UNIVERZITET U NOVOM SADU
FAKULTET TEHNIČKIH NAUKA
U NOVOM SADU

Srđan Milićević

Algorithms for computing the optimal Geršgorin-type localizations

-doctoral dissertation-

Algoritmi za računanje optimalnih lokalizacija Geršgorinovog tipa

-doktorska disertacija-

Novi Sad, 2020.

УНИВЕРЗИТЕТ У НОВОМ САДУ ● **ФАКУЛТЕТ ТЕХНИЧКИХ НАУКА**
21000 НОВИ САД, Трг Доситеја Обрадовића 6

КЉУЧНА ДОКУМЕНТАЦИЈСКА ИНФОРМАЦИЈА

Редни број, **РБР**:		
Идентификациони број, **ИБР**:		
Тип документације, **ТД**:		Монографска документација
Тип записа, **ТЗ**:		Текстуални штампани материјал
Врста рада, **ВР**:		Докторска теза
Аутор, **АУ**:		Срђан Милићевић
Ментор, **МН**:		др Владимир Костић
Наслов рада, **НР**:		Алгоритми за рачунање оптималних локализација Гершгориновог типа
Језик публикације, **ЈП**:		Енглески
Језик извода, **ЈИ**:		Српски / Енглески
Земља публиковања, **ЗП**:		Србија
Уже географско подручје, **УГП**:		Војводина
Година, **ГО**:		2020.
Издавач, **ИЗ**:		Ауторски репринт
Место и адреса, **МА**:		Нови Сад, Факултет техничких наука, Трг Доситеја Обрадовића 6
Физички опис рада, **ФО**: (поглавља/страна/цитата/табела/слика/графика/прилога)		4/118/72/5/18/0/0
Научна област, **НО**:		Примењена математика
Научна дисциплина, **НД**:		Нумеричка математика
Предметна одредница/Кључне речи, **ПО**:		Примена линеарне алгебре, локализација карактеристичних корена, нумерички поступци, Гершгоринови кругови, минимални Гершгоринов скуп, нумерички распон, конвексни полигон
УДК		
Чува се, **ЧУ**:		У библиотеци Факултета техничких наука, Трг Доситеја Обрадовића 6, Нови Сад
Важна напомена, **ВН**:		
Извод, **ИЗ**:		Постоје бројни начини за локализацију карактеристичних корена. Један од најчувенијих резултата је да се спектар дате матрице $A \in C^{n,n}$ налази у скупу који представља унију кругова са центрима у дијагоналним елементима матрице и полупречницима који су једнаки суми модула вандијагоналних елемената одговарајуће врсте у матрици. Овај резултат (Гершгоринова теорема, 1931.), сматра се једним од најзначајнијих и најелегантнијих начина за локализацију карактеристичних корена ([61]). Међу свим локализацијама Гершгориновог типа, минимални Гершгоринов скуп даје најпрецизнију локализацију спектра ([39]). У овој дисертацији, приказани су нови алгоритми за одређивање тачне и поуздане апроксимације минималног Гершгориновог скупа.
Датум прихватања теме, **ДП**:		5.9.2019.
Датум одбране, **ДО**:		
Чланови комисије, **КО**:	Председник:	др Љиљана Цветковић, редовни професор
	Члан:	др Ксенија Дорословачки, ванредни професор
	Члан:	др Маја Недовић, доцент
	Члан:	др Зоран Томљановић, ванредни професор
	Члан, ментор:	др Владимир Костић, ванредни професор

Потпис ментора

UNIVERSITY OF NOVI SAD ● **FACULTY OF TECHNICAL SCIENCES**
21000 NOVI SAD, Trg Dositeja Obradovića 6

KEY WORDS DOCUMENTATION

Accession number, **ANO**:		
Identification number, **INO**:		
Document type, **DT**:		Monograph type
Type of record, **TR**:		Printed text
Contents code, **CC**:		PhD thesis
Author, **AU**:		Srđan Milićević
Mentor, **MN**:		Professor Vladimir Kostić, PhD
Title, **TI**:		Algorithms for computing the optimal Geršgorin-type localizations
Language of text, **LT**:		Serbian
Language of abstract, **LA**:		Serbian / English
Country of publication, **CP**:		Serbia
Locality of publication, **LP**:		Vojvodina
Publication year, **PY**:		2020
Publisher, **PB**:		Authors reprint
Publication place, **PP**:		Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6
Physical description, **PD**: (chapters/pages/ref./tables/pictures/graphs/appendixes)		4/118/72/5/18/0/0
Scientific field, **SF**:		Applied mathematics
Scientific discipline, **SD**:		Numerical mathematics
Subject/Key words, **S/KW**:		Applied linear algebra, localization of eigenvalues, numerical procedures, Geršgorin circles, minimal Geršgorin set, numerical range, convex polygon
UC		
Holding data, **HD**:		Library of the Faculty of Technical Sciences, Trg Dositeja Obradovića 6, Novi Sad
Note, **N**:		
Abstract, **AB**:		There are numerous ways to localize eigenvalues. One of the best known results is that the spectrum of a given matrix $A \in C^{n,n}$ is a subset of a union of discs centered at diagonal elements whose radii equal to the sum of the absolute values of the off-diagonal elements of a corresponding row in the matrix. This result (Geršgorin's theorem, 1931) is one of the most important and elegant ways of eigenvalues localization ([63]). Among all Geršgorin-type sets, the minimal Geršgorin set gives the sharpest and the most precise localization of the spectrum ([39]). In this thesis, new algorithms for computing an efficient and accurate approximation of the minimal Geršgorin set are presented.
Accepted by the Scientific Board on, **ASB**:		05.09.2019
Defended on, **DE**:		
Defended Board, **DB**:	President:	Ljiljana Cvetković, PhD, Full Professor
	Member:	Ksenija Doroslovački, PhD, Associate Professor
	Member:	Maja Nedović, PhD, Assistant Professor
	Member:	Zoran Tomljanović, PhD, Associate Professor
	Member, Mentor:	Vladimir Kostić, PhD, Associate Professor

Menthor's sign

Apstrakt

Jedna od najznačajnijih i najprimenljivijih oblasti matrične analize je proučavanje karakterističnih korena. Na osnovu njihovog položaja u kompleksnoj ravni, možemo da dobijemo korisne informacije o mnogim svojstvima matrice.

Termine „karakteristični koren“, „karakteristični vektor“ i „spektralna teorija“ uveo je Hilbert (u knjizi sa Kuranom publikovanoj 1924. godine). Reč „karakteristični“ vodi poreklo od nemačke reči „eigen“ što znači „odgovarajući“, „specifični“, „sopstveni“. Iako dati termini postaju standardni u literaturi početkom dvadesetog veka, upotreba karakterističnih korena datira još iz osamnaestog veka, prilikom rešavanja sistema diferencijalnih jednačina oblika $y' = Ay$. Potom su korišćeni i u radu Furijea kod rešavanja parcijalnih diferencijalnih jednačina i kasnije u radu mnogih drugih naučnika kao što su: Poason, Veber, Hilbert, Šmit, Nejman i mnogi drugi.

Karakteristični koreni imaju značajnu primenu u mnogim naučnim oblastima, a neke od njih su: akustika, ekologija, mehanika fluida, hemija, ekonomija, analiza vibracija, kvantna mehanika, obrada slika, lanci Markova, parcijalne diferencijalne jednačine, funkcionalna analiza, itd ([24], [30], [32], [52], [58], [61]).

Spektralna analiza se može koristiti u algoritamske i fizičke svrhe. U algoritmima, korišćenje karakterističnih korena može da pojednostavi i ubrza rešavanje problema na taj način što ih redukuje na skup skalarnih problema. Što se tiče upotrebe u fizičke svrhe, karakteristični koreni mogu da opišu oblike ponašanja evolucionih sistema određenih matricama, kao što su rezonanca, stabilnost i asimptotsko ponašanje.

Neki naučnici ističu i psihološki efekat koji daju karakteristčni koreni. Pomoću njih možemo da dobijemo vizuelizaciju matrice kao sliku u kompleksnoj ravni. „Karakteristični koreni daju osobenost matrici“ ([60]).

Postoje brojni načini za lokalizaciju karakterističnih korena. Jedan od najčuvenijih rezultata je da se spektar date matrice $A \in \mathbb{C}^{n,n}$ nalazi u skupu koji predstavlja uniju krugova sa centrima u dijagonalnim elementima matrice i poluprečnicima koji su jednaki sumi modula vandijagonalnih elemenata odgovarajuće vrste u matrici. Ovaj rezultat (Geršgorinova teorema, 1931.), smatra se jednim od najznačajnih i najelegantnijih načina za lokalizaciju karakterističnih korena ([63]). Među svim lokalizacijama Geršgorinovog tipa, minimalni Geršgorinov skup daje najprecizniju lokalizaciju spektra ([39]). U ovoj disertaciji, prikazani su novi algoritmi za određivanje precizne i pouzdane aproksimacije minimalnog Geršgorinovog skupa. Teza se sastoji iz četiri poglavlja.

U prvom poglavlju, prikazan je pregled poznatih rezultata u literaturi i motivacija za istraživanjem. Najpre su detaljno objašnjeni karakteristični koreni i njihove osobine. Potom sledi deo posvećen nenegativnim matricama i Peron-Frobenijusovoj teoriji. Zatim, opisane su neke klase matrica koje će biti kasnije korišćene u tezi. Takođe, predstavljeni su i rezultati o vezama između lokalizacija Geršgorinovog tipa sa odgovarajućim klasama matrica. Poseban akcenat je stavljen na minimalni Geršgorinov skup, njegove osobine i karakterizaciju. Na kraju poglavlja je deo posvećen numeričkom rasponu, njegovim osobinama i položaju u kompleksnoj ravni.

Glavni rezultati disertacije prikazani su u drugom poglavlju - novi algoritmi za računanje minimalnog Geršgorinovog skupa. Prvo je dat pregled do sada korišćenih algoritama i njihove karakteristike (**gMGS**, **bMGS** i **eMGSs**). Potom, prikazane su dve karakterizacije minimalnog Geršgorinovog skupa i tri pristupa obilaženja njegovog ruba. Prvi algoritam u disertaciji koji je novi rezultat je implicitni algoritam za računanje minimalnog Geršgorinovog skupa (**iMGSs**). On predstavlja poboljšanje eksplicitnog algoritma (**eMGSs**) koji je davao najbolje rezultate od svih poznatih algoritama ([40]). Brža konvergencija je dobijena korišćenjem rešavanja sistema linearnih jednačina umesto računanja karakterističnih korena. Takođe, prikazani su algoritmi koji koriste predictor-corrector metod (**eMGSp** i **iMGSp**) i algoritmi bazirani na korišćenju trougaone mreže (**eMGSt** i **iMGSt**). Navedeni algoritmi takođe predstavljaju originalne rezultate i dodatno su ubrzali izračunavanje i smanjili računsku složenost, posebno za matrice velikih formata.

U trećem poglavlju, govori se o odnosu između minimalnog Geršgorinovog skupa i numeričkog raspona. Predstavljeni su originalni rezultati za računanje apscise minimalnog Geršgorinovog skupa i konstrukciju konveksnog poligona koji

ga sadrži. Numerički postupak za određivanje konveksnog poligona je znatno brži i praktičniji nego algoritmi za računanje minimalnog Geršgorinovog skupa. U nekim slučajevima, aproksimacija konveksnim poligonom je veoma blizu granice minimalnog Geršgorinovog skupa, što je prikazano kroz primere.

Konačno, u poslednjem poglavlju teze, predstavljeni su numerički eksperimenti i implementacija. Novi algoritmi su testirani na brojnim primerima i rezultati su upoređeni sa do sada poznatim algoritmima. Sve prednosti i poboljšanja novih algoritama su na kraju sumirane u vidu kratkog zaključka.

Abstract

Research of eigenvalues is one of the most important and applicable areas of matrix analysis. If we know the position of eigenvalues of a given matrix in the complex plane, we can obtain useful information about many properties of that matrix.

The terms "eigenvalue", "eigenvector" and "spectral theory" were introduced by Hilbert (book by Hilbert and Courant published in 1924). The prefix "eigen" comes from the same German word which means "proper", "specific" or "characteristic". Although the words "eigenvalue" and "eigenvector" became standard in literature at the beginning of the 20^{th} century, their first known usage was in the 18^{th} century in solving differential equations of the form $y' = Ay$. Furthermore, they appeared in Fourier's work on partial differential equations and later, in the works of other mathematicians: Poisson, Weber, Hulbert, Schmidt, Neumann and many others.

Some of the fields where eigenvalues have very important roles are: acoustics, ecology, fluid mechanics, chemistry, economics, vibration analysis, quantum mechanics, image processing, Markov chains, partial differential equations, functional analysis, etc ([24], [30], [32], [52], [58], [61]).

Eigenvalues analysis can be used for algorithmic and physical purposes. In algorithms, the usage of eigenvalues can simplify solutions of some problems by reducing them to a collection of scalar problems. On the other hand, in physics, eigenvalues can be used to describe behavior of evolving system given by a system of linear equations. Some examples are studies of resonance, asymptotics and stability.

Also, some mathematicians emphasize a psychological usage of eigenvalues. They help us to perceive an abstraction of a matrix as a picture in the complex

plane. "Eigenvalues give a personality to a matrix" ([60]).

There are numerous ways to localize eigenvalues. One of the best known results is that the spectrum of a given matrix $A \in \mathbb{C}^{n,n}$ is a subset of a union of discs centered at diagonal elements whose radii equal to the sum of the absolute values of the off-diagonal elements of a corresponding row in the matrix. This result (Geršgorin's theorem, 1931) is one of the most important and elegant ways of eigenvalues localization ([63]). Among all Geršgorin-type sets, the minimal Geršgorin set gives the sharpest and the most precise localization of the spectrum ([39]). In this thesis, new algorithms for computing an efficient and accurate approximation of the minimal Geršgorin set are presented. The thesis consists of four chapters.

The introductory chapter presents an overview of the relevant results in the literature and the motivation for research. Firstly, eigenvalues and their properties are explained thoroughly. Then, a part dedicated to non-negative matrices and Perron-Frobenius theory follows. Moreover, classes of matrices that will be used later in the thesis are described. In addition, the results about localizations of Geršgorin-type sets and their relations with proper classes of matrices are given. Special emphasis is put on the minimal Geršgorin set, its properties and its characterizations. At the end of the chapter, there is a section about the numerical range, its properties and position in the complex plane.

The main results of the thesis - new algorithms for computing the minimal Geršgorin set are presented in the second chapter. First of all, a review of existing algorithms and their properties is given (**gMGS**, **bMGS** and **eMGSs**). Then, two characterizations of the minimal Geršgorin set and three approaches of curve tracing are presented. The first algorithm in the thesis which is a new result is an implicit algorithm for computing the minimal Geršgorin set (**iMGSs**). It represents an improvement of an explicit algorithm (**eMGSs**) which provided the best results among all previously known algorithms ([40]). A faster convergence is obtained by solving a system of linear equations instead of computing eigenvalues. In addition, algorithms which use predictor-corrector method (**eMGSp** and **iMGSp**) and algorithms based on triangular grid (**eMGSt** and **iMGSt**) are given. These algorithms also present new results and additionally accelerate computing and decrease computational complexity, especially for matrices of large size.

In the third chapter, the topic is a relation between the minimal Geršgorin set

and the numerical range. The original results for computing the abscissa of the minimal Geršgorin set and the construction of a convex polygon which contains the minimal Geršgorin set are given. Numerical algorithm for determination of the convex polygon is significantly faster and more practical than algorithms for computing the minimal Geršgorin set. In some cases, an approximation by the convex polygon is very close to the boundary of the minimal Geršgorin set which will be shown through various examples.

Finally, in the last part of thesis, some numerical experiments and implementations are shown. New algorithms are tested in several examples and the results are compared with the results of well-known algorithms. All advantages and improvements are summarized in a brief conclusion.

Acknowledgement

This thesis is the result of a multi-year research and mathematical learning and, in some way, it is the verification of work and effort during all my schooling. The thesis is dedicated to my family, friends, teachers, students and all the people who believed in me and gave me strength and support.

First of all, I would like to thank my mentor, Professor Vladimir Kostić, for the knowledge, ideas, energy and patience; for numerous readings of the thesis which significantly improved its quality. Furthermore, I need to mention Professor Ljiljana Cvetković who has had a lot of influence on my academic work in general. Nevertheless, I would like to thank her for the confidence and honor to be a member of her research group SC:ALA. Moreover, I want to thank my first advisor, Professor Zagorka Lozanov-Crvenković, for introducing me to the world of Ph.D. studies. Additionally, I am thankful to my defense committee members: Professor Ksenija Doroslovački, Professor Maja Nedović and Professor Zoran Tomljanović for their effectiveness, constructive comments and suggestions.

I am very grateful to Professors Mila Stojaković, the head of Chair of Mathematics at the Faculty of Technical Sciences, and Ljiljana Teofanov, the head of Doctoral Studies of Mathematics in Engineering, for their understanding, useful tips and wholehearted conversations.

I am thankful to my dear colleagues: Professor Zorica Uzelac, Professor Ilija Kovačević, Maria, Marija, Jovana, Buda, Dunja, Gabrijela, Lela, Miroslav, Đole and many others for their collaboration, unselfish help and positive energy. I thank them for the great atmosphere which made my job easier.

I am extending my thanks to Ivana, Aca and Jovan for motivation, friendship and technical support during writing the thesis. I thank all my other friends for being a part of my life!

Finally, I would like to express the deepest gratitude to my parents Ljubisav and Milena, to my sister Marina and fiancé Radosna for endless love, sincere trust and unconditional support.

February, 2020

Notation

$\mathbb{N}$ - the set of positive integers

$\mathbb{N}_0$ - the set of non-negative integers ($\mathbb{N}_0 = \mathbb{N} \cup \{0\}$)

$\mathbb{R}$ - the set of real numbers

$\mathbb{C}$ - the set of complex numbers

i - the imaginary unit

$N := \{1, 2, ..., n\}$ - the set of integers from 1 to $n,\ n \in \mathbb{N}$

$\text{dist}(x, S) = \min\limits_{s \in S} |x - s|$ - the distance from the point x to the set S

$\text{card}(S)$ - the cardinality of the set S

$\text{co}(S)$ - the convex hull of the set S

$\text{int}(S)$ - the interior of the set S

∂S - the boundary of the set S

$v = [v_1, v_2, ..., v_n]^T \in \mathbb{C}^{n,1}$ - a column vector, $n \in \mathbb{N}$

$\sphericalangle(\vec{u}, \vec{v})$ - the angle between the vectors $\vec{u}$ and $\vec{v}$

For a square matrix $A \in \mathbb{C}^{n,n},\ n \in \mathbb{N}$:

$A = [a_{ij}],\ a_{ij}$ - entries of $A,\ i, j \in N$

$\text{diag}(A)$ - a vector of diagonal elements of A

A^T - the transpose of A

A^* - the conjugate transpose of A

A^{-1} - the inverse of A

A^+ - the Moore-Penrose inverse (pseudoinverse) of A

$\det(A)$ - the determinant of A

$\mathrm{tr}(A)$ - the trace of A

$||\cdot||_p$ - the p-norm, $1 \leq p \leq \infty$

$\sigma(A)$ - the spectrum of A

$\rho(A)$ - the spectral radius of A

$\alpha(A)$ - the spectral abscissa of A

$\Gamma_i(A)$ - the i^{th}-Geršgorin disk of A, $i \in N$

$\Gamma(A)$ - the Geršgorin set of A

$\mathcal{K}(A)$ - the Brauer set of A

$\Gamma^{\mathcal{R}}(A)$ - the minimal Geršgorin set of A

$\gamma(A)$ - the abscissa of the Geršgorin set of A

$\mu(A)$ - the abscissa of the minimal Geršgorin set of A

$W(A)$ - the numerical range of A

$\omega(A)$ - the abscissa of the numerical range of A

gMGS(A) - the griding algorithm for computing the minimal Geršgorin set of A

bMGS(A) - the bisection algorithm for computing the minimal Geršgorin set of A

eMGSs(A) - the explicit star-shaped algorithm for computing the minimal Geršgorin set of A

iMGSs(A) - the implicit star-shaped algorithm for computing the minimal Geršgorin set of A

eMGSp(A) - the explicit predictor-corrector algorithm for computing the minimal Geršgorin set of A

iMGSp(A) - the implicit predictor-corrector algorithm for computing the minimal Geršgorin set of A

eMGSt(A) - the explicit triangular algorithm for computing the minimal Geršgorin set of A

iMGSt(A) - the implicit triangular algorithm for computing the minimal Geršgorin set of A

For $f = f^{\xi,\theta}(t) = f(\xi + te^{\mathrm{i}\theta}), \xi \in \mathbb{C},\ \theta \in [0, 2\pi),\ t \in \mathbb{R}$:

$\dfrac{\partial}{\partial t} f^{\xi,\theta}(t)$ - the first order derivative of f

$\dfrac{\partial^2}{\partial t^2} f^{\xi,\theta}(t)$ - the second order derivative of f

For $f = f(x, y) = f(x + \mathrm{i}y),\ x, y \in \mathbb{R}$:

$f_x = \dfrac{\partial}{\partial x} f(x, y),\ f_y = \dfrac{\partial}{\partial y} f(x, y)$ - the first order derivatives of f

$f_{xx} = \dfrac{\partial^2}{\partial x^2} f(x, y),\ f_{xy} = \dfrac{\partial^2}{\partial x \partial y} f(x, y),\ f_{yy} = \dfrac{\partial^2}{\partial y^2} f(x, y)$ - the second order derivatives of f

Contents

Chapter 1

Introduction and preliminaries

"Eigenvalues are among the most successful tools of applied mathematics."

Lloyd Trefethen[1]

In this chapter, we summarize well-known theoretical results that represent the bases of this thesis. First, in Section 1.1, a motivation for research is given. Next, in Sections 1.2 and 1.3, the basic terms in the theory of eigenvalues, norms and non-negative matrices are introduced. Then, in Sections 1.4 and 1.5, the results about Geršgorin sets are given with a special emphasis on the minimal Geršgorin set. Finally, Section 1.5 discuses the relationships between nonsingularity of a given matrix and localization of its spectrum while in Section 1.6, the results about the numerical range are given.

1.1 Main motivations

Since matrices occur in problem-solving processes in engineering and many scientific disciplines, localization of their eigenvalues represents a powerful tool in solving those problems. Therefore, approaches to localizations of eigenvalues

[1]Lloyd Nicholas Trefethen (1955) is an American mathematician, professor of numerical analysis and head of the Numerical Analysis Group at the Mathematical Institute, University of Oxford.

occupy a central place in numerical linear algebra, and numerous such results can be applied successfully in many branches of science.

For a given square matrix $A = [a_{ij}] \in \mathbb{C}^{n,n},\ n \in \mathbb{N}$, one of the best known results for localization of its spectrum is Geršgorin's[2] circle theorem. This result, published in 1931, presents how to simply localize eigenvalues of a given matrix by the Geršgorin set which represents the union of n discs in the complex plane whose centers are a_{ii} and radii $\sum_{i \neq j}^{n} |a_{ij}|,\ i \in N := \{1, 2, ..., n\}$. Among many eigenvalue localizations that were developed since then, the minimal Geršgorin set (MGS) plays permanent role. It gives, in a certain sense, the sharpest inclusion set for eigenvalues of A, with respect to all positive diagonal similarity transformations ([39]). So, it represents a kind of "optimal" localization for the spectrum of a given matrix. Although modern computers can successfully compute eigenvalues, this still presents a challenge for matrices of large size. Therefore, the computation of MGS can provide useful information about the position of the spectrum of large matrices in the complex plane.

Next, we elaborate more on "optimality" of the minimal Geršgorin set. First, as it was shown in [63], every point of the minimal Geršgorin set of an arbitrary matrix A is, in fact, an eigenvalue of some matrix that has the same diagonal entries as the given one, while its off-diagonal entries are bounded in moduli by the corresponding off-diagonal entries of A. So, the minimal Geršgorin set is the union of all the spectra of all matrices bounded by the given one in the sense explained above.

On the other hand, the minimal Geršgorin set is also "optimal" in terms of diagonally dominant matrices. Namely, as it is well established, there is an equivalence between the eigenvalue localization results and the matrix nonsingularity results, see [63]. In that sense, the minimal Geršgorin set corresponds to the class of H-matrices, while the Geršgorin set corresponds to the strictly diagonally dominant matrices. Inspired by this, the concept of Geršgorin-type sets was developed in [39], where it was shown that the minimal Geršgorin set is the minimal element among all Geršgorin-type sets.

The research on the minimal Geršgorin set provided many interesting theoretical results, while the algorithms for its computation remained less well developed

[2]Semen Aranovich Geršgorin (1901-1933) was a Soviet mathematician, who researched in partial differential equations and localizations of eigenvalues.

in the literature. Unlike the Geršgorin set, it is not easy to numerically determine MGS. The main setback often lies in the complexity of its computation ([63], [67]) because MGS is defined as an intersection of infinitely many sets. Luckily, as we will see shortly, it can be characterized via appropriate eigenvalue problems.

In [40], to decrease numerical cost of previous known algorithms (**gMGS**, **bMGS** of [69]), the authors developed the algorithm **eMGSs**. Although **eMGSs** performed well on tests, its application was limited to matrices of small and medium size.

The main results of this thesis are new algorithms for computing the minimal Geršgorin set that can successfully be applied to large matrices. We combine two approaches for the characterization of MGS (explicit and implicit) and three methods for curve tracing (star-shaped, predictor-corrector and triangular). In that way, six efficient procedures are developed: **eMGSs, iMGSs, eMGSp, iMGSp, eMGSt** and **iMGSt**.

First new algorithm presented in this thesis is **iMGSs**. It represents an extension and improvement of **eMGSs**. It is developed by using solving of the system of linear equations instead of computation of eigenvalues, the idea adopted from [25]. Next, we constructed the algorithms **eMGSp** and **iMGSp**. These algorithms are based on two steps: predictor (find approximate point in the direction of tracing) and corrector (find point on curve using predicted point and numerical methods). The motivation for the predictor-corrector method is found in [2]. Then, the algorithms **eMGSt** and **iMGSt** are constructed. These algorithms are based on path following approach via triangular grid and in that way, they reduce numerical computations and can be adapted for matrices of large size. This method is originally developed in [47], for the computations of the pseudospectra.

Benefits of MGS are not only based on the fact that it represents optimal localization of eigenvalues. Beside theoretical, MGS also has practical importance.

Let us consider time-dependent linear dynamical system which appears in models which describe oscillatory systems:

$$\dot{x}(t) = A(t)x(t),\ t \geq 0, \tag{1.1}$$

where matrix $A(t)$ has time invariant diagonal elements, i.e., $a_{ii}(t) = \alpha_{ii} \in \mathbb{C},\ i \in N$, while other elements $a_{ij}(t)$ are complex analytical functions bounded by $\alpha_{ij} > 0,\ i, j \in N,\ i \neq j$. Using the infinity norm of matrix and Coppel inequality ([60]), we obtain that system (1.1) is exponentially asymptotically stable if there exists $\mu > 0$ such that

$$\max_{i \in N} \left\{ \mathrm{Re}(\alpha_{ii}) + \sum_{j \neq i}^{n} \alpha_{ij} \right\} < -\mu,$$

i.e., the Geršgorin set of $[\alpha_{ij}]_{n \times n}$ lies in the open left half-plane of the complex plane.
Analogously, if we use a vector norm $||X(\cdot)||_\infty$, where X is any diagonal matrix with positive diagonal entries, we have that system (1.1) is exponentially asymptotically stable if the minimal Geršgorin set is situated in the open left half-plane of the complex plane, which represents a stronger result.

One more interesting item for determining the minimal Geršgorin set is its relation with the numerical range. For a given matrix $A \in \mathbb{C}^{n,n}$, its numerical range $W(A)$ is a subset of the convex envelope of union of the Geršgorin set of A and the Geršgorin set of A^T (Theorem 3.2.2). Moreover, it is easy to show that set

$$\bigcap_{X \in \mathbb{D}_n} W(X^{-1}AX)$$

is a subset of the the convex envelope of the minimal Geršgorin set of A, where $\mathbb{D}_n$ is the set of all diagonal matrices whose diagonal entries are positive (Corollary 3.2.3). Therefore, using the position of the convex envelope of the minimal Geršgorin set in the complex plane, we can get useful information about the numerical range after a proper scaling.

As we can see, the research of the minimal Geršgorin set provided many useful results. They served as a motivation for the construction of new algorithms for its computing.

1.2 Eigenvalues and norms

Let $A \in \mathbb{C}^{n,n},\ n \in \mathbb{N}$, be a square matrix. A vector $x \in \mathbb{C}^n,\ x \neq 0$, is an eigenvector of A and $\lambda \in \mathbb{C}$ is its corresponding eigenvalue if

$$Ax = \lambda x. \tag{1.2}$$

The equation (1.2) can be equivalently written as

$$(\lambda I - A)x = 0, \tag{1.3}$$

where I is the $n \times n$ identity matrix. The equation (1.3) has a non-zero solution x if and only if a determinant of matrix $\lambda I - A$ is equal to zero, i.e.,

$$\det(\lambda I - A) = 0. \tag{1.4}$$

The left side of (1.4) is a polynomial function of λ denoted by $p_A(\lambda)$. It is called the characteristic polynomial of A and the equation (1.4) is the characteristic equation of A.

Point $\lambda \in \mathbb{C}$ is an eigenvalue of $A \in \mathbb{C}^{n,n}$ if and only if $p_A(\lambda) = 0$. We can notice that even if A is a real matrix, its eigenvalues can be complex numbers.

The set of all eigenvalues of a matrix A is the spectrum of A, denoted by $\sigma(A)$, i.e.,

$$\sigma(A) := \{\lambda \in \mathbb{C} : \det(\lambda I - A) = 0\}. \tag{1.5}$$

An equivalent definition for the spectrum is that it is the set of points $z \in \mathbb{C}$ for which a resolvent matrix $(zI - A)^{-1}$ does not exist.

The spectral radius $\rho(A)$ of $A \in \mathbb{C}^{n,n}$ is defined by

$$\rho(A) := \{\max |\lambda| : \lambda \in \sigma(A)\}. \tag{1.6}$$

The spectral abscissa $\alpha(A)$ of $A \in \mathbb{C}^{n,n}$ is defined by

$$\alpha(A) := \{\max(\mathrm{Re}(\lambda)) : \lambda \in \sigma(A)\}. \tag{1.7}$$

Let λ be an eigenvalue of $A \in \mathbb{C}^{n,n}$. The algebraic multiplicity of λ is its multiplicity as a root of the characteristic polynomial of A. If the algebraic multiplicity of λ is equal to 1, λ is called a simple eigenvalue.

The geometric multiplicity of λ is a dimension of its corresponding eigenspace $E_\lambda = \{v \in \mathbb{C}^n : (\lambda I - A)v = 0\}$.

The algebraic multiplicity of eigenvalue is at least as great as its geometric multiplicity. If λ is an eigenvalue of $A \in \mathbb{C}^{n,n}$ whose algebraic multiplicity is greater

than its geometric multiplicity, then λ is a defective eigenvalue. A matrix with at least one defective eigenvalue is a defective matrix. Otherwise, it is a non-defective matrix.

In the following, we review well-known properties of eigenvalues that will be used in this thesis.

- Matrix $A \in \mathbb{C}^{n,n}$ has n eigenvalues, counted with multiplicity.
- Matrix $A \in \mathbb{C}^{n,n}$ is nonsingular if and only if its every eigenvalue is non-zero.
- The trace $\mathrm{tr}(A)$ of matrix $A \in \mathbb{C}^{n,n}$ (the sum of its diagonal elements) is equal to the sum of its all eigenvalues, i.e., $\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$.
- The determinant of $A \in \mathbb{C}^{n,n}$ is equal to the product of its eigenvalues, i.e., $\det(A) = \prod_{i=1}^{n} \lambda_i = \lambda_1 \lambda_2 \cdots \lambda_n$.
- Eigenvalues of a diagonal matrix are its diagonal elements.
- Matrix $A \in \mathbb{C}^{n,n}$ and its transpose $A^T \in \mathbb{C}^{n,n}$ have the same spectrum.
- Matrices $A \in \mathbb{C}^{n,n}$ and $XAX^{-1} \in \mathbb{C}^{n,n},\ X \in \mathbb{C}^{n,n},\ \det(\mathrm{X}) \neq 0$ (XAX^{-1} is a similar matrix to A), have the same spectrum.
- If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of the matrix $A \in \mathbb{C}^{n,n}$, then $\lambda_1^k, \lambda_2^k, ...,$ λ_n^k are the eigenvalues of the matrix $A^k, k \in \mathbb{N}$.
- If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of the nonsingular matrix $A \in \mathbb{C}^{n,n}$, then $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, ..., \frac{1}{\lambda_n}$ are the eigenvalues of its inverse matrix $A^{-1} \in \mathbb{C}^{n,n}$ and each algebraic and geometric multiplicity of the corresponding eigenvalues are identical.
- If $A \in \mathbb{C}^{n,n}$ is a Hermitian matrix (i.e., $A = A^*$), then each its eigenvalue is real. So, in a special case, if A is a symmetric real matrix (i.e., $A = A^T \in \mathbb{R}^{n,n}$), its eigenvalues must be real.
- If $A \in \mathbb{R}^{n,n}$ and n is odd, then A has at least one real eigenvalue.

- If $\lambda_1, \lambda_2, ..., \lambda_n$ are the eigenvalues of a unitary matrix $A \in \mathbb{C}^{n,n}$ (i.e., $AA^* = A^*A = I$), then $|\lambda_j| = 1, j \in N$.

Term "eigenvector" is typically used for "right eigenvector". It is a vector column that satisfies the equation (1.2). However, we can also observe the problem

$$u^T A = \tilde{\lambda} u^T, \tag{1.8}$$

where $A \in \mathbb{C}^{n,n}$, $\tilde{\lambda} \in \mathbb{C}$ is scalar and $u \in \mathbb{C}^{n,1}$. Any non-zero vector u satisfying (1.8) is a left eigenvector of matrix A and $\tilde{\lambda}$ is its corresponding eigenvalue. Taking the transpose of the equation (1.8), we get $A^T u = \tilde{\lambda} u$. Therefore, any left eigenvector of A is the right eigenvector of A^T.

