Introduction.

We are all familiar with vector spaces and their basic properties. We understand linear operators, dot products, and maybe even norms. What the reader may not be as familiar with, however, is the concept of numerical range. For any operator $T : (V, < \cdot, \cdot >) \to V$, where V is a vector space with inner product $< \cdot, \cdot >$, we define the numerical range of T as:

$$W(T) = \{< Tv, v >: v \in V, \|v\| = 1\}.$$

This definition holds in both finite and infinite dimensions. One should note that the numerical range of T is a collection of scalars, unlike the range of T, which is a collection of vectors.

This thesis mainly concerns numerical ranges of linear operators acting on Hilbert spaces. Informally speaking, a Hilbert space H is an inner product space in which Cauchy sequences converge. We will give a rigorous definition later in the paper. Numerical ranges have many interesting properties. For instance, if $T : \mathbb{C}^2 \to \mathbb{C}^2$, $W(T)$ will always be an ellipse! Perhaps even more interestingly, if T happens to be self-adjoint, $W(T)$ will be a line segment regardless of the space on which T operates.

But really, who cares? Beyond mathematical fun, why should we spend time studying numerical range? It can be quite difficult to calculate, often seemingly impossible. Is it worth the bother? Actually, it's a matter of national security! While that is a bit of an overstatement, it is nonetheless true that numerical range has ties to quantum computing. In particular, it informs the idea of quantum error correction (see, e.g. [4, Thm. 10.1]).

In 2001, Professor John Terilla proposed a problem based on his work on quantum error correction: if one is given a set of linear operators $S := \{T | T : H \to H\}$, can one find a subspace M of large dimension such that when $T \in S$ is compressed to M, the resulting operator on M will have a single point numerical range? Further, how can one maximize the dimension of such a subspace M? It is this question that interests those studying quantum error correction.

In this paper, we will not address this question in terms of sets of operators. Rather, we will study the case of the numerical range of a single operator $T : H \to H$ when compressed to a subspace M. By "compressed to a subspace M," we mean that T takes M as its domain, and for $m \in M$, Tm is projected onto M, resulting in an operator mapping M into itself. We denote this compression by T_M. To rephrase our problem using this new notation, we are studying operators $T : H \to H$ and subspaces $M \subseteq H$ such that $W(T_M)$ is a single point. We will scrutinize the properties of T that affect $W(T_M)$.

In particular, we are going to study what effect the eigenvalues of T have on $W(T_M)$. Eigenvalues are very closely related to the study of numerical range. For example, if $M = H$ and $W(T_M) = \{\alpha\}$, we find that $T = \alpha I$, so α is an eigenvalue of T. Even if M does not equal H, so long as $W(T_M) = \{\alpha\}$, we find that α is an eigenvalue of T_M.

These are fine observations, but how do they help us find appropriate subspaces M of large dimension? It turns out that knowledge of the eigenvalues of T helps us put upper bounds on the dimension of any M such that $W(T_M) = \{\alpha\}$. We will find that if $W(T_M) = \{\alpha\}$ and $\dim M > \frac{1}{2} \dim H$, then α must be an eigenvalue of T. Further, if $W(T_M) = \{\alpha\}$ for α an eigenvalue, the multiplicity of α helps us determine a bound on the dimension of M.

While bounding is helpful, it does not always confirm when we have found an M of largest dimension for which $W(T_M)$ is a single point. At the conclusion of this paper, we will prove a theorem that, once we have found one M, will help us augment that M to a subspace of even larger dimension. In this way, we can approach the upper bound.

We will begin the paper with a review of concepts from linear algebra. These concepts lay the groundwork for our more complex results. We will also introduce the concept of operator norm, since it will be useful in a study of numerical range. For $T : H \to H$, we define the norm of T as

$$\|T\| = \sup\{\|Tv\| : v \in H, \|v\| = 1\}.$$

We will study this concept further in the body of the paper. From there, we will move into a general introduction of numerical range, with analysis of some of its important properties. Finally, we will address Professor Terilla's problem and conclude with an example in which we use our augmentation theorem.

Background.

Before encountering the real work of the paper, the reader should recall a few simple definitions and ideas from linear algebra:

Definition 1. *An inner product space $(V, < \cdot, \cdot >)$ is a vector space V over the complex field on which there is a function $< \cdot, \cdot >: V \times V \to \mathbb{C}$, called an inner product, satisfying the following five conditions:*

1. $< v, v > \geq 0 \; \forall \, v \in V$;
2. $< v, v >= 0$ *if and only if* $v = 0$;
3. $< \lambda v, w >= \lambda < v, w > \forall \, v, w \in V$ *and* $\lambda \in \mathbb{C}$;
4. $< v, w >= \overline{< w, v >} \; \forall \, v, w \in V$;
5. $< v + w, u >=< v, u > + < w, u > \forall \, u, v, w \in V$.

We have chosen to work with complex inner product spaces, since this thesis will focus in that area. The reader should be aware, however, that one can also work in real inner product spaces. The only differences are that $< \cdot, \cdot >$ maps into $\mathbb{R}$ and condition (4) becomes $< v, w >=< w, v >$ for all $v, w \in V$.

As a specific example of a complex inner product, recall that in $\mathbb{C}^n$, we define the inner product of two vectors $v = (v_1, v_2, \cdots, v_n)$ and $w = (w_1, w_2, \cdots, w_n)$ as

$$< v, w >= v_1\overline{w_1} + v_2\overline{w_2} + \cdots + v_n\overline{w_n}.$$

We can also define an inner product on the space of complex-valued continuous functions on [0,1] as

$$< f, g >= \int_0^1 f(x)\overline{g(x)}dx.$$

While there are many ways to define an inner product, they all share certain characteristics. For instance, they all induce norms:

Definition 2. *A normed linear space $(V, \| \cdot \|)$ is a linear space V together with a function $\| \cdot \|: V \to \mathbb{R}$ called a norm satisfying the following four properties:*

1. $\| v \| \geq 0 \; \forall \, v \in V$;
2. $\| v \| = 0$ *if and only if* $v = 0$;
3. $\| \lambda v \| = |\lambda| \, \| v \|$ *for all* $v \in V$ *and* $\lambda \in \mathbb{C}$;
4. $\| v + w \| \leq \| v \| + \| w \|$ *for all* $v, w \in V$ (triangle inequality).

Theorem 1. *An inner product will always give rise to a norm. If V is an inner-product space, then we can define a norm on V by $\| v \| = \sqrt{< v, v >}$ for all $v \in V$.*

It is easy to show that the norm arising from the inner product satisfies the first three properties of the norm. To see that it satisfies the last property, we will need the following theorem:

Theorem 2. *Cauchy-Schwarz Inequality: Let $v, w \in V$, where V is an inner product space. Then*

$$|< v, w >| \leq \| v \| \| w \|,$$

where equality holds if and only if one vector is a scalar multiple of the other.

Proof. We will first prove the inequality, relying on a proof in [5]. Let $w \in V$. If $w = 0$, the result follows immediately. Suppose $w \neq 0$ and suppose for now that $\| w \| = 1$. We have

$$\begin{aligned}
0 &\leq \| v - <v,w> w \|^2 \\
&= <v - <v,w> w, v - <v,w> w> \\
&= <v,v> - <v,w> \overline{<v,w>} - <v,w> \overline{<v,w>} + <v,w> \overline{<v,w>} <w,w> \\
&= <v,v> - <v,w> \overline{<v,w>} \\
&= \| v \|^2 - |<v,w>|^2 . \quad (*)
\end{aligned}$$

Therefore, we have $|<v,w>|^2 \leq \| v \|^2$. Now pick an arbitrary $p \in V \backslash \{0\}$. Therefore, $\| p \| \neq 0$, so we can let $u = \frac{p}{\|p\|}$. Thus, $\| u \| = 1$. By the first part of the proof, we then have $|<v,u>| \leq \| v \|$. Since $|<v,u>| = \frac{|<v,p>|}{\|p\|}$, the result follows.

To prove the equality clause, suppose there exists $\alpha \in \mathbb{C}$ such that $v = \alpha w$. Then

$$\begin{aligned}
|<v,w>| &= |<\alpha w, w>| \\
&= |\alpha <w,w>| \\
&= |\alpha| \| w \|^2 \\
&= |\alpha| \| w \| \| w \| \\
&= \| \alpha w \| \| w \| \\
&= \| v \| \| w \|,
\end{aligned}$$

as desired.

Now suppose $v, w \in V$ are such that $| <v,w> | = \| v \| \| w \|$. Without loss of generality, assume $v, w \in V \backslash \{0\}$, since if either were 0, one vector would be the scalar 0 times the other. For now, also assume that $\| w \| = 1$. By equation (*) above, we have

$$\begin{aligned}
0 &\leq \|v - <v,w> w\|^2 \\
&= \| v \|^2 - |<v,w>|^2 \\
&= \|v\|^2 - \|v\|^2 \|w\|^2 \quad \text{(by hypothesis)} \\
&= \|v\|^2 - \|v\|^2 \\
&= 0,
\end{aligned}$$

where the second to last inequality comes from the assumption that w is a unit vector. Thus, we know that

$$0 = \|v - <v,w> w\|^2,$$

which implies

$$v = <v,w> w. \tag{1}$$

Therefore, v is a scalar multiple of w. Now, let $p \in V \backslash \{0\}$ and suppose $| <v,p> | = \|v\| \|p\|$. Note $\|p\| \neq 0$. We have

$$\begin{aligned}
| <v,p> | &= \|v\| \|p\| \\
\frac{| <v,p> |}{\|p\|} &= \|v\| \\
| <v, \frac{p}{\|p\|} > | &= \|v\| \| \frac{p}{\|p\|} \|,
\end{aligned}$$

where the last equality comes from the fact that $\frac{p}{\|p\|}$ is unit vector. We know by Equation 1 that

$$v = <v, \frac{p}{\|p\|}> \frac{p}{\|p\|}.$$

Rewriting this equation slightly, we see that

$$v =< v, \frac{p}{\|p\|^2} > p,$$

so that v is indeed a scalar multiple of p. □

Now we are ready to prove that the norm arising from the inner product satisfies the triangle inequality.

Proof. For $v, w \in V$, we have

$$\begin{aligned} \| v + w \| &= \sqrt{< v+w, v+w >} \\ &= \sqrt{< v,v > + < v,w > + < w,v > + < w,w >} \\ &= \sqrt{\| v \|^2 + \| w \|^2 + < v,w > + < w,v >} \\ &\leq \sqrt{\| v \|^2 + 2 \| v \| \| w \| + \| w \|^2} \quad (\text{Cauchy} - \text{Schwarz}) \\ &= \| v \| + \| w \| . \end{aligned}$$

□

The reader is familiar with several normed spaces: $\mathbb{R}^n$ and $\mathbb{C}^n$, for instance. Much of the work in this paper will take place on a larger class of normaed spaces, known as Hilbert spaces:

Definition 3. *A Hilbert space is an inner product space that is complete in the norm induced by the inner product.*

Some of these terms may be unfamiliar. "Complete in the norm" means norm-Cauchy sequences converge. We will now define norm-Cauchy sequences:

Definition 4. *The sequence $(\vec{v}_n)$ of vectors is norm-Cauchy provided that $\forall\ \epsilon > 0, \exists\ K \in N$ such that if $n, m \geq K$, then $\| \vec{v}_n - \vec{v}_m \| < \epsilon$.*

For example, $\mathbb{C}^n$ is a Hilbert space. To see this, let $X = (\vec{x}_k)$ be a sequence in $\mathbb{C}^n$. Note that each term $\vec{x}_k$ of X is really a vector: $\vec{x}_k = (v_{k_1}, v_{k_2}, \cdots, v_{k_n})$, where $v_{k_j} \in \mathbb{C}, j \in \{1, 2, \cdots, n\}$. Suppose X is Cauchy. We can break X down into component sequences. For example, consider X_1, the sequence of all first components of the elements of X. It is easy to see that this sequence must be Cauchy: Let $\epsilon > 0$. Because X is Cauchy, $\exists\ K > 0$ such that if $m, n \geq K$, then $\| \vec{x}_m - \vec{x}_n \| < \epsilon$. It will also be true, then, that if v_{m_1} and v_{n_1} are terms of X_1 and $m, n \geq K$, then $\| v_{m_1} - v_{n_1} \| \leq \| \vec{x}_m - \vec{x}_n \| < \epsilon$. Thus X_1 is also Cauchy. This argument holds for all X_j, where $j \in \{1, 2, \cdots, n\}$. Now, we can think of a sequence X_j in $\mathbb{C}$ as two sequences in $\mathbb{R}$. An argument similar to the one above shows these sequences will also be Cauchy. They converge since Cauchy sequences in $\mathbb{R}$ converge. This forces X_j to converge for $j \in \{1, 2, \cdots, n\}$, which in turn forces X to converge.

