A definition of numerical range of rectangular matrices*

Ch. Chorianopoulos, S. Karanasios and P. Psarrakos†
Department of Mathematics, National Technical University of Athens
Zografou Campus, 15780 Athens, Greece

September 8, 2008

Abstract

Bonsall and Duncan (1973) observed that the numerical range of a bounded linear operator can be written as an infinite intersection of closed circular discs. Motivated by this interesting property (which does not seem to be very popular to people working on numerical ranges), we propose a definition of numerical range of rectangular complex matrices. The new range is always compact and convex, and satisfies basic properties of the standard numerical range. Our analysis is based on the properties of norms and the Birkhoff-James orthogonality.

Keywords: numerical range; norm; Birkhoff-James orthogonality
AMS Subject Classifications: 15A60; 47A12

1 Introduction

Let $(\mathcal{X}, \|\cdot\|)$ be a complex Banach space, $(\mathcal{X}^*, \|\cdot\|)$ be its dual space, and $\mathcal{B}(\mathcal{X})$ be the algebra of all bounded linear operators acting on $\mathcal{X}$. Define the set of *normalized states* $\Omega = \{\omega \in \mathcal{B}(\mathcal{X})^* : \omega(I) = \|\omega\| = 1\}$, where I denotes the identity operator. For any operator $A \in \mathcal{B}(\mathcal{X})$, the *(algebraic) numerical range* (also known as *field of values*) of A is defined by

$$F(A) = \{\omega(A) : \omega \in \Omega\}. \tag{1}$$

In the finite-dimensional case $(\mathcal{X}, \|\cdot\|) = (\mathbb{C}^n, \|\cdot\|_2)$, where $\|\cdot\|_2$ is the spectral norm, the numerical range of a square matrix $A \in \mathbb{C}^{n\times n}$ is also written

$$F(A) = \{x^*Ax \in \mathbb{C} : x \in \mathbb{C}^n,\, x^*x = 1\}. \tag{2}$$

The suggested references on numerical ranges of operators and matrices are [5, 6, 8, 10, 12].

It is known that $F(A)$ in (1) is *compact* and *convex* (this follows readily from the properties of states [15]), and contains the *spectrum* of A [17], that is, $\sigma(A) =$

*Research supported by a grant of the EPEAEK project PYTHAGORAS II. The project is co-funded by the European Social Fund (75%) and Greek National Resources (25%).
†Corresponding author. E-mail: `ppsarr@math.ntua.gr`.

$\{\mu \in \mathbb{C} : A - \mu I$ is not invertible in $\mathcal{X}\}$. Moreover, $F(A)$ coincides with the set [6, Lemma 6.22.1]

$$\begin{aligned} W_{\|\cdot\|}(A) &= \{\mu \in \mathbb{C} : \|A - \lambda I\| \geq |\mu - \lambda|, \forall\, \lambda \in \mathbb{C}\} \\ &= \bigcap_{\lambda \in \mathbb{C}} \{\mu \in \mathbb{C} : |\mu - \lambda| \leq \|A - \lambda I\|\}\,, \end{aligned} \tag{3}$$

i.e., it is written as an intersection of closed (circular) discs $\mathcal{D}(\lambda, \|A - \lambda I\|) = \{\mu \in \mathbb{C} : |\mu - \lambda| \leq \|A - \lambda I\|\}$ $(\lambda \in \mathbb{C})$. In this way, it is confirmed once again that $F(A)$ is a compact and convex subset of the complex plane that lies in the closed disc $\mathcal{D}(0, \|A\|) = \{\mu \in \mathbb{C} : |\mu| \leq \|A\|\}$. The set $W_{\|\cdot\|}(A)$ is also known as the *polynomial numerical hull of first degree* of A [7, 16].

Comparing (2) with (3), we observe that the definition of $W_{\|\cdot\|}(A)$ in (3) is based on the norm instead of the inner product. Thus, it is natural to use a formula analogous to (3) to propose a definition of the numerical range of rectangular matrices. In particular, for any $A, B \in \mathbb{C}^{n\times m}$ and any matrix norm $\|\cdot\|$, we define the *numerical range of A with respect to B* as the compact and convex set

$$\begin{aligned} W_{\|\cdot\|}(A; B) &= \{\mu \in \mathbb{C} : \|A - \lambda B\| \geq |\mu - \lambda|, \forall\, \lambda \in \mathbb{C}\} \\ &= \bigcap_{\lambda \in \mathbb{C}} \mathcal{D}(\lambda, \|A - \lambda B\|)\,. \end{aligned} \tag{4}$$

Apparently, this definition is also applicable to vectors.

In Sections 2 and 3, we comment on the new definition and obtain some basic properties of $W_{\|\cdot\|}(A; B)$, extending known properties of the standard numerical range. Furthermore, we construct explicitly $W_{\|\cdot\|}(A; B)$ when the matrix norm $\|\cdot\|$ is induced by an inner product of matrices. It is remarkable that most results of these two sections are also valid for bounded linear operators $A, B \in \mathcal{B}(\mathcal{X}, \mathcal{Y})$ from a Banach space $\mathcal{X}$ to a Banach space $\mathcal{Y}$, since their proofs are based on elementary properties of norms. The only exception is Corollary 7.

In Section 4, we derive an intersection result that explains the use of B instead of the matrix

$$I_{n,m} = \begin{cases} I_n, & n = m \\ [\,I_n \;\; 0\,], & n < m \\ \begin{bmatrix} I_m \\ 0 \end{bmatrix}, & n > m \end{cases}\,,$$

where I_n denotes the $n \times n$ identity matrix. In Section 5, we consider the eigenvalues of $A \in \mathbb{C}^{n\times m}$ with respect to $B \in \mathbb{C}^{n\times m}$ and the case of diagonal matrices. Moreover, in Section 6, for square matrices $A, B \in \mathbb{C}^{n\times n}$, we investigate the relation between $W_{\|\cdot\|_2}(A; B)$ and the numerical range of the *linear pencil* $A - \lambda B$. Several simple examples are also given to illustrate our results.

For our discussion, it is necessary to recall orthogonality of operators in the Birkhoff-James sense. For two elements u and v of a (complex) normed linear space, u is said to be *Birkhoff-James orthogonal* to v, denoted by $u \perp v$, if $\|u + \lambda v\| \geq \|u\|$ for all $\lambda \in \mathbb{C}$. This orthogonality is neither symmetric nor additive [13]. However, it is homogeneous, i.e., $u \perp v$ if and only if $au \perp bv$ for any nonzero $a, b \in \mathbb{C}$.

2 Numerical range of rectangular matrices

Let $A, B \in \mathbb{C}^{n\times m}$ and $\|\cdot\|$ be a matrix norm. Consider the numerical range of A with respect to B defined by (4), $W_{\|\cdot\|}(A;B)$, and recall that it is a compact and convex subset of the complex plane that lies in the closed disc $\mathcal{D}(0,\|A\|)$. Clearly, for $n=m$, $B=I_n$ and $\|\cdot\|=\|\cdot\|_2$, the set $W_{\|\cdot\|_2}(A;I_n)=W_{\|\cdot\|_2}(A)$ coincides with the classical numerical range $F(A)$ in (2) (see also Figure 1 below).

Note that since $W_{\|\cdot\|}(A;B)=\bigcap_{\lambda\in\mathbb{C}}\mathcal{D}\left(\lambda,\|A-\lambda B\|\right)$, one can estimate $W_{\|\cdot\|}(A;B)$ by drawing a sufficiently large number of circles of the form $\partial\mathcal{D}\left(\lambda,\|A-\lambda B\|\right)=\{\mu\in\mathbb{C}:|\mu-\lambda|=\|A-\lambda B\|\}$, as illustrated in our examples herein (see also [1]). To confirm the effectiveness of this procedure, consider the matrix $A=\begin{bmatrix}1&2&0\\0&2&0\\0&0&\mathrm{i}\end{bmatrix}$. The boundary of the numerical range $F(A)$ is sketched in the left part of Figure 1.

