BENEFITS FROM THE GENERALIZED DIAGONAL DOMINANCE

By

Vladimir Kostić

SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN MATHEMATICAL SCIENCES AT
DEPARTMENT OF MATHEMATICS AND INFORMATICS
FACULTY OF SCIENCE AT THE
UNIVERSITY OF NOVI SAD
TRG D. OBRADOVIĆA 3
NOVI SAD
SERBIA
MARCH 2010

UNIVERSITY OF NOVI SAD

Date: **March 2010**

Author: **Vladimir Kostić**

Title: **Benefits from the Generalized Diagonal Dominance**

Department: **Department of Mathematics and Informatics**

Degree: **Ph.D.** Convocation: **May/June** Year: **2010**

Permission is herewith granted to University of Novi Sad to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

To all those who find beauty and joy
in discovering new spaces of knowledge and conscience,
and find the pleasure in sharing it with others.

V.K.

Acknowledgements

I wish to express my sincere gratitude to everyone who supported me, and helped me during the years of my scientific work that resulted in this thesis.

First, I am grateful to Professor Ljiljana Cvetković for her support, endless motivation, creativity, and an exceptional willingness to work with me from the very beginning of my scientific life. Thanks to her, I have decided to study applied linear algebra, and to give my scientific contribution to this field. I would also like to thank her for the wonderful friendship, and her help that she was, and still is, unselfishly giving.

My deep and sincere gratitude, I owe to Professor Richard Varga, who has always been there to work with Professor Cvetković and me. I am truly happy that I have a chance to work with such extraordinary mathematician, and I am deeply grateful for his inspiration, his open-heartedness, and his endless help. Working with him made me certainly a better scientist, and undoubtedly a better man.

I am infinitely grateful to my parents, my brother and sisters who, through the unconditional love that they carry inside, are my life's inspiration, and continuous support in everything I do.

I also thank my friends, who have offered me all their support, in mind and spirit, and helped me progress in so many ways.

Finally, I would like to thank all my former teachers and current colleagues, with whom I have discovered the beauty of mathematics and, especially, to all my students with whom I have an opportunity to do it now.

Novi Sad, Serbia — Vladimir Kostić
March 1, 2010

Abstract

The matrix property of diagonal dominance has been applied in many different ways, and proved truly beneficial in diverse areas of research. Therefore, the main motivation of this thesis is to make a unifying framework to the subject of diagonal dominance and its generalizations, that will alow some new insights in the already known facts, and open some new areas of research. Due to space limitations, the main focus of the thesis will be on the author's most recent published results, and on the new and unpublished material that is obtained through the work on this thesis. Some other areas in which diagonal, and generalized diagonal dominance can be applied and produce significant benefits, will be just briefly mentioned together with some of the author's references on it.

The outline of the thesis is the following. In the first chapter we introduce the basic concepts of diagonal dominance, its extensions and its generalization. While the first section consists of the traditional nonsingularity results, the second one includes some the original contributions of the author, which could be found in Theorems 1.2.13, 1.2.17, 1.2.23, 1.2.25, 1.2.26 and 1.2.27, which have already been published in [19], [13] and [11]. Third section focuses on the new concept of a *DD-type* class of matrices, given in Definition 1.3.7. This term is defined here for the first time, in order to make a unifying framework for all the classes of matrices that define some kind of diagonal dominance. The main result is Theorem 1.3.9, which can, besides to the classes presented in the thesis, be applied to many others, found in the literature. The utility of the new concept is closely related to the following chapters, where it will be used extensively, to produce new practical results. Final section of the first chapter covers the scaling technique, concept that was extensively developed by Ljiljana Cvetković, and the author, in several papers on different subjects. The main benefits of this technique, beside the ones presented in detail in this thesis, lie in the application to the convergence theory of the (pointwise and block) iterative methods for solving systems of linear equations: for more detail see [15, 16], then to the treatment of matrix properties, such as invariance of the Schur, and diagonal Schur complements,

[17], invariance under subdirect sums, [6, 12], and in obtaining new nonsingularity results, [18]. The main part of this section was published in the papers [19] and [14].

The second chapter is treatment of the benefits of the various generalizations of diagonal dominance in the field of eigenvalue localization. This material is, in most of the parts, present in author's M.Sc. thesis. Apart form the Varga's Equivalence Principle, which was for the first time explicitly written in [34], the main new contribution is the Isolation Principle, given in Theorem 2.2.3, which generalizes some very well known results on the disjoint localization sets, while at the same time produces new facts that were not published up to now: Theorems 2.2.5, 2.2.10, 2.2.16 and 2.2.22. Also, a full treatment of the important subject of computation of the minimal Geršgorin set, published in [52] together with R. S. Varga and Lj. Cvetković, is given in the third section of this chapter.

Third chapter consists, almost in full, of completely new material. The starting point is author's very recent joint work with Lj. Cvetković and R. S. Varga, published in [35], which covers a part of Section 3.2 and Section 3.5. The rest of the chapter consists of completely new material, which represents the development of the Geršgorin-type localization theory for the generalized eigenvalues. The backbone of the approach that is used is generalized diagonal dominance, presented in the concepts of the *DD-type* and *SDD-type* classes of matrices, from which the fundamental principles were proven: *Varga's Equivalence Principle*, *Isolation Principle*, *Boundedness Principle*, and *Approximation Principle*. Coupling these principles together with the results of the first chapter, theorems of this theory are obtained, and each of them is illustrated through numerical examples. In addition, a new parameter dependent approximation technique was proposed for future research.

The forth, and the final chapter, discuses the application of generalized diagonal dominance coupled with noncooperative game theory in the modeling of ad-hoc multihop wireless sensor networks. A certain power consumption optimization problem was proposed by Yuan and Yu in [55], and solved in the case when the network satisfies, in a certain way, strict diagonal dominance. In the last chapter of this thesis, a review of the basics of the proposed model is given, and then the main result is generalized to a wider set of realistic network setups. In addition, possible benefits from the scaling technique, developed in the first chapter, were presented, and some new insights of the application of the matrix iterative methods for solving systems of linear equations in wireless network power control algorithms were given. The material of the last chapter is published in this thesis for the first time.

Apstrakt

Matrične osobine tipa dijagonalne dominacije primenjivane su na mnoge različite načine i pokazale se istinski korisnim u različitim oblastima linearne algebre i njenih primena. Osnovna motivacija ove teze je da stvori jedinstveni radni okvir za temu dijagonalne dominacije i njenih uopštenja, koji treba da omogući nove uvide u već poznate činjenice, kao i da otvori neke nove oblasti istraživanja. Usled prostornog ograničenja, u fokusu ove teze biće autorovi najnoviji objavljeni rezultati, kao i novi neobjavljen materijal koji je nastao tokom rada na ovoj tezi. Neke od ostalih oblasti u kojima se dijagonalna i generalizovana dijagonalna dominacija mogu upotrebiti i dovesti do značajnih poboljšanja, kratko će biti pomenuti, zajedno sa nekim od autorovih referenci na tu temu.

Sastav teze je sledeći. U prvom poglavlju uvodimo osnovni koncept dijagonalne dominacije, njegova proširenja i uopštenja. Dok se prva sekcija sastoji od tradicionalnih rezultata o regularnosti, druga sadrži neke od autorovih originalnih doprinosa, koji se mogu naći u Teoremama 1.2.13, 1.2.17, 1.2.23, 1.2.25, 1.2.26 i 1.2.27, a koji su već objavljeni u radovima [19], [13] i [11]. Treća sekcija se fokusira na novi koncept klasa matrica *DD-tipa*, dat u Definiciji 1.3.7. Taj pojam je definisan po prvi put u ovoj tezi, sa ciljem da pruži jedinstveni okvir za sve klase matrica koje su definisane pomoću nekog oblika dijagonalne dominacije. Glavni rezultat ovog poglavlja je Teorema 1.3.9, koja se može, pored klasa prezentovanih u okviru ove teze, primeniti na mnoge klase matrica koje se nalaze u literaturi. Korisnost ovog novog koncepta je blisko povezana sa narednim poglavljima, gde će on biti intenzivno korišćen u cilju dobijanja novih praktičnih rezultata. Poslednja sekcija prvog poglavlja se bavi tehnikom skaliranja, konceptom koji su Lj. Cvetković i autor razvili i iscrpno koristili u nekoliko radova na različite teme. Osnovne prednosti ove tehnike, pored onih koje su iznete u ovoj tezi, nalaze se u primenama u teoriji konvergencije (tačkastih i paralelnih) iterativnih postupaka za rešavanje sistema linearnih jednačina, za više detalja pogledati [15, 16], u ispitivanju matričnih osobina, kako što su invarijantnost Šurovog komplementa i dijagnoalnog Šurovog komplementa, [17], invarijantnost pri subdirektnom

sabiranju, [6, 12], kao i primene u ostvarivanju novih rezultata o regularnosti matrica, [18]. Najveći deo poslednje sekcije prvog poglavlja obljavljen je u [19] i [14].

Drugo poglavlje se bavi prednostima različitih uopštenja dijagonalne dominacije u oblasti lokalizacije krakterističnih korena matrica. Ovaj materijal je u svom najvećem delu prisutan u autorovoj magistarskoj tezi. Pored Varginog principa ekvivalencije, koji je po prvi put eksplicitno zapisan u [34], osnovni doprinosi su Princip izolacije, dat u Teoremi 2.2.3, koji uopštava dobro poznate rezultate o disjunktnim oblastima lokalizacje, istovremeno dajući nova tvrdjenja koja nisu objavljena do sada: Teoreme 2.2.5, 2.2.10, 2.2.16 i 2.2.22. Pored toga, u poslednjoj sekciji ovog poglavlja, u potpunosti je obradjena važna tema o izračunavanju minimalnog Geršgorinovog skupa, objavljena u [52], zajedno sa R. S. Vargom i Lj. Cvetković.

Treće poglavlje se sastoji, gotovo u potpunosti, od novog materijala. Polazna tačka je autorov skorašnji rad sa Lj. Cvetković i R. S. Vargom, objavljen u [35], koji je prezentovan u Sekciji 3.2 i Sekciji 3.5. Ostatak poglavlja se sastoji od novog materijala, koji predstavlja razvoj teorije lokalizacije generalizovanih karakterističnih korena Geršgorinovog tipa. Osnova pristupa koji je korišćen je generalizovana dijagonalna dominacija, prisutna u konceptima klasa matrica *DD-tipa* i *SDD-tipa*, na osnovu kojih su dokazani osnovni principi: *Varгin princip ekvivalencije*, *princip izolacije*, *princip ograničenosti* i *princip aproksimacije.* Povezujući ove principe sa rezultatima iz prvog poglavlja, dokazane su odgovarajuće teoreme su dobijene i svaka od njih je ilustrovana numeričkim primerima. Pored toga, tehnika aproksimacije, koja zavisi od slobodnih parametara, predložena je za dalja istraživanja.

Četvrto i poslednje poglavlje razmatra primenu generalizovane dijagonalne dominacije, zajedno sa teorijom nekooperativnih igara u modelovanju ad-hoc multihop bežičnih senzor mreža. Polazna tačka je odredjeni problem optimizacije potrošnje struje, koji je predložen od strane Yuan i Yu u [55], gde je i rešen u slučaju kada mreža zadovoljava, u odredjenom smislu, strogu dijagonalnu dominaciju. Nakon pregleda osnova predloženog modela, glavni rezultat je uopšten na širi skup relističnih mrežnih postavki. Pored toga, prikazana je i moguća korist od tehnike skaliranja, razvijene u prvom poglavlju, kao i novi uvidi u primenu matričnih ierativnh postupaka za rešavanje sistema linearnih jednačina u okviru algoritma za optimizaciju kontrole napona u bežičnim senzor mrežama. Materijal ovog posledenjeg poglavlja je, takodje, originalni doprinos ove teze.

Table of Contents

Acknowledgements **vii**

Abstract **ix**

Apstrakt **xi**

Table of Contents **xv**

Introduction **1**

1 Diagonal and Generalized Diagonal Dominance **5**
1.1 Diagonally and Strictly Diagonally Dominant Matrices 7
1.1.1 Irreducibility 9
1.1.2 Alternatives to Irreducibility 13
1.2 Extensions of (Strictly) Diagonal Dominant Matrices 15
1.2.1 Extensions by multiplication 15
1.2.2 Extensions via graph theory 17
1.2.3 Extensions by partitions 22
1.2.4 Extensions by column sums 26
1.3 Generalized Diagonally Dominant Matrices 33
1.4 Scaling Approach 37

2 Eigenvalue Localization **41**
2.1 Geršgorin's Theorem 43
2.1.1 Geršgorin's theorem and diagonal similarities 47
2.1.2 Geršgorin's theorem and matrix transpose 49
2.1.3 Geršgorin's theorem and irreducibility 51
2.2 Geršgorin-type Theorems 55
2.2.1 Brauer's ovals of Cassini 58
2.2.2 Brualdi sets 61
2.2.3 Cvetković-Kostić-Varga Sets 64
2.2.4 Ostrowski sets 70
2.3 Minimal Geršgorin Sets 77
2.3.1 Sharpness and geometry of the minimal Geršgorin set 78
2.3.2 Computation of the minimal Geršgorin set 80

3 Localization of Generalized Eigenvalues **85**
3.1 Generalized Eigenvalues 87
3.2 Geršgorin's Theorem for the Generalized Eigenvalues 91
3.3 Geršgorin-type Localizations for Generalized Eigenvalues 101
3.3.1 Brauer set for Generalized Eigenvalues 105
3.3.2 Brualdi set for Generalized Eigenvalues 112
3.3.3 Cvetković-Kostić-Varga set for Generalized Eigenvalues 116
3.3.4 Ostrowski sets for Generalized Eigenvalues 125
3.4 Improved Approximations of the Generalized Geršgorin-type Sets 129
3.5 Minimal Geršgorin Set for the Generalized Eigenvalues 137

4 Application of Generalized Diagonal Dominance in Wireless Sensor Network Optimization Problems **153**

Bibliography **167**

Short Biography **173**

Kratka biografija **175**

Introduction

Although this may seem a paradox, all exact science is dominated by the idea of approximation.

Bertrand Russell
(1872-1970)

Diagonally dominant matrices drew the attention of several great mathematicians of the nineteenth century, and inspired important research in the mathematics of twentieth century. Actually, from the very beginning of the matrix theory, when Sylvester[1] in 1850 distinguished the notion of 'matrix' from the concept of 'determinant', the question of nonsingularity was one of the main topics. Various properties were studied in order to catch the connection between the value and the structure of the entries in a matrix, and the value of its determinant. The beautiful idea of investigating the dominance of the diagonal in a matrix, in order to ensure a nonzero determinant, can be traced back to the paper of Lévy[2] in 1881, [36]. Nevertheless, the fact that strictly diagonally dominant matrices are nonsingular has been repeatedly rediscovered, in several different equivalent formulations, and not always explicitly, by quite a few mathematicians, throughout 1881-1949. In the work of Desplanques, [22] it can be found for all real matrices, while in the work of Minkowski[3] in 1900, [38], and, independently Hadamard[4] in 1903, [27], it is stated for any complex matrix. In the work of Nekrasov[5] in 1892 it was implicitly stated through the convergence of Gauss-Seidel iteration, while Hopf[6] obtained it, most probably independently, in 1929, in order to prove a fixed point theorem. It is interesting

[1] James Joseph Sylvester (1814 - 1897), an English mathematician who made fundamental contributions to matrix theory, invariant theory, number theory, partition theory and combinatorics.

[2] L. Lévy, a French mathematician, who discovered, under certain limitations on the sign of the matrix entries, that a matrix whose diagonal entries are, by absolute value, strictly greater than the sum of the absolute values of the off-diagonal entries from the same row, has a nonzero determinant.

[3] Hermann Minkowski (1864 - 1909), a German mathematician, who created, and developed the geometry of numbers, and used geometrical methods to solve difficult problems in number theory, mathematical physics, and the theory of relativity.

[4] Jacques Salomon Hadamard (1865 - 1963), a French mathematician, who made major contributions in number theory, complex function theory, differential geometry and partial differential equations.

[5] Pavel Alekseevich Nekrasov (1853 - 1924), a Russian mathematician, who made important contributions to algebra, analysis, probability and mechanics.

[6] Heinz Hopf (1894 - 1971) a German mathematician who made a major influence in algebraic topology.

to note, also, that Bankwitz in 1930 proved it, by referring to Perron[7]-Frobenius[8] theory (1907-1912), and applied it to knot theory.

But, definitely, one of the most important formulation of this theorem is due to Geršgorin[9], who stated in his, paper [25] from 1931, that the eigenvalues of an arbitrary matrix of order $n \in N$, can be localized in the complex plane using n simple circles. This result was a sensation of that time, and in a scientific community, it initiated the lot of enthusiasm in the direction of "simple approximations" of the spectra of matrices. The reason, of course, lies in the elegance of the result and its wide applicability, both theoretical, and practical. By this theorem, and its generalizations obtained during the years, many matrix properties that depend of the structure of its eigenvalues, found an alternative formulations and effective testing methods. Until today, this subject remaines an active area of research, and it still attracts by its elegance and simplicity.

But, the work of Geršgorin had another important influence to the theory of matrices. Namely, Geršgorin, who at that time worked in the Leningrad Mechanical Institute, luckily had published his result in his only paper written in German. But, nevertheless, it did not attract the full attention of the scientific community until the beginning of World War II, and the growing need for the applications of mathematics in engineering. Predecessors of the first computers motivated research in the matrix theory via numerical computation. And, at that time, in the period between the 1943 and 1949, Olga Taussky[10] was working in the National Physics Laboratory in Teddington, near London. She was doing research on the problem of an airplane design that guarantees stability of the aircraft. In order to speed up the necessary calculations of the eigenvalues of a certain square 6×6 matrix, she uses the elegant result of a Russian mathematician. Charmed with its elegance and simplicity, she promotes it through her further studies, corrections and applications. In this period she publishes her results, first for the British Ministry of Aviation, and then in the American Mathematical Monthly, and, thus, opens a new exciting chapter of the use of diagonal dominance and the localization of eigenvalues.

Techniques introduced by Olga Taussky became an inspiration to many, and this field of research grows, and develops even today. Works of Ostrowski, Brauer, Brualdi and Varga are just some of many contributions. It is interesting to note that, although in her paper "A recurring theorem in determinants", from the American Mathematical Monthly, Olga Taussky makes a link between Geršgorin's theorem and the theorem on the nonsingularity that dates from the end of the nineteenth century, the relationship of the diagonal dominance and the localization of eigenvalues, although in the continuous use, remained without a precise and complete formulation until the end of the twentieth century, when

[7]Oskar Perron (1880 - 1975), a German mathematician, who made numerous contributions concerning differential equations and partial differential equations, and was also famous for his encyclopedic book on continued fractions.

[8]Ferdinand Georg Frobenius (1849 - 1917), a German mathematician, best-known for his contributions to the theory of differential equations and to group theory, also gave the first full proof for the CayleyHamilton theorem.

[9]Semyon Aranovich Gershgorin (1901 -1933), a Russian mathematician, famous for his circle theorem.

[10]Olga Taussky Tod (1906 - 1995) an Austrian mathematician famous for her work in matrix theory, algebraic number theory and differential equations, also known as a torchbearer for the matrix theory.

it was stated in the book of Richard Varga, "*Geršgorin and His Circles*".

As we could notice, the idea of diagonal dominance has been, from the very beginning, applied in many different ways, and proved truly beneficial in many different areas of research. So, the main motivation of this thesis is to make a unifying framework to the subject of diagonal dominance and its generalizations, that will alow some new insights in the already known facts, and open some new areas of research. Due to space limitations, the main focus of the thesis will be on the author's most recent published results, and on the new and unpublished material that is obtained through the work on this thesis. Therefore, some other areas in which diagonal, and generalized diagonal dominance can produce significant benefits, will be just briefly mentioned in the following of this introduction, together with some of the author's references on it.

The outline of the thesis is the following. In the first chapter we introduce the basic concepts of diagonal dominance, its extensions and its generalization. While the first section consists of the traditional nonsingularity results, the second one includes some the original contributions of the author, which could be found in Theorems 1.2.13, 1.2.17, 1.2.23, 1.2.25, 1.2.26 and 1.2.27, which have already been published in [19], [13] and [11]. Third section focuses on the new concept of a *DD-type* class of matrices, given in Definition 1.3.7. This term is defined here for the first time, in order to make a unifying framework for all the classes of matrices that define some kind of diagonal dominance. The main result is Theorem 1.3.9, which can, besides to the classes presented in the thesis, be applied to many others, found in the literature. The utility of the new concept is closely related to the following chapters, where it will be used extensively, to produce new practical results. Final section of the first chapter covers the scaling technique, concept that was extensively developed by Ljiljana Cvetković, and the author, in several papers on different subjects. The main benefits of this technique, beside the ones presented in detail in this thesis, lie in the application to the convergence theory of the (pointwise and block) iterative methods for solving systems of linear equations; for more detail see [15, 16], then to the treatment of matrix properties, such as invariance of the Schur, and diagonal Schur complements, [17], invariance under subdirect sums, [6, 12], and in obtaining new nonsingularity results, [18]. The main part of this section was published in the papers [19] and [14].

The second chapter is treatment of the benefits of the various generalizations of diagonal dominance in the field of eigenvalue localization. This material is, in most of the parts, present in author's M.Sc. thesis. Apart form the Varga's Equivalence Principle, which was for the first time explicitly written in [34], the main new contribution is the Isolation Principle, given in Theorem 2.2.3, which generalizes some very well known results on the disjoint localization sets, while at the same time produces new facts that were not published up to now: Theorems 2.2.5, 2.2.10, 2.2.16 and 2.2.22. Also, a full treatment of the important subject of computation of the minimal Geršgorin set, published in [52] together with R. S. Varga and Lj. Cvetković, is given in the third section of this chapter.

Third chapter consists, almost in full, of completely new material. The starting point is author's very recent joint work with Lj. Cvetković and R. S. Varga, published in [35], which covers a part of Section 3.2 and Section 3.5. The rest of the chapter consists of completely new material, which represents the development of the Geršgorin-type localization theory for the generalized eigenvalues. The backbone of the approach that is used is generalized diagonal dominance, presented in the concepts of the *DD-type* and *SDD-type* classes of matrices, from which the fundamental principles were proven: *Varga's Equivalence Principle*, *Isolation Principle*, *Boundedness Principle*, and *Approximation Principle*. Coupling these principles together with the results of the first chapter, theorems of this theory are obtained, and each of them is illustrated through numerical examples. In addition, a new parameter dependent approximation technique was proposed for future research.

The forth, and the final chapter, discuses the application of generalized diagonal dominance coupled with noncooperative game theory in the modeling of ad-hoc multihop wireless sensor networks. A certain power consumption optimization problem was proposed by Yuan and Yu in [55], and solved in the case when the network satisfies, in a certain way, strict diagonal dominance. In the last chapter of this thesis, a review of the basics of the proposed model is given, and then the main result is generalized to a wider set of realistic network setups. In addition, possible benefits from the scaling technique, developed in the first chapter, were presented, and some new insights of the application of the matrix iterative methods for solving systems of linear equations in wireless network power control algorithms were given. The material of the last chapter is published in this thesis for the first time.

Chapter 1

Diagonal and Generalized Diagonal Dominance

Starting from the concepts of diagonal and strict diagonal dominance, [36], [38], [22] and [27], given in the first Section of this Chapter, in the following one we will investigate different extensions of these two concepts that have been developed during the years. The third Section will introduce the class of generalized diagonally dominant matrices that coincides with the class of H-matrices, and, by that fact, has a strong connection with nonnegative matrices, [3]. The final Section of this Chapter will introduce the original technique that was developed and used throughout the dissertation in order to obtain different improvements of known facts in various fields of applied mathematics, and to discover new ones.

1.1 Diagonally and Strictly Diagonally Dominant Matrices

Throughout this dissertation, for an arbitrary n from the set of positive integers $\mathbb{N}$, by $\mathbb{C}^n$, we denote a complex n-dimensional vector space of column vectors $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$, where $x_i \in \mathbb{C}$, $i = 1, 2, \dots, n$, and, for arbitrary $m, n \in \mathbb{N}$, by $\mathbb{C}^{m,n}$, we denote the collection of all $m \times n$ matrices with complex entries. A matrix $A \in \mathbb{C}^{m,n}$ is denoted by $A = [a_{i,j}]$, or by

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}, \tag{1.1.1}$$

where $a_{i,j} := (A)_{i,j} \in \mathbb{C}$, for all $1 \leq i \leq m$ and $1 \leq j \leq n$.

We denote the set of indices by $N := \{1, 2, \dots, n\}$, and

$$r_i(A) := \sum_{j \in N \setminus \{i\}} |a_{i,j}| \quad (i \in N) \tag{1.1.2}$$

is called the **i-th deleted absolute row sum**[1] of the matrix A. If the matrix A is of the dimension $n = 1$, we define $r_1(A) := 0$.

In a similar way, by $\mathbb{R}^n$ and $\mathbb{R}^{m,n}$, we denote, respectively, the real n-dimensional vector space of vector columns and the collection of rectangular matrices with real entries.

With I_n we denote the $n \times n$ **identity matrix**, i.e., a matrix whose diagonal entries are all equal to one, while its off-diagonal entries are all zero.

With the previous notations, we state the famous first nonsingularity result, which appeared in the early paper of Lévy in 1881, [36], and then latter, independently, in the work of Minkowski in 1900, [38], (in both cases stated only for *real* matrices). The complex case was covered in a paper of Desplanque in 1887, [22], and the book of Hadamard in 1903, [27]. In contemporary notation, basically, all of them stated the following.

Theorem 1.1.1. (Lévy-Desplanques)[2] *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ be an arbitrary matrix. If*

$$|a_{i,i}| > r_i(A) \quad \textit{for all } i \in N, \tag{1.1.3}$$

then A is nonsingular.

Proof. Let us suppose the contrary, i.e., that the matrix A is singular. Then, there exists a nonzero vector $\mathbf{x} = [x_1, x_2, \dots, x_n] \in \mathbb{C}^n$, such that $A\mathbf{x} = \mathbf{0}$, or, equivalently, $\sum_{j \in N} a_{i,j} x_j = 0$, for each $i \in N$. Since $\mathbf{x} \neq \mathbf{0}$, there exists an index $k \in N$, so that $0 < |x_k| = \max\{|x_j| : j \in N\}$. For this index k we have that

$$\sum_{j \in N \setminus \{k\}} a_{k,j} x_j = -a_{k,k} x_k,$$

[1] In the following, we will use the term "i-th row sum", or simply, "row sum", whenever there is no possible confusion.

[2] Since the theorem was published in the well-known book [27] it is also know as "Hadamard's Theorem".

which, on taking absolute values, and applying the triangle inequality, gives

$$|a_{k,k}||x_k| \leq \sum_{j \in N\setminus\{k\}} |a_{k,j}||x_j| \leq |x_k| \sum_{j \in N\setminus\{k\}} |a_{k,j}|.$$

By dividing the last inequality by $|x_k| > 0$, we get

$$|a_{k,k}| \leq r_k(A),$$

which obviously contradicts the assumption (1.1.3). Thus, A is a nonsingular matrix. □

Based on the condition (1.1.3) which defines them, matrices that satisfy the previous theorem are called **strictly diagonally dominant matrices**, or, briefly, **SDD** matrices. Their beauty lies in the fact that, while it takes a lot of computation to determine whether $det(A)$ is equal to zero or not, the condition (1.1.3) is rather easy to verify. On the other hand, condition (1.1.3) is not a necessary condition for nonsingularity, as the subsequent example easily shows, so we are working with a subclass of the class of all nonsingular matrices. How one can use its nonsingularity in various ways, knowing that a matrix is an SDD matrix, will be studied in the following pages.

For a square matrix of dimension n, it suffices to check n inequalities, i.e., to check if, in each row, the diagonal entry has an absolute value which is strictly greater than the sum of the absolute values of off-diagonal entries. The rows which fulfill this condition will be called the **SDD rows** of a matrix. So, SDD matrices are matrices that have *all* rows SDD. A natural question to ask is the: Is it necessary to have all the rows to be SDD, in order to guarantee nonsingularity? The next example gives us first insights.

Example 1.1.2. *For each of the following matrices, their determinant is computed and their SDD rows are shown in boldface.*

$$A_1 = \begin{bmatrix} 1 & 1 \\ \mathbf{i} & \mathbf{2i} \end{bmatrix}, \quad det(A_1) = i \neq 0,$$

$$A_2 = \begin{bmatrix} 1 & -i & 0 \\ 2 & 1 & 0 \\ \mathbf{1} & \mathbf{1} & \mathbf{3} \end{bmatrix}, \quad det(A_2) = 3 + 6i \neq 0,$$

$$A_3 = \begin{bmatrix} i & 1 & 0 & 1 & 0 & 1 \\ 1 & -i & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 2 & i & i \\ 0 & 0 & 0 & 1 & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{i} & \mathbf{i} & \mathbf{-3} \end{bmatrix}, \quad det(A_3) = 0.$$

$$A_4 = \begin{bmatrix} \mathbf{1} & \mathbf{0.2} & \mathbf{0.5} & \mathbf{0.2} \\ \mathbf{0} & \mathbf{1} & \mathbf{0.4} & \mathbf{0.5} \\ 0 & 0 & 1 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}, \quad det(A_4) = 0,$$

$$A_5 = \begin{bmatrix} \mathbf{1} & \mathbf{0.2} & \mathbf{0.5} & \mathbf{0.2} \\ \mathbf{0} & \mathbf{1} & \mathbf{0.4} & \mathbf{0.5} \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad det(A_5) = 1.32 \neq 0.$$

After a closer look at the matrices of the previous example, it is interesting to note that, first, even one non-SDD row can be sufficient to cause singularity of a matrix. On the other hand, even one SDD row can lead to nonsingularity, too. So, it is of interest to discover at which point less then all SDD rows will assure nonsingularity of a matrix, and what are the other factors that can interfere. From the first steps in this direction, up to know, many different extensions of the notion of strict diagonal dominance were developed. Here, in the second section of this chapter we will cover some of them.

Before we continue any further, a first observation that arises is what can happen if the strict inequalities of (1.1.3) are replaced by non-strict ones. In another words, is the matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, for which

$$|a_{i,i}| \geq r_i(A) \quad \text{for all } i \in N, \tag{1.1.4}$$

nonsingular, or not.

Since we are dealing with diagonal dominance, a natural restriction to the inequalities (1.1.4) is to have at least one of them to be strict. Such matrices we will call **diagonally dominant matrices**. Otherwise, if all inequalities in (1.1.4) are in fact equalities, there is no reason to call the matrix in any way "diagonally dominant".

More precisely we use the following definition.

Definition 1.1.3. An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called **diagonally dominant matrix**[3] if

$$|a_{i,i}| \geq r_i(A) \quad \text{for all } i \in N, \tag{1.1.5}$$

and for at least one index $k \in N$,

$$|a_{k,k}| > r_k(A). \tag{1.1.6}$$

Question of nonsingularity of DD matrices arose first with the works of Geršgorin in 1931, [25], and Olga Taussky-Todd in 1948, [45]. Both of these papers dealt with the localizations of eigenvalues that were directly related to the nonsingularity of diagonally dominant matrices, which we will present in detail in the second chapter. In the sense of eigenvalue localization, Geršgorin prematurely concluded that DD matrices are always nonsingular, but as matrices A_4 and A_5 of the previous example imply, that is not, in general, true. Interested in his work, and especially intrigued by this case, Olga Taussky-Todd noticed, in her famous paper [45], the Geršgorin error. In order to fix it, she opened a beautiful chapter in the theory of matrices, by introducing graph theory and the notion of irreducibility.

1.1.1 Irreducibility

As we have already mentioned, for a DD matrix, it is not sufficient to observe only the number of strict inequalities and equalities in (1.1.4), in order to determine nonsingularity

[3] Or, briefly, **DD matrices**.

criteria. In fact, we need to take into account a matrix structure, as one can see from matrices A_4 and A_5 of Example 1.1.2. Namely, the singular matrix A_4 has a block of zero entries in lower left corner, while that is not the case for nonsingular matrix A_5.

Motivated by this observation, we introduce the following definitions.

Definition 1.1.4. A matrix $P \in \mathbb{R}^{n,n}$ is said to be a **permutation matrix** if there is a permutation π, i.e., an one-to-one mapping $\pi : N \to N$, such that $P = [p_{i,j}] := [\delta_{i,\pi(j)}] \in \mathbb{R}^{n,n}$, where $\delta_{k,l}$ is a familiar **Kronecker delta function**, i.e.,

$$\delta_{k,l} := \begin{cases} 1, & k = l, \\ 0, & k \neq l. \end{cases}$$

Definition 1.1.5. Given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, then A is **reducible** if there exists a permutation matrix $P \in \mathbb{R}^{n,n}$ and an integer r, $1 \leq r < n$, such that

$$P^T A P = \begin{bmatrix} A_{1,1} & A_{1,2} \\ O & A_{2,2} \end{bmatrix}, \tag{1.1.7}$$

where $A_{1,1} \in \mathbb{C}^{r,r}$, and $A_{2,2} \in \mathbb{C}^{n-r,n-r}$. Otherwise, if such a permutation matrix does not exist, we say that matrix A is **irreducible**. If $A \in \mathbb{C}^{1,1}$, then A is irreducible if its single entry is nonzero, and reducible otherwise.

The matrix property of irreducibility, as a tool in linear algebra, appears in a clear form, at first, in early papers of Olga Taussky in 1949, [46], Theorem 1.12. It illustrates an extremely close relation between matrix theory and graph theory. Namely, we can observe matrices as weighted directed graphs, or **weighted digraphs**, where the indices of rows, or columns, represent vertices, while nonzero matrix entries represent directed edges with respected weights.

However, in our study of irreducibility, only the zero/nonzero structure of a matrix is important. Thus, instead of observing all the matrix data represented by a weighted digraph, we can restrict attention to the standard **digraph of a matrix** A. More precisely, given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $\{v_1, v_2, \ldots, v_n\}$, we denote n distinct objects, or points, which we call **vertices**. For every nonzero entry $a_{i,j}$ of a matrix A, we create a directed edge $\overrightarrow{v_i v_j}$ which connects the vertex v_i with a vertex v_j. In the case when $i = j$, i.e., when $a_{i,i} \neq 0$, the edge $\overrightarrow{v_i v_i}$ is called a *loop*.

The set of all such directed edges $\mathbb{E}(A) := \{\overrightarrow{v_i v_j} : a_{i,j} \neq 0,\ i, j, \in N\}$, together with the set of vertices $\mathbb{V}(A) := \{v_1, v_2, \ldots, v_n\}$ makes the **digraph** $\mathbb{G}(A)$ **attributed to the matrix**[4] A. To illustrate this, let us observe the matrices A_1, A_2 and A_3, of Example 1.1.2, whose graphs $\mathbb{G}(A_1)$ and $\mathbb{G}(A_2)$ are given in Figure 1.1.1, and $\mathbb{G}(A_3)$ in Figure 1.1.2.

Based on these examples, we can note a few interesting facts. First, concerning the connectivity of the vertices in graphs, the edges can abut one another and form, in such a way, **paths** which connect one vertex with another. So, in the graph of the matrix A,

[4]Throughout the text, we will also use the term "**graph of a matrix**".

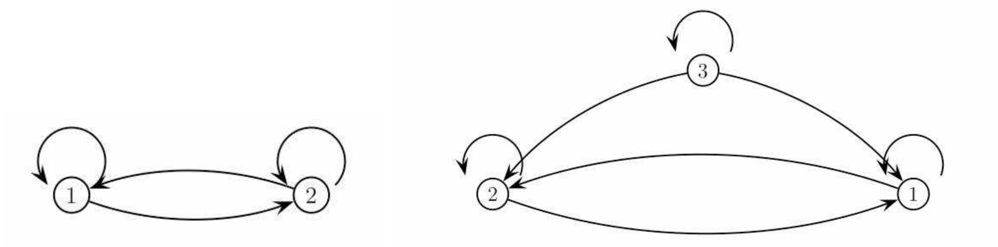

Figure 1.1.1: From left to right, graphs of matrices A_1 and A_2 of the Example 1.1.2
(*Sa leva na desno, grafovi matrica* A_1 *and* A_2 *iz Primera 1.1.2)*

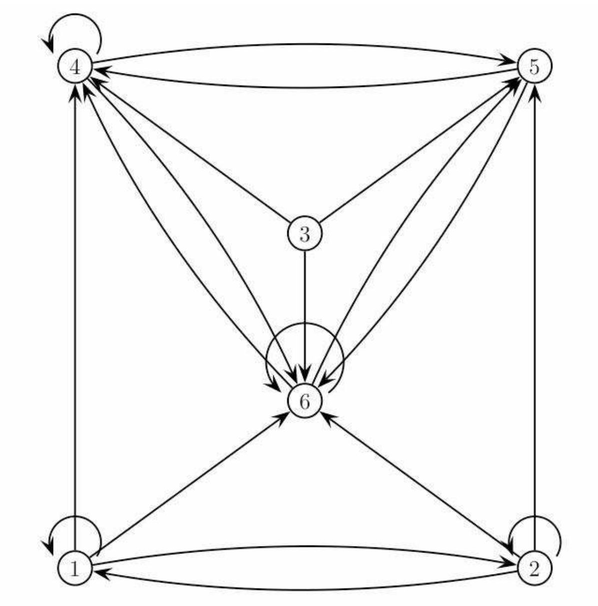

Figure 1.1.2: Graph of a matrix A_3 of the Example 1.1.2
(Graf matrice A_3 *iz Primera 1.1.2)*

there exists a path $\overrightarrow{v_{i_0}v_{i_1}}, \overrightarrow{v_{i_1}v_{i_2}}, \ldots, \overrightarrow{v_{i_{\ell-1}}v_{i_\ell}}$, if and only if the sequence of matrix indices $\{i_j\}_{j=0}^{\ell} \subseteq N$ is such that $a_{i_{j-1},i_j} \neq 0$, for all $1 \leq j \leq \ell$, which can also be written as

$$\prod_{j=1}^{\ell} a_{i_{j-1},i_j} \neq 0.$$

Second, the structure of the graph of a matrix can be such that it consists of a "block" in which all vertices are connected with each other, as in the case of graph $\mathbb{G}(A_1)$, or, in turn, may consists of several such "blocks", which are not mutually connected to each other, as in the case[5] of graphs of matrices A_2 and A_3. Inspired by this, we can observe the following definition.

Definition 1.1.6. The digraph $\mathbb{G}(A)$ of the matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is **strongly connected** if, for each ordered pair of distinct vertices $\langle v_i, v_j \rangle$, there exists a directed path in $\mathbb{G}(A)$ from the vertex v_i to the vertex v_j.

Third, notice that a permutation of the matrix entries leaves the graph structure unchanged, i.e., the graph $\mathbb{G}(P^T AP)$, for an arbitrary permutation matrix P, is just $\mathbb{G}(A)$ after an adequate enumeration of its vertices. This observation leads to the following proposition.

Theorem 1.1.7. ([3]) *An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is irreducible if and only if its graph $\mathbb{G}(A)$ is strongly connected.*

Now, we are ready to deal with the question of diagonal dominance. Going back to our example, we see that the singular matrix A_4 is reducible, while the nonsingular matrix A_5 is irreducible. Thus, one could suppose that irreducibility was needed to be coupled with the diagonal dominance in order to achieve nonsingularity of a matrix.

That this observation is true, is, in fact, the result of the Theorem 1.4 of Olga Taussky in [46], which we present here as it was given in [51].

Theorem 1.1.8. (Olga Taussky) *Every irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is diagonally dominant is nonsingular.*

Proof. Since the case $n = 1$ directly follows from the Definition 1.1.3, we can assume that $n \geq 2$. Having $n \geq 2$, we suppose, on the contrary, that the matrix A is singular. Then, zero is an eigenvalue of the matrix A, and there exists its (nonzero) eigenvector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{C}^n$ such that $A\mathbf{x} = \mathbf{0}$. Since the last equation is homogenous in $\mathbf{x}$, we can normalize it so that $\max\{|x_i| : i \in N\} = 1$. Letting $S := \{j \in N : |x_j| = 1\}$, then S is a nonempty subset of N. Now, the equation $A\mathbf{x} = \mathbf{0}$ implies that $\sum_{j \in N} a_{k,j} x_j = 0$, for all $k \in N$, or equivalently,

$$-a_{k,k} x_k = \sum_{j \in N \setminus \{k\}} a_{k,j} x_j \quad (k \in N).$$

[5]In $\mathbb{G}(A_3)$, from the "block" $\{4, 5, 6\}$ there is no directed edge to the "block" $\{1, 2\}$.

Then, considering $i \in S$ and applying the triangle inequality, after taking absolute values, we get, from the previous equation for $k = i$, that

$$|a_{i,i}| \leq \sum_{j \in N \setminus \{i\}} |a_{i,j}||x_j| \leq r_i(A) \quad (i \in S).$$

Because the reverse inequality of the above must hold by hypothesis from (1.1.5), then

$$|a_{i,i}| = \sum_{j \in N \setminus \{i\}} |a_{i,j}||x_j| = r_i(A) \quad (i \in S). \tag{1.1.8}$$

Now, consider an arbitrary $i \in S$. The terms in the sum of (1.1.8) cannot, due to irreducibility, all vanish. So, for an arbitrary $a_{i,j} \neq 0$, where $j \in N$, $j \neq i$, from the last equality in (1.1.8), it follows that $|x_j| = 1$, i.e., $j \in S$.

On the other hand, since strict inequality must hold, by hypothesis, in (1.1.6), for at least one index i, it follows that S has to be a proper subset of N, i.e., $\emptyset \neq S \subsetneq N$. Thus, we can take $i_0 \in S$ and $i_\ell \in N \setminus S$. Since the graph $\mathbb{G}(A)$ is strongly connected by Theorem 1.1.7, there exists a path of directed edges from the vertex v_{i_0} to the vertex v_{i_ℓ}, i.e., there exist a series of nonzero entries $a_{i_0,i_1}, a_{i_1,i_2}, \ldots, a_{i_{\ell-1},i_\ell}$. But, applying the previous conclusions consecutively on the series, we obtain that $i_\ell \in S$, which is an obvious contradiction. Thus, the matrix A is nonsingular. □

Matrices defined by the Taussky theorem will be simply called **irreducibly diagonally dominant**, or, briefly, **iDD** matrices.

1.1.2 Alternatives to Irreducibility

We will finish this section with a few interesting results on DD matrices that can assure nonsingularity and avoid irreducibility, which is, in practice, rather difficult to check.

Given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, consider the standard splitting $A = D - L - U$, where D is a **diagonal matrix**, while L and U are, respectively, **strictly lower** and **strictly upper triangular** matrices. More precisely, let $D = \text{diag}(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$, $L = [l_{i,j}]$, where

$$l_{i,j} = \begin{cases} -a_{i,j}, & j < i, \\ 0, & \text{otherwise}, \end{cases}$$

and $U = [u_{i,j}]$, where

$$u_{i,j} = \begin{cases} -a_{i,j}, & j > i, \\ 0, & \text{otherwise}. \end{cases}$$

We define the quantities $l_i(A) := r_i(L)$ and $u_i(A) := r_i(U)$, or, equivalently $l_i(A) := \sum_{j<i} |a_{i,j}|$ and $u_i(A) := \sum_{j>i} |a_{i,j}|$. Using these notations, we present the result of Beauwens, [2].

Definition 1.1.9. A matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called **lower semistrictly diagonally dominant** (**lsDD**) if $|a_{i,i}| \geq r_i(A)$ for all $i \in N$, and $|a_{i,i}| > l_i(A)$ for all $i \in N$.

Definition 1.1.10. A matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called **semistrictly diagonally dominant** (**sDD**) if there exists a permutation matrix P such that $P^T AP$ is lower semistrictly diagonally dominant.

Theorem 1.1.11. (Beauwens) *A matrix is irreducibly diagonally dominant if and only if it is semistrictly diagonally dominant.*

Since every permutation matrix is nonsingular, this means that lsDD matrices are nonsingular, too. So, the real practical benefit from this result lies in the fact that we do not need to have a special structure of a matrix, or to investigate the connectivity of its graph in order to obtain nonsingularity. Instead, we need to check additional n inequalities.

Another interesting result is due to Cvetković, [51], Exercise 1 on page 17. It covers a special case of DD matrices, when only one row is not an SDD one. In this case too, no irreducibility is needed to conclude nonsingularity.

Theorem 1.1.12. (Cvetković) *Given a matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, *such that* $|a_{i,i}| > r_i(A)$, *while* $1 \leq i \leq n-1$, *and* $|a_{n,n}| = r_n(A)$, *then* A *is nonsingular if* $r_n(A) > 0$, *and singular if* $r_n(A) = 0$.

Proof. The case when $r_n(A) = 0$ is obvious, since the whole n-th row of a matrix is zero. So, we assume that $|a_{n,n}| = r_n(A) > 0$. Since for every $i = 1, 2, \ldots, n-1$, we have that $|a_{i,i}| > r_i(A)$, or equivalently

$$|a_{i,i}| > \sum_{j=1}^{n-1} |a_{i,j}| + |a_{i,n}|,$$

there exists sufficiently small $\varepsilon > 0$, such that

$$|a_{i,i}| > \sum_{j=1}^{n-1} |a_{i,j}| + (1+\varepsilon)|a_{i,n}|, \quad \text{for all } i = 1, 2, \ldots, n-1. \tag{1.1.9}$$

On the other hand, for this $\varepsilon > 0$ we have that

$$(1+\varepsilon)|a_{n,n}| = r_n(A) + \varepsilon|a_{n,n}| > r_n(A). \tag{1.1.10}$$

Now, consider the nonsingular diagonal matrix $W = \text{diag}(1, 1, \ldots, 1, 1+\varepsilon) \in \mathbb{R}^{n,n}$. From (1.1.9) and (1.1.10), it follows that the matrix $AW = [a_{i,j}w_j]$ is SDD; thus, it is nonsingular, implying that the matrix A is nonsingular, too. □

While simple, the proof of this result gives us a first insight in what is to become a very useful technique and an excellent tool for improving known results and producing new ones. Since the main idea was to construct a nonsingular diagonal matrix that will scale, from the right hand side, the original matrix into an SDD one, we will address this as a **scaling technique**. More precisely, this technique will be presented in the last section of this chapter, while its full potential will be explored in the consecutive ones.

1.2 Extensions of (Strictly) Diagonal Dominant Matrices

The simplicity and the beauty of the SDD matrices are obvious. But, on the other hand, the condition that defines them is quite restrictive, i.e., the corresponding class of matrices is not "sufficiently large". So, in this section, we will give several extensions of the nonsingularity result of SDD matrices. Until now, through iDD, nDD and sDD matrices, we could obtain only equalities, instead of strict inequalities in each row, which is an improvement. But, as the examples show, it a natural thing to expect to obtain nonsingularity, despite the fact that in one, or few rows, the SDD property could be made worst than being an equality. In each of the following subsections we will present an approach used to extend nonsingularity result from SDD matrices to "SDD-like" matrices. In the next section, the precise definition of the term "SDD-like" will be given.

1.2.1 Extensions by multiplication

The following result is due to Ostrowski, [41], and it represents one of the first generalizations of strict diagonal dominance. We state it here in the form as it was given in [51].

Theorem 1.2.1. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, be an arbitrary matrix. If*

$$|a_{i,i}||a_{j,j}| > r_i(A)r_j(A), \tag{1.2.1}$$

where $r_i(A)$ is given by (1.1.2), holds for every two distinct indices $i, j \in N$, then A is nonsingular.

Proof. Suppose, on the contrary, that A satisfies (1.2.1) is singular. As before, this implies that there exists a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$, or, equivalently,

$$-a_{i,i}x_i = \sum_{j \in N \setminus \{i\}} a_{i,j}x_j, \quad (i \in N). \tag{1.2.2}$$

Similar to the proof of Lévy-Desplanques theorem, since $x \neq \mathbf{0}$, there exist indices $k, \ell \in N$, such that $|x_k| \geq |x_\ell| \geq \max\{|x_i| : i \in N \setminus \{k, \ell\}\}$, where the last quantity is defined to be zero if $n = 2$, which, according to (1.2.2) and triangle inequality, implies that

$$|a_{k,k}||x_k| \leq \sum_{j \in N \setminus \{k\}} |a_{k,j}||x_j| \leq |x_\ell| r_k(A). \tag{1.2.3}$$

In the case when $|x_\ell| = 0$, all the quantities in the last relation reduce to zero, which is a clear contradiction with the fact that (1.2.1) holds for all indices. Thus, $|x_\ell| > 0$. Taking $i = \ell$, after applying again the triangle inequality, (1.2.2) becomes

$$|a_{\ell,\ell}|x_\ell| \leq \sum_{j \in N \setminus \{\ell\}} |a_{\ell,j}||x_j| \leq |x_k| r_\ell(A). \tag{1.2.4}$$

Finally, after multiplying (1.2.3) and (1.2.4), we get

$$|a_{k,k}||a_{\ell,\ell}||x_k||x_\ell| \leq r_k(A)r_\ell(A))|x_k||x_\ell|,$$

Since $|x_k||x_\ell| > 0$, the last inequality implies that $|a_{k,k}||a_{\ell,\ell}| \leq r_k(A)r_\ell(A))$, which contradicts (1.2.1). □

In the literature, these matrices are called doubly diagonally dominant matrices, or briefly doubly DD matrices. Here, in order to point out the difference between strict inequalities and non-strict ones, the matrices from the previous theorem we will call **doubly strictly diagonally dominant** matrices, or, briefly, **doubly SDD**.

Note that if A is an SDD matrix, then (1.2.1) is valid for each two different indices. Thus, every SDD matrix is doubly SDD, too. The reverse is not true in general, as Example 1.2.3 illustrates below.

On the other hand, if a matrix is not an SDD matrix, at least one row is not an SDD row. Obviously, in order to apply Theorem 1.2.1, there has to be not more than one non SDD row, and, thus, we need to check additional $n-1$ inequalities, namely combinations of non SDD row with all the others.

Similar to the notion of DD matrices, we can define doubly diagonally dominant matrices as follows.

Definition 1.2.2. An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called **doubly diagonally dominant** matrix[6] if

$$|a_{i,i}||a_{j,j}| \geq r_i(A)r_j(A) \quad \text{for all } i, j \in N, \; i \neq j, \tag{1.2.5}$$

and for at least one pair of indices $k, \ell \in N$, $k \neq \ell$,

$$|a_{k,k}||a_{\ell,\ell}| > r_k(A)r_\ell(A). \tag{1.2.6}$$

It comes as no surprise that doubly DD matrices are not, in general, nonsingular; see matrix A_7 of the following example.

Example 1.2.3. *Let*

$$A_6 = \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 1 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix},$$

$$A_7 = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}.$$

As before, the SDD rows in given matrices are shown in boldface. We notice that the nonsingular matrix A_6, $\det(A_6) = 1$, is not an SDD matrix, while it is a doubly SDD matrix. On the other hand, the singular matrix A_7 is neither SDD nor doubly SDD. But, it is a singular doubly DD matrix.

In the same fashion, in which nonsingularity for iDD matrices was proven in Theorem of Taussky, Ostrowski proved the following theorem.

[6] Or briefly **doubly DD** matrix.

Theorem 1.2.4. (Ostrowski) *Every irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is doubly diagonally dominant is nonsingular.*

Matrices defined by the Theorem 1.2.4 will simply be called **doubly irreducibly diagonally dominant**, or **doubly iDD**) matrices.

A natural question, of course, is to ask whether we can continue in this way and generate nonsingular classes of strictly triple diagonally dominant matrices, and so on. This is not so, as the matrix A_7 of the previous example shows. There, we have that $|(A_7)_{1,1}(A_7)_{2,2}(A_7)_{3,3}| = 1 > 0 = r_1(A_7)r_2(A_7)r_3(A_7)$, while $\det(A_7) = 0$. Actually, this counter-example was first given by Morris Newman. Thus, other properties must come into the play in order to ensure nonsingularity, which is the subject of the following subsection.

1.2.2 Extensions via graph theory

This direction of generalization of the results on diagonal dominance was championed by Brualdi in his paper in 1982, [7], in the field of eigenvalue localization. Developing the connection between graph theory and theory of matrices, he introduces the notion of the cycle in the graph of a matrix, which is the missing link leading to nonsingularity of matrices in the fashion of SDD and doubly SDD. Here, we will, actually, present Brualdi's result extended by the work of Varga, who defined the notion of weak cycles and formulated the corresponding nonsingularity result as it is now generally known.

Given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 1$, let $\mathbb{G}(A)$ be its graph, as was described in Subsection 1.1.1: n vertices $\{v_1, v_2, \ldots, v_n\}$, we connect with directed edges, in such a way that each nonzero entry $a_{i,j} \neq 0$ makes an edge $\overrightarrow{v_i v_j}$, which, in the case when $i = j$, is called a loop.

In the graph $\mathbb{G}(A)$, a **strong cycle** γ, of length $p \geq 2$, is a p-tuple of integers $\gamma := (i_1, i_2, \ldots, i_p)$, such that $\overrightarrow{v_{i_j}, v_{i_{j+1}}}$ is an edge, for all $j = 1, 2, \ldots, p$, where we take $i_{p+1} := i_1$.

In other words, an ordered set $\gamma := (i_1, i_2, \ldots, i_p)$ is called a strong cycle in $\mathbb{G}(A)$ if the entries $a_{i_1,i_2}, a_{i_2,i_3}, \ldots, a_{i_p,i_1}$ of the matrix A are nonzero. Obviously, the sequence of indices is what defines a cycle in a graph, while it is not of importance which one is the beginning and the end. The following example illustrates that.

Example 1.2.5. *Let*

$$A_8 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

In Figure 1.2.1, which represents the graph $\mathbb{G}(A_8)$ attributed to the given matrix, we can observe that there are only two strong cycles, and they are made of the edges $\overrightarrow{v_1, v_2}$, $\overrightarrow{v_2, v_3}$ and $\overrightarrow{v_3, v_1}$, and of the edges $\overrightarrow{v_3, v_4}$ and $\overrightarrow{v_4, v_3}$. So, first of those two cycles we can write as $\gamma_1 = (1, 2, 3)$, or $\gamma_1 = (2, 3, 1)$, or as $\gamma_1 = (3, 1, 2)$, while the other one can be written as

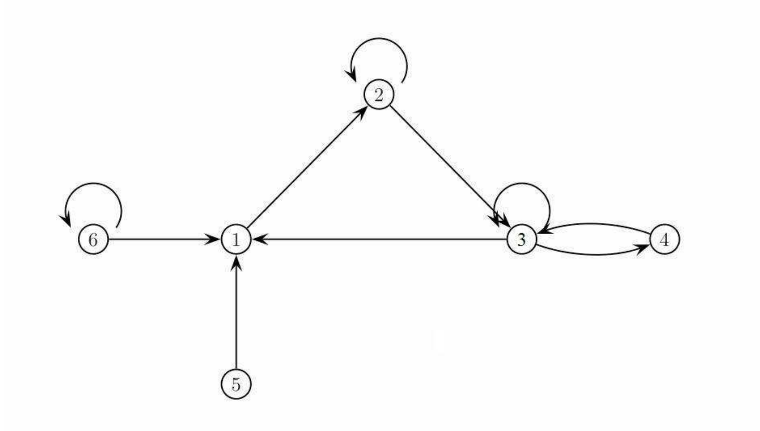

Figure 1.2.1: Graph of a matrix A_8 of the Example 1.2.5
(Graf matrice A_8 iz Primera 1.2.5)

$\gamma_1 = (3,4)$, *or* $\gamma_1 = (4,3)$. *We see that neither one of the two strong cycles of the matrix* A_8 *contains indices* 5 *and* 6. *In that case we will define the weak cycles* $\gamma_3 = (5)$ *and* $\gamma_4 = (6)$, *as described below.*

For an arbitrary index $i \in N$, such that there is no strong cycle which passes through the vertex v_i, we define, regardlessly of whether a loop of $\mathbb{G}(A)$ in the vertex v_i exists or not, a **weak cycle** $\gamma = (i)$, i.e., we define a weak cycle, regardlessly of whether $a_{i,i} \neq 0$ or $a_{i,i} = 0$. In such a way, we have accomplished that, for each index $i \in N$, at least one cycle, weak or strong, passes through the vertex v_i. We denote by $C(A)$ the **set of all cycles in the graph** $\mathbb{G}(A)$, strong ones as well as the weak ones.

Now, if we get back to the notion of irreducibility that was introduced in Subsection 1.1.1 by Definition 1.1.5, we can observe that, for an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, either A is irreducible, or it can be written as

$$P^T A P = \begin{bmatrix} A_{1,1} & A_{1,2} \\ O & A_{2,2} \end{bmatrix},$$

where $A_{1,1}$ and $A_{2,2}$ are square matrices, and P is an $n \times n$ permutation matrix. If we continue with the same reasoning, and apply it to the blocks $A_{1,1}$ and $A_{2,2}$, and further to their progenies, by consecutive permutations, we ultimately obtain a permutation matrix $\widetilde{P} \in \mathbb{R}^{n,n}$, and a positive integer m, $2 \leq m \leq n$, such that

$$\widetilde{P}^T A \widetilde{P} = \begin{bmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,m} \\ O & R_{2,2} & \cdots & R_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & R_{m,m} \end{bmatrix}, \tag{1.2.7}$$

where

$$\begin{cases} R_{i,i} \in \mathbb{C}^{p_i,p_i}, & \text{is irreducible matrix with } p_i \geq 2, \text{ or} \\ R_{i,i} = [a_{k,k}] \in \mathbb{C}^{1,1}, & \text{is } 1 \times 1 \text{ matrix for some } k \in N. \end{cases} \tag{1.2.8}$$

for every $1 \leq i \leq m$.

The permutation matrix $\widetilde{P}$ is obtained as a product of individual permutation matrices generated at each step of the process.

The form (1.2.7) is called the **normal reduced form** of the matrix A.

For a matrix A_8 of the Example 1.2.5, the normal reduced form is given by

$$\widetilde{P}^T A_8 \widetilde{P} = \left[\begin{array}{c|c|cccc} 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right], \tag{1.2.9}$$

where $\widetilde{P}$ is a permutation matrix that corresponds to permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 5 & 6 & 1 & 2 \end{pmatrix}.$$

In addition, we can note that in this case, the matrices $R_{1,1}$ and $R_{2,2}$ are both 1×1, while $R_{3,3}$ is an irreducible matrix of the size 4×4. During this transformation, we observe that the structure of the graph $\mathbb{G}(A_8)$ remained the same, while the vertices have interchanged their names. In this new notation, according to the permutation π, the cycles are given by $\gamma_1 = (3, 5, 4)$, $\gamma_2 = (5, 6)$, $\gamma_3 = (2)$ and $\gamma_4 = (1)$. Here, we can observe that the weak cycles correspond to the 1×1 blocks, while the strong cycles occur in irreducible diagonal blocs of the size at least 2.This is not only the case in this particular example. It holds in general, as a consequence of the definition of the normal reduced form, and the definition of reducibility.

Before we give a nonsingularity result that corresponds to Brualdi's theorem from 1982, as Varga's work suggests, we can point out that, according to the form given in (1.2.7), an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is nonsingular if and only if all of the matrices $R_{i,i}$, $1 \leq i \leq m$ are nonsingular. So, instead of using row sums of the matrix A, one can take only that part of the sum which lies in the corresponding diagonal block. Namely, we define the **reduced row sums**

$$\widetilde{r}_i(A) := r_j(R_{k,k}), \tag{1.2.10}$$

where the i-th row of the matrix A corresponds, in fact, the j-th row of a matrix $R_{k,k}$ in the form (1.2.7). An easy consequence of such a definition is the fact that $\widetilde{r}_i(A) = 0$ if and only if a weak cycle of the graph $\mathbb{G}(A)$ passes through the vertex v_i.

Theorem 1.2.6. (Brualdi)[7] *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, if*

$$\prod_{i\in\gamma} |a_{i,i}| > \prod_{i\in\gamma} \widetilde{r}_i(A) \tag{1.2.11}$$

[7] *The theorem presented here is, in fact, Varga's generalization of the original Brualdi's theorem from [7].*

holds for every, either strong or weak, cycle[8] $\gamma \in C(A)$, *where* $\widetilde{r}_i(A)$ *is given by (1.2.10), then the matrix* A *is nonsingular.*

Proof. As before, we start with the assumption that a matrix fulfilling the conditions of the theorem is, on the contrary, singular. From normal reduced form (1.2.7), it follows that there exists a diagonal block $R_{k,k}$ that is singular.

Let us, first, consider the case when $R_{k,k} = [a_{j,j}] \in \mathbb{C}^{1,1}$. Then, $\gamma = (j)$ is a weak cycle, and, according to (1.2.11), it follows that $|a_{j,j}| > 0$, which contradicts the fact that $R_{k,k}$ is singular.

Thus, $R_{k,k} = [\widetilde{a}_{i,j}] \in \mathbb{C}^{p_k,p_k}$ has to be an irreducible matrix of size at least 2×2. Here, with tilde we are denoting the entries of the original matrix A which have been permuted to obtain the normal reduced form. To simplify the notation, and without any loss of generality, we can now assume that $R_{k,k} = A$. Then, $\widetilde{r}_i(A) = r_i(A)$, for all $i \in N$, and, since A is irreducible, through every index in N passes at least one strong cycle from $C(A)$, implying that, according to (1.2.11), every diagonal entry of the matrix A is nonzero.

