GERSHGORIN TYPE SETS FOR EIGENVALUES OF MATRIX POLYNOMIALS*

CHRISTINA MICHAILIDOU† AND PANAYIOTIS PSARRAKOS†

Abstract. New localization results for polynomial eigenvalue problems are obtained, by extending the notions of the Gershgorin set, the generalized Gershgorin set, the Brauer set and the Dashnic-Zusmanovich set to the case of matrix polynomials.

Key words. Eigenvalue, Matrix polynomial, The Gershgorin set, The generalized Gershgorin set, The Brauer set, The Dashnic-Zusmanovich set.

AMS subject classifications. 15A18, 15A22.

1. Introduction. For an arbitrary square complex matrix $A \in \mathbb{C}^{n\times n}$ with standard spectrum $\sigma(A)$, the celebrated Gershgorin (Geršgorin) circle theorem [8, 21] implies n easily computable disks, in the complex plane, centered at the diagonal entries of the matrix, whose union contains $\sigma(A)$. The simplicity and the applications of the Gershgorin circle theorem have inspired further research in this area, resulting in hundreds of papers on Gershgorin disks and related sets such as the generalized Gershgorin set (known also as the $\mathcal{A}$-Ostrowski set), the Brauer set, the Dashnic-Zusmanovich set and others (see [3, 5, 6, 18, 21] and the references therein), which are widely used in the analysis of ordinary eigenvalue problems.

In this paper, we consider $n \times n$ *matrix polynomials* of the form

$$P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \cdots + A_1\lambda + A_0, \tag{1.1}$$

where λ is a complex variable, $A_0, A_1, \ldots, A_m \in \mathbb{C}^{n\times n}$ with $A_m \neq 0$, and the determinant $\det P(\lambda)$ *is not identically zero.* The study of matrix polynomials, especially with regard to their spectral analysis and the solution of higher order linear differential (or difference) systems with constant coefficients, has a long history and important applications; see [9, 14, 15, 16, 17] and the references therein.

A scalar $\mu \in \mathbb{C}$ is called an *eigenvalue* of $P(\lambda)$ if the system $P(\mu)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$. This solution x_0 is known as an *eigenvector* of $P(\lambda)$ corresponding to the eigenvalue μ. The set of all finite eigenvalues of $P(\lambda)$,

$$\sigma(P) = \{\mu \in \mathbb{C} : \det P(\mu) = 0\} = \{\mu \in \mathbb{C} : 0 \in \sigma(P(\mu))\},$$

is the *finite spectrum* of $P(\lambda)$. The *algebraic multiplicity* of an eigenvalue $\mu \in \sigma(P)$ is the multiplicity of μ as a root of the (scalar) polynomial $\det P(\lambda)$, and it is always greater than or equal to the *geometric multiplicity* of μ, that is, the dimension of the null space of matrix $P(\mu)$. Furthermore, it is said that $\mu = \infty$ is an eigenvalue of $P(\lambda)$ exactly when 0 is an eigenvalue of the *reverse* matrix polynomial

$$\hat{P}(\lambda) = \lambda^m P(1/\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_{m-1}\lambda + A_m.$$

In this case, the algebraic multiplicity and the geometric multiplicity of the eigenvalue $\mu = \infty$ of $P(\lambda)$ are defined as the algebraic multiplicity and the geometric multiplicity of the eigenvalue 0 of $\hat{P}(\lambda)$, respectively.

*Corresponding Author: Panayiotis Psarrakos.

†Department of Mathematics, School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece (xristina.michailidou@gmail.com, ppsarr@math.ntua.gr). The research of the first author was supported by the Research Committee of the National Technical University of Athens (ELKE PhD Scholarships).

In the next sections, we extend the notions of the Gershgorin set (Section 2), the generalized Gershgorin set (Section 3), the Brauer set (Section 4) and the Dashnic-Zusmanovich set (Section 5) to the case of matrix polynomials; see also [1, 12, 13]. In particular, in each section, we define an eigenvalues' inclusion set, present basic topological and geometrical properties of this set, and give illustrative examples to verify our results. Our approach is direct, does not involve any kind of linearizations, and it is motivated by the following simple observation: Suppose that F is a set-valued function of $n \times n$ matrices which associates each matrix $A \in \mathbb{C}^{n\times n}$ to a region $F(A) \subseteq \mathbb{C}$ that contains the spectrum $\sigma(A)$. If, for an $n \times n$ matrix polynomial $P(\lambda)$ as in (1.1), we define the set $F(P) = \{\mu \in \mathbb{C} : 0 \in F(P(\mu))\}$, then $\sigma(P) = \{\mu \in \mathbb{C} : 0 \in \sigma(P(\mu))\} \subseteq \{\mu \in \mathbb{C} : 0 \in F(P(\mu))\} = F(P)$.

2. The Gershgorin set.

2.1. Definitions. Consider a square complex matrix $A \in \mathbb{C}^{n\times n}$, define the set $\mathcal{N} = \{1, 2, \dots, n\}$, and let $(A)_{i,j}$ denote the (i, j)-th entry of A, $i, j \in \mathcal{N}$. For any $i \in \mathcal{N}$, define also the nonnegative quantity $r_i(A) = \sum_{j \in \mathcal{N}\setminus\{i\}} |(A)_{i,j}|$ and the *i-th row Gershgorin disk* $G_i(A) = \{\mu \in \mathbb{C} : |\mu - (A)_{i,i}| \le r_i(A)\}$. The Gershgorin circle theorem [8, 21] states that the spectrum $\sigma(A)$ lies in the union $G(A) = \bigcup_{i \in \mathcal{N}} G_i(A)$, which is known as the *Gershgorin set* of A.

Consider now an $n \times n$ matrix polynomial $P(\lambda)$ as in (1.1), and define the nonnegative functions

$$r_i(P(\lambda)) = \sum_{j \in \mathcal{N}\setminus\{i\}} |(P(\lambda))_{i,j}|, \quad i \in \mathcal{N},$$

the *i-th row Gershgorin sets* of $P(\lambda)$

$$G_i(P) = \{\mu \in \mathbb{C} : 0 \in G_i(P(\mu))\} = \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| \le r_i(P(\mu))\}, \quad i \in \mathcal{N},$$

and the *Gershgorin set* of $P(\lambda)$

$$G(P) = \{\mu \in \mathbb{C} : 0 \in G(P(\mu))\} = \bigcup_{i \in \mathcal{N}} G_i(P).$$

It is easy to see that

$$G_i(\hat{P})\setminus\{0\} = \left\{\mu \in \mathbb{C}\setminus\{0\} : 0 \in G_i\left(P\left(\mu^{-1}\right)\right)\right\} = \left\{\mu \in \mathbb{C}\setminus\{0\} : \mu^{-1} \in G_i(P)\right\}, \quad i \in \mathcal{N},$$

and

$$G(\hat{P})\setminus\{0\} = \left\{\mu \in \mathbb{C}\setminus\{0\} : 0 \in G\left(P\left(\mu^{-1}\right)\right)\right\} = \left\{\mu \in \mathbb{C}\setminus\{0\} : \mu^{-1} \in G(P)\right\}.$$

As in the case of the eigenvalues of $P(\lambda)$, we say that $\mu = \infty$ lies in $G_i(P)$ (resp., in $G(P)$) exactly when 0 lies in $G_i(\hat{P})$ (resp., in $G(\hat{P})$). The Gershgorin circle theorem [8, 21] is directly generalized to the case of matrix polynomials.

THEOREM 2.1. *All the (finite and infinite) eigenvalues of the matrix polynomial $P(\lambda)$ lie in the Gershgorin set $G(P)$.*

Proof. For any finite eigenvalue $\mu \in \sigma(P)$, it is apparent that $0 \in \sigma(P(\mu)) \subseteq G(P(\mu))$, and hence, μ lies in $G(P)$. Moreover, if $\mu = \infty$ is an eigenvalue of $P(\lambda)$, then $0 \in \sigma(\hat{P}) \subseteq G(\hat{P})$, and consequently, $\mu = \infty$ lies in $G(P)$. ∎

By the original proof of the Gershgorin circle theorem [8] (see also [21, Theorem 1.1]), it follows that if λ_0 is an eigenvalue of a matrix $A \in \mathbb{C}^{n\times n}$ and $x_0 = [\,x_1 \;\; x_2 \;\; \cdots \;\; x_n\,]^T$ is an eigenvector of A corresponding to λ_0, with $0 < |x_k| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$ for some $k \in \mathcal{N}$, then λ_0 lies in the k-th row Gershgorin disk $G_k(A)$. As a consequence, if μ_0 is an eigenvalue of a matrix polynomial $P(\lambda)$ and $x_0 = [\,x_1 \;\; x_2 \;\; \cdots \;\; x_n\,]^T$ is an eigenvector of $P(\lambda)$ corresponding to μ_0, with $0 < |x_k| = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$ for some $k \in \mathcal{N}$, then 0 is an eigenvalue of the matrix $P(\mu_0)$ with x_0 as an associated eigenvector, and thus, $0 \in G_k(P(\mu_0))$, or equivalently, $\mu_0 \in G_k(P)$.

2.2. Basic properties. For an $n \times n$ matrix polynomial $P(\lambda)$ as in (1.1), it is easy to verify the following properties.

Proposition 2.2. *Let $i \in \mathcal{N}$.*

(i) *$G_i(P)$ is a closed subset of $\mathbb{C}$.*
(ii) *For any scalar $b \in \mathbb{C}\backslash\{0\}$, consider the matrix polynomials $Q_1(\lambda) = P(b\lambda)$, $Q_2(\lambda) = bP(\lambda)$ and $Q_3(\lambda) = P(\lambda + b)$. Then $G_i(Q_1) = b^{-1}G_i(P)$, $G_i(Q_2) = G_i(P)$ and $G_i(Q_3) = G_i(P) - b$.*
(iii) *If the (i,i)-th entry of $P(\lambda)$, $(P(\lambda))_{i,i}$, is (identically) zero, then $G_i(P) = \mathbb{C}$.*
(iv) *If all the coefficient matrices $A_0, A_1, \ldots, A_m$ have their i-th rows real, then $G_i(P)$ is symmetric with respect to the real axis.*

Proof. (i) Consider a scalar $\mu \notin G_i(P)$. Then $|(P(\mu))_{i,i}| > r_i(P(\mu))$, and continuity implies that for any scalar $\hat{\mu}$ sufficiently close to μ, $|(P(\hat{\mu}))_{i,i}| > r_i(P(\hat{\mu}))$, i.e., $\hat{\mu} \notin G_i(P)$. Thus, the set $\mathbb{C}\backslash G_i(P)$ is open.

(ii) Recall that $\mu \in G_i(P)$ if and only if $0 \in G_i(P(\mu))$, and observe that

$$G_i(Q_1) = \{\mu \in \mathbb{C} : 0 \in G_i(P(b\mu))\} = \left\{\frac{\mu}{b} \in \mathbb{C} : 0 \in G_i(P(\mu))\right\},$$

$$G_i(Q_2) = \{\mu \in \mathbb{C} : 0 \in G_i(bP(\mu))\} = \{\mu \in \mathbb{C} : 0 \in G_i(P(\mu))\}$$

and

$$G_i(Q_3) = \{\mu \in \mathbb{C} : 0 \in G_i(P(\mu + b))\} = \{\mu - b \in \mathbb{C} : 0 \in G_i(P(\mu))\}.$$

(iii) If the (i,i)-th entry of $P(\lambda)$ is identically zero, then for any $\mu \in \mathbb{C}$, the origin is the center of the Gershgorin disk $G_i(P(\mu))$. Hence, the inequality $|(P(\mu))_{i,i}| \leq r_i(P(\mu))$ is satisfied trivially for all $\mu \in \mathbb{C}$.

(iv) Suppose that all the coefficient matrices $A_0, A_1, \ldots, A_m$ have their i-th rows real. If $\mu \in G_i(P)$, then

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\mu^k\right| \leq \sum_{j\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m}(A_k)_{i,j}\mu^k\right|,$$

or equivalently,

$$\left|\overline{\sum_{k=0}^{m}(A_k)_{i,i}\mu^k}\right| \leq \sum_{j\in\mathcal{N}\backslash\{i\}} \left|\overline{\sum_{k=0}^{m}(A_k)_{i,j}\mu^k}\right|,$$

or equivalently,

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\overline{\mu}^k\right| \leq \sum_{j\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m}(A_k)_{i,j}\overline{\mu}^k\right|.$$

This means that $\overline{\mu} \in G_i(P)$. □

PROPOSITION 2.3. *For any $i \in \mathcal{N}$, the set $\{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| < r_i(P(\mu))\}$ lies in the interior of $G_i(P)$; as a consequence,*

$$\partial G_i(P) \subseteq \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| = r_i(P(\mu))\} = \{\mu \in \mathbb{C} : 0 \in \partial G_i(P(\mu))\}.$$

Proof. Consider a scalar $\mu \in \mathbb{C}$ such that $|(P(\mu))_{i,i}| < r_i(P(\mu))$. By continuity, there is a real $\varepsilon > 0$ such that for every $\hat{\mu} \in \mathbb{C}$ with $|\mu - \hat{\mu}| \leq \varepsilon$, $|(P(\hat{\mu}))_{i,i}| \leq r_i(P(\hat{\mu}))$. Thus, μ is an interior point of $G_i(P)$. □

By applying the proof technique of Theorem 3.1 in [1], we obtain a necessary condition for the isolated points.