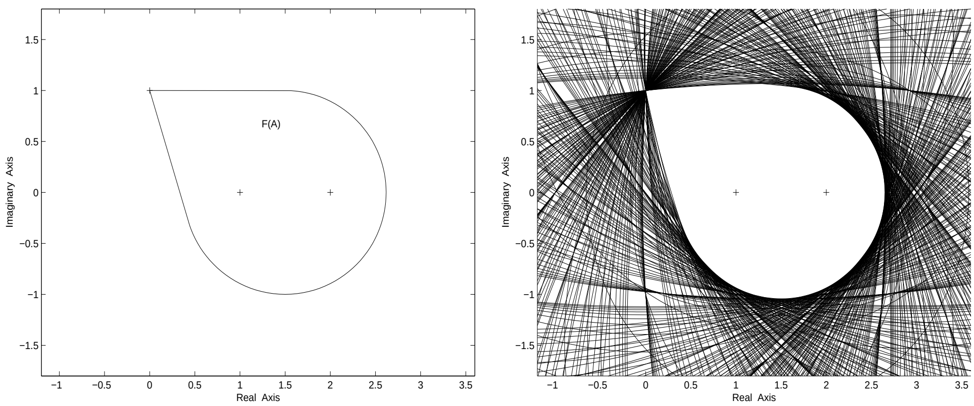

Figure 1: The numerical range $F(A)$ (left) coincides with $W_{\|\cdot\|_2}(A;I_3)$ (right).

In the right part of this figure, 400 circles of the form $\{\mu\in\mathbb{C}:|\mu-\lambda|=\|A-\lambda I_3\|_2\}$ ($\lambda\in\mathbb{C}$) containing $W_{\|\cdot\|_2}(A;I_3)$, are drawn. In both parts of the figure, the eigenvalues of A are marked with $+$'s. The unshaded region in the right part is a satisfactory estimation of $F(A)=W_{\|\cdot\|_2}(A;I_3)$ that contains $F(A)$.

Assume that $\|B\|=\beta<1$ and $\mu\in W_{\|\cdot\|}(A;B)$. Then $|\mu|\le\|A\|$, and for every $\lambda\in\mathbb{C}$, it holds that

$$||\mu|-|\lambda||\;\le\;|\mu-\lambda|\;\le\;|\mu|+|\lambda|\quad\text{and}\quad\|A-\lambda B\|\;\le\;\|A\|+|\lambda|\beta.$$

Hence, for $|\lambda|>|\mu|$, it follows $|\lambda|-|\mu|\le|\mu-\lambda|\le\|A-\lambda B\|\le\|A\|+|\lambda|\beta$, and thus, $(1-\beta)|\lambda|\le 2\|A\|$. For $|\lambda|>2\|A\|/(1-\beta)$, the latter inequality is not true, i.e., we have a contradiction. Consequently, for every $B\in\mathbb{C}^{n\times m}$ with $\|B\|<1$, $W_{\|\cdot\|}(A;B)=\emptyset$. Thus, from this point and in the remainder of the paper, we assume that $\|B\|\ge 1$. In this case, $W_{\|\cdot\|}(A;B)$ is always nonempty, as it is noticed in the sequel (see Corollary 4).

As expected from the discussion in [2], the numerical range $W_{\|\cdot\|}(A;B)$ can be expressed in terms of the Birkhoff-James orthogonality.

Theorem 1. *For any* $A, B \in \mathbb{C}^{n\times m}$ *with* $\|B\| = 1$, *it holds that*

$$W_{\|\cdot\|}(A;B) \;=\; \{\mu \in \mathbb{C} : B \perp (A-\mu B)\}\,.$$

Proof. We have that $B \perp A$ if and only if

$$\|B-\lambda A\| \;\geq\; \|B\| \;=\; 1, \quad \forall\, \lambda \in \mathbb{C},$$

or equivalently, if and only if

$$\|A-\lambda^{-1}B\| \;\geq\; |\lambda|^{-1}, \quad \forall\, \lambda \in \mathbb{C}\backslash\{0\},$$

or equivalently, if and only if $0 \in W_{\|\cdot\|}(A;B)$. Moreover, we see that

$$\begin{aligned}
W_{\|\cdot\|}(A;B) &= \{\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\} \\
&= \{\mu \in \mathbb{C} : \|A-\mu B + (\mu-\lambda)B\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\} \\
&= \left\{\mu \in \mathbb{C} : \left\|B + \frac{1}{\mu-\lambda}(A-\mu B)\right\| \geq 1,\, \forall\, \lambda \in \mathbb{C}\backslash\{\mu\}\right\} \\
&= \{\mu \in \mathbb{C} : \|B + \lambda(A-\mu B)\| \geq \|B\|,\, \forall\, \lambda \in \mathbb{C}\} \\
&= \{\mu \in \mathbb{C} : B \perp (A-\mu B)\}\,,
\end{aligned}$$

and the proof is complete. □

If we replace the matrix B by bB for some nonzero $b \in \mathbb{C}$, then

$$\begin{aligned}
W_{\|\cdot\|}(A;bB) &= \{\mu \in \mathbb{C} : \|A-\lambda(bB)\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\} \\
&= \left\{\mu \in \mathbb{C} : \|A-(b\lambda)B\| \geq |b|^{-1}\,|b\mu-b\lambda|,\, \forall\, \lambda \in \mathbb{C}\right\} \\
&= \left\{\mu \in \mathbb{C} : \|A-\lambda B\| \geq |b|^{-1}|b\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\right\}.
\end{aligned}$$

We consider three cases:

(a) If $|b| = 1$, then

$$W_{\|\cdot\|}(A;bB) \;=\; \left\{b^{-1}\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\right\} \;=\; b^{-1}W_{\|\cdot\|}(A;B).$$

(b) If $|b| < 1$, then

$$W_{\|\cdot\|}(A;bB) \;\subseteq\; \left\{b^{-1}\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\right\} \;=\; b^{-1}W_{\|\cdot\|}(A;B).$$

(c) If $|b| > 1$, then

$$W_{\|\cdot\|}(A;bB) \;\supseteq\; \left\{b^{-1}\mu \in \mathbb{C} : \|A-\lambda B\| \geq |\mu-\lambda|,\, \forall\, \lambda \in \mathbb{C}\right\} \;=\; b^{-1}W_{\|\cdot\|}(A;B).$$

Since the Birkhoff-James orthogonality is homogeneous, Theorem 1 and the observation (b) above yield directly the following corollary.

Corollary 2. *For any* $A, B \in \mathbb{C}^{n\times m}$ *with* $\|B\| > 1$,

$$\{\mu \in \mathbb{C} : B \perp (A-\mu B)\} \;=\; \|B\|^{-1}W_{\|\cdot\|}(A;\|B\|^{-1}B) \;\subseteq\; W_{\|\cdot\|}(A;B).$$

The next lemma was proved by James [13] for real normed linear spaces. Following the arguments of his proof, one can easily verify that this result is also valid for complex normed linear spaces.

Lemma 3. *If u and v are any two elements of a complex normed linear space, then there exists a scalar $\mu \in \mathbb{C}$ such that $u \perp (v + \mu u)$.*

By Theorem 1, Corollary 2, Lemma 3 and the relative discussion at the beginning of the section, it is apparent that the proposed definition of numerical range of rectangular matrices is nontrivial.

Corollary 4. *For any $A, B \in \mathbb{C}^{n\times m}$, $W_{\|\cdot\|}(A;B)$ is nonempty if and only if $\|B\| \geq 1$.*

3 Basic properties

We present some basic properties of the numerical range $W_{\|\cdot\|}(A;B)$, which are direct consequences of the properties of norms.

Proposition 5. *If $A = bB$ (with $\|B\| \geq 1$) for some $b \in \mathbb{C}$, then $W_{\|\cdot\|}(bB;B) = \{b\}$.*

Proof. For $A = bB$ (with $\|B\| \geq 1$ and $b \in \mathbb{C}$), we have

$$\begin{aligned} W_{\|\cdot\|}(bB;B) &= \{\mu \in \mathbb{C} : \|bB - \lambda B\| \geq |\mu - \lambda|, \forall\, \lambda \in \mathbb{C}\} \\ &= \{\mu \in \mathbb{C} : \|B\|\,|b - \lambda| \geq |\mu - \lambda|, \forall\, \lambda \in \mathbb{C}\}\,. \end{aligned}$$

Since $\|B\| \geq 1$, $\|B\|\,|b-\lambda| \geq |b-\lambda|$ for every $\lambda \in \mathbb{C}$, and hence, $b \in W_{\|\cdot\|}(bB;B)$. Furthermore, if we assume that $\hat{b} \in W_{\|\cdot\|}(bB;B)\backslash\{b\}$, then in the interior of the line segment with endpoints b and $\hat{b}$, there exists a λ_0 such that $\|B\|\,|b-\lambda_0| < |\hat{b} - \lambda_0|$, i.e., $\hat{b} \notin W_{\|\cdot\|}(bB;B)$. This is a contradiction, and thus, $W_{\|\cdot\|}(bB;B) = \{b\}$. $\square$

Proposition 6. *Consider a linear map $f : (\mathbb{C}^{n_1\times m_1}, \|\cdot\|) \to (\mathbb{C}^{n_2\times m_2}, |||\cdot|||)$ such that $|||f(M)||| =$ (resp., $\geq, \leq$) $\|M\|$ for every $M \in \mathbb{C}^{n_1\times m_1}$. Then for any $A, B \in \mathbb{C}^{n_1\times m_1}$, it holds that $W_{|||\cdot|||}(f(A);f(B)) =$ (resp., $\supseteq, \subseteq$) $W_{\|\cdot\|}(A;B)$.*

Proof. Suppose that $\|M\| \leq |||f(M)|||$ for every $M \in \mathbb{C}^{n_1\times m_1}$, and let $\mu \in W_{\|\cdot\|}(A;B)$. Then for every $\lambda \in \mathbb{C}$,

$$|\mu - \lambda| \,\leq\, \|A - \lambda B\| \,\leq\, |||f(A-\lambda B)||| \,=\, |||f(A) - \lambda f(B))|||,$$

and hence, $\mu \in W_{|||\cdot|||}(f(A);f(B))$. The remaining two cases can be treated similarly. $\square$

This proposition and properties of norms [19] yield immediately (i)–(iii) of the next corollary. Parts (iv)–(vi) also follow readily.