Now, since A is singular, there exists a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{C}^n$, $\mathbf{x} \neq \mathbf{0}$, such that $A\mathbf{x} = \mathbf{0}$, or, equivalently,

$$-a_{i,i}x_i = \sum_{j \in N \setminus \{i\}} a_{i,j}x_j, \quad (i \in N). \tag{1.2.12}$$

Letting $\ell_1 \in N$ be such that $|x_{\ell_1}| := \max\{|x_j| : j \in N\}$, we have, as a consequence of $\mathbf{x} \neq \mathbf{0}$, that $|x_{\ell_1}| > 0$, and, from (1.2.12) and triangle inequality, that

$$0 < |a_{\ell_1,\ell_1}||x_{\ell_1}| \leq \sum_{j \in N \setminus \{\ell_1\}} |a_{\ell_1,j}||x_j|. \tag{1.2.13}$$

Hence, there is an $\ell_2 \in N \setminus \{\ell_1\}$, such that $|a_{\ell_1,\ell_2}||x_{\ell_2}| > 0$. Choosing an ℓ_2 so that

$$|x_{\ell_2}| = \max\{|x_j| : |a_{\ell_1,j}| > 0, \text{ and } j \in N \setminus \{x_{\ell_1}\}\},$$

from (1.2.12), we have that

$$|a_{\ell_2,\ell_2}||x_{\ell_2}| \leq \sum_{j \in N \setminus \{\ell_2\}} |a_{\ell_2,j}||x_j| \leq |x_{\ell_1}| r_{\ell_2}(A). \tag{1.2.14}$$

Again, from (1.2.14) we conclude that there is $\ell_3 \in N \setminus \{\ell_2\}$, such that

$$|x_{\ell_3}| = \max\{|x_j| : |a_{\ell_2,j}| > 0, \text{ and } j \in N \setminus \{x_{\ell_2}\}\},$$

and we obtain

$$|a_{\ell_3,\ell_3}||x_{\ell_3}| \leq \sum_{j \in N \setminus \{\ell_3\}} |a_{\ell_3,j}||x_j| \leq |x_{\ell_2}| r_{\ell_3}(A). \tag{1.2.15}$$

Repeating this procedure, we obtain a sequence of indices

$$\ell_1, \ell_2, \ldots, \ell_n, \ell_{n+1} \tag{1.2.16}$$

such that, for every $i = 1, 2, \ldots, n$, $\ell_i \neq \ell_{i+1}$, $|a_{\ell_i,\ell_{i+1}}||x_{\ell_{i+1}}| > 0$ and

$$|a_{\ell_{i+1},\ell_{i+1}}||x_{\ell_{i+1}}| \leq \sum_{j \in N \setminus \{\ell_{i+1}\}} |a_{\ell_{i+1},j}||x_j| \leq |x_{\ell_i}| r_{\ell_{i+1}}(A). \tag{1.2.17}$$

[8] *In the notation* $i \in \gamma$, *the cycle* γ *is considered as the* set *of the corresponding indices.*

But, the set of indices of the matrix A is of cardinality n, so, in the sequence (1.2.16), at least one index has to repeat. Having that every two consecutive indices are distinct, there must exist a subsequence $\ell_k, \ell_{k+1} \dots, \ell_{k+p}$, where $p \geq 1$, and $\ell_{k+p+1} = \ell_k$. Multiplying inequalities (1.2.17) for $i = k, k+1, \dots, k+p$ and dividing the result by $|x_{\ell_k}||x_{\ell_{k+1}}| \cdots |x_{\ell_{k+p}}| > 0$, we get

$$|a_{\ell_k,\ell_k}||a_{\ell_{k+1},\ell_{k+1}}| \cdots |a_{\ell_{k+p},\ell_{k+p}}| \leq r_{\ell_k}(A) r_{\ell_{k+1}}(A) \cdots r_{\ell_{k+p}}(A). \tag{1.2.18}$$

But, from the construction of the sequence (1.2.16), it follows that $|a_{\ell_i,\ell_{i+1}}| \neq 0$, for all $i = k, k+1, \dots, k+p$. Thus, $\gamma = (k, k+1, \dots, k+p) \in C(A)$, and (1.2.18) contradicts (1.2.11), leading to the conclusion that the matrix A is nonsingular. □

Matrices from the previous theorem we will call **Brualdi strictly diagonally dominant**, or just **Brualdi SDD** matrices.

Example 1.2.7. *Observing the normal reduced form (1.2.9) of the matrix A_8 of the Example 1.2.5, we conclude that $C(A_8) = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$, where $\gamma_1 = (3, 5, 4)$, $\gamma_2 = (5, 6)$, $\gamma_3 = (2)$, and $\gamma_4 = (1)$. Thus, in this case we need to check four inequalities:*

$$|a_{3,3}||a_{5,5}||a_{4,4}| > \widetilde{r}_3(A)\widetilde{r}_5(A)\widetilde{r}_4(A), \; |a_{5,5}||a_{6,6}| > \widetilde{r}_5(A)\widetilde{r}_6(A), \; |a_{2,2}| > 0 \text{ and } |a_{1,1}| > 0.$$

One can ask if the new class of matrices is bigger, smaller or neither, comparing it to the class of doubly SDD matrices. While it is rather easy to conclude that every SDD matrix is Brualdi SDD, to compare doubly SDD matrices to Brualdi SDD matrices is not so obvious. Nevertheless, Brualdi SDD matrices include doubly SDD matrices, as the following result of Varga states.

Theorem 1.2.8. *Every doubly strictly diagonally dominant matrix is Brualdi strictly diagonally dominant.*

For proof, see Theorem 2.9 in [51].

Converse of the Theorem 1.2.8 is not true, in general, as shown by the following example.

Example 1.2.9. *Given an irreducible matrix*

$$A_9 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ \mathbf{0.9} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix},$$

observe that $\gamma = (1, 2, 3, 4)$,is the only cycle in $\mathbb{G}(A)$. Thus, since

$$|a_{1,1}||a_{2,2}||a_{3,3}||a_{4,4}| = 1 > 0.9 = r_1(A_9)r_2(A_9)r_3(A_9)r_4(A_9),$$

we conclude that A_9 is a Brualdi SDD matrix. But, since there are more than one non SDD rows, it's obvious that it is not a doubly SDD matrix.

As we have mentioned, to determine if a matrix is doubly SDD, it is necessary to check $\frac{n(n-1)}{2}$ inequalities. Regarding Brualdi SDD matrices, since they depend on the associated graph structure, their number can vary, from only one inequality, as in the case of matrix A_9 of Example 1.2.9, to a few of them, Example 1.2.7, or to a number that far exceeds $n-1$. The ultimate case is when a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, of the size $n \geq 2$, has all nonzero off-diagonal entries. Then, each choice of two or more indices is a strong cycle, so their number is $\sum_{k=2}^{n} \frac{n!}{k!}$. It is interesting to note that in that case, for each cycle $\gamma = (i,j)$ of the length 2, inequality (1.2.11) becomes (1.2.1), implying that such Brualdi SDD matrices are, in fact, doubly SDD. In other words, if we are considering only such matrices, Brualdi SDD becomes the same as doubly SDD.

So, in case of matrices without any zero off-diagonal entry, most of the cycles of the graph $\mathbb{G}(A)$ do not affect the condition of diagonal dominance. That, of course, raises the question if the set of cycles in the graph $C(A)$ can be reduced, while the condition that defines Brualdi SDD matrices remains unchanged. The answer is yes, and the reduced set of cycles is given in [51], Theorem 2.10.[9]

To conclude this subsection, we give an analog of the Taussky's nonsingularity result of Theorem 1.1.8, which also appears in [7].

Definition 1.2.10. An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called a **Brualdi diagonally dominant matrix**[10] if

$$\prod_{i\in\gamma} |a_{i,i}| \geq \prod_{i\in\gamma} r_i(A), \quad (\text{all } \ \gamma \in C(A)), \tag{1.2.19}$$

with strict inequality holding for at least one cycle $\gamma \in C(A)$.

Theorem 1.2.11. (Brualdi) *Every irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is Brualdi diagonally dominant is nonsingular.*

Matrices defined by the Theorem 1.2.11 we will call **Brualdi irreducibly diagonally dominant** (**Brualdi iDD**) matrices.

1.2.3 Extensions by partitions

In this subsection we start with a motivating result by Daschnic and Zusmanovich from 1970, [21], and we give its generalization by Cvetković, Kostić and Varga from 2004, which is done by the use of partitioning of the set of indices. The original technique which was used in both papers is in essence different from the ones presented up to now, and we will study it, in detail, in the fourth Section of this Chapter. Here, we will give new proofs of these results, using an approach similar to the ones that we were using throughout this section. We should also mention that extensions by partitions of the set of indices were

[9] Actually, Theorem 2.10 covers the same question in an equivalent form of eigenvalue localization, a topic covered in the next chapter.

[10] Or, briefly, **Brualdi DD matrices**.

done also by Gao and Wang in [26], and Huang in [30]. While the result of Gao and Wang is in fact a special case of the result of Cvetković, Kostić and Varga, Huang gave more general result. But, the drawback of his result lies in the complexity of the conditions to be checked, which limits the application of his theorem, especially in the fields that are covered in the later chapters of this theses.

Theorem 1.2.12. (**Dashnic-Zusmanovich**) *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be an arbitrary matrix, and let $r_i(A)$ be given by (1.1.2). If there exists an index $\ell \in N$, such that for every $j \in N \setminus \{\ell\}$*

$$|a_{\ell,\ell}| \cdot (|a_{j,j}| - r_j(A) + |a_{j,\ell}|) > r_\ell(A)|a_{j,\ell}|, \tag{1.2.20}$$

then A is nonsingular matrix.

Proof. Suppose, on the contrary, that A whose elements satisfy (1.2.20) is singular, i.e., there exists a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$. Equivalently,

$$-a_{i,i}x_i = \sum_{j \in N \setminus \{i\}} a_{i,j}x_j, \quad (i \in N). \tag{1.2.21}$$

Similar to the proof of Ostrowski's theorem on doubly SDD matrices, take $k \in N$, $k \neq \ell$, so that $|x_k| = \max\{|x_i| : i \in N \setminus \{\ell\}\}$. Then, (1.2.21) and the triangle inequality imply that

$$|a_{k,k}||x_k| \leq \sum_{j \in N \setminus \{k\}} |a_{k,j}||x_j| \leq |x_k|(r_k(A) - |a_{k,\ell}|) + |a_{k,\ell}||x_\ell|, \tag{1.2.22}$$

i.e.,

$$(|a_{k,k}| - r_k(A) + |a_{k,\ell}|)|x_k| \leq |a_{k,\ell}||x_\ell|. \tag{1.2.23}$$

Assuming that $|x_\ell| = 0$, the right hand size of the last relation reduces to zero. But, since $\mathbf{x} \neq \mathbf{0}$, then $|x_k| > 0$, implying that $|a_{k,k}| - r_k(A) + |a_{k,\ell}| < 0$, which contradicts the fact that (1.2.20) holds for each index $j \in N \setminus \{\ell\}$. Therefore, $|x_\ell| \neq 0$.

Now, taking $i = \ell$, and applying again the triangle inequality to (1.2.21), we obtain

$$|a_{\ell,\ell}||x_\ell| \leq \sum_{j \in N \setminus \{\ell\}} |a_{\ell,j}||x_j| \leq |x_k|r_\ell(A). \tag{1.2.24}$$

Finally, after multiplying (1.2.23) and (1.2.24), and dividing by $|x_k||x_\ell| > 0$, we get

$$(|a_{k,k}| - r_k(A) + |a_{k,\ell}|)|a_{\ell,\ell}| \leq |a_{k,\ell}|r_\ell(A)),$$

which contradicts (1.2.20). Hence, A is a nonsingular matrix. □

Again, it is easy to see that the matrices that satisfy the previous theorem, which we will call **Dashnic-Zusmanovich-SDD** matrics, or **DZ-SDD** matrices, include SDD matrices. Namely, if each row of a matrix is SDD, then (1.2.20) directly follows.

The following theorem is a direct generalization of the previous one.

Theorem 1.2.13. (**Cvetković-Kostić-Varga**) *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be an arbitrary matrix, and let $S \subseteq N$ be a nonempty subset of indices. If, for every two indices $i \in S$, and $j \in \overline{S} := N \setminus S$, there holds*

$$|a_{i,i}| > r_i^S(A), \quad \text{and} \tag{1.2.25}$$

$$(|a_{i,i}| - r_i^S(A)) \cdot (|a_{j,j}| - r_j^{\overline{S}}(A)) > r_i^{\overline{S}}(A) r_j^S(A), \tag{1.2.26}$$

where $r_i^S(A) := \sum_{j \in S \setminus \{i\}} |a_{i,j}|$, then A is nonsingular.

Remark 1.2.14. Before we give the proof, it is interesting to note that taking any $i \in S$ and applying (1.2.26) for all $j \in \overline{S}$, we obtain that

$$|a_{j,j}| > r_j^{\overline{S}}(A) \quad (j \in \overline{S}). \tag{1.2.27}$$

Thus, the condition (1.2.27) is a necessary one, and it is implicitly stated within (1.2.25) and (1.2.26). For the same reason, to check if the conditions for this theorem are valid, it *suffice* to check if (1.2.25) holds for *at least one* index $i \in S$, instead for all of them.

Proof. First, observe that in the case when $S = N$, condition (1.2.26) vanishes leaving the condition (1.2.25), which becomes the same as (1.1.3), implying that the matrix is SDD, and hence, nonsingular. So, we assume, for the nonempty subset of indices S, that $S \subsetneq N$, i.e., that its complement $\overline{S}$ is nonempty, too.

As before, we start by assuming, on the contrary, A that satisfies (1.2.25) and (1.2.26), and that A is singular. Then, we can take a vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \neq \mathbf{0}$ such that $A\mathbf{x} = \mathbf{0}$, or equivalently,

$$-a_{i,i}x_i = \sum_{j \in N \setminus \{i\}} a_{i,j}x_j = \sum_{j \in S \setminus \{i\}} a_{i,j}x_j + \sum_{j \in \overline{S} \setminus \{i\}} a_{i,j}x_j, \quad (i \in N). \tag{1.2.28}$$

Since S and $\overline{S}$ are nonempty, we can take $k \in S$, so that $|x_k| = \max\{|x_i| : i \in S\}$, and $\ell \in \overline{S}$, so that $|x_\ell| = \max\left\{|x_j| : j \in \overline{S}\right\}$. Calling $i = k$ in (1.2.28), and taking absolute values and applying the triangle inequality, we obtain

$$|a_{k,k}||x_k| \leq \sum_{j \in N \setminus \{k\}} |a_{k,j}||x_j| \leq |x_k| r_k^S(A) + |x_\ell| r_k^{\overline{S}}(A), \tag{1.2.29}$$

i.e.,

$$(|a_{k,k}| - r_k^S(A))|x_k| \leq r_k^{\overline{S}}(A)|x_\ell|. \tag{1.2.30}$$

Similar, we get

$$(|a_{\ell,\ell}| - r_\ell^{\overline{S}}(A))|x_\ell| \leq r_\ell^S(A)|x_k|. \tag{1.2.31}$$

Now, since $\mathbf{x} \neq \mathbf{0}$, at least one of the following is true: $|x_k| > 0$ or $|x_\ell| > 0$. Without any loss of generality, we can assume that $|x_k| > 0$. Then, according to (1.2.25), (1.2.30) implies that $|x_\ell| > 0$, too. Thus, after multiplying (1.2.30) and (1.2.31), and dividing by $|x_k||x_\ell| > 0$, we get a direct contradiction to (1.2.26). Hence, A is a nonsingular matrix. □

In the literature, matrices defined in the previous theorem are called **S-strictly diagonally dominant** matrices, or just **S-SDD** matrices, where S is the fixed set of indices for which the conditions (1.2.25) and (1.2.26) hold.

Clearly, as S is an arbitrary non-empty subset of indices, we can define a larger class of matrices by letting S vary. The result is the class of matrices known in literature as $\mathcal{S}$-SDD, [10], or Σ-SDD, [6]. Here, for a uniform notation, we will call them **Cvetković-Kostić-Varga SDD** matrices, or, briefly, **CKV-SDD** matrices. More precisely, the matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is a CKV-SDD matrix if and only if there is a non-empty set of indices S such that for every $i \in S$ and every $j \in \overline{S}$ conditions (1.2.25) and (1.2.26) hold.

Taking S to be a singleton, (1.2.26) transforms into (1.2.20), and therefore, every DZ-SDD matrix is a CKV-SDD matrix.

It is interesting to compare these classes of matrices to the previously defined ones. In fact the following theorem holds.

Theorem 1.2.15. *Every doubly strictly diagonally dominant matrix is a Dashnic-Zusmanovich strictly diagonally dominant, and every Dashnic-Zusmanovich strictly diagonally dominant matrix is Cvetković-Kostić-Varga strictly diagonally dominant one.*

That CKV-SDD matrices and Brualdi SDD matrices stand in a general position, one can find in detail in [10] and [51].

To conclude this subsection, we give Taussky's analog of the Theorem 1.2.13, originally obtained in [13]. Although they can be proved by following the same idea as in Theorem 1.1.8, the original proof in [13] was obtained by the scaling technique, which we will give in the last section of this chapter.

Definition 1.2.16. Given a nonempty subset $S \subset N$, an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called an **S-diagonally dominant** [11] matrix, if

$$|a_{i,i}| \geq r_i^S(A) \quad (\text{all} \ \ i \in S), \tag{1.2.32}$$

and

$$(|a_{i,i}| - r_i^S(A))(|a_{j,j}| - r_j^{\overline{S}}(A)) \geq r_i^{\overline{S}}(A) r_j^S(A) \quad (\text{all} \ \ i \in S, \ \ \text{and all} \ \ j \in \overline{S}), \tag{1.2.33}$$

with strict inequality holding for at least one pair of indices $i \in S$, and $j \in \overline{S}$. If a matrix is S-daigonally dominant for some nonempty subset $S \subset N$, then it is **CKV-diagonally dominant** matrix [12].

Theorem 1.2.17. (Cvetković-Kostić) *Every irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is S-diagonally dominant is nonsingular, and consequently, every irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is CKV-diagonally dominant, is nonsingular.*

As before, matrices defined by the Theorem 1.2.17 we call **S-irreducibly diagonally dominant** (**S-iDD**) matrices, and **CKV-irreducibly diagonally dominant** (**CKV-iDD**) matrices.

[11] Or, briefly, **S-DD**.
[12] Or, briefly, **CKV-DD**.

1.2.4 Extensions by column sums

It is well-known that an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is nonsingular if and only if its **transpose**, i.e.,

$$A^T := [\widetilde{a}_{i,j}] \in \mathbb{C}^{n,n}, \quad \text{where } \widetilde{a}_{i,j} := a_{j,i}, \quad \text{for all } i,j \in N,$$

is nonsingular. So, a direct corollary of the Lévy-Desplanques Theorem is the following nonsingularity result.

Theorem 1.2.18. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ be an arbitrary matrix. If*

$$|a_{i,i}| > c_i(A) := r_i(A^T) = \sum_{j \in N \setminus \{i\}} |a_{j,i}| \quad \text{for all } i \in N, \tag{1.2.34}$$

then A is nonsingular.

Such matrices are known in the literature as **column strictly diagonally dominant matrices**, while SDD matrices are sometimes called by row strictly diagonally dominant matrices.

After collecting this simple extension together with the original result on SDD matrices, one can easily see that the following proposition is true.

Proposition 1.2.19. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ be an arbitrary matrix. If*

$$|a_{i,i}| > r_i(A), \quad \text{for all } i \in N, \quad \text{or } |a_{i,i}| > c_i(A), \quad \text{for all } i \in N, \tag{1.2.35}$$

then, the matrix A is nonsingular.

While trivial, this observation opens a very interesting question. Namely, (1.2.35) is, definitely, more restrictive than the following property

$$|a_{i,i}| > r_i(A) \text{ or } |a_{i,i}| > c_i(A), \quad \text{for all } i \in N,$$

which can, also, be written as

$$|a_{i,i}| > \min\{r_i(A), c_i(A)\}, \quad \text{for all } i \in N. \tag{1.2.36}$$

In other words, with (1.2.36), we have obtained a class of matrices larger then both column SDD and row SDD classes, so, the question about their nonsingularity arises.

Example 1.2.20. *Given a matrix*

$$A_{10} = \begin{bmatrix} 2 & 6 \\ 1 & 3 \end{bmatrix},$$

observe that $r_1(A_{10}) = c_2(A_{10}) = 6$, and $r_2(A_{10}) = c_1(A_{10}) = 1$.

Then, $\min\{r_1(A_{10}), c_1(A_{10})\} = \min\{r_2(A_{10}), c_2(A_{10})\} = 1$, and, hence, given matrix fulfills (1.2.36). But, $\det(A_{10}) = 0$, so the matrix is singular.

As the example suggests, (1.2.36) is "too loose", meaning that it is not a sufficient condition for nonsingularity, while the stronger condition (1.2.35) is. So, a suitable motivation would be to find another condition, as nonrestrictive as it can be, which can be added to (1.2.36) in order to insure nonsingularity, while being weaker than (1.2.35).

Before we give an answer to this problem, we will focus on two well-known extensions of the SDD property, due to Ostrowski in 1951, [39], which make use of both row and column sums, for each diagonal entry. Their proof can be found in [51].

Theorem 1.2.21. (**Ostrowski**) *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be an arbitrary matrix, and let $r_i(A)$ and $c_i(A)$ be defined by (1.1.2) and (1.2.34), respectively. If there exists a parameter $\alpha \in [0,1]$, such that*

$$|a_{i,i}| > \alpha r_i(A) + (1-\alpha)c_i(A) \ \ \textit{for all} \ \ i \in N, \tag{1.2.37}$$

then A is nonsingular.

Theorem 1.2.22. (**Ostrowski**) *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be an arbitrary matrix, and let $r_i(A)$ and $c_i(A)$ be defined by (1.1.2) and (1.2.34), respectively. If there exists a parameter $\alpha \in [0,1]$, such that*

$$|a_{i,i}| > (r_i(A))^{\alpha}(c_i(A))^{1-\alpha} \ \ \textit{for all} \ \ i \in N, \tag{1.2.38}$$

then, A is nonsingular.

Matrices that fulfill the conditions of the previous two theorem are known in the literature as **Ostrowski matrices**. In particular, they are known, respectively, as α_1-**matrices** and α_2-**matrices**, [10] and [11]. Here, we will call them α_1-**strictly diagonally dominant matrices**, or just α_1-**SDD** matrices, and α_2-**strictly diagonally dominant matrices**, or, briefly, α_2-**SDD** matrices.

For different values of the parameter α, in both cases, we obtain different nonsingularity conditions. In particular, taking $\alpha = 1$, conditions (1.2.37) and (1.2.38) become the same as (1.1.3), while, taking $\alpha = 0$, they transform in (1.2.34). Thus, α_1- and α_2-SDD matrices include, both, row and column SDD matrices.

As we have seen, the parameter α gave us an opportunity to enlarge the classes of nonsingular matrices, which is good. But, on the other hand, when we really want to apply these results on a particular matrix, we encounter the problem of finding a suitable parameter, for which the matrix fulfills nonsingularity conditions (1.2.37) and/or (1.2.38). Since it is not, in general, a trivial thing to do, it seems to be an interesting open problem.

There are two closely related approaches, developed in [11] and [6], which address this problem. Here we present them in the form of nonsingularity theorems, while their equivalent forms will be used in later chapters that concern different applications.

Actually, the nonsingularity of α_1-SDD matrices follows directly from nonsingularity of α_2-SDD matrices, by the generalized arithmetic-geometric mean inequality:

$$\alpha a + (1-\alpha)b \geq a^{\alpha}b^{1-\alpha}, \tag{1.2.39}$$

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

Nevertheless, we will consider both of these classes, since they can play separate roles in different applications. For more detail, see [11] and [12].

Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, observe that, in the case when $r_i(A) = c_i(A)$, (1.2.37) and (1.2.38) become $|a_{i,i}| > r_i(A) = c_i(A)$, independently of the value of α. In general, there are two more possibilities: $r_i(A) > c_i(A)$ or $r_i(A) < c_i(A)$. Accordingly, we partition the set of indices N into three subsets:

$$\begin{aligned} \mathcal{R}(A) &:= \{i \in N : \quad r_i(A) > c_i(A)\}, \\ \mathcal{C}(A) &:= \{i \in N : \quad c_i(A) > r_i(A)\}, \\ \mathcal{E}(A) &:= \{i \in N : \quad r_i(A) = c_i(A)\}, \end{aligned} \tag{1.2.40}$$

and for each $i \in N \setminus \mathcal{E}(A)$, we define the quantity

$$\phi_i^{(1)}(A) = \frac{|a_{i,i}| - c_i(A)}{r_i(A) - c_i(A)} \in \mathbb{R}, \tag{1.2.41}$$

which we shall use to obtain the set of feasible values of a parameter α.

Calling

$$U^{(1)}(A) = (-\infty, \min_{i \in \mathcal{R}(A)} \phi_i^{(1)}(A)) \quad \bigcap \quad (\max_{i \in \mathcal{C}(A)} \phi_i^{(1)}(A), +\infty), \tag{1.2.42}$$

where, by convention, we take

$$\min_{i \in \mathcal{R}(A)} \phi_i^{(1)}(A) = +\infty \quad \text{if } \mathcal{R}(A) := \emptyset, \quad \text{and} \quad \max_{i \in \mathcal{C}(A)} \phi_i^{(1)}(A) := -\infty \quad \text{if } \mathcal{C}(A) = \emptyset,$$

we can give the following characterization of the class of α_1-SDD matrices.

Theorem 1.2.23. *Given an arbitrary $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, A is an α_1-SDD matrix if and only if the following two conditions hold:*

(i) $U^{(1)}(A) \cap [0, 1] \neq \emptyset$,

(ii) $|a_{i,i}| > r_i(A)$, *for all* $i \in \mathcal{E}(A)$.

Proof. First, let us assume that A is an α_1-SDD matrix. Consider $i \in \mathcal{R}(A)$. From equation (1.2.41) we have

$$|a_{i,i}| = \phi_i^{(1)}(A)(r_i(A) - c_i(A)) + c_i(A), \tag{1.2.43}$$

where $r_i(A) - c_i(A) > 0$. Since A is an α_1-SDD matrix, there exists $\alpha \in [0, 1]$ such that

$$|a_{i,i}| > \alpha(r_i(A) - c_i(A)) + c_i(A), \quad \text{for all } i \in N. \tag{1.2.44}$$

Therefore, from (1.2.43) and (1.2.44), we conclude that $\phi_i^{(1)}(A) > \alpha$ for all $i \in \mathcal{R}(A)$, and thus we have

$$\min_{i \in \mathcal{R}(A)} \phi_i^{(1)}(A) > \alpha. \tag{1.2.45}$$

In an analogous way, it is easy to show that

$$\alpha > \max_{i\in\mathcal{C}(A)} \phi_i^{(1)}(A). \tag{1.2.46}$$

Note that (1.2.45) and (1.2.46) still hold if $\mathcal{R}(A)$ or $\mathcal{C}(A)$ are empty sets. Hence,

$$\max_{i\in\mathcal{R}(A)} \phi_i^{(1)}(A) < \alpha < \min_{i\in\mathcal{C}(A)} \phi_i^{(1)}(A). \tag{1.2.47}$$

By (1.2.42), we have

$$U^1(A) = (\max_{i\in\mathcal{C}(A)} \phi_i^{(1)}(A),\ \min_{i\in\mathcal{R}(A)} \phi_i^{(1)}(A)), \tag{1.2.48}$$

and, therefore, we conclude that $\alpha \in U^{(1)}(A) \cap [\,0,1\,]$.

If $i \in \mathcal{E}(A)$, then $r_i(A) = c_i(A)$, and, since A is an α_1-SDD matrix, the condition (ii) follows.

Conversely, assume that the conditions (i) and (ii) hold. From expression (1.2.42), we have $U^{(1)}(A) = (\max_{i\in\mathcal{C}(A)} \phi_i^{(1)}(A),\ \min_{i\in\mathcal{R}(A)} \phi_i^{(1)}(A))$.

Now, we show that A is an α_1-SDD matrix. More precisely, let us prove that (1.2.37) holds for each $i \in N$, and for some $\alpha \in U^{(1)}(A) \cap [\,0,1\,]$, which is not empty by the first condition (i). If $i \in \mathcal{C}(A)$, from equation (1.2.41), we have

$$|a_{i,i}| = \phi_i^{(1)}(A)(r_i(A) - c_i(A)) + c_i(A).$$

Since $\alpha < \min_{i\in\mathcal{R}(A)} \phi_i^{(1)}(A)$, we obtain

$$|a_{i,i}| > \alpha(r_i(A) - c_i(A)) + c_i(A). \tag{1.2.49}$$

Therefore, the expression (1.2.49) holds for all $\alpha < \mathop{min}_{i\in\mathcal{R}(A)} \phi_i^{(1)}(A)$. In the same way, if $i \in \mathcal{C}(A)$, we have

$$|a_{i,i}| > \alpha(r_i(A) - c_i(A^T)) + c_i(A^T), \tag{1.2.50}$$

for all $\alpha > \phi_i^{(1)}(A)$, so, the expression (1.2.50) holds for all $\alpha > \max_{i\in\mathcal{C}(A)} \phi_i^{(1)}(A)$. If $i \in \mathcal{C}(A)$, we have, by the condition (ii), that $|a_{i,i}| > c_i(A)$. So, A is an α_1-matrix for all $\alpha \in U^{(1)}(A) \cap [0,1]$, which is nonempty by the condition (i). □

Note that, in general, the set $U^{(1)}(A)$ and the interval $[0,1]$ need not intersect. The following example shows that such a matrix, in general, may have a determinant equal to zero, and, hence, is nonsingular.

Example 1.2.24. *Given the matrix*

$$A_{11} = \begin{bmatrix} \mathbf{4} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 1 & 2 & 0 & 1 \\ 1 & 2 & 1 & 2 \\ 2 & 2 & 0 & 2 \end{bmatrix}.$$

calculate $U^{(1)}(A_{11}) = [3, +\infty)$ *which has no mutual points with the interval* $[0,1]$*, while* $\det(A_{11}) = 0$.

To characterize α_2-SDD matrices, we can use a similar approach.

Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, for each $i \in N \setminus \mathcal{E}(A)$, we define the quantity

$$\phi_i^{(2)}(A) = \frac{\log |a_{i,i}| - \log c_i(A)}{\log r_i(A) - \log c_i(A)} \quad \in \mathbb{R}, \tag{1.2.51}$$

and the set

$$U^{(2)}(A) = (-\infty, \min_{i \in \mathcal{R}(A)} \phi_i^{(2)}(A)) \quad \bigcap \quad (\max_{i \in \mathcal{C}(A)} \phi_i^{(2)}(A), +\infty), \tag{1.2.52}$$

where, by convention,

$$\mathit{min}_{i \in \mathcal{R}(A)} \phi_i^{(2)}(A) = +\infty \quad \text{if } \mathcal{R}(A) = \emptyset \quad \text{and} \quad \mathit{max}_{i \in \mathcal{C}(A)} \phi_i^{(2)}(A) = -\infty \quad \text{if } \mathcal{C}(A) = \emptyset.$$

The proof of the following result is analogous to that of Theorem 1.2.23, the only difference is that we are working with $U^{(2)}(A)$ and $\phi_i^{(2)}(A)$, instead of $U^{(1)}(A)$ and $\phi_i^{(1)}(A)$, respectively.

Theorem 1.2.25. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, A is an α_2-SDD matrix if and only if the two following conditions hold:*

(i) $U^{(2)}(A) \cap [0, 1] \neq \emptyset$,

(ii) $|a_{i,i}| > r_i(A)$, *for all* $i \in \mathcal{E}(A)$.

As we have mentioned, the class of α_1-SDD matrices is a subclass of α_2-SDD matrices. Thus, for any α_1-SDD matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, clearly, $U^{(1)}(A) \cap U^{(2)}(A) = U^{(1)}(A)$.

Apart from this characterization of α_1-SDD matrices, that gave us the values for the parameter α for which nonsingularity is assured, we will prove another one, due to Cvetković, Bru, Kostić and Pedroche in 2009, [11], that is more suitable for the applications in eigenvalue localization theory.

Theorem 1.2.26. *A matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, is an α_1-SDD matrix if and only if the following two conditions hold:*

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

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

Proof. First, let us assume that A is an α_1-matrix. Then, there exists $\alpha \in [0, 1]$ such that

$$|a_{i,i}| > \alpha(r_i(A) - c_i(A)) + c_i(A), \quad \text{for all } i \in N. \tag{1.2.53}$$

Therefore, for every $i \in \mathcal{R}(A)$, we conclude that $\frac{|a_{i,i}| - c_i(A)}{r_i(A) - c_i(A)} > \alpha$, and, for every $j \in \mathcal{C}(A)$, $\frac{c_j(A) - |a_{j,j}|}{c_j(A) - r_j(A)} < \alpha$. Thus, (ii) obviously holds. Condition (i) follows directly from (1.2.53) and the fact that $\alpha \in [0, 1]$.

Conversely, assume that the conditions (i) and (ii) hold. For every index $i \in \mathcal{E}(A)$, condition (i) directly implies (1.2.53), so it remains to prove that (1.2.53) holds for indices from the set $\mathcal{R}(A) \cup \mathcal{C}(A)$.

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

$$\frac{|a_{i,i}| - c_i(A)}{r_i(A) - c_i(A)} > 0. \tag{1.2.54}$$

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

$$\frac{c_j(A) - |a_{j,j}|}{c_j(A) - r_j(A)} < 1. \tag{1.2.55}$$

Now, gathering conditions (ii), (1.2.54) and (1.2.55), we have that there exists a parameter α such that, for every $i \in \mathcal{R}(A)$ and every $j \in \mathcal{C}(A)$,

$$\max\{0, \frac{c_j(A) - |a_{j,j}|}{c_j(A) - r_j(A)}\} < \alpha < \min\{\frac{|a_{i,i}| - c_i(A)}{r_i(A) - c_i(A)}, 1\}.$$

Starting from the left inequality, we obtain that $|a_{j,j}| > \alpha(r_j(A) - c_j(A)) + c_j(A)$ for every $j \in \mathcal{C}(A)$, while from the right one we get the same for indices $i \in \mathcal{R}(A)$. Thus, (1.2.53) holds for the chosen parameter $\alpha \in [0,1]$ and every index $i \in \mathcal{R}(A) \cup \mathcal{C}(A)$, which completes the proof. □

Calling

$$\mathcal{R}^*(A) := \mathcal{R}(A) \setminus \{i : c_i(A) = 0\}, \tag{1.2.56}$$

and

$$\mathcal{C}^*(A) := \mathcal{C}(A) \setminus \{i : r_i(A) = 0\}, \tag{1.2.57}$$

we prove the similar characterization for α_2 matrices.

Theorem 1.2.27. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, A is an α_2-SDD matrix if and only if the following two conditions hold*

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

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

Proof. First, we assume that A is an α_2-SDD matrix, i.e., that there exists $\alpha \in [0,1]$ such that, for each index $i \in N$,

$$|a_{i,i}| > (r_i)^\alpha (c_i)^{1-\alpha}. \tag{1.2.58}$$

Now, since $(r_i(A))^\alpha (c_i(A))^{1-\alpha} \geq \min\{r_i(A), c_i(A)\}$ is true, for all $i \in N$, and all $0 \leq \alpha \leq 1$, we have that the condition (i) holds.

Consider an arbitrary $i \in \mathcal{R}^*(A)$. Then, the condition (1.2.58) can be written as

$$\frac{|a_{i,i}|}{c_i(A)} > (\frac{r_i(A)}{c_i(A)})^\alpha.$$

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

$$\log_{\frac{r_i}{c_i}} \frac{|a_{i,i}|}{c_i(A)} > \alpha. \tag{1.2.59}$$

Similar, for an arbitrary index $j \in \mathcal{C}^*(A)$, we obtain that

$$\log_{\frac{c_j(A)}{r_j(A)}} \frac{c_j(A)}{|a_{j,j}|} < \alpha, \tag{1.2.60}$$

which, together with (1.2.59), implies the condition (ii).

Conversely, let us assume that A satisfies (i) and (ii).
For an arbitrary index $i \in \mathcal{E}(A)$, (i) directly implies (1.2.58). For $i \in \mathcal{R}(A)$, for which $c_i(A) = 0$, and for $j \in \mathcal{C}(A)$, such that $r_j(A) = 0$, (1.2.58) follows immediately. Thus, it remains to show that (1.2.58) holds for indices from the set $\mathcal{R}^*(A)$ and the set $\mathcal{C}^*(A)$.

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

$$\log_{\frac{r_i(A)}{c_i(A)}} \frac{|a_{i,i}|}{c_i(A)} > 0.$$

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

$$\max\{0, \log_{\frac{c_j(A)}{r_j(A)}} \frac{c_j(A)}{|a_{j,j}|}\} < \alpha < \min\{\log_{\frac{r_i(A)}{c_i(A)}} \frac{|a_{i,i}|}{c_i(A)}, 1\}.$$

Starting from the right inequality, for every $i \in \mathcal{R}^*(A)$ we have that

$$\frac{|a_{i,i}|}{c_i(A)} > \left(\frac{r_i(A)}{c_i(A)}\right)^\alpha,$$

implying that (1.2.59) holds. In the same way, from the left inequality, we obtain that (1.2.59) is true for every index from the set $\mathcal{C}^*(A)$. Since $\alpha \in [0, 1]$, this concludes the proof. $\square$

1.3 Generalized Diagonally Dominant Matrices

The term generalized diagonal dominance dates back to early seventies, when the convergence theory of iterative methods was a highly attractive set of research. In the work of James and Riha from 1974, [31], this term was used in a sense that is present also nowadays. They defined a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ to be generalized diagonal dominant if there exists an entrywise positive vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$, such that

$$|a_{i,i}|x_i > \sum_{j \in N \setminus \{i\}} |a_{i,j}|x_j \quad (i \in N). \tag{1.3.1}$$

This notion, of course, generalizes the notion of SDD matrices, which are a special case of (1.3.1) for $\mathbf{x} = [1, 1, \ldots, 1]^T$. But, this idea dates back much earlier than 1974. Actually, it can be found implicitly even in the famous work of Geršgorin, published in 1931, on localization of eigenvalues, [25], as we will see in the next chapter. The basic idea is the following: starting with an SDD matrix, which by the Lev́y-Desplanques theorem is nonsingular, and multiplying it, on the right hand side, by a nonsingular diagonal matrix, its product remains nonsingular, while the SDD row inequalities can be rewritten as in (1.3.1).

Although it may seem that this is a fairly easy and simple observation, it showed up as a great tool in the field, nowadays known as M-matrix theory. The name M-matrices originates form the work of Ostrowski in 1937, who, starting from the already mentioned result of Minkowski on nonsingularity, referred to a class of matrices, that have all principal minors positive, as Minkowski-matrices, or **M-matrices**. Since then, more than seventy different equivalent definitions of M-matrix were discovered, and many famous mathematicians have given their contributions in this direction. Thus, connections with the Perron-Frobenius theory of nonnegative matrices, positive definiteness, positive stability and diagonal dominance are just some of them. Here, we will use one of the basic definitions of M-Matrices, condition (N_{38}) of Theorem 6.2.3 in [3].

Definition 1.3.1. A real matrix $A \in \mathbb{R}^{n,n}$ is called **nonsingular M-matrix** if all of the following conditions hold:

1. $a_{i,i} > 0$, for all $i \in N$,
2. $a_{i,j} \leq 0$, for all $i, j \in N$, $i \neq j$,
3. A is nonsingular, i.e., A^{-1} exists, and
4. A is inverse nonnegative, i.e., $A^{-1} \geq O$.

Actually, this class of matrices came directly from economic models in the form of the well-known Hawkins-Simon condition. So, from the early beginnings, up to now, this theory has been doubly motivated and conducted. On one hand, mathematicians developed a wide range of applications in establishing bounds on the eigenvalues of nonnegative matrices, establishing convergence criteria for iterative methods for solving large sparse systems of linear equations, localizing eigenvalues, while on the other hand, economy researchers have studied gross substitutability, stability of general equilibrium and Leontief's input/output analysis[13] of economic systems. An excellent survey on this subject can be found in the famous book of Berman and Plemons on nonnegative matrices, [3]. If one goes thoroughly through contemporary research in mathematics, engineering, robotics, ecology, pharmaceutical modeling, economics, and many others, one can find all sorts of applications of the M-matrix theory.

Among many other aspects of M-matrix theory, diagonal dominance has always been strongly present. Since a lot of progress has been made in this subject in recent years, a need for a systematic approach in applying different types of diagonal dominance, especially by the researchers who are not mainly mathematicians, is somewhat evident. Thus, the main motivation of this dissertation is to give clear and useful insight to contemporary facts on diagonal dominance and their different applications.

Until now, we have given some results, classical ones and some new ones, that generalize the basic concept, given in first Section of this Chapter. In this Section we continue, and give the unifying framework through the next definition and the theorems that will follow.

Definition 1.3.2. Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, if there exists an (entrywise) positive vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$, such that AX is strictly diagonally dominant matrix, where $X := \mathrm{diag}(x_1, x_2, \ldots, x_n)$, then, the matrix A is **generalized diagonally dominant**, or, briefly, a **GDD matrix**.

How this definition is related to the above mentioned theory of matrices, for the first time was published in famous paper of Fiedler and Pták, from 1962, [23]. It was proved that a matrix that fulfills conditions (1.) and (2.) of the Definition 1.3.1 is an M-matrix if and only if it is generalized diagonally dominant. Thus, a natural extension to the complex case case followed.

Definition 1.3.3. Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, its **comparison matrix** $\langle A \rangle := [m_{i,j}] \in \mathbb{R}^{n,n}$ is defined by

$$m_{i,j} := \begin{cases} |a_{i,i}|, & i = j, \\ -|a_{i,j}|, & \text{otherwise.} \end{cases} \tag{1.3.2}$$

Definition 1.3.4. A matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is called a **nonsingular H-matrix** if its comparison matrix is a nonsingular M-matrix, i.e. if $\langle A \rangle$ is nonsingular, and $\langle A \rangle^{-1} \geq O$.

Thus, from the result of Fiedler and Pták in 1962, [23], we have one of the key theorems in the theory of H-matrices.

Theorem 1.3.5. (Fiedler-Pták) *An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is a nonsingular H-matrix if and only if it is a generalized diagonally dominant matrix, i.e., there exists a positive diagonal matrix X, such that AX is a strictly diagonally dominant matrix.*

[13] Actually the origin of the Leontief's input/output analysis dates back in 1936, and the term, the Hawkins-Simon condition, became later widely used in the economy.

Therefore, the classes of nonsingular H-matrices and GDD matrices are the same, and we will denote them by $\mathbb{H}$.

We will also use another characterization of nonsingular H-matrices that can be found, in a slightly different form, in [3], as conditions (L_{32}).

Theorem 1.3.6. (Beauwens-Neumann) *An arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is a nonsingular H-matrix if and only if there exists a positive diagonal matrix X, such that AX is a semistrictly diagonally dominant matrix, i.e., an irreducible diagonally dominant matrix.*

Here it is interesting to note that the letter H in H-matrices (that are generalizations of M-matrices, where matrix entries can be complex numbers) comes from the name of Jacques Hadamard, and that in his book [27], we have the nonsingularity for the complex SDD matrices.

As we have seen, $\mathbb{H}$ includes SDD matrices, and the first obvious question is whether all the other classes of nonsingular matrices, that we have presented up to now: iDD, doubly SDD, doubly iDD, Brualdi SDD, Brualdi iDD, DZ-SDD, S-SDD, S-iDD and CKV-iDD, are included in $\mathbb{H}$. In order to answer this question, let us first find a unified frame for all these classes of matrices.

Definition 1.3.7. Let $\mathbb{K}$ be a nonempty class of square matrices of an arbitrary size. If $\mathbb{K}$ is such that:

- for any $A \in \mathbb{K}$, diagonal entries of A are nonzero,
- for any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $A \in \mathbb{K}$ if and only if $|A| \in \mathbb{K}$, where $|A| := [|a_{i,j}|]$,
- for any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $A \in \mathbb{K}$ implies that, for every B such that $\langle B \rangle \geq \langle A \rangle$, $B \in \mathbb{K}$,

then we say that $\mathbb{K}$ is a **diagonally dominant-type**, or briefly **DD-type**, class of matrices.

It is not so difficult to see that all the classes of matrices mentioned above are DD-type classes. An addition, a useful property that the DD-type classes possesses, is given in the following proposition.

Proposition 1.3.8. *If $\mathbb{K}$ is a diagonally dominant-type class of matrices, then, for any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $A \in \mathbb{K}$ implies that $|A| + D \in \mathbb{K}$, for every nonnegative diagonal matrix D.*

Proof. Given any $A \in \mathbb{K}$ and an arbitrary nonnegative diagonal matrix D, then $|A| \in \mathbb{K}$, and, obviously $\langle |A| + D \rangle = \langle A \rangle + D \geq \langle A \rangle$. Thus, by definition, $|A| + D \in \mathbb{K}$. □

Now, we prove the main theorem of this section.

Theorem 1.3.9. *If a diagonally dominant-type class of matrices $\mathbb{K}$ is a class of nonsingular matrices, then it is a subclass of nonsingular H-matrices, i.e., $\mathbb{K} \subseteq \mathbb{H}$.*

Proof. Take an arbitrary $A \in \mathbb{K}$. Since $|\langle A\rangle| = |A| \in \mathbb{K}$, we have that $\langle A\rangle \in \mathbb{K}$, hence, $\langle A\rangle$ is nonsingular. We need to prove that $\langle A\rangle^{-1}$ is nonnegative. Take a splitting of $\langle A\rangle = D_A - B_A$, where $D_A := \mathrm{diag}(\langle A\rangle) = \mathrm{diag}(|a_{1,1}|, |a_{2,2}|, \ldots, |a_{n,n}|)$. Obviously, D_A is a diagonal matrix with positive diagonal entries, so, we can write $\langle A\rangle = D_A(I_n - D_A^{-1}B_A)$, which implies that $I_n - D_A^{-1}B_A$ is nonsingular, and $\langle A\rangle^{-1} = (I_n - D_A^{-1}B_A)^{-1}D_A^{-1}$.

Let us show that $\rho(D_A^{-1}B_A) := \max\left\{|\lambda| : \lambda \in \sigma(D_A^{-1}B_A)\right\} < 1$. Assume, on the contrary, that there exists $\lambda \in \sigma(D_A^{-1}B_A)$, such that $|\lambda| \geq 1$. Then, $\lambda I_n - D_A^{-1}B_A = D_A^{-1}(\lambda D_A - B_A)$ is singular. But, since $|\lambda| \geq 1$, we can write $|\lambda D_A - B_A| = |\lambda| D_A + B_A = D_A + B_A + (|\lambda| - 1)D_A = |A| + D$, where $D := (|\lambda| - 1)D_A$ is nonnegative diagonal matrix. Hence, $\lambda D_A - B_A \in \mathbb{K}$, and, therefore, nonsingular, which is an obvious contradiction.

Now, since, $\rho(D_A^{-1}B_A) < 1$, geometric series $\sum_{k=0}^{\infty}(D_A^{-1}B_A)^k$ converges to $(I_n - D_A^{-1}B_A)^{-1}$. Having that $D_A^{-1}B_A$ is nonnegative, the limit of the series is nonnegative, which completes the proof. □

Now, applying the previous criteria to the classes of nonsingular matrices from two previous sections, we easily obtain the following result.

Theorem 1.3.10. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, if A is either*

- *irreducible diagonally dominant, given by Theorem 1.1.8, or*
- *doubly strictly diagonally dominant, given by Theorem 1.2.1, or*
- *doubly irreducibly diagonally dominant, given by Theorem 1.2.4, or*
- *Brualdi strictly diagonally dominant, given by Theorem 1.2.6, or*
- *Brualdi irreducibly diagonally dominant, given by Theorem 1.2.11, or*
- *α_1-strictly diagonally dominant, given by Theorem 1.2.21, or*
- *α_2-strictly diagonally dominant, given by Theorem 1.2.22, or*
- *Dashnic-Zusmanovich strictly diagonally dominant, given by Theorem 1.2.12, or*
- *S-strictly diagonally dominant, given by Theorem 1.2.13, or*
- *S-irreducibly diagonally dominant, given by Theorem 1.2.17, or*
- *Cvetković-Kostić-Varga strictly diagonally dominant, given by Theorem 1.2.13,*
- *Cvetković-Kostić-Varga irreducibly diagonally dominant, given by Theorem 1.2.17,*

then A is a nonsingular H-matrix, or equivalently, a generalized diagonally dominant matrix.

1.4 Scaling Approach

As we have seen in the previous section, GDD matrices are, actually, obtained from SDD matrices, by multiplication on the right hand side with an arbitrary positive diagonal matrix. Since each entry of the diagonal matrix X multiplies the corresponding column of the A in the product AX, we will name this operation **scaling**. Also, we will use the terms **scaled matrix** and **scaling matrix** referring, respectively, to such a product, and to the positive diagonal matrix being used.

In order to investigate this approach thoroughly, we introduce some additional notation.

First, by $\mathbb{D}$, we denote the set of all positive diagonal matrices, i.e.,

$$\mathbb{D} := \{X = \mathrm{diag}(x_1, x_2, \ldots, x_n) \in \mathbb{R}^{n,n} : \; x_i > 0, \; i \in N, \;\; \text{and } n \in \mathbb{N}\}. \tag{1.4.1}$$

Then, for an arbitrary subclass of GDD matrices $\mathbb{K} \subseteq \mathbb{H}$, with $\mathbb{X}^{\mathbb{K}}$ we denote the family of all diagonal matrices that scale matrices from $\mathbb{K}$ to SDD ones, i.e.,

$$\mathbb{X}^{\mathbb{K}} := \{X \in \mathbb{D} : AX \text{ is SDD}, \text{ for some } A \in \mathbb{K}\}. \tag{1.4.2}$$

On the other hand, we can start with an arbitrary nonempty family of positive diagonal matrices $X \in \mathbb{D}$, and define $\mathbb{K}^{\mathbb{X}}$ as the class of all matrices that are scaled by a matrix from $\mathbb{X}$ into SDD one, i.e.,

$$\mathbb{K}^{\mathbb{X}} := \{A \in \mathbb{H} : AX \text{ is SDD}, \text{ for some } X \in \mathbb{X}\}. \tag{1.4.3}$$

It is an easy thing to see that $\mathbb{X} \subseteq \mathbb{X}^{\mathbb{K}^{\mathbb{X}}}$ and $\mathbb{K} \subseteq \mathbb{K}^{\mathbb{X}^{\mathbb{K}}}$. That equality doesn't hold in general, is obvious, since $\mathbb{K}^{\mathbb{X}^{\{I\}}} = \mathbb{H}$, and $\{I\} \subsetneq \mathbb{X}^{\mathbb{K}^{\{I\}}}$.

Now, having in mind the result of Theorem 1.3.10, one can be motivated to find suitable scaling matrices for each of the mentioned classes. In the literature, results of Theorem 1.2.12 and Theorem 1.2.13 on DZ-SDD and S-SDD matrices were proved in [21] and [19], respectively, using this scaling approach, their explicit form was constructed. Since this technique will play a significant role in some of the applications in later chapters, in the reminder of this section, we give the original proofs.

Proof of the Theorem 1.2.12:

First, we recall that a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, that fulfills the condition (1.2.20) of Theorem 1.2.12, has all diagonal entries nonzero. Therefore, for a suitable index $\ell \in N$, and every $j \in N \setminus \{\ell\}$, we can write inequality (1.2.20) as

- $\dfrac{r_\ell(A)}{|a_{\ell,\ell}|} < \dfrac{|a_{j,j}| - r_j(A) + |a_{j,\ell}|}{|a_{j,\ell}|}$, when $a_{j,i} \neq 0$, and
- $|a_{j,j}| > r_j(A)$, when $a_{j,i} = 0$.

Next, calling

$$\alpha_\ell(A) := \frac{r_\ell(A)}{|a_{\ell,\ell}|}, \tag{1.4.4}$$

and

$$\beta_\ell(A) := \min\left\{\frac{|a_{j,j}| - r_j(A) + |a_{j,\ell}|}{|a_{j,\ell}|} \; : \; j \in N \setminus \{\ell\}\,, a_{j,\ell} \neq 0\right\}, \tag{1.4.5}$$

where, by the convention, $\beta_\ell(A) := +\infty$ in the case when for all $j \in N \setminus \{\ell\}$ $a_{j,\ell} = 0$, we conclude that an arbitrary matrix is a DZ-SDD matrix if and only if there exist an index $\ell \in N$ such that $\alpha_\ell(A) < \beta_\ell(A)$, or, in another words, the interval $(\alpha_\ell(A), \beta_\ell(A))$ is nonempty. Now, take $\gamma \in (\alpha_\ell(A), \beta_\ell(A))$.

Then, from $\alpha_\ell(A) < \gamma$ it follows that

$$\gamma|a_{\ell,\ell}| > r_\ell(A), \tag{1.4.6}$$

and from $\gamma < \beta_\ell(A)$,

$$|a_{j,j}| > r_j(A) - |a_{j,\ell}| + \gamma|a_{j,\ell}| = \sum_{k\in N\setminus\{\ell,j\}} |a_{j,k}| + \gamma|a_{j,\ell}|, \;\text{ for all } j \in N \setminus \{\ell\}\,. \tag{1.4.7}$$

In another words, if, for a chosen parameter γ, we construct the diagonal matrix $X = \operatorname{diag}(x_1, x_2, \ldots, x_n)$, so that $x_\ell = \gamma$, and $x_j = 1$, for all $j \in N \setminus \{\ell\}$, then (1.4.6) and (1.4.7) imply that AX is an SDD matrix, and, hence, nonsingular. Using definition 1.3.2, we obtain that it is a GDD matrix. □

Proof of the Theorem 1.2.13:

Since the case $S = N$ is obvious, we assume that S and $\overline{S}$ are nonempty sets of indices.

We construct a diagonal matrix $X = \operatorname{diag}(x_1, x_2, \ldots, x_n)$, with $x_i > 0$, for all $i \in N$, that will scale A into an SDD matrix. For the entries of the matrix X, we choose:

$$x_i = \begin{cases} \gamma > 0, & \text{for } i \in S, \\ 1, & \text{for } i \in \overline{S}, \end{cases} \tag{1.4.8}$$

where $\gamma > 0$ is arbitrary.

Then, entries of the matrix $AX = [\widetilde{A}_{i,j}] \in \mathbb{C}^{n,n}$ are given by

$$a_{i,j} = \begin{cases} \gamma a_{i,j}, & \text{for } j \in S, \\ a_{i,j}, & \text{for } j \in \overline{S}, \end{cases} \tag{1.4.9}$$

and, hence, its the row sums are $r_\ell(AX) = r_\ell^S(AX) + r_\ell^{\overline{S}}(AX) = \gamma r_\ell^S(A) + r_\ell^{\overline{S}}(A)$, for all $\ell \in N$. Therefore, AX is an SDD matrix if and only if

$$\begin{cases} \gamma|a_{i,i}| > \gamma r_i^S(A) + r_i^{\overline{S}}(A), & \text{for all } i \in S, \text{ and} \\ |a_{j,j}| > \gamma r_j^S(A) + r_j^{\overline{S}}(A), & \text{for all } j \in \overline{S}, \end{cases}$$

or, equivalently,

$$\begin{cases} \gamma(|a_{i,i}| - r_i^S(A)) > r_i^{\overline{S}}(A), & \text{for all } i \in S, \text{ and} \\ |a_{j,j}| - r_j^{\overline{S}}(A) > \gamma r_j^S(A), & \text{for all } j \in \overline{S}. \end{cases}$$

But, having (1.2.25) and (1.2.27), this implies that AX is an SDD matrix if and only if

$$\frac{r_i^{\overline{S}}(A)}{(|a_{i,i}| - r_i^S(A))} < \gamma, \quad \text{for all } \ i \in S, \tag{1.4.10}$$

$$\gamma < \frac{|a_{j,j}| - r_j^{\overline{S}}(A)}{r_j^S(A)}, \quad \text{for all } \ j \in \overline{S}, \ \text{ such that } \ r_j^S(A) \neq 0, \tag{1.4.11}$$

and

$$|a_{j,j}| > r_j^{\overline{S}}(A), \quad \text{for all } \ j \in \overline{S}, \ \text{ such that } \ r_j^S(A) = 0. \tag{1.4.12}$$

Since (1.4.10) and (1.4.11) give, respectively, lower and upper bounds for the parameter γ, we can take the greatest lower and smallest upper bound, which leads us to the interval of feasible values for γ, that assure AX to be an SDD matrix:

$$0 \leq \alpha_S(A) := \min_{i \in S} \frac{r_i^{\overline{S}}(A)}{(|a_{i,i}| - r_i^S(A))} < \gamma < \max_{j \in \overline{S},\, r_j^S(A) \neq 0} \frac{|a_{j,j}| - r_j^{\overline{S}}(A)}{r_j^S(A)} =: \beta_S(A). \tag{1.4.13}$$

Here, we define $\beta_S(A) = +\infty$ in the case when $r_j^S(A) = 0$, for all $j \in \overline{S}$.

Having (1.2.25), condition (1.2.26) is equivalent to the fact that $\alpha_S(A) < \beta_S(A)$. Hence, we can choose $\gamma > 0$ such that, according to (1.4.13), AX is *SDD*. Thus, having AX and X nonsingular, A is nonsingular. □

Now, using these ideas, we formulate two theorems on the characterization of DZ-SDD matrices, S-SDD and CKV-SDD matrices. Their proofs include the original proofs of Theorems 1.2.12 and 1.2.13, and additional results obtained in [20] and [14], respectively.

Theorem 1.4.1. *Given*

$$\mathbb{X}_k := \{X = \operatorname{diag}(x_1, x_2, ..., x_n) \in \mathbb{D} \,:\, x_j = 1,\, j \in N \setminus \{k\}\}\,, \tag{1.4.14}$$

where $k \in N$, *and*

$$\mathbb{X}_{DZ} := \bigcup_{k \in N} \mathbb{X}_k, \tag{1.4.15}$$

the class $\mathbb{K}^{\mathbb{X}_{DZ}}$ *is the class of Dashnic-Zusmanovich strictly diagonally dominant matrices. Moreover,* $\mathbb{X}^{\mathbb{K}^{\mathbb{X}_{DZ}}} = \mathbb{X}_{DZ}$ *holds.*

Proof. From the proof with the scaling technique of the Theorem 1.2.12, it obviously holds that every DZ-SDD matrix belongs to the class $\mathbb{K}^{X_{DZ}}$. We now prove the converse. Assume that $A \in \mathbb{K}^{X_{DZ}}$, i.e., there exists $X \in \mathbb{X}_{DZ}$ such that AX is SDD. Set $\ell \in N$ so that $x_\ell = \gamma > 0$ and $x_j = 1$, for all $j \in N \setminus \{\ell\}$. Then,

$$\gamma |a_{\ell,\ell}| > r_\ell(A), \tag{1.4.16}$$

and, for all $j \in N \setminus \{\ell\}$,

$$|a_{j,j}| > \sum_{k \in N \setminus \{\ell, j\}} |a_{j,k}| + \gamma |a_{j,\ell}| = r_j(A) - |a_{j,\ell}| + \gamma |a_{j,\ell}|. \tag{1.4.17}$$

Hence, from (1.4.4) and (1.4.5), it follows that $\alpha_\ell(A) < \gamma < \beta_\ell(A)$, which is, in the case when $\beta_\ell(A) < +\infty$, equivalent to (1.2.20), i.e., to the fact that A is a DZ-SDD matrix. If $\beta_\ell(A) = +\infty$, i.e., if $a_{j,\ell} = 0$ for all $j \in N \setminus \{\ell\}$, then, for all $j \in N \setminus \{\ell\}$, from (1.4.17), we have $|a_{j,j}| > r_j(A) = r_j(A) - |a_{j,\ell}|$, which, together with the fact that right hand side of (1.2.20) is equal to zero, implies that A is DZ-SDD. Therefore, every matrix from $\mathbb{K}^{\mathbb{X}_{DZ}}$ is DZ-SDD, and thus, $\mathbb{K}^{\mathbb{X}_{DZ}}$ is exactly the class of DZ-SDD matrices.

In general, we have that $\mathbb{X}_{DZ} \subseteq \mathbb{X}^{\mathbb{K}^{\mathbb{X}_{DZ}}}$. The opposite inclusion is not so obvious, and it is the main result of the paper of [20]. □

Theorem 1.4.2. *Given a nonempty set $S \subseteq N$, define*

$$\mathbb{X}_S := \left\{X = \mathrm{diag}(x_1, x_2, ..., x_n) \in \mathbb{D} \,:\, x_i = x_k, \text{ for all } i, k \in S, \text{ and } x_j = 1, \text{ for all } j \in \overline{S}\right\}, \tag{1.4.18}$$

and

$$\mathbb{X}_{CKV} := \bigcup_{S \subseteq N} \mathbb{X}_S. \tag{1.4.19}$$

Then, for an arbitrary nonempty $S \subseteq N$, $\mathbb{K}^{X_S}$ and $\mathbb{K}^{X_{CKV}}$ are the classes of S-strictly diagonally dominant and Cvetković-Kostić-Varga strictly diagonally dominant matrices, respectively. Moreover, $\mathbb{X}^{\mathbb{K}^{\mathbb{X}_S}} = \mathbb{X}_S$ and $\mathbb{X}^{\mathbb{K}^{\mathbb{X}_{CKV}}} = \mathbb{X}_{CKV}$.

