NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2012; **19**:728–741
Published online 3 November 2011 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/nla.801

Localization of generalized eigenvalues by Cartesian ovals

V. Kostić [1,*,†], R. S. Varga [2] and L. Cvetković [1]

[1] *Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad, Novi Sad, Serbia*
[2] *Department of Mathematical Sciences, Kent State University, Cleveland, OH, USA*

SUMMARY

In this paper, we consider the localization of generalized eigenvalues, and we discuss ways in which the Gersgorin set for generalized eigenvalues can be approximated. Earlier, Stewart proposed an approximation using a chordal metric. We will obtain here an improved approximation, and using the concept of generalized diagonal dominance, we prove that the new approximation has some of the basic properties of the original Geršgorin set, which makes it a handy tool for generalized eigenvalue localization. In addition, an isolation property is proved for both the generalized Geršgorin set and its approximation. Copyright © 2011 John Wiley & Sons, Ltd.

Received 1 May 2010; Revised 30 May 2011; Accepted 20 June 2011

KEY WORDS: Geršgorin sets; generalized eigenvalues; H-matrices

1. INTRODUCTION

The paper consists of four sections. The first one gives preliminary material. The second reviews the Geršgorin set for generalized eigenvalues, in the form it was obtained by Stewart in [1] and by Kostić, Cvetković, and Varga in [2]. In the same section, we establish an important property that can be used to isolate the generalized eigenvalues using the Geršgorin set, and in the third section, we derive a new approximation of the obtained generalized Geršgorin set, which is more suitable for practical use than the original set and gives better results than the one known from the pioneering work of Stewart and Sun [3, Corollary VI.2.5]. Finally, in the last section, using sets dependent on an arbitrary complex parameter, we discuss improvements in the approximation.

2. GENERALIZED EIGENVALUES

Given two arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geqslant 1$, the family of matrices $A - zB$, parameterized by the complex number z, is called a **matrix pencil**. This matrix pencil $A - zB$ will also be considered as a matrix pair (A, B). By $\mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, we denote the set of all matrix pairs of square matrices of the size n.

Definition 2.1. Given arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geqslant 1$, then a matrix pair (A, B) is called **singular** if $\det(A - zB) = 0$ for all $z \in \mathbb{C}$. Otherwise, the pair (A, B) is called **regular**.

A sufficient condition for the singularity of a matrix pair (A, B) is that the matrices A and B have overlapping null spaces, meaning that there exists a nonzero vector **x** in the null spaces of both A

*Correspondence to: V. Kostić, Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad, Novi Sad, Serbia.
†E-mail: vkostic@dmi.uns.ac.rs

and B, so that for an arbitrary $z \in \mathbb{C}$, it follows that $(A - zB)\mathbf{x} = \mathbf{0}$. This sufficient condition is not a necessary one, as the following example shows.[‡]

Example 2.2
With

$$A := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then A and B do not have overlapping null spaces, but the pair (A, B) is singular.

We proceed with regular matrix pairs, and we define the concept of an eigenvalue of a matrix pair.

Definition 2.3. Given a regular matrix pair (A, B), if there exists a nonzero vector $\mathbf{v} \in \mathbb{C}^n$ and a scalar $\lambda \in \mathbb{C}$ such that $A\mathbf{v} = \lambda B\mathbf{v}$, then $\mathbf{v}$ is called an **eigenvector** of the matrix pair (A, B), and λ is called a **finite eigenvalue** of the matrix pair (A, B). Furthermore, if there exists a nonzero vector $\mathbf{v} \in \mathbb{C}^n$ such that $B\mathbf{v} = \mathbf{0}$, but $A\mathbf{v} \neq \mathbf{0}$, then we define $\lambda := \infty$ and write, by convention, $A\mathbf{v} = \lambda B\mathbf{v}$. In this case, λ is called an **infinite eigenvalue** of a matrix pair (A, B), and $\mathbf{v}$ is, again, the corresponding **eigenvector**. The term **eigenvalue** is used for both finite and infinite eigenvalues of the given matrix pair.

Having a regular matrix pair (A, B), then λ is a finite eigenvalue of the pair (A, B) if and only if $A - \lambda B$ is a singular matrix, that is, if $\det(A - \lambda B) = 0$. Thus, the previous definition can be expressed in terms of determinants.

Given a regular matrix pair (A, B), then

$$\det(A - zB) =: p(z), \tag{1}$$

where $p(z) \not\equiv 0$ is a polynomial in z, with degree at most n. From [3], it is known that the degree of the polynomial $p(z)$ is n if and only if B is nonsingular. This implies that if B is singular, then $p(z)$ is of degree r with $r < n$, so the number of the finite eigenvalues of the matrix pair (A, B) is r, and again, by convention, the remaining $n - r$ eigenvalues are set equal to ∞.

Having a regular matrix pair (A, B) and taking $B = I_n$, where I_n is the identity matrix of order n, we have that the polynomial $p(z)$ in Equation (1) is $p(z) = \det(A - zI_n)$, so all of its n zeros are the eigenvalues of the matrix A. Therefore, in the literature, the eigenvalues of matrix pairs are often called **generalized eigenvalues**, and the corresponding eigenvectors are called **generalized eigenvectors**.

Because generalized eigenvalues can be infinite, we have to work in the one-point compactification of the complex plane, called the **extended complex plane** $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$. In this topological space, for every closed unbounded subset $U \subseteq \mathbb{C}$, the standard compactification is given by $\widetilde{U} := U \cup \{\infty\}$, and for every compact subset $U \subset \mathbb{C}$, $\widetilde{U} = U$. While working in the extended complex plane, whenever there is no possibility of confusion, we will identify the set $\widetilde{U} \subseteq \mathbb{C}_\infty$ with the set U.

Although this approach is intuitive for graphing purposes, it has a few drawbacks when it is used in theoretical considerations. Indeed, a more convenient setup for dealing with infinite eigenvalues can be found in [3], where the main idea is to rewrite the equation $A\mathbf{x} = zB\mathbf{x}$ in its cross-product form

$$\beta A\mathbf{x} = \alpha B\mathbf{x} \tag{2}$$

Then, infinite eigenvalues correspond to nonzero pairs (α, β) for which $\beta = 0$, a case that is not essentially different from $\alpha = 0$ (i.e., $z = 0$). Obviously, if the pair (α, β) satisfies Equation (2), then so does $\tau(\alpha, \beta)$, for any $\tau \in \mathbb{C}$. Consequently, the generalized eigenvalues are characterized by

$$\langle \alpha, \beta \rangle := \{\tau(\alpha, \beta) : \tau \in \mathbb{C} \setminus \{0\}\}, \tag{3}$$

where $(\alpha, \beta) \neq (0, 0)$.

