Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

A NOTE ON THE BOUNDARY OF THE BIRKHOFF–JAMES ε-ORTHOGONALITY SETS*

GEORGIOS KATSOULEAS†, VASILIKI PANAGAKOU†, AND PANAYIOTIS PSARRAKOS†

Abstract. The Birkhoff–James ε-sets of vectors and vector-valued polynomials (in one complex variable) have recently been introduced as natural generalizations of the standard numerical range of (square) matrices or operators and matrix or operator polynomials, respectively. Corners on the boundary curves of these sets are of particular interest, not least because of their importance in visualizing these sets. In this paper, we provide a characterization for the corners of the Birkhoff-James ε-sets of vectors and vector-valued polynomials, completing and expanding upon previous exploration of the geometric properties of these sets. We also propose a randomized algorithm for approximating their boundaries.

Key words. Birkhoff–James orthogonality, Norm, Vector-valued polynomial, Numerical range, Matrix polynomials, Randomized algorithms.

AMS subject classifications. 15A60, 47A12, 68W20, 68W25.

1. Introduction. The standard numerical range has been extensively studied for more than a century and is a useful tool in understanding (square) matrices and operators; see [2, 3, 13, 14, 30] and the references therein. During the last few decades, applications on higher-order linear systems of differential (or difference) equations and stability theory have motivated generalizations to matrix and operator polynomials [10, 12, 22, 23]. More precisely, the *numerical range* of a matrix polynomial:

$$P(z) = A_m z^m + A_{m-1} z^{m-1} + \cdots + A_1 z + A_0,$$

where $z \in \mathbb{C}$ is a complex variable and $A_j \in \mathbb{C}^{n\times n}$ $(j = 0, 1, \ldots, m)$ with $A_m \neq 0$, is defined (see e.g., [22]) by:

$$W(P) = \{\mu \in \mathbb{C} : x^* P(\mu) x = 0 \text{ for some nonzero } x \in \mathbb{C}^n\}.$$

For $P(z) = zI - A$, the set $W(P)$ reduces to the classical numerical range $F(A) = \{x^* A x : x \in \mathbb{C}^n, \|x\|_2 = 1\}$ of the matrix A (also known as *field of values*) [14].

Extensions beyond the realm of square matrices to general *rectangular* matrices and vectors (using some norm instead of the inner product) have been inspired by the notion of the Birkhoff–James ε-orthogonality [1, 8, 16]. Based on this concept, Chorianopoulos and Psarrakos [7] and Karamanlis and Psarrakos [18] introduced the Birkhoff–James ε-orthogonality sets of rectangular matrices and elements of a complex normed linear space, respectively. These sets are direct generalizations of the standard numerical range of matrices and operators, and their interesting geometrical properties were studied in [6, 7, 18, 27]. Then, the corresponding Birkhoff–James ε-set of vector-valued polynomials (in one complex variable) was recently introduced by Panagakou, Psarrakos, and Yannakakis in [27] and also demonstrates a rich structure. The

*Received by the editors on August 31, 2021. Accepted for publication on December 18, 2021. Handling Editor: Ilya Spitkovsky. Corresponding Author: Georgios Katsouleas.

†Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Greece (gekats@mail.ntua.gr, vpanagakou@mail.ntua.gr, ppsarr@math.ntua.gr).

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

analysis in [27] was based on an equivalent description of the Birkhoff–James ε-set which involves continuous linear functionals, while a preliminary definition for rectangular matrix polynomials had already appeared in [7].

The continued interest in such sets is readily explained by the potential of their geometry to capture and reveal information about the algebraic and analytic properties of the corresponding linear transformations. Pride of place must go to the study of their boundary, not least because of its importance for visualization purposes. Particular problems in this direction may arise in cases where the boundary is nonsmooth. In fact, the (non)-smoothness of the boundary of the numerical range is well known to reveal spectral information. In particular, *corners* of the numerical range boundary curve are (normal and semisimple) eigenvalues of the corresponding matrix (or, under some weak conditions, of the corresponding matrix polynomial). Related results regarding the shape, the boundary, and the corners of the numerical range of matrix polynomials can be found in [7, 20, 24, 26].

In this note, we focus on corners of the boundary of the Birkhoff–James ε-orthogonality set of vectors and the Birkhoff–James ε-set of vector-valued polynomials. Specifically, in Section 2, we briefly recall the relevant background; namely, the definitions and the various equivalent descriptions of these two sets, along with their basic properties, revealing their interesting structure. In Section 3, we formally introduce the notion of a corner and provide a characterization for the corners of the Birkhoff–James ε-orthogonality sets, completing (in some sense) the geometrical exploration of these sets which was initiated in [27] and furthering our understanding of vector-valued polynomials. Finally, in Section 4, we propose a randomized algorithm for drawing (approximately) the boundaries of the Birkhoff–James ε-orthogonality sets, which is quite faster than the standard grid method.

2. Definitions and basic properties. Consider a complex normed linear space $(\mathcal{X}, \|\cdot\|)$ (henceforth, for simplicity abbreviated as $\mathcal{X}$). For two elements x and y of $\mathcal{X}$, x is said to be *Birkhoff–James orthogonal* to y, denoted by $x \perp_{BJ} y$, if $\|x + \lambda y\| \geq \|x\|$ for all $\lambda \in \mathbb{C}$ [1, 16]. The Birkhoff–James orthogonality is homogeneous, but it is neither symmetric nor additive [16]. Moreover, for any $\varepsilon \in [0, 1)$, x is said to be *Birkhoff–James ε-orthogonal* to y, denoted by $x \perp^{\varepsilon}_{BJ} y$, if [5, 8]

$$\|x + \lambda y\| \geq \sqrt{1 - \varepsilon^2}\, \|x\|, \quad \forall\, \lambda \in \mathbb{C}. \tag{2.1}$$

The Birkhoff–James ε-orthogonality is also homogeneous. If the norm $\|\cdot\|$ is induced by an inner product $\langle \cdot, \cdot \rangle$, then an $x \in \mathcal{X}$ is said to be *ε-orthogonal* to a $y \in \mathcal{X}$, denoted by $x \perp^{\varepsilon} y$, if $|\langle x, y\rangle| \leq \varepsilon \|x\| \|y\|$; apparently, for $\varepsilon = 0$, we have the standard orthogonality. Furthermore, for any $\varepsilon \in [0, 1)$, $x \perp^{\varepsilon} y$ if and only if $x \perp^{\varepsilon}_{BJ} y$ [5, 8].

For any $x, y \in \mathcal{X}$, with $y \neq 0$, and any $\varepsilon \in [0, 1)$, the *Birkhoff–James ε-orthogonality set of x with respect to y* is defined and denoted by:

$$F^{\varepsilon}_{\|\cdot\|}(x; y) = \{\mu \in \mathbb{C} : \, y \perp^{\varepsilon}_{BJ} (x - \mu y)\}.$$

Keeping in mind (2.1), it is straightforward to see that

$$\begin{aligned} F^{\varepsilon}_{\|\cdot\|}(x; y) &= \left\{\mu \in \mathbb{C} : \, \|y - \lambda(x - \mu y)\| \geq \sqrt{1 - \varepsilon^2}\, \|y\|, \ \forall\, \lambda \in \mathbb{C}\right\} \\ &= \left\{\mu \in \mathbb{C} : \, \left\|y - \frac{1}{\lambda}(x - \mu y)\right\| \geq \sqrt{1 - \varepsilon^2}\, \|y\|, \ \forall\, \lambda \in \mathbb{C} \setminus \{0\}\right\} \\ &= \left\{\mu \in \mathbb{C} : \, \frac{1}{|\lambda|}\, \|\lambda y - (x - \mu y)\| \geq \sqrt{1 - \varepsilon^2}\, \|y\|, \ \forall\, \lambda \in \mathbb{C} \setminus \{0\}\right\} \end{aligned}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

$$= \left\{\mu \in \mathbb{C} : \|x - (\mu+\lambda)y\| \geq \sqrt{1-\varepsilon^2}\,\|y\|\,|\lambda|,\ \forall\,\lambda \in \mathbb{C}\right\}$$
$$= \left\{\mu \in \mathbb{C} : \|x - \lambda y\| \geq \sqrt{1-\varepsilon^2}\,\|y\|\,|\mu-\lambda|,\ \forall\,\lambda \in \mathbb{C}\right\}$$
$$= \bigcap_{\lambda\in\mathbb{C}} \mathcal{D}\left(\lambda, \frac{\|x-\lambda y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}\right), \tag{2.2}$$

where $\mathcal{D}\left(\lambda, \frac{\|x-\lambda y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}\right)$ denotes the closed circular disk with center at λ and radius $\frac{\|x-\lambda y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}$. Since the set $F^{\varepsilon}_{\|\cdot\|}(x;y)$ is an infinite intersection of closed circular disks, its compactness and convexity are apparent. Moreover, it is non-empty by Corollary 2.2 of [16].

