Numerical Ranges of Operators

Paul Skoufranis

August 21, 2014

Abstract

The purpose of this document is to define and develop various notions of numerical ranges of operators. This document is designed to include the some of the basic and most important properties of each numerical range which appear as important technical details in the proofs of many approximation properties. In particular, we shall prove one of the most essential results about the numerical range - the Toeplitz-Hausdorff Theorem.

The reader of these notes must have a basic knowledge of the bounded linear maps on a Hilbert space. For some of the more advanced topics, an understanding of approximate unitarily equivalence of operators and spectral theorems for normal operators will be useful. Note that all inner products in this document are linear in the first variable. Moreover $\mathcal{H}$ will denote a Hilbert space, $\mathcal{B}(\mathcal{H})$ will denote the bounded linear maps on $\mathcal{H}$, and $\sigma(T)$ will denote the spectrum of an element $T \in \mathcal{B}(\mathcal{H})$.

This document is for educational purposes and should not be referenced. Please contact the author of this document if you need aid in finding the correct reference. Comments, corrections, and recommendations on these notes are always appreciated and may be e-mailed to the author (see his website for contact info).

Contents

1 Numerical Radius

We shall begin with the basic concept of the numerical radius of an operator. One approach to the theory of the numerical radius would be to first develop the numerical range of an operator which will be done in Section 2. However we shall avoid this approach to make the concept of the numerical radius as simple as possible. Since the computations of the numerical radius of an operator are identical to the computations of the numerical range of an operator, we shall postpone examples until Section 2. Thus we begin with the definition of the numerical radius of an operator.

Definition 1.1. Let $T \in \mathcal{B}(\mathcal{H})$. The numerical radius of T is

$$nr(T) := \sup\{|\langle T\xi, \xi\rangle| \mid \xi \in \mathcal{H}, \|\xi\| \leq 1\} \in [0, \infty).$$

As mentioned in the opening of this section, we shall postpone specific examples of computations of the numerical radius until Section 2. However, in certain situations, the numerical radius is easy to compute. For example, the following shows the numerical radius of a self-adjoint operator is the norm of the operator and Theorem 1.7 will show the same for normal operators.

Proposition 1.2. *Let $T \in \mathcal{B}(\mathcal{H})$ be a self-adjoint operator. Then $nr(T) = \|T\|$.*

Proof. Clearly

$$nr(T) \leq \{\|T\xi\| \, \|\xi\| \mid \xi \in \mathcal{H}, \|\xi\| \leq 1\} \leq \|T\|.$$

To show the other inequality we recall that

$$\|T\| = \{|\langle T\xi, \eta\rangle| \mid \xi, \eta \in \mathcal{H}, \|\xi\|, \|\eta\| \leq 1\}.$$

The goal of the proof is 'to change η into ξ' in the above expression. To begin, fix ξ and $\eta \in \mathcal{H}$ with $\|\xi\| \leq 1$ and $\|\eta\| \leq 1$. For a reason that will become apparent later, choose $\theta \in [0, 2\pi)$ such that

$$|\langle T\xi, \eta\rangle| = e^{i\theta}\langle T\xi, \eta\rangle.$$

Let $\eta' := e^{-i\theta}\eta \in \mathcal{H}$ so that $\|\eta'\| \leq 1$ and $\langle T\xi, \eta'\rangle = |\langle T\xi, \eta\rangle| \in \mathbb{R}$.

Next recall the polarization identity which states

$$\langle T\xi, \eta'\rangle = \frac{1}{4}\left(\langle T(\xi + \eta'), \xi + \eta'\rangle - \langle T(\xi - \eta'), \xi - \eta'\rangle + i\langle T(\xi + i\eta')\xi + i\eta'\rangle - i\langle T(\xi - i\eta'), \xi - i\eta'\rangle\right).$$

(if you are not familiar with this, simply expand out the right hand side and you will get the left). Notice that each inner product in the above expression is a real number as T is self-adjoint. Since each inner product is a real number and, by our choice of η', $\langle T\xi, \eta'\rangle$ is real we must have the complex terms sum to zero and thus

$$\langle T\xi, \eta'\rangle = \frac{1}{4}\left(\langle T(\xi + \eta'), \xi + \eta'\rangle - \langle T(\xi - \eta'), \xi - \eta'\rangle\right).$$

By applying the definition of the numerical radius and a simple scaling argument, we see that $|\langle T\zeta, \zeta\rangle| \leq \|\zeta\|^2 nr(T)$ for all $\zeta \in \mathcal{H}$. Whence

$$|\langle T\xi, \eta\rangle| \leq \frac{1}{4}\left(|\langle T(\xi + \eta'), \xi + \eta'\rangle| + |\langle T(\xi - \eta'), \xi - \eta'\rangle|\right) \leq \frac{1}{4}nr(T)\left(\|\xi + \eta\|^2 + \|\xi - \eta\|^2\right).$$

Thus, by applying the Parallelogram Law (a rare use of the law!), we obtain that

$$|\langle T\xi, \eta\rangle| \leq \frac{1}{4}nr(T)\left(2\|\xi\|^2 + 2\|\eta'\|^2\right) \leq nr(T)$$

as claimed. As $\eta, \xi \in \mathcal{H}$ were arbitrary elements with norm at most one, the expression for $\|T\|$ shows that $\|T\| \leq nr(T)$. $\square$

Our next goal is to show that the numerical radius is a norm on $\mathcal{B}(\mathcal{H})$ that is equivalent to the operator norm. However, the numerical radius need not equal the norm of an operator which will be demonstrated in Example 2.12. To show that the numerical radius is a norm, we begin with the following simple lemma.

Lemma 1.3. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $nr(Re(T)) \leq nr(T)$ and $nr(Im(T)) \leq nr(T)$.*

Proof. Simply applying definitions gives

$$\begin{array}{rcl} nr(Re(T)) & = & \sup\left\{\left|\left\langle\left(\frac{T+T^*}{2}\right)\xi,\xi\right\rangle\right| \mid \xi \in \mathcal{H}, \|\xi\| \leq 1\right\} \\ & \leq & \sup\left\{\frac{1}{2}\left|\langle T\xi,\xi\rangle\right| + \frac{1}{2}\left|\langle \xi,T\xi\rangle\right| \mid \xi \in \mathcal{H}, \|\xi\| \leq 1\right\} \\ & = & nr(T). \end{array}$$

The proof that $nr(Im(T)) \leq nr(T)$ is identical. ☐

Theorem 1.4. *For all $T \in \mathcal{B}(\mathcal{H})$, $nr(T) \leq \|T\| \leq 2nr(T)$. Moreover $T \mapsto nr(T)$ defines a norm on $\mathcal{B}(\mathcal{H})$.*

Proof. Clearly
$$nr(T) \leq \{\|T\xi\|\,\|\xi\| \mid \xi \in \mathcal{H}, \|\xi\| \leq 1\} \leq \|T\|.$$
To prove the second inequality write $T = Re(T) + iIm(T)$. Then $Re(T)$ and $Im(T)$ are self-adjoint elements so combining Proposition 1.2 and Lemma 1.3 gives $\|Re(T)\| = nr(Re(T)) \leq nr(T)$ and $\|Im(T)\| = nr(Im(T)) \leq nr(T)$. Thus
$$\|T\| \leq \|Re(T)\| + \|Im(T)\| \leq 2nr(T)$$
as claimed.

It is trivial to verify that $T \mapsto nr(T)$ is a non-negative valued map, $nr(\lambda T) = |\lambda| nr(T)$, and $nr(T+S) \leq nr(T) + nr(S)$ for all $S, T \in \mathcal{B}(\mathcal{H})$. Lastly, if $nr(T) = 0$ then $\|T\| \leq 2nr(T) = 0$ so that $T = 0$. Whence nr defines a norm on $\mathcal{B}(\mathcal{H})$. ☐

To complete this section we desire to show that the numerical radius of a normal operator is equal to the norm of the operator. To do this, we will need to appeal to a theorem of Weyl, von Neumann, and Berg which states that any normal operator is approximately unitarily equivalent to a diagonal normal operator. Thus it is necessary to demonstrate that the numerical radius is invariant under approximate unitary equivalence. First we demonstrate the triviality that the numerical radius is invariant under unitary equivalence.

Lemma 1.5. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $nr(UTU^*) = nr(T)$ for every unitary operator $U \in \mathcal{B}(\mathcal{H})$.*

Proof. If U is a unitary operator, it is clear that ξ is a unit vector in $\mathcal{H}$ if and only if $U^*\xi$ is a unit vector. Therefore, since
$$\langle UTU^*\xi, \xi\rangle = \langle TU^*\xi, U^*\xi\rangle$$
for all $\xi \in \mathcal{H}$, the result clearly follows from the definition of the numerical radius. ☐

With the above and Theorem 1.4, we easily obtain that the numerical radius is invariant under approximate unitary equivalence.

Lemma 1.6. *Let $T, S \in \mathcal{B}(\mathcal{H})$ be approximately unitarily equivalent (that is, there exists a sequence $(U_n)_{n\geq 1}$ of unitaries such that $\lim_{n\to\infty} \|U_nSU_n^* - T\| = 0$). Then $nr(S) = nr(T)$.*

Proof. Suppose there exists a sequence $(U_n)_{n\geq 1}$ of unitaries such that $\lim_{n\to\infty} \|U_nSU_n^* - T\| = 0$. Therefore, since the numerical radius is a norm by Theorem 1.4 and by Lemma 1.5,

$$nr(T) \leq nr(T - U_nSU_n^*) + nr(U_nSU_n^*) \leq \|T - U_nSU_n^*\| + nr(S).$$

Therefore, since $\lim_{n\to\infty} \|U_nSU_n^* - T\| = 0$, we obtain that $nr(T) \leq nr(S)$. Since $(U_n^*)_{n\geq 1}$ is also a sequence of unitaries such that $\lim_{n\to\infty} \|U_n^*TU_n - S\| = 0$, by repeating the above argument we obtain that $nr(S) \leq nr(T)$. Hence $nr(S) = nr(T)$ as desired. ☐

Hence we easily obtain the following using the Weyl-von Neumann-Berg Theorem.

Theorem 1.7. *Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator. Then $nr(N) = \|N\|$.*

Proof. From Theorem 1.4 it is clear that $nr(N) \leq \|N\|$. Moreover, since N is a normal operator, it is elementary to show that $\|N\| = \sup\{|\lambda| \mid \lambda \in \sigma(N)\}$. Choose $\lambda \in \sigma(N)$ such that $|\lambda| = \|N\|$. By the Weyl-von Neumann-Berg Theorem N is approximately unitarily equivalent to a diagonal normal operator D such that $\langle D\xi, \xi\rangle = \lambda$ for some unit vector $\xi \in \mathcal{H}$. Therefore, by Lemma 1.6,

$$nr(N) = nr(D) \geq |\lambda| = \|N\|$$

which completes the proof. $\square$

2 Numerical Range - Basic Properties and Examples

In this section we will develop the basic properties of the numerical range of an operator. As the numerical range and radius of an operator are intimately connected, we will draw more information about the numerical radius in this section. We shall also demonstrate several examples of the numerical ranges of operators that will be vital examples. In particular, we shall completely describe the numerical range of any 2×2 matrix in terms of its eigenvalues and eigenvectors which is surprisingly the vital ingredient in the proof of the Toeplitz-Hausdorff Theorem in Section 3.

