A simple generalization of Geršgorin's theorem

Ljiljana Cvetković[1] Vladimir Kostić[1] Rafael Bru[2]
Francisco Pedroche[2]

1 Department of Mathematics and Informatics, Faculty of Science, University of Novi Sad. Serbia, 21000 Novi Sad. {lila,vkostic@im.ns.ac.yu}

2 Institut de Matemàtica Multidisciplinar, Universitat Politècnica de València. Camí de Vera s/n. 46022 València. Spain. {rbru,pedroche@mat.upv.es}

Keywords. Eigenvalue localization, H-matrices, α-matrices.
AMS Subject Classification. 15A18, 65F15

Abstract

It is well known that the spectrum of a given matrix A belongs to the Geršgorin set $\Gamma(A)$, as well as to the Geršgorin set applied to the transpose of A, $\Gamma(A^T)$. So, the spectrum belongs to their intersection. But, if we first intersect i-th Geršgorin disk $\Gamma_i(A)$ with the corresponding disk $\Gamma_i(A^T)$, and then we make union of such intersections, which are, in fact, the smaller disks of each pair, what we get is not an eigenvalue localization area. The question is what should be added in order to catch all the eigenvalues, while, of course, staying within the set $\Gamma(A) \cap \Gamma(A^T)$. The answer lies in the appropriate characterization of some subclasses of nonsingular H-matrices. In this paper we give two such characterizations, and then we use them to prove localization areas that answer this question.

1 Introduction

We start by recalling the very well known theorem of Geršgorin, [4].

Theorem 1. (Geršgorin) *Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, and let λ be its eigenvalue. Then, there exists an index $i \in N := \{1, 2, ..., n\}$ such that $|\lambda - a_{ii}| \leq r_i(A)$, where r_i is defined to be the i-th deleted row sum of A, i.e.*

$$r_i(A) := \sum_{j \in N \setminus \{i\}} |a_{ij}|.$$

Thus, denoting $\sigma(A)$ to be the spectrum of the matrix A, we have that

$$\sigma(A) \subset \Gamma(A) := \bigcup_{i \in N} \Gamma_i(A), \tag{1}$$

where

$$\Gamma_i(A) := \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i(A)\}. \tag{2}$$

The set $\Gamma(A)$ is called the *Geršgorin set* of the matrix A, and the set $\Gamma_i(A)$ is called *i-th Geršgorn disk* of the matrix A.

Since matrices A and A^T have the same spectra, we have that $\sigma(A) \subseteq \Gamma(A^T)$, and thus $\sigma(A) \subseteq \Gamma(A) \cap \Gamma(A^T)$.

Let us now consider the pair of i-th Geršgorin disks for A and A^T, namely $\Gamma_i(A)$ and $\Gamma_i(A^T)$. They have the same centers, and so, their intersection is the smaller one:

$$\overline{\Gamma}_i(A) := \left\{z \in \mathbb{C} : |z - a_{ii}| \leq \min\{r_i(A), r_i(A^T)\}\right\}.$$

The following simple example shows that the union of such disks

$$\overline{\Gamma}(A) := \bigcup_{i \in N} \overline{\Gamma}_i(A) \tag{3}$$

is not an eigenvalue localization area.

Example 1. *The eigenvalues of the matrix*

$$A_1 = \begin{pmatrix} 10 & 9 \\ 1 & 10 \end{pmatrix}$$

are $\sigma(A_1) = \{7, 13\}$, while $\overline{\Gamma}_1(A_1)$ and $\overline{\Gamma}_2(A_1)$ are the same disks with the center 10 and radius 1, i.e., $\overline{\Gamma}(A_1) = \overline{\Gamma}_1(A_1) = \overline{\Gamma}_2(A_1) = \{z \in \mathbb{C} : |z - 10| \leq 1\}$, see Figure 5.

In the next section we will give new characterizations of two special classes of nonsingular matrices, named $\alpha 1$ and $\alpha 2$ in [1], and introduced by Ostrowski in [5]. These characterizations will be used in Section 3 for proving new eigenvalue localizations, which both answer the question what should be added to the set (3) in order to capture all the eigenvalues of A, while staying within the set $\Gamma(A) \cap \Gamma(A^T)$.

2 Characterizations of $\alpha 1$ and $\alpha 2$ matrices

Concerning nonsingularity of matrices, there are two well known results that combine the information about a matrix and its transpose, see [5] and [6]. Before we state them, let us, for the sake of simplicity, denote $r_i := r_i(A)$, $i \in N$, and $c_i := r_i(A^T)$, $i \in N$.