Corollary 7. *For any $A, B \in \mathbb{C}^{n\times m}$, the following hold:*

(i) *If the norm $\|\cdot\|$ is unitarily invariant, then for any unitary matrices $U \in \mathbb{C}^{n\times n}$ and $V \in \mathbb{C}^{m\times m}$, $W_{\|\cdot\|}(UAV;UBV) = W_{\|\cdot\|}(A;B)$.*

(ii) *If the norm $\|\cdot\|$ is unitarily invariant, and $\hat{A}$ and $\hat{B}$ are submatrices of A and B, respectively, formed by the same rows and the same columns, then $W_{\|\cdot\|}(\hat{A};\hat{B}) \subseteq W_{\|\cdot\|}(A;B)$.*

(iii) *If the norm $\|\cdot\|$ is invariant under the transpose operation $\cdot^T$, then $W_{\|\cdot\|}(A^T;B^T) = W_{\|\cdot\|}(A;B)$.*

(iv) *If the norm $\|\cdot\|$ is invariant under the conjugate operation $\overline{\cdot}$, then $W_{\|\cdot\|}(\overline{A};\overline{B}) = \overline{W_{\|\cdot\|}(A;B)}$.*

(v) *If the norm $\|\cdot\|$ is invariant under the conjugate transpose operation $\cdot^*$, then $W_{\|\cdot\|}(A^*;B^*) = \overline{W_{\|\cdot\|}(A;B)}$.*

(vi) *If the norm $\|\cdot\|$ is invariant under the conjugate transpose operation $\cdot^*$ (resp., the conjugate operation $\overline{\cdot}$) and the matrices A,B are $n \times n$ hermitian (resp., $n \times m$ real), then $W_{\|\cdot\|}(A;B)$ is symmetric with respect to the real axis.*

Proposition 8. *For any scalars $a,b \in \mathbb{C}$, $W_{\|\cdot\|}(aA+bB;B) = aW_{\|\cdot\|}(A;B)+b$.*

Proof. Suppose that $a \neq 0$. Then a complex number $a\mu+b$ belongs to the numerical range $W_{\|\cdot\|}(aA+bB;B)$ if and only if

$$\|aA+bB-\lambda B\| \geq |a\mu+b-\lambda|, \quad \forall\, \lambda \in \mathbb{C},$$

or equivalently, if and only if

$$|a| \left\| A + \frac{b-\lambda}{a} B \right\| \geq |a| \left| \mu + \frac{b-\lambda}{a} \right|, \quad \forall\, \lambda \in \mathbb{C},$$

or equivalently, if and only if

$$\|A-\lambda B\| \geq |\mu-\lambda|, \quad \forall\, \lambda \in \mathbb{C}.$$

Hence, $a\mu+b \in W_{\|\cdot\|}(aA+bB;B)$ if and only if $\mu \in W_{\|\cdot\|}(A;B)$.

If $a=0$, then $W_{\|\cdot\|}(aA+bB;B) = W_{\|\cdot\|}(bB;B) = \{b\}$, keeping in mind that $\|B\| \geq 1$. □

For example, consider the 3×4 matrices $A = \begin{bmatrix} 5+\mathrm{i} & 0.2 & 0 & -0.1 \\ 0 & 1-\mathrm{i}\,5 & -\mathrm{i}\,0.1 & 0 \\ 0 & 0 & 0.1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1.1 & 0 & 0 & 0 \\ 0 & 1.2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, and let $\|\cdot\| = \|\cdot\|_1$. The ranges $W_{\|\cdot\|_1}(A;B)$ and $W_{\|\cdot\|_1}(\mathrm{i}A-4B;B)$ are illustrated in the left and right parts of Figure 2, respectively, confirming Proposition 8.

The interior of $W_{\|\cdot\|}(A;B)$, $\mathrm{Int}\left[W_{\|\cdot\|}(A;B)\right]$, can be characterized by using strict inequality in the definition (4).

Proposition 9. *For any $A,B \in \mathbb{C}^{n\times m}$,*

$$\mathrm{Int}\left[W_{\|\cdot\|}(A;B)\right] \subseteq \{\mu \in \mathbb{C} : \|A-\lambda B\| > |\mu-\lambda|, \forall\, \lambda \in \mathbb{C}\}.$$

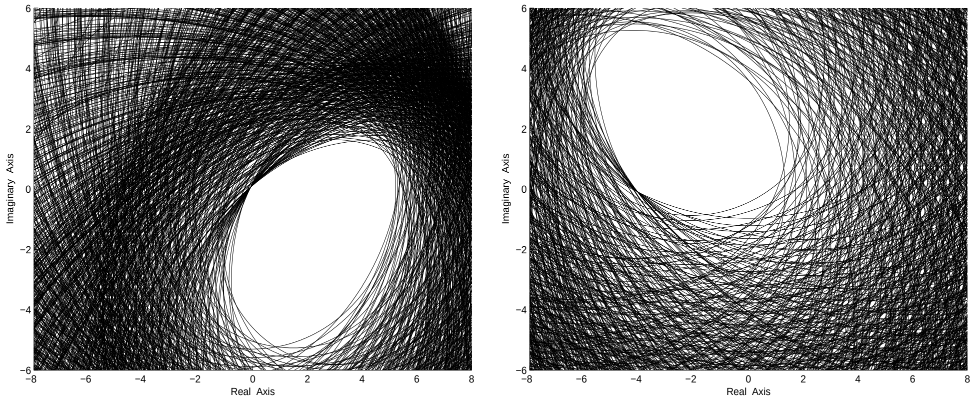

Figure 2: The numerical ranges $W_{\|\cdot\|_1}(A;B)$ (left) and $W_{\|\cdot\|_1}(\mathrm{i}A-4B;B)$ (right).

Proof. Let $\mu \in \mathrm{Int}\left[W_{\|\cdot\|}(A;B)\right]$ (recalling that $\|B\| \geq 1$). Then there is a $\rho > 0$ such that $\mu + e^{\mathrm{i}\theta}\rho \in W_{\|\cdot\|}(A;B)$ for all $\theta \in [0, 2\pi]$. Hence, it follows

$$\|A - \lambda B\| \geq \left|\left(\mu + e^{\mathrm{i}\theta}\rho\right) - \lambda\right|, \quad \forall\, \lambda \in \mathbb{C},\, \theta \in [0, 2\pi].$$

But apparently, for any $\lambda \in \mathbb{C}$, there exists an angle $\theta(\lambda) \in [0, 2\pi]$ such that $\left|\left(\mu + e^{\mathrm{i}\theta(\lambda)}\rho\right) - \lambda\right| > |\mu - \lambda|$. As a consequence, $\|A - \lambda B\| > |\mu - \lambda|$ for all $\lambda \in \mathbb{C}$. $\square$

Next we see that the relation between $W_{\|\cdot\|}(A;B)$ and $W_{\|\cdot\|}(B;A)$ is strong.