Using eigenvalues, we can write some matrices in the form of a factorization. An eigenvalue decomposition of matrix $A \in \mathbb{C}^{n,n}$ is the factorization

$$A = X \Lambda X^{-1}, \tag{1.9}$$

where $X \in \mathbb{C}^{n,n}$ is nonsingular and $\Lambda \in \mathbb{C}^{n,n}$ is a diagonal matrix. As we see, the eigenvalue decomposition is a similarity transformation. If A has the eigenvalue decomposition, we can say that it is a diagonalizable matrix. From the definition of the eigenvalue decomposition, we get $AX = X\Lambda$, i.e.,

$$A \begin{bmatrix} x_1 | x_2 | \cdots | x_n \end{bmatrix} = \begin{bmatrix} x_1 | x_2 | \cdots | x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

So, $Ax_j = \lambda_j x_j, j \in N$, where x_j is j^{th} column of X and the eigenvector of A and λ_j is j^{th} diagonal entry of Λ and the eigenvalue of A corresponding to the eigenvector x_j.

Matrix $A \in \mathbb{C}^{n,n}$ has a complete set of eigenvectors $X = \{x_1, x_2, ..., x_n\}$ if the vectors $x_1, x_2, ..., x_n$ are linearly independent and $Ax_j = \lambda_j x_j, j \in N$. So, each matrix with a complete set of eigenvectors has the eigenvalue decomposition (i.e., A is a diagonalizable matrix). If this decomposition can be performed in such a way that diagonalization is orthonormal one, i.e., $A = Q\Lambda Q^*$, where Q is unitary, the matrix is called normal matrix. It is well-known that matrix A is normal if and only if $A^*A = AA^*$. It can be shown that a matrix $A \in \mathbb{C}^{n,n}$ is a non-defective matrix if and only if it has the eigenvalue decomposition.

In applications, eigenvalues are often used for computing powers $A^k, k \in \mathbb{N}$, or exponential $e^{tA} = I + tA + \frac{1}{2!}(tA)^2 + ...$, of a given matrix $A \in \mathbb{C}^{n,n}$. If A is a diagonalizable matrix, then we can compute A^k and e^{tA} as: $A^k = (V\Lambda V^{-1})^k = V\Lambda^k V^{-1}$ and $e^{tA} = Ve^{t\Lambda}V^{-1}$.

If $A \in \mathbb{C}^{n,n}$ is singular, we can use Moore[3]-Penrose[4] inverse of it, denoted by A^+. It is the unique matrix $X \in \mathbb{C}^{n,n}$ satisfying the following Penrose equations:

$$(i)\ AXA = A,\ (ii)\ XAX = X,\ (iii)\ (AX)^* = AX,\ (iv)\ (XA)^* = XA.$$

A purpose of the next theorem is to provide explicit formulas for the derivatives of eigenvalues and eigenvectors. These formulas are useful in the analysis of systems of dynamic equations and have many other applications.

Theorem 1.2.1 *(Theorem 2, [45]) Let λ_0 be a simple eigenvalue of a matrix $A_0 \in \mathbb{C}^{n,n}$, and let v_0 be an associated eigenvector, so that $A_0 v_0 = \lambda_0 v_0$. Then a (complex) function λ and a (complex) vector function v are defined for all A in some neighborhood $\mathcal{O}(A_0) \in \mathbb{C}^{n,n}$ of A_0, such that*

$$\lambda(A_0) = \lambda_0,\ v(A_0) = v_0$$

and

$$Av = \lambda v,\ v_0^* v = 1,\ A \in \mathcal{O}(A_0).$$

Moreover, the functions λ and v are smooth on $\mathcal{O}(A_0)$ and the differentials at A_0 are

$$d\lambda = u_0^*(dA)v_0/u_0^* v_0 \tag{1.10}$$

and

$$dv = (\lambda_0 I - A_0)^+(I - v_0 u_0^*/u_0^* v_0)(dA)v_0, \tag{1.11}$$

where u_0 is the left eigenvector of A_0 associated with the eigenvalue λ_0.

---

[3]Eliakim Hastings Moore (1862-1932) was an American mathematician, who researched in abstract algebra, geometry, number theory and integral equations.

[4]Sir Roger Penrose (1931) is an English mathematical physicist, mathematician and philosopher of science, Emeritus Rouse Ball Professor of Mathematics in the University of Oxford.

Since we will repeatedly work on construction of mappings with a complex argument $f : \mathbb{C} \to \mathbb{C}^{m,n}$, where $m, n \in \mathbb{N}$, without possible confusion to simplify notations, we will use abbreviations:

$$f = f(z), \text{ for } z \in \mathbb{C};$$

$$f^{\xi,\theta}(t) = f(\xi + te^{\mathrm{i}\theta}), \text{ for } \xi \in \mathbb{C},\ \theta \in [0, 2\pi),\ t \in \mathbb{R};$$

$$f(x, y) = f(x + \mathrm{i}y), \text{ for } x, y \in \mathbb{R}.$$

In that context, derivatives in the corresponding arguments are denoted as:

$$\frac{\partial}{\partial t} f^{\xi,\theta}(t),\ f_x = \frac{\partial}{\partial x} f(x, y),\ f_y = \frac{\partial}{\partial y} f(x, y);$$

$$\frac{\partial^2}{\partial t^2} f^{\xi,\theta}(t),\ f_{xx} = \frac{\partial^2}{\partial x^2} f(x, y),\ f_{xy} = \frac{\partial^2}{\partial x \partial y} f(x, y),\ f_{yy} = \frac{\partial^2}{\partial y^2} f(x, y).$$

Next, some preliminaries about norms will be presented. A standard inner product of vectors $x = [x_1, x_2, ..., x_n]^T \in \mathbb{C}^n$ and $y = [y_1, y_2, ..., y_n]^T \in \mathbb{C}^n$ is

$$y^* x := \sum_{i=1}^{n} x_i \overline{y}_i.$$

A norm is a function $|| \cdot || : \mathbb{C}^n \to \mathbb{R},\ n \in \mathbb{N}$, which satisfies the following conditions:

1) $||x|| \geq 0$,

2) $||x|| = 0$ if and only if $x = 0$,

3) $||x + y|| \leq ||x|| + ||y||$,

4) $||\alpha x|| = |\alpha| ||x||$,

for all vectors $x, y \in \mathbb{C}^n$ and scalars $\alpha \in \mathbb{C}$.
Some of the most important norms are the p-norms:

$$||x||_1 = \sum_{i=1}^{n} |x_i|,$$

$$||x||_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}} = \sqrt{x^* x},$$

$$||x||_\infty = \max_{i \in N} |x_i|,$$

$$||x||_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}},\ 1 \leq p < \infty,$$

for $x = [x_1, x_2, ..., x_n]^T \in \mathbb{C}^n$.

Beside p-norms, widely used are the weighted norms. For any norm $||\cdot||$ and nonsingular matrix $W \in \mathbb{C}^{n,n}$, the weighted norm of vector $x \in \mathbb{C}^n$ is:

$$||x||_W = ||Wx||,$$

where $W \in \mathbb{C}^{n,n}$ is a nonsingular matrix.

For a given norm $||\cdot||$ on $\mathbb{C}^n$, the induced matrix norm of A is defined by:

$$||A|| := \sup_{x \in \mathbb{C}^n \setminus \{0\}} \frac{||Ax||}{||x||} = \sup_{||x||=1, x \in \mathbb{C}^n} ||Ax||,\ A \in \mathbb{C}^{n,n}.$$

The associated induced matrix norms of a given matrix $A \in \mathbb{C}^{n,n}$ for norms $||\cdot||_1$, $||\cdot||_2$ and $||\cdot||_\infty$ are:

$$||A||_1 = \max_{j \in N} \sum_{i=1}^{n} |a_{ij}|,$$

$$||A||_2 = \sqrt{\rho(AA^*)},$$

$$||A||_\infty = \max_{i \in N} \sum_{j=1}^{n} |a_{ij}|.$$

The associated induced matrix norm of a given matrix $A \in \mathbb{C}^{n,n}$ for the weighted norm $||\cdot||_W$ is

$$||A||_W = ||WAW^{-1}||,$$

where $W \in \mathbb{C}^{n,n}$ is a nonsingular matrix.

For any $A \in \mathbb{C}^{n,n}$ and any norm $||\cdot||$ on $\mathbb{C}^{n,n}$, holds:

$$\rho(A) \leq ||A||.$$

Vectors $x = [x_1, x_2, ..., x_n]^T \in \mathbb{C}^n$ and $y = [y_1, y_2, ..., y_n]^T \in \mathbb{C}^n$ are orthogonal if $x^*y = 0$. Additionally, x and y are orthonormal if they are orthogonal and $||x||_2 = ||y||_2 = 1$.

1.3 Perron-Frobenious theory

In linear algebra, the Perron[5]-Frobenius[6] theorem asserts that the largest, by moduli, eigenvalue of a real non-negative square matrix is real and its corresponding eigenvector can be chosen to have all non-negative components. Also, this theorem has important applications to probability theory, to the theory of dynamical systems, to economics, to demography, etc.

In the following, we will use the next definitions.

- A matrix $A = [a_{ij}] \in \mathbb{R}^{n.n}$ is positive if $a_{ij} > 0,\ i, j \in N$, and we write it as $A > 0$.
- A matrix $A = [a_{ij}] \in \mathbb{R}^{n.n}$ is non-negative if $a_{ij} \geq 0,\ i, j \in N$, and we write it as $A \geq 0$.
- A matrix $A = [a_{ij}] \in \mathbb{R}^{n.n}$ is essentially non-negative if $a_{ij} \geq 0,\ i, j \in N,\ i \neq j$.
- A matrix $A = [a_{ij}] \in \mathbb{R}^{n.n}$ is a Z-matrix if $a_{ij} \leq 0,\ i, j \in N,\ i \neq j$.

First, Perron formulated theorem for non-negative matrices in 1905. Inspired by his work, Frobenius proved extended version of the theorem (introduced irreducibility) in 1912.

Theorem 1.3.1 *(Perron-Frobenius) If $A = [a_{ij}] \in \mathbb{R}^{n.n}$ and $A \geq 0$, then:*

1. *A has a non-negative real eigenvalue equal to its spectral radius $\rho(A)$;*
2. *there corresponds an eigenvector $x = [x_1, x_2, ..., x_n]^T \geq 0,\ x \neq 0$, to $\rho(A)$;*

[5]Oscar Perron (1880-1975) was a German mathematician, professor at the University of Heidelberg and University of Munich.

[6]Ferdinand Georg Frobenius (1849-1917) was a German mathematician, best known for his contributions to differential equations, number theory and group theory.

3. *$\rho(A)$ may be a multiple eigenvalue of A;*

4. *$\rho(A)$ does not decrease when any entry of A increases;*

5. *the eigenvalue $\rho(A)$ satisfies*

$$\rho(A) = \inf_{x>0,\ x\in\mathbb{R}^{n,n}} \left\{ \max_{i\in N} \left[\frac{\sum_{j=1}^{n} a_{ij}x_j}{x_i} \right] \right\}.$$

A matrix $P \in \mathbb{R}^{n,n}$ is a permutation matrix if there is a permutation $\pi : N \to N$ such that $P = [p_{ij}] = [\delta_{i,\pi(j)}]$, where

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

(Kronecker delta function), $k, l \in N$.

Definition 1.3.2 *If $A \in \mathbb{C}^{n,n},\ n \geq 2$, it is reducible if there exists a permutation matrix $P \in \mathbb{R}^{n,n}$ and $r \in \mathbb{N},\ 1 \leq r < n$, for which*

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where $A_{11} \in \mathbb{C}^{r,r}$ and $A_{22} \in \mathbb{C}^{(n-r),(n-r)}$.
A matrix $A \in \mathbb{C}^{1,1}$ is reducible if it is zero matrix.

If a matrix is not reducible, it is irreducible.

For a given matrix $A \in \mathbb{C}^{n,n}$ and set of vertices $\{v_1, v_2, ..., v_n\}$, define directed arc $\overrightarrow{v_i v_j}$ from vertex v_i to vertex v_j if $a_{ij} \neq 0,\ i, j \in N$. The collection of all such directed arcs is the direct graph $G(A)$ of A. A directed path in $G(A)$ from vertex v_i to vertex v_j is a collection of directed arcs connecting v_i as an initial and v_j as a terminal vertex.

The directed graph $G(A)$ of a given matrix $A \in \mathbb{C}^{n,n}$ is strongly connected if for each ordered pair v_i and v_j of vertices there exists a direct path in $G(A)$ with an initial vertex v_i and a terminal vertex $v_j,\ i, j \in N$.

Theorem 1.3.3 *([63], Theorem 1.9.) A matrix $A \in \mathbb{C}^{n,n}$ is irreducible if and only if its direct graph $G(A)$ is strongly connected.*

For a given arbitrary matrix $A \in \mathbb{C}^{n,n}, n \geq 2$, its normal reduced form is

$$PAP^T = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ & A_{22} & \dots & A_{2m} \\ & & \ddots & \vdots \\ & & & A_{mm} \end{bmatrix}, \tag{1.12}$$

where P is a permutation matrix and diagonal blocks $A_{ii} \in \mathbb{C}^{n_i,n_i}$ are either 1×1 or irreducible $n_i \times n_i$ matrices, $n_i \geq 2, i \in \{1, 2, ..., m\}$.

Theorem 1.3.4 *(Perron-Frobenius, irreducible matrix) If $A = [a_{ij}] \in \mathbb{R}^{n.n}$ is irreducible and essentially non-negative, then:*

1. *A has a real eigenvalue equal to its spectral abscissa $\alpha(A)$ with the corresponding right $x = [x_1, x_2, ..., x_n]^T > 0$ and left eigenvector $y = [y_1, y_2, ..., y_n] > 0$;*

2. *the only right (left) eigenvectors whose components are all positive are multiples of x (y);*

3. *$\alpha(A)$ is a simple eigenvalue of A;*

4. *$\alpha(A)$ increases when any entry of A increases;*

5. $\min\limits_{i \in N} \sum\limits_{j=1}^{n} a_{ij} \leq \alpha(A) \leq \max\limits_{i \in N} \sum\limits_{j=1}^{n} a_{ij}$;

6. *(Collatz[7]-Wielandt[8] formula) the eigenvalue $\alpha(A)$ satisfies*

$$\alpha(A) = \sup_{x>0,\ x \in \mathbb{R}^{n,n}} \left\{ \min_{i \in N} \left[\frac{\sum\limits_{j=1}^{n} a_{ij} x_j}{x_i} \right] \right\} = \inf_{x>0,\ x \in \mathbb{R}^{n,n}} \left\{ \max_{i \in N} \left[\frac{\sum\limits_{j=1}^{n} a_{ij} x_j}{x_i} \right] \right\}.$$

[7]Lothar Collatz (1910-1990) was a German mathematician, founder of spectral graph theory.

[8]Helmut Wielandt (1910-2001) was a German mathematician who worked on permutation groups.

Originally, this theorem is given in the literature for irreducible non-negative matrices.

The spectral abscissa $\alpha(A)$ of an irreducible and essentially non-negative matrix $A \in \mathbb{R}^{n,n}$ is called the Perron root or the Perron-Frobenius root. Its corresponding right eigenvector x is the Perron eigenvector or the Perron-Frobenius eigenvector and the pair $(\alpha(A), x)$ is the Perron pair or the Perron-Frobenius pair.

The algorithm **Noda iteration** for computing the spectral radius of non-negative irreducible matrices is presented in [50]. We use it to formulate an algorithm for computing the spectral abscissa of essentially non-negative matrix.

Input parameters of the algorithm are: an irreducible and essentially non-negative matrix $A \in \mathbb{R}^{n,n}$, a scalar $\lambda_0 \geq \alpha(A)$ and positive normalized vector $x^0 \in \mathbb{R}^n$ (i.e., $x^0 > 0$, $||x^0||_2 = 1$). For example, for parameter λ_0, we can choose the value $\max\limits_{i \in N} \sum\limits_{j=1}^{n} a_{ij}$ (Theorem 1.3.4, item 5.).

Algorithm Noda iteration

Input: A, λ_0, x^0, $\epsilon > 0$

1: **for** $k = 0, 1, 2, ...$
2: Solve $(\lambda_k I - A)y^{k+1} = x^k$;
3: Normalize $x^{k+1} = \dfrac{y^{k+1}}{||y^{k+1}||_2}$;
4: Compute $\lambda_{k+1} = \lambda_k - \min\limits_{i \in N} \dfrac{x_i^k}{y_i^{k+1}}$;
5: **until** $||\lambda_{k+1} x^{k+1} - A x^{k+1}||_2 < \epsilon$;

Output: $\alpha(A) \leftarrow \lambda_{k+1}$, $x \leftarrow x^{k+1}$

As we see, the presented algorithm is similar to the Rayleigh quotient iteration ([61]), but with different shifts. By analyzing the step 4 in the algorithm, we obtain:

$$\lambda_{k+1} = \lambda_k - \min_{i \in N} \frac{x_i^k}{y_i^{k+1}} = \max_{i \in N} \frac{(\lambda_k y^{k+1} - x^k)_i}{y_i^{k+1}} = \max_{i \in N} \frac{(Ay^{k+1})_i}{y_i^{k+1}} = \max_{i \in N} \frac{(Ax^{k+1})_i}{x_i^{k+1}},$$

i.e.,

$$\lambda_{k+1} = \max_{i \in N} \frac{(Ax^{k+1})_i}{x_i^{k+1}}. \tag{1.13}$$

As $\lambda_0 > \alpha(A)$, $\lambda_0 I - A$ is nonsingular and $(\lambda_0 I - A)^{-1} > 0$ (Proposition 1.6.4). Then, from the steps 2 and 3 in the algorithm, we have $y^1 = (\lambda_0 I - A)^{-1} x^0 > 0$ and $x^1 = \dfrac{y^1}{||y^1||_2} > 0$. Consequently, from the step 4, we obtain $\lambda_1 < \lambda_0$. In addition, from the (1.13) and Theorem 1.3.4 (item 6), it follows $\lambda_1 > \alpha(A)$.

Analogously, by induction, it can be shown that given algorithm generates a decreasing sequence $\{\lambda_k\}$ bounded below by $\alpha(A)$ and a sequence of positive vectors $\{x^k\},\ k \in \mathbb{N}_0$.

Theorem 1.3.5 *([50], Theorem 4.1) Let $A \in \mathbb{R}^{n,n}$ be essentially non-negative irreducible matrix. Then, the pair of sequences $\{\{\lambda_k\}, \{x^k\}\},\ k \in \mathbb{N}_0$, generated by the algorithm* ***Noda iteration****, converges to the Perron-Frobenius pair of A.*

1.4 Geršgorin circles

The sum

$$r_i(A) := \sum_{j \neq i}^{n} |a_{ij}|$$

is called the i^{th} deleted absolute row sum of $A \in \mathbb{C}^{n,n},\ i \in N$, where $r_i(A) := 0$ for $n = 1$. Denote sets in the complex plane:

$$\Gamma_i(A) := \Big\{ z \in \mathbb{C} : |z - a_{ii}| \leq r_i(A) \Big\},\ i \in N.$$

The set $\Gamma_i(A)$ is called i^{th} Geršgorin disk and the set $\Gamma(A) := \bigcup_{i=1}^{n} \Gamma_i(A)$ is called the Geršgorin set. It is a well-known result (Geršgorin's theorem) that the Geršgorin set of a given matrix contains its spectrum.

Theorem 1.4.1 *(A. S. Geršgorin, 1931) For any $A \in \mathbb{C}^{n,n}$ and any $\lambda \in \sigma(A)$, there exists $k \in N$ such that*

$$|\lambda - a_{kk}| \leq r_k(A).$$

Consequently, $\lambda \in \Gamma_k(A) \subseteq \Gamma(A)$. *As this is true for each* $\lambda \in \sigma(A)$, *then*

$$\sigma(A) \subseteq \Gamma(A).$$

Proof: Let $\lambda \in \sigma(A)$. There exists $x = [x_1, x_2, ..., x_n]^T \in \mathbb{C}^n,\ x \neq 0$, such that $Ax = \lambda x$, i.e., $\sum_{j=1}^{n} a_{ij}x_j = \lambda x_i,\ i \in N$. Let $|x_k| = \max\{|x_1|, |x_2|, ..., |x_n|\}$. Thus, $|x_k| > 0$ and $(\lambda - a_{kk})x_k = \sum_{j \neq k}^{n} a_{kj}x_j$. Using the absolute values and the triangle inequality, we have:

$$|\lambda - a_{kk}||x_k| \leq \sum_{j \neq k}^{n} |a_{kj}||x_j| \leq \sum_{j \neq k}^{n} |a_{kj}||x_k| \leq |x_k| r_k(A).$$

Dividing the above expression by $|x_k| > 0$, it follows $\lambda \in \Gamma_k(A)$, i.e., $\lambda \in \Gamma(A)$. As it is true for each eigenvalue, we obtain $\sigma(A) \subseteq \Gamma(A)$. ❑

While the Geršgorin's theorem guaranties that all eigenvalues lie in the union of the Geršgorin circles, it does not reveal how many eigenvalues, if any, are located in every circle. For that purpose, we have the following statement usually known as The second Geršgorin's theorem.
Let S be a proper non-empty subset of N and $n \geq 2$. The cardinality of S is denoted by $\text{card}(S)$ and the complement of S with respect to N is denoted by $\bar{S}$, i.e. $\bar{S} := N \backslash S$. Also, we denote:

$$\Gamma_S(A) := \bigcup_{i \in S} \Gamma_i(A) \text{ and } \Gamma_{\bar{S}}(A) := \bigcup_{i \in \bar{S}} \Gamma_i(A).$$

Theorem 1.4.2 *(A. S. Geršgorin, 1931) For any* $A \in \mathbb{C}^{n,n},\ n \geq 2$, *for which* $\Gamma_S(A) \bigcap \Gamma_{\bar{S}}(A) = \emptyset$, *for some proper non-empty set* S *of* N, *the set* $\Gamma_S(A)$ *contains exactly card*(S) *eigenvalues of* A.

An other important theorem is due to O. Taussky[9], which states that for an irreducible matrix an eigenvalue can be located on the boundary of the Geršgorin set only if all Geršgorin circles pass through it. More precisely, the following holds.

[9] Olga Taussky-Todd (1906-1995) was a Czech-American mathematician who researched in number theory, integral matrices and matrices in algebra and analysis.

Theorem 1.4.3 *(O. Taussky, 1948) Let* $A = [a_{ij}] \in \mathbb{C}^{n.n}$ *be irreducible. If* $\lambda \in \sigma(A)$ *is such that* $\lambda \notin int(\Gamma_i(A))$, *then for all* $i \in N$,

$$|\lambda - a_{ii}| = r_i(A). \tag{1.14}$$

While the Geršgorin set is an elegant way to localize eigenvalues, sometimes it may be crude. In the years that followed Geršgorin's original publications, many improved localization sets were obtain ([11], [14], [18], [19], [35], [39], [63], [64], [67]). Arguable, the most famous one is due to Brauer[10]. Namely, for a given matrix A, define the set

$$K_{i,j}(A) := \Big\{ z \in \mathbb{C} : |z - a_{ii}||z - a_{jj}| \leq r_i(A) r_j(A) \Big\}$$

and the set $\mathcal{K}(A) := \bigcup_{i,j=1}^{n} K_{i,j}(A)$. The following theorem holds.

Theorem 1.4.4 *([63], Theorem 2.2.) For any* $A \in \mathbb{C}^{n,n}$, $n \geq 2$, *and any* $\lambda \in \sigma(A)$, *there exist* $k, l \in N$ *such that*

$$|\lambda - a_{kk}||\lambda - a_{ll}| \leq r_k(A) r_l(A).$$

Consequently, $\lambda \in K_{k,l}(A) \subseteq \mathcal{K}(A)$. *As this is true for each* $\lambda \in \sigma(A)$, *then*

$$\sigma(A) \subseteq \mathcal{K}(A).$$

The set $\mathcal{K}(A)$ is called *the Brauer set*, and for $i, j \in N, i \neq j$, $K_{i,j}(A)$ is called $(i,j) - th$ Brauer-Cassini[11] oval. This set is in fact an improvement of the the Geršgorin set, i.e., we have the following.

Theorem 1.4.5 *([63], Theorem 2.3.) For any* $A \in \mathbb{C}^{n,n}$, $n \geq 2$, *holds*

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

The relationship between individual ovals and the Geršgorin's circles is given in the following.

[10] Alfred Theodor Brauer (1894-1985) was a German-American mathematician who worked in number theory.

[11] Giovanni Domenico Cassini (1625-1712) was an Italian mathematician, astronomer and engineer.

Theorem 1.4.6 *(R. Varga[12], 1999) For any $A \in \mathbb{C}^{n,n},\ n \geq 2$ holds:* $K_{i,j}(A) = \Gamma_i(A) \cup \Gamma_j(A)$ *if and only if* $r_i(A) = r_j(A) = 0$ *or* $r_i(A) = r_j(A) > 0,\ a_{ii} = a_{jj},\ i, j \in N,\ i \neq j$.

If all the off-diagonal row sums and diagonal entries of the matrix are fixed, the best eigenvalue localization we can obtain is the Brauer set. This optimality result was obtained in [63]. More precisely, define the families of matrices:

$$\tilde{\omega}(A) := \{B = [b_{ij}] \in \mathbb{C}^{n,n} : b_{ii} = a_{ii} \text{ and } r_i(B) = r_i(A),\ i \in N\}$$

and

$$\hat{\omega}(A) := \{B = [b_{ij}] \in \mathbb{C}^{n,n} : b_{ii} = a_{ii} \text{ and } r_i(B) \leq r_i(A),\ i \in N\}.$$

Using the notation:

$$\sigma(\tilde{\omega}(A)) = \bigcup_{B \in \tilde{\omega}(A)} \sigma(B) \text{ and } \sigma(\hat{\omega}(A)) = \bigcup_{B \in \hat{\omega}(A)} \sigma(B),$$

we have the following result.

Theorem 1.4.7 *([63], Theorem 2.4.) For any $A \in \mathbb{C}^{n,n},\ n \geq 2$ holds:*

$$\sigma(\tilde{\omega}(A)) = \begin{cases} \partial\mathcal{K}(A) = \partial K_{1,2}(A),\ n = 2 \\ \mathcal{K}(A),\ n \geq 3 \end{cases}$$

and

$$\sigma(\hat{\omega}(A)) = \mathcal{K}(A).$$

While many improvements of the Geršgorin set were obtained and used in the applications of linear algebra, the central role is definetely played by the minimal Geršgorin set. In the following, we provide all its important properties before we continue to the analysis of its computations.

[12]Richard Varga (1928) is an American mathematician who specialized in numerical analysis and linear algebra, emeritus professor at Kent State University and adjunct professor at Case Western Reserve University.

1.5 Minimal Geršgorin set

Given a positive vector $x = [x_1, x_2, \ldots, x_n]^T > 0$ and a diagonal matrix $X := \mathrm{diag}(x) \in \mathbb{R}^{n,n}$, the Geršgorin disks for the matrix $X^{-1}AX$ are given by:

$$\Gamma_i^{r^x}(A) := \Big\{ z \in \mathbb{C} : |z - a_{ii}| \leq r_i^x(A) := \sum_{j \neq i}^{n} \frac{|a_{ij}| x_j}{x_i} \Big\}, \quad \text{for} \quad i \in N. \tag{1.15}$$

Moreover, the associated Geršgorin set is defined as:

$$\Gamma^{r^x}(A) := \bigcup_{i=1}^{n} \Gamma_i^{r^x}(A). \tag{1.16}$$

The set

$$\Gamma^{\mathcal{R}}(A) := \bigcap_{x \in \mathbb{R}^n,\, x>0} \Gamma^{r^x}(A) \tag{1.17}$$

is called the minimal Geršgorin set and it gives the sharpest inclusion set for $\sigma(A)$, with a respect to all positive diagonal similarity transformations $X^{-1}AX$ of A, i.e.,

$$\sigma(A) \subseteq \Gamma^{\mathcal{R}}(A) \subseteq \Gamma(A). \tag{1.18}$$

Note that, generally, when $X \in \mathbb{C}^{n,n}$ is not necessarily diagonal, the intersection of sets $\Gamma(X^{-1}AX)$, over the family of all nonsingular X, is $\sigma(A)$.

To state the optimality of this set, given any matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$ and the complex number $z \in \mathbb{C}$, define the matrix $Q_A(z) = [q_{ij}(z)]$, $Q_A : \mathbb{C} \to \mathbb{R}^{n,n}$ by:

$$q_{ii}(z) := -|z - a_{ii}| \quad \text{and} \quad q_{ij}(z) := |a_{ij}|, \quad \text{for } i \neq j,\ i, j \in N. \tag{1.19}$$

The right-most eigenvalue $\nu_A(z)$ of the essentially non-negative matrix $Q_A(z) = [q_{ij}(z)]$ is real (Theorem 1.3.1) and it can be characterized by:

$$\nu_A(z) = \inf_{x>0} \max_{i \in N} ((Q_A(z)x)_i / x_i). \tag{1.20}$$

i.e.,

$$\nu_A(z) = \inf_{x>0} \max_{i \in N} (r_i^x(A) - |z - a_{ii}|). \tag{1.21}$$

The corresponding eigenvector to $\nu_A(z)$ is non-negative. In case when A (and thus $Q_A(z)$) is irreducible, the corresponding eigenvector is positive (Theorem 1.3.4).

Using this notation, we obtain the following characterization of the minimal Geršgorin set in the complex plane.

Theorem 1.5.1 *([63], Preposition 4.3.) For any $A \in \mathbb{C}^{n.n}, n \geq 2$:*

$$z \in \Gamma^{\mathcal{R}}(A) \text{ if and only if } \nu_A(z) \geq 0. \tag{1.22}$$

Proof: ($\Longrightarrow$) Assume $z \in \Gamma^{\mathcal{R}}(A)$. By the definition of the minimal Geršgorin set, $z \in \Gamma^{r^x}(A)$ for all $x \in \mathbb{R}^n,\ x > 0$. Equivalently, for each $x > 0$ there exists $k \in N$, such that $|z - a_{kk}| \leq r_k^x(A)$, i.e., $r_k^x(A) - |z - a_{kk}| \geq 0$. Hence, $(Q_A(z)x)_k / x_k \geq 0$, i.e., $\max\limits_{i \in N}((Q_A(z)x)_i / x_i) \geq 0$, for each $x > 0$. So, $\nu_A(z) \geq 0$.
($\Longleftarrow$) Assume $\nu_A(z) \geq 0$. Then, for each $x \in \mathbb{R}^n,\ x > 0$, there exists $k \in N$ such that $0 \leq \nu_A(z) \leq (Q_A(z)x)_k / x_k = r_k^x(A) - |z - a_{kk}|$. It implies $|z - a_{kk}| \leq r_k^x(A)$, i.e., $z \in \Gamma^{r^x}(A)$, for each $x > 0$, and therefore, $z \in \Gamma^{\mathcal{R}}(A)$. ❑

From the previous theorem and the continuity of $\nu_A(z)$, as a function of z, the following characterization of the boundary of the minimal Geršgorin set theorem holds.

Theorem 1.5.2 *For any $A \in \mathbb{C}^{n.n}, n \geq 2$:*
$z \in \partial\Gamma^{\mathcal{R}}(A)$ *if and only if $\nu_A(z) = 0$ and there is a sequence of complex numbers $\{z_j\}_{j=1}^{\infty}$ with $\lim\limits_{j \to \infty} z_j = z$, for which $\nu_A(z_j) < 0$, for all $j \in \mathbb{N}$.*

Next, we review the optimality of the minimal Geršgorin set. For a given matrix $A \in \mathbb{C}^{n,n}$, the family of matrices

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

is called the equimodular family of A and the family of matrices

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

is called the extended equimodular family of A.
We will use notation:

$$\sigma(\Omega(A)) := \bigcup_{B \in \Omega(A)} \sigma(B) \text{ and } \sigma(\hat{\Omega}(A)) := \bigcup_{B \in \hat{\Omega}(A)} \sigma(B).$$

So, it is obvious that

$$\sigma(\Omega(A)) \subseteq \sigma(\hat{\Omega}(A)) \subseteq \Gamma^{\mathcal{R}}(A). \tag{1.23}$$

The next theorem gives the sharpness of the inclusion in (1.23), i.e., it shows that the minimal Geršgorin set is optimal spectral localization for the extended equimodular family of matrices.

Theorem 1.5.3 *([63], Theorem 4.4.) For any* $A = [a_{ij}] \in \mathbb{C}^{n,n}, n \geq 2$:

$$\partial\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\Omega(A)) \subseteq \sigma(\hat{\Omega}(A)) = \Gamma^{\mathcal{R}}(A). \tag{1.24}$$

Proof: Because of (1.23), it is enough to show that $\partial\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\Omega(A))$ and $\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\hat{\Omega}(A))$.

First we show that $\partial\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\Omega(A))$. To that end, we prove that for any $A = [a_{ij}] \in \mathbb{C}^{n,n}, n \geq 2$, and $z \in \mathbb{C}$ with $\nu_A(z) = 0$, there exists a matrix $B = [b_{ij}] \in \Omega(A)$ for which z is an eigenvalue of B.
Let $z \in \partial\Gamma^{\mathcal{R}}(A)$. So, $\nu_A(z) = 0$ and there exists $y \in \mathbb{R}^n,\ y \geq 0,\ y \neq 0$, such that $Q_A(z)y = \nu_A(z)y = 0$. It follows that $|z - a_{ii}|y_i = \sum_{j \neq i}^{n} |a_{ij}|y_j,\ i \in N$. Let ψ_i be the numbers such that $z - a_{ii} = |z - a_{ii}|e^{\mathrm{i}\psi_i}, i \in N$, and define the matrix $B = [b_{ij}] \in \mathbb{C}^{n,n}$ by:

$$b_{ii} := a_{ii} \ \text{ and } \ b_{ij} := |a_{ij}|e^{\mathrm{i}\psi_i},\ i \neq j,\ i, j \in N.$$

Now, it is easy to see that $B \in \Omega(A)$. Namely, by computing $(By)_i$, we find:
$(By)_i = \sum_{j=1}^{n} b_{ij}y_j = a_{ii}y_i + \sum_{j \neq i}^{n} |a_{ij}|e^{\mathrm{i}\psi_i}y_j = (z - |z - a_{ii}|e^{\mathrm{i}\psi_i})y_i + \sum_{j \neq i}^{n} |a_{ij}|e^{\mathrm{i}\psi_i}y_j =$
$zy_i + e^{\mathrm{i}\psi_i}(\sum_{j \neq i}^{n} |a_{ij}|y_j - |z - a_{ii}|y_i) = zy_i,$ for all $i \in N$. It follows $By = zy$, i.e., $z \in \sigma(B)$. So, each point of $\partial\Gamma^{\mathcal{R}}(A)$ is in $\sigma(\Omega(A))$, i.e., $\partial\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\Omega(A))$.