Theorem 2. (Ostrowski) *Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be such that*

$$|a_{ii}| > \alpha r_i + (1-\alpha)c_i \ \text{ for some } \ \alpha \in [0,1], \ \text{ and for each } \ i \in N. \tag{4}$$

Then, A is a nonsingular matrix.

Theorem 3. (Ostrowski) *Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be such that*

$$|a_{ii}| > (r_i)^{\alpha}(c_i)^{1-\alpha} \ \text{ for some } \ \alpha \in [0,1], \ \text{ and for each } \ i \in N. \tag{5}$$

Then, A is a nonsingular matrix.

The matrices that fulfill condition (4) are known as $\alpha 1$-*matrices*, while $\alpha 2$-*matrices* are the matrices that fulfill (5). Actually, the nonsingularity of $\alpha 1$ matrices follows directly from nonsingularity of $\alpha 2$ matrices, by the generalized arithmetic-geometric mean inequality:

$$\alpha a + (1-\alpha)b \geq a^{\alpha} b^{1-\alpha}, \tag{6}$$

where $a, b \geq 0$ and $0 \leq \alpha \leq 1$.

Nevertheless, we will consider both of these classes, since they will both play their role in the concluding section.

Given a matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, we partition set of indices N into three sets

$$\begin{array}{rcll} \mathcal{R} & := & \{i \in N : & r_i > c_i\}, \\ \mathcal{C} & := & \{i \in N : & c_i > r_i\}, \\ \mathcal{E} & := & \{i \in N : & r_i = c_i\}. \end{array} \tag{7}$$

Then, the following theorem holds.

Theorem 4. *A matrix $A = [a_{ij}] \in \mathbb{C}^{n\times n}$, with $n \geq 2$, is an $\alpha 1$-matrix if and only if the following two conditions hold*

(i) $|a_{ii}| > \min\{r_i, c_i\}$, *for all $i \in N$,*

(ii) $\dfrac{|a_{ii}| - c_i}{r_i - c_i} > \dfrac{c_j - |a_{jj}|}{c_j - r_j}$, *for all $i \in \mathcal{R}$, and all $j \in \mathcal{C}$.*

Proof. First, let us assume that A is an $\alpha 1$-matrix. Then, there exists $\alpha \in [0,1]$ such that

$$|a_{ii}| > \alpha(r_i - c_i) + c_i, \quad \text{for all } \ i \in N. \tag{8}$$

Therefore, for every $i \in \mathcal{R}$, we conclude that $\frac{|a_{ii}|-c_i}{r_i-c_i} > \alpha$, and for every $j \in \mathcal{C}$, $\frac{c_j-|a_{jj}|}{c_j-r_j} < \alpha$. Thus, (ii) obviously holds. Condition (i) follows directly from (4) and the fact that $\alpha \in [0,1]$.

Conversely, assume that the conditions (i) and (ii) hold. For every index $i \in \mathcal{E}$, condition (i) directly implies (4), so, it remains to prove that (4) holds for indices from the set $\mathcal{R} \cup \mathcal{C}$, too.

First, observe that for every $i \in \mathcal{R}$, we have $r_i - c_i > 0$, and, thus, by condition (i), $|a_{ii}| - c_i > 0$. This, obviously, implies that

$$\frac{|a_{ii}| - c_i}{r_i - c_i} > 0. \tag{9}$$

Similarly, for every $j \in \mathcal{C}$, $|a_{jj}| > r_j$ and, thus, $c_j - |a_{jj}| < c_j - r_j$. Since $c_j - r_j > 0$, this implies that

$$\frac{c_j - |a_{jj}|}{c_j - r_j} < 1. \tag{10}$$

Now, gathering condition (ii), (9) and (10), we have that there exists a parameter α such that, for every $i \in \mathcal{R}$ and every $j \in \mathcal{C}$,

$$\max\{0, \frac{c_j - |a_{jj}|}{c_j - r_j}\} < \alpha < \min\{\frac{|a_{ii}| - c_i}{r_i - c_i}, 1\}.$$

Starting from the left inequality we obtain that $|a_{jj}| > \alpha(r_j - c_j) + c_j$ for every $j \in \mathcal{C}$, while from the right one we get the same for indices $i \in \mathcal{R}$. Thus, (4) holds for the chosen parameter $\alpha \in [0,1]$ and every index $i \in \mathcal{R} \cup \mathcal{C}$, which completes the proof. □

We prove the similar characterization for $\alpha 2$ matrices.

