ON THE NUMERICAL RANGE WITH RESPECT TO A FAMILY OF PROJECTIONS

WAED DADA, JOACHIM KERNER, AND NAZIFE ERKURŞUN-ÖZCAN

Abstract. In this note we introduce the concept of a numerical range of a bounded linear operator on a Hilbert space with respect to a family of projections. We give a precise definition and elaborate on its connection to the classical numerical range as well as to generalizations thereof such as the quadratic numerical range, block numerical range, and product numerical range. In general, the importance of this new notion lies within its unifying aspect.

1. Introduction

In this paper, we establish an abstract notion of a numerical range which forms a direct generalization of the well-known (classical) numerical range of a bounded linear operator, defined on a (separable, complex) Hilbert space, as well as of other existing versions of a numerical range. For a given operator \( A \), the classical numerical range is defined as

\[
W(A) := \{ \langle Ax, x \rangle : x \in H, \|x\| = 1 \} \subset \mathbb{C}.
\]

Originally introduced for linear operators on \( \mathbb{C}^n \) (i.e., matrices) by Toeplitz [19] and Hausdorff [8], it was later extended to more general operators by Stone [16]. Unlike the spectrum, the numerical range is a unitary invariant but in general not invariant under similarity transformations and hence provides additional information about the operator [17]. In particular, the numerical range allows to localize the spectrum and to estimate the resolvent. More generally, one has the following:

1. If \( H = \mathbb{C}^2 \) then \( W(A) \) is a (possibly degenerate) ellipse.
2. If \( H \) is finite-dimensional then \( W(A) \) is compact.
3. \( \sigma(A) \subseteq \overline{W(A)} \) and \( \sigma_p(A) \subseteq W(A) \). (Spectral Inclusion)
4. \( W(A) \subset \mathbb{C} \) is convex. (Toeplitz-Hausdorff Theorem)
5. \( \| (\lambda - A)^{-1} \| \leq \frac{1}{\text{dist}(\lambda, W(A))} \), \( \lambda \notin \overline{W(A)} \).

Furthermore, as shown in [4], if \( A \) is compact on an infinite-dimensional \( H \) then \( W(A) \) is closed if \( 0 \in W(A) \). If, in addition, \( A \) is self-adjoint then \( W(A) \) is the convex hull of the point spectrum of \( A \) (see also Proposition 3.4).

The great advantage of the numerical range, when compared to the spectrum, is that it is relatively easy to compute (certainly in the case of matrices). It was for this reason why it became a useful tool in many applications in physics [2] and different branches of pure and applied mathematics such as control theory [15], numerical analysis [7] and operator theory [13]. Most importantly, however, the concept of the numerical range turned out to be much more flexible and adaptable to applications which consequently led to various different versions of a numerical range: quadratic and block numerical range [10],

\[2010 \text{ Mathematics Subject Classification.} \ 47A12, 47A10, 47A46, 15A60, 65Fxx.\]
\[\text{Key words and phrases.} \ \text{Numerical range, bounded operator, spectrum.}\]
c-numerical range [11], essential numerical range [5, 1], joint numerical range [18] and higher numerical range [3, 12] whose motivation came from quantum error correction.

In this paper, besides the connection to the classical numerical range, we are particularly interested in the connection of the abstract notion of a numerical range to the quadratic and block numerical range since these concepts proved particularly powerful in localizing the spectrum of the operator and sharpening resolvent estimates [9, 10]. In addition, we establish a connection to the so-called product numerical range which plays an important role in quantum information theory [14, 6].