The Birkhoff–James ε-orthogonality set is a direct generalization of the standard numerical range of matrices. In particular, for a square complex matrix $A \in \mathbb{C}^{n\times n}$ and $\|\cdot\| = \|\cdot\|_2$, the Birkhoff–James 0-orthogonality set $F^{0}_{\|\cdot\|}(A; I_n)$ (where I_n is the $n\times n$ identity matrix) coincides [7] with the numerical range $F(A)$ of A.

Let $\mathcal{X}^*$ denote the dual space of $\mathcal{X}$, that is, the complex linear space of all continuous linear functionals of $\mathcal{X}$ (using the induced operator norm). Consider also two elements x and y of $\mathcal{X}$, with $y \neq 0$. For any $\varepsilon \in [0,1)$, we recall the next definition [27]:

$$L_{\varepsilon}(y) = \left\{f \in \mathcal{X}^* \,:\, f(y) = \sqrt{1-\varepsilon^2}\,\|y\|,\ \|f\| \leq 1\right\}.$$

This set is non-empty, closed and convex [27, Lemma 2.1]. Using this definition, an alternative characterization of the Birkhoff–James ε-orthogonality set, namely,

$$F^{\varepsilon}_{\|\cdot\|}(x;y) = \left\{\frac{f(x)}{\sqrt{1-\varepsilon^2}\,\|y\|} \,:\, f \in L_{\varepsilon}(y)\right\}, \tag{2.3}$$

was obtained in [27, Theorem 2.4]. Next, for clarity and reader's convenience, we summarize results of [18, 27] (see also [6, 7] for rectangular matrices), describing basic properties of the Birkhoff–James ε-orthogonality set.

(P_1) For any $a, b \in \mathbb{C}$ and any $\varepsilon \in [0,1)$, $F^{\varepsilon}_{\|\cdot\|}(ax+by;y) = a\,F^{\varepsilon}_{\|\cdot\|}(x;y) + b$.

(P_2) For any nonzero $b \in \mathbb{C}$ and any $\varepsilon \in [0,1)$, $F^{\varepsilon}_{\|\cdot\|}(x;by) = \frac{1}{b}\,F^{\varepsilon}_{\|\cdot\|}(x;y)$.

(P_3) If x is a nonzero element of $\mathcal{X}$, then for any $\varepsilon \in [0,1)$,

$$\left\{\mu^{-1} \in \mathbb{C} : \mu \in F^{\varepsilon}_{\|\cdot\|}(x;y),\ |\mu| \geq \frac{\|x\|}{\|y\|}\right\} \subseteq F^{\varepsilon}_{\|\cdot\|}(y;x).$$

(P_4) $x = ay$ for some $a \in \mathbb{C}$ if and only if $F^{\varepsilon}_{\|\cdot\|}(x;y) = \{a\}$ for every $\varepsilon \in [0,1)$.

(P_5) If x is not a scalar multiple of y, then for any $0 \leq \varepsilon_1 < \varepsilon_2 < 1$, $F^{\varepsilon_1}_{\|\cdot\|}(x;y)$ lies in the interior of $F^{\varepsilon_2}_{\|\cdot\|}(x;y)$. As a consequence, for any $\varepsilon \in (0,1)$, $F^{\varepsilon}_{\|\cdot\|}(x;y)$ has a non-empty interior.

(P_6) If x is not a scalar multiple of y, then for any bounded region $\Omega \subset \mathbb{C}$, there is an $\varepsilon_\Omega \in [0,1)$ such that $\Omega \subseteq F^{\varepsilon_\Omega}_{\|\cdot\|}(x;y)$ (i.e., $F^{\varepsilon}_{\|\cdot\|}(x;y)$ can be arbitrarily large for ε sufficiently close to 1).

(P_7) A $\mu_0 \in F^{\varepsilon}_{\|\cdot\|}(x;y)$ lies on the boundary $\partial F^{\varepsilon}_{\|\cdot\|}(x;y)$ if and only if $\inf\limits_{\lambda\in\mathbb{C}} \left\{\|x-\lambda y\| - \sqrt{1-\varepsilon^2}\,\|y\|\,|\mu_0-\lambda|\right\} = 0$, where '$\inf$' can be replaced by '$\min$' when $\varepsilon > 0$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

(P_8) If the norm $\|\cdot\|$ is induced by an inner product $\langle\cdot,\cdot\rangle$, then for any $\varepsilon \in [0,1)$,

$$F^{\varepsilon}_{\|\cdot\|}(x;y) = \mathcal{D}\left(\frac{\langle x,y\rangle}{\|y\|^2}\,,\ \left\|x - \frac{\langle x,y\rangle}{\|y\|^2}\,y\right\| \frac{\varepsilon}{\sqrt{1-\varepsilon^2}\,\|y\|}\right).$$

(P_9) For any $x_1, x_2 \in \mathcal{X}$ and $\varepsilon \in [0,1)$, it holds that $F^{\varepsilon}_{\|\cdot\|}(x_1+x_2;y) \subseteq F^{\varepsilon}_{\|\cdot\|}(x_1;y) + F^{\varepsilon}_{\|\cdot\|}(x_2;y)$.

Based on the definition of the Birkhoff–James ε-orthogonality set of vectors, the Birkhoff–James ε-orthogonality set of vector-valued polynomials (in one complex variable) was introduced in [27]. Consider a vector-valued polynomial:

$$P(z) = x_m z^m + x_{m-1}z^{m-1} + \cdots + x_1 z + x_0, \tag{2.4}$$

with vector coefficients $x_i \in \mathcal{X}$ $(i = 0,1,\ldots,m)$, $x_m \neq 0$, and a scalar variable $z \in \mathbb{C}$. Vector-valued polynomials appear in the approximation of vector-valued functions [28]. Moreover, special cases of vector-valued polynomials such as square matrix polynomials [9, 10, 11, 17, 19], rectangular matrix polynomials [9, 17], and operator polynomials [12, 15, 21, 23] arise in many applications.

Let $P(z)$ be a vector-valued polynomial as in (2.4), $\varepsilon \in [0,1)$, and $y \in \mathcal{X}$ be a nonzero vector such that $F^{\varepsilon}_{\|\cdot\|}(x_m;y) \neq \{0\}$. The *Birkhoff–James ε-orthogonality set of $P(z)$ with respect to y* is defined and denoted by:

$$\begin{aligned} W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big) &= \Big\{\mu \in \mathbb{C} :\ 0 \in F^{\varepsilon}_{\|\cdot\|}\big(P(\mu);y\big)\Big\} \\ &= \big\{\mu \in \mathbb{C} :\ f\big(P(\mu)\big) = 0,\ f \in L_{\varepsilon}(y)\big\} \\ &= \{\mu \in \mathbb{C} :\ y \perp^{\varepsilon}_{BJ} P(\mu)\} \\ &= \Big\{\mu \in \mathbb{C} :\ \|P(\mu) - \lambda y\| \geq \sqrt{1-\varepsilon^2}\,\|y\|\,|\lambda|,\ \forall\,\lambda \in \mathbb{C}\Big\}. \end{aligned} \tag{2.5}$$

It is worth noting that for $x_m \neq 0$ and $\varepsilon \in (0,1)$, the condition $F^{\varepsilon}_{\|\cdot\|}(x_m;y) \neq \{0\}$ is always satisfied; see Properties (P_4), (P_5), and (P_6).

Since the set $L_{\varepsilon}(y)$ is non-empty and closed, it follows readily that $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ is also non-empty and closed. Moreover, for any $0 \leq \varepsilon_1 < \varepsilon_2 < 1$, $W^{\varepsilon_1}_{\|\cdot\|}\big(P(z);y\big) \subseteq W^{\varepsilon_2}_{\|\cdot\|}\big(P(z);y\big)$. When $P(z)$ is a square matrix polynomial, note that the Birkhoff–James 0-orthogonality set $W^{0}_{\|\cdot\|_2}\big(P(z);I_n\big)$ reduces to the usual numerical range $W(P)$ of $P(z)$, in complete analogy to the equality $F^{0}_{\|\cdot\|_2}(A;I_n) = F(A)$ for $A \in \mathbb{C}^{n\times n}$.

Consider a vector-valued polynomial $P(z)$ as in (2.4), the vector-valued polynomial $P'(z) = mx_m z^{m-1} + (m-1)x_{m-1}z^{m-2} + \cdots + 2x_2 z + x_1$, an $\varepsilon \in [0,1)$, and a nonzero vector $y \in \mathcal{X}$ with $F^{\varepsilon}_{\|\cdot\|}(x_m;y) \neq \{0\}$. Next, again for clarity and reader's convenience, we summarize results of [27] (see also [7, 20, 22, 26] for matrix polynomials), describing basic properties of the Birkhoff–James ε-orthogonality set of $P(z)$.

