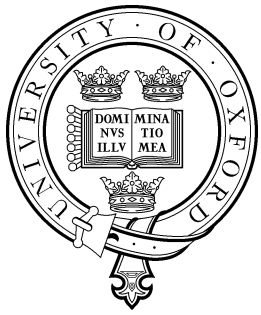
Numerical Computation of Eigenfunctions of Planar Regions



Timo Betcke

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2005

Numerical Computation of Eigenfunctions of Planar Regions

Timo Betcke

Keble College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas 2005

In 1967 Fox, Henrici and Moler published a beautiful article describing the Method of Particular Solutions (MPS) for the Laplace eigenvalue problem with zero Dirichlet boundary conditions on planar regions. The idea is to use particular solutions that satisfy the eigenvalue equation but not necessarily the zero boundary conditions to approximate the eigenfunctions. Unfortunately, their method becomes unstable for more complicated regions including regions with several corner singularities, which led to a decline of interest in such methods in the numerical analysis community.

In this thesis we return to the original idea of Fox, Henrici and Moler and devise a modification based on angles between subspaces that avoids the problems of their method. Our new "subspace angle method" has close links to the generalized singular value decomposition (GSVD). We use this to show the stability of our method and explain why the GSVD is a natural framework for methods based on particular solutions.

Classical error bounds for the MPS were derived by Moler and Payne. We extend these bounds to our method and verify the first eigenvalue on the L-shaped region to 13 rounded digits of accuracy.

The approximation theory of the MPS goes back to results by Vekua. We use his theory and analytic continuation of eigenfunctions to prove exponential convergence of our method on regions with zero or one corner singularity. Using conformal mapping techniques we compute the exact asymptotic exponential rate on several regions. For regions with multiple corner singularities we propose a choice of basis functions that seems to lead to better than algebraic convergence rates.

We then show how to extend the GSVD approach to a domain decomposition method by Descloux and Tolley and improve their original convergence estimates using Vekua's theory.

Finally, we present eigenvalue and eigenfunction computations on many planar regions including the L-shaped region, isospectral drums and some multiply connected regions.

to Marta
kocham Cię

Acknowledgements

This thesis would not have been possible without a number of people. My foremost thanks goes to my supervisor Professor Nick Trefethen who has been a great inspiration throughout the last three years. His constant encouragement, enthusiasm for interesting problems and also criticism have deeply influenced my way of thinking.

This was supported by a unique atmosphere in the Numerical Analysis Group. I will always remember it with great joy to have studied with Zachary Battles, Sue Dollar, Andris Lasis and all the others in the group.

The thesis also profited very much from many interesting discussions and suggestions from Alex Barnett, Toby Driscoll, Cleve Moler, Stan Eisenstat and many others who contributed ideas. Especially the discussions with Alex Barnett led to many new ideas which have shaped this thesis. I also thank him for the great time I had visiting him at the Courant Institute New York.

In my second year I spent six months as a visitor at the University of Queensland at the Advanced Computational Modelling Centre under Professor Kevin Burrage. I want to thank him and the members of his group for this unforgettable time in Australia.

All this would not have been possible without one very special person in my life, my fiancée Marta Markiewicz. Her understanding, love and compassion throughout these sometimes not easy years gave me the strength to pursue this research. My third year when we both were in Oxford was one of the happiest in my life so far and I am looking forward for many more happy years to come with her.

I will also always remember the Institute of Mathematics at the Hamburg University of Technology under Professor Voss. The support of this group encouraged me to do numerical analysis and to apply for the D.Phil program in Oxford.

During my time in Oxford I was always in contact with the Rotary Foundation who gave me the opportunity to take part at many interesting events. In Germany I want thank Dr. Herbert Schwiegk from the Rotary Club of Winsen/Luhe for his support.

Work is only one part of the life in Oxford. Equally important are the friends from my college. I am deeply grateful for my very good friends Justin Walker, Olympia Bobou, Christopher Guyver, our two chaplains Mark Butchers and Allen Shin and all the other members of the MCR and the chapel community.

But all this would not have been such a great period of my life without my family and especially my mother who were always there with their help and understanding.

Certainly, studying in Oxford always has a financial side. I am grateful for the support of the Scatcherd European Scholarship throughout this time.

Thank you all !

Contents

List of Figures

Chapter 1

Introduction

1.1 The Dirichlet eigenvalue problem

This thesis is about the accurate numerical solution of the Laplace eigenvalue problem with Dirichlet boundary conditions, defined by

$$-\Delta u = \lambda u \quad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1.1b}$$

where Ω is a bounded planar region. One of the early roots of the great mathematical interest in this problem is the work of Chladni at the end of the 18th and the beginning of the 19th century. He used sand to make the nodal lines in vibrating plates visible. Napoleon was so excited by these experiments that he set out a price of 3000 francs for anyone who could explain the mathematical theory behind these figures. This price was awarded in 1816 to Sophie Germain, who managed to partially explain them by finding the fourth order PDE describing vibrations of a plate but did not state the boundary conditions correctly. Although the mathematical theory behind Chladni's figures differs from the membrane eigenvalue problem (1.1), his work can be seen as the key starting point in the investigation of both phenomena.