Theorem 5. *Let $A = [a_{ij}] \in \mathbb{C}^{n \times n}$, $n \geq 2$. Then A is an $\alpha 2$-matrix if and only if the following two conditions hold*

(i) $|a_{ii}| > \min\{r_i, c_i\}$, *for all* $i \in N$,

(ii) $\log_{\frac{r_i}{c_i}} \frac{|a_{ii}|}{c_i} > \log_{\frac{c_j}{r_j}} \frac{c_j}{|a_{jj}|}$, *for all* $i \in \mathcal{R}$, *for which* $c_i \neq 0$, *and for all* $j \in \mathcal{C}$, *for which* $r_j \neq 0$.

Proof. First, we assume that A is an $\alpha 2$ matrix, i.e., that there exists $\alpha \in [0,1]$ such that for each index $i \in N$

$$|a_{ii}| > (r_i)^{\alpha}(c_i)^{1-\alpha}. \tag{11}$$

Now, for each $i \in N$, from $0 \leq \alpha \leq 1$, we have that $(r_i)^{\alpha}(c_i)^{1-\alpha} \geq \min\{r_i, c_i\}$, so, the condition (i) obviously holds.

Consider an arbitrary $i \in \mathcal{R}$, such that $c_i \neq 0$. Then, the condition (11) can be written as

$$\frac{|a_{ii}|}{c_i} > (\frac{r_i}{c_i})^{\alpha}.$$

Since $r_i > c_i$, taking the logarithm of the above inequality for the base $\frac{r_i}{c_i} > 1$, and using the monotonicity, we obtain that

$$\log_{\frac{r_i}{c_i}} \frac{|a_{ii}|}{c_i} > \alpha. \tag{12}$$

Similarly, for an arbitrary index $j \in \mathcal{C}$, such that $r_j \neq 0$, we obtain that

$$\log_{\frac{c_j}{r_j}} \frac{c_j}{|a_{jj}|} < \alpha, \tag{13}$$

which, together with (12), implies the condition (ii).

Conversely, let us assume that A satisfies (i) and (ii).

For an arbitrary index $i \in \mathcal{E}$, (i) directly implies (11). For $i \in \mathcal{R}$, for which $c_i = 0$, and for $j \in \mathcal{C}$, such that $r_j = 0$, (11) follows immediately, as well. Thus, it remains to show that (11) holds for indices from the set $\mathcal{R} \setminus \{i \ : \ c_i = 0\}$ and the set $\mathcal{C} \setminus \{j \ : \ r_j = 0\}$.

First, let us note that for every $i \in \mathcal{R}$, we have $r_i > c_i$, thus, by condition (i), $|a_{ii}| > c_i$. Now, using the properties of the log function for the base greater than one, we obtain

$$\log_{\frac{r_i}{c_i}} \frac{|a_{ii}|}{c_i} > 0.$$

Similarly, for every $j \in \mathcal{C}$, we obtain that $\log_{\frac{c_j}{r_j}} \frac{c_j}{|a_{jj}|} < 1$, which, from the strict inequality of (ii), ensures that there exists a parameter α, such that for an arbitrary index $i \in \mathcal{R}$ and arbitrary $j \in \mathcal{C}$,

$$\max\{0, \log_{\frac{c_j}{r_j}} \frac{c_j}{|a_{jj}|}\} < \alpha < \min\{\log_{\frac{r_i}{c_i}} \frac{|a_{ii}|}{c_i}, 1\}.$$

Starting from the right inequality, for every $i \in \mathcal{R} \setminus \{i \ : \ c_i = 0\}$ we have that

$$\frac{|a_{ii}|}{c_i} > (\frac{r_i}{c_i})^\alpha,$$

implying that (11) holds. In the same way from the left inequality we obtain that (11) is true for every index from the set $\mathcal{C} \setminus \{j \ : \ r_j = 0\}$. Since $\alpha \in [0, 1]$, this concludes the proof. □

3 Eigenvalue localizations

Having the results of the previous section that characterize $\alpha 1$ and $\alpha 2$ classes, independently of the value of the parameter α, we are ready to give the corresponding eigenvalue localizations of Geršgorin's type.