Proof. Similar to the proof of the previous theorem, it is easy to show that if $A \in \mathbb{K}^{\mathbb{X}_S}$, then A is an S-SDD matrix, and that if $A \in \mathbb{K}^{\mathbb{X}_{CKV}}$, then A is a CKV-SDD matrix, which implies that $\mathbb{K}^{X_S}$ and $\mathbb{K}^{X_{CKV}}$ are exactly the classes of S-SDD and CKV-SDD matrices, respectively. The proof that $\mathbb{X}^{\mathbb{K}^{\mathbb{X}_S}} = \mathbb{X}_S$ and $\mathbb{X}^{\mathbb{K}^{\mathbb{X}_{CKV}}} = \mathbb{X}_{CKV}$ is given in [14]. □

Note that using the same idea with Beauwens-Neumann's theorem in the place of Fiedler-Pták's theorem, the proof of the Theorem 1.2.17 follows directly.

Chapter 2

Eigenvalue Localization

In this chapter we will present one of the most popular applications of diagonal dominance - the theory of eigenvalue localization of Geršgorin type. First, we start with the well-known result of Geršgorin from [25], accompanied with its very well-known generalization by diagonal similarities, and the use of irreducibility in determining additional information about eigenvalues on the boundary of the localization set. The second section will focus on the equivalence between nonsingularity results and eigenvalue localization results. There, the term *localization of Geršgorin type*, will be introduced, which will give a unifying framework for all of the theorems that follow. A special focus will be on the scaling technique, which reveals the relationships between certain subclasses of H-matrices and the concept of minimal Geršgorin sets. Finally, the third section will be dealing with minimal Geršgorin sets. Its theoretical properties, and recent contributions to its effective computation, will be presented. Most of the material presented in this chapter could be find, in more detail, in author's master thesis *Eigenvalue localization by Geršgorin type theorems*, (in serbian), [34].

2.1 Geršgorin's Theorem

Starting from the original Geršgorin result on the set in the complex plane which includes the spectra of a given matrix, [25], we will present its first generalization, with a short historical overview. To conclude this section, we will consider the conditions for an eigenvalue to be on the boundary of the Geršgorin set, the importance of irreducibility, and the role played by Olga Taussky, [45] and [46].

Given an arbitrary square matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, the set of its eigenvalues is called the **spectrum**, and is denoted by $\sigma(A)$, i.e.,

$$\sigma(A) := \{\lambda \in \mathbb{C} : det(\lambda I_n - A) = 0\}. \tag{2.1.1}$$

Additionally, we define

$$\begin{cases} \Gamma_i(A) := \{z \in \mathbb{C} \, : \, |z - a_{i,i}| \leq r_i(A)\}, \quad (i \in N), \\ \Gamma(A) := \bigcup_{i \in N} \Gamma_i(A). \end{cases} \tag{2.1.2}$$

The following proposition is the famous Geršgorin's theorem [25]. In order to be in the agreement with contemporary notation, we give it in the form it was stated in Theorem 1.1 of [51].

Theorem 2.1.1. (Geršgorin's first theorem) *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, let λ be an eigenvalue. Then, there exists an index $k \in N$, such that*

$$|\lambda - a_{k,k}| \leq r_k(A), \tag{2.1.3}$$

implying that $\lambda \in \Gamma_k(A)$, and, therefore $\lambda \in \Gamma(A)$. Since $\lambda \in \sigma(A)$ is arbitrary, consequently, it follows that

$$\sigma(A) \subseteq \Gamma(A). \tag{2.1.4}$$

Proof. Given $\lambda \in \sigma(A)$, let $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{C}^n$ be its associated nonzero eigenvector, i.e., $A\mathbf{x} = \lambda \mathbf{x}$. In other words, we have that

$$\sum_{j \in N} a_{i,j} x_j = \lambda x_i, \quad (i \in N),$$

or, equivalently,

$$(\lambda - a_{i,i}) x_i = \sum_{j \in N \setminus \{i\}} a_{i,j} x_j, \quad (i \in N). \tag{2.1.5}$$

Since $\mathbf{x} \neq \mathbf{0}$, there exists an index $k \in N$, so that $0 < |x_k| = \max\{|x_i| \, : \, i \in N\}$. Hence, (2.1.5), together with the triangle inequality, implies that

$$|\lambda - a_{k,k}||x_k| \leq \sum_{j \in N \setminus \{k\}} |a_{k,j}||x_j| \leq |x_k| r_k(A),$$

where $r_k(A) := \sum_{j \in N \setminus \{k\}} |a_{k,j}|$, as in (1.1.2). Finally, on dividing the last inequality by $|x_k| > 0$, we obtain (2.1.3). Thus, $\lambda \in \Gamma_k(A)$, and, consequently, $\lambda \in \Gamma(A)$.

Since λ is an arbitrary eigenvalue, (2.1.4) holds. □

Observe that the set $\Gamma_i(A)$, which we will call the ***i*-th Geršgorin disk** of the matrix A, is a closed disk in complex plane, centered in $a_{i,i}$, and with radius $r_i(A)$. The set $\Gamma(A)$ is, therefore, the union of all Geršgorin disks, and, consequently, it is closed and bounded, i.e., it is a compact subset of $\mathbb{C}$, which includes the spectrum of the matrix. The geometrical structure of Geršgorin set, as well as the way it captures eigenvalues, can vary, as we shall see at the end of this subsection.

While going through the proof, one could notice that we had an almost identical reasoning before. Namely, calling $\lambda = 0$, we practically obtain the proof of Lévy-Desplanques theorem.

In fact, the relationship between these two famous theorems is very close. More precisely, there exists an ***equivalence*** between the Geršgorin's theorem and the Lévy-Deslanques's theorem, which is high-lighted in the book of Varga [51].

In other words, starting from the fact that Geršgorin's theorem holds, we can show that every SDD matrix is nonsingular, i.e., the theorem of Lévy-Desplanques holds. And, vice versa, assuming that every SDD matrix is nonsingular, we can deduce that all eigenvalues of an arbitrary matrix are inside of the Geršgorin set.

The idea, in fact, is very simple: it consists of the fact that a square *matrix is singular if and only if at least one of its eigenvalues is equal to zero.*

More precisely, while assuming that Geršgorin's theorem is true, take an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ that is SDD, i.e., it fulfills (1.1.3), and, contrary to the theorem of Lévy-Desplanques, assume that it is singular, i.e., that one of its eigenvalues is zero. But, Geršgorin's theorem implies that $0 \in \Gamma(A)$, i.e., there exists an index $k \in N$, so that $|0 - a_{k,k}| = |a_{k,k}| \leq r_k(A)$, which directly contradicts (1.1.3). Therefore, A is a nonsingular matrix.

In the opposite direction, assuming that the theorem of Lévy-Desplanques holds, i.e., that every SDD matrix is nonsingular, we take an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ and one of its eigenvalues λ. Since $\lambda I_n - A$ is singular, under that assumption, it cannot be that $\lambda I_n - A$ is an SDD matrix, i.e., condition (1.1.3) cannot hold, which means that for at least one $k \in N$, $|\lambda - a_{i,i}| = |(\lambda I_n - A)_{i,i}| \leq r_i(\lambda I_n - A) = r_i(A)$, and, hence, (2.1.3) implies that $\lambda \in \Gamma_k(A) \subseteq \Gamma(A)$.

As we have seen in the first chapter, there are many different extensions, or, generalizations, of the Lévy-Desplanques theorem, so one would expect that our mentioned equivalence would occur in these cases, too, producing new regions in complex plane that contain eigenvalues of a given matrix.

It is interesting that, although it was implicitly used in many papers on eigenvalue localization and on diagonal dominance, this equivalence was stated explicitly for the first time in the book of Varga in 2004, [51], fairly late, considering that the topic emerged in matrix theory at the beginning of the 20th century. For that reason we will call it **Varga's Equivalence Principle**, and give its exact mathematical formulation in Section 2 of this chapter, where we will use it more extensively. But, before that, in the subsections that follow, we will consider some of the first extensions of Geršgorin's theorem, and some interesting questions that arose from them.

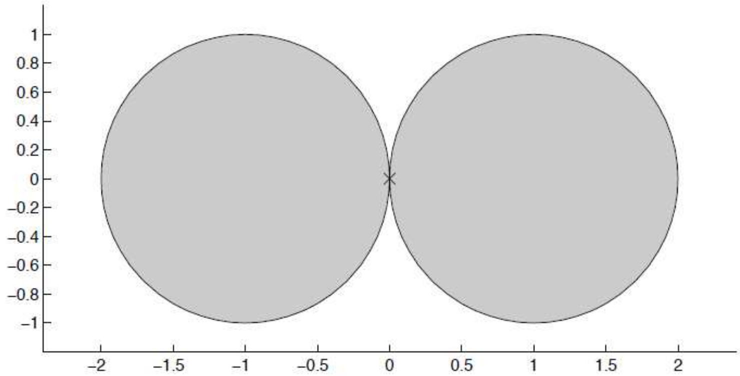

Figure 2.1.1: The Geršgorin set of the matrix A_1 of the Example 2.1.2
(Geršgorinov skup matrice A_1 iz Primera 2.1.2)

The beauty of Geršgorin's theorem lies in its simplicity. Namely, given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, by simple arithmetic, we obtain the row sums $\{r_i(A)\}_{i \in N}$, that are the radii of n disks whose union contains n eigenvalues of the matrix A. Nevertheless, from the following example, we can see that this information about eigenvalues, obtained in such a way, is not necessarily of great practical value. As Figure 2.1.1 indicates, the Geršgorin set may not give a good estimate of the spectrum, while Figure 2.1.3 illustrates the opposite case.

Example 2.1.2. *For matrices A_1, A_2 and A_3, given below, their spectra are calculated and their Geršgorin sets are plotted in Figures 2.1.1 - 2.1.3. The exact eigenvalues for these matrices are marked by "$\times$".*

$$A_1 = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, \quad \sigma(A_1) = \{0; 0\}.$$

$$A_2 = \begin{bmatrix} 2 & 2i & 0 \\ 1 & 8 & 1 \\ 2 & 0 & 14 \end{bmatrix}, \quad \sigma(A_2) = \{1.99 - 0.28i; 8.01 + 022i; 14.00 + 0.06i\}.$$

$$A_3 = \begin{bmatrix} i & 1 & 0 & 1 & 0 & 1 \\ 1 & -i & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 2 & i & i \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & i & i & -2 \end{bmatrix}, \; \sigma(A_3) = \{0; 0; 0; 1.61 + 0.77i; 0.43 - 0.33i; -2.03 - 0.44i\}.$$

Next, given $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, we observe that the same Geršgorin set $\Gamma(A)$ can be used to localize eigenvalues of the matrix $B = [b_{i,j}] \in \mathbb{C}^{n,n}$, such that, for every $i \in N$, $b_{i,i} = a_{i,i}$ and $r_i(B) = r_i(A)$ can significantly differ from the eigenvalues of the matrix A. This situation is illustrated by the following example.

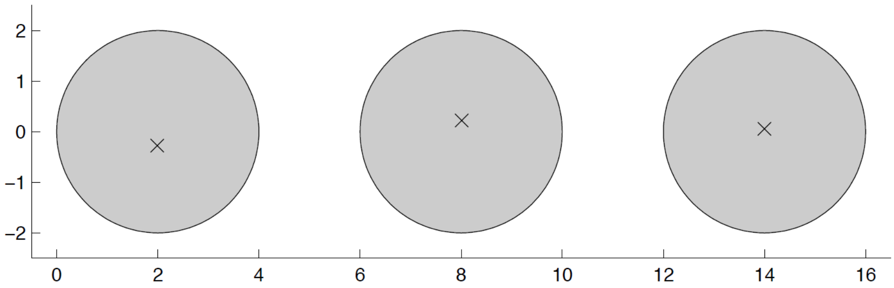

Figure 2.1.2: The Geršgorin set of the matrix A_2 of the Example 2.1.2
(Geršgorinov skup matrice A_2 iz Primera 2.1.2)

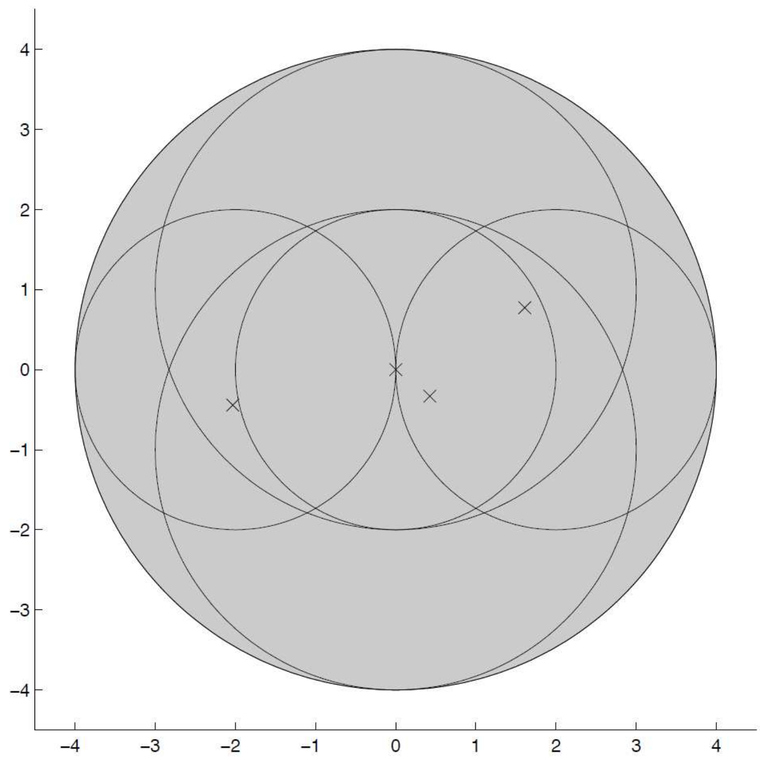

Figure 2.1.3: The Geršgorin set of the matrix A_3 of the Example 2.1.2
(Geršgorinov skup matrice A_3 iz Primera 2.1.2)

Example 2.1.3. *Let*

$$A_4 = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 8 & 0 \\ 2 & 0 & 14 \end{bmatrix}.$$

The Geršgorin set of the matrix A_4 is the same as Geršgorin set of the matrix A_2 from Example 2.1.2, and it is given in Figure 2.1.2. However, $\sigma(A_4) = \{1.53; 8.30; 14.17\}$, while $\sigma(A_2) = \{1.99 - 0.28i; 8.01 + 022i, 14.00 + 0.06i\}$.

Observing Figure 2.1.2 for the matrix A_2 of the Example 2.1.2, we notice that this Geršgorin set consists of three disks, each containing one eigenvalue. That it is not the general case, as the other two figures of the same example show. So, the question arises, when is it possible to claim that each disk contains exactly one eigenvalue. This is the second result of Geršgorin in the same paper from 1931, [25], which gives us the possibility to isolate an eigenvalue if we succeed to make one Geršgorin disk disjoint from the others. An extreme case is when all the disks are disjoint, implying that each one of them contains exactly one eigenvalue.

Given $n \geq 2$ and $S \subseteq N$, by $|S|$, we denote the **cardinality** of the set S, i.e., the number of its elements, and by $\overline{S} := N \setminus S$, as before, its complement. Furthermore, given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, then $\Gamma_S(A) := \bigcup_{i \in S} \Gamma_i(A)$ denote the part of Geršgorin set that "corresponds" to the indices from the set S. Then, the following theorem holds.

Theorem 2.1.4. (Geršgorin's second theorem) *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, and set of indices $S \subsetneq N$, if*

$$\Gamma_S(A) \bigcap \Gamma_{\overline{S}}(A) = \emptyset, \tag{2.1.6}$$

then $\Gamma_S(A)$ contains exactly $|S|$ eigenvalues of the matrix A, and, consequently, $\Gamma_{\overline{S}}(A)$ contains the reminder of the spectrum of A.

Instead of giving the proof, which can be found in [34], Theorem 1.1.3, following the same idea, in Section 2 of this chapter, we will prove a more general result.

2.1.1 Geršgorin's theorem and diagonal similarities

The first generalization of Geršgorin's theorem was considered in his original paper in 1931.

For a given $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$, such that $x_i > 0$, for every $i \in N$, we write $\mathbf{x} > \mathbf{0}$, and define a corresponding diagonal matrix $X := \text{diag}(x_1, x_2, \ldots, x_n)$, which is obviously nonsingular, i.e., X^{-1} exists. The set of all such positive diagonal matrices we denote by $\mathbb{D}$.

Take an arbitrary $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, and consider the matrix $X^{-1}AX = [\frac{a_{i,j}x_j}{x_j}]$. According to the similarity of matrices, their spectra are the same, i.e., $\sigma(A) = \sigma(X^{-1}AX)$. So, in order to localize eigenvalues of the matrix A, we can apply Geršgorin's theorem to the matrix $X^{-1}AX$, where we have n positive parameters which we can arbitrarily choose, and, hence, influence the shape and the size of the localization set.

In an analogous way to the notation given in (1.1.2) and (2.1.2), by

$$r_i^{\mathbf{x}}(A) := r_i(X^{-1}AX) = \sum_{j\in N\backslash\{i\}} \frac{|a_{i,j}|x_j}{x_i} \quad (i \in N,\ \mathbf{x} > \mathbf{0}), \tag{2.1.7}$$

we denote **i-th weighted row sum** of the matrix A, and by

$$\begin{cases} \Gamma_i^X(A) := \{z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq r_i(X^{-1}AX) = r_i^{\mathbf{x}}(A)\}, \quad (i \in N), \\ \Gamma^X(A) := \bigcup_{i\in N} \Gamma_i^X(A). \end{cases} \tag{2.1.8}$$

respectively, the ***i*-th scaled Geršgorin disk** of the matrix A, and the **scaled Geršgorin set** of the matrix A.

A direct corollary of Theorem 2.1.1 is:

Corollary 2.1.5. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \neq 2$, and positive diagonal matrix $X \in \mathbb{D}$,*

$$\sigma(A) \subseteq \Gamma^X(A), \tag{2.1.9}$$

holds, and, consequently,

$$\sigma(A) \subseteq \bigcap_{X\in\mathbb{D}} \Gamma^X(A). \tag{2.1.10}$$

The localization set given in (2.1.10) is obviously the best possible one that can be achieved through diagonal similarities. In the literature, the set in (2.1.10) is denoted by $\Gamma^{\mathfrak{R}}(A) := \bigcap_{\mathbf{x}>\mathbf{0}} \Gamma^X(A)$ and it is called the *minimal Geršgorin set* of the matrix A. It was, for the first time, considered in the paper of Varga in 1965, [47], while a detailed review of its different theoretical properties was given in the book of Varga in 2004, [51]. Here, following the unified notation of this dissertation, we will denote the **minimal Geršgorin set** of the matrix A by $\Gamma^{\mathbb{D}}(A) := \bigcap_{X\in\mathbb{D}} \Gamma^X(A)$. This set will be the topic of the last section of this chapter, where we will summarize its many theoretical properties, mainly investigated in [47], and reviewed in [51], and present new results on its effective calculation, published very recently in [52].

Of course, in order to make a similarity transformation, the only limitation on the matrix X is that it is nonsingular. So, instead of taking only nonsingular diagonal matrices, one could use any nonsingular matrix. In that way we get the best possible result, but as we can see in the following corollary, it is of purely theoretical value, since its calculation would use Jordan canonical forms. The detailed proof of this corollary could be found in [34], Corollary 1.2.2.

Corollary 2.1.6. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, and a nonsingular matrix $X \in \mathbb{C}^{n,n}$, then $\sigma(A) \subseteq \Gamma(X^{-1}AX)$ holds. Moreover,*

$$\sigma(A) = \bigcap_{\det(X)\neq 0} \Gamma(X^{-1}AX). \tag{2.1.11}$$

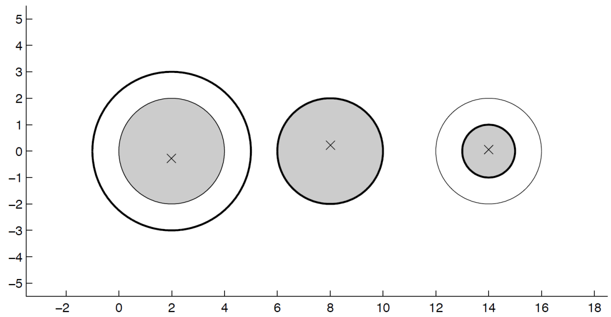

Figure 2.1.4: The Geršgorin set of the matrix A_4 of the Example 2.1.3
(Geršgorinov skup matrice A_4 iz Primera 2.1.3)

2.1.2 Geršgorin's theorem and matrix transpose

As we have seen in the previous chapter, the concept of SDD matrices can be applied to both rows and columns. Equivalently, we can apply the Geršgorin set either to the matrix A, or to A^T, and obtain the localization of their spectra $\sigma(A) = \sigma(A^T)$.

So, a very simple extension of the Geršgorin's theorem is the following:

Theorem 2.1.7. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 1$,*

$$\sigma(A) \subseteq \Gamma(A) \cap \Gamma(A^T) \tag{2.1.12}$$

holds.

Example 2.1.8. *To illustrate the localization result (2.1.12), consider the matrix A_2 of the Example 2.1.2, and the matrix*

$$A_5 = \begin{bmatrix} 1 & 1 & 2 \\ i & 1 & 0 \\ 0 & 1 & i \end{bmatrix}.$$

In Figure 2.1.4, the thicker lines show the Geršgorin circles for matrix A_4, and the thinner lines Geršgorin circles for A_4^T. The intersection $\Gamma(A_4) \cap \Gamma(A_4^T)$ is shaded, and the eigenvalues are, as before, denoted by "$\times$". In the same way, Figure 2.1.5 shows the localization given in (2.1.12) for the matrix A_5.

Looking at Figure 2.1.4, we see that, using information about the column sums of a matrix, and taking an intersection with the initial Geršgorin set, we can improve the localization set. The intersection in (2.1.12) in this case corresponds to the union of the "smaller" disks. Namely,

$$\Gamma(A_4) \cap \Gamma(A_4^T) = \bigcup_{i=1}^{3} [\Gamma_i(A_4) \cap \Gamma_i(A_4^T)].$$

That this equality doesn't hold in general, is clearly shown in Figure 2.1.5.

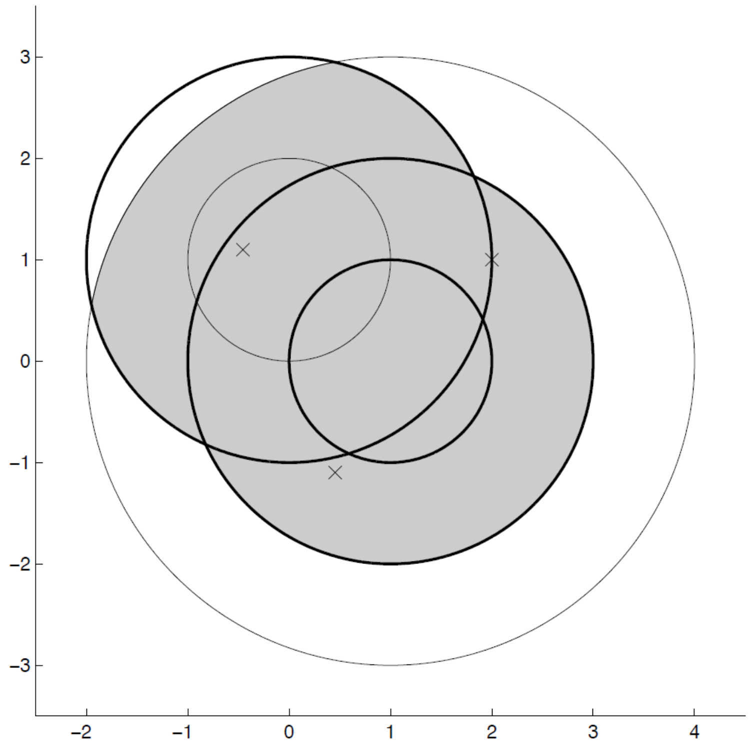

Figure 2.1.5: The Geršgorin set of the matrix A_5 of the Example 2.1.8
(Geršgorinov skup matrice A_5 iz Primera 2.1.3)

Inspired by this remark, for a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, define:

$$\begin{cases} \Gamma_i^m(A) := \Gamma_i(A) \cap \Gamma_i(A^T) = \{z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq \min\{r_i(A), c_i(A)\}\}, \quad (i \in N), \\ \Gamma^m(A) := \bigcup_{i \in N} \Gamma_i^m(A). \end{cases} \tag{2.1.13}$$

In general, $\Gamma^m(A) \subseteq \Gamma(A) \cap \Gamma(A^T)$, so, the question arises, whether inclusion $\sigma(A) \subseteq \Gamma^m(A)$ is true or not. The answer is no, in general, as the following example illustrates.

Example 2.1.9. *Let*

$$A_6 = \begin{bmatrix} 5 & 4 \\ 1 & 5 \end{bmatrix}.$$

For a given matrix, Figure 2.1.6 illustrates the set $\Gamma(A_6) \cap \Gamma(A_6^T)$, marked with thicker lines, while the set $\Gamma^m(A_6)$ is shaded. Eigenvalues are, as before, denoted by "$\times$".

Therefore, it remains an open question at what point in between (2.1.13), and (2.1.12) the set captures spectrum of a matrix.

The answer of this question will be given in the next section, using Varga's Equivalence Principle, and results of Cvetković, Bru, Kostić and Pedroche [11], presented in the previous chapter on diagonal dominance.

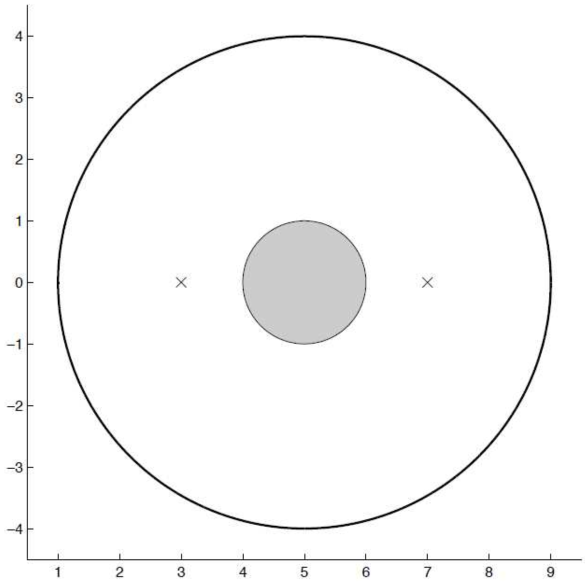

Figure 2.1.6: The Geršgorin set of the matrix A_6 of the Example 2.1.9
(Geršgorinov skup matrice A_6 iz Primera 2.1.9)

2.1.3 Geršgorin's theorem and irreducibility

As we have seen in the first chapter, the difference between singularity/nonsingularity of SDD matrices and DD matrices lies in the matrix being reducibile/irreducibile. More precisely, we have presented the result of Olga Taussky, partly from 1948 [45], and partly from 1949 [46], on diagonal dominance of an irreducible matrix. Now, we give an eigenvalue localization analogue, also published in the same papers.

To that end, denote $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$ to be the extended complex plane, and, for a given set $T \in \mathbb{C}$, let $cl(T)$ be the closure of the set T in $\mathbb{C}_\infty$, $\partial T := cl(T) \cap cl(\mathbb{C}_\infty \setminus T)$ the boundary, and $int(T) := T \setminus \partial T$ the interior of the set T.

Theorem 2.1.10. (Taussky's theorem) *Given an irreducible matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, if $\lambda \in \sigma(A)$ is such that, for every $i \in N$, $\lambda \notin int(\Gamma_i(A))$, i.e., $|\lambda - a_{i,i}| \geq r_i(A)$, then*

$$|\lambda - a_{i,i}| = r_i(A), \quad \textit{for all } \; i \in N.$$

In another words, all Geršgorin circles[1] *pass through λ. Therefore, if λ is an eigenvalue of the matrix A which lies on the boundary of Geršgorin set $\Gamma(A)$, then it lies in the intersection of all Geršgorin circles.*

[1] *Where, by the Geršgorin circle, we consider the boundary of Geršgorin disk, i.e., $\{z \in \mathbb{C} : |z - a_{i,i}| = r_i(A)\}$.*

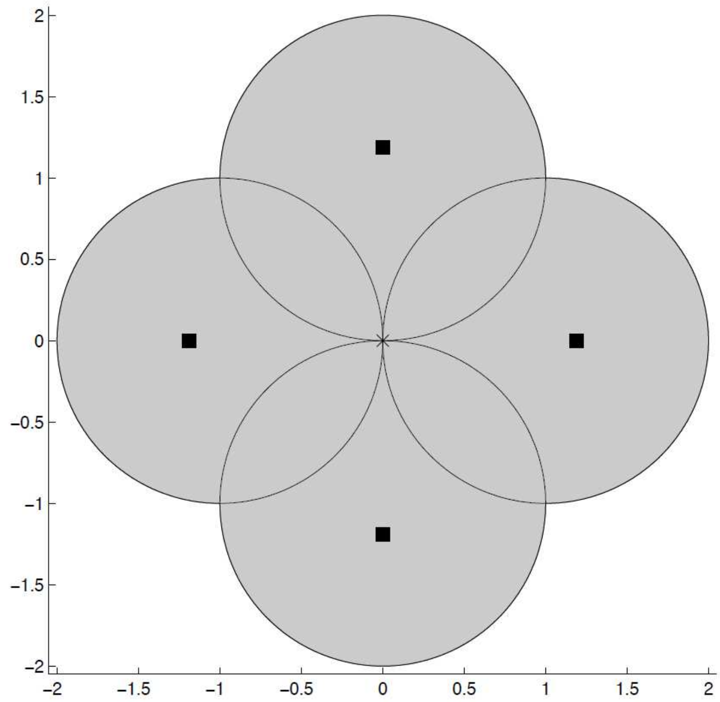

Figure 2.1.7: The Geršgorin disks of the matrix $A_7(\theta)$ given in (2.1.14)
(Geršgorinovi diskovi za matricu $A_7(\theta)$ datu u (2.1.14))

As an illustration of this theorem, observe the matrix

$$A_7(\theta) = \begin{bmatrix} 1 & e^{i\theta} & 0 & 0 \\ 0 & i & e^{i\theta} & 0 \\ 0 & 0 & -1 & e^{i\theta} \\ e^{i\theta} & 0 & 0 & -i \end{bmatrix}, \tag{2.1.14}$$

where $\theta \in [0, 2\pi)$ is a free parameter. In Figure 2.1.7, the Geršgorin disks $\Gamma_i(A_7(\theta))$, $1 \leq i \leq 4$, don't depend of the value of the parameter θ.

Matrix $A_7(\pi/4)$ is irreducible and singular, with $\sigma(A_7(\pi/4)) = \{0, 0, 0, 0\}$. As for any $i = 1, 2, 3, 4$, zero is not in the interior of the circle $\Gamma_i(A_7(\pi/4))$, according to Theorem 2.1.10, all Geršgorin circles pass through it. This is illustrated in Figure 2.1.7, where zero is marked by "$\times$".

Note, on the other hand, that zero is an eigenvalue of the irreducible matrix $A_7(\pi/4)$ that is *not* on the boundary of Geršgorin set, although all Geršgorin circles pass through it. Therefore, the second part of the theorem 2.1.10 gives us only a *necessary condition* for an eigenvalue of irreducible matrix to be found on the boundary of Geršgorin set, and not a sufficient one.

More details about the matrix $A_7(\theta)$ can be found in [51], Exercise 6, p. 18.

Necessary and sufficient conditions are given in [33], Theorem 3.2, where the problems of singularity/nonsingularity and eigenvalue localization are considered for block matrices.

As Theorem 2.1.10 gives conditions which insure that an eigenvalue of an irreducible matrix is in the intersection of all Geršgorin circles, one can ask the following: given an irreducible matrix, is the point through which all Geršgorin circles pass (if such a point exists) an eigenvalue? That, in general, is *not* valid, and can be illustrated by matrix $A_7(0)$, whose eigenvalues are marked with ■ in Figure 2.1.7.

2.2 Geršgorin-type Theorems

Following the concepts and open questions presented in the first section of this chapter, we focus now on the theorems of Geršgorin type that provide many answers. Unlike the original Geršgorin set, which consists of n disks in the complex plane, theorems that follow introduce localizations that consist of "more complicated" sets, so we review the possibilities of their implementation and necessary costs of their calculation, too, as well the justification for each one of them.

We start with by defining what precisely is a Geršgorin-type theorem. We will use, in previous section already announced, Varga's Equivalence Principle, and, thus, provide a clear framework for this chapter. Subsections that will follow after that, focus on specific results derived from the Section 1.2, and their generalizations. Thus, Subsection 2.2.1 starts from the result of Ostrowski given in Theorem 1.2.1, introduces the Brauer localization, [5], and deals with issues of its applicability and accuracy. In Subsection 2.2.2 we use, as in Subsection 1.2.2, graph theory to obtain Brualdi's results [7, 8], and Varga's, [49, 51], generalizations of Brualdi's results. Subsection 2.2.3 refers to the application of a scaling technique, introduced in Section 1.4, in order to improve localization sets, [19, 14]. Finally, Subsection 2.2.4 discusses the Ostrowski localization sets, based on the nonsingularity of α_1 and α_2-matriccs, and their characterizations.

As it was the case with Geršgorin's theorem, generally, there is a clear connection (more precisely, equivalence) between the propositions on the localization of eigenvalues and propositions about the matrix nonsingularity.

To investigate this in detail, we continue by stating this equivalence in form of Theorem 2.1.1 from [34], which we will later apply to all of the previously mentioned classes of nonsingular matrices, in order to produce eigenvalue localization sets.

It is interesting to note that this equivalence, although published for Geršgorin's theorem in the independent work of Rohrbach in 1931, too, [42], was relatively recently fully recognized. In the book of R.S. Varga, "Geršgorin and His Circles" [51], it is one of the recurring themes, as we have mentioned before. To emphasize the importance of this recurring theme, we will formulate it, for the first time, explicetely as the following Theorem.

Theorem 2.2.1. (Varga's Equivalence Principle) *Given a class of square complex matrices of an arbitrary size, denoted by $\mathbb{K}$, for an arbitrary square matrix A, define the set of complex numbers*

$$\Theta^{\mathbb{K}}(A) := \{z \in \mathbb{C} \,:\, zI - A \notin \mathbb{K}\}\,. \tag{2.2.1}$$

Then, the following two conditions are equivalent:

- *All matrices from $\mathbb{K}$ are nonsingular, and*
- *Given an arbitrary square matrix A, the set $\Theta^{\mathbb{K}}(A)$ contains all its eigenvalues, i.e., $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$.*

Proof. Assuming that all matrices in $\mathbb{K}$ are nonsingular, taking an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ and $\lambda \in \sigma(A)$, the matrix $\lambda I_n - A$ is singular, and, hence, $\lambda I_n - A \notin \mathbb{K}$. Therefore, $\lambda \in \Theta^{\mathbb{K}}(A)$, and, consequently, $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$.

For proving the opposite implication, assume that for every $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$. Now, assume that $A \in \mathbb{K}$ is singular. Then, $0 \in \sigma(A)$, and, consequently, $0 \in \Theta^{\mathbb{K}}(A)$. But, this is equivalent to the fact that $0I_n - A = A \notin \mathbb{K}$, which is an obvious contradiction. Thus, every $A \in \mathbb{K}$ is nonsingular. □

In the extreme case when $\mathbb{K}$ is taken to be the class of *all* nonsingular matrices, then for every A, $\Theta^{\mathbb{K}}(A) = \sigma(A)$ holds. Narrowing the class $\mathbb{K}$, we are obtaining the set $\Theta^{\mathbb{K}}(A)$ which becomes, in general, "wider", and, thus, we are obtaining an "approximation" of the spectrum, i.e., we get a certain localization set for the spectrum. More precisely, given that $\mathbb{K}_1 \subseteq \mathbb{K}_2$, then by definition follows that $\Theta^{\mathbb{K}_2}(A) \subseteq \Theta^{\mathbb{K}_1}(A)$.

But, the question is how "interesting" is the obtained localization? In other words, are we able to construct it in the complex plane, and, whether the cost of this is significantly less then calculating the actual eigenvalues themselves?

In the case of Geršgorin's theorem, where $\Theta^{\mathbb{K}}(A) = \Gamma(A)$, we have seen that the class $\mathbb{K}$ is the class of all SDD matrices. In a similar way, we define the term *Geršgorin-type theorem.*

First, we say that set $\Theta^{\mathbb{K}}(A) = \{z \in \mathbb{C} : zI - A \notin \mathbb{K}\}$ is a **Geršgorin-type set**, if $\mathbb{K}$ is a diagonally dominant-type class of nonsingular matrices. In another words, due to Theorem 1.3.9, it is a subclass of nonsingular H-matrices.

Hence, a **Geršgorin-type theorem** is a statement that claims that a certain Geršgorin-type set contains the spectra of a given matrix. So, we refer to such sets also as **Geršgorin-type localization sets**, or **Geršgorin-type eigenvalue inclusion sets**.

Now, we extend the concept of diagonal scaling, originally present in Geršgorin's paper [25], to other Geršgorin-type eigenvalue inclusion sets. As in Subsection 2.1.1, we consider the scaled Geršgorin sets, given in (2.1.8), and for a family of positive diagonal matrices $\mathbb{X} \subseteq \mathbb{D}$, we define the corresponding set in complex plane:

$$\Gamma^{\mathbb{X}}(A) := \bigcap_{X \in \mathbb{X}} \Gamma^{X}(A), \tag{2.2.2}$$

which we all call it the **minimal Geršgorin set attributed to the family** $\mathbb{X}$.

Obviously, when $\mathbb{X} = \mathbb{D}$, we have the minimal Geršgorin set, as it was defined by Varga in [47]. In that case, according to the Fiedler-Pták's theorem, $\mathbb{K}^{\mathbb{D}} = \mathbb{H}$, so, for an arbitrary matrix A, $\Gamma^{\mathbb{D}}(A) = \Theta^{\mathbb{H}}(A)$. Therefore, the minimal Geršgorin set is the *best possible* Geršgorin-type eigenvalue inclusion set. But, as we will see in detail in the last section of this chapter, to obtain this set, for a given matrix, it is not an easy task. Therefore, one could be motivated to determine the cases when the minimal Geršgorin set, attributed to a specific family $\mathbb{X}$, can be *explicitly expressed*, of course with the tendency for $\mathbb{X}$ to be as much "closer" to $\mathbb{H}$ as possible. Here, explicitly means without any use of the scaling parameters, entries of the scaling matrices.

Before we continue with giving particular results of Geršgorin-type, we will generalize another important property of Geršgorin sets. Namely, Geršgorin's second theorem is a key ingredient if one wants to find an accurate approximation to one, or a few eigenvalues, by giving the possibility to be certain of the number of eigenvalues that are contained in the disjointed parts of the localization set. Therefore, we generalize this concept to "practically all" Geršgorin-type theorems.

Definition 2.2.2. A given class of matrices $\mathbb{K}$ is said to be **positively homogenous**, if $A \in \mathbb{K}$ implies $\alpha A \in \mathbb{K}$, for arbitrary $\alpha > 0$.

Theorem 2.2.3. (Isolation Principle) *Given a Geršgorin-type set*

$$\Theta^{\mathbb{K}}(A) = \{z \in \mathbb{C} \, : \, zI - A \notin \mathbb{K}\}, \tag{2.2.3}$$

where $\mathbb{K} \subset \mathbb{H}$ is a positively homogenous diagonally dominant-type subclass of H-matrices, for any matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and

$$\Theta^{\mathbb{K}}(A) = U \cup V, \tag{2.2.4}$$

then, the set U contains exactly $|\{i \in N \, : \, a_{i,i} \in U\}|$ eigenvalues of the matrix A.

Proof. Let $D_A := \text{diag}(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$. Take the splitting of the matrix $A = D_A - F_A$, and consider the family of matrices $A(t) := D_A - tB_A$, for $0 \leq t \leq 1$.

First, let $t \in (0, 1]$, and take $z \in \Theta^{\mathbb{K}}(A(t))$, i.e., $zI - A(t) \notin \mathbb{K}$. Assume, on the contrary that $z \notin \Theta^{\mathbb{K}}(A)$, i.e., that $zI - A \in \mathbb{K}$. Since,

$$t|zI - A| = t(|zI - D_A| + |F_A|) \pm |zI - D_A| = |zI - A(t)| - (1 - t)|zI - D_A|,$$

then,

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

According to the assumption that $\mathbb{K}$ is positively homogenous and a DD-type class of matrices, from the previous equality we conclude that $|zI - A(t)| \in \mathbb{K}$. Therefore, $zI - A(t) \in \mathbb{K}$, which is an obvious contradiction. Thus, $z \notin \Theta^{\mathbb{K}}(A)$, and, consequently, $\Theta^{\mathbb{K}}(A(t)) \subseteq \Theta^{\mathbb{K}}(A)$, for all $t \in (0, 1]$.

Let us consider the case when $t = 0$. Then, $A(0) = D_A$, and $z \in \Theta^{\mathbb{K}}(A(0))$ if and only if $zI - D_A \notin \mathbb{K}$. Obviously, if $z = a_{i,i}$, for some $i \in N$, then $zI - D_A$ has a zero on diagonal. Thus it can not be in $\mathbb{K}$ which is the DD-type class of matrices. Therefore, $a_{i,i} \in \Theta^{\mathbb{K}}(A(0))$, for all $i \in N$. For the same reasons, $a_{i,i} \in \Theta^{\mathbb{K}}(A)$, $i \in N$. On the other hand, when $z \neq a_{i,i}$, for all $i \in N$, $zI - D_A$ is a nonsingular diagonal matrix, implying that $zI - D_A \in \mathbb{K}$, i.e., $z \notin \Theta^{\mathbb{K}}(A(0))$. So, $\Theta^{\mathbb{K}}(A(0)) = \{a_{1,1}, a_{2,2}, \ldots, a_{n,n}\} \subseteq \Theta^{\mathbb{K}}(A)$. In another words, we have obtained that $\Theta^{\mathbb{K}}(A(0)) = \sigma(A(0)) = \{a_{1,1}, a_{2,2}, \ldots, a_{n,n}\}$, and that $\Theta^{\mathbb{K}}(A(t)) \subseteq \Theta^{\mathbb{K}}(A)$, for all $t \in [0, 1]$.

Now, since the sets $U, V \subseteq \mathbb{C}$ are disjoint, and $\{a_{1,1}, a_{2,2}, \ldots, a_{n,n}\} \subseteq \Theta^{\mathbb{K}}(A) = U \cup V$, if the number of diagonal entries that lie in the set U is denoted by m, then, of course, $n - m$ diagonal entries of the matrix $A(0)$ lie in the set V.

Let us with $\lambda(t)$ denote the eigenvalue of the matrix $A(t)$ that, for $t = 0$, becomes a diagonal entry that lies in the set $U \subseteq \mathbb{C}$. Since the eigenvalues are continuous functions of matrix entries, [40], we can consider $\{\lambda(t) \, : \, t \in [0, 1]\}$, as a continuous curve in the complex plane, such that $\{\lambda(t) \, : \, t \in [0, 1]\} \subseteq \Theta^{\mathbb{K}}(A) = U \cup V$. Again, since U and V are disjoint, the whole curve must be contained in exactly one of them. Therefore, since $\lambda(0) \in U$, $\lambda(1) \in U$. Consequently, the number of eigenvalues of $A(1) = A$ that are in the set U is m. □

Since many of the corollaries of this theorem are knowen, we could say that this result is essentially known, but until know it hasn't been stated in such a general way.

2.2.1 Brauer's ovals of Cassini

In this subsection, we start from Theorem 1.2.1, which claims the nonsingularity of doubly SDD matrices. Observing the condition (1.2.1), which defines them, it is clear that the class of doubly SDD matrices, denoted here by $\mathbb{K}_{dSDD}$, is a positively homogenous class and a DD-type class of matrices.

Namely, given $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $\alpha > 0$ and any matrix B such that $\langle B\rangle \geq \langle A\rangle$, we have that, for all $i \in N$,

$$(\langle B\rangle)_{i,i} = |b_{i,i}| \geq |a_{i,i}|, \text{ and } r_i(\langle B\rangle) = r_i(B) \leq r_i(A), \tag{2.2.5}$$

and

$$|(\alpha A)_{i,i}| = \alpha|a_{i,i}|, \text{ and } r_i(\alpha A) = \alpha r_i(A). \tag{2.2.6}$$

Clearly, if $A \in \mathbb{K}_{dSDD}$, then $\langle B\rangle \in \mathbb{K}_{dSDD}$ too, and, since $\alpha > 0$,

$$|(\alpha A)_{i,i}||(\alpha A)_{j,j}| > r_i(\alpha A) r_j(\alpha A)$$

is equivalent to (1.2.1), for all $i, j \in N$, $i \neq j$. Thus, $\mathbb{K}_{dSDD}$ is, by definition, a positively homogenous DD-type class of matrices.

Therefore, $\Theta^{\mathbb{K}_{dSDD}}(A)$ is a Geršgorin-type set, to which we can apply our Isolation Principle. So, it remains to consider the possibilities and the costs of constructing this set in complex plane for a given matrix A. First, we give an explicit form of this set and its corresponding Isolation Principle.

Theorem 2.2.4. (Brauer) *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, for every $\lambda \in \sigma(A)$, there exists a pair of indices $i, j \in N$, $i \neq j$, so that*

$$\lambda \in K_{i,j}(A) := \{z \in \mathbb{C} : |z - a_{i,i}||z - a_{j,j}| \leq r_i(A) r_j(A)\}, \tag{2.2.7}$$

and, consequently,

$$\sigma(A) \subseteq \mathcal{K}(A) := \bigcup_{i\in N} \bigcup_{j=1}^{i-1} K_{i,j}(A). \tag{2.2.8}$$

Proof. Consider the class of doubly SDD matrices $\mathbb{K}_{dSDD}$. According to Theorem 1.2.1, it is a class of nonsingular matrices. Thus, Varga's Equivalence Principle implies that $\sigma(A) \subseteq \Theta^{\mathbb{K}_{dSDD}}(A)$, for any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$. We prove that $\Theta^{\mathbb{K}_{dSDD}}(A) = \mathcal{K}(A)$. Start with any $z \in \Theta^{\mathbb{K}_{dSDD}}(A)$. This means that $zI_n - A \notin \mathbb{K}_{dSDD}$, or, equivalently, that there exist indices $i, j \in N$, $i \neq j$, such that

$$|(zI_n - A)_{i,i}||(zI_n - A)_{j,j}| \leq r_i(zI_n - A) r_j(zI_n - A). \tag{2.2.9}$$

But, since for all $i \in N$, $|(zI_n - A)_{i,i}| = |z - a_{i,i}|$, and $r_i(zI_n - A) = r_i(A)$, according to (2.2.7), the condition (2.2.9) is equivalent to the fact that $z \in K_{i,j}(A) = K_{j,i}(A)$. Therefore, $z \in \Theta^{\mathbb{K}_{dSDD}}(A)$ is equivalent to $z \in \mathcal{K}(A)$, implying that $\Theta^{\mathbb{K}_{dSDD}}(A) = \mathcal{K}(A)$. □

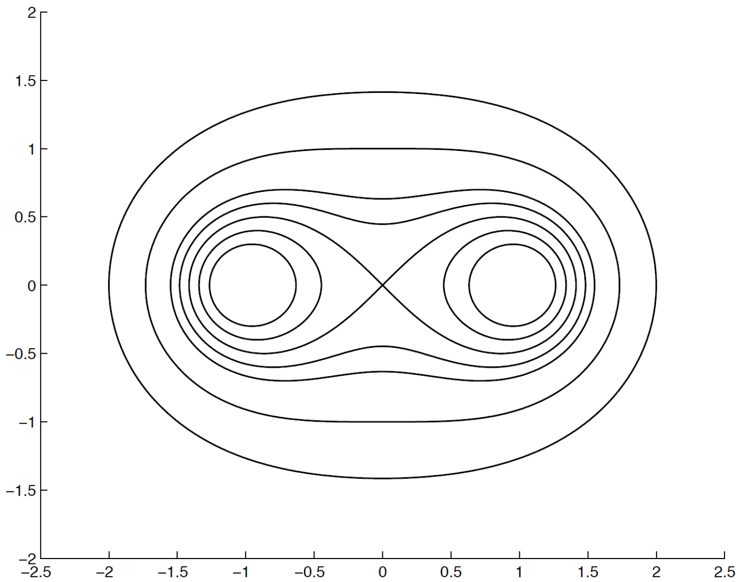

Figure 2.2.1: Ovals of Cassini with foci -1 and 1, and radius $\eta = 0.6, 0.8, 1, 1.2, 1.4, 2, 3$
(Kazinijevi ovali sa žižama u -1 *i* 1*, i polupre
čnika* $\eta = 0.6, 0.8, 1, 1.2, 1.4, 2, 3$*)*

Theorem 2.2.5. *Given an arbitrary matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U, V \subseteq \mathbb{C}$*, such that* $U \cap V = \emptyset$*, and* $\mathcal{K}(A) = U \cup V$*, then, the set* U *contains exactly* $|\{i \in N : a_{i,i} \in U\}|$ *eigenvalues of the matrix* A*.*

First of all, note that, **Brauer set** $\mathcal{K}(A)$, given in (2.2.8), unlike Geršgorin set, is made of $\frac{n(n-1)}{2}$ compact sets in the complex plane. They depend on diagonal entries and row sums, generally they are not disks. Namely, the boundary of $K_{i,j}(A)$ can be described with the following equation:

$$|z - \xi_1||z - \xi_2| = \eta, \tag{2.2.10}$$

with ξ_1 and ξ_2 being complex numbers, and $\eta \geq 0$. Numbers ξ_1 and ξ_2 are called foci of the curve (2.2.10) whose radius is η.

The shape of the curve depends of these values. Namely, for $0 < \eta < \frac{(\xi_1 - \xi_2)^2}{4}$, the curve (2.2.10), consists of two disjoint parts that asymptotically tend to *circles*, centered in ξ_1 and ξ_2, and of radius $\frac{\eta}{|\xi_1 - \xi_2|}$, when $\eta \to 0$.

For $\eta = \frac{(\xi_1 - \xi_2)^2}{4}$, (2.2.10) is second-order *lemniscate*, with foci ξ_1 and ξ_2, while, for the values $\eta > \frac{(\xi_1 - \xi_2)^2}{4}$, (2.2.10) becomes smooth curve without self-intersections, which tends to circle centered in $\frac{\xi_1 + \xi_2}{2}$ and radius $\frac{\eta}{|\xi_1 - \xi_2|}$, when $\eta \to \infty$.

Figure 2.2.1 illustrates these cases for the values $\xi_1 = -1$, $\xi_2 = 1$, and $\eta \in \{0.6, 0.8, 1, 1.2, 1.4, 2, 3\}$.

Generally, this curve is called the **Cassini[2] oval**, although its shape is oval only in the case when $\eta \geq \sqrt{|\xi_1 - \xi_2|}$. For that reason, we will call the set $K_{i,j}(A)$ (i, j)**-th Brauer oval of Cassini** of the matrix A.

[2]The name comes from Italian astronomer Giovanni Domenico Cassini (1625-1712), who spend all his professional life in Paris.

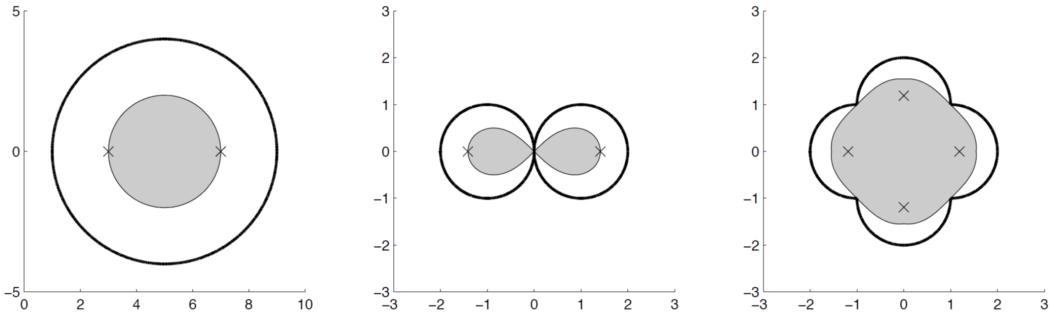

Figure 2.2.2: From left to right, Brauer set (thick boundary) and Geršgorin set (shaded) for matrices A_1, A_2 and A_3 of the Example 2.2.7
(Sa leva na desno, Brauerov skup (debljom linijom) i Geršgorinov skup (osenčen) za matrice A_1, A_2 i A_3 iz Primera 2.2.7)

Since, in order to determine Brauer set, we need to make significantly more computation (compared to the Geršgorin set), and, moreover, curves that define Brauer set are already more complicated than the simple Geršgorin disks, a natural question is if it is appropriate to use the Brauer eigenvalue localization. Namely, this localization is justified only if it gives significantly better results that the original Geršgorin one. Luckily, for the Brauer ovals of Cassini, the following statement is true, Theorem 2.3 of [51], and [53].

Theorem 2.2.6. *For any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, and for any two indices $i, j \in N$, $i \neq j$,*

$$K_{i,j}(A) \subseteq \Gamma_i(A) \cap \Gamma_j(A), \tag{2.2.11}$$

where equality holds if and only if $r_i(A) = r_j(A) = 0$, or if $r_i(A) = r_j(A) > 0$ in the case when $a_{i,i} = a_{j,j}$. Therefore, consequently,

$$\mathcal{K}(A) \subseteq \Gamma(A). \tag{2.2.12}$$

How significant this improvement can be, the following example shows.

Example 2.2.7. *Consider the following matrices*

$$A_1 = \begin{bmatrix} 5 & 4 \\ 1 & 5 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \text{and} \quad A_3 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & -i \end{bmatrix}.$$

Figure 2.2.2 illustrates, respectively, from the left to the right, Brauer sets $\mathcal{K}(A_1)$, $\mathcal{K}(A_2)$ and $\mathcal{K}(A_3)$, shaded, while the thick lines represent boundaries of the corresponding Geršgorin sets $\Gamma(A_1)$, $\Gamma(A_2)$ and $\Gamma(A_3)$.

As mentioned before, both the Brauer set and the Geršgorin set of a given matrix depend of the same collection of $2n$ data, while the Brauer set gives a localization set that is not greater then the Geršgorin one. The question is, how sharp is it.

Results of Varga and Krautstengl obtained 1999, in [53], show that Brauer localization sets are, indeed, sharp. Basically, their result states that the Brauer set for a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ gives the best possible estimate of eigenvalues, based on the data collections $\{a_{i,i}\}_{1 \in N}$ and $\{r_i(A)\}_{1 \in N}$.

To explain that, for a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, we define the **equiradial set** of A as

$$\omega(A) := \{B = [b_{i,j}] \in \mathbb{C} \,:\, b_{i,i} = a_{i,i} \text{ and } r_i(B) = r_i(A), \;\; i \in N\}, \tag{2.2.13}$$

and we also define

$$\widehat{\omega}(A) := \{B = [b_{i,j}] \in \mathbb{C} \,:\, b_{i,i} = a_{i,i} \text{ and } r_i(B) \leq r_i(A), \;\; i \in N\} \tag{2.2.14}$$

as the **extended equiradial set** for A.

Obviously, $\omega(A) \subseteq \widehat{\omega}(A)$, and, according to the definition of the Brauer ovals of Cassini, given in (2.2.7), it is clear that for each matrix from $\omega(A)$ the Brauer set remains the same as it is for the matrix A, while for each matrix from $\widehat{\omega}(A)$, it gets smaller and lies in $\mathcal{K}(A)$. Therefore, the collection of all eigenvalues of all matrices from these two sets of matrices is contained in the Brauer set of the matrix A, i.e.,

$$\sigma(\omega(A)) \subseteq \sigma(\widehat{\omega}(A)) \subseteq \mathcal{K}(A), \tag{2.2.15}$$

where

$$\sigma(\mathfrak{K}) := \bigcup_{A \in \mathfrak{K}} \sigma(A), \tag{2.2.16}$$

for an arbitrary family of matrices $\mathfrak{K}$.

Furthermore, according to Theorem 2.4 in [51], the equalities between these three sets of complex numbers are given in the next result.

Theorem 2.2.8. (Varga-Krautstengl) *Given any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, then*

$$\sigma(\omega(A)) = \begin{cases} \partial\mathcal{K}(A) = \partial K_{1,2}(A) & \text{if } n = 2, \text{ and,} \\ \mathcal{K}(A) & \text{if } n \geq 3, \end{cases} \tag{2.2.17}$$

and, in general, for any $n \geq 2$,

$$\sigma(\widehat{\omega}(A)) = \mathcal{K}(A). \tag{2.2.18}$$

So, for any matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ of size $n \geq 3$, the Brauer set $\mathcal{K}(A)$ perfectly estimates the spectrum of all matrices that are equiradial with the matrix A, and all matrices that are from the extended set $\widehat{\omega}(A)$. This is not the case with the Geršgorin set, as it is shown in Figure 2.2.2.

2.2.2 Brualdi sets

Moving forward to the product over the three and more rows in order to obtain nonsingularity results, as we have seen in Subsection 1.2.2, has led us to the concept of cycles in a graph of a matrix. Here, we review the original Brualdi result on the localization of eigenvalues, together with the theorem of Varga on the sharpness of Brualdi lemiscate sets.

Again, we start with the positively homogenous DD-type class of matrices, the class of Brualdi-SDD matrices, and construct the corresponding Geršgorin type set. Then, we deduce its explicit form given in the following theorem. We omit the proof, since it is completely analogous to the one in the previous subsection.

Theorem 2.2.9. (Brualdi)[3] *Given an arbitrary matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, *for every* $\lambda \in \sigma(A)$, *there exists a cycle* $\gamma \in C(A)$, *either strong or weak, so that*

$$\lambda \in \mathcal{B}_\gamma(A) := \left\{ z \in \mathbb{C} \,:\, \prod_{i\in\gamma} |z - a_{i,i}| \leq \prod_{i\in\gamma} \widetilde{r}_i(A) \right\}, \tag{2.2.19}$$

if the cycle $\gamma \in C(A)$ *is strong, or,*

$$\lambda \in \mathcal{B}_\gamma(A) := \{ z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq \widetilde{r}_i(A) = 0 \} = \{a_{i,i}\}, \tag{2.2.20}$$

if the cycle $\gamma = \{i\} \in C(A)$ *is weak. Consequently,*

$$\sigma(A) \subseteq \mathcal{B}(A) := \bigcup_{\gamma\in C(A)} \mathcal{B}_\gamma(A). \tag{2.2.21}$$

Theorem 2.2.10. *Given an arbitrary matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, *if there exist sets* $U, V \subseteq \mathbb{C}$, *such that* $U \cap V = \emptyset$, *and* $\mathcal{B}(A) = U \cup V$, *then, set* U *contains exactly* $|\{i \in N \,:\, a_{i,i} \in U\}|$ *eigenvalues of the matrix* A.

This result on the disjoint subsets of the Brualdi's set, is a new one, and was not stated anywhere before.

The set of complex numbers $\mathcal{B}$, defined in (2.2.21), we call the **Brualdi set** of a matrix A, while the set $\mathcal{B}_\gamma$, from (2.2.19)-(2.2.20), we call the **Brualdi lemniscate**.

To illustrate this set, let us consider matrix A_8 of the Example 1.2.5 in its normal reduced form (1.2.9). There, we have seen that $C(A_8) = \{\gamma_1, \gamma_2, \gamma_3, \gamma_4\}$, where $\gamma_1 = (3, 5, 4)$, $\gamma_2 = (5, 6)$, $\gamma_3 = (2)$, and $\gamma_4 = (1)$, so that the corresponding Brualdi lemniscates are:

$$\begin{aligned} \mathcal{B}_{\gamma_1}(A_8) &= \left\{ z \in \mathbb{C} \,:\, |z||z-1|^2 \leq 2 \right\}, \\ \mathcal{B}_{\gamma_2}(A_8) &= \{ z \in \mathbb{C} \,:\, |z||z-1| \leq 2 \}, \\ \mathcal{B}_{\gamma_3}(A_8) &= \{1\}, \quad \text{and} \quad \mathcal{B}_{\gamma_4}(A_8) = \{0\}, \end{aligned}$$

and we have that $\{0, 1, 2\} = \sigma(A_8) \subseteq \mathcal{B}(A_8)$, which is illustrated in the Figure 2.2.3.

Using Varga's Equivalence Principle, the fact that every doubly SDD matrix is a Brualdi SDD matrix, stated in Theorem 1.2.8, gives us the relationship between Brauer sets and Brualdi sets. The original proof of this theorem is given in [51], Theorem 2.9.