A possibly less familiar Hilbert space is ℓ^2:

Definition 5. *ℓ^2 is the space of all square summable sequences of complex numbers:*

$$\{(a_n)_{n=0}^{\infty} : a_n \in \mathbb{C}\ \forall\ n \text{ and} \sum_{n=0}^{\infty} |a_n|^2 < \infty\}$$

For example, $(\frac{1}{n+1})_{n=0}^{\infty}$ is a square summable sequence, as is easily verified using, say, the integral test. Thus, the sequence is an element of ℓ^2.

Propostion 1. *If we define $< \vec{a}, \vec{b} >= \sum_{j=0}^{\infty} a_j \overline{b_j}$ for $\vec{a} = (a_0, a_1, a_2, \cdots)$ and $\vec{b} = (b_0, b_1, b_2, \cdots)$ in ℓ^2, then $(\ell^2, < \cdot, \cdot >)$ is a Hilbert space.*

Proof. First, we'll show that ℓ^2 is a vector space. Let $\vec{a}$ and $\vec{b}$ be in ℓ^2, and let $c \in \mathbb{C}$. We have $c\vec{a} = (ca_0, ca_1, \cdots)$, and

$$\sum_{n=0}^{\infty} |ca_n|^2 = |c^2| \sum_{n=0}^{\infty} |a_n|^2,$$

which converges, so $c\vec{a} \in \ell^2$. Next, observe

$$\begin{aligned}
\sum_{n=0}^{\infty} |a_n + b_n|^2 &\leq \sum_{n=0}^{\infty} (|a_n| + |b_n|)^2 \\
&= \sum_{n=0}^{\infty} (|a_n|^2 + 2|a_n||b_n| + |b_n|^2) \\
&\leq \sum_{n=0}^{\infty} (|a_n|^2 + 2|a_n||b_n| + |b_n|^2 + (|a_n| - |b_n|)^2) \\
&= \sum_{n=0}^{\infty} (2|a_n|^2 + 2|b_n|^2),
\end{aligned}$$

which converges, so $\vec{a} + \vec{b} \in \ell^2$. Since $\vec{a}$ and $\vec{b}$ are arbitrary, ℓ^2 is a vector space.

Next, we'll show ℓ^2 is an inner product space. To do this, we'll begin by showing $\sum_{j=0}^{\infty} a_j \overline{b_j} \in \mathbb{C}$ when $\vec{a}, \vec{b} \in \ell^2$. Claim $(\sum_{j=0}^{k} a_j \overline{b_j})_{k=0}^{\infty}$ is a Cauchy sequence. To prove the claim, let $\epsilon > 0$. Choose $K \in N$ such that $\sum_{j=K}^{\infty} |a_j|^2 < \epsilon$ and $\sum_{j=K}^{\infty} |b_j|^2 < \epsilon$. For $m > n \geq K$, we have

$$\begin{aligned}
\left| \sum_{j=0}^{m} a_j \overline{b_j} - \sum_{j=0}^{n} a_j \overline{b_j} \right| &= \left| \sum_{j=n+1}^{m} a_j \overline{b_j} \right| \\
&\leq \sum_{j=n+1}^{m} |a_j \overline{b_j}| \\
&\leq \sqrt{\sum_{j=n+1}^{m} |a_j|^2} \sqrt{\sum_{j=n+1}^{m} |b_j|^2} \\
&\leq \sqrt{\sum_{j=n+1}^{\infty} |a_j|^2} \sqrt{\sum_{j=n+1}^{m} |b_j|^2} \\
&< \epsilon
\end{aligned}$$

by choice of K. Hence the sequence is Cauchy and therefore converges to a complex number. It is easy to verify that the five properties of inner products hold.

Finally, we'll show ℓ^2 is complete. Suppose $(\vec{a}_n)$ is Cauchy in ℓ^2. Let

$$\begin{aligned}
\vec{a}_1 &= (a_{10}, a_{11}, \cdots), \\
\vec{a}_2 &= (a_{20}, a_{21}, \cdots), \\
\vec{a}_3 &= (a_{30}, a_{31}, \cdots), \\
&\vdots
\end{aligned}$$

and so on. Let $k \in \mathbb{Z}_{\geq 0}$. We claim that $(a_{jk})_{j=1}^{\infty}$ is Cauchy. (For example, $(a_{10}, a_{20}, \cdots)$ is Cauchy.) To prove the claim, observe that since $(\vec{a}_n)$ is Cauchy, $\forall \epsilon > 0 \; \exists \, L > 0$ such that if $m, n \geq L, \| \vec{a}_n - \vec{a}_m \| < \epsilon$.

Hence, we have

$$\sqrt{<\vec{a}_n - \vec{a}_m, \vec{a}_n - \vec{a}_m>} \quad < \quad \epsilon \quad \text{so that}$$
$$\sqrt{\sum_{i=0}^{\infty} |a_{ni} - a_{mi}|^2} \quad < \quad \epsilon.$$

This implies that $\sqrt{(a_{nk} - a_{mk})^2} < \epsilon$, which means $|a_{nk} - a_{mk}| < \epsilon$. The claim follows.

By the completeness of $\mathbb{C}$, for each k, $\exists\ a_k \in \mathbb{C}$ such that $\lim_{j\to\infty}(a_{jk}) = a_k$. Let $\vec{a} = (a_0, a_1, \cdots) = (a_k)_{k=0}^{\infty}$. Claim $\vec{a} \in \ell^2 and \lim(\vec{a}_n) = \vec{a}$. To prove the claim, let $\epsilon > 0$. Since $(\vec{a}_n)$ is Cauchy in ℓ^2, $\exists\ L$ such that if $m, n \geq L$, $\| \vec{a}_m - \vec{a}_n \| < \epsilon$. Observe that this is the same as writing $\sqrt{\sum_{j=0}^{\infty} |a_{mj} - a_{nj}|^2} < \epsilon$. Observe that for each integer q, we have

$$\begin{aligned}\sqrt{\sum_{j=0}^{q} |a_{mj} - a_j|^2} &= \lim_{n\to\infty} \sqrt{\sum_{j=0}^{q} |a_{mj} - a_{nj}|^2} \\ &\leq \sqrt{\sum_{j=0}^{\infty} |a_{mj} - a_{nj}|^2} \\ &< \epsilon.\end{aligned}$$

By the monotone convergence theorem, $\sqrt{\sum_{j=0}^{\infty} |a_{mj} - a_j|^2} \leq \epsilon$. The work above shows that $(\vec{a}_m - \vec{a})$ is square summable, which means $(\vec{a}_m - \vec{a}) \in \ell^2$. In fact, $\| \vec{a}_m - \vec{a} \| \leq \epsilon$. Since $\vec{a} = \vec{a}_m - (\vec{a}_m - \vec{a}) \in \ell^2, \vec{a} \in \ell^2$. By definition of limit of a sequence, $(\vec{a}_n)$ converges to $\vec{a}$. $\square$

We will now look at some examples of operators on ℓ^2 and other Hilbert spaces.

Definition 6. *A linear operator on a Hilbert space H is a function $T : H \to H$ such that*

1. *$T(x + y) = Tx + Ty$ for all x, y in H, and*
2. *$T(cx) = cTx$ for all c in $\mathbb{C}$ and $x \in H$.*

For example, any $n \times n$ matrix mapping $\mathbb{C}^n$ to $\mathbb{C}^n$ is a linear operator. In fact, any linear operator from $\mathbb{C}^n$ to C^n may be realized as an $n \times n$ matrix. We now introduce two basic linear operators on ℓ^2:

Definition 7. *Let $S : \ell^2 \to \ell^2$ be given by $Sa = (0, a_0, a_1, \cdots)$, where $a = (a_0, a_1, a_2, \cdots) \in \ell^2$. We call S the forward shift operator.*

Definition 8. *Let $B : \ell^2 \to \ell^2$ be given by $Ba = (a_1, a_2, \cdots)$, where a is as in the preceding definition. We call B the backward shift operator.*

We can think of S and B as infinite dimensional matrices:

$$S = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}, \quad \text{while}$$

$$B = \begin{bmatrix} 0 & 1 & 0\cdots & \\ 0 & 0 & 1\cdots & \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Propostion 2. *Spse $T : H \to H$ is linear and $< Tv, v >= 0 \; \forall \; v \in H$. Then $T = 0$.*

Proof. For all $v, w \in H$, we have

$$\begin{aligned} 0 &= < T(v+w), (v+w) > \\ &= < Tv, w > + < Tw, v > .(*) \end{aligned}$$

Similarly, $\forall \; v, w \in H$,

$$\begin{aligned} 0 &= < T(v+iw), (v+iw) > \\ &= < Tv, iw > + < Tiw, v > \\ &= -i < Tv, w > +i < Tw, v > \\ &= i(< Tw, v > - < Tv, w >), \end{aligned}$$

which implies $< Tw, v > - < Tv, w >= 0$. Adding this equation to equation (*), we have that $< Tw, v >= 0$. Since this holds for all v, it holds when $v = Tw$. This implies $Tw = 0$ for all w, so T must be the zero operator. □

Definition 9. *The operator $T : H \to H$ is bounded on H provided $\exists \; C \geq 0$ such that*

$$\| Tv \| \leq C \| v \|$$

$\forall \; v \in H$.

Definition 10. *The operator $T : H \to H$ is continuous at a point $v_0 \in H$ provided that $\forall \; \epsilon > 0, \exists \; \delta > 0$ such that if $v \in H$ is such that $\|v - v_0\| < \delta$, then $\| Tv - Tv_0 \| < \epsilon$. We say T is continuous on H if T is continuous at every point of H.*

The following is an important theorem linking boundedness to continuity:

Theorem 3. *Spse $T : H \to H$ is linear on the Hilbert space H. The following are equivalent:*

1. *T is continuous on H;*
2. *T is continuous at $\vec{0}$;*
3. *T is bounded.*

Proof. Clearly, (1) $\implies$ (2). To see that (3) $\implies$ (1), let $\epsilon > 0$, and fix $v_0 \in H$. Since T is bounded, there exists $C > 0$ such that for all $v \in V$, $\| Tv \| \leq C \| v \|$. Choose $\delta > 0$ such that $\delta < \frac{\epsilon}{C}$. Let $v \in H$ be arbitrary such that $\| v - v_0 \| < \delta$. We have

$$\begin{aligned} \| Tv - Tv_0 \| &= \| T(v_0 - v) \| \\ &\leq C \| v - v_0 \| \\ &< \epsilon. \end{aligned}$$

Since v is arbitrary, this implies T is continuous at v_0. Since v_0 is in turn arbitrary, T is continuous on H. In fact, T is uniformly continuous on H—the preceding argument shows that $\|Tv - Tw\| \leq C\|v - w\|$ for all $v, w \in H$, so that T is in fact Lipschitz.

To see that (2) $\implies$ (3), let $\epsilon > 0$. Because T is continuous at $0, \exists \; \delta > 0$ such that whenever $v \in H$ satisfies $\| v \| < \delta$, then $\| Tv \| < \epsilon$. Let $w \in H \backslash \{0\}$. Observe by our choice of δ, $\| T(\frac{\delta}{2} \frac{w}{\|w\|}) \| < \epsilon$. Hence, we may choose $C = \frac{2\epsilon}{\delta}$, and we will have $\| Tw \| \leq C \| w \|$, making T bounded. □

We are now in a position to introduce the concept of an operator norm.

Definition 11. *Spse $T : H \to H$ is linear. Then the norm of T, denoted $\| T \|$, is given by*

$$\| T \| = \sup\{\| Tv \| : v \in (H)_1\},$$

where $(H)_1 = \{v \in H : \| v \| = 1\}$.

Example 1. *Find the norm of $T : \mathbb{C}^2 \to \mathbb{C}^2$, where $T = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$.*

Let $v \in (\mathbb{C}^2)_1$. Thus, $v = (v_1, v_2)$ such that $\sqrt{|v_1|^2 + |v_2|^2} = 1$. We have

$$\begin{aligned} \left\| \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\| &= \sqrt{|v_1 + 2v_2|^2 + |2v_1 + v_2|^2} \\ &\leq \sqrt{(|v_1| + 2|v_2|)^2 + (2|v_1| + |v_2|)^2} \quad \text{(triangle inequality)} \\ &= \sqrt{5|v_1|^2 + 8|v_1||v_2| + 5|v_2|^2} \\ &= \sqrt{5 + 8|v_1||v_2|}. \end{aligned}$$

To obtain an upper bound on the norm, we will find $\sup\{|v_1||v_2| : |v_1|^2 + |v_2|^2 = 1\}$. Substituting $|v_1| = \sqrt{1 - |v_2|^2}$, we have

$$|v_1||v_2| = \sqrt{1 - |v_2|^2}|v_2|,$$

and a little calculus shows $f(r) = r\sqrt{1 - r^2}$ has maximum value $\frac{1}{2}$ on $[0, 1]$. Thus, we see that 3 is an upper bound for $\|T\|$. To see that $\|T\| = 3$, note $T(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}) = 3$.