Proposition 10. *For any* $A, B \in \mathbb{C}^{n\times m}$, *it holds that*

$$\left\{\mu^{-1} \in \mathbb{C} : \mu \in W_{\|\cdot\|}(A;B),\, |\mu| \geq 1\right\} \subseteq W_{\|\cdot\|}(B;A).$$

Proof. Let $\mu \in W_{\|\cdot\|}(A;B)$ with $|\mu| \geq 1$. Then

$$\|A - \lambda B\| \geq |\mu - \lambda|, \quad \forall \lambda \in \mathbb{C},$$

or

$$|\lambda| \left\|\frac{1}{\lambda}A - B\right\| \geq |\lambda| \left|\frac{\mu}{\lambda} - 1\right|, \quad \forall\, \lambda \in \mathbb{C}\backslash\{0\},$$

or

$$\|B - \lambda A\| \geq |\mu\,\lambda - 1| = |\mu|\,|\mu^{-1} - \lambda|, \quad \forall\, \lambda \in \mathbb{C}.$$

Since $|\mu| \geq 1$, it follows that $\mu^{-1} \in W_{\|\cdot\|}(B;A)$. $\square$

Consider the 3×3 diagonal matrix $A = diag\{2+\mathrm{i}\,3, 2, 4\}$. The numerical range $F(A) = W_{\|\cdot\|_2}(A;I_3)$ coincides with the convex hull of the diagonal entries of A and lies outside the unit disc $\mathcal{D}(0,1)$, as one can see in the left part of Figure 3. The range $W_{\|\cdot\|_2}(I_3;A)$, estimated by the unshaded region in the right part of the figure, contains the set $\left\{\mu^{-1} \in \mathbb{C} : \mu \in F(A)\right\}$, verifying Proposition 10. By this example, it is also clear that, in general, $W_{\|\cdot\|}(B;A) \neq \left\{\mu^{-1} \in \mathbb{C} : \mu \in W_{\|\cdot\|}(A;B),\, |\mu| \geq 1\right\}$ (note

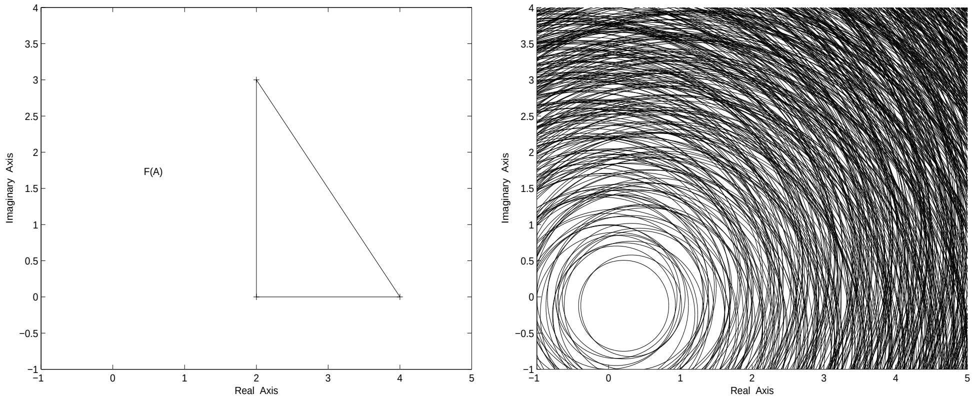

Figure 3: The ranges $F(A) = W_{\|\cdot\|_2}(A; I_3)$ (left) and $W_{\|\cdot\|_2}(I_3; A)$ (right).

that in the right part of Figure 3, the origin lies in the interior of $W_{\|\cdot\|_2}(I_3; A)$). Of course, for any $B \in \mathbb{C}^{n\times m}$ and $b \in \mathbb{C}$ such that $\|B\| \geq 1$ and $|b| = 1$, we have

$$W_{\|\cdot\|_2}(bB; B) \;=\; \{b\} \quad \text{and} \quad W_{\|\cdot\|_2}(B; bB) \;=\; W_{\|\cdot\|_2}(b^{-1}(bB); bB) \;=\; \{b^{-1}\}.$$

Now we turn our attention to matrix norms induced by inner products of matrices; this is the case for the Frobenius norm.

Lemma 11. *Suppose that the matrix norm $\|\cdot\|$ is induced by the inner product $(\cdot,\cdot)$, and let $A, B \in \mathbb{C}^{n\times m}$. Then $A \perp B$ if and only if $(A, B) = 0$.*

Proof. We have that $A \perp B$ if and only if

$$\|A + \lambda B\| \;\geq\; \|A\|, \quad \forall\, \lambda \in \mathbb{C},$$

or equivalently, if and only if

$$|\lambda|^2 \|B\|^2 \;\geq\; -2\,\mathrm{Re}\left[\overline{\lambda}\,(A, B)\right], \quad \forall\, \lambda \in \mathbb{C}.$$

Letting λ be real implies

$$|\lambda|^2 \|B\|^2 \;\geq\; -2\,\mathrm{Re}\left[\overline{\lambda}\,(A, B)\right] \;=\; -2\,\lambda\,\mathrm{Re}[(A, B)], \quad \forall\, \lambda \in \mathbb{R}$$

and

$$|\lambda|^2 \|B\|^2 \;=\; |\mathrm{i}\lambda|^2 \|B\|^2 \;\geq\; -2\,\mathrm{Re}\left[\overline{\mathrm{i}\lambda}\,(A, B)\right] \;=\; -2\,\lambda\,\mathrm{Im}[(A, B)], \quad \forall\, \lambda \in \mathbb{R}.$$

As a consequence, $\mathrm{Re}[(A, B)] = \mathrm{Im}[(A, B)] = 0$, and the proof is complete. ☐

Lemma 12. *Suppose that the matrix norm $\|\cdot\|$ is induced by the inner product $(\cdot,\cdot)$, and let $A, B \in \mathbb{C}^{n\times m}$ with $\|B\| \geq 1$. Then $(A, B)/\|B\|^2 \in W_{\|\cdot\|}(A; B)$. Moreover, if $\|B\| = 1$, then $W_{\|\cdot\|}(A; B) = \{(A, B)\}$.*

Proof. By Lemma 11, a scalar $\mu \in \mathbb{C}$ satisfies the orthogonality $B \perp (A - \mu B)$ if and only if $(B, A - \mu B) = 0$, or equivalently, if and only if $\mu = (A, B)/\|B\|^2$. By Theorem 1 and Corollary 2, the proof is completed. ☐

Let $\|\cdot\|$ be a matrix norm induced by the inner product $(\cdot,\cdot)$, and let B be an $n\times m$ matrix with $\|B\|=1$. Then for any $A\in\mathbb{C}^{n\times m}$, the numerical range $W_{\|\cdot\|}(A;B)$ is a singleton, although A is not necessarily a scalar multiple of B. This means that the converse of Proposition 5 is not true in general.

Furthermore, for $\|B\|>1$ and A not a scalar multiple of B, the numerical range $W_{\|\cdot\|}(A;B)$ is a (nontrivial) closed disc centered at $(A,B)/\|B\|^2$, and hence, it has a nonempty interior.

Proposition 13. *Suppose that the matrix norm $\|\cdot\|$ is induced by the inner product $(\cdot,\cdot)$, and let $A,B\in\mathbb{C}^{n\times m}$ with $\|B\|\geq 1$. Then it holds that*

$$W_{\|\cdot\|}(A;B) \;=\; \mathcal{D}\left(\frac{(A,B)}{\|B\|^2}\,,\;\left\|A-\frac{(A,B)}{\|B\|^2}\,B\right\|\,\frac{\sqrt{\|B\|^2-1}}{\|B\|}\right).$$

Proof. If $\|B\|=1$, then by Lemma 12, $W_{\|\cdot\|}(A;B)=\{(A,B)\}=\mathcal{D}((A,B),0)$.