THEOREM 2.4. *If a scalar $\mu \in \mathbb{C}$ is an isolated point of $G_i(P)$, then μ is a common root of all polynomials $(P(\lambda))_{i,j}$, $j \in \mathcal{N}$; as a consequence, μ is an eigenvalue of $P(\lambda)$.*

Proof. For the sake of contradiction, assume that $(P(\mu))_{i,i} \neq 0$. Since μ is an isolated point of $G_i(P)$, it belongs to the boundary $\partial G_i(P)$, and there exists an $\varepsilon > 0$ such that the disk $D(\mu, \varepsilon) = \{\lambda \in \mathbb{C} : |\lambda - \mu| \leq \varepsilon\}$ contains no other point of $G_i(P)$. The i-th row Gershgorin set can be written as

$$\begin{aligned} G_i(P) &= \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| \leq r_i(P(\mu))\} \\ &= \left\{\mu \in \mathbb{C} : \frac{r_i(P(\mu))}{|(P(\mu))_{i,i}|} \geqslant 1\right\} \\ &= \left\{\mu \in \mathbb{C} : \log \frac{r_i(P(\mu))}{|(P(\mu))_{i,i}|} \geqslant 0\right\}. \end{aligned}$$

Define the function

$$\begin{aligned} \varphi(\lambda) &= \log \frac{r_i(P(\lambda))}{|(P(\lambda))_{i,i}|}, \quad \lambda \in D(\mu,\varepsilon) \\ &= \log \left\| \left[\frac{(P(\lambda))_{i,1}}{(P(\lambda))_{i,i}}, \ldots, \frac{(P(\lambda))_{i,i-1}}{(P(\lambda))_{i,i}}, \frac{(P(\lambda))_{i,i+1}}{(P(\lambda))_{i,i}}, \ldots, \frac{(P(\lambda))_{i,n}}{(P(\lambda))_{i,i}} \right] \right\|_1, \quad \lambda \in D(\mu,\varepsilon), \end{aligned} \tag{2.2}$$

and observe that it is subharmonic and satisfies the maximum principle [2, 4]. By definition, $\varphi(\lambda)$ is equal to zero on the boundary of $G_i(P)$, nonnegative in the interior of $G_i(P)$ and negative elsewhere; see Proposition 2.3. Since $\mu \in \partial G_i(P)$, it follows

$$|(P(\mu))_{i,i}| = r_i(P(\mu)) = \sum_{j \in \mathcal{N} \setminus \{i\}} |(P(\mu))_{i,j}|.$$

So, the function $\varphi(\lambda)$ is equal to zero at the center μ of $D(\mu, \varepsilon)$ and negative in the rest of the disk. Since $\varphi(\lambda)$ satisfies the maximum principle, it should take it's maximum value on the boundary of the disk; this is a contradiction. As a consequence,

$$0 = (P(\mu))_{i,i} = \sum_{j \in \mathcal{N} \setminus \{i\}} |(P(\mu))_{i,j}|,$$

and μ is a common root of all polynomials $(P(\lambda))_{i,j}$, $j \in \mathcal{N}$. □

Let Ω be a closed subset of $\mathbb{C}$, and let $\mu \in \Omega$. The *local dimension* of the point μ in Ω is defined as the limit $\lim_{h \to 0^+} \dim\{\Omega \cap D(\mu, h)\}$ $(h \in \mathbb{R},\ h > 0)$, where $\dim\{\cdot\}$ denotes the topological dimension [11]. Any isolated point in Ω has local dimension 0. Any non-isolated point in Ω has local dimension 2 if it belongs to the closure of the interior of Ω, and 1 otherwise.

THEOREM 2.5. *Any point of the i-th row Gershgorin set $G_i(P)$ has local dimension either 2 or 0 (isolated point); in other words, the Gershgorin set of $P(\lambda)$ cannot have parts with (nontrivial) curves.*

Proof. For the sake of contradiction, assume that the local dimension of a point $\mu \in G_i(P)$ is equal to 1. Then $\mu \in \partial G_i(P)$ and there is an $\varepsilon > 0$ such that

$$G_i(P) \cap \{\lambda \in \mathbb{C} : |\lambda - \mu| \leq \varepsilon\} \subset \partial G_i(P).$$

Thus, the function $\varphi(\lambda)$ in (2.2), defined in the disk $|\lambda - \mu| \leq \varepsilon$, takes its maximum value 0 in (infinitely many) interior points of the disk. Since $\varphi(\lambda)$ is subharmonic and satisfies the maximum principle, this is a contradiction. Hence, the local dimension of μ is either 2 or 0. ∎

Next we obtain necessary and sufficient conditions for the row Gershgorin sets to be bounded. For any $i \in \mathcal{N}$, we define the sets

$$\beta_i = \{j \in \mathcal{N} : (A_m)_{i,j} \neq 0\} \quad \text{and} \quad \overline{\beta}_i = \mathcal{N} \backslash \beta_i = \{j \in \mathcal{N} : (A_m)_{i,j} = 0\}. \tag{2.3}$$

It is worth mentioning that a scalar $\mu \in \mathbb{C}$ is a common root of all polynomials $(P(\lambda))_{i,1}, (P(\lambda))_{i,2}, \ldots, (P(\lambda))_{i,n}$ if and only if the i-th row of matrix $P(\mu)$ is zero, or equivalently, if and only if $G_i(P(\mu)) = \{0\}$. By Theorem 2.4, if the origin is an isolated point of the row Gershgorin set $G_i(P)$, then the i-th row of the coefficient matrix A_0 is zero. As a consequence, if $\beta_i \neq \emptyset$, then the origin is not an isolated point of $G_i(\hat{P})$, i.e., $G_i(P)$ is not the union of a bounded set and ∞.

THEOREM 2.6. *Suppose that for an $i \in \mathcal{N}$, $\beta_i \neq \emptyset$.*

(i) *If $i \in \beta_i$, then the i-th row Gershgorin set $G_i(P)$ is unbounded if and only if $0 \in G_i(A_m)$.*
(ii) *If $i \in \overline{\beta}_i$, then the i-th row Gershgorin set $G_i(P)$ is unbounded and $0 \in G_i(A_m)$.*

Proof. (i) Suppose that $i \in \beta_i$, i.e., $(A_m)_{i,i} \neq 0$.

Let $G_i(P)$ be unbounded. Since $\beta_i \neq \emptyset$, the origin is not an isolated point of $G_i(\hat{P})$ and there is a sequence $\{\mu_l\}_{l\in\mathbb{N}}$ in $G_i(P)\backslash\{0\}$ such that $|\mu_l| \to +\infty$. Then, for every $l \in \mathbb{N}$,

$$\left|\sum_{k=0}^{m} (A_k)_{i,i}\mu_l^k\right| \leq \sum_{j\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m} (A_k)_{i,j}\mu_l^k\right|,$$

or

$$\left|\sum_{k=0}^{m} (A_k)_{i,i}\frac{\mu_l^k}{\mu_l^m}\right| - \sum_{j\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m} (A_k)_{i,j}\frac{\mu_l^k}{\mu_l^m}\right| \leq 0.$$

As $l \to +\infty$, it follows

$$|(A_m)_{i,i}| - \sum_{j\in\beta_i\backslash\{i\}} |(A_m)_{i,j}| \leq 0,$$

and hence, $0 \in G_i(A_m)$.

For the converse, suppose that $0 \in G_i(A_m)$ (or equivalently, $|(A_m)_{i,i}| \leq r_i(A_m)$). Then $0 \in G_i(\hat{P})$ and, by definition, $\infty \in G_i(P)$ (where by hypothesis, 0 is a non-isolated point of $G_i(\hat{P})$, and hence, ∞ is a non-isolated point of $G_i(P)$).

(ii) Suppose that $i \in \overline{\beta}_i$, i.e., $(A_m)_{i,i} = 0$.

It is clear that $0 \in G_i(A_m)$, and

$$G_i(P)\backslash\{0\} = \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\mu^k\right| \leq \sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\mu^k\right|\right\}$$
$$= \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\frac{\mu^k}{\mu^{m-1}}\right| - \sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\frac{\mu^k}{\mu^{m-1}}\right| \leq 0\right\},$$

where at least one of the coefficients $(A_m)_{i,j}$, $j \in \mathcal{N}\backslash\{i\}$, is nonzero. As a consequence, for "large enough" $|\mu|$, μ lies in $G_i(P)$. Thus, there exists a real number $M > 0$ such that every scalar $\mu \in \mathbb{C}$ with $|\mu| \geq M$ lies in $G_i(P)$, i.e., $\{\mu \in \mathbb{C} : |\mu| \geq M\} \subseteq G_i(P)$. ◻

Remark 2.7. Suppose that $0 \in G_i(A_m)$, i.e., $0 \in G_i(\hat{P})$. If the origin is a non-isolated point of $G_i(\hat{P})$, then there exists a sequence $\{\mu_l\}_{l\in\mathbb{N}} \subset G_i(\hat{P})\backslash\{0\}$ that converges to 0. This means that the sequence $\{\mu_l^{-1}\}_{l\in\mathbb{N}} \subset G_i(P)$ is unbounded. Hence, the i-th row Gershgorin set $G_i(P)$ is also unbounded and does not have the infinity as an isolated point. On the other hand, if the origin is an isolated point of $G_i(\hat{P})$ (which yields $\beta_i = \emptyset$), then $G_i(P)\backslash\{0\} = \left\{\mu \in \mathbb{C}\backslash\{0\} : \mu^{-1} \in G_i(\hat{P})\backslash\{0\}\right\} \cup \{\infty\}$, where the set $\left\{\mu \in \mathbb{C}\backslash\{0\} : \mu^{-1} \in G_i(\hat{P})\backslash\{0\}\right\}$ is bounded.

Remark 2.8. Suppose that $i \in \beta_i$ and the origin is an interior point of $G_i(A_m)$ (or equivalently, $(A_m)_{i,i} \neq 0$ and $|(A_m)_{i,i}| - r_i(A_m) < 0$). Then, for any sequence $\{\mu_l\}_{l\in\mathbb{N}}$ in $\mathbb{C}\backslash\{0\}$ such that $|\mu_l| \to +\infty$, and for sufficiently large l,

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\frac{\mu_l^k}{\mu_l^m}\right| - \sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\frac{\mu_l^k}{\mu_l^m}\right| \leq 0,$$

or

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\mu_l^k\right| - \sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\mu_l^k\right| \leq 0,$$

or

$$\mu_l \in G_i(P).$$

As a consequence, $G_i(P)$ is unbounded, and there exists a real number $M > 0$ such that $\{\mu \in \mathbb{C} : |\mu| \geq M\} \subseteq G_i(P)$.

The next result, as well as Theorem 3.7 at the end of the next section, is a special case of Theorem 3.1 in the work of D. Bindel and A. Hood [1] (see also Theorem 2 in [13]). For clarity, we give an elementary proof.

Theorem 2.9. *Suppose that A_m has no zero rows, and the Gershgorin set $G(P)$ is bounded. Then the number of the connected components of $G(P)$ is less than or equal to nm. Moreover, if $\mathcal{G}$ is a connected component of $G(P)$ constructed by ξ connected components of the row Gershgorin sets, then the total number of roots of the polynomials $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \ldots, (P(\lambda))_{n,n}$ in $\mathcal{G}$ is at least ξ and it is equal to the number of eigenvalues of $P(\lambda)$ in $\mathcal{G}$, counting multiplicities.*

Proof. Since $G(P)$ is bounded, Theorem 2.6 (i) implies that for every $i \in \mathcal{N}$, $0 \notin G_i(A_m)$, $i \in \beta_i$ and the polynomial $(P(\lambda))_{i,i}$ is of degree m (i.e., it has exactly m roots, counting multiplicities).

For all coefficient matrices of $P(\lambda)$, $A_0, A_1, \ldots, A_m$, consider the splitting

$$A_j = D_j + F_j, \quad j = 0, 1, \ldots, m,$$

where D_j is the diagonal part of A_j (i.e., D_j is a diagonal matrix and its diagonal coincides with the diagonal of A_j) for every $j = 0, 1, \dots, m$. Define also the family of matrix polynomials

$$\begin{aligned} P_t(\lambda) &= (D_m + tF_m)\lambda^m + (D_{m-1} + tF_{m-1})\lambda^{m-1} + \cdots + (D_1 + tF_1)\lambda + (D_0 + tF_0) \\ &= D_m\lambda^m + D_{m-1}\lambda^{m-1} + \cdots + D_1\lambda + D_0 + t\left(F_m\lambda^m + F_{m-1}\lambda^{m-1} + \cdots + F_1\lambda + F_0\right) \end{aligned}$$

for $t \in [0, 1]$. We observe that

$$G(P_{t_1}) \subseteq G(P_{t_2}), \quad 0 \le t_1 \le t_2 \le 1. \tag{2.4}$$

Indeed, if $0 < t_1 \le t_2 \le 1$ and a scalar $\mu \in \mathbb{C}$ lies in $G_i(P_{t_1})$ for some $i \in \mathcal{N}$, then $|(P_{t_2}(\mu))_{i,i}| = |(P_{t_1}(\mu))_{i,i}| \le r_i(P_{t_1}(\mu)) = \frac{t_1}{t_2}\, r_i(P_{t_2}(\mu)) \le r_i(P_{t_2}(\mu))$, and thus, $\mu \in G_i(P_{t_2})$. Moreover, it is apparent that if $0 = t_1 \le t_2 \le 1$ and $\mu \in G_i(P_0)$, then μ is a root of $(P(\lambda))_{i,i}$ and $|(P_{t_2}(\mu))_{i,i}| = |(P_0(\mu))_{i,i}| = r_i(P_0(\mu)) = 0 \le r_i(P_{t_2}(\mu))$.