[3] *This theorem, like the one in Subsection 1.2.2, is also Varga's generalization of the original Brualdi's theorem from [7].*

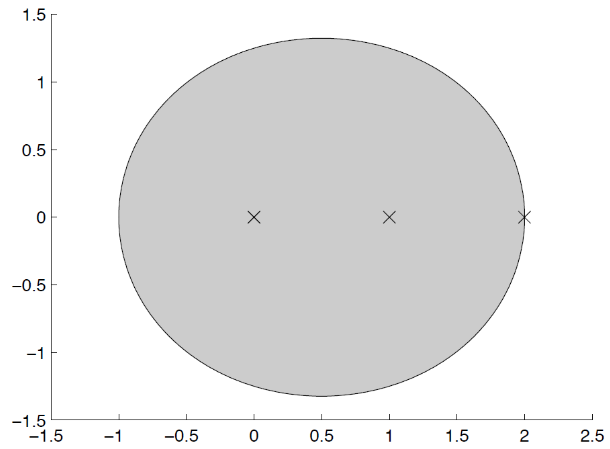

Figure 2.2.3: The Brualdi set for the matrix A_8 of the Example 1.2.5
(Brualdijev skup za matricu A_8 iz Primera 1.2.5)

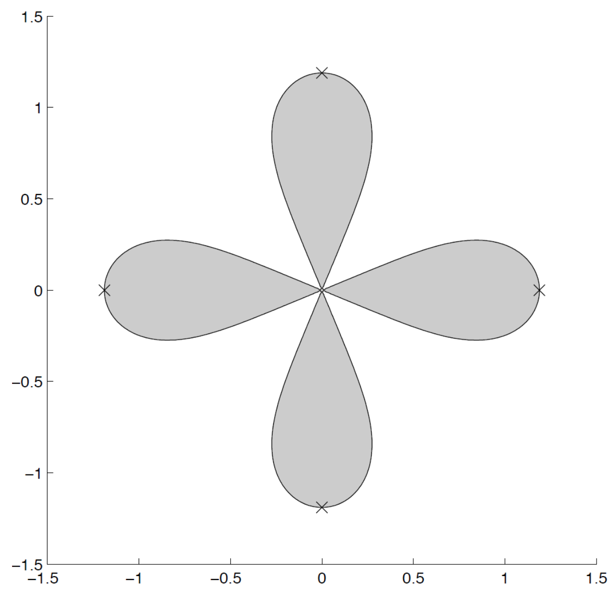

Figure 2.2.4: The Brualdi set for the matrix A_7 given in (2.1.14)
(Brualdijev skup za matricu A_7 datu sa (2.1.14))

Theorem 2.2.11. *Given an arbitrary matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$*,* $n \geq 2$*, then*

$$\mathcal{B}(A) \subseteq \mathcal{K}(A), \tag{2.2.22}$$

with $\mathcal{K}(A)$ *and* $\mathcal{B}(A)$ *given in (2.2.8) and (2.2.21), respectively.*

As we have discussed in the previous chapter, for a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, the Brauer set consists of $\frac{n(n-1)}{2}$ Cassini ovals, while the number of Brualdi lemniscates depends on the associated graph structure, and can vary quite a lot, from only one lemniscate, as in the case of matrix A_9 of the Example 1.2.9, to a few of them, Example 1.2.7, or up to the the number that far exceeds $\frac{n(n-1)}{2}$. The ultimate case is, again, when a matrix A has all nonzero off-diagonal entries, when each choice of two or more indices is a strong cycle, and their number is $\sum_{k=2}^{n} \frac{n!}{k!}$. In that case, for each cycle $\gamma = (i,j)$ of the length 2, (i,j)-Brauer's oval of Cassini becomes the same as Brauldi lemniscate associated with the cycle γ, i.e., $K_{i,j}(A) = \mathcal{B}_\gamma(A)$, implying that, in this case $\mathcal{K}(A) = \mathcal{B}(A)$.

So, in case of matrices without any zero off-diagonal entry, most of the cycles of the graph $\mathbb{G}(A)$ do not affect the Brualdi set, so, it opens a search for reducing the number of cycles that construct Brualdi localization set. Again, the reduced set of cycles is given in Theorem 2.10 from [51].

It is interesting to remark (Varga, private note) that, for a matrix $A \in \mathbb{C}^{n,n}$, with all nonzero off diagonal entries, the number of all Brualdi lemniscates, that are to be calculated, can be reduced to $2^n - (n+1)$, where $n \geq 2$. For example, for $n = 10$, the above number is $1,013$, while the number of all lemniscates is $1,012,073$.

As for the Brauer set, there is a corresponding theorem on the sharpness of the Brualdi sets. More details of this can be found in [51].

2.2.3 Cvetković-Kostić-Varga Sets

Here we start with the results on nonsingularity given in Theorems 1.2.12 and 1.2.13 and their characterizations given in Theorems 1.4.1 and 1.4.2, respectively.

As a consequence of Varga's Equivalence Principle, we have the following eigenvalue localization theorems.

Theorem 2.2.12. (**Dashnic-Zusmanovich**) *Let* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$*, with* $n \geq 2$*, be an arbitrary matrix, and* $\lambda \in \sigma(A)$ *be its arbitrary eigenvalue. Then, for every index* $i \in N$*, there exists an index* $j \in N \setminus \{i\}$*, such that*

$$\lambda \in \mathcal{D}_{i,j}(A) := \{z \in \mathbb{C} : |z - a_{i,i}| \cdot (|a_{j,j}| - r_j(A) + |a_{j,i}|) \leq r_i(A)|a_{j,i}|\}. \tag{2.2.23}$$

Consequently,

$$\sigma(A) \subseteq \mathcal{D}(A) := \bigcap_{i \in N} \mathcal{D}_i(A), \tag{2.2.24}$$

where

$$\mathcal{D}_i(A) := \bigcup_{j \in N \setminus \{i\}} \mathcal{D}_{i,j}(A). \tag{2.2.25}$$

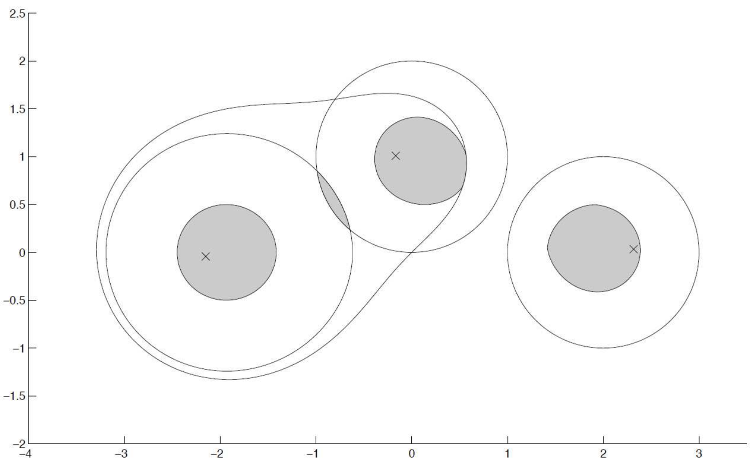

Figure 2.2.5: The Dashnic-Zusmanovich's set for the matrix A_4 of the Example 2.2.13
(Dašnjic-Zusmanovič skup za matricu A_4 iz Primera 2.2.13)

The following example illustrates this localization set. As it is illustrated, we can consider this **Dashnic-Zusmanovich's set**, given in (2.2.24), to be made up of the perturbed Brauer's Ovals of Cassini.

Example 2.2.13. *Let*

$$A_4 = \begin{bmatrix} 2 & 1 & 0 \\ 1 & -2 & 1 \\ 1 & 0 & i \end{bmatrix}, \quad \textit{and} \quad A_5 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & -i \end{bmatrix}.$$

In Figure 2.2.5, the sets $\mathcal{D}_{i,j}(A_4)$, where $i, j \in \{1, 2, 3\}$, $i \neq j$, are shown by thin lines, while Dashnic-Zusmanovich's set $\mathcal{D}(A_4)$ is shaded. Figure 2.2.6 illustrates the same localization set for the matrix A_5. The set $\mathcal{D}(A_5)$ is again shaded, while the boundary of the set $\mathcal{D}_1(A_5)$ is given with a thick line. Finally, Figure 2.2.7 illustrates the relationship between Dashnic-Zusmanovich's set $\mathcal{D}(A_5)$ (shaded) and Brauer's set $\mathcal{K}(A_5)$ (with a thick boudary).

As before, using Varga's Equivalence Principle we can state the result of Theorem 1.4.1, in terms of eigenvalue inclusion sets. This leads us to the concept of the minimal Geršgorin set, attributed to the family of positive diagonal matrices, in this case $\mathbb{X}_{DZ}$.

Theorem 2.2.14. (Dashnic-Zusmanovich) *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, then the minimal Geršgorin set attributed to the family $\mathbb{X}_{DZ}$, given by (1.4.15), is equal to the Dashnic-Zusmanovich's set $\mathcal{D}(A)$, i.e.,*

$$\mathcal{D}(A) = \bigcap_{X \in \mathbb{X}_{DZ}} \Gamma^X(A) = \Gamma^{\mathbb{X}_{DZ}}(A). \tag{2.2.26}$$

So, for the class of DZ-SDD matrices, it is possible to explicitly express the minimal Geršgorin set attributed to this class, given in (2.2.26). In other words, although the minimal Geršgorin set attributed to the family $\mathbb{X}_{DZ}$ is, by definition, an intersection of

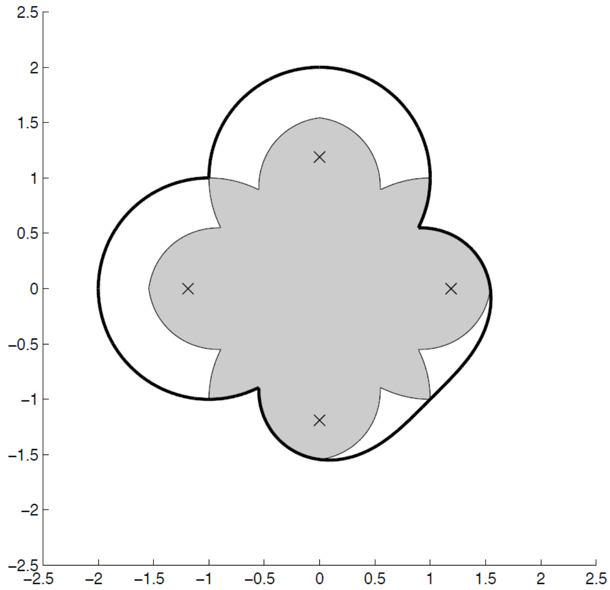

Figure 2.2.6: Sets $\mathcal{D}(A_5)$ and $\mathcal{D}_1(A_5)$ for the matrix A_5 of the Example 2.2.13
(Skupovi $\mathcal{D}(A_5)$ i $\mathcal{D}_1(A_5)$ za matricu A_5 iz Primera 2.2.13)

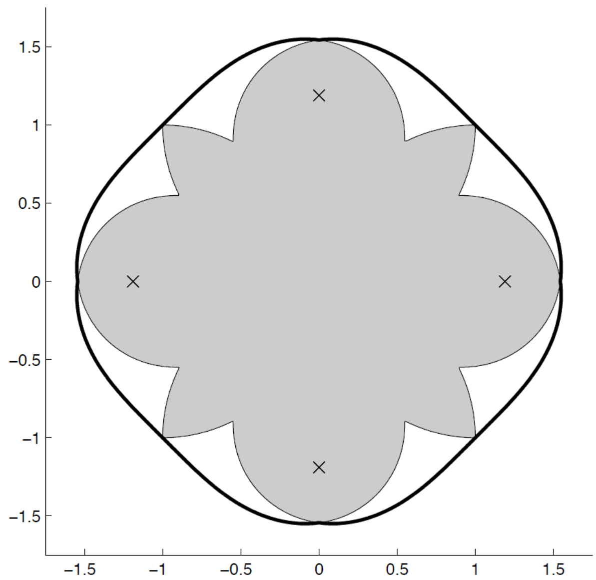

Figure 2.2.7: The sets $\mathcal{D}(A_5)$ and $\mathcal{K}(A_5)$ for the matrix A_5 of the Example 2.2.13
(Skupovi $\mathcal{D}(A_5)$ and $\mathcal{K}(A_5)$ za matricu A_5 iz primera 2.2.13)

continuum many compact sets, it can be expressed as a *finite* intersection of some compact sets in the complex plane. These sets depend solely of the matrix entries. More details concerning the geometrical interpretation of this result could be found in [20].

The eigenvalue localization result based on the class of S-SDD matrices, i.e., CKV-SDD matrices, was given in the Cvetković, Kostić and Varga [19]gave in the same paper [19], together with a nonsingularity result.

Theorem 2.2.15. (**Cvetković-Kostić-Varga**) *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, be an arbitrary matrix, and $\lambda \in \sigma(A)$ be its arbitrary eigenvalue. Then, for every nonempty subset of indices $S \subseteq N$, there exist indices $i \in S$, and $j \in \overline{S} := N \setminus S$, such that*

$$\lambda \in \Gamma_i^S(A) := \left\{ z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq r_i^S(A) \right\}, \tag{2.2.27}$$

or

$$\lambda \in V_{i,j}^S(A) := \left\{ z \in \mathbb{C} \,:\, (|z - a_{i,i}| - r_i^S(A)) \cdot (|z - a_{j,j}| - r_j^{\overline{S}}(A)) > r_i^{\overline{S}}(A) r_j^S(A) \right\}. \tag{2.2.28}$$

Therefore, for every nonempty subset of indices $S \subseteq N$,

$$\sigma(A) \subseteq \mathcal{C}^S(A) := [\bigcup_{i \in S} \bigcup_{j \in \overline{S}} V_{i,j}^S(A)] \bigcup [\bigcup_{i \in S} \Gamma_i^S(A)], \tag{2.2.29}$$

and consequently

$$\sigma(A) \subseteq \mathcal{C}(A) := \bigcap_{\emptyset \neq S \subseteq N} \mathcal{C}^S(A). \tag{2.2.30}$$

Having that the class of S-SDD matrices is a positively homogenous DD-type class, we can, as before, apply the Isolation Principle.

Theorem 2.2.16. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and $\mathcal{C}_S(A) = U \cup V$, for some nonempty $S \subseteq N$, then the set U contains exactly $|\{i \in N \,:\, a_{i,i} \in U\}|$ eigenvalues of the matrix A.*

Again, we have the equivalent form of Theorem 1.4.2 in terms of eigenvalue localization, which leads us to the concept of the minimal Geršgorin set. Originally, the following result was proved in [14].

Theorem 2.2.17. (Cvetković-Kostić) *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ and an arbitrary nonempty subset of indices $S \subseteq N$, then the minimal Geršgorin set attributed to the family $\mathbb{X}_S$, given by (1.4.18), is equal to the set $\mathcal{C}^S(A)$, i.e.,*

$$\mathcal{C}^S(A) = \bigcap_{X \in \mathbb{X}_S} \Gamma^X(A) = \Gamma^{\mathbb{X}_S}(A), \tag{2.2.31}$$

and, consequently, the minimal Geršgorin set, attributed to the family $\mathbb{X}_{CKV}$, given by (1.4.19), is equal to the set $\mathcal{C}(A)$, i.e.,

$$\mathcal{C}(A) = \bigcap_{X \in \mathbb{X}_{CKV}} \Gamma^X(A) = \Gamma^{\mathbb{X}_{CKV}}(A). \tag{2.2.32}$$

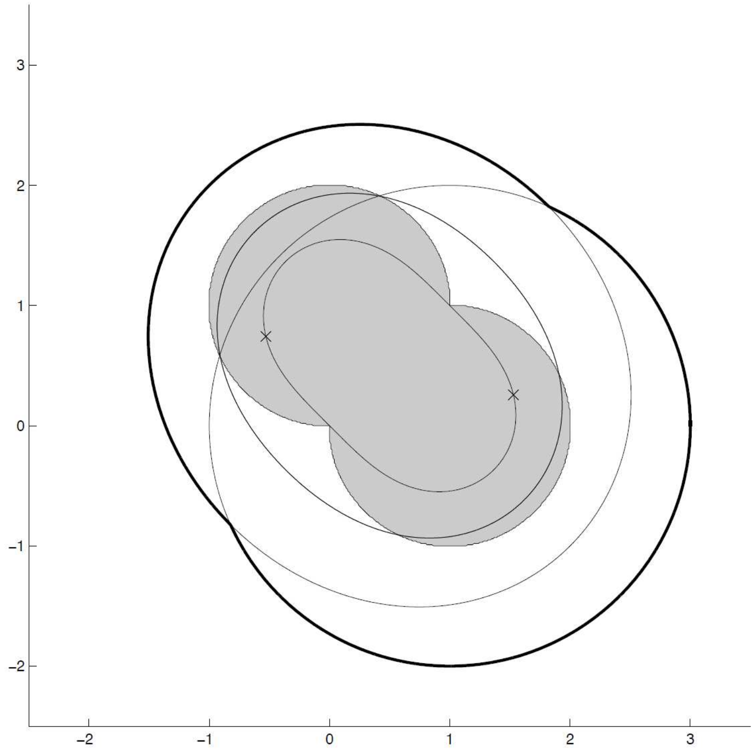

Figure 2.2.8: Dashnic-Zusmanovich's set for the matrix A_6 of the Example 2.2.18
(Dašnjic-Zusmanovič skup za matricu A_6 iz Primera 2.2.18)

The following example illustrates this localization set.

Example 2.2.18. *Let*

$$A_6 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & i & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & i \end{bmatrix}.$$

In Figure 2.2.8, the set $\mathcal{C}(A_6)$ is shaded, the set $\mathcal{C}^S(A_6)$, for $S = \{1,3\}$, has the thick boundary, while the sets $V_{i,j}^S(A_6)$, for $i \in \{1,3\}$ and $j \in \{2,4\}$, and $\Gamma_1^S(A_6)$ have thin boundaries. Eigenvalues of the matrix A_6 are marked by "$\times$". Figure 2.2.9 illustrates the relationship between CKV set, given in (2.2.30) with other localizations. There, the set $\mathcal{C}(A_6)$ is shaded, while the Dashnic-Zusmanovich set, the Brauer set and the Geršgorin set have, respectively, thicker and thicker boundaries. Again, eigenvalues are marked by "$\times$".

First, observe that the form of the set $V_{i,j}^S(A)$ may be asymmetric, contrary to Brauer's Ovals of Cassini which are invariant to the interchange of the foci. Thus, it is expected that this freedom leads to an improvement of the localization set. Apart from this, in order to construct the set $\mathcal{C}(A)$, we use more information than in the case of the set $\mathcal{K}(A)$. The following theorem establishes the general relationship between mentioned localizations; for more details, see [19] and [51], Chapter 3.3.

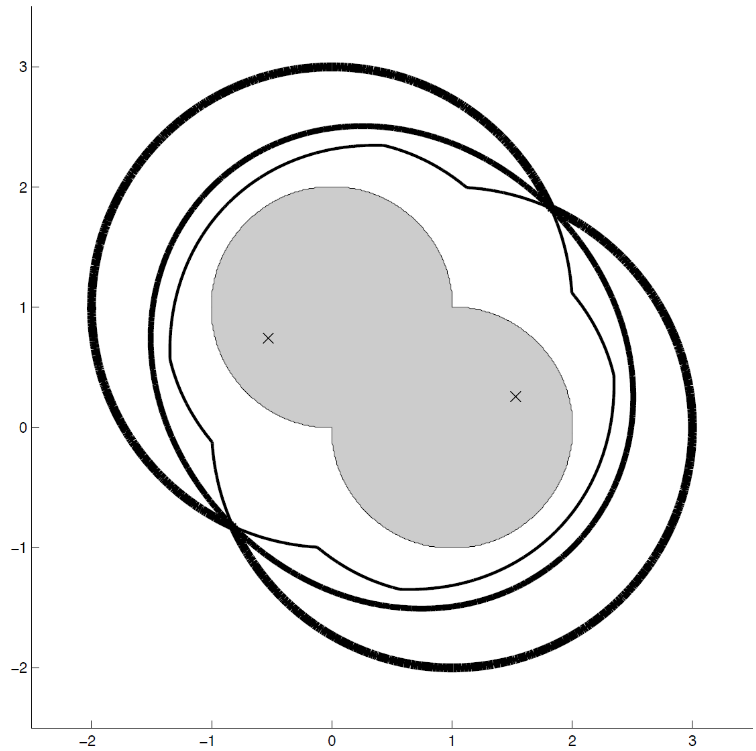

Figure 2.2.9: Dashnic-Zusmanovich's sets $\mathcal{D}(A_5)$ and $\mathcal{D}_1(A_5)$ for the matrix A_5 of the Example 2.2.13

(Dašnjic-Zusmanovič skupovi $\mathcal{D}(A_5)$ i $\mathcal{D}_1(A_5)$ za matricu A_5 iz Primera 2.2.18)

Theorem 2.2.19. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, be an arbitrary matrix, and let the set $\Gamma(A)$ be given by (2.1.2), set $\mathcal{K}(A)$ by (2.2.8), the set $\mathcal{D}_i(A)$ by (2.2.25), the set $\mathcal{D}(A)$ by (2.2.24), the set $\mathcal{C}^S(A)$ by (2.2.29), and the set $\mathcal{C}(A)$ by (2.2.30). Then,*

- $\mathcal{C}^{\{i\}}(A) = \mathcal{D}_i(A) \subseteq \Gamma(A), \quad (i \in N)$,
- $\mathcal{C}^S(A) \subseteq \Gamma(A), \quad (S \subseteq N)$, *and, consequently,*
- $\mathcal{C}(A) \subseteq \mathcal{D}(A) \subseteq \mathcal{K}(A) \subseteq \Gamma(A)$.

Moreover, there exist matrices $P, Q, R \in \mathbb{C}^{n,n}$, so that

- $\mathcal{C}(P) \nsubseteq \mathcal{B}(P)$, *and* $\mathcal{B}(P) \nsubseteq \mathcal{C}(P)$,
- $\mathcal{D}_i(Q) \nsubseteq \mathcal{K}(Q)$, *and* $\mathcal{K}(Q) \nsubseteq \mathcal{D}_i(Q)$, *for some* $i \in N$,
- $\mathcal{C}^S(R) \nsubseteq \mathcal{K}(R)$, *and* $\mathcal{K}(R) \nsubseteq \mathcal{C}^S(R)$, *for some* $S \subseteq N$.

2.2.4 Ostrowski sets

Here we start with nonsingularity results given in Subsection 1.2.4. As a consequence of Varga's Equivalence Principle, we have two rather well-known eigenvalue localization theorems.

To conclude this section on Geršgorin-type theorems, we consider eigenvalue localization sets that are derived by Varga's Equivalence Principle from the corresponding nonsingularity results from Subsection 1.2.4: Theorems 1.2.21 and 1.2.22, and their characterizations given in Theorems 1.2.26 and 1.2.27, respectively.

Theorem 2.2.20. *Given an arbitrary $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, let λ be one of its eigenvalues. Then, for an arbitrary $\alpha \in [0,\, 1]$, there exists an index $i \in N$ such that $|\lambda - a_{i,i}| \leq \alpha r_i(A) + (1-\alpha)c_i(A)$. In other words, for an arbitrary $\alpha \in [0,\, 1]$,*

$$\sigma(A) \subseteq \mathcal{A}^1_\alpha(A) := \bigcup_{i \in N} \mathcal{A}^1_{\alpha,i}(A), \tag{2.2.33}$$

where $\mathcal{A}^1_{\alpha,i}(A) := \{z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq \alpha r_i(A) + (1-\alpha)c_i(A)\}$. Consequently,

$$\sigma(A) \subseteq \mathcal{A}^1(A) := \bigcap_{\alpha \in [0\,,\,1]} \mathcal{A}^1_\alpha(A). \tag{2.2.34}$$

Theorem 2.2.21. *Given an arbitrary $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, let λ be one of its eigenvalues. Then, for an arbitrary $\alpha \in [0,\, 1]$, there exists an index $i \in N$ such that $|\lambda - a_{i,i}| \leq (r_i(A))^\alpha (c_i(A))^{1-\alpha}$. In other words, for an arbitrary $\alpha \in [0,\, 1]$,*

$$\sigma(A) \subseteq \mathcal{A}^2_\alpha(A) := \bigcup_{i \in N} \mathcal{A}^2_{\alpha,i}(A), \tag{2.2.35}$$

where $\mathcal{A}^2_{\alpha,i}(A) := \{z \in \mathbb{C} \,:\, |z - a_{i,i}| \leq (r_i(A))^\alpha (c_i(A))^{1-\alpha}\}$. Consequently,

$$\sigma(A) \subseteq \mathcal{A}^2(A) := \bigcap_{\alpha \in [0\,,\,1]} \mathcal{A}^2_\alpha(A). \tag{2.2.36}$$

We will refer to the localization sets introduced in the previous two theorems as α_1**-sets** and α_2**-sets**, respectively. To be more precise, these terms will refer to localization sets given in (2.2.33) and (2.2.35), while the sets in (2.2.34) and (2.2.36) we call α_1**-minimal set** and α_2**-minimal set**, respectively.

As before, since the classes of α_1-SDD and α_2-SDD matrices are both positively homogenous, the Isolation Principle holds in both cases.

Theorem 2.2.22. *Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, and $k \in \{1, 2\}$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and $\mathcal{A}^k(A) = U \cup V$, then, the set U contains exactly $|\{i \in N : a_{i,i} \in U\}|$ eigenvalues of the matrix A.*

As we have seen, (1.2.39) implies that every α_1-SDD matrix is also an α_2-SDD matrix. In the same way, it is clear that for an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $\mathcal{A}^2(A) \subseteq \mathcal{A}^1(A)$. Thus, if we want to obtain a better localization set for the same "price", the proper choice is always $\mathcal{A}^2(A)$!

Example 2.2.23. *For the given matrix*

$$A_7 = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & -1 \end{bmatrix},$$

Figure 2.2.10 shows, from upper left to lower right corner, α_2-localization set $\mathcal{A}^2_\alpha(A_7)$ for the following values of the parameter $\alpha = 1,\ 0.8,\ 0.6,\ 0.4,\ 0.2,\ 0$, respectively. In fact, this represents a "step-by-step" transformation from $\Gamma(A_7)$ to $\Gamma(A_7^T)$. Eigenvalues of the matrix A_7 are everywhere marked by "$\times$".

Naturally, for different values of parameter α, we obtain different localization sets, and, as this example shows, α_2-sets for different values of the parameter α stand in the general position, i.e., neither one of them is a subset of the other one. Taking the intersection (2.2.36) over all possible values of the parameter, we obtain, of course, the best possible localization set in this direction.

While the α_2-set, for a fixed value of the parameter, is essentially as easy to draw as the Geršgorin set, this is not the case for α_2-minimal set. Given in the form of intersection of a continuum of many sets, it obviously raises a question how can we compute it, in general. But, starting from Theorem 1.2.27, on the characterization of α_2-matrices, and using Varga's Equivalence Principle, we obtain a different form of this set, which is much more useful.

Theorem 2.2.24. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, and let λ be one of its eigenvalues. Then, there exists an index $i \in N$ such that $|\lambda - a_{i,i}| \leq \min\{r_i(A), c_i(A)\}$, or, there exist $i \in \mathcal{R}^*(A)$ and $j \in \mathcal{C}^*(A)$, where the sets $\mathcal{R}^*(A)$ and $\mathcal{C}^*(A)$ are given in (1.2.56) and (1.2.57), respectively, such that*

$$\frac{|\lambda - a_{i,i}|}{c_i(A)} \left(\frac{|\lambda - a_{j,j}|}{c_j(A)} \right)^{\log_{\frac{c_j(A)}{r_j(A)}} \frac{r_i(A)}{c_i(A)}} \leq 1. \tag{2.2.37}$$

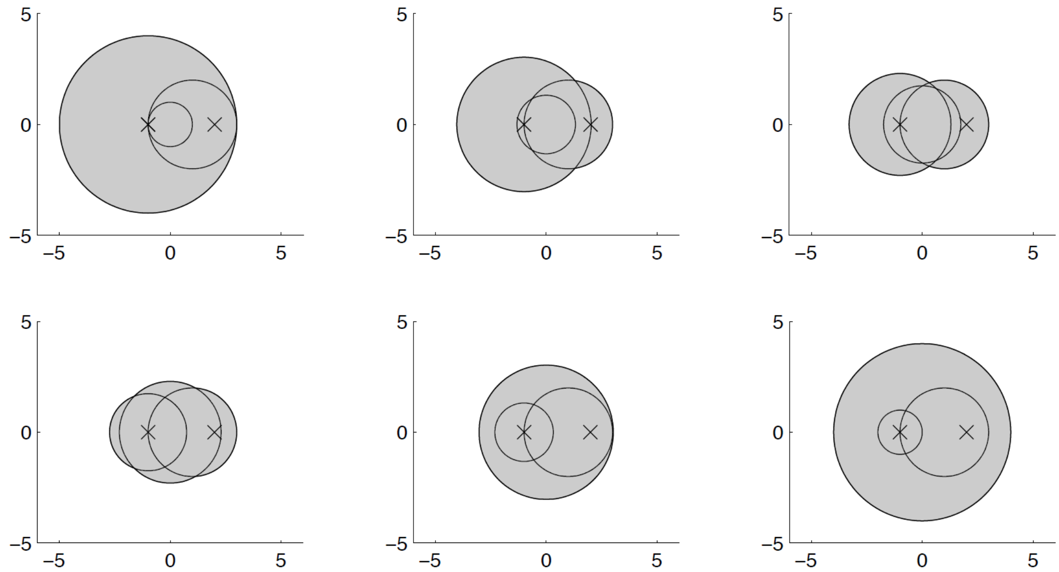

Figure 2.2.10: α_2-localization set $\mathcal{A}^2_\alpha(A_7)$ for the matrix A_7 of the Example 2.2.23 for the following values of the parameter $\alpha = 1,\ 0.8,\ 0.6,\ 0.4,\ 0.2,\ 0$, from upper left to lower right corner, respectively

(α_2-skupovi za lokalizaciju $\mathcal{A}^2_\alpha(A_7)$ matrice A_7 iz Primera 2.2.23 za sledeće vrednosti parametra $\alpha = 1,\ 0.8,\ 0.6,\ 0.4,\ 0.2,\ 0$, od gornjeg levog ugla do donjeg desnog ugla, redom)

Thus, we have that

$$\sigma(A) \subset \mathcal{A}^2(A) := \Gamma^m(A) \bigcup \Lambda^2(A), \tag{2.2.38}$$

where $\Gamma^m(A)$ is given by (2.1.13),

$$\Lambda^2(A) := \bigcup_{\substack{i \in \mathcal{R}^*(A) \\ j \in \mathcal{C}^*(A)}} \Lambda^2_{i,j}(A), \quad \textit{and} \tag{2.2.39}$$

$$\Lambda^2_{i,j}(A) := \{ z \in \mathbb{C} : \frac{|z - a_{i,i}|}{c_i(A)} \left(\frac{|z - a_{j,j}|}{c_j(A)} \right)^{\log_{\frac{c_j(A)}{r_j(A)}} \frac{r_i(A)}{c_i(A)}} \leq 1 \}, \tag{2.2.40}$$

for $i \in \mathcal{R}^*$, and $j \in \mathcal{C}^*$.

As we can see, Theorem 1.2.27 allows us to represent the α_2-minimal set of an arbitrary matrix as the *finite* union of compact sets of the complex plane, and, therefore, it *is possible* to compute it, in general.

How useful this localization can be is illustrated in the following example.

Example 2.2.25. *Let*

$$A_8 = \begin{bmatrix} 5 & 4 \\ 1 & 5 \end{bmatrix}, \quad A_9 = \begin{bmatrix} 1 & 0 & 0.3 & 0.4 \\ 0 & 0 & 0 & 0.6 \\ 0.5 & 0 & -1 & 0.4 \\ 0 & 0.5 & 0 & -i \end{bmatrix}, \quad \textit{and } A_{10} = \begin{bmatrix} 1 & 0 & 0.3 & 0.65 \\ 0 & i & 0.5 & 0.3 \\ 0.5 & 0 & -1 & 0 \\ 0 & 0.5 & 0 & -i \end{bmatrix}.$$

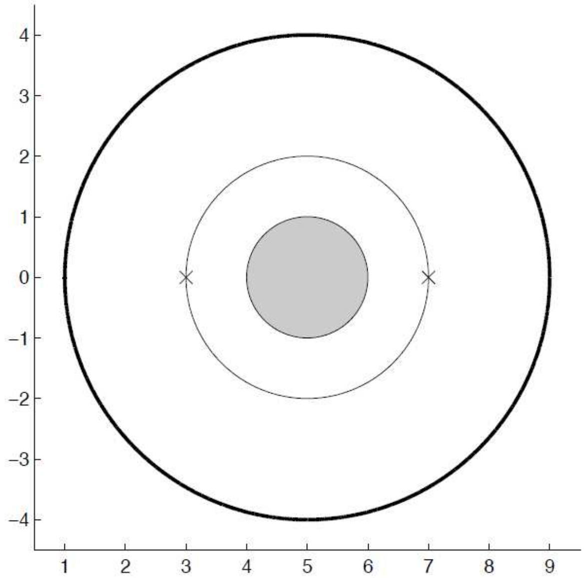

Figure 2.2.11: The α_2-minimal set for the matrix A_8 of the Example 2.2.25
(α_2-minimalni skup za matricu A_8 iz Primera 2.2.25)

Figures 2.2.11, 2.2.12 and 2.2.13, show the sets $\Gamma(A_i) \cap \Gamma(A_i^T)$ by the thick line, sets $\Lambda^2(A_i)$ with a thin line, and the sets $\Gamma^m(A_i)$ shaded, for $i = 8, 9, 10$, respectively. Again, the eigenvalues are marked by "$\times$".

As noted in Subsection 2.1.2, the matrix A_8, of the previous example, shows that, in general, given a matrix A, the set $\Gamma^m(A)$ does *not* have to contain some, or even all of the spectra of A. Thus, the role of the set $\Lambda^2(A)$ is essential. Actually, this set answers the question proposed at the end of Subsection 2.1.2.

Figures 2.2.12 and 2.2.13 illustrate that, sometimes, the improvement obtained by the use of the α_2-minimal set, can be really significant.

Besides the α_2-minimal set, we can also use α_1-minimal set to answer the question of Subsection 2.1.2. Namely, a consequence of the Theorem 1.2.26 is the following one.

Theorem 2.2.26. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, with $n \geq 2$, and let λ be one of its eigenvalues. Then, there exists an index $i \in N$ such that $|\lambda - a_{i,i}| \leq \min\{r_i(A), c_i(A)\}$, or, there exist $i \in \mathcal{R}(A)$ and $j \in \mathcal{C}(A)$, so that*

$$|\lambda - a_{i,i}|\,(c_j(A) - r_j(A)) + |\lambda - a_{j,j}|\,(r_i(A) - c_i(A)) \leq c_j(A)r_i(A) - c_i(A)r_j(A). \quad (2.2.41)$$

Thus, we have that

$$\sigma(A) \subset \mathcal{A}^1(A) := \Gamma^m(A) \bigcup \Lambda^1(A), \quad (2.2.42)$$

where $\Gamma^m(A)$ is given by (2.1.13),

$$\Lambda^1(A) := \bigcup_{\substack{i\in\mathcal{R}(A)\\ j\in\mathcal{C}(A)}} \Lambda^1_{i,j}(A), \quad \textit{and} \quad (2.2.43)$$

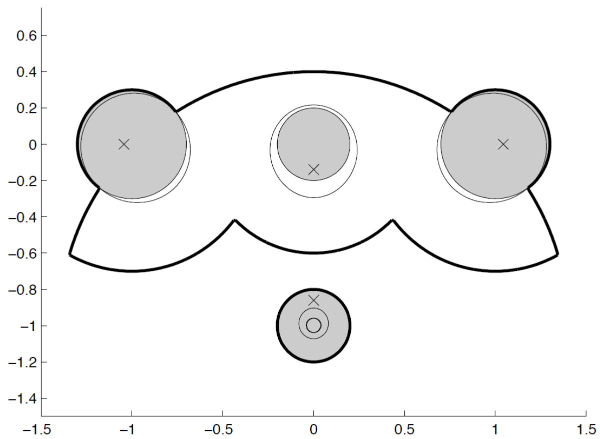

Figure 2.2.12: The α_2-minimal set for the matrix A_9 of the Example 2.2.25
(α_2-minimalni skup za matricu A_9 iz Primera 2.2.25)

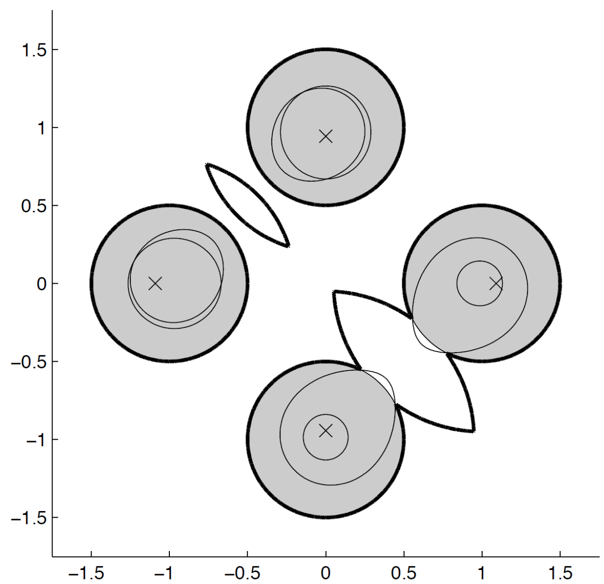

Figure 2.2.13: The α_2-minimal set for the matrix A_{10} of the Example 2.2.25
(α_2-minimalni skup za matricu A_{10} iz Primera 2.2.25)

$$\Lambda^1_{i,j}(A) := \{z \in \mathbb{C} : |z - a_{i,i}|\,(c_j(A) - r_j(A)) + \tag{2.2.44}$$
$$|z - a_{j,j}|\,(r_i(A) - c_i(A)) \leq c_j(A)r_i(A) - c_i(A)r_j(A)\},$$

for $i \in \mathcal{R}$, and $j \in \mathcal{C}$.

Of course, we have seen that, in general, the set (2.2.43) contains the set of (2.2.39), so, it gives worse estimates of spectra then the previous one.

For both of these localization sets, it remains to analyze the "cost" one has pay to obtain them. Obviously, given a matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, in both cases we start with n disks. The number of the other sets that have to be plotted depends of the structure of the matrix. Namely, if the matrix is real symmetric, this number is zero, but then, both of these sets collapse to the original Geršgorin set. On the other hand, if the symmetry is ruined in only one place, this number is one. Depending of the number of indices where row sums dominate over column sums, and vice versa, this number can be at most $0.25n^2$, if n is even, or $0.25(n^2 - 1)$, if n is odd. Thus, it could be wise to use this set when matrices have "unbalanced" dominance of row sums over column sums, or vice versa.

For the matrices A_9 and A_{10}, the additional number of the sets that has to be determined is 3 and 4, respectively.

As before, it is interesting to compare these eigenvalue localizations to the ones defined in the previous sections. The first part of the following proposition can be easily obtained from the fact that, taking $\alpha = 1$, the set in (2.2.35) becomes the Geršgorin set of the corresponding matrix, and, taking $\alpha = 0$, it becomes the Geršgorin set of the transpose of the corresponding matrix. The second part follows directly by simple examples.

Theorem 2.2.27. *Let $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, be an arbitrary matrix, and let the set $\Gamma(A)$ be given by (2.1.2), the set $\mathcal{B}(A)$ by (2.2.21), the set $\mathcal{C}^S(A)$ by (2.2.29), the set $\mathcal{C}(A)$ by (2.2.30), the set $\mathcal{A}^1(A)$ by (2.2.34), and the set $\mathcal{A}^2(A)$ by (2.2.36). Then,*

$$\mathcal{A}^2(A) \subseteq \mathcal{A}^1(A) \subseteq \Gamma(A) \cap \Gamma(A^T). \tag{2.2.45}$$

Moreover, there exist matrices $P, Q \in \mathbb{C}^{n,n}$, so that

- $\mathcal{A}^2(P) \nsubseteq \mathcal{B}(P)$, *and* $\mathcal{B}(P) \nsubseteq \mathcal{A}^1(P)$, *and*
- $\mathcal{A}^2(Q) \nsubseteq \mathcal{C}(Q)$, *and* $\mathcal{C}(Q) \nsubseteq \mathcal{A}^1(Q)$.

Therefore, α_1 and α_2-(minimal) sets stand each in a general position, with the observed extensions of the Geršgorin set.

2.3 Minimal Geršgorin Sets

In the last section of the second chapter, we will discus, in more detail, the minimal Geršgorin set, which was introduced at the beginning of this chapter. The name *minimal Geršgorin set* was first used by Varga in 1965, [47], where he proved many of its properties. Here, briefly, we will present some of them, together with the results from [53] and [50], and in the form in which they appear in [51]. We will conclude the chapter with the recent results on the computation of the minimal Geršgorin set, which are due to Varga, Cvetković and Kostić, [52].

The minimal Geršgorin set for the matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ is a compact set in the complex plane given by:

$$\Gamma^{\mathbb{D}}(A) := \bigcap_{X \in \mathbb{D}} \Gamma^X(A), \tag{2.3.1}$$

where

$$\begin{cases} \Gamma_i^X(A) := \{z \in \mathbb{C} \,:\, |z - a_{i,i}| \le r_i^{\mathbf{x}}(A)\}, \quad (i \in N), \\ \Gamma^X(A) := \bigcup_{i \in N} \Gamma_i^X(A), \end{cases} \tag{2.3.2}$$

and

$$r_i^{\mathbf{x}}(A) := r_i(X^{-1}AX) = \sum_{j \in N \setminus \{i\}} \frac{|a_{i,j}| x_j}{x_i} \quad (i \in N,\ \mathbf{x} > \mathbf{0}). \tag{2.3.3}$$

As Corollary 2.1.5 states, the minimal Geršgorin set is an eigenvalue inclusion set, i.e., for an arbitrary matrix A, $\sigma(A) \subseteq \Gamma^{\mathbb{D}}(A)$.

As we have seen in the introduction of Section 2.2, having a subclass $\mathbb{K}$ of nonsingular H-matrices, i.e., GDD matrices, and using the Varga's Equivalence Principle, for a given matrix A, we derive an eigenvalue localization set $\Theta^{\mathbb{K}}(A)$, given in 2.2.1. Moreover, we have seen that this set is the minimal Geršgorin set, attributed to the family $\mathbb{X}^{\mathbb{K}}$, given in (1.4.2), i.e., $\Theta^{\mathbb{K}}(A) = \Gamma^{\mathbb{X}^{\mathbb{K}}}(A) := \bigcap_{X \in \mathbb{X}^{\mathbb{K}}} \Gamma^X(A)$.

On the other hand, we have also seen that the family of positive diagonal matrices $\mathbb{X} \subseteq \mathbb{D}$ generates the subclass of H-matrices $\mathbb{K}^{\mathbb{X}}$, and then, again, we have that, for a given A, the corresponding localization set $\Theta^{\mathbb{K}^{\mathbb{X}}}(A)$ is the minimal Geršgorin set attributed to the family $\mathbb{X}$, i.e., $\Theta^{\mathbb{K}^{\mathbb{X}}}(A) = \Gamma^{\mathbb{X}}(A)$.

Since, $\mathbb{K}^{\mathbb{D}} = \mathbb{H}$ is the class of all nonsingular H-matrices, we have that $\Theta^{\mathbb{K}^{\mathbb{D}}}(A) = \Gamma^{\mathbb{D}}(A)$. Thus, the proposition that all GDD matrices are nonsingular is equivalent to the proposition that the minimal Geršgorin set is the eigenvalue inclusion set. Moreover, this property implies that the minimal Geršgorin set is the best possible Geršgorin-type localization set, as it was defined in Section 2.2. In other words, for every $\mathbb{K} \subseteq \mathbb{H}$, $\Gamma^{\mathbb{D}}(A) \subseteq \Theta^{\mathbb{K}}(A)$, where A is an arbitrary matrix. In fact, this conclusion unifies, in a simple way, some known results, like the one in Theorem 2.14 in [51], about the relationship of the minimal Geršgorin set and other known eigenvalue inclusion sets.

In fact, apart from this property, as we will see, there is also another reason to call the set from (2.3.1) *minimal*.

2.3.1 Sharpness and geometry of the minimal Geršgorin set

Given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, we define

$$\Omega(A) := \{B = [b_{i,j}] \in \mathbb{C}^{n,n} : b_{i,i} = a_{i,i} \text{ and } |b_{i,j}| = |a_{i,j}| \text{ for } i \neq j \ (i, j \in N)\}, \qquad (2.3.4)$$

and call it the **equimodular family** of matrices attributed to the matrix A. In a similar way, we define

$$\widehat{\Omega}(A) := \{B = [b_{i,j}] \in \mathbb{C}^{n,n} : b_{i,i} = a_{i,i} \text{ and } |b_{i,j}| \leq |a_{i,j}| \text{ for } i \neq j \ (i, j \in N)\}, \qquad (2.3.5)$$

and call it the **extended equimodular family** of matrices attributed to the matrix A. Then, obviously,

$$\sigma(\Omega(A)) \subseteq \sigma(\widehat{\Omega}(A)) \subseteq \Gamma^{\mathbb{D}}(A), \qquad (2.3.6)$$

where A is an arbitrary matrix, and $\sigma(\Omega(A))$ and $\sigma(\widehat{\Omega}(A))$ are defined in (2.2.16).

That inclusions in (2.3.6) are in fact, equalities, it was proved by Varga in [47].

Theorem 2.3.1. (Varga) *For any* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$*, then*

$$\partial\Gamma^{\mathbb{D}}(A) \subseteq \sigma(\Omega(A)) \subseteq \sigma(\widehat{\Omega}(A)) = \Gamma^{\mathbb{D}}(A). \qquad (2.3.7)$$

This sharpness can be also expressed in the following way: Given a matrix A, every point on the boundary of the minimal Geršgorin set $\Gamma^{\mathbb{D}}(A)$ is an eigenvalue of a matrix from the family $\Omega(A)$, and every point of the minimal Geršgorin set $\Gamma^{\mathbb{D}}(A)$ is an eigenvalue of the matrix from the family $\widehat{\Omega}(A)$. In other words, the eigenvalues of the matrices from $\Omega(A)$ *fill up* the boundary of the minimal Geršgorin set, while the the eigenvalues of the matrices from $\widehat{\Omega}(A)$ *fill up* the entire minimal Geršgorin set.

In order to prove the previous theorem, Varga introduced some useful concepts for exploring the properties of the minimal Geršgorin set. Namely, given an arbitrary matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, and a complex number $z \in \mathbb{C}$, we define the associated matrix $Q(z) = [q_{i,j}(z)] \in \mathbb{R}^{n,n}$ by

$$q_{i,i}(z) := -|z - a_{i,i}|, \text{ and } q_{i,j}(z) := |a_{i,j}|, \text{ for } i \neq j \ (i, j \in N). \qquad (2.3.8)$$

Taking

$$\mu(z) := \max_{i \in N} |z - a_{i,i}|, \qquad (2.3.9)$$

we obtain that the matrix $B(z) := [b_{i,j}(z)] \in \mathbb{R}^{n,n}$, defined by

$$b_{i,i}(z) := \mu(z) - |z - a_{i,i}|, \text{ and } b_{i,j}(z) := |a_{i,j}|, \ i \neq j \ (i, j \in N), \qquad (2.3.10)$$

satisfies

$$B(z) = Q(z) + \mu(z) I_n. \qquad (2.3.11)$$

Here, $B(z)$ is a nonnegative matrix in $\mathbb{R}^{n,n}$. Therefore, by the Perron-Frobenius theory of nonnegative matrices, (c.f. Theorem C.2 from [51]), the spectral radius of the matrix

$B(z)$, $\rho\big(B(z)\big)$ is its nonnegative eigenvalue, and there exists a nonnegative eigenvector $\mathbf{y} \geq \mathbf{0}$, such that $B\mathbf{y} = \rho(B)\mathbf{y}$. Even more, $\rho\big(B(z)\big)$ can be characterized as

$$\rho\big(B(z)\big) = \inf_{\mathbf{x}>0} \left\{ \max_{i\in N} \left\{ \big(B(z)\mathbf{x}\big)_i / x_i \right\} \right\}. \tag{2.3.12}$$

Thus, if we set

$$\nu_A(z) := \rho(B(z)) - \mu(z) \ \ (\text{all} \ \ z \in \mathbb{C}), \tag{2.3.13}$$

then $\nu_A(z)$ is a real-valued function, defined for all $z \in \mathbb{C}$. Moreover, it can be expressed as

$$\nu_A(z) = \inf_{\mathbf{x}>0} \left\{ \max_{i\in N} \left\{ \big(Q(z)\mathbf{x}\big)_i / x_i \right\} \right\} = \inf_{\mathbf{x}>0} \left\{ \max_{i\in N} \left\{ r_i^{\mathbf{x}}(A) - |z - a_{i,i}| \right\} \right\}. \tag{2.3.14}$$

Now, using the equality (2.3.14), the following connection of the function $\nu_A(z)$ to the minimal Geršgorin set, $\Gamma^{\mathbb{D}}(A)$, from the work of Varga, [47] and [51], can be easily proved.

Theorem 2.3.2. *For any $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, then*

$$z \in \Gamma^{\mathbb{D}}(A) \text{ if and only if } \nu_A(z) \geq 0. \tag{2.3.15}$$

Furthermore, if $z \in \partial\Gamma^{\mathbb{D}}(A)$, then $\nu_A(z) = 0$, and, conversely, if $\nu_A(z) = 0$, and there exists a sequence of complex numbers $\{z_k\}_{k\in\mathbb{N}}$, with $\lim_{k\to\infty} z_k = z$, for which $\nu_A(z_k) < 0$, for all $k \in \mathbb{N}$, then $z \in \partial\Gamma^{\mathbb{D}}(A)$.

Another interesting consequence of the equality (2.3.14) is that the real-valued function ν_A of the complex argument z is *uniformly continuous* on $\mathbb{C}$, i.e., we can prove that,

$$\text{for any } z \text{ and } z' \text{ in } \mathbb{C},\ |\nu_A(z) - \nu_A(z')| \leq |z - z'|. \tag{2.3.16}$$

Therefore, the minimal Geršgorin set of the given matrix A can be expressed in terms of the uniformly continuous real-valued function ν_A. This property turned out to be very useful in many ways, as we shall present in the reminder of this section.

Before we continue, we will discuss the case when a given matrix A is irreducible. As it is widely known, the Perron-Frobenious theory for irreducible nonnegative matrices gives us stronger results. Concerning our discussion, the following properties of the matrices $B(z)$ and $Q(z)$ can be obtained from Theorem C.1 of [51].

Assuming that $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, $n \geq 2$, is irreducible, nonnegative matrix $B(z)$ is irreducible, too. Thus it possesses a *positive* real eigenvalue, $\rho(B(z))$, called the *Perron root* of $B(z)$, which is characterized as follows:

For any $\mathbf{x} > \mathbf{0}$ in $\mathbb{R}^{n,n}$, either

$$\min_{i\in N}\{\big(B(z)\mathbf{x}\big)_i / x_i\} < \rho(B(z)) < \max_{i\in N}\{\big(B(z)\mathbf{x}\big)_i / x_i\}, \tag{2.3.17}$$

or

$$B(z)\mathbf{x} = \rho(B(z))\mathbf{x}. \tag{2.3.18}$$

Moreover, from (2.3.17) and (2.3.18), for any $\mathbf{x} > \mathbf{0}$ in $\mathbb{R}^n$ and any $z \in \mathbb{C}$, either

$$\min_{i\in N}\{\big(Q(z)\mathbf{x}\big)_i/x_i\} < \nu_A(z) < \max_{i\in N}\{\big(Q(z)\mathbf{x}\big)_i/x_i\}, \tag{2.3.19}$$

or

$$Q(z)\mathbf{x} = \nu_A(z)\mathbf{x}, \tag{2.3.20}$$

where the last equation gives us that $\nu_A(z)$ is an eigenvalue of $Q(z)$.

To conclude this subsection, we give another interesting result of Varga, obtained in [51], that clarifies the geometrical structure of the minimal Geršgorin set.

Theorem 2.3.3. (Varga) *Given an irreducible matrix* $A = [a_{i,j}] \in \mathbb{C}^{n,n}$*, with* $n \geq 2$*, then, for every* $k \in N$*, and every* θ *with* $0 \leq \theta \leq 2\pi$*, there exists an* $\hat{\varrho}_k(\theta) \geq 0$ *such that the entire complex interval* $[a_{k,k} + te^{i\theta}]_{t=0}^{\hat{\varrho}_k(\theta)}$ *is contained in* $\Gamma^{\mathbb{D}}(A)$*, and, consequently,*

$$\bigcup_{\theta=0}^{2\pi}[a_{k,k} + te^{i\theta}]_{t=0}^{\hat{\varrho}_k(\theta)} \subseteq \Gamma^{\mathbb{D}}(A). \tag{2.3.21}$$

Namely, from the assumption that A is irreducible, it can be deduced that

$$\nu_A(a_{i,i}) > 0, \text{ for all } i \in N, \tag{2.3.22}$$

and, further, that, given any real number θ, with $0 \leq \theta < 2\pi$, there exists the largest number $\hat{\varrho}_i(\theta) > 0$, such that

$$\nu_A(a_{i,i} + \hat{\varrho}_i(\theta)e^{i\theta}) = 0, \text{ and } \nu_A(a_{i,i} + te^{i\theta}) \geq 0, \text{ for all } 0 \leq t < \hat{\varrho}_i(\theta), \tag{2.3.23}$$

so that the entire complex interval $[a_{i,i} + te^{i\theta}]_{t=0}^{\hat{\varrho}_i(\theta)}$ lies in $\Gamma^{\mathbb{D}}(A)$.

This implies that the set

$$\bigcup_{\theta=0}^{2\pi}[a_{i,i} + te^{i\theta}]_{t=0}^{\hat{\varrho}_i(\theta)} \tag{2.3.24}$$

is a **star-shaped** subset of $\Gamma^{\mathbb{D}}(A)$, for each $i \in N$, with

$$\nu_A(a_{i,i} + \hat{\varrho}_i(\theta)e^{i\theta}) \in \partial\Gamma^{\mathbb{D}}(A). \tag{2.3.25}$$

The set given in 2.3.21 we call the **star-shaped subset** of $\Gamma^{\mathbb{D}}(A)$, with respect to the point $a_{k,k}$. So, as expected, the minimal Geršgorin set is in a way "centered" in diagonal entries of the given matrix. How this concept can be used, is illustrated in the following subsection, where we discus the computation of the minimal Geršgorin set.

2.3.2 Computation of the minimal Geršgorin set

Unlike the Geršgorin set $\Gamma(A)$ of (2.1.2), or $\Gamma^X(A)$ of (2.3.2), the minimal Geršgorin set $\Gamma^{\mathbb{D}}(A)$ of (2.3.1) is not, in general, easy to determine numerically. The aim of this subsection is to find a *reasonable approximation* of $\Gamma^{\mathbb{D}}(A)$, with a finite number of calculations, which contains $\Gamma^{\mathbb{D}}(A)$, and for which a limited number of boundary points of this approximation are actual boundary points of $\Gamma^{\mathbb{D}}(A)$. The determination of these latter boundary points are then related to a famous sharpening, by Olga Taussky, Theorem 2.1.10, of the Geršgorin set of (2.1.2).

First, we can observe that if $\nu_A(z) = 0$, then $\det Q(z) = 0$, which follows directly from (2.3.20), since $\nu_A(z)$ is an eigenvalue of $Q(z)$. But, from Theorem 2.3.2, we can see that in order to determine the boundary points of the minimal Geršgorin set of the matrix A, we need to compute the complex values z, in which the determinant $Q(z)$ vanishes. Of course, computing the determinant is not, in general, an easy task, so, we will explore a different way to obtain the values on the boundary of the minimal Geršgorin set.

With the given irreducible matrix $A = [a_{i,j}]$ in $\mathbb{C}^{n,n}$, choose any j in N, and set $z = a_{j,j}$. Next, we assume that the nonnegative irreducible matrix $B(a_{j,j})$, which has at least one zero diagonal entry from (2.3.10), is a *primitive matrix*, as it is defined in Section 2.2. of [48].

This is, certainly, the case if some diagonal entry of $B(a_{j,j})$ is positive. More generally, if $B(a_{j,j})$ is not primitive, i.e., $B(a_{j,j})$ is cyclic of some index $p \geq 2$, then any simple shift of $B(a_{j,j})$, into $B(a_{j,j}) + \varepsilon I_n$, *is* primitive for each $\varepsilon > 0$.

With $B(a_{j,j})$ primitive, then, starting with an $\mathbf{x}^{(0)} > \mathbf{0}$ in $\mathbb{R}^n$, the power method gives convergent upper and lower estimates for $\rho(B(a_{j,j}))$, i.e., if $\mathbf{x}^{(m)} := B^m(a_{j,j})\mathbf{x}^{(0)}$ for all $m \geq 1$, then with $\mathbf{x}^{(m)} := [x_1^{(m)}, x_2^{(m)}, ..., x_n^{(m)}]^T$, we have

$$\underline{\lambda_m} := \min_{i \in N}\{\frac{x_i^{(m+1)}}{x_i^{(m)}}\} \leq \rho(B(a_{j,j})) \leq \max_{i \in N}\{\frac{x_i^{(m+1)}}{x_i^{(m)}}\} =: \overline{\lambda_m} \tag{2.3.26}$$

for all $m \geq 1$, with

$$\lim_{m\to\infty} \underline{\lambda_m} = \rho(B(a_{j,j})) = \lim_{m\to\infty} \overline{\lambda_m}. \tag{2.3.27}$$

In this way, using (2.3.11), (2.3.13) and (2.3.19), convergent upper and lower estimates of $\nu_A(a_{j,j})$ can be numerically obtained.

These estimations of $\nu_A(a_{j,j})$ do not need great accuracy for graphing purposes.

Next, assume, for convenience, that $\nu_A(a_{j,j}) > 0$ is accurately known, and select any real θ, with $0 \leq \theta < 2\pi$. The numerical goal now is to estimate the largest $\hat{\varrho}_j(\theta)$, with sufficient accuracy, where, from (2.3.13),

$$\nu_A(a_{j,j} + \hat{\varrho}_j(\theta)e^{i\theta}) = 0, \text{ with } \nu_A(a_{j,j} + (\hat{\varrho}_j(\theta) + \varepsilon)e^{i\theta}) < 0 \tag{2.3.28}$$

for all sufficiently small $\varepsilon > 0$. By definition, we then have that

$$a_{j,j} + \hat{\varrho}_j(\theta)e^{i\theta} \text{ is a boundary point of } \Gamma^{\mathbb{D}}(A). \tag{2.3.29}$$

This means, from the min-max conditions (2.3.19)-(2.3.20), that there is an $\mathbf{x} > \mathbf{0}$, in $\mathbb{R}^n$, such that

$$Q(a_{j,j} + \hat{\varrho}_j(\theta)e^{i\theta})\mathbf{x} = \mathbf{0}, \text{ where } \mathbf{x} = [x_1, x_2, ..., x_n]^T > \mathbf{0}. \tag{2.3.30}$$

Equivalently, on calling $a_{j,j}+\hat{\varrho}_j(\theta)e^{i\theta} =: z_j(\theta)$, we can express (2.3.30), using the definition of (2.3.8), as

$$|z_j(\theta) - a_{i,i}| = \sum_{k\in N\backslash\{i\}} |a_{i,k}|x_k/x_i, \quad (\text{all } i \in N), \tag{2.3.31}$$

which can be interpreted, from Theorem 2.1.10, as Olga Taussky's boundary result. What is, perhaps, more interesting, is that it is geometrically *unnecessary* to determine the components of the vector $\mathbf{x} > \mathbf{0}$ in $\mathbb{R}^n$, for which (2.3.31) is valid. This follows from the fact that knowing the boundary point $z_j(\theta)$ of $\Gamma^{\mathbb{D}}(A)$, and knowing each of the centers, $\{a_{i,i}\}_{i\in N}$, of the associated n Geršgorin disks, then all the circles of (2.3.31) can be directly drawn, without knowing the components of the vector $\mathbf{x}$.

We return to the numerical estimation of $\hat{\varrho}_j(\theta)$, which satisfies (2.3.28) - (2.3.30). Setting $z := a_{j,j}$ and $z' := a_{j,j} + \hat{\varrho}_j(\theta)e^{i\theta}$, we know, from (2.3.16), that

$$\hat{\varrho}_j(\theta) \geq \nu_A(a_{j,j}) > 0. \tag{2.3.32}$$

Consider, now, the number $\nu_A(a_{j,j} + \nu_A(a_{j,j})e^{i\theta})$. If this number is positive, then increase the number $\nu_A(a_{j,j})$ to $\nu_A(a_{j,j}) + \Delta, \Delta > 0$, until $\nu_A(a_{j,j} + (\nu_A(a_{j,j}) + \Delta)e^{i\theta})$ is negative, and apply a bisection search to the interval $[\nu_A(a_{j,j}), \nu_A(a_{j,j}) + \Delta]$ to determine $\hat{\varrho}_j(\theta)$, satisfying (2.3.28). (Again, as in the estimation of $\nu_A(a_{j,j})$, estimates of $\hat{\varrho}_j(\theta)$ do not need great accuracy for graphing purposes.) We remark that a similar bisection search, on z, can be directly applied to

$$\det Q\big(\nu_A(a_{j,j} + \hat{\varrho}_j(\theta)e^{i\theta})\big) = 0, \tag{2.3.33}$$

but this requires, however, the evaluation of an $n \times n$ determinant.