According to Lord Rayleigh [70], the mathematical analysis of the membrane eigenvalue problem was first considered by Poisson, who investigated vibrations on a rectangle. Important 19th century contributions were also made by Lamé, Clebsch,

Weber, Rayleigh, Schwarz and Pockels[1].

In the 20th century the membrane eigenvalue problem gained large interest in the context of Schrödinger's equation. It was shown that (1.1) governs quantum states of a particle trapped in a two-dimensional well. Nowadays this equation plays an important role in the field of quantum chaos, which has emerged in the last twenty to thirty years. Physicists in this field are interested in the behavior of eigenfunctions for very high energies, i.e. large values of λ [33].

Among mathematicians the membrane eigenvalue problem gained a lot of attention in the second half of the $20th$ century with Kac's famous article from 1966 "Can one hear the shape of a drum?" [41]. The question asks whether there are two distinct planar regions which have the same spectrum. This was first answered in 1992 by Gordon, Webb and Wolpert [32] who were able to construct such isospectral regions.

1.2 Drum computations and a famous logo

The computation of eigenvalues and eigenfunctions of (1.1) is a nontrivial problem. General purpose methods are for example finite differences, boundary element, or finite element methods. Another more specialized approach is the Method of Particular Solutions (MPS), which was introduced by Fox, Henrici and Moler in 1967. It uses particular solutions that satisfy the eigenvalue equation (1.1a) but not necessarily the zero boundary conditions. The idea is to find values of λ for which there exist linear combinations of the basis functions which are small on a given set of boundary collocation points. This method was successfully applied by Fox, Henrici and Moler to compute the first eigenvalues on the L-shaped region to up to eight digits of accuracy. Most numerical analysts will have seen an example of this method without actually knowing it. The famous Matlab logo is derived from applying this method to the L-shaped region and is an approximation of the first eigenfunction on this region. Apparently, it does not satisfy the zero boundary conditions. For aesthetic reasons Moler chose the image in Figure 1.1 instead of the correct eigenfunction. The correct eigenfunction can be obtained with the Matlab command `membrane(1,15,9,4)`. This

[1] An extensive bibliography for the membrane eigenvalue problem can be found in the beautiful review article by Kuttler and Sigillito [44].

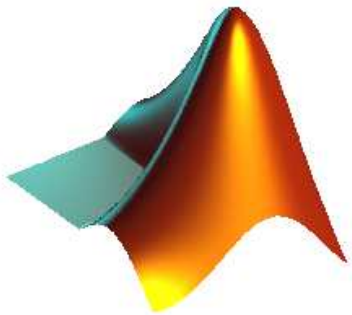


Figure 1.1: The famous Matlab logo.

does not compute the first eigenvalue but uses a stored constant and only computes the eigenfunction. The function used in the Matlab logo is obtained with the command `membrane`, which is equivalent to calling `membrane(1,15,9,2)`. The command `logo` internally calls the `membrane` function but formats the results such that it is the Matlab logo in the familiar form shown in Figure 1.1.

Unfortunately, the original MPS by Fox, Henrici and Moler fails for more complicated regions. This led to a decline of interest in this idea in the Numerical Analysis community. Until recently, the most successful method for (1.1) has been a domain decomposition approach by Descloux and Tolley [18], which was later improved by Driscoll [21].

While there was a decline of interest in the MPS among numerical analysists there has been a growing interest in such methods among physicists in the last twenty years under the name of "point matching methods". One of the original works is due to Heller [34, 35] who developed a method very similar to the MPS without knowing the work of Fox, Henrici and Moler. His method and its generalizations are nowadays frequently used by physicists working in quantum chaos and related fields.

1.3 The structure of this thesis

The starting point of this thesis was the original paper by Fox, Henrici and Moler from 1967. It bothered us that such a beautiful idea should fail for more complicated regions such as polygons with several corner singularities. Since there are only a finite

number of known singularities on such regions there should be a beautiful and robust method which computes the eigenvalues and eigenfunctions to high accuracy.

This thesis can be roughly divided into four different areas:

1. Efficient tools from linear algebra
2. Accuracy bounds
3. Approximation theory
4. Eigenfunction computations

Chapters 2 to 4 and the first half of Chapter 7 belong to the first area. In Chapter 2 and 3 we investigate the failure of the Method of Particular Solutions of Fox, Henrici and Moler and introduce two tools from linear algebra, subspace angles and the generalized singular value decomposition (GSVD). With these tools we devise a new method in Chapter 3 which we call the "subspace angle method" and show a first example of it on the L-shaped region. Parts of Chapters 2 and 4 are also published in [15]. The robustness of our method is investigated in Chapter 4, where we take a close look at the condition numbers of certain generalized singular values, which are just the tangents of subspace angles that we compute in our method. It turns out that our new approach even admits highly accurate computations of eigenvalues and eigenfunctions of (1.1) if the basis of particular solutions is highly ill-conditioned. In the first half of Chapter 7 we extend the idea of using generalized singular values to a certain class of domain decomposition methods for (1.1).