Next, we show that $\Gamma^{\mathcal{R}}(A) \subseteq \sigma(\hat{\Omega}(A))$. Let $z \in \Gamma^{\mathcal{R}}(A)$. So, $\nu_A(z) \geq 0$ and there exists $y \in \mathbb{R}^n,\ y \geq 0,\ y \neq 0$, such that $Q_A(z)y = \nu_A(z)y$. It follows that $|z - a_{ii}|y_i = \sum_{j \neq i}^{n} |a_{ij}|y_j - \nu_A(z)y_i,\ i \in N$. Define the matrix $C = [b_{ij}] \in \mathbb{C}^{n,n}$ by:

$$c_{ii} := a_{ii} \ \text{ and } \ c_{ij} := \mu_i a_{ij},\ i \neq j,\ i, j \in N,$$

where μ_i is defined by:

$$\mu_i = \begin{cases} \dfrac{\sum\limits_{j\neq i}^{n} |a_{ij}|y_j - \nu_A(z)y_i}{\sum\limits_{j\neq i}^{n} |a_{ij}|y_j}, & \sum\limits_{j\neq i}^{n} |a_{ij}|y_j > 0 \\ 1, & \sum\limits_{j\neq i}^{n} |a_{ij}|y_j = 0 \end{cases}.$$

It is easy to see that $0 \leq \mu_i \leq 1$ and, consequently, $C \in \hat{\Omega}(A)$. But since for all $i \in N$, $|z - c_{ii}|y_i = |z - a_{ii}|y_i = \sum\limits_{j\neq i}^{n} |a_{ij}|y_j - \nu_A(z)y_i = \mu_i \sum\limits_{j\neq i}^{n} |a_{ij}|y_j = \sum\limits_{j\neq i}^{n} |c_{ij}|y_j$, i.e., $|z - c_{ii}|y_i = \sum\limits_{j\neq i}^{n} |c_{ij}|y_j$, we have that $z \in \partial\Gamma^{\mathcal{R}}(C)$.
Finally, according to the the first part of the proof of this theorem, we have that there exists $D \in \Omega(C) \subseteq \hat{\Omega}(A)$ such that $z \in \sigma(D)$. It follows $z \in \sigma(\hat{\Omega}(A))$ and consequently, $\sigma(\hat{\Omega}(A)) = \Gamma^{\mathcal{R}}(A)$. ❑

Therefore, $\sigma(\hat{\Omega}(A))$ completely fills out $\Gamma^{\mathcal{R}}(A)$, i.e., the minimal Geršgorin set is "optimal" because it exactly determines the spectrum of the family of matrices $\sigma(\hat{\Omega}(A))$.

To conclude this section, we give some useful properties of the minimal Geršgorin set.

Theorem 1.5.4 *For any $A \in \mathbb{C}^{n,n}$ and $c \in \mathbb{C}$ holds:*

1) $\Gamma^{\mathcal{R}}(A)$ *is a compact set in* $\mathbb{C}$;

2) $\Gamma^{\mathcal{R}}(A) = \Gamma^{\mathcal{R}}(A^T)$;

3) $\Gamma^{\mathcal{R}}(cA) = c\Gamma^{\mathcal{R}}(A)$;

4) $\Gamma^{\mathcal{R}}(A + cI) = \Gamma^{\mathcal{R}}(A) + c$.

Proof: 1) As the Geršgorin set is a compact set (the union of n bounded and closed sets) in $\mathbb{C}$, it follows that $\Gamma^{\mathcal{R}}(A)$ is also compact.

2) For $n = 1$, the statement holds trivially. For a given $z \in \mathbb{C},\ n \geq 2$, matrices $Q_A(z)$ and $Q_{A^T}(z)$ have the same eigenvalues. It follows $\nu_A(z) = \nu_{A^T}(z)$, i.e., $z \in \Gamma^{\mathcal{R}}(A)$ if and only if $z \in \Gamma^{\mathcal{R}}(A^T)$.

3) For $c = 0$ the statement holds trivially. For $c \neq 0$, we have: $z \in \Gamma^{\mathcal{R}}(cA) \Longleftrightarrow \inf\limits_{x>0} \max\limits_{i \in N}(|c| r_i^x(A) - |z - ca_{ii}|) \geq 0 \Longleftrightarrow \inf\limits_{x>0} \max\limits_{i \in N}(r_i^x(A) - |\frac{z}{c} - a_{ii}|) \geq 0 \Longleftrightarrow \frac{z}{c} \in \Gamma^{\mathcal{R}}(A) \Longleftrightarrow z \in c\Gamma^{\mathcal{R}}(A)$.

4) $z \in \Gamma^{\mathcal{R}}(A+cI) \Longleftrightarrow \inf\limits_{x>0} \max\limits_{i \in N}(r_i^x(A) - |z - (a_{ii}+c)|) \geq 0 \Longleftrightarrow \inf\limits_{x>0} \max\limits_{i \in N}(r_i^x(A) - |(z-c) - a_{ii}|) \Longleftrightarrow z - c \in \Gamma^{\mathcal{R}}(A) \Longleftrightarrow z \in \Gamma^{\mathcal{R}}(A) + c$.

❑

Theorem 1.5.5 *Given an arbitrary matrix* $A = [a_{ij}] \in \mathbb{C}^{n,n}, n_i \geq 1,\ i \in \{1, 2, ..., m\}$ *and its normal reduced form defined by* (1.12). *Then,*

$$\Gamma^{\mathcal{R}}(A) = \bigcup_{i=1}^{m} \Gamma^{\mathcal{R}}(A_{ii}).$$

Proof: Firstly, we can notice that diagonal elements and their corresponding deleted absolute row and column sums are the same for a matrix and its normal reduced form, i.e., the minimal Geršgorin set is invariant under simultaneous permutations of rows and columns. Let's define positive diagonal matrices $X_k = \text{diag}(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)})$, where $k \in \mathbb{N}$ and $x_j^{(k)} := \frac{1}{i^k}$, for all j from the set of indices corresponding to the block $A_{ii},\ i \in \{1, 2, ..., m\}$. From

$$\lim_{k \to \infty} X_k^{-1} A X_k = \begin{bmatrix} A_{11} & & & \\ & A_{22} & & \\ & & \ddots & \\ & & & A_{mm} \end{bmatrix},$$

it follows $\Gamma^{\mathcal{R}}(A) \subseteq \bigcap\limits_{k \in \mathbb{N}} \Gamma(X_k^{-1} A X_k) = \bigcup\limits_{i=1}^{m} \Gamma(A_{ii}) \subseteq \Gamma^{\mathcal{R}}(A)$. So, $\Gamma^{\mathcal{R}}(A) = \bigcup\limits_{i=1}^{m} \Gamma^{\mathcal{R}}(A_{ii})$.

❑

In some cases, a component of $\Gamma^{\mathcal{R}}(A),\ A \in \mathbb{C}^{n,n}$, can be isolated point (for example when all off-diagonal entries in one row of A are zero). To avoid such situations in the algorithms, we can use the previous lemma and consider only irreducible matrices.

1.6 Nonsingularity of matrices vs. localization of eigenvalues

In this section, a brief review of few classes of matrices is given and to be used in the rest of the thesis. Also, a relation between these classes of matrices and the corresponding results on localization of eigenvalues is presented.

A matrix $A = [a_{ij}] \in \mathbb{C}^{n.n}$ is diagonally dominant (DD) if

$$|a_{ii}| \geq r_i(A),\ i \in N.$$

A matrix $A = [a_{ij}] \in \mathbb{C}^{n.n}$ is strictly diagonally dominant (SDD) if

$$|a_{ii}| > r_i(A),\ i \in N.$$

Theorem 1.6.1 *([63], Theorem 1.4.) If $A \in \mathbb{C}^{n.n}$ is strictly diagonally dominant, then A is nonsingular matrix.*

A matrix $A \in \mathbb{C}^{n,n}$ is an irreducibly diagonally dominant matrix if it is irreducible and diagonally dominant ($|a_{ii}| \geq r_i(A), i \in N$) and if strict inequality ($|a_{kk}| > r_k(A)$) holds for at least one $k \in N$.

Theorem 1.6.2 *(O. Taussky, 1949) If $A = [a_{ij}] \in \mathbb{C}^{n,n}$ is an irreducibly diagonally dominant, then A is a nonsingular matrix.*

A matrix $A = [a_{ij}] \in \mathbb{C}^{n.n}$ is an Ostrowski[13] matrix if $|a_{ii}||a_{jj}| > r_i(A)r_j(A)$, $i, j \in N,\ i \neq j,\ n \geq 2$.

Theorem 1.6.3 *([63], Theorem 2.1.) If $A \in \mathbb{C}^{n.n}$ is an Ostrowski matrix, then A is a nonsingular matrix.*

[13] Alexander Ostrowski (1893-1986) was a Ukrainian mathematician, professor at University of Basel.

A given Z-matrix $A = [a_{ij}] \in \mathbb{R}^{n,n}$ is an M-matrix if there exist $\mu \in \mathbb{R}$ and $B \in \mathbb{R}^{n,n}$ such that $A = \mu I - B,\ B \geq 0$ and $\rho(B) \leq \mu$.

Proposition 1.6.4 *([5]) For a Z-matrix $A = [a_{ij}] \in \mathbb{R}^{n,n}$, the following statements are equivalent:*

- *A is a nonsingular M-matrix;*
- *$A^{-1} \geq 0$;*
- *there exists vector $x \in \mathbb{R}^n,\ x > 0$, such that $Ax > 0$;*
- *the real part of each eigenvalue of A is positive.*

For a given matrix $A \in \mathbb{C}^{n,n}$, define its comparison matrix $\langle A \rangle := [\alpha_{ij}] \in \mathbb{R}^{n,n}$ such that:

$$\alpha_{ij} := \begin{cases} |a_{ii}|, & i = j \\ -|a_{ij}|, & i \neq j \end{cases},$$

where $i, j \in N$. We can notice that $\langle A \rangle$ is a Z-matrix.

A given matrix $A \in \mathbb{C}^{n,n}$ is an H-matrix if $\langle A \rangle$ is an M-matrix. If $A \in \mathbb{C}^{n,n}$ and $\langle A \rangle$ is a nonsingular M-matrix, then A is a nonsingular H-matrix. The class of H-matrices plays a special role in the theory of the localization of eigenvalues since it is superclass of many classes of nonsingular matrices used to construct localizations. Among many such classes we mentioned just a few of them.

Theorem 1.6.5 *Classes of strictly diagonally dominant matrices, irreducible diagonally dominant matrices and Ostrowski matrices are subclasses of nonsingular H-matrices.*

While the relationship between nonsingularity results and eigenvalue localization sets has been implicitly present since the early years of the matrix theory, their equivalence was explicitly formulated for the first time in [63] and formalized in [39]. Here, we state it in the form of the following theorem.

Theorem 1.6.6 *(Varga's principle of equivalence) Let $\mathbb{K}$ be a class of matrices from $\mathbb{C}^{n,n}$ and for any $A \in \mathbb{C}^{n,n}$, the set $\Theta^{\mathbb{K}}(A)$ is defined by:*

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

The following statements are equivalent:

- *all matrices from* $\mathbb{K}$ *are nonsingular;*
- $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$, *for any* $A \in \mathbb{C}^{n,n}$.

Proof: ($\Longrightarrow$) Assume that all matrices from $\mathbb{K}$ are nonsingular. Let $A \in \mathbb{C}^{n,n}$ and let $\lambda \in \sigma(A)$ be its arbitrary eigenvalue. Then, $A - \lambda I$ is singular and because of that $A - \lambda I \notin \mathbb{K}$. It follows $\lambda \in \Theta^{\mathbb{K}}(A)$, i.e, $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$.
($\Longleftarrow$) Assume $\sigma(A) \subseteq \Theta^{\mathbb{K}}(A)$, for all $A \in \mathbb{C}^{n,n}$. Suppose there exists a matrix $\tilde{A} \in \mathbb{K}$ which is singular. So, $\lambda = 0$ is its eigenvalue and $0 \in \Theta^{\mathbb{K}}(\tilde{A})$, i.e., $\tilde{A} - 0 \cdot I = \tilde{A} \notin \mathbb{K}$, which gives a contradiction. Thus, all matrices from $\mathbb{K}$ are nonsingular. ❑

As expected, wider classes produce better localization sets. Namely, we have the following.

Theorem 1.6.7 *(Monotonicity principle, [39]) Let* $\mathbb{K}_1$ *and* $\mathbb{K}_2$ *be classes of matrices from* $\mathbb{C}^{n,n}$ *and* $A \in \mathbb{C}^{n,n}$. *If* $\mathbb{K}_1 \subseteq \mathbb{K}_2$, *then* $\Theta^{\mathbb{K}_2}(A) \subseteq \Theta^{\mathbb{K}_1}(A)$.

Due to many scientific papers on the localization of eigenvalues similar to the Geršgorin set, the term "Geršgorin-type" was vaguely used. We will adapt the formalization of [39] from where one can conclude that the set $\Theta^{\mathbb{K}}(A)$ is a set of Geršgorin-type if the class $\mathbb{K}$ is a subset of the set of nonsingular H-matrices.

One can easily verify the following statements.

- If $\mathbb{K}$ is a class of SDD matrices and $A \in \mathbb{C}^{n,n}$, then $\Theta^{\mathbb{K}}(A) = \Gamma(A)$.
- If $\mathbb{K}$ is a class of Ostrowski matrices and $A \in \mathbb{C}^{n,n}$, then $\Theta^{\mathbb{K}}(A) = \mathcal{K}(A)$.
- If $\mathbb{K}$ is a class of a nonsingular H-matrices and $A \in \mathbb{C}^{n,n}$, then $\Theta^{\mathbb{K}}(A) = \Gamma^{\mathcal{R}}(A)$.

Therefore, for a given $A \in \mathbb{C}^{n,n}$, if $z \in \Gamma^{\mathcal{R}}(A)$, then $A - zI$ is not a nonsingular H-matrix, i.e, $\langle A - zI \rangle$ is not a nonsingular M-matrix. Also, using monotonicity pinciple (Theorem 1.6.7), it follows that the minimal Geršgorin set is the smallest of all here mentioned sets of Geršgorin-type, i.e., the following result holds.

Corollary 1.6.8 *For any matrix* $A \in \mathbb{C}^{n,n}$, $n \geq 2$, *holds:*

$$\Gamma^{\mathcal{R}}(A) \subseteq \mathcal{K}(A) \subseteq \Gamma(A).$$

In a special case for $n = 2$, the minimal Geršgorin set is equal to the Brauer set, i.e., $\Gamma^{\mathcal{R}}(A) = \mathcal{K}(A)$.

This property holds for many other localization sets. For details on the other subclasses of H-matrices and their corresponding eigenvalue localization sets see [63].

1.7 Numerical range

Beside the Geršgorin's circles, there exist many other localization sets of the spectrum. One of them is a numerical range or a field of values.

The numerical range $W(A)$ of a given matrix $A \in \mathbb{C}^{n,n}$ is a set defined by:

$$W(A) := \left\{ \frac{x^*Ax}{x^*x} : x \in \mathbb{C}^n,\ x \neq 0 \right\} = \{y^*Ay : y \in \mathbb{C}^n,\ ||y||_2 = 1\}. \qquad (1.25)$$

In another words, the numerical range is the range of the Rayleigh quotients $\frac{x^*Ax}{x^*x}$, for all $x \neq 0$. In the following, we mention some well-known properties of the numerical range.

Theorem 1.7.1 *$W(A)$ is a non-empty, compact and connected set in $\mathbb{C}$.*

Theorem 1.7.2 *(Toeplitz[14]-Hausdorff[15]) $W(A)$ is a convex subset of $\mathbb{C}$.*

Theorem 1.7.3 *For matrices $A, B \in \mathbb{C}^{n,n}$, an identity matrix $I \in \mathbb{R}^{n,n}$, an unitary matrix $U \in \mathbb{C}^{n,n}$ and scalar $c \in \mathbb{C}$, the following properties hold:*

1) $W(cA) = cW(A)$,

2) $W(cI + A) = c + W(A)$,

3) $W(A + B) \subseteq W(A) + W(B)$,

4) $W(U^*AU) = W(A)$.

[14] Otto Toeplitz (1881-1940) was a German mathematician working in functional analysis.

[15] Felix Hausdorff (1868-1942) was a German mathematician who is considered to be one of the founders of modern topology.

Proof: 1) If $c = 0$, the statement holds trivially. Let $z \in W(cA),\ c \neq 0$. There exists $x \in \mathbb{C}^n$ such that $z = x^*(cA)x,\ x^*x = 1$, i.e., $\frac{z}{c} = x^*Ax \Longleftrightarrow \frac{z}{c} \in W(A) \Longleftrightarrow z \in cW(A)$.

2) Let $z \in W(cI+A)$. There exists $x \in \mathbb{C}^n$ such that $z = x^*(cI+A)x,\ x^*x = 1$, i.e., $z = c + x^*Ax \Longleftrightarrow z - c \in W(A) \Longleftrightarrow z \in c + W(A)$.

3) Let $z \in W(A+B)$. There exists $x \in \mathbb{C}^n$ such that $z = x^*(A+B)x,\ x^*x = 1$, i.e., $z = x^*(A+B)x = x^*Ax + x^*Bx \Longrightarrow z \in W(A) + W(B)$.

4) Let $z \in W(U^*AU)$. There exists $x \in \mathbb{C}^n$ such that $z = x^*(U^*AU)x,\ x^*x = 1 \Longleftrightarrow z = y^*Ay,\ y^*y = 1 \Longleftrightarrow z \in W(A)$, where $y = Ux$.

❑

Theorem 1.7.4 *For an arbitrary matrix $A \in \mathbb{C}^{n,n}$, it holds that*

$$\sigma(A) \subseteq W(A). \tag{1.26}$$

Proof: Let $\lambda \in \mathbb{C}$ be an arbitrary eigenvalue of A and $x \in \mathbb{C}^n,\ x \neq 0$, is its associated eigenvector. Without loss of generality, assume that x is normalized, i.e., $||x||_2 = 1$. Then, $\lambda = \lambda(x^*x) = x^*\lambda x = x^*Ax$, i.e., $\lambda \in W(A)$. As this holds for each eigenvalue of A, it follows that $\sigma(A) \subseteq W(A)$. ❑

Theorem 1.7.5 *([71], Theorem 2) $W(A)$ is a real line segment $[a, b]$ if and only if $A \in \mathbb{C}^{n,n}$ is a Hermitian matrix with its the smallest eigenvalue a and the largest eigenvalue b.*

The abscissa of the numerical range $W(A)$ of matrix $A \in \mathbb{C}^{n,n}$ is

$$\omega(A) := \max\{\mathrm{Re}(z) : z \in W(A)\}. \tag{1.27}$$

So, $W(A)$ lies in a closed half-plane, i.e.,

$$W(A) \subseteq \{z \in \mathbb{C} : \mathrm{Re}(z) \leq \omega(A)\}.$$

For a given set S in the complex plane, its convex hull $\mathrm{co}(S)$ is

$$\mathrm{co}(S) = \bigcap_T \{T : T \text{ is convex and } S \subseteq T\}.$$

Theorem 1.7.6 *([71], Theorem 3) If $A \in \mathbb{C}^{n,n}$ is a normal matrix, the set $W(A)$ is the convex hull of its eigenvalues, i.e.,*

$$W(A) = co(\sigma(A)) = \{t\lambda_1 + (1-t)\lambda_2 : \lambda_1, \lambda_2 \in \sigma(A), t \in [0,1]\}.$$

A Hermitian part of a given matrix $A \in \mathbb{C}^{n,n}$ is:

$$H(A) := \frac{A + A^*}{2}.$$

It is obvious that $H(A)$ is a Hermitian matrix, implying that all its eigenvalues are real numbers. Also, because of Theorem 1.7.5, the numerical range of $H(A)$ is closed interval $[\lambda_{min}, \lambda_{max}]$, where λ_{min} and λ_{max} are its the smallest and largest eigenvalue, respectively.

Lemma 1.7.7 *([63], Lemma 3.8.) For a given matrix $A \in \mathbb{C}^{n,n}$ holds:*

$$W(H(A)) = Re(W(A)) := \{Re(z) : z \in W(A)\}.$$

First connection between Geršgorin's localizations and numerical range was established by C. R. Johnson[16] in [36]. Here, we state the result with its proof.

Theorem 1.7.8 *([63], Theorem 3.9.) For a given matrix $A \in \mathbb{C}^{n,n}$ holds:*

$$W(A) \subseteq J(A) := co\Big(\bigcup_{i=1}^{n}\Big\{z \in \mathbb{C} : |z - a_{ii}| \leq \frac{r_i(A) + r_i(A^T)}{2}\Big\}\Big).$$

Proof: First, we will prove the statement "*If* $J(A) \subseteq \{z \in \mathbb{C} : \ \mathrm{Re}(z) > 0\}$, *then* $W(A) \subseteq \{z \in \mathbb{C} : Re(z) > 0\}$".

From $J(A) \subseteq \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$, it is obvious that $\mathrm{Re}(a_{ii}) > \dfrac{r_i(A) + r_i(A^T)}{2}$, $i \in N$. For $H(A) = \dfrac{A + A^*}{2}$, we obtain:

$$r_i(H(A)) = \sum_{j \neq i}^{n} \frac{|a_{ij} + a_{ij}^*|}{2} \leq \frac{\sum\limits_{j \neq i}^{n} |a_{ij}| + \sum\limits_{j \neq i}^{n} |a_{ij}^*|}{2} = \frac{r_i(A) + r_i(A^T)}{2},$$

[16] Charles Royal Johnson (1948) is an American mathematician specializing in linear algebra.

$i \in N$. Thus, the whole Geršgorin set of $H(A)$ lies in the open right half-plane. As $H(A)$ is Hermitian, it has only positive real eigenvalues and $W(H(A))$ is a real line segment (Theorem 1.7.5). So, $W(A) \subseteq \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$ (Theorem 1.7.7).
Next, we prove the second statement: "*If* $0 \notin J(A)$, *then* $0 \notin W(A)$".
Assume that $0 \notin J(A)$. As $J(A)$ is a convex set by definition, there exists $\theta \in [0, 2\pi)$ such that $J(e^{\mathrm{i}\theta}A) \subseteq \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$ (Separating hyperplane theorem, Horn[17] and Johnson, 1985). Using the first statement, this implies that $W(e^{\mathrm{i}\theta}A) \subseteq \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$. As $W(A) = e^{-\mathrm{i}\theta}W(e^{\mathrm{i}\theta}A)$ (Theorem 1.7.3), it follows that $0 \notin W(A)$.
Finally, we prove the third statement: "*If* $z \notin J(A)$, *then* $z \notin W(A)$".
If $z \notin J(A)$, then $0 \notin J(A - zI)$. From the second statement, we obtain $0 \notin W(A - zI) = W(A) - z$, which gives that $z \notin W(A)$.
Therefore, if $z \in W(A)$, then $z \in J(A)$, i.e., $W(A) \subseteq J(A)$. ❑

Motivated by this, in the following, we will consider the relationship between the numerical range and the minimal Geršgorin set.

[17] Roger Alan Horn (1942) is an American mathematician specializing in matrix analysis.

Chapter 2

Algorithms for computing the minimal Geršgorin set

"The beauty and simplicity of Geršgorin's theorem has undoubtedly inspired further research, resulting in hundreds of papers in which the name Geršgorin appears."

Richard Varga

In this chapter, in Sections 2.1 and 2.2, we review existing algorithms for computing the minimal Geršgorin set. In Section 2.3, we present two different characterizations - explicit and implicit. The implicit characterization is a new result published in [49]. Then, in Section 2.4, we introduce three different approaches for tracing of the boundary of the minimal Geršgorin set (star-shaped, predictor-corrector, triangular), that are combined with two characterizations resulting in six algorithms.

2.1 Grid based algorithm

The griding MGS algorithm **(gMGS)** is the simplest algorithm for computing the minimal Geršgorin set. It is constructed by computing values of the function $\nu_A(z)$, for all z in a given rectangular grid.

As the minimal Geršgorin set is a subset of the Geršgorin set of a given matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, we can easily construct the rectangular grid $[l_{re}, u_{re}] \times [l_{im}, u_{im}]$, where:

$$l_{re} = \min_{i \in N}(\mathrm{Re}(a_{ii}) - r_i(A)), \quad u_{re} = \max_{i \in N}(\mathrm{Re}(a_{ii}) + r_i(A)), \tag{2.1}$$

$$l_{im} = \min_{i \in N}(\mathrm{Im}(a_{ii}) - r_i(A)), \quad u_{im} = \max_{i \in N}(\mathrm{Im}(a_{ii}) + r_i(A)). \tag{2.2}$$

The approximation of $\partial\Gamma^{\mathcal{R}}(A)$ is obtained as a zero-level curve on that grid.

Therefore, the algorithm **gMGS** is simple for implementation. For input parameters, it uses a given matrix A and a parameter n_g, which determines the number of points in the grid. However, in cases when we use a fine grid in order to get a better localization, this algorithm requires a large number of eigenvalue computations. So, the algorithm **gMGS** is numerically expensive, even for matrices of small sizes.

Algorithm gMGS

Input: A, n_g

1: Compute l_{re}, u_{re}, l_{im} and u_{im} using (2.1) and (2.2);
2: Set $\delta_x = \frac{u_{re}-l_{re}}{n_g}$ and $\delta_y = \frac{u_{im}-l_{im}}{n_g}$;
3: **for** $k_x = 0 : n_g$ **do**
4: **for** $k_y = 0 : n_g$ **do**
5: Set $z := (l_{re} + k_x\delta_x) + \mathrm{i}(l_{im} + k_y\delta_y)$;
6: Compute $G(k_x, k_y) := \nu_A(z)$ as the rightmost eigenvalue (r.m.e) of $Q_A(z)$;
7: **end for**
8: **end for**
9: Compute the zero level curve $\mathcal{C}$ of G;

Output: $\mathcal{C}$

2.2 Bisection based approximation

An improvement to the straight-forward **gMGS** algorithm was introduced by R. Varga, Lj. Cvetković and V. Kostić in [69]. Here we call it the bisection MGS algorithm **(bMGS)**, The approximation of $\Gamma^{\mathcal{R}}(A)$ is obtained by computing

several points on the boundary and through them constructing the desired set using Theorem 1.4.3. In order to obtain the points on the boundary, the proposed approach was to use the fact that the minimal Geršgorin set is star shaped.

Definition 2.2.1 *A set $S \in \mathbb{C}$ is called star-shaped (or star-convex or star-domain) if there exists an $x_0 \in S$ such that for all $x \in S$ the line segment from x_0 to x is in S.*

Theorem 2.2.2 *([63], Theorem 4.6.) If $A = [a_{ij}] \in \mathbb{C}^{n,n}$, $n \geq 2$, is an arbitrary irreducible matrix, then $\nu_A(a_{ii}) > 0$ for each $i \in N$. Moreover, for each $i \in N$ and each real $\theta \in [0, 2\pi)$, there exists the largest number $\widehat{\rho}_i(\theta) > 0$ such that*

$$\nu_A(a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta}) = 0 \ \text{ and } \ \nu_A(a_{ii} + te^{\mathrm{i}\theta}) \geq 0, \ \text{ for all } \ 0 \leq t < \widehat{\rho}_i(\theta), \tag{2.3}$$

i.e., the entire complex interval $[a_{ii} + te^{\mathrm{i}\theta}]_{t=0}^{\widehat{\rho}_i(\theta)}$ is contained in $\Gamma^{\mathcal{R}}(A)$. Consequently, for each $i \in N$, the set

$$\bigcup_{\theta=0}^{2\pi} [a_{ii} + te^{\mathrm{i}\theta}]_{t=0}^{\widehat{\rho}_i(\theta)} \tag{2.4}$$

is a star-shaped subset of the miminal Geršgorin set with respect to a_{ii} and

$$a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A). \tag{2.5}$$

Proof: Irreducibility of A implies that $Q_A(z)$ is also an irreducible matrix, for any choice of $z \in \mathbb{C}$. Hence, for any $z \in \mathbb{C}$ there exists $y = [y_1, y_2, ..., y_n]^T > 0$, such that $\nu_A(z) = (Q_A(z)y)_j/y_j$ (Theorem 1.3.4), for all $j \in N$. If a_{ii} is an arbitrary diagonal element of A, $i \in N$, and $x = [x_1, x_2, ..., x_n]^T > 0$, is an associated vector for which $\nu_A(a_{ii}) = (Q_A(a_{ii})x)_i/x_i$, we obtain:

$$\nu_A(a_{ii}) = r_i^x(A) - |a_{ii} - a_{ii}| = r_i^x(A).$$

So, $\nu_A(a_{ii}) > 0$ ($r_i^x(A) > 0$ because A is irreducible).

Now, we consider a semi-infinite complex line $a_{ii} + te^{\mathrm{i}\theta}$ which begins in a_{ii}, for $t \geq 0$ and fixed $\theta \in [0, 2\pi)$. For $t = 0$, $\nu_A(a_{ii}) > 0$ and $\nu_A(z) < 0$ for $z \notin \Gamma^{\mathcal{R}}(A)$. Because of a continuity of the function ν_A, there is the smallest number $\widehat{\rho}_i(\theta) > 0$ such that $a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$. Therefore, each line segment

that joins a diagonal element of a_{ii} and point $a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$ is a subset of $\Gamma^{\mathcal{R}}(A)$. As it is true for each $\theta \in [0, 2\pi)$, we have

$$\bigcup_{\theta=0}^{2\pi} [a_{ii} + te^{\mathrm{i}\theta}]_{t=0}^{\widehat{\rho}_i(\theta)} \subseteq \Gamma^{\mathcal{R}}(A),$$

i.e., the set $\bigcup_{\theta=0}^{2\pi} [a_{ii} + te^{\mathrm{i}\theta}]_{t=0}^{\widehat{\rho}_i(\theta)}$ is a star-shaped subset of the minimal Geršgorin set with a respect to $a_{ii},\ i \in N$. ❑

According to the previous theorem, we can start with an arbitrary diagonal entry a_{ii}, choose an arbitrary angle θ and search for the boundary point of the form $a_{ii} + te^{\mathrm{i}\theta}$ for $t \geq 0$. The use of the bisection method is grounded due to the following property of ν_A.

Theorem 2.2.3 *For any $A \in \mathbb{C}^{n,n}$, the function ν_A is uniformly continuous on $\mathbb{C}$, i.e.,*

$$|\nu_A(z) - \nu_A(z')| \leq |z - z'|,$$

for all z and z' in $\mathbb{C}$.

Proof: For given $z, z' \in \mathbb{C}$, without loss of generality, let assume $\nu_A(z) \geq \nu_A(z')$. From (1.21), it follows that there exist positive vectors $x_z, x_{z'} \in \mathbb{R}^n$ and $j, k \in N$ such that:
$|\nu_A(z) - \nu_A(z')| = \nu_A(z) - \nu_A(z') =$
$\max_{i \in N}(r_i^{x_z}(A) - |z - a_{ii}|) - \max_{i \in N}(r_i^{x_{z'}}(A) - |z' - a_{ii}|) \leq$
$\max_{i \in N}(r_i^{x_{z'}}(A) - |z - a_{ii}|) - \max_{i \in N}(r_i^{x_{z'}}(A) - |z' - a_{ii}|) =$
$(r_j^{x_{z'}}(A) - |z - a_{jj}|) - (r_k^{x_{z'}}(A) - |z' - a_{kk}|) \leq$
$(r_j^{x_{z'}}(A) - |z - a_{jj}|) - (r_j^{x_{z'}}(A) - |z' - a_{jj}|) =$
$|z' - a_{jj}| - |z - a_{jj}| \leq |z' - a_{jj} - z + a_{jj}| = |z - z'|.$ ❑

Given an irreducible matrix $A \in \mathbb{C}^{n,n}$ and its diagonal element $a_{ii},\ i \in N$. From the uniform continuity of ν_A, we have:

$$|\nu_A(z) - \nu_A(z')| \leq |z - z'|.$$

Replacing z and z' with $a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$ and a_{ii}, respectively, and using the fact that $\nu_A(a_{ii}) > 0$ (Theorem 2.2.2), we obtain:

$$|\nu_A(a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta}) - \nu_A(a_{ii})| \leq |a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta} - a_{ii}|,$$

i.e.,

$$0 < \nu_A(a_{ii}) \leq \widehat{\rho}_i(\theta).$$

Therefore, it is possible to find numbers $\delta > 0$ and $l_{ik} \in \mathbb{N}$, $k \in \{1, 2, ..., m\},\ m \in \mathbb{N}$ such that

$$\widehat{\rho}_i(\theta) \in [\nu_A(a_{ii}) + (l_{ik} - 1)\delta, \nu_A(a_{ii}) + l_{ik}\delta],$$

where m is a given number of different angular directions. Using the bisection method and the fact that $\nu_A(z)$ is a uniformly continuous function, we can construct the procedure **bSearch** that produces points $\{\omega_{j,k}\}$ which lie on the boundary of the minimal Geršgorin set, $j \in N,\ k \in \{1, 2, ..., m\}$.