(P_{10}) For any scalar $a \in \mathbb{C}\setminus\{0\}$, it holds $W^{\varepsilon}_{\|\cdot\|}\big(P(az);y\big) = a^{-1}W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ and $W^{\varepsilon}_{\|\cdot\|}\big(P(z+a);y\big) = W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big) - a$.

(P_{11}) If $R(z) = x_0 z^m + x_1 z^{m-1} + \cdots + x_{m-1}z + x_m = z^m P\big(z^{-1}\big)$ is the *reverse vector-valued polynomial* of $P(z)$, then

$$W^{\varepsilon}_{\|\cdot\|}\big(R(z);y\big)\setminus\{0\} = \Big\{\mu \in \mathbb{C} :\ \mu^{-1} \in W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)\setminus\{0\}\Big\}.$$

(P_{12}) If there exists a continuous linear functional $f \in L_{\varepsilon}(y)$ such that $f(x_m) = f(x_{m-1}) = \cdots = f(x_0) = 0$, then $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big) = \mathbb{C}$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

(P_{13}) The set $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ is bounded if and only if $0 \notin F^{\varepsilon}_{\|\cdot\|}(x_m;y)$.

(P_{14}) Let $y \in \mathcal{X}$ be a nonzero vector such that $0 \notin F^{\varepsilon}_{\|\cdot\|}(x_m;y)$ (or equivalently, $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ is bounded), and suppose that $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ has r connected components. If κ is the minimum number of distinct zeros of the scalar polynomial $f(P(z)) = f(x_m)z^m + f(x_{m-1})z^{m-1} + \cdots + f(x_1)z + f(x_0)$ over all $f \in L_{\varepsilon}(y)$, then $r \leq \kappa \leq m$.

(P_{15}) If $z_0 \in \partial W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$, then $0 \in \partial F^{\varepsilon}_{\|\cdot\|}\big(P(z_0);y\big)$.

(P_{16}) Let $z_0 \in W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$ such that $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0);y\big) \neq \{0\}$ and $0 \notin F^{\varepsilon}_{\|\cdot\|}\big(P'(z_0);y\big)$. If $0 \in \partial F^{\varepsilon}_{\|\cdot\|}\big(P(z_0);y\big)$, then $z_0 \in \partial W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$.

(P_{17}) Let $y \in \mathcal{X}$ be a nonzero vector such that $0 \notin F^{\varepsilon}_{\|\cdot\|}(x_m;y)$. If z_0 is an isolated point of $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$, then $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0);y\big) = \{0\}$. If, in addition, $\varepsilon > 0$, then $P(z_0) = 0$.

3. Corners of the Birkhoff–James ε-orthogonality sets. In this section, we investigate the corners of the (closed) Birkhoff–James ε-orthogonality sets $F^{\varepsilon}_{\|\cdot\|}(x;y)$ and $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$. A boundary point z_0 of a closed subset Ω of the complex plane is called a *corner* of Ω if there exist some $\delta > 0$ and two angles $\phi_1, \phi_2 \in [0, 2\pi]$, with $\phi_2 - \phi_1 \in [0, \pi)$, such that

$$\phi_1 \leq \arg(z - z_0) \leq \phi_2, \quad \forall\, z \in \Omega \cap \mathcal{D}(z_0, \delta),$$

where $\mathcal{D}(z_0, \delta)$ is the closed circular disk with center z_0 and radius δ.

Consider a complex normed linear space $(\mathcal{X}, \|\cdot\|)$ (for simplicity $\mathcal{X}$) and let $x, y \in \mathcal{X}$ with $y \neq 0$. For arbitrary $\varepsilon \in [0, 1)$, we will be making use of the set:

$$A_{\varepsilon}(x;y) = \{f \in L_{\varepsilon}(y) : \ f(x) = 0\},$$

which is convex by [27, Lemma 4.3]. Moreover, removing the restriction $\|f\| \leq 1$ in $L_{\varepsilon}(y)$, we introduce the set:

$$\widehat{L}_{\varepsilon}(y) = \Big\{f \in \mathcal{X}^* : \ f(y) = \sqrt{1-\varepsilon^2}\,\|y\|\Big\},$$

and provide the analog definition of

$$\widehat{A}_{\varepsilon}(x;y) = \Big\{f \in \widehat{L}_{\varepsilon}(y) : \ f(x) = 0\Big\}.$$

It is straightforward to see that $\widehat{A}_{\varepsilon}(x;y)$ is also convex. In the remainder of the paper, the ball centered at a linear functional $f_0 \in \mathcal{X}^*$ with radius $\delta \geq 0$ is denoted by $\mathcal{B}(f_0, \delta)$.

LEMMA 3.1. *Let $f_0 \in A_{\varepsilon}(x;y)$ be a linear functional (i.e., $f_0 \in L_{\varepsilon}(y)$ with $f_0(x) = 0$) such that*

$$\phi_1 \leq \arg\big(f(x)\big) \leq \phi_2, \quad \forall\, f \in \widehat{L}_{\varepsilon}(y) \cap \mathcal{B}(f_0, \delta),$$

for some radius $\delta > 0$ and a pair of angles $\phi_1, \phi_2 \in [0, 2\pi]$, with $\phi_2 - \phi_1 \in [0, \pi)$. Then, for arbitrary $g_0 \in \widehat{A}_{\varepsilon}(x;y)$, we have

$$\phi_1 \leq \arg\big(g(x)\big) \leq \phi_2, \quad \forall\, g \in \widehat{L}_{\varepsilon}(y) \cap \mathcal{B}(g_0, \delta).$$

Proof. Let $f_0 \in L_{\varepsilon}(y)$ be a linear functional such that $f_0(x) = 0$, and consider scalars $\delta > 0$ and $\phi_1, \phi_2 \in [0, 2\pi]$ with $\phi_2 - \phi_1 \in [0, \pi)$, such that

$$\phi_1 \leq \arg\big(f(x)\big) \leq \phi_2, \quad \forall\, f \in \widehat{L}_{\varepsilon}(y) \cap \mathcal{B}(f_0, \delta).$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

Then, for any linear functional $g_0 \in \widehat{L}_\varepsilon(y)$ with $g_0(x) = 0$ and for every $g \in \mathcal{B}(g_0, \delta) \cap \widehat{L}_\varepsilon(y)$, it holds that

$$f_g = g + f_0 - g_0 \in \mathcal{B}(f_0, \delta),$$

because $\|f_g - f_0\| = \|g - g_0\| \leq \delta$, and

$$f_g = g + f_0 - g_0 \in \widehat{L}_\varepsilon(y),$$

because $f_g(y) = g(y) + f_0(y) - g_0(y) = g(y) = \sqrt{1-\varepsilon^2}\,\|y\|$. This means that $f_g = g + f_0 - g_0 \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta)$, and by hypothesis, we have

$$\begin{aligned}\phi_1 \leq \arg\big(f_g(x)\big) \leq \phi_2 &\Leftrightarrow \phi_1 \leq \arg\big(g(x) + f_0(x) - g_0(x)\big) \leq \phi_2 \\ &\Leftrightarrow \phi_1 \leq \arg\big(g(x)\big) \leq \phi_2,\end{aligned}$$

completing the proof. ∎

Under the assumptions of Lemma 3.1, an alternative way to express its conclusion would be by saying that for every linear functional:

$$g \in \widehat{A}_\varepsilon(x; y) + \mathcal{B}(0, \delta) = \Big\{ f \in \mathcal{X}^* : \|f - f_0\| \leq \delta \text{ for some } f_0 \in \widehat{A}_\varepsilon(x; y) \Big\},$$

the argument of the complex scalar $g(x)$ is bounded by:

$$\phi_1 \leq \arg\big(g(x)\big) \leq \phi_2.$$

Proposition 3.2. *Let $f_0 \in A_\varepsilon(x; y)$ be a linear functional (i.e., $f_0 \in L_\varepsilon(y)$ with $f_0(x) = 0$). Suppose there exist a $\delta > 0$ and two angles $\phi_1, \phi_2 \in [0, 2\pi]$, with $\phi_2 - \phi_1 \in [0, \pi)$, such that*

$$\phi_1 \leq \arg\big(f(x)\big) \leq \phi_2, \quad \forall\, f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta). \tag{3.6}$$

Then, the origin is a corner of the set $F^\varepsilon_{\|\cdot\|}(x; y)$.

Proof. By the discussion following Lemma 3.1, the assumption (3.6) on $f_0 \in A_\varepsilon(x; y)$ ensures

$$\phi_1 \leq \arg\big(g(x)\big) \leq \phi_2,$$

for every $g \in \widehat{A}_\varepsilon(x; y) + \mathcal{B}(0, \delta)$. If we assume that the origin is not a corner of $F^\varepsilon_{\|\cdot\|}(x; y)$, then there exists some point z_0 of $F^\varepsilon_{\|\cdot\|}(x; y)$ that lies outside the cone $\{z \in \mathbb{C} : \phi_1 \leq \arg(z) \leq \phi_2\}$. By the characterization of the Birkhoff–James ε-orthogonality set in (2.3), there exists a linear functional $h_0 \in L_\varepsilon(y)$ such that $\dfrac{h_0(x)}{\sqrt{1-\varepsilon^2}\,\|y\|} = z_0$.