[‡]We thank a referee for this comment.

DOI: 10.1002/nla

Note that in this setting, if $\beta \neq 0$, $z = \alpha\beta^{-1}$ is a finite eigenvalue, whereas the case $\beta = 0$ corresponds to an infinite eigenvalue $z = \infty$.

It is easy to see that the set of all such objects can also be represented on the Riemann sphere.

Because we are interested in obtaining graphable areas in the complex plane, in the sequel we will use the first approach based on a one-point compactification.

Definition 2.4. Given a regular matrix pair (A, B), the collection of all eigenvalues of the pair (A, B) is called the **spectrum** of the matrix pair (A, B), and it is denoted by

$$\sigma(A, B) := \begin{cases} \{z \in \mathbb{C} : \det(zB - A) = 0\}, & \text{if } B \text{ is nonsingular,} \\ \{z \in \mathbb{C} : \det(zB - A) = 0\} \cup \{\infty\}, & \text{if } B \text{ is singular.} \end{cases} \quad (4)$$

The set $\sigma_F(A, B) := \{z \in \mathbb{C} : \det(zB - A) = 0\}$ is called the **finite spectrum** of the pair (A, B).

Clearly, if $B = I_n$, then the spectrum of the matrix pair (A, B) reduces to the standard spectrum of A, that is, $\sigma(A, I_n) = \sigma(A)$. So again, the word **generalized spectrum** is used to refer to the spectrum of a matrix pair.

If B is a nonsingular matrix, it is easy to see that $0 = \det(A - zB) = \det(B^{-1}A - zI_n)$, so that in this case, the spectrum of the matrix pair (A, B) is equal to its finite spectrum, and it reduces to the ordinary spectrum of the matrix $B^{-1}A$. This concept of conversion of the generalized spectrum into the ordinary spectrum is used as a basis for many numerical methods. But as a consequence of rounding errors when B is ill conditioned, these methods can fail. Thus, the direct treatment of the generalized spectra is an important topic [3].

3. GERŠGORIN'S THEOREM FOR GENERALIZED EIGENVALUES

The localization of the eigenvalues of a given matrix, by means of Geršgorin-type methods, as in [4, 5], has been a useful tool in determining information about the *actual* spectrum of the given matrix, *before* it is calculated. One would expect that a similar approach could be used for the spectrum of a given matrix pair (A, B), but this topic has not been actively considered, until very recently. Actually, an extension of Geršgorin's theorem to the concept of matrix pairs was done by Stewart in 1975 in [1], but since then, not much has been done to examine the behavior of the obtained sets and to improve localization results. A study of generalized eigenvalue localization sets in terms of perturbations, through spectral value sets, was done by Karow in [6]. But the simplicity and elegance one finds in the original Geršgorin set has somehow been lost in the generalized eigenvalue case. The very recent paper by Kostić, Cvetković and Varga [2] examined the Geršgorin set for generalized eigenvalues, and we continue with a review of these results.

Definition 3.1. Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the set $\Gamma(A, B)$, defined as $\Gamma(A, B) = \bigcup_{i \in N} \Gamma_i(A, B)$, where

$$\Gamma_i(A, B) := \{z \in \mathbb{C} : |zb_{i,i} - a_{i,i}| \leqslant \sum_{j \in N \setminus \{i\}} |zb_{i,j} - a_{i,j}|\}, \quad (\text{all } i \in N := \{1, 2, ..., n\}), \quad (5)$$

is called the **generalized Geršgorin set** of the matrix pair (A, B).

The generalized Geršgorin set, as was the case for the Geršgorin set, could be derived from the class of strictly diagonally dominant (SDD) matrices, defined below, using the equivalence principle stated in [5].

Definition 3.2. Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ be an arbitrary matrix. If

$$|a_{i,i}| > r_i(A) := \sum_{j \in N \setminus \{i\}} |a_{i,j}|, \quad (\text{all } i \in N), \quad (6)$$

then A is called an **SDD** matrix.

DOI: 10.1002/nla

That SDD matrices are nonsingular is a result of the famous Lévy–Desplanques theorem [5, Theorem 1.4], which is in fact equivalent, from [2, Theorem 7], to the following Theorem:

Theorem 3.3
Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the spectrum of the matrix pair (A, B) belongs to the generalized Geršgorin set of the matrix pair (A, B), that is, the following inclusion holds:

$$\sigma(A, B) \subseteq \Gamma(A, B),^{\S} \tag{7}$$

where $\Gamma(A, B)$ is given in Definition 3.1.

On inspecting the form of the generalized Geršgorin 'disks' of Equation (5), the following properties were established in [2].

Theorem 3.4
Let $A, B \in \mathbb{C}^{n,n}$, with $n \geqslant 2$. Then, the following statements hold:

1. Let $i \in N$ be such that, for at least one $j \in N$, $b_{i,j} \neq 0$. Then, the ith generalized Geršgorin set $\Gamma_i(A, B)$ of Equation (5) is a bounded set in the complex plane $\mathbb{C}$ if and only if $|b_{i,i}| > r_i(B)$.
2. The generalized Geršgorin set $\Gamma(A, B)$ is a compact set in $\mathbb{C}$ if and only if B is an SDD matrix.
3. The ith generalized Geršgorin set $\Gamma_i(A, B)$, given in Equation (5), contains zero if and only if $|a_{i,i}| \leqslant r_i(A)$.
4. The generalized Geršgorin set $\Gamma(A, B)$ contains zero if and only if A is not an SDD matrix.
5. If there exists an $i \in N$ such that both $b_{i,i} = 0$ and $|a_{i,i}| \leqslant r_i(A)$, then $\Gamma_i(A, B)$, and consequently, $\Gamma(A, B)$ are the entire complex plane.

A very useful tool in eigenvalue localization is the fact that, having k Geršgorin disks of a given matrix whose union is a set disjoint from all of the remaining disks, exactly k eigenvalues are contained in this union of k disks. For the generalized Geršgorin set, this property has not been stated. We continue by giving the analog, of this result, for generalized eigenvalues.