Example 2. *For S the forward shift on ℓ^2, $\|S\| = 1$.*

For any unit vector $\vec{v} \in \ell^2$, $\|Sv\| = 1$. To see this, let $\vec{v} = (v_0, v_1, v_2, \cdots) \in (\ell^2)_1$. Then

$$\begin{aligned} \|Sv\| &= \|(0, v_0, v_1, v_2, \cdots)\| \\ &= \sqrt{\sum_{j=0}^{\infty} | v_j |^2} \\ &= \|v\| \\ &= 1. \end{aligned}$$

Since $v \in (\ell^2)_1$ was arbitrary, we have that $\|S\| = 1$.

Example 3. *For B the backward shift operator on ℓ^2, $\|B\| = 1$.*

First, note that for any unit vector $\vec{v}$ in ℓ^2, $\|Bv\| \leq 1$. To see this, let $\vec{v} = (v_0, v_1, v_2, \cdots) \in \ell^2$. We have

$$\begin{aligned} \|Bv\| &= \|(v_1, v_2, v_3, \cdots)\| \\ &= \sqrt{\sum_{j=1}^{\infty} | v_j |^2} \\ &\leq \sqrt{\sum_{j=0}^{\infty} | v_j |^2} \\ &= \|v\| \\ &= 1. \end{aligned}$$

Since v is arbitrary, we have that $\|Bv\| \leq 1$, implying $\|B\| \leq 1$. Now, observe that $\vec{w} = (0, 1, 0, 0, \cdots)$ is a unit vector in ℓ^2 and that $\|Bw\| = 1$. Thus, 1 is the least upper bound for $\{\|Bv\| : v \in (\ell^2)_1\}$. Thus, $\|B\| = 1$.

Lemma 1. *Let $T : H \to H$, where H is an n-dimensional Hilbert space. Then $\| T \|$ is finite.*

Proof. Since H is n-dimensional, we can think of T as an $n \times n$ matrix:

$$T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}.$$

We will let $\vec{r_j} = (a_{j1}, a_{j2}, \cdots, a_{jn})$ for $j \in \{1, 2, \cdots, n\}$. Now, let $\vec{v} \in (H)_1$. We can write $\vec{v} = (v_1, v_2, \cdots, v_n)$. We have

$$\begin{aligned} \| Tv \| &= \left\| \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \right\| \\ &= \left\| \begin{bmatrix} < \vec{r_1}, v > \\ < \vec{r_2}, v > \\ \vdots \\ < \vec{r_n}, v > \end{bmatrix} \right\| \\ &= \sqrt{|< \vec{r_1}, v >|^2 + |< \vec{r_2}, v >|^2 + \cdots + |< \vec{r_n}, v >|^2} \\ &\leq \sqrt{\| \vec{r_1} \|^2 \| v \|^2 + \| \vec{r_2} \|^2 \| v \|^2 + \cdots + \| \vec{r_n} \|^2 \| v \|^2} \\ &= \sqrt{\| \vec{r_1} \|^2 + \| \vec{r_2} \|^2 + \cdots + \| \vec{r_n} \|^2}. \end{aligned}$$

Since $\vec{v} \in (H)_1$ was arbitrary, we have found an upper bound for $\|Tv\|$. Thus, $\| T \|$ is finite. □

Lemma 2. *Let $T : H \to H$ be linear and have finite norm. Then for all $v \in H$, $\| Tv \| \leq \| T \| \| v \|$.*

Proof. Since the inequality of the lemma clearly holds if $v = 0$, let $v \in H \setminus \{0\}$. Then

$$\begin{aligned} \| Tv \| &= \| Tv \| \frac{\| v \|}{\| v \|} \\ &= \left\| T\left(\frac{v}{\| v \|}\right) \right\| \| v \| \\ &\leq \| T \| \| v \|, \end{aligned}$$

where the last equality comes from the definition of operator norm. □

Observe this lemma shows us finiteness of norm is equivalent to boundedness, which is equivalent to continuity by Theorem 3. Thus, we have the following theorem:

Theorem 4. *Let H be a finite dimensional Hilbert Space, and suppose $T : H \to H$ is linear. Then T is continuous on H.*

However, one should note that linear mappings on infinite dimensional spaces need not be continuous. For instance, let $h : C([0, 1], < \cdot, \cdot >) \to \mathbb{C}$ be given by $h(f) = f(\frac{1}{2})$. Here, for f,g in C([0,1]), we will, as before, define the inner product:

(2) $$< f, g >= \int_0^1 f(x)\overline{g(x)}dx.$$

Thus,

$$\| f \| = \sqrt{\int_0^1 |f(x)|^2 dx}.$$

Note C([0,1]) is infinite dimensional (e.g. the set $\{x \mapsto 1, x \mapsto x, x \mapsto x^2, \cdots\}$ is a linearly independent set of functions in the space) and that h is clearly linear. However, h is neither bounded nor continuous. To see this, we need only look at unit vectors in C([0,1]). These are vectors f such that the area under the curve f^2 is equal to 1. We can make h(f) as large as we please by choosing the unit vector f correctly. Let's look at a simple example. Consider a graph of f that is 0 on all of $[0,1]$ except near $\frac{1}{2}$. On either side of $\frac{1}{2}$, the graph slopes linearly upward, forming a triangle with upper vertex at $(\frac{1}{2}, y)$. We can make y whatever we please by making the triangle wider or narrower. Thus, h is not bounded. By Theroem 3, this implies h is not continuous.

The space of continuous functions on $[0,1]$ with the inner product (2) is not a Hilbert space because it is not complete. It is possible to construct non-continuous linear operators on infinite dimensional Hilbert spaces; see, e.g., [6, p.10].

We now define numerical range:

Definition 12. *Spse $T : V \to V$, where V is an inner product space. Then the numerical range of T, denoted $W(T)$, is given by*

$$W(T) = \{< Tv, v >: v \in (V)_1\}.$$

Example 4. *Spse $I : \mathbb{C}^n \to \mathbb{C}^n$ is the identity mapping. Let $v \in (\mathbb{C}^n)_1$. Then*

$$\begin{aligned} < Iv, v > &= < v, v > \\ &= < v, v > \\ &= 1. \end{aligned}$$

Since v was an arbitrary unit vector, we have that $W(I) = \{1\}$.

Thus, we see that the numerical range of the identity mapping is $\{1\}$.

Example 5. *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be given by*

$$T = \begin{bmatrix} 5 & 4 \\ -4 & 5 \end{bmatrix}.$$

Note that we have moved into the real setting for this example.

Let $v \in (\mathbb{R}^2)_1$, where $v = (v_1, v_2)$. Thus, $1 = \| v \| = \sqrt{| v_1 |^2 + | v_2 |^2}$. We have

$$\begin{aligned} < Tv, v > &= < \begin{bmatrix} 5v_1 + 4v_2 \\ -4v_1 + 5v_2 \end{bmatrix}, \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} > \\ &= 5 | v_1 |^2 + 4v_1v_2 - 4v_1v_2 + 5 | v_2 |^2 \\ &= 5(| v_1 |^2 + | v_2 |^2) = 5 \end{aligned}$$

Since $v \in (\mathbb{R}^2)_1$ is arbitrary, we have that $W(T) = \{5\}$.

Note that if the matrix of Example 5 mapped $\mathbb{C}^2 \to \mathbb{C}^2$, it would have a different numerical range because we would have to use a different inner product. In fact, by Theorem 7 below, it would be an ellipse, with foci at the eigenvalues of T.

For the remainder of the paper, we will consider only complex Hilbert spaces.

Example 6. *Let $a \in \mathbb{C}$, and let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be given by*

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

Then $W(T)$ is a closed disk of radius $\frac{|a|}{2}$.

Choose $\alpha \in \mathbb{R}$ such that $a = \mid a \mid e^{i\alpha}$. Let (z_1, z_2) be an arbitrary unit vector in C^2 and set $r = \mid z_1 \mid$. It follows that there are real numbers θ_1 and θ_2 such that $z_1 = re^{i\theta_1}$ and $z_2 = \sqrt{1-r^2}e^{i\theta_2}$, since $\mid z_1 \mid^2 + \mid z_2 \mid^2 = 1$. Thus, we have

$$\begin{aligned} < T(z_1, z_2), (z_1, z_2) > &= < (az_2, 0), (z_1, z_2) > \\ &= az_2\overline{z}_1 \\ &= \mid a \mid r\sqrt{1-r^2}e^{i(\alpha+\theta_2-\theta_1)}. \end{aligned}$$

Because (z_1, z_2) is arbitrary, we see that

$$W(T) = \{\mid a \mid r\sqrt{1-r^2}e^{i\theta} : 0 \leq r \leq 1, 0 \leq \theta \leq 2\pi\}. \tag{3}$$

Let's find the maximum value of the absolute values of the elements of $W(T)$. To do this, since $\mid e^{i\theta} \mid = 1$, take the derivative of $r \mapsto r\sqrt{1-r^2}$ and solve to find that it obtains an extreme value on $[0,1]$ at $r = \frac{\sqrt{2}}{2}$. This tells us that the maximum value of the absolute value of an element of $W(T)$ is $\frac{|a|}{2}$. Note $\frac{|a|}{2} \in W(T)$ by choosing $r = \frac{\sqrt{2}}{2}$ and $\theta = 0$ in equation (3). Also, note $0 \in W(T)$, occurring when $r = 0$ or $r = 1$. Because $r\sqrt{1-r^2}$ is continuous on $[0,1]$, which is a connected set, the image of $[0,1]$ is connected under $r \mapsto (r\sqrt{1-r^2})$. Therefore, the line segment connecting 0 and $\frac{|a|}{2}$ is in $W(T)$. Letting θ vary between 0 and 2π, we see that $W(T)$ is the closed unit disk of radius $\frac{a}{2}$.

We will see more examples later. However, we will first pause to establish some general properties of the numerical range.

Theorem 5. *Suppose $T : H \to H$, and suppose T is bounded and linear. Then $W(T)$ is bounded; in fact, $W(T) \subseteq \{z : \mid z \mid \leq \parallel T \parallel\}$.*

Proof. Let $v \in (H)_1$. We have

$$\begin{aligned} \mid < Tv, v > \mid &\leq \parallel Tv \parallel \parallel v \parallel \text{ (Cauchy − Schwarz)} \\ &\leq \|T\| \parallel v \parallel^2 \quad \text{(Lemma 2)} \\ &= \|T\|. \end{aligned}$$

Since $v \in (H)_1$ is arbitrary, we have that $\mid < Tv, v > \mid \leq \|T\|$ for all $v \in (H)_1$. Thus, $W(T)$ is bounded; in fact, it is contained in the closed disk of radius $\|T\|$ centered at the origin. □

Theorem 6. *Suppose $T : \mathbb{C}^n \to \mathbb{C}^n$ is linear. Then $W(T)$ is closed.*

Proof. Since T acts on finite dimensions, T is continuous by Theorem 4. Now, let w be a cluster point of $W(T)$. By definition, there exists a sequence $(w_n) \in W(T)$ converging to w. Observe $\forall n \in \mathbb{N}, w_n = < Tv_n, v_n >$ where $v_n \in (\mathbb{C}^n)_1$. Because (v_n) is bounded, it has a subsequence (v_{r_n}) of unit vectors that converges to some unit vector, call it v. Because T is continuous, $T(v_{r_n})$ converges to Tv. Claim $< Tv_{r_n}, v_{r_n} >$ converges to $< Tv, v >$. To prove the claim, let $\epsilon > 0$. Then $\exists\ k_1 > 0$ such that if $n > k_1$, $\parallel v_{r_n} - v \parallel < \frac{\epsilon}{2(\|Tv\|+1)}$. Similarly, $\exists\ k_2 > 0$ such that if $n > k_2$, then $\parallel Tv_{r_n} - Tv \parallel < \frac{\epsilon}{2}$. Let $K = \sup\{k_1, k_2\}$. Then if $n > K$, we have

$$\begin{aligned} \|< Tv_{r_n}, v_{r_n} > - < Tv, v >\| &= \|< Tv_{r_n}, v_{r_n} > - < Tv, v_{r_n} > + < Tv, v_{r_n} > - < Tv, v >\| \\ &= \|< Tv_{r_n} - Tv, v_{r_n} > + < Tv, v_{r_n} - v >\| \\ &\leq \|< Tv_{r_n} - Tv, v_{r_n} >\| + \|< Tv, v_{r_n} - v >\| \text{ (Triangle Inequality)} \\ &\leq \parallel Tv_{r_n} - Tv \parallel \parallel v_{r_n} \parallel + \parallel Tv \parallel \parallel v_{r_n} - v \parallel \text{ (Cauchy − Schwarz)} \\ &\leq \frac{\epsilon}{2} + \parallel Tv \parallel \left(\frac{\epsilon}{2(\parallel Tv \parallel + 1)}\right) \\ &< \epsilon. \end{aligned}$$