Suppose that $\|B\|>1$. Denote $\hat{A}=A-\frac{(A,B)}{\|B\|^2}\,B$, and observe that

$$W_{\|\cdot\|}(\hat{A};B) \;=\; W_{\|\cdot\|}(A;B)-\frac{(A,B)}{\|B\|^2} \quad\text{and}\quad (\hat{A},B)\;=\;0.$$

Then for any $r>0$,

$$\|\hat{A}-re^{\mathrm{i}\theta}B\| \;=\; \sqrt{\|\hat{A}\|^2+r^2\|B\|^2}\,,\quad \forall\,\theta\in[0,2\pi],$$

and as a consequence,

$$\bigcap_{\theta\in[0,2\pi]}\mathcal{D}\left(re^{\mathrm{i}\theta},\left\|\hat{A}-re^{\mathrm{i}\theta}B\right\|\right) \;=\; \mathcal{D}\left(0\,,\sqrt{\|\hat{A}\|^2+r^2\|B\|^2}-r\right).$$

It is also easy to see that

$$\min_{r\geq 0}\left\{\sqrt{\|\hat{A}\|^2+r^2\|B\|^2}-r\right\} \;=\; \|\hat{A}\|\frac{\sqrt{\|B\|^2-1}}{\|B\|}\,,$$

where the minimum is attained at $r=\|\hat{A}\|/(\|B\|\sqrt{\|B\|^2-1})$. Thus,

$$W_{\|\cdot\|}(\hat{A};B) \;=\; \bigcap_{r\geq 0}\mathcal{D}\left(0\,,\sqrt{\|\hat{A}\|^2+r^2\|B\|^2}-r\right) \;=\; \mathcal{D}\left(0\,,\,\|\hat{A}\|\frac{\sqrt{\|B\|^2-1}}{\|B\|}\right),$$

and the proof is complete. □

Remark. It is straightforward to see that if the norm $\|\cdot\|$ is induced by the inner product $(\cdot,\cdot)$, then for any $A,\hat{A},B\in\mathbb{C}^{n\times m}$ with $\|B\|\geq 1$,

$$\begin{aligned}
W_{\|\cdot\|}(A+\hat{A};B) \;&=\; \mathcal{D}\left(\frac{(A+\hat{A},B)}{\|B\|^2}\,,\;\left\|A+\hat{A}-\frac{(A+\hat{A},B)}{\|B\|^2}\,B\right\|\,\frac{\sqrt{\|B\|^2-1}}{\|B\|}\right)\\
&\subseteq\; \mathcal{D}\left(\frac{(A,B)}{\|B\|^2}\,,\;\left\|A-\frac{(A,B)}{\|B\|^2}\,B\right\|\,\frac{\sqrt{\|B\|^2-1}}{\|B\|}\right)\\
&\qquad+\mathcal{D}\left(\frac{(\hat{A},B)}{\|B\|^2}\,,\;\left\|\hat{A}-\frac{(\hat{A},B)}{\|B\|^2}\,B\right\|\,\frac{\sqrt{\|B\|^2-1}}{\|B\|}\right)\\
&=\; W_{\|\cdot\|}(A;B)+W_{\|\cdot\|}(\hat{A};B),
\end{aligned}$$

i.e., the subadditivity property holds. Furthermore, our experiments lead to the conjecture that subadditivity of numerical ranges holds for all norms. Unfortunately, such a result is not obtained and it is still an open question. In the case of an affirmative answer, the subadditivity will provide us with the continuity of the mapping $A \longrightarrow W_{\|\cdot\|}(A;B)$. In particular, for any matrix $E \in \mathbb{C}^{n\times m}$, we would have

$$W_{\|\cdot\|}(A+E;B) \subseteq W_{\|\cdot\|}(A;B) + W_{\|\cdot\|}(E;B) \subseteq W_{\|\cdot\|}(A;B) + \mathcal{D}(0,\|E\|).$$

As a consequence, for every $\mu \in W_{\|\cdot\|}(A;B)$, there would be a $\hat{\mu} \in W_{\|\cdot\|}(A+E;B)$ such that $|\mu-\hat{\mu}| \le \|E\|$, and conversely, for every $\hat{\mu} \in W_{\|\cdot\|}(A+E;B)$, there would be a $\mu \in W_{\|\cdot\|}(A;B)$ with $|\mu-\hat{\mu}| \le \|E\|$. Thus, with regard to the Hausdorff metric, the distance between the ranges $W_{\|\cdot\|}(A;B)$ and $W_{\|\cdot\|}(A+E;B)$ would be

$$\max\left\{ \max_{\mu\in W_{\|\cdot\|}(A;B)} \; \min_{\hat{\mu}\in W_{\|\cdot\|}(A+E;B)} |\mu-\hat{\mu}|, \; \max_{\hat{\mu}\in W_{\|\cdot\|}(A+E;B)} \; \min_{\mu\in W_{\|\cdot\|}(A;B)} |\mu-\hat{\mu}| \right\} \le \|E\|.$$

4 An intersection result and an explanation for B

For a square matrix $A \in \mathbb{C}^{n\times n}$, we have seen that

$$F(A) = W_{\|\cdot\|_2}(A) = W_{\|\cdot\|_2}(A;I_n) = \left\{\mu \in \mathbb{C} : \|A-\lambda I_n\|_2 \ge |\mu-\lambda|, \forall\, \lambda \in \mathbb{C}\right\}.$$

So, one may question the use of the (general) matrix B for defining the numerical range of rectangular matrices instead of the matrix $I_{n,m}$. In this section, we explain why the numerical range $W_{\|\cdot\|_2}(A;I_{n,m})$ cannot be considered as an appropriate generalization.

Without loss of generality, we assume that $n>m$, $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ with $A_1 \in \mathbb{C}^{m\times m}$ and $A_2 \in \mathbb{C}^{(n-m)\times m}$, and $I_{n,m} = \begin{bmatrix} I_m \\ 0 \end{bmatrix}$. Then

$$\begin{aligned} W_{\|\cdot\|_2}(A;I_{n,m}) &= \left\{\mu\in\mathbb{C} : \left\| \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} - \lambda \begin{bmatrix} I_m \\ 0 \end{bmatrix} \right\|_2 \ge |\mu-\lambda|, \forall\,\lambda\in\mathbb{C} \right\} \\ &= \left\{\mu\in\mathbb{C} : \left\| \begin{bmatrix} A_1 & 0 \\ A_2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} \right\|_2 \ge |\mu-\lambda|, \forall\,\lambda\in\mathbb{C} \right\}, \end{aligned}$$

where the matrices $\begin{bmatrix} A_1 & 0 \\ A_2 & 0 \end{bmatrix}$ and $\begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}$ are $n\times n$. Moreover, it is easy to verify that for every $\lambda\in\mathbb{C}$, $M_1 \in \mathbb{C}^{m\times(n-m)}$ and $M_2 \in \mathbb{C}^{(n-m)\times(n-m)}$,

$$\left\| \begin{bmatrix} A_1 & 0 \\ A_2 & 0 \end{bmatrix} - \lambda \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} \right\|_2 \le \left\| \begin{bmatrix} A_1 & M_1 \\ A_2 & M_2 \end{bmatrix} - \lambda \begin{bmatrix} I_m & 0 \\ 0 & I_{n-m} \end{bmatrix} \right\|_2$$

(see also Theorem 4.3.15 of [11]). Thus, denoting $M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$,

$$\begin{aligned} W_{\|\cdot\|_2}(A;I_{n,m}) &\subseteq \bigcap_{M\in\mathbb{C}^{n\times(n-m)}} \left\{\mu\in\mathbb{C} : \|[\,A\ \ M\,] - \lambda I_n\|_2 \ge |\mu-\lambda|, \forall\,\lambda\in\mathbb{C}\right\} \\ &= \bigcap_{M\in\mathbb{C}^{n\times(n-m)}} F\left([\,A\ \ M\,]\right), \end{aligned}$$

where the intersection is taken over all $n \times (n-m)$ matrices M. Furthermore, this intersection coincides with the numerical range $F(A_1)$.

Proposition 14. *For any $n \times m$ ($n > m$) matrix $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ with $A_1 \in \mathbb{C}^{m\times m}$ and $A_2 \in \mathbb{C}^{(n-m)\times m}$, it holds that*

$$F(A_1) = \bigcap_{M\in\mathbb{C}^{n\times(n-m)}} F\left([\, A \;\; M \,]\right).$$

Proof. For every $M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} \in \mathbb{C}^{n\times(n-m)}$, $F(A_1) \subseteq F\left([\, A \;\; M \,]\right)$, and thus,

$$F(A_1) \subseteq \bigcap_{M\in\mathbb{C}^{n\times(n-m)}} F\left([\, A \;\; M \,]\right).$$

By the convexity of the standard numerical range of square matrices, it is enough to prove that for every $\theta \in [0, 2\pi]$, there is an $n\times n$ matrix $[\, A \;\; M_\theta \,] = \begin{bmatrix} A_1 & M_1 \\ A_2 & M_2 \end{bmatrix}$ such that the numerical ranges $F\left(e^{\mathrm{i}\theta}A_1\right) = e^{\mathrm{i}\theta}F(A_1)$ and $F\left(e^{\mathrm{i}\theta}[\, A \;\; M_\theta \,]\right) = e^{\mathrm{i}\theta}F\left([\, A \;\; M_\theta \,]\right)$ have exactly the same projection on the real axis.