Theorem 3.5
(Isolation property) Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geqslant 2$, if there exist closed sets $U, V \subseteq \mathbb{C}$ such that

$$\widetilde{U} \cap \widetilde{V} = \emptyset \text{ and } \Gamma(A, B) = U \cup V, \tag{8}$$

then one of them, say U, is compact in $\mathbb{C}$, and it contains exactly

$$\left|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i} b_{i,i}^{-1} \in U\right\}\right|$$

finite eigenvalues of the matrix pair (A, B).

Proof
First note that from Equation (8), both sets $\widetilde{U}$ and $\widetilde{V}$ cannot contain the point at infinity. Next, let $D_A := \mathrm{diag}(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$ and $D_B := \mathrm{diag}(b_{1,1}, b_{2,2}, \ldots, b_{n,n})$. With the splittings of the matrices A and B, defined by

$$A = D_A - F_A \text{ and } B = D_B - F_B,$$

consider the family of matrices

$$A(t) := D_A - tF_A \text{ and } B(t) := D_B - tF_B, \text{ for any } t \text{ with } 0 \leqslant t \leqslant 1.$$

§Here, the set $\Gamma(A, B)$ is identified with its one-point compactification.

DOI: 10.1002/nla

First, we prove that $\Gamma(A(t), B(t)) \subseteq \Gamma(A, B)$, for all $t \in [0, 1]$. Take an arbitrary $t \in [0, 1]$ and an arbitrary $z \in \Gamma(A(t), B(t))$. Then, $zB(t) - A(t)$ is not an SDD matrix. Assuming, on the contrary, that $z \notin \Gamma(A, B)$, we have that $zB - A$ is SDD. Because

$$t|zB - A| = t(|zD_B - D_A| + |zF_B - F_A|) \pm |zD_B - D_A| = |zB(t) - A(t)| - (1-t)|zB - D_A|,$$

then

$$|zB(t) - A(t)| = t|zB - A| + (1-t)|zD_B - D_A|.$$

Because we assumed that $zB - A$ is SDD, implying that $zD_B - D_A$ is nonsingular, from the previous equality, we conclude that $|zB(t) - A(t)|$ has also to be an SDD matrix, which is an obvious contradiction. Thus, $z \in \Gamma(A, B)$ and, consequently, $\Gamma(A(t), B(t)) \subseteq \Gamma(A, B)$, for all $t \in [0, 1]$.

Considering the case when $t = 0$, we have that $A(0) = D_A$ and $B(0) = D_B$. Then, $z \in \Gamma(A(0), B(0))$ if and only if $zD_B - D_A$ is not an SDD matrix. Obviously, the only case when this is true is when, for some $i \in N$, we have $z = a_{i,i}b_{i,i}^{-1}$, with $b_{i,i} \neq 0$, or when $a_{i,i} = b_{i,i} = 0$. Therefore, for all $i \in N$ such that $b_{i,i} \neq 0$, then $a_{i,i}b_{i,i}^{-1} \in \Gamma(A(0), B(0))$. That this set is not empty follows from the fact that the case when $\Gamma(A, B) = \mathbb{C}$ cannot occur, because the sets U and V are disjoint, and $\widetilde{U} \cap \widetilde{V} = \emptyset$. Therefore, according to Theorem 3.4, for at least one $i \in N$, $b_{i,i} \neq 0$.

To summarize, we have obtained that

$$\Gamma(A(0), B(0)) = \sigma_F(A(0), B(0)) = \left\{a_{i,i}b_{i,i}^{-1} : b_{i,i} \neq 0,\ i \in N\right\} \neq \emptyset$$

and that

$$\Gamma(A(t), B(t)) \subseteq \Gamma(A, B), \quad \text{for all } t \in [0, 1]. \tag{9}$$

Let us now suppose that $a_{i,i}b_{i,i}^{-1} \in U$, and with $\lambda(t)$, let us denote the generalized eigenvalue of the matrix pair $(A(t), B(t))$ that, for $t = 0$, becomes $a_{i,i}b_{i,i}^{-1}$. Because the eigenvalues (considered as points on the Riemann sphere) are continuous functions of the entries of both matrices, as in [3], we can consider $\{\lambda(t) : t \in [0, 1]\}$ to be a continuous curve on the Riemann sphere, that is, on the geometrical representation of the extended complex plane $\mathbb{C}_\infty$. Following Theorem 3.3 and Equation (9), we have that $\{\lambda(t) : t \in [0, 1]\} \subseteq \Gamma(A, B) = U \cup V$. But because $\lambda(0) \in U$ and U is a closed set in $\mathbb{C}_\infty$, it follows that $\lambda(1) \in U$.

Consequently, having any generalized eigenvalue curve $\{\lambda(t) : t \in [0, 1]\}$, both of its ends have to belong to the same set, either U or V.

To conclude the proof, we observe that at least one of the sets U and V needs to be bounded in $\mathbb{C}$. Namely, if they are both unbounded, because they are both closed, then their one-point compactifications contain ∞, and therefore, $\widetilde{U} \cap \widetilde{V} \neq \emptyset$. So without loss of generality, let us suppose that U is bounded, that is, $\infty \notin \widetilde{U}$. Therefore, all the generalized eigenvalues of a matrix pair (A, B) that lie in U must be finite, and according to the previous conclusions, their number is exactly $|\left\{i \in N : b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$. □

4. APPROXIMATIONS OF GERŠGORIN SETS FOR GENERALIZED EIGENVALUES

One of the major drawbacks of the generalized Geršgorin set is that it is not as elegant as the original Geršgorin set. Namely, the sets $\Gamma_i(A, B)$, in general, are not disks, and moreover, they are more difficult to calculate in order to obtain useful plots, as is discussed in the following section. So, one can be motivated to approximate the generalized Geršgorin set, in order to obtain, although larger, more practical localization sets for generalized eigenvalues. The first attempt was done by Stewart in [1], where, for the first time, the set (5) was defined. Using the concept of a *chordal metric*, he obtained the following result, which we state in an equivalent form using our notation.