Another alternative to compute the values $\hat{\varrho}_j(\theta)$, i.e., the values $z_j(\theta)$, is to use the fact that ν_A is uniformly continuous, and to construct a sequence $\{\xi_k^\theta\}_{k\in\mathbb{N}}$ that will converge to the value $z_j(\theta)$.

Namely, given an index $j \in N$, we define the first element of a sequence to be the center of j-th star-shaped subset, i.e., $\xi_1 := a_{j,j}$. Then, for a fixed direction $0 \leq \theta < 2\pi$, we define the recursion

$$\xi_{k+1}^\theta := \xi_k^\theta + \nu_A(\xi_k^\theta)e^{i\theta}, \tag{2.3.34}$$

where $k \in \mathbb{N}$.

Since $\nu_A(\xi_k^\theta) > 0$, for $k = 1$, by induction we can prove, using the continuity of ν_A, that $\nu_A(\xi_k^\theta) \geq 0$, for all $k \in \mathbb{N}$, and hence, obtain that (2.3.34) is the monotone sequence on the direction θ. On the other hand, for $k \in \mathbb{N}$ $\nu_A(\xi_k^\theta) \geq 0$ implies, by Theorem 2.3.2, that $\xi_k^\theta \in \Gamma^{\mathbb{D}}(A)$, which is bounded set in $\mathbb{C}$. Therefore, we have obtained the convergent sequence $\{\xi_k^\theta\}_{k\in\mathbb{N}}$. It remains to assure that the limit will be the point z_j^θ. With $\lim_{k\to\infty} \xi_k^\theta =: \xi^\theta$, letting $k \to \infty$ in (2.3.34), we directly obtain that $\nu_A(\xi^\theta) = 0$. Thus, we only need to check if $\nu_A(\xi^\theta + \varepsilon e^{i\theta}) < 0$, for a reasonably small $\varepsilon > 0$. If this is true, than $z_j^\theta = \xi^\theta$. If not, we restart the sequence (2.3.34) by taking $\xi_1^\theta := \xi^\theta$.

To summarize, given an irreducible matrix $A = [a_{i,j}]$ in $\mathbb{C}^{n,n}$, our procedure for approximating its minimal Geršgorin set, $\Gamma^{\mathbb{D}}(A)$, is to first determine, with reasonable accuracy, the positive numbers $\{\nu_A(a_{j,j})\}_{j\in N}$, and then, again with reasonable accuracy, to determine a few boundary points $\{\omega_k\}_{k=1}^m$ of $\Gamma^{\mathbb{D}}(A)$. For each such boundary point ω_k of $\Gamma^{\mathbb{D}}(A)$, there is an associated Geršgorin set, consisting of the union of the n Geršgorin disks, namely,

$$\Gamma^{\omega_k}(A) := \bigcup_{i\in N}\{z \in \mathbb{C} : |z - a_{i,i}| \le |\omega_k - a_{i,i}|\}, \tag{2.3.35}$$

and their intersection,

$$\bigcap_{k=1}^{m} \Gamma^{\omega_k}(A), \tag{2.3.36}$$

gives an approximation to $\Gamma^{\mathbb{D}}(A)$, for which $\Gamma^{\mathbb{D}}(A)$ is a *subset*, and for which m points, of the boundary of $\bigcap_{k=1}^m \Gamma^{\omega_k}(A)$, are *boundary points* of $\Gamma^{\mathbb{D}}(A)$.

Consider the irreducible 3×3 matrix

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \tag{2.3.37}$$

whose minimal Geršgorin set, $\Gamma^{\mathbb{D}}(C)$, is shown with the most inner boundary in Figure 2.3.1. For the vector $\mathbf{x}^0 = [1,1,1]^T \in \mathbb{R}^3$, the associated Geršgorin set $\Gamma^{X^0}(A)$, turns out to be simply

$$\Gamma^{X^0}(A) = \{z \in \mathbb{C} : |z - 2| \le 2\}. \tag{2.3.38}$$

The boundary of this set is the most outer circle in Figure 2.3.1.

Next, starting with the diagonal entry, $z = 2$, of the matrix C, we estimate $\nu_A(2)$, which is positive, from (2.3.22). As $\mu(2) = 1$, from (2.3.9), the associated nonnegative irreducible matrix $B(2)$, from (2.3.10), is

$$B(2) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix},$$

and a few power method iterations (see (2.3.26) - (2.3.27)), starting with $\mathbf{x}^0 = [2,1,2]^T$, gives that $\rho(B(2)) \doteq 2.2$. More precisely[4], $\rho(B(2)) = 2.24697$, so that from (2.3.13), we have $\nu(2) = 1.24697$.

Now, we search on the ray $2+t$, with $t \ge 0$, for the largest value $\hat{t}$, for which $\nu_A(2+\hat{t}) = 0$, and $\nu(2+t) \ge 0$ for all $0 \le t \le \hat{t}$. Using the inequality of (2.3.16), it follows that $\hat{t} \ge \nu_A(2) = 1.24697$, but in this particular case, it happens that $\hat{t} = 1.24697$, so that $z_1 = 3.24697$ is such that $\nu_A(z_1) = 0$, with $z_1 \in \partial\Gamma^{\mathbb{D}}(A)$. Similar, on considering the diagonal entry $1 = a_{2,2}$, we approximate $\nu_A(1)$, which turns out to be $\nu_A(1) = 0.80194$, and, then, searching on the ray $1-t$, $t \ge 0$, we Similar obtain $\nu_A(z_2) = 0$ with $z_2 = 0.19806$,

[4]All such numbers are truncated after five decimal digits.

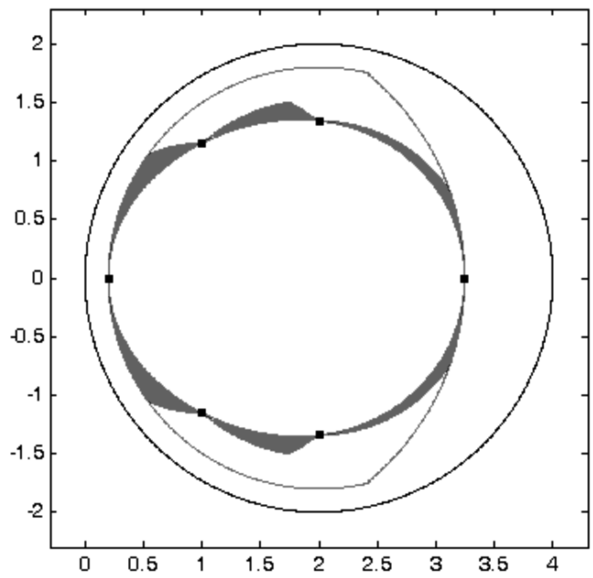

Figure 2.3.1: Approximations of the minimal Geršgorin set of the matrix A given in (2.3.37) *(Aproksimacija minimalnog Geršgorinovog skupa za matricu A datu sa (2.3.37))*

and with $z_2 \in \partial\Gamma^{\mathbb{D}}(A)$. Calling $\Gamma^{X_1}(A)$, and $\Gamma^{X_2}(A)$ the associated Geršgorin sets, then the intersection of the three sets, $\bigcap_{j=0}^{2}\Gamma^{X_j}(A)$, is shown in Figure 2.3.1 with the boundary that is the closest to the boundary of Geršgorin set of the matrix A.

We see, from Figure 2.3.1, that this set contains $\Gamma^{\mathbb{D}}(A)$, and that it has two real boundary points, shown as the black squares z_1 and z_2, in common with $\Gamma^{\mathbb{D}}(A)$.

We continue, and knowing $\nu_A(a_{1,1} = a_{3,3} = 2) = 1.24697$ and $\nu(a_{2,2} = 1) = 0.80194$, we look for four additional points of $\partial\Gamma^{\mathbb{D}}(A)$, which are found on the four rays: $2 \pm it$, $t \geq 0$, and $1 \pm it$, $t \geq 0$. This gives us the following four points $\{z_j\}_{j=3}^{6}$ of $\Gamma^{\mathbb{D}}(A)$:

$$z_3 = 1 + i(1.150963),\ z_4 = \overline{z_3},\ z_5 = 2 + i(1.34236),\ z_6 = \overline{z_5}.$$

The intersection of the above associated six Geršgorin sets is shown in Figure 2.3.1 with the boundary that is closest to the $\Gamma^{\mathbb{D}}(A)$ and has six boundary points in common with $\partial\Gamma^{\mathbb{D}}(A)$, shown as solid black squares. The region between the obtained approximation, and the boundary of $\Gamma^{\mathbb{D}}(A)$ is shaded, and can be seen as the set of small "roofs", composed of segments of circular arcs.

The amount of numerical calculations to obtain good approximation to $\Gamma^{\mathbb{D}}(A)$ is moderate, and it is evident that *better* approximations to $\Gamma^{\mathbb{D}}(A)$, having more points in common with $\partial\Gamma^{\mathbb{D}}(A)$, can be Similar constructed.

How Brauer ovals of Cassini, and other Geršgorin-type localizations sets, can be used in order to improve the obtained approximations of the minimal Geršgorin set, it was also discussed in [52].

Chapter 3

Localization of Generalized Eigenvalues

In this chapter we give the very recent contribution to the theory of localization of generalized eigenvalues. The material is divided into three sections. The first one gives the introduction and defines the problem. The second introduces the Geršgorin set for generalized eigenvalues in the form it was obtained by Kostić, Cvetković and Varga in [35]. In the same section, we derive a new approximation of the obtained generalized Geršgorin set, which is more suitable for practical use then the original set, and gives better results then the one known from the pioneering work of Stewart and Sun, [44, Corollary VI.2.5]. The third section introduces completely new results on the localization of Generalized Eigenvalues, which are obtained by using the concepts previously developed in this theses. The general principles will be defined, and then used as a unifying framework for the several consecutive results that concern specific generalized Geršgorin-type theorems. Fourth section generalizes the approximated sets through the use of complex parameter, and suggests the direction for further research. The last section is the treatment of the generalized minimal Geršgorin set as it was done in [35].

3.1 Generalized Eigenvalues

Given two arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geq 1$, the family of matrices $A - zB$, parameterized by the complex number z, is called **a matrix pencil**. So, a matrix pencil $A - zB$ can 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 complex matrices of the size n.

Definition 3.1.1. Given arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geq 1$, 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 **regular**.

The case when singularity of matrix pairs occurs can be expressed in the terms of eigenvectors of the matrices A and B. Namely, if A and B have an overlapping null spaces, meaning that there exists a vector $\mathbf{x}$ that is in the null space of A, and in the null space of B, then, for an arbitrary $z \in \mathbb{C}$, we can write $(A - zB)\mathbf{x} = \mathbf{0}$. This, of course, implies that $\det(A - zB) = 0$, for all $z \in \mathbb{C}$.

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

Definition 3.1.2. 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 $A\mathbf{v} = \lambda B\mathbf{v}$, $\mathbf{v}$ is called an **eigenvector** of the pair (A, B), and λ is called a **finite eigenvalue** of the pair (A, B). Furthermore, if there exists a nonzero vector $\mathbf{v} \in \mathbb{C}^n$, such that $B\mathbf{v} = \mathbf{0}$, then $A\mathbf{v} \neq \mathbf{0}$, and 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), λ is a finite eigenvalue of the pair (A, B), if and only if $A - \lambda B$ is singular matrix, i.e., if $\det(A - \lambda B) = 0$. So, the previous definition can be expressed in terms of determinants.

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

$$\det(A - zB) =: p(z), \tag{3.1.1}$$

where $p(z)$ is a polynomial in z, with a degree at most n. From [44], 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 $B = I_n$, we have that polynomial $p(z)$ in (3.1.1) 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**.

Since Generalized Eigenvalues may be infinite, in order to cover them, we will have to work in the **extended complex plane** $\mathbb{C}_\infty := \mathbb{C} \cup \{\infty\}$. Namely, for every subset $U \in \mathbb{C}$

such that there exists a sequence of complex numbers $\{z_k\}_{k\in\mathbb{N}}$, such that $|z_k| \to \infty$, when $k \to \infty$, and $z_k \in U$, for all $k \in \mathbb{N}$, we define the set $\widetilde{U}$ as $\widetilde{U} := U \cup \{\infty\}$. While working in the extended complex plane, we will always identify the set $\widetilde{U} \subseteq \mathbb{C}_\infty$ with the set U. So, it is easy to see that, with the convention $\infty^{-1} = 0$, for every set $U \in \mathbb{C}_\infty$, $\infty \in U$ if and only if $0 \in U^{-1}$, where the set U^{-1} is defined as $U^{-1} := \{z \in \mathbb{C}_\infty : z^{-1} \in U\}$.

Definition 3.1.3. Given a regular matrix pair (A, B), the collection of all eigenvalues of the pair (A, B) is called the **spectrum** of the pair (A, B), and it is denoted by

$$\sigma(A,B) := \begin{cases} \{z \in \mathbb{C} : \det(zB - A) = 0\}, & B \text{ is nonsingular}, \\ \{z \in \mathbb{C} : \det(zB - A) = 0\} \cup \{\infty\}, & B \text{ is singular}. \end{cases} \tag{3.1.2}$$

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 $B = I_n \in \mathbb{C}^{n,n}$, then the spectrum of the matrix pair (A, B) reduces to the standard spectrum of A, i.e., $\sigma(A, I_n) = \sigma(A)$. So, again, the word **generalized spectrum** is used to refer to the spectrum of a matrix pair.

An interesting property of the spectrum of a matrix pair is that for an arbitrary regular $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $\sigma(A, B)^{-1} = \sigma(B, A)$. Namely, if $\lambda \neq 0$, then $\lambda \in \sigma_F(A, B)$ is equivalent to $0 = \det(\lambda B - A) = \lambda^{-n} \det(B - \lambda^{-1}A)$, i.e., $\lambda^{-1} \in \sigma_F(B, A)$. Next, if $\lambda = 0$, then $\lambda B - A = -A$, and $\lambda \in \sigma_F(A, B)$ is equivalent to the fact that A is singular. Since $\lambda^{-1} = \infty$, again we have that $\lambda^{-1} \in \sigma(B, A)$. The last case is $\lambda = \infty$; then if $\lambda \in \sigma(A, B)$ the matrix B is singular, and $\det(B - \lambda^{-1}A) = \det(B) = 0$, so, again, $\lambda^{-1} \in \sigma(B, A)$. Since in the each of the three cases we had the equivalences, we conclude that $\sigma(A, B)^{-1} = \sigma(B, A)$.

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 base for many numerical methods. But, as a consequence of rounding errors when B is ill-conditioned, these methods fail. Thus, the treatment of the generalized spectra directly is an important topic. As the localization of eigenvalues of Geršgorin type, covered in the previous chapter, has found its value as a tool for determining information about the spectra of a matrix before it is calculated, one can be motivated to perform the similar thing with the generalized spectrum. But, this topic wasn't considered actively until very recently. Actually, an extension of the Geršgorin's theorem to the concept of matrix pairs was done by Stewart in 1975 in [43], but since then, there wasn't much done to examine the behavior of the obtained sets, or to improve localization results. A study of generalized eigenvalue localization sets in terms of perturbations, through spectral value sets, was done by Karow in [32]. But, the simplicity and elegance one can find in the original Geršgorin set, was somehow lost in the generalized eigenvalue case. The very recent paper by Kostić, Cvetković and Varga, [35], examines thoroughly the Geršgorin set for generalized eigenvalues, and introduces the minimal Geršgorin set, together with an important treatment of its calculation. Since the main idea lies in generalized diagonal dominance, this topic is an excellent example of how the concepts of the first and second chapter can be applied. Giving the review of the results obtained in [35], and using the results of the previous chapters, we will obtain several new eigenvalue inclusion sets for generalized eigenvalues that are suitable for computation, and we will discuss their relationships.

3.2 Geršgorin's Theorem for the Generalized Eigenvalues

As the Geršgorin set could be derived from the class of SDD matrices using Varga's Equivalence Principle, we derive the generalized Geršgorin set.

Definition 3.2.1. Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the set $\Gamma(A, B)$ defined as:

$$\Gamma(A, B) := \{z \in \mathbb{C} : A - zB \text{ is not an SDD matrix}\}, \tag{3.2.1}$$

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

Theorem 3.2.2. *Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the spectrum of the pair (A, B) belongs to the generalized Geršgorin set of the matrix pair (A, B), i.e., the following inclusion holds:*

$$\sigma(A, B) \subseteq \Gamma(A, B). \tag{3.2.2}$$

The proof of this theorem follows in the same way as the corresponding part of the proof of Varga's Equivalence Principle.

We continue with the characterization of this set, in terms of entries of the two corresponding matrices. It is easy to see, from (3.2.1), that $\Gamma(A, B) = \bigcup_{i \in N} \Gamma_i(A, B)$, where

$$\Gamma_i(A, B) := \{z \in \mathbb{C} : |b_{i,i}z - a_{i,i}| \leq \sum_{j \in N \setminus \{i\}} |b_{i,j}z - a_{i,j}|\}, \quad (\text{all } i \in N). \tag{3.2.3}$$

Moreover, if we denote, by $r_i^S(A) := \sum_{j \in S \setminus \{i\}} |a_{i,j}|$, the part of a row sum that corresponds to the columns given by the set of indices $S \subseteq N$, and if we denote the particular sets of indices $\beta_i := \{j \in N : b_{i,j} \neq 0\}$ and $\overline{\beta}_i := \{j \in N : b_{i,j} = 0\}$, for all $i \in N$, then we can write

$$\Gamma_i(A, B) = \{z \in \mathbb{C} : |z - \frac{a_{i,i}}{b_{i,i}}||b_{i,i}| - \sum_{j \in \beta_i \setminus \{i\}} |z - \frac{a_{i,j}}{b_{i,j}}||b_{i,j}| \leq r_i^{\overline{\beta}_i}(A)\}, \tag{3.2.4}$$

whenever $i \in \beta_i$,

$$\Gamma_i(A, B) = \{z \in \mathbb{C} : |a_{i,i}| - r_i^{\overline{\beta}_i}(A) \leq \sum_{j \in \beta_i} |z - \frac{a_{i,j}}{b_{i,j}}||b_{i,j}|\}, \tag{3.2.5}$$

whenever $i \in \overline{\beta}_i$.

We remark that a set $\Gamma_i(A, B)$, as defined in (3.2.3), can be an *empty set*, which can occur when $\beta_i = \emptyset$, i.e., when all entries of the i-th row of the matrix B are zero. Then, the i-th generalized Geršgorin set has the following form

$$\Gamma_i(A, B) = \{z \in \mathbb{C} : |a_{i,i}| \leq r_i^{\overline{\beta}_i}(A) = r_i(A)\}; \tag{3.2.6}$$

thus,

$$\Gamma_i(A, B) = \begin{cases} \emptyset, & \text{if } |a_{i,i}| > r_i(A), \\ \mathbb{C}, & \text{if } |a_{i,i}| \leq r_i(A). \end{cases} \tag{3.2.7}$$

Of course, when the second case of (3.2.7) occurs, the matrix B is singular, and $p(z) = \det(A - zB)$ has degree less then n. Since we are considering regular matrix pairs, the degree of the polynomial $p(z)$ has to be at least one; thus, at least one of the sets $\Gamma_i(A, B)$ has to be nonempty, implying that the generalized Geršgorin set of a regular matrix pencil is always nonempty.

On inspecting the form of the generalized Geršgorin "disks" of (3.2.4) and (3.2.5), the following properties were established in [35].

Theorem 3.2.3. *Let $A, B \in \mathbb{C}^{n,n}$, with $n \geq 2$. 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 i-th generalized Geršgorin set, $\Gamma_i(A, B)$, as defined in (3.2.4) and (3.2.5), 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 i-th generalized Geršgorin set $\Gamma_i(A, B)$, given in (3.2.3), contains zero if and only if $|a_{i,i}| \leq r_i(A)$.*

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 $|a_{i,i}| \leq r_i^{\overline{\beta}_i}(A)$, then $\Gamma_i(A, B)$, and consequently $\Gamma(A, B)$, are the entire complex plane.*

Proof. First, it is evident that 2. and 4. follow directly from 1. and 3., respectively. Second, 3. is easy to obtain, by putting $z = 0$ in the inequalities of (3.2.3), and 5. follows directly from (3.2.5). Then, it remains to prove 1.

If $i \in \overline{\beta}_i$, then $|b_{i,i}| = 0 \leq r_i(B)$ and $\Gamma_i(A, B)$ is unbounded from (3.2.3). Thus, let $i \in \beta_i$. If we suppose that $\Gamma_i(A, B)$ is unbounded, then, there is a sequence $\{z_k\}_{k\in\mathbb{N}}$ of complex numbers, such that $|z_k| \to \infty$, as $k \to \infty$, and $z_k \in \Gamma_i(A, B)$. But, for a sufficiently large $k \in \mathbb{N}$, from (3.2.4), we have

$$|z_k|(|b_{i,i}| - r_i(B)) \leq r_i^{\overline{\beta}_i}(A). \tag{3.2.8}$$

Now, if $|b_{i,i}| > r_i(B)$, then taking the limit as $k \to \infty$ in (3.2.8), we obtain a contradiction. Conversely, let $|b_{i,i}| \leq r_i(B)$, and let $\{z_k\}_{k\in\mathbb{N}}$ be a a sequence of complex numbers such that $|z_k| \to \infty$, when $k \to \infty$. Then, it is easy to see that, for a sufficiently large $k \in \mathbb{N}$,

$$|z_k - \frac{a_{i,i}}{b_{i,i}}||b_{i,i}| - \sum_{j\in\beta_i\setminus\{i\}} |z_k - \frac{a_{i,j}}{b_{i,j}}||b_{i,j}| \leq 0,$$

and thus, $z_k \in \Gamma_i(A, B)$ from (3.2.5). $\square$

The following example illustrates the generalized Geršgorin set.

Example 3.2.4.

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

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

By inspection, B_1 and B_2 are SDD matrices, and, according to the item (2) of Theorem 3.2.3, sets $\Gamma(A_1, B_1)$ and $\Gamma(A_2, B_2)$ are compact in the complex plane. Similar, from Theorem 3.2.6, sets $\widehat{\Gamma}(A_1, B_1)$ and $\widehat{\Gamma}(A_2, B_2)$ are compact, too.

This is shown in Figures 3.2.1 and 3.2.2, 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, since 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, matrix A_2 is not SDD, and $0 \in \Gamma(A_2, B_2) \subseteq \widehat{\Gamma}(A_2, B_2)$.

Figure 3.2.3 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 reminder of the generalized Geršgorin set of the matrix pair (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 doesn't reflect this situation.

One of the major drawbacks of this set is that it is not as elegant as the original Geršgorin set. Namely, as we have seen, the sets $\Gamma_i(A, B)$, in general, are not circles, and moreover, they are sufficiently hard to calculate, in order to obtain useful plots. So, one can be motivated to approximate the generalized Geršgorin set in order to obtain, although larger, more practical localization sets for generalized eigenvalues. First attempt was done by Stewart in [43], where for the first time set (3.2.3) was defined. Here, we will use a different approach, and obtain approximations of the generalized Geršgorin sets that are, in the worst case as good as the ones in [43].

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 $i \in N$,

$$r_i(zB - A) = \sum_{j \in N \setminus \{i\}} |zb_{i,j} - a_{i,j}|. \tag{3.2.9}$$

Then, we proceed by approximating the right hand side by the triangle inequality, and obtain

$$r_i(zB - A) \leq |z| r_i(B) + r_i(A). \tag{3.2.10}$$

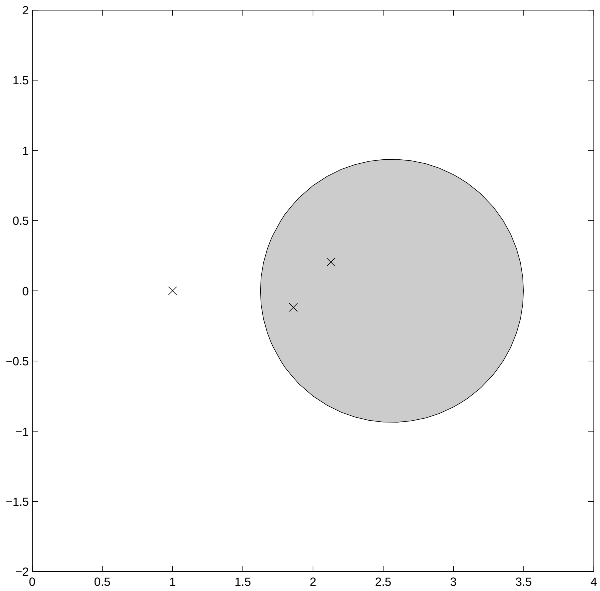

Figure 3.2.1: Generalized Geršgorin set of the matrix pair (A_1, B_1) of the Example 3.2.4
(Generalizovani Geršgorinov skup za matrični par (A_1, B_1) *iz Primera 3.2.4)*

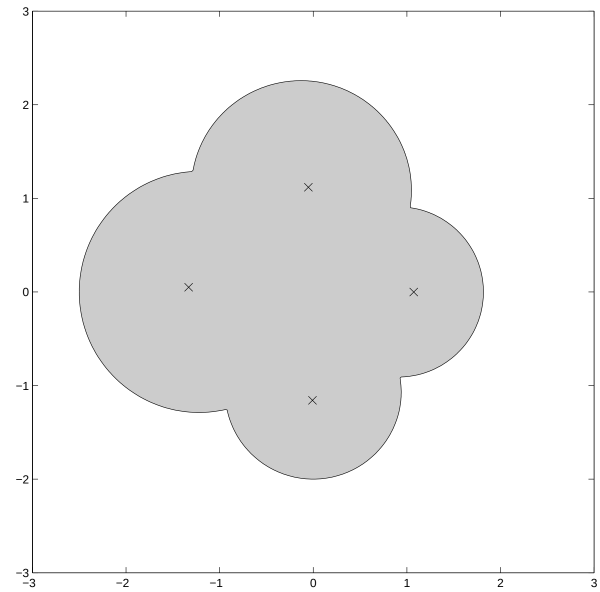

Figure 3.2.2: Generalized Geršgorin set of the matrix pair (A_2, B_2) of the Example 3.2.4
(Generalizovani Geršgorinov skup za matrični par (A_2, B_2) *iz Primera 3.2.4)*

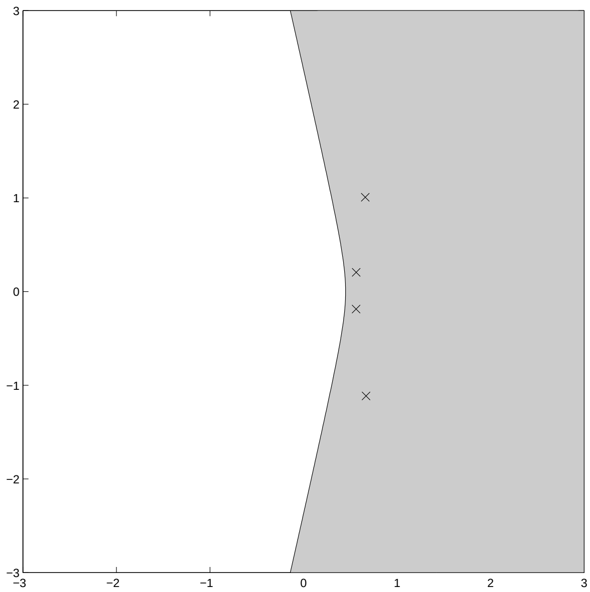

Figure 3.2.3: Generalized Geršgorin set of the matrix pair (A_1, A_2) of the Example 3.2.4
(Generalizovani Geršgorinov skup za matrični par (A_1, A_2) iz Primera 3.2.4)

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

$$\begin{cases} \widehat{\Gamma}_i(A, B) := \{z \in \mathbb{C} \, : \, |zb_{i,i} - a_{i,i}| \leq |z| r_i(B) + r_i(A)\}, \quad (i \in N), \\ \widehat{\Gamma}(A, B) := \bigcup_{i \in N} \widehat{\Gamma}_i(A, B), \end{cases} \tag{3.2.11}$$

which we call the **i-th approximated generalized Geršgorin set** and the **approximated generalized Geršgorin set**.

It is interesting to note that, taking $B = I$, our approximation $\widehat{\Gamma}(A, B)$ reduces to ordinary Geršgorin set $\Gamma(A)$, which is not the case for the approximation of Stewart, obtained in [43].

Assuming that $b_{i,i} \neq 0$ and $a_{i,i} \neq 0$, the i-th approximated generalized Geršgorin set is an area in the complex plane whose boundary is the curve which can be represented as

$$|z - \xi| = \beta |z| + \alpha, \tag{3.2.12}$$

where $\alpha, \beta \geq 0$ and $\xi \in \mathbb{C}$. After some analysis, the following classification can be obtained:

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

[1] Apollonius of Perga (262 BC190 BC), a Greek geometer and astronomer noted for his writings on conic sections.

- If $\alpha > 0$ and $\beta = 0$, then the curve is a circle centered in ξ with radius α;
- If $\alpha > 0$, and $\beta \geq 0$, then the curve is a **Cartesian Oval**[2] with foci in 0 and $\xi \in \mathbb{C}$, and linear factors $-\beta\alpha^{-1}$ and α^{-1};

To illustrate this, Figure 3.2.4 shows the curves (3.2.12) 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, respectively.

As a consequence of (3.2.10) and Theorem 3.2.2, the following theorem obviously holds.

Theorem 3.2.5. *Given a regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the finite spectrum of the pair (A, B) belongs to the approximated generalized Geršgorin set of the matrix pair (A, B). Moreover, the following inclusions hold:*

$$\sigma(A, B) \subseteq \Gamma(A, B) \subseteq \widehat{\Gamma}(A, B), \tag{3.2.13}$$

where $\widehat{\Gamma}(A, B)$ is given in 3.2.11.

Considering the form of the i-th approximated generalized Geršgorin set, given in 3.2.3, as in Theorem 3.2.3, we obtain a somewhat expected result.

Theorem 3.2.6. *Let $A, B \in \mathbb{C}^{n,n}$, with $n \geq 2$. 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 i-th approximated generalized Geršgorin set, $\widehat{\Gamma}_i(A, B)$, as defined in (3.2.11), 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 i-th approximated generalized Geršgorin set $\widehat{\Gamma}_i(A, B)$, given in (3.2.3), contains zero if and only if $|a_{i,i}| \leq r_i(A)$.*
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 $|a_{i,i}| \leq r_i(A)$, then $\widehat{\Gamma}_i(A, B)$, and consequently, $\widehat{\Gamma}(A, B)$, are the entire complex plane.*

That this approximation can be sufficiently good, is shown in the Figures 3.2.6 and and 3.2.7, where the original generalized Geršgorin set is shaded, while the boundary of the approximated one is given by the thick black line. On the other hand, due to the application of the triangle inequality on the absolute value of the difference of two close values, significant deviations can occur, which can easily be seen in the Figure 3.2.5.

[2]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 oval, which can occur here, are limaçon of Pascal (for $\alpha = 1$) and hyperbola (for $\beta = 1$)

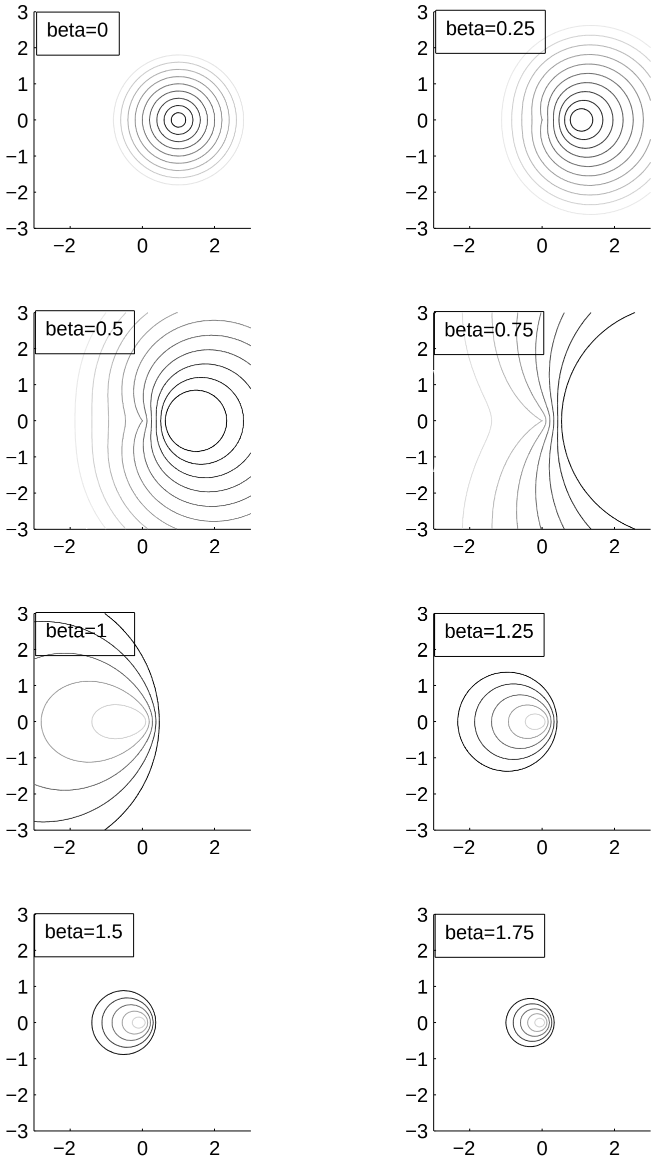

Figure 3.2.4: The curves (3.2.12) plotted for $\xi = 1$ and $\alpha \in \{0, 0.2, 0.4, \ldots, 1.8, 2\}$, setting $\beta = 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$, for each plot, from upper left to lower right corner, respectively

(Kriva data sa (3.2.12) je nacrtana za $\xi = 1$ *i* $\alpha \in \{0, 0.2, 0.4, \ldots, 1.8, 2\}$*, fiksirajući* $\beta = 0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$*, za svaki crtež, redom, od gornjeg levog do donjeg desnog ugla)*

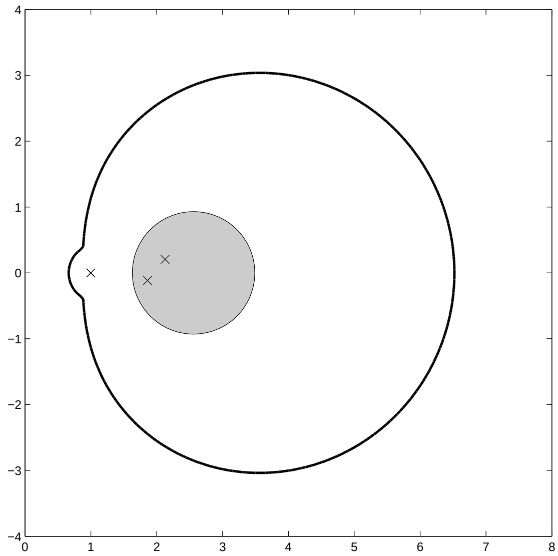

Figure 3.2.5: Approximated generalized Geršgorin set of the matrix pair (A_1, B_1) of the Example 3.2.4
(Aproksimirani generalizovani Geršgorinov skup za matrični par (A_1, B_1) *iz Primera 3.2.4)*

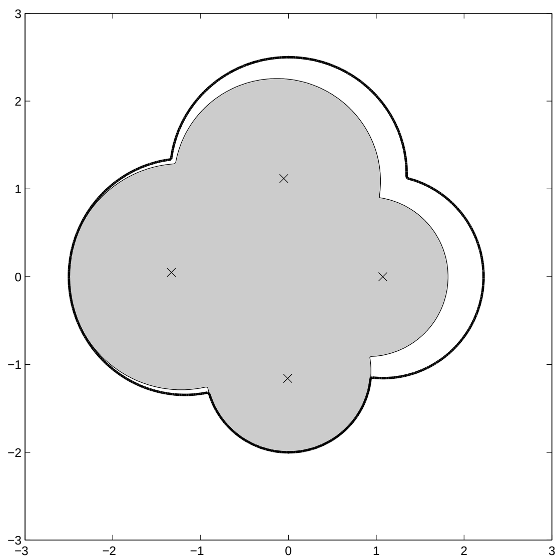

Figure 3.2.6: Approximated generalized Geršgorin set of the matrix pair (A_2, B_2) of the Example 3.2.4
(Aproksimirani generalizovani Geršgorinov skup za matrični par (A_2, B_2) *iz Primera 3.2.4)*

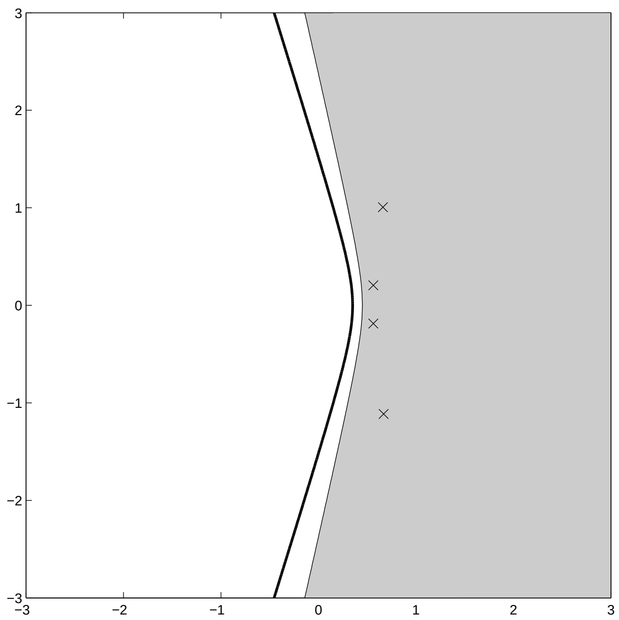

Figure 3.2.7: Approximated generalized Geršgorin set of the matrix pair (A_1, A_2) of the Example 3.2.4

(Aproksimirani generalizovani Geršgorinov skup za matrični par (A_1, A_2) *iz Primera 3.2.4)*

3.3 Geršgorin-type Localizations for Generalized Eigenvalues

Here, our object is to *estimate* the spectra of regular matrix pairs, as the union of the Geršgorin disks of a given matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$ *estimates* the eigenvalues of A. As in the second chapter, we use Varga's Equivalence Principle adapted to generalized eigenvalues.

Theorem 3.3.1. (Varga's Equivalence Principle) *Given a class of square matrices of an arbitrary size, denoted by $\mathbb{K}$, for any regular matrix pair (A, B), define the set of complex numbers $\Theta^{\mathbb{K}}(A, B) := \{z \in \mathbb{C} : zB - A \notin \mathbb{K}\}$. Then, the following two conditions are equivalent:*

- *All matrices from $\mathbb{K}$ are nonsingular,*
- *Given any regular matrix pair (A, B), the set $\Theta^{\mathbb{K}}(A, B)$ contains all its eigenvalues, i.e., $\sigma(A, B) \subseteq \Theta^{\mathbb{K}}(A, B)$.*

Proof. The proof of the theorem is similar to the corresponding one of the Section 2.2. The only difference lies in the case when $\lambda = \infty \in \sigma(A, B)$. To prove that in this case $\lambda \in \Theta^{\mathbb{K}}(A, B)$, too, observe that the generalized Geršgorin-type set has the same property as the spectrum when the matrices in the corresponding pair exchange their roles. More precisely, given any regular pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $(\Theta^{\mathbb{K}}(A, B))^{-1} = \Theta^{\mathbb{K}}(B, A)$, for every class of nonsingular matrices $\mathbb{K}$. This follows in the same way as the property was proved for the spectrum of the matrix pair. $\square$

As before, for Geršgorin-type sets, we define the term *generalized Geršgorin-type set.* We say that set $\Theta^{\mathbb{K}}(A, B) = \{z \in \mathbb{C} : zB - A \notin \mathbb{K}\}$ is a **generalized Geršgorin-type set**, if $\mathbb{K}$ is a diagonally dominant-type class[3] of nonsingular matrices. Since Theorem 3.3.1 holds, we refer to such sets also as **generalized Geršgorin-type localization sets**, or **generalized Geršgorin-type inclusion sets**.

While the sets obtained in this way in the previous chapter were practical enough for the localization of the spectrum of the given matrix, now, even generalized Geršgorin set can be sufficiently difficult to manage. Therefore, we need to obtain approximates in order to be able to apply these results in practice. So, we continue with an extension of the concept of the *approximated* set (3.2.11).

Given arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, define the operation $\langle A, B\rangle =: M = [m_{i,j}]$, so that

$$m_{i,j} := \begin{cases} |a_{i,i} - b_{i,i}|, & i = j, \\ -|a_{i,j}| - |b_{i,j}|, & \text{otherwise}, \end{cases} \tag{3.3.1}$$

for all $i, j \in N$. Then, obviously, the following inequality hold:

$$\langle A - B\rangle \geq \langle A, B\rangle. \tag{3.3.2}$$

[3]Given in Definition 1.3.7.

Theorem 3.3.2. (Approximation Principle) *Given a diagonally dominant-type class of matrices, denoted by* $\mathbb{K}$*, for any regular matrix pair* (A,B)*, define the set of complex numbers* $\widehat{\Theta}^{\mathbb{K}}(A,B) := \{z \in \mathbb{C} : \langle zB, A\rangle \notin \mathbb{K}\}$*. Then,*

$$\Theta^{\mathbb{K}}(A,B) \subseteq \widehat{\Theta}^{\mathbb{K}}(A,B).$$

Proof. Given a DD-type class $\mathbb{K}$, take an arbitrary regular pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, and assume $z \in \Theta^{\mathbb{K}}(A,B)$. Then, $zB - A \notin \mathbb{K}$, so there exists at least one matrix M such that $\langle M\rangle \geq \langle zB - A\rangle$ and $M \notin \mathbb{K}$. But, since $\langle zB - A\rangle \geq \langle zB, A\rangle$, then $\langle M\rangle \geq \langle zB, A\rangle$, and $\langle zB, A\rangle \notin \mathbb{K}$. In other words, $z \in \widehat{\Theta}^{\mathbb{K}}(A,B)$, which completes the proof. □

Given a regular matrix pair (A,B), we will call the set $\widehat{\Theta}^{\mathbb{K}}(A,B)$ the **approximated generalized Geršgorin-type set** attributed to the class $\mathbb{K}$. Together with Theorem 3.3.1, this gives the following corollary.

Corollary 3.3.3. *Given a diagonally dominant-type class of nonsingular matrices, denoted by* $\mathbb{K}$*, for any regular matrix pair* (A,B)*,*

$$\sigma(A,B) \subseteq \Theta^{\mathbb{K}}(A,B) \subseteq \widehat{\Theta}^{\mathbb{K}}(A,B).$$

An obvious example of this set is (3.2.11).

A useful property in the second chapter was also *Isolation Principle*, so, we extend this concept to generalized eigenvalues, too.

Theorem 3.3.4. (Isolation Principle) *Given a generalized Geršgorin-type set*

$$\Theta^{\mathbb{K}}(A,B) = \{z \in \mathbb{C} : zB - A \notin \mathbb{K}\}, \tag{3.3.3}$$

where $\mathbb{K} \subset \mathbb{H}$ *is a positively homogenous* [4] *diagonally dominant-type subclass of H-matrices, for any regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U, V \subseteq \mathbb{C}$*, such that*

$$U \cap V = \emptyset \quad \text{and} \quad \Theta^{\mathbb{K}}(A,B) = U \cup V, \tag{3.3.4}$$

then, the set U *contains exactly* $|\left\{i \in N : b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *finite eigenvalues of the pair* (A,B)*, and, if* U *is unbounded, exactly* $|\{i \in N : b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N : a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *eigenvalues of the pair* (A,B)*, where, by convention,* $z \cdot \infty := \infty$*.*

Proof. Let $D_A := \operatorname{diag}(a_{1,1}, a_{2,2}, \ldots, a_{n,n})$ and $D_B := \operatorname{diag}(b_{1,1}, b_{2,2}, \ldots, b_{n,n})$. Take the splittings of the matrices $A = D_A - F_A$ and $B = D_B - F_B$, and consider the families of matrices $A(t) := D_A - tF_A$ and $B(t) := D_B - tF_B$, for $0 \leq t \leq 1$.

As in the proof of Theorem 2.2.3, for all $t \in (0,1]$, we obtain $\Theta^{\mathbb{K}}(A(t),B(t)) \subseteq \Theta^{\mathbb{K}}(A,B)$.

So, considering the case when $t = 0$, we have that $A(0) = D_A$, and $z \in \Theta^{\mathbb{K}}(A(0),B(0))$ if and only if $zD_B - D_A \notin \mathbb{K}$. Obviously, if, for some $i \in N$, $z = a_{i,i}b_{i,i}^{-1}$, when $b_{i,i} \neq 0$, then $zD_B - D_A$ has a zero on its diagonal. Thus, it can not be in $\mathbb{K}$, which is a DD-type class of matrices. Therefore, $a_{i,i}b_{i,i}^{-1} \in \Theta^{\mathbb{K}}(A(0),B(0))$, for all $i \in N$, such that $b_{i,i} \neq 0$. For the same reason, for every $i \in N$, such that $b_{i,i} \neq 0$, $a_{i,i}b_{i,i}^{-1} \in \Theta^{\mathbb{K}}(A,B)$.

[4] Given in Definition 2.2.2.

On the other hand, when $z \neq a_{i,i}b_{i,i}^{-1}$, for all $i \in N$ where $b_{i,i} \neq 0$, $zD_B - D_A$ is a nonsingular diagonal matrix, implying that $zD_B - D_A \in \mathbb{K}$, i.e., $z \notin \Theta^{\mathbb{K}}(A(0), B(0))$. So, $\Theta^{\mathbb{K}}(A(0), B(0)) = \left\{a_{i,i}b_{i,i}^{-1} \,:\, b_{i,i} \neq 0,\ i \in N\right\} \subseteq \Theta^{\mathbb{K}}(A, B)$.

In another words, we have obtained that

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

and that $\Theta^{\mathbb{K}}(A(t), B(t)) \subseteq \Theta^{\mathbb{K}}(A, B)$, for all $t \in [0,\ 1]$.

Now, since the finite eigenvalues are continuous functions of the entries of both matrices, [44], as in the proof of Theorem 2.2.3, we obtain the desired result.

To prove the second part of the theorem, let $\infty \in U$, and let $r = |\,\{i \in N \,:\, b_{i,i} = 0\}\,|$. Since $\det(zD_B - D_A) = \prod_{i \in N} |zb_{i,i} - a_{i,i}|$ is the polynomial of the degree $n - r$, the pair $(A(0), B(0))$ has exactly r infinite eigenvalues. As we have proved, the number of the finite eigenvalues in U remains unchanged for all $t \in [0, 1]$. Hence, since $\infty \notin V$, we have that all of the r infinite eigenvalues of the pair (A, B) belong to the set U. □

As the obvious corollary, we state the Isolation Principle applied to the generalized Geršgorin set. Moreover, it is not hard to see that Isolation Principle is *hereditable* to the approximations of the localization sets. Therefore, the corresponding theorem is true for the approximated generalized Geršgorin set.

Theorem 3.3.5. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that*

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

then, the set U contains exactly $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ finite eigenvalues of the pair (A, B) and, if U is unbounded it has exactly $|\,\{i \in N \,:\, b_{i,i} = 0\}\,|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ eigenvalues of the pair (A, B).

Theorem 3.3.6. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that*

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

then the set U contains exactly $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ finite eigenvalues of the pair (A, B), and, if U is unbounded, exactly $|\,\{i \in N \,:\, b_{i,i} = 0\}\,|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ eigenvalues of the pair (A, B).

Inspired by Theorem 3.2.3, published in [35], which states that the generalized Geršgorin set is working "well" when one of the two corresponding matrices is an SDD matrix, we will now prove its generalization. Thus, the idea is to show what happens with the localization set $\Theta^{\mathbb{K}}(A, B)$ when $A \in \mathbb{K}$, or $B \in \mathbb{K}$. We first give some definitions.

Definition 3.3.7. A given class of matrices $\mathbb{K}$ is said to be **open** if, for every matrix $A \in \mathbb{K}$, there exists an arbitrary small $\varepsilon > 0$, so that for every matrix $B \in \mathbb{C}^{n,n}$, $|(A - B)_{i,j}| < \varepsilon$, for all $i, j \in N$, implies $B \in \mathbb{K}$.

Definition 3.3.8. An opened diagonally dominant-type class of matrices is called **strictly diagonally dominant-type** class, or briefly, **SDD-type** class.

It is interesting to note that, while SDD, Brauer SDD, Brualdi SDD, S-SDD and CKV-SDD matrices are SDD-type matrices, this is not the case with iDD, Brauer iDD, Brualdi iDD, S-iDD, matrices, which are only DD-type classes.

Theorem 3.3.9. (Boundedness Principle) *Given a positively homogenous SDD-type class of matrices* $\mathbb{K}$*, for any regular matrix pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \Theta^{\mathbb{K}}(A, B)$ *if and only if* $A \in \mathbb{K}$*, and*
- $\infty \notin \Theta^{\mathbb{K}}(A, B)$ *if and only if* $B \in \mathbb{K}$*.*

Proof. Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, let $z = 0$. Then $zB - A \in \mathbb{K}$ becomes $-A \in \mathbb{K}$, and, since $\mathbb{K}$ is a DD-type class, $A \in \mathbb{K}$. Thus, equivalently, $0 \notin \Theta^{\mathbb{K}}(A, B)$ if and only if $A \in \mathbb{K}$.
For the second item, start by assuming that $z = \infty \notin \Theta^{\mathbb{K}}(A, B)$. Then, for every sequence $\{z_k\}_{k \in \mathbb{N}} \subseteq \mathbb{C}$, such that $|z_k| \to \infty$, when $k \to \infty$, then $z_k B - A \in \mathbb{K}$, for a sufficiently large $k \in \mathbb{N}$. For such $k \in \mathbb{N}$ consider the matrix $M_k := B - (z_k)^{-1} A$. Since $\mathbb{K}$ is a positively homogeneous DD-type class, for a sufficiently large $k \in \mathbb{N}$, $|z_k||M_k| = |z_k B - A| \in \mathbb{K}$, and hence, $M_k \in \mathbb{K}$. But, for a sufficiently large $k \in \mathbb{N}$, we can make $|B - M_k| = |z_k|^{-1}|A|$ arbitrarily small. Thus, since $\mathbb{K}$ is an opened class of matrices, $M_k \in \mathbb{K}$ implies $B \in \mathbb{K}$.

To prove the converse, assume that $B \in \mathbb{K}$, and again, for an arbitrary sequence $\{z_k\}_{k \in \mathbb{N}} \subseteq \mathbb{C}$, such that $|z_k| \to \infty$, when $k \to \infty$, consider the matrix $M_k := B - (z_k)^{-1} A$. As before, from the fact that $\mathbb{K}$ is an opened positively homogenous DD-type class of matrices, we obtain that, for sufficiently large $k \in \mathbb{N}$, $z_k M_k = z_k B - A \in \mathbb{K}$, and, hence $z_k \notin \Theta^{\mathbb{K}}(A, B)$. Since the sequence was arbitrary, this implies that $z = \infty \notin \Theta^{\mathbb{K}}(A, B)$. □

An interesting corollary is the following: given a regular pair (A, B), such that one of the matrices is from the positively homogenous SDD-type class $\mathbb{K}$, using the corresponding generalized Geršgorin-type localization, we can always obtain a bounded localization set to estimate Generalized eigenvalues. Namely, if in the regular pair (A, B), $B \in \mathbb{K}$, and $\Theta^{\mathbb{K}}(A, B)$ is bounded set in $\mathbb{C}$. On the other hand, if $A \in \mathbb{K}$, then $\Theta^{\mathbb{K}}(B, A)$ is bounded set. But, since, $(\sigma_F(A, B))^{-1} = \sigma_F(B, A) \subseteq \Theta^{\mathbb{K}}(B, A)$, we have obtained the bounded localization of the reciprocal values of the finite eigenvalues.

As we have seen, instead of working with actual generalized Geršgorin type sets, it is much more suitable to work with their approximations. Therefore, it is an interesting question whether they also satisfy Boundedness Principle. The following theorem positively answers this question.

Theorem 3.3.10. *Given a positively homogenous SDD-type class of matrices* $\mathbb{K}$*, for any regular matrix pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$ *if and only if* $A \in \mathbb{K}$*, and*
- $\infty \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$ *if and only if* $B \in \mathbb{K}$*.*

Proof. First, given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, let $z = 0$. Then $\langle zB, A\rangle \in \mathbb{K}$ means $\langle A\rangle \in \mathbb{K}$, and since $\mathbb{K}$ is a DD-type class, $A \in \mathbb{K}$. Thus, equivalently, $0 \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$ if and only if $A \in \mathbb{K}$.

For the second item, start by assuming that $z = \infty \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$, then, for every sequence $\{z_k\}_{k\in\mathbb{N}} \subseteq \mathbb{C}$, such that $|z_k| \to \infty$, when $k \to \infty$, and $\langle z_k B, A\rangle \in \mathbb{K}$, for a sufficiently large $k \in \mathbb{N}$. Now, for such $k \in \mathbb{N}$, consider the matrix $M_k := \langle B, (z_k)^{-1}A\rangle$. Then, for a sufficiently large $k \in \mathbb{N}$, $|z_k|M_k = \langle z_k B, A\rangle \in \mathbb{K}$, and, since $\mathbb{K}$ is a positively homogeneous, $M_k \in \mathbb{K}$. But, for a sufficiently large $k \in \mathbb{N}$, we can make $|M_k - \langle B\rangle| \leq |z_k|^{-1}|A|$ arbitrarily small. Thus, since $\mathbb{K}$ is an SDD-type class of matrices, $M_k \in \mathbb{K}$ implies $\langle B\rangle \in \mathbb{K}$, and hence, $\langle B\rangle \in \mathbb{K}$.

To prove the converse, assume that $B \in \mathbb{K}$, and again, for an arbitrary sequence $\{z_k\}_{k\in\mathbb{N}} \subseteq \mathbb{C}$, such that $|z_k| \to \infty$, when $k \to \infty$, consider the matrix $M_k := \langle B, (z_k)^{-1}A\rangle$. As before, from the fact that $\mathbb{K}$ is an opened positively homogenous DD-type class of matrices, we obtain that, for sufficiently large $k \in \mathbb{N}$, $|z_k|M_k = \langle z_k B, A\rangle \in \mathbb{K}$, and, hence $z_k \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$. Since the sequence was arbitrary, this implies that $z = \infty \notin \widehat{\Theta}^{\mathbb{K}}(A, B)$. □

Therefore, having a generalized eigenvalue problem $A\mathbf{x} = \lambda B\mathbf{x}$, $\mathbf{x} \in \mathbb{C}^n$, if one of the matrices is an H-matrix, more precisely, if it is from one of the positively homogenous SDD-type subclasses of H-matrices, we can localize the solutions using the (suitable) bounded sets in the complex plane. So, in the following sections we proceed by giving the generalized Geršgorin sets that correspond to the SDD-type classes, covered in the first chapter. Furthermore, we have developed an approximation technique which gives more practical localization areas, while the important properties of the original localizations are not lost.

In what follows, we give the generalized Brauer set, generalized Brualdi set, generalized CKV-set, and generalized alpha-sets, together with their approximations. We will, in particular, establish their relationships, and present illustrative examples, as it was done for the corresponding localizations in the second chapter.

3.3.1 Brauer set for Generalized Eigenvalues

As we have seen, the class of doubly SDD matrices is a class of *nonsingular* matrices that is *positively homogenous* and a *DD-type*. This suffices, by Theorem 3.3.1, to conclude that the corresponding generalized Geršgorin type set is a localization area for generalized eigenvalues. Moreover, by Theorem 3.3.4, we can use the Isolation Principle, too. But, in order to apply Boundedness Principle, it remains to prove that this class of matrices is *opened*, as it was suggested in the previous section. Now, since this class is defined exclusively by the strict inequalities, it is easy to see that, for a given doubly SDD matrix $A = [a_{i,j}] \in \mathbb{C}^{n,n}$, one can always choose a sufficiently small $\varepsilon > 0$, so that perturbations of the matrix entries, which are by absolute value smaller then ε, produce the matrix B that is also doubly SDD. Therefore, the following three theorems are direct corollaries of the mentioned three principles for generalized eigenvalues.

Theorem 3.3.11. (Brauer) *Given an arbitrary regular pair $A, B \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, for every $\lambda \in \sigma(A, B)$, there exists a pair of indices $i, j \in N$, $i \neq j$, so that*

$$\lambda \in K_{i,j}(A, B) := \{z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}||zb_{i,i} - a_{j,j}| \leq r_i(zB - A)r_j(zB - A)\}\,, \tag{3.3.5}$$

where $r_i(zB - A)$, for all $i \in N$ is given in (3.2.9). Consequently,

$$\sigma(A, B) \subseteq \mathcal{K}(A, B) := \bigcup_{i \in N} \bigcup_{j=1}^{i-1} K_{i,j}(A, B). \tag{3.3.6}$$

Theorem 3.3.12. *Given an arbitrary pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and $\mathcal{K}(A) = U \cup V$, then the set U contains exactly $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ finite eigenvalues of the pair (A, B) and, if U is unbounded, exactly $|\{i \in N \,:\, b_{i,i} = 0\}|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ eigenvalues of the pair (A, B).*

Theorem 3.3.13. *For any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following two conditions hold:*

- *$0 \notin \mathcal{K}(A, B)$ if and only if A is doubly SDD, and*
- *$\infty \notin \mathcal{K}(A, B)$ if and only if B is doubly SDD.*

First, we illustrate the generalized Brauer set using the matrices of the Example 3.2.4. In Figures 3.3.1, 3.3.2 and 3.3.3, generalized Brauer sets $\mathcal{K}(A_1, B_1)$, $\mathcal{K}(A_2, B_2)$ and $\mathcal{K}(A_1, A_2)$ are shaded, respectively, while the boundary of the corresponding generalized Geršgorin sets $\Gamma(A_1, B_1)$, $\Gamma(A_2, B_2)$ and $\Gamma(A_1, A_2)$ are given by the thick black line.

To see that, sometimes, the improvement obtained by using the generalized Brauer sets instead of the generalized Geršgorin sets, can be truly significant, we consider the following example.

Example 3.3.14.

$$A_3 = \begin{pmatrix} 1 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0 & 0 & 0 & -i \end{pmatrix}, \quad B_3 = \begin{pmatrix} 1 & 0.8 & 0 & 0 \\ 0 & -1 & 0.8 & 0 \\ 0 & 0 & i & 0.8 \\ 1 & 0 & 0 & -i \end{pmatrix}.$$

By inspection, B_3 is not an SDD matrix, but it is doubly SDD. Therefore, its generalized Geršgorin set $\Gamma(A_3, B_3)$ is unbounded; actually, in this case it is a part of right half-plane with the boundary on the line $x = 0.5$, while the generalized Brauer set is a compact one. This is shown in the Figure 3.3.4. Again, Generalized Eigenvalues are marked with "$\times$".

As it was the case with generalized Geršgorin set, this localization is sufficiently hard to calculate, too. So, we use the approximation obtained by Theorem 3.3.2:

$$\begin{cases} \widehat{\mathcal{K}}_{i,j}(A, B) := \{z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}| \cdot |zb_{j,j} - a_{j,j}| \leq \\ \qquad (|z|r_i(B) + r_i(A)) \cdot (|z|r_j(B) + r_j(A))\}, \quad (i, j \in N), (j \neq i), \\ \widehat{\mathcal{K}}(A, B) := \bigcup_{i \in N} \bigcup_{j=1}^{i-1} \widehat{\mathcal{K}}_{i,j}(A, B), \end{cases} \tag{3.3.7}$$

and the next three theorems follow immediately.