2. Preliminaries and definitions

Since we are interested in the definition of the numerical range of a linear bounded operator \( A \) with respect to families of orthogonal projections we define, for \( k \in \mathbb{N} \),

\[
\mathcal{P} := \{ P \in \mathcal{L}(H) : P \text{ orthogonal projection in } H \},
\]

\[
\mathcal{P}_k := \{ P \in \mathcal{P} : \dim(\text{ran}(P)) = k \},
\]

as well as

\[
\mathcal{P}_A := \{ P \in \mathcal{P} : PA = AP, \dim(\text{ran}(P)) < \infty \}.
\]

**Proposition 2.1.** The sets \( \mathcal{P}, \mathcal{P}_k \subset \mathcal{L}(H) \) are closed with respect to the operator norm.

**Definition 2.2.** For \( A \in \mathcal{L}(H) \) and \( P \in \mathcal{P} \) we define an operator \( A_P \) on the range \( \text{ran}(P) \) by

\[
A_P : \text{ran}(P) \to \text{ran}(P), \quad x \mapsto A_Px := PAx.
\]

The relation between \( A_P \) and \( A \) can be expressed by

\[
A_P P = PAP.
\]

The operator \( A_P \) is called the compression of \( A \) to \( \text{ran}(P) \) and \( A \) is called a dilation of \( A_P \) to \( H \).

**Remark 2.3.** We have \( W(A_P) \subset W(A) \): For any \( \lambda \in W(A_P) \) there exists \( x \in \text{ran}(P) \) with \( \|x\| = 1, Px = x \) and

\[
\lambda = \langle A_Px, x \rangle = \langle PAPx, x \rangle = \langle APx, Px \rangle = \langle Ay, y \rangle,
\]

where \( y := Px \). Since \( \|y\| = \|Px\| = \|x\| = 1 \) we conclude \( \lambda \in W(A) \).

Due to the spectral inclusion (see 5. in the list above) Remark 2.3 implies, in particular, that each \( \lambda \in \sigma(A_P) \) is contained in \( W(A) \) given that \( \dim(\text{ran}(P)) < \infty \). This motivates the following definition.

**Definition 2.4** (Numerical range with respect to a family of projections). Let \( A \in \mathcal{L}(H) \) be a bounded operator and \( \mathcal{P} \subseteq \mathbb{P} \). Then

\[
W_P(A) := \bigcup_{P \in \mathcal{P}} \sigma(A_P)
\]

is called the numerical range of \( A \) with respect to the family of orthogonal projections \( \mathcal{P} \) or \( \mathcal{P} \)-numerical range of \( A \) for short.

**Remark 2.5.**

(1) \( W_P(A) \) is in general not closed.

(2) \( W_P(A) \subset \{ \lambda \in \mathbb{C} : |\lambda| \leq \|A\| \} \).

(3) \( W_P(A^*) = (W_P(A))^* \).

(4) For \( A \in \mathcal{L}(H) \) self-adjoint one has \( W_P(A) \subset \mathbb{R} \).
3. Main Results

3.1. Connection to the classical numerical range, higher-rank numerical range and the (point) spectrum. The first result establishes the connection with the classical numerical range. For the proof note that, for each $P \in \mathcal{P}_k$ and for any orthonormal basis $\{f_i\}_{i=1}^k$ of $\text{ran}(P)$, one has

$$Px = \sum_{i=1}^k (x, f_i) f_i \quad \forall x \in H .$$

**Theorem 3.1.** For $A \in \mathcal{L}(H)$ we have $W_{\mathcal{P}_1}(A) = W(A)$.

**Proof.** Let $\lambda \in W_{\mathcal{P}_1}(A)$. Then there exist $P \in \mathcal{P}_1$ and $f \in \text{ran}(P)$ with $\|f\| = 1$ such that $PAPf = \lambda f$. Therefore

$$\langle Af, f \rangle = \langle APf, P f \rangle = \langle PAPf, f \rangle = \langle \lambda f, f \rangle = \lambda$$

and hence $\lambda \in W(A)$.

If $\lambda \in W(A)$ then there exists $f \in H$ with $\|f\| = 1$ such that $\lambda = \langle Af, f \rangle$. Let $P$ denote the orthogonal projection onto span$\{f\}$. Then, according to (3.1),

$$PAPf = PAf = \langle Af, f \rangle f = \lambda f$$

and hence $\lambda \in \sigma(A_P)$. Thus $\lambda \in W_{\mathcal{P}_1}(A)$.