Theorem 4.1
(Stewart) Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following inclusion holds:

$$\sigma(A, B) \subseteq \Gamma(A, B) \subseteq \mathcal{G}(A, B) := \bigcup_{i \in N} \mathcal{G}_i(A, B), \tag{10}$$

DOI: 10.1002/nla

where

$$\mathcal{G}_i(A,B) := \{z \in \mathbb{C} : |zb_{i,i} - a_{i,i}| \leqslant (1+|z|^2)\sqrt{[r_i(A)]^2 + [r_i(B)]^2}\}, \ \ (i \in N). \tag{11}$$

To discuss our approach to this problem, we start with the observation that, having a regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$ and an arbitrary point $z \in \mathbb{C}$, for every $i \in N$, we set

$$r_i(zB - A) = \sum_{j \in N \setminus \{i\}} |zb_{i,j} - a_{i,j}|. \tag{12}$$

Then, we proceed by applying the triangular inequality to Equation (12), giving

$$r_i(zB - A) \leqslant |z| r_i(B) + r_i(A). \tag{13}$$

Thus, for a given regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, we define the sets

$$\begin{cases} \widehat{\Gamma}_i(A,B) := \{z \in \mathbb{C} \ : \ |zb_{i,i} - a_{i,i}| \leqslant |z| r_i(B) + r_i(A)\}, \ \ (i \in N), \\ \widehat{\Gamma}(A,B) := \bigcup_{i \in N} \widehat{\Gamma}_i(A,B), \end{cases} \tag{14}$$

and we call the first set of Equation (14) the ***i*th approximated generalized Geršgorin set** and the second set of Equation (14) the **approximated generalized Geršgorin set**.

As a consequence of Equation (13) and Theorem 3.3, the following theorem then holds.

Theorem 4.2

Given any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the finite spectrum of the matrix pair (A,B) belongs to the approximated generalized Geršgorin set of the matrix pair (A,B). Moreover, the following inclusion holds:

$$\sigma(A,B) \subseteq \Gamma(A,B) \subseteq \widehat{\Gamma}(A,B) \subseteq \mathcal{G}(A,B). \tag{15}$$

where $\widehat{\Gamma}(A,B)$ and $\mathcal{G}(A,B)$ are given in Equations (14) and (10), respectively.

Proof

Because the first inclusion in Equation (15) was proved in [2] and the second one is an obvious consequence of Equation (13), we need only to discuss the relationship between the sets $\widehat{\Gamma}(A,B)$ and $\mathcal{G}(A,B)$. After a closer look, we see that for each matrix pair (A,B) and each $i \in N$, $(|z|r_i(B) + r_i(A))^2 \leqslant (1+|z|^2)([r_i(A)]^2 + [r_i(B)]^2)$, implying that $\widehat{\Gamma}_i(A,B) \subseteq \mathcal{G}_i(A,B)$. Therefore, for every matrix pair (A,B), we have $\widehat{\Gamma}(A,B) \subseteq \mathcal{G}(A,B)$. □

It is interesting to note that on taking $B = I_n$, our approximation $\widehat{\Gamma}(A,B)$ reduces to the ordinary Geršgorin set $\Gamma(A)$, which is not true for the approximation $\mathcal{G}(A,B)$ of Stewart as given in Equation (11).

The ith approximated generalized Geršgorin set is a set in the complex plane whose boundary is the curve that can be represented as

$$|z - \xi| = \beta|z| + \alpha, \tag{16}$$

where $\alpha, \beta \geqslant 0$ and $\xi \in \mathbb{C} \setminus \{0\}$, when $b_{i,i} \neq 0$ and $a_{i,i} \neq 0$.

After some analysis, the following classification can be obtained:

- If $\alpha = 0$ and $\beta = 0$, then the curve of Equation (16) is actually a single point ξ;
- If $\alpha = 0$ and $\beta = 1$, then the curve is a perpendicular bisection line of the complex line segment $[0, \xi]$;
- If $\alpha = 0$ and $0 < \beta \neq 1$, then the curve is a **circle of Apollonius**[¶] with foci in 0 and ξ, with the ratio β;

[¶]Apollonius of Perga (262–190 BC), a Greek geometer and astronomer noted for his writings on conic sections.

DOI: 10.1002/nla

- If $\alpha > 0$ and $\beta = 0$, then the curve is a circle centered in ξ with radius α;
- If $\alpha > 0$ and $\beta > 0$, then the curve is a **Cartesian oval**[‖] with foci in 0 and $\xi \in \mathbb{C}$ and with linear factors $-\beta\alpha^{-1}$ and α^{-1}.

To illustrate this, Figure 1 shows the curves (16) plotted for $\xi = 1$ and $\alpha \in \{0, 0.2, 0.4, \dots, 1.8, 2\}$, going from black to light gray, respectively. The parameter β is fixed for each plot, and it takes values $\beta = 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$, from upper left to lower right corner in 1, respectively.

Considering the form of the ith approximated generalized Geršgorin set, given in Equation (14), as in Theorem 3.4, we obtain the somewhat expected result.

Theorem 4.3
Let $A, B \in \mathbb{C}^{n,n}$, with $n \geqslant 2$. Then, the following statements hold:

1. Let $i \in N$ be such that for at least one $j \in N$, $b_{i,j} \neq 0$. Then, the ith approximated generalized Geršgorin set $\widehat{\Gamma}_i(A, B)$, as defined in Equation (14), is a bounded set in the complex plane $\mathbb{C}$ if and only if $|b_{i,i}| > r_i(B)$.
2. The approximated generalized Geršgorin set $\widehat{\Gamma}(A, B)$ is a compact set in $\mathbb{C}$ if and only if B is an SDD matrix.
3. The ith approximated generalized Geršgorin set $\widehat{\Gamma}_i(A, B)$, given in Equation (14), contains zero if and only if $|a_{i,i}| \leqslant r_i(A)$.
4. The approximated generalized Geršgorin set $\widehat{\Gamma}(A, B)$ contains zero if and only if A is not an SDD matrix.
5. If there exists an $i \in N$ such that both $b_{i,i} = 0$ and $|a_{i,i}| \leqslant r_i(A)$, then $\widehat{\Gamma}_i(A, B)$, and consequently, $\widehat{\Gamma}(A, B)$ are the entire complex plane.

Proof
First, it is evident that statements 2 and 4 follow directly from statements 1 and 3, respectively. Second, statements 3 and 5 are easy to obtain, by putting $z = 0$ in the inequalities of Equation (14). Thus, it remains to prove statement 1.

First, we suppose that $|b_{i,i}| > r_i(B)$. Because for each $z \in \widehat{\Gamma}_i(A, B)$, we have from Equation (14) that

$$|z|(|b_{i,i}| - r_i(B)) \leqslant r_i(A) + |a_{i,i}|,$$

so that

$$|z| \leqslant \frac{r_i(A) + |a_{i,i}|}{|b_{i,i}| - r_i(B)}, \tag{17}$$

showing that the set $\widehat{\Gamma}_i(A, B)$ is bounded.