Consider now the line segment:

$$f_t = (1-t)h_0 + tf_0, \quad t \in [0, 1],$$

in the convex set $L_\varepsilon(y)$, which joins h_0 with f_0. Then, the line segment:

$$f_t(x) = (1-t)h_0(x) + tf_0(x) = (1-t)\sqrt{1-\varepsilon^2}\,\|y\|z_0, \quad t \in [0, 1],$$

joins the point $\sqrt{1-\varepsilon^2}\,\|y\|z_0$ with 0. Since f_t lies in the interior of $\widehat{A}_\varepsilon(x; y) + \mathcal{B}(0, \delta)$, there exists a $t_0 \in (0, 1)$ such that $f_t \in \widehat{A}_\varepsilon(x; y)$ for $t \geq t_0$. Consequently,

$$f_t(x) = (1-t)\sqrt{1-\varepsilon^2}\,\|y\|z_0 \in \{z \in \mathbb{C} : \phi_1 \leq \arg(z) \leq \phi_2\},$$

which is a contradiction. Hence, the origin is a corner of $F^\varepsilon_{\|\cdot\|}(x; y)$. ∎

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

The converse of Proposition 3.2 is obvious. Indeed, supposing that the origin is a corner of $F^{\varepsilon}_{\|\cdot\|}(x;y)$ and the (convex) set $F^{\varepsilon}_{\|\cdot\|}(x;y)$ lies in a cone $\{z \in \mathbb{C} : \phi_1 \leq \arg(z) \leq \phi_2,\ \phi_2 - \phi_1 \in [0,\pi)\}$, then for any $f \in L_\varepsilon(y)$, it is clear from (2.3) that $\phi_1 \leq \arg(f(x)) \leq \phi_2$, and this observation can be extended to elements $f \in \widehat{L}_\varepsilon(y)$. Hence, we may state

COROLLARY 3.3. *Let $\mathcal{X}$ be a normed linear space. Consider two vectors $x, y \in \mathcal{X}$ with $y \neq 0$ and some $\varepsilon \in [0,1)$. The following are equivalent:*

(i) *The origin is a corner of $F^{\varepsilon}_{\|\cdot\|}(x;y)$.*
(ii) *There exist a linear functional $f_0 \in A_\varepsilon(x;y)$ (i.e., $f_0 \in L_\varepsilon(y)$ with $f_0(x) = 0$), some radius $\delta > 0$, and two angles $\phi_1, \phi_2 \in [0, 2\pi]$, with $\phi_2 - \phi_1 \in [0,\pi)$, such that*

$$\phi_1 \leq \arg\big(f(x)\big) \leq \phi_2, \quad \forall\, f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta).$$

Now, we turn our attention to vector-valued polynomials and corners of $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$. As the first step, we proceed to investigate the linear polynomial case.

THEOREM 3.4. *Let $P(z) = x_1 z - x_0$, and let z_0 be a corner of the Birkhoff–James ε-orthogonality set $W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big)$. Then, the origin is a corner of $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0);y\big)$.*

Proof. Let

$$Q(z) = P(z + z_0) = x_1 z + (x_1 z_0 - x_0).$$

The fact that z_0 is a corner of $W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ implies that the origin is a corner of

$$W^{\varepsilon}_{\|\cdot\|}\big(Q(z);y\big) = W^{\varepsilon}_{\|\cdot\|}\big(P(z+z_0);y\big) = W^{\varepsilon}_{\|\cdot\|}\big(P(z);y\big) - z_0,$$

by the translation Property (P_{10}). Since $0 \in W^{\varepsilon}_{\|\cdot\|}\big(Q(z);y\big)$, the corresponding definition ensures the existence of some linear functional $f_0 \in L_\varepsilon(y)$, such that the origin is root of the polynomial:

$$p(z) = f_0(x_1)z + f_0(x_1 z_0 - x_0).$$

Note that $f_0(x_1 z_0 - x_0) = f_0(x_1) = 0$ cannot hold simultaneously, since then Property (P_{12}) would imply $W^{\varepsilon}_{\|\cdot\|}(Q(z);y) = \mathbb{C}$. Hence, $f_0(x_1 z_0 - x_0) = 0$ and $f_0(x_1) \neq 0$. Additionally, since the origin is in fact a *corner* of $W^{\varepsilon}_{\|\cdot\|}\big(Q(z);y\big)$, there exist radii $r, \rho > 0$ and angles $\phi_1, \phi_2 \in [0, 2\pi]$ with $\phi_2 - \phi_1 \in [0,\pi)$, such that for every $f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \rho)$, the corresponding root:

$$\mu_f = -\frac{f(x_1 z_0 - x_0)}{f(x_1)} \in W^{\varepsilon}_{\|\cdot\|}\big(Q(z);y\big) \cap \mathcal{D}(0,r),$$

lies in a neighborhood of 0 and its argument is bounded by:

$$\phi_1 \leq \arg(\mu_f) \leq \phi_2. \tag{3.7}$$

By the continuity of the linear functionals, for any $d > 0$, there exists a neighborhood $\mathcal{B}(f_0, \delta)$ such that for every $f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta)$,

$$f(x_1) \in \mathcal{D}\big(f_0(x_1), d\big) \quad \text{and} \quad \mu_f \in \mathcal{D}(0,r).$$

Then, by the equation:

$$\arg\big(f(x_1)\big) + \arg(\mu_f) = \arg\big(f(x_1 z_0 - x_0)\big) + \pi = \arg\big(f\left(P\left(z_0\right)\right)\big) + \pi,$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

for d small enough, there exist suitable angles $\theta_1, \theta_2 \in [0, 2\pi]$, with $\theta_2 - \theta_1 < \pi$, such that

$$\arg\big(f\left(P\left(z_0\right)\right)\big) \in [\theta_1, \theta_2], \quad \forall\, f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta).$$

By Proposition 3.2, it is clear that the origin is a corner of $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0); y\big)$. ∎

Application of Theorem 3.4 leads to a corresponding statement for general m-th degree vector-valued polynomials $P(z) = x_m z^m + x_{m-1} z^{m-1} + \cdots + x_1 z + x_0$, as in (2.4). We will also be making use of the notation $P'(z) = m x_m z^{m-1} + (m-1) x_{m-1} z^{m-2} + \cdots + 2x_2 z + x_1$.

THEOREM 3.5. *Let $P(z)$ be the m-th degree vector-valued polynomial in (2.4) and z_0 be a corner of $W^{\varepsilon}_{\|\cdot\|}\big(P(z); y\big)$. If there exists a linear functional $f_0 \in L_\varepsilon(y)$ such that $f_0\big(P(z_0)\big) = 0$, $f_0\big(P'(z_0)\big) \neq 0$, and $f_0(x_m) \neq 0$, then the origin is a corner of $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0); y\big)$.*

Proof. Defining the vector-valued polynomial

$$Q(z) = P(z + z_0) = x_m z^m + \cdots + P'(z_0) z + P(z_0),$$

the translation Property (P_{10}) yields

$$W^{\varepsilon}_{\|\cdot\|}\big(Q(z); y\big) = W^{\varepsilon}_{\|\cdot\|}\big(P(z + z_0); y\big) = W^{\varepsilon}_{\|\cdot\|}\big(P(z); y\big) - z_0,$$

whereby the origin is a corner of $W^{\varepsilon}_{\|\cdot\|}\big(Q(z); y\big)$. Since $f_0(x_m) \neq 0$ and by the continuity of the linear functionals of the set $L_\varepsilon(y)$, there exists a neighborhood $\mathcal{B}(f_0, \delta)$ for suitable $\delta > 0$, such that $f(x_m) \neq 0$ for every linear functional $f \in \mathcal{B}(f_0, \delta)$. Moreover, note that $f_0\big(Q(0)\big) = f_0\big(P(z_0)\big) = 0$ and the assumption $f_0\big(Q'(0)\big) = f_0\big(P'(z_0)\big) \neq 0$ implies that the origin is in fact a simple root of the equation:

$$f_0\big(Q(z)\big) = 0.$$

Hence, restricting our attention to $f \in \mathcal{B}(f_0, \delta) \cap \widehat{L}_\varepsilon(y)$, we denote $z_1(f)$, $z_2(f)$, …, $z_m(f)$ the roots of the m–th degree ($f(x_m) \neq 0$) polynomial equation:

$$f\big(Q(z)\big) = f(x_m) z^m + \cdots + f\big(P'(z_0)\big) z + f(P(z_0)) = 0,$$

indexed so that $z_m(f_0) = 0$ and the product:

$$pr_{m-1}(f) = z_1(f) \cdots z_{m-1}(f),$$

is nonzero for $f = f_0$. Note that $z_i(f)$ $(i = 1, 2, \ldots, m)$ are continuous functions of $f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \delta)$ and their product is (up to sign) given by the ratio:

$$\frac{f\big(P(z_0)\big)}{f(x_m)} = (-1)^m \prod_{j=1}^{m} z_j(f) = (-1)^m pr_{m-1}(f) z_m(f). \tag{3.8}$$