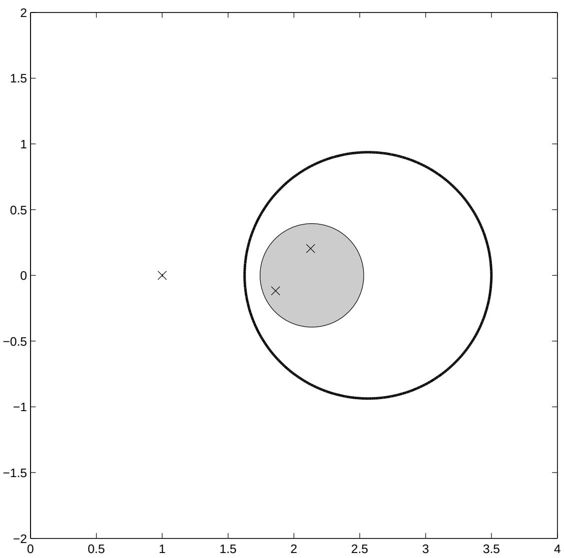

Figure 3.3.1: Generalized Brauer set (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_1, B_1) of the Example 3.2.4
(Generalizovani Brauerov skup (osenčen) i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_1, B_1) iz Primera 3.2.4)

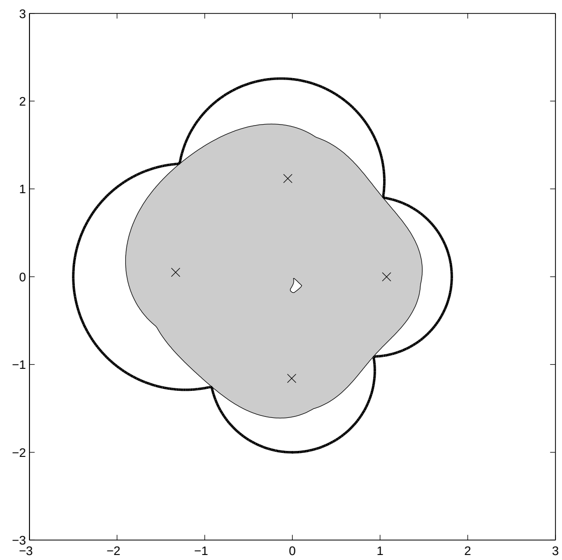

Figure 3.3.2: Generalized Brauer set (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_2, B_2) of the Example 3.2.4
(Generalizovani Brauerov skup (osenčen) i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_2, B_2) iz Primera 3.2.4)

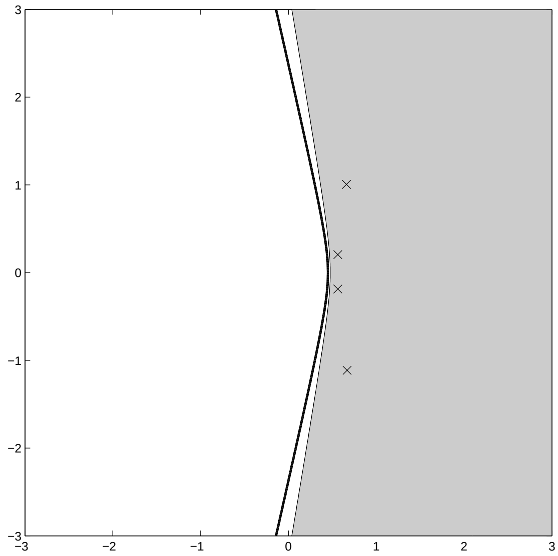

Figure 3.3.3: Generalized Brauer set (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_1, A_2) of the Example 3.2.4

(Generalizovani Brauerov skup (osenčen) i generalizovani Geršgorinov skup (deblja linija) za matrični par (A_1, A_2) iz Primera 3.2.4)

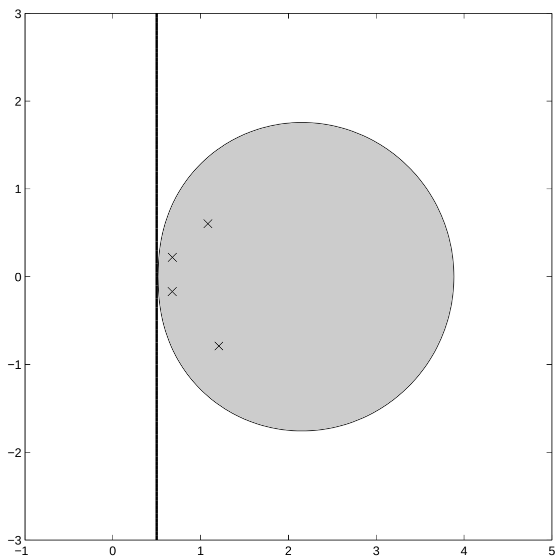

Figure 3.3.4: Generalized Brauer set (shaded) and the generalized Geršgorin set (thick black line) of the matrix pair (A_3, B_3) of the Example 3.3.14

(Generalizovani Brauerov skup (osenčen) i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_3, B_3) iz Primera 3.3.14)

Theorem 3.3.15. *Given a regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the spectrum of the pair* (A,B) *belongs to the* ***approximated generalized Brauer set*** *(3.3.7) of the matrix pair* (A,B)*. Moreover, the following inclusion holds:*

$$\sigma(A,B) \subseteq \mathcal{K}(A,B) \subseteq \widehat{\mathcal{K}}(A,B). \tag{3.3.8}$$

Theorem 3.3.16. *Given an arbitrary pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U, V \subseteq \mathbb{C}$*, such that* $U \cap V = \emptyset$*, and* $\widehat{\mathcal{K}}(A) = U \cup V$*, then the set* U *contains exactly* $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *finite eigenvalues of the pair* (A,B) *and, if* U *is unbounded, exactly* $|\{i \in N \,:\, b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *eigenvalues of the pair* (A,B)*.*

Theorem 3.3.17. *For any regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \widehat{\mathcal{K}}(A,B)$ *if and only if* A *is doubly SDD, and*
- $\infty \notin \widehat{\mathcal{K}}(A,B)$ *if and only if* B *is doubly SDD.*

Obviously, the complexity of the generalized Brauer sets is greater then the complexity of generalized Geršgorin sets. So, the question is when is necessary to use this localization. The first hint can give us the Boundedness Principle. Namely, as we have seen in the first chapter, the class of doubly SDD matrices is wider then the class of SDD matrices. Thus, there are many cases when for a given regular pair (A,B), one of the matrices, let's say B, will *not* be an SDD matrix, while it *may be* in the class of doubly SDD matrices. Then, by the Boundedness Principle, generalized Geršgorin set for this pair will be unbounded, meaning that it will contain ∞. On the other hand, Theorem 1.2.1 implies that matrix B is nonsingular, and therefore $\infty \notin \sigma(A,B)$. So, the generalized Geršgorin set for this pair is *"unbounded without a reason"*, meaning that there is no infinite eigenvalue to be localized. At the same time, Boundedness Principle implies that generalized Brauer set *is bounded*. Thus, in this particular situation, extra work that is needed to construct a Brauer set is worthwhile. The same reasoning can be applied for the approximated sets, too, as Theorem 3.3.10 implies.

We illustrate the approximated generalized Brauer sets, using the matrices of the Examples 3.2.4 and 3.3.14. In the Figures 3.3.5, 3.3.6, 3.3.7 and 3.3.8, generalized Brauer sets $\mathcal{K}(A_1,B_1)$, $\mathcal{K}(A_2,B_2)$, $\mathcal{K}(A_1,A_2)$ and $\mathcal{K}(A_3,B_3)$ are shaded, respectively, while the boundary of the corresponding approximated generalized Brauer set, $\widehat{\mathcal{K}}(A_1,B_1)$, $\widehat{\mathcal{K}}(A_2,B_2)$, $\widehat{\mathcal{K}}(A_1,A_2)$ and $\widehat{\mathcal{K}}(A_3,B_3)$, is given by the thick black line.

The relationship of the ordinary Geršgorin set, and Brauer set that is given in Theorem 2.2.6, can be easily extended to the (approximated) generalized Brauer set and the (approximated) generalized Geršgorin set.

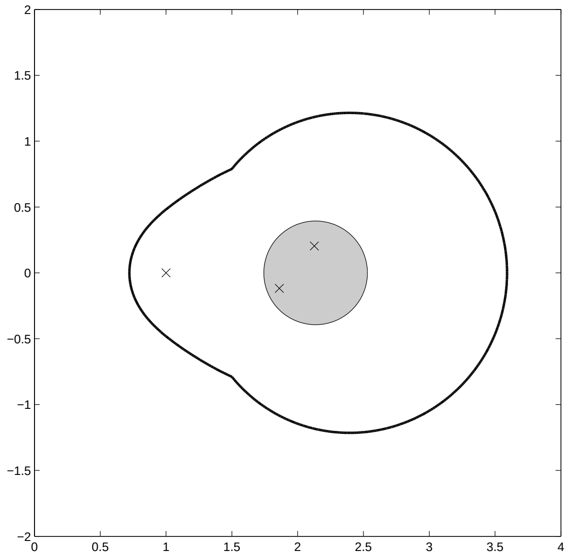

Figure 3.3.5: Generalized Brauer set (shaded), and the approximated generalized Brauer set (thick black line) of the matrix pair (A_1, B_1) of the Example 3.2.4
(Generalizovani Brauerov skup (osenčen) i aproksimirani generalizovani Brauerov skup (debljom linijom) za matrični par (A_1, B_1) *iz Primera 3.2.4)*

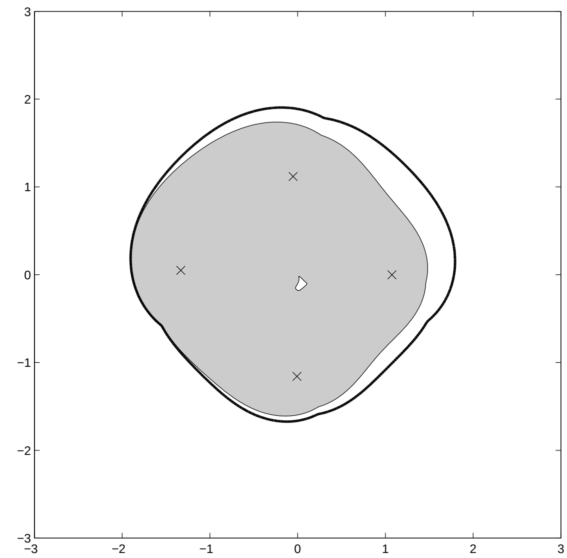

Figure 3.3.6: Generalized Brauer set (shaded), and the approximated generalized Brauer set (thick black line) of the matrix pair (A_2, B_2) of the Example 3.2.4
(Generalizovani Brauerov skup (osenčen) i aproksimirani generalizovani Brauerov skup (debljom linijom) za matrični par (A_2, B_2) *iz Primera 3.2.4)*

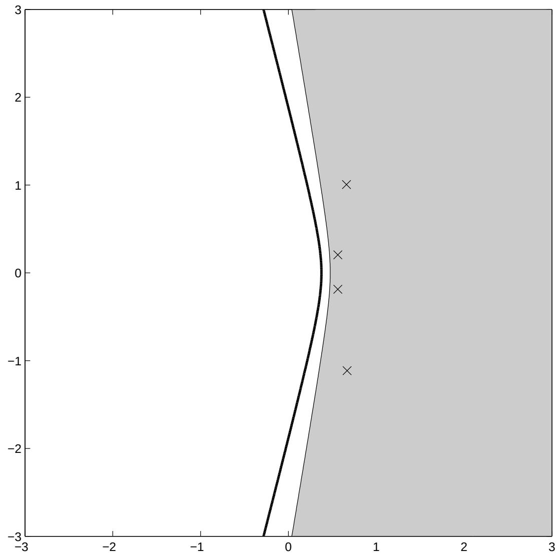

Figure 3.3.7: Generalized Brauer set (shaded), and the approximated generalized Brauer set (thick black line) of the matrix pair (A_1, A_2) of the Example 3.2.4
(Generalizovani Brauerov skup (osenčen) i aproksimirani generalizovani Brauerov skup (debljom linijom) za matrični par (A_1, A_2) *iz Primera 3.2.4)*

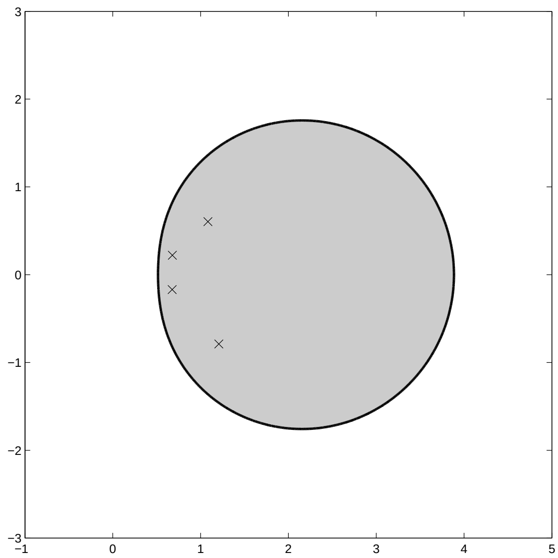

Figure 3.3.8: Generalized Brauer set (shaded), and the approximated generalized Brauer set (thick black line) of the matrix pair (A_3, B_3) of the Example 3.3.14
(Generalizovani Brauerov skup (osenčen) i aproksimirani generalizovani Brauerov skup (debljom linijom) za matrični par (A_3, B_3) *iz Primera 3.3.14)*

Theorem 3.3.18. *For any regular pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, *and any two indices* $i, j \in N$, $i \neq j$,

$$K_{i,j}(A,B) \subseteq \Gamma_i(A,B) \cap \Gamma_j(A,B), \quad \text{and}$$
$$\widehat{K}_{i,j}(A,B) \subseteq \widehat{\Gamma}_i(A,B) \cap \widehat{\Gamma}_j(A,B).$$

Therefore, consequently,

$$\mathcal{K}(A,B) \subseteq \Gamma(A,B), \quad \text{and}$$
$$\widehat{\mathcal{K}}(A,B) \subseteq \widehat{\Gamma}(A,B).$$

The proof is almost identical as in the case of standard eigenvalue localizations, and can be found in [51].

3.3.2 Brualdi set for Generalized Eigenvalues

In order to construct Brualdi sets for generalized eigenvalues, we must take the graph structure of a "suitable" matrix into account. In the ordinary case, since the complex number z was influencing only on the diagonal of a given matrix A, the graph structure of $zI - A$ and A were the same, i.e., independent of z. Of course, this is not the case with generalized eigenvalues. Therefore, first, given a regular pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, we extract the points in the complex plane where the graph structure of $zB - A$ is changing due to value of z. We denote

$$\zeta(A,B) := \left\{ \frac{a_{i,j}}{b_{i,j}} \in \mathbb{C} \,:\, b_{i,j} \neq 0,\ i,j \in N,\ i \neq j \right\}. \tag{3.3.9}$$

Then, obviously, for every $z \in \mathbb{C} \setminus \zeta(A,B)$, $\mathbb{G}(zB - A)$ is independent of z, and, hence, we can define a **graph attributed to the matrix pair** $\mathbb{G}(A,B) := \mathbb{G}(zB - A)$, where $z \in \mathbb{C} \setminus \zeta(A,B)$.

As before, $C(A,B)$ will denote the set of all cycles, weak or strong, in the graph $\mathbb{G}(A,B)$.

Example 3.3.19. *Let*

$$A_4 = \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0.5 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad B_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0.5 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i & 0 & 0 \\ 0.5 & 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

In the Figure 3.3.9, which represents the graph $\mathbb{G}(A_4, B_4)$, *attributed to the given matrix pair, we can observe that* $\zeta(A_4, B_4) = \{0.5, 1\}$, *and that the strong cycles are* $\gamma_1 = (1,2,3)$ *and* $\gamma_2 = (3,4)$, *while the weak cycles are* $\gamma_3 = (5)$ *and* $\gamma_4 = (6)$.

An important concept used in Varga's generalization of the original Brualdi result was the concept of the normal reduced form and irreducibility. So, we continue by extending them to the matrix pairs. An elegant way to do it is through the graph theoretic approach.

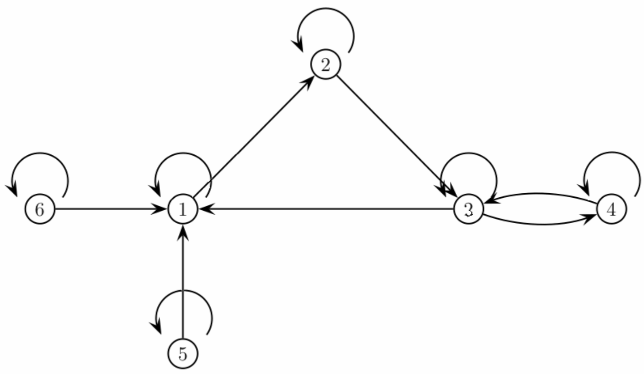

Figure 3.3.9: Graph of a matrix pair (A_4, B_4) of the Example 3.3.19

Definition 3.3.20. An arbitrary $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$ is said to be an **irreducible matrix pair** if the attributed graph $\mathbb{G}(A, B)$ is strongly connected.

Inspired by the case of a single matrix, we define a normal reduced form of the matrix pair. Given a matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, we observe the graph $\mathbb{G}(A, B)$. If the graph is strongly connected, then the matrix pair is irreducible. Otherwise, we can distinguish the partition of the set of vertices into V_1 and V_2, such that for every directed edge $\overrightarrow{v_i, v_j}$, if $v_i \in V_2$, then $v_j \notin V_1$, i.e., there is no path from the vertices of V_2 to the vertices of V_1. But, this implies that there exists a permutation matrix P, such that

$$P^T AP = \begin{bmatrix} A_{1,1} & A_{1,2} \\ O & A_{2,2} \end{bmatrix}, \quad \text{and} \quad P^T BP = \begin{bmatrix} B_{1,1} & B_{1,2} \\ O & B_{2,2} \end{bmatrix}, \tag{3.3.10}$$

where $A_{1,1}$, $A_{2,2}$, $B_{1,1}$ and $B_{2,2}$ are square matrices.

Now, if we continue with this reasoning, and apply it to the sets of vertices V_1 and V_2, and further to their progenies, by consecutive permutations, we ultimately obtain a permutation matrix $\widetilde{P} \in \mathbb{R}^{n,n}$, and a positive integer m, $2 \leq m \leq n$, such that

$$\widetilde{P}^T A\widetilde{P} = \begin{bmatrix} \widetilde{A}_{1,1} & \widetilde{A}_{1,2} & \cdots & \widetilde{A}_{1,m} \\ O & \widetilde{A}_{2,2} & \cdots & \widetilde{A}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \widetilde{A}_{m,m} \end{bmatrix}, \quad \text{and} \quad \widetilde{P}^T B\widetilde{P} = \begin{bmatrix} \widetilde{B}_{1,1} & \widetilde{B}_{1,2} & \cdots & \widetilde{B}_{1,m} \\ O & \widetilde{B}_{2,2} & \cdots & \widetilde{B}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ O & O & \cdots & \widetilde{B}_{m,m} \end{bmatrix}, \tag{3.3.11}$$

where

$$\begin{aligned} &\widetilde{A}_{i,i}, \widetilde{B}_{i,i} \in \mathbb{C}^{p_i,p_i}, \quad \text{are irreducible matrices for } p_i \geq 2, \text{ or} \\ &\widetilde{A}_{i,i} = [a_{k,k}], \widetilde{B}_{i,i} = [b_{k,k}] \in \mathbb{C}^{1,1}, \quad \text{for some } k \in N. \end{aligned} \tag{3.3.12}$$

for every $1 \leq i \leq m$.

The permutation matrix $\widetilde{P}$ is obtained as a product of individual permutation matrices, which corresponds, at each step, to a splitting of diagonal blocks into the form (3.3.10).

Similar to the single matrix case, the form (3.3.11) we call the **normal reduced form of the matrix pair** (A, B).

For the matrix pair (A_4, B_4) of the Example 3.3.19, the normal reduced form is given by

$$\widetilde{P}^T A_4 \widetilde{P} = \left[\begin{array}{c|c|cccc} 0 & 0 & 0 & 0.5 & 0 & 0 \\ \hline 0 & 1 & 0 & 0.5 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right], \text{ and } \widetilde{P}^T B_4 \widetilde{P} = \left[\begin{array}{c|c|cccc} -i & 0 & 0 & 0.5 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & i \end{array}\right], \tag{3.3.13}$$

where $\widetilde{P}$ is a permutation matrix that corresponds to the permutation:

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 3 & 5 & 6 & 1 & 2 \end{pmatrix}.$$

As in single matrix case, we can observe that, for any $z \in \mathbb{C}$, matrix $zB - A$ is singular if and only if, for some $1 \leq i \leq m$, matrix $z\widetilde{B}_{i,i} - \widetilde{A}_{i,i}$ is singular. Therefore, we can use the reduced row sums, as defined in (1.2.10). Having this, by restricting from the whole complex plane to the set $\mathbb{C} \setminus \zeta(A, B)$, since the class of Brualdi SDD matrices is a positively homogenous SDD-type class, we can apply Theorems 3.3.1-3.3.4, Theorem 3.3.9 and Theorem 3.3.10.

Theorem 3.3.21. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, for every $\lambda \in \sigma(A, B) \setminus \zeta(A, B)$, there exists a cycle $\gamma \in C(A, B)$, either strong or weak, so that*

$$\lambda \in \mathcal{B}_\gamma(A, B) := \left\{ z \in \mathbb{C} \,:\, \prod_{i \in \gamma} |zb_{i,i} - a_{i,i}| \leq \prod_{i \in \gamma} \widetilde{r}_i(zB - A) \right\}, \tag{3.3.14}$$

if the cycle $\gamma \in C(A, B)$ is strong, or,

$$\lambda \in \mathcal{B}_\gamma(A, B) := \{ z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}| \leq \widetilde{r}_i(zB - A) = 0 \} = \left\{ a_{i,i} b_{i,i}^{-1} \right\}, \tag{3.3.15}$$

with the convention that $0^{-1} = \infty$, if the cycle $\gamma = \{i\} \in C(A, B)$ is weak.

Consequently,

$$\sigma(A, B) \subseteq \mathcal{B}(A, B) := \bigcup_{\gamma \in C(A,B)} \mathcal{B}_\gamma(A, B) \bigcup \zeta(A, B). \tag{3.3.16}$$

Theorem 3.3.22. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, for every $\lambda \in \sigma(A, B) \setminus \zeta(A, B)$, there exists a cycle $\gamma \in C(A, B)$, either strong or weak, so that*

$$\lambda \in \widehat{\mathcal{B}}_\gamma(A, B) := \left\{ z \in \mathbb{C} \,:\, \prod_{i \in \gamma} |zb_{i,i} - a_{i,i}| \leq \prod_{i \in \gamma} (|z| \widetilde{r}_i(B) + \widetilde{r}_i(A)) \right\}, \tag{3.3.17}$$

if the cycle $\gamma \in C(A,B)$ *is strong, or,*

$$\lambda \in \widehat{\mathcal{B}}_\gamma(A,B) := \{z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}| \leq |z|\widetilde{r}_i(B) + \widetilde{r}_i(A) = 0\} = \left\{a_{i,i}b_{i,i}^{-1}\right\}, \qquad (3.3.18)$$

with the convention that $0^{-1} = \infty$*, if the cycle* $\gamma = \{i\} \in C(A,B)$ *is weak.*
Consequently,

$$\sigma(A,B) \subseteq \mathcal{B}(A,B) \subseteq \widehat{\mathcal{B}}(A,B) := \bigcup_{\gamma \in C(A,B)} \widehat{\mathcal{B}}_\gamma(A,B) \bigcup \zeta(A,B). \qquad (3.3.19)$$

As before, the Isolation and Boundedness Principles hold for both localization areas.

Theorem 3.3.23. *Given an arbitrary pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U,V \subseteq \mathbb{C}$*, such that* $U \cap V = \emptyset$*, and* $\mathcal{B}(A,B) = U \cup V$*, then the set* U *contains exactly* $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *finite eigenvalues of the pair* (A,B) *and, if* U *is unbounded, exactly* $|\{i \in N \,:\, b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *eigenvalues of the pair* (A,B)*.*

Theorem 3.3.24. *For any regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \mathcal{B}(A,B)$ *if and only if* A *is Brualdi SDD, and*
- $\infty \notin \mathcal{B}(A,B)$ *if and only if* B *is Brualdi SDD.*

Theorem 3.3.25. *Given an arbitrary pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U,V \subseteq \mathbb{C}$*, such that* $U \cap V = \emptyset$*, and* $\widehat{\mathcal{B}}(A) = U \cup V$*, then the set* U *contains exactly* $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *finite eigenvalues of the pair* (A,B) *and, if* U *is unbounded, exactly* $|\{i \in N \,:\, b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N \,:\, a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *eigenvalues of the pair* (A,B)*.*

Theorem 3.3.26. *For any regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \widehat{\mathcal{B}}(A,B)$ *if and only if* A *is Brualdi SDD, and*
- $\infty \notin \widehat{\mathcal{B}}(A,B)$ *if and only if* B *is Brualdi SDD.*

To illustrate the generalized Brualdi set and its approximation, consider the matrix pair (A_4, B_4) of the Example 3.3.19 in its normal reduced form (3.3.13). The set of cycles in the graph $\mathbb{G}(A_4,B_4)$ is $C(A_4,B_4) = \{\gamma_1,\gamma_2,\gamma_3,\gamma_4\}$, where $\gamma_1 = (3,5,4)$, $\gamma_2 = (5,6)$, $\gamma_3 = (2)$, and $\gamma_4 = (1)$. Thus, we have that the generalized Brualdi set is given by $\mathcal{B}(A_4,B_4) = \bigcup_{i=1}^4 \mathcal{B}_{\gamma_i}(A_4,B_4)$, where:

$$\begin{aligned}
\mathcal{B}_{\gamma_1}(A_4,B_4) &= \left\{z \in \mathbb{C} \,:\, |z||z^2-1| \leq 0.125|2z-1|(|z|+1)\right\} \\
\mathcal{B}_{\gamma_2}(A_4,B_4) &= \{z \in \mathbb{C} \,:\, 2|z||z+1| \leq |z-1|(|z|+1)\} \\
\mathcal{B}_{\gamma_3}(A_4,B_4) &= \{1\}, \quad \text{and} \quad \mathcal{B}_{\gamma_4}(A_4,B_4) = \{0\}.
\end{aligned}$$

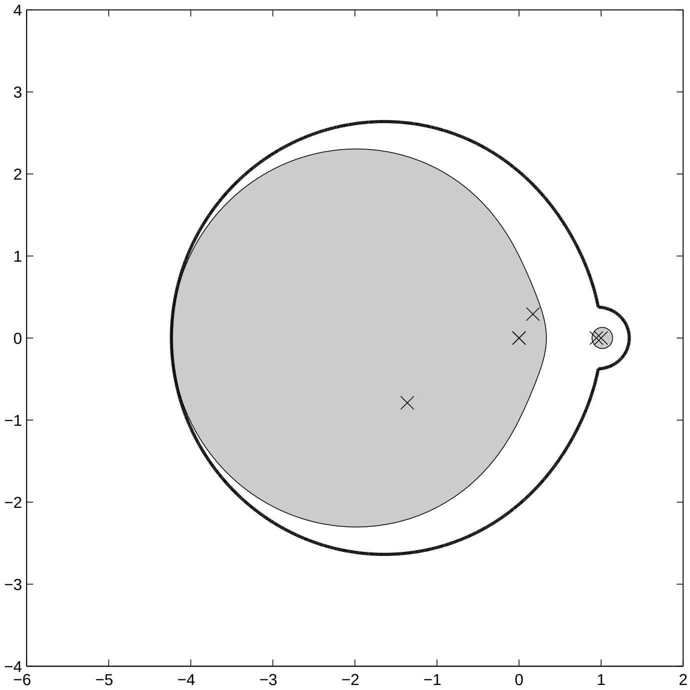

Figure 3.3.10: Generalized Brualdi set $\mathcal{B}(A_4, B_4)$ of the matrix pair (A_4, B_4) of the Example 3.3.19 (shaded), and the approximated generalized Brualdi set $\widehat{\mathcal{B}}(A_4, B_4)$ (thick black line) *(Generalizovani Brualdijev skup $\mathcal{B}(A_4, B_4)$ (osenčen) i aproksimirani generalizovani Brualdijev skup $\widehat{\mathcal{B}}(A_4, B_4)$ (debljom linijom) za matrični par (A_4, B_4) iz Primera 3.3.19)*

Similar to that, the approximated generalized Brualdi set is $\widehat{\mathcal{B}}(A_4, B_4) = \bigcup_{i=1}^{4} \widehat{\mathcal{B}}_{\gamma_i}(A_4, B_4)$, where:

$$\begin{aligned}
\widehat{\mathcal{B}}_{\gamma_1}(A_4, B_4) &= \left\{z \in \mathbb{C} \,:\, |z||z^2-1| \le 0.125(2|z|+1)(|z|+1)\right\}, \\
\widehat{\mathcal{B}}_{\gamma_2}(A_4, B_4) &= \left\{z \in \mathbb{C} \,:\, 2|z||z+1| \le (|z|+1)^2\right\}, \\
\widehat{\mathcal{B}}_{\gamma_3}(A_4, B_4) &= \{1\}, \quad \text{and} \quad \widehat{\mathcal{B}}_{\gamma_4}(A_4, B_4) = \{0\}.
\end{aligned}$$

By inspection, B_4 is a Brualdi SDD matrix and, according to Theorem 3.3.24, the set $\mathcal{B}(A_4, B_4)$ is compact in the complex plane. The same holds for the approximated generalized Brualdi set $\widehat{\mathcal{B}}(A_4, B_4)$, from the Theorem 3.3.26. These sets are shown in Figure 3.3.10, where the generalized Brualdi set is shaded, its approximation has thick black boundary, and Generalized Eigenvalues are marked with "$\times$".

3.3.3 Cvetković-Kostić-Varga set for Generalized Eigenvalues

Here we introduce the generalized CKV-set and its approximation. As in Section 2.2.3, for the case of the "ordinary" eigenvalues, we show their connection with the generalized minimal Geršgorin set, which will be considered in detail in the concluding section of this chapter.

We start with the observation that the class of S-SDD matrices, and consequently, the class of CKV-SDD matrices, is a positively homogenous SDD-type class. Therefore, Theorems 3.3.1, 3.3.4, and 3.3.9 imply the following localization areas for Generalized Eigenvalues.

Theorem 3.3.27. *Let $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, with $n \geq 2$, be any regular matrix pair, and let $\lambda \in \sigma(A, B)$ be an arbitrary, either finite or infinite, eigenvalue. Then, for every nonempty subset of indices $S \subseteq N$, there exist indices $i \in S$, and $j \in \overline{S} := N \setminus S$, such that*

$$\lambda \in \Gamma_i^S(A, B) := \left\{z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}| \leq r_i^S(zB - A)\right\}, \tag{3.3.20}$$

or

$$\begin{aligned}\lambda \in V_{i,j}^S(A, B) := \{z \in \mathbb{C} \,:\, \\ (|zb_{i,i} - a_{i,i}| - r_i^S(zB - A)) \cdot (|zb_{j,j} - a_{j,j}| - r_j^{\overline{S}}(zB - A)) \\ \leq r_i^{\overline{S}}(zB - A) r_j^S(zB - A)\}.\end{aligned} \tag{3.3.21}$$

Therefore, for every nonempty subset of indices $S \subseteq N$,

$$\sigma(A, B) \subseteq \mathcal{C}^S(A, B) := [\bigcup_{i \in S} \bigcup_{j \in \overline{S}} V_{i,j}^S(A, B)] \bigcup [\bigcup_{i \in S} \Gamma_i^S(A, B)], \tag{3.3.22}$$

and consequently

$$\sigma(A, B) \subseteq \mathcal{C}(A, B) := \bigcap_{\emptyset \neq S \subseteq N} \mathcal{C}^S(A, B). \tag{3.3.23}$$

Theorem 3.3.28. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, an arbitrary nonempty set of indices $S \subseteq N$, and sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, if $\mathcal{C}^S(A, B) = U \cup V$, or if $\mathcal{C}(A, B) = U \cup V$, then the set U contains exactly $|\left\{i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i} b_{i,i}^{-1} \in U\right\}|$ finite eigenvalues of the pair (A, B) and, if U is unbounded, exactly $|\{i \in N \,:\, b_{i,i} = 0\}|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N \,:\, a_{i,i} b_{i,i}^{-1} \in U\right\}|$ eigenvalues of the pair (A, B).*

Theorem 3.3.29. *For any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following conditions hold:*

- $0 \notin \mathcal{C}^S(A, B)$ *if and only if A is an S-SDD matrix, where $\emptyset \neq S \subseteq N$,*
- $\infty \notin \mathcal{C}^S(A, B)$ *if and only if B is an S-SDD matrix, again $\emptyset \neq S \subseteq N$,*
- $0 \notin \mathcal{C}(A, B)$ *if and only if A is a CKV-SDD matrix, and*
- $\infty \notin \mathcal{C}(A, B)$ *if and only if B is a CKV-SDD matrix.*

As before, we illustrate the generalized CKV-sets using the matrices of the Examples 3.2.4 and 3.3.14. In the Figures 3.3.11, 3.3.12, 3.3.13 and 3.3.14, generalized CKV-sets $\mathcal{C}^{\{1,4\}}(A_1, B_1)$, $\mathcal{C}^{\{1\}}(A_2, B_2)$, $\mathcal{C}^{\{1,4\}}(A_1, A_2)$, and $\mathcal{C}^{\{1\}}(A_3, B_3)$ are shaded, respectively, while the boundaries of the corresponding approximated generalized Geršgorin sets, $\Gamma(A_1, B_1)$, $\Gamma(A_2, B_2)$, $\Gamma(A_1, A_2)$ and $\Gamma(A_3, B_3)$, are given by the thick black line.

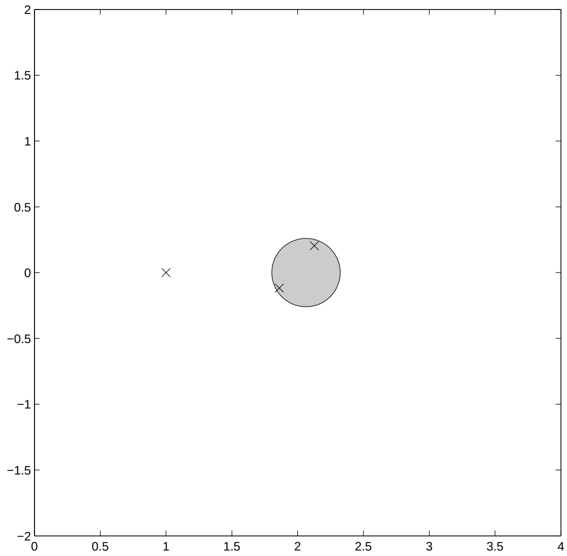

Figure 3.3.11: Generalized CKV-set for fixed $S = \{1, 4\}$ (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_1, B_1) of the Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{1, 4\}$ (osenčen), i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_1, B_1) iz Primera 3.2.4)

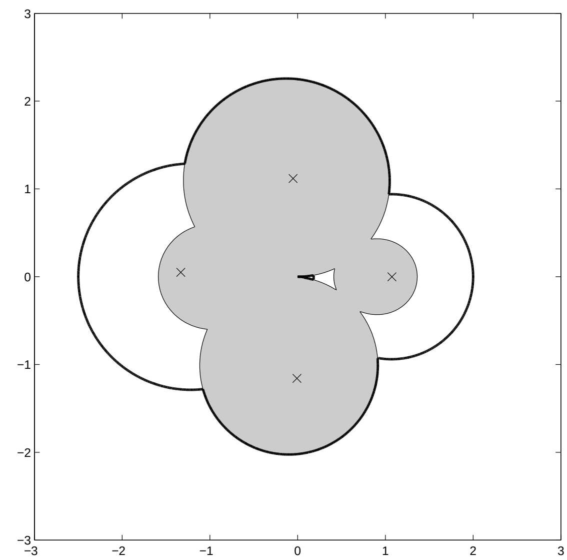

Figure 3.3.12: Generalized CKV-set for fixed $S = \{1\}$ (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_2, B_2) of the Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{2\}$ (osenčen), i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_2, B_2) iz Primera 3.2.4)

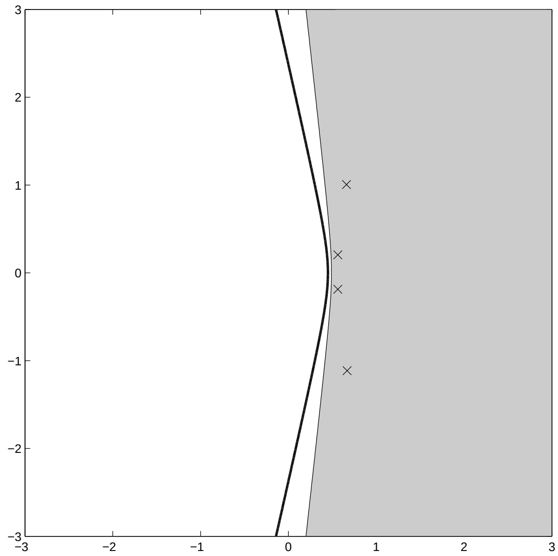

Figure 3.3.13: Generalized CKV-set for fixed $S = \{1, 4\}$ (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_1, A_2) of the Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{1, 4\}$ (osenčen), i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_1, A_2) iz Primera 3.2.4)

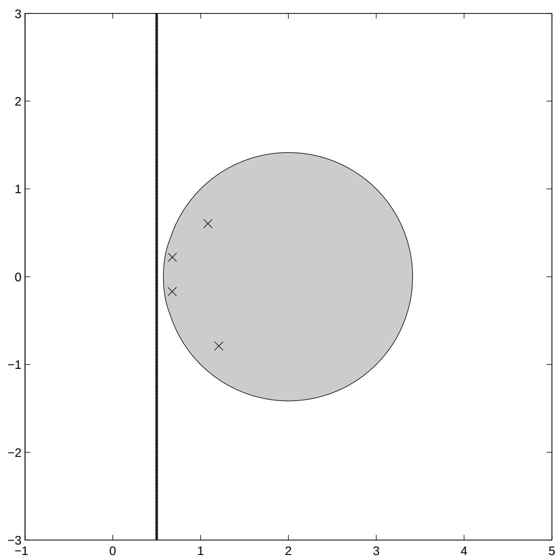

Figure 3.3.14: Generalized CKV-set for fixed $S = \{1\}$ (shaded), and the generalized Geršgorin set (thick black line) of the matrix pair (A_3, B_3) of the Example 3.3.14
(Generalizovani CKV-skup za fiksiran skup $S = \{1\}$ (osenčen), i generalizovani Geršgorinov skup (debljom linijom) za matrični par (A_3, B_3) iz Primera 3.3.14)

Using Theorem 3.3.10 we introduce the approximation of the CKV-set that is suitable for practical application. In addition, we state its properties which follow from Theorems 3.3.4 and 3.3.10, and, finally, we compare the set and its approximation through illustrative examples.

Theorem 3.3.30. *Let $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, with $n \geq 2$, be any regular matrix pair, and let $\lambda \in \sigma(A,B)$ be an arbitrary, either finite or infinite, eigenvalue. Then, for every nonempty subset of indices $S \subseteq N$, there exist indices $i \in S$, and $j \in \overline{S} := N \setminus S$, such that*

$$\lambda \in \widehat{\Gamma}_i^S(A,B) := \left\{ z \in \mathbb{C} \,:\, |zb_{i,i} - a_{i,i}| \leq |z| r_i^S(B) + r_i^S(A) \right\}, \tag{3.3.24}$$

or

$$\begin{aligned} \lambda \in \widehat{V}_{i,j}^S(A,B) := \{z \in \mathbb{C} \,:\, & \\ (|zb_{i,i} - a_{i,i}| - |z| r_i^S(B) - r_i^S(A)) \cdot (|zb_{j,j} - a_{j,j}| - |z| r_j^{\overline{S}}(B) - r_j^{\overline{S}}(A)) & \\ \leq (|z| r_i^{\overline{S}}(B) + r_i^{\overline{S}}(A)) \cdot (|z| r_j^S(B) + r_j^S(A))\}. & \end{aligned} \tag{3.3.25}$$

Therefore, for every nonempty subset of indices $S \subseteq N$,

$$\sigma(A,B) \subseteq \widehat{\mathcal{C}}^S(A,B) := \Big[\bigcup_{i\in S} \bigcup_{j \in \overline{S}} \widehat{V}_{i,j}^S(A,B)\Big] \bigcup \Big[\bigcup_{i \in S} \widehat{\Gamma}_i^S(A,B)\Big], \tag{3.3.26}$$

and consequently

$$\sigma(A,B) \subseteq \widehat{\mathcal{C}}(A,B) := \bigcap_{\emptyset \neq S \subseteq N} \widehat{\mathcal{C}}^S(A,B). \tag{3.3.27}$$

Theorem 3.3.31. *Given any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, an arbitrary nonempty set of indices $S \subseteq N$, and sets $U, V \subseteq \mathbb{C}$. such that $U \cap V = \emptyset$, if $\widehat{\mathcal{C}}^S(A,B) = U \cup V$, or if $\widehat{\mathcal{C}}(A,B) = U \cup V$, then the set U contains exactly $|\left\{ i \in N \,:\, b_{i,i} \neq 0 \text{ and } a_{i,i} b_{i,i}^{-1} \in U \right\}|$ finite eigenvalues of the pair (A,B) and, if U is unbounded, exactly $|\{i \in N \,:\, b_{i,i} = 0\}|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{ i \in N \,:\, a_{i,i} b_{i,i}^{-1} \in U \right\}|$ eigenvalues of the pair (A,B).*

Theorem 3.3.32. *For any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following conditions hold:*

- $0 \notin \widehat{\mathcal{C}}^S(A,B)$ *if and only if A is S-SDD matrix, where $\emptyset \neq S \subseteq N$,*
- $\infty \notin \widehat{\mathcal{C}}^S(A,B)$ *if and only if B is S-SDD matrix, again $\emptyset \neq S \subseteq N$,*
- $0 \notin \widehat{\mathcal{C}}(A,B)$ *if and only if A is CKV-SDD matrix, and*
- $\infty \notin \widehat{\mathcal{C}}(A,B)$ *if and only if B is CKV-SDD matrix.*

How the approximated generalized CKV-sets compare with the original generalized CKV-sets is shown in the Figures 3.3.15, 3.3.16, 3.3.17 and 3.3.18, where the generalized CKV-sets $\mathcal{C}^{\{1,4\}}(A_1,B_1)$, $\mathcal{C}^{\{1\}}(A_2,B_2)$, $\mathcal{C}^{\{1,4\}}(A_1,A_2)$ and $\mathcal{C}^{\{1\}}(A_3,B_3)$ are shaded, respectively, while the boundaries of the corresponding approximated generalized CKV-sets, $\widehat{\mathcal{C}}^{\{1,4\}}(A_1,B_1)$, $\widehat{\mathcal{C}}^{\{1\}}(A_2,B_2)$, $\widehat{\mathcal{C}}^{\{1,4\}}(A_1,A_2)$ and $\widehat{\mathcal{C}}^{\{1\}}(A_3,B_3)$, are given by the thick black

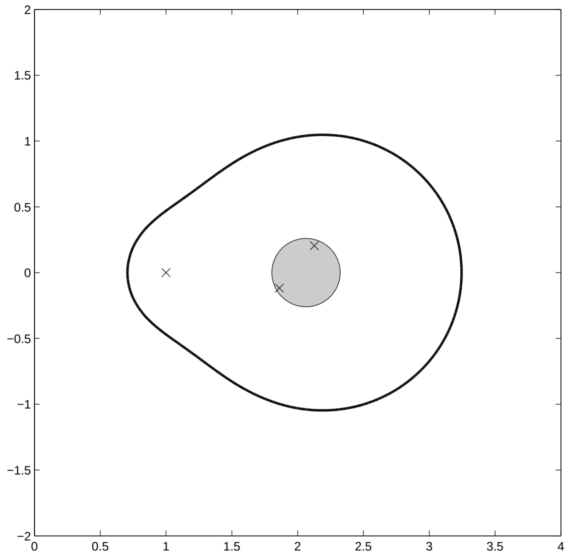

Figure 3.3.15: Generalized CKV-set for fixed $S = \{1, 4\}$ (shaded), and its approximation (thick black line) of the matrix pair (A_1, B_1) of Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{1, 4\}$ (osenčen) i njegova aproksimacija (debljom linijom) za matrični par (A_1, B_1) iz Primera 3.2.4)

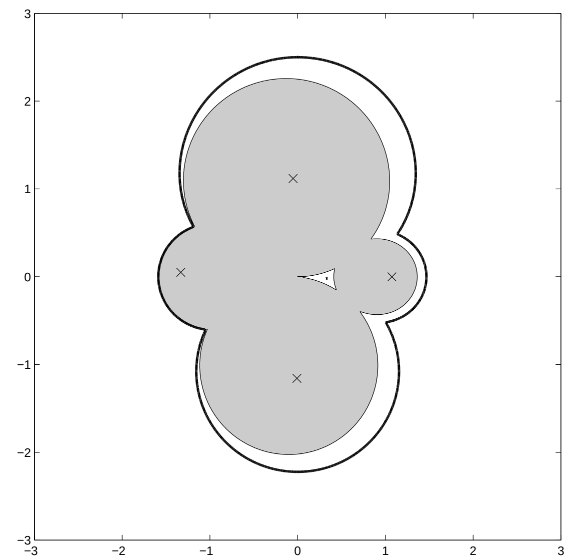

Figure 3.3.16: Generalized CKV-set for fixed $S = \{1\}$ (shaded), and its approximation (thick black line) of the matrix pair (A_2, B_2) of Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{1\}$ (osenčen) i njegova aproksimacija (debljom linijom) za matrični par (A_2, B_2) iz Primera 3.2.4)

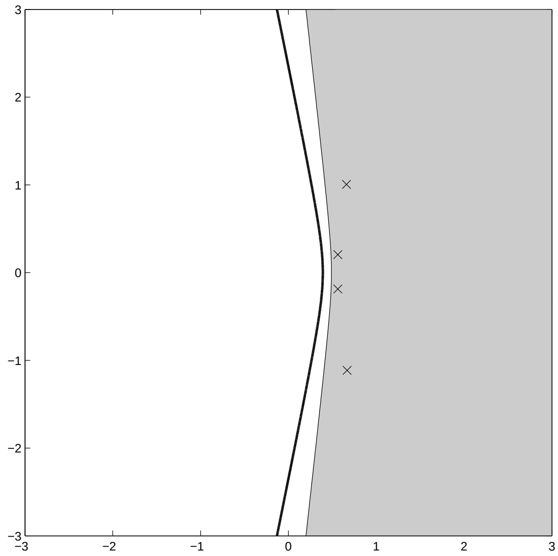

Figure 3.3.17: Generalized CKV-set for fixed $S = \{1, 4\}$ (shaded), and its approximation (thick black line) of the matrix pair (A_1, A_2) of Example 3.2.4
(Generalizovani CKV-skup za fiksiran skup $S = \{1, 4\}$ (osenčen) i njegova aproksimacija (debljom linijom) za matrični par (A_1, A_2) iz Primera 3.2.4)

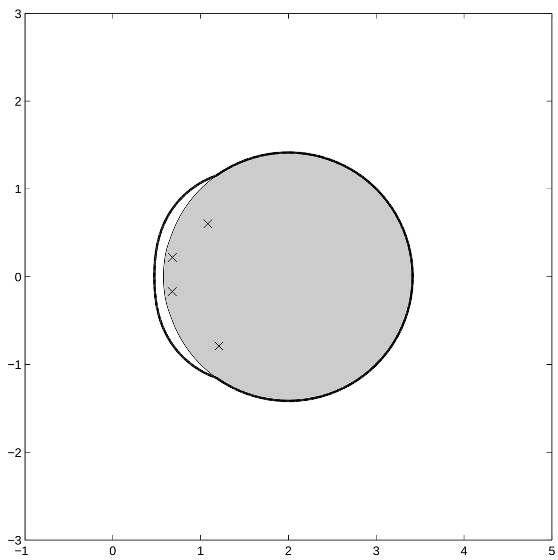

Figure 3.3.18: Generalized CKV-set for fixed $S = \{1\}$ (shaded), and its approximation (thick black line) of the matrix pair (A_3, B_3) of Example 3.3.14
(Generalizovani CKV-skup za fiksiran skup $S = \{1\}$ (osenčen) i njegova aproksimacija (debljom linijom) za matrični par (A_3, B_3) iz Primera 3.3.14)

line. The matrices A_1-A_3 and B_1-B_3 are the same as in Examples 3.2.4 and 3.3.14, respectively.

Similar to the Example 3.3.14, if matrix B in the regular matrix pair (A, B), is S-SDD, for some fixed set of indices S, while it is not doubly SDD, then the (approximated) generalized CKV-set, for such S, will be bounded in the complex plane, while the (approximated) generalized Brauer set will be unbounded. The similar thing stands for generalized Brualdi sets.

Observing that, for an arbitrary nonempty $S \subseteq N$, $\mathbb{K}^{\mathbb{X}_S}$ and $\mathbb{K}^{\mathbb{X}_{CKV}}$ are the classes of S-SDD and CKV-SDD matrices, respectively. Using the fact that these classes are also of SDD-type, Theorem 3.3.1 produces different equivalent form of Theorem 1.4.2 in terms of the localization sets of generalized eigenvalues. Moreover, we give the general idea of the *generalized minimal Geršgorin set attributed to the family* $\mathbb{X} \subseteq \mathbb{D}$. Namely, given a family of positive diagonal matrices $\mathbb{X} \subseteq \mathbb{D}$, we define:

$$\Gamma^{\mathbb{X}}(A, B) := \bigcap_{X \in \mathbb{X}} \Gamma(X^{-1}AX, X^{-1}BX), \tag{3.3.28}$$

and call it the **generalized minimal Geršgorin set attributed to the family** $\mathbb{X}$.

Theorem 3.3.33. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, and an arbitrary nonempty subset of indices $S \subseteq N$, then*

$$\mathcal{C}^S(A, B) = \Gamma^{\mathbb{X}_S}(A, B), \tag{3.3.29}$$

and consequently,

$$\mathcal{C}(A, B) = \Gamma^{\mathbb{X}_{CKV}}(A, B), \tag{3.3.30}$$

where $\mathbb{X}_S$ and $\mathbb{X}_{CKV}$ are given in (1.4.18) and (1.4.19), respectively.

In other words, the theorem states that the generalized minimal Geršgorin set attributed to the family $\mathbb{X}_S$ is equal to the set $\mathcal{C}^S(A)$, and that the generalized minimal Geršgorin set attributed to the family $\mathbb{X}_{CKV}$ is equal to the set $\mathcal{C}(A)$.

Now, since $X^{-1}\langle A, B\rangle X = \langle X^{-1}AX, X^{-1}BX\rangle$, for any $X \in \mathbb{D}$, we have that the same thing follows for the approximated sets. More precisely, given a family of positive diagonal matrices $\mathbb{X} \subseteq \mathbb{D}$, we define:

$$\widehat{\Gamma}^{\mathbb{X}}(A, B) := \bigcap_{X \in \mathbb{X}} \widehat{\Gamma}(X^{-1}AX, X^{-1}BX), \tag{3.3.31}$$

and call it the **approximated generalized minimal Geršgorin set attributed to the family** $\mathbb{X}$. Then, the following holds.

Theorem 3.3.34. *Given any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, and an arbitrary nonempty subset of indices $S \subseteq N$, then*

$$\widehat{\mathcal{C}}^S(A, B) = \widehat{\Gamma}^{\mathbb{X}_S}(A, B), \tag{3.3.32}$$

and consequently,

$$\widehat{\mathcal{C}}(A, B) = \widehat{\Gamma}^{\mathbb{X}_{CKV}}(A, B), \tag{3.3.33}$$

where $\mathbb{X}_S$ and $\mathbb{X}_{CKV}$ are given in (1.4.18) and (1.4.19), respectively.

To complete this section we establish the relationship of these generalized eigenvalue localization sets, which is, in fact, the consequence of the relationship between the subclasses of H-matrices which have generated them. Therefore, as for Theorem 2.2.19, we have the following.

Theorem 3.3.35. *Let $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, be any regular matrix pair, and let the set $\Gamma(A,B)$ be given by (3.2.1), the set $\mathcal{K}(A,B)$ by (3.3.6), the set $\mathcal{B}(A,B)$ by (3.3.16),the set $\mathcal{C}^S(A,B)$ by (3.3.22), and the set $\mathcal{C}(A,B)$ by (3.3.23). Then,*

- $\mathcal{C}^{\{i\}}(A,B) \subseteq \Gamma(A,B), \quad (i \in N)$,
- $\mathcal{C}^S(A,B) \subseteq \Gamma(A,B), \quad (S \subseteq N)$, *and, consequently,*
- $\mathcal{C}(A,B) \subseteq \mathcal{K}(A,B) \subseteq \Gamma(A,B)$.

Moreover, there exist regular matrix pairs $(P_1,P_2),(Q_1,Q_2),(R_1,R_2) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, so that

- $\mathcal{C}(P_1,P_2) \nsubseteq \mathcal{B}(P_1,P_2)$, *and* $\mathcal{B}(P_1,P_2) \nsubseteq \mathcal{C}(P_1,P_2)$,
- $\mathcal{C}^{\{i\}}(Q_1,Q_2) \nsubseteq \mathcal{K}(Q_1,Q_2)$, *and* $\mathcal{K}(Q_1,Q_2) \nsubseteq \mathcal{C}^{\{i\}}(Q_1,Q_2)$, *for some* $i \in N$,
- $\mathcal{C}^S(R_1,R_2) \nsubseteq \mathcal{K}(R_1,R_2)$, *and* $\mathcal{K}(R_1,R_2) \nsubseteq \mathcal{C}^S(R_1,R_2)$, *for some* $S \subseteq N$.

But, as we are interested in calculation of the localization areas, it is preferable to use the approximated sets. Therefore, we are, in fact, more interested in the relationship between the approximated generalized Geršgorin-type sets. Fortunately, from the definition of the the approximated sets (c.f. Theorem 3.3.2) the same relationship holds for them, too.

Theorem 3.3.36. *Let $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, be any regular matrix pair, and let the set $\widehat{\Gamma}(A,B)$ be given by (3.2.11), the set $\widehat{\mathcal{K}}(A,B)$ by (3.3.7), the set $\widehat{\mathcal{B}}(A,B)$ by (3.3.19), the set $\widehat{\mathcal{C}}^S(A,B)$ by (3.3.26), and the set $\widehat{\mathcal{C}}(A,B)$ by (3.3.27). Then,*

- $\widehat{\mathcal{C}}^{\{i\}}(A,B) \subseteq \widehat{\Gamma}(A,B), \quad (i \in N)$,
- $\widehat{\mathcal{C}}^S(A,B) \subseteq \widehat{\Gamma}(A,B), \quad (S \subseteq N)$, *and, consequently,*
- $\widehat{\mathcal{C}}(A,B) \subseteq \widehat{\mathcal{K}}(A,B) \subseteq \widehat{\Gamma}(A,B)$.

Moreover, there exist regular matrix pairs $(P_1,P_2),(Q_1,Q_2),(R_1,R_2) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, so that

- $\widehat{\mathcal{C}}(P_1,P_2) \nsubseteq \widehat{\mathcal{B}}(P_1,P_2)$, *and* $\widehat{\mathcal{B}}(P_1,P_2) \nsubseteq \widehat{\mathcal{C}}(P_1,P_2)$,
- $\widehat{\mathcal{C}}^{\{i\}}(Q_1,Q_2) \nsubseteq \widehat{\mathcal{K}}(Q_1,Q_2)$, *and* $\widehat{\mathcal{K}}(Q_1,Q_2) \nsubseteq \widehat{\mathcal{C}}^{\{i\}}(Q_1,Q_2)$, *for some* $i \in N$,
- $\widehat{\mathcal{C}}^S(R_1,R_2) \nsubseteq \widehat{\mathcal{K}}(R_1,R_2)$, *and* $\widehat{\mathcal{K}}(R_1,R_2) \nsubseteq \widehat{\mathcal{C}}^S(R_1,R_2)$, *for some* $S \subseteq N$.

3.3.4 Ostrowski sets for Generalized Eigenvalues

Since we can, without any obstacle, apply the above principles to the class of α_1 and α_2-matrices, we will just briefly introduce the generalized α_2-localization set, and its approximation, with some illustrative examples. Although we have given the characterization of the minimal α_2-set in Theorem 2.2.24, to extend it to the generalized case is not an easy task. The reason lies in the fact that sets of indices used to define the localization sets, vary, dependently of the values z..

Theorem 3.3.37. *Given any regular pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, with* $n \geq 2$*, let* $\lambda \in \sigma(A, B)$*. Then, for an arbitrary* $\alpha \in [0, 1]$*, there exists an index* $i \in N$ *such that* $|\lambda b_{i,i} - a_{i,i}| \leq (r_i(\lambda B - A))^\alpha (c_i(\lambda B - A))^{1-\alpha}$*. In other words, for an arbitrary* $\alpha \in [0, 1]$*,*

$$\sigma(A, B) \subseteq \mathcal{A}^2_\alpha(A, B) := \bigcup_{i \in N} \mathcal{A}^2_{\alpha,i}(A, B), \tag{3.3.34}$$

where $\mathcal{A}^2_{\alpha,i}(A, B) := \{z \in \mathbb{C} \, : \, |zb_{i,i} - a_{i,i}| \leq (r_i(zB - A))^\alpha (c_i(zB - A))^{1-\alpha}\}$,
and, consequently,

$$\sigma(A) \subseteq \mathcal{A}^2(A, B) := \bigcap_{\alpha \in [0\,,\,1]} \mathcal{A}^2_\alpha(A, B). \tag{3.3.35}$$

Theorem 3.3.38. *Given any regular pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*,* $n \geq 2$*, if there exist sets* $U, V \subseteq \mathbb{C}$*, such that* $U \cap V = \emptyset$*, and* $\mathcal{A}^2(A, B) = U \cup V$*, then the set* U *contains exactly* $|\left\{i \in N \, : \, b_{i,i} \neq 0 \text{ and } a_{i,i}b^{-1}_{i,i} \in U\right\}|$ *finite eigenvalues of the pair* (A, B) *and, if* U *is unbounded, exactly* $|\{i \in N \, : \, b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N \, : \, a_{i,i}b^{-1}_{i,i} \in U\right\}|$ *eigenvalues of the pair* (A, B)*.*

Theorem 3.3.39. *For any regular matrix pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, the following two conditions hold:*

- $0 \notin \mathcal{A}^2(A, B)$ *if and only if* A *is an* α_2*-matrix, and*
- $\infty \notin \mathcal{A}^2(A, B)$ *if and only if* B *is an* α_2*-matrix.*

Theorem 3.3.40. *Given any regular pair* $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$*, with* $n \geq 2$*, let* $\lambda \in \sigma(A, B)$*. Then, for an arbitrary* $\alpha \in [0, 1]$*, there exists an index* $i \in N$ *such that* $|\lambda b_{i,i} - a_{i,i}| \leq (|\lambda| r_i(B) + r_i(A))^\alpha (|\lambda| c_i(B) + c_i(A))^{1-\alpha}$*. In other words, for an arbitrary* $\alpha \in [0, 1]$*,*

$$\sigma(A, B) \subseteq \widehat{\mathcal{A}}^2_\alpha(A, B) := \bigcup_{i \in N} \widehat{\mathcal{A}}^2_{\alpha,i}(A, B), \tag{3.3.36}$$

where

$$\widehat{\mathcal{A}}^2_{\alpha,i}(A, B) := \left\{z \in \mathbb{C} \, : \, |zb_{i,i} + a_{i,i}| \leq (|z| r_i(B) + r_i(A))^\alpha (|z| c_i(B) + c_i(A))^{1-\alpha}\right\},$$

and, consequently,

$$\sigma(A) \subseteq \widehat{\mathcal{A}}^2(A, B) := \bigcap_{\alpha \in [0\,,\,1]} \widehat{\mathcal{A}}^2_\alpha(A, B). \tag{3.3.37}$$

Theorem 3.3.41. *Given any regular pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, *if there exist sets* $U, V \subseteq \mathbb{C}$, *such that* $U \cap V = \emptyset$, *and* $\widehat{\mathcal{A}}^2(A) = U \cup V$, *then the set* U *contains exactly* $|\left\{i \in N : b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *finite eigenvalues of the pair* (A,B) *and, if* U *is unbounded, exactly* $|\{i \in N : b_{i,i} = 0\}|$ *infinite eigenvalues. Consequently, the set* U *contains exactly* $|\left\{i \in N : a_{i,i}b_{i,i}^{-1} \in U\right\}|$ *eigenvalues of the pair* (A,B).

Theorem 3.3.42. *For any regular matrix pair* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, *the following two conditions hold:*

- $0 \notin \widehat{\mathcal{A}}^2(A,B)$ *if and only if* A *is an an* α_2*-matrix, and*
- $\infty \notin \widehat{\mathcal{A}}^2(A,B)$ *if and only if* B *is an* α_2*-matrix.*

The following example illustrates the generalized α_2-set and its approximation.

Example 3.3.43.

$$A_5 = \begin{pmatrix} 1 & 1 & 0 & 0.2 \\ 0 & -1 & 0.4 & 0 \\ 0 & 0 & i & 1 \\ 0.2 & 0 & 0 & -i \end{pmatrix} \quad \text{and } B_5 = \begin{pmatrix} 0.5 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0.1 & 0 & 0 & -0.5i \end{pmatrix}.$$

Figure 3.3.19 shows the generalized α_2*-set for fixed* $\alpha = 0.5$, *shaded, and its approximation with the thick black line. Generalized Eigenvalues are marked, as always, with "×".*

To conclude the section, following the same reasoning as before, we establish the relationship of this generalized localization set with the previous ones.

Theorem 3.3.44. *Let* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, *be any regular matrix pair, and let the set* $\Gamma(A,B)$ *be given by (3.2.1), the set* $\mathcal{B}(A,B)$ *by (3.3.16), the set* $\mathcal{C}^S(A,B)$ *by (3.3.22), the set* $\mathcal{C}(A,B)$ *by (3.3.23), and the set* $\mathcal{A}^2(A,B)$ *by (3.3.35). Then,*

$$\mathcal{A}^2(A,B) \subseteq \Gamma(A,B) \cap \Gamma(A^T,B^T). \tag{3.3.38}$$

Moreover, there exist matrix pairs $(P_1,P_2),(Q_1,Q_2) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, *so that*

- $\mathcal{A}^2(P_1,P_2) \nsubseteq \mathcal{B}(P_1,P_2)$, *and* $\mathcal{B}(P_1,P_2) \nsubseteq \mathcal{A}^2(P_1,P_2)$, *and*
- $\mathcal{A}^2(Q_1,Q_2) \nsubseteq \mathcal{C}(Q_1,Q_2)$, *and* $\mathcal{C}(Q_1,Q_2) \nsubseteq \mathcal{A}^2(Q_1,Q_2)$.