We begin with the definition of the numerical range.

Definition 2.1. Let $T \in \mathcal{B}(\mathcal{H})$. The numerical range of T, denoted $W(T)$, is the non-empty set

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

Remarks 2.2. It is clear from the definitions of the numerical range and radius that

$$nr(T) = \sup\{|\lambda| \mid \lambda \in W(T)\}.$$

Thus these two topic are intimately related.

Before we move onto to compute the numerical ranges of certain operators, we demonstrate how the numerical range contains more information than the spectrum. In particular, there are several operators that have a single point as spectrum (e.g. quasinilpotent operators) yet these is only one operator whose numerical range is a given singleton.

Proposition 2.3. *Let $T \in \mathcal{B}(\mathcal{H})$ and let $\lambda \in \mathbb{C}$. Then $W(T) = \{\lambda\}$ if and only if $T = \lambda I_{\mathcal{H}}$.*

Proof. It is clear that $W(\lambda I_{\mathcal{H}}) = \{\lambda\}$ for any $\lambda \in \mathbb{C}$.

Suppose $W(T) = \{\lambda\}$ for some $T \in \mathcal{B}(\mathcal{H})$. Then for all $\xi \in \mathcal{H}$ with $\|\xi\| = 1$,

$$\langle(\lambda I_{\mathcal{H}} - T)\xi, \xi\rangle = \lambda - \langle T\xi, \xi\rangle = \lambda - \lambda = 0.$$

Hence $\langle(\lambda I_{\mathcal{H}} - T)\xi, \xi\rangle = 0$ for all $\xi \in \mathcal{H}$ and thus $\lambda I_{\mathcal{H}} - T = 0$. □

Note that the above gives the trivial example that $nr(\lambda I_{\mathcal{H}}) = |\lambda|$.

Before computing specific examples of numerical ranges of operators, it is also useful to develop some of the basic machinery of the numerical range as this will ease in the computations.

Proposition 2.4. *Let $T, S \in \mathcal{B}(\mathcal{H})$. Then*

1. $W(T^*) = \overline{W(T)}$.
2. *$W(T)$ contains all of the eigenvalues of T.*
3. *$W(T)$ is contained in the closed disk of radius $\|T\|$ around the origin.*
4. *If $a, b \in \mathbb{C}$ then $W(aT + bI_{\mathcal{H}}) = aW(T) + b$.*
5. *If $U \in \mathcal{B}(\mathcal{H})$ is a unitary then $W(UTU^*) = W(T)$.*
6. *$W(T) \subseteq \mathbb{R}$ if and only if T is self-adjoint.*
7. *If $\mathcal{H}$ is finite dimensional, $W(T)$ is closed and thus compact.*
8. $W(T + S) \subseteq W(T) + W(S)$.

Proof. Property (1) clearly follows from the fact that $\langle T^*\xi, \xi\rangle = \overline{\langle T\xi, \xi\rangle}$ for all $\xi \in \mathcal{H}$.

Property (2) follows from the fact that if λ is an eigenvalue of T with non-zero eigenvector ξ_0, then $\xi := \frac{1}{\|\xi_0\|}$ is a unit eigenvector for T with eigenvalue λ and thus $\lambda = \langle T\xi, \xi\rangle \in W(T)$ as desired.

Property (3) trivially follows from the fact that $|\langle T\xi, \xi\rangle| \leq \|T\xi\| \, \|\xi\| \leq \|T\|$ for all $\xi \in \mathcal{H}$ with norm one.

Property (4) follows trivially from the fact that $\langle (aT + bI_{\mathcal{H}})\xi, \xi\rangle = a\langle T\xi, \xi\rangle + b$ for all $\xi \in \mathcal{H}$ with norm one.

Property (5) follows from the fact that $\langle UTU^*\xi, \xi\rangle = \langle T(U^*\xi), U^*\xi\rangle$ for all $\xi \in \mathcal{H}$ and ξ has norm one if and only if $U^*\xi$ has norm one.

Property (6) follows from the fact that $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint if and only if $\langle T\xi, \xi\rangle \in \mathbb{R}$ for every $\xi \in \mathcal{H}$ which is equivalent to $\langle T\xi, \xi\rangle \in \mathbb{R}$ for every $\xi \in \mathcal{H}$ with norm one.

To see that Property (7) holds, we notice that compactness follows if $W(T)$ is closed by part (3). Suppose $(\lambda_n)_{n\geq 1}$ is a sequence in $W(T)$ that converges to $\lambda \in \mathbb{C}$. For each $n \in \mathbb{N}$, choose $\xi_n \in \mathcal{H}$ with norm one such that $\lambda_n = \langle T\xi_n, \xi_n\rangle$. Since $\mathcal{H}$ is a finite dimensional Hilbert space, the unit ball of $\mathcal{H}$ is compact and thus there exists a sequence $(\xi_{n_k})_{k\geq 1}$ that converges to a unit vector $\xi \in \mathcal{H}$. This implies that

$$\langle T\xi, \xi\rangle = \lim_{n\to\infty} \langle T\xi_n, \xi_n\rangle = \lim_{n\to\infty} \lambda_n = \lambda$$

and thus $\lambda \in W(T)$ as desired.

Property (8) clearly follows from the definition of the numerical range. $\square$

To begin our examples of numerical ranges of operators, we will first discuss the unilateral backward shift.

Example 2.5. Let $\mathcal{H}$ be a separable Hilbert space with an orthonormal basis $\{e_n\}_{n\geq 1}$. Let $T \in \mathcal{B}(\mathcal{H})$ be the unilateral backward shift operator; that is $T(e_1) = 0$ and $T(e_n) = e_{n-1}$ for all $n \geq 2$. Then $W(T)$ is the open unit disk centred at the origin. To begin, we notice that if $\xi \in \mathcal{H}$ has norm one, then

$$|\langle T\xi, \xi\rangle| \leq \|T\xi\| \, \|\xi\| \leq 1$$

with equality if and only if $T\xi$ and ξ are multiples of each other and $\|T\xi\| = 1$. This would imply that $T\xi = \lambda\xi$ for some $\lambda \in \mathbb{C}$ with $|\lambda| = 1$. However, if $\xi = \sum_{n\geq 1} a_n e_n$, the equation $T\xi = \lambda\xi$ would imply that $\lambda a_n = a_{n+1}$ for all $n \in \mathbb{N}$. Thus $|a_n| = |a_1|$ for all $n \in \mathbb{N}$ which is impossible as $1 = \|\xi\|^2 = \sum_{n\geq 1} |a_n|^2$. Thus $W(T)$ is a subset of the open unit disk.

To see that $W(T)$ is the open unit disk, let $\lambda \in \mathbb{C}$ be such that $|\lambda| < 1$. Let $\xi_0 := \sum_{n\geq 1} \lambda^n e_n \in \mathcal{H}$ which exists as $\sum_{n\geq 1} |\lambda^n|^2$ converges. Thus $T\xi_0 = \lambda\xi_0$. Hence λ is an eigenvalue for T and thus $\lambda \in W(T)$ by Proposition 2.4 part (2). Hence $W(T)$ is the open unit disk.

Remarks 2.6. Notice that if T is the unilateral backward shift then $W(T)$ is open and not closed. This provides an example of where Proposition 2.4 part (7) fails. Moreover this demonstrates that $nr(T) = 1 = \|T\|$.

Next we will demonstrate the important example of the numerical range of diagonal operators.

Example 2.7. Let $\mathcal{H}$ be a separable Hilbert space and let $T \in \mathcal{B}(\mathcal{H})$ be a diagonal operator; that is, there exists an orthonormal basis $\{e_n\}_{n\geq 1}$ of $\mathcal{H}$ and a bounded set $\{a_n\}_{n\geq 1}$ of scalars such that $Te_n = a_n e_n$ for all $n \in \mathbb{N}$. Let $\xi = \sum_{n\geq 1} c_n e_n$ be an arbitrary unit vector. Thus $\sum_{n\geq 1} |c_n|^2 = 1$ and

$$\begin{aligned} \langle T\xi, \xi\rangle &= \left\langle \textstyle\sum_{n\geq 1} a_n c_n e_n, \sum_{n\geq 1} c_n e_n \right\rangle \\ &= \textstyle\sum_{n\geq 1} a_n |c_n|^2. \end{aligned}$$

Hence

$$W(T) = \left\{ \sum_{n\geq 1} a_n b_n \;\middle|\; b_n \geq 0, \sum_{n\geq 1} b_n = 1 \right\}.$$

We claim that $W(T) = conv(\{a_n\}_{n\geq 1})$; the convex hull of $\{a_n\}_{n\geq 1}$. It is clear from the above expression that $W(T)$ is convex and $conv(\{a_n\}_{n\geq 1}) \subseteq W(T)$.

Suppose $\lambda \in W(T)$. Then, either $\lambda \in conv(\{a_n\}_{n\geq 1})$ or, by a corollary to the Hahn-Banach theorem (or low-dimensional topology), there exists a closed half-plane with λ on the boundary that contains $conv(\{a_n\}_{n\geq 1})$. By Proposition 2.4 part (4) and the fact $conv(a\{a_n\}_{n\geq 1} + b) = aconv(\{a_n\}_{n\geq 1}) + b$ for all $a, b \in \mathbb{C}$, by performing a translation and rotation we can assume that $\lambda = 0$ and $conv(\{a_n\}_{n\geq 1})$ is contained in the closed upper half-plane. Since $\lambda \in W(T)$, there exists $\{b_n\}_{n\geq 1}$ such that $b_n \geq 0$ for all $n \in \mathbb{N}$, $\sum_{n\geq 1} b_n = 1$, and $0 = \sum_{n\geq 1} a_n b_n$. Since each b_n is positive and each a_n is contained in the closed upper half-plane, $b_n = 0$ whenever a_n contains an imaginary part. Therefore, since $\sum_{n\geq 1} b_n = 1$ and $0 = \sum_{n\geq 1} a_n b_n$, either $a_m = 0$ and $b_m = 1$ for some $m \in \mathbb{N}$, or there exists $m_1, m_2 \in \mathbb{N}$ such that $a_{m_1} > 0$ and $a_{m_2} < 0$. It is then apparent that $0 \in conv(\{a_n\}_{n\geq 1})$. Hence $W(T) = conv(\{a_n\}_{n\geq 1})$ as desired.