As it is numerically expensive to compute boundary points of $\Gamma^{\mathcal{R}}(A)$ using the bisection method, we construct an approximation of the minimal Geršgorin set. Namely, consider the sets:

$$\Gamma^{\omega_{j,k}}(A) := \bigcup_{i=1}^{n} \{z \in \mathbb{C} : |z - a_{ii}| \leq |\omega_{j,k} - a_{ii}|\}$$

and

$$\hat{\Gamma}^{\mathcal{R}}(A) := \Gamma(A) \cap \bigcap_{k=1}^{m} \bigcap_{j=1}^{n} \Gamma^{\omega_{j,k}}(A).$$

Using Olga Taussky's boundary result (Theorem 1.4.3), for a given $\omega_{j,k} \in \partial\Gamma^{\mathcal{R}}(A)$, there is $x = [x_1, x_2, ..., x_n]^T > 0$ such that $Q_A(\omega_{j,k})x = 0$. Equivalently, for all $i \in N$,

$$|\omega_{j,k} - a_{ii}| = \sum_{k \neq i}^{n} \frac{|a_{ik}|x_k}{x_i}.$$

It is interesting that knowing the boundary point $\omega_{j,k}$, all discs in $\Gamma^{\omega_{j,k}}(A)$ with centers in a_{ii} can be directly drawn, without knowing the components of x. So, their intersection gives the approximation of $\Gamma^{\mathcal{R}}(A)$, i.e.,

$$\Gamma^{\mathcal{R}}(A) \subseteq \hat{\Gamma}^{\mathcal{R}}(A) \subseteq \Gamma(A).$$

bSearch

Input: A, ξ, θ, tol

Set $z := \xi$ and compute $f := \nu_A(z)$ as the rightmost eigenvalue of $Q_A(z)$;
Set $\delta := f$ and $l := 0$;
while $f > 0$ **do**
  $l := l + 1$;
  Set $z := \xi + l\delta e^{\mathrm{i}\theta}$ and compute $f := \nu_A(z)$ as the rightmost eigenvalue of $Q_A(z)$;
end while
Set $a := (l-1)\delta$ and $b := l\delta$;
while $b - a > tol$ **do**
  Set $z := \xi + \frac{a+b}{2} e^{\mathrm{i}\theta}$ and compute $f := \nu_A(z)$ as the rightmost eigenvalue of $Q_A(z)$;
  if $f > 0$ **then**
    $a := \frac{a+b}{2}$;
  else
    $b := \frac{a+b}{2}$;
  end if
end while
Set $\omega = \xi + \frac{a+b}{2} e^{\mathrm{i}\theta}$;

Output: ω

Now, we construct the procedure **gerApprox** in order to obtain boundary of the set $\hat{\Gamma}^{\mathcal{R}}(A)$. Apart from the sequence of points $\{\omega_{j,k}\}$, the procedure **gerApprox** uses also parameter $n_g \in \mathbb{N}$ as an input. In that way, the approximation of $\partial\Gamma^{\mathcal{R}}(A)$ is constructed on an equally spaced rectangular grid.

Furthermore, we present the algorithm **bMGS**. For parameters $A \in \mathbb{C}^{n,n}$, m, $n_g \in \mathbb{N}$ and a given tolerance $tol > 0$, this algorithm computes $m \cdot n$ points on the boundary of $\partial\Gamma^{\mathcal{R}}(A)$.

The algorithm **bMGS** has several drawbacks. The main drawback is a usage of the bisection method on $m \cdot n$ points, resulting with a large number of eigenvalues computations even for matrices of medium sizes. In addition, it needs pre-computations in order to find a proper starting bisection interval. In some cases, if the parameter δ is not chosen in a proper way, the first zero of ν_A can

gerApprox

Input: $A, n_g, \{\omega_{j,k}\}_{j\in N,\ k\in\{1,2,\ldots,m\}}$

Compute l_{re}, u_{re}, l_{im} and u_{im} using (2.1) and (2.2);
Set $\delta_x = \frac{u_{re}-l_{re}}{n_g}$ and $\delta_y = \frac{u_{im}-l_{im}}{n_g}$;
for $k_x = 0 : n_g$ **do**
  for $k_y = 0 : n_g$ **do**
    Set $z := (l_{re} + k_x\delta_x) + \mathrm{i}(l_{im} + k_y\delta_y)$;
    Compute $G(k_x, k_y) := \min_{i\in N}\{|z - a_{ii}| - r_i(A)\}$;
    for $j = 1 : n$ and $k = 1 : m$ **do**
      Set $G_{jk}(k_x, k_y) = |z - a_{11}| - |\omega_{j,k} - a_{11}|$;
      for $i = 2 : n$ **do**
        Update $G_{jk}(k_x, k_y) = \min\{G_{jk}(k_x, k_y), |z - a_{ii}| - |\omega_{j,k} - a_{ii}|\}$;
      end for
      Update $G(k_x, k_y) = \max\{G(k_x, k_y), G_{jk}(k_x, k_y)\}$;
    end for
  end for
end for
Compute the zero level curve $\mathcal{C}$ of G;

Output: $\mathcal{C}$

Algorithm bMGS

Input: A, n_g, m, tol

for $j = 1 : n$ **do**
  for $k = 1 : m$ **do**
    Set $\theta_k = k\frac{2\pi}{m}$;
    Run the procedure **bSearch**$(A, a_{jj}, \theta_k, tol)$ to compute $\{\omega_{j,k}\}_{j\in N,\ k\in\{1,2,\ldots,m\}}$;
  end for
end for
Run the procedure **gerApprox**$(A, n_g, \{\omega_{j,k}\}_{j\in N,\ k\in\{1,2,\ldots,m\}})$;

Output: $\mathcal{C}$

be skipped on the ray $\{a_{ii} + te^{\mathrm{i}\theta}\},\ t > 0,\ i \in N$.

Finally, notice that the original version of this algorithm ([69]) was used only for irreducible matrices although it can be adaptable for reducible matrices, too

(Lemma 1.5.5). Also, note that the approximation can be improved, at the expense of more computations, if we use the Brauer set instead of the Geršgorin set for boundary points.

2.3 Explicit vs. implicit characterization of the minimal Geršgorin set

In this section, two characterizations of the minimal Geršgorin set are presented: explicit and implicit. The explicit characterization is given in [40], while the implicit characterization is a new result which is the basis for the algorithms applicable to large sparse matrices ([49]).

2.3.1 Explicit characterization

According to Theorems 1.5.1 and 1.5.2, the minimal Geršgorin set of a given matrix $A \in \mathbb{C}^{n,n}$ is completely characterized by:

$$z \in \Gamma^{\mathcal{R}}(A) \Longleftrightarrow \nu_A(z) \geq 0,$$

while its boundary is characterized by:

$$z \in \partial\Gamma^{\mathcal{R}}(A) \Longleftrightarrow \begin{cases} \nu_A(z) = 0 \text{ and there exists a sequence } \{z_j\} \text{ such that} \\ \lim\limits_{j\to\infty} z_j = z \text{ and } \nu_A(z_j) < 0,\ j \in \mathbb{N}. \end{cases}$$

Using the fact that the boundary points of $\Gamma^{\mathcal{R}}(A)$ are of the form $a_{ii} + \widehat{\rho}_i(\theta)e^{\mathrm{i}\theta}$, here, for a given $A \in \mathbb{C}^{n,n}, \xi \in \mathbb{C}$ and $\theta \in [0, 2\pi)$, we define a function $f_A^{\xi,\theta} : [0, \infty) \to \mathbb{R}$ by:

$$f_A^{\xi,\theta}(t) := \nu_A(\xi + te^{\mathrm{i}\theta}). \tag{2.6}$$

Using this function, we have the following theorem whose proof follows directly from Theorems 1.5.1 and 1.5.2.

Theorem 2.3.1 *Given an arbitrary irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, a complex number ξ, a real number $\theta \in [0, 2\pi)$ and the function $f_A^{\xi,\theta}$ defined by (2.6). Then,*

1. *$\xi + te^{\mathrm{i}\theta} \in \Gamma^{\mathcal{R}}(A)$ if and only if $f_A^{\xi,\theta}(t) \geq 0$;*

2. $\xi + te^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$ *if and only if* $f_A^{\xi,\theta}(t) = 0$ *and for* $\epsilon > 0$ *there exists* $\theta_\epsilon \in [0, 2\pi)$ *such that* $f_A^{\xi+te^{\mathrm{i}\theta},\theta_\epsilon}(\epsilon) < 0$.

The statement of the previous theorem we will call the explicit characterization.

The usefulness of the function $f_A^{\xi,\theta}$ lies in the fact that it is differentiable everywhere except at the most n points. Moreover, one can obtain its first order derivative, as it is expressed in the following theorem.

Theorem 2.3.2 *Given an arbitrary irreducible matrix* $A = [a_{ij}] \in \mathbb{C}^{n,n}$, *a complex number* ξ, *a real number* $\theta \in [0, 2\pi)$ *and the function* $f_A^{\xi,\theta}$ *defined by (2.6). Then,* $f_A^{\xi,\theta}$ *is* $\infty-$*differentiable in* $t \notin \{(\xi - a_{ii})e^{\mathrm{i}(\pi-\theta)} : i \in N\}$ *and its first derivative is given as:*

$$\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) = -\frac{u(t)^T D_A^{\xi,\theta}(t) v(t)}{u(t)^T v(t)}, \tag{2.7}$$

where $v(t)$ *and* $u(t)$ *are right and left eigenvectors of* $Q_A(\xi+te^{\mathrm{i}\theta})$ *corresponding to the eigenvalue* $f_A^{\xi,\theta}(t)$, *respectively, where diagonal matrix* $D_A^{\xi,\theta}(t) := diag([d_1(t), d_2(t), \ldots, d_n(t)])$ *is given by:*

$$d_i(t) := \frac{Re[(\xi - a_{ii})\, e^{-\mathrm{i}\theta}] + t}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|}, i \in N. \tag{2.8}$$

Proof: As A is irreducible, it follows that $Q_A(\xi + te^{\mathrm{i}\theta})$ is also an irreducible matrix. Then, $f_A^{\xi,\theta}(t)$ is a simple eigenvalue of $Q_A(\xi + te^{\mathrm{i}\theta})$ and denote its corresponding right and left eigenvectors by $v(t)$ and $u(t)$, respectively. Using a well-known result on differentiability of simple eigenvalues (Theorem 1.2.1), we obtain:

$$\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) = \frac{u(t)^T \frac{\partial}{\partial t} Q_A(\xi + te^{\mathrm{i}\theta}) v(t)}{u(t)^T v(t)}, \tag{2.9}$$

i.e., $f_A^{\xi,\theta}$ is a ∞-differentiable in t, for $t \neq (\xi - a_{ii})e^{\mathrm{i}(\pi-\theta)}, i \in N$.
From $q_{ii}(\xi+te^{\mathrm{i}\theta}) = -|\xi+te^{\mathrm{i}\theta} - a_{ii}| = -|e^{\mathrm{i}\theta}||(\xi-a_{ii})e^{-\mathrm{i}\theta}+t| = -|(\xi-a_{ii})e^{-\mathrm{i}\theta}+t|$ and $q_{ij}(\xi + te^{\mathrm{i}\theta}) = |a_{ij}|$, we have:

$$\frac{\partial}{\partial t} q_{ii}(\xi + te^{\mathrm{i}\theta}) = -\frac{\mathrm{Re}[(\xi - a_{ii})\, e^{-\mathrm{i}\theta}] + t}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|} \text{ and}$$

$$\frac{\partial}{\partial t} q_{ij}(\xi + te^{\mathrm{i}\theta}) = 0,$$

for $t \neq (\xi - a_{ii})e^{\mathrm{i}(\pi-\theta)}$, $i, j \in N,\ i \neq j$.
Therefore, $\frac{\partial}{\partial t} Q_A(\xi + te^{\mathrm{i}\theta}) = -D_A^{\xi,\theta}(t)$ and from (2.9) it follows:

$$\frac{\partial}{\partial t} f_A^{\xi,\theta}(t) = -\frac{u(t)^T D_A^{\xi,\theta}(t) v(t)}{u(t)^T v(t)}.$$

❑

2.3.2 Implicit characterization

Now, we present a novel characterization of $\Gamma^{\mathcal{R}}(A)$ that is the result of this thesis, published in [49]. First, we define functions $g_A^{\xi,\theta}$ and $h_A^{\xi,\theta}$ that will play role of $f_A^{\xi,\theta}$.

Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ and a real number $\theta \in [0, 2\pi)$, let us fix a vector $c \in \mathbb{R}^n$, $c > 0$, and for every $t \geq 0$ construct a system of linear equations

$$\underbrace{\begin{bmatrix} -Q_A(\xi + t\,e^{\mathrm{i}\theta}) & -c \\ -c^T & 0 \end{bmatrix}}_{M_A^{\xi,\theta}(t)} \begin{bmatrix} w_A^{\xi,\theta}(t) \\ g_A^{\xi,\theta}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \tag{2.10}$$

Assuming that $M_A^{\xi,\theta}(t)$ is nonsingular, (2.10) can be uniquely solved, and the Cramer's rule provides that

$$g_A^{\xi,\theta}(t) := -\frac{\det(-Q_A(\xi + t\,e^{\mathrm{i}\theta}))}{\det(M_A^{\xi,\theta}(t))}. \tag{2.11}$$

The defined function becomes zero whenever matrix $Q_A(z)$ becomes singular in the point $z = \xi + t\,e^{\mathrm{i}\theta}$. In the following, we see how $g_A^{\xi,\theta}$ can be used instead of a ν_A to characterize the boundary of the minimal Geršgorin set.

Next, we will use well-known properties of the Schur complement. Namely, if a matrix M is partitioned into four blocks, i.e., $M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, where matrices

A and D are square and A and $E := D - CA^{-1}B$ are nonsingular, using the Schur complement, we have:

$$M^{-1} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}BE^{-1}CA^{-1} & -A^{-1}BE^{-1} \\ -E^{-1}CA^{-1} & E^{-1} \end{bmatrix}, \tag{2.12}$$

and

$$\det(M) = \det(A)\det(E)\,. \tag{2.13}$$

Theorem 2.3.3 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ, a real $0 \leq \theta < 2\pi$ and $c \in \mathbb{R}^n$ arbitrary positive vector, let $\widehat{t} > 0$ be maximal such that $\xi + t\,e^{\mathrm{i}\theta} \in \Gamma^{\mathcal{R}}(A)$, for all $t \in [0, \widehat{t}]$. Then, there exists $\varepsilon > 0$ such that $M_A^{\xi,\theta}(t)$ is a nonsingular matrix for all $t \in [\widehat{t} - \varepsilon, \widehat{t} + \varepsilon]$. Consequently, (2.10) defines an ∞-differentiable functions $g_A^{\xi,\theta}$ and $w_A^{\xi,\theta}$ on $[\widehat{t} - \varepsilon, \widehat{t} + \varepsilon]$.*

Proof: First let us show that the matrix $M_A^{\xi,\theta}(\widehat{t})$ is nonsingular. Assume that $M_A^{\xi,\theta}(\widehat{t})$ is singular. As the off-diagonal zero pattern is the same for A and $Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta})$ and $c > 0$, the irreducibility of A implies the irreducibility of matrices $Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta})$ and $M_A^{\xi,\theta}(\widehat{t})$. Let $[w\ \alpha]^T$ be the right eigenvector of the matrix $M_A^{\xi,\theta}(\widehat{t})$ corresponding to the zero eigenvalue, i.e.,

$$\begin{bmatrix} -Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta}) & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} w \\ \alpha \end{bmatrix} = 0 \begin{bmatrix} w \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.14}$$

Hence, we obtain

$$Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta})w + c\alpha = 0 \tag{2.15}$$

and

$$c^T w = 0. \tag{2.16}$$

Using the Perron-Frobenius theorem for essentially non-negative irreducible matrices (Theorem 1.3.4), an eigenvalue $\nu_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta}) = 0$ of the matrix $Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta})$ has a positive right and left eigenvector $\widehat{x}$ and $\widehat{y}$, respectively. Moreover, every right (left) eigenvector corresponding to the eigenvalue zero will be a scalar multiple of $\widehat{x}$ ($\widehat{y}$). Multiplying equation (2.15) by $\widehat{y}^T$, we obtain

$$\widehat{y}^T Q_A(\xi + \widehat{t}\,e^{\mathrm{i}\theta})w + \widehat{y}^T c\alpha = 0 \implies \widehat{y}^T c\alpha = 0 \implies \alpha = 0, \tag{2.17}$$

which together with (2.15) implies that $Q_A(\xi + \widehat{t}\, e^{\mathrm{i}\theta})w = 0$, i.e., $w \neq 0$ is the right eigenvector corresponding to the eigenvalue zero. Hence, there exists $\beta \neq 0$ that $w = \beta\widehat{x}$. Then, from (2.16), we obtain

$$\beta c^T \widehat{x} = 0 \implies c^T \widehat{x} = 0, \tag{2.18}$$

which is a contradiction. Therefore, $M_A^{\xi,\theta}(\widehat{t})$ has to be a nonsingular matrix. Moreover, using the continuity of $M_A^{\xi,\theta}(t)$ in parameter t, we conclude that there exists a sufficiently small $\varepsilon > 0$ such that for all $t \in [\widehat{t} - \varepsilon, \widehat{t} + \varepsilon]$, $M_A^{\xi,\theta}(t)$ is nonsingular and $g_A^{\xi,\theta}(t)$ and $w_A^{\xi,\theta}$ are ∞-differentiable functions for all $t \in [\widehat{t} - \varepsilon, \widehat{t} + \varepsilon]$. ❑

Theorem 2.3.4 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ, a real $0 \leq \theta < 2\pi$ and $c \in \mathbb{R}^n$ arbitrary positive vector, let $\widehat{t} > 0$ be maximal such that $\xi + t\, e^{\mathrm{i}\theta} \in \Gamma^{\mathcal{R}}(A)$, for all $t \in [0, \widehat{t}]$. Then, there exists $\varepsilon > 0$ such that:*

1. *for every $z = \xi + t\, e^{\mathrm{i}\theta} \notin \Gamma^{\mathcal{R}}(A)$, $g_A^{\xi,\theta}(t)$ and $w_A^{\xi,\theta}(t)$ are well defined and positive;*
2. *$g_A^{\xi,\theta}(t) > 0$ for all $t \in (\widehat{t}, \widehat{t} + \varepsilon]$;*
3. *$g_A^{\xi,\theta}(\widehat{t}) = 0$;*
4. *$g_A^{\xi,\theta}(t) < 0$ for all $t \in [\widehat{t} - \varepsilon, \widehat{t})$;*
5. *the first and the second derivatives $\frac{\partial}{\partial t} g_A^{\xi,\theta}$, $\frac{\partial}{\partial t} w_A^{\xi,\theta}$ and $\frac{\partial^2}{\partial t^2} g_A^{\xi,\theta}$, $\frac{\partial^2}{\partial t^2} w_A^{\xi,\theta}$ are defined via the linear systems:*

$$\begin{bmatrix} -Q_A(\xi + t\, e^{\mathrm{i}\theta}) & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial t} w_A^{\xi,\theta}(t) \\ \frac{\partial}{\partial t} g_A^{\xi,\theta}(t) \end{bmatrix} = \begin{bmatrix} -D_A^{\xi,\theta}(t) w_A^{\xi,\theta}(t) \\ 0 \end{bmatrix}, \tag{2.19}$$

$$\begin{bmatrix} -Q_A(\xi + t\, e^{\mathrm{i}\theta}) & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} \frac{\partial^2}{\partial t^2} w_A^{\xi,\theta}(t) \\ \frac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(t) \end{bmatrix} = \begin{bmatrix} -S_A^{\xi,\theta}(t) w_A^{\xi,\theta}(t) - 2D_A^{\xi,\theta}(t) \frac{\partial}{\partial t} w_A^{\xi,\theta}(t) \\ 0 \end{bmatrix}, \tag{2.20}$$

where the vector function $w_A^{\xi,\theta}(t)$ is defined as the solution of the linear system (2.10), $D_A^{\xi,\theta}(t) := diag([d_1(t), d_2(t), \ldots, d_n(t)])$ and $S_A^{\xi,\theta}(t) := diag([s_1(t), s_2(t), \ldots, s_n(t)])$, where:

$$d_i(t) := \frac{Re[(\xi - a_{ii})\, e^{-\mathrm{i}\theta}] + t}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|}, \tag{2.21}$$

and

$$s_i(t) := \frac{(Im[(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t])^2}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|^3}, \tag{2.22}$$

for $t \neq (\xi - a_{ii})e^{\mathrm{i}(\pi-\theta)}$, $i \in N$.

Proof: 1. Let $z = \xi + t\, e^{\mathrm{i}\theta} \notin \Gamma^{\mathcal{R}}(A)$, then $-Q_A(\xi + t\, e^{\mathrm{i}\theta})$ is a nonsingular M-matrix, implying $(-Q_A(\xi + t\, e^{\mathrm{i}\theta}))^{-1} \geq 0$ and $\det(-Q_A(\xi + t\, e^{\mathrm{i}\theta})) > 0$. So, from (2.13),

$$\det(M_A^{\xi,\theta}(t)) = -\det(-Q_A(\xi + t\, e^{\mathrm{i}\theta}))c^T(-Q_A(\xi + t\, e^{\mathrm{i}\theta}))^{-1}c < 0 \tag{2.23}$$

and we obtain that $M_A^{\xi,\theta}(t)$ is nonsingular. Therefore, using a formula for the inversion of a block matrix (2.12), we have that

$$(M_A^{\xi,\theta}(t))^{-1} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} (-Q_A(\xi + t\, e^{\mathrm{i}\theta}))^{-1}c(c^T(-Q_A(\xi + t\, e^{\mathrm{i}\theta}))^{-1}c)^{-1} \\ (c^T(-Q_A(\xi + t\, e^{\mathrm{i}\theta}))^{-1}c)^{-1} \end{bmatrix} > 0,$$

implying that $g_A^{\xi,\theta}(t)$ and $w_A^{\xi,\theta}(t)$ are well-defined and positive.

Items 2., 3. and 4. follow from the continuity of $\det(M_A^{\xi,\theta}(t))$ and the fact that $\det(-Q_A(\xi + \widehat{t}\, e^{\mathrm{i}\theta})) = 0$.

5. Finally, we get the expressions for the derivatives. If $t \neq (\xi - a_{ii})e^{\mathrm{i}(\pi-\theta)}$ for $i \in N$, the entries of $Q_A(\xi + t\, e^{\mathrm{i}\theta})$ are ∞-differentiable functions in t and their first and second derivatives are given as:

$$\frac{\partial}{\partial t} q_{ii}(\xi + t\, e^{\mathrm{i}\theta}) = \frac{Re[(\xi - a_{ii})\, e^{-\mathrm{i}\theta}] + t}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|}, \quad \frac{\partial}{\partial t} q_{ij}(\xi + t\, e^{\mathrm{i}\theta}) = 0, \quad \text{for} \quad j \neq i, \tag{2.24}$$

and

$$\frac{\partial^2}{\partial t^2} q_{ii}(\xi + t\, e^{\mathrm{i}\theta}) = \frac{(Im[(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t])^2}{|(\xi - a_{ii})\, e^{-\mathrm{i}\theta} + t|^3}, \quad \frac{\partial^2}{\partial t^2} q_{ij}(\xi + t\, e^{\mathrm{i}\theta}) = 0, \quad \text{for} \quad j \neq i, \tag{2.25}$$

for all $i, j \in N$. By differentiating (2.10) and using (2.24) and (2.25), we obtain (2.19) and (2.20). ❑

Definition 2.3.5 *For a fixed* $\xi \in \mathbb{C}$ *and* $0 \leq \theta < 2\pi$, *for* $t \geq 0$, *define the functions:*

$$\chi_A^{\xi,\theta}(t) := \min\left\{(w_A^{\xi,\theta}(t))_i \,:\, 1 \leq i \leq n\right\}$$

and

$$h_A^{\xi,\theta}(t) := \min\left\{g_A^{\xi,\theta}(t), \chi_A^{\xi,\theta}(t)\right\}. \tag{2.26}$$

Then, the following characterization of the minimal Geršgorin set holds.

Theorem 2.3.6 *Let be given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$ and its arbitrary diagonal entry $\xi = a_{kk}, k \in N$ and $z = \xi + t\,e^{\mathrm{i}\theta}$, where $t \geq 0$ and $0 \leq \theta < 2\pi$. Then,*

- *$z \notin \Gamma^{\mathcal{R}}(A)$ if and only if $h_A^{\xi,\theta}(t) > 0$;*
- *$z \in \partial\Gamma^{\mathcal{R}}(A)$ such that $t = \widehat{\rho}_k(\theta)$ if and only if*

  i) $h_A^{\xi,\theta}(t) = g_A^{\xi,\theta}(t) = 0$,

  ii) $h_A^{\xi,\theta}(s_1) \leq 0$ holds for all $0 \leq s_1 \leq t$, and

  iii) for every $\varepsilon > 0$ there exists $s_2 \geq t$ such that $s_2 - t < \varepsilon$ and $h_A^{\xi,\theta}(s_2) > 0$;

- *if c is chosen to be a positive normalized ($\|c\|_2 = 1$) eigenvector of the Perron eigenvalue $\nu_A(\xi) > 0$ of $Q_A(\xi)$ and $M_A^{\xi,\theta}(0)$ is a nonsingular matrix, then $g_A^{\xi,\theta}(0) < 0$ and $w_A^{\xi,\theta}(0) > 0$.*

Proof: For the first item, we prove the equivalence. Assume that $z \notin \Gamma^{\mathcal{R}}(A)$, then, as it is shown in the item 1. of Theorem 2.3.4, $h_A^{\xi,\theta}(t) > 0$. On the other hand, assume that $h_A^{\xi,\theta}(t) > 0$, then from the system (2.10), we get $-Q_A(z)w_A^{\xi,\theta}(t) = g_A^{\xi,\theta}(t)c > 0$, while $w_A^{\xi,\theta}(t) > 0$. However, this implies that $-Q_A(z)$ is a nonsingular M-matrix. So, $z \notin \Gamma^{\mathcal{R}}(A)$.

For the second item, first observe that $i) - iii)$ imply that $t = \widehat{\rho}_k(\theta)$ as defined in Theorem 2.2.2. So, assume $z \in \partial\Gamma^{\mathcal{R}}(A)$ such that $t = \widehat{\rho}_k(\theta)$ and let $\varepsilon > 0$ and $0 \leq s_1 \leq t \leq s_2$ and $s_2 - t < \varepsilon$. Then Theorem 2.3.4, item 3. gives $g_A^{\xi,\theta}(t) = 0$. Item 1. states that $w_A^{\xi,\theta}(s_2) > 0$, which with continuity implies $w_A^{\xi,\theta}(t) \geq 0$. So, we conclude $h_A^{\xi,\theta}(t) = 0$. Obviously, $ii)$ follows from the previous item and $iii)$ from the definition of $\widehat{\rho}_k(\theta)$.

For the third item, if c is a positive normalized eigenvector of the Perron eigenvalue $\nu_A(\xi) > 0$ of $Q_A(\xi)$, then $Q_A(\xi)c = \nu_A(\xi)c$ and $c^T c = 1$. From the system (2.10) for $t = 0$, we obtain $-Q_A(\xi)w_A^{\xi,\theta}(0) = g_A^{\xi,\theta}(0)c$ and $-c^T w_A^{\xi,\theta}(0) = -1$. Finally, we get $g_A^{\xi,\theta}(0) = -\nu_A(\xi) < 0$ and $w_A^{\xi,\theta}(0) = c > 0$ because of the nonsingularity of $M_A^{\xi,\theta}(0)$. ❑

The statement of the previous theorem we will call the implicit characterization of the minimal Geršgorin set.

2.4 Curve tracing approaches

In this section, three different curve tracing approaches are presented. Combining them with explicit and implicit characterization of the minimal Geršgorin set, six algorithms are developed: **eMGSs**, **iMGSs**, **eMGSp**, **iMGSp**, **eMGSt** and **iMGSt** (Figure 2.1).

Figure 2.1: Algorithms for computing MGS.

All presented algorithms are performed for irreducible matrices. However, this is not a setback (theoretically) for their application to reducible matrices (Lemma 1.5.5).

2.4.1 Star-shaped curve tracing

In this subsection, two algorithms for computing the minimal Geršgorin set are presented- **eMGSs** and **iMGSs**. They are based on the computation of points that lie on the boundary of the minimal Geršgorin set using Newton-like method with a given accuracy, function $f_A^{\xi,\theta}$ (2.6) for **eMGSs** and functions $g_A^{\xi,\theta}$ (2.11) and $h_A^{\xi,\theta}$ (2.26) for **iMGSs**.

First, we describe construction of the explicit algorithm (**eMGSs**). It is developed in [40] and is based on the modified Newton's method.

Using the previous results (Theorem 2.2.2 and Lemma 1.5.5), we can construct the minimal Geršgorin set starting from each diagonal entry of a given matrix A and computing points on its boundary for m angles $\theta,\ m \in \mathbb{N},\ \theta \in [0, 2\pi)$.

Now, we formulate the modified Newton's method for computing zero of the function $f_A^{\xi,\theta}$ defined as (2.6). Define $t_0 := f_A^{\xi,\theta}(0) = \nu_A(\xi) > 0$ and the sequence $\{t_k\},\ k \in \mathbb{N}_0$, with

$$t_{k+1} := t_k + \gamma_k \Delta_k, \tag{2.27}$$

where Δ_k is defined as

$$\Delta_k := \begin{cases} -\dfrac{f_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}, & \text{if } \frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k) < 0 \\ \\ f_A^{\xi,\theta}(t_k), & \text{otherwise} \end{cases}, \tag{2.28}$$

and

$$\gamma_k := \begin{cases} 1, & \text{if } f_A^{\xi,\theta}(t_{k+1}) \geq 0 \\ \\ \tau^{q_k}, & \text{otherwise} \end{cases},$$

with $\tau \in (0, 1)$ and the smallest $q_k \in \mathbb{N}$ such that

$$f_A^{\xi,\theta}(t_k + \tau^{q_k}\, \Delta_k) > 0 \quad \text{and} \quad f_A^{\xi,\theta}(t_k + \tau^{q_k - 1}\, \Delta_k) < 0. \tag{2.29}$$

Lemma 2.4.1 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ and a real $0 \leq \theta < 2\pi$. The sequence $\{t_k\},\ k \in \mathbb{N}_0$, is well defined by (2.27), and it converges to $\widehat{t} > 0$ such that $f_A^{\xi,\theta}(\widehat{t}\,) = 0$.*

Proof: Firstly, observe that the uniform continuity of ν_A (2.2.3) implies that for every $t > 0$ and $\varepsilon > 0$

$$|f_A^{\xi,\theta}(t + \varepsilon) - f_A^{\xi,\theta}(t)| \leq \varepsilon, \tag{2.30}$$

and consequently, whenever $f_A^{\xi,\theta}$ is differentiable such that

$$\left| \frac{\partial}{\partial t} f_A^{\xi,\theta}(t) \right| \leq 1. \tag{2.31}$$

Therefore, for all $k \in \mathbb{N}_0$, $\Delta_k \geq f_A^{\xi,\theta}(t_k)$.
Now, we show that the sequence $\{t_k\}$ is well-defined. Since $f_A^{\xi,\theta}(t_0) = \nu_A(\xi + \nu_A(\xi)\, e^{\mathrm{i}\theta})$,

$$|f_A^{\xi,\theta}(t_0) - f_A^{\xi,\theta}(0)| = |\nu_A(\xi + \nu_A(\xi)\, e^{\mathrm{i}\theta}) - \nu_A(\xi)| \leq |e^{\mathrm{i}\theta}\nu_A(\xi)| = \nu_A(\xi),$$

i.e.,

$$-\nu_A(\xi) \leq f_A^{\xi,\theta}(t_0) - f_A^{\xi,\theta}(0) \leq \nu_A(\xi),$$

it follows that $f_A^{\xi,\theta}(t_0) \geq 0$ and consequently, $\Delta_0 \geq 0$. Thus, $t_0 + \Delta_0 \geq t_0 > 0$, and $f_A^{\xi,\theta}(t_0 + \Delta_0)$ is well-defined. If $f_A^{\xi,\theta}(t_0 + \Delta_0) < 0$, the continuity of $f_A^{\xi,\theta}$, implies that there exists $q_0 \in \mathbb{N}$, such that $\gamma_0 = \tau^{q_0}$. Otherwise, if $f_A^{\xi,\theta}(t_0 + \Delta_0) \geq 0$, then $\gamma_0 = 1$. By this, we obtain $t_1 := t_0 + \gamma_0\Delta_0 \geq t_0 > 0$, such that $f_A^{\xi,\theta}(t_1) \geq 0$. Continuing in the same way, we obtain that the sequence $\{t_k\}$ is well-defined, monotonically nondecreasing and $f_A^{\xi,\theta}(t_k) \geq 0$, for all $k \in \mathbb{N}_0$.
Let's assume that $\{t_k\}$ is unbounded. Then, for some $\ell \in \mathbb{N}$, there exists a subsequence $\{t_{k_\ell}\}$, such that $\lim\limits_{\ell\to\infty} t_{k_\ell} = \infty$. From the construction of the sequence, we have that $f_A^{\xi,\theta}(t_{k_\ell}) \geq 0$. It follows that $z_\ell := \xi + t_{k_\ell}\, e^{\mathrm{i}\theta} \in \Gamma^{\mathcal{R}}(A)$, for all $\ell \in \mathbb{N}$, (Theorem 2.2.2). However, since $|z_\ell| \to \infty$, it is a contradiction because the minimal Geršgorin set is a compact set in $\mathbb{C}$. As a result, the sequence $\{t_k\}$ is convergent and we denote its limit as $\widehat{t} = \lim\limits_{k\to\infty} t_k$.
Now, we distinguish two cases.
First, assume that $\lim\limits_{k\to\infty} \inf\gamma_k > 0,\ k \in \mathbb{N}$. Then,

$$0 \leq f_A^{\xi,\theta}(t_k) \leq \Delta_k = \frac{t_{k+1} - t_k}{\gamma_k}.$$

Taking the limit for $k \to \infty$, it follows $f_A^{\xi,\theta}(\widehat{t}\,) = 0$.

Second, assume $\lim\limits_{\ell\to\infty} \gamma_{k_\ell} = 0$. Then, obviously, $\gamma_{k_\ell} = \tau^{q_{k_\ell}} < 1$, for $\ell \in \mathbb{N}$, and $\lim\limits_{\ell\to\infty} q_{k_\ell} = \infty$. It implies that $0 \leq f_A^{\xi,\theta}(\widehat{t}\,) \leq 0$, which completes the proof.

❑

Using the previous lemma, we construct the procedure **eSearch**.