Theorem 3.3.45. *Let* $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, *be any regular matrix pair, and let the set set* $\widehat{\Gamma}(A,B)$ *be given by (3.2.11), the set* $\widehat{\mathcal{B}}(A,B)$ *by (3.3.19), the set* $\widehat{\mathcal{C}}^S(A,B)$ *by (3.3.26), the set* $\widehat{\mathcal{C}}(A,B)$ *by (3.3.27) and the set* $\mathcal{A}^2(A,B)$ *by (3.3.37). Then,*

$$\widehat{\mathcal{A}}^2(A,B) \subseteq \widehat{\Gamma}(A,B) \cap \widehat{\Gamma}(A^T,B^T). \tag{3.3.39}$$

Moreover, there exist matrix pairs $(P_1,P_2),(Q_1,Q_2) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, *so that*

- $\widehat{\mathcal{A}}^2(P_1,P_2) \nsubseteq \widehat{\mathcal{B}}(P_1,P_2)$, *and* $\widehat{\mathcal{B}}(P_1,P_2) \nsubseteq \widehat{\mathcal{A}}^2(P_1,P_2)$, *and*
- $\widehat{\mathcal{A}}^2(Q_1,Q_2) \nsubseteq \widehat{\mathcal{C}}(Q_1,Q_2)$, *and* $\widehat{\mathcal{C}}(Q_1,Q_2) \nsubseteq \widehat{\mathcal{A}}^2(Q_1,Q_2)$.

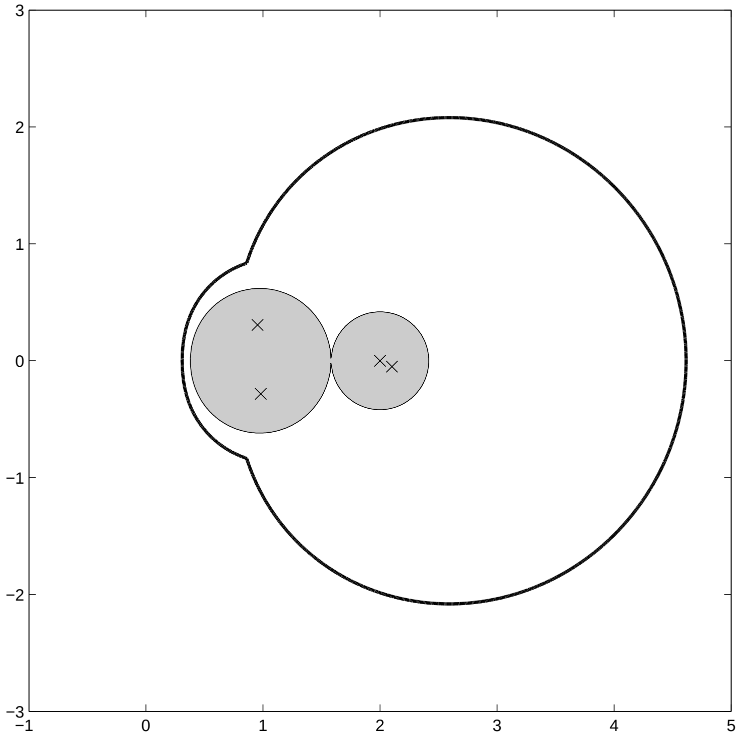

Figure 3.3.19: Generalized α_2-set (shaded) and its approximation (thick boundary) of the matrix pair (A_5, B_5) of the Example 3.3.43
(Generalizovani α_2-skup (osenčen) i njegova aproksimacija (debljom linijom) za matrični par (A_5, B_5) iz Primera 3.3.43)

3.4 Improved Approximations of the Generalized Geršgorin-type Sets

The goal of this section is to introduce a technique through which we can obtain better approximations of the generalized Geršgorin-type sets. In the previous section for the DD-type class of matrices $\mathbb{K}$, we have obtained a set $\widehat{\Theta}^{\mathbb{K}}(A, B)$ in complex plane, which contains the eigenvalues of the matrix pair (A, B), and we called this set the *approximated* generalized Geršgorin-type set. Although these approximations have had some mutual properties with the original generalized Geršgorin-type sets $\Theta^{\mathbb{K}}(A, B)$, in many examples throughout the previous section, we could see that sometimes they are *not* sufficiently close to the their originals. Therefore, it is an interesting problem to develop an improved technique which will allow us to approach closer to the original generalized Geršgorin-type sets, of course with a sufficiently less computation.

Given an arbitrary matrices $A, B \in \mathbb{C}^{n,n}$, take an arbitrary complex number $\xi \in \mathbb{C}$ and define the matrix $T_z^{\xi}(A, B) =: M = [m_{i,j}]$, so that

$$m_{i,j} := \begin{cases} |zb_{i,i} - a_{i,i}|, & i = j, \\ -|z - \xi||b_{i,j}| - |\xi b_{i,j} - a_{i,j}|, & \text{otherwise}, \end{cases} \tag{3.4.1}$$

for all $i, j \in N$.

Now, for an arbitrary $\xi \in \mathbb{C}$, and for every distinct indices $i, j \in N$,

$$|zb_{i,j} - a_{i,j}| = |zb_{i,j} - \xi b_{i,j} + \xi b_{i,j} - a_{i,j}| \leq |z - \xi||b_{i,j}| + |\xi b_{i,j} - a_{i,j}|,$$

and, hence, $\langle zB - A\rangle \geq T_z^{\xi}(A, B)$.

Therefore, given an arbitrary complex number $\xi \in \mathbb{C}$, and any diagonally dominant-type class of matrices, denoted by $\mathbb{K}$, for every regular matrix pair (A, B), we can define the set of complex numbers $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B) := \{z \in \mathbb{C} : T_z^{\xi}(A, B) \notin \mathbb{K}\}$. Then, as in Theorem 3.3.2, $\Theta^{\mathbb{K}}(A, B) \subseteq \widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B)$. So, we have obtained the approximation of the generalized Geršgorin-type set that depends on the free complex parameter $\xi \in \mathbb{C}$. Moreover, the following theorem holds.

Theorem 3.4.1. *Given a diagonally dominant-type class of nonsingular matrices, denoted by $\mathbb{K}$, for any regular matrix pair (A, B),*

$$\sigma(A, B) \subseteq \Theta^{\mathbb{K}}(A, B) \subseteq \widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B), \quad (\xi \in \mathbb{C}).$$

Moreover,

$$\Theta^{\mathbb{K}}(A, B) = \bigcap_{\xi \in \mathbb{C}} \widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B). \tag{3.4.2}$$

Proof. The first part is obvious, so we need to prove the equality in (3.4.2), more precisely, we need to prove that for each $z \in \Theta^{\mathbb{K}}(A, B)$, there exists $\xi \in \mathbb{C}$, such that $z \in \widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B)$. So, assuming that $z \in \Theta^{\mathbb{K}}(A, B)$, $zB - A \notin \mathbb{K}$. But, taking $\xi := z$, $T_z^{z}(A, B) = \langle zB - A\rangle$. So, from the fact that $\mathbb{K}$ is a DD-type class, $T_z^{z}(A, B) \notin \mathbb{K}$, and, hence, $z \in \widehat{\Theta}_{z}^{\mathbb{K}}(A, B) \subseteq \bigcap_{\xi \in \mathbb{C}} \widehat{\Theta}_{\xi}^{\mathbb{K}}(A, B)$. □

Of course, this principle we can now apply to every DD-type class of nonsingular matrices, the ones we have covered in this theses, and many other which could be found in the literature. Before we illustrate this approach on the generalized Geršgorin set, we will give the useful concepts of the Isolation Principle and Boundedness Principle for these approximated sets. Their proofs are obvious. First one follows form the fact that the isolation is the property that approximations inherit, while the second one can be proved the same way as Theorem 3.3.10, only by taking M_k to be $M_k := |z_k|^{-1}T^{\xi}_{z_k}(A, B)$.

Theorem 3.4.2. (Isolation Principle) *Given a positively homogenous DD-type class of nonsingular matrices $\mathbb{K}$, and an arbitrary complex number $\xi \in \mathbb{C}$, for any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and*

$$\widehat{\Theta}^{\mathbb{K}}_{\xi}(A, B) = U \cup V, \tag{3.4.3}$$

then, the set U contains exactly $|\left\{i \in N : b_{i,i} \neq 0 \text{ and } a_{i,i}b^{-1}_{i,i} \in U\right\}|$ finite eigenvalues of the pair (A, B), and, if U is unbounded, exactly $|\{i \in N : b_{i,i} = 0\}|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N : a_{i,i}b^{-1}_{i,i} \in U\right\}|$ eigenvalues of the pair (A, B).

Theorem 3.4.3. (Boundedness Principle) *Given a positively homogenous SDD-type class of matrices $\mathbb{K}$, and an arbitrary complex number $\xi \in \mathbb{C}$, for any regular matrix pair $(A, B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following two conditions hold:*

- *$0 \notin \widehat{\Theta}^{\mathbb{K}}_{\xi}(A, B)$ if and only if $A \in \mathbb{K}$, and*
- *$\infty \notin \widehat{\Theta}^{\mathbb{K}}_{\xi}(A, B)$ if and only if $B \in \mathbb{K}$.*

Clearly, taking $\xi := 0$, $\widehat{\Theta}^{\mathbb{K}}_{\xi}(A, B) = \widehat{\Theta}^{\mathbb{K}}(A, B)$, so, it seems to be an interesting topic to explore how the changes in ξ influence the form of the approximation set. Since ξ can be any complex number, we can experiment and vary only the absolute value, and, on the other hand, we can vary only the argument, i.e., the complex sign. In the following we present some examples, while the general question, how to chose the parameter ξ, remains an open problem.

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

$$\begin{cases} \widehat{\Gamma}^{\xi}_i(A, B) := \{z \in \mathbb{C} : |zb_{i,i} - a_{i,i}| \leq |z - \xi| r_i(B) + r_i(\xi B - A)\}, \quad (i \in N), \\ \widehat{\Gamma}^{\xi}(A, B) := \bigcup_{i \in N} \widehat{\Gamma}^{\xi}_i(A, B). \end{cases} \tag{3.4.4}$$

Now, let, as in the previous section,

$$A_1 = \begin{pmatrix} 1 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0.1 & 0 & 0 & -i \end{pmatrix}, \text{ and } B_1 = \begin{pmatrix} 0.5 & 0.1 & 0.1 & 0.1 \\ 0 & -1 & 0.1 & 0.1 \\ 0 & 0 & i & 0.1 \\ 0.1 & 0 & 0 & -0.5i \end{pmatrix}. \tag{3.4.5}$$

In Figure 3.4.1, set $\widehat{\Gamma}^{\xi}_i(A_1, B_1)$ is shown, for the values $\xi = -2, -1.75, -0.5, \ldots, 1$, from the upper left corner to the bottom right corner, respectively.

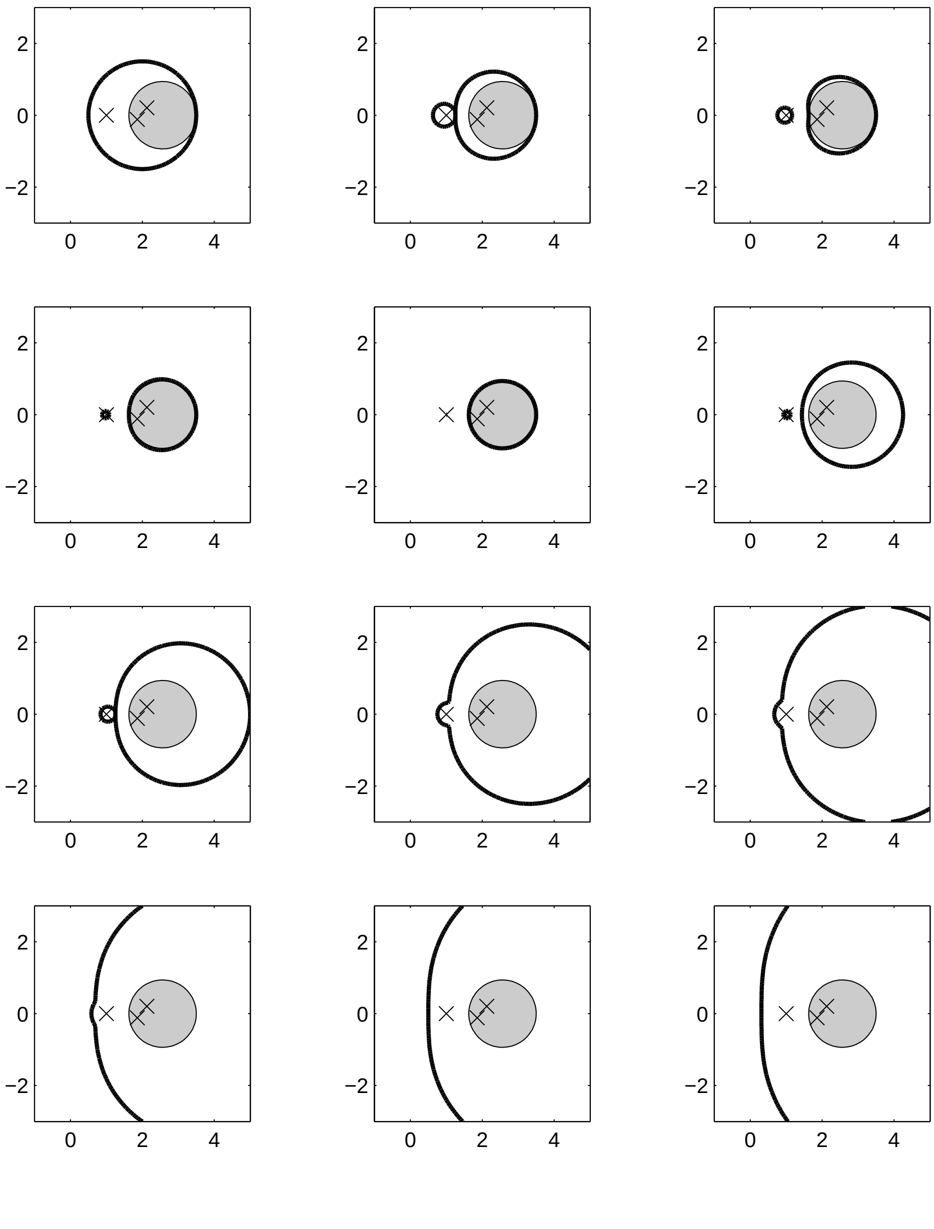

Figure 3.4.1: The set $\widehat{\Gamma}_i^{\xi}(A_1, B_1)$, for the values $\xi = -2, -0.75, -0.5, \ldots, 1$, from the upper left corner to the bottom right corner, respectively, for the matrix pair (A_1, B_1) from (3.4.5) *(Skup* $\widehat{\Gamma}_i^{\xi}(A_1, B_1)$*, za vrednosti* $\xi = -2, -0.75, -0.5, \ldots, 1$*, redom, od gornjeg levog ka donjem desnom uglu, za matrični par* (A_1, B_1) *dat sa (3.4.5))*

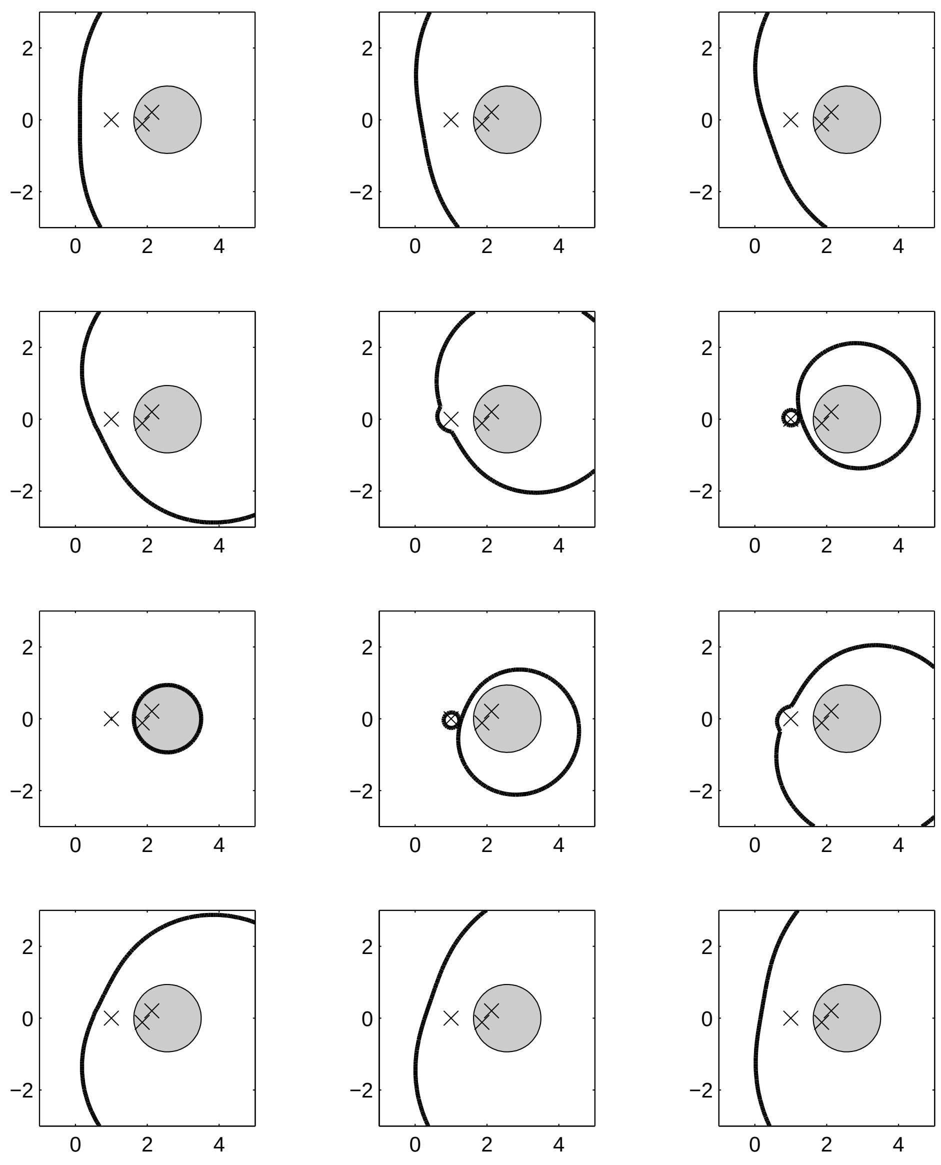

Figure 3.4.2: The set $\widehat{\Gamma}_i^{\xi}(A_1, B_1)$, for the values $\xi = e^{i\frac{2k\pi}{12}}$, where $k = 0, 1, \ldots, 12$, from the upper left corner to the bottom right corner, respectively, for the matrix pair (A_1, B_1) from (3.4.5)

(Skup $\widehat{\Gamma}_i^{\xi}(A_1, B_1)$, za vrednosti $\xi = e^{i\frac{2k\pi}{12}}$, gde je $k = 0, 1, \ldots, 12$, redom, od gornjeg levog ka donjem desnom uglu, za matrični par (A_1, B_1) dat sa (3.4.5))

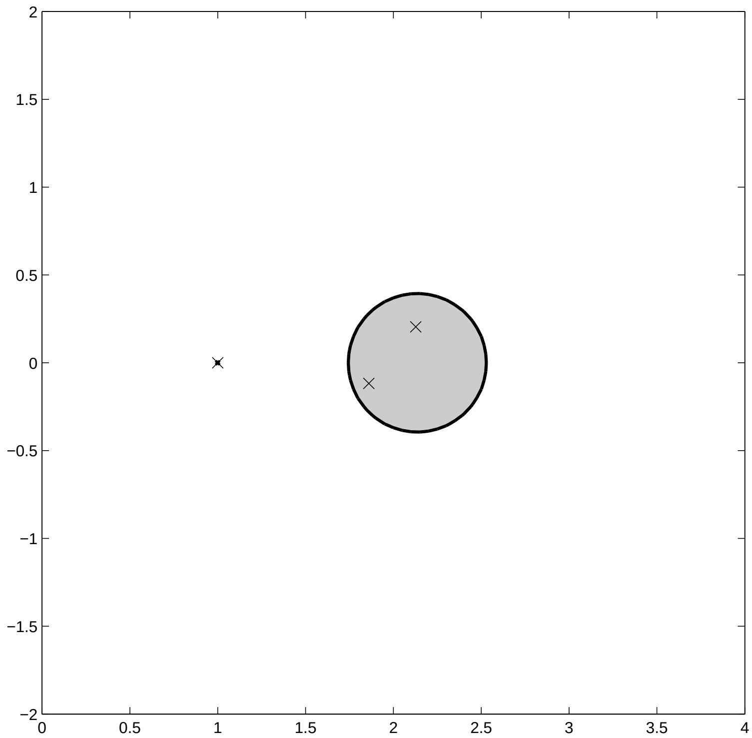

Figure 3.4.3: The set $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_1, B_1)$, for the value $\xi = -1$, and the matrix pair (A_1, B_1) from Example 3.2.4, where $\mathbb{K}$ is the class of doubly SDD matrices
(Skup $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_1, B_1)$, za vrednost $\xi = -1$ i za matrični par (A_1, B_1) iz Primera 3.2.4, gde je $\mathbb{K}$ klasa dvostruko SDD matrica)

In Figure 3.4.2, set $\widehat{\Gamma}_i^{\xi}(A_1, B_1)$ is shown, for the values $\xi = e^{i\frac{2k\pi}{12}}$, for $k = 0, 1, \ldots, 12$, from the upper left corner to the bottom right corner, respectively.

Of course, this technique can be used in all of the mentioned generalized Geršgorin-type localizations. While omitting the details concerning their explicit form, in Figures 3.4.3 - 3.4.5 we show the possible improvements for the various localization sets for the matrices of Examples 3.2.4, 3.3.19, and 3.3.43.

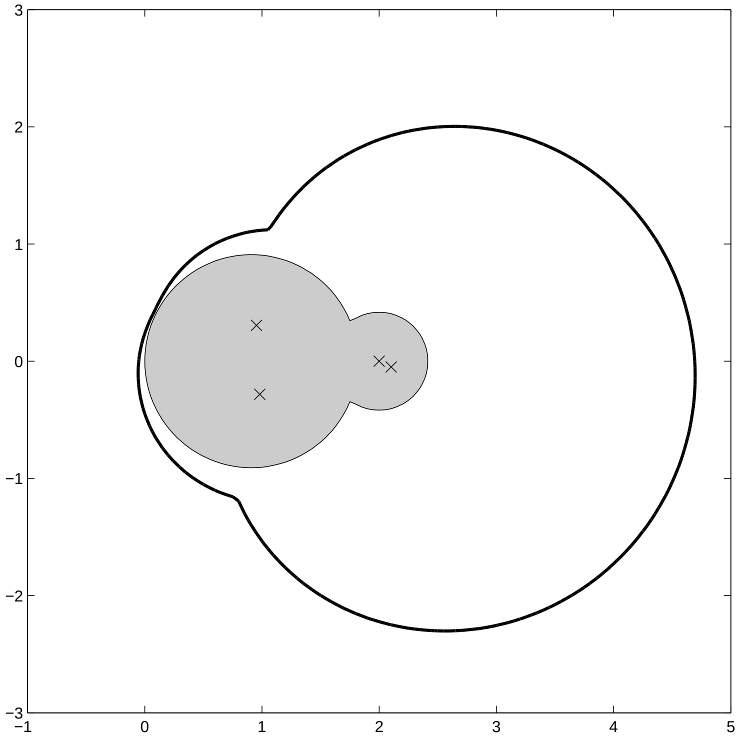

Figure 3.4.4: The set $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_5, B_5)$, for the value $\xi = -0.5i$, and the matrix pair (A_5, B_5) from the Example 3.3.43, where $\mathbb{K}$ is S-SDD class of matrices for $S = \{4\}$
(Skup $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_5, B_5)$, za vrednost $\xi = -0.5$ i za matrični par (A_5, B_5) iz Primera 3.3.43, gde je $\mathbb{K}$ klasa S-SDD matrica za skup $S = \{4\}$)

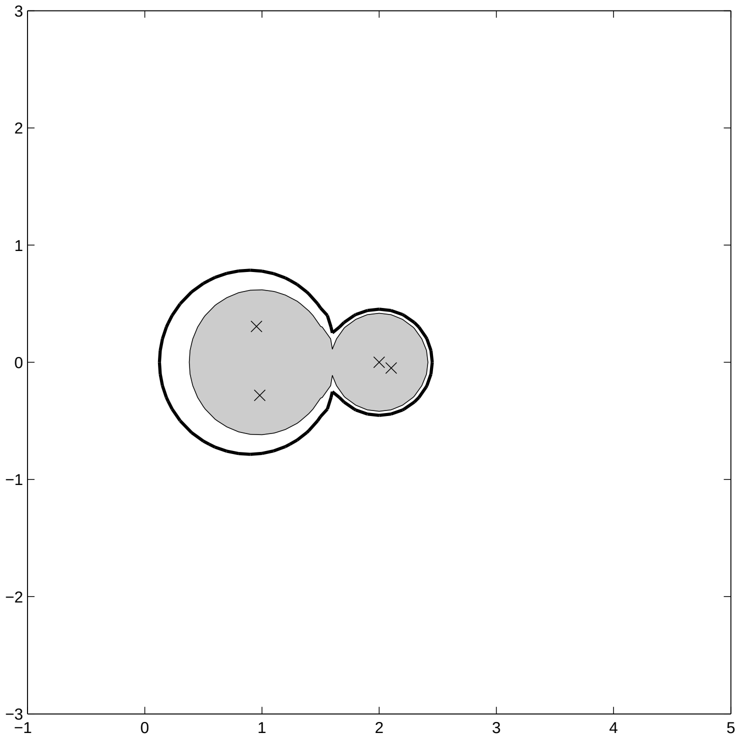

Figure 3.4.5: The set $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_5, B_5)$, for the value $\xi = -2$, and the matrix pair (A_5, B_5) from the Example 3.3.43, where $\mathbb{K}$ is the class of α_2-SDD matrices

(Skup $\widehat{\Theta}_{\xi}^{\mathbb{K}}(A_5, B_5)$, za vrednost $\xi = -2$ i za matrični par (A_5, B_5) iz Primera 3.3.43, gde je $\mathbb{K}$ klasa α_2-SDD matrica)

3.5 Minimal Geršgorin Set for the Generalized Eigenvalues

In the last section of this chapter we consider the generalized minimal Geršgorin set, as it was introduced by Kostić, Cvetković and Varga in [35]. While many new concepts have been developed in this theses for the the application of the generalized Geršgorin set, the material that is given in [35] on the topic of minimal Geršgorin set is, more or less, self contained, and up to date with the current knowledge. Therefore, here we give a review of this result.

We begin, as in the previous section, with

Definition 3.5.1. The set $\Gamma^{\mathbb{D}}(A,B)$, defined as

$$\Gamma^{\mathbb{D}}(A,B) := \{z \in \mathbb{C} \ : \ A - zB \ \text{ is not a nonsingular H-matrix}\},$$

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

This time we have weakened the singularity property of a matrix pencil, in the point z, to be the property that $A - zB$ is not a nonsingular H-matrix, in order to "enlarge" spectrum up to the generalized minimal Geršgorin set. Since the class of nonsingular H-matrices fulfils all the conditions to be positively homogenous SDD-type of matrices, we can apply Theorems 3.3.1, 2.2.3 and 3.3.9.

Theorem 3.5.2. *Given any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, the generalized spectrum of the matrix pair (A,B) belongs to the generalized minimal Geršgorin set of the matrix pair (A,B), i.e., the following inclusion holds:*

$$\sigma(A,B) \subseteq \Gamma^{\mathbb{D}}(A,B). \tag{3.5.1}$$

Theorem 3.5.3. *Given any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, $n \geq 2$, if there exist sets $U, V \subseteq \mathbb{C}$, such that $U \cap V = \emptyset$, and*

$$\Gamma^{\mathbb{D}}(A,B) = U \cup V,$$

then, the set U contains exactly $|\left\{i \in N \ : \ b_{i,i} \neq 0 \text{ and } a_{i,i}b_{i,i}^{-1} \in U\right\}|$ finite eigenvalues of the pair (A,B) and, if U is unbounded, exactly $|\{i \in N \ : \ b_{i,i} = 0\}|$ infinite eigenvalues. Consequently, the set U contains exactly $|\left\{i \in N \ : \ a_{i,i}b_{i,i}^{-1} \in U\right\}|$ eigenvalues of the pair (A,B).

Theorem 3.5.4. *For any regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, the following two conditions hold:*

- $0 \notin \Gamma^{\mathbb{D}}(A,B)$ *if and only if A is a nonsingular H-matrix, and*
- $\infty \notin \Gamma^{\mathbb{D}}(A,B)$ *if and only if B is a nonsingular H-matrix.*

Moreover, given a regular matrix pair $(A,B) \in \mathbb{C}^{n,n} \times \mathbb{C}^{n,n}$, with $n \geq 2$, if there exists $i \in N$ such that $b_{i,i} = 0$, then $a_{i,i} = 0$ if and only if $\Gamma^{\mathbb{D}}(A,B) = \mathbb{C}_\infty$.

In the following, we address the problem of computing and plotting the generalized minimal Gešgorin set for a given matrix pair. First, we remark that, in the case of the original minimal Geršgorin set, the progress has recently been made in computing its tight approximation with an iterative approach in [52], as we have presented it in Section 2.3. Here, we will develop an analogue of this, in the sense of our generalized minimal Geršgorin set. Thus, we will need the necessary tools, derived from the Perron-Frobenius theory of nonnegative matrices.

For a given matrix pencil $zB - A \in \mathbb{C}^{n,n}$ and a given $z \in \mathbb{C}$, we define the matrix $Q_z := -\langle zB - A\rangle$, where the comparison matrix operator $\langle\cdot\rangle$ is defined in (1.3.2). Defining $\delta(z) := \max\{|a_{i,i} - zb_{i,i}| \ : \ i \in N\}$, and putting

$$P_z := Q_z + \delta(z)I, \tag{3.5.2}$$

we obtain the nonnegative matrix P_z which, by the Perron-Frobenius theory of nonnegative matrices [3], possesses a real, nonnegative eigenvalue $\rho(P_z)$, called the *Perron root* of P_z.

Now, by setting $\nu_{(A,B)}(z) := \rho(P_z) - \delta(z)$, we have, from Theorem C.2 in [51], that

$$\nu_{(A,B)}(z) = \inf_{\mathbf{x}>\mathbf{0}}\{\max_{i\in N}[(Q_z\mathbf{x})_i/x_i]\}, \tag{3.5.3}$$

or equivalently,

$$\nu_{(A,B)}(z) = \inf_{\mathbf{x}>\mathbf{0}}\{\max_{i\in N}[x_i^{-1}\cdot \sum_{j\in N\setminus\{i\}} |zb_{i,j} - a_{i,j}|x_j - |zb_{i,i} - a_{i,i}|]\}. \tag{3.5.4}$$

Thus, the following characterization of the generalized minimal Geršgorin set holds.

Theorem 3.5.5. *Given any two matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geq 2$, then*

$$z \in \Gamma^{\mathbb{D}}(A,B) \ \textit{ if and only if } \ \nu_{(A,B)}(z) \geq 0. \tag{3.5.5}$$

The proof of this theorem follows in the same way as in the proof of Proposition 4.3 of [51], which characterizes the minimal Geršgorin set. In addition, following the same idea as in (2.3.16), the real-valued complex function $\nu_{(A,B)}$ is continuous, and the generalized minimal Geršgorin set is a closed set in the extended complex plain $\mathbb{C}_\infty$, so, we obtain that

$$z \in \partial\Gamma^{\mathbb{D}}(A,B) \ \text{ if and only if } \begin{cases} i) \quad \nu_{(A,B)}(z) = 0, \ \text{ and} \\ ii) \ \text{there exists a sequence of complex} \\ \quad\ \text{numbers } \ \{z_j\}_{j=1}^{\infty} \ \text{ such that } \ \lim_{j\to\infty} z_j = z, \\ \quad\ \text{and } \ \nu_{(A,B)}(z_j) < 0 \ \text{ for all } \ j \geq 1. \end{cases} \tag{3.5.6}$$

This brings us, as before, to the notion of a star-shaped set, needed in the next result. The set $U \subseteq \mathbb{C}_\infty$ is said to be a **star-shaped with a respect to** a given point z_0, if for every z in U, the entire line segment between z_0 and z lies in U, i.e., $\{\alpha z_0 + (1-\alpha)z \ : \ 0 \leq \alpha \leq 1\} \subseteq U$.

Theorem 3.5.6. *For any two matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geq 2$, such that B is a nonsingular H-matrix, then*

$$\nu_{(A,B)}(\frac{a_{k,k}}{b_{k,k}}) \geq 0$$

for each $k \in N$. Moreover, for each $k \in N$ and for each θ with $0 \leq \theta \leq 2\pi$, there exists an $\hat{\varrho}_k(\theta) \geq 0$ such that the entire complex interval $[\frac{a_{k,k}}{b_{k,k}} + te^{i\theta}]_{t=0}^{\hat{\varrho}_k(\theta)}$ is contained in $\Gamma^{\mathbb{D}}(A, B)$, and, consequently, the set

$$\bigcup_{\theta=0}^{2\pi}[\frac{a_{k,k}}{b_{k,k}} + te^{i\theta}]_{t=0}^{\hat{\varrho}_k(\theta)} \tag{3.5.7}$$

is star-shaped subset *of* $\Gamma^{\mathbb{D}}(A, B)$ *with respect to the point* $\frac{a_{k,k}}{b_{k,k}}$.

Proof. Since B is a nonsingular H-matrix, then, for every $k \in N$, $b_{k,k} \neq 0$, and on taking $z = \frac{a_{k,k}}{b_{k,k}}$ in (3.5.4), we obtain

$$\nu_{(A,B)}(\frac{a_{k,k}}{b_{k,k}}) \geq \inf_{\mathbf{x}>\mathbf{0}}\{x_k^{-1} \cdot \sum_{j\in N\backslash\{k\}} |\frac{a_{k,k}b_{k,j}}{b_{k,k}} - a_{k,j}|x_j\} \geq 0.$$

Thus, $\frac{a_{k,k}}{b_{k,k}}$ lies in the set $\Gamma^{\mathbb{D}}(A, B)$. Now, for a fixed θ in $0 \leq \theta \leq 2\pi$, consider the ray $[\frac{a_{k,k}}{b_{k,k}} + te^{i\theta}]$, $t \geq 0$. Its starting point lies in $\Gamma^{\mathbb{D}}(A, B)$, which is, according to the Theorem 3.5.4, a compact set in $\mathbb{C}$. Thus, there exists a point $\frac{a_{k,k}}{b_{k,k}} + \hat{\varrho}_k(\theta)e^{i\theta}$ which lies on the boundary of the $\Gamma^{\mathbb{D}}(A, B)$. Taking the smallest $\hat{\varrho}_k(\theta)$ of such points we obtain a star-shaped subset 3.5.7. □

Now, for a fixed θ, with $0 \leq \theta \leq 2\pi$, it is interesting to note that if $\nu_{(A,B)}(\frac{a_{k,k}}{b_{k,k}}) = 0$ and if $\hat{\varrho}_k(\theta) = 0$, then $\frac{a_{k,k}}{b_{k,k}}$ actually lies on the boundary of $\Gamma^{\mathbb{D}}(A, B)$. In addition, if $\hat{\varrho}_k(\theta) = 0$ for each θ with $0 \leq \theta \leq 2\pi$, then $\frac{a_{k,k}}{b_{k,k}}$ is a generalized eigenvalue of the pair (A, B). This brings us to

Theorem 3.5.7. *Given two matrices $A, B \in \mathbb{C}^{n,n}$, with $n \geq 2$, for which there exists a $k \in N$ such that $b_{k,k} = 0$, then for every sequence of complex numbers $\{z_k\}_{k=1}^{\infty}$ such that $|z_k| \to \infty$, as $k \to \infty$, there exists an $\alpha \geq 0$ such that $\nu_{(A,B)}(z_k) \to \alpha$, which we, by convention, write as $\nu_{(A,B)}(\infty) > 0$. Moreover, if A is a nonsingular H-matrix, then, for each θ with $0 \leq \theta \leq 2\pi$, there exists a $\hat{\varrho}_k(\theta) > 0$ such that the whole complex interval $[\hat{\varrho}_k(\theta)e^{i\theta} + t]_{t=0}^{\infty}$ is contained in $\Gamma^{\mathbb{D}}(A, B)$, and, consequently, the set*

$$\bigcup_{\theta=0}^{2\pi}[\hat{\varrho}_k(\theta)e^{i\theta} + t]_{t=0}^{\infty} \tag{3.5.8}$$

is a star-shaped *subset of* $\Gamma^{\mathbb{D}}(A, B)$ *with respect to* ∞.

Proof. The idea of this proof is the following. For any $z \neq 0$ in $\sigma(A, B)$, from (3.1.2) we have that $\det(A - zB) = 0$, from which it follows that $\det(B - \frac{1}{z}A) = 0$. This necessarily implies that $\frac{1}{z} \in \sigma(B, A)$. So, for a $k \in N$ such that $b_{k,k} = 0$ and $a_{k,k} \neq 0$, we have that the star-shaped subset of (3.5.7) for the matrix pair (B, A), corresponding to the center $\frac{b_{k,k}}{a_{k,k}} = 0$, transforms to the set (3.5.8). □

Since the generalized minimal Geršgorin set is obtained as a intersection of all "weighted" generalized Geršgorin sets, it is, in a sense, the minimal set that contains the generalized spectrum. On the other hand, we see from (3.2.4) and (3.2.5) that the generalized Geršgorin set is defined uniquely from the following data: $|a_{i,j}|, |b_{i,j}|$ and $\frac{a_{i,j}}{b_{i,j}}$, when $b_{i,j} \neq 0$, where $i, j \in N$, and thus, for every matrix pair which leaves this data set unchanged, its generalized spectrum will be included in the same generalized Geršgorin set, and, consequently, in the generalized minimal Geršgorin set.

What we are going to show here is that the generalized minimal Geršgorin set is, in a way, the *best possible localization set* for such matrix pairs. We start by introducing the equimodular set of matrix pairs $\Omega(A, B)$ and the extended equimodular set $\widehat{\Omega}(A, B)$, Similar as in [51], Chapter 4.

$$\Omega(A, B) := \{(\tilde{A}, \tilde{B}) \ : \ |\tilde{a}_{i,j}| = |a_{i,j}|, |\tilde{b}_{i,j}| = |b_{i,j}|, \text{ and if } b_{i,j} \neq 0, \frac{\tilde{a}_{i,j}}{\tilde{b}_{i,j}} = \frac{a_{i,j}}{b_{i,j}} \ \ i, j \in N\}, \tag{3.5.9}$$

$$\widehat{\Omega}(A, B) := \{(\tilde{A}, \tilde{B}) \ : \ |\tilde{a}_{i,j}| \leq |a_{i,j}|, |\tilde{b}_{i,j}| \leq |b_{i,j}|, \text{ and if } b_{i,j} \neq 0, \frac{\tilde{a}_{i,j}}{\tilde{b}_{i,j}} = \frac{a_{i,j}}{b_{i,j}} \ \ i, j \in N\}. \tag{3.5.10}$$

Now, as is natural, we take the spectrum of these sets to be the union of all the spectra of their elements:

$$\sigma(\Omega(A, B)) := \bigcup_{(\tilde{A}, \tilde{B}) \in \Omega(A,B)} \sigma(\tilde{A}, \tilde{B}), \ \text{ and } \ \sigma(\widehat{\Omega}(A, B)) := \bigcup_{(\tilde{A}, \tilde{B}) \in \widehat{\Omega}(A,B)} \sigma(\tilde{A}, \tilde{B}). \tag{3.5.11}$$

It is evident from their definitions that

$$\sigma(\Omega(A, B)) \subseteq \sigma(\widehat{\Omega}(A, B)) \subseteq \Gamma^{\mathbb{D}}(A, B). \tag{3.5.12}$$

How tight these inclusions are, is described by the next two theorems.

Theorem 3.5.8. *For any pair of matrices (A, B) from $\mathbb{C}^{n,n}$, and given an arbitrary $z \in \mathbb{C}$, such that $\nu_{(A,B)}(z)$, of (3.5.3), satisfies $\nu_{(A,B)}(z) = 0$, there exists a matrix pair $(\tilde{A}, \tilde{B}) \in \Omega(A, B)$ such that z is a generalized eigenvalue of the matrix pair $(\tilde{A}, \tilde{B})$. Thus, the following inclusions hold:*

$$\partial\Gamma^{\mathbb{D}}(A, B) \subseteq \sigma(\Omega(A, B)) \subseteq \sigma(\widehat{\Omega}(A, B)) \subseteq \Gamma^{\mathbb{D}}(A, B). \tag{3.5.13}$$

Proof. Let $z \in \mathbb{C}$ be such that $\nu_{(A,B)}(z) = 0$. Then, from the facts leading to (3.5.4), we have that there exists a nonzero $\mathbf{y} = [y_1, \ y_2, \ \ldots, y_n]^T$, with $\mathbf{y} \geq \mathbf{0}$, such that $Q_z\mathbf{y} = \mathbf{0}$, or, equivalently,

$$\sum_{j \in N \setminus \{k\}} |b_{k,j}z - a_{k,j}|y_j = |b_{k,k}z - a_{k,k}|y_k, \ \text{ for all } \ k \in N,$$

which, according to (3.2.4) and (3.2.5), for every $k \in N$ can be written as

$$\sum_{j\in\beta(k)\backslash\{k\}} |z-\frac{a_{k,j}}{b_{k,j}}||b_{k,j}|y_j + \sum_{j\in\overline{\beta}(k)\backslash\{k\}} |a_{k,j}|y_j = |z-\frac{a_{k,k}}{b_{k,k}}||b_{k,k}|y_k, \text{ when } k\in\beta(k) \quad (3.5.14)$$

or

$$\sum_{j\in\beta(k)} |z-\frac{a_{k,j}}{b_{k,j}}||b_{k,j}|y_j + \sum_{j\in\overline{\beta}(k)\backslash\{k\}} |a_{k,j}|y_j = |a_{k,k}|y_k, \text{ otherwise.} \quad (3.5.15)$$

Now, let the real numbers $\{\phi_{k,j}\}_{k,j=1}^{n}$ satisfy

$$|z-\frac{a_{k,j}}{b_{k,j}}| = (z-\frac{a_{k,j}}{b_{k,j}})e^{i\phi_{k,j}}, \quad (3.5.16)$$

for each $k \in N$ and each $j \in \beta(k)$. Having these numbers, we define the matrices $\tilde{A} = [\tilde{a}_{k,j}]$ and $\tilde{B} = [\tilde{b}_{k,j}]$, both in $\mathbb{C}^{n,n}$, by means of

$$\tilde{a}_{k,j} := \begin{cases} a_{k,j}b_{k,j}^{-1}|b_{k,j}|e^{i\phi_{k,j}}, & j\in\beta(k), \\ |a_{k,j}|, & j\in\overline{\beta}(k), \end{cases} \quad (3.5.17)$$

and

$$\tilde{b}_{k,j} := \begin{cases} |b_{k,j}|e^{i\phi_{k,j}}, & j\in\beta(k), \\ 0, & j\in\overline{\beta}(k), \end{cases} \quad (3.5.18)$$

where $j,k \in N$. After a closer look, we can see that $(\tilde{A},\tilde{B}) \in \Omega(A,B)$, so that from (3.5.14) and (3.5.15) respectively, it follows that

$$\sum_{j\in\beta(k)\backslash\{k\}} (z-\frac{\tilde{a}_{k,j}}{\tilde{b}_{k,j}})\tilde{b}_{k,j}y_j + \sum_{j\in\overline{\beta}(k)\backslash\{k\}} \tilde{a}_{k,j}y_j = (z-\frac{\tilde{a}_{k,k}}{\tilde{b}_{k,k}})\tilde{b}_{k,k}y_k, \text{ when } k\in\beta(k),$$

and

$$\sum_{j\in\beta(k)} (z-\frac{\tilde{a}_{k,j}}{\tilde{b}_{k,j}})\tilde{b}_{k,j}y_j + \sum_{j\in\overline{\beta}(k)\backslash\{k\}} \tilde{a}_{k,j}y_j = \tilde{a}_{k,k}y_k, \text{ otherwise.}$$

This leads us to the conclusion that $(A - zB)\mathbf{y} = \mathbf{0}$, i.e., $A\mathbf{y} = zB\mathbf{y}$. Thus, z is a (generalized) eigenvalue of a matrix pair $(\tilde{A},\tilde{B})$, and consequently, from (3.5.11), $z \in \sigma(\Omega(A,B))$. □

As the first inequalities in (3.5.12) and (3.5.13) turn out to be *equalities* for the usual minimal Geršgorin sets (see Theorem 4.5 of [51]), the same is true here.

Theorem 3.5.9. *For any pair of matrices* (A,B) *from* $\mathbb{C}^{n,n}$,

$$\sigma(\widehat{\Omega}(A,B)) = \Gamma^{\mathbb{D}}(A,B). \quad (3.5.19)$$

Proof. Let z be any point of $\Gamma^{\mathbb{D}}(A,B)$. Then, $\nu_{(A,B)}(z) \geq 0$, and, from (3.5.3), there exists a nonzero vector $\mathbf{y} \in \mathbb{R}^n$, with $\mathbf{y} \geq \mathbf{0}$, such that $Q_z\mathbf{y} = \nu_{(A,B)}(z)\mathbf{y}$. Writing the last expression by components, we have

$$\sum_{j\in N\backslash\{k\}} |b_{k,j}z - a_{k,j}|y_j = (|b_{k,k}z - a_{k,k}| + \nu_{(A,B)}(z))y_k, \text{ for all } k\in N. \quad (3.5.20)$$

Now, we define real numbers $\{\delta_k\}_{k=1}^n$ as

$$\delta_k := \begin{cases} \frac{\sum_{j\in N\setminus\{k\}} |b_{k,j}z-a_{k,j}|y_j-\nu_{(A,B)}(z)y_k}{\sum_{j\in N\setminus\{k\}} |b_{k,j}z-a_{k,j}|y_j}, & \text{if } \sum_{j\in N\setminus\{k\}} |b_{k,j}z-a_{k,j}|y_j > 0, \\ 1, & \text{if } \sum_{j\in N\setminus\{k\}} |b_{k,j}z-a_{k,j}|y_j = 0. \end{cases} \tag{3.5.21}$$

Obviously, from (3.5.20), (3.5.21), and the fact that $\nu_{(A,B)}(z)y_k \geq 0$ for each $k \in N$, it follows that $0 \leq \delta_k \leq 1$, and we can construct matrices $\tilde{A} = [\tilde{a}_{jk}]$ and $\tilde{B} = [\tilde{b}_{jk}]$ such that $(\tilde{A}, \tilde{B}) \in \widehat{\Omega}(A, B)$, in the following way: for all $k \in N$, $\tilde{a}_{k,k} = a_{k,k}$ and $\tilde{b}_{k,k} = b_{k,k}$, while for every $j \in N \setminus \{k\}$, $\tilde{a}_{k,j} = \delta_k a_{k,j}$ and $\tilde{b}_{k,j} = \delta_k b_{k,j}$.

Now, it is readily verified that

$$|\tilde{b}_{k,k}z - \tilde{a}_{k,k}|y_k = |b_{k,k}z - a_{k,k}|y_k = \sum_{j\in N\setminus\{k\}} |b_{k,j}z - a_{k,j}|y_j - \nu_{(A,B)}(z)y_k =$$

$$\delta_k \sum_{j\in N\setminus\{k\}} |b_{k,j}z - a_{k,j}|y_j = \sum_{j\in N\setminus\{k\}} |\tilde{b}_{k,j}z - \tilde{a}_{k,j}|y_j, \text{ for all } k \in N,$$

which is the same as the starting point of the proof of Theorem 3.5.8. As before, we can proceed and obtain the pair of matrices $(\hat{A}, \hat{B}) \in \Omega(\tilde{A}, \tilde{B})$, such that $z \in \sigma(\hat{A}, \hat{B})$. Similar, it follows that $z \in \widehat{\Omega}(A, B)$, which completes the proof. □

Finally, we approach the problem of graphing the generalized minimal Geršgorin set. Having the properties given above, this problem becomes the problem of graphing the subset of $\mathbb{C}_\infty$, for which the function $\nu_{(A,B)}(z)$ is nonnegative. In order to resolve this problem, we need to find a way to compute the value of the function $\nu_{(A,B)}(z)$, for different values of z. Using the concept of *irreducibility*, we have the following result.

Theorem 3.5.10. *Given matrices $A, B \in \mathbb{C}^{n,n}$ with $n \geq 2$, let the matrix pencil $A - zB$, at the point $z \in \mathbb{C}$, be irreducible. Then, for each $\mathbf{x} > \mathbf{0}$ in $\mathbb{R}^n$, either*

$$\min_{i\in N}\{(Q_z\mathbf{x})_i/x_i\} < \nu_{(A,B)}(z) < \max_{i\in N}\{(Q_z\mathbf{x})_i/x_i\}, \tag{3.5.22}$$

or

$$Q_z\mathbf{x} = \nu_{(A,B)}(z)\mathbf{x}. \tag{3.5.23}$$

As (3.5.22) and (3.5.23) suggest, we can use the power method as a tool to compute the eigenvalue $\nu_{(A,B)}(z)$. We start with the nonnegative matrix P_z, given in (3.5.2), which we assume to be *irreducible*. Then, either P_z is primitive or it can be shifted to a primitive matrix $P_z + \epsilon I$, $\epsilon > 0$, (see Section 2.2 of [48]). Thus, either way, we can apply power iterations to compute $\rho(P_z)$.

Starting with an $\mathbf{x}^{(0)} > \mathbf{0}$ in $\mathbb{R}^n$, the power iteration gives convergent upper and lower estimates for $\rho(P_z)$, i.e., if $\mathbf{x}^{(m)} := P_z^m\mathbf{x}^{(0)}$ for all $m \geq 1$, then with $\mathbf{x}^{(m)} := [x_1^{(m)}, x_2^{(m)}, ..., x_n^{(m)}]^T$, we have that

$$\underline{\lambda_m} := \min_{i\in N}\{\frac{x_i^{(m+1)}}{x_i^{(m)}}\} \leq \rho(P_z) \leq \max_{i\in N}\{\frac{x_i^{(m+1)}}{x_i^{(m)}}\} =: \overline{\lambda_m} \tag{3.5.24}$$

for all $m \geq 1$, and

$$\lim_{m\to\infty} \underline{\lambda_m} = \rho(P_z) = \lim_{m\to\infty} \overline{\lambda_m}. \tag{3.5.25}$$

Thus,

$$\underline{\lambda_m} - \delta(z) \le \nu_{(A,B)}(z) \le \overline{\lambda_m} - \delta(z).$$

It is important to say that, from (3.5.24), we do not need to find the value of $\nu_{(A,B)}(z)$ with great accuracy, and it is sufficient to iterate, until either one of the next two conditions is fulfilled:

1. $\underline{\lambda_m} > \delta(z)$, implying that $\nu_{(A,B)}(z) > 0$ and, thus, $z \in \Gamma^{\mathbb{D}}(A, B)$, or
2. $\overline{\lambda_m} < \delta(z)$, implying that $\nu_{(A,B)}(z) < 0$ and, thus, $z \in \mathbb{C}_\infty \setminus \Gamma^{\mathbb{D}}(A, B)$.

If neither one is fulfilled until we achieve a certain accuracy $\epsilon > 0$, i.e., $\overline{\lambda_m} - \underline{\lambda_m} < \epsilon$, we conclude that z lies in the ϵ-neighborhood of a boundary point of $\Gamma^{\mathbb{D}}(A, B)$.

So, the simplest way to plot an approximation of the generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A, B)$ is to introduce the coarse grid, say $n_x \times n_y$, of the $[-L, L]^2 \subset \mathbb{C}$, for sufficiently large $L > 0$. For this grid, we will have $n_x n_y$ complex nodes, and we determine for each node either to "color" it or not. Namely, each of them will be either "colored" to be in the $\Gamma^{\mathbb{D}}(A, B)$, if either $\underline{\lambda_m} > \delta(z)$ (case 1.) or $\overline{\lambda_m} - \underline{\lambda_m} < \epsilon$, where ϵ represents the coarseness of the grid. If $\underline{\lambda_m} < \delta(z)$ (case 2.) occurs, the point is left "uncolored", as it is in the exterior of the $\Gamma^{\mathbb{D}}(A, B)$.

Example 3.5.11. *Let*

$$A_6 = \begin{pmatrix} 1 & 1 & 0 & 0.5 \\ 0 & -1 & 0.5 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & -i \end{pmatrix}, \quad B_6 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0.5 & 0 & 0 & 1 \end{pmatrix}$$

$$A_7 = \begin{pmatrix} 0.5 & 0 & 0 & 0.3 \\ 0 & 0.5 & 0.1 & 0 \\ 0 & 0 & 0.7 & 0.1 \\ 1.2 & 0 & 0 & 0.7 \end{pmatrix}, \quad \textit{and} \;\; B_7 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & i & 1 \\ 2 & 0 & 0 & -2i \end{pmatrix}.$$

Generalized minimal Geršgorin sets of the matrix pairs (A_1, B_1), (A_2, B_2) and (A_1, A_2), of the Example 3.2.4, are shown in Figures 3.5.1 - 3.5.3, and the generalized minimal Geršgorin sets of the matrix pairs (A_3, B_3), of the Example 3.3.14, (A_4, B_4), of the Example 3.3.19, (A_5, B_5), of the Example 3.3.43, (A_6, B_6) and (A_7, B_7), of the Example 3.5.11, are shown in Figures 3.5.4-3.5.8, respectively. The eigenvalues are, as always marked by "$\times$".