Remarks 2.8. Example 2.7 has many interesting consequences. To begin, let $T \in \mathcal{B}(\mathcal{H})$ be a diagonal self-adjoint operator with spectrum $[0, 1]$. Then, depending on whether 0 and 1 appear along the diagonal of T, it is apparent from Example 2.7 that $W(T)$ is either $[0, 1]$, $(0, 1]$, $[0, 1)$, or $(0, 1)$ and any of these occur for some such self-adjoint operator. Hence, unlike the numerical range as demonstrated in Lemma 1.6, the numerical range of an operator is not invariant under approximate unitary equivalence. Moreover this gives additional examples that the numerical range need not be closed.

Another interesting consequence is the following: if $N \in \mathcal{B}(\mathcal{H})$ is a normal operator, it need not be the case that $\sigma(N) \subseteq W(N)$ nor $W(N) \subseteq \sigma(N)$. Indeed we have seen that a self-adjoint diagonal operator with spectrum $[0, 1]$ can have $(0, 1)$ as its numerical radius so $\sigma(N) \subseteq W(N)$ may not occur. Furthermore, if N is a diagonal normal operator with diagonal entries $\{a_n\}_{n\geq 1}$, then $\sigma(N) = \overline{\{a_n\}_{n\geq 1}}$ yet $W(N) = conv(\{a_n\}_{n\geq 1})$ so, for certain choices of a_n, $W(N)$ need not be a subset of $\sigma(N)$.

Remarks 2.8 show that the numerical range does not behave nicely as it need not be closed and does not appear to have any direct relation to the spectrum. These problems occur directly because the numerical range is not closed. The following then shows that the spectrum and concepts of approximate unitary equivalence behave nicely with respect to the closure of the numerical range.

Theorem 2.9. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $\sigma(T) \subseteq \overline{W(T)}$.*

Proof. Let $\lambda \in \sigma(T)$. Then either $\lambda I_{\mathcal{H}} - T$ is not bounded below, or $\lambda I_{\mathcal{H}} - T$ is bounded below but not onto. Suppose $\lambda I_{\mathcal{H}} - T$ is not bounded below. Then there exists a sequence of unit vectors $\xi_n \in \mathcal{H}$ such that $\lim_{n\to\infty} \|(\lambda I_{\mathcal{H}} - T)\xi_n\| = 0$. Thus

$$0 = \lim_{n\to\infty} \langle(\lambda I_{\mathcal{H}} - T)\xi_n, \xi_n\rangle = \lim_{n\to\infty} \lambda - \langle T\xi_n, \xi_n\rangle.$$

As each $\langle T\xi_n, \xi_n\rangle \in W(T)$, $\lambda \in \overline{W(T)}$ in this case.

Now suppose $\lambda I_{\mathcal{H}} - T$ is bounded below but not onto. Therefore $ker(\overline{\lambda} I - T^*) = Im(\lambda I_{\mathcal{H}} - T)^{\perp}$ is non-empty so $\overline{\lambda}$ is an eigenvalue of T^*. Hence $\overline{\lambda} \in W(T^*)$ Proposition 2.4 part (2) so $\lambda \in W(T)$ by Proposition 2.4 part (1). $\square$

Theorem 2.10. *Let $T, S \in \mathcal{B}(\mathcal{H})$ be approximately unitarily equivalent. Then $\overline{W(T)} = \overline{W(S)}$.*

Proof. Suppose there exists a sequence $(U_n)_{n\geq 1}$ of unitaries such that $\lim_{n\to\infty} \|U_n S U_n^* - T\| = 0$. Fix $\lambda \in W(T)$ and choose a unit vector $\xi \in \mathcal{H}$ such that $\lambda = \langle T\xi, \xi\rangle$. Therefore

$$\lim_{n\to\infty} |\lambda - \langle S(U_n^*\xi), U_n^*\xi\rangle| = \lim_{n\to\infty} |\langle(U_n S U_n^* - T)\xi, \xi\rangle| = 0.$$

Therefore, since $U_n^*\xi$ is a unit vector for all $n \in \mathbb{N}$, $\langle S(U_n^*\xi), U_n^*\xi\rangle \in W(S)$ for all $n \in \mathbb{N}$ so $\lambda \in \overline{W(S)}$. Therefore, since $\lambda \in W(T)$ was arbitrary, $\overline{W(T)} \subseteq \overline{W(S)}$. The result then follows by symmetry. $\square$

Combining Theorem 2.10, Example 2.7, and the Weyl-von Neumann-Berg Theorem, we have the following result relating the spectrum and numerical range of a normal operator.

Theorem 2.11. *Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator. Then $\overline{W(N)} = conv(\sigma(N))$.*

Proof. By the Weyl-von Neumann-Berg Theorem, there exists a diagonal normal operator D such that N and D are approximately unitarily equivalent and $\sigma(N) = \sigma(D)$. Therefore, by Theorem 2.10, $\overline{W(N)} = \overline{W(D)}$. However, by Example 2.7, $W(D)$ is the convex hull of the eigenvalues $\{a_n\}_{n\geq 1}$ of D. Since the eigenvalues of D are dense in the spectrum of D, it is clear that

$$\overline{W(D)} = \overline{conv(\{a_n\}_{n\geq 1})} = \overline{conv(\sigma(N))} = conv(\sigma(N))$$

as $\sigma(N)$ is compact subset of $\mathbb{C}$ so its convex hull is closed. Hence the result is complete. □

Our next example may seem simplistic, but we shall see in Remarks 2.13 and in the proof of Theorem 2.14 that the following example may be the most important example given yet.

Example 2.12. Consider

$$M := \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \in \mathcal{M}_2(\mathbb{C}).$$

Then $W(M)$ is the closed disk of radius $\frac{1}{2}$ centred at the origin. To see this we note that it is elementary to show that $\xi \in \mathbb{C}^2$ is a unit vector if and only if we can write $\xi = (\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2})$ for some $\theta_j \in [0, 2\pi)$ and $\theta \in [0, \frac{\pi}{2}]$. However

$$\langle M(\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2}), (\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2})\rangle = \cos(\theta)\sin(\theta)e^{i(\theta_2-\theta_1)}.$$

By ranging over all possible $\theta_j \in [0, 2\pi)$ and $\theta \in [0, \frac{\pi}{2}]$ and using the fact that the range of $\cos(\theta)\sin(\theta) = \frac{1}{2}\sin(2\theta)$ over $\theta \in [0, \frac{\pi}{2}]$ is $[0, \frac{1}{2}]$, we see that $W(M)$ is precisely the closed disk of radius $\frac{1}{2}$ centred at the origin.

Remarks 2.13. Note Example 2.12 shows that $nr(M) = \frac{1}{2} \neq 1 = \|M\|$ thus demonstrating that the numerical range and operator norm are not equal norms (even though they are equivalent by Theorem 1.4). Moreover this demonstrates that $\|M\| = 2nr(M)$ thus showing the non-trivial inequality in Theorem 1.4 is strict.

To complete this section, we shall demonstrate the most important use of Example 2.12 which gives a complete description of the numerical ranges of element of $\mathcal{M}_2(\mathbb{C})$. Surprisingly this is the main step in the proof of the all-important Toeplitz-Hausdorff Theorem which will be given in Section 3

Theorem 2.14. *For $A \in \mathcal{M}_2(\mathbb{C})$, either*

1. *if $A = \lambda I_2$, then $W(A) = \{\lambda\}$,*
2. *if the eigenvalues of A are equal and A is not a multiple of the identity, $W(A)$ is a non-trivial closed disk centred at the eigenvalues of A, or*
3. *if the eigenvalues λ_1 and λ_2 of A are distinct, $W(A)$ is a possibly degenerate ellipse with foci λ_1 and λ_2. Moreover, if ξ_i is any unit eigenvector for λ_i, then the eccentricity of $W(A)$ is $|\langle \xi_1, \xi_2\rangle|^{-1}$ and the length of the major axis is $\frac{|\lambda_1-\lambda_2|}{\sqrt{1-|\langle \xi_1,\xi_2\rangle|^2}}$.*

Proof. It is clear that (1) follows from Proposition 2.4 part (4).

To see that (2) holds, suppose that the eigenvalues of A are equal and A is not a multiple of the identity. Let $\lambda \in \mathbb{C}$ be the eigenvalue of A. Then, by basic theory of matrices, there exists a unitary $U \in \mathcal{M}_2(\mathbb{C})$ such that $A = U(\lambda I_2 + aM)U^*$ where $a \in \mathbb{C}$ is non-zero and M is the matrix in Example 2.12. Therefore, by Proposition 2.4 parts (4) and (5),

$$W(A) = W(U(\lambda I_2 + aM)U^*) = W(\lambda I_2 + aM) = \lambda + aW(M)$$

so the result follows from Example 2.12.

Finally, we arrive at the significant case. Suppose $A \in \mathcal{M}_2(\mathbb{C})$ has two distinct eigenvalues. Since $W(A + aI_2) = W(A) + a$ for all $a \in \mathbb{C}$, it is easy to see that we may assume that there exists a $\lambda \in \mathbb{C}$ such that the eigenvalues of A are λ and $-\lambda$ as the conclusions of this case are invariant under translations (and as the eigenvectors of translations are the eigenvectors of the original matrix). Since the eigenvalues of A are $\pm\lambda$, it is clear that $tr(A) = 0$.

Let ξ_1 be a unit eigenvector for λ and let ξ_2 be a unit eigenvector for $-\lambda$. If ξ_1 and ξ_2 are orthogonal, it is easy to see using Example 2.7 that $W(A)$ is the line segment connecting λ to $-\lambda$. As a line segment is an ellipse with foci at the endpoints, with infinite eccentricity, and with a major axis of length $2\lambda = \frac{|\lambda-(-\lambda)|}{\sqrt{1-|\langle \xi_1,\xi_2\rangle|^2}}$, the proof is complete in this setting.

Hence we may assume that $\langle \xi_1, \xi_2\rangle \neq 0$. Choose $\theta \in [0, 2\pi)$ such that $e^{-i\theta}\langle \xi_1, \xi_2\rangle$ is real. Therefore, if $\eta_1 := \xi_1 + e^{i\theta}\xi_2$ then

$$\langle A\eta_1, \eta_1\rangle = \langle \lambda\xi_1 - \lambda e^{i\theta}\xi_2, \xi_1 + e^{i\theta}\xi_2\rangle = \lambda - \lambda + \lambda e^{-i\theta}\langle \xi_1, \xi_2\rangle - \lambda e^{i\theta}\langle \xi_2, \xi_1\rangle = 2i\lambda Im(e^{-i\theta}\langle \xi_1, \xi_2\rangle) = 0.$$

Let η_2 be any unit vector that is orthogonal to η_1. Since $tr(A) = 0$ and $\langle A\eta_1, \eta_1\rangle = 0$, we easily obtain that $\langle A\eta_2, \eta_2\rangle = 0$. Therefore, A is unitarily equivalent to a matrix B of the form

$$B = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$$

for some $a, b \in \mathbb{C}$. Therefore, as the numerical range is invariant under unitary conjugation by Proposition 2.4 part (5), $W(A) = W(B)$. Moreover, notice that the eigenvalues of B are $\pm\sqrt{ab}$ for some choice of the square root function of complex numbers. Since the eigenvalues of A and B are equal, we must have that $\sqrt{ab}$ and $-\sqrt{ab}$ are distinct.