**Remark 3.2.** Theorem 3.1 is interesting from the following point of view: In general the spectrum forms only a “small” subset of $W(A)$ (for example, think of a matrix $A \in \mathbb{C}^2$ for which the spectrum consists of two points whereas $W(A)$ is a (possibly degenerate) ellipse). However, by considering the union of all $\sigma(A_P)$ for $P \in \mathcal{P}_1$ instead, the whole classical numerical range is obtained by “filling it up” with spectral values, see also Proposition 3.4.

**Remark 3.3.** By Theorem 3.1, $W_{\mathcal{P}_1}(A)$ is a convex set.

In order to illustrate Remark 3.2 even more, we present the following statement which is due to [4].

**Proposition 3.4.** Let $A \in \mathcal{L}(H)$ be a compact, self-adjoint operator. Then one has

$$W_{\mathcal{P}_1}(A) = co\{\sigma_p(A)\} .$$

**Proof.** “$\supseteq$” Since $co\{\sigma_p(A)\} \subseteq co\{W(A)\} = W(A)$, $W_{\mathcal{P}_1}(A)$.

“$\subseteq$” We follow [4]: Let $\lambda \in W_{\mathcal{P}_1}(A) \backslash co\{\sigma_p(A)\}$ be given. Since $\lambda \in W_{\mathcal{P}_1}(A)$ there exists $x \in \text{ran}(P)$ with (w.l.o.g.) $\|x\| = 1$ such that $PAPx = \lambda x$. Hence

$$\lambda = \langle \lambda x, x \rangle = \langle PAPx, x \rangle = \langle APx, Px \rangle = \langle Ax, x \rangle .$$

Employing the spectral theorem one obtains $\lambda = \sum_{n=1}^{\infty} \mu_n \langle x, e_n \rangle^2$, $\{\mu_n\}$ being the eigenvalues of $A$ with associated eigenvectors $\{e_n\}$, and consequently we obtain the relation

$$\sum_{n=1}^{\infty} (\lambda - \mu_n) \langle x, e_n \rangle^2 = 0 .$$

Without loss of generality one can assume that $\langle x, e_n \rangle \neq 0$ for all $n$ and hence we conclude that there exist $n_1, n_2 \in \mathbb{N}$ such that $(\lambda - \mu_{n_1}), (\lambda - \mu_{n_2})$ have opposite sign. However, this readily implies that $\lambda \in [\mu_{n_1}, \mu_{n_2}]$ and consequently $\lambda \in co\{\sigma_p(A)\}$ which is a contradiction.

The following statement is a direct generalization of Theorem 3.1.

**Lemma 3.5.** For $A \in \mathcal{L}(H)$ and the family $\mathcal{P}_k$ the following holds:
we conclude that (possibly after restricting to a subsequence) 
\[ \lambda f = \lim_{n \to \infty} (\lambda_n f_n) = \lim_{n \to \infty} (P_n A P_n f_n) = P A P f \]
for some (normalized) \( \lambda \) that, for arbitrary \( A \in \mathcal{L}(H) \), 

\[ \lambda \in \sigma_r(A) \]

By Proposition 3.4 and Theorem 3.6 we see that \( \sigma_r(A) = \sigma(A) \).