The claim follows. Because (w_n) has a subsequence converging to $< Tv, v >$, (w_n) must also converge to $< Tv, v >$. Since v is a unit vector, $< Tv, v > \in W(T)$. Since sequences converge uniquely, $w =< Tv, v >$. Because w is arbitrary, $W(T)$ contains all its cluster points. Therefore, $W(T)$ is closed. $\square$

Lemma 3. *Let $T : H \to H$ be linear, and suppose λ is an eigenvalue of T. Then $\lambda \in W(T)$.*

Proof. Let λ be an eigenvalue of T, and let $\frac{v}{\|v\|}$ be a corresponding normalized eigenvector. Observe $\frac{v}{\|v\|} \in (H)_1$. We have

$$\begin{aligned} < T\frac{v}{\| v \|}, \frac{v}{\| v \|} > &= < \lambda\frac{v}{\| v \|}, \frac{v}{\| v \|} > \\ &= \lambda < \frac{v}{\| v \|}, \frac{v}{\| v \|} > \\ &= \lambda. \end{aligned}$$

Hence $\lambda \in W(T)$. Since λ is arbitrary, this is true for all eigenvalues of T. $\square$

Recall the following definition:

Definition 13. *Let $T : H \to H$ be a bounded linear operator on the Hilbert space H. Then the adjoint operator for T, denoted T^*, is the unique bounded operator on H such that $< Tv, w >=< v, T^*w >$ for all $v, w \in H$.*

For a proof of the existence of T^*, see [3]. One can easily check that if one has a matrix representation for T, T^* is the conjugate transpose of T.

Note that B is the adjoint operator of S. One can verify this property using matrix representations for S and B. We choose instead to use a direct proof: Let $\vec{v}, \vec{w} \in \ell^2$, where $\vec{v} = (v_0, v_1, v_2, \cdots)$ and $\vec{w} = (w_0, w_1, w_2, \cdots)$. Without loss of generality, assume neither $\vec{v}$ nor $\vec{w}$ is the zero vector. By definition of adjoint operator, we have

$$\begin{aligned} < \vec{v}, S^*\vec{w} > &= < S\vec{v}, \vec{w} > \\ &= < (0, v_0, v_1, \cdots), (w_0, w_1, w_2, \cdots) > \\ &= v_0\overline{w}_1 + v_1\overline{w}_2 + v_2\overline{w}_3 + \cdots \end{aligned}$$

Now, note that we also have

$$\begin{aligned} < \vec{v}, B\vec{w} > &= < (v_0, v_1, v_2, \cdots), (w_1, w_2, w_3, \cdots) > \\ &= v_0\overline{w}_1 + v_1\overline{w}_2 + v_2\overline{w}_3 + \cdots . \end{aligned}$$

Since $\vec{v}, \vec{w}$ are arbitrary, we have that for all $\vec{v}, \vec{w} \in \ell^2$,

$$\begin{aligned} < \vec{v}, S^*\vec{w} > &= < \vec{v}, B\vec{w} > \\ < \vec{v}, (S^* - B)\vec{w} > &= 0. \end{aligned}$$

Since $\vec{w}$ and $\vec{v}$ are arbitrary, we have that $S^* = B$. For the remainder of the paper, we will refer to B as S^*.

Lemma 4. *Let $T : H \to H$ be a bounded linear operator on the Hilbert space H, and let T^* be the adjoint operator of T. Then $W(T^*)$ is the reflection of $W(T)$ across the real axis.*

Proof. Note that $W(T) = \{< Tv, v >: v \in (H)_1\}$ and $W(T^*) = \{< T^*v, v >: v \in (H)_1\} = \{< v, Tv >: v \in (H)_1\}$. Therefore, $W(T^*)$ will be made up of the conjugates of the elements of $W(T)$. Since the conjugate of any complex number is the reflection of the number across the real axis, we see that $W(T^*)$ is the reflection of $W(T)$ across the real axis. $\square$

Lemma 5. *For all $\alpha \in (\mathbb{C})_1$, α is an eigenvalue of S^*.*

Proof. Observe that for $| \alpha |< 1, \alpha \in \mathbb{C}$, we have $(\alpha^j)_{j=0} \in \ell^2$. Hence, we have

$$S^*(1, \alpha, \alpha^2, \cdots) = (\alpha(\alpha^j)_{j=0}) \in \ell^2.$$

Therefore, α is an eigenvalue for S^*. □

We now present two lemmas that will help us determine the numerical ranges of S and S^*.

Lemma 6. *S has no eigenvalues.*

Proof. Spse, in order to obtain a contradiction, that λ is an eigenvalue of S. Then $\exists$ a vector $v = (v_0, v_1, \cdots)$ such that $Sv = \lambda v$. By definition of S, this implies that

$$\begin{aligned} 0 &= \lambda v_0 \\ v_0 &= \lambda v_1 \\ v_1 &= \lambda v_2 \\ &\vdots \end{aligned}$$

For this to be true, v must be the zero vector, which cannot be an eigenvector. Therefore, λ is not an eigenvalue for S. Since λ is arbitrary, S has no eigenvalues. □

Propostion 3. *$W(S) = W(S^*)$ and is the open unit disk in $\mathbb{C}$.*

Proof. By Lemma 5, we have that $\forall \alpha \in \mathbb{C}$ such that $| \alpha |< 1$, α is an eigenvalue of S^*. Thus, by Lemma 3 above, $\alpha \in W(S^*)$. Since α is arbitrary, we see by Lemma 3 that the open unit disk of $\mathbb{C}$ is entirely contained in $W(S^*)$. By Lemma 4, this means it is also contained in $W(S)$. Now suppose, in order to obtain a contradiction, that $\exists v \in (\ell^2)_1$ such that $|< Sv, v >|= 1$. Then we have

$$\begin{aligned} 1 &= |< Sv, v >| \\ &\leq \| Sv \| \| v \| \\ &= \| Sv \| \\ &\leq 1. \end{aligned}$$

This implies $|< Sv, v >|=\| Sv \| \| v \|$, which can happen only when $Sv = \lambda v$ for some constant λ. This makes λ an eigenvalue of S, contradicting Lemma 6. Therefore, there exists no unit vector in ℓ^2 such that $|< Sv, v >|= 1$. The same holds for S^* by Lemma 4. By Theorem 5 and example 2, $W(S)$ has no point of modulus greater than 1, which implies that $W(S^*)$ does not have one. Therefore, $W(S) = W(S^*) =$ the open unit disk in $\mathbb{C}$. □

We will now consider the numerical ranges of normal operators. First, a few definitions.

Definition 14. *Let $T : H \to H$ be a bounded linear operator on the Hilbert space H. Then T is normal provided $T^*T = TT^*$.*

Theorem 7. *Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be linear and suppose T is normal. Then $\mathbb{C}^h$ has an orthnonormal basis of eigenvectors of T.*

For a proof, see [1].

Example 7. *Let $T : \mathbb{C}^2 \to \mathbb{C}^2$ be normal. Then $W(T)$ is the line segment connecting the eigenvalues of T.*

Because T is normal, $\mathbb{C}^2$ has an orthonormal basis of eigenvectors of T. Let this basis be given by $\{\vec{b_1}, \vec{b_2}\}$, where $\vec{b_1}$ has associated eigenvalue λ_1 and $\vec{b_2}$ has associated eigenvalue λ_2. Let $v \in (\mathbb{C}^2)_1$. We can write $v = \alpha_1\vec{b_1} + \alpha_2\vec{b_2}$, where $\alpha_1, \alpha_2 \in \mathbb{C}$, and $| \alpha_1 |^2 + | \alpha_2 |^2= 1$. We have

$$\begin{aligned} < Tv, v > &= < T(\alpha_1\vec{b_1} + \alpha_2\vec{b_2}), \alpha_1\vec{b_1} + \alpha_2\vec{b_2} > \\ &= < \lambda_1\alpha_1\vec{b_1} + \lambda_2\alpha_2\vec{b_2}, \alpha_1\vec{b_1} + \alpha_2\vec{b_2} > \\ &= \lambda_1 | \alpha_1 |^2 + \lambda_2 | \alpha_2 |^2 . \end{aligned}$$

Since v was arbitrary, we see that

$$W(T) = \{\lambda_1 \mid \alpha_1 \mid^2 + \lambda_2 \mid \alpha_2 \mid^2 : \alpha_1, \alpha_2 \in \mathbb{C},\ \mid \alpha_1 \mid^2 + \mid \alpha_2 \mid^2 = 1\}.$$

Alternatively, we can let $\mid \alpha_1 \mid^2 = t$ so that $\mid \alpha_2 \mid^2 = 1 - t$. Thus, we have

$$\begin{aligned} W(T) &= \{\lambda_1 t + \lambda_2(1-t) : 0 \le t \le 1\} \\ &= \{\lambda_2 + (\lambda_1 - \lambda_2)t : 0 \le t \le 1\}. \end{aligned}$$

Thus, $W(T)$ is the line segment joining λ_1 and λ_2.

To tackle the more general case of a normal operator $T : \mathbb{C}^n \to \mathbb{C}^n$, we need background on convexity.

Definition 15. *A set $A \subseteq H$ is called convex provided that whenever x and y are points in A, the line connecting x and y is in A as well.*

Definition 16. *The convex hull of a set $A \subseteq H$ is the set of all convex combinations of elements of A; that is, it is the set of finite sums $t_1x_1 + t_2x_2 + \cdots + t_nx_n$ where $x_i \in A$, $t_j \ge 0$, $\sum t_j = 1$, and n is an arbitrary positive integer.*

Geometrically speaking, if $A \subseteq H$, the convex hull of A may be viewed as the smallest convex subset of H that contains A. The next proposition will make this intuitive definition more precise.

Propostion 4. *Let $A \subseteq H$. Then*

$$\text{convex hull}(A) = \cap\{E : E \subseteq H \text{ is convex and } A \subseteq E\}.$$

Proof. First, we'll show convex hull(A) is itself a convex subset of H. Let $x, y \in$ convex hull(A). Then for $x_i, y_i \in A, x = t_1x_1 + t_2x_2 + \cdots + t_nx_n$ and $y = s_1y_1 + s_2y_2 + \cdots + s_my_m$, where $t_i > 0, s_i > 0, \sum t_i = 1$, and $\sum s_i = 1$. We wish to show that the line connecting x and y is itself in convex hull(A). Note than any point on this line is of the form $rx + (1-r)y$ for $r \in [0,1]$. Now, let $r \in [0,1]$ be arbitrary. We have

$$rx + (1-r)y = r(t_1x_1 + t_2x_2 + \cdots + t_nx_n) + (1-r)(s_1y_1 + s_2y_2 + \cdots + s_my_m)$$

If the sum of the coefficients of the x_j's and y_j's is 1, this point will be in convex hull(A). We have

$$\begin{aligned} rt_1 + \cdots rt_n + (1-r)s_1 + (1-r)s_2 + \cdots + (1-r)s_m &= r(t_1 + \cdots + t_n) - (1-r)(s_1 + \cdots + s_m) \\ &= r + (1-r) \\ &= 1. \end{aligned}$$

Therefore, $[rx + (1-r)y] \in$ convex hull(A). Since $r \in [0,1]$ is arbitrary , we see that the line connecting x and y is in convex hull(A). Therefore, since x and y were arbitrary, convex hull(A) is convex. Thus, we have

$$\cap\{E : E \subseteq H \text{ is convex and } A \subseteq E\} \subseteq \text{convex hull}(A).$$

Now, let E be an arbitrary convex subset of H such that $A \subseteq E$. We wish to show that convex hull$(A) \subseteq E$. We proceed by induction. First, consider $\{a_1, a_2\} \subseteq A$. Note that any convex combination of a_1 and a_2 will be of the form $ta_1 + (1-t)a_2$, where $0 \le t \le 1$. This is simply the line segment linking a_1 and a_2. By convexity of E, then, all convex combinations of a_1 and a_2 are in E. Because a_1 and a_2 are arbitrary, this holds for any two element subset of A.