Write $a = |a|e^{i\alpha}$ and $b = |b|e^{i\beta}$ where $\alpha, \beta \in [0, 2\pi)$. Consider the matrix

$$V := \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{1}{2}i(\alpha-\beta)} \end{bmatrix} \in \mathcal{M}_2(\mathbb{C}).$$

Since $|e^{\frac{1}{2}i(\alpha-\beta)}| = 1$, is it easy to see that V is a unitary matrix such that

$$V^* = \begin{bmatrix} 1 & 0 \\ 0 & e^{\frac{1}{2}i(-\alpha+\beta)} \end{bmatrix}$$

and

$$C := VBV^* = \begin{bmatrix} 0 & ae^{\frac{1}{2}i(-\alpha+\beta)} \\ be^{\frac{1}{2}i(\alpha-\beta)} & 0 \end{bmatrix} = e^{\frac{1}{2}i(\alpha+\beta)} \begin{bmatrix} 0 & |a| \\ |b| & 0 \end{bmatrix}.$$

Moreover, if $|a| < |b|$, we can conjugate by the matrix

$$W := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

so that we may assume that $0 \leq |b| \leq |a|$. Moreover if $|b| = 0$ then the eigenvalues of B and thus A are both zero which is a contradiction. Hence we may assume $0 < |b|$. Furthermore, if $|b| = |a|$, then A is unitarily equivalent to a self-adjoint operator which implies the eigenvectors of A are orthogonal and thus we are done by our first case. Hence we may assume that $0 < |b| < |a|$. Therefore, if

$$D := \begin{bmatrix} 0 & |a| \\ |b| & 0 \end{bmatrix}$$

we obtain by Proposition 2.4 parts (4) and (5) that $W(A) = e^{\frac{1}{2}i(\alpha+\beta)}W(D)$, the eigenvalues of D (specifically $\pm\sqrt{|a||b|}$) are a complex multiple of modulus one of the eigenvalues of A (so the absolute value of the difference of the eigenvalues of D and the absolute value of the difference of the eigenvalues of A agree), and the eigenvectors of D correspond to the eigenvectors of A via a unitary operator (so the inner product of the unit eigenvectors of the two eigenvalues of D and the inner product of the unit eigenvectors of the two eigenvalues of A agree). Hence it suffices to prove the result for D where $0 < |b| < |a|$.

To compute $W(D)$, we note that it is elementary to show that $\xi \in \mathbb{C}^2$ is a unit vector if and only if we can write $\xi = (\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2})$ for some $\theta_j \in [0, 2\pi)$ and $\theta \in [0, \frac{\pi}{2}]$. However

$$\begin{aligned}\langle D(\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2}), (\cos(\theta)e^{i\theta_1}, \sin(\theta)e^{i\theta_2})\rangle &= |a|\sin(\theta)e^{i\theta_2}\cos(\theta)e^{-i\theta_1} + |b|\cos(\theta)e^{i\theta_1}\sin(\theta)e^{-i\theta_2} \\ &= \cos(\theta)\sin(\theta)\left(|a|e^{i(\theta_2-\theta_1)} + |b|e^{-i(\theta_2-\theta_1)}\right).\end{aligned}$$

Therefore, as any $\theta_j \in [0, 2\pi)$ are possible,

$$\begin{aligned}W(D) &= \{\cos(\theta)\sin(\theta)(|a|e^{i\phi} + |b|e^{-i\phi}) \mid \theta \in [0, \tfrac{\pi}{2}], \phi \in [0, 2\pi)\} \\ &= \{\cos(\theta)\sin(\theta)((|a| + |b|)\cos(\phi) + (|a| - |b|)\sin(\phi)) \mid \theta \in [0, \tfrac{\pi}{2}], \phi \in [0, 2\pi)\}.\end{aligned}$$

It is an exercise in elementary geometry that $W(D)$ is then an ellipse centred as 0 with major axis of length $2\left(\frac{1}{2}(|a| + |b|)\right) = |a| + |b|$ (as $\cos(\theta)\sin(\theta)$ obtains every value between 0 and $\frac{1}{2}$ as θ varies from 0 to $\frac{\pi}{2}$), with minor axis of length $|a| - |b|$, and with foci at $\pm\sqrt{\left(\frac{|a|+|b|}{2}\right)^2 - \left(\frac{|a|-|b|}{2}\right)^2} = \pm\sqrt{|a||b|}$ which are the eigenvalues of D. Thus it remains only to demonstrate the eccentricity and length of the major axis are of the above form.

However, it is easy to verify that the only possible unit eigenvectors of $\sqrt{|a||b|}$ and $-\sqrt{|a||b|}$ for D are

$$\zeta_\theta = \frac{e^{i\theta}}{\sqrt{1 + \frac{|b|}{|a|}}}\begin{bmatrix} 1 \\ \sqrt{\frac{|b|}{|a|}} \end{bmatrix} \quad \text{and} \quad \omega_\phi = \frac{e^{i\phi}}{\sqrt{1 + \frac{|b|}{|a|}}}\begin{bmatrix} 1 \\ -\sqrt{\frac{|b|}{|a|}} \end{bmatrix}$$

where $\theta, \phi \in [0, 2\pi)$. Therefore, as

$$|\langle \zeta_\theta, \omega_\phi\rangle| = \frac{1}{1 + \frac{|b|}{|a|}}\left(1 - \frac{|b|}{|a|}\right) = \frac{|a| - |b|}{|a| + |b|}$$

which is the reciprocal of the eccentricity of $W(D)$ and

$$\frac{|\sqrt{|a||b|} - (-\sqrt{|a||b|})|}{\sqrt{1 - |\langle \zeta_\theta, \omega_\phi\rangle|^2}} = \frac{2\sqrt{|a||b|}}{\sqrt{1 - \frac{|a|^2 - 2|a||b| + |b^2}{|a|^2 + 2|a||b| + |b|^2}}} = \frac{2\sqrt{|a||b|}}{\sqrt{\frac{4|a||b|}{(|a|+|b|)^2}}} = |a| + |b|$$

which is the length of the major axis, the proof is complete. ☐

3 Numerical Range - The Toeplitz-Hausdorff Theorem

In this section we shall prove Theorem 3.2, the all-important Toeplitz-Hausdorff Theorem. The proof of this theorem is a surprisingly simple result given Theorem 2.14 as proved in the previous section. Once the Toeplitz-Hausdorff Theorem has been completed, we shall demonstrate two interesting results (of the many) that immediately follow. We also note that the Toeplitz-Hausdorff Theorem will be use in subsequent sections.

We begin with the following trivial lemma that essentially completes the proof of Theorem 3.2.

Lemma 3.1. *Let $\mathcal{H}$ be a Hilbert space and let $\mathcal{L}$ be a closed linear subspace of $\mathcal{H}$. If $P_{\mathcal{L}}$ is the projection onto $\mathcal{L}$ and $T \in \mathcal{B}(\mathcal{H})$, then $W(P_{\mathcal{L}}T|_{\mathcal{L}}) \subseteq W(T)$.*

Proof. It is clear that if $\xi \in \mathcal{L}$ is a unit vector, then $\xi \in \mathcal{H}$ is also a unit vector and

$$\langle P_{\mathcal{L}}T|_{\mathcal{L}}\xi, \xi\rangle_{\mathcal{L}} = \langle T\xi, \xi\rangle_{\mathcal{H}} \in W(T).$$

Hence $W(P_{\mathcal{L}}T|_{\mathcal{L}}) \subseteq W(T)$ as desired. □

Theorem 3.2 (Toeplitz-Hausdorff)**.** *Let $T \in \mathcal{B}(\mathcal{H})$. Then $W(T)$ is convex.*

Proof. Suppose $\alpha, \beta \in W(T)$ are distinct scalars. Then there exists two unit vectors $\xi, \eta \in \mathcal{H}$ such that $\langle T\xi, \xi\rangle = \alpha$ and $\langle T\eta, \eta\rangle = \beta$. If $\xi = \lambda\eta$ for some $\lambda \in \mathbb{C}$, then $|\lambda| = 1$ as $\|\xi\| = 1 = \|\eta\|$ and thus

$$\alpha = \langle T\xi, \xi\rangle = \langle \lambda T\eta, \lambda\eta\rangle = \beta$$

which is a contradiction. Similarly $\eta \neq \lambda\xi$ for any $\lambda \in \mathbb{C}$ so ξ and η are linearly independent.

Let $\mathcal{L} := span\{\xi, \eta\}$ which is a two-dimensional subspace of $\mathcal{H}$. Since $\xi, \eta \in \mathcal{L}$ and by Lemma 3.1,

$$\alpha, \beta \in W(P_{\mathcal{L}}T|_{\mathcal{L}}) \subseteq W(T).$$

However, since $\mathcal{L}$ is two-dimensional, $W(P_{\mathcal{L}}T|_{\mathcal{L}})$ is convex by Theorem 2.14. Hence

$$t\alpha + (1-t)\beta \in W(P_{\mathcal{L}}T|_{\mathcal{L}}) \subseteq W(T)$$

for all $0 < t < 1$. Hence, as $\alpha, \beta \in W(T)$ were arbitrary, $W(T)$ is convex as desired. □

With the proof of Theorem 3.2 complete, we note the following two interesting results that demonstrate the power of the theorem.

Theorem 3.3. *Let $A \in \mathcal{M}_n(\mathbb{C})$ have trace zero. Then A is unitarily equivalent to an matrix whose diagonal entries are all zero.*

Proof. It suffices to show that $\mathbb{C}^n$ has an orthonormal basis $\{\eta_k\}_{k=1}^n$ such that $\langle A\eta_k, \eta_k\rangle = 0$ for all $1 \leq k \leq n$. Let $\{e_k\}_{k=1}^n$ be the standard orthonormal basis of $\mathbb{C}^n$. Therefore, as the trace of A is zero, we obtain that

$$0 = \frac{1}{n}\sum_{k=1}^{n}\langle Ae_k, e_k\rangle.$$

Therefore, since $\langle Ae_k, e_k\rangle \in W(A)$ for all $1 \leq k \leq n$ and since $W(A)$ is convex by Theorem 3.2, $0 \in W(A)$. Hence there exists a unit vector $f_1 \in \mathbb{C}^n$ such that $\langle Af_1, f_1\rangle = 0$.