**Remark 3.7.** By Proposition 3.4 and Theorem 3.6 we see that \( \sigma_0(A) = \sigma(A) \).

Furthermore, regarding Theorem 3.6 we note the following: if \( A \) is such that \( \sigma_0(A) = 0 \) then there cannot exist an orthogonal projection \( P \) with finite-dimensional range commuting with \( A \). As an example, one might think of the right-shift operator \( R \) on \( \ell^2(\mathbb{N}) \) which has no eigenvalues. This implies that \( R \) has no non-trivial finite-dimensional invariant subspaces. On the other hand, the left-shift operator \( L \) on \( \ell^2(\mathbb{N}) \) is known to have the open unit disk as point spectrum. However, since the adjoint of \( L \) is the right-shift operator \( R \), there cannot exist any finite-dimensional reducing subspaces as asked after in (2.3) since they would be invariant under \( L \) as well as under \( R \). Consequently, \( \mathcal{P}_L(L) = \emptyset \) and the inclusion in Theorem 3.6 is strict.
In order to formulate an analogue of Theorem 3.6 for normal operators we introduce
\begin{equation}
\mathcal{P}_A := \{ P \in \mathbb{P} : PA = A^*P, \dim(\text{ran}(P)) < \infty \}.
\end{equation}

Based on this family of projections we then obtain the following result.

**Theorem 3.8.** Let \( A \in \mathcal{L}(H) \) be a normal operator. Then \( W_{\mathcal{P}_A}(A) = \sigma_A(A^*) = (\sigma_P(A))^* \).

**Proof.** The proof follows the same line as the proof of Theorem 3.6. We only remark that line (3.7) now reads, \( \lambda \in \sigma_A(A^*) \),
\[
PAx = (Ax, f)f = (x, A^* f)f = (x, A^* f)\lambda f = A^*Pxf.
\]
Also note that \( \sigma_P(A^*) = (\sigma_P(A))^* \) holds for all normal operators. \( \square \)

### 3.2. Connection to the quadratic and block numerical range.

As defined in [9] (and discussed in detail in [10]), the quadratic numerical range of a \( 2 \times 2 \)-block operator matrix
\begin{equation}
\mathcal{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix},
\end{equation}
with \( \mathcal{A} \) acting as an operator on \( H_1 \oplus H_2 \), is the set of all eigenvalues of all \( 2 \times 2 \)-matrices
\begin{equation}
A_{f, g} = \begin{pmatrix} \langle Af, f \rangle & \langle Bg, f \rangle \\ \langle Cf, g \rangle & \langle Dg, g \rangle \end{pmatrix}
\end{equation}
with \( f \in H_1, g \in H_2 \) and \( \|f\| = \|g\| = 1 \). The quadratic numerical range of \( \mathcal{A} \) will be denoted by \( W_{H_1, H_2}(\mathcal{A}) \).

In order to relate the quadratic numerical range to a family of projections, one considers the set of all projections \( P \in \mathcal{P}_2 \) such that \( \text{ran}(P) \) has dimension two and is spanned by two vectors in \( H_1 \oplus H_2 \) of the form \( F_1 := f_1 \oplus 0, F_2 := 0 \oplus f_2 \) with (non-zero) \( f_1 \in H_1 \) and \( f_2 \in H_2 \). We will denote this family of projections by \( \mathcal{P}_{H_1, H_2} \).

For any such \( P \in \mathcal{P}_{H_1, H_2} \) we obtain
\begin{equation}
\mathcal{A}P F_1 := PA F_1 = \langle AF_1, F_1 \rangle F_1 + \langle AF_1, F_2 \rangle F_2,
\end{equation}
and
\begin{equation}
\mathcal{A}P F_2 := PA F_2 = \langle AF_2, F_1 \rangle F_1 + \langle AF_2, F_2 \rangle F_2.
\end{equation}

Accordingly, \( \mathcal{A}P \) can be represented by a \( 2 \times 2 \) matrix with respect to this basis as
\begin{equation}
\mathcal{A}P := \begin{pmatrix} \langle AF_1, F_1 \rangle & \langle AF_1, F_2 \rangle \\ \langle AF_2, F_1 \rangle & \langle AF_2, F_2 \rangle \end{pmatrix} \in M_{2 \times 2}(\mathbb{C}).
\end{equation}

Furthermore, a direct calculation shows that
\begin{equation}
\mathcal{A}P = \begin{pmatrix} \langle Af_1, f_1 \rangle & \langle Cf_1, f_2 \rangle \\ \langle Bf_2, f_1 \rangle & \langle Df_2, f_2 \rangle \end{pmatrix} = A_{f_1, f_2}^T.
\end{equation}

This allows us to establish the following result.