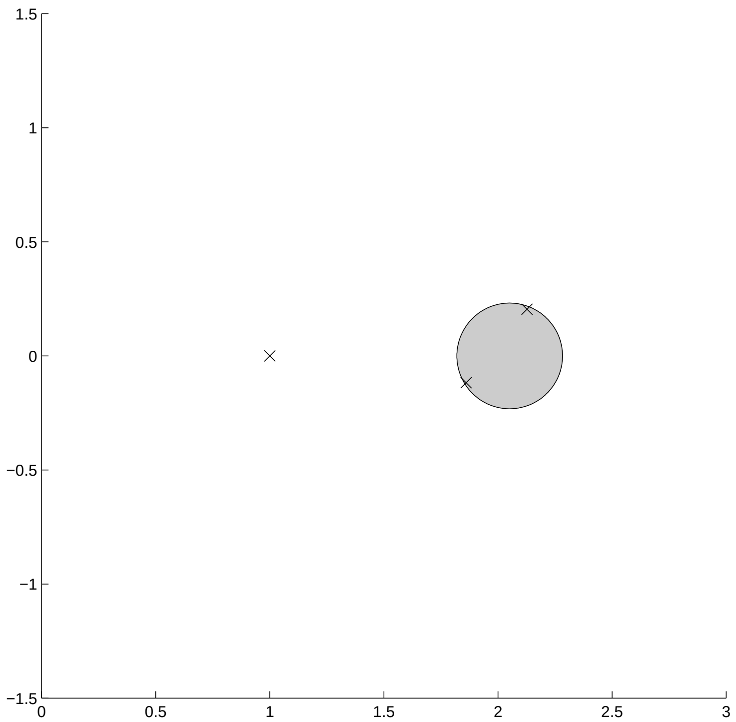

Figure 3.5.1: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_1, B_1)$ of the matrix pair (A_1, B_1) of Example 3.2.4

(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_1, B_1)$ za matrični par (A_1, B_1) iz Primera 3.2.4)

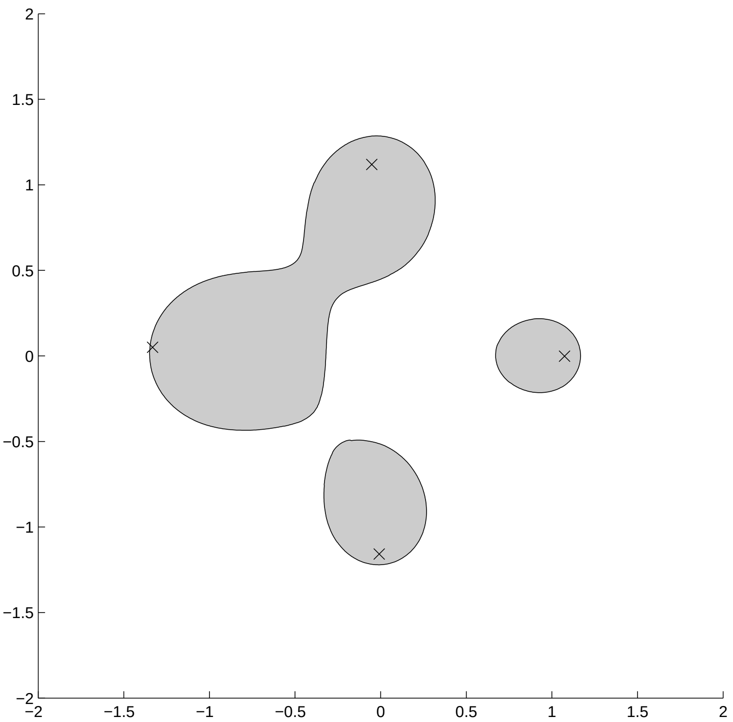

Figure 3.5.2: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_2, B_2)$ of the matrix pair (A_2, B_2) of Example 3.2.4

(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_2, B_2)$ za matrični par (A_2, B_2) iz Primera 3.2.4)

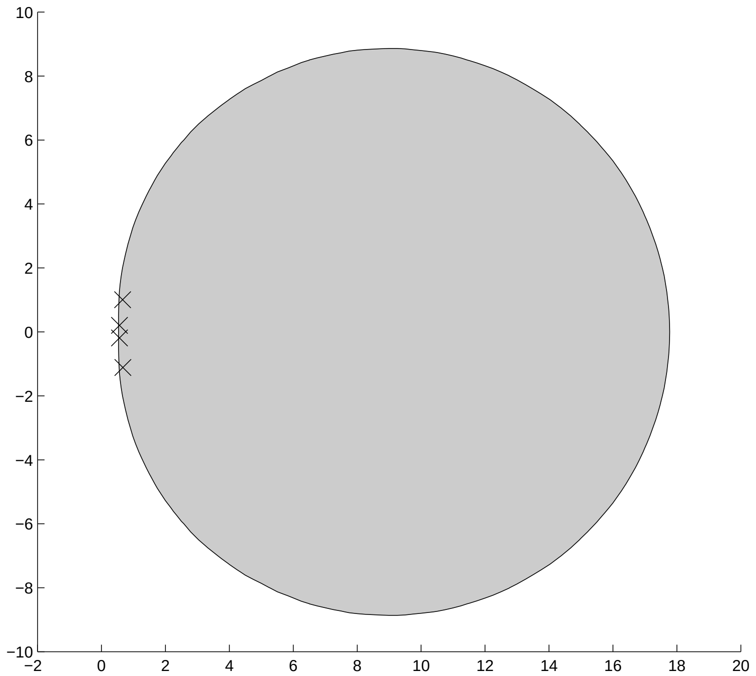

Figure 3.5.3: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_1, A_2)$ of the matrix pair (A_1, A_2) of Example 3.2.4

(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_1, A_2)$ *za matrični par* (A_1, A_2) *iz Primera 3.2.4)*

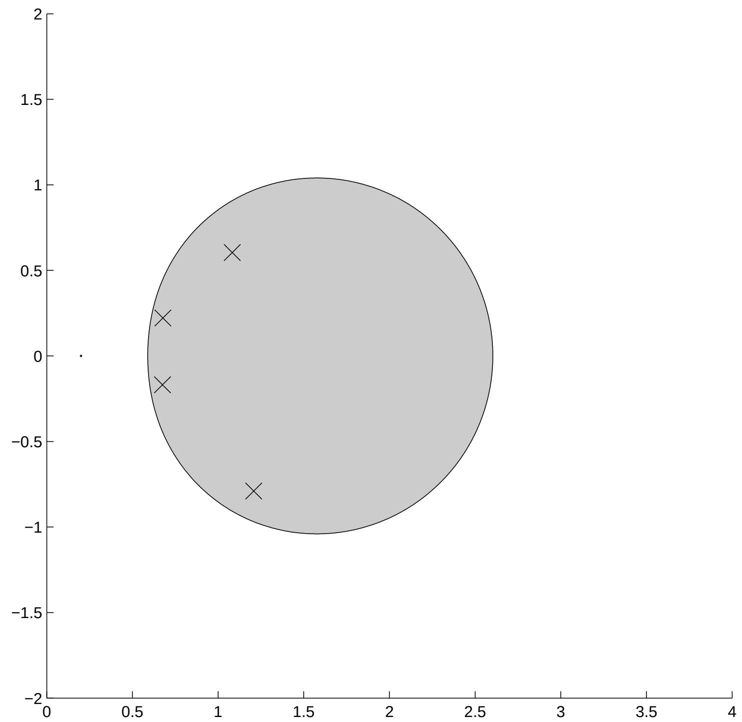

Figure 3.5.4: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_3, B_3)$ of the matrix pair (A_3, B_3) of Example 3.3.14

(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_3, B_3)$ za matrični par (A_3, B_3) iz Primera 3.3.14)

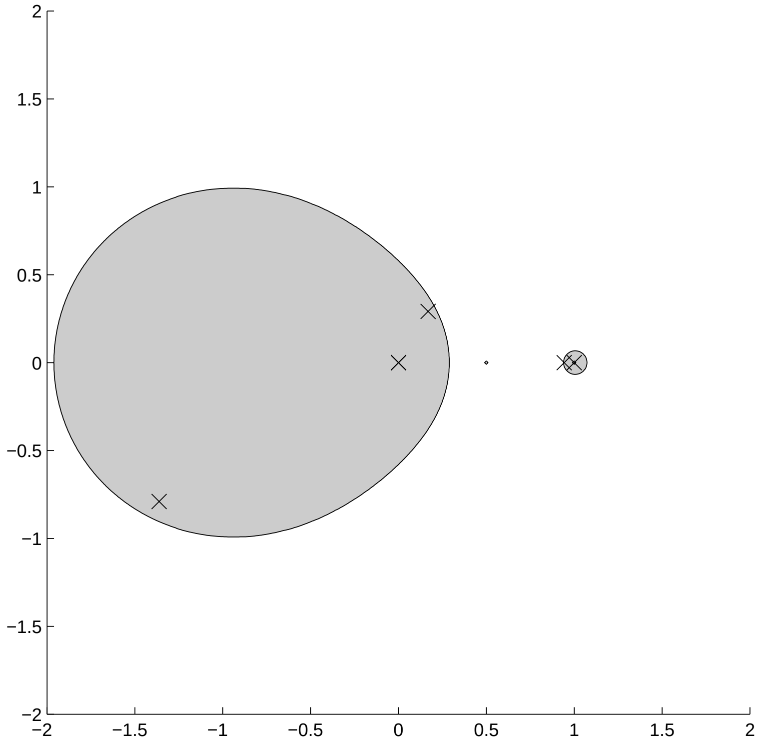

Figure 3.5.5: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_4, B_4)$ of the matrix pair (A_4, B_4) of Example 3.3.19
(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_4, B_4)$ za matrični par (A_4, B_4) iz Primera 3.3.19)

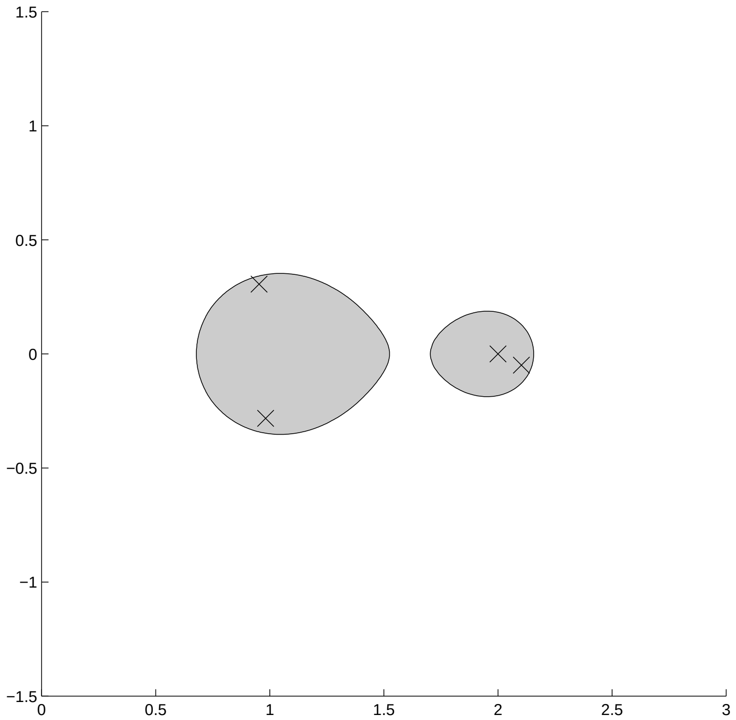

Figure 3.5.6: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_5, B_5)$ of the matrix pair (A_5, B_5) of Example 3.3.43
(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_5, B_5)$ za matrični par (A_5, B_5) iz Primera 3.3.43)

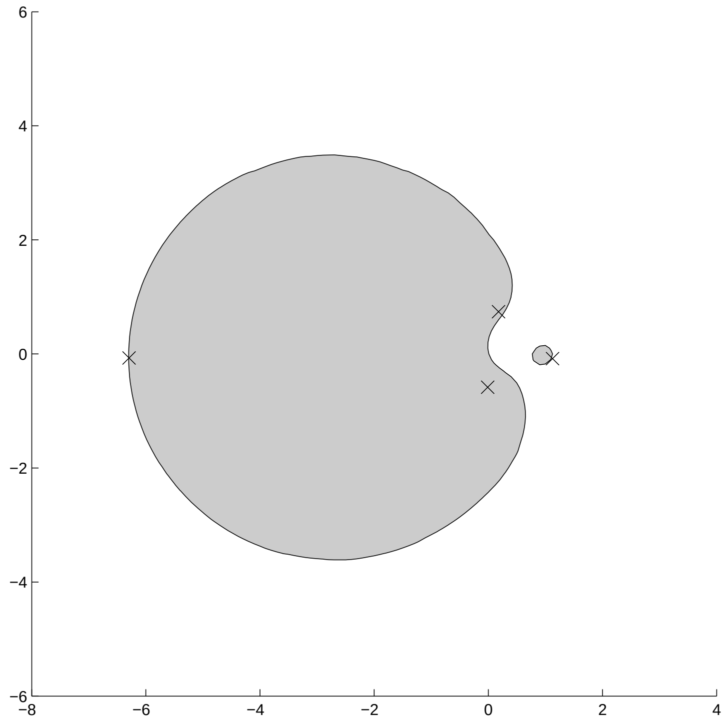

Figure 3.5.7: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_6, B_6)$ of the matrix pair (A_6, B_6) of Example 3.5.11

(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_6, B_6)$ za matrični par (A_6, B_6) iz Primera 3.5.11)

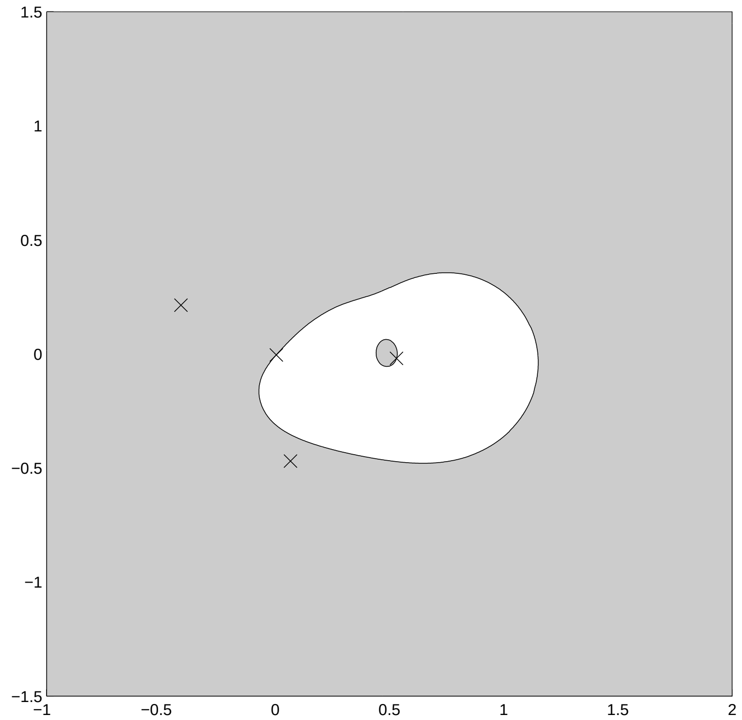

Figure 3.5.8: Generalized minimal Geršgorin set $\Gamma^{\mathbb{D}}(A_7, B_7)$ of the matrix pair (A_7, B_7) of Example 3.5.11
(Generalizovani minimalni Geršgorinov skup $\Gamma^{\mathbb{D}}(A_7, B_7)$ za matrični par (A_7, B_7) iz Primera 3.5.11)

Another way to plot the generalized minimal Geršgorin set is to compute its approximation. A way to do it is a modification of the approach presented in the Subsection 2.3.2. This approach could be useful here if H-matrices are involved, since the result of Theorem 3.5.6 gives the main motivation.

Namely, we are interested in determining the approximations of the star-shaped subsets (3.5.7) and (3.5.8) of the generalized minimal Geršgorin set. Thus, assuming that B is a nonsingular H-matrix, and that $A - zB$ is an irreducible matrix for each $z \in \mathbb{C}$, we start by fixing an index $k \in N$, and the corresponding center $\frac{a_{k,k}}{b_{k,k}}$ of the star-shaped subset of (3.5.7). For each $\theta \in [0, 2\pi]$, since $\nu_{(A,B)}(\frac{a_{k,k}}{b_{k,k}}) > 0$, we can, with a few trial steps, find Δ, $\Delta > 0$, such that $\nu_{(A,B)}(\frac{a_{k,k}}{b_{k,k}} + \Delta) < 0$. Then, we can apply the bisection search to the interval $[\frac{a_{k,k}}{b_{k,k}}, \frac{a_{k,k}}{b_{k,k}} + \Delta]$ to determine $\hat{\varrho}_k(\theta)$. As a result, we have approximated boundary point $\frac{a_{k,k}}{b_{k,k}} + \hat{\varrho}_k(\theta)e^{i\theta}$ of the generalized minimal Geršgorin set.

Finally, moving the angles $\theta \in [0, 2\pi]$, we obtain the approximation of the set of (3.5.7).

Chapter 4

Application of Generalized Diagonal Dominance in Wireless Sensor Network Optimization Problems

In this chapter we present the recent application of diagonal dominance in the development of the optimization algorithms in the wireless sensor networks design, done by J. Yuan and W. Yu in [55], extended in [54], and surveyed in [37]. In their work, authors address the cross-layer optimization problem, that can be decomposed into two subproblems, each corresponding to a separate layer of the overall system (the physical, and the application layer). In order to solve the nonconvex and nonlinear source coding subproblem at the application layer, and the power-allocation subproblem at the physical layer, both in a distributed manner, theory of noncooperative games was used. In that setting, under certain conditions, solutions to both subproblems were obtained as unique and stabile Nash equilibria. The physical-layer power-allocation subproblem is modeled as a power control game, and an iterative algorithm is designed that converges to the desirable solution under the assumption that the certain matrix is strictly diagonally dominant. Here, we will first use generalized diagonal dominance, to improve obtained result on the existence and uniqueness of the Nash equilibrium, and briefly discuss the applicability of such improvement. Then, using the theory of iterative methods for solving systems of linear equations, we will introduce new techniques for the power control algorithm, and prove that they globally converge to asymptotically stable unique Nash equilibrium of the observed power control game.

Wireless sensor networks have a wide range of applications, such as military security, traffic control, and environmental monitoring. A sensor network consists of a large number of sensors, deployed in a field. Each sensor makes a local observation of some underlying physical phenomenon, quantizes its observation, and transfers the data back to a central estimation office (i.e., CEO). Due to the limited transmission power, sensors that are far away from the CEO, deliver their quantization data through a multi-hop network, as shown in Figure 4.0.1. The goal of the sensor network design is to measure and estimate the underlying physical phenomenon, as accurately as possible, under the network resource limitation. Therefore, the sensor network problem is an optimization problem, in which

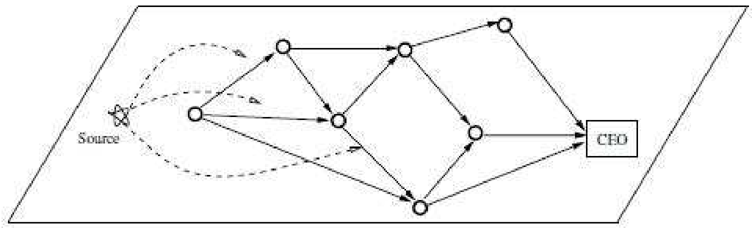

Figure 4.0.1: Wireless sensor network with one CEO
(Bežična senzor mreža sa jednim CEO)

the objective is to the minimize the overall distortion, i.e., the difference between the true underlying field and its estimation at the CEO. But, due to the partial observation at each sensor, the overall estimation error at CEO is a coupled and nonseparable function of all sensors's data rates. In addition, due to the shared nature of the wireless medium, geographically close transmissions often interfere with each other. Thus, in order to manage the above issues, one should consider the fundamental performance limits of sensor networks.

In [55], the authors adopted a separate source-channel coding model and used information theoretical concepts, such as rate-distortion region and capacity region, to explore the fundamental tradeoffs in wireless sensor network design.

They used the set of dual variables to coordinate the interaction between the layers, and they decomposed the overall network optimization problem in the dual domain into two disjoint subproblems: a power control subproblem at the physical layer, and a source coding subproblem at the application layer. Both subproblems are, inherently, nonlinear and nonconvex, and, hence, difficult to solve. And, additionally, since the realistic sensor network deployment often encounters significant variations in source statistics and physical layer channel characteristics, real-time and distributed solving algorithms are needed.

Although in [55] a game-theoretic approach was used to solve two subproblems, we will focus only to the physical layer, and the algorithm used for the solution of the corresponding optimization subproblem.

We start with some preliminaries. Since wireless ad hoc network is characterized by a distributed, dynamic, self-organizing architecture, each node in the network is capable of independently adapting its operation based on the current environment, and according to predetermined algorithms and protocols. Therefore, due to the distributed and dynamic nature of ad hoc networks, the analytical models to evaluate their performance have been infrequent.

Game theory offers a suite of tools that may be used effectively in modeling the interaction among independent nodes in an ad-hoc network, [37]. It is a field of applied mathematics that describes and analyzes interactive decision situations. It provides analytical tools to predict the outcome of complex interactions among rational entities. The entities are *rational* in a way that they have a strict strategy, based on perceived or measured results. The main areas of the application of game theory are economics, political science, biology and sociology. From the early 1990s, engineering and computer science have been added to this list.

A **game** has three components: a **set of players**, a **set of possible actions** for each player, and a **set of strategies**. A **player's strategy** is a complete plan of actions to be taken, when the game is actually played. In a game, players can act selfishly to maximize their gains, and, hence, a distributed strategy for players can provide an optimized solution to the game. In any game, utility represents the motivation of players. A **utility function** describes the players preferences for a given action. It assigns a number for every possible outcome of the game, and it has the property that a higher number implies the outcome that is more preferred. The **payoff function** is the utility function minus the penalty price, also called a **tax**, that a player has to pay for each action. A **Nash equilibrium** is a set of actions of the players, such that any other action, chosen by a single player, results in less favorable utility for every player in a game.

The game formulation that we will use here is **non-cooperative game**, a game where players act selfishly, to maximize their individual payoffs in a distributed decision-making environment. This is in contrast to a **cooperative game** where players agree on premediated strategies to maximize their payoffs.

In a wireless sensor network, the design goal is to minimize the total distortion, by jointly optimizing source coding and power allocation, which can be formulated as follows:

$$\boxed{\begin{array}{ll} \text{minimize} & \alpha^T d \\ \text{subject to} & s \in \mathcal{R}(d),\ c \in \mathcal{C}(p),\ Ac \geq s \end{array}} \tag{4.0.1}$$

where α is a vector, representing the relative emphasis on different elements of the distortion vector d; s is a set of source rates at each node; c is a set of link capacities; and p is the power consumption vector. $\mathcal{R}(d)$ is a fundamental concept in source coding, called the rate-distortion region. The constraint $s \in \mathcal{R}(d)$ models the inter-dependence of the distortion on the source rates. $\mathcal{C}(p)$ is a fundamental concept in channel coding, called the capacity region. The constraint $c \in \mathcal{C}(p)$ models the inter-dependence of the link capacity vector on the power consumption. The last inequality $Ac \geq s$ reflects the fact that the source rate at each node must be less than the link capacity support. Here, A is an $m \times n$ node-incident matrix with m nodes and n links, which, using the *multi-commodity flow routing model*, [4], can be characterized with:

$$a_{i,j} = \begin{cases} 1, & \text{if } i \text{ is a starting node for the link } j, \\ -1, & \text{if } i \text{ is an end node for the link } j, \\ 0, & \text{otherwise.} \end{cases} \tag{4.0.2}$$

Applying the dual decomposition technique, [55], the joint optimization problem (4.0.1) can be, further, decoupled into two distinct subproblems. A power control subproblem at the physical layer:

$$\boxed{\begin{array}{ll} \text{maximize} & \mu^T c \\ \text{subject to} & c \in \mathcal{C}(p), \end{array}} \tag{4.0.3}$$

and a source coding subproblem at the application layer:

$$\boxed{\begin{array}{ll} \text{minimize} & \alpha^T d + \lambda^T s \\ \text{subject to} & s \in \mathcal{R}(d), \end{array}} \tag{4.0.4}$$

where μ is related to the dual variable λ by the *link price consistency equations* $\mu^T = \lambda^T A$. The Lagrange multipliers λ and μ have the interpretation of being the *shadow prices* coordinating the application layer *demand* and the physical layer *supply.*

Now, using the game theory setup, we address the physical layer subproblem that concerns the transmission interference among nearby sensors. Interference management is one of the main challenges in the physical layer design of wireless networks. A key concept at the physical layer is the achievable capacity region, which characterizes a tradeoff between achievable capacities at different links. We consider a network, where, for every link $i \in N$, $g_{i,i}$, p_i, and ξ_i are the **link gain**, **the power action**, and the **noise of the link**, respectively. By $g_{i,j}$ we denote the **gain of the interference** from link j to link i. The values of the **gain matrix** $G = [g_{i,j}] \in \mathbb{R}^{n,n}$ and the **noise vector** ξ are generally obtained through some estimation techniques, and they characterize the channel[1] statistics. Further, we assume that each node has a certain power budget, such that the power action of the link i is limited by p_i^{max}, i.e., $p \leq p^{max} := [p_1^{max}, p_2^{max}, \ldots, p_n^{max}]^T$. Thus, the power control subproblem (4.0.3) with a physical-layer interference model may be formulated as:

$$\boxed{\begin{array}{ll} \text{find} & \mathbf{0} \leq p \leq p^{max} \ \text{ that maximizes } \ \sum_{i \in N} \mu_i c_i, \\ \text{where} & c_i := \log(1 + \text{SINR}_i), \\ & \text{SINR}_i = \dfrac{g_{i,i} p_i}{\sum_{j \in N \setminus \{i\}} g_{i,j} p_j + \xi_i}, \end{array}} \tag{4.0.5}$$

where c_i is the **capacity** of the link $i \in N$, and SINR_i is its **signal to interference and noise ratio**. Because of the interference, the power control subproblem (4.0.5) is a nonconvex optimization problem that is inherently difficult to solve. We use game theory to approach this problem iteratively, and to solve it. In a power control game that is defined in the sequel, each link is modeled as a player with an aim of maximizing its payoff function. In the conventional game theoretic approaches, each link uses its own achievable rate as the payoff function. Competitive equilibria in such a game may not correspond to desirable operating points, especially when the interference level is high. Thus, the payoff function proposed in [55] is such that each player's (i.e. link's) payoff includes not only its achievable rate, but also the interference effect to other links. So, a tax mechanism was introduced into the game, so that the players will have an incentive to intelligently

[1] For simplicity, we consider the case when each link has one channel. The realistic case can be modeled in the same way.

avoid interference, by keeping the signal to interference and noise ratio as high as possible, while at the same time tending to minimize the overall power usage. Mathematically, such a **power control game** consists from the i-th player strategy to maximize its payoff function Q_i, while paying the tax rate t_i, and performing the action p_i. It can be can be expressed as follows:

$$\boxed{\begin{array}{ll} \text{maximize} & Q_i := \mu_i \log\big(1 + \frac{g_{i,i}p_i}{\sum_{j\in N\setminus\{i\}} g_{i,j}p_j + \xi_i}\big) - t_i p_i, \\ \text{by changing} & 0 \le p_i \le p_i^{max}, \end{array}} \tag{4.0.6}$$

where the tax rate for the link $i \in N$, that was proposed in [55], is the rate at which other users' achievable data rates decrease, with an additional amount of power of the link i, i.e.,

$$t_i := \Big|\frac{\partial \sum_{j\in N\setminus\{i\}} \mu_j c_j}{\partial p_i}\Big| = \sum_{k\in N\setminus\{i\}} \frac{\mu_k g_{k,i} g_{k,k} p_k}{(g_{k,k}p_k + \sum_{j\in N\setminus\{k\}} g_{k,j}p_j + \xi_k)(\sum_{j\in N\setminus\{k\}} g_{k,j}p_j + \xi_k)}. \tag{4.0.7}$$

Here, the more power link i uses, the more interference it will produce to others, and, therefore, more tax (i.e., $t_i p_i$) it has to pay.

The power vector p that solves the optimization problem (4.0.5) is, exactly, the Nash equilibrium of the power control game (4.0.6). Since, in general, not every game has a Nash equilibrium, and neither is the equilibrium necessarily stable, at first, the goal is to prove the existence, uniqueness and stability of the Nash equilibrium for the power control game. Then, the aim is to design a distributed iterative algorithm that will converge to that equilibrium. In [55], a power control game algorithm is proposed. It consists of two phases: the power update, and the tax update. The power update is based upon the fact that, at each step, every player $i \in N$ tries to maximize its own payoff Q_i, while assuming that the power levels of all other players and the taxes are fixed. The expression for such an optimal p_i^* is, then, obtained by setting the derivative Q_i with respect to p_i to zero, i.e., $\frac{\partial Q_i}{\partial p_i} = 0$, and it is called the **best response function** of the player i, denoted by $B_i(p)$. In such a way we have obtained a locally optimal power vector p^*, with the property that for every $i \in N$, p_i^* strikes a balance between maximizing its own rate and minimizing its interference to other links (which is taken into account via t_i). For example, a large value of tax rate t_i indicates that the link i is producing severe interference to other links. This is reflected in the power update, as the larger t_i leads to a lower p_i. Although each player appears to be selfish in maximizing only its own payoff, since the payoff function incorporates social welfare, the Nash equilibrium of this game is, in fact, a *cooperative social optimum.*

Therefore, calculating the locally optimal power vector p^* consists in solving the system of equations $\frac{\partial Q_i}{\partial p_i} = 0$, for $i \in N$, which can be expressed in an equivalent form as $p = B(p) := [B_1(p), B_2(p), \ldots, B_n(p)]^T$. Thus, p^* can be seen as a fixed point of the best response vector function. This approach was used in [55] to obtain the existence, uniqueness and the dynamical stability of the power control game. Here, we will write the system

$\frac{\partial Q_i}{\partial p_i} = 0$, for $i \in N$, in a matrix form:

$$Gp = D_G D_t^{-1} \mu - \xi, \tag{4.0.8}$$

where $G = [g_{i,j}] \in \mathbb{R}^{n,n}$ is the gain matrix of the links in the wireless network, $D_G :=$ diag$(g_{1,1}, g_{2,2}, \ldots, g_{n,n})$ its diagonal part, $D_t :=$ diag$(t_1, t_2, \ldots, t_n)$ diagonal matrix of the tax rates, $\xi \in \mathbb{R}^n$ the noise vector, and $\mu \in \mathbb{R}^n$ is the vector of dual variables. This formulation of the problem will alow us to generalize the work of Yuan and Yu, as it will be shortly presented.

Once locally optimal power vector is obtained, the algorithm proceeds with the tax rate update, using the formula (4.0.7). As it was given in [55], tax rate can be expressed through the signal to noise ratios in the form that is convenient for the *distributed implementation.* Here distributed implementation signifies that the tax rate update is directly calculated from the information that is received through each individual link. Namely,

$$t_i = \sum_{j \in N \setminus \{i\}} g_{i,j} b_j, \quad (i \in N) \tag{4.0.9}$$

where

$$b_i = \mu_i \frac{\text{SINR}_i}{g_{i,i} p_i} \frac{\text{SINR}_i}{1 + \text{SINR}_i}, \tag{4.0.10}$$

is the broadcast message of the link $i \in N$.

Although it is clear that the tax rate vector t is calculated from the actual power vector p, in each power update step this vector is fixed, and therefore the system (4.0.8) is the system of linear equations with the system matrix G.

Now we give the general framework for the power control game algorithm:

Power Control Game Algorithm

1. Initialize $p^{(0)}$ and $t^{(0)}$, and set $l = 0$.

2. Iteratively determine p^*, such that

$$Gp = D_G D_{t^{(l)}}^{-1} \mu - \xi,$$

   where $D_{t^{(l)}} :=$ diag$(t_1^{(l)}, t_2^{(l)}, \ldots, t_n^{(l)})$, and set $p^{(l+1)} := p^*$.

3. For each link $i \in N$, calculate signal to noise ratio

$$\text{SINR}_i^{(l+1)} = \frac{g_{i,i} p_i^{(l+1)}}{\sum_{j \in N \setminus \{i\}} g_{i,j} p_j^{(l+1)} + \xi_i},$$

   and the broadcast message

$$b_i^{(l+1)} = \mu_i \frac{\text{SINR}_i^{(l+1)}}{g_{i,i} p_i^{(l+1)}} \frac{\text{SINR}_i^{(l+1)}}{1 + \text{SINR}_i^{(l+1)}}.$$

4. For each link $i \in N$, update the tax rate

$$t_i^{(l+1)} = \sum_{j \in N \setminus \{i\}} g_{i,j} b_j^{(l+1)}$$

5. Set $l := l + 1$, and return to step 2. until convergence.

In [55], one of the main results is that, under the condition that the gain matrix G is an SDD matrix, the power control game is asymptotically stable, and it (more precisely, its power control game algorithm) always converges to the unique Nash equilibrium. But, through the simulations, the authors have noticed that, even if this condition is *not* satisfied, the power control game could converge nicely. Namely, although it may seem as a natural condition, SDD property of the overall gain matrix G could be ruined, due to the stronger interferences inherent to the link topology, or to the state of the medium, through which the carrier wave is propagated. But, sometimes, while the augmented interference rates, that come from the specific links, are such that, for some link $i \in N$, the link gain $g_{i,i}$ is dominated by the interferences from other links $\sum_{j \in N \setminus \{i\}} g_{i,j}$, these interferences will not cross the point at which the power control algorithm fails. Therefore, a natural question is whether we can improve our theoretical results, in order to guarantee the convergence and the stability of the power control game, and its iterative procedure, in wider range of real situations.

As we have seen in the previous chapters, the "core" properties that have strictly diagonally dominant matrices can be obtained by the use of the generalized diagonally dominant matrices, where the problem of determining if the given matrix has desired property could be solved through the use of the numerous subclasses of H-matrices. So, here we formulate the generalization of the Yuan-You's theorem on the stability and the convergence of the power control game.

Theorem 4.0.12. *Given a wireless sensor network with certain link topology, let* $G = [g_{i,j}] \in \mathbb{R}^{n,n}$*,* $G \geq O$*, be the overall gain matrix, and* $\xi \in \mathbb{R}^n$*,* $\xi \geq \mathbf{0}$*, the overall link noise vector. If* G *is a generalized diagonally dominant matrix, then the power control game, given by (4.0.6), where the tax rate vector* t *is defined by (4.0.7), has a unique stable Nash equilibrium* p^**. Moreover, the game is asymptotically stable, and the power control game algorithm converges to* p^**, for each starting nonnegative vectors* $p^{(0)}, t^{(0)} \in \mathbb{R}^n$*.*

Proof. First, for every $i \in N$, the payoff function Q_i, given in (4.0.6), of the link (player) i is continuous in p, and strictly concave in p_i (which can be verified by computing its Hessian). Hence, since the i-th link action profile $[0, p_i^{max}]$ is a compact convex set, by Theorem 4.3 in [1], it follows that the power control game has at least one pure Nash equilibrium, which can be found as an intersection point of the reaction curves of all the players. Namely, if by p^* we denote Nash equilibrium of the game (4.0.6), p^* satisfies the system of linear equations (4.0.8), i.e., $Gp^* = D_G D_t^{-1}\mu - \xi$. But, since G is the GDD matrix, it is nonsingular, and, thus, $p^* = G^{-1}(D_G D_t^{-1}\mu - \xi)$ is the unique Nash equilibrium of the power control game (4.0.6).

In order to prove that this (local) Nash equilibrium is stable, as in [55], we prove the asymptotic stability of the game (4.0.6). We will use the concept of the best response function $B(p)$, and the **dynamic stability matrix** $\Delta := [\Delta_{i,j}]$, where $\Delta_{i,j} := \frac{\partial B_i(p)}{\partial p_j}$, for $i, j \in N$. According to [24], the game is asymptotically stable if all the eigenvalues of the dynamic stability matrix lie in the unit circle, i.e., if $\rho(\Delta) < 1$.

In our case, for $i \in N$, the best response of the link i is:

$$B_i(p) = \frac{\mu_i}{t_i} - \frac{1}{g_{i,i}}\Big(\sum_{j \in N\setminus\{i\}} g_{i,j}p_j + \xi_i\Big),$$

and, thus,

$$\Delta i, j = \frac{\partial B_i(p)}{\partial p_j} = -\frac{g_{i,j}}{g_{i,i}},$$

for $j \in N \setminus \{i\}$, while $\Delta_{i,i} = 0$.

Now, since $G \geq O$ is a GDD matrix, there exists a matrix $X = \text{diag}(x_1, x_2, \ldots, x_n) \in \mathbb{D}$, such that GX is an SDD matrix, i.e.,

$$g_{i,i}x_i > \sum_{j \in N\setminus\{i\}} g_{i,j}x_j, \quad (i \in N),$$

or, equivalently,

$$r_i^{\mathbf{x}}(\Delta) = \sum_{j \in N\setminus\{i\}} \frac{g_{i,j}x_j}{g_{i,i}x_i} < 1, \quad (i \in N),$$

where $r_i^{\mathbf{x}}(\Delta)$ is defined in 2.1.7.

On the other hand, by the Corollary 2.1.5,

$$\sigma(\Delta) \subseteq \Gamma^X(\Delta) := \Gamma(X^{-1}\Delta X) = \bigcup_{i \in N} \Gamma_i(X^{-1}\Delta X).$$

So, for every eigenvalue $\lambda \in \sigma(\Delta)$, there exists $i \in N$, such that

$$|\lambda - \Delta_{i,i}| \leq r_i^{\mathbf{x}}(\Delta),$$

and consequently $|\lambda| < 1$. To complete the proof, we observe that the sequence of tax rates is convergent, and therefore, the power control algorithm converges for every starting vectors $p^{(0)}, t^{(0)} \in \mathbb{R}^n$. □

A simple corollary of the previous theorem is the following one.

Corollary 4.0.13. *Given any nonsingular DD-type class of matrices $\mathbb{K}$, and the wireless sensor network with the prescribed link topology, let $G = [g_{i,j}] \in \mathbb{R}^{n,n}$, $G \geq O$, be the overall gain matrix, and $\xi \in \mathbb{R}^n$, $\xi \geq \mathbf{0}$, be the overall link noise vector. If $G \in \mathbb{K}$, then the power control game, given by (4.0.6), where the tax rate vector t is defined by (4.0.7), has a unique stable Nash equilibrium p^*. Moreover, the game is asymptotically stable, and the power control game algorithm converges to p^*, for each starting nonnegative vectors $p^{(0)}, t^{(0)} \in \mathbb{R}^n$.*

As we have seen throughout the first chapter of this thesis, there are quite a few subclasses of H-matrices that are significantly wider that the SDD class of matrices. Thus, the improvement we have made by the Theorem 4.0.12 opens a new possibilities of different network setups, for which we can guarantee that the solution of the power control game, and, thus, of the optimization problem, is approximated by the iterative algorithm. To check whether the obtained sufficient conditions are fulfilled, according to the Corollary 4.0.13, we can use Theorems 1.2.1, 1.2.6, 1.2.13, 1.2.21 or 1.2.22, of this thesis, as well as as many different results on subclasses of H-matrices that could be found in the literature, [13, 28, 29, 30, 33, 51].

It is interesting to note that the topology of the wireless network can lead to a specific structure of the matrix G. Namely, knowing that the interferences between the links occur if the links are "close" to each other, for certain network topologies we can have specific patterns of matrix entries. Therefore, matrix properties, like block forms and reducibility, could be used, in order to obtain different improvements in modeling wireless sensor networks.

Another interesting application of the generalized diagonal dominance lies in the usage of S-SDD matrices, of Theorem 1.2.13, and the underlying scaling technique. Namely, if the overall gain matrix G is an SDD matrix, it is an S-SDD matrix, too, for an arbitrary set of links $S \subseteq N$. Therefore, we can use the information contained in the adequate scaling matrix $X \in \mathbb{X}_S$, in order to introduce more freedom in the management of the power resources. Namely, given the multi-hop wireless sensor network with the SDD overall gain matrix G, assume that several nodes work with the severe power constraints, but due to the their location in the network topology, they have to be deployed for measuring and/or transmitting. In this case, we would like to prolong the life time of such relays, and the overall optimization of such network should *additionally* minimize the power action of such links, while achieving the Nash equilibrium of the power control game, which maximizes the overall network capacity. To address this issue, we will use the scaling technique developed in the Section 1.4 of this thesis.

First, let M_0 be the set of nodes that are having restrictive power consumption. Having the node-link incidence matrix $A = [a_{i,j}]$, given by (4.0.2), we define the set of power restricted links $L := \{j \in N : a_{i,j} \neq 1, \ i \in M_0\}$. Since the gain matrix G is an SDD matrix, then it is also an S-SDD, where $S = L$, i.e., meaning that for each $i \in L$, and every $j \in \overline{L} := N \setminus L$,

$$(g_{i,i} - r_i^L(G))(g_{j,j} - r_j^{\overline{L}}(G)) > r_i^{\overline{L}}(G) r_j^L(G), \text{ and}$$

$$g_{i,i} > r_i^L(G),$$

where $r_i^L(A) := \sum_{j \in L \setminus \{i\}} g_{i,j}$. But, using the quantities $\alpha_L(G)$ and $\beta_L(G)$, as defined in (1.4.13), we can see that

$$0 \leq \alpha_L(G) := \min_{i \in L} \frac{r_i^{\overline{L}}(G)}{(g_{i,i} - r_i^L(G))} < 1 < \max_{j \in \overline{L},\, r_j^L(G) \neq 0} \frac{g_{j,j} - r_j^{\overline{L}}(G)}{r_j^L(G)} =: \beta_L(G),$$

and that for each $\gamma \in (\alpha_L(G), \beta_L(G))$, the matrix $\widetilde{G} := GX$ is SDD, where $X = \mathrm{diag}(x_1, x_2, \ldots, x_n)$, with

$$x_j = \begin{cases} \gamma, & \text{if } j \in L \\ 1, & \text{otherwise.} \end{cases}$$

Thus, by setting $\widetilde{p}_i := \frac{p_i}{x_i}$, for $i \in N$, the power control game (4.0.6) becomes

$$\boxed{\begin{array}{ll} \text{maximize} & Q_i := \mu_i \log\left(1 + \frac{\widetilde{g}_{i,i}\widetilde{p}_i}{\sum_{j\in N\setminus\{i\}} \widetilde{g}_{i,j}\widetilde{p}_j + \xi_i}\right) - \widetilde{t}_i\widetilde{p}_i, \\ \text{by changing} & 0 \le \widetilde{p}_i \le p_i^{max}, \end{array}} \tag{4.0.11}$$

where the tax rate $\widetilde{t}_i := t_i x_i$, for each link $i \in N$, satisfies

$$\widetilde{t}_i := \sum_{k\in N\setminus\{i\}} \frac{\mu_k \widetilde{g}_{k,i}\widetilde{g}_{k,k}\widetilde{p}_k}{(\widetilde{g}_{k,k}\widetilde{p}_k + \sum_{j\in N\setminus\{k\}} \widetilde{g}_{k,j}\widetilde{p}_j + \xi_k)(\sum_{j\in N\setminus\{k\}} \widetilde{g}_{k,j}\widetilde{p}_j + \xi_k)}. \tag{4.0.12}$$

Now, having that the matrix $\widetilde{G}$ is SDD, using the Theorem 4.0.12, we have that the game (4.0.11) is asymptotically stable, and we obtain the unique Nash equilibrium $\widetilde{p}^*$ that satisfies the equality $\widetilde{p}^* = X^{-1}p^*$. Thus,

$$\widetilde{p}_i^* = \begin{cases} \gamma^{-1}p_i^*, & \text{if } i \in L \\ p_i^*, & \text{otherwise.} \end{cases}$$

The obtained relation shows the connection between the originally obtained vector of the link power action that is Nash equilibrium of the power control game, and the the new one. Since the links $i \in L$ are such that it is desirable to have as small as possible power action, we wish to adjust the parameter γ to be bigger than 1, and as big as possible. But, since we have that $1 < \beta_L(G)$, by choosing $1 < \gamma < \beta_L(G)$ to be sufficiently close to the value of $\beta_L(G)$, for each $i \in L$, we have that $\widetilde{p}_i^* < p_i^*$, while, for $i \in \overline{L}$, $\widetilde{p}_i^* = p_i^*$. Therefore, we have obtained the unique Nash equilibrium that better suites the power constraints of the given wireless sensor network.

In the sequel, we focus on the power control algorithm (briefly **PCA**), and we are interested to improve its convergence speed. Since one part of the overall energy consumption in wireless sensor network is spent in calculation, used to implement the PCA, the complexity of calculation, and the speed of convergence are issues that should be treated. We start with the observation that PCA consists of the inner iteration and outer iteration. The inner iteration is the power allocation vector update (step 2.), performed at each step l. The outer iteration consists of the tax rate update through broadcast message vector. The original algorithm that was given in [55], in step 2. of PCA performed, at each link, a fixed point iteration, using the best response function of the concerned link. Namely, the given inner iterative procedure is given by $p^{(k+1)} := B(p^{(k)}) = [B_1(p^{(k)}), B_2(p^{(k)}), \ldots, B_n(p^{(k)})]^T$, for any $p^{(0)}$, and all $l \in \mathbb{N}$. Equivalently, this can be written as:

$$\boxed{p_i^{(k+1)} := \frac{\mu_i}{t_i^{(l)}} - \frac{1}{g_{i,i}}\left(\sum_{j\in N\setminus\{i\}} g_{i,j}p_j^{(k)} + \xi_i\right), \quad (i \in N)\,(k \in \mathbb{N}),} \tag{4.0.13}$$

where $p^{(0)}$ is arbitrary. When, at some step k_0, the iterative approximation is satisfactory, the link power allocation at that step is, then, forwarded to the outer iteration, i.e., to the tax rate update. The convergence of this procedure was obtained through the argument that the best response function of the link i is a contraction. The described procedure can be seen as distributed one, meaning that the power update for link i is obtained by the calculation that can be implemented using exclusively the information that link i is capable to measure. Therefore, each link is capable to make its own power update, using the actual power consumption vector of the overall network. The similar argument stands for the tax rate update, too. Finally, under the assumption that the gain matrix is SDD, the authors proved the asymptotic convergence of the game, and, thus, the convergence of PCA.

Here we will address only the inner iteration. We propose new iterative procedures, discuss their implementation and convergence. The main idea is based upon the fact that the locally optimal equilibrium p^* can be obtained as the solution of the linear system (4.0.8). Under the assumption that the overall gain matrix is an H-matrix, in the previous considerations, we have proven the asymptotic stability of the game (4.0.6), and, hence, obtained the convergence of the PCA. Therefore, if we obtain, under the same condition, the convergence of the new procedures for the inner iteration, the modified PCA will also converge to the Nash equilibrium of the power control game (4.0.6).

Given wireless network with the certain topology of n links and the overall gain matrix $G = [g_{i,j}] \in \mathbb{R}^{n,n}$, by $D_G := \operatorname{diag}(g_{1,1}, g_{2,2}, \ldots, g_{n,n})$ denote the diagonal part of G, and write $B_G := G - D_G$. For the fixed tax rates $t = [t_1, t_2, \ldots, t_n]^T$ of the power control game (4.0.6), define $D_t := \operatorname{diag}(t_1, t_2, \ldots, t_n)$. Let $\xi \in \mathbb{R}^n$ be the noise vector, and $\mu \in \mathbb{R}^n$ the vector of dual variables in the power control subproblem (4.0.3) of the cross-layer optimization problem (4.0.1). If G is an H-matrix, then, the locally optimal vector of the link power allocation p^* is the unique solution of the system of linear equations

$$Gp = D_G D_t^{-1}\mu - \xi. \tag{4.0.14}$$

If we use the splitting of the matrix $G = D_G - B_G$, then we can write (4.0.14) in the fixed point form $p = D_G^{-1}(B_G p - \xi) + D_t^{-1}\mu$, and define the iteration procedure $p^{(k+1)} = D_G^{-1}(B_G p^{(k)} - \xi) + D_t^{-1}\mu$. Since we took the Jacobi splitting of the system matrix G, the iterative method is the famous **Jacobi iteration**. On the other hand, it is easy to see that this procedure is exactly (4.0.13), the one proposed by Yuan and Yu.

The other fundamental procedure in the theory of iterative methods is, of course, the Gauss-Seidel iteration scheme. Given a matrix $G = [g_{i,j}]$, consider the standard splitting $G = D_G - L_G - U_G$, where D_G is a diagonal matrix, while L_G and U_G are, respectively, strictly lower and strictly upper triangular matrices. More precisely, let $D = \operatorname{diag}(g_{1,1}, g_{2,2}, \ldots, g_{n,n})$, $L = [l_{i,j}]$, where

$$l_{i,j} = \begin{cases} -g_{i,j}, & j < i, \\ 0, & \text{otherwise,} \end{cases}$$

and $U = [u_{i,j}]$, where

$$u_{i,j} = \begin{cases} -g_{i,j}, & j > i, \\ 0, & \text{otherwise.} \end{cases}$$

Then, **Gauss-Seidel** iterative method for the system (4.0.14) can be written as:

$$D_G p^{(k+1)} = L_G p^{(k+1)} + U_G p^{(k)} - \xi + D_G D_t^{-1} \mu, \quad (k \in \mathbb{N}).$$

In the form of the power update procedure for each link $i \in N$, we obtain

$$\boxed{p_i^{(k+1)} := \frac{\mu_i}{t_i^{(l)}} - \frac{1}{g_{i,i}} \big(\textstyle\sum_{j=1}^{i-1} g_{i,j} p_j^{(k+1)} + \sum_{j=i+1}^{n} g_{i,j} p_j^{(k)} + \xi_i \big), \quad (i \in N)\, (k \in \mathbb{N}).} \qquad (4.0.15)$$

Here, we assume that the link 1 is first to update its power, then link 2, link 3, and, finally, the link n. Then, using the Jacobi iteration, at the time we compute the i-th link power, the updated powers from all of the previous links are already available. Thus, the natural thing to do is to use them. In this way we, naturally, obtain the Gauss-Seidel iteration (4.0.15). Since the system matrix is an H-matrix, it is well known that Gauss-Seidel iterative method is globally convergent. But, although there are many examples where the Gauss-Seidel iteration is preferable than the Jacobi iteration, we cannot state that in general iterative procedure (4.0.15) works faster than (4.0.13).

While often performing faster then Jacobi iteration, Gauss-Seidel iteration has, in this case, a significant drawback. Namely, due to the sequentiality, the link that has to update its power has often to wait its turn. But this is not necessary, since each link is updating its power with the data it has already collected. Therefore, the Gauss-Seidel procedure, while in theory good, behaves rather poorly in the wireless sensor networks due to the link's computational stand-by time in the iterative procedure in the step 2. of PCA. The answer to this drawback of Gauss-Seidel is the **chaotic asynchronous relaxation**, developed by Chazan and Miranker in [9]. Without going into detail notation, we remark that this algorithm uses the same rule as (4.0.13) while the power levels on the right hand side are not necessarily from the same iteration step. Namely, each link uses the most recent powers of other links to update its own. The main value of this algorithm is that, in a wireless sensor network, it behaves very good, in a way that it avoids the link stand-by time due to asynchronous computations in power update iterations, while it allows the distributed implementation. The only issue that needs to be addressed is the convergence. But, the fundamental theorem on the chaotic asynchronous relaxation states that this iterative method converges if all the eigenvalues of the modulus of the Jacobi iteration matrix lie inside the open unit disk, in our case, if $\rho(|D_G^{-1}B_G|) < 1$. But, if the overall gain matrix G is an H-matrix, this is true. To prove it, assume that $\lambda \in \sigma(|D_G^{-1}B_G|)$. Since G is an H-matrix, then there exists $X = \text{diag}(x_1, x_2, \dots, x_n) \in \mathbb{D}$, such that GX is an SDD. But, from (2.1.9), there exists $i \in N$, such that $\lambda \in \Gamma_i^X(|D_G^{-1}B_G|)$, and, hence, $|\lambda| \leq \sum_{j \in N \setminus \{i\}} \frac{g_{i,j} x_j}{g_{i,i} x_i} < 1$. So, implementing the chaotic asynchronous relaxation procedure in the step 2. of PCA, the algorithm will converge to the unique and stable Nash equilibrium of the power control game.

Bibliography

[1] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM, Philadelphia, 1999.

[2] R. Beauwens, *Semistrict diagonal dominance*, SIAM J. Numer Anal. **13** (1976), 109–112.

[3] A. Berman and R. J. Plemons, *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, vol. 9, SIAM, Philadelohia, 1994.

[4] D. Bertsekas and R. G. Gallager, *Data Networks*, Prentice Hall, New Jersey, 1991.

[5] A. Brauer, *Limits for the characteristic roots of a matrix II*, Duke Math. J. **14** (1947), 21–26.

[6] R. Bru, Lj. Cvetković, V. Kostić and F. Pedroche, *Sums of σ-strictly diagonally dominant matrices*, Linear and Multilinear Algebra **58** (2009), no. 1, 75–78.

[7] R. Brualdi, *Matrices, eigenvalues and directed graphs*, Liner and Multilinear Algebra **11** (1982), 143–165.

[8] ———, *The symbiotic relationship of combinatorics and matrix theory*, Linear Algebra Appl. **162/164** (1992), 65–105.

[9] D. Chazan and W. L. Miranker, *Chaotic relaxation*, Linear Algebra Appl. **2** (1969), 199–222.

[10] Lj. Cvetković, *H-matrix theory vs. eigenvalue localization*, Numerical Algorithms **42** (2006), no. 3-4.

[11] Lj. Cvetković, R. Bru, V. Kostić and F. Pedroche, *A simple generalization of Geršgorin's theorem*, Advances in Computational Mathematics, (in print).

[12] ______, *Characterization of aplha1 and alpha2-matrices*, Central European Journal of Mathematics **8** (2010), 32–40.

[13] Lj. Cvetković and V. Kostić, *New criteria for identifying H-matrices*, J. Comput Appl. Math. (2005), 465–478.

[14] ______, *Between Geršgorin and the the minimal Geršgorin set*, J. Comput Appl. Math. **196** (2006), no. 2, 452–458.

[15] ______, *New subclasses of block H-matrices with applications to parallel decomposition-type relaxation methods*, Numerical Algorithms **42** (2006), no. 3-4, 325–334.

[16] ______, *A note on the convergence of the AOR method*, Appl. Math. Comput. **194** (2007), no. 2, 394–399.

[17] Lj. Cvetković, V. Kostić, M. Kovačević and T. Szulc, *Further results on H-matrices and their Schur complements*, Appl. Math. Comput. **198** (2008), no. 2, 506–510.

[18] Lj. Cvetković, V. Kostić and S. Rauški, *A new subclass of H-matrices*, Appl. Math. Comput. **208** (2009), 206–210.

[19] Lj. Cvetković, V. Kostić and R. S. Varga, *A new Geršgorin-type eigenvalue inclusion set*, ETNA (Elec. Trans. on Numer. An.) **18** (2004), 73–80.

[20] L. S. Dashnic and M. S. Zusmanovich, *K voprosu o lokalizacii harakteristicheskih chisel matricy*, Zh. vychisl, matem, i matem, fiz. **10** (1970), no. 6, 1321–1327.

[21] ______, *O nektoryh kriteryah regulyarnosti matric i lokalizacii ih spectra*, Zh. vychisl, matem, i matem, fiz. **10** (1970), no. 5, 1092–1097.

[22] J. Desplanques, *Thèorém d'algébre*, J. de Math. Spec. **9** (1887), 12–13.

[23] M. Fiedler and V. Pták, *On matrices with nonpositive diagonal elements and positive principal minors*, Czech. Math. J. **12** (1962), 382–400.

[24] D. Fudenberg, *The Theory of Learning in Games*, MIT Press, Massachusetts, 1998.

[25] S. Geršgorin, *Über die Abgrenzung der Eigenwerte einer Matrix*, Izv. Akad. Nauk SSSR Ser. Mat. **1** (1931), 749–754.

[26] Y. Gao and X. Wang, *Criteria for generalized diagonally dominant matrices*, Linear Algebr. Appl. **169** (1992), 257–268.

[27] J. Hadamard, *Leçones sur la propagations des ondes*, Hermann et fils, Paris, 1903, reprinted in 1949 by Chelsea, New York.

[28] A. J. Hoffman, *Gersgorin variations I: on a theme of Pupkov and Solovev*, Linear Algebra Appl. **304** (2000), 173–177.

[29] ______, *Gersgorin variations II: On themes of Fan and Gudkov*, Advances in Computational Mathematics **25** (2006), 1–6.

[30] T. Z. Huang, *A note on generalized diagonally dominant matrices*, Linear Algebra Appl. **225** (1995), 237–242.

[31] K. James and W. Riha, *Convergence criteria for successive overrelaxation.*

[32] M. Karow, *Geometry of Spectral Value Sets*, Ph.D. thesis, University of Bremen, Bremen, Germany, 2003.

[33] L. Yu. Kolotilina, *Generalizations of Ostrowski-Brauer theorem*, Linear Algebra Appl. **364** (2003), 65–80.

[34] V. Kostić, *Lokalizacija karakterističnih korena teoremama geršgorinovog tipa*, Master's thesis, University of Novi Sad, Faculty of Science, Department of Mathematics and Informatics, November 2009, (in serbian).

[35] V. Kostić, Lj. Cvetković and R. S. Varga, *Geršgorin-type localizations of generalized eigenvalues*, Numerical Linear Algebra with Applications **16** (2009), no. 11-12, 883–898.

[36] L. Lévy, *Sur le possibilitè du l'equlibre èlectrique*, C. R. Acad. Paris **93** (1881), 706–708.

[37] R. Machado and S. Tekinay, *A survey of game-theoretic approaches in wireless sensor networks*, Computer Networks **52**, no. 16.

[38] H. Minkowski, *Zur Theorie der Einheiten in den algebraischen Zahlkörpern*, Nachr. Königlichen Ges. Wiss. Göttingen Math. - Phys. Kl. **90-93**, no. 1.

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

[40] ______, *Solution of equations and systems of equations*, Academic Press, New York, 1960.

[41] ______, *Über die Determinanten mit Überwiegender Hauptdiagonale*, Coment. Math. Helv. **10** (37), 69–96.

[42] H. Rohrbach, *Bemerkungen zu einem Determinantensatz von Minkowski*, Jahresber. Deutch. Math. Verein. **40** (1931), 49–53.

[43] R. Stewart, *Geršgorin theory for generalized eigenvalue problem* $ax = \lambda x$, Math. Comput. **29** (1975), 600–606.

[44] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[45] O. Taussky, *Bounds for the characteristic roots of matrices*, Duke Math. J. **15** (1948), 1043–1044.

[46] ______, *A recurring theorem on determinants*, Amer. Math. Month. **56** (1949), 672.

[47] R. S. Varga, *Minimal Gerschgorin sets*, Pac. J. Math. **15** (1965), 719–729.

[48] ______, *Matrix Iterative Analysis, Second revised and expanded edition*, Springer-Verlag, Berlin, 2000.

[49] ______, *Gerschgorin disks, Brauer ovals of Cassini (a vindication), and Brualdi sets*, Information **4** (2001), 171–178.

[50] ______, *Gerschgorin-type eigenvalue inclusion theorems and their sharpness*, ETNA (El. Trans. on Numer. An.) **12** (2001), 113–133.

[51] ______, *Geršgorin and His Circles*, Springer-Verlag, New York, 2004.

[52] R. S. Varga, Lj. Cvetković and V. Kostić, *Aproximation of the minimal Geršgorin set of a square complex matrix*, ETNA (Elec. Trans. on Numer. An.) **40** (2008), 398–405.

[53] R. S. Varga and A. Kraustengl, *On Gerschgorin-type problems and ovals of Cassini*, ETNA (El. Trans. on Numer. An.) **8** (1999), 15–20.

[54] J. Yuan, Z. Li, W. Yu and B. Li, *A cross-layer optimization framework for multihop multicast in wireless mesh networks*, Journal on Selected Areas in Communications, **24** (2006), 2092-2103.

[55] J. Yuan and W. Yu, *Distributed cross-layer optimization of wireless sensor networks: A game theoretic approach*, Proc. GLOBECOM2006 (San Francisco, CA, U.S.A.), Nov. 2006.

Short Biography

Vladimir Kostić received a B.S. in Mathematics in 2003, and MSc in Mathematics in 2009, from the Department of Mathematics and Informatics, Faculty of Science at the University of Novi Sad.

From 2004 he is working as teaching assistant to Professor Ljiljana Cvetković in Numerical Linear Algebra at the Department of Mathematics and Informatics, Mathematics and Statistics and Software packages for data processing at the Department of Biology and Ecology.

For his scientific and academic achievements, he received *1st Young Scientists Award* for a talk given on MAT-TRIAD 2007 Conference, Bedlewo, Poland, *The Best Student of the University of Novi Sad Award* in 2004, *The Honor Mileva Marić-Einstein* in 2004, *Fellowship of the Fund of the Royal House of Karadjordjević* in 2002/03, *Award for high academic achievements* by the Embassy of the Kingdom of Norway 2002 and *The Annual University Honors* in 2001, 2002 and 2003.

He has published 19 scientific papers in high rated international journals, and was coauthor on 2 student manuals and one monograph.

In his scientific work he had one invited lecture at Banach Center, Warszawa, Poland, in 2009, one plenary lecture at the conference MAT-TRIAD 2009, Bedlewo, Poland, and he had 13 short contributions on international conferences in the field of Applied and Numerical Linear Algebra and Scientific Computing.

He was secretary of the Organizing Committee of two international conferences.

He is a member of *International Linear Algebra Society*, *GAMM (Gesselshaft fur Angewandte Mathematik und Mechanik)*, and a member of *GAMM Activity Group on Applied and Numerical Linear Algebra*.

Novi Sad, Serbia
March 1, 2010

Vladimir Kostić

Kratka biografija

Vladimir Kostić je diplomirao iz oblasti primenjene matematike 2003. godine na Departmanu za matematiku i informatiku Prirodno-matematičkog fakulteta Univerziteta u Novom Sadu, gde je i magistrirao 2009. godine.

Od 2004. godine radi kao asistent pripravnik kod Prof. dr Ljiljane Cvetković na predmetu *numeričke metode linearne algebre* na Departmanu za matematiku i informatiku i predmetima *matematika sa statistikom* i *programski paketi za obradu podataka* na Departmanu za biologiju i ekologiju.

Za svoja naučna i akademska ostvarenja primio je *1st Young Scientists Award* za izlaganje na naučnom skupu MAT-TRIAD 2007, Bedlewo, Poland, nagradu *najbolji student Univerziteta u Novom Sadu* 2004. godine, priznanje *Mileva Marić-Einstein* 2004. godine, *nagradnu stipendiju fonda Kraljevskog doma Karadjordjević* 2002/03. godine, nagradu *za visoka akademska dostignuća* od strane ambasade Kraljevine Norveške 2002. godine i *godišnja Univerzitetska priznanja* u 2001., 2002. i 2003. godini.

Objavio je 19 naučnih radova u visoko rangiranim medjunarodnim časopisima, koautor je 2 priručnika za studente i jedne monografije.

Tokom svog naučnog rada održao je jedno predavanje po pozivu u Banah Centru, Varšava, Poljska, 2009. godine, jedno plenarno predavanje na konferenciji MAT-TRIAD 2009, Bedlewo, Poljska i imao je 13 kratkih saopštenja u oblasti primenjene i numeričke linearne algebre.

Bio je sekretar Organizacionog Odbora dve medjunarodne konferencije.

Član je društva *ILAS (International Linear Algebra Society)*, društva *GAMM (Gesselshaft fur Angewandte Mathematik und Mechanik)* i aktivni je član radne grupe *GAMM Activity Group on Applied and Numerical Linear Algebra.*

Novi Sad, Srbija — Vladimir Kostić
1. mart 2010.

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCE
WORDS DOCUMENTATION

Accession number:
ANO
Identification number:
INO
Document type: Monograph type
DT
Type of record: Printed text
TR
Contents code: PhD thesis
CC
Author: Vladimir Kostić, MSc
AU
Mentor: Professor Ljiljana Cvetković, PhD
MN
Title: Benefits from the Generalized Diagonal Dominance
TI
Language of text: English
LT
Language of abstract: English/Serbian
LA
Country of publication: Serbia
CP
Locality of publication: Vojvodina
LP
Publication year: 2010
PY
Publisher: Authors reprint
PU
Publication place: Novi Sad, Faculty of Science, Dositeja Obradovića 4
PP
Physical description: 4/196/55/0/62/3/0
(chapters/pages/literature/tables/pictures/graphics/appendices)
PD
Scientific field: Mathematics
SF
Scientific discipline: Numerical Mathematics
SD
Subject / Key words: Applied Linear Algebra, Diagonal Dominance, H-matrices, Nonsingularity of matrices, Eigenvalues of matrices, Geršgorins theorem, Generalized Eigenvalues

SKW
Holding data: library of the Department of Mathematics and Informatics, Novi Sad
HD
Note:
N
Abstract: This theses is dedicated to the study of generalized diagonal dominance and its various benefits. The starting point is the well known nonsingularity result of strictly diagonally dominant matrices, from which generalizations were formed in different directions. In theses, after a short overview of very well known results, special attention was turned to contemporary contributions, where overview of already published original material is given, together with new obtained results. Particulary, Geršgorin-type localization theory for matrix pencils is developed, and application of the results in wireless sensor networks optimization problems is shown.
AB
Accepted by Scientific Board on:
ASB
Defended:
DE
Thesis defend board:
President: Zoran Stojaković, PhD, Full Professor, Faculty of Sciences, University of Novi Sad, Serbia
Member: Ljiljana Cvetković, PhD, Full Professor, Faculty of Sciences, University of Novi Sad, Serbia
Member: Richard S. Varga, PhD, Professor Emeritus, Department of Mathematical Sciences, Kent State University, Ohio, USA
DB

UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI FAKULTET
KLJUČNA DOKUMANTACIJSKA INFORMACIJA

Redni broj:
RBR
Identifikacioni broj:
IBR
Tip dokumentacije: Monografska dokumentacija
TD
Tip zapisa: Tekstualni štampani materijal
TZ
Vrsta rada: Doktorska teza
VR
Autor: mr Vladimir Kostić
AU
Mentor: Prof dr Ljiljana Cvetković
MN
Naslov rada: Prednosti generalizovane dijagonalne dominacije
NR
Jezik publikacije: engleski
JP
Jezik izvoda: engleski/srpski
JI
Zemlja publikovanja: Srbija
ZP
Uže geografsko područje: Vojvodina
UGP
Godina: 2010
GO
Izdavač: Autorski reprint
IZ
Mesto i adresa: Novi Sad, Prirodno-matematički fakultet, Trg Dositeja Obradovića 4
MA
Fizički opis rada: 4/196/55/0/62/3/0
(broj poglavlja/strana/lit. citata/tabela/slika/grafika/priloga)
FO
Naučna oblast: Matematika
NO
Naučna disciplina: Numerička matematika
ND

Predmetna odrednica/Ključne reči: Primenjena linearna algebra, dijagonalna dominacija, H-matrice, regularnost, karakteristični koreni matrica, Geršgorinova teorema, generalizovani karakteristični koreni
PO
UDK:
Čuva se: u biblioteci Departmana za matematiku i informatiku, Novi Sad
ČU
Važna napomena:
VN
Izvod: Ova teza je posvećena izučavanju generalizovane dijagonalne dominacije i njenih brojnih prednosti. Osnovu čini poznati rezultat o regularnosti strogo dijagonalnih matrica, čija su uopštenja formirana u brojnim pravcima. U tezi, nakon kratkog pregleda dobro poznatih rezultata, posebna pažnja je posvećena savremenim doprinosima, gde je dat i pregled već objavljenih autorovih rezultata, kao i detaljan tretman novih dobijenih rezultata. Posebno je razvijena teorija lokalizacije Geršgorinovog tipa generalizovanih karakterističnih korena i pokazana je primena rezultata u problemima optimizacije bežičnih senzor mreža.
IZ
Datum prihvatanja teme od strane NN Veća:
DP
Datum odbrane:
DO
Članovi komisije:
Predsednik: Prof dr Zoran Stojaković, redovni profesor, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
Član: Prof dr Ljiljana Cvetković, redovni profesor, Prirodno-matematički fakultet, Univerzitet u Novom Sadu
Član: Prof dr Richard S. Varga, profesor emeritus, Departman za Matematičke nauke, Državni Univerzitet u Kentu, Ohajo, SAD
KO