For any $\theta \in [0, 2\pi]$, consider the matrix

$$[\, A \;\; M_\theta \,] = \begin{bmatrix} A_1 & e^{\mathrm{i}(\pi-2\theta)}A_2^* \\ A_2 & \mu I_{n-m} \end{bmatrix}$$

for some $\mu \in F(A_1)$. Then we have

$$\begin{aligned} \frac{1}{2}\left(e^{\mathrm{i}\theta}[\, A \;\; M_\theta \,] + e^{-\mathrm{i}\theta}[\, A \;\; M_\theta \,]^*\right) &= \frac{1}{2}\begin{bmatrix} e^{\mathrm{i}\theta}A_1 + e^{-\mathrm{i}\theta}A_1^* & e^{\mathrm{i}(\pi-\theta)}A_2^* + e^{-\mathrm{i}\theta}A_2^* \\ e^{\mathrm{i}\theta}A_2 + e^{-\mathrm{i}(\pi-\theta)}A_2 & e^{\mathrm{i}\theta}\mu I_{n-m} + e^{-\mathrm{i}\theta}\overline{\mu} I_{n-m} \end{bmatrix} \\ &= \begin{bmatrix} \frac{e^{\mathrm{i}\theta}A_1 + e^{-\mathrm{i}\theta}A_1^*}{2} & 0 \\ 0 & \mathrm{Re}\{e^{\mathrm{i}\theta}\mu\} I_{n-m} \end{bmatrix}, \end{aligned}$$

where

$$\mathrm{Re}\{e^{\mathrm{i}\theta}\mu\} \in \mathrm{Re}\{F\left(e^{\mathrm{i}\theta}A_1\right)\} = F\left(\frac{e^{\mathrm{i}\theta}A_1 + e^{-\mathrm{i}\theta}A_1^*}{2}\right).$$

Hence, the numerical ranges

$$F\left(e^{\mathrm{i}\theta}A_1\right) \quad \text{and} \quad F\left(e^{\mathrm{i}\theta}\begin{bmatrix} A_1 & e^{\mathrm{i}(\pi-2\theta)}A_2^* \\ A_2 & \mu I_{n-m} \end{bmatrix}\right)$$

have exactly the same projection on the real axis. $\square$

By the above proof, the next corollary follows readily.

Corollary 15. *Let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ with $A_1 \in \mathbb{C}^{m\times m}$ and $A_2 \in \mathbb{C}^{(n-m)\times m}$. If $\mu \in F(A_1)$, then*

$$F(A_1) = \bigcap_{\theta\in[0,2\pi]} F\left(\begin{bmatrix} A_1 & e^{\mathrm{i}\theta}A_2^* \\ A_2 & \mu I_{n-m} \end{bmatrix}\right).$$

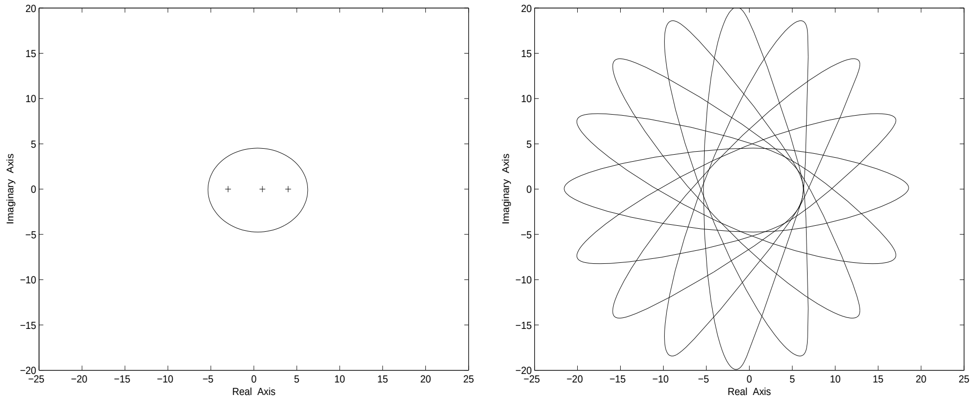

Figure 4: The numerical ranges $F(A_1)$ (left), and $F(C_k)$ for $k = 0, 1, \dots, 7$ (right).

For a stronger result than Proposition 14 (though not covering Corollary 15), see Lemma 1 of [9].

To illustrate numerically this intersection result, consider the matrix

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ \mathrm{i} & 9 & 4 \\ -14 & 13 & -5 \end{bmatrix}.$$

The numerical range of the 3×3 submatrix $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -3 & 0 \\ \mathrm{i} & 9 & 4 \end{bmatrix}$ is indicated in the left part of Figure 4, where the eigenvalues of A_1 are marked with $+$'s. In the right part of the figure, the boundaries of the numerical ranges of matrices $C_k = \begin{bmatrix} A_1 & e^{\mathrm{i}\frac{k\pi}{4}} A_2^* \\ A_2 & 0 \end{bmatrix}$ for $k = 0, 1, \dots, 7$, are sketched. Keeping in mind that $0 \in F(A_1)$, Corollary 15 is apparently confirmed.

Next we obtain that the set $W_{\|\cdot\|_2}(A; I_{n,m})$ coincides with the standard numerical range of the submatrix A_1, $F(A_1)$. This makes the use of $I_{n,m}$ having no practical value and explains the appearance of B in our generalization for rectangular matrices.

Theorem 16. *For any $n \times m$ $(n > m)$ matrix $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ with $A_1 \in \mathbb{C}^{m\times m}$ and $A_2 \in \mathbb{C}^{(n-m)\times m}$, it holds that $W_{\|\cdot\|_2}(A; I_{n,m}) = F(A_1)$.*

Proof. We have already seen that

$$W_{\|\cdot\|_2}(A; I_{n,m}) \subseteq \bigcap_{M \in \mathbb{C}^{n\times(n-m)}} F\left([\,A \;\; M\,]\right) = F(A_1),$$

and hence, it is enough to prove the inverse inclusion. Let $\mu_0 \in F(A_1)$. Since $W_{\|\cdot\|_2}(A - \mu_0 I_{n,m}; I_{n,m}) = W_{\|\cdot\|_2}(A; I_{n,m}) - \mu_0$ and $F(A_1 - \mu_0 I_m) = F(A_1) - \mu_0$, without loss of generality, we can assume that $\mu_0 = 0$. Then we have

$$0 \in F(A_1) = W_{\|\cdot\|_2}(A_1; I_m),$$

or equivalently,

$$\|I_m - \lambda A_1\|_2 \geq 1, \quad \forall \lambda \in \mathbb{C}.$$

Since $\|I_{n,m} - \lambda A\|_2 \geq \|I_m - \lambda A_1\|_2$, it follows

$$\|I_{n,m} - \lambda A\|_2 \geq 1, \quad \forall \lambda \in \mathbb{C},$$

or

$$\left\|A - \lambda^{-1} I_{n,m}\right\|_2 \geq |\lambda|^{-1}, \quad \forall \lambda \in \mathbb{C}\backslash\{0\},$$

or

$$\|A - \lambda I_{n,m}\|_2 \geq |\lambda|, \quad \forall \lambda \in \mathbb{C}.$$

As a consequence, $0 \in W_{\|\cdot\|_2}(A; I_{n,m})$, and the proof is complete. □

5 Eigenvalues and diagonal rectangular matrices

Let A, B be $n \times m$ matrices with $n \geq m$. A scalar $\mu_0 \in \mathbb{C}$ is said to be an *eigenvalue of A with respect to B* if $(A - \mu_0 B)\, x_0 = 0$ for some nonzero vector $x_0 \in \mathbb{C}^m$. The vector x_0 is called an *eigenvector of A with respect to B* corresponding to μ_0, and the set of all eigenvalues of A with respect to B is denoted by $\sigma(A; B)$.

The above definition of eigenvalues has been used in [3, 4, 18, 20], but it has not become a mainstream concept in linear algebra. The reason is that most of non-square matrices have no eigenvalues at all, and for those that do, a random perturbation will in general remove them. On the other hand, this definition can be considered as an appropriate generalization since for $n = m$ and $B = I_n$, it coincides with the standard definition of eigenvalues and eigenvectors.

In the remainder of this section, we assume that the matrix norm $\|\cdot\|$ is induced by a vector norm (acting on $\mathbb{C}^n$ and $\mathbb{C}^m$).

Proposition 17. *Let $A, B \in \mathbb{C}^{n\times m}$ $(n \geq m)$. Any eigenvalue μ_0 of A with respect to B, with an associated unit eigenvector $x_0 \in \mathbb{C}^m$ such that $\|Bx_0\| \geq 1$, lies in the numerical range $W_{\|\cdot\|}(A; B)$.*

Proof. For the unit eigenvector $x_0 \in \mathbb{C}^m$, we know that $(A - \mu_0 B)x_0 = 0$ and $\|Bx_0\| \geq 1$. Then for every $\lambda \in \mathbb{C}$, $(A - \lambda B)x_0 = (\mu_0 - \lambda)Bx_0$ and thus,

$$\|A - \lambda B\| \geq \|(A - \lambda B)x_0\| = |\mu_0 - \lambda|\, \|Bx_0\| \geq |\mu_0 - \lambda|.$$

As a consequence, $\mu_0 \in W_{\|\cdot\|}(A; B)$. □

In several experiments, we have met eigenvalues of a matrix A with respect to a matrix B that lie in $W_{\|\cdot\|}(A; B)$, although $\|Bx_0\| < 1$ for any associated unit eigenvector x_0. Hence, in Proposition 17, the condition $\|Bx_0\| \geq 1$ is sufficient but not necessary. Furthermore, for the eigenvalues on the boundary of the numerical range, we have the next result.