Theorem 6. *Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, and let λ be its eigenvalue. Then, there exists an index $i \in N$ such that $|\lambda - a_{ii}| \leq \min\{r_i, c_i\}$, or, there exist $i \in \mathcal{R}$ and $j \in \mathcal{C}$, such that*

$$|\lambda - a_{ii}|\,(c_j - r_j) + |\lambda - a_{jj}|\,(r_i - c_i) \leq c_j r_i - c_i r_j. \tag{14}$$

Thus, we have that

$$\sigma(A) \subset \mathcal{A}_1(A) := \overline{\Gamma}(A) \bigcup \widehat{\Gamma}(A), \tag{15}$$

where $\overline{\Gamma}(A)$ is given by (3),

$$\widehat{\Gamma}(A) := \bigcup_{\substack{i \in \mathcal{R} \\ j \in \mathcal{C}}} \widehat{\Gamma}_{ij}(A), \quad \text{and} \tag{16}$$

$$\begin{aligned} \widehat{\Gamma}_{ij}(A) := \{z \in \mathbb{C} : |z - a_{ii}|\,(c_j - r_j) + |z - a_{jj}|\,(r_i - c_i) \\ \leq c_j r_i - c_i r_j\}, \quad (i \in \mathcal{R}) \ \ (j \in \mathcal{C}). \end{aligned} \tag{17}$$

Proof. Let λ be an arbitrary eigenvalue of the matrix A. Then, $\lambda I - A$, where I stands for the identity matrix, is singular. Thus, by Theorem 2, we have that it cannot be an $\alpha 1$ matrix. Now, applying Theorem 4, we obtain that the conditions (i) and (ii) cannot simultaneously hold for the matrix $\lambda I - A$. Since off-diagonal entries of the matrices A and $\lambda I - A$ are the same, we have that the sets $\mathcal{R}$ and $\mathcal{C}$ for the matrix $\lambda I - A$ remain the same.

Therefore, either there exists an index $i \in N$ such that $|\lambda - a_{ii}| \leq \min\{r_i, c_i\}$, or, there exist indices $i \in \mathcal{R}$ and $j \in \mathcal{C}$, such that

$$\frac{|\lambda - a_{ii}| - c_i}{r_i - c_i} \leq \frac{c_j - |\lambda - a_{jj}|}{c_j - r_j}.$$

Obviously, the last inequality can be rewritten as (14), which completes the proof. □

It is interesting to note that, if we start from the original definition of $\alpha 1$ matrices, given in Theorem 2, and derive directly eigenvalue localization area, see [1], we get

$$\mathcal{A}_1(A) = \bigcap_{0 \leq \alpha \leq 1} \bigcup_{i \in N} \{z \in \mathbf{C} : |z - a_{ii}| \leq \alpha r_i + (1 - \alpha) c_i\}. \tag{18}$$

Of course, the form of the set $\mathcal{A}_1(A)$ we have obtained in (15) is much more convenient than the one from (18), which is not of much practical use.

Since $\mathcal{A}_1(A) = \mathcal{A}_1(A^T)$, we have that $\mathcal{A}_1(A) \subseteq \Gamma(A) \cap \Gamma(A^T)$.

Example 2. *Let*

$$A_2 = \begin{pmatrix} 0 & 2 & 0.1 & 0.1 \\ 0 & 0 & 0.2 & 0.1 \\ 0.1 & 0.2 & 0 & 0 \\ 0.1 & 0.1 & 0 & 2 \end{pmatrix}, \quad and \quad A_3 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0.45 & 0 \\ 0 & 0 & i & 1 \\ 0.45 & 0 & 0 & -i \end{pmatrix}.$$

In Figure 1, on the left hand side, the set $\mathcal{A}_1(A_2)$ is shown, where the set $\overline{\Gamma}(A_2)$ is filled, and the set $\widehat{\Gamma}(A_2)$ is represented by the thick boundary. The exact eigenvalues are plotted with asterisks. In the same figure, on the right hand side the intersection of the Geršgorin sets $\Gamma(A_2)$ and $\Gamma(A_2^T)$ is drawn. As we can see, intersecting Geršgorin disks one by one and taking the union fails to capture the eigenvalues, so, the necessity of the set $\widehat{\Gamma}(A_2)$ is evident.

In Figure 2, the same was done for the matrix A_3. It is worthwhile to note that, although the intersection of the Geršgorin sets of a given matrix and its transpose in this case fails to give disjoint localization areas, the set $\mathcal{A}_1(A_3)$ succeeds in that.

In the following, we start from Theorems 3 and 5, in order to derive new localization area.