Now, suppose it is true that the set of convex combinations of any n-element subset of elements of A is a subset of E. We will show that the same is true for an arbitrary $(n+1)$-element subset of A. Let $\{a_1, a_2, \cdots, a_{n+1}\}$ be such a subset. We want to show that

$$t_1a_1 + t_2a_2 + \cdots + t_{n+1}a_{n+1}$$

is in E for $\sum_{j=0}^{n+1} t_j = 1$,$t_j \geq 0$. If $t_1 = 1$, we have nothing to prove. Otherwise, we have that $t_2 + t_3 + \cdots + t_{n+1} > 0$ and observe we can rewrite the previous quantity as:

(4)

$$t_1 a_1 + (t_2 + t_3 + \cdots + t_{n+1})(\frac{t_2}{t_2 + t_3 + \cdots + t_{n+1}} a_2 + \frac{t_3}{t_2 + t_3 + \cdots + t_{n+1}} a_3 + \frac{t_{n+1}}{t_2 + t_3 + \cdots + t_{n+1}} a_{n+1}).$$

Note that now the sum of the coefficients inside the parantheses is 1, so that the sum inside the parantheses is a convex combination of n elements of A. By hypothesis, it is then in E. Since

$$t_1 + (t_2 + t_3 + \cdots + t_{n+1}) = 1,$$

equation 4 is just a convex combination of two elements of E. By the convexity of E, this is in E. Therefore,

$$t_1 a_1 + t_2 a_2 + \cdots + t_{n+1} a_{n+1}$$

is in E. We therefore see by induction that convex hull$(A) \subseteq E$. Therefore,

$$\text{convex hull}(A) \subseteq \cap\{E : E \subseteq H \text{ is convex and } A \subseteq E\}.$$

Finally, we have

$$\text{convex hull}(A) = \cap\{E : E \subseteq H \text{ is convex and } A \subseteq E\}.$$

□

Our only interest in convex hulls will be in those of finite subcollections of points in $\mathbb{C}$. In general, to construct the convex hull of a finite set, join each point to every other point with a single line segment. Then the convex hull will be all these line segments plus the bounded regions they enclose. For example, the convex hull of two points will be a line, while the convex hull of three points will be a triangle and the region it bounds.

We will now consider some of the connections between convex hulls and numerical range.

Example 8. *Let $T : \mathbb{C}^n \to \mathbb{C}^n$ be normal. Then $W(T)$ is the convex hull of the eigenvalues of T.*

Since T is normal, $\mathbb{C}^n$ has an orthonormal basis of eigenvectors of T, call it $\{\vec{b_1}, \cdots, \vec{b_n}\}$. Now, $v \in (\mathbb{C}^n)_1$ if and only if $v = \alpha_1 \vec{b_1} + \cdots + \alpha_n \vec{b_n}$, where $| \alpha_1 |^2 + \cdots + | \alpha_n |^2 = 1$. We have

$$\begin{aligned} < Tv, v > &= < T(\alpha_1 \vec{b_1} + \cdots + \alpha_n \vec{b_n}), \alpha_1 \vec{b_1} + \cdots + \alpha_n \vec{b_n} > \\ &= < \lambda_1 \alpha_1 \vec{b_1} + \cdots + \lambda_n \alpha_n \vec{b_n}, \alpha_1 \vec{b_1} + \cdots + \alpha_n \vec{b_n} > \\ &= \lambda_1 | \alpha_1 |^2 + \cdots + \lambda_2 | \alpha_2 |^2 . \end{aligned}$$

Hence, we have

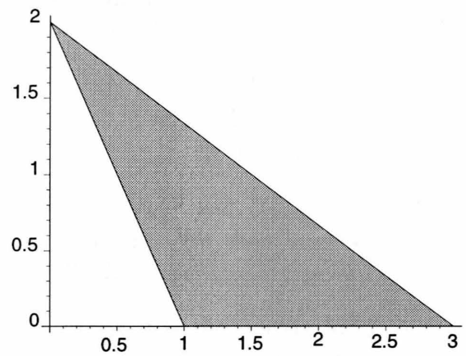
$$W(T) = \{\lambda_1 | \alpha_1 |^2 \cdots + \lambda_n | \alpha_n |^2 : | \alpha_1 |^2 + \cdots + | \alpha_n |^2 = 1\}$$

Therefore, by definition, $W(T)$ is the convex hull of the eigenvalues of T.

Example 9. *Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by $T = \begin{bmatrix} 2i & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$. Then by Example 8 above, we have*

$$W(T) = \{2it_1 + t_2 + 3t_3 : t_1 + t_2 + t_3 = 1, t_j \geq 0\}.$$

This is the triangle whose vertices are at $1, 3$ and $2i$, the eigenvalues of T:



We have seen that the eigenvalues of an operator T belong to $W(T)$. Recall also that eigenvalues are "similarity invariant."

Definition 17. *The $n \times n$ matrices A and B are similar provided there exists an $n \times n$ invertible matrix S such that $S^{-1}AS = B$.*

Similarity preserves certain nice properties of A and B. For example, as mentioned above, if A and B are similar, they have the same eigenvalues. Although this result might lead one to think that similar matrices have the same numerical range, this is not the case.

Example 10. *Let $A : \mathbb{C}^2 \to \mathbb{C}^2$ be given by $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, let $B : \mathbb{C}^2 \to \mathbb{C}^2$ be given by $B = \begin{bmatrix} -1 & -2 \\ 0 & 1 \end{bmatrix}$, and let $S : \mathbb{C}^2 \to \mathbb{C}^2$ be given by $S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then $S^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$, and $S^{-1}AS = B$. Thus, A and B are similar. However, $W(A) \neq W(B)$.*

To see this, we will first find $W(A)$. First, note that since $A^*A = AA^*$, A is normal with eigenvalues 1 and -1. Therefore, by Example 7 above, we have that $W(A)$ is merely the line segment connecting 1 and -1. However, this is not $W(B)$. For example, let $\vec{v} \in (\mathbb{C}^2)_1$ be given by $\vec{v} = (\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}})$. Then we have $< B\vec{v}, \vec{v} >= -i$. Since this is in $W(B)$ but is clearly not in $W(A)$, $W(A) \neq W(B)$. In fact, as will be seen in the next lemma, $W(B)$ is an ellipse with foci at its eigenvalues: 1 and -1. Its minor axis extends from -1 to 1.

We have now seen that the equivalence relation of similarity does not preserve numerical range. However, the stronger relation of unitary equivalence does:

Definition 18. *The $n \times n$ matrix U is unitary provided $U^*U = 1$.*

Example 11. *Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by $T = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{i}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & -\frac{i}{\sqrt{2}} \end{bmatrix}$. Then T is unitary.*

Since we have a matrix representation for T, T^* is the conjugate transpose of T. It is easy to check that $T^*T = I = TT^*$.

Definition 19. *The $n \times n$ matrices A and B are unitarily equivalent provided that there is an $n \times n$ unitary matrix U such that $U^*AU = B$*

Propostion 5. *Suppose A and B are unitarily equivalent $n \times n$ matrices. Then*

$$W(A) = W(B).$$

Proof. Since A and B are unitarily equivalent, there exists a unitary matrix U such that $B = U^*AU$. Let $w \in W(B)$. Then there exists $v \in (\mathbb{C}^n)_1$ such that $< Bv, v >= w$. We have

$$\begin{aligned} w &= < Bv, v > \\ &= < U^*AUv, v > \\ &= < AUv, Uv > . \end{aligned}$$

Since unitary matrices act as isometries ($\|Uv\|^2 =< Uv, Uv >=< U^*Uv, v >=< v, v >= \|v\|^2$), Uv is a unit vector since v is. Therefore, $w \in W(A)$. We then have $W(B) \subseteq W(A)$. A similar argument shows $W(A) \subseteq W(B)$. Therefore, $W(A) \ = \ W(B)$. □

Suppose $T : \mathbb{C}^n \longrightarrow \mathbb{C}^n$ is linear, and that A is the matrix for T relative to one orthonormal basis for $\mathbb{C}^n$ while T is the matrix for T relative to another orthonormal basis for $\mathbb{C}^n$. Recall from linear algebra that these matrices A and B are unitarily equivalent. Thus, by Proposition 5, if we seek $W(T)$, we may work with a matrix for T relative to any orthonormal basis of $\mathbb{C}^n$ we wish.

The proof of the following lemma is an adaptation of an argument appearing in [2].

Lemma 7. *Let $T : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ be linear. Then $W(T)$ is a "filled-in" ellipse whose foci are the eigenvalues of T.*

Proof. Let u_1 be a unit eigenvector for T corresponding to the eigenvalue λ_1. Take any unit vector u_2 orthogonal to u_1 to yield an orthonormal basis $\{u_1, u_2\}$ for $\mathbb{C}^2$. Then the matrix for T relative to this basis is

$$M_1 = \begin{bmatrix} \lambda_1 & \alpha_1 \\ 0 & \alpha_2 \end{bmatrix}.$$

Now set $c = (\lambda_1 + \alpha_2)/2$ and note $M_2 = M_1 - cI$ has the form

$$M_2 = \begin{bmatrix} -b & \alpha_1 \\ 0 & b \end{bmatrix},$$

where $b = \frac{\alpha_2}{2} - \frac{\lambda_1}{2}$. If we show $(W(M_2))$ is an ellipse E, then $W(M_1)$ will be the ellipse $E + c$, that is, the ellipse that results from translating E by c. Now, let $b = de^{i\theta}$ and let

$$M_3 = e^{-i\theta}M_2 = \begin{bmatrix} -d & \beta \\ 0 & d \end{bmatrix},$$

where $d = |b| \geq 0$ and $\beta = e^{-i\theta}\alpha_1$. If $W(M_3)$ is an ellipse E, then $W(M_2)$ is the ellipse $\{e^{i\theta}z : z \in E\}$ (i.e., the ellipse that results when E is rotated through the angle θ). Now, let $\beta = se^{i\phi}$, where $s \geq 0$ is $|\beta|$. Set

$$M_4 = \begin{bmatrix} -d & s \\ 0 & d \end{bmatrix}$$

and note the M_4 is unitarily equivalent to M_3:

$$M_4 = \begin{bmatrix} e^{-i\phi} & 0 \\ 0 & 1 \end{bmatrix} M_3 \begin{bmatrix} e^{i\phi} & 0 \\ 0 & 1 \end{bmatrix}.$$

By the proposition above, $W(M_4) = W(M_3)$. Thus if $W(M_4)$ is an ellipse, so are $W(M_3)$, $W(M_2)$, $W(M_1)$, and $W(T)$.

Now, letting $b = \frac{1}{2}s$ and choosing a such that $d = \sqrt{a^2 - b^2}$, we have

$$M_4 = \begin{bmatrix} -\sqrt{a^2 - b^2} & 2b \\ 0 & \sqrt{a^2 - b^2} \end{bmatrix}.$$

Let $M_4 = M$. Claim that $W(M)$ is the ellipse with boundary equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. Now, letting $v = (\lambda_1, \lambda_2)$ be on the unit sphere of $\mathbb{C}^2$, we have

$$\begin{aligned} < Mv, v > &= -\sqrt{a^2-b^2}|\lambda_1|^2 + 2b\lambda_2\overline{\lambda_1} + \sqrt{a^2-b^2}|\lambda_2|^2 \\ &= \sqrt{a^2-b^2}(-|\lambda_1|^2 + |\lambda_2|^2) + 2b\lambda_2\overline{\lambda_1} \\ &= \sqrt{a^2-b^2}(-|\lambda_1|^2 - |\lambda_2|^2 + 2|\lambda_2|^2) + 2b\lambda_2\overline{\lambda_1} \\ &= -\sqrt{a^2-b^2} + 2b\lambda_2\overline{\lambda_1} + 2|\lambda_2|^2\sqrt{a^2-b^2}. \end{aligned}$$

Letting $|\lambda_2| = r$, we have

$$< Mv, v >= -\sqrt{a^2-b^2} + 2br\sqrt{1-r^2}e^{it} + 2r^2\sqrt{a^2-b^2},$$

where $t = \arg(\lambda_2) - \arg(\lambda_1)$. We will show that as r ranges from 0 to 1 and t ranges from 0 to 2π,

$$f(r,t) := -\sqrt{a^2-b^2} + 2br\sqrt{1-r^2}e^{it} + 2r^2\sqrt{a^2-b^2}$$

covers the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ and its interior.