By Proposition 2.2 (i), Theorem 2.6 and (2.4), it follows that the family

$$G(P_t), \quad t \in [0, 1]$$

is a family of nondecreasing compact sets (with respect to $t \in [0, 1]$). Furthermore, $G(P_1) = G(P)$, and $G(P_0)$ coincides with the roots of the polynomials $(P(\lambda))_{1,1}$, $(P(\lambda))_{2,2}, \dots, (P(\lambda))_{n,n}$ (which are eigenvalues of the diagonal matrix polynomial $D_m\lambda^m + D_{m-1}\lambda^{m-1} + \cdots + D_1\lambda + D_0$). Keeping in mind that all polynomials $\det P_t(\lambda)$, $t \in [0, 1]$, are of degree nm, the continuity of the eigenvalues of $P_t(\lambda)$ in $t \in [0, 1]$ implies that each root of the polynomials $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \dots, (P(\lambda))_{n,n}$ is connected to an eigenvalue of $P(\lambda)$ with a continuous curve in $G(P)$. Hence, any (compact) connected component of $G(P)$ constructed by ξ connected components of the row Gershgorin sets, $\mathcal{G}$, contains at least ξ roots of the polynomials $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \dots, (P(\lambda))_{n,n}$, and each root of $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \dots, (P(\lambda))_{n,n}$ in $\mathcal{G}$ is connected to an eigenvalue of $P_t(\lambda)$, $t \in [0, 1]$, with a continuous curve in $\mathcal{G}$. ∎

2.3. Examples. The Gershgorin set is simple and gives better results than known bounds based on norms. In the next example, we compare it to some bounds given in [10]; in particular, we draw and compare the Gershgorin set to the bounds given by Lemmas 2.2 and 2.3 of [10].

For a matrix polynomial $P(\lambda)$ as in (1.1), consider the $n \times n$ matrices $U_i = A_m^{-1}A_i$ $(i = 0, 1, \dots, m-1)$ and $L_i = A_0^{-1}A_i$ $(i = 1, 2, \dots, m)$. Then, by Lemma 2.2 of [10], every eigenvalue λ_0 of $P(\lambda)$ satisfies

$$\left(1 + \sum_{j=1}^{m} \|L_j\|_p\right)^{-1} \le |\lambda_0| \le 1 + \sum_{j=0}^{m-1} \|U_j\|_p, \quad 1 \le p \le \infty. \tag{2.5}$$

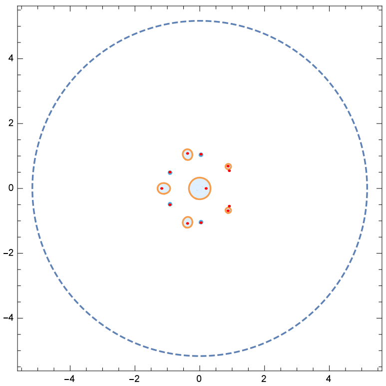
Define also the $n \times nm$ matrices $U = [\, U_0 \ U_1 \ \cdots \ U_{m-1} \,]$ and $L = [\, L_m \ L_{m-1} \ \cdots \ L_1 \,]$. Then, by Lemma 2.3 of [10], we have the following bounds:

$$\max\left\{\|L_m\|_1, 1 + \max_{i=1,2,\dots,m-1} \|L_i\|_1\right\}^{-1} \le |\lambda| \le \max\left\{\|U_0\|_1, 1 + \max_{i=1,2,\dots,m-1} \|U_i\|_1\right\}, \tag{2.6}$$

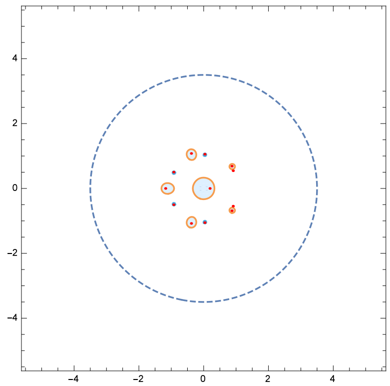
$$\max\{\|L\|_\infty, 1\}^{-1} \le |\lambda| \le \max\{\|U\|_\infty, 1\}) \tag{2.7}$$

and

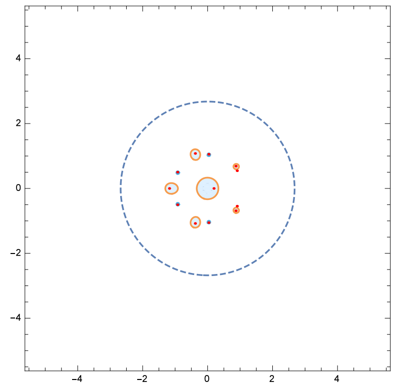
$$\|I + LL^*\|_2^{-1/2} \le |\lambda| \le \|I + UU^*\|_2^{1/2}. \tag{2.8}$$



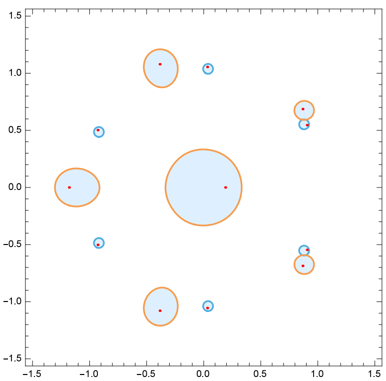
(a) The set $G(P)$ and the bounds of [10, Lemma 2.2] for norm-2.



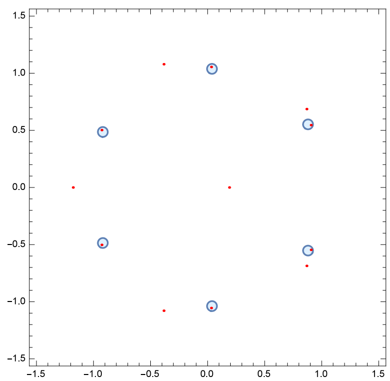
(b) The set $G(P)$ and the bounds of [10, Lemma 2.3] for norm-∞.



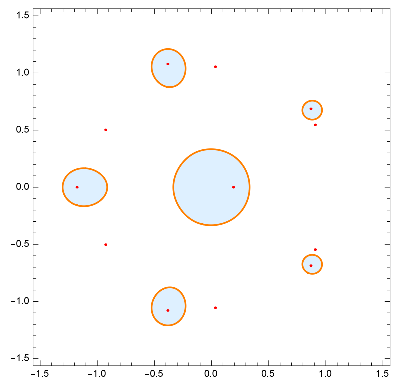
(c) The set $G(P)$ and the bounds of [10, Lemma 2.3] for norm-2.



(d) The set $G(P)$.



(e) The set $G_1(P)$.



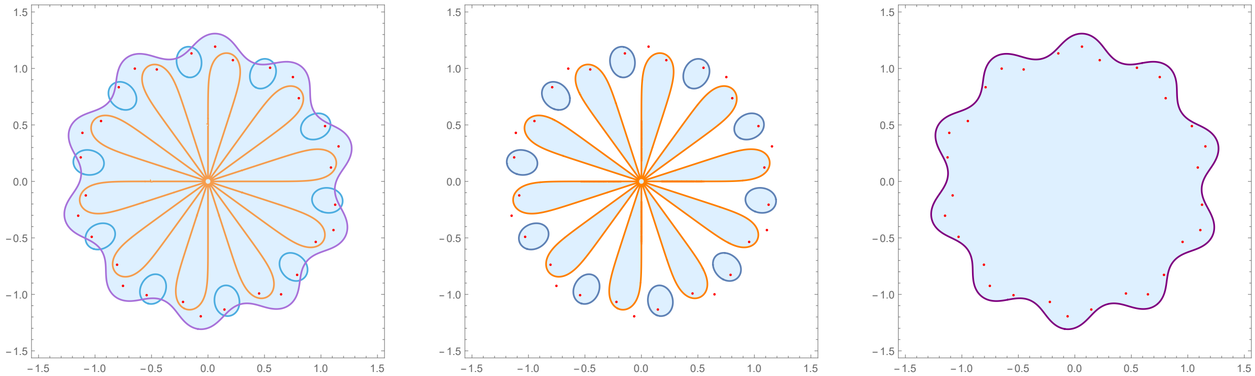
(f) The set $G_2(P)$.

Figure 1: Comparing the Gershgorin set to the bounds of Lemmas 2.2 and 2.3 of [10].

Example 2.10. Consider the (real) matrix polynomial

$$P(\lambda) = \begin{bmatrix} 5\lambda^6 + \lambda^3 + 7 & 2\lambda^6 + 4 \\ \lambda^6 + 1 & 2\lambda^6 + 3\lambda \end{bmatrix}.$$

The upper bound given by (2.5) for norm-2 is approximately equal to 5.1673 and can be seen in Figure 1 (a) (dashed line) compared to the Gershgorin set $G(P)$. The lower bound of (2.5) is equal to 0.09 and since it is relatively small, its illustration is omitted in the figure. The lower and upper bounds given by (2.6)–(2.8) are 0.1081 and 3.6250 for norm-1, 0.1111 and 3.50 for norm-∞, and 0.1394 and 2.6780 for norm-2. The upper bounds for norm-∞ and norm-2 compared to the Gershgorin set $G(P)$ are illustrated in parts (b) and (c) of the figure, respectively. The sets $G(P)$, $G_1(P)$ (with exactly six connected components) and $G_2(P)$ (with exactly six connected components) are given (magnified) in parts (d), (e) and (f) of Figure 1, respectively. Here and elsewhere, the eigenvalues are marked with dots. Note that Proposition 2.2 (iv) (symmetry with respect to the real axis), Theorem 2.6 (boundedness conditions) and Theorem 2.9 (distribution of eigenvalues in the connected components) are clearly verified.



(a) The set $G(Q)$. (b) The sets $G_1(Q)$ and $G_2(Q)$. (c) The set $G_3(Q)$.

Figure 2: The Gershgorin set of $Q(\lambda) = I\lambda^m - A$.

For a matrix $A \in \mathbb{C}^{n\times n}$, consider the matrix polynomial $Q(\lambda) = I\lambda^m - A$. Then, for any $i \in \mathcal{N}$, $G_i(Q) = \{\mu \in \mathbb{C} : |\lambda^m - (A)_{i,i}| \le r_i(A)\} = \left\{\mu \in \mathbb{C} : \mu^{1/m} \in G_i(A)\right\}$ is bounded and its shape depends on the location of the origin with respect to $G_i(A)$; in particular, we have three cases illustrated in the next example.

Example 2.11. Consider the (complex) matrix polynomial

$$Q(\lambda) = I\lambda^{10} - A = I\lambda^{10} - \begin{bmatrix} -3.5\mathrm{i} & -3\mathrm{i} & 0 \\ -2 & 2\mathrm{i} & 0 \\ -0.5-2\mathrm{i} & 7\mathrm{i} & -5+3\mathrm{i} \end{bmatrix} = \begin{bmatrix} \lambda^{10}+3.5\mathrm{i} & 3\mathrm{i} & 0 \\ 2 & \lambda^{10}-2\mathrm{i} & 0 \\ 0.5+2\mathrm{i} & -7\mathrm{i} & \lambda^{10}+5-3\mathrm{i} \end{bmatrix}.$$

The Gershgorin set $G(Q) = G_1(Q) \cup G_2(Q) \cup G_3(Q)$, the union of the sets $G_1(Q)$ and $G_2(Q)$, and the set $G_3(Q)$ are illustrated in parts (a), (b) and (c) of Figure 2, respectively. Observe that $0 \notin G_1(A)$ (since $|-3.5\mathrm{i}| > |-3\mathrm{i}|$), and thus, $G_1(Q)$ consists of ten connected components located symmetrically around the origin. Concerning the second row, notice that $|2\mathrm{i}| = |-2|$ and $0 \in \partial G_2(A)$, and hence, the set $G_2(Q)$ is connected and "daisy" shaped. Finally, in the third row, we have $|-5+3\mathrm{i}| < |-0.5-2\mathrm{i}| + |7\mathrm{i}|$ (i.e., the origin lies in the interior of $G_3(A)$), and $G_3(Q)$ is connected, with smooth boundary.

3. The generalized Gershgorin set. In this section, we extend the notion of the generalized Gershgorin set (or the $\mathcal{A}$-Ostrowski set) of matrices to the case of matrix polynomials.