Extend $\{f_1\}$ to an orthonormal basis $\{f_k\}_{k=1}^n$. Let $\mathcal{L}_1 := span\{f_2, f_3, \ldots, f_n\}$ which is an $(n-1)$-dimensional space. Moreover, since A has trace zero and $\langle Af_1, f_1\rangle = 0$, $P_{\mathcal{L}_1}A|_{\mathcal{L}_1}$ also has trace zero when viewed as an $(n-1)\times(n-1)$ matrix. The result then follows by induction on n by selecting an orthonormal basis $\{\eta_k\}_{k=2}^n$ for $\mathcal{L}_1$ such that $\langle A\eta_k, \eta_k\rangle = 0$ for all $2 \leq k \leq n$. □

Theorem 3.4 (Folk)**.** *Let $T \in \mathcal{B}(\mathcal{H})$ be such that $\lambda \in \partial W(T)$. If no closed disk of $W(T)$ contains λ, then λ is an eigenvalue of T.*

Proof. Let $\lambda \in \partial W(T)$ be such that no closed disk of $W(T)$ contains λ. Choose a unit vector $\xi \in \mathcal{H}$ such that $\langle T\xi, \xi \rangle = \lambda$. Notice if ξ were an eigenvector of T with eigenvalue α, then

$$\alpha = \langle \alpha\xi, \xi \rangle = \langle T\xi, \xi \rangle = \lambda.$$

Suppose ξ is not an eigenvector of T. Then $\mathcal{L} := span\{\xi, T\xi\}$ is a two-dimensional subspace of $\mathcal{H}$. Let $A := P_{\mathcal{L}}T|_{\mathcal{L}}$. Therefore, since $\xi \in \mathcal{L}$,

$$\lambda \in W(A) \subseteq W(T).$$

Moreover, since $\lambda \in \partial W(T)$, $\lambda \in \partial W(A)$. However A is not a multiple of the identity or else $\lambda \in W(A)$ implies $A = \lambda I_{\mathcal{L}}$ by Proposition 2.3 which implies $T\xi = A\xi = \lambda\xi$ which is a contradiction. Hence $W(A)$ is either as described in conclusion (2) or conclusion (3) of Theorem 2.14. Either way, every point of $W(A)$ is then contained in a closed disk contained in $W(A)$. Since $\lambda \in W(A) \subseteq W(T)$, $W(A)$ and thus $W(T)$ contains a closed disk containing λ which is a contradiction. Hence ξ is an eigenvector with eigenvalue λ as desired. $\square$

4 Numerical Range - Hildebrandt's Theorem

In this section we will provide a proof of Hildebrandt's Theorem which relates the spectrum of an operator with the numerical ranges of all operators in the similarity orbit. The proof requires a result due to Rota and this is where we start.

Proposition 4.1 (Rota)**.** *Let $T \in \mathcal{B}(\mathcal{H})$ be such that $\sigma(T)$ is contained in the open unit disk. Then $V := \sum_{k=0}^{\infty}(T^*)^k T^k$ is an invertible positive operator with $\left\| V^{\frac{1}{2}} T V^{-\frac{1}{2}} \right\| < 1$; that is T is similar to a contraction.*

Proof. To see that the series is norm convergent, we notice that

$$\limsup_{k\to\infty} \left\|(T^*)^k T^k\right\|^{\frac{1}{k}} \leq \limsup_{k\to\infty} \left\|T^k\right\|^{\frac{2}{k}} = \rho(T)^2$$

where $\rho(T)$ is the spectral radius of T. Since $\sigma(T)$ is a subset of the open unit disk, $\rho(T) < 1$ so $\limsup_{k\to\infty} \left\|(T^*)^k T^k\right\|^{\frac{1}{k}} < 1$. Hence the sum converges absolutely by the root test.

Clearly V is a positive operator being a sum of positive operators. Moreover, we notice that $T^*VT = V - I_{\mathcal{H}}$. Thus, since $T^*VT \geq 0$, $V \geq I_{\mathcal{H}}$ so V is invertible. Furthermore

$$\left\| V^{\frac{1}{2}} T V^{-\frac{1}{2}} \right\|^2 = \left\| V^{-\frac{1}{2}} T^* V T V^{-\frac{1}{2}} \right\| = \left\| V^{-\frac{1}{2}} (V - I) V^{-\frac{1}{2}} \right\| = \left\| I - V^{-1} \right\|.$$

However, since $V \geq I$, $0 < V^{-1} \leq I$. Thus $\sigma(V^{-1}) \subseteq (0,1]$ so $\sigma(I - V^{-1}) \subseteq [0,1)$ and thus $\left\|I - V^{-1}\right\| < 1$. Hence $\left\| V^{\frac{1}{2}} T V^{-\frac{1}{2}} \right\| < 1$ as desired. $\square$

Theorem 4.2 (Hildebrandt)**.** *Let $T \in \mathcal{B}(\mathcal{H})$. Then*

$$conv(\sigma(T)) = \bigcap \left\{ \overline{W(VTV^{-1})} \mid V \in \mathcal{B}(\mathcal{H}), V \textit{ an invertible operator} \right\}.$$

Proof. We note that $\sigma(VTV^{-1}) = \sigma(T)$ for every invertible operator $V \in \mathcal{B}(\mathcal{H})$. Therefore $\sigma(T) = \sigma(VTV^{-1}) \subseteq \overline{W(VTV^{-1})}$ for all invertible operators $V \in \mathcal{B}(\mathcal{H})$ by Theorem 2.9. Therefore, since the numerical range of any operator is convex by Theorem 3.2, we obtain that $conv(\sigma(T)) \subseteq \overline{W(VTV^{-1})}$ for all invertible operators $V \in \mathcal{B}(\mathcal{H})$. Hence

$$conv(\sigma(T)) \subseteq \bigcap \left\{ \overline{W(VTV^{-1})} \mid V \in \mathcal{B}(\mathcal{H}), V \text{ an invertible operator} \right\}.$$

Suppose to the contrary that the above inclusion is strict. Therefore there exists a $\lambda \in \mathbb{C}$ such that $\lambda \in \overline{W(VTV^{-1})}$ for all invertible operator yet $\lambda \notin conv(\sigma(T))$. Therefore, by translating and scaling using Proposition 2.4 part (4), we may assume that $conv(\sigma(T))$ is a subset of the open unit disk and $|\lambda| \geq 1$. However, since $\sigma(T) \subseteq conv(\sigma(T))$ is a subset of the open unit disk, Proposition 4.1 implies that there exists an invertible element $V \in \mathcal{B}(\mathcal{H})$ such that $\left\|VTV^{-1}\right\| < 1$. Therefore Proposition 2.4 part (3) implies that $\overline{W(VTV^{-1})}$ is a subset of the open unit disk which contradicts the fact that $\lambda \in \overline{W(VTV^{-1})}$ and $|\lambda| \geq 1$. Hence the proof is complete. $\square$

5 Essential Numerical Range

In this section we will develop the essential numerical range of an operator. The role of the essential numerical range in comparison with the essential spectrum mimics the role of the numerical range in comparison with the spectrum. We shall begin with the definition of the essential numerical range and subsequently show that the essential numerical range has an alternate definition that is similar in flavour to the numerical range of an operator (Theorem 5.7), that the essential numerical range is a non-empty, closed, convex set (Corollary 5.8), that the essential numerical range is invariant under approximate unitary equivalence (Lemma 5.9), and that the essential spectrum of an operator is contained within the essential numerical range (Theorem 5.10).

In this section $\mathcal{H}$ will denote an infinite dimensional Hilbert space (the concept of the essential numerical range does not make sense in the finite dimensional setting), $\mathfrak{K}(\mathcal{H})$ will denote the set of compact operators on $\mathcal{H}$, $\sigma_e(T)$ will denote the essential spectrum of an operator $T \in \mathcal{B}(\mathcal{H})$, and $\pi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})/\mathfrak{K}(\mathcal{H})$ will denote the canonical quotient map.

We begin with the definition of the essential numerical range of an operator.

Definition 5.1. Let $T \in \mathcal{B}(\mathcal{H})$. The essential numerical range of T, denoted $W_e(T)$, is the set

$$W_e(T) := \bigcap_{K \in \mathfrak{K}(\mathcal{H})} \overline{W(T+K)}.$$

Remarks 5.2. It is clear that the essential numerical range of an operator is contained in the closure of the numerical range. However, unlike with the numerical range, it is not apparent that the essential numerical range is non-empty.

In order to give examples and develop the theory of the essential numerical range, we need the following result pertaining to compact operators.

Lemma 5.3. *Let $T \in \mathcal{B}(\mathcal{H})$. Then T is compact if and only if $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = 0$ for every orthonormal set $\{\xi_n\}_{n\geq 1}$.*

Proof. Let $T \in \mathcal{B}(\mathcal{H})$ be compact and let $\{\xi_n\}_{n\geq1}$ be an orthonormal set. Let P_n be the projection onto $span\{\xi_1, \ldots, \xi_n\}$. Since T is compact, $\lim_{n\to\infty} \|(I_\mathcal{H} - P_n)T(I_\mathcal{H} - P_n)\| = 0$. Therefore, since

$$|\langle T\xi_{n+1}, \xi_{n+1}\rangle| = |\langle (I_\mathcal{H} - P_n)T(I_\mathcal{H} - P_n)\xi_{n+1}, \xi_{n+1}\rangle| \leq \|(I_\mathcal{H} - P_n)T(I_\mathcal{H} - P_n)\|,$$

we obtain that $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = 0$.

Suppose that $T \in \mathcal{B}(\mathcal{H})$ is such that $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = 0$ for every orthonormal set $\{\xi_n\}_{n\geq1}$. Note that there exists unit vectors $\xi, \eta \in \mathcal{H}$ such that

$$|\langle T\xi, \eta\rangle| \geq \frac{\|T\|}{2}.$$

However, the polarization identity implies that

$$\frac{\|T\|}{2} \leq |\langle T\xi, \eta\rangle| \leq \frac{1}{4}\sum_{k=1}^{4} |i^k \langle T(\xi + i^k\eta), \xi + i^k\eta\rangle|.$$

Hence there exists a k such that

$$|\langle T(\xi + i^k\eta), \xi + i^k\eta\rangle| \geq \frac{\|T\|}{2}.$$

Since $\left\|\xi + i^k\eta\right\| \leq 2$, we obtain that there exists a unit vector $\zeta_1 \in \mathcal{H}$ such that

$$|\langle T\zeta_1, \zeta_1\rangle| \geq \frac{\|T\|}{8}.$$

Let P_1 be the orthogonal projection onto the span of ζ_1. By repeating the above procedure, there exists a unit vector ζ_2 orthogonal to ζ_1 such that

$$|\langle T\zeta_2, \zeta_2\rangle| \geq \frac{\|(I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1)\|}{8}.$$

Whence, by proceeding by recursion, there exists an orthonormal set $\{\zeta_n\}_{n\geq 1}$ such that if P_n is the projection onto the span of $\{\zeta_1, \ldots, \zeta_n\}$ then

$$|\langle T\zeta_n, \zeta_n\rangle| \geq \frac{\|(I_{\mathcal{H}} - P_{n-1})T(I_{\mathcal{H}} - P_{n-1})\|}{8}.$$

Since $\lim_{n\to\infty}\langle T\zeta_n, \zeta_n\rangle = 0$ by assumption,

$$\lim_{n\to\infty} \|(I_{\mathcal{H}} - P_n)T(I_{\mathcal{H}} - P_n)\| = 0$$

and thus T is compact (being the limit of the finite rank operators $-P_nT - TP_n + P_nTP_n$). $\square$

With the above result pertaining to compact operators, we have our first example.

