, \\ s_2 &= \sup\{\cos\theta|z||w|(|z|^{p-2} - |w|^{p-2}) : \\ &\qquad |z|^p + |w|^p = 1, |z|^p - |w|^p + \sin\theta|z||w|(|z|^{p-2} - |w|^{p-2}) = 0\}.\end{aligned}$$

For some concrete values of $z_1, z_2, w_1, w_2 \in \mathbb{C}$

$$s_1 = |z_1||w_1|(|z_1|^{p-2} - |w_1|^{p-2}), \quad s_2 = \cos\theta |z_2||w_2|(|z_2|^{p-2} - |w_2|^{p-2}).$$

Let us suppose that $p \neq 2$. Then $s_1 > 0$, and $s_2 < s_1$ when $\cos\theta \neq 1$. If $\cos\theta = 1$, then $|z|^{p-2} = |w|^{p-2}$, $s_2 = 0$, and again $s_2 < s_1$. $\boldsymbol{V}(T)$ is symmetric about the real axis, if $y \in \boldsymbol{V}(T)$, then $\overline{y} \in \boldsymbol{V}(T)$ as well. Therefore s_1 is attained on both sides of the real axis, and $\boldsymbol{V}(T)$ is not convex.

Definition 4.17. Let the *norm* $\times$ *weak-* topology* be the product topology on $X \times X'$ where the norm topology is given on X and the weak-* topology is given on X'.

Lemma 4.18. *If $E \subseteq \Pi(X)$ is relatively closed in $\Pi(X)$ according to the norm $\times$ weak-* topology, then $\pi_1 E$ is a norm-closed subset of X.*

Proof. Let the sequence $(x_n) \subseteq \pi_1 E$ be such that $x_n \to x$ for some $x \in X$. For (x_n) we have another sequence $(f_n) \subseteq S(X')$ such that $(x_n, f_n) \in E$. The closed unit ball is weak-* compact in X', so (f_n) has a cluster point f such that $\|f\| \leq 1$. We can decompose $f(x)$ as

$$f(x) = (f - f_n)(x) + f_n(x - x_n) + f_n(x_n).$$

Therefore we have the bound

$$\begin{aligned} |f(x) - 1| &= |f(x) - f_n(x_n)| \\ &= |(f - f_n)(x) + f_n(x - x_n)| \\ &\leq |(f - f_n)(x)| + |f_n(x - x_n)| \\ &\leq |(f - f_n)(x)| + \|x - x_n\|. \end{aligned}$$

For appropriate n the right hand side of the above inequality vanishes, so $f(X) = 1$ and $(x, f) \in \Pi(X)$. Since E is relatively closed in $\Pi(X)$, $(x, f) \in E$ and $x \in \pi_1 E$. □

Lemma 4.19. *According to the norm $\times$ weak-* topology $\Pi(X)$ is a connected subset of $X \times X'$, except when X is one-dimensional over $\mathbb{R}$.*

Proof. Let us suppose that $\Pi(X)$ is made up of two relatively closed disconnected components A and B: $A \cup B = \Pi(X)$ and $A \cap B = \emptyset$. By the previous lemma $\pi_1 A$ and $\pi_1 B$ are closed subsets of X, and $S(X) = \pi_1 A \cup \pi_1 A$. If there is an common element $x \in \pi_1 A \cap \pi_1 A$, there is also $f, g \in X'$ for which $(x, f) \in A$ and $(x, g) \in B$. In this case for $0 \leq t \leq 1$ we would have $(x, tf + (1-t)g) \in \Pi(X)$, which is impossible because A and B are disjoint. Therefore $\pi_1 A$ and $\pi_1 B$ have to be disjoint as well.
First we will look at the case when X is not one-dimensional over $\mathbb{R}$. If $x, y \in S(X)$ and $x + y \neq 0$, then $z_t := tx + (1-t)y$. For every $0 \leq t \leq 1$ we have $z_t \neq 0$ and $\frac{1}{\|z_t\|} z_t \in S(X)$. So x and y are in the same component. In the case when $x + y = 0$, let us pick an element

$z \in S(X)$ linearly independent from x. We have just seen that z must be in the same component as x. On the other hand z has to be linearly independent from y too, so they are in the same component as well. Therefore x and y share the same component, and $S(X)$ is connected.
If X is one-dimensional and real, then let us suppose that $u \in X$ and $\|u\| = 1$. Every $x \in X$ can be written as $x = ku$ where $k \in \mathbb{R}, |k| = \|x\|$. Similarly, every $f \in X'$ can be written as $f(x) = kf(u)$ where $\|f\| = |f(u)|$. Let g be a functional defined as $g(x) = g(ku) = k$. $S(X) = \{u, -u\}$, and $\Pi(X)$ is made up of two disconnected points, (u, g) and $(-u, -g)$. □