Regarding the argument of this quantity for f in a neighborhood of f_0, we proceed to obtain bounds for the arguments of individual factors $\arg\big(z_m(f)\big)$ and $\arg[(-1)^m pr_{m-1}(f)]$ separately. Since the origin is a corner of $W^{\varepsilon}_{\|\cdot\|}\big(Q(z); y\big)$, there exist radii $\rho < \delta$ and $\eta > 0$ such that for every $f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \rho)$, the corresponding Birkhoff–James ε-orthogonality set point:

$$z_m(f) \in W^{\varepsilon}_{\|\cdot\|}\big(Q(z); y\big) \cap \mathcal{D}(0, \eta),$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

lies in a neighborhood of the origin and its argument satisfies the inclusion:

$$\arg\big(z_m(f)\big) \in [\phi_1, \phi_2], \quad \text{for suitable } \phi_1, \phi_2 \in [0, 2\pi] \text{ with } \phi_2 - \phi_1 < \pi. \tag{3.9}$$

Moreover, by the continuity of $pr_{m-1}(f)$, the radius $\eta > 0$ may be chosen to be small enough, so that

$$\arg[(-1)^m pr_{m-1}(f)] \in [\theta_1, \theta_2], \quad \forall\, f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \rho), \tag{3.10}$$

with $\theta_2 - \theta_1 < \pi - (\phi_2 - \phi_1)$. Then, combining (3.9) and (3.10), expression (3.8) leads to

$$\arg\left(\frac{f\big(P(z_0)\big)}{f(x_m)}\right) \in [\omega_1, \omega_2], \quad \forall\, f \in \widehat{L}_\varepsilon(y) \cap \mathcal{B}(f_0, \rho),$$

where $\omega_i = \phi_i + \theta_i$ $(i = 1, 2)$ satisfy $\omega_2 - \omega_1 = (\phi_2 - \phi_1) + (\theta_2 - \theta_1) < \pi$. This means that (3.7) holds true for the linear pencil $x_m z - P(z_0)$. Invoking the subsequent arguments in the proof of Theorem 3.4, we conclude that the origin is a corner of $F^{\varepsilon}_{\|\cdot\|}\big(P(z_0); y\big) = F^{\varepsilon}_{\|\cdot\|}\big(Q(0); y\big)$. ∎

The following example demonstrates Theorem 3.5 in action.

Example 3.6. Consider the normed linear space $(\mathcal{X}, \|\cdot\|) = (\mathbb{C}^3, \|\cdot\|_\infty)$ and the vector-valued polynomial:

$$P(z) = \begin{bmatrix} 3 \\ 2 \\ 6 \end{bmatrix} z^2 + \begin{bmatrix} -1 \\ 1 \\ \frac{1}{2} \\ -2 \end{bmatrix} z + \begin{bmatrix} -2 \\ -2 \\ -5 \end{bmatrix}.$$

The corresponding ε-set of $P(z)$ with respect to the unit vector $y = \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$ for $\varepsilon = \dfrac{\sqrt{3}}{2}$ is depicted in Figure 1. Evidently, the boundary point 1 in the rightmost component is a corner of $W^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}\big(P(z); y\big)$.

Now, considering the (unit) vector $P(1) = \begin{bmatrix} 0 \\ 1 \\ \frac{1}{2} \\ -1 \end{bmatrix}$, note that

$$\|y + \lambda P(1)\|_\infty = \left\| \begin{bmatrix} \frac{1}{2} \\ \frac{\lambda}{2} \\ 1 - \lambda \end{bmatrix} \right\|_\infty = \max\left\{\frac{1}{2}, \frac{|\lambda|}{2}, |1 - \lambda|\right\}$$
$$\geq \frac{1}{2} = \|y\|_\infty \sqrt{1 - \left(\frac{\sqrt{3}}{2}\right)^2}, \quad \forall\, \lambda \in \mathbb{C}.$$

Equivalently, $0 \in F^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}(P(1); y)$. In fact, for $\lambda = 1$, we get the equality $\|y + P(1)\|_\infty = \dfrac{1}{2}$, and thus, $0 \in \partial F^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}(P(1); y)$, confirming Property (P_{15}). The corresponding set is depicted in Figure 2, where 0 is a corner of $F^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}\big(P(1); y\big)$, verifying Theorem 3.5.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

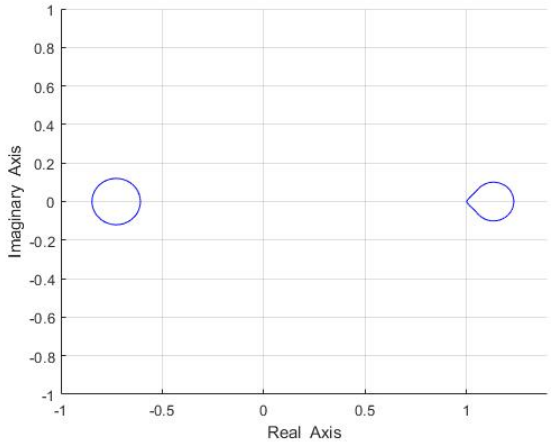


FIGURE 1. *The point* 1 *is a corner of* $W^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}(P(z);y)$.

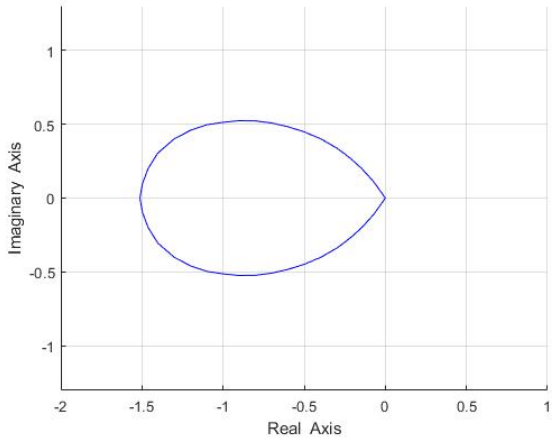


FIGURE 2. *The origin is a corner of* $F^{\frac{\sqrt{3}}{2}}_{\|\cdot\|_\infty}(P(1);y)$.

4. A randomized algorithm. Numerical determination of the Birkhoff–James ε-orthogonality set $F^{\varepsilon}_{\|\cdot\|}(x;y)$ is straightforward enough by its characterization (2.2) as an infinite intersection of closed disks. Indeed, for most applications, it suffices to plot some hundreds of circles $\partial\mathcal{D}\left(\lambda, \dfrac{\|x-\lambda y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}\right)$ centered at random points $\lambda \in \mathbb{C}$ on the complex plane and then estimate $F^{\varepsilon}_{\|\cdot\|}(x;y)$ as the region which remains unshaded.

On the other hand, estimating the set $W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ of an m-th degree vector-valued polynomial $P(z)$ in (2.4) is a much more daunting task. As suggested by the definition of this set in (2.5), it is crucial to develop an efficient procedure to check for $\mu \in \mathbb{C}$ whether inclusion $0 \in F^{\varepsilon}_{\|\cdot\|}(P(\mu);y)$ holds true. In this direction, denoting for $x, y \in \mathcal{X}$ $(y \neq 0)$ the quantity:

$$d(x,y) = \max_{\lambda\in\mathbb{C}} \left\{ |\lambda| - \frac{\|x-\lambda y\|}{\sqrt{1-\varepsilon^2}\,\|y\|} \right\},$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

we immediately obtain the equivalences:

$$\begin{aligned} d(x,y) < 0 &\;\Leftrightarrow\; 0 \in \operatorname{Int}\left(F^{\varepsilon}_{\|\cdot\|}(x;y)\right), \\ d(x,y) = 0 &\;\Leftrightarrow\; 0 \in \partial F^{\varepsilon}_{\|\cdot\|}(x;y), \\ d(x,y) > 0 &\;\Leftrightarrow\; 0 \notin F^{\varepsilon}_{\|\cdot\|}(x;y). \end{aligned}$$

Hence, the sign of $d(x,y)$ is indicative of the relative position of the origin with respect to the boundary of the set $F^{\varepsilon}_{\|\cdot\|}(x;y)$ and may be used to characterize $W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ as:

$$W^{\varepsilon}_{\|\cdot\|}(P(z);y) = \{\mu \in \mathbb{C} : d(P(\mu),y) \leq 0\}.$$

Moreover, under the assumption $F^{\varepsilon}_{\|\cdot\|}(x_m;y) \neq \{0\}$, Property (P_{15}) implies the inclusion:

$$\partial W^{\varepsilon}_{\|\cdot\|}(P(z);y) \subseteq \left\{\mu \in \mathbb{C} : 0 \in \partial F^{\varepsilon}_{\|\cdot\|}(P(\mu);y)\right\},$$

whereby a numerical approximation of $W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ is obtained by plotting the 0-level-set:

$$\{\mu \in \mathbb{C} : d\,(P(\mu),y) = 0\}.$$

This observation leads us to Algorithm 1.