Theorem 2.4.2 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ and real numbers $0 \leq \theta < 2\pi$ and $\tau \in (0, 1)$, the procedure* ***eSearch***

eSearch

Input: A, ξ, θ, tol

1: Set $z = \xi$ and compute r.m.e. (the rightmost eigenpair) (f, x, y) of $Q_A(z)$;
2: Set $t := f$, $\Delta = 2\,tol$ and $\beta = 0.9$;
3: **while** $\Delta > tol$ **do**
4: Compute $df := -\frac{y(t)^T \frac{\partial}{\partial t} Q_A(\xi + te^{\mathrm{i}\theta})x(t)}{y(t)^T x(t)}$;
5: **if** $df > 0$ **then**
6: Set $df := -1$;
7: **end if**
8: Set $\Delta := -f/df$ and $\gamma := 1$;
9: Set $z = \xi + (t + \gamma\Delta)e^{\mathrm{i}\theta}$ and compute r.m.e. (f, x, y) of $Q_A(z)$;
10: **while** $f < 0$ **do**
11: $\gamma := \gamma \cdot \beta$;
12: Set $z = \xi + (t + \gamma\Delta)e^{\mathrm{i}\theta}$ and compute r.m.e. (f, x, y) of $Q_A(z)$;
13: **end while**
14: $t := t + \gamma\Delta$;
15: **end while**
16: Set $z = \xi + (t + \gamma\Delta)e^{\mathrm{i}\theta}$ and compute r.m.e. (f, x, y) of $Q_A(z)$;
17: **if** $f > 0$ **then**
18: Set $t := t + 2\,tol$ and $\Delta := 2\,tol$;
19: **go to 3**
20: **end if**
21: Set $\omega = \xi + t\,e^{\mathrm{i}\theta}$;

Output: ω

produces $\omega = \xi + \widehat{t}\,e^{\mathrm{i}\theta} \in \mathbb{C}$ *such that* $\omega \in \partial\Gamma^{\mathcal{R}}(A)$. *Furthermore, if* $\dfrac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(\widehat{t}) > 0$, *the convergence is locally quadratic and otherwise, the convergence is linear with the convergence rate* $\lim\limits_{k\to\infty} \sup(1 - \tau^{q_k})$.

Proof: The lines (1–15) of the procedure **eSearch**, generate a sequence $\{t_k\}$, $k \in \mathbb{N}$, such that $\lim\limits_{t\to\infty} t_k = \hat{t}$ and $f(\hat{t}) = 0$. However, the obtained point $\xi + \hat{t}e^{\mathrm{i}\theta}$ does not have to lie on the boundary of the minimal Geršgorin set. In order to find a proper point, the lines $(16 - 20)$ are added in the procedure to check if $f_A^{\xi,\theta}$ changes the sign in $\hat{t}$. A repeated restart will occur until this condition is satisfied. In that way, we obtain the point $\omega = \xi + \hat{t}e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$. So, there exists

$\epsilon > 0$ such that $\dfrac{\partial}{\partial t} f_A^{\xi,\theta}(t) < 0$, for $t \in (\hat{t} - \epsilon, \hat{t})$. Consequently, there is $k_0 \in \mathbb{N}$, such that $\Delta_k = -\dfrac{f_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}$, for all $k \geq k_0$.

Now, we prove the rate of local convergence. There are two cases.

If $\dfrac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t) > 0$ for $t \in (\hat{t} - \epsilon, \hat{t})$ and $k \geq k_0$, then $f_A^{\xi,\theta}$ is locally convex function and $\gamma_k = 1$. It implies the quadratic convergence of modified Newton's method.

If $\dfrac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t) \leq 0$ for $t \in (\hat{t} - \epsilon, \hat{t})$ and $k \geq k_0$, from (2.27) and (2.28), we obtain:

$$\hat{t} - t_{k+1} = \hat{t} - t_k - \gamma_k \Delta_k = \hat{t} - t_k + \gamma_k \frac{f_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}.$$

Using quadratic Taylor polynomial for $f_A^{\xi,\theta}$, there exists $t^* \in (\hat{t} - t_k, \hat{t})$ such that

$$0 = f_A^{\xi,\theta}(\hat{t}) = f_A^{\xi,\theta}(t_k) + \frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)(\hat{t} - t_k) + \frac{1}{2} \frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t^*)(\hat{t} - t_k)^2.$$

So,

$$\hat{t} - t_{k+1} = \hat{t} - t_k - \gamma_k \frac{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)(\hat{t} - t_k) + \frac{1}{2} \frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t^*)(\hat{t} - t_k)^2}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)} =$$

$$(\hat{t} - t_k)(1 - \gamma_k - \frac{\frac{1}{2}\gamma_k \frac{\partial^2}{\partial t^2} f_A^{\xi,\theta}(t^*)(\hat{t} - t_k)}{\frac{\partial}{\partial t} f_A^{\xi,\theta}(t_k)}).$$

an consequently,

$$\lim_{k \to \infty} \sup \frac{\hat{t} - t_{k+1}}{\hat{t} - t_k} = \lim_{k \to \infty} \sup (1 - \tau^{q_k}).$$

❑

Now, we present the original results that are used for a construction of new algorithm for the numerical approximation of the minimal Geršgorin set. This new algorithm will be called an implicit one, abbreviated as **iMGSs** ([49]), since the main idea is to avoid explicit computation of the Perron eigentriplets within the algorithm **eMGSs**, by replacing the function $f_A^{\xi,\theta}$ with a new function $h_A^{\xi,\theta}$ that reveals the Perron eigenvalue implicitly through the solution of a structured

system of linear equations. The motivation for this approach to significantly reduce the overall number of expensive eigenvalue computations can be found in the idea of the implicit determinant method for pseudospectra given in [25].

Results of Theorems 2.3.4 and 2.3.6 are the basis for the procedure **iSearch** which is the core of the implicit algorithm for computing the minimal Geršgorin set.

We formulate the new modified Newton's method for the computing zeros of the function h_A. First, it will be defined the sequence $\{t_k\}$, with

$$t_{k+1} := t_k + \gamma_k \Delta_k,\ k \in \mathbb{N}_0, \tag{2.32}$$

where $t_0 := 0$ and Δ_k is given as

$$\Delta_k := \begin{cases} -\dfrac{g_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k)}, & \text{if } \frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k) > 0 \\ \\ \Delta, & \text{otherwise} \end{cases}, \tag{2.33}$$

where $\Delta > 0$ is given parameter and

$$\gamma_k := \begin{cases} 1, & \text{if } h_A^{\xi,\theta}(t_{k+1}) \leq 0 \\ \\ \tau^{q_k}, & \text{otherwise} \end{cases}, \tag{2.34}$$

with parameter $\tau \in (0,1)$ arbitrarily fixed and $q_k \in \mathbb{N}$ being the smallest number such that

$$h_A^{\xi,\theta}(t_k + \tau^{q_k}\,\Delta_k) < 0 \quad \text{and} \quad h_A^{\xi,\theta}(t_k + \tau^{q_k - 1}\,\Delta_k) > 0. \tag{2.35}$$

For c, we choose vector defined in the third item in Theorem 2.3.6. Additionally, if the convergence is achieved in $\widetilde{t}$, then we check if $h_A^{\xi,\theta}(\widetilde{t} + \varepsilon) > 0$ for a small tolerance $\varepsilon > 0$ and if not, we restart the sequence taking $t_0 := \widetilde{t} + \varepsilon$.

Theorem 2.4.3 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$, a complex number ξ and a real $0 \leq \theta < 2\pi$, a sequence $\{t_k\}_{k\in\mathbb{N}_0}$ defined by (2.32) is monotonically non-decreasing and it converges to $\widehat{t} > 0$, such that $\xi + \widehat{t}e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$. Furthermore, if $\dfrac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(\widehat{t}) > 0$, the convergence is locally quadratic and otherwise, the convergence is linear with the convergence rate $\lim\limits_{k\to\infty} \sup(1 - \tau^{q_k})$.*

Proof: First we show that the sequence $\{t_k\}$, $k \in \mathbb{N}_0$, is well-defined. From Theorem 2.3.6 follows $h_A^{\xi,\theta}(0) < 0$. From the definition of Δ_k, we have $\Delta_0 > 0$. Thus $t_0+\Delta_0 > 0$, and the continuity of $h_A^{\xi,\theta}$ together with (2.35) implies that there exists $0 < \gamma_0 \leq 1$ such that $h_A^{\xi,\theta}(t_0+\gamma_0\Delta_0) < 0$. So, we obtain $t_1 := t_0+\gamma_0\Delta_0 > t_0$, such that $h_A^{\xi,\theta}(t_1) < 0$. By induction, we obtain that the sequence $\{t_k\}_{k\in\mathbb{N}}$ is well-defined and that $h_A^{\xi,\theta}(t_k) < 0$ with $t_k > t_{k-1}$, for all $k \in \mathbb{N}$.

To prove the convergence of monotonically increasing sequence $\{t_k\}$, it is enough to show that it is bounded above. Let's assume that $\{t_k\}$ is unbounded. Then for some $m \in \mathbb{N}$, there exists a subsequence $\{t_{k_m}\}$, such that $\lim\limits_{m\to\infty} t_{k_m} = \infty$. Also, the fact $h_A^{\xi,\theta}(t_{k_m}) < 0$ implies that $z_m := \xi + t_{k_m}\, e^{\mathrm{i}\theta} \in \Gamma^{\mathcal{R}}(A)$, for all $m \in \mathbb{N}$ and $\lim\limits_{m\to\infty} |z_m| = \infty$. However, this is a contradiction because the minimal Geršgorin set is a compact set in $\mathbb{C}$. Therefore, the sequence is convergent and we denote its limit by $\widehat{t} = \lim\limits_{k\to\infty} t_k$.

From the construction of the sequence we have that $g_A^{\xi,\theta}(\widehat{t}) = 0$, when $\frac{\partial}{\partial t} g_A^{\xi,\theta}(\widehat{t})$ exists and it is positive, or that $h_A^{\xi,\theta}(\widehat{t}) = 0$, otherwise. Finally, due to restarts we obtain that $z = \xi + \widehat{t}e^{\mathrm{i}\theta}$ fulfills the second item of Theorem 2.3.6, and, therefore $z = \xi + \widehat{t}e^{\mathrm{i}\theta} \in \partial\Gamma^{\mathcal{R}}(A)$.

Now, we prove the rate of local convergence. There are two cases.

If $\dfrac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(\widehat{t}) > 0$, then $g_A^{\xi,\theta}$ is a locally convex function, and for sufficiently large $k \in \mathbb{N}$, $\gamma_k = 1$. This implies the quadratic convergence of modified Newton's method.

If $\dfrac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(\widehat{t}) \leq 0$, then for sufficiently large $k \in \mathbb{N}$, $\dfrac{\partial}{\partial t} g_A^{\xi,\theta}(t_k) > 0$, and, thus, from (2.32) and (2.33), we obtain:

$$\widehat{t} - t_{k+1} = \widehat{t} - t_k - \gamma_k\Delta_k = \widehat{t} - t_k + \gamma_k \frac{g_A^{\xi,\theta}(t_k)}{\frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k)}.$$

Using quadratic Taylor expansion for $g_A^{\xi,\theta}$, there exists $t^* \in (\widehat{t} - t_k, \widehat{t})$ such that

$$0 = g_A^{\xi,\theta}(\widehat{t}) = g_A^{\xi,\theta}(t_k) + \frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k)(\widehat{t} - t_k) + \frac{1}{2}\frac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(t^*)(\widehat{t} - t_k)^2.$$

So,

$$\widehat{t} - t_{k+1} = \hat{t} - t_k - \gamma_k \frac{\frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k)(\widehat{t} - t_k) + \frac{1}{2}\frac{\partial^2}{\partial t^2} g_A^{\xi,\theta}(t^*)(\widehat{t} - t_k)^2}{\frac{\partial}{\partial t} g_A^{\xi,\theta}(t_k)} =$$

$$(\widehat{t}-t_k)(1-\gamma_k-\frac{\frac{1}{2}\gamma_k\frac{\partial^2}{\partial t^2}g_A^{\xi,\theta}(t^*)(\widehat{t}-t_k)}{\frac{\partial}{\partial t}g_A^{\xi,\theta}(t_k)}),$$

and consequently,

$$\lim_{k\to\infty}\sup\frac{\widehat{t}-t_{k+1}}{\widehat{t}-t_k}=\lim_{k\to\infty}\sup(1-\tau^{q_k}).$$

❑

iSearch

Input: A, ξ, θ, tol

Set $z := \xi$ and compute the rightmost eigenpair $(\nu_A(z), c)$ of $Q_A(z)$;
Set $t := 0,\ \Delta := 2\,tol$ and $\beta = 0.9$;
while $\Delta > tol$ **do**
  Compute $g = g_A^{\xi,\theta}(t)$ by solving (2.10);
  Compute $dg = \frac{\partial}{\partial t}g_A^{\xi,\theta}(t)$ by solving (2.19);
  if $dg < 0$ **then**
    Set $dg := -g/\Delta$;
  end if
  Set $\Delta := -g/dg$ and $\gamma := 1$;
  Compute $h = h_A^{\xi,\theta}(t+\gamma\,\Delta)$ using (2.10);
  while $h > 0$ **do**
    $\gamma := \gamma \cdot \beta$;
    Compute $h = h_A^{\xi,\theta}(t+\gamma\,\Delta)$ using (2.10);
  end while
  Set $t := t + \gamma\,\Delta$;
end while
Compute $h = h_A^{\xi,\theta}(t+\gamma\,\Delta)$ using (2.10);
if $h < 0$ **then**
  Set $t := t + 2\,tol$ and $\Delta := 2\,tol$;
  go to 3
end if
Set $\omega = \xi + te^{\mathrm{i}\theta}$;

Output: ω

Finally, we present the algorithms ***MGSs**, where $* \in$ {**e,i**}, here and in the following in the thesis. For a given irreducible matrix A of a size $n \in \mathbb{N}$, denote its

set of different diagonal elements as $\mathcal{D} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$, where elements of $\mathcal{D}$ are sorted from the leftmost to the rightmost (i.e., $\mathrm{Re}(a_{i_1 i_1}) \leq \mathrm{Re}(a_{i_2 i_2}) \leq ... \leq \mathrm{Re}(a_{i_{\tilde{n}} i_{\tilde{n}}})$), $\tilde{n} \in N$. Also, we denote the points that represent the approximation of the points of the boundary of the minimal Geršgorin set with $\{\omega_{i,j}\}_{j=1}^{m_i},\ i \leq s,\ m_i \in \mathbb{N}$. The allowed maximal distance between that nodes and the boundary of the minimal Geršgorin set is less than ε_1 and the distance between successive points has to be less than ε_2, i.e., $\mathrm{dist}(\omega_{i,j}, \partial\Gamma^{\mathcal{R}}(A)) := \min\limits_{z \in \partial\Gamma^{\mathcal{R}}(A)} |\omega_{i,j} - z| < \varepsilon_1$, and $|\omega_{i,j} - \omega_{i,j+1}| < \varepsilon_2$, where $\omega_{i,m_i+1} := \omega_{i,1}$.

Starting with values $\xi := a_{i_1 i_1},\ \varphi_1 = -\pi$ and using the procedure ***Search**, we obtain the point $\omega_{1,1} := \xi + t_1 e^{\mathrm{i}\varphi_1}$, where $t_1 :=$ ***Search**$(A, \xi, \varphi_1, \varepsilon_1)$. Then, we change the angle φ_1 to $\varphi_2 := \varphi_1 + \arctan\frac{l}{t_1}$, where $l := \tilde{\tau}\varepsilon_2$ and $\tilde{\tau} > 0$ is a given parameter. In that way, we get the new point $\omega_{1,2} := \xi + t_2 e^{\mathrm{i}\varphi_2}$, where $t_2 :=$ ***Search**$(A, \xi, \varphi_2, \varepsilon_1)$. Analogously, as long as the angle $\varphi_j \leq \pi$, we construct a sequence of points $\{\omega_{1,j}\}, j \in \{1, 2, ..., m_1\}$, such that $|\omega_{1,j} - \omega_{1,j+1}| < \varepsilon_2$. The obtained polygon $\{\omega_{1,j}\}_{j=1}^{m_1}$ approximates the boundary of one component of the set $\Gamma^{\mathcal{R}}(A)$.

In some steps of the algorithm ***MGSs**, because of the geometry of the minimal Geršgorin set, it can be impossible to find the next point. We use the following lemma to overcome that difficulty and change the center point ξ.

Lemma 2.4.4 *Given an arbitrary irreducible matrix $A = [a_{ij}] \in \mathbb{C}^{n,n}$, for every point $\omega \in \partial\Gamma^{\mathcal{R}}(A)$, sufficiently small $\varepsilon > 0$ and $z \in \partial\Gamma^{\mathcal{R}}(A)$ satisfying $|z - \omega| < \varepsilon$ and $arg(z - a_{ii}) > arg(\omega - a_{ii})$, there exists an index $i \in N$ such that for all $\alpha \in [0, 1]$ holds $\alpha z + (1 - \alpha)a_{ii} \in \Gamma^{\mathcal{R}}(A)$,*

Proof: If $A = D - B$, where $D = \mathrm{diag}(a_{11}, a_{22}, ..., a_{nn})$, let's consider the family of matrices $A(t) := D - tB$, for $t \in [0, 1]$. Clearly, $A(0) = D$ and $A(1) = A$. Without loss of generality, let's assume that diagonal elements of A are distinct. As $t \to \nu_{A(t)}(z),\ z \in \mathbb{C}$, is a continuous function in t ([40]), $\Gamma^{\mathcal{R}}(A(t))$ grows continuously from the set of n points $\{a_{ii},\ i \in N\}$ to $\Gamma^{\mathcal{R}}(A)$. For the sufficiently small $t > 0$, $\Gamma^{\mathcal{R}}(A(t))$ consists of n disjoint sets around points $a_{ii},\ i \in N$, and the statement of lemma holds trivially. Also, because of continuity, that property is preserved when disjointed components merge. ❑

Therefore, for each point $z \in \partial\Gamma^{\mathcal{R}}(A)$, there is diagonal entry $a_{ii},\ i \in N$, such that $a_{ii} + te^{\mathrm{arg}(z - a_{ii})} \in \Gamma^{\mathcal{R}}(A),\ t \in [0, \widehat{\rho}_i(\mathrm{arg}(z - a_{ii}))]$.

After the construction of the first component of the minimal Geršgorin set, we check which entries from $\mathcal{D}$ are in that component and denote the set of that entries with $\mathcal{S}_1$. If $\mathcal{S}_1 \neq \mathcal{D}$, we construct a new polygon $\{\omega_{2,j}\}_{j=1}^{m_2}$ which represents the approximation of the next disjoint component of the minimal Geršgorin set. Then, we test which entries from the set $\mathcal{D}\backslash\mathcal{S}_1$ are in that component and denote the set of those entries with $\mathcal{S}_2$. We stop with that procedure when all elements of $\mathcal{D}$ are included in some component, i.e., $\mathcal{D} = \mathcal{S}_1 \dot{\cup} \mathcal{S}_2 \dot{\cup} ... \dot{\cup} \mathcal{S}_k$, where k is the number of disjoint components of $\Gamma^{\mathcal{R}}(A)$. Therefore, we have the following result on **eMGSs** and **iMGSs**.

Theorem 2.4.5 *The algorithm* ****MGSs*** *produces the numerical approximation of the minimal Geršgorin set.*

***MGSs**

Input: A, N_s, $\tilde{\tau}$, tol

Set $\bar{\mathcal{S}} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ and initialize $i = 1$;

Set $\varepsilon_2 = \dfrac{\max\{u_{re} - l_{re}, u_{im} - l_{im}\}}{N_s}$;

while $\bar{\mathcal{S}} \neq \emptyset$ **do**

Set $\xi_0 = \bar{\mathcal{S}}(1),\ \mathcal{S}_i = \{\xi_0\},\ \xi = \xi_0, \varphi = \theta_0 = \theta_1 = -\pi, l = \dfrac{\varepsilon_2}{\tilde{\tau}},\ j = 1$;

Run ***Search**(A, ξ, φ, tol) to compute t_j and $\omega_{i,j}$;

while $|\theta_j - \theta_{j-1}| < \pi$ **do**

Update $\varphi \leftarrow \varphi + \arctan(\frac{l}{t_j})$ and $j \leftarrow j + 1$;

Run ***Search**(A, ξ, φ, tol) to compute t_j and $\omega_{i,j}$ and set $\theta_j = \arg(\frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|})$;

while $|\omega_{i,j} - \omega_{i,j-1}| < \varepsilon_2$ and $|\theta_j - \theta_{j-1}| < \pi$ **do**

Update $\varphi \leftarrow \varphi + \arctan(\frac{l}{t_j})$ and $j \leftarrow j + 1$;

Run ***Search**(A, ξ, φ, tol) to compute t_j and $\omega_{i,j}$ and set $\theta_j = \arg(\frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|})$;

end while

if $|\theta_j - \theta_{j-1}| < \pi$ **then**

Set $k = 1$ and $\tilde{\mathcal{S}}$ to $\bar{\mathcal{S}}$ ordered with respect to the distance $\frac{\omega_{i,j} + \omega_{i,j-1}}{2}$;

repeat

if $\tilde{S}(k) \neq \xi$ **then**

Run ***Search**$(A, \tilde{\mathcal{S}}(k), \arg(\frac{\omega_{i,j-1} - \tilde{\mathcal{S}}(k)}{|\omega_{i,j-1} - \tilde{\mathcal{S}}(k)|}), tol)$ to compute t_{old} and ω_{old};

if $|\omega_{old} - \omega_{i,j-1}| < \varepsilon_2$ **then**

Run ***Search**$(A, \tilde{\mathcal{S}}(k), \arg(\frac{\omega_{i,j} - \tilde{\mathcal{S}}(k)}{|\omega_{i,j} - \tilde{\mathcal{S}}(k)|}), tol)$ to compute t_{new} and ω_{new};

end if

end if

Update $k \leftarrow k + 1$;

until $(|\omega_{old} - \omega_{i,j-1}| < \varepsilon_2$ and $|\omega_{old} - \omega_{new}| < \varepsilon_2)$ or $k > \text{card}(\tilde{\mathcal{S}})$;

if $(|\omega_{old} - \omega_{i,j-1}| < \varepsilon_2$ and $|\omega_{old} - \omega_{new}| < \varepsilon_2)$ **then**

Set $\xi = \tilde{\mathcal{S}}(k)$ and update $\mathcal{S}_i \leftarrow \mathcal{S}_i \cup \{\xi\}$;

Set $t_j = t_{new},\ \omega_{i,j} = \omega_{new}$ and set $\varphi = \arg(\frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|})$;

else

Update $l \leftarrow l\tilde{\tau}$ and $j \leftarrow j - 1$ and set $\varphi = \arg(\frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|})$;

end if

end if

end while

Update $\bar{\mathcal{S}} \leftarrow \bar{\mathcal{S}} \backslash \mathcal{S}_i$, $i \leftarrow i + 1$;

Update $\bar{\mathcal{S}}$ to exclude all elements inside of the polygon $\{\omega_{i,j}\}_{1 \leq j \leq m_i}$;

end while

Output: $\{\{\omega_{1,j}\}_{1 \leq j \leq m_1}, \{\omega_{2,j}\}_{1 \leq j \leq m_2}, ..., \{\omega_{s,j}\}_{1 \leq j \leq m_s}\}$

2.4.2 Predictor-corrector curve tracing

One of the typical path following methods to numerically trace the curve $\mathcal{C}$ in the complex plane is a generic predictor-corrector method. It uses a combination of two different steps.

Let $\mathcal{C}$ be a solution curve of the equation $H(\omega) = 0$, where $H : \mathbb{C} \to \mathbb{R}$ is a smooth map and $0 \in \text{range}(H)$.

In the first step (predictor step), an approximation along the curve is used, usually in the direction of the tangent of the curve.

In the second step (corrector step), iterations for solving $H(\omega) = 0$ are used. Typically corrections are of Newton or gradient type. In that way, the predicted point "brings back" to the curve.

Generic predictor-corrector method

Input: $\omega_0 \in \mathbb{C}$, $H(\omega_0) \approx 0$ (initial point), $h > 0$ (initial step length)

1: **for** $k = 1 : m$ **do**

2: **(Predictor step)** Predict $z_i \in \mathbb{C}$ such that $||z_i - \omega_{i-1}|| \approx h$ in the direction of tracing;

3: **(Corrector step)** Let $\omega_i \in \mathbb{C}$ approximately solve $\min_{\omega}\{||z_i - \omega|| : H(\omega) = 0\}$ and choose a new step-length $h > 0$;

4: **end for**

Output: $\omega_i \in \mathbb{C}, i \in \{0, 1, 2, ..., m\}$

First, we consider the explicit characterization of the minimal Geršgorin set. In the following theorem, we present derivatives of the first and second order of f_A, with respect to x and y.

Theorem 2.4.6 *For a given an irreducible matrix $A \in \mathbb{C}^{n,n}$ and $z = x + \mathrm{i}y \in \mathbb{C}$, let's $v(x + \mathrm{i}y)$ and $u(x + \mathrm{i}y)$ be right and left eigenvector of $Q_A(x + \mathrm{i}y)$, corresponding to $f_A(x + \mathrm{i}y)$, where $Q_A(x + \mathrm{i}y)$ and $f_A(x + \mathrm{i}y)$ are defined by (1.19) and (2.6), respectively. Then, the first and second derivatives of f_A are defined by:*

$$f_x = -\frac{u^T D_x v}{u^T v}, \tag{2.36}$$

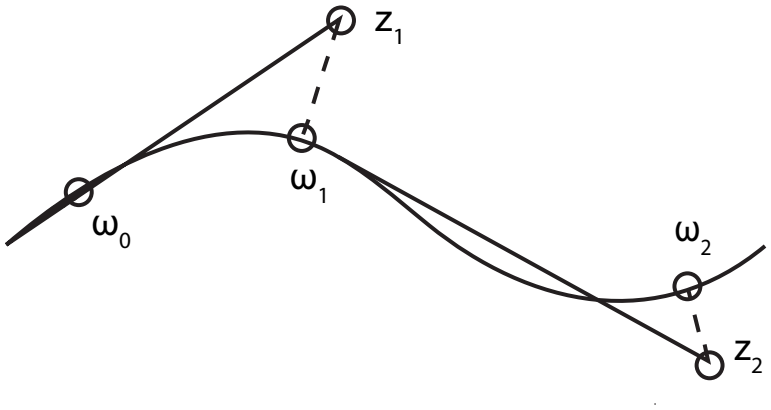


Figure 2.2: Predictor-corrector step.

$$f_y = -\frac{u^T D_y v}{u^T v}, \tag{2.37}$$

$$f_{xx} = -\frac{u^T D_{xx} v + 2u^T D_x v_x + 2f_x u^T v_x}{u^T v}, \tag{2.38}$$

$$f_{xy} = -\frac{u^T D_{xy} v + u^T D_x v_y + u^T D_y v_x + f_x u^T v_y + f_y u^T v_x}{u^T v} \tag{2.39}$$

and

$$f_{yy} = -\frac{u^T D_{yy} v + 2u^T D_y v_y + 2f_y u^T v_y}{u^T v}, \tag{2.40}$$

where: $D_x = diag\Big(\frac{Re(x+\mathrm{i}y-a_{11})}{|x+\mathrm{i}y-a_{11}|}, \frac{Re(x+\mathrm{i}y-a_{22})}{|x+\mathrm{i}y-a_{22}|}, ..., \frac{Re(x+\mathrm{i}y-a_{nn})}{|x+\mathrm{i}y-a_{nn}|}\Big)$,
$D_y = diag\Big(\frac{Im(x+\mathrm{i}y-a_{11})}{|x+\mathrm{i}y-a_{11}|}, \frac{Im(x+\mathrm{i}y-a_{22})}{|x+\mathrm{i}y-a_{22}|}, ..., \frac{Im(x+\mathrm{i}y-a_{nn})}{|x+\mathrm{i}y-a_{nn}|}\Big)$,
$D_{xx} = diag\Big(\frac{(Im(x+\mathrm{i}y-a_{11}))^2}{|x+\mathrm{i}y-a_{11}|^3}, \frac{(Im(x+\mathrm{i}y-a_{22}))^2}{|x+\mathrm{i}y-a_{22}|^3}, ..., \frac{(Im(x+\mathrm{i}y-a_{nn}))^2}{|x+\mathrm{i}y-a_{nn}|^3}\Big)$,
$D_{xy} = diag\Big(\frac{-Re(x+\mathrm{i}y-a_{11})Im(x+\mathrm{i}y-a_{11})}{|x+\mathrm{i}y-a_{11}|^3}, \frac{-Re(x+\mathrm{i}y-a_{22})Im(x+\mathrm{i}y-a_{22})}{|x+\mathrm{i}y-a_{22}|^3}, ..., \frac{-Re(x+\mathrm{i}y-a_{nn})Im(x+\mathrm{i}y-a_{nn})}{|x+\mathrm{i}y-a_{nn}|^3}\Big)$
and $D_{yy} = diag\Big(\frac{(Re(x+\mathrm{i}y-a_{11}))^2}{|x+\mathrm{i}y-a_{11}|^3}, \frac{(Re(x+\mathrm{i}y-a_{22}))^2}{|x+\mathrm{i}y-a_{22}|^3}, ..., \frac{(Re(x+\mathrm{i}y-a_{nn}))^2}{|x+\mathrm{i}y-a_{nn}|^3}\Big)$,
for $z = x + \mathrm{i}y \neq a_{ii},\ i \in N$.

Proof: From the definition of f_A, we have:

$$Q_A(x+\mathrm{i}y)v(x,y) = f_A(x+\mathrm{i}y)v(x,y) \tag{2.41}$$

and

$$(u(x,y))^T Q_A(x+\mathrm{i}y) = f_A(x+\mathrm{i}y)(u(x,y))^T. \tag{2.42}$$

Differentiating the equation (2.41) with respect to x and y, we obtain

$$-D_x v + Q_A v_x = f_x v + f v_x \tag{2.43}$$

and

$$-D_y v + Q_A v_y = f_y v + f v_y. \tag{2.44}$$

Multiplying the equations (2.43) and (2.44) by u^T and using (2.41) and (2.42), we obtain (2.36) and (2.37).

Using Theorem 1.2.1, we have

$$v_x = -(fI - Q_A)^+ (I - \frac{vu^T}{u^T v}) D_x v \tag{2.45}$$

and

$$v_y = -(fI - Q_A)^+ (I - \frac{vu^T}{u^T v}) D_y v. \tag{2.46}$$

Differentiating the equation (2.43) with respect to x and y, and the equation (2.44) with respect to y, we obtain equations:

$$-D_{xx} v - 2D_x v_x + Q_A v_{xx} = f_{xx} v + 2f_x v_x + f v_{xx}, \tag{2.47}$$

$$-D_{xy} v - D_x v_y - D_y v_x + Q_A v_{xy} = f_{xy} v + f_x v_y + f_y v_x + f v_{xy}, \tag{2.48}$$

and

$$-D_{yy} v - 2D_y v_y + Q_A v_{yy} = f_{yy} v + 2f_y v_y + f v_{yy}, \tag{2.49}$$

respectively.

Multiplying the equations (2.47), (2.48) and (2.49) by u^T and using (2.41) and (2.42), we obtain expressions (2.38), (2.39) and (2.40). ❑

Now, let's consider the implicit characterization of the minimal Geršgorin set given by the system:

$$\begin{bmatrix} -Q_A(x+\mathrm{i}y) & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} w_A(x,y) \\ g_A(x,y) \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \tag{2.50}$$

Theorem 2.4.7 *Given an irreducible matrix $A \in \mathbb{C}^{n,n}$, a vector $c > 0$, $c \in \mathbb{R}^n$ and w_A and g_A defined by the system (2.50). Then, the first and second derivatives of g_A are given by systems:*

$$\begin{bmatrix} -Q_A & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} w_x & w_y \\ g_x & g_y \end{bmatrix} = \begin{bmatrix} -D_x w & -D_y w \\ 0 & 0 \end{bmatrix} \tag{2.51}$$

and

$$\begin{bmatrix} -Q_A & -c \\ -c^T & 0 \end{bmatrix} \begin{bmatrix} w_{xx} & w_{xy} & w_{yy} \\ g_{xx} & g_{xy} & g_{yy} \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} D_1 & D_2 & D_3 \\ 0 & 0 & 0 \end{bmatrix}, \tag{2.52}$$

where $D_1 = -2D_x w_x - D_{xx} w$, $D_2 = -D_x w_y - D_y w_x - D_{xy} w$ and $D_3 = -2D_y w_y - D_{yy} w$.

Proof: Differentiating the system of equations

$$\begin{aligned} -Q_A w - cg &= 0 \\ -c^T w &= -1 \end{aligned} \tag{2.53}$$

with respect to x and y, we obtain systems:

$$\begin{aligned} -Q_A w_x - cg_x &= -D_x w \\ -c^T w_x &= 0 \end{aligned} \tag{2.54}$$

and

$$\begin{aligned} -Q_A w_y - cg_y &= -D_y w \\ -c^T w_y &= 0, \end{aligned} \tag{2.55}$$

respectively.
Writing (2.54) and (2.55) in a matrix form, we obtain (2.51).

Differentiating the system (2.54) with respect to x and y, and the system (2.55) with respect to y, we obtain systems:

$$\begin{aligned} -Q_A w_{xx} - cg_{xx} &= -2D_x w_x - D_{xx} w \\ -c^T w_{xx} &= 0, \\ -Q_A w_{xy} - cg_{xy} &= -D_x w_y - D_y w_x - D_{xy} w \end{aligned} \tag{2.56}$$

$$-c^T w_{xy} = 0 \tag{2.57}$$

and

$$\begin{aligned} -Q_A w_{yy} - c g_{yy} &= -2D_y w_y - D_{yy} w \\ -c^T w_{yy} &= 0, \end{aligned} \tag{2.58}$$

respectively. Finally, using (2.56), (2.57) and (2.58), it follows (2.52). ❑

Now, we construct the algorithm **eMGSp**. The boundary of the minimal Geršgorin set is given by $\partial\Gamma^{\mathcal{R}}(A) = \{z = x + \mathrm{i}y \in \mathbb{C} : f_A(x + \mathrm{i}y) = 0\}$. Starting with the point $\omega_0 \in \partial\Gamma^{\mathcal{R}}$, which we can obtain by the procedure **eSearch**, we want to find the next point on the boundary of $\Gamma^{\mathcal{R}}(A)$, named ω_1.