To prove the converse, suppose that $\widehat{\Gamma}_i(A, B)$ is bounded. Then, having that $|zb_{i,i} - a_{i,i}| \leqslant |z||b_{i,i}| + |a_{i,i}|$, the set described by $|z|(|b_{i,i}| - r_i(B)) \leqslant r_i(A) - |a_{i,i}|$ is bounded, too. Obviously, $|b_{i,i}| \leqslant r_i(B)$ cannot hold, and thus, it remains to discuss the case when $|b_{i,i}| = r_i(B)(> 0)$. Define the sequence of complex numbers $\{z_k\}_{k\in\mathbb{N}}$ such that $z_k := k\mathrm{e}^{i\theta}$, for $k \in \mathbb{N}$, with $\theta := \arg(a_{i,i}/b_{i,i})$. Obviously, there exists an integer $k_0 \in \mathbb{N}$ such that for all $k \geqslant k_0$,

$$0 \leqslant 1 - \frac{1}{k}\frac{|a_{i,i}|}{|b_{i,i}|} \leqslant 1 + \frac{1}{k}\frac{r_i(A)}{|b_{i,i}|}$$

holds. Because these inequalities imply that

$$|1 - \frac{1}{z_k}\frac{a_{i,i}}{b_{i,i}}| \leqslant 1 + \frac{1}{|z_k|}\frac{r_i(A)}{|b_{i,i}|},$$

[‖]The Cartesian oval is a curve, defined as a collection of points for which the distances to two foci are related linearly. Some special cases of Cartesian ovals, which occur here, are limaçons of Pascal (for $\alpha = 1$) and hyperbolas (for $\beta = 1$). Cartesian ovals are also called ovals of Descartes. For a nice treatment, see the description of the Circle of Apollonius in Wikipedia.

DOI: 10.1002/nla

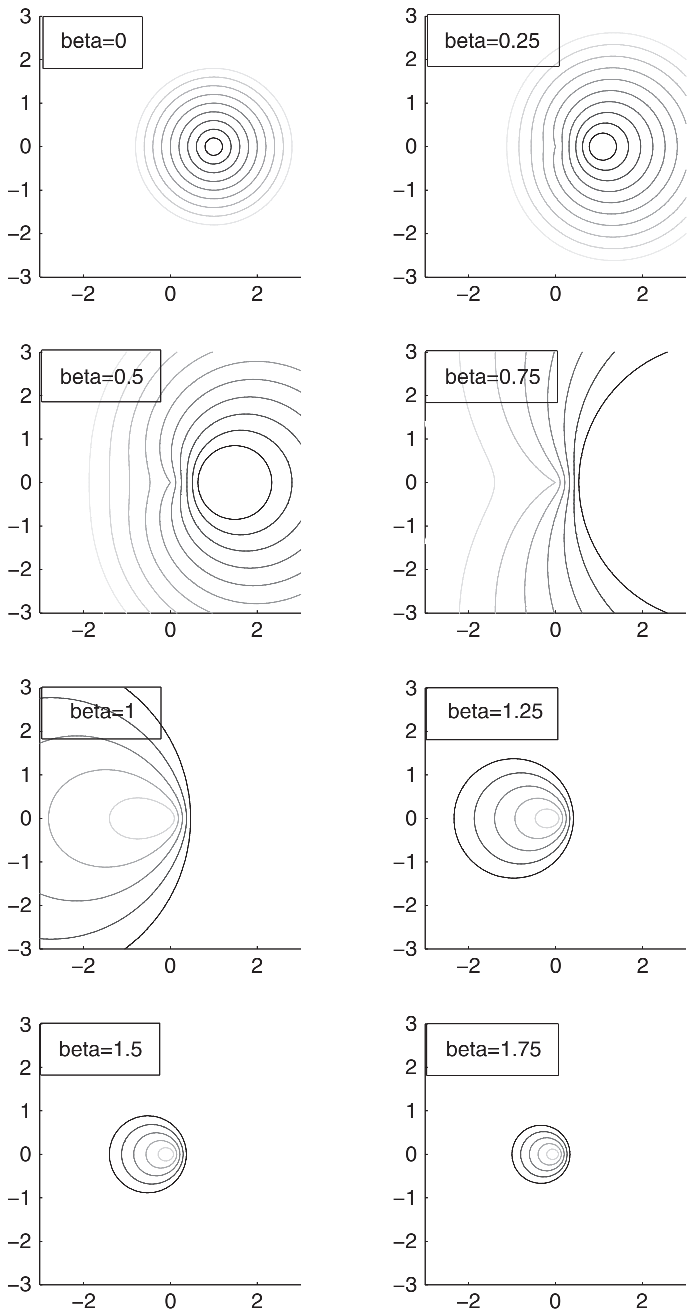

Figure 1. Curves (16) plotted for $\xi = 1$ and for $\alpha \in \{0, 0.2, 0.4, \ldots, 1.8, 2\}$, setting $\beta = 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$, for each plot, from upper left to lower right corner, respectively.

one can conclude that $z_k \in \widehat{\Gamma}_i(A, B)$, for all $k \geqslant k_0$. Therefore, $\widehat{\Gamma}_i(A, B)$ is unbounded. □

Thus, our approximation of the generalized Geršgorin set as a union of Cartesian ovals keeps the same property of *boundedness* as the original set. Because the *isolation property* is inherited by the supersets of the localization set, the following theorem follows directly.

Theorem 4.4
(Isolation property) Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geqslant 2$, if there exist closed sets $U, V \subseteq \mathbb{C}$ such that

DOI: 10.1002/nla

$$\widetilde{U} \cap \widetilde{V} = \emptyset \text{ and } \widehat{\Gamma}(A,B) = U \cup V, \tag{18}$$

then one of them, say U, is compact in $\mathbb{C}$, and it contains exactly

$$|\{i \in N : b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\}|$$

finite eigenvalues of the matrix pair (A, B).

How those two sets in the complex plane relate to each other will be illustrated by the following examples.