Algorithm 1 Standard grid method for $W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ estimation

Input: $x_m, x_{m-1}, \ldots, x_1, x_0, y \in \mathcal{X}$, $\varepsilon \in [0,1)$
$[s_{\min}, s_{\max}]$: range of real parts of μ
$[t_{\min}, t_{\max}]$: range of imaginary parts of μ
Construct a grid $\Gamma \subset \Omega = [s_{\min}, s_{\max}] \times [t_{\min}, t_{\max}]$;
for *every* $\mu \in \Gamma$ **do**
  Compute $P(\mu) = x_m\mu^m + x_{m-1}\mu^{m-1} + \cdots + x_2\mu^2 + x_1\mu + x_0$;
  $\eta = \dfrac{\|2P(\mu)\|}{\|y\|}$;
  Construct a grid $\Gamma_\mu \subset \Omega_\mu = [-\eta, \eta] \times [-\eta, \eta]$;
  $D = [\ \ ]$;
  for *every* $\zeta \in \Gamma_\mu$ **do**
    $r = \dfrac{\|P(\mu) - \zeta y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}$;
    $D = \begin{bmatrix} D & |\zeta| - r \end{bmatrix}$;
  end
  Compute $d\,(P(\mu),y) = \max D$;
end
Draw the set $\{\mu \in \mathbb{C} : d\,(P(\mu),y) = 0\}$;

Algorithm 1 discretizes an initial region $\Omega = [s_{\min}, s_{\max}] \times [t_{\min}, t_{\max}] \supset W^{\varepsilon}_{\|\cdot\|}(P(z);y)$ and estimates $d\,(P(\mu),y)$ for every complex scalar $\mu \in \Gamma$ on the corresponding grid $\Gamma \subset \Omega$. This task necessitates the introduction of another grid Γ_μ for each $\mu \in \Gamma$. The computation and storage of the relevant data greatly increases memory requirements and markedly compromises the efficiency of the procedure.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

A substantially more efficient approach would be to introduce randomization at both levels: initially for sampling μ in the region $\Omega \supset W^{\varepsilon}_{\|\cdot\|}(P(z); y)$ and then for selecting the random circles $\partial\mathcal{D}\left(\zeta, \dfrac{\|P(\mu) - \zeta y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}\right)$ used to estimate $d\,(P(\mu), y)$, without ever forming the corresponding grids on Ω, Ω_μ. This approach is outlined in Algorithm 2.

Algorithm 2 Randomized method for $W^{\varepsilon}_{\|\cdot\|}(P(z); y)$ estimation

Input: $x_m, x_{m-1}, \ldots, x_1, x_0, y \in \mathcal{X}$, $\varepsilon \in [0, 1)$
$[s_{\min}, s_{\max}]$: range of real parts of μ
$[t_{\min}, t_{\max}]$: range of imaginary parts of μ
$n_1 \in \mathbb{N}$: the number of random points μ for which $d\,(P(\mu), y)$ is computed
$n_2 \in \mathbb{N}$: the number of random points ζ used to estimate $d\,(P(\mu), y)$ for each μ
for $k = 1 : n_1$ **do**
  $s = (s_{\max} - s_{\min}) \cdot rand + s_{\min}$ with $rand \in [0, 1]$ a random number;
  $t = (t_{\max} - t_{\min}) \cdot rand + t_{\min}$ with $rand \in [0, 1]$ a random number;
  $\mu = s + \mathbf{i}\, t$;
  Compute $P(\mu) = x_m\mu^m + x_{m-1}\mu^{m-1} + \cdots + x_2\mu^2 + x_1\mu + x_0$;
  $\eta = \dfrac{\|2P(\mu)\|}{\|y\|}$;
  $D = [\ ]$;
  for $\ell = 1 : n_2$ **do**
    $w = 2\eta \cdot rand + \eta$ with $rand \in [0, 1]$ a random number;
    $v = 2\eta \cdot rand + \eta$ with $rand \in [0, 1]$ a random number;
    $\zeta = w + \mathbf{i}\, v$;
    $r = \dfrac{\|P(\mu) - \zeta y\|}{\sqrt{1-\varepsilon^2}\,\|y\|}$;
    $D = \begin{bmatrix} D & |\zeta| - r \end{bmatrix}$;
  end
  Compute $d\,(P(\mu), y) = \max D$;
end
Draw the set $\{\mu \in \mathbb{C} : d\,(P(\mu), y) = 0\}$;

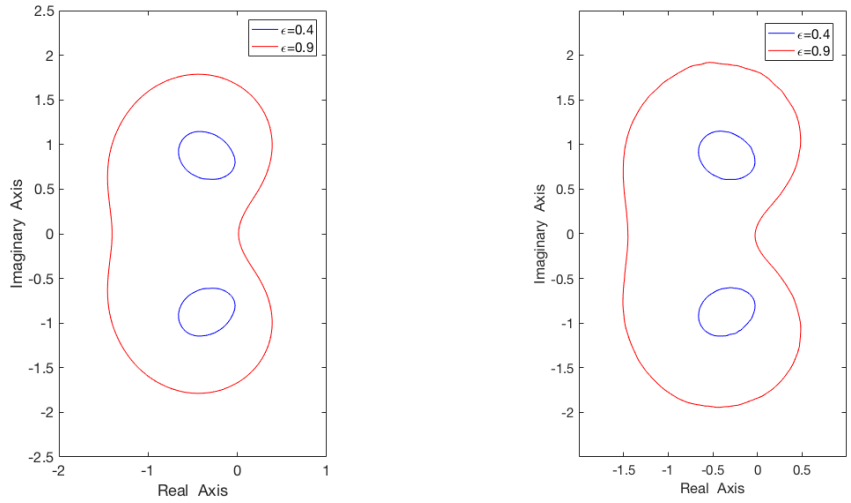
The following numerical experiments were implemented in a python environment and were performed on an 1.7 GHz Intel Core i7 processor with 8 GB of RAM.

Example 4.1. Consider the normed linear space $(\ell^1(\mathbb{N}), \|\cdot\|_1)$ and the vector-valued polynomial:

$$P(z) = \left(\frac{1}{4^n}\right)_{n=0}^{\infty} z^2 + \left[\left(\frac{0.8}{3^n}\right)_{n=0}^{\infty} \otimes e_1\right] z + \left(\frac{0.6}{2^n}\right)_{n=0}^{\infty},$$

with $e_1 \in \mathbb{C}^5$ the standard basis vector. Figure 3(a) depicts the set $W^{\varepsilon}_{\|\cdot\|_1}(P(z); y)$ with respect to the vector $y = \left(\frac{1}{3^n}\right)_{n=0}^{\infty} \in \ell^1(\mathbb{N})$ for the parameter pair $\varepsilon = 0.4, 0.9$ and was obtained via Algorithm 1 using 100×100 grids Γ and Γ_μ on $\Omega = [-2, 1] \times [-2.5, 2.5]$ and the auxiliary regions Ω_μ ($\mu \in \Gamma$), respectively. This procedure took 3248.83 s to complete. Algorithm 2 with $n_1 = 10000$ random points in the same region and then $n_2 = 800$ points for each $d(P(\mu), y)$-estimation yields the remarkably similar Figure 3(b) after 291.70 s, that is, a speedup factor of 11.13 was observed.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.