3.1. Definitions. For a square complex matrix $A \in \mathbb{C}^{n\times n}$, denote $c_i(A) = r_i(A^T)$, $i \in \mathcal{N}$, and for any $a \in [0,1]$, define the disks $\mathcal{A}_{i,a}(A) = \left\{\mu \in \mathbb{C} : |\mu - (A)_{i,i}| \le r_i(A)^a c_i(A)^{1-a}\right\}$, $i \in \mathcal{N}$. Then the spectrum of matrix A lies in the union $\mathcal{A}_a(A) = \bigcup_{i\in\mathcal{N}} \mathcal{A}_{i,a}(A)$, and consequently, it lies in the intersection $\mathcal{A}(A) = \bigcap_{a\in[0,1]} \mathcal{A}_a(A)$ [18, 21], which is known as the *generalized Gershgorin set* (or the *$\mathcal{A}$-Ostrowski set*) of A.

Similarly, for a matrix polynomial $P(\lambda)$ as in (1.1), we define

$$\mathcal{A}_{i,a}(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{A}_{i,a}(P(\mu))\} = \left\{\mu \in \mathbb{C} : |P(\mu)_{i,i}| \le r_i(P(\mu))^a c_i(P(\mu))^{1-a}\right\}$$

for any $i \in \mathcal{N}$ and $a \in [0,1]$,

$$\mathcal{A}_a(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{A}_a(P(\mu))\} = \bigcup_{i \in \mathcal{N}} \mathcal{A}_{i,a}(P), \quad a \in [0,1],$$

and the *generalized Gershgorin set* of $P(\lambda)$

$$\mathcal{A}(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{A}(P(\mu))\} = \bigcap_{a \in [0,1]} \mathcal{A}_a(P).$$

As before, we say that $\mu = \infty$ lies in $\mathcal{A}_{i,a}(P)$ (resp., in $\mathcal{A}_a(P)$, or in $\mathcal{A}(P)$) exactly when 0 lies in $\mathcal{A}_{i,a}(\hat{P})$ (resp., in $\mathcal{A}_a(\hat{P})$, or in $\mathcal{A}(\hat{P})$).

THEOREM 3.1. *The generalized Gershgorin set of a matrix polynomial $P(\lambda)$ is a subset of the Gershgorin set of $P(\lambda)$ that contains all the (finite and infinite) eigenvalues of $P(\lambda)$.*

Proof. By definition, a scalar $\mu \in \mathbb{C}$ lies in $\mathcal{A}(P)$ (resp., in $\sigma(P)$, or in $G(P)$) if and only if 0 lies in $\mathcal{A}(P(\mu))$ (resp., in $\sigma(P(\mu))$, or in $G(P(\mu))$). Since $\sigma(P(\mu)) \subseteq \mathcal{A}(P(\mu)) \subseteq G(P(\mu))$ and $\sigma(\hat{P}(\mu)) \subseteq \mathcal{A}(\hat{P}(\mu)) \subseteq G(\hat{P}(\mu))$ [18, 21], the proof follows readily. ∎

3.2. Basic properties. Let $P(\lambda)$ be an $n \times n$ matrix polynomial as in (1.1). It is easy to verify the following properties.

PROPOSITION 3.2. *Let $a \in [0,1]$ and $i \in \mathcal{N}$.*

- **(i)** *$\mathcal{A}_{i,a}(P)$ is a closed subset of $\mathbb{C}$.*
- **(ii)** *For any scalar $b \in \mathbb{C}\backslash\{0\}$, consider the matrix polynomials $Q_1(\lambda) = P(b\lambda)$, $Q_2(\lambda) = bP(\lambda)$ and $Q_3(\lambda) = P(\lambda + b)$. Then $\mathcal{A}_{i,a}(Q_1) = b^{-1}\mathcal{A}_{i,a}(P)$, $\mathcal{A}_{i,a}(Q_2) = \mathcal{A}_{i,a}(P)$ and $\mathcal{A}_{i,a}(Q_3) = \mathcal{A}_{i,a}(P) - b$.*
- **(iii)** *If the (i,i)-th entry of $P(\lambda)$, $(P(\lambda))_{i,i}$, is (identically) zero, then $\mathcal{A}_{i,a}(P) = \mathbb{C}$.*
- **(iv)** *If $a \in (0,1)$, and all the coefficient matrices $A_0, A_1, \ldots, A_m$ have their i-th rows and i-th columns real, then $\mathcal{A}_{i,a}(P)$ is symmetric with respect to the real axis (for $a = 1$ or 0, see Proposition 2.2* (iv) *applied to $P(\lambda)$ or its transpose, respectively).*

Proof. The proof is similar to the proof of Proposition 2.2. ∎

Proposition 2.3 for the row Gershgorin sets can be readily generalized for $\mathcal{A}_{i,a}(P)$.

PROPOSITION 3.3. *For any $a \in [0,1]$ and $i \in \mathcal{N}$, the set $\left\{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| < r_i(P(\mu))^a c_i(P(\mu))^{1-a}\right\}$ lies in the interior of $\mathcal{A}_{i,a}(P)$; as a consequence,*

$$\partial\mathcal{A}_{i,a}(P) \subseteq \left\{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| = r_i(P(\mu))^a c_i(P(\mu))^{1-a}\right\} = \{\mu \in \mathbb{C} : 0 \in \partial\mathcal{A}_{i,a}(P(\mu))\}.$$

Theorems 2.4 and 2.5 can also be extended.

THEOREM 3.4. (For $a = 1$ or 0, see Theorem 2.4 applied to $P(\lambda)$ or its transpose, respectively.) *Let $a \in (0,1)$ and $i \in \mathcal{N}$. If μ is an isolated point of $\mathcal{A}_{i,a}(P)$, then μ is a common root of all polynomials $(P(\lambda))_{i,j}$, $j \in \mathcal{N}$, or a common root of all polynomials $(P(\lambda))_{p,i}$, $p \in \mathcal{N}$; as a consequence, μ is an eigenvalue of $P(\lambda)$.*

Proof. For the sake of contradiction, assume that $(P(\mu))_{i,i} \neq 0$. Since μ is an isolated point, it belongs to the boundary of $\mathcal{A}_{i,a}(P)$, and there exists an $\varepsilon > 0$ such that the disk $D(\mu, \varepsilon) = \{\lambda \in \mathbb{C} : |\lambda - \mu| \leq \varepsilon\}$

contains no other point of $\mathcal{A}_{i,a}(P)$. The set $\mathcal{A}_{i,a}(P)$ can be written as

$$\begin{aligned}\mathcal{A}_{i,a}(P) &= \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| \leq r_i(P(\mu))^a c_i(P(\mu))^{1-a}\} \\ &= \left\{\mu \in \mathbb{C} : \frac{r_i(P(\mu))^a c_i(P(\mu))^{1-a}}{|(P(\mu))_{i,i}|} \geqslant 1\right\} \\ &= \left\{\mu \in \mathbb{C} : \log \frac{r_i(P(\mu))^a c_i(P(\mu))^{1-a}}{|(P(\mu))_{i,i}|} \geqslant 0\right\} \\ &= \left\{\mu \in \mathbb{C} : a \log \frac{r_i(P(\mu))}{|(P(\mu))_{i,i}|} + (1-a) \log \frac{c_i(P(\mu))}{|(P(\mu))_{i,i}|} \geqslant 0\right\}.\end{aligned}$$

Consider the function

$$\phi(\lambda) = a \log \frac{r_i(P(\lambda))}{|(P(\lambda))_{i,i}|} + (1-a) \log \frac{c_i(P(\lambda))}{|(P(\lambda))_{i,i}|}, \quad \lambda \in D(\mu, \varepsilon), \tag{3.9}$$

and observe that it is subharmonic (as a sum of subharmonic functions) and satisfies the maximum principle [2, 4]. By definition, $\phi(\lambda)$ is equal to zero on the boundary of $\mathcal{A}_{i,a}(P)$, nonnegative in the interior of $\mathcal{A}_{i,a}(P)$ and negative elsewhere; see Proposition 3.3. Furthermore, since $\mu \in \partial\mathcal{A}_{i,a}(P)$, it follows that the function $\phi(\lambda)$ is equal to zero at the center μ of $D(\mu, \varepsilon)$ and negative in the rest of the disk; this is a contradiction because $\phi(\lambda)$ satisfies the maximum principle. As a consequence,

$$0 = (P(\mu))_{i,i} = \left(\sum_{j \in \mathcal{N}\setminus\{i\}} |(P(\mu))_{i,j}|\right)^a \left(\sum_{p \in \mathcal{N}\setminus\{i\}} |(P(\mu))_{p,i}|\right)^{1-a},$$

and the result follows readily. ∎

Theorem 3.5. *Let $a \in [0,1]$ and $i \in \mathcal{N}$. Any point of the i-th generalized Gershgorin set $\mathcal{A}_{i,a}(P)$ has local dimension either 2 or 0 (isolated point); in other words, the generalized Gershgorin set of $P(\lambda)$ cannot have parts with (nontrivial) curves.*

Proof. For $a = 1$ or 0, see Theorem 2.5 applied to $P(\lambda)$ or its transpose, respectively. For $a \in (0,1)$, the proof is similar to the proof of Theorem 2.5, using the (subharmonic) function $\phi(\lambda)$ in (3.9) instead of $\varphi(\lambda)$ in (2.2). ∎

Recalling the definition of β_i and $\overline{\beta}_i$ in (2.3), we define similarly the sets

$$\gamma_j = \{p \in \mathcal{N} : (A_m)_{p,j} \neq 0\} \quad \text{and} \quad \overline{\gamma}_j = \mathcal{N}\setminus\gamma_j = \{p \in \mathcal{N} : (A_m)_{p,j} = 0\}, \tag{3.10}$$

and we obtain (as in Theorem 2.6) necessary and sufficient conditions for the sets $\mathcal{A}_{i,a}(P)$ to be bounded.

Note that, by Theorem 3.4, if $\mu \in \mathbb{C}$ is an isolated point of $\mathcal{A}_{i,a}(P)$ for some $a \in (0,1)$ and $i \in \mathcal{N}$, then μ is a common root of all polynomials $(P(\lambda))_{i,j}$, $j \in \mathcal{N}$, or a common root of all polynomials $(P(\lambda))_{p,i}$, $p \in \mathcal{N}$. As a consequence, if the sets β_i and γ_i are nonempty, then the origin is not an isolated point of $\mathcal{A}_{i,a}(\hat{P})$, i.e., $\mathcal{A}_{i,a}(P)$ is not the union of a bounded set and ∞.

Theorem 3.6. (For $a = 1$ or 0, see Theorem 2.6 applied to $P(\lambda)$ or its transpose, respectively.) *Let $a \in (0,1)$, and suppose that for an $i \in \mathcal{N}$, the sets β_i and γ_i are nonempty.*

(i) *If $i \in \beta_i$ (thus, $i \in \gamma_i$), then $\mathcal{A}_{i,a}(P)$ is unbounded if and only if $0 \in \mathcal{A}_{i,a}(A_m)$.*
(ii) *If $i \in \overline{\beta}_i$ (thus, $i \in \overline{\gamma}_i$), then $\mathcal{A}_{i,a}(P)$ is unbounded and $0 \in \mathcal{A}_{i,a}(A_m)$.*

Proof. (i) Suppose that $i \in \beta_i$, i.e., $(A_m)_{i,i} \neq 0$.

Let $\mathcal{A}_{i,a}(P)$ be unbounded. Since the sets β_i and γ_i are nonempty, the origin is not an isolated point of $\mathcal{A}_{i,a}(\hat{P})$ and there is a sequence $\{\mu_l\}_{l\in\mathbb{N}}$ in $\mathcal{A}_{i,a}(P)\backslash\{0\}$ such that $|\mu_l| \to +\infty$. Then, for every $l \in \mathbb{N}$,

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\mu_l^k\right| \le \left(\sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\mu_l^k\right|\right)^a \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{p,i}\mu_l^k\right|\right)^{1-a},$$

or

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\frac{\mu_l^k}{\mu_l^m}\right| \le \left(\sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\frac{\mu_l^k}{\mu_l^m}\right|\right)^a \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{p,i}\frac{\mu_l^k}{\mu_l^m}\right|\right)^{1-a},$$

As $l \to +\infty$, it follows

$$|(A_m)_{i,i}| - \left(\sum_{j\in\beta_i\backslash\{i\}}|(A_m)_{i,j}|\right)^a \left(\sum_{p\in\gamma_i\backslash\{i\}}|(A_m)_{p,i}|\right)^{1-a} \le 0,$$

and hence, $0 \in \mathcal{A}_{i,a}(A_m)$.

For the converse, suppose that $0 \in \mathcal{A}_{i,a}(A_m)$ (or equivalently, $|(A_m)_{i,i}| \le r_i(A_m)^a c_i(A_m)^{1-a}$). Then $0 \in \mathcal{A}_{i,a}(\hat{P})$ and, by definition, $\infty \in \mathcal{A}_{i,a}(P)$.

(ii) Suppose that $i \in \overline{\beta}_i$, i.e., $(A_m)_{i,i} = 0$.