Firstly, in the predictor step, we obtain the point

$$z_1 := \omega_0 + hdl, \tag{2.59}$$

where h is a given length of a step and $dl := \pm\dfrac{-f_y + \mathrm{i}f_x}{|-f_y + \mathrm{i}f_x|}$ (we choose the sign in the direction of curve tracing), where f_x and f_y are computed in ω_0.
Then, in the corrector step, we want to find the point $\omega_1 \in \partial\Gamma^{\mathcal{R}}$, which is the nearest to z_1. To that end, we solve the problem:

$$||\omega - z_1||_2^2 \to \min,\ f_A(\omega) = 0.$$

Forming a function

$$L(x, y, \lambda) := (x - \mathrm{Re}(z_1))^2 + (y - \mathrm{Im}(z_1))^2 + \lambda f_A(x, y)$$

and differentiating it with respect to x and y, we obtain the following iterations:

$$\begin{bmatrix} x_1^{(k)} \\ y_1^{(k)} \\ \lambda^{(k)} \end{bmatrix} = \begin{bmatrix} \mathrm{Re}(\omega_1^{(k)}) \\ \mathrm{Im}(\omega_1^{(k)}) \\ 1 \end{bmatrix} - \begin{bmatrix} 2 + \lambda f_{xx} & \lambda f_{xy} & f_x \\ \lambda f_{xy} & 2 + \lambda f_{yy} & f_y \\ f_x & f_y & 0 \end{bmatrix}^{-1} \begin{bmatrix} 2\mathrm{Re}(\omega_1^{(k)} - z_1) + \lambda f_x \\ 2\mathrm{Im}(\omega_1^{(k)} - z_1) + \lambda f_y \\ f \end{bmatrix}, \tag{2.60}$$

where $\omega_1^{(0)} := z_1,\ \omega_1^{(k)} := x_1^{(k-1)} + \mathrm{i}y_1^{(k-1)},\ k \in \mathbb{N}$, and $f,\ f_x,\ f_y,\ f_{xx},\ f_{xy}$ and f_{yy} are computed in $\omega_1^{(k)}$.

Computation of these iterations will stop when $|f| \leq tol$, for some $l \in \mathbb{N}$ and a given accuracy $tol > 0$. In practice, as z_1 is near to border of $\Gamma^{\mathcal{R}}(A)$, it is

sufficient to compute just a few iterations. In that way, we get $\omega_1 := \omega_1^{(l)}$.

Analogously, we find a sequence of points $\{\omega_j\}_{j=0}^{m}$, which approximate the boundary of the one component of minimal Geršgorin set. In the same way, we can find approximation of all other components of $\Gamma^{\mathcal{R}}(A)$.

Finally, using function $g_A(x, y)$ instead of $f_A(x, y)$ for the characterization of the boundary of the minimal Geršgorin set, we construct the implicit predictor-corrector method for computing the minimal Geršgorin set- **iMGSp**. In that case, we use the characterizations of derivatives of g_A given in Theorem 2.4.7.

---

eMGSp

Input: $A,\ h,\ tol$

Set $\mathcal{D} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ and initialize $i = 1$;
while $\mathcal{D} \neq \emptyset$ **do**
  Initialize $\xi = \mathcal{D}(1),\ \theta = -\pi,\ \theta_1 = -3\pi$ and $j = 0$;
  Set $\omega =$**eSearch** $(A, \xi, -\pi, tol)$, and $\omega_{i,0} := \omega$;
  while $\theta - \theta_1 > -\pi$ **do**
    Compute f_x and f_y in $\omega_{i,j}$ by (2.36) and (2.37);
    Compute $z_{i,j+1}$ using (2.59)
    Set $w = z_{i,j+1}$ and compute $f = f(w)$ as the Perron-Frobenius eigenvalue of $Q_A(w)$;
    while $|f| > tol$ **do**
      Compute $f_{xx},\ f_{xy}$ and f_{yy} in w by (2.38), (2.39) and (2.40);
      Compute w by solving the system (2.60);
      Compute $f = f(w)$ as the Perron-Frobenius eigenvalue of $Q_A(w)$;
    end while
    Update $j \leftarrow j + 1$ and $\omega_{i,j} \leftarrow w$;
    Set $\theta_1 := \theta,\ \theta := -\mathrm{i} \ln \frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|}$;
  end while
  Update $i \leftarrow i + 1$;
  Update $\mathcal{D}$ to exclude all elements inside of the polygon $\{\omega_{i,j}\}_{0 \leq j \leq m_i}$;
end while

Output: $\{\{\omega_{1,j}\}_{0 \leq j \leq m_1}, \{\omega_{2,j}\}_{0 \leq j \leq m_2}, ..., \{\omega_{s,j}\}_{0 \leq j \leq m_s}\}$

---

iMGSp

Input: $A,\ h,\ tol$

Set $\mathcal{D} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ and initialize $i = 1$;
while $\mathcal{D} \neq \emptyset$ **do**
  Initialize $\xi = \mathcal{D}(1),\ \theta = -\pi,\ \theta_1 = -3\pi$ and $j = 0$;
  Set $\omega =$**iSearch** $(A, \xi, -\pi, tol),\ \omega_{i,0} := \omega$;
  while $\theta - \theta_1 > -\pi$ **do**
    Compute g_x and g_y in $\omega_{i,j}$ by solving the system (2.51);
    Compute $z_{i,j+1}$ using (2.59);
    Set $w = z_{i,j+1}$ and compute $g = g(w)$ by solving the system (2.50);
    while $|g| > tol$ **do**
      Compute $g_{xx},\ g_{xy}$ and g_{yy} in w by solving the system (2.52);
      Compute w by solving the system (2.60);
      Compute $g = g(w)$ by solving the system (2.50);
    end while
    Update $j \leftarrow j + 1$ and $\omega_{i,j} \leftarrow w$;
    Set $\theta_1 := \theta,\ \theta := -\mathrm{i} \ln \frac{\omega_{i,j} - \xi}{|\omega_{i,j} - \xi|}$;
  end while
  Update $i \leftarrow i + 1$;
  Update $\mathcal{D}$ to exclude all elements inside of the polygon $\{\omega_{i,j}\}_{0 \le j \le m_i}$;
end while

Output: $\{\{\omega_{1,j}\}_{0 \le j \le m_1}, \{\omega_{2,j}\}_{0 \le j \le m_2}, ..., \{\omega_{s,j}\}_{0 \le j \le m_s}\}$

2.4.3 Triangular curve tracing

In this subsection, two new algorithms for computing the minimal Geršgorin set are constructed. For a given matrix $A \in \mathbb{C}^{n,n}$, we combine the triangular grid approach presented in [47] with characterizations of $\Gamma^{\mathcal{R}}(A)$ by functions f_A and h_A to develop algorithms **eMGSt** and **iMGSt**, respectively.

Given any $(z_i, z_e) \in \mathbb{C}^2$ such that $z_i \neq z_e$, for $k \neq l$, define the following points:

$$\mathcal{L}_{k,l} = z_i + k(z_e - z_i) + l(z_e - z_i)e^{\frac{\mathrm{i}\pi}{3}},$$

to obtain a uniform lattice of vertices

$$\mathcal{L}(z_i, z_e) = \{\mathcal{L}_{k,l} : (k,l) \in \mathbb{Z}^2\},$$

satisfying

$$|\mathcal{L}_{k,l+1} - \mathcal{L}_{k,l}| = |\mathcal{L}_{k+1,l} - \mathcal{L}_{k,l}| = |z_i - z_e|.$$

Next, we define a triangular mesh, see Figure 2.3, as:

$$\Omega(z_i, z_e) = \Psi(z_i, z_e) \cup \widetilde{\Psi}(z_i, z_e),$$

where

$$\Psi(z_i, z_e) = \{T_{kl} = \{\mathcal{L}_{k,l}, \mathcal{L}_{k+1,l}, \mathcal{L}_{k,l+1}\} : (k,l) \in \mathbb{Z}^2\},$$

and

$$\widetilde{\Psi}(z_i, z_e) = \{\tilde{T}_{kl} = \{\mathcal{L}_{k,l}, \mathcal{L}_{k+1,l}, \mathcal{L}_{k+1,l-1}\} : (k,l) \in \mathbb{Z}^2\}.$$

For a given matrix $A \in \mathbb{C}^{n,n}$, let us denote by $\mathcal{T}$ the subset of $\Omega(z_i, z_e)$, where $T \in \mathcal{T}$ if and only if T has at least one vertex in $\Gamma^{\mathcal{R}}(A)$ and at least one outside of $\Gamma^{\mathcal{R}}(A)$.
Let the pivot $p(T)$ be the vertex of a triangle $T \in \mathcal{T}$ which is situated on the opposite side of the border of $\Gamma^{\mathcal{R}}(A)$ to other two vertices, e.g., vertex $\tilde{z}_{i,0}$ in the triangle $\{\tilde{z}_{i,0}, \tilde{z}_{i,1}, \tilde{z}_{i,2}\}$ in Figure 2.4. Define a transformation:

$$F(T) = \rho(p(T), \mathrm{sgn}(\nu_A(p(T))) \cdot \frac{\pi}{3})(T),$$

where $\mathrm{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases}$ and $\rho(z, \theta)(\omega)$ denotes the rotation of $\omega \in \mathbb{C}$ centered at $z \in \mathbb{C}$ with angle θ, i.e.,

$$\rho(z, \theta)(\omega) := z + (\omega - z)e^{\mathrm{i}\theta}.$$

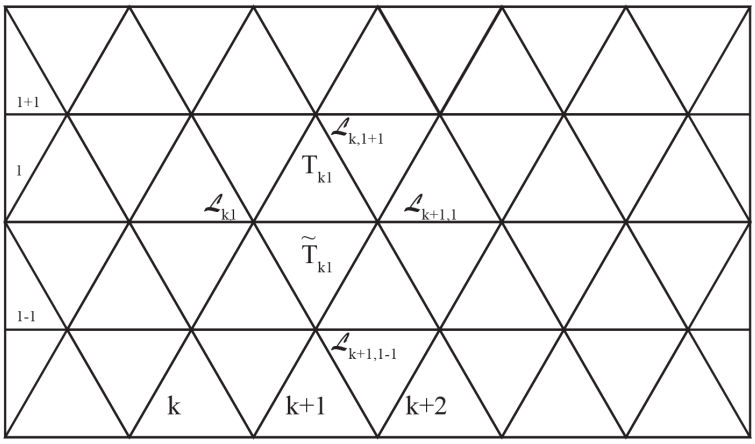


Figure 2.3: Triangular grid.

Now, we state some useful properties of triangular grids and mapping F defined on them.

Proposition 2.4.8 *For $z_i \neq z_e$, $\mathcal{T}$ is a finite set.*

Proof: As $\Gamma^{\mathcal{R}}(A)$ is bounded, then $\Gamma^{\mathcal{R}}(A) \cap \Omega(z_i, z_e)$ is a finite set. From $\text{card}(\mathcal{T}) \leq \text{card}(\Gamma^{\mathcal{R}}(A) \cap \Omega(z_i, z_e))$, it follows that $\mathcal{T}$ is a finite set. ❑

Proposition 2.4.9 *For a given triangle $T \in \mathcal{T}$, the following statements hold:*

1. *$F(T) \neq T$;*
2. *$p(T)$ is a vertex of $F(T)$;*
3. *T and $F(T)$ are adjacent;*
4. *$F(T) \in \mathcal{T}$;*
5. *$p(F(T))$ is a vertex of T;*
6. *$F^2(T) \neq T$;*
7. *if $T \in \Psi(z_i, z_e)$, then $F(T) \in \tilde{\Psi}(z_i, z_e)$ and if $T \in \tilde{\Psi}(z_i, z_e)$, then $F(T) \in \Psi(z_i, z_e)$.*

Proof: 1. The statement follows directly from the definition of F.

2. As $p(T)$ is a center of the rotation of T, it is a common vertex of T and $F(T)$.

3. Since T is an equilateral triangle and $F(T)$ is also an equilateral triangle obtained by a rotation of T centered at a vertex of T with the angle $\theta = \pm\frac{\pi}{3}$, T and $F(T)$ are adjacent.

4. $F(T) \in \mathcal{T}$ because T and $F(T)$ have a common edge whose one vertex is in $\Gamma^{\mathcal{R}}(A)$ and the other vertex is outside of $\Gamma^{\mathcal{R}}(A)$.

5. The vertex $p(F(T))$ lies on the common edge of T and $F(T)$.

6. Triangles T and $F(T)$ are adjacent and $p(T) \in T \cap F(T)$ and $p(F(T)) \in T \cap F(T)$. There are two cases.
   If $p(T) = p(F(T))$, then $F^2(T) = \rho(p(T), \pm\frac{2\pi}{3})(T) \neq T$.
   If $p(T) \neq p(F(T))$, then $F^2(T) = \rho(p(F(T)), -\theta)(\rho(p(T), \theta)(T))$
   So, $F^2(T) = p(F(T)) + ((p(T) + (T - p(T))e^{\mathrm{i}\theta}) - p(F(T)))e^{-\mathrm{i}\theta}$, implying that $F^2(T) = (p(F(T)) - p(T))(1 - e^{-\mathrm{i}\theta}) + T \neq T$.

7. It immediately follows.

❑

Proposition 2.4.10 *F is a bijection from $\mathcal{T}$ onto $\mathcal{T}$.*

Proof: It is enough to show that F is a one-to-one mapping. Assume that there exist two triangles T and T', such that $F(T) = F(T')$. From the previous proposition, we know that $p(F(T)) \in T$, $p(F(T)) \in F(T)$, $p(F(T)) \in F^2(T)$ and $p(F(T)) \in T'$, i.e., $p(F(T)) \in T \cap F(T) \cap F^2(T) \cap T'$.
Furthermore, as T, $F^2(T)$ and T' are adjacent to $F(T)$, it has to be $T' = F^2(T)$. Therefore, $F(T') = F(F^2(T)) = F^2(F(T))$. Using the previous proposition (item 6.), we obtain $F(T) \neq F(T')$, which is a contradiction with starting assumption. Hence, $F : \mathcal{T} \to \mathcal{T}$ is one-to-one mapping, and, thus, bijection onto $\mathcal{T}$. ❑

For any given $T \in \mathcal{T}$ define $T_k := F^k(T), k \in \mathbb{N}$, and set $O(T) := \{T_k,\ k \in \mathbb{N}\}$, where $T_0 := T$.

Proposition 2.4.11 *For a given triangle $T \in \mathcal{T}$, the following statements hold:*

1. *the set $O(T)$ is finite;*

2. *if $n = card(O(T))$, then n is even and the smallest positive integer such that $T_n = T$;*

3. $\sum_{i=0}^{n-1} \theta_i = 2\pi m,\ m \in \mathbb{N}_0$, *where θ_i is the rotation angle of F for the triangle T_i;*

4. *for a given triangle T', either $O(T) = O(T')$ or $O(T) \cap O(T') = \emptyset$.*

Proof: 1. As $\mathcal{T}$ is a finite set and $O(T) \subseteq \mathcal{T}$, it follows that $O(T)$ is a finite set, too.

2. Assume that exists an integer $k,\ 0 < k < n$, such that $T_n = T_k = T$. But, then, $F^n(T) = F^k(T)$ and, consequently, $F^{n-k}(T) = T$, implying that $\text{card}(O(T)) < n$, which is a contradiction.
Since T and $F(T)$ belong to disjoint sets ($\Psi(z_i, z_e)$ and $\tilde{\Psi}(z_i, z_e)$), n has to be even.

3. So, $O(T) = \{T_0, T_1, ..., T_{n-1}\}$, $T_0 := T$ and let $\{P_i, S_{1,i}, S_{2,i}\}$ be the vertices of the triangle T_i for $i \in \{0, 1, ..., n-1\}$ such that P_i is the pivot of T_i, $\sphericalangle(\overrightarrow{P_iS_{1,i}}, \overrightarrow{P_iS_{2,i}}) = \theta_i$, and $T_{i+1} = F(T_i) = \rho(P_i, \theta_i)(T_i)$. Using the previous results, we have:

$$T_i \cap T_{i+1} = [P_i, S_{2,i}] = [P_{i+1}, S_{1,i+1}].$$

We consider two cases:

- if $P_{i+1} = P_i$ and $S_{1,i+1} = S_{2,i}$, then

$$\overrightarrow{P_{i+1}S_{1,i+1}} = \overrightarrow{P_iS_{2,i}};$$

- if $P_{i+1} = S_{2,i}$ and $P_i = S_{1,i+1}$, then

$$\overrightarrow{P_{i+1}S_{1,i+1}} = -\overrightarrow{P_iS_{2,i}}. \tag{2.61}$$

Therefore,

$$\sum_{i=0}^{n-1}\theta_i=\sum_{i=0}^{n-1}\sphericalangle(\overrightarrow{P_iS_{1,i}},\overrightarrow{P_iS_{2,i}})=\sphericalangle(\overrightarrow{P_0S_{1,0}},\overrightarrow{P_0S_{1,0}})+l\pi,$$

where $l\in\mathbb{N}$ is a number of triangles that satisfy (2.61).
As $F_n(T)=T_0$, l is even. Finally, we obtain $\sum_{i=0}^{n-1}\theta_i=2\pi m,\ m\in\mathbb{N}_0$.

4. Let assume that $O(T)\cap O(T')\neq\emptyset$ and $n\in\mathbb{N}$ is a cardinality of $O(T)$. Then, there exist a triangle T'' and $i,\ j\in\mathbb{N}$, such that $T''\in O(T)\cap O(T')$ and $T''=F^i(T)=F^{n+i}(T)=F^j(T')$. It follows that $F^{n+i-j}(T)=T'$, leading to $O(T')\subseteq O(T)$.
Analogously, we can prove $O(T)\subseteq O(T')$, i.e., $O(T)=O(T')$.

❑

Using the prepositions above and the function $f_A(z)$, we present the algorithm **eMGSt**.

As before, given irreducible matrix $A\in\mathbb{C}^{n,n}$, the set of its different diagonal elements is $\mathcal{D}=\{a_{i_1i_1},a_{i_2i_2},...,a_{i_{\tilde{n}}i_{\tilde{n}}}\}$, where $\mathrm{Re}(a_{i_1i_1})\leq\mathrm{Re}(a_{i_2i_2})\leq...\leq\mathrm{Re}(a_{i_{\tilde{n}}i_{\tilde{n}}})$, $\tilde{n}\in N$, and let s be the number of disjoint components of $\Gamma^{\mathcal{R}}(A)$. We denote m_i points representing the approximation of the boundary of the i^{th} component of the minimal Geršgorin set by $\{z_{i,j}\}_{j=1}^{m_i},\ i\in\{1,2,...,s\}$. Starting with $\xi=a_{i_1i_1},\ \varphi=-\pi$ and given accuracy $\varepsilon>0$ (e.g., $\varepsilon=10^{-12}$), we use the procedure **eSearch** from [40] to obtain $\omega_1:=\xi+t_1e^{\mathrm{i}\varphi}$, with $t_1:=$ **eSearch**$(A,\xi,\varphi,\varepsilon)$.

Then, we get the points $\tilde{z}_{1,0}:=\omega_1-\frac{\tau}{2}$ and $\tilde{z}_{1,1}:=\omega_1+\frac{\tau}{2}$, where τ is a given length of edge of equilateral triangles which form triangular grid and $z_{1,1}:=\tilde{z}_{1,1}$. Furthermore, we obtain the point $\tilde{z}_{1,2}:=\tilde{z}_{1,0}+(\tilde{z}_{1,1}-\tilde{z}_{1,0})e^{\frac{\mathrm{i}\pi}{3}}$. As a result, the triangle $T_1=\{\tilde{z}_{1,0},\tilde{z}_{1,1},\tilde{z}_{1,2}\}$ is an element of $\mathcal{T}$ which generates the set $O(T_1)$. If $f_A(\tilde{z}_{1,2})<0$, we define $z_{1,2}:=\tilde{z}_{1,2}$ and choose as pivot $\tilde{z}_{piv}$ the point $\tilde{z}_{1,0}$ to get $\tilde{z}_{1,3}:=\tilde{z}_{piv}+(\tilde{z}_{1,2}-\tilde{z}_{piv})e^{\frac{\mathrm{i}\pi}{3}}$. Otherwise, we choose $\tilde{z}_{piv}:=\tilde{z}_{1,1}$ and $\tilde{z}_{1,3}:=\tilde{z}_{piv}+(\tilde{z}_{1,2}-\tilde{z}_{piv})e^{-\frac{\mathrm{i}\pi}{3}}$. Analogously, we construct a sequence of points $\{\tilde{z}_{1,l}\},\ l\in\mathbb{N}_0$, as long as $\tilde{z}_{1,l}=\tilde{z}_{1,0}$. From that set of points, we choose a subset $\{z_{1,j}\}_{j=1}^{m_1}$, such that $f_A(z_{1,j})<0$. The obtained polygon $\{z_{1,j}\}_{j=1}^{m_1}$ contains

one component of the set minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$, see Figure 2.4, and $\mathrm{dist}(z_{i,j}, \partial\Gamma^{\mathcal{R}}(A)) \leq \tau$. Notice, that we could give also an inner approximation of the boundary $\partial\Gamma^{\mathcal{R}}(A)$ simply by taking a subset of points with non-negative values of f_A.

After completing the construction of the first component of the minimal Geršgorin set, we check which entries from $\mathcal{D}$ are in that component and denote the set of these diagonal entries by $\mathcal{S}_1$. If $\mathcal{S}_1 \neq \mathcal{D}$, choosing for ξ the leftmost element of $\mathcal{D}\backslash\mathcal{S}_1$, we construct a new polygon $\{z_{2,j}\}_{j=1}^{m_2}$ that represents the approximation of next disjoint component of the minimal Geršgorin set. Then, we again test which entries from the set $\mathcal{D}\backslash\mathcal{S}_1$ are in that component and denote the set of these entries by $\mathcal{S}_2$. We stop with that procedure when all elements of $\mathcal{D}$ are included in some component of the minimal Geršgorin set.

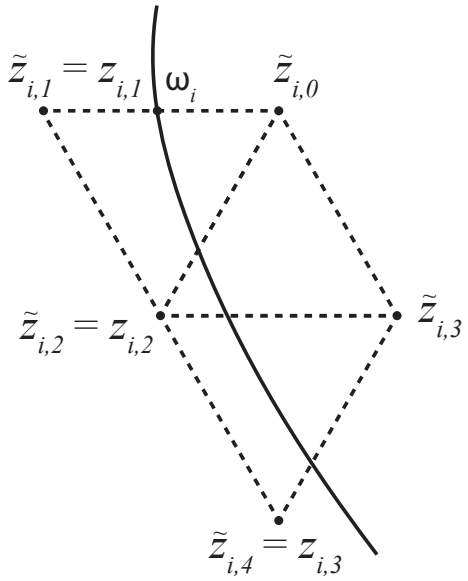


Figure 2.4: Construction of the polygon $\{z_{i,j}\}_{j=1}^{m_i}$.

Finally, we present **iMGSt** algorithm. In fact, we improve algorithm **eMGSt** by replacing the function f_A with the function h_A. So, using the idea of the implicit determinant method [25], we achieve to reduce significantly the overall number of expensive eigenvalue computations. This algorithm gives excellent results, especially for matrices of large sizes, which will be shown through examples in the last chapter.

eMGSt

Input: $A,\ N_t,\ tol$

Set $\tau = \dfrac{2d(A)\sqrt{3}}{3N_t}$, where $d(A) = \max\{u_{re} - l_{re}, u_{im} - l_{im}\}$;
Set $\mathcal{D} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ and initialize $i = 1$;
while $\mathcal{D} \neq \emptyset$ **do**
  Set $\xi = \mathcal{D}(1)$ and $\theta = -\pi$;
  Run **eSearch**(A, ξ, θ, tol) to compute $\omega_i \in \mathbb{C}$;
  Compute $\tilde{z}_{i,0} = \omega_i - \frac{\tau}{2}$ and $\tilde{z}_{i,1} = \omega_i + \frac{\tau}{2}$;
  Compute $\tilde{z}_{i,2} = \tilde{z}_{i,0} + (\tilde{z}_{i,1} - \tilde{z}_{i,0})e^{\frac{\mathrm{i}\pi}{3}}$;
  Set $z_{i,start} = \tilde{z}_{i,0}$ and $z_{i,1} = \tilde{z}_{i,1}$;
  Initialize $j = 2$;
  while $\tilde{z}_{i,2} \neq z_{i,start}$ **do**
    if $f_A(\tilde{z}_{i,2}) < 0$ **then**
      $z_{i,j} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,0} = \tilde{z}_{i,0}$;
      $\tilde{z}_{i,1} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,2} = \tilde{z}_{i,0} + (\tilde{z}_{i,1} - \tilde{z}_{i,0})e^{\frac{\mathrm{i}\pi}{3}}$;
      Update $j \leftarrow j + 1$;
    else
      $\tilde{z}_{i,0} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,1} = \tilde{z}_{i,1}$;
      $\tilde{z}_{i,2} = \tilde{z}_{i,1} + (\tilde{z}_{i,0} - \tilde{z}_{i,1})e^{-\frac{\mathrm{i}\pi}{3}}$;
    end if
  end while
  Update $i \leftarrow i + 1$;
  Update $\mathcal{D}$ to exclude all elements inside of the polygon $\{z_{i,j}\}_{1 \leq j \leq m_i}$;
end while

Output: $\{\{z_{1,j}\}_{1 \leq j \leq m_1}, \{z_{2,j}\}_{1 \leq j \leq m_2}, ..., \{z_{s,j}\}_{1 \leq j \leq m_s}\}$

iMGSt

Input: $A,\ N_t,\ tol$

Set $\tau = \dfrac{2d(A)\sqrt{3}}{3N_t}$, where where $d(A) = \max\{u_{re} - l_{re}, u_{im} - l_{im}\}$;
Set $\mathcal{D} = \{a_{i_1 i_1}, a_{i_2 i_2}, ..., a_{i_{\tilde{n}} i_{\tilde{n}}}\}$ and initialize $i = 1$;
while $\mathcal{D} \neq \emptyset$ **do**
  Set $\xi = \mathcal{D}(1)$ and $\theta = -\pi$;
  Run **iSearch**(A, ξ, θ, tol) to compute $\omega_i \in \mathbb{C}$;
  Compute $\tilde{z}_{i,0} = \omega_i - \frac{\tau}{2}$ and $\tilde{z}_{i,1} = \omega_i + \frac{\tau}{2}$;
  Compute $\tilde{z}_{i,2} = \tilde{z}_{i,0} + (\tilde{z}_{i,1} - \tilde{z}_{i,0})e^{\frac{\mathrm{i}\pi}{3}}$;
  Set $z_{i,start} = \tilde{z}_{i,0}$ and $z_{i,1} = \tilde{z}_{i,1}$;
  Initialize $j = 2$;
  while $\tilde{z}_{i,2} \neq z_{i,start}$ **do**
    if $h_A(\tilde{z}_{i,2}) < 0$ **then**
      $z_{i,j} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,0} = \tilde{z}_{i,0}$;
      $\tilde{z}_{i,1} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,2} = \tilde{z}_{i,0} + (\tilde{z}_{i,1} - \tilde{z}_{i,0})e^{\frac{\mathrm{i}\pi}{3}}$;
      Update $j \leftarrow j + 1$;
    else
      $\tilde{z}_{i,0} = \tilde{z}_{i,2}$;
      $\tilde{z}_{i,1} = \tilde{z}_{i,1}$;
      $\tilde{z}_{i,2} = \tilde{z}_{i,1} + (\tilde{z}_{i,0} - \tilde{z}_{i,1})e^{-\frac{\mathrm{i}\pi}{3}}$;
    end if
  end while
  Update $i \leftarrow i + 1$;
  Update $\mathcal{D}$ to exclude all elements inside of the polygon $\{z_{i,j}\}_{1 \leq j \leq m_i}$;
end while

Output: $\{\{z_{1,j}\}_{1 \leq j \leq m_1}, \{z_{2,j}\}_{1 \leq j \leq m_2}, ..., \{z_{s,j}\}_{1 \leq j \leq m_s}\}$

Chapter 3

Algorithms for approximating convex hull of the minimal Geršgorin set

"More data beats clever algorithms,
but better data beats more data."

Peter Norving[18]

In this chapter, in Section 3.1, results about the abscissa of the minimal Geršgorin set are given. Then, in Section 3.2, we compare two eigenvalue localization sets, the minimal Geršgorin set and the numerical range. Next, in Section 3.3, a well-known result about computing the numerical range of a given matrix is presented and finally, in Section 3.4, new algorithms for a construction of a convex polygon that contains the minimal Geršgorin set are developed.

3.1 Characterization of the abscissa of the minimal Geršgorin set

The abscissa of the Geršgorin set $\Gamma(A)$ of a matrix $A \in \mathbb{C}^{n,n}$ is

$$\gamma(A) := \max\Big\{\mathrm{Re}(z) : z \in \Gamma(A)\Big\}. \tag{3.1}$$

[16]Peter Norvig (1956) is an American computer scientist and director of research at Google Inc.

In the same way, we have the abscissa of the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$:

$$\mu(A) := \max\{\mathrm{Re}(z) : z \in \Gamma^{\mathcal{R}}(A)\}. \tag{3.2}$$

For a given matrix $A \in \mathbb{C}^{n,n}$, let's define an essentially non-negative matrix $M_{re}(A) = [m_{ij}] \in \mathbb{R}^{n,n}$ such that

$$m_{ij} := \begin{cases} \mathrm{Re}(a_{ii}), & i = j \\ |a_{ij}|, & i \neq j \end{cases},$$

where $i, j \in N$, that allows us to characterize $\mu(A)$. We start with the following lemma.

Lemma 3.1.1 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$. If $\hat{\mu}$ is the Perron-Frobenius eigenvalue of a matrix $M_{re}(A)$, then $\hat{\mu} \geq Re(a_{ii}),\ i \in N$.*

Proof: If $M_{re}(A) = D + B$, where $D = \mathrm{diag}(M_{re}(A))$, consider the family of matrices $M_{re}(A)(t) := D + tB,\ t \in [0,1]$. For $t = 0$, we get a diagonal matrix $M_{re}(A)(0) = D$ and $\alpha(M_{re}(A)(0)) \geq \mathrm{Re}(a_{ii})$ holds trivially, $i \in N$. As t increases, at least one off-diagonal entry of $M_{re}(A)(t)$ increase because of irreducibility. Using Theorem 1.3.4, item 4., it follows that $\alpha(M_{re}(A)(t)) \geq \mathrm{Re}(a_{ii}),\ t \in [0,1],\ i \in N$. Therefore, $\hat{\mu} = \alpha(M_{re}(A)(1)) \geq \mathrm{Re}(a_{ii}),\ i \in N$. ❑

Theorem 3.1.2 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$. If μ is the Perron-Frobenius eigenvalue of a matrix $M_{re}(A)$, then $\mu(A) \leq \hat{\mu}$. Moreover, the set $\Gamma(W^{-1}AW)$ intersects with the line $\{z \in \mathbb{C} : Re(z) = \hat{\mu}\}$, where w is the Perron-Frobenius eigenvector of $M_{re}(A)$ and $W = diag(w)$.*

Proof: Let $(\hat{\mu}, w)$ be the Perron-Frobenius pair of essentially non-negative matrix $M_{re}(A)$. Then, $\hat{\mu}$ is the right-most eigenvalue of matrix $M_{re}(A)$ and $w = [w_1, w_2, ..., w_n] > 0$. Since $|\hat{\mu} + \mathrm{i}t - a_{ii}| \geq |\mathrm{Re}(\hat{\mu} - a_{ii})| = \hat{\mu} - \mathrm{Re}(a_{ii})$, for $i \in N$ and $t \in \mathbb{R}$, we obtain $Q_A(\hat{\mu} + \mathrm{i}t) \leq -\hat{\mu}I + M_{re}(A)$. Therefore, the right-most eigenvalue of matrix $Q_A(\hat{\mu}+\mathrm{i}t)$ is non-positive, i.e., $\nu_A(\hat{\mu}+\mathrm{i}t) \leq 0$. This, however, implies that points on the line $\hat{\mu}+\mathrm{i}t$ are either on the boundary or outside of the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$. Using the previous lemma, it follows $\mu(A) \leq \hat{\mu}$. From $M_{re}(A)w = \hat{\mu}w$, we obtain $\hat{\mu} = \mathrm{Re}(a_{ii}) + \sum_{j \neq i}^{n} \frac{|a_{ij}|w_j}{w_i}$, $i \in N$. Hence, the set $\Gamma(W^{-1}AW)$ intersects with the line $\{z \in \mathbb{C} : \mathrm{Re}(z) = \hat{\mu}\}$. ❑

Theorem 3.1.3 *Given an arbitrary irreducible matrix $A \in \mathbb{C}^{n,n}$ and $a_{ii} \in \mathbb{R}$, $i \in N$, $\hat{\mu} = \mu(A)$ if and only if $\hat{\mu}$ is the Perron-Frobenius eigenvalue of a matrix $M_{re}(A)$. Moreover, all Geršgorin discs of the matrix $W^{-1}AW$ pass through $\hat{\mu}$, where w is the Perron-Frobenius eigenvector of $M_{re}(A)$ and $W = diag(w)$.*

Proof: ($\Rightarrow$) Assume that $\hat{\mu} = \mu(A)$. Hence, $\nu_A(\hat{\mu}) = 0$ is the right-most eigenvalue of matrix $Q_A(\hat{\mu})$. Using the Perron-Frobenius theorem for essentially non-negative irreducible matrices, there exists a positive eigenvector $\tilde{w}$ associated with the eigenvalue $\nu_A(\hat{\mu})$. Since $\hat{\mu} \geq a_{ii}, i \in N$, it follows $Q_A(\hat{\mu}) = -\hat{\mu}I + M_{re}(A)$. Therefore, $Q_A(\hat{\mu})\tilde{w} = (-\hat{\mu}I + M_{re}(A))\tilde{w} = 0$, i.e., $M_{re}(A)\tilde{w} = \hat{\mu}\tilde{w}$. Hence, μ is the Perron-Frobenius eigenvalue of matrix $M_{re}(A)$.
($\Leftarrow$) Let $(\hat{\mu}, w)$ be the Perron-Frobenius eigenpair of essentially non-negative matrix $M_{re}(A)$. Then, $\hat{\mu}$ is the right-most eigenvalue of matrix $M_{re}(A)$ and $w = [w_1, w_2, \ldots, w_n] > 0$. Since for $\alpha \geq a_{ii}, |\alpha - a_{ii}| = \alpha - a_{ii},\ i \in N$, then $Q_A(\alpha) = -\alpha I + M_{re}(A)$ and $\nu_A(\alpha) = \hat{\mu} - \alpha$ is the right-most eigenvalue of matrix $Q_A(\alpha)$. Therefore, if $\alpha > \hat{\mu}$, then $\nu_A(\alpha) < 0$ and α is not in the minimal Geršgorin set of matrix A. However, if $\alpha = \hat{\mu}$ then $\nu_A(\alpha) = 0$ and consequently, $\hat{\mu} = \mu(A)$.