Example 5.4. Notice $W_e(I_{\mathcal{H}}) = \{1\}$. To see this, recall from Example 2.3 that $W(I_{\mathcal{H}}) = \{1\}$ so $W_e(I_{\mathcal{H}}) \subseteq \{1\}$. To verify the other inclusion, let $K \in \mathfrak{K}(\mathcal{H})$ be compact and let $\{e_n\}_{n\geq 1}$ be any orthonormal set. Thus

$$1 = 1 + \lim_{n\to\infty} \langle Ke_n, e_n\rangle = \lim_{n\to\infty} \langle (I_{\mathcal{H}} + K)e_n, e_n\rangle$$

so $1 \in \overline{W(I + K)}$. Since $K \in \mathfrak{K}(\mathcal{H})$ was arbitrary, $W_e(I_{\mathcal{H}}) = \{1\}$.

Using Lemma 5.3 and Proposition 2.4, we obtain the analog of Proposition 2.4 for the essential numerical range.

Proposition 5.5. *Let $T \in \mathcal{B}(\mathcal{H})$. Then*

1. *$W_e(T + K) = W_e(T)$ for all $K \in \mathfrak{K}(\mathcal{H})$.*
2. *$W_e(T^*) = \overline{W_e(T)}$.*
3. *$W_e(T)$ is contained the the close disk of radius $\|\pi(T)\|$ centred around the origin.*
4. *If $a, b \in \mathbb{C}$, $W_e(aT + bI_{\mathcal{H}}) = aW_e(T) + b$.*
5. *If $U \in \mathcal{B}(\mathcal{H})$ is a unitary, then $W_e(UTU^*) = W_e(T)$.*
6. *$W_e(T)$ contains all eigenvalues of T of infinite multiplicity.*

Proof. Property (1) follows trivially from the definition of the essential numerical range.

For Property (2), we notice for every compact operator $K \in \mathfrak{K}(\mathcal{H})$ that $W(T + K) = \overline{W(T^* + K^*)}$ by Proposition 2.4 part (1). Since the adjoint map is an bijection on the set of compact operators, the property easily follows.

For Property (3), we notice for every compact operator K that $W(T + K)$ is contained in the closed disk of radius $\|T + K\|$ centred at the origin by Proposition 2.4 part (3). Hence the property follows.

Property (4) follows trivially from part (4) of Proposition 2.4.

Property (5) follows trivially from part (5) of Proposition 2.4 and the fact that UKU^* is compact if and only if K is compact for any fixed unitary U.

To see Property (6), let λ be an eigenvalue of T with infinite multiplicity. Therefore there exists an orthonormal set $\{e_n\}_{n\geq 1}$ such that $Te_n = \lambda e_n$ for all $n \in \mathbb{N}$. Thus, for any compact operator K,

$$\lambda = \lambda + \lim_{n\to\infty} \langle Ke_n, e_n\rangle = \lim_{n\to\infty} \langle (T + K)e_n, e_n\rangle$$

so $\lambda \in \overline{W(T + K)}$ for every compact operator K. Hence $\lambda \in W_e(T)$ as desired. $\square$

To extend our example set of essential numerical ranges of operators, we have the following corollary that is the essential version of Proposition 2.3.

Corollary 5.6 (Anderson-Stampi)**.** *Let $T \in \mathcal{B}(\mathcal{H})$. Then T is of the form $T = \lambda I_{\mathcal{H}} + K$ where K is compact if and only if $W_e(T) = \{\lambda\}$.*

Proof. It is clear from Example 5.4 and parts (1) and (4) of Proposition 5.5 that

$$W_e(\lambda I_{\mathcal{H}} + K) = W_e(\lambda I_{\mathcal{H}}) = \lambda W_e(I_{\mathcal{H}}) = \{\lambda\}$$

which completes one direction.

Suppose that $T \in \mathcal{B}(\mathcal{H})$ is such that $W_e(T) = \{\lambda\}$ for some $\lambda \in \mathbb{C}$. Then,

$$W_e(T - \lambda I_{\mathcal{H}}) = \{0\}$$

by part (4) of Proposition 5.5. Therefore, by part (3) of Proposition 5.5, $\|\pi(T - \lambda I_{\mathcal{H}})\| = 0$. Therefore $T - \lambda I_{\mathcal{H}}$ is compact and thus there exists a compact operator $K \in \mathfrak{K}(\mathcal{H})$ such that $T = \lambda I_{\mathcal{H}} + K$ as desired. □

To continue our study of the essential numerical range of an operator, it is vital to develop the following alternate definition of the essential numerical range of an operator. Note how this corresponding definition relates to the definition of the numerical range.

Theorem 5.7. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $\lambda \in W_e(T)$ if and only if there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = \lambda$.*

Proof. Suppose there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = \lambda$. Then, for every compact operator $K \in \mathfrak{K}(\mathcal{H})$,

$$\lambda = \lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = \lim_{n\to\infty}\langle (T+K)\xi_n, \xi_n\rangle \in \overline{W(T+K)}$$

by Lemma 5.3. Therefore, since K was arbitrary, $\lambda \in W_e(T)$.

Let $\lambda \in W_e(T)$. Therefore $\lambda \in \overline{W(T)}$ so there exists a unit vector ξ_1 such that

$$|\langle T\xi_1, \xi_1\rangle - \lambda| \leq \frac{1}{2}.$$

Let $\mathcal{L}_1 := span\{\xi_1\}$, let P_1 be the orthogonal projection onto $\mathcal{L}_1$, let $\mu_1 \in W((I_{\mathcal{H}} - P_1)T|_{\mathcal{L}_1^\perp})$, and let

$$F_1 := \mu_1 P_1 - P_1 T P_1 - P_1 T(I_{\mathcal{H}} - P_1) - (I_{\mathcal{H}} - P_1)TP_1.$$

Clearly F_1 is a finite rank operator on $\mathcal{H}$ and thus is compact. Therefore

$$\lambda \in \overline{W(T + F_1)} = \overline{W(\mu_1 P_1 + (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1))}.$$

However, it is clear that

$$\begin{aligned}
& W(\mu_1 P_1 + (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1)) \\
= \;& \{\langle(\mu_1 P_1 + (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1))\eta, \eta\rangle \mid \eta \in \mathcal{H}, \|\eta\| = 1\} \\
= \;& \{\mu_1 \|\eta_1\|^2 + \langle (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1))\eta_2, \eta_2\rangle \mid \eta_1 \in \mathcal{L}_1, \eta_2 \in \mathcal{L}_1^\perp, \|\eta_1\|^2 + \|\eta_2\|^2 = 1\} \\
= \;& \{\mu_1 \|\eta_1\|^2 + \|\eta_2\|^2 \langle (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1))\tfrac{1}{\|\eta_2\|}\eta_2, \tfrac{1}{\|\eta_2\|}\eta_2\rangle \mid \eta_1 \in \mathcal{L}_1, \eta_2 \in \mathcal{L}_1^\perp, \|\eta_1\|^2 + \|\eta_2\|^2 = 1\}.
\end{aligned}$$

However, since $\mu_1 \in W((I_{\mathcal{H}} - P_1)T|_{\mathcal{L}_1^\perp})$ and $W((I_{\mathcal{H}} - P_1)T|_{\mathcal{L}_1^\perp})$ is convex by Theorem 3.2 we obtain that $W(\mu_1 P_1 + (I_{\mathcal{H}} - P_1)T(I_{\mathcal{H}} - P_1)) = W((I_{\mathcal{H}} - P_1)T|_{\mathcal{L}_1^\perp})$. Hence $\lambda \in \overline{W((I_{\mathcal{H}} - P_1)T|_{\mathcal{L}_1^\perp})}$ so there exists a unit vector $\xi_2 \in \mathcal{L}_1^\perp$ (so ξ_2 is orthogonal to ξ_1) such that

$$|\langle T\xi_2, \xi_2\rangle - \lambda| \leq \frac{1}{2^2}.$$

If $\xi_1, \dots, \xi_n$ are orthonormal vectors such that

$$|\langle T\xi_n, \xi_n\rangle - \lambda| \leq \frac{1}{2^n}$$

we can repeat the above procedure with $\mathcal{L}_n := span\{\xi_1, \dots, \xi_n\}$, P_n the orthogonal projection onto $\mathcal{L}_n$, $\mu_n \in W((I_\mathcal{H} - P_n)T|_{\mathcal{L}_n^\perp})$, and

$$F_n := \mu_n P_n - P_n T P_n - P_n T(I_\mathcal{H} - P_n) - (I_\mathcal{H} - P_n)TP_n$$

to obtain a unit vector ξ_{n+1} orthogonal to each ξ_k for $1 \leq k \leq n$ such that

$$|\langle T\xi_{n+1}, \xi_{n+1}\rangle - \lambda| \leq \frac{1}{2^{n+1}}.$$

Hence, by recursion, there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ in $\mathcal{H}$ such that $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = \lambda$ as desired. $\square$

With the above alternate definition of the essential numerical range of an operator, we obtain the following results that are the essential versions of Theorem 3.2, Theorem 2.10, Theorem 2.9, and Theorem 2.11 respectively.