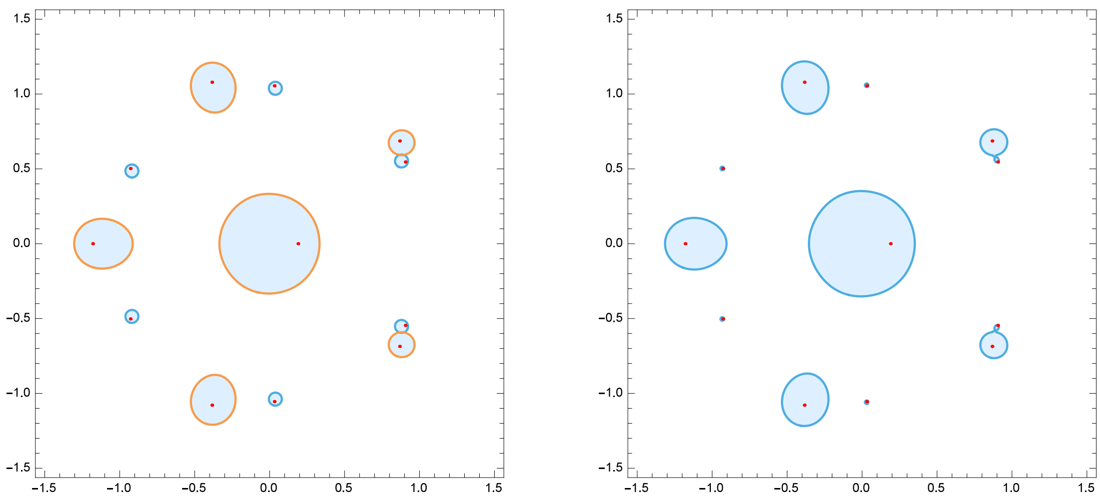
It is clear that $0 \in \mathcal{A}_{i,a}(A_m)$. Moreover,

$$\begin{aligned}
&\mathcal{A}_{i,a}(P)\backslash\{0\}\\
&= \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\mu^k\right| \le \left(\sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\mu^k\right|\right)^a \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{p,i}\mu^k\right|\right)^{1-a}\right\}\\
&= \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\frac{\mu^k}{\mu^m}\right| \le \left(\sum_{j\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,j}\frac{\mu^k}{\mu^m}\right|\right)^a \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{p,i}\frac{\mu^k}{\mu^m}\right|\right)^{1-a}\right\},
\end{aligned}$$

where at least one of the coefficients $(A_m)_{i,j}$, $j \in \mathcal{N}\backslash\{i\}$, and one of the coefficients $(A_m)_{p,i}$, $p \in \mathcal{N}\backslash\{j\}$, are nonzero. As a consequence, for "large enough" $|\mu|$, μ lies in $\mathcal{A}_{i,a}(P)$. Thus, there exists a real number $M > 0$ such that every scalar $\mu \in \mathbb{C}$ with $|\mu| \ge M$ lies in $\mathcal{A}_{i,a}(P)$, i.e., $\{\mu \in \mathbb{C} : |\mu| \ge M\} \subseteq \mathcal{A}_{i,a}(P)$. ◻

As mentioned in the paragraph before Theorem 2.9, we conclude this section with a result which is a special case of Theorem 3.1 of [1].

THEOREM 3.7. (For $a = 1$ or 0, see Theorem 2.9 applied to $P(\lambda)$ or its transpose, respectively.) *Let $a \in (0, 1)$, and suppose that for every $i \in \mathcal{N}$, the sets β_i and γ_i are nonempty. If $\mathcal{A}_a(P)$ is bounded, then the number of connected components of $\mathcal{A}_a(P)$ is less than or equal to nm. Moreover, if $\mathcal{G}$ is a connected component of $\mathcal{A}_a(P)$ constructed by ξ connected components of the sets $\mathcal{A}_{1,a}(P), \mathcal{A}_{2,a}(P), \ldots, \mathcal{A}_{n,a}(P)$, then the total number of roots of the polynomials $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \ldots, (P(\lambda))_{n,n}$ in $\mathcal{G}$ is at least ξ and it is equal to the number of eigenvalues of $P(\lambda)$ in $\mathcal{G}$, counting multiplicities.*



(a) The Gershgorin set.

(b) The generalized Gershgorin set.

Figure 3: Comparing the Gershgorin set and the generalized Gershgorin set.

Proof. Since $\mathcal{A}_a(P)$ is bounded, Theorem 3.6 (i) implies that for every $i \in \mathcal{N}$, $i \in \beta_i \cap \gamma_i$ and the polynomial $(P(\lambda))_{i,i}$ is of degree m (i.e., it has exactly m roots, counting multiplicities). Based on the continuity of the eigenvalues of a matrix polynomial with respect to its matrix coefficients and the boundedness of $\mathcal{A}_a(P)$, one can complete the proof, following the arguments of the proof of Theorem 2.9. ∎

3.3. Examples.

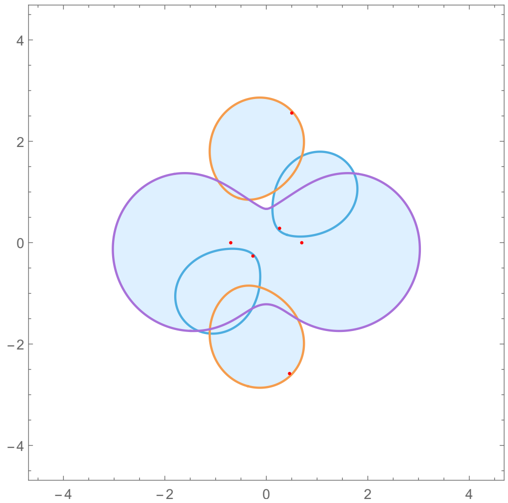
Example 3.8. Consider the matrix polynomial of Example 2.10. In Figure 3, one can see the improvement of the Gershgorin set due to the use of column sums in addition to row sums[1]. Moreover, Proposition 3.2 (iv) (symmetry with respect to the real axis), Theorem 3.6 (boundedness conditions) and Theorem 3.7 (distribution of eigenvalues in the connected components) are verified.

Example 3.9. Consider the matrix polynomial

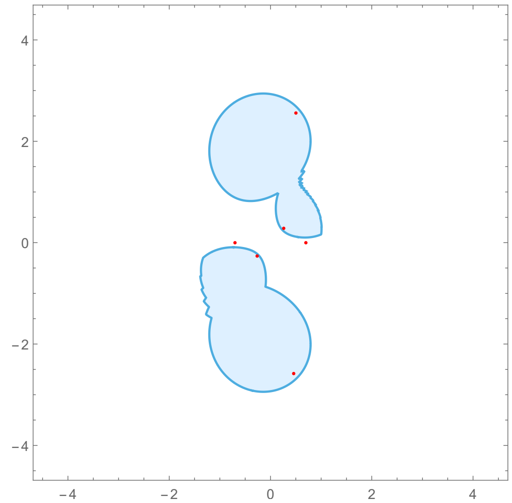
$$P(\lambda) = \begin{bmatrix} 4.2\lambda^2 - \mathrm{i} & 4\lambda^2 & 0 \\ \lambda - 3 & \lambda^2 + 4 & 0 \\ -2\lambda + \mathrm{i} & \lambda^2 + 2 & 2\lambda^2 - 1 \end{bmatrix}.$$

In Figure 4, the Gershgorin and generalized Gershgorin sets of $P(\lambda)$ are illustrated. It is worth mentioning that $\mu_1 = \frac{\sqrt{2}}{2}$ and $\mu_2 = -\frac{\sqrt{2}}{2}$ are isolated points of $\mathcal{A}(P)$, and the third columns of the matrices $P(\mu_1) = P\left(\frac{\sqrt{2}}{2}\right)$ and $P(\mu_2) = P\left(-\frac{\sqrt{2}}{2}\right)$ are zero; see Theorems 2.4 and 3.4.

[1]In Figures 3 (b) and 4 (b), the generalized Gershgorin set $\mathcal{A}(P)$ is estimated by the intersection of the sets $\mathcal{A}_a(P)$, $a = 0, 0.05, 0.10, \ldots, 0.95, 1$.



(a) The Gershgorin set.



(b) The generalized Gershgorin set.

Figure 4: Comparing the Gershgorin set and the generalized Gershgorin set.

4. The Brauer set. In this section, we introduce the Brauer set of matrix polynomials, similarly to the Gershgorin set and the generalized Gershgorin set.

4.1. Definitions. The *Brauer set* of a matrix $A \in \mathbb{C}^{n\times n}$ is defined as $\mathcal{B}(A) = \bigcup_{i\in\mathcal{N}} \bigcup_{j=1}^{i-1} \mathcal{B}_{i,j}(A)$, where $\mathcal{B}_{i,j}(A) = \{\mu \in \mathbb{C} : |\mu - (A)_{i,i}|\,|\mu - (A)_{j,j}| \le r_i(A)r_j(A)\}$. The Brauer set consists of $\frac{n(n-1)}{2}$ Cassini ovals, lies in the Gershgorin set, and contains all the eigenvalues of the matrix [3, 21].

For a matrix polynomial $P(\lambda)$ as in (1.1) and $i, j \in \mathcal{N}$ with $j < i$, we define the (i,j)*-th Brauer set* of $P(\lambda)$

$$\mathcal{B}_{i,j}(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{B}_{i,j}(P(\mu))\} = \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| \le r_i(P(\mu))r_j(P(\mu))\},$$

and the *Brauer set* of $P(\lambda)$

$$\mathcal{B}(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{B}(P(\mu))\} = \bigcup_{i\in\mathcal{N}} \bigcup_{j=1}^{i-1} \mathcal{B}_{i,j}(P).$$

We also say that $\mu = \infty$ lies in $\mathcal{B}_{i,j}(P)$ (resp., in $\mathcal{B}(P)$) exactly when 0 lies in $\mathcal{B}_{i,j}(\hat{P})$ (resp., in $\mathcal{B}(\hat{P})$). The Brauer theorem for matrices [3, 21] is generalized to the case of matrix polynomials as well as the Gershgorin circle theorem.

THEOREM 4.1. *The Brauer set of a matrix polynomial $P(\lambda)$ is a subset of the Gershgorin set of $P(\lambda)$ that contains all the (finite and infinite) eigenvalues of $P(\lambda)$.*

Proof. By definition, a scalar $\mu \in \mathbb{C}$ lies in $\mathcal{B}(P)$ (resp., in $\sigma(P)$, or in $G(P)$) if and only if 0 lies in $\mathcal{B}(P(\mu))$ (resp., in $\sigma(P(\mu))$, or in $G(P(\mu))$). Since $\sigma(P(\mu)) \subseteq \mathcal{B}(P(\mu)) \subseteq G(P(\mu))$ and $\sigma(\hat{P}(\mu)) \subseteq \mathcal{B}(\hat{P}(\mu)) \subseteq G(\hat{P}(\mu))$ [3, 21], the proof follows readily. ∎

4.2. Basic properties. Let $P(\lambda)$ be an $n \times n$ matrix polynomial as in (1.1). It is easy to verify the following properties.

PROPOSITION 4.2. *Consider two integers $i, j \in \mathcal{N}$ with $j < i$.*

(i) *$\mathcal{B}_{i,j}(P)$ is a closed subset of $\mathbb{C}$.*
(ii) *For any scalar $b \in \mathbb{C}\backslash\{0\}$, consider the matrix polynomials $Q_1(\lambda) = P(b\lambda)$, $Q_2(\lambda) = bP(\lambda)$ and $Q_3(\lambda) = P(\lambda + b)$. Then $\mathcal{B}_{i,j}(Q_1) = b^{-1}\mathcal{B}_{i,j}(P)$, $\mathcal{B}_{i,j}(Q_2) = \mathcal{B}_{i,j}(P)$ and $\mathcal{B}_{i,j}(Q_3) = \mathcal{B}_{i,j}(P) - b$.*
(iii) *If the (i,i)-th or the (j,j)-th entry of $P(\lambda)$ is (identically) zero, then $\mathcal{B}_{i,j}(P) = \mathbb{C}$.*
(iv) *If all the coefficient matrices $A_0, A_1, \ldots, A_m$ have their i-th and j-th rows real, then $\mathcal{B}_{i,j}(P)$ is symmetric with respect to the real axis.*

Proof. The proof is similar to the proof of Proposition 2.2. ∎

Propositions 2.3 and 3.3 for the Gershgorin set and generalized Gershgorin set, respectively, can be readily extended to the case of the Brauer set.

PROPOSITION 4.3. *For any $i, j \in \mathcal{N}$ with $j < i$, the set $\{\mu \in \mathbb{C} : |(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| < r_i(P(\mu)) \cdot r_j(P(\mu))\}$ lies in the interior of $\mathcal{B}_{i,j}(P)$; as a consequence,*

$$\partial\mathcal{B}_{i,j}(P) \subseteq \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| = r_i(P(\mu))r_j(P(\mu))\} = \{\mu \in \mathbb{C} : 0 \in \partial\mathcal{B}_{i,j}(P(\mu))\}.$$

As in Theorem 2.4 for the Gershgorin set and Theorem 3.4 for the generalized Gershgorin set, we obtain a necessary condition for the isolated points of $\mathcal{B}_{i,j}(P)$.