Example 4.5
Let

$$A_1 = \begin{pmatrix} 1 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0.1 & 0 & 0 & -i \end{pmatrix}, \quad B_1 = \begin{pmatrix} 0.5 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0.1 & 0 & 0 & -0.5i \end{pmatrix},$$

$$A_2 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & i & 1 \\ 0.8 & 0 & 0 & -i \end{pmatrix}, \text{ and } B_2 = \begin{pmatrix} 1 & 0.1 & 0 & 0 \\ 0.1 & 1 & 0.1 & 0 \\ 0 & 0.1 & 1 & 0.1 \\ 0 & 0 & 0.1 & 1 \end{pmatrix}.$$

By inspection, B_1 and B_2 are SDD matrices, and according to item 2 of Theorem 3.4, the sets $\Gamma(A_1, B_1)$ and $\Gamma(A_2, B_2)$ are compact in the complex plane. Similarly, from Theorem 4.3, the sets $\widehat{\Gamma}(A_1, B_1)$ and $\widehat{\Gamma}(A_2, B_2)$ are compact, too.

This is shown in Figures 2 and 3, where the original generalized Geršgorin set is shaded, while the boundary of the approximation is given by the thick black line. The actual generalized eigenvalues are marked with '×'.

We also remark that because the matrix A_1 is SDD, zero is not contained in the sets $\Gamma(A_1, B_1)$ and $\widehat{\Gamma}(A_1, B_1)$, while on the other hand, the matrix A_2 is not SDD, and $0 \in \Gamma(A_2, B_2) \subseteq \widehat{\Gamma}(A_2, B_2)$.

Figure 4 shows the generalized Geršgorin sets $\Gamma(A_1, A_2)$ and its approximation $\widehat{\Gamma}(A_1, A_2)$, which are unbounded, as a consequence of the fact that A_2 is not an SDD matrix.

Observing the structure of matrices A_1 and B_1, we can see that $\Gamma_2(A_1, B_1) = \Gamma_3(A_1, B_1) = \{1\}$, a single and isolated point, out of the remainder of the generalized Geršgorin set of the matrix pair

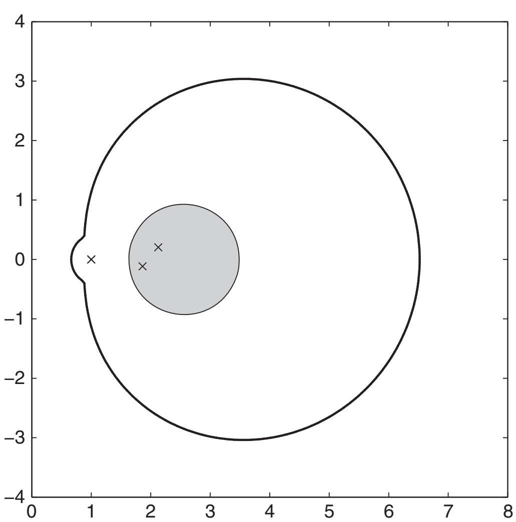

Figure 2. Approximated generalized Geršgorin set of the matrix pair (A_1, B_1).

DOI: 10.1002/nla

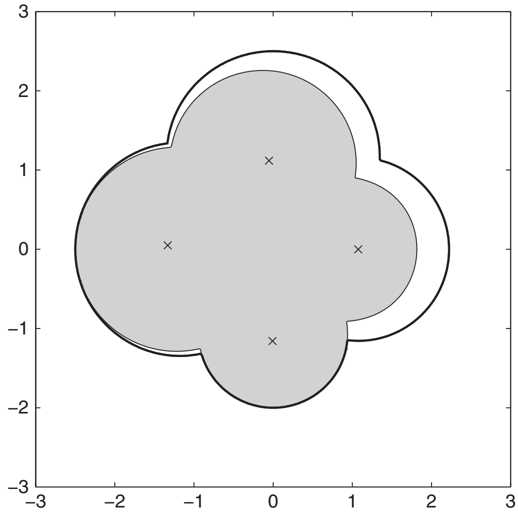

Figure 3. Approximated generalized Geršgorin set of the matrix pair (A_2, B_2).

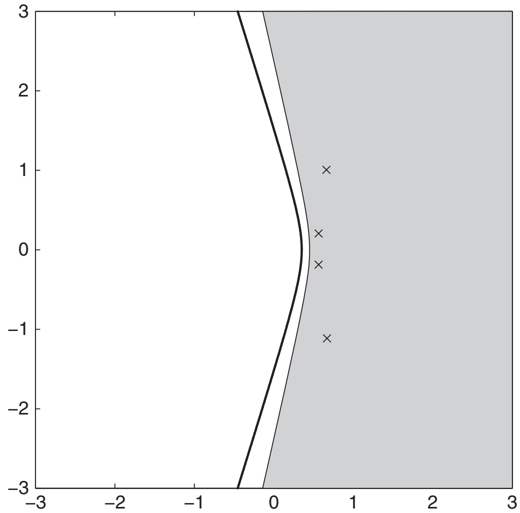

Figure 4. Approximated generalized Geršgorin set of the matrix pair (A_1, A_2).

(A_1, B_1). As a consequence of the isolation property, we see that 1 is a generalized eigenvalue of multiplicity 2 of the pair (A_1, B_1). On the other hand, the approximated generalized Geršgorin set does not reflect this situation.

5. NUMERICAL PROCEDURE FOR OBTAINING APPROXIMATED GENERALIZED GERŠGORIN SET

In this Section, we will give, in more detail, the main motivation one has in order to construct and use the approximated, instead of the original, generalized Geršgorin set. As mentioned, the main reason lies in the different costs one must pay in order to plot them effectively. Namely, given a matrix pair $A, B \in \mathbb{C}^{n,n}$, the original ith generalized Geršgorin set

$$\Gamma_i(A, B) := \{z \in \mathbb{C} : |zb_{i,i} - a_{i,i}| \leqslant \sum_{j \in N\setminus\{i\}} |zb_{i,j} - a_{i,j}|\} \tag{19}$$

DOI: 10.1002/nla

has as a boundary the curve with, in general, n foci, which is very difficult to parameterize in a useful way. Thus, in order to plot it, one should use a different technique. A relatively standard way is to generate a mesh of nodes (x, y) in the desired part of the complex plane and then putting $z = x + iy$ to calculate the values

$$|zb_{i,i} - a_{i,i}| - \sum_{j \in N \setminus \{i\}} |zb_{i,j} - a_{i,j}|$$

in each node and to 'color' the nodes, for which this value is nonpositive. Because in each node, we need to sum up n calculated values, this means that for matrices of size $n \times n$ and mesh of size $m \times m$, the cost of construction of the generalized Geršgorin set is of the order $(nm)^2$ operations. Because, of course, the size of a mesh implies the precision of the generated plot, m is often a very large number.