Now, from $M_{re}(A)w = \mu(A)w$, we obtain $\mu(A) = a_{ii} + \sum_{j \neq i}^{n} \frac{|a_{ij}|w_j}{w_i},\ i \in N$. Therefore, all Geršgorin discs of matrix $W^{-1}AW$ pass through $\hat{\mu}$, i.e., $\hat{\mu} \in \partial\Gamma_i(W^{-1}AW),\ i \in N$. ❑

Beside theoretical importance, the abscissa of the minimal Geršgorin set has an important role in practice. For example, it can be used in the theory of dynamical systems.

First order time varying dynamical systems can be represented by the system of ordinary differential equations:

$$\frac{d}{dt}x(t) = F(t, x(t)),\ x(t_0) = x_0, \tag{3.3}$$

where variable $x(t) \in \mathbb{R}^n$ represents a state of the system in time $t \geq 0$.
If $F(t, x(t)) = F(x(t))$, the system is autonomous. A point $x^* \in \mathbb{R}^n$ is the equilibrium point if $F(x^*) = 0$.

If dynamical system is linear, system (3.3) can be written as:

$$\dot{x}(t) = Ax(t),\ x(t_0) = x_0, \tag{3.4}$$

where $A \in \mathbb{R}^{n,n}$ is a constant matrix. While dynamical systems in general do not have closed-form solutions, linear dynamical systems can be solved exactly and its solution is given by:

$$x(t) = e^{At}x_0. \tag{3.5}$$

If A is a nonsingular matrix, equilibrium point $x^* = 0$ is unique and its dynamical properties are determined by evolution function $\phi_A(t) = ||e^{At}||,\ t \geq 0$, where $||\cdot||$ is induced matrix norm.

Reactivity of the observed linear dynamical system represents initial growth rate of $||e^{At}||$, i.e.,

$$\frac{d^+}{dt}\Big[\phi_A(t)\Big]\Big|_{t=0} = \lim_{t\to 0}\frac{||e^{At}|| - ||e^0||}{t} = \lim_{t\to 0}\frac{||I + tA|| - 1}{t}. \tag{3.6}$$

In addition, if the reactivity is less than zero, then dynamical system is exponentially stable. If $||\cdot|| = ||\cdot||_\infty$, we obtain:

$$\lim_{t\to 0}\frac{||I + tA||_\infty - 1}{t} = \lim_{t\to 0}\frac{\max\limits_{k\in N}\Big\{|1 + ta_{kk}| - 1 + t\sum\limits_{j\neq k}^{n}|a_{kj}|\Big\}}{t} \tag{3.7}$$

i.e., the reactivity is:

$$\max_{k\in N}\Big\{\mathrm{Re}(a_{kk}) + \sum_{j\neq k}^{n}|a_{kj}|\Big\} = \gamma(A). \tag{3.8}$$

Therefore, the reactivity is equal to the abscissa of the Geršgorin set. From [60], we have $||e^{At}||_\infty \leq e^{\gamma(A)t}$.

Using the induced a weighted norms $||\cdot||_X$, where $X \in \mathbb{R}^{n,n}$ is diagonal matrix whose diagonal elements are positive and the fact $\mu(A) = \inf\limits_X \gamma(X^{-1}AX)$, we obtain:

$$\inf_X ||e^{At}||_X \leq e^{\mu(A)t}.$$

Therefore, if the abscissa of the minimal Gersšgorin set of A is negative, there exists a weighted norm $||\cdot||_X$ such that the system (3.4) is exponentially stable.

For example, if $A = \begin{bmatrix} -12 & 2 & 5 \\ 1 & -10 & 2 \\ 0 & -4 & -2 \end{bmatrix}$, then $\gamma(A) = 2$ and we do not have proper information whether the system determined by A is exponentially stable or not. However, $\mu(A) = -1.7574$ and it implies that the system is exponentially stable (Figure 3.1).

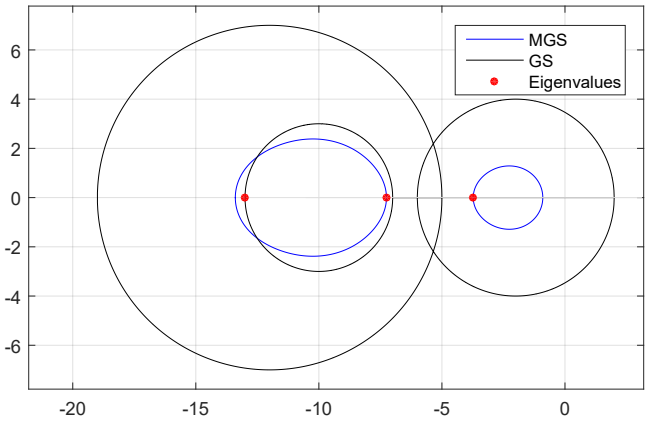


Figure 3.1: The Geršgorin (GS) vs. the minimal Geršgorin set (MGS).

3.2 Minimal Geršgorin set vs. numerical range

Given a matrix $A \in \mathbb{C}^{n,n}$, $n \in \mathbb{N}$. Analogies between the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ and the numerical range $W(A)$ of A are:

1) non-empty and localization sets for $\sigma(A)$ (it follows from (1.18) and Theorem 1.7.4);

2) compact sets (bounded and closed) in $\mathbb{C}$ (Theorems 1.5.4 and 1.7.1);

3) both are homogeneous, i.e., $\Gamma^{\mathcal{R}}(cA) = c\Gamma^{\mathcal{R}}(A)$ and $W(cA) = cW(A)$, $c \in \mathbb{C}$ (Theorems 1.5.4 and 1.7.3);

4) $\Gamma^{\mathcal{R}}(cI + A) = c + \Gamma^{\mathcal{R}}(A)$ and $W(cI + A) = c + W(A)$, $c \in \mathbb{C}$, I is the $n \times n$ identity matrix (Theorems 1.5.4 and 1.7.3).

Differences between the minimal Geršgorin set and the numerical range are:

i) $W(A)$ is a connect set for all $A \in \mathbb{C}^{n,n}$ (Theorem 1.7.1), but in general, $\Gamma^{\mathcal{R}}(A)$ may not be (Example 3.2.1, Figure 3.2);

ii) $W(A)$ is a convex set for all $A \in \mathbb{C}^{n,n}$ (Theorem 1.7.2), but in general, $\Gamma^{\mathcal{R}}(A)$ may not be (Example 3.2.1, Figure 3.2);

iii) $W(A+B) \subseteq W(A) + W(B)$ for all $A,\ B \in \mathbb{C}^{n,n}$ (Theorem 1.7.3), but in general, $\Gamma^{\mathcal{R}}(A+B) \not\subseteq \Gamma^{\mathcal{R}}(A) + \Gamma^{\mathcal{R}}(B)$ (e.g., for $A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, $\Gamma^{\mathcal{R}}(A+B) = \{1, 2\}$, but $\Gamma^{\mathcal{R}}(A) + \Gamma^{\mathcal{R}}(B) = \{3\}$);

iv) $W(A)$ is invariant under unitary transformation (i.e., $W(A) = W(UAU^*)$), where $U \in \mathbb{C}^{n,n}$ is arbitrary unitary matrix (Theorem 1.7.3), but in general, $\Gamma^{\mathcal{R}}(A)$ may not be (Example 3.2.1, Figure 3.4);

v) $\Gamma^{\mathcal{R}}(A)$ is invariant under similarity transformation by nonsingular and non-negative diagonal matrices (i.e., $\Gamma^{\mathcal{R}}(A) = \Gamma^{\mathcal{R}}(XAX^{-1})$, where $X \in \mathbb{R}^{n,n}$ is an arbitrary diagonal matrix such that $\det(X) \neq 0$, and $X \geq 0$), but in general, $W(A)$ may not be (Example 3.2.1, Figure 3.3);

vi) $\Gamma^{\mathcal{R}}(A)$ and $W(A)$ are in general relation, i.e., $\Gamma^{\mathcal{R}}(A) \not\subseteq W(A)$ and $W(A) \not\subseteq \Gamma^{\mathcal{R}}(A)$ (Example 3.2.1, Figure 3.2).

Example 3.2.1 *There are given matrices:*

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 10 \end{bmatrix},\ X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ and } U = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The minimal Geršgorin set (blue line), the numerical range (green line) and eigenvalues (red dots) of A are presented in Figure 3.2.
The minimal Geršgorin set (blue line), the numerical range (green line) and eigenvalues (red dots) of XAX^{-1} are presented in Figure 3.3.
The minimal Geršgorin set (blue line), the numerical range (green line) and eigenvalues (red dots) of UAU^* are presented in Figure 3.4.

The minimal Geršgorin set and the numerical range are both localization sets for the spectrum of a given matrix. As it is presented, they are incomparable

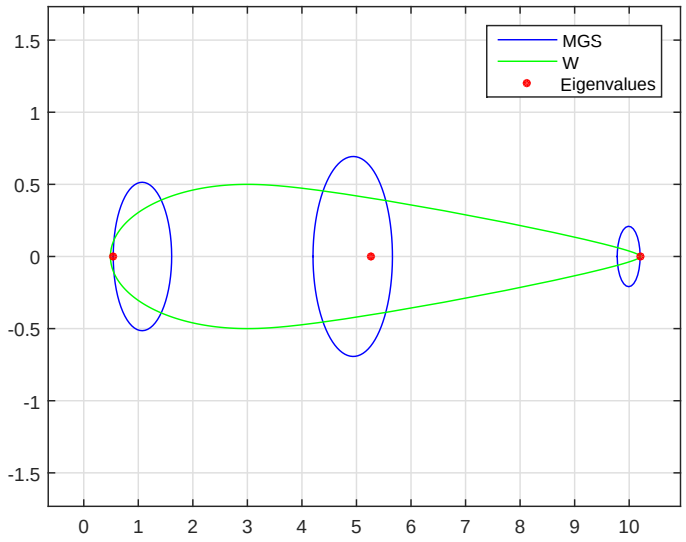


Figure 3.2: Comparison of $\Gamma^{\mathcal{R}}(A)$ and $W(A)$.

sets, but we can compare them after diagonal scaling by nonsingular non-negative diagonal matrices. Namely, from Theorem 1.7.8, we have the following.

Theorem 3.2.2 *For a given matrix $A \in \mathbb{C}^{n,n}$ holds:*

$$W(A) \subseteq co(\Gamma(A) \cup \Gamma(A^T)).$$

Proof: Using an inequality $\dfrac{r_i(A) + r_i(A^T)}{2} \leq \max\{r_i(A), r_i(A^T)\}$ and Theorem 1.7.8, it follows:

$$\left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \frac{r_i(A) + r_i(A^T)}{2} \right\} \subseteq \Gamma_i(A) \cup \Gamma_i(A^T), \ i \in N.$$

So, $W(A) \subseteq \text{co}(\Gamma(A) \cup \Gamma(A^T))$. ❑

Although it is simple to prove, up to the author's knowledge, the following result is not well-known in a literature.

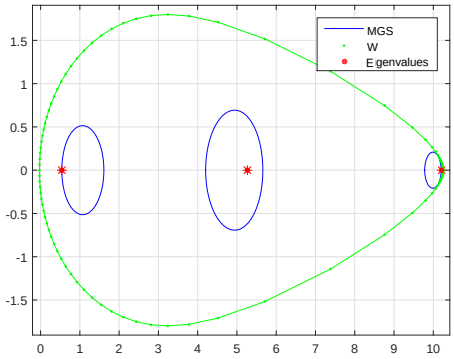


Figure 3.3: Comparison of $\Gamma^{\mathcal{R}}(XAX^{-1})$ and $W(XAX^{-1})$.

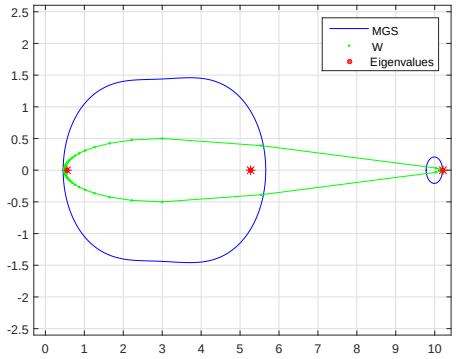


Figure 3.4: Comparison of $\Gamma^{\mathcal{R}}(UAU^*)$ and $W(UAU^*)$.

Corollary 3.2.3 *For a given matrix $A \in \mathbb{C}^{n,n}$ holds:*

$$\bigcap_{X \in D_n} W(X^{-1}AX) \subseteq co(\Gamma^{\mathcal{R}}(A)),$$

where D_n is the family of nonsingular and non-negative diagonal $n \times n$ matrices.

Proof: Using the previous theorem and the fact $\Gamma^{\mathcal{R}}(A) = \Gamma^{\mathcal{R}}(A^T)$, (Theorem 1.5.4), the statement holds trivially. ❑

While the computation of $W(A)$ can be performed efficiently using the algorithm presented in the next section, the computation of scaled $W(A)$ is more complex task. So, the previous corollary represents also an interesting motivation for drawing a set that contains a convex envelope of the minimal Geršgorin set whose position in the complex plane gives an information of the impact of proper diagonal scaling on the numerical range of a given matrix.

Finally, as it is presented in the previous section, the reactivity of a dynamical system is equal to the abscissa of the minimal Geršgorin set in optimal weighted infinity norm. Analogously, by using norm $||\cdot||_2$, it can be shown that the reactivity is equal to the abscissa of numerical range, i.e., $\omega(A)$, ([31], Corollary 5.5.26).

3.3 Algorithm for computing the numerical range

As $W(A)$ is a compact and connected set in $\mathbb{C}$ (Theorem 1.7.1), the set $\mathrm{Re}(W(A))$ is closed real interval $[a, b]$, where a and b are the smallest and largest eigenvalue of a Hermitian matrix $H(A) = \dfrac{A + A^*}{2}$, respectively (Theorem 1.7.5). If $y \in \mathbb{C}^n$ is a normalized eigenvector (i.e., $y^*y = 1$) corresponding to the maximal eigenvalue b of $H(A)$, then by the definition, the point $\omega := y^*Ay$ is in $W(A)$. Moreover, ω lies on the boundary of $W(A)$ because $\mathrm{Re}(\omega) = b$.

Let consider Hermitian matrices

$$H(e^{\mathrm{i}\theta_j}A) = \frac{e^{\mathrm{i}\theta_j}A + e^{-\mathrm{i}\theta_j}A^*}{2},$$

where $\{\theta_0, \theta_1, ..., \theta_{m-1}\}$ is the set of distinct angles, $m \in \mathbb{N}$. Now, for each matrix $H(e^{\mathrm{i}\theta_j}A)$, we can find the largest eigenvalue $\lambda(\theta_j)$ and its corresponding normalized eigenvector y_j, i.e.,

$$y_j^* H(e^{\mathrm{i}\theta_j}A)y_j = \lambda(\theta_j),\ y_j^*y_j = 1,\ j \in \{0, 1, ..., m-1\}. \tag{3.9}$$

Again, it is obvious that a complex number $y_j^*Ay_j$ belongs to $W(A)$, but as $\lambda(\theta_j)$ is the largest eigenvalue of $H(e^{\mathrm{i}\theta_j}A)$, it follows that

$$\omega_j := y_j^*Ay_j \in \partial W(A),\ j \in \{0, 1, ..., m-1\}. \tag{3.10}$$

So, using the set of points $\{\omega_0, \omega_1, ..., \omega_{m-1}\}$, we obtain a discrete approximation of the boundary of $W(A)$. Notice that the associated tangent lines through points $\omega_j, j \in \{0, 1, ..., m-1\}$, make angles $\dfrac{\pi}{2} - \theta_j$ with the real axis. It is a corollary of the fact that

$$W(H(e^{\mathrm{i}\theta_j}A)) = \mathrm{Re}(W(e^{\mathrm{i}\theta_j}A)) = \mathrm{Re}(e^{\mathrm{i}\theta_j}W(A)).$$

3.4 Algorithms for computing a convex approximation of MGS

In this section, we present two algorithms for computing an approximation of a convex hull of the minimal Geršgorin set. We will use the similar technique as it was used for computing the numerical range.

Algorithm for calculating the numerical range $W(A)$

Input: A, m

1: **for** $j = 0 : (m-1)$ **do**
2: Set $\theta_j = \frac{2\pi j}{m}$;
3: Compute $\lambda(\theta_j)$ and y_j as (3.9);
4: Compute ω_j as (3.10);
5: **end for**

Output: $\{\omega_0, \omega_1, ..., \omega_{m-1}\}$

Let an irreducible matrix $A \in \mathbb{C}^{n,n}$ and $m \in \mathbb{N},\ m \geq 3$, be given. Following Theorem 3.1.2, we can find upper limit for the abscissa of the minimal Geršgorin set as the Perron-Frobenius eigenvalue of $M_{re}(A)$. Furthermore, we can find the Perron-Frobenius eigenvalues of all matrices $M_{re}(Ae^{-\mathrm{i}\theta_k})$, where $\theta_k = \dfrac{2k\pi}{m}$, $k \in \{0, 1, ..., m-1\}$. In that way, we get points z_k as:

$$z_k := \alpha(M_{re}(Ae^{-\mathrm{i}\theta_k}))e^{\mathrm{i}\theta_k}. \tag{3.11}$$

Let the line n_k orthogonal to vector $\vec{z}_k,\ z_k \in \mathbb{C},\ k \in \{0, 1, ..., m-1\}$, be given as

$$n_k : \begin{cases} y = \mathrm{Im}(z_k), & \text{if } \mathrm{Re}(z_k) = 0, \\ x = \mathrm{Re}(z_k), & \text{if } \mathrm{Im}(z_k) = 0, \\ y - \mathrm{Im}(z_k) = -\frac{1}{\tan(\theta_k)}(x - \mathrm{Re}(z_k)), & \text{otherwise.} \end{cases} \tag{3.12}$$

We obtain points ω_k as intersections of lines n_k and n_{k+1}, for $k \in \{0, 1, ..., m-1\}$, with $\theta_m := 0,\ z_m := z_0$ and $n_m := n_0$, i.e.,

$$\omega_k := \begin{cases} \mathrm{Re}(z_{k+1}) - \tan(\theta_{k+1})\mathrm{Im}(z_k - z_{k+1}) + \mathrm{i}\mathrm{Im}(z_k), & \text{if } \mathrm{Re}(z_k) = 0, \\ \mathrm{Re}(z_k) - \tan(\theta_k)\mathrm{Im}(z_{k+1} - z_k) + \mathrm{i}\mathrm{Im}(z_{k+1}), & \text{if } \mathrm{Re}(z_{k+1}) = 0, \\ \mathrm{Re}(z_k) - \tan(\theta_k)(y_k - \mathrm{Im}(z_k)) + \mathrm{i}y_k, & \text{otherwise,} \end{cases} \tag{3.13}$$

where

$$y_k = \frac{\mathrm{Re}(z_{k+1} - z_k) + \tan(\theta_{k+1})\mathrm{Im}(z_{k+1}) - \tan(\theta_k)\mathrm{Im}(z_k)}{\tan(\theta_{k+1}) - \tan(\theta_k)},\ k \in \{0, 1, ..., m-1\}.$$

A polygon $P^{(m)}(A)$ with vertices $\omega_0, \omega_1, ..., \omega_{m-1}$ is convex and contains the minimal Geršgorin set.

If we use **Noda iteration** for computing the spectral abscissa, we obtain an additionally a sequence of polygons, $P^{(m)}_{(l)}(A)$, where l is a given number of iterations, that monotonically improve. Clearly, the following theorem holds.

Theorem 3.4.1 *Given an irreducible matrix $A \in \mathbb{C}^{n,n}$ and $m, l \in \mathbb{N},\ m \geq 3$, then:*

$$\sigma(A) \subseteq \Gamma^{\mathcal{R}}(A) \subseteq P^{(m)}(A) \subseteq P^{(m)}_{(l+1)}(A) \subseteq P^{(m)}_{(l)}(A).$$

We notice that $P^{(m)}_{(l)}(A) \to P^{(m)}(A),\ l \to \infty$.

Algorithm for calculating vertices of the polygon $P^{(m)}(A)$

Input: A, m

1: Set $\theta_0 = 0$;
2: Compute $z_0 = \alpha(M_{re}(A))$ and set $\omega_0 = z_0$;
3: **for** $k = 1 : m$ **do**
4: Set $\theta_k = \frac{2k\pi}{m}$;
5: Compute z_k as in (3.11);
6: Compute ω_{k-1} as in (3.13);
7: **end for**

Output: $\{\omega_0, \omega_1, ..., \omega_{m-1}\}$

Algorithm for calculating vertices of the polygon $P^{(m)}_{(l)}(A)$

Input: A, m, l

1: Set $\theta_0 = 0$;
2: Compute $z_0 = \alpha(M_{re}(A))$ using l **Noda iterations** and set $\omega_0 = z_0$;
3: **for** $k = 1 : m$ **do**
4: Set $\theta_k = \frac{2k\pi}{m}$;
5: Compute z_k as in (3.11) using l **Noda iterations**;
6: Compute ω_{k-1} as in (3.13);
7: **end for**

Output: $\{\omega_0, \omega_1, ..., \omega_{m-1}\}$

Applications of the results from this section will be given in Chapter 4.

Chapter 4

Numerical results and conclusion

"In theory, theory and practice are the same. In practice, they are not."

Albert Einstein[19]

This chapter is organized as follows. In Section 4.1, algorithms for computing the minimal Geršgorin set are compared through several examples. Then, in Section 4.2, the numerical results about convex polygon that contains the minimal Geršgorin set are presented. Finally, in Section 4.3, benefits and improvements of new results are given in a short conclusion. All algorithms are implemented in MATLAB version R2018b and tested on 2.7 GHz Intel® Core™ i7 machine.

4.1 Curve tracing algorithms

In this section, we test all algorithms ({**e,i**}**MGS**{**s,p,t**}) on six examples and compare the performances of the novel ones with the performance of **eMGSs** algorithm that was the state of art. We notice that the performances of new approaches significantly accelerate convergence.

Example 4.1.1 *In the first example we test algorithms on the cyclic matrix of*

[17] Albert Einstein (1879-1955) was a German-born theoretical physicist who developed the theory of relativity, received the Nobel Prize in 1921.

a size $n = 4$*:*

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & i & 1 \\ 1 & 0 & 0 & -i \end{bmatrix},$$

setting the parameters of the algorithms to be: $tol = 10^{-12}$, $\tilde{\tau} = 2$, $h = 0.0254$, $N_s = 40$ *and* $N_t = 500$. *CPU times for all algorithms are presented in Table 4.1. The number of computed points for* ***eMGSs*** *and* ***iMGSs*** *is* 430, *for* ***eMGSp*** *and* ***iMGSp*** 436 *and for* ***eMGSt*** *and* ***iMGSt*** *it is* 2086. *Figure* 4.1 *shows the minimal Geršgorin set of* A *using all three approaches. Also, their corresponding zoomed versions around the orgin are presented. Comparing them, we notice that the algorithms* ***eMGSt*** *and* ***iMGSt*** *give more reliable approximation (zero belongs to the minimal Geršgorin set of* A*).*

MGS	**s**	**p**	**t**
e	$1.4667s$	$0.2728s$	$0.1102s$
i	$0.3721s$	$0.1203s$	$\mathbf{0.0282}\boldsymbol{s}$

Table 4.1: CPU times for Example 4.1.1.

Example 4.1.2 *In the second example we consider a parameter dependent triangular matrix* T_μ *of a size* $n = 20$ *defined as follows:*

$$T_\mu = \begin{bmatrix} \mu & 1 & 0 & \dots & 0 \\ 1 & 2\mu & 1 & \ddots & \vdots \\ 0 & 1 & 3\mu & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & 1 & 20\mu \end{bmatrix}.$$

For $\mu = 2$, *the results obtained by the algorithms with the parameters* $tol = 10^{-12}$, $\tilde{\tau} = 2$, $h = 0.0362$, $N_s = 600$ *and* $N_t = 2500$ *are presented in Table* 4.2. *Figure* 4.2 *shows the approximation of the minimal Geršgorin set of the matrix* T_2 *obtained by (a)* ***eMGSs/iMGSs*** *algorithm (*2211 *points), (b)* ***eMGSp*** *algorithm (*2202 *points) and (c)* ***eMGSt/iMGSt*** *algorithm (*4716 *points). The algorithm* ***iMGSp*** *does not give any results in the observed period.*

MGS	s	p	t
e	$1.5552s$	$3.6399s$	$\mathbf{0.1091s}$
i	$3.3026s$	/	$0.3738s$

Table 4.2: CPU times for Example 4.1.2.

Example 4.1.3 *In this example we test the algorithms on the matrix twisted_ demo(n) of a size* $n = 50$ *from the Matrix Market repository ([6]). The results of the algorithms with parameters* $tol = 10^{-12}$, $\tilde{\tau} = 2$, $h = 0.0431$, $N_s = 100$ *and* $N_t = 500$ *are presented in Table* 4.3. *Figure* 4.3 *shows the results obtained by (a)* ***eMGSs*** *algorithm (*462 *points), (b)* ***eMGSp*** *algorithm (*464 *points) and (c)* ***eMGSt/iMGSt*** *algorithm (*1200 *points).*

MGS	s	p	t
e	$2.9748s$	$1.0106s$	$\mathbf{0.1705s}$
i	/	/	$0.2055s$

Table 4.3: CPU times for Example 4.1.3.

Example 4.1.4 *In this example we implement the algorithms on the Leslie matrix:* $L = diag(b \cdot (1:n-1).\hat{}(-1), -1) + a \cdot [\xi.\hat{}(1:n); zeros(n-1, n)], L(1,1) = 0$, *for values* $a = 0.1$, $b = 0.2$, $\xi = 0.95$ *and* $n = 70$. *The results obtained with parameters* $tol = 10^{-12}$, $\tilde{\tau} = 2$, $h = 0.0036, N_s = 100$ *and* $N_t = 200$ *are presented in Table* 4.4. *Figure* 4.4 *represents the approximation of the minimal Geršgorin set of the Leslie matrix obtained by (a)* ***eMGSs/iMGSs*** *algorithm (*315 *points), (b)* ***iMGSp*** *algorithm (*315 *points) and (c)* ***eMGSt/iMGSt*** *algorithm (*602 *points).*

MGS	s	p	t
e	$38.9456s$	/	$2.3955s$
i	$1.0306s$	$0.2796s$	$\mathbf{0.2576s}$

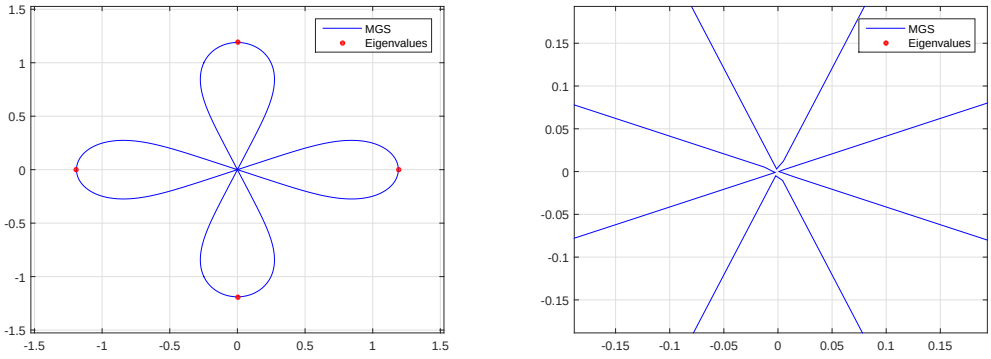
Table 4.4: CPU times for Example 4.1.4.

Example 4.1.5 *In this example we use the Tolosa matrix of a size* $n = 340$ *from the Matrix Market repository ([6]). This matrix is sparse, highly non-normal of a medium size. The parameters are set as:* $tol = 10^{-12}$, $\tilde{\tau} = 2$, $h = 9.381$, $N_s = 50$ *and* $N_t = 300$, *which produces* 295 *points for* ***eMGSs/iMGSs*** *algorithm,* 302 *for* ***iMGSp*** *and* 866 *points for* ***eMGSt/iMGSt****. The results of testing are presented in Figure 4.5 and corresponding CPU times are given in Table 4.5.*

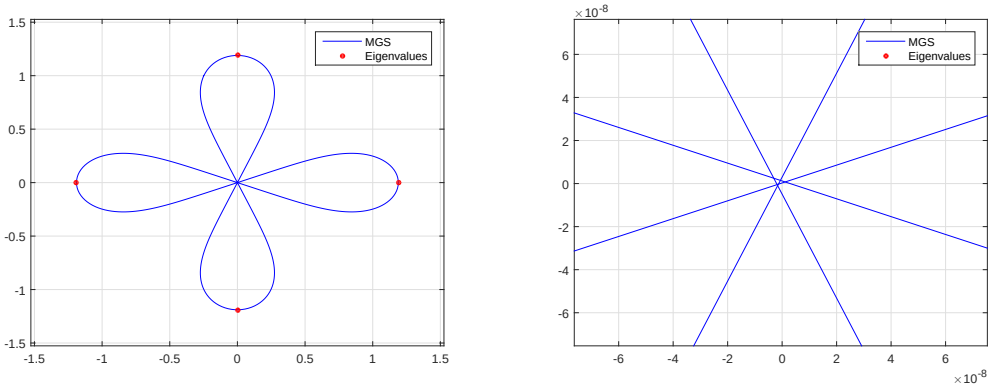
MGS	**s**	**p**	**t**
e	$86.6712s$	/	$11.5196s$
i	$19.9532s$	$2.7571s$	$\mathbf{2.4453s}$

Table 4.5: CPU times for Example 4.1.5.

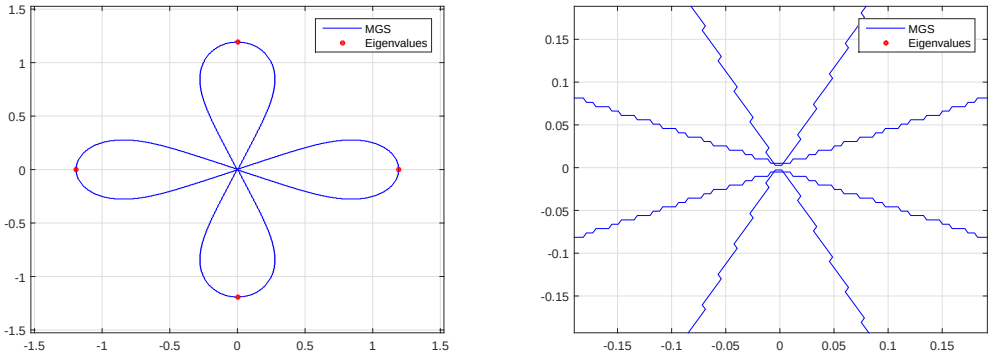
Example 4.1.6 *Finally, in the last example, we test all algorithms on two matrices of a large size. For the Orr-Sommerfeld matrix of a size* $n = 1000$ *([6]) with* $N_t = 400$, *the CPU time for* ***iMGSt*** *is* $51.41335s$ (546 *points, see Figure* 4.6). *For the Poisson matrix of a size* $n = 2500$ *(MATLAB gallery) with* $N_t = 300$, *CPU for* ***iMGSt*** *is* $240.1962s$ (902 *points, see Figure* 4.7). *Other algorithms do not give results in the observed period.*



(a) ***MGSs**

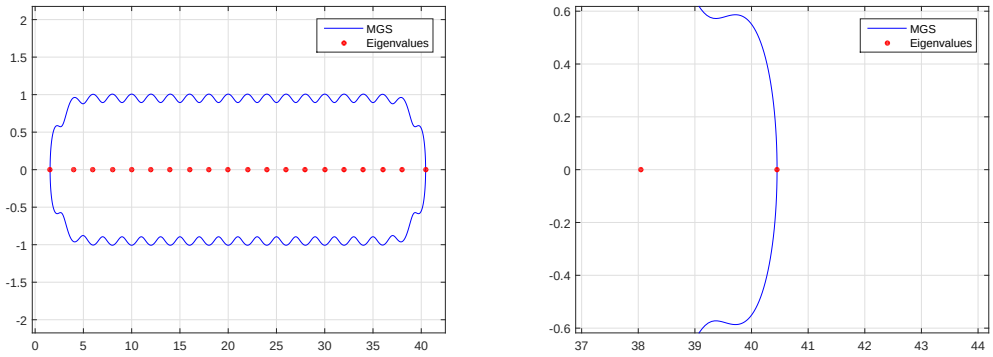


(b) ***MGSp**

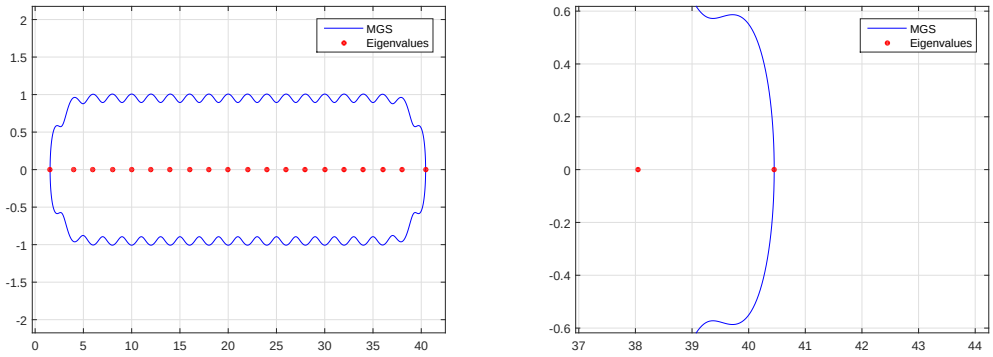


(c) ***MGSt**

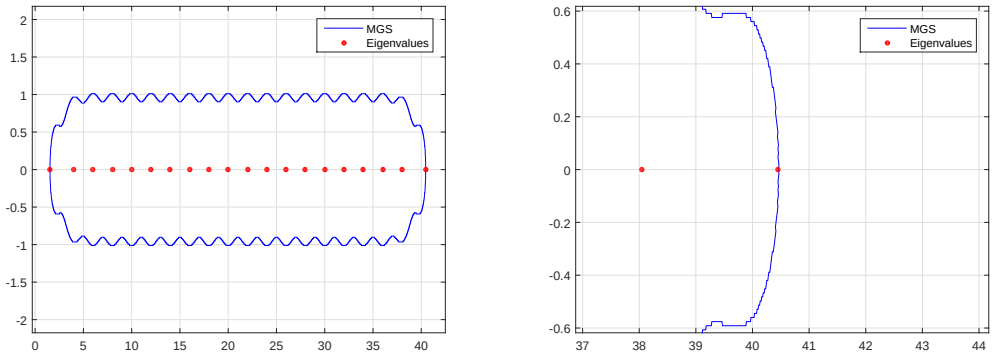
Figure 4.1: The results of the algorithms for the the cyclic matrix A from Example 4.1.1: complete plot and plot zoomed around the origin.