Plugging the real and imaginary parts of $f(r,t)$ into $\frac{x^2}{a^2} + \frac{y^2}{b^2}$, we obtain

$$\frac{(-\sqrt{a^2-b^2} + 2br\sqrt{1-r^2}\cos(t) + 2r^2\sqrt{a^2-b^2})^2}{a^2} + \frac{(2br\sqrt{1-r^2}\sin(t))^2}{b^2}.$$

A calculation shows that the preceding quantity equals

$$\frac{a^2 - ((2r\sqrt{a^2-b^2}\cos(t) + b\sqrt{1-r^2})\sqrt{1-r^2} - br^2)^2}{a^2}, \tag{5}$$

which is less than or equal to one. Thus every point in the image of f belongs to the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ or its interior. Note now from (5) that a point on the boundary of the ellipse is taken on precisely when

$$((2r\sqrt{a^2-b^2}\cos(t) + b\sqrt{1-r^2})\sqrt{1-r^2} - br^2)^2 = 0;$$

that is, when

$$\cos(t) = \frac{b(2r^2-1)}{2r\sqrt{1-r^2}\sqrt{a^2-b^2}}. \tag{6}$$

Now a little calculation shows that the right hand side of the preceding display increases from -1 to 1 as r varies from $\sqrt{\frac{(a-\sqrt{a^2-b^2})}{2a}}$ to $\sqrt{\frac{(a+\sqrt{a^2-b^2})}{2a}}$. Thus for each r in this closed interval, $\{f(r,t) : t \in [0, 2\pi]\}$ is a circle that hits the boundary of the ellipse for t satisfying (6). The point on the ellipse hit when $r = \sqrt{\frac{(a-\sqrt{a^2-b^2})}{2a}}$ is the point $(-a, 0)$; the point hit when $r = \sqrt{\frac{(a+\sqrt{a^2-b^2})}{2a}}$ is $(a, 0)$. A preservation-of-connectedness argument shows every other point on the ellipse's boundary is hit.

To see that every point in the interior of the ellipse is also in $W(T)$, observe that $f(0,t) = -\sqrt{a^2-b^2}$ and $f(1,t) = \sqrt{a^2-b^2}$ for any t. Thus both foci are hit. Take a point P in the interior of the ellipse that is not one of the foci. Draw a vertical ray from the x-axis through P; this line hits the ellipse at a point s. As we have shown, s belongs to the circle $C_{r_0} := \{f(r_0, t) : t \in [0, 2\pi]\}$ for some r_0 between $\sqrt{\frac{(a-\sqrt{a^2-b^2})}{2a}}$ and $\sqrt{\frac{(a+\sqrt{a^2-b^2})}{2a}}$. Because C_{r_0} has center on the x-axis, an elementary argument using the Pythagorean Theorem shows that P is in the interior of C_{r_0}. Let $r_* = \sup\{r \in [0, r_0] : P \text{ is not inside the circle } C_r\}$. Note 0 is in the set since C_0 is the left focus of the ellipse. By continuity of f, the point P belongs to $\{f(r_*, t) : t \in [0, 2\pi]\}$.

□

We have now seen that the numerical range of a linear operator of $\mathbb{C}^2$ is a filled-in ellipse, which is a convex set. We have also seen that the numerical range of a normal operator on $\mathbb{C}^n$ is convex. Our goal is to show that that the numerical range of any bounded, linear operator is convex. To do so, we will first need a few definitions and lemmas.

Definition 20. *Let $M \subseteq H$ be a subspace. Then $M^{\perp} = \{h \in H : h \perp m \,\forall\, m \in M\}$.*

Theorem 8. *Suppose $M \subseteq H$ is a closed subspace and let $h \in H$. Then there exist unique $h_M \in M$ and $h_{M^\perp} \in M^\perp$ such that $h = h_M + h_{M^\perp}$.*

Proof. We limit ourselves to the case in which H is finite-dimensional. This of course implies that M is finite-dimensional as well. Therefore, we can construct an orthonormal basis $m_1, m_2, \cdots, m_n$ for M and extend it to an orthonormal basis for $H : \{m_1, \cdots, m_n, e_1, e_2, \cdots, e_k\}$ see [1]. Now, let $h \in H$. Then we can write

$$\begin{aligned} h &= <h, m_1> m_1 + <h, m_2> m_2 + \cdots + <h, m_n> m_n + <h, e_1> e_1 + \cdots + <h, e_k> e_k \\ &= \sum_{i=1}^{n} <h, m_i> m_i + \sum_{j=1}^{k} <h, e_j> e_j. \end{aligned}$$

Since $\sum_{i=1}^{n} <h, m_i> m_i \in M$ while $\sum_{j=1}^{k} <h, e_j> e_j \in M^\perp$, we have shown that an arbitrary element of H can be decomposed into the sum of a vector in M and a vector in $M^\perp$. To see that this decomposition is unique, first suppose that it is not. Then there exist $h_M, v_M \in M$ and $h_{M^\perp}, v_{M^\perp} \in M^\perp$ such that $h = h_M + h_{M^\perp}$ and $h = v_M + v_{M^\perp}$. We have

$$\begin{aligned} h_M + h_{M^\perp} &= v_M + v_{M^\perp} \\ h_M - v_M &= v_{M^\perp} - h_{M^\perp}. \end{aligned}$$

But since $h_M - v_M \in M$ and $v_{M^\perp} - h_{M^\perp} \in M^\perp$ by the closure of subspaces under addition, we have that $h_M - v_M$ is perpendicular to itself; that is, we have that

$$<h_M - v_M, h_M - v_M >= 0.$$

But by the second property of Definition 1, this implies $h_M - v_M = 0$. Therefore, $h_M = v_M$. A similar argument shows $v_{M^\perp} = h_{M^\perp}$. The decomposition is unique. Observe too that our proof of uniqueness does not depend on our assumption that H is finite dimensional. □

Definition 21. *Let $M \subseteq H$ be a closed subspace of H. Define $P_M : H \to M$ by $P_M h = h_M$, where h_M is the unique element in M such that $h - h_M$ is in $M^\perp$. P_M is called the orthogonal projection of H onto M.*

Propostion 6. *Let M be a closed subspace of H. Suppose $P_M : H \to M$. P_M is self-adjoint. That is, $P_M^* = P_M$.*

Proof. Let $v, w \in H$. By Theorem 8, there exist $v_M, w_M \in M$ and $v_{M^\perp}, w_{M^\perp} \in M^\perp$ such that $v = v_M + v_{M^\perp}$ and $w = w_M + w_{M^\perp}$. We have

$$\begin{aligned} <P_M v, w> &= <P_M(v_M + v_{M^\perp}), w_M + w_{M^\perp}> \\ &= <v_M, w_M + w_{M^\perp}> \\ &= <v_M, w_M>. \end{aligned}$$

Similarly,

$$\begin{aligned} <v, P_M w> &= <v_M + v_{M^\perp}, P_M(w_M + w_{M^\perp})> \\ &= <v_M + v_{M^\perp}, w_M> \\ &= <v_M, w_M>. \end{aligned}$$

Therefore, $P_M^* = P_M$. □

Theorem 9. *Let M be a proper closed subspace of H such that $\{0\} \neq M$, and suppose $P_M : H \to M$. Then $W(P_M) = [0,1]$.*

Proof. Let $h \in (H)_1$. By Theorem 8, there exist $h_M \in M$ and and $h_{M^\perp} \in M^\perp$ such that $h = h_M + h_{M^\perp}$. Note that

$$\begin{aligned} < P_M h, h > &= < P_M(h_M + h_{M^\perp}), h_M + h_{M^\perp} > \\ &= < h_M, h_M + h_{M^\perp} > \\ &= \|h_M\|^2. \end{aligned}$$

Since $\|h_M\|^2 \leq \|h_M\|^2 + \|h_{M^\perp}\|^2 = 1$, it is clear that $W(T)$ has no element greater than 1. Also, note that $1 \in W(T)$ by choosing $h \in M$. Note as well that $0 \in W(T)$ by choosing $h \in M^\perp$. To see that $W(T) = [0,1]$, let $\alpha \in (0,1)$. Choose $h_M \in M$. Observe then that $\frac{\sqrt{\alpha}}{\|h_M\|} h_M \in M$. Now choose $h_{M^\perp} \in M^\perp$ and note that $\frac{\sqrt{1-\alpha}}{\|h_{M^\perp}\|} h_{M^\perp} \in M^\perp$. Let $h_1 = \frac{\sqrt{\alpha}}{\|h_M\|} h_M + \frac{\sqrt{1-\alpha}}{\|h_{M^\perp}\|} h_{M^\perp}$ and observe that h_1 is a unit vector. We have

$$\begin{aligned} < P_M h_1, h_1 > &= < \frac{\sqrt{\alpha}}{\|h_M\|} h_M, \frac{\sqrt{\alpha}}{\|h_M\|} h_M + \frac{\sqrt{1-\alpha}}{\|h_{M^\perp}\|} h_{M^\perp} > \\ &= \|\frac{\sqrt{\alpha}}{\|h_M\|} h_M\|^2 \\ &= \alpha. \end{aligned}$$

Thus $\alpha \in W(T)$. Therefore, since α was arbitrary, we see that $W(T) = [0,1]$. □

We might also have given a proof of the preceding theorem based on Proposition 6 and Example 8.

Definition 22. *Suppose $M \subseteq H$ is closed and that $T : H \to H$ is bounded and linear. Let $T_M : M \to M$ be the linear operator given by $T_M g = P_M T g$ for all $g \in M$. T_M is called the compression of T onto M.*

It's easy to check that $\|T_M\| \leq \|T\|$. Also, note $T_M = P_M T|_M$.

Lemma 8. *Suppose $M \subseteq H$ is closed and that $T : H \to H$ is bounded and linear. Then $W(T_M) \subseteq W(T)$.*

Proof. Let $\alpha \in W(T_M)$. Then there exists a $v \in (M)_1$ such that $\alpha =< T_M v, v >$. We have

$$\begin{aligned} \alpha &= < T_M v, v > \\ &= < P_M T v, v > \\ &= < T v, P_M v > \quad (P_M \text{ is self}-\text{adjoint}) \\ &= < T v, v >, \end{aligned}$$

where the last inequality comes from the fact that $v \in M$. Thus, $\alpha \in W(T)$. Therefore, $W(T_M) \subseteq W(T)$. □

Theorem 10. *Toeplitz-Hausdorff Theorem. Suppose $T : H \to H$ is bounded and linear. Then $W(T)$ is convex.*

Proof. Let α and β be distinct points in $W(T)$. Then there exist $h_1, h_2 \in (H)_1$ such that $\alpha =< Th_1, h_1 >$ and $\beta =< Th_2, h_2 >$. Let $M = \text{span}\{h_1, h_2\}$. Recall that this means M is the space of all linear combinations of h_1 and h_2. Clearly h_1 and h_2 in particular are in M; in fact they are in $(M)_1$. We will show that $\alpha, \beta \in W(T_M)$, where $T_M : M \to M$ as in Definition 22. We have

$$\begin{aligned} < T_M h_1, h_1 > &= < P_M T h_1, h_1 > \\ &= < T h_1, P_M h_1 > \quad (P_M \text{is self}-\text{adjoint}) \\ &= < T h_1, h_1 > \quad (h_1 \in M) \\ &= \alpha. \end{aligned}$$

Thus, $\alpha \in W(T_M)$. A similar argument using h_2 shows $\beta \in W(T_M)$. We will now use Lemma 7 to show that $W(T_M)$ is a solid ellipse. To see this, we first show $\{h_1, h_2\}$ is linearly independent. Suppose there exist non-zero $c_1, c_2 \in \mathbb{C}$ such that $c_1h_1 + c_2h_2 = 0$. Then $c_1h_1 = -c_2h_2$. Therefore,

$$\begin{aligned} |c_1|^2 &= |c_1|^2\|h_1\|^2 \\ &= \|c_1h_1\|^2 \\ &= \|c_2h_2\|^2 \\ &= |c_2|^2. \end{aligned}$$

We have

$$\begin{aligned} |c_1|^2\alpha &= < Tc_1h_1, c_1h_1 > \\ &= < Tc_2h_2, c_2h_2 > \\ &= |c_2|^2 < Th_2, h_2 > \\ &= |c_2|^2\beta \\ &= |c_1|^2\beta, \end{aligned}$$

so that $(\alpha - \beta)|c_1|^2 = 0$. Since $\alpha \neq \beta$, we conclude $c_1 = 0 = c_2$, so h_1 and h_2 are linearly independent. Since $\{h_1, h_2\}$ spans M, $\{h_1, h_2\}$ is a basis for M. Thus, we see that M is two-dimensional. Let A be the matrix for T_M relative to an orthonormal basis $\{e_1, e_2\}$ of M. We can then view A as a linear operator on $\mathbb{C}^2$ in the usual way. As we discussed after the proof of Proposition 5, $W(T_M) = W(A)$. By Lemma 7, we see then that $W(A)$ is a solid ellipse, so $W(T_M)$ is as well. Therefore, the line segment joining α and β is in $W(T_M)$. Since $W(T_M) \subseteq W(T)$ by Lemma 8, this line segment is also in $W(T)$. Therefore, since α and β were arbitrary, we have that $W(T)$ is convex. □

Principal Results.

Recall that given a linear operator $T : H \rightarrow H$, we wish to find a subspace M of largest dimension such that $W(T_M)$ is a single point. Unless otherwise stated, for the remainder of this paper we will let $T : H \rightarrow H$ be linear, where H is finite-dimensional, and M is a subspace of H.