THEOREM 4.4. *If μ is an isolated point of the (i,j)-th Brauer set $\mathcal{B}_{i,j}(P)$, $1 \le j < i \le n$, then*

(a) *μ is a root of $(P(\lambda))_{i,i}$ or $(P(\lambda))_{j,j}$, and*
(b) *μ is a common root of all polynomials $(P(\lambda))_{i,p}$, $p \in \mathcal{N}\backslash\{i\}$, or a common root of all polynomials $(P(\lambda))_{j,p}$, $p \in \mathcal{N}\backslash\{j\}$.*

Proof. For the sake of contradiction, assume that $(P(\mu))_{i,i} \neq 0$ and $(P(\mu))_{j,j} \neq 0$. Since μ is an isolated point, it belongs to the boundary of the Brauer set, and also there exists an $\varepsilon > 0$ such that the disk $D(\mu, \varepsilon) = \{\lambda \in \mathbb{C} : |\lambda - \mu| \le \varepsilon\}$ contains no other point of $\mathcal{B}_{i,j}(P)$. The Brauer set can be written as

$$\begin{aligned}\mathcal{B}_{i,j}(P) &= \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| \le r_i(P(\mu))r_j(P(\mu))\} \\ &= \left\{\mu \in \mathbb{C} : \frac{r_i(P(\mu))r_j(P(\mu))}{|(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}|} \ge 1\right\} \\ &= \left\{\mu \in \mathbb{C} : \log\frac{r_i(P(\mu))r_j(P(\mu))}{|(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}|} \ge 0\right\} \\ &= \left\{\mu \in \mathbb{C} : \log\frac{r_i(P(\mu))}{|(P(\mu))_{i,i}|} + \log\frac{r_j(P(\mu))}{|(P(\mu))_{j,j}|} \ge 0\right\}.\end{aligned}$$

Consider the function

$$\vartheta(\lambda) = \log\frac{r_i(P(\lambda))}{|(P(\lambda))_{i,i}|} + \log\frac{r_j(P(\lambda))}{|(P(\lambda))_{j,j}|}, \quad \lambda \in D(\mu, \varepsilon), \tag{4.11}$$

and observe that it is subharmonic (as a sum of subharmonic functions) and satisfies the maximum principle [2, 4]. By definition, $\vartheta(\lambda)$ is equal to zero on the boundary of $\mathcal{B}_{i,j}(P)$, nonnegative in the interior of $\mathcal{B}_{i,j}(P)$

and negative elsewhere. Since $\mu \in \partial\mathcal{B}_{i,j}(P)$, it follows

$$|(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| = r_i(P(\mu))r_j(P(\mu)) = \left(\sum_{p\,\in\mathcal{N}\backslash\{i\}} |(P(\mu))_{i,p}|\right)\left(\sum_{p\,\in\mathcal{N}\backslash\{j\}} |(P(\mu))_{j,p}|\right).$$

Hence, the function $\vartheta(\mu)$ is equal to zero at the center μ of $D(\mu,\varepsilon)$ and negative in the rest of the disk; this is a contradiction because $\vartheta(\lambda)$ satisfies the maximum principle. As a consequence,

$$0 = |(P(\mu))_{i,i}|\,|(P(\mu))_{j,j}| = \left(\sum_{p\,\in\mathcal{N}\backslash\{i\}} |(P(\mu))_{i,p}|\right)\left(\sum_{p\,\in\mathcal{N}\backslash\{j\}} |(P(\mu))_{j,p}|\right),$$

and the proof is complete. □

THEOREM 4.5. *Any point of the (i,j)-th Brauer set $\mathcal{B}_{i,j}(P)$, $1 \le j < i \le n$, has local dimension either 2 or 0 (isolated point); in other words, the Brauer set cannot have parts with (nontrivial) curves.*

Proof. The proof is similar to the proof of Theorem 2.5, using the (subharmonic) function $\vartheta(\lambda)$ in (4.11) instead of $\varphi(\lambda)$ in (2.2). □

Similarly to Theorem 2.6 for the Gershgorin set and Theorem 3.6 for the generalized Gershgorin set, and recalling the definition of β_i and $\overline{\beta}_i$ in (2.3), we obtain necessary and sufficient conditions for $\mathcal{B}_{i,j}(P)$ to be bounded.

By Theorem 4.4, if $\mu \in \mathbb{C}$ is an isolated point of $\mathcal{B}_{i,j}(P)$ for some $i,j \in \mathcal{N}$ with $j < i$, then μ is a root of $(P(\lambda))_{i,i}$ or $(P(\lambda))_{j,j}$, and a common root of polynomials $(P(\lambda))_{i,p}$, $p \in \mathcal{N}\backslash\{i\}$, or of polynomials $(P(\lambda))_{j,p}$, $p \in \mathcal{N}\backslash\{j\}$. As a consequence, if $i \in \beta_i$ and $j \in \beta_j$, or $\beta_i\backslash\{i\} \neq \emptyset$ and $\beta_j\backslash\{j\} \neq \emptyset$, then the origin is not an isolated point of $\mathcal{B}_{i,j}(\hat{P})$, i.e., $\mathcal{B}_{i,j}(P)$ is not the union of a bounded set and ∞.

THEOREM 4.6. *Suppose that for some integers $i,j \in \mathcal{N}$ with $j < i$, it holds that $i \in \beta_i$ and $j \in \beta_j$, or $\beta_i\backslash\{i\} \neq \emptyset$ and $\beta_j\backslash\{j\} \neq \emptyset$.*

(i) *If $i \in \beta_i$ and $j \in \beta_j$, then $\mathcal{B}_{i,j}(P)$ is unbounded if and only if $0 \in \mathcal{B}_{i,j}(A_m)$.*
(ii) *If $i \in \overline{\beta}_i$ and $j \in \overline{\beta}_j$, then $\mathcal{B}_{i,j}(P)$ is unbounded and $0 \in \mathcal{B}_{i,j}(A_m)$.*

Proof. (i) Suppose that $i \in \beta_i$, i.e., $(A_m)_{i,i} \neq 0$, and $j \in \beta_j$, i.e., $(A_m)_{j,j} \neq 0$.

Let $\mathcal{B}_{i,j}(P)$ be unbounded. By hypothesis, the origin is not an isolated point of $\mathcal{B}_{i,j}(\hat{P})$, and there is a sequence $\{\mu_l\}_{l\in\mathbb{N}}$ in $\mathcal{B}_{i,j}(P)\backslash\{0\}$ such that $|\mu_l| \to +\infty$. Then, for every $l \in \mathbb{N}$,

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\mu_l^k\right|\left|\sum_{k=0}^{m}(A_k)_{j,j}\mu_l^k\right| \le \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\mu_l^k\right|\right)\left(\sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\mu_l^k\right|\right),$$

or

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\frac{\mu_l^k}{\mu_l^m}\right|\left|\sum_{k=0}^{m}(A_k)_{j,j}\frac{\mu_l^k}{\mu_l^m}\right| \le \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\frac{\mu_l^k}{\mu_l^m}\right|\right)\left(\sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\frac{\mu_l^k}{\mu_l^m}\right|\right).$$

As $l \to +\infty$, it follows

$$|(A_m)_{i,i}|\;|(A_m)_{j,j}| - \left(\sum_{p\in\beta_i\backslash\{i\}}|(A_m)_{i,p}|\right)\left(\sum_{p\in\beta_j\backslash\{j\}}|(A_m)_{j,p}|\right) \le\; 0,$$

and hence, $0 \in \mathcal{B}_{i,j}(A_m)$.

For the converse, suppose that $0 \in \mathcal{B}_{i,j}(A_m)$ (or equivalently, $|(A_m)_{i,i}|\,|(A_m)_{j,j}| \le r_i(A_m) r_j(A_m)$). Then $0 \in \mathcal{B}_{i,j}(\hat{P})$ and, by definition, $\infty \in \mathcal{B}_{i,j}(P)$.

(ii) Suppose that $i \in \overline{\beta}_i$ and $j \in \overline{\beta}_j$, i.e., $(A_m)_{i,i} = (A_m)_{j,j} = 0$. Then, it is clear that $0 \in \mathcal{B}_{i,j}(A_m)$. Moreover,

$$\begin{aligned}
\mathcal{B}_{i,j}(P)\backslash\{0\} = \Bigg\{ \mu \in \mathbb{C}\backslash\{0\} : & \left|\sum_{k=0}^{m-1} (A_k)_{i,i}\mu^k\right| \left|\sum_{k=0}^{m-1} (A_k)_{j,j}\mu^k\right| \\
& \le \left(\sum_{p\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m} (A_k)_{i,p}\mu^k\right| \right) \left(\sum_{p\in\mathcal{N}\backslash\{j\}} \left|\sum_{k=0}^{m} (A_k)_{j,p}\mu^k\right| \right) \Bigg\} \\
= \Bigg\{ \mu \in \mathbb{C}\backslash\{0\} : & \left|\sum_{k=0}^{m-1} (A_k)_{i,i}\frac{\mu^k}{\mu^m}\right| \left|\sum_{k=0}^{m-1} (A_k)_{j,j}\frac{\mu^k}{\mu^m}\right| \\
& - \left(\sum_{p\in\mathcal{N}\backslash\{i\}} \left|\sum_{k=0}^{m} (A_k)_{i,p}\frac{\mu^k}{\mu^m}\right| \right) \left(\sum_{p\in\mathcal{N}\backslash\{j\}} \left|\sum_{k=0}^{m} (A_k)_{j,p}\frac{\mu^k}{\mu^m}\right| \right) \le 0 \Bigg\},
\end{aligned}$$

where at least one of the coefficients $(A_m)_{i,p}$ ($p \in \mathcal{N}\backslash\{i\}$) and one of the coefficients $(A_m)_{j,p}$ ($p \in \mathcal{N}\backslash\{j\}$) are nonzero. As a consequence, for "large enough" $|\mu|$, μ lies in $\mathcal{B}_{i,j}(P)$. Thus, there exists a real number $M > 0$ such that every scalar $\mu \in \mathbb{C}$ with $|\mu| \ge M$ lies in $\mathcal{B}_{i,j}(P)$, i.e., $\{\mu \in \mathbb{C} : |\mu| \ge M\} \subseteq \mathcal{B}_{i,j}(P)$. ◻

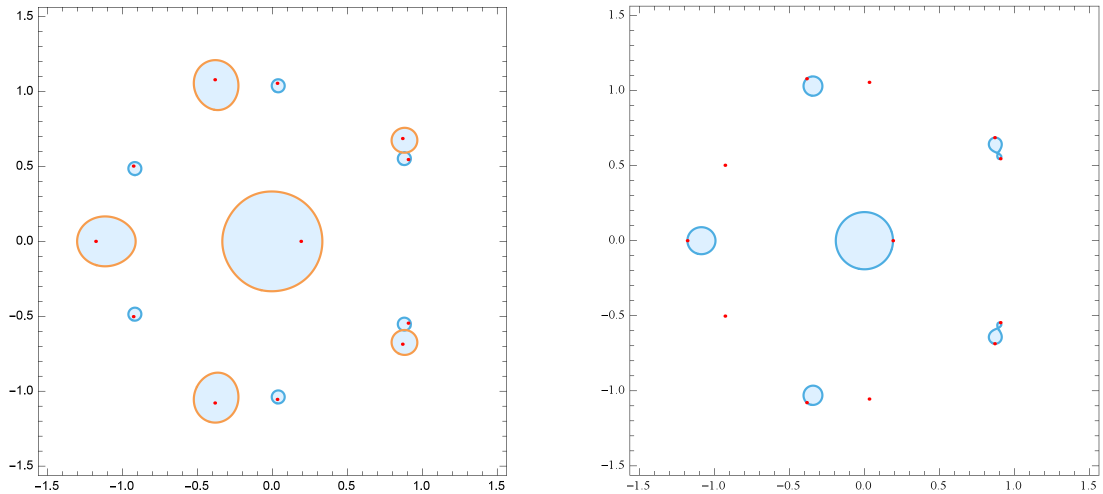
Theorems 2.9 and 3.7 can be easily extended to the case of the Brauer set.