Proposition 18. *Let $A, B \in \mathbb{C}^{n\times m}$ $(n \geq m)$ such that $W_{\|\cdot\|}(A; B)$ is not a singleton, and let μ_0 be an eigenvalue of A with respect to B on the boundary of $W_{\|\cdot\|}(A; B)$. Then for any associated unit eigenvector $x_0 \in \mathbb{C}^m$, it holds that $\|Bx_0\| \leq 1$.*

Proof. Suppose that the eigenvalue μ_0 lies on the boundary of $W_{\|\cdot\|}(A;B)$. Then for any $\varepsilon > 0$, there is a $\lambda_\varepsilon \in \mathbb{C}$, $\lambda_\varepsilon \neq \mu_0$, such that

$$\mathcal{D}(\mu_0,\varepsilon) \not\subseteq \mathcal{D}(\lambda_\varepsilon, \|A-\lambda_\varepsilon B\|),$$

or

$$|\mu_0-\lambda_\varepsilon| + \varepsilon > \|A-\lambda_\varepsilon B\| \geq \|(A-\lambda_\varepsilon B)x_0\|.$$

Since $Ax_0 = \mu_0 Bx_0$, it follows

$$|\mu_0-\lambda_\varepsilon| + \varepsilon > \|(\mu_0-\lambda_\varepsilon)Bx_0\| = |\mu_0-\lambda_\varepsilon|\,\|Bx_0\|$$

or

$$|\mu_0-\lambda_\varepsilon|\,(\|Bx_0\|-1) \leq \varepsilon. \tag{5}$$

From the fact that $W_{\|\cdot\|}(A;B)$ is not a singleton and lies in $\mathcal{D}(\lambda_\varepsilon, \|A-\lambda_\varepsilon B\|)$, we see that λ_ε cannot be arbitrarily close to μ_0. Thus, if we assume that $\|Bx_0\| > 1$, then the inequality (5) yields a contradiction. As a consequence, $\|Bx_0\| \leq 1$. □

Suppose that the matrices $A, B \in \mathbb{C}^{n\times m}$ $(n \geq m)$ are diagonal, i.e., all their off-diagonal entries are zero, and denote $A = diag\{a_1, a_2, \ldots, a_m\}$ and $B = diag\{b_1, b_2, \ldots, b_m\}$. Assume also that all the diagonal entries of B are nonzero. Then the ratios a_1/b_1, a_2/b_2, …, a_m/b_m are the eigenvalues of A with respect to B, with corresponding eigenvectors the vectors of the standard basis of $\mathbb{C}^m$. Moreover, Proposition 17 implies the following.

Corollary 19. *Let $A = diag\{a_1, a_2, \ldots, a_m\}$ and $B = diag\{b_1, b_2, \ldots, b_m\}$ be $n\times m$ $(n \geq m)$ diagonal matrices with $b_i \neq 0$ $(i = 1, 2, \ldots, m)$. Then every eigenvalue $a_i/b_i \in \sigma(A;B)$ with $|b_i| \geq 1$ lies in $W_{\|\cdot\|}(A;B)$.*

Denote by $\|\cdot\|_d$ a matrix norm such that for any $n\times m$ $(n \geq m)$ diagonal matrix $D = diag\{d_1, d_2, \ldots, d_m\}$, $\|D\|_d = \max\{|d_j| : j = 1, 2, \ldots, m\}$; this is the case for matrix norms induced by *absolute* and *monotone* vector norms [11, Theorem 5.6.37], such as $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$.

Proposition 20. *Let $A = diag\{a_1, a_2, \ldots, a_m\}$ and $B = diag\{b_1, b_2, \ldots, b_m\}$ be two $n\times m$ $(n \geq m)$ diagonal matrices. If $|b_1| = |b_2| = \cdots = |b_m| = 1$, then the numerical range $W_{\|\cdot\|_d}(A;B)$ coincides with the convex hull of the eigenvalues of A with respect to B, $a_1/b_1, a_2/b_2, \ldots, a_m/b_m$.*

Proof. By Corollary 19 and the convexity of $W_{\|\cdot\|_d}(A;B)$, it follows that $co\,\{\sigma(A;B)\} \subseteq W_{\|\cdot\|_d}(A;B)$. Consider now a $\mu \notin co\,\{\sigma(A;B)\}$. Then there exist a $\lambda_\mu \in \mathbb{C}$ and a real $r_\mu > 0$ such that the closed disc $\mathcal{D}(\lambda_\mu, r_\mu)$ contains $co\,\{\sigma(A;B)\}$ but not μ. As a consequence,

$$\left|\lambda_\mu - \frac{a_i}{b_i}\right| \leq r_\mu < |\lambda_\mu - \mu|\,, \quad i = 1, 2, \ldots, m,$$

or

$$|a_i - \lambda_\mu b_i| < |\lambda_\mu - \mu|\,, \quad i = 1, 2, \ldots, m,$$

or

$$\|A - \lambda_\mu B\|_d < |\lambda_\mu - \mu|.$$

Hence, $\mu \notin W_{\|\cdot\|_d}(A;B)$, and the proof is complete. □

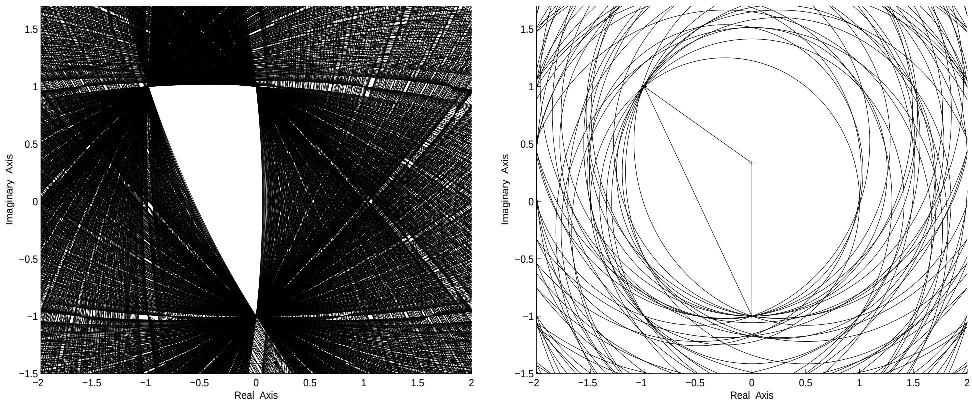

Figure 5: The numerical ranges $W_{\|\cdot\|_2}(A;B)$ (left) and $W_{\|\cdot\|_2}(A;\hat{B})$ (right).

For the 4×3 diagonal matrices $A = diag\{\mathrm{i}, 1, 1+\mathrm{i}\}$ and $B = diag\{1, \mathrm{i}, -\mathrm{i}\}$, the numerical range $W_{\|\cdot\|_2}(A;B)$ is estimated by the unshaded region in the left part of Figure 5, and it coincides with the convex hull of $\sigma(A;B) = \{\mathrm{i}, -\mathrm{i}, -1+\mathrm{i}\}$, confirming Proposition 20. If we replace B by the 4×3 matrix $\hat{B} = diag\{3, \mathrm{i}, -\mathrm{i}\}$, then the convex hull of the eigenvalues $\mathrm{i}/3, -\mathrm{i}, -1+\mathrm{i} \in \sigma(A;\hat{B})$ (marked with +'s) is a strict subset of $W_{\|\cdot\|_2}(A;\hat{B})$, as we see in the right part of Figure 5. It is worth noting that the eigenvalues that correspond to the diagonal entries of $\hat{B}$ of modulus one remain on the boundary of the numerical range.