Corollary 5.8. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $W_e(T)$ is a non-empty, compact, convex set.*

Proof. To see that $W_e(T)$ is non-empty, let $(\xi_n)_{n\geq 1}$ be an orthonormal sequence. Therefore $(\langle T\xi_n, \xi_n\rangle)_{n\geq 1}$ is a sequence of complex numbers bounded by $\|T\|$ and thus has a convergent subsequence. Hence $W_e(T)$ is non-empty by Theorem 5.7. Moreover, it is clear that $W_e(T)$ is closed and convex being the intersection of closed, convex sets. $\square$

Lemma 5.9. *Let $T, S \in \mathcal{B}(\mathcal{H})$ be two operators that are approximately unitarily equivalent. Then $W_e(S) = W_e(T)$.*

Proof. Since S and T are approximately unitarily equivalent, there exists a sequence $(U_n)_{n\geq 1}$ of unitaries such that

$$\lim_{n\to\infty} \|U_n T U_n^* - S\| = 0 = \lim_{n\to\infty} \|T - U_n^* S U_n\|.$$

Let $\lambda \in W_e(T)$. Then, by Theorem 5.7, there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lambda = \lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle$. Since $(\langle U_m^* S U_m \xi_n, \xi_n\rangle)_{n\geq 1}$ is a bounded sequence of complex numbers, for each fixed $m \in \mathbb{N}$ there exists a convergent subsequence $(\langle (U_m^* S U_m - T)\xi_{n_k}, \xi_{n_k}\rangle)_{k\geq 1}$ of $(\langle U_m^* S U_m \xi_n, \xi_n\rangle)_{n\geq 1}$ that converges to a complex number μ_m with absolute value at most $\|U_m^* S U_m - T\|$. Thus

$$\lambda + \mu_m = \lim_{k\to\infty}\langle T\xi_{n_k}, \xi_{n_k}\rangle + \langle (U_m^* S U_m - T)\xi_{n_k}, \xi_{n_k}\rangle = \lim_{k\to\infty}\langle U_m^* S U_m \xi_{n_k}, \xi_{n_k}\rangle \in W_e(U_m^* S U_m).$$

However, $W_e(U_m^* S U_m) = W_e(S)$ by Proposition 5.5 part (5). Therefore, since $\lim_{m\to\infty}\mu_m = 0$ and $W_e(S)$ is closed, $\lambda \in W_e(S)$. Hence $W_e(T) \subseteq W_e(S)$. The reverse inclusion the follows by symmetry. $\square$

Theorem 5.10. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $\sigma_e(T) \subseteq W_e(T)$.*

Proof. Let $\lambda \in \sigma_e(T)$ and let $S := \lambda I_\mathcal{H} - T$. There are three cases: the range of S is not closed, the kernel of S is infinite dimensional, or the kernel of S^* is infinite dimensional.

If the range of S is not closed, then S is not bounded below on the orthogonal complement of $ker(S)$. Let $\mathcal{L} := ker(S)^\perp$. Then there exists a unit vector $\xi_1 \in \mathcal{L}$ such that $\|S\xi_1\| \leq 1$. Then, since S is not bounded below, there must exist a unit vector $\xi_2 \in \mathcal{L}$ orthogonal to ξ_1 such that $\|S\xi_2\| \leq \frac{1}{2}$. By repeating this process ad nauseum, we obtain an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lim_{n\to\infty}\|S\xi_n\| = 0$. Thus $\lambda \in W_e(T)$ by Theorem 5.7.

If the kernel of S is infinite dimensional, it is easy to construct an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\langle S\xi_n, \xi_n\rangle = 0$ for all n. Hence $\lambda \in W_e(T)$ by Theorem 5.7. Similarly if the kernel of S^* is infinite dimensional then $\overline{\lambda} \in W_e(T^*)$ by Theorem 5.7. Thus $\lambda \in W_e(T)$ by Proposition 5.5 part (2) as desired. $\square$

Corollary 5.11. *Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator on an infinite dimensional Hilbert space $\mathcal{H}$. Then $W_e(N) = conv(\sigma_e(N))$.*

Proof. By Theorem 5.10, $\sigma_e(N) \subseteq W_e(N)$. Thus $conv(\sigma_e(N)) \subseteq W_e(N)$ by Corollary 5.8.

To see the other inclusion, we recall that there exists a diagonal normal operator D with $\sigma(D) = \sigma(N)$ and $\sigma_e(D) = \sigma_e(N)$ such that N and D are approximately unitarily equivalent. Hence $W_e(D) = W_e(N)$ and $conv(\sigma_e(D)) = conv(\sigma_e(N))$. Hence it suffices to show that $W_e(D) \subseteq conv(\sigma_e(D))$.

Recall from Example 2.7 that the numerical range of a diagonal operator is the convex hull of the diagonal entries. For each $\epsilon > 0$, the number diagonal entries of D that lie outside $\sigma_e(D)$ is finite and thus there exists a diagonal compact operator K_ϵ such that $\overline{W(D + K_\epsilon)}$ is the closed convex hull of the diagonal entries of D that lie within ϵ of $\sigma_e(D)$. Thus, as $W_e(D)$ is contained in the intersection of all such $\overline{W(D + K_\epsilon)}$, $W_e(D)$ is contained in the closed convex hull of the diagonal entries of D that are in $\sigma_e(D)$. Hence $W_e(D)$ is contained in the closed convex hull of $\sigma_e(D)$. As $\sigma_e(D)$ is compact, $\overline{conv(\sigma_e(D))} = conv(\sigma_e(D))$ and thus $W_e(D) \subseteq \sigma_e(D)$ as desired. $\square$

6 Essential Numerical Radius

In this section we will use the essential numerical range from Section 5 to obtain a notion of the essential numerical radius of an operator. The results for the essential numerical radius will mimic those for the numerical radius by our results in Section 5.

In this section $\mathcal{H}$ will denote an infinite dimensional Hilbert space, $\mathfrak{K}(\mathcal{H})$ will denote the set of compact operators on $\mathcal{H}$, $\sigma_e(T)$ will denote the essential spectrum of an operator $T \in \mathcal{B}(\mathcal{H})$, and $\pi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})/\mathfrak{K}(\mathcal{H})$ will denote the canonical quotient map.

We begin with the definition of the essential numerical radius.

Definition 6.1. Let $T \in \mathcal{B}(\mathcal{H})$. The essential numerical radius of T, denoted $nr_e(T)$, is

$$nr_e(T) := \sup\{|\lambda| \mid \lambda \in W_e(T)\}.$$

For our first example of the essential numerical radius of an operator, we appeal to Corollary 5.11 to obtain the essential version of Theorem 1.7.

Example 6.2. Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator. Then $nr_e(N) = \|\pi(N)\|$. Indeed by Corollary 5.11 $W_e(N) = conv(\sigma_e(N))$. Thus

$$nr_e(T) = \sup\{|\lambda| \mid \lambda \in conv(\sigma_e(N))\} = \sup\{|\lambda| \mid \lambda \in \sigma_e(N)\} = \|\pi(N)\|$$

(as $\pi(N)$ is normal) as desired.

Our main goal of this section is to show that the essential numerical radius is a norm on the Calkin algebra equivalent to the essential norm as done with the numerical radius and the operator norm in Theorem 1.4. This is done via the following lemma that mimics Lemma 1.3.

Lemma 6.3. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $nr_e(Re(T)) \leq nr_e(T)$ and $nr_e(Im(T)) \leq nr_e(T)$.*

Proof. Let $\lambda \in W_e(Re(T))$. By Theorem 5.7 there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lambda = \lim_{n\to\infty}\langle Re(T)\xi_n, \xi_n\rangle = \lim_{n\to\infty} Re(\langle T\xi_n, \xi_n\rangle) \in \mathbb{R}$. Since $(Im(\langle T\xi_n, \xi_n\rangle))_{n\geq 1}$ is a sequence bounded by $\|T\|$, there exists a subsequence $(Im(\langle T\xi_{n_k}, \xi_{n_k}\rangle))_{k\geq 1}$ that converges to some real number μ. Therefore

$$\lambda + i\mu = \lim_{k\to\infty}\langle T\xi_{n_k}, \xi_{n_k}\rangle \in nr(T)$$

by Theorem 5.7. Therefore, as $\lambda, \mu \in \mathbb{R}$, $|\lambda| \leq |\lambda + i\mu| \leq nr_e(T)$. Hence, as $\lambda \in W_e(Re(T))$ was arbitrary, $nr_e(Re(T)) \leq nr_e(T)$.

The proof that $nr_e(Im(T)) \leq nr_e(T)$ is identical. □

Proposition 6.4. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $nr_e(T) \leq \|\pi(T)\| \leq 2nr_e(T)$. Moreover, $\pi(T) \mapsto nr_e(T)$ is a norm on the Calkin algebra.*

Proof. Fix $T \in \mathbb{B}(\mathcal{H})$. Then, by Example 6.2 and Lemma 6.3,

$$\|\pi(T)\| \leq \|\pi(Re(T))\| + \|\pi(Im(T))\| \leq nr_e(Re(T)) + nr_e(Im(T)) \leq 2nr_e(T).$$

The other inequality follows trivially from Proposition 5.5 part (3).

It is trivial to verify that $\pi(T) \mapsto nr_e(T)$ is a well-defined as $W_e(T) = W_e(T + K)$ for all compact operators $K \in \mathfrak{K}(\mathcal{H})$ and all $T \in \mathcal{B}(\mathcal{H})$ by Proposition 5.5 part (1). Clearly $nr_e(T) \geq 0$ for all $T \in \mathcal{B}(\mathcal{H})$. Furthermore $nr_e(\lambda T) = |\lambda| nr_e(T)$ for all $T \in \mathcal{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$ by Proposition 5.5 part (4).

To see that $nr_e(T + S) \leq nr_e(T) + nr_e(S)$ for all $S, T \in \mathcal{B}(\mathcal{H})$, suppose $\lambda \in W_e(T + S)$. By Theorem 5.7 there exists an orthonormal sequence $(\xi_n)_{n\geq 1}$ such that $\lambda = \lim_{n\to\infty}\langle (T+S)\xi_n, \xi_n\rangle$. Since $(\langle T\xi_n, \xi_n\rangle)_{n\geq 1}$ and $(\langle S\xi_n, \xi_n\rangle)_{n\geq 1}$ are bounded sequences, there exists a subsequence $(n_k)_{k\geq 1}$ such that $\lambda_T := \lim_{k\to\infty}\langle T\xi_{n_k}, \xi_{n_k}\rangle$ and $\lambda_S := \lim_{k\to\infty}\langle S\xi_{n_k}, \xi_{n_k}\rangle$ exist. Hence $\lambda_T \in W_e(T)$ and $\lambda_S \in W_e(S)$ by Theorem 5.7 are such that

$$|\lambda| = |\lambda_T + \lambda_S| \leq |\lambda_T| + |\lambda_S| \leq nr_e(T) + nr_e(S).$$

Therefore, since the above holds for all $\lambda \in W_e(T+S)$, $nr_e(T+S) \leq nr_e(T) + nr_e(S)$ as desired.

Finally, if $nr_e(T) = 0$ then $\|\pi(T)\| \leq 2nr_e(T) = 0$ so $\pi(T) = 0$. Hence nr_e defines a norm on the Calkin algebra. $\square$

Note the above proof also shows that $W_e(T+S) \subseteq W_e(T) + W_e(S)$ for all $T, S \in \mathcal{B}(\mathcal{H})$.

7 Maximal Numerical Range

In this section we will develop the notion of the maximal numerical range of an operator. In particular, we will show that the maximal numerical range is non-empty, closed, and convex. In some sense, the maximal numerical range contains only asymptotic extremes of the numerical range which can be seen in the following definition.