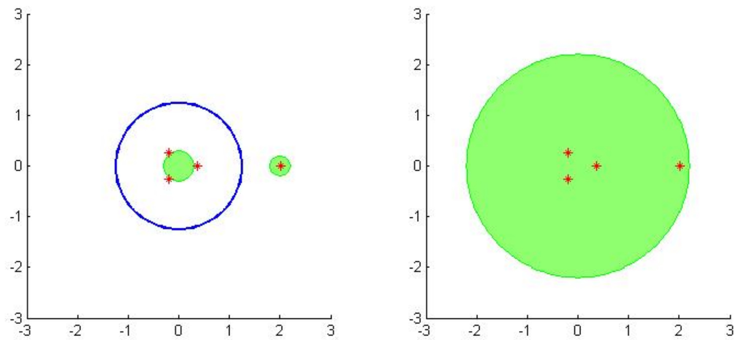

Figure 1: Inclusion regions for the matrix A_2 of the Example 2

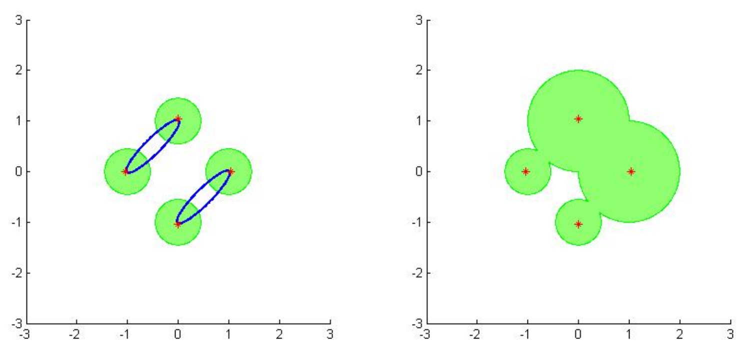

Figure 2: Inclusion regions for the matrix A_3 of the Example 2

Theorem 7. *Let $A = [a_{ij}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, and let λ be its eigenvalue. Then, there exists an index $i \in N$ such that $|\lambda - a_{ii}| \leq \min\{r_i, c_i\}$, or, there exist $i \in \mathcal{R}$ and $j \in \mathcal{C}$, such that $c_i \neq 0$, $r_j \neq 0$, and*

$$\frac{|\lambda - a_{ii}|}{c_i} \left(\frac{|\lambda - a_{jj}|}{c_j} \right)^{\log_{\frac{c_j}{r_j}} \frac{r_i}{c_i}} \leq 1. \tag{19}$$

Thus, we have that

$$\sigma(A) \subset \mathcal{A}_2(A) := \overline{\Gamma}(A) \bigcup \widetilde{\Gamma}(A), \tag{20}$$

where $\overline{\Gamma}(A)$ is given by (3),

$$\widetilde{\Gamma}(A) := \bigcup_{\substack{i \in \mathcal{R} \,:\, c_i \neq 0 \\ j \in \mathcal{C} \,:\, r_j \neq 0}} \widetilde{\Gamma}_{ij}(A), \quad and \tag{21}$$

$$\widetilde{\Gamma}_{ij}(A) := \{z \in \mathbb{C} : \frac{|z - a_{ii}|}{c_i} \left(\frac{|z - a_{jj}|}{c_j} \right)^{\log_{\frac{c_j}{r_j}} \frac{r_i}{c_i}} \leq 1\}, \tag{22}$$

for $i \in \mathcal{R}$, such that $c_i \neq 0$, and $j \in \mathcal{C}$, such that $r_j \neq 0$.

Proof. Let λ be an arbitrary eigenvalue of the matrix A. Then, $\lambda I - A$ is singular. Similarly to the proof of the previous theorem, we obtain that, either $\lambda \in \overline{\Gamma}(A)$, or, there exist indices $i \in \mathcal{R}$ and $j \in \mathcal{C}$, such that $c_i \neq 0$, $r_j \neq 0$, and

$$\log_{\frac{r_i}{c_i}} \frac{|\lambda - a_{ii}|}{c_i} \leq \log_{\frac{c_j}{r_j}} \frac{c_j}{|\lambda - a_{jj}|}.$$

Since, $\frac{r_i}{c_i} > 1$, and $\frac{c_j}{r_j} > 1$, we can rewrite the right hand side of the above inequality as

$$\log_{\frac{c_j}{r_j}} \frac{r_i}{c_i} \log_{\frac{r_i}{c_i}} \frac{c_j}{|\lambda - a_{jj}|},$$

and, thus, obtain that

$$\frac{|\lambda - a_{ii}|}{c_i} \left(\frac{|\lambda - a_{jj}|}{c_j} \right)^{\log_{\frac{c_j}{r_j}} \frac{r_i}{c_i}} \leq 1,$$

which completes the proof. □

In Figures 3 and 4, the localization areas (20) for the matrices A_2 and A_3 from Example 2 are given, respectively.