Before moving into more advanced work, we will look at a few simple cases. First, suppose $M = H$. The following proposition holds:

Propostion 7. *Let $T : H \rightarrow H$. Then $W(T) = \{\alpha\}$ if and only if $T = \alpha I$, where $\alpha \in \mathbb{C}$.*

Proof. First, suppose $W(T) = \{\alpha\}$. Then $\forall\, v \in (H)_1$,

$$\begin{aligned} < Tv, v > &= \alpha \\ &= \alpha < v, v > . \end{aligned}$$

Thus, we have, $\forall\, v \in (H)_1$,

$$\begin{aligned} < Tv, v > - < \alpha v, v > &= 0 \\ < Tv - \alpha v, v > &= 0 \\ < (T - \alpha I)v, v > &= 0, \end{aligned}$$

which implies $T - \alpha I = 0$ by Proposition 2 since the final equality holds for all $v \in H$ once it holds for all $v \in (H)_1$. Thus, $T = \alpha I$.

Now, spse $T = \alpha I$ and let $v \in (H)_1$. We have

$$\begin{aligned} < Tv, v > &= < \alpha v, v > \\ &= \alpha < v, v > \\ &= \alpha. \end{aligned}$$

Since $v \in (H)_1$ was arbitrary, we have $W(T) = \{\alpha\}$. $\square$

Thus, we can fully understand the case in which $M = H$. We have learned that only a scalar multiple of the identity can have a one-point numerical range on all of H. Now we want to move on to studying proper subspaces. Probably the simplest case to understand is that in which M is an eigenspace. If λ is an eigenvalue of T and M is the eigenspace associated with λ, then clearly $W(T_M) = \{\lambda\}$.

We now move on to consider situations in which the subspace is neither the entire space nor necessarily an eigenspace.

Lemma 9. *Suppose $W(T_M) = \{\alpha\}$ for $\alpha \in \mathbb{C}$. Then $T_M v = \alpha v \ \forall \ v \in (M)_1$. (Observe that this implies α is an eigenvalue of T_M.)*

Proof. Let $v \in (M)_1$. We have $< T_M v, v >= \alpha$, which can be rewritten as $< T_M v, v >=< \alpha v, v >$, which implies $T_M v = \alpha v$ by Proposition 2. $\square$

Note that this does not necessarily imply that α is an eigenvalue of T:

Example 12. *Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be given by*

$$T = \begin{bmatrix} 3 & 2 & 0 & 1 \\ 4 & 1 & 3 & 2 \\ 0 & 1 & 3 & 4 \\ 3 & 2 & 3 & 2 \end{bmatrix},$$

and let $M = span\{e_1, e_3\}$, where $\{e_1, e_2, e_3, e_4\}$ are the natural basis vectors of $\mathbb{C}^4$. Observe $W(T_M) = \{3\}$. While this implies that 3 is an eigenvalue of T_M, a calculation shows that 3 is not an eigenvalue of T.

We would like to move on to bigger, more general results. For example, let's find an M of larger dimension than we have previously such that $W(T_M)$ is a single point.

Example 13. *Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be given by*

$$T = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \tag{7}$$

and let

$$M = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}. \tag{8}$$

Then $W(T_M) = \{2\}$.

One can see this result holds by letting $v \in (M)_1$. Then $v = (v_1, v_2, v_2, v_3)$, where $|v_1|^2 + 2|v_2|^2 + |v_3|^2 = 1$. We have

$$\begin{aligned} < Tv, v > &= < (2v_1, 8v_2, -4v_2, 2v_3), (v_1, v_2, v_2, v_3) > \\ &= 2|v_1|^2 + 8|v_2|^2 - 4|v_2|^2 + 2|v_3|^2 \\ &= 2(|v_1|^2 + 2|v_2|^2 + |v_3|^2) \\ &= 2. \end{aligned}$$

Since $v \in (M)_1$ is arbitrary, we see that indeed $W(T_M) = \{2\}$. Moreover, note that the vector $(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0) \in M$, yet is not an eigenvector of T. Thus, M is not an eigenspace.

Thus, it is possible to construct a subspace M such that $W(T_M)$ is an eigenvalue of T, but M is not an eigenspace. Would it be possible to construct a 3-dimensional M for our T such that $W(T_M)$ is not an eigenvalue? As we will see in the next theorem, the answer is no.

Theorem 11. *Let $T : H \to H$, and let $M \subseteq H$ be a subspace such that $\dim M > \frac{1}{2} \dim H$. Suppose $W(T_M) = \{\alpha\}$. Then α is an eigenvalue of T. Moreover, α has multiplicity no less than* $\dim M - \dim M^\perp$.

Proof. Let $v \in M$. Observe

$$\begin{aligned} Tv &= P_M Tv + P_{M^\perp} Tv \quad \text{(Theorem 8)} \\ &= T_M v + P_{M^\perp} Tv \\ &= \alpha v + P_{M^\perp} Tv, \end{aligned}$$

where the final equality comes from Lemma 9. If we can find a vector $v \in (M)_1$ such that $P_{M^\perp} Tv = 0$, we will have proven α is an eigenvalue of T. Consider $P_{M^\perp} T|_M : M \to M^\perp$. By a theorem from linear algebra, we know

(9) $$\dim M = \dim \ker P_{M^\perp} T|_M + \dim \text{range } P_{M^\perp} T|_M.$$

Note, however, that since M and $M^\perp$ intersect only at 0, $\dim M + \dim M^\perp = \dim H$. Since $\dim M > \frac{1}{2}\dim H$, $\dim M > \dim M^\perp$. Thus, by Equation (9), we have:

$$\begin{aligned} \dim \ker P_{M^\perp} T|_M &= \dim M - \dim \text{range } P_{M^\perp} T|_M \\ &\geq \dim M - \dim M^\perp > 0. \end{aligned}$$

Therefore, there exists a non-zero vector $v \in M$ such that $P_{M^\perp} Tv = 0$. Therefore, α is an eigenvalue of T.

To see that α has multiplicity greater than or equal to $\dim M - \dim M^\perp$, let n be the multiplicity of α. Next, note that every basis vector of $\ker P_{M^\perp} T|_M$ is an eigenvector for T corresponding to α. We have

$$\begin{aligned} n &\geq \dim \ker P_{M^\perp} T|_M \\ &\geq \dim M - \dim M^\perp. \end{aligned}$$

Therefore, α is of multiplicity no less than $\dim M - \dim M^\perp$. □

Corollary 1. *Suppose α is an eigenvalue of T of multiplicity n. Then any subspace M such that $W(T_M) = \{\alpha\}$ can have dimension no larger than $\frac{n+\dim H}{2}$.*

Proof. Let M be a subspace of H such that $W(T_M) = \{\alpha\}$. If $\dim M \leq \frac{1}{2} \dim H$, the conclusion obviously holds. Next, suppose $\dim M > \frac{1}{2} \dim H$. By Theorem 11 above, we have

$$\begin{aligned} \dim M &\leq n + \dim M^\perp \\ &= n + (\dim H - \dim M); \quad \text{thus,} \\ 2 \dim M &\leq n + \dim H, \quad \text{or} \\ \dim M &\leq \frac{n + \dim H}{2}, \end{aligned}$$

as desired. $\square$

Corollary 2. *Let $T : H \to H$, and let α be an eigenvalue of T. Suppose that $M$is the subspace of H of largest dimension such that $W(T_M) = \{\alpha\}$ and $\dim M > \frac{\dim H}{2}$. Then there is no subspace D of H of larger dimension such that T_D has a one-point numerical range.*

Proof. We will argue by contradiction. Let α and M be as described above, with $\dim M = n$. Also, let $q = \dim H$. Suppose there exists a subspace D of dimension p, with $p > n$ such that $W(T_D) = \beta$, where $\beta \neq \alpha$. Let $\{m_1, m_2, \cdots, m_n\}$ be a basis for M, and let $\{d_1, d_2, \cdots, d_p\}$ be a basis for D. Since $n + p > q$, there exist $c_1, c_2, \cdots, c_{n+p}$, not all zero, such that

$$c_1 m_1 + \cdots + c_n m_n + c_{n+1} d_1 + \cdots + c_{n+p} d_p = 0.$$

This tells us

$$c_1 m_1 + \cdots + c_n m_n = -(c_{n+1} d_1 + \cdots + c_{n+p} d_p),$$

which tells us M and D have a common vector. Note that this vector is non-zero because if it were zero, $\{c_1, c_2, \cdots c_{n+p}\}$ would all be zero by the linear independence of $\{m_1, m_2, \cdots, m_n\}$ and $\{d_1, d_2, \cdots, d_p\}$. Normalizing this vector shows us that $(M)_1$ and $(D)_1$ have a common vector, call it v. Therefore, we have

$$\begin{aligned} \alpha &= < P_M T v, v > \\ &= < Tv, P_M v > \\ &= < Tv, P_D v > \\ &= < P_D T v, v > \\ &= \beta, \end{aligned}$$

a contradiction. Therefore, M is the largest subspace of H such that T_M has a one-point numerical range. $\square$

Sometimes it happens, however, that T has no eigenvalues with an associated M of dimension greater than $\frac{1}{2} \dim H$. For instance, consider Example 12 above. All the eigenvalues of T in this example have multiplicity 1, which means their associated spaces M can have dimension no larger than 2. Can we use this information to put a general bound on the dimension of subspaces M such that $W(T_M) = \{\alpha\}$?

Corollary 3. *Let $\dim H = n$ and suppose $T : H \to H$ has n distinct eigenvalues. Then H has no subspace M of dimension larger than $\frac{\dim H+1}{2}$ such that T_M has a one-point numerical range.*

Proof. We will argue by contradiction. Suppose there exists a subspace M of H such that $\dim M > \frac{\dim H+1}{2}$ and $W(T_M) = \{\alpha\}$ for $\alpha \in \mathbb{C}$. Then by Theorem 11, α is an eigenvalue of T. Since T has n eigenvalues, α has multiplicity 1. Thus, by Corollary 1, M can have dimension no larger than $\frac{\dim H+1}{2}$, contrary to our hypothesis. Therefore, H can have no subspace M of dimension larger than $\frac{\dim H+1}{2}$ such that $W(T_M) = \{\alpha\}$. $\square$

Let's look at a few examples using this idea of bounding the size of M. To assist us, we will want another theorem.

Definition 23. *Let $T : H \to H$. We say α is a corner point of $W(T)$ provided that $W(T)$ is contained in the area bounded by two rays forming an angle of 180 degrees at their common vertex α.*

Lemma 10. *Spse $T : H \to H$ and $W(T)$ is a line segment. Then T is normal.*

Proof. Spse $W(T)$ is the line segment with endpoints α and β, for α, $\beta \in \mathbb{C}$. Then $W(T - \alpha I)$ is the line segment connecting 0 and $(\beta - \alpha)$. Now, let θ be the angle this line segment makes with the real axis, and let $Q = e^{-i\theta}(T - \alpha I)$. Then $W(Q)$ is a line segment on the real axis with one endpoint at 0 and length

$|\beta - \alpha|$. Thus, $W(Q) \subseteq \mathbb{R}$. Therefore, for all $v \in (H)_1$, we have $< Qv, v > = < v, Qv >$. Claim that this implies Q is self-adjoint. To prove the claim, observe that for all $v \in V$, we have

$$\begin{aligned} < Qv, v > &= < v, Qv > \\ < Qv, v > - < Q^*v, v > &= 0 \\ < (Q - Q^*)v, v > &= 0, \end{aligned}$$

which implies $Q = Q^*$ by Propsition 2. Thus, Q is self-adjoint. Since

$$Q = e^{-i\theta}(T - \alpha I)$$

and

$$Q^* = e^{i\theta}(T^* - \overline{\alpha} I),$$

we have

$$e^{i\theta}(T^* - \overline{\alpha} I) = e^{-i\theta}(T - \alpha I),$$

which implies

$$T^* = e^{-2i\theta}(T - \alpha I) + \overline{\alpha} I.$$

Clearly, then, $TT^* = T^*T$, so T is normal. □

Theorem 12. *Suppose c is a corner point of $W(T)$, and suppose $v \in (H)_1$ is such that $c = < Tv, v >$. Then $Tv = cv$; that is, c is an eigenvalue of T with associated eigenvector v.*

Proof. By Theorem 8 and Lemma 9, we know that there exist $\gamma \in \mathbb{C}$ and $w \in (H)_1 \cap < v >^\perp$ such that

$$Tv = cv + \gamma w. \tag{10}$$

(Here, $< v >$ is the one-dimensional subspace spanned by v. Let $M = \text{span}\{v, w\}$. Thus, we know $W(T_M)$ is an ellipse contained in $W(T)$, and we know $c \in W(T_M)$. Since c is a corner point, we know $W(T_M)$ must be a degenerate ellipse, i.e, a line segment. Thus, T_M is normal by Lemma 10. This means that since c is an endpoint of $W(T_M)$, c must be an eigenvalue of T_M. Now, either $W(T_M)$ is a degenerate line segment or it has positive length. If it is degenerate, we have $W(T_M) = \{c\}$. By Lemma 9, we have $T_M v = cv$. But by Equation 10, we now have

$$\begin{aligned} cv &= T_M v \\ &= P_M T v \\ &= P_M(cv + \gamma w) \\ &= cv + \gamma w, \end{aligned}$$

since $(cv + \gamma w) \in M$. Thus, we have $\gamma = 0$, so by Equation 10, we have $Tv = cv$. Thus, c is an eigenvalue of T with associated eigenvector v.