Theorem 4.20. *If $T \in \mathscr{B}(X)$, then $\mathbf{V}(T)$ is a connected set.*

Proof. If $(x, f), (y, g) \in \Pi(X)$, then

$$\begin{aligned} |f(Tx) - g(Ty)| &= |f(Tx) - g(Tx) + g(Tx - Ty)| \\ &\le |(f - g)(Tx)| + |g(T(x - y))| \\ &\le \|f - g\|\|T\| + \|x - y\|\|T\|. \end{aligned}$$

Therefore $(x, f) \mapsto f(Tx)$ is a continuous mapping from $\Pi(X)$ to $\boldsymbol{V}(T)$ (with the norm $\times$ weak-* topology). By the previous lemma $\Pi(X)$ is connected, so $\boldsymbol{V}(T)$ is connected too. The only exception could be when X is one-dimensional and real, but in that case $\boldsymbol{V}(T)$ consists of a single point, $g(Tu)$. □

Bibliography

[1] F. F. Bonsall and J. Duncan. *Numerical Ranges II.* Cambridge University Press, 1973. ISBN: 9780521202275.

[2] F. F. Bonsall and J. Duncan. *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras.* Cambridge University Press, 1971. ISBN: 9780521079884.

[3] K. Boyko et al. "Numerical index of Banach spaces and duality". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 142 (2007), pp. 93–102. DOI: 10.1017/S0305004106009650.

[4] T. Constantinescu. *Schur Parameters, Factorization and Dilation Problems.* Birkhauser Verlag, 1996. ISBN: 9783034899109.

[5] J. B. Conway. *A Course in Functional Analysis.* Springer, 1990. ISBN: 9783540960423.

[6] J. Duncan et al. "The Numerical Index of a Normed Space". In: *Journal of the London Mathematical Society* 2.3 (1970), pp. 481–488. DOI: 10.1112/jlms/2.Part_3.481.

[7] K. E. Gustafson and D. K. M. Rao. *Numerical Range. The Field of Values of Linear Operators and Matrices.* Springer, 1997. ISBN: 9780387948355.

[8] P. Halmos. *A Hilbert Space Problem Book.* Springer-Verlag, 1982. ISBN: 978-0-3879-0685-0.

[9] J. A. R. Holbrook. "Multiplicative properties of the numerical radius in operator theory." In: *Journal für die reine und angewandte Mathematik* 237 (1969), pp. 166–174. URL: http://eudml.org/doc/150949.

[10] Molnár L. *Banach-algebrák, C*-algebrák és Neumann-algebrák.* Debreceni Egyetem Matematikai és Informatikai Intézete, 2000.

[11] G. Lumer. "Semi-inner-product spaces". In: *Transactions of the American Mathematical Society* 100.1 (1971), pp. 29–43. DOI: https://doi.org/10.1090/S0002-9947-1961-0133024-2.

[12] M. Martín. "A Survey on the Numerical Index of a Banach Space". In: *Extracta mathematicae* 15.2 (2000), pp. 265–276. ISSN: 0213-8743.

[13] T. W. Palmer. *Banach Algebras and The General Theory of *-Algebras I.* Cambridge University Press, 1994. ISBN: 9780521366373.

[14] C. Pearcy. "An elementary proof of the power inequality for the numerical radius." In: *Michigan Math. J.* 13.3 (1966), pp. 289–291. DOI: 10.1307/mmj/1031732779. URL: https://projecteuclid.org/euclid.mmj/1031732779.

[15] A. M. Sinclair. "The Norm of a Hermitian Element in a Banach Algebra". In: *Proceedings of the American Mathematical Society* 28.2 (1971), pp. 446–450. URL: http://www.jstor.org/stable/2037989.

[16] P. G. Spain. "Characterizations of Hilbert space and the Vidav-Palmer theorem". In: *Rocky Mountain Journal of Mathematics* 43.4 (2013), pp. 1337–1354. DOI: 10.1216/RMJ-2013-43-4-1337.

[17] Tarcsay Zs. *Funkcionálanalízis.* Eötvös Loránd Tudományegyetem TTK, 2018.