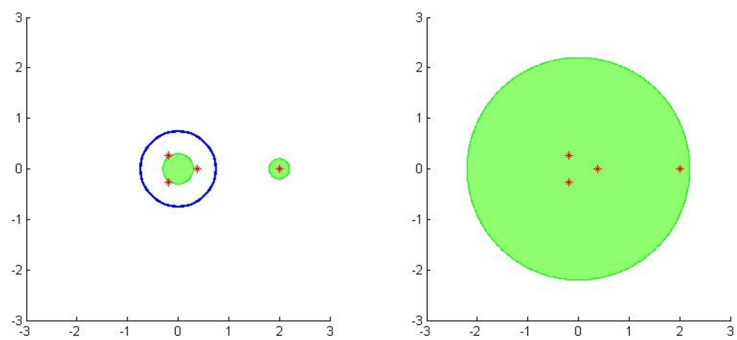

Figure 3: Inclusion regions for the matrix A_2 of the Example 2

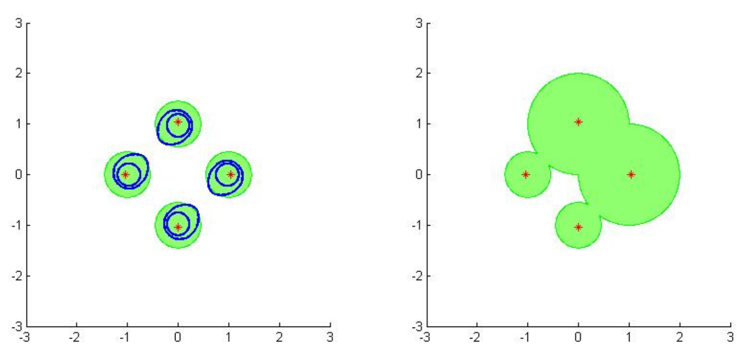

Figure 4: Inclusion regions for the matrix A_3 of the Example 2

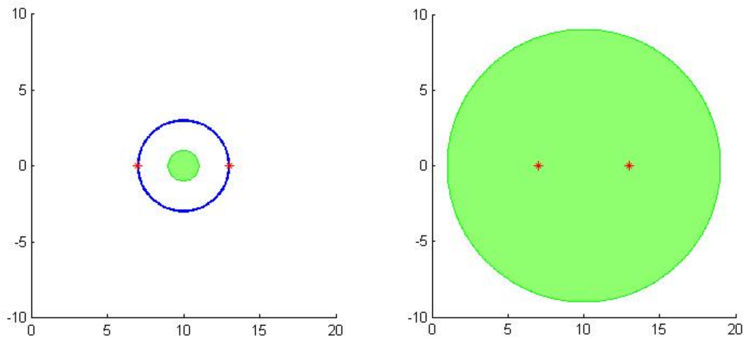

Figure 5: Inclusion regions for the matrix A_1 of the Example 1

Acknowledgments

This work is supported by an agreement of research cooperation between University of Novi Sad (Serbia) and Universitat Politècnica de València (Spain), by Spanish DGI grant MTM2007-64477, by the Provincial Secretariat of Science and Technological Development of Vojvodina, Serbia, grant 0708 and by the Ministry of Science and Environmental Protection of Serbia, grant 144025.

References

[1] Cvetković, L., H-matrix theory vs. eigenvalue localization *Numerical Algorithms*, 42: 229-245, 2006.

[2] Cvetković, Lj., and Kostić, V., New criteria for identifying H-matrices, *J. Comput. Appl. Math.*, 180:265–278, 2005.

[3] Cvetković, L., Kostić, V. and Varga, R. S., A new Geršgoring-type eigenvalue inclusion set, *Electron. Trans. Numer. Anal.*, 18:73–80, 2004.

[4] Geršgorin, S., Über die Abgrenzung der Eigenwerte einer Matrix, *Izv. Akad. Nauk SSSR Ser. Mat.* 1:749-754, 1931.

[5] Ostrowski, A. M. Über die Determinanten mit überwiegender Hauptdiagonale, *Comment. Math. Helv.* 10:69-96, 1937

[6] Varga, R. S., *Geršgorin and his circles*, Springer Series in Computational Mathematics, vol. 36. Springer, Berlin, Heidelberg, 2004.