**Theorem 3.9.** Let \( H = H_1 \oplus H_2 \) be a Hilbert space and \( \mathcal{A} \in \mathcal{L}(H) \) a block operator of the form (3.9). Then
\begin{equation}
W_{\mathcal{P}_{H_1, H_2}}(\mathcal{A}) = W_{H_1, H_2}(\mathcal{A}).
\end{equation}

**Proof.** For \( \lambda \in W_{H_1, H_2}(\mathcal{A}) \) and by definition of the quadratic numerical range there exist (normalized) \( f_1 \in H_1, f_2 \in H_2 \) and a (non-zero) \( h \in \mathbb{C}^2 \) such that \( A_{f_1, f_2}h = \lambda h \). Let \( P \) denote the orthogonal projection onto \( \text{span}\{f_1 \oplus 0, 0 \oplus f_2\} \). Then \( P \in \mathcal{P}_{H_1, H_2} \) and, according to (3.14), \( \sigma(\mathcal{A}P) = \sigma(A_{f_1, f_2}^T) \) which implies \( \lambda \in W_{\mathcal{P}_{H_1, H_2}}(\mathcal{A}) \).
Now, let $\lambda \in W_{PH_1H_2}(A)$ be given. Then there exists a projection $P \in \mathcal{P}_{H_1H_2}$ and a (non-zero) element $h \in \text{ran}(P)$ such that $A_P h = \lambda h$. Employing relation (3.14) again yields $\lambda \in W_{H_1H_2}(A)$. Note that the existence of corresponding (normalized) vectors $f_1 \in H_1$, $f_2 \in H_2$ follows from the definition of the family $\mathcal{P}_{H_1H_2}$.

Theorem 3.9 can be directly generalized to $k$-block operators $A$ acting on a Hilbert space of the form $H = \bigoplus_{i=1}^{n} H_i$ by defining the family $\mathcal{P}_{H_1,...,H_k}$ of projections in analogy to the case of $k = 2$. Furthermore, in analogy to the quadratic numerical range one introduces the block numerical range $W_{H_1,...,H_k}(A)$ (see also [21, 20]) and can obtain the following result.

**Theorem 3.10.** Let $H = \bigoplus_{i=1}^{n} H_i$ be a Hilbert space and $A \in \mathcal{L}(H)$ a block operator on $H$, i.e., $(A)_{1 \leq i,j \leq k} = A_{ij}$ with $A_{ij} : H_j \to H_i$ bounded linear operators. Then

$$W_{PH_1,...,H_k}(A) = W_{H_1,...,H_k}(A).$$

**Remark 3.11.** Regarding Theorem 3.9 and Theorem 3.10 we observe the following: If $A$ is a $n \times n$-matrix acting on $\mathbb{C}^n$, we can divide it into blocks as to obtain a $k$-block operator acting on $\mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_k}$ with $n_1 + \cdots + n_k = n$. Also, dividing each $\mathbb{C}^{n_i}$ further and hence obtaining a refined partition of $\mathbb{C}^n$ yields a $p$-block operator acting on $\mathbb{C}^{n_1} \oplus \cdots \oplus \mathbb{C}^{n_p}$ with $n_1 + \cdots + n_p = n$, $p > k$. Denoting all three operators by $A$, the definition of the families $\mathcal{P}_{H_1,...,H_k}$ from above allows us to obtain the inclusion

$$W_{\mathcal{P}_{C^{n_1},...,C^{n_k}}}(A) \subseteq W_{\mathcal{P}_{C^{n_1},...,C^{n_p}}}(A) \subseteq W_{\mathcal{P}_{C^{n}}}(A) = W(A),$$

where the last equality is due to Proposition 3.1. Regarding the block numerical ranges we therefore obtain the inclusion

$$W_{C^{n_1},...,C^{n_k}}(A) \subseteq W_{C^{n_1},...,C^{n_p}}(A) \subseteq W(A).$$

It is interesting to note that equation (3.18) was already obtained in [21] by different methods.