THEOREM 4.7. *Suppose that for every $i, j \in \mathcal{N}$ with $j < i$, $i \in \beta_i$ and $j \in \beta_j$. If the Brauer set $\mathcal{B}(P)$ is bounded, then the number of connected components of $\mathcal{B}(P)$ is less than or equal to nm. Moreover, if $\mathcal{G}$ is a connected component of $\mathcal{B}(P)$ constructed by ξ connected components of (i, j)-th Brauer sets, then the total number of roots of the polynomials $(P(\lambda))_{1,1}, (P(\lambda))_{2,2}, \ldots, (P(\lambda))_{n,n}$ in $\mathcal{G}$ is at least ξ and it is equal to the number of eigenvalues of $P(\lambda)$ in $\mathcal{G}$, counting multiplicities.*

Proof. By hypothesis, for every $i, j \in \mathcal{N}$ with $j < i$, the polynomial $(P(\lambda))_{i,i}(P(\lambda))_{j,j}$ is of degree $2m$ (i.e., it has exactly $2m$ roots, counting multiplicities). Based on the continuity of the eigenvalues of a matrix polynomial with respect to its matrix coefficients and the boundedness of $\mathcal{B}(P)$, one can complete the proof, following the arguments of the proof of Theorem 2.9. ◻

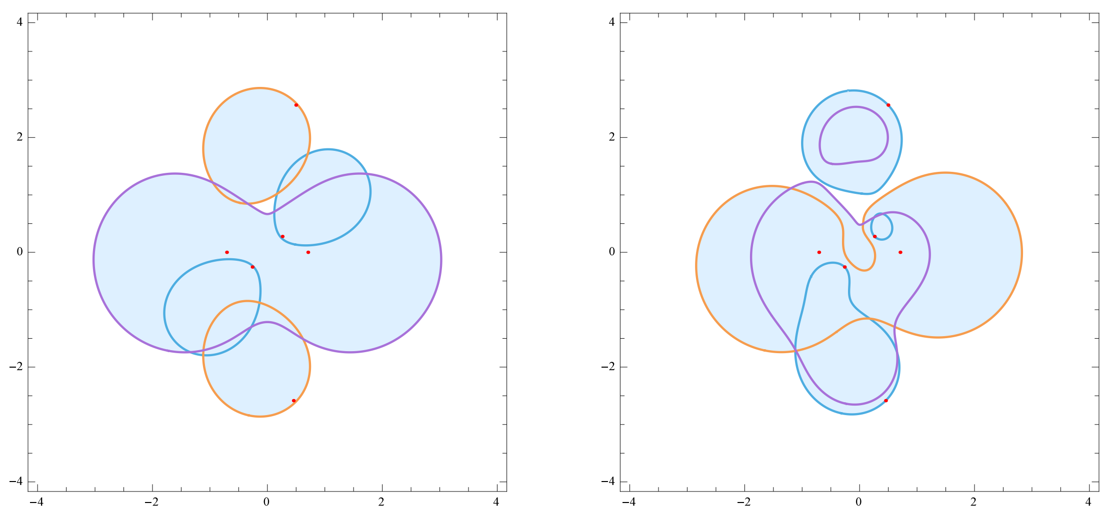
4.3. Examples. In the case of a 2×2 matrix polynomial $P(\lambda)$, the relation $\det P(\lambda) = 0$ apparently yields $|(P(\lambda))_{1,1}|\,|(P(\lambda))_{2,2}| = |(P(\lambda))_{1,2}|\,|(P(\lambda))_{2,1}|$, and consequently, the eigenvalues appear on the boundary of the Brauer set.

EXAMPLE 4.8. Consider the 2×2 matrix polynomial $P(\lambda)$ of Example 2.10. In Figure 5, the Gershgorin set $G(P)$ is illustrated in part (a) and the Brauer set $\mathcal{B}(P) = \mathcal{B}_{2,1}(P)$ is illustrated in part (b). It is clearly verified that the Brauer set contains the eigenvalues of $P(\lambda)$ and lies in the Gershgorin set. As expected, the eigenvalues lie on the boundary of the Bruer set. Moreover, Proposition 4.2 (iv) (symmetry with respect to the real axis), Theorem 4.4 (isolated points -0.9305 ± 0.5028i and 0.0298 ± 1.0573i), Theorem 4.6 (boundedness conditions) and Theorem 4.7 (distribution of eigenvalues in the connected components) are also confirmed.



(a) The Gershgorin set. (b) The Brauer set.

Figure 5: Comparing the Gershgorin set and the Brauer set.



(a) The Gershgorin set. (b) The Brauer set.

Figure 6: Comparing the Gershgorin set and the Brauer set.

Example 4.9. Consider the 3×3 matrix polynomial $P(\lambda)$ of Example 3.9. In Figure 6, the Gershgorin set $G(P)$ is illustrated in part (a) and the Brauer set $\mathcal{B}(P)$ is illustrated in part (b). It is once again verified that the Brauer set contains the eigenvalues of $P(\lambda)$ and lies in the Gershgorin set.

5. The Dashnic-Zusmanovich set. In this section, we introduce the Dashnic-Zusmanovich set of matrix polynomials.

5.1. Definitions. The *Dashnic-Zusmanovich set* of a matrix $A \in \mathbb{C}^{n\times n}$ is defined as $\mathcal{D}(A) = \bigcap_{i\in\mathcal{N}} \bigcup_{j\in\mathcal{N}\backslash\{i\}} \mathcal{D}_{i,j}(A)$, where $\mathcal{D}_{i,j}(A) = \{\mu \in \mathbb{C} : |\mu - (A)_{i,i}| (|\mu - (A)_{j,j}| - r_j(A) + |(A)_{j,i}|) \leq r_i(A)|(A)_{j,i}|\}$ for distinct integers $i, j \in \mathcal{N}$. The Dashnic-Zusmanovich set $\mathcal{D}(A)$ is determined by $n(n-1)$ oval sets, lies in the Brauer set $\mathcal{B}(A)$, and contains all the eigenvalues of matrix A [6, 21].

For a matrix polynomial $P(\lambda)$ as in (1.1), we define

$$\begin{aligned}\mathcal{D}_{i,j}(P) &= \{\mu \in \mathbb{C} : 0 \in \mathcal{D}_{i,j}(P(\mu))\}\\ &= \{\mu \in \mathbb{C} : |(P(\mu))_{i,i}| (|(P(\mu))_{j,j}| - r_j(P(\mu)) + |(P(\mu))_{j,i}|) \leq r_i(P(\mu))|(P(\mu))_{j,i}|\}\end{aligned}$$

for distinct integers $i, j \in \mathcal{N}$, the union

$$\mathcal{D}_i(P) = \bigcup_{j\in\mathcal{N}\backslash\{i\}} \mathcal{D}_{i,j}(P), \quad i \in \mathcal{N},$$

and the *Dashnic-Zusmanovich set* of $P(\lambda)$

$$\mathcal{D}(P) = \{\mu \in \mathbb{C} : 0 \in \mathcal{D}(P(\mu))\} = \bigcap_{i\in\mathcal{N}} \bigcup_{j\in\mathcal{N}\backslash\{i\}} \mathcal{D}_{i,j}(P) = \bigcap_{i\in\mathcal{N}} \mathcal{D}_i(P).$$

We also say that $\mu = \infty$ lies in $\mathcal{D}_{i,j}(P)$ (resp., in $\mathcal{D}_i(P)$, or in $\mathcal{D}(P)$) exactly when 0 lies in $\mathcal{D}_{i,j}(\hat{P})$ (resp., in $\mathcal{D}_i(\hat{P})$, or in $\mathcal{D}(\hat{P})$).

THEOREM 5.1. *The Dashnic-Zusmanovich set of a matrix polynomial $P(\lambda)$ is a subset of the Brauer set of $P(\lambda)$ that contains all the (finite and infinite) eigenvalues of $P(\lambda)$.*

Proof. By definition, a scalar $\mu \in \mathbb{C}$ lies in $\mathcal{D}(P)$ (resp., in $\sigma(P)$, or in $\mathcal{B}(P)$) if and only if 0 lies in $\mathcal{D}(P(\mu))$ (resp., in $\sigma(P(\mu))$, or in $\mathcal{B}(P(\mu))$). Since $\sigma(P(\mu)) \subseteq \mathcal{D}(P(\mu)) \subseteq \mathcal{B}(P(\mu))$ and $\sigma(\hat{P}(\mu)) \subseteq \mathcal{D}(\hat{P}(\mu)) \subseteq \mathcal{B}(\hat{P}(\mu))$ [6, 21], the proof follows readily. ∎

5.2. Basic properties. Let $P(\lambda)$ be an $n \times n$ matrix polynomial as in (1.1). It is easy to verify the following properties.

PROPOSITION 5.2. *Consider two distinct integers $i, j \in \mathcal{N}$.*

(i) *$\mathcal{D}_{i,j}(P)$ and $\mathcal{D}_i(P)$ are closed subsets of $\mathbb{C}$.*
(ii) *For any scalar $b \in \mathbb{C}\backslash\{0\}$, consider the matrix polynomials $Q_1(\lambda) = P(b\lambda)$, $Q_2(\lambda) = bP(\lambda)$ and $Q_3(\lambda) = P(\lambda + b)$. Then $\mathcal{D}_{i,j}(Q_1) = b^{-1}\mathcal{D}_{i,j}(P)$, $\mathcal{D}_{i,j}(Q_2) = \mathcal{D}_{i,j}(P)$ and $\mathcal{D}_{i,j}(Q_3) = \mathcal{D}_{i,j}(P) - b$.*
(iii) *If the (i, i)-th or the the (j, j)-th entry of $P(\lambda)$ is (identically) zero, then $\mathcal{D}_{i,j}(P) = \mathcal{D}_i(P) = \mathbb{C}$.*
(iv) *If all the coefficient matrices $A_0, A_1, \ldots, A_m$ have their i-th and j-th rows real, then $\mathcal{D}_{i,j}(P)$ is symmetric with respect to the real axis.*

Proof. The proof is similar to the proof of Proposition 2.2. ∎

Similarly to Theorems 2.6, 3.6 and 4.6, we obtain necessary and sufficient conditions for the Dashnic-Zusmanovich set to be bounded.

THEOREM 5.3. *Suppose that for some distinct $i, j \in \mathcal{N}$, the sets β_i and β_j are nonempty, and the origin is not an isolated point of $\mathcal{D}_{i,j}(\hat{P})$, i.e., $\mathcal{D}_{i,j}(P)$ is not the union of a bounded set and ∞.*

(i) *If $i \in \beta_i$ and $j \in \beta_j$, then $\mathcal{D}_{i,j}(P)$ is unbounded if and only if $0 \in \mathcal{D}_{i,j}(A_m)$.*
(ii) *If $i \in \overline{\beta}_i \cap \beta_j$ and $j \in \overline{\beta}_j$, then $\mathcal{D}_{i,j}(P)$ is unbounded and $0 \in \mathcal{D}_{i,j}(A_m)$.*

Proof. (i) Suppose that $i \in \beta_i$, i.e., $(A_m)_{i,i} \neq 0$ and $j \in \beta_j$, i.e., $(A_m)_{j,j} \neq 0$.

Let $\mathcal{D}_{i,j}(P)$ be unbounded. Since the origin is not an isolated point of $\mathcal{D}_{i,j}(\hat{P})$, there is a sequence $\{\mu_l\}_{l\in\mathbb{N}}$ in $\mathcal{D}_{i,j}(P)\backslash\{0\}$ such that $|\mu_l| \to +\infty$. Then, for every positive integer l,

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\mu_l^k\right| \left(\left|\sum_{k=0}^{m}(A_k)_{j,j}\mu_l^k\right| - \sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\mu_l^k\right| + \left|\sum_{k=0}^{m}(A_k)_{j,i}\mu_l^k\right|\right)$$
$$\leq \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\mu_l^k\right|\right)\left|\sum_{k=0}^{m}(A_k)_{j,i}\mu_l^k\right|,$$

or

$$\left|\sum_{k=0}^{m}(A_k)_{i,i}\frac{\mu_l^k}{\mu_l^m}\right| \left(\left|\sum_{k=0}^{m}(A_k)_{j,j}\frac{\mu_l^k}{\mu_l^m}\right| - \sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\frac{\mu_l^k}{\mu_l^m}\right| + \left|\sum_{k=0}^{m}(A_k)_{j,i}\frac{\mu_l^k}{\mu_l^m}\right|\right)$$
$$\leq \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\frac{\mu_l^k}{\mu_l^m}\right|\right)\left|\sum_{k=0}^{m}(A_k)_{j,i}\frac{\mu_l^k}{\mu_l^m}\right|,$$

As $l \to +\infty$, it follows

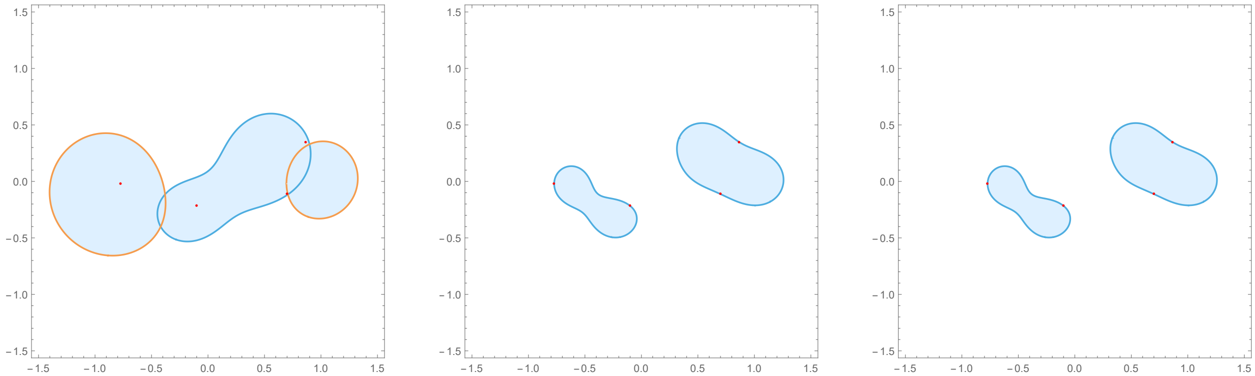
$$|(A_m)_{i,i}|\left(|(A_m)_{j,j}| - \sum_{p\in\beta_i\backslash\{i\}}|(A_m)_{i,p}| + |(A_m)_{j,i}|\right) \leq \left(\sum_{p\in\beta_i\backslash\{i\}}|(A_m)_{i,p}|\right)|(A_m)_{j,i}|,$$

and thus, $0 \in \mathcal{D}_{i,j}(A_m)$.