Therefore, especially for matrix pairs of a larger size, one should consider 'cheaper' ways to obtain a localization set for generalized eigenvalues. The proposed set, given in the previous section, is the approximated generalized Geršgorin set, which, like the original one, consists of n sets from Equation (14):

$$\widehat{\Gamma}_i(A, B) := \{z \in \mathbb{C} \, : \, |zb_{i,i} - a_{i,i}| - |z| r_i(B) \leqslant r_i(A)\}, \quad (i \in N). \tag{20}$$

First, if we consider the same plotting technique as explained above, the values we need to calculate in each node, that is, $|zb_{i,i} - a_{i,i}| - |z| r_i(B) - r_i(A)$, where $i \in N$, use only precalculated data for a given matrix pair and, thus, demand only three summands (instead of n summands in the case of the original generalized Geršgorin set). So, using the same technique as before, the cost of plotting the approximated generalized Geršgorin set would be of order nm^2, which is, of course, more favorable if one is dealing with matrix pairs of a large size.

On the other hand, inequalities that define Equation (20) imply that the boundaries of such sets are curves for which we can obtain suitable parameterizations and, consequently, obtain the approximated generalized Geršgorin set by plotting n curves. If we take m nodes in each parameterization, the cost of obtaining such a set is only of the order of nm operations.

Whereas the previously explained technique required the use of meshgrid and contour functions in MATLAB (MathWorks, Natick, MA, USA), this one can be implemented using only a simple plot function.

In the following, we will give the guidelines for the parameterizations of Equation (20) that can be used for plotting the approximated generalized Geršgorin set.

For a given matrix pair $A, B \in \mathbb{C}^{n,n}$ and given $k \in N$, we consider the inequality of Equation (14) that defines kth approximated generalized Geršgorin set,

$$|zb_{k,k} - a_{k,k}| - |z| r_k(B) \leqslant r_k(A), \tag{21}$$

and discuss the form of this set and its boundary in the specific cases that can occur.

- Case 1. If $|b_{k,k}| = r_k(B) = 0$, then the set given in Equation (21) is either an empty set, if $|a_{k,k}| > r_k(A)$, or the entire complex plane, otherwise.
- Case 2. If $|b_{k,k}| = 0$ and $r_k(B) > 0$, then inequality (21) becomes

$$|z| \geqslant \frac{|a_{k,k}| - r_k(A)}{r_k(B)},$$

  implying that the corresponding set is the exterior of the circle centered in the origin.
- Case 3. If $|b_{k,k}| > 0$ and $r_k(B) = 0$, then inequality (21) becomes

$$|z - \frac{a_{k,k}}{b_{k,k}}| \leqslant \frac{r_k(A)}{|b_{k,k}|},$$

  implying that the corresponding set is the interior of the circle centered in $a_{k,k}/b_{k,k}$.

- Case 4. If $|b_{k,k}| = r_k(B) > 0$, then inequality (21) becomes

$$|z - \frac{a_{k,k}}{b_{k,k}}| - |z| \leqslant \frac{r_k(A)}{|b_{k,k}|},$$

and we have two subcases. First, if $|a_{k,k}| \leqslant r_k(A)$, then after a closer look, we deduce that the corresponding set is the entire complex plane. Otherwise, we have that $|a_{k,k}| > r_k(A)$, and taking the last inequality to be an equality, we obtain the definition of the left branch of a hyperbola, with foci in the origin and $a_{k,k}/b_{k,k}$. The parameterization of such a curve is well known. Here, the set given in Equation (21) is the unbounded part of the complex plane that does not contain the origin and has the obtained hyperbola as a boundary.
- Case 5. If $0 < |b_{k,k}| < r_k(B)$ and $|a_{k,k}| > r_k(A)$, then inequality (21) becomes

$$|z - \frac{a_{k,k}}{b_{k,k}}| - |z|\frac{r_k(B)}{b_{k,k}} \leqslant \frac{r_k(A)}{b_{k,k}},$$

implying that the corresponding set is an interior of the inner part of a Cartesian oval. The parameterization of such an oval is much less known and can be obtained in the following way:

$$x(t) := \rho(t)\cos(t) \quad \text{and} \quad y(t) := \rho(t)\sin(t),$$

where

$$\rho(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta - \sqrt{(\cos(t) + \alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2},$$

$$\alpha := \frac{r_k(A)}{|a_{k,k}|}, \quad \beta := \frac{r_k(B)}{|b_{k,k}|}, \quad \text{and} \quad t \in (-\pi, \pi).$$

Here, we should comment that if $\rho(t) < 0$, for all $t \in (-\pi, \pi)$, inequality (21) defines an empty set.
- Case 6. If $|b_{k,k}| > r_k(B) > 0$ and $|a_{k,k}| \leqslant r_k(A)$, then inequality (21) is again

$$|z - \frac{a_{k,k}}{b_{k,k}}| - |z|\frac{r_k(B)}{b_{k,k}} \leqslant \frac{r_k(A)}{b_{k,k}},$$

implying that the corresponding set is an exterior of the outer part of a Cartesian oval parameterized by

$$x(t) := \rho(t)\cos(t) \quad \text{and} \quad y(t) := \rho(t)\sin(t),$$

where

$$\rho(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta + \sqrt{(\cos(t) + \alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2},$$

$$\alpha := \frac{r_k(A)}{|a_{k,k}|}, \quad \beta := \frac{r_k(B)}{|b_{k,k}|}, \quad \text{and} \quad t \in (-\pi, \pi).$$

Similar to the previous case, we should comment that if $\rho(t) < 0$, for all $t \in (-\pi, \pi)$, inequality (21) defines the entire complex plane.
- Case 7. If $|b_{k,k}| > r_k(B) > 0$ and $|a_{k,k}| > r_k(A)$, then inequality (21) is again

$$|z - \frac{a_{k,k}}{b_{k,k}}| - |z|\frac{r_k(B)}{b_{k,k}} \leqslant \frac{r_k(A)}{b_{k,k}}.$$

Using the polar equation of Cartesian ovals, for the values of the parameter $t \in (-\pi, \pi)$ for which the value

$$\Delta := (\cos(t) + \alpha\beta)^2 - (1-\alpha^2)(1-\beta^2),$$

DOI: 10.1002/nla

where again

$$\alpha := \frac{r_k(A)}{|a_{k,k}|} \quad \text{and} \quad \beta := \frac{r_k(B)}{|b_{k,k}|}$$

are nonnegative, one obtains that the set defined by Equation (21) is the interior of the corresponding Cartesian oval, which is parameterized by two branches:

$$\rho_1(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta - \sqrt{(\cos(t)+\alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2},$$

$$\rho_2(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta + \sqrt{(\cos(t)+\alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2}.$$

Here, it is important to mention that Equation (21) can be an empty set, when both $\rho_1(t)$ and $\rho_2(t)$ are negative, for every $t \in (-\pi, \pi)$.