(a) Algorithm 1 (b) Algorithm 2

FIGURE 3. *Comparison of ε-set estimations for a vector-valued polynomial with coefficients in $\ell^1(\mathbb{N})$ with respect to some vector $y \in \ell^1(\mathbb{N})$, the norm $\|\cdot\| = \|\cdot\|_1$, and the pair of parameters $\varepsilon = 0.4, 0.9$ using the grid and the randomized methods (100×100 grids and $(n_1, n_2) = (10000, 800)$ random points in the regions $\Omega = [-2, 1] \times [-2.5, 2.5]$ and Ω_μ, respectively).*

EXAMPLE 4.2. We consider the 64×64 quartic matrix polynomial:

$$P(z) = M_4 z^4 + M_3 z^3 + M_2 z^2 + M_1 z + M_0,$$

with alternating coefficients:

$$M_i = I_8 \otimes \hat{M}_i + \hat{M}_i \otimes I_8 \quad (i = 0, 1, 2, 3, 4),$$

defined by

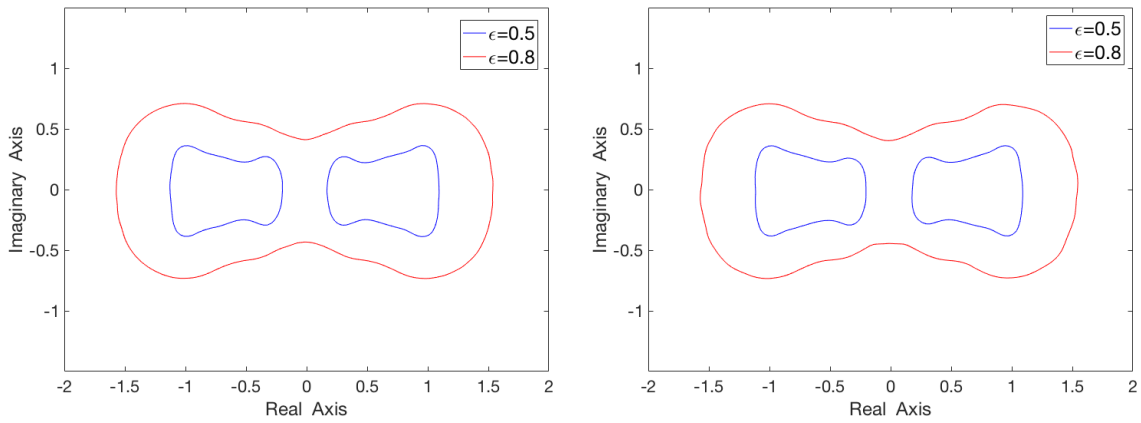
$$\hat{M}_0 = \frac{1}{6}\left(4I_8 + N + N^T\right), \quad \hat{M}_1 = \hat{M}_3 = N - N^T, \quad \hat{M}_2 = -\hat{M}_4 = -\left(2I_8 - N - N^T\right),$$

where N is the 8×8 nilpotent matrix having ones in its subdiagonal and zeros elsewhere [25]. Visualizations of the set $W^{\varepsilon}_{\|\cdot\|_\infty}(P(z); M_4)$ for the parameter pair $\varepsilon = 0.5, 0.8$ using the two approaches on the initial region on $\Omega = [-2, 2] \times [-1.5, 1.5]$ can be found in Figure 4. Algorithm 1 (using 100×100 grids) took 2237.01 s to complete, while Algorithm 2 (with $(n_1, n_2) = (1000, 140)$ random points) performed considerably faster (35.68 s), speeding up the procedure by a factor of 62.71.

Now let the rectangular matrix polynomial $Q(z) = \tilde{M}_4 z^4 + \tilde{M}_3 z^3 + \tilde{M}_2 z^2 + \tilde{M}_1 z + \tilde{M}_0$, where $\tilde{M}_j \in \mathbb{C}^{32 \times 64}$ is the submatrix of M_j $(j = 0, 1, \ldots, 4)$ consisting of its leading 32 rows. Repeating the procedure for $W^{\varepsilon}_{\|\cdot\|_\infty}(Q(z); \tilde{M}_3)$ $(\varepsilon = 0.5, 0.8)$, Algorithm 2 completed the visualizations in 29.85 s, performing 56.41 times faster than Algorithm 1 (1684.17 s). The corresponding plots can be found in Figure 5.

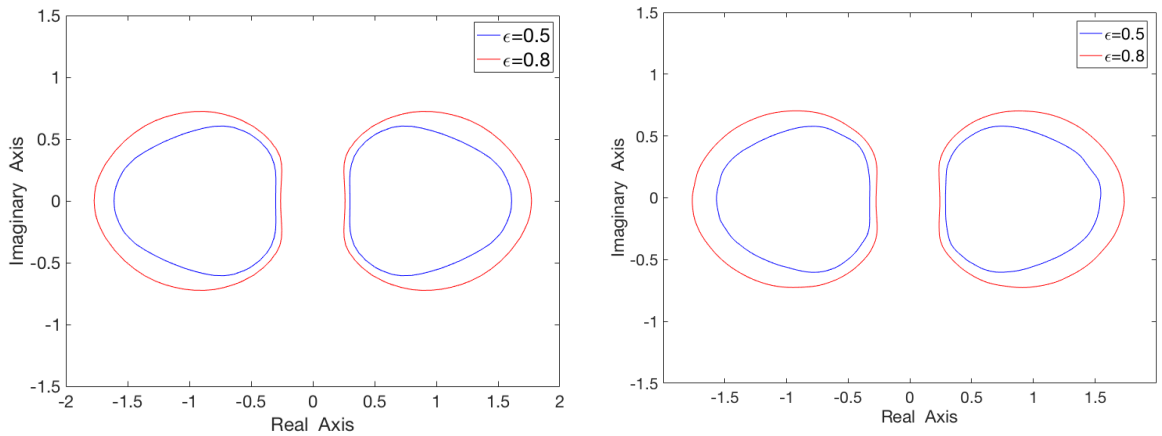
We conclude this example, verifying that Algorithms 1 and 2 approximate the boundary of the standard numerical range of a square matrix polynomial (or some square matrix), when applied for $\varepsilon = 0$, $\|\cdot\| = \|\cdot\|_2$ and $\psi = \chi_m = I_n$. To compare with the inclusion–exclusion procedure for numerical range estimation for monic matrix polynomials in [29], consider $R(z) = M_4^{-1} P(z)$. Approximations of $W^0_{\|\cdot\|_2}(R(z); I_{64}) = W(R)$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.



(a) Algorithm 1 (b) Algorithm 2

FIGURE 4. *Comparison of ε-set estimations for a quartic matrix polynomial with respect to its leading coefficient, the norm $\|\cdot\| = \|\cdot\|_\infty$, and the pair of parameters $\varepsilon = 0.5, 0.8$ using the grid and the randomized methods (100×100 grids and $(n_1, n_2) = (2500, 140)$ random points in the regions $\Omega = [-2, 2] \times [-1.5, 1.5]$ and Ω_μ, respectively).*

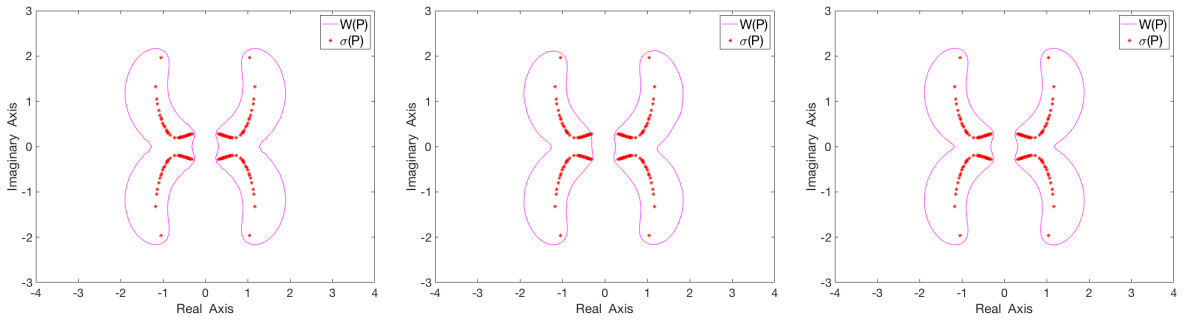


(a) Algorithm 1 (b) Algorithm 2

FIGURE 5. *Comparison of ε-set estimations for $Q(z)$ ($\varepsilon = 0.5, 0.8$) using the grid and the randomized methods (100×100 grids and $(n_1, n_2) = (2500, 140)$ random points in the regions $\Omega = [-2, 2] \times [-1.5, 1.5]$ and Ω_μ, respectively).*

via the standard grid (50×50 grids on $\Omega = [-4.4] \times [-3, 3]$ and then on Ω_μ) and the randomized approaches ($n_1 = n_2 = 2500$ random points on each region Ω and Ω_μ) can be found in Figures 6(a) and (b), respectively. Both of these estimations in fact virtually coincide with the visualization in Figure 6(c) obtained via the procedure in [29] (using grid lengths $h_x = h_y = 0.01$); Algorithm 2 (252.80 s) proved to be the most efficient approach, completing the visualization 20.82 times faster than Algorithm 1 (5263.30 s) and 12.95 times faster than the standard inclusion–exclusion procedure in [29] (3272.39 s). It is noteworthy that Algorithms 1 and 2 resulted in remarkably similar output to [29], in spite of using much sparser grids/less random points on

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.



(a) Algorithm 1 (b) Algorithm 2 (c) Inclusion–exclusion procedure in [29]

FIGURE 6. *Numerical range visualization* ($W(P) = W^0_{\|\cdot\|_2}(P(z); I_{64})$) *via the standard grid and randomized methods, as well as the inclusion–exclusion procedure in* [29].

Ω, respectively. As expected, application of Algorithms 1 and 2 for $\|\cdot\| = \|\cdot\|_2$ is far more computationally intensive than for $\|\cdot\| = \|\cdot\|_\infty$. Contrary to [29], an additional advantage of the present method is that Algorithms 1 and 2 remain relevant in the general non-monic matrix polynomial case.