Now suppose $W(T_M)$ has positive length. Let δ be the other endpoint of the line segment. By Example 8, c and δ are eigenvalues of T_M. Now suppose that for $v_1 \in (M)_1$, we have

$$T_M v_1 = cv_1,$$

and that for $w_1 \in (M)_1$, we have

$$T_M w_1 = \delta w_1.$$

Because eigenvectors corresponding to distinct eigenvalues of normal operators are orthogonal, $\{v_1, w_1\}$ is an orthonormal basis of M. Then there exist $\alpha, \beta \in \mathbb{C}$ such that

$$v = \alpha v_1 + \beta w_1.$$

We have

$$\begin{aligned} c &= < Tv, v > \\ &= < T(\alpha v_1 + \beta w_1), P_M(\alpha v_1 + \beta w_1) > \\ &= < T_M(\alpha v_1 + \beta w_1), \alpha v_1 + \beta w_1 > \\ &= < \alpha c v_1 + \beta \delta w_1, \alpha v_1 + \beta w_1 > \\ &= |\alpha|^2 c + |\beta|^2 \delta. \end{aligned}$$

For this equality to hold, $|\alpha|^2$ must equal 1, so that $\beta = 0$. So, we have

$$\begin{aligned} v &= \alpha v_1 + \beta w_1 \\ &= e^{i\theta} v_1. \end{aligned}$$

Thus, we have

$$\begin{aligned} T_M v &= T_M(e^{i\theta} v_1) \\ &= e^{i\theta} T_M(v_1) \\ &= c(e^{i\theta} v_1) \\ &= cv. \end{aligned}$$

But

$$\begin{aligned} T_M v &= P_M(cv + \gamma w) \\ &= (cv + \gamma w), \end{aligned}$$

so we have that $\gamma = 0$. Thus, $Tv = cv$, as desired. □

We are now in a position to look at a few specialized examples of bounding the size of M.

Example 14. *Let $T : \mathbb{C}^4 \to \mathbb{C}^4$ be given by*

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

Then the subspace M of largest dimension such that $W(T_M)$ is a single point does not correspond to an eigenvalue of T.

Clearly, the eigenvalues of T are $1, -1, i$, and $-i$. Let α be any of these, and suppose there exists a two-dimensional subspace $M = \text{span}\{h_1, h_2\}$ for $h_1, h_2 \in (M)_1$ such that $W(T_M) = \{\alpha\}$. Clearly, then, we have

$$\begin{aligned} \alpha &= < T_M h_1, h_1 > \\ &= < Th_1, h_1 > . \end{aligned}$$

Now, since T is normal, we have by Example 8 that $W(T)$ is the convex hull of the eigenvalues of T. It is easy to verify that α is a corner point of $W(T)$. Therefore, by Theorem 12, h_1 is an eigenvector with associated eigenvalue α. A similar argument shows that h_2 is as well. Note that this means h_1 and h_2 are linearly independent. They are also linearly independent from the eigenvectors associated with the other three eigenvalues of T. This means we have at least 5 linearly independent vectors in $\mathbb{C}^4$, a contradiction. Therefore, an eigenvalue can have an associated subspace M of dimension no greater than 1.

We will now show that there exists a two-dimesional M such that $W(T_M) = \{0\}$. Let $M = \text{span}\{h_3, h_4\}$, where $h_3 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0)$ and $h_4 = (0, 0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. It is easy to check that $< Th_3, h_3 >= 0 =< Th_4, h_4 >$. Now, let $(mh_3 + nh_4) \in (\mathbb{C}^4)_1$ for $m, n \in \mathbb{C}$. We have

$$\begin{aligned} < T(mh_3 + nh_4), (mh_3 + nh_4) > &= < Tmh_3, mh_3 > + < Tmh_3, nh_4 > + < Tnh_4, mh_3 > + < Tnh_4, nh_4 > \\ &= |m|^2 < Th_3, h_3 > + |n|^2 < Th_4, h_4 > + m\bar{n} < Th_3, h_4 > + n\bar{m} < Th_4, h_3 > \\ &= m\bar{n} < Th_3, h_4 > + n\bar{m} < Th_4, h_3 > \\ &= 0, \end{aligned}$$

since it is easy to check that $Tmh_3 \in < h_4 >^{\perp}$ and vice versa. Thus, $W(T_M) = \{0\}$, as desired.

Note, then, that the subspace M of largest dimension such that $W(T_M) = \{\alpha\}$ need not be such that α is an eigenvalue. We now turn to an example with an extremely small M.

Example 15. *Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by*

$$T = \begin{bmatrix} 2+i & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2-57i \end{bmatrix}.$$

Then if M is a subspace of $\mathbb{C}^3$ such that $W(T_M) = \{\alpha\}$, $\dim M = 1$.

This can be easily verified without doing any calculations. Note that if M is such that $W(T_M) = \{\alpha\}$, and if $\dim M > 1$, α must be an eigenvalue of T by Thereom 11. But, by the paragraph immediately following Example ?? above, we see that since α has multiplicity 1, $\dim M = 1$.

We have now seen a few ways to put upper bounds on the size of the dimension of a subspace M yielding a compression with a one-point numerical range. The next question is, if one has a subspace M smaller than the bounding dimension, how can one augment this M to obtain, if possible, a subspace of larger dimension with the same numerical range? We now address ourselves to finding answers. First, a definition.

Definition 24. *Let M be a proper subspace of H. Then for $w \in H$ such that w is not in M, we define $M \vee \{w\}$ to be the subspace generated by M and w so that $M \vee \{w\} = \{\alpha w + \beta v : \alpha, \beta \in \mathbb{C}$ and $v \in M\}$.*

One convenient way to augment a vector space in this fashion is to choose $w \in M^{\perp}$.

Theorem 13. *Let M be a proper subspace of H such that $W(T_M) = \{\alpha\}$. Suppose $w \in (M^{\perp})_1$ such that $< Tw, w >= \alpha$. Let $D = M \vee \{w\}$. Then $W(T_D)$ is a single point if and only if $< Tv, w >= 0 =< Tw, v >$ for all $v \in M$.*

Proof. Suppose $< Tv, w >=< Tw, v >= 0$ for all $v \in M$. Let $c_1 v + c_2 w \in (D)_1$ for some $v \in M$. We have

$$\begin{aligned} < T_D(c_1v + c_2w), c_1v + c_2w > &= < T(c_1v + c_2w), c_1v + c_2w > \\ &= < Tc_1v, c_1v > + < Tc_2w, c_2w > + c_1\bar{c}_2 < Tv, w > + c_2\bar{c}_1 < Tw, v > \\ &= \alpha(|c_1|^2 + |c_2|^2) \quad \text{(by hypothesis)} \\ &= \alpha \quad (\text{since } \mathrm{c_1v + c_2w \in (D)_1}) \end{aligned}$$

Since $c_1v + c_2w$ is arbitrary, we have that $W(T_D) = \{\alpha\}$.

Now, suppose $W(T_D) = \{\alpha\}$. Let $v \in M$ and $c_1, c_2 \in \mathbb{C}\backslash\{0\}$ be arbitrary such that $c_1v + c_2w \in (D)_1$. We have

$$\begin{aligned} \alpha &= < T_D(c_1v + c_2w), (c_1v + c_2w) > \\ &= \alpha + c_1\bar{c}_2 < Tv, w > + c_2\bar{c}_1 < Tw, v >, \end{aligned}$$

which implies

$$c_1\bar{c}_2 < Tv, w > + c_2\bar{c}_1 < Tw, v >= 0. \tag{11}$$

Now, observe that for $c_1v + ic_2w \in (D)_1$, we have

$$\begin{aligned} \alpha &= < T_D c_1 v + ic_2 w, c_1 v + ic_2 w > \\ &= \alpha - c_1\bar{c}_2 i < Tv, w > + ic_2\bar{c}_1 < Tw, v >, \end{aligned}$$

which implies

$$\bar{c}_1 c_2 < Tw, v > -\bar{c}_2 c_1 < Tv, w >= 0.$$

Adding this equation to equation 11, we see that

$$2c_2\bar{c}_1 < Tw, v >= 0,$$

or $< Tw, v >= 0$. Plugging this value into equation 11, we see that $< Tv, w >= 0$ as well. □

We now have a way to work toward finding subspaces M of larger dimension. Used in conjunction with Theorem 11 and its corollaries, it will often allow us to find, beyond any doubt, the subspace M of largest dimension such that $W(T_M)$ is a single point.

Example 16. *Let $T : \mathbb{C}^3 \to \mathbb{C}^3$ be given by*

$$T = \begin{bmatrix} 3 & i & 0 \\ \frac{5}{2}i & 0 & -\frac{5}{2}i \\ 2 & 3i & 1 \end{bmatrix}.$$

Find the subspace M of largest dimension such that $W(T_M)$ is a single point.

Observe that 3 is an eigenvalue for T with associated eigenvector $v = (\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$. Moreover, it is easily verified that the eigenspace associated with 3 is simply the space $M = \text{span}\{v\}$. Clearly $W(T_M) = \{3\}$. We can use Theorem 14 to augment M to find a two-dimensional space D such that $W(T_D) = \{3\}$. We must find a vector w such that the following conditions hold for all $m \in M$:

1. $\|w\| = 1$;

2. $< m, w >= 0$;

3. $< Tw, w >= 3$;

4. $< Tm, w >= 0$; and

5. $< Tw, m >=0$.

For Equation 2 to hold, it is easy to see that $w = (w_1, w_2, -w_1)$ for $w_1, w_2 \in \mathbb{C}$. Since $\forall\ m \in M$, m is an eigenvector of T, this will also ensure that Equation 4 holds. Observe that we may rewrite Equation 5 as $< w, T^*m >= 0$. Now, let $m \in M \setminus \{0\}$, so $m = (b, 0, b)$, where $b \neq 0$. We have

$$\begin{aligned} 0 &= < w, T^*m > \\ &= < \begin{pmatrix} w_1 \\ w_2 \\ -w_1 \end{pmatrix}, \begin{bmatrix} 3 & -\frac{5}{2}i & 2 \\ -i & 0 & -3i \\ 0 & \frac{5i}{2} & 1 \end{bmatrix} \begin{pmatrix} b \\ 0 \\ b \end{pmatrix} > \\ &= < \begin{pmatrix} w_1 \\ w_2 \\ -w_1 \end{pmatrix}, \begin{pmatrix} 5b \\ -4ib \\ b \end{pmatrix} > \\ &= \bar{b}(4w_1 + 4iw_2), \end{aligned}$$

which implies $w_1 = -iw_2$. We now have $w = (-iw_2, w_2, iw_2)$. Because w is a unit vector, we have

$$1 = \sqrt{3|w_2|^2},$$

or $|w_2| = \frac{1}{\sqrt{3}}$. If we let $w_2 = \frac{1}{\sqrt{3}}$, we have

$$w = (\frac{-i}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{i}{\sqrt{3}}).$$

It is easy to verify that this w does indeed satisfy all five equations. Let $D = M \vee \text{span}\{\text{w}\}$. By Theorem 13, we have that $W(T_D) = \{3\}$. Can we find a subspace E of larger dimension such that $W(T_E) = \{3\}$? It easily verified that 3 is an eigenvalue of multiplicity 1. Therefore, by Corollary 1, there exists no such subspace E. In fact, we can go further and say there exists no subspace E of larger dimension such that $W(T_E)$ is a single point. If $\dim M = 3$, then $T = 3I$ by Proposition 7, which is clearly false.

References

[1] Sheldon Axler, *Linear Algebra Done Right*, Springer, New York, 1996.

[2] Paul Bourdon, letter to Joel Shapiro, May 1999.

[3] Ronald G. Douglas, *Banach Algebra Techniques in Operator Theory*, Academic Press, New York, 1972.

[4] Michael A. Nielsen and Isaac L. Chuan, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[5] Karen Saxe, *Beginning Functional Analysis*, Springer, New York, 2002.

[6] Christina Spofford, "Norms of Composition Operators on the Hardy Space," Thesis, Washington and Lee University, 2003.