- Case 8. If $0 < |b_{k,k}| < r_k(B)$ and $|a_{k,k}| \leqslant r_k(A)$, then inequality (21) is again

$$|z - \frac{a_{k,k}}{b_{k,k}}| - |z|\frac{r_k(B)}{b_{k,k}} \leqslant \frac{r_k(A)}{b_{k,k}}.$$

Using the polar form of Cartesian ovals, for the values of the parameter $t \in (-\pi, \pi)$ for which the value

$$\Delta := (\cos(t) + \alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)$$

is nonnegative, where again

$$\alpha := \frac{r_k(A)}{|a_{k,k}|} \quad \text{and} \quad \beta := \frac{r_k(B)}{|b_{k,k}|},$$

one obtains that the set defined by Equation (21) is the exterior of the corresponding Cartesian oval parameterized by two branches:

$$\rho_1(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta - \sqrt{(\cos(t)+\alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2},$$

$$\rho_2(t) := \frac{|a_{k,k}|}{|b_{k,k}|} \cdot \frac{\cos(t) + \alpha\beta + \sqrt{(\cos(t)+\alpha\beta)^2 - (1-\alpha^2)(1-\beta^2)}}{1-\beta^2}.$$

Here, it is important to mention that the set (21) can be the entire complex plane, when both $\rho_1(t)$ and $\rho_2(t)$ are nonpositive, for every $t \in (-\pi, \pi)$.

For matrices given in Example 4.5, Figures 5, 6, and 7, which correspond to Figures 2, 3, and 4, respectively, show the plots obtained using MATLAB's plot function and the parameterizations given above. In these figures, every ith approximated generalized Geršgorin set is shown by its boundary. In Figures 5 and 6, these sets are bounded, whereas in Figure 7, three of them are unbounded. For these unbounded sets, the boundaries are shown by heavy lines, and the corresponding sets are the right-hand side parts of the complex plane (the parts that do not contain the origin).

ACKNOWLEDGEMENTS

We thank the anonymous referees for their helpful suggestions and comments that improved the quality of this paper. This work was partially supported by the Ministry of Science of Republic of Serbia, grant 174019, and Autonomous Province of Vojvodina, Secretariat for Science and Technological Development, grants 0702 and 2211.

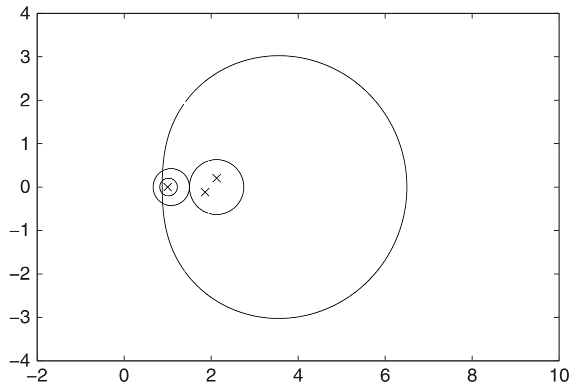

Figure 5. Approximated generalized Geršgorin set of the matrix pair (A_1, B_1).

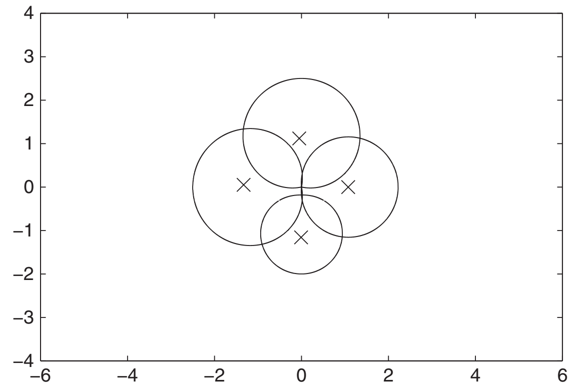

Figure 6. Approximated generalized Geršgorin set of the matrix pair (A_2, B_2).

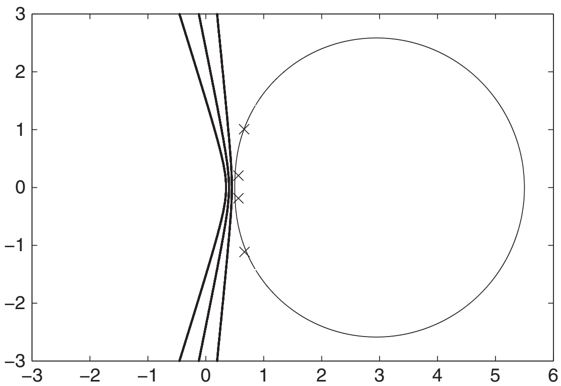

Figure 7. Approximated generalized Geršgorin set of the matrix pair (A_1, A_2).

REFERENCES

1. Stewart GW. Geršgorin theory for generalized eigenvalue problem $ax = \lambda x$. *Mathematics of Computation* 1975; **29**:600–606.
2. Cvetković L, Kostić V, Varga RS. Geršgorin-type localizations of generalized eigenvalues. *Numerical Linear Algebra with Applications* 2009; **16**(11–12):883–898.
3. Stewart GW, Sun JG. *Matrix Perturbation Theory*. Academic Press: New York, 1990.
4. Cvetković L. H-matrix theory vs. eigenvalue localization. *Numerical Algorithms* 2006; **42**(3–4):229–245.
5. Varga RS. *Geršgorin and His Circles*. Springer-Verlag: New York, 2004.
6. Karow M. Geometry of Spectral Value Sets. *PhD thesis*, University of Bremen, Bremen, Germany, 2003.

DOI: 10.1002/nla