For the converse, suppose that $0 \in \mathcal{D}_{i,j}(A_m)$ (or equivalently, $|(A_m)_{i,i}|(|(A_m)_{j,j}| - r_j(A_m) + |(A_m)_{j,i}|) \leq |(A_m)_{j,i}|r_j(A_m)$). Then $0 \in \mathcal{D}_{i,j}(\hat{P})$ and, by definition, $\infty \in \mathcal{D}_{i,j}(P)$.

(ii) Suppose that $i \in \overline{\beta}_i$, $j \in \overline{\beta}_j$ and $i \in \beta_j$, i.e., $(A_m)_{i,i} = 0$, $(A_m)_{j,j} = 0$ and $(A_m)_{j,i} \neq 0$. Then, it is clear that $0 \in \mathcal{D}_{i,j}(A_m)$, and

$$\mathcal{D}_{i,j}(P)\backslash\{0\} = \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\mu^k\right| \left(\left|\sum_{k=0}^{m-1}(A_k)_{j,j}\mu^k\right| - \sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\mu^k\right|\right.\right.$$
$$\left.\left. + \left|\sum_{k=0}^{m}(A_k)_{j,i}\mu^k\right|\right) \leq \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\mu^k\right|\right)\left|\sum_{k=0}^{m}(A_k)_{j,i}\mu^k\right|\right\}$$
$$= \left\{\mu \in \mathbb{C}\backslash\{0\} : \left|\sum_{k=0}^{m-1}(A_k)_{i,i}\frac{\mu^k}{\mu^{m-1}}\right| \left(\left|\sum_{k=0}^{m-1}(A_k)_{j,j}\frac{\mu^k}{\mu^{m-1}}\right| - \sum_{p\in\mathcal{N}\backslash\{j\}}\left|\sum_{k=0}^{m}(A_k)_{j,p}\frac{\mu^k}{\mu^{m-1}}\right|\right.\right.$$
$$\left.\left. + \left|\sum_{k=0}^{m}(A_k)_{j,i}\frac{\mu^k}{\mu^{m-1}}\right|\right) \leq \left(\sum_{p\in\mathcal{N}\backslash\{i\}}\left|\sum_{k=0}^{m}(A_k)_{i,p}\frac{\mu^k}{\mu^{m-1}}\right|\right)\left|\sum_{k=0}^{m}(A_k)_{j,i}\frac{\mu^k}{\mu^{m-1}}\right|\right\},$$



(a) The Gershgorin set. (b) The Brauer set. (c) The D.-Z. set.

Figure 7: Comparing the Gershgorin set, the Brauer set and the D.-Z. set.

where $(A_m)_{j,i}$ and at least one of the coefficients $(A_m)_{i,p}$, $p \in \mathcal{N}\backslash\{i\}$, is nonzero. As a consequence, for "large enough" $|\mu|$, μ lies in $\mathcal{D}_{i,j}(P)$. Thus, there exists a real number $M > 0$ such that every scalar $\mu \in \mathbb{C}$ with $|\mu| \geq M$ lies in $\mathcal{D}_{i,j}(P)$, i.e., $\{\mu \in \mathbb{C} : |\mu| \geq M\} \subseteq \mathcal{D}_{i,j}(P)$. □

5.3. Examples.

Example 5.4. Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 8\mathrm{i}\lambda^2 - 2\mathrm{i}\lambda + 2 & 2\mathrm{i}\lambda^2 + \mathrm{i}\lambda + (1+2\mathrm{i}) \\ -3\mathrm{i}\lambda^2 + 5\mathrm{i}\lambda + (1+\mathrm{i}) & -6\mathrm{i}\lambda^2 + 3\mathrm{i}\lambda + 4\mathrm{i} \end{bmatrix}.$$

In Figure 7, the Gershgorin set, the Brauer set and the Dashnic-Zusmanovich (D.-Z.) set of $P(\lambda)$ are drawn. Notice that the Brauer set and the Dashnic-Zusmanovich set are the same (because the matrix polynomial has only 2 rows) and lie in the Gershgorin set.

In the following two examples, we consider two 3×3 quadratic matrix polynomials, with bounded (Example 5.5) and unbounded (Example 5.6) eigenvalues' inclusion sets.

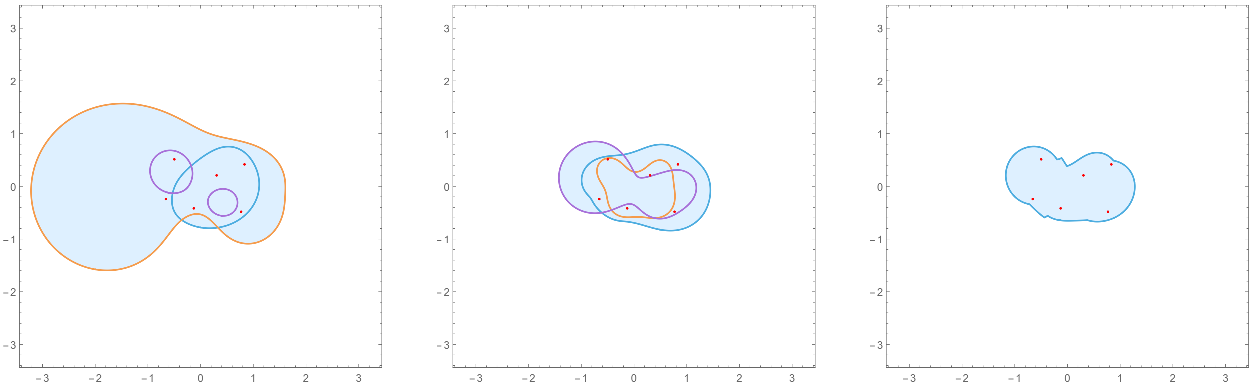
Example 5.5. Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 8\mathrm{i}\lambda^2 - 2\mathrm{i}\lambda + 2 & 2\mathrm{i}\lambda^2 + \mathrm{i}\lambda + (1+2\mathrm{i}) & (-1+\mathrm{i})\lambda^2 + \lambda + 2 \\ -3\mathrm{i}\lambda^2 + 5\mathrm{i}\lambda + (1+\mathrm{i}) & -8\mathrm{i}\lambda^2 + 3\mathrm{i}\lambda + 4\mathrm{i} & (2-2\mathrm{i})\lambda^2 - 4\lambda - 5\mathrm{i} \\ (0.8-\mathrm{i})\lambda^2 + \mathrm{i}\lambda + (1-\mathrm{i}) & 0.6\mathrm{i}\lambda^2 - \mathrm{i}\lambda & (6-2\mathrm{i})\lambda^2 + 2\mathrm{i} \end{bmatrix}.$$

The Gershgorin set, the Brauer set and the Dashnic-Zusmanovich set are illustrated in Figure 8. It is clear that these three sets are bounded, confirming Theorems 2.6, 4.6 and 5.3, and $\mathcal{D}(P) \subseteq \mathcal{B}(P) \subseteq G(P)$.

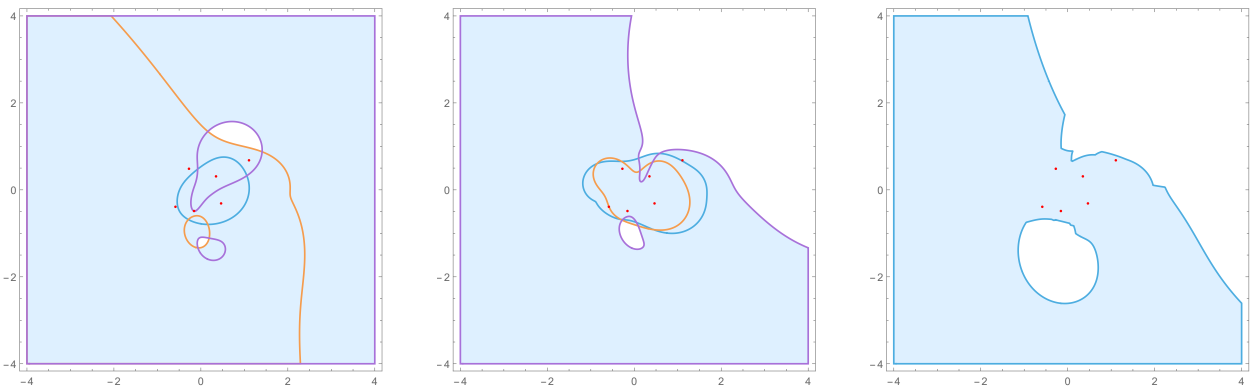
Example 5.6. Consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 8\mathrm{i}\lambda^2 - 2\mathrm{i}\lambda + 2 & 2\mathrm{i}\lambda^2 + \mathrm{i}\lambda + (1+2\mathrm{i}) & (-1+\mathrm{i})\lambda^2 + \lambda + 2 \\ -3\mathrm{i}\lambda^2 + 5\mathrm{i}\lambda + (1+\mathrm{i}) & -6\mathrm{i}\lambda^2 + 3\mathrm{i}\lambda + 4\mathrm{i} & (2-2\mathrm{i})\lambda^2 - 4\lambda - 5\mathrm{i} \\ (0.8-\mathrm{i})\lambda^2 + \mathrm{i}\lambda + (1-\mathrm{i}) & 6\mathrm{i}\lambda^2 - \mathrm{i}\lambda & (6-2\mathrm{i})\lambda^2 + 2\mathrm{i} \end{bmatrix}.$$



(a) The Gershgorin set. (b) The Brauer set. (c) The D.-Z. set.

Figure 8: Comparing the Gershgorin set, the Brauer set and the D.-Z. set.



(a) The Gershgorin set. (b) The Brauer set. (c) The D.-Z. set.

Figure 9: Comparing the Gershgorin set, the Brauer set and the D.-Z. set.

The Gershgorin set, the Brauer set and the Dashnic-Zusmanovich set are drawn in Figure 9, and they are unbounded, verifying Theorems 2.6, 4.6 and 5.3. Notice once again that $\mathcal{D}(P) \subseteq \mathcal{B}(P) \subseteq G(P)$.

Acknowledgments. The authors wish to thank the anonymous referees for valuable comments and suggestions.

REFERENCES

[1] D. Bindel and A. Hood. Localization theorems for nonlinear eigenvalue problems. *SIAM Journal on Matrix Analysis and Applications*, 34(6):1728–1749, 2013.

[2] S. Boyd and C.A. Desoer. Subharmonic functions and performance bounds on linear time-invariant feedback systems. *IMA Journal of Mathematical Control and Information*, 2:153–170, 1985.

[3] A. Brauer. Limits for the characteristic roots of a matrix II. *Duke Mathematical Journal*, 14:21–26, 1947.

[4] J.B. Conway. *Functions of One Complex Variable*. Springer-Verlag, New York, 1973.

[5] L. Cvetkovic, V. Kostic, and R.S. Varga. A new Geršgorin-type eigenvalue inclusion set. *Electronic Transactions on Numerical Analysis*, 18:73–80, 2004.

[6] L.S. Dashnic and M.S. Zusmanovich. K voprosu o lokalizacii harakteristicheskih chisel matricy. *Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki*, 10(6):1321–1327, 1970.

[7] T.W. Gamelin. *Complex Analysis*. Springer Science & Business Media, Springer-Verlag, New York, 2003.

[8] S. Gershgorin. Über die abgrenzung der eigenwerte einer matrix. *Izvestiya Akademii Nauk SSSR, Seriya Fizicheskaya-Matematicheskaya*, 6:749–754, 1931.

[9] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Society for Industrial and Applied Mathematics, Philadelphia, 2009.

[10] N.J. Higham and F. Tisseur. Bounds for eigenvalues of matrix polynomials. *Linear Algebra and its Aplications*, 358:5–22, 2003.

[11] W. Hurewicz and H. Wallman. *Dimension Theory* (revised edition). Princeton Mathematical Series, Vol. 4, Princeton University Press, Princeton, 1996.

[12] V. Kostic, L.J. Cvetkovic, and R.S. Varga. Geršgorin-type localizations of generalized eigenvalues. *Numerical Linear Algebra with Applications*, 16(11/12):883–898, 2009.

[13] V. Kostić and D. Gardašević. On the Geršgorin-type localizations for nonlinear eigenvalue problems. *Applied Mathematics and Computation*, 337:179–189, 2018.

[14] P. Lancaster. *Lambda-Matrices and Vibrating Systems*. Dover Publications, 2002.

[15] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*, second edition. Academic Press, Orlando, 1985.

[16] A.S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils*. Translations of Mathematical Monographs, Vol. 71, American Mathematical Society, Providence, 1988.

[17] M.F. Neuts. *Matrix-Geometric Solutions in Stochastic Models*. Springer, Berlin, 1984.

[18] A.M. Ostrowski. Über das Nichtverschwinden einer Klasse von Determinanten und die Lokalisierung der charakteristischen Wurzeln von Matrizen. *Compositio Mathematica*, 9:209–226, 1951.

[19] O. Taussky. A recurring theorem on determinants. *American Mathematical Monthly*, 56:672–676, 1949.

[20] O. Taussky. Bounds for characteristic roots of matrices. *Duke Mathematical Journal*, 15:1043–1044, 1948.

[21] R.S. Varga. *Geršgorin and His Circles*. Springer Series in Computational Mathematics, Vol. 36, Springer-Verlag, Berlin, 2004.