(a) ***MGSs**

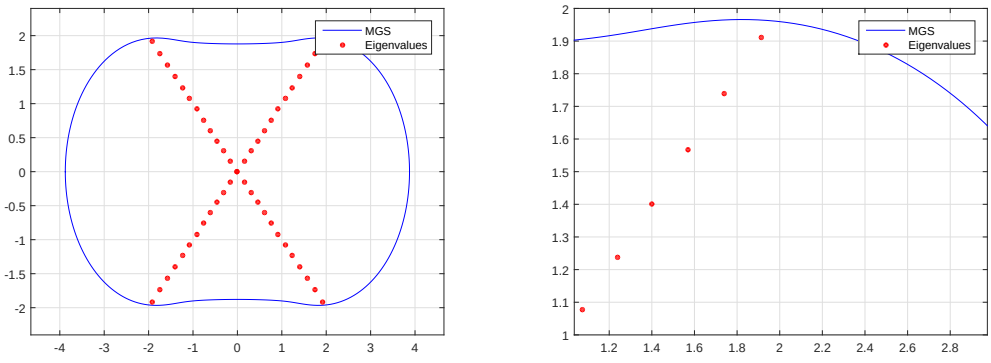


(b) **eMGSp**

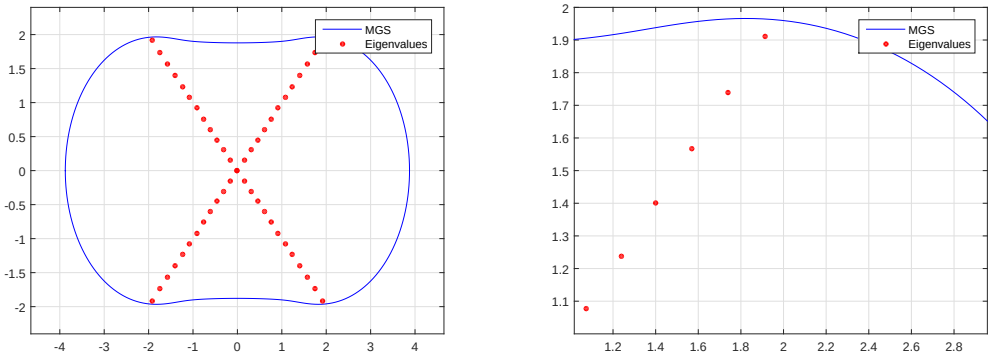


(c) ***MGSt**

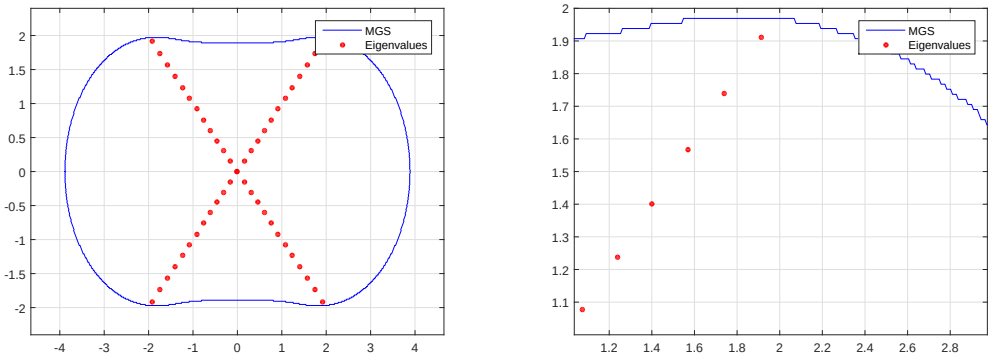
Figure 4.2: The results of the algorithms for the tridiagonal matrix T_2: complete plot and plot zoomed around the rightmost eigenvalue.



(a) **eMGSs**

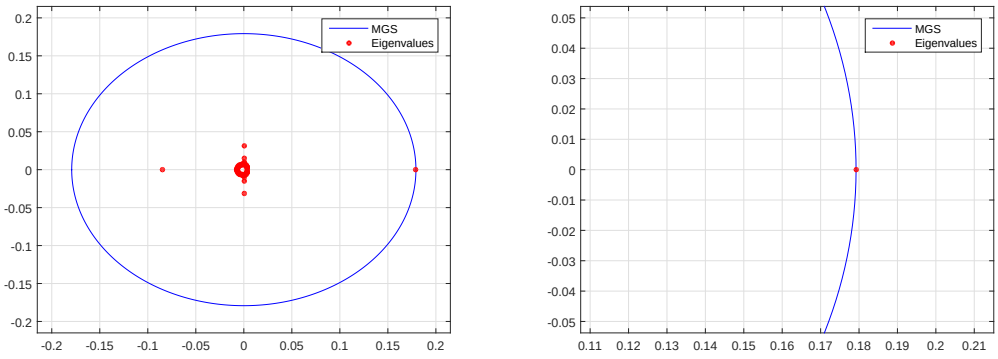


(b) **eMGSp**

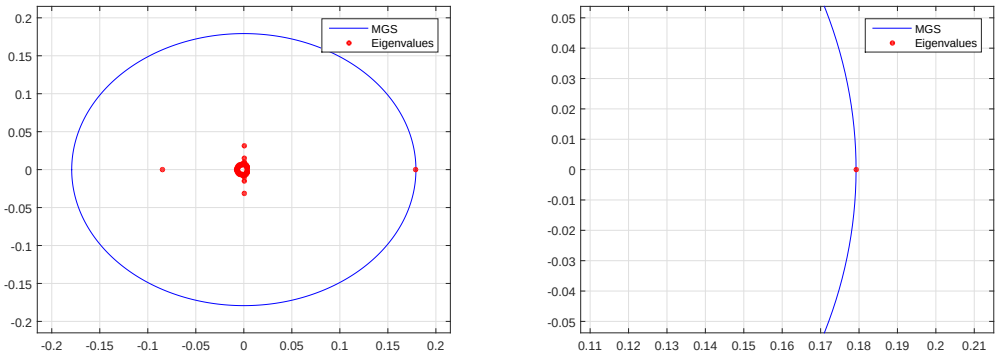


(c) ***MGSt**

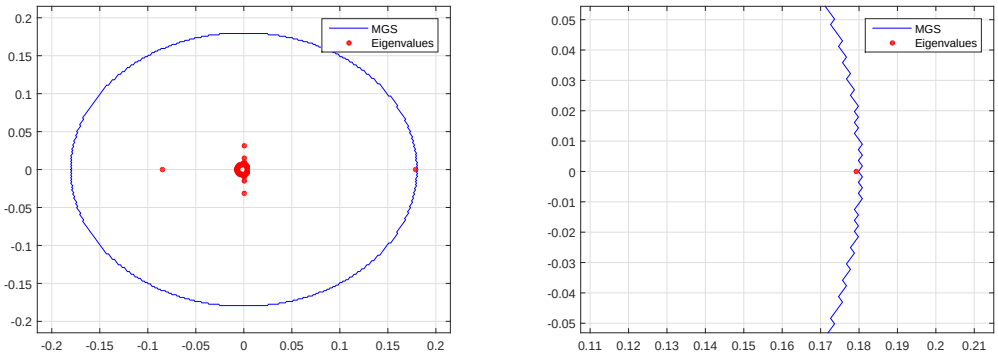
Figure 4.3: The results of the algorithms for the twisted matrix of a size $n = 50$: complete plot and plot zoomed around the point $z = 2 + 2\mathrm{i}$.



(a) ***MGSs**

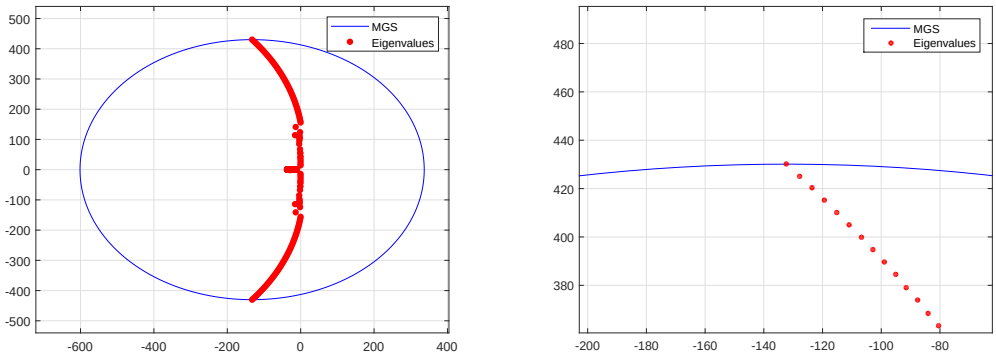


(b) **iMGSp**

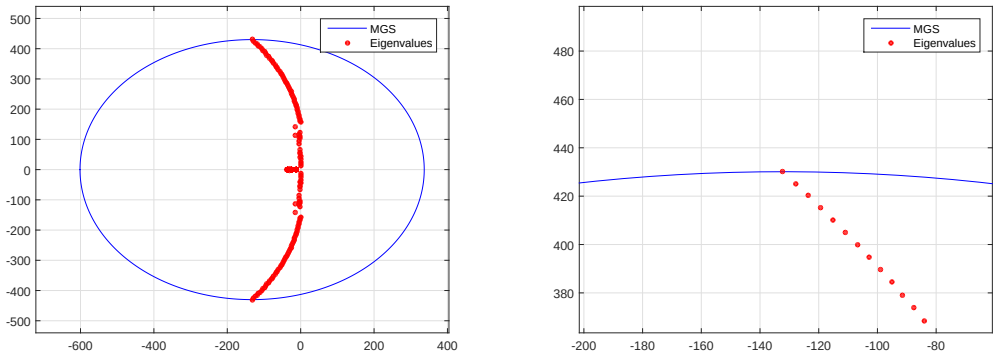


(c) ***MGSt**

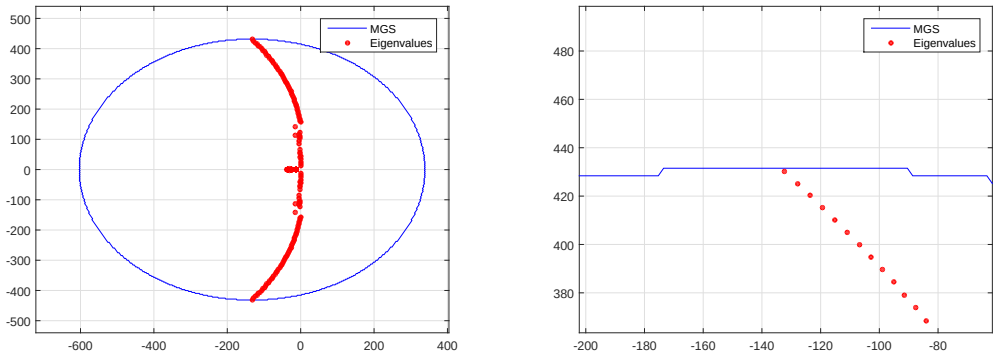
Figure 4.4: The results of the algorithms for the Leslie matrix of a size $n = 70$: complete plot and plot zoomed around the rightmost eigenvalue.



(a) ***MGSs**



(b) **iMGSp**



(c) ***MGSt**

Figure 4.5: The results of the algorithms for the Tolosa matrix of a size $n = 340$: complete plot and plot zoomed around the point $z = -130 + 430\mathrm{i}$.

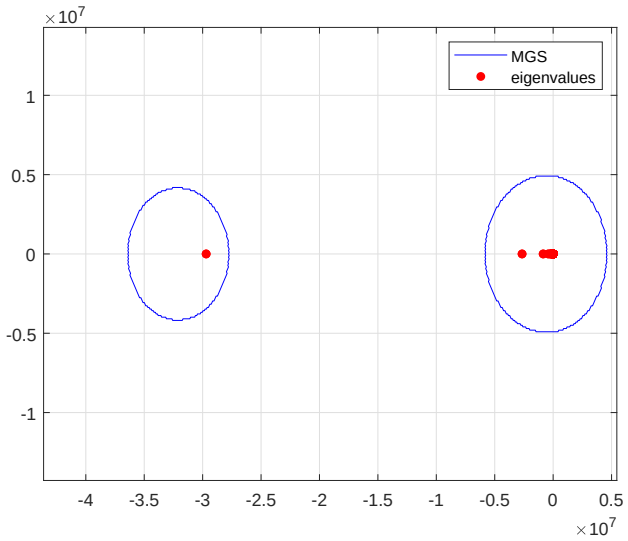


Figure 4.6: The result of the algorithm **iMGSt** for the Orr-Sommerfeld matrix of a size $n = 1000$.

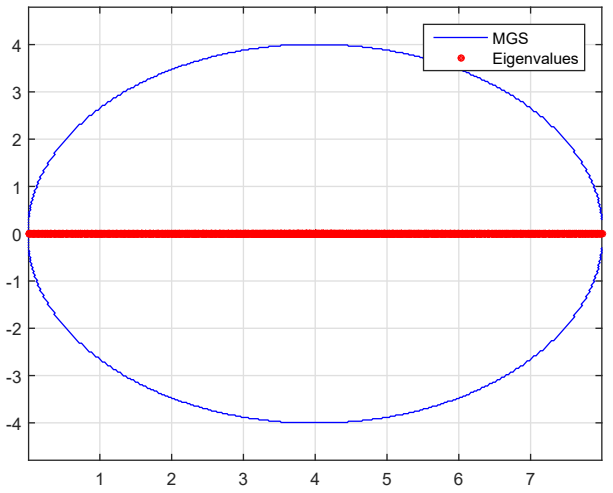
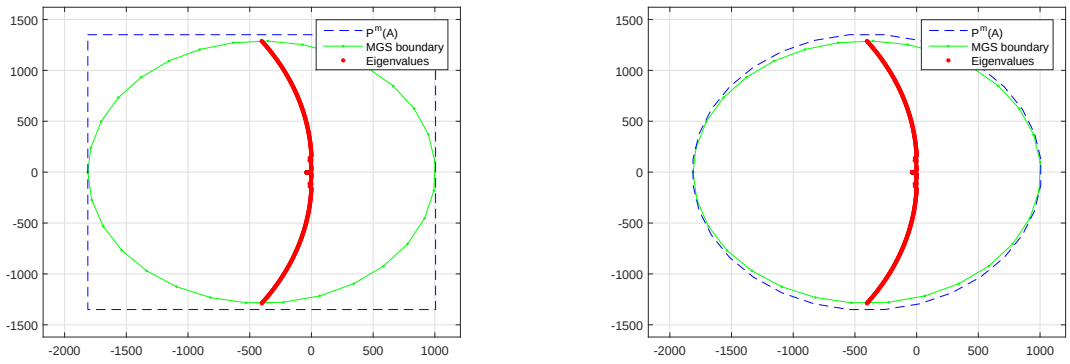


Figure 4.7: The result of the algorithm **iMGSt** for the Poisson matrix of a size $n = 2500$.

4.2 Convex polygon

In this section, we illustrate an interesting and practical result, specially for matrices of large sizes: localization of eigenvalues of a given matrix by the convex polygon that contains MGS. This approach is significantly numerically cheaper than computing the minimal Geršgorin set itself. Here, we compare it to the algorithm **eMGSs**.

Example 4.2.1 *For the Tolosa matrix A of a dimension $n = 1090$, the algorithm **eMGSs** ([40]) with the parameters $tol = 10^{-12}, \tilde{\tau} = 2$ and $N_s = 5$, gives the approximation of the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$ by 30 points in 114.7494s. For $m = 4$, the polygons $P^{(m)}(A)$ and $P^{(m)}_{(3)}(A)$ are rectangles, see Figure 4.8(a), and they can be found in 0.7643s and 0.3696s, respectively. For $m = 32$, the approximations by the convex polygons $P^{(m)}(A)$ and $P^{(m)}_{(3)}(A)$ are very close to the minimal Geršgorin set $\Gamma^{\mathcal{R}}(A)$, see Figure 4.8(b), and require only 4.3261s and 1.5987s, respectively.*

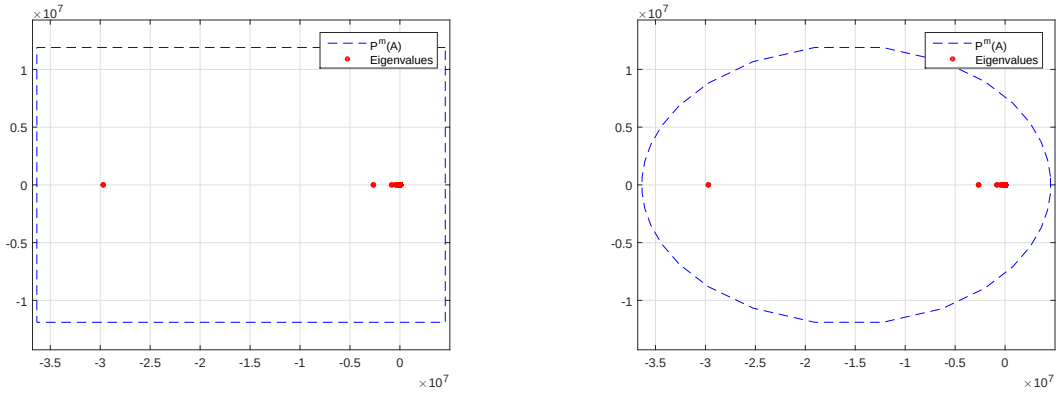


(a) $m = 4$ (b) $m = 32$

Figure 4.8: Polygon $P^{(m)}(A)$, the minimal Geršgorin set and eigenvalues of the Tolosa matrix of a size $n = 1090$.

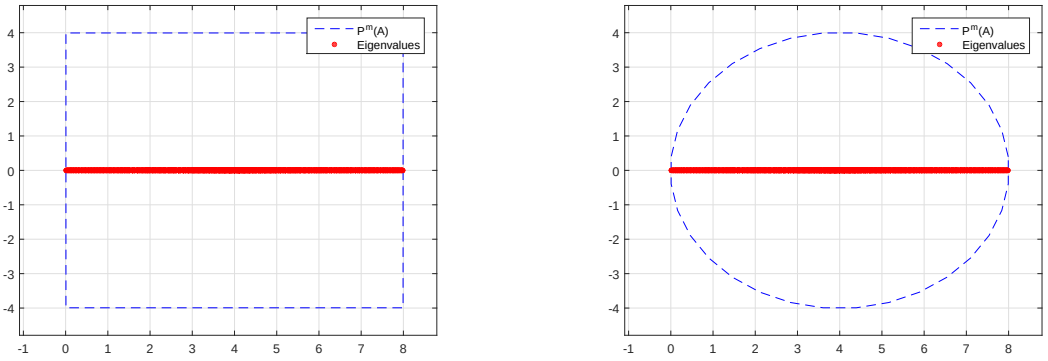
Example 4.2.2 *In this example, we test the algorithms $P^{(m)}(A)$ and $P^{(m)}_{(3)}(A)$ on the Orr-Sommerfeld matrix of a dimension $n = 1000$ ([6]) and the Poisson matrix of a dimension $n = 2500$ (MATLAB gallery). CPU times for the Orr-Somerfeld matrix are: 1.5185s and 0.2668s ($m = 4$, Figure 4.9(a)) and 8.6671s*

and 1.9416s *(*$m = 32$*, Figure* 4.9*(b)). For the Poisson matrix, CPU times are:* 0.5906s *and* 0.0564s *(*$m = 4$*, Figure* 4.10*(a)) and* 3.4579s *and* 0.2074s *(*$m = 32$*, Figure* 4.10*(b)). The algorithm* ***eMGSs*** *does not give results for both matrices in the observed period of* 10 *minutes.*



(a) $m = 4$ (b) $m = 32$

Figure 4.9: Polygon $P^{(m)}(A)$ and eigenvalues of the Orr-Sommerfeld matrix of a size $n = 1000$.



(a) $m = 4$ (b) $m = 32$

Figure 4.10: Polygon $P^{(m)}(A)$ and eigenvalues of the Poisson matrix of a size $n = 2500$.

4.3 Conclusion

In this paper, new algorithms for computing the minimal Geršgorin set are presented, together with convex polygon that contains it. Improvements of new results are illustrated through numerical examples.

New algorithms for computing the minimal Geršgorin set have several important adventages. First, new methods are significantly faster. As it is presented in the examples, the run time of new algorithms outperforms the existing algorithms. We can see that for all tested matrices, the overall best results were obtained by using the algorithms **eMGSt** and **iMGSt**. Furthermore, for some test matrices of large sizes, **eMGSs** did not produce any result in the observed period of time (**Example 4.1.6**).

Second, new algorithms are simpler for an implementation. For example, the algorithms which use triangular approach for curve tracing are straightforward since they do not depend on many parameters (the only required information is accuracy and the number of triangular grid points). All other necessary information is computed automatically.

Third, new approaches are more reliable. The algorithms **eMGSt** and **iMGSt** produce the polygons that always contain a desired localization set. In **Example 4.1.1**, belonging of the origin to the minimal Geršgorin set of the observed cyclic matrix A is not correctly detected by **eMGSs**. Furthermore, the fact that in this example $\Gamma^{\mathcal{R}}(A)$ consists of only one connected component (which can be an important information for counting the number of eigenvalues in the localization set) is not satisfied.

Finally, in addition to the algorithms for computing the minimal Geršgorin set, the algorithms for construction of its convex hull are developed. They are easy and very practical for usage, specially for matrices of large sizes and can provide useful information in applications. Also, in some cases, the convex polygon can be a very precise approximation of the minimal Geršgorin set (**Example 4.2.1** (b)).

Bibliography

[1] T. Akenine-Möller, E. Haines and N. Hoffman, *Real-Time Renering 3rd Edition*, A. K. Peters, Ltd, Natick, MA, USA, 2008.

[2] E. L. Allgower and K. Georg, *Continuation and path following,* Acta numerica, $1 - 64, 1993$.

[3] E. L. Allgower and K. Georg, *Numerical path following,* Colorado State University, 1994.

[4] C. Bekes and E. Gallopoulos, *COBRA: A hybrid method for computing the matrix pseudo spectrum*, In Copper Mountain Conference on Iterative Methods, 1998.

[5] A. Berman and R. Plemmons, *Non-negative Matrices in the Mathematical Sciences*, Volume 9 of Classics in Applied Mathematics, Society of Industrial and Applied Mathematics (SIAM), Philadelphia, 1994.

[6] B. Boisvert, R. Pozzo, K. Remington, B. Miller and R. Lipman, *Matrix Market repository. http://math.nist.gov/MatrixMarket/.*

[7] F. F. Bonsall and J. Duncan, *Numerical Ranges of Operators on Normed Spaces and of Elements of Normed Algebras*, Cambridge University Press, ISBN $978 - 0 - 521 - 07988 - 4,$ 1971.

[8] F. F. Bonsall and J. Duncan, *Numerical Ranges II*, Cambridge University Press, ISBN $978 - 0 - 521 - 20227 - 5,$ 1971.

[9] R. Brualdi, *Matrices, eigenvalues and directed graphs*, Linear and Multilinear Algebra, $11 : 143 - 165,$ 1982.

[10] X. S. Chen, S. W. Vong, W. Li and H. Xu, *Noda iterations for generalized eigenproblems following Perron-Frobenius theory*, Numerical Algorithms, 2018.

[11] Lj. Cvetković, *H-matrix theory vs. eigenvalue localization*, Numerical Algorithms, $42(3-4): 229-245,$ 2006.

[12] Lj. Cvetković, *NUMELA I*, Department of mathematics and Informatics, Faculty of science at the University of Novi Sad, 2014.

[13] Lj. Cvetković and V. Kostić, *Between Geršgorin and the minimal Geršgorin set*, Journal of Computational and Applied Mathematics, $196(2): 452-458,$ 2006.

[14] Lj. Cvetković and V. Kostić, *New criteria for identifying H-matrices*, Journal of Computational and Applied Mathematics, $465-478,$ 2005.

[15] Lj. Cvetković and V. Kostić, *New subclasses of block H-matrices with applications to parallel decomposition-type relaxation methods*, Numerical Algorithms, $42(3-4): 325-334,$ 2006.

[16] Lj. Cvetković, V. Kostić, R. Bru and F. Pedorche, *A simple generalization of Geršgorin's theorem*, Advances in Computational Mathematics, 2011.

[17] Lj. Cvetković, V. Kostić and D. Cvetković, *Pseudospectra localizations and their applications*, Numerical Linear Algebra with Applications, Vol. 23, No $2: 356-372,$ 2016.

[18] Lj. Cvetković, V. Kostić and R. S. Varga, *A new Geršgorin type eigenvalue inclusion set*, ETNA (Electronic Transactions on Numerical Analysis), $18: 73-80,$ 2004.

[19] Lj. Cvetković and J. Pena, *Minimal sets alternative to minimal Geršgorin set*, Applied Numerical Mathematics, $60: 442-451,$ 2010.

[20] B. N. Datta, *Numerical Linear Algebra and Applications,* Brooks/Cole, Pacific Grove, CA, 1995.

[21] J. W. Demmel, *Applied linear algebra*, SIAM, 1997.

[22] C. Desoer, *Feedback System: Input-output Properties*, Elsevier Science, 2012.

[23] L. Elsner, *Inverse iteration for calculating the spectral radius of a non-negative irreducible matrix*, Linear Algebra and Its Applications, 15 : 235 – 242, 1976.

[24] R. A. Frazer, W. J. Duncan and A. R. Collar, *Elementary Matrices and some applications to dynamics and differential equations*, Cambridge University Press, 1938.

[25] M. A. Freitag and A. Spence, *A Newton-based method for the calculation of the distance to instability*, Linear Algebra and Its Applications, 435(12) : 3189 – 3205, 2011.

[26] S. Geršgorin, *Über die Abgrenzung der Eigenwerte einer Matrix*, Bulletin de l'Académie des Sciences de l'URSS, Classe des sciences mathématiques et na, No. 6, 749 – 754, 1931.

[27] G. H. Golub and C. F. Loan, *Matrix Computations*, 2nd edition, The John Hopkins University Press, 1989.

[28] J. Hefferon, *Linear Algebra*, Online book, St Michael's College, Colchester, Vermont, USA, 2001.

[29] D. Hershkowitz, *Recent directions in matrix stability*, Linear Algebra and Its Applications, 171, 161 – 186, 1992.

[30] H. M. Higham and H. M. Kim, *Numerical solution of quadratic matrix equation*, IMA Journal of Numerical Analysis, 499 – 519, 2000.

[31] D. Hinrichsen and B. Kelb, *Spectral value sets: a graphical tool for robustness analysis*, Systems & Control Letters, 127 – 136, 1993.

[32] M. E. Hochstenbach, *Subspace methods for eigenvalue problems*, Phd thesis at Faculty of Mathematics and Computer Science, Urecht University, Holland, 2003.

[33] A. S. Householder, *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York, 1964.

[34] J. Huitfieldt and A. Ruhe, *A new algorithm for numerical path following applied to an example from hydrodynamical flow*, SIAM, 11, 1181 – 1192, 1990.

[35] C. R. Johnson, *A Geršgorin inclusion set for the field of values of a finite matrix*, JSTOR, $41 : 57 - 60,$ 1973.

[36] C. R. Johnson, *Functional Characterizations of the Field of Values and the Convex Hull of the Spectrum*, Proceedings of the American Mathematical Society, $61(2) : 201 - 204,$ 1976.

[37] H. B. Keller, *Practical procedures in path following near limit points*, Computing Methods in Applied Sciences and Engineering, 1982.

[38] V. Kostić, *Benefits from the generalized diagonal dominance*, Phd dissertation at Department of mathematics and Informatics, Faculty of science at the University of Novi Sad, 2010.

[39] V. Kostić, *On general principles of eigenvalue localizations via diagonal dominance*, Advances in Computational Mathematics, $41 : 55 - 75,$ 2015.

[40] V. Kostić, A. Miedlar and Lj. Cvetković, *An algorithm for computing minimal Geršgorin sets*, Numerical linear algebra with applications, $00 : 1 - 19,$ 2015.

[41] V. Kostić, Lj. Cvetković and L. Krukier, *Matrix nonsingularity and diagonal dominance property* (in Russian), Izvestiya Vuzov. Severo Kavkazskiy region. Estestvennie nauki 3, pp. $16 - 18,$ 2013.

[42] P. Lancaster, *Matrix theory*, Moscow: Science Publishers, 1973.

[43] P. Lancaster, *On eigenvalues of matrices dependent on a parameter*, Numerische Mathematik, $6 : 377 - 387,$ 1964.

[44] B. W. Levinger and R. S. Varga, *Minimal Gershgorin sets II*, Pacific Journal of Mathematics, $17 : 199 - 210,$ 1966.

[45] J. R. Magnus, *On differentiating eigenvalues and eigenvectors*, Econometric Theory, $1 : 179 - 191,$ 1985.

[46] C. Mayer, *Matrix analysis and applied linear algebra*, SIAM, $978-0-89871-454 - 8,$ 2000.

[47] D. Mehzer and B. Philippe, *PAT-a Reliable Path Following Algorithm*, Numerical Algorithms, 2002.

[48] S. Milićević and V. Kostić, *Detecting stability via the abscissa of the minimal Gershgorin set*, The third Conference on Mathematics in Engineering: Theory and Applications, $40-45$, Faculty of Technical Sciences, Novi Sad, 2018.

[49] S. Milićević, V. Kostić, Lj. Cvetković and A. Miedlar, *An implicit algorithm for computing the minimal Geršgorin set*, FILOMAT, $13: 4229-4238$, $0354-5180$, 2019.

[50] T. Noda, *Note on the computation of the maximal eigenvalue of a non-negative irreducible matrix*, Numerische Mathematik, $17: 382-386$, 1971.

[51] B. Philippe and M. Sadkane, *Computation of the Fundamental Singular Subspace of a Large Matrix*, Linear Algebra and Its Application 257, $77-104$, 1997.

[52] M. Reed and B. Simon, *Methods of modern mathematical physics*, Academic Press Inc., San Diego, 1978.

[53] W. Rudin, *Real and Complex analysis*, 3rd Edition, McGraw-Hill International Editions, 1986.

[54] Y. Sead, *Numerical methods for large eigenvalue problems*, Manchester University Press Series in Algorithms and Architectures for advanced scientific computing, 1992.

[55] L. A. Steen, *Highlights in the history of spectral theory*, Saint Olaf College, 1973.

[56] G. W. Stewart, *Matrix algorithms*, SIAM, Philadelphia, 2001.

[57] Z. Stojaković and D. Herceg, *Linearna algebra i analitička geometrija*, Symbol, Novi Sad, 2005.

[58] A. Takayama, *Mathematical Economics* 2^{nd} *edition*, Cambridge University Press, 1985.

[59] O. Taussky, *Bounds for the characteristic roots of matrices*, Duke Math. J, $15: 1043-1044$, 1948.

[60] L. N. Trefethen and D. Bau, *Spectra and pseudospectra*, SIAM, 1997.

[61] L. N. Trefethen and M. Embree, *Numerical linear algebra*, Princeton University Press, 2005.

[62] C. W. Uberhuber, *Numerical Computation: Methods, Software and Analysis*, Volumes 1 and 2, Springer-Verlag, New York, 1997.

[63] R. S. Varga, *Geršgorin and His Circles*, Springer-Verlag, New York, 2004.

[64] R. S. Varga, *Geršgorin disks, Brauer ovals of Cassini (a vindication) and Brualdi sets,* Information, $4(2) : 171 - 178,$ 2001.

[65] R. S. Varga, *Geršgorin-type eigenvalue inclusion theorems and their sharpness*, ETNA (Electronic Transactions on Numerical Analysis), $12 : 113 - 133,$ 2001.

[66] R. S. Varga, *Matrix iterative analysis*, Springer, Berlin, 1962.

[67] R. S. Varga, *Minimal Gershgorin sets,* Pacific J. Math., $15 : 719-729,$ 1965.

[68] R. S. Varga, *On recurring theorems on diagonal dominance*, Linear Algebra and Its Application, $13 : 1 - 9,$ 1976.

[69] R. S. Varga, Lj. Cvetković and V. Kostić, *Approximation of the minimal Geršgorin set of a square complex matrix*, ETNA (Electronic Transactions on Numerical Analysis), $40 : 308 - 405,$ 2008.

[70] R. S. Varga and A. Krautstengl, *On Gersgorin-type problems and ovals of Cassini*, ETNA (Electronic Transactions on Numerical Analysis), $8 : 15-20,$ 1999.

[71] P. F. Zachlin and M. E. Hochstenbach, *On the numerical range of a matrix*, Mathematische Nachrichten $6 : 193 - 228,$ 1951.

[72] X. Z. Zhan, *Matrix theory*, American Mathematical Society Providence, Rhode Island, 2013.