EXAMPLE 4.3. Consider the harmonic wave equation $\Delta p - \left(\frac{2\pi f}{c}\right)^2 p = 0$ for the acoustic pressure p in the domain $\Omega = [0,1] \times [0,1]$ with Dirichlet boundary conditions on the sides $x = 0$, $y = 0$, $y = 1$ and impedance condition $\frac{\partial p}{\partial n} + \frac{2\pi \mathbf{i} f}{\zeta} p = 0$ on the right boundary ($x = 1$), as studied in [4]. The scalar parameters f, c, ζ denote frequency, speed of sound in the medium, and impedance, respectively. Discretization with mesh size $h = 1/45$ and $f = c = \zeta = 1$ leads to a quadratic matrix polynomial $P(z) = z^2 M + zD + K$ of order 1980 $\left(= \frac{1}{h}\left(\frac{1}{h} - 1\right)\right)$ with mass, damping, and stiffness matrices given by:

$$M = -4\pi^2 h^2 I_{44} \otimes \left(I_{45} - \frac{1}{2} e_{45} e_{45}^T\right), \quad D = \frac{2\pi \mathbf{i} h}{\zeta} I_{44} \otimes \left(e_{45} e_{45}^T\right), \quad K = I_{44} \otimes C + T \otimes \left(-I_{45} + \frac{1}{2} e_{45} e_{45}^T\right),$$

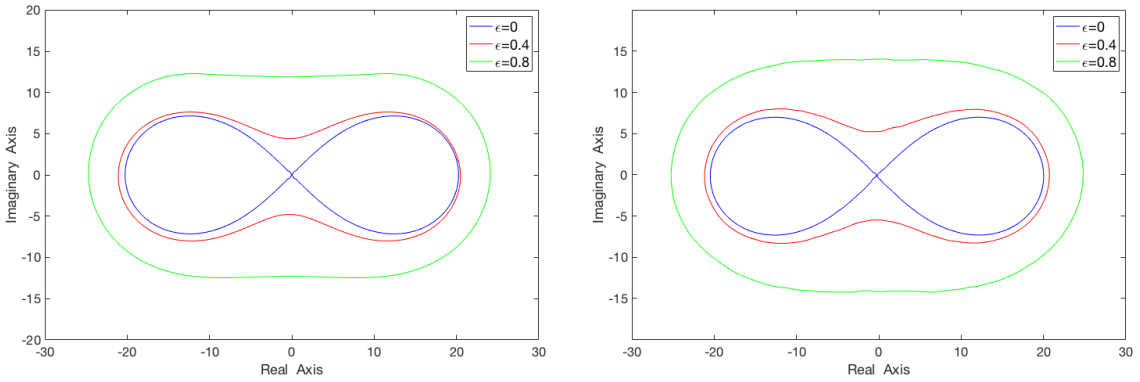
where

$$C = \begin{bmatrix} 4 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & 4 & -1 \\ & & -1 & 2 \end{bmatrix} \in \mathbb{R}^{45\times 45}, \quad T = \text{tridiag}\{1, 0, 1\} \in \mathbb{R}^{44\times 44},$$

and e_{45} is the last column of I_{45}.

Approximations of the set $W^\varepsilon_{\|\cdot\|_\infty}(P(z); M)$ for the parameter triplet $\varepsilon = 0, 0.4, 0.8$ using the grid and randomized approaches are plotted in Figure 7. Computations in both cases were restricted to the region $\mathcal{R} = [-30, 30] \times [-20, 20]$, using 100×100 grids and $(n_1, n_2) = (10000, 280)$ random points, respectively. Algorithm 2 took 1508.27 s to complete, while the far more computationally intensive Algorithm 1 required as many as 52569.89 s. This amounts to significant computational savings and a speedup factor of 34.85.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.



(a) Algorithm 1 (b) Algorithm 2

FIGURE 7. *Comparison of ε-set estimations for a quadrartic matrix polynomial arising in acoustics and the parameter triplet $\varepsilon = 0, 0.4, 0.8$ using the grid and the randomized methods.*

Acknowledgment. This research is carried out/funded in the context of the project 'Approximation algorithms and randomized methods for large-scale problems in computational linear algebra' (MIS 5049095) under the call for proposals '*Researchers' support with an emphasis on young researchers-2nd Cycle*'. The project is co-financed by Greece and the European Union (European Social Fund-ESF) by the Operational Programme Human Resources Development, Education and Lifelong Learning 2014–2020.

REFERENCES

[1] G. Birkhoff. Orthogonality in linear metric spaces. *Duke Math. J.*, 1:169–172, 1935.
[2] F.F. Bonsall and J. Duncan. *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*. London Math. Soc. Lecture Note Ser., Cambridge University Press, New York, 1971.
[3] F.F. Bonsall and J. Duncan. *Numerical Ranges II*. London Math. Soc. Lecture Note Ser., Cambridge University Press, New York, 1973.
[4] F.C. Chaitin–Chatelin and M.B. van Gijzen. Analysis of parameterized quadratic eigenvalue problems in computational acoustics with homotopic deviation theory. *Numer. Linear Algebra Appl.*, 13:487–512, 2006.
[5] J. Chmieliński. On an ε-Birkhoff orthogonality. *J. Inequal. Pure Appl. Math.*, 6:Article no. 79, 2005.
[6] Ch. Chorianopoulos, S. Karanasios, and P. Psarrakos. A definition of numerical range of rectangular matrices. *Linear Multilinear Algebra*, 57:459–475, 2009.
[7] Ch. Chorianopoulos and P. Psarrakos. Birkhoff-James approximate orthogonality sets and numerical ranges. *Linear Algebra Appl.*, 434:2089–2108, 2011.
[8] S.S. Dragomir. On approximation of continuous linear functionals in normed linear spaces. *An. Univ. Timişoara Ser. Ştiinţ. Mat.*, 29:51–58, 1991.
[9] F.R. Gantmacher. *The Theory of Matrices*. Chelsea Publishing Company, New York, 1959.
[10] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982.
[11] I. Gohberg, P. Lancaster, and L. Rodman. Spectral analysis of selfadjoint matrix polynomials. *Ann. Math.*, 112:31–71, 1980.
[12] I. Gohberg, P. Lancaster, and L. Rodman. Representation and divisibility of operator polynomials. *Canad. J. Math.*, 30: 1045–1069, 1978.
[13] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1990.
[14] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
[15] N. Ito and H.K. Wimmer. Self-inversive Hilbert space operator polynomials with spectrum on the unit circle. *J. Math. Anal. Appl.*, 436:683–691, 2016.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 38, pp. 32-48, January 2022.

[16] R.C. James. Orthogonality and linear functionals in normed linear spaces. *Trans. Amer. Math. Soc.*, 61:265–292, 1947.
[17] T. Kailath. *Linear Systems*. Prentice Hall, Englewood Cliffs, 1980.
[18] M. Karamanlis and P.J. Psarrakos. Birkhoff-James ϵ-orthogonality sets in normed linear spaces. *Textos Mat.*, University of Coimbra, 44:81–92, 2013.
[19] P. Lancaster. *Lambda-matrices and Vibrating Systems*. Pergamon Press, Oxford, 1966.
[20] P. Lancaster and P. Psarrakos. Normal and seminormal eigenvalues of analytic matrix functions. *Integral Equations Operator Theory*, 41:331–342, 2001.
[21] L. Lerer, L. Rodman, and M. Tismenetsky. Bezoutian and Schur-Cohn problem for operator polynomials. *J. Math. Anal. Appl.*, 103:83–102, 1984.
[22] C-K. Li and L. Rodman Numerical range of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 15:1256–1265, 1994.
[23] A.S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*. Translations of Mathematical Monographs, Vol. 71. American Mathematical Society, Providence, 1988.
[24] J. Maroulas and P. Psarrakos. The boundary of numerical range of matrix polynomials. *Linear Algebra Appl.*, 267:101–111, 1997.
[25] V. Mehrmann and D. Watkins. Polynomial eigenvalue problems with Hamiltonian structure. *Electron. Trans. Numer. Anal.*, 13:106–118, 2002.
[26] H. Nakazato and P. Psarrakos. On the shape of numerical range of matrix polynomials. *Linear Algebra Appl.*, 338:105–123, 2001.
[27] V. Panagakou, P. Psarrakos, and N. Yannakakis. Birkhoff-James epsilon-orthogonality sets of vectors and vector-valued polynomials. *J. Math. Anal. Appl.*, 454:59–78, 2017.
[28] I.B. Prolla. *Approximation of Vector Valued Functions*. Notas de Matematica, Vol. 62, North-Holland Publishing Company, Amsterdam, 1977.
[29] P. Psarrakos. On the estimation of the q–numerical range of monic matrix polynomials. *Electron. Trans. Numer. Anal.*, 17:1–10, 2004.
[30] J.G. Stampfli and J.P. Williams. Growth conditions and the numerical range in a Banach algebra. *Tôhoku Math. J.*, 20:417–424, 1968.