Definition 7.1. Let $T \in \mathcal{B}(\mathcal{H})$. The maximal numerical range of T, denoted $W_0(T)$, is the set

$$W_0(T) := \left\{ \lambda \in \mathbb{C} \mid \exists (\xi_n)_{n\geq 1} \subseteq \mathcal{H} \text{ such that } \|\xi_n\| = 1, \lim_{n\to\infty} \langle T\xi_n, \xi_n\rangle = \lambda, \lim_{n\to\infty} \|T\xi_n\| = \|T\| \right\}.$$

Remarks 7.2. It is clear that $W_0(T)$ is contained in the closed disk of radius $\|T\|$ centred at the origin, $W_0(\lambda I_{\mathcal{H}}) = \{\lambda\}$, and that $W_0(\lambda T) = \lambda W_0(T)$ for all $\lambda \in \mathbb{C}$ and $T \in \mathcal{B}(\mathcal{H})$. Moreover, it is clear that $W_0(T) \subseteq \overline{W(T)}$.

We shall split the proof that the maximal numerical range is a non-empty, closed, convex set into two parts. Again the fact that the maximal numerical range is convex follows from Theorem 3.2.

Lemma 7.3. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $W_0(T)$ is non-empty and closed.*

Proof. To see that $W_0(T)$ is non-empty, we note by the definition of the operator norm that there exists a sequence of unit vectors $(\xi_n)_{n\geq 1}$ such that

$$\lim_{n\to\infty} \|T\xi_n\| = \|T\|.$$

Therefore, since $(\langle T\xi_n, \xi_n\rangle)_{n\geq 1}$ is a bounded sequence, there exists a subsequence that converges to some $\lambda \in \mathbb{C}$. Thus $\lambda \in W_0(T)$ and thus $W_0(T)$ is non-empty.

Let $(\lambda_n)_{n\geq 1}$ be a sequence in $W_0(T)$ that converges to some $\lambda \in \mathbb{C}$. By the definition of $W_0(T)$, for each $n \in \mathbb{N}$ there exists a unit vector ξ_n such that $\|T\| \leq \|T\xi_n\| + \frac{1}{n}$ and $|\lambda_n - \langle T\xi_n, \xi_n\rangle| < \frac{1}{n}$. Therefore it is easy to see that $\lim_{n\to\infty} \|T\xi_n\| = \|T\|$ and $\lim_{n\to\infty}\langle T\xi_n, \xi_n\rangle = \lambda$. Hence $\lambda \in W_0(T)$. $\square$

Theorem 7.4. *Let $T \in \mathcal{B}(\mathcal{H})$. Then $W_0(T)$ is convex.*

Proof. Without loss of generality we may assume that $\|T\| = 1$ by Remarks 7.2. To see that $W_0(T)$ is convex, let $\lambda, \mu \in W_0(T)$ be distinct. Therefore there exists sequences of unit vectors $(\xi_n)_{n\geq 1}$ and $(\eta_n)_{n\geq 1}$ such that

$$\lim_{n\to\infty} \|T\xi_n\| = 1 = \lim_{n\to\infty} \|T\eta_n\|, \lim_{n\to\infty} \langle T\xi_n, \xi_n\rangle = \lambda, \text{ and } \lim_{n\to\infty} \langle T\eta_n, \eta_n\rangle = \mu.$$

First we claim that there exists a constant $K \in \mathbb{R}$ such that $|\langle \xi_n, \eta_n\rangle| \leq K < 1$ for all n sufficiently large. To see this, suppose otherwise. Then, by replacing $(\xi_n)_{n\geq 1}$ and $(\eta_n)_{n\geq 1}$ with subsequences, we may assume that $\lim_{n\to\infty} |\langle \xi_n, \eta_n\rangle| = 1$. However, if for each $n \in \mathbb{N}$ we write $\eta_n = \alpha_n \xi_n + \zeta_n$ where ζ_n is orthogonal to ξ_n, then $\lim_{n\to\infty} |\langle \xi_n, \eta_n\rangle| = 1$ implies that $\lim_{n\to\infty} |\alpha_n| = 1$ and thus $\lim_{n\to\infty} \|\zeta_n\| = 0$ as $\|\xi_n\| = 1 = \|\eta_n\|$ for all $n \in \mathbb{N}$. Therefore

$$\begin{aligned}
\mu &= \lim_{n\to\infty}\langle T\eta_n, \eta_n\rangle \\
&= \lim_{n\to\infty}\langle T(\alpha_n\xi_n + \zeta_n), (\alpha_n\xi_n + \zeta_n)\rangle \\
&= \lim_{n\to\infty} |\alpha_n|^2\langle T\xi_n, \xi_n\rangle + \alpha_n\langle T\xi_n, \zeta_n\rangle + \overline{\alpha_n}\langle T\zeta_n, \xi_n\rangle + \langle T\zeta_n, \zeta_n\rangle \\
&= \lambda + 0 + 0 + 0 = \lambda
\end{aligned}$$

which is a contradiction. Hence the claim has been proved.

Let β be any complex number that is a convex combination of λ and μ. For each $n \in \mathbb{N}$ let P_n be the projection onto $span\{\xi_n, \eta_n\}$. Therefore, by the proof of Theorem 3.2 and since

$$\lim_{n\to\infty} \langle T\xi_n, \xi_n\rangle = \lambda, \text{ and } \lim_{n\to\infty} \langle T\eta_n, \eta_n\rangle = \mu,$$

we can select unit vectors $\omega_n \in span\{\xi_n, \eta_n\}$ such that $\langle T\omega_n, \omega_n\rangle$ is a convex combination of $\langle T\xi_n, \xi_n\rangle$ and $\langle T\eta_n, \eta_n\rangle$ and $\lim_{n\to\infty}\langle T\omega_n, \omega_n\rangle = \beta$. Thus, to show that $\beta \in W_0(T)$, it suffices to show that $\lim_{n\to\infty} \|T\omega_n\| = \|T\| = 1$.

Since $\omega_n \in span\{\xi_n, \eta_n\}$ for each $n \in \mathbb{N}$, there exists scalars $a_n, b_n \in \mathbb{N}$ such that $\omega_n = a_n\xi_n + b_n\eta_n$. Let $M := \frac{1}{\sqrt{1-K^2}}$. Then we claim that $|a_n|, |b_n| \leq M$ for sufficiently large n. Indeed we notice that

$$1 = \|\omega_n\|^2 = |a_n|^2 + |b_n|^2 + 2Re(a_n\overline{b_n}\langle\xi_n, \eta_n\rangle)$$

so

$$1 \geq |a_n|^2 + |b_n|^2 - 2K|a_n||b_n|$$

for sufficiently large n. This implies that a triangle of side lengths $|a_n|$ and $|b_n|$ with an angle of $\cos^{-1}(K)$ in between those sides has a third side length of at most 1. In terms of an upper bound for $|a_n|$, this will occur when the side opposite to the angle $\cos^{-1}(K)$ is perpendicular to the third side and of length 1. In this case, $\sin(\cos^{-1}(K)) = \frac{1}{|a_n|}$ so the maximum of $|a_n|$ is $\frac{1}{\sin(\cos^{-1}(K))} = \frac{1}{\sqrt{1-K^2}} = M$ via some elementary geometry. Similarly $|b_n| \leq M$ as desired.

However, this implies that

$$\begin{aligned} 1 &\geq \|T\omega_n\|^2 \\ &= \langle T^*T\omega_n, \omega_n\rangle \\ &= \|\omega_n\|^2 - \langle(I_{\mathcal{H}} - T^*T)\omega_n, \omega_n\rangle \\ &= 1 - a_n\langle(I_{\mathcal{H}} - T^*T)\xi_n, \omega_n\rangle - b_n\langle(I_{\mathcal{H}} - T^*T)\eta_n, \omega_n\rangle. \end{aligned}$$

However, since

$$\begin{aligned} 0 &\leq \|(I_{\mathcal{H}} - T^*T)\xi_n\|^2 \\ &= \|\xi_n\| - 2\|T\xi_n\|^2 + \|T^*T\xi_n\|^2 \\ &\leq 1 - \|T\xi_n\|^2 \end{aligned}$$

and similarly $0 \leq \|(I_{\mathcal{H}} - T^*T)\eta_n\|^2 \leq 1 - \|T\eta_n\|^2$, we obtain that

$$\lim_{n\to\infty} \|(I_{\mathcal{H}} - T^*T)\xi_n\| = 0 = \lim_{n\to\infty} \|(I_{\mathcal{H}} - T^*T)\eta_n\|.$$

However, since each ω_n is a unit vector and $|a_n|, |b_n| \leq M$ for sufficiently large n, the above implies that $\lim_{n\to\infty} \|T\omega_n\| = 1$ as desired. $\square$

8 C*-Numerical Range

To complete this document, we will briefly outline the generalization of the numerical range of an operator in $\mathcal{B}(\mathcal{H})$ to an operator in a C*-algebra. As the only proof follows trivially from C*-algebra theory, the proof is omitted.

Definition 8.1. Let $\mathfrak{A}$ be a C*-algebra and let $\mathcal{S}(\mathfrak{A})$ be the state space of $\mathfrak{A}$. If $A \in \mathfrak{A}$, the numerical range of A in $\mathfrak{A}$, denoted $V(A)$, is the set

$$V(A) := \{\varphi(A) \mid \varphi \in \mathcal{S}(\mathfrak{A})\}.$$

Remarks 8.2. Recall that $\varphi \in \mathcal{S}(\mathfrak{A})$ if and only if there exists a non-degenerate *-homomorphism $\pi : \mathfrak{A} \to \mathcal{B}(\mathcal{H})$ and a unit vector $\xi \in \mathcal{H}$ such that $\varphi(A) = \langle \pi(A)\xi, \xi \rangle$. Therefore

$$V(A) = \bigcup_{\substack{\pi : \mathfrak{A} \to \mathcal{B}(\mathcal{H}) \\ \pi \text{ a non-degenerate } ^*\text{-homomorphism}}} W(\pi(A))$$

for every $A \in \mathfrak{A}$.

Proposition 8.3. *Let $\mathfrak{A}$ be a C*-algebra and let $A \in \mathfrak{A}$. Then $V(A)$ is a non-empty, convex, compact set contained in the ball of radius $\|A\|$ centred at the origin.*

Proof. This result follows trivially from the theory of the state space of a C*-algebra. $\square$

References

[1] J. H. Anderson and J. G. Stamepfli, *Commutators and Compressions*, Israel J. Math., 10 (1970), 433–441

[2] D. A. Herrero, *Approximation of Hilbert Space Operators, Volume I*, Research Notes in Mathematics, Pitman Books Limited, (1982).

[3] J. Shapiro, *Notes on the Numerical Range*, http://www.mth.msu.edu/~shapiro/pubvit/downloads/numrangenotes/numrange_notes.pdf, (2003).

[4] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, Cambridge Studies in Advanced Mathematics, (2002).