6 Numerical range of linear pencils

In the case of square matrices, i.e., for $n = m$, a question that arises in a natural way is how the range $W_{\|\cdot\|_2}(A;B)$ is related to the numerical range of the linear pencil $A - \lambda B$, that is,

$$W(A - \lambda B) \,=\, \{\mu \in \mathbb{C} : x^*(A - \mu B)x = 0,\, x \in \mathbb{C}^n,\, x^*x = 1\}\,.$$

It is known that $W(A - \lambda B)$ is a closed subset of the complex plane, but it is not necessarily convex, and it is bounded if and only if $0 \notin F(B)$ [14]. Moreover, if the matrices A and B have a common isotropic vector, i.e., a nonzero vector $y \in \mathbb{C}^n$ such that $y^*Ay = y^*By = 0$, then $W(A - \lambda B) = \mathbb{C}$. Thus, it is clear that in general, the ranges $W_{\|\cdot\|_2}(A;B)$ and $W(A - \lambda B)$ are different. On the other hand, it is worth mentioning that for $B = I_n$, $W_{\|\cdot\|_2}(A;I_n) = F(A) = W(A - \lambda I_n)$.

Proposition 21. *Suppose $A, B \in \mathbb{C}^{n\times n}$ (with $\|B\|_2 \geq 1$) and $\mu_0 \in W(A - \lambda B)$. If there is a unit (with respect to the spectral norm) vector $x_0 \in \mathbb{C}^n$ such that $x_0^*(A - \mu_0 B)x_0 = 0$ and $|x_0^* B x_0| \geq 1$, then $\mu_0 \in W_{\|\cdot\|_2}(A;B)$.*

Proof. Let $\mu_0 \in W(A - \lambda B)$ and $x_0 \in \mathbb{C}^n$ such that $x_0^*(A - \mu_0 B)x_0 = 0$, $x_0^* x_0 = 1$ and $|x_0^* B x_0| \geq 1$. Then for every $\lambda \in \mathbb{C}$,

$$x_0^*(A - \lambda B)x_0 \,=\, -\,(\lambda - \mu_0)\, x_0^* B x_0,$$

and hence,

$$\|A - \lambda B\| \;\geq\; |\lambda - \mu_0|\; |x_0^* B x_0| \;\geq\; |\mu_0 - \lambda|\,.$$

As a consequence, $\mu_0 \in W_{\|\cdot\|_2}(A; B)$. $\square$

Corollary 22. *If the standard numerical range of $B \in \mathbb{C}^{n\times n}$, $F(B)$, does not contain interior points of the unit disc $\mathcal{D}(0,1)$, then $W(A - \lambda B) \subseteq W_{\|\cdot\|_2}(A; B)$.*

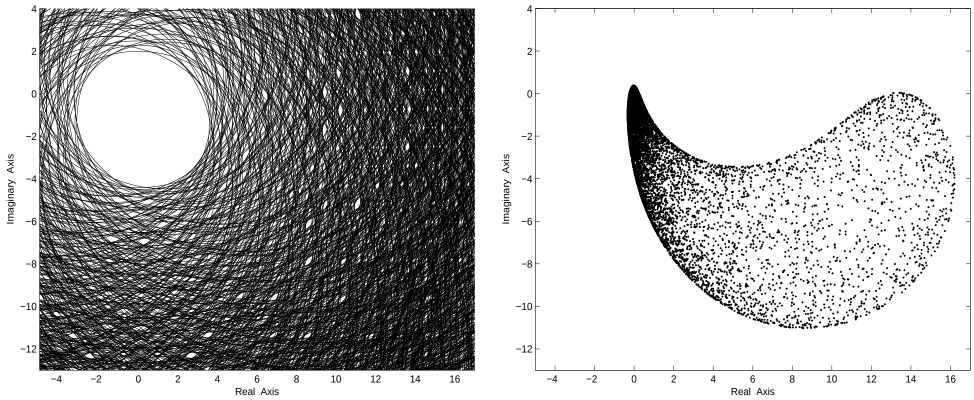

Figure 6: The numerical ranges $W_{\|\cdot\|_2}(A; B)$ (left) and $W(A - \lambda B)$ (right).

Note that for sufficiently small $|x^*Bx| \neq 0$ $(x^*x = 1)$, the ratio $(x^*Ax)/(x^*Bx)$ lies in $W(A - \lambda B)$ but not in $W_{\|\cdot\|_2}(A; B)$. For example, consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} \mathrm{i}\,1.5 & 0 \\ 0 & 0.3 \end{bmatrix}$. The unshaded region in the left part of Figure 6 and the shaded region in the right part of the figure are estimations of the numerical ranges $W_{\|\cdot\|_2}(A; B)$ and $W(A - \lambda B)$, respectively. It is obvious that $W(A - \lambda B) \not\subseteq W_{\|\cdot\|_2}(A; B)$.

If $\|B\|_2 = 1$, then $F(B)$ lies in $\mathcal{D}(0,1)$ and Corollary 22 is not applicable. In this case, we have the next inclusion result.

Proposition 23. *For any $A, B \in \mathbb{C}^{n\times n}$ such that B is nonsingular with spectral norm $\|B\|_2 = 1$, it holds that $W_{\|\cdot\|_2}(A; B) \subseteq F(AB^{-1}) \cap F(B^{-1}A)$.*

Proof. Since B is nonsingular and $\|B\|_2 = 1$,

$$\begin{aligned} W_{\|\cdot\|_2}(A; B) &= \left\{\mu \in \mathbb{C} : \left\|(AB^{-1} - \lambda I_n)B\right\|_2 \geq |\mu - \lambda|, \,\forall\, \lambda \in \mathbb{C}\right\} \\ &\subseteq \left\{\mu \in \mathbb{C} : \left\|AB^{-1} - \lambda I_n\right\|_2 \geq |\mu - \lambda|, \,\forall\, \lambda \in \mathbb{C}\right\} \\ &= F(AB^{-1}). \end{aligned}$$

Similarly, we verify that $W_{\|\cdot\|_2}(A; B) \subseteq F(B^{-1}A)$. $\square$

Acknowledgement. The authors acknowledge with thanks an anonymous referee for his/her useful suggestions.

References

[1] A. Abdollahi, The polynomial numerical hull of a matrix and algorithms for computing the numerical range, *Appl. Math. Comput.*, **180** (2006) 635–640.

[2] R. Bhatia and P. Šemrl, Orthogonality of matrices and some distance problems, *Linear Algebra Appl.*, **287** (1999) 77–85.

[3] D. Boley, Estimating the sensitivity of the algebraic structure of pencils with simple eigenvalue estimates, *SIAM J. Matrix Anal. Appl.*, **11** (1990) 632–643.

[4] G. Boutry, M. Elad, G.H. Golub and P. Milanfar, The generalized eigenvalue problem for nonsquare pencils using a minimal perturbation approach, *SIAM J. Matrix Anal. Appl.*, **27** (2005) 582–601.

[5] F.F. Bonsall and J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, London Mathematical Society Lecture Note Series, Cambridge University Press, New York, 1971.

[6] F.F. Bonsall and J. Duncan, *Numerical Ranges II*, London Mathematical Society Lecture Notes Series, Cambridge University Press, New York, 1973.

[7] A. Greenbaum, Generalizations of the field of values useful in the study of polynomial functions of a matrix, *Linear Algebra Appl.*, **347** (2002) 233–249.

[8] K.E. Gustafson and D.K.M. Rao, *Numerical Range. The Field of Values of Linear Operators and Matrices*, Springer-Verlag, New York, 1997.

[9] D.W. Hadwin, K.J. Harrison and J.A. Ward, Numerical ranges and matrix completions, *Linear Algebra Appl.*, **315** (2000) 145–154.

[10] P.R. Halmos, *A Hilbert Space Problem Book*, 2nd edition, Springer-Verlag, New York, 1982.

[11] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge 1985.

[12] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge 1991.

[13] R.C. James, Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.*, **61** (1947) 265–292.

[14] C.-K. Li, and L. Rodman, Numerical range of matrix polynomials, *SIAM J. Matrix Anal. Appl.*, **15** (1994) 1256–1265.

[15] G. Lumer, Semi-inner product spaces, *Trans. Amer. Math. Soc.*, **100** (1961) 29–43.

[16] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser, Basel, 1993.

[17] J.G. Stamfli and J.P. Williams, Growth condition and the numerical range in a Banach algebra, *Tohoku Math. Journ.*, **20** (1968) 417–424.

[18] G.W. Stewart, Perturbation theory for rectangular matrix pencils, *Linear Algebra Appl.*, **208/209** (1994) 297–301.

[19] G.W. Stewart and J.-Q. Sun, *Matrix Perturbation Theory*, Academic Press, 1991.

[20] G.L. Thompson and R.L. Weil, The roots of matrix pencils ($Ay = \lambda By$): Existence, calculations, and relations to game theory, *Linear Algebra Appl.*, **5** (1972) 207–226.