### 3.3. Connection to the product numerical range

In quantum physics and quantum information theory in particular, another notion of numerical range has proven interesting, namely, the so-called product numerical range; see [14, 6] and references therein.

To introduce this, assume that the underlying Hilbert space $H$ is given as a tensor product of two (separable) Hilbert spaces $H_k$ and $H_l$, i.e.,

$$H = H_k \otimes H_l.$$

Note that, in quantum information theory, one is often interested in the case where $H$ is finite-dimensional of composite dimension $n = kl$ where $\dim H_k = k$ and $\dim H_l = l$.

In any case, for $A \in \mathcal{L}(H_k \otimes H_l)$, the product numerical range is defined as

$$\Lambda^\otimes(A) := \{(f_k \otimes f_l, A(f_k \otimes f_l)) : f_k \in H_k \text{ and } f_l \in H_l\}$$

with $\|f_k\|_{H_k} = 1$ as well as $\|f_l\|_{H_l} = 1$. In order to identify (3.19) as a numerical range with respect to a family of projections we introduce

$$\tilde{\mathcal{P}} := \{P \in \mathcal{P} : \exists f_k \in H_k, f_l \in H_l \text{ such that } P = f_k \otimes f_l \cdot \langle f_k \otimes f_l, \cdot \rangle \},$$

again with $\|f_k\|_{H_k} = 1$ as well as $\|f_l\|_{H_l} = 1$. In other words, for any $P \in \tilde{\mathcal{P}}$ one has

$$Ph = f_k \otimes f_l \cdot \langle f_k \otimes f_l, h \rangle$$

for some normalized elements $f_k \in H_k$ and $f_l \in H_l$ and all $h \in H$. Furthermore, we can establish the following statement.

**Theorem 3.12.** For $A \in \mathcal{L}(H_k \otimes H_l)$ we have $W_{\tilde{\mathcal{P}}}(A) = \Lambda^\otimes(A)$. 
Proof. In a first step, assume that \( \lambda \in \Lambda^\otimes(A) \): then there exist normalized \( f_k \in H_k \) and \( f_l \in H_l \) with

\[
\lambda = \langle (f_k \otimes f_l), A(f_k \otimes f_l) \rangle.
\]

Then, define \( P \) to be projection onto the one-dimensional subspace spanned by \( f_k \otimes f_l \), i.e., \( P = f_k \otimes f_l : (f_k \otimes f_l, \cdot) \). Clearly, \( P \in \mathcal{P} \). In addition,

\[
(PA)(f_k \otimes f_l) = \lambda(f_k \otimes f_l)
\]

and hence \( \lambda \in W^\otimes_{\mathcal{P}}(A) \).

In a next step, assume that \( \lambda \in W^\otimes_{\mathcal{P}}(A) \): again there exist normalized vectors \( g_k \in H_k \) and \( g_l \in H_l \) such that

\[
(PA)(g_k \otimes g_l) = \lambda(g_k \otimes g_l)
\]

with \( P = g_k \otimes g_l : (g_k \otimes g_l, \cdot) \). However, the structure of \( P \) also immediately implies that \( \lambda = \langle (g_k \otimes g_l), A(g_k \otimes g_l) \rangle \) which concludes the statement. \( \square \)

Acknowledgments. The authors would like to thank R. Nagel for helpful comments and suggestions. We also want to thank the referee for various suggestions which helped to improve the manuscript.

References

Received 08/05/2018; Revised 06/08/2018