Classical a posteriori accuracy bounds for the Method of Particular Solutions are discussed in Chapter 5 and extended to the subspace angle method. In the second half of the chapter we use these bounds to verify thirteen rounded digits of the first eigenvalue on the L-shaped region. This is the most accurate computation of the first eigenvalue on the L-shaped region that we are aware of.

The approximation theory of the Method of Particular Solutions is investigated in Chapter 6. We show how to use results from Vekua and Garabedian to establish exponential convergence estimates of the MPS for regions with zero or one singular

corner and how one can compute the exact asymptotic convergence rates on these regions using conformal mapping techniques and analytic continuation of eigenfunctions. These ideas are extended to domain decomposition methods in the second half of Chapter 7. For regions with multiple singular corners we devise in the last part of Chapter 6 an approach that seems to deliver faster than algebraic convergence rates.

Computations of eigenvalues and eigenfunctions of several regions are presented in Chapter 8. We also take a closer look at the concept of higher subspace angles and avoidance phenomena between them. Parts of Chapter 8 are published in [14, 75].

1.4 Notation

Most of the notation used in this thesis is standard. Everything else will be defined when appropriate. Here we summarize some of the notation used throughout the thesis.

By a region Ω we understand an open connected set in $\mathbb{R}^2$. In some sections (especially in Chapter 7) we also use the term domain for a region. Often we will identify the complex plane $\mathbb{C}$ with the set $\mathbb{R}^2$ by the identity $z = x + iy$. We also frequently use polar coordinates (r, θ) to denote a point $z = re^{i\theta}$. The closure of a region Ω is denoted by $\overline{\Omega}$. The complex conjugate of a complex number z is denoted by $\overline{z}$ and Ω^* is the set of all complex numbers whose complex conjugate is in Ω, i.e. $\Omega^* := \{\overline{z} : z \in \Omega\}$. The area $|\Omega|$ of a region Ω is defined as

$$|\Omega| := \int_\Omega 1 dxdy.$$

For a scalar real or complex variable x we denote by $|x|$ its absolute value. If $x \in \mathbb{R}^n$ then $|x| := (\sum_{k=1}^n |x_k|^2)^{1/2}$, where x_k is the kth component of the vector x. For $x, y \in \mathbb{R}^n$ we denote by $\langle x, y\rangle = x^T y$ the standard Euclidian inner product and by $\|x\|_2 := |x|$ its Euclidian norm. The maximum norm of a vector $x \in \mathbb{R}^n$ is defined by $\|x\|_\infty := \max_n |x_n|$.

Sometimes we use Matlab notation to denote parts of a matrix. Hence, if A is a matrix the first column of A is $A(:, 1)$. For real matrices we will use two norms. The

spectral norm of $A \in \mathbb{R}^{m\times n}$ is defined as $\|A\|_2 := \sqrt{\lambda_{max}(A^TA)}$, where $\lambda_{max}(A^TA)$ is the largest eigenvalue of A^TA and the Frobenius norm $\|A\|_F$ is defined as $\|A\|_F := \sqrt{tr(A^TA)}$, where tr is the trace operator. It follows directly that $\|A\|_2 \leq \|A\|_F \leq \sqrt{\text{rank}(A)}\|A\|_2$.

The L^2-inner product $\langle u, v\rangle$ in Ω is defined by

$$\langle u, v\rangle_\Omega := \int_\Omega u(x,y)v(x,y)dxdy.$$

The associated norm is defined by $\|u\|_\Omega := \langle u, u\rangle_\Omega^{1/2}$. We will also need the inner product of two functions on the boundary $\partial\Omega$. This is defined as the path integral

$$\langle u, v\rangle_{\partial\Omega} := \int_{\partial\Omega} u(s)v(s)ds.$$

Furthermore, we let $\|u\|_{\partial\Omega} := \langle u, u\rangle_{\partial\Omega}^{1/2}$. The sup-norm $\|u\|_{\infty,S}$ of a function u in a set S is defined as

$$\|u\|_{\infty,S} := \sup_{x\in S} |u(x)|.$$

Sometimes we need the relative machine accuracy ϵ_{mach}, which is defined as the distance from 1 to the next larger floating point number. In IEEE double precision the value of this number is 2^{-52}. We will also encounter the unit roundoff u which is 2^{-53} in IEEE double precision arithmetic[2].

In most chapters we use the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ which are defined as

$$\mathcal{A}(\lambda) := \{u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega) : -\Delta u = \lambda u \text{ in } \Omega\} \tag{1.2}$$

and

$$\mathcal{D}_0 := \{u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega) : u|_{\partial\Omega} = 0\}. \tag{1.3}$$

Hence, $\mathcal{A}(\lambda)$ is the space of all particular solutions which are continuous in $\overline{\Omega}$ and $\mathcal{D}_0$ is the space of functions which are twice continuously differentiable in Ω and zero on $\partial\Omega$. Depending on the section the symbol $\mathcal{A}(\lambda)$ can also denote a subspace of the space of particular solutions or the space spanned by a basis of particular solutions evaluated on a set of discretization points. Similarly, $\mathcal{D}_0$ can also mean the space of functions which are zero on a given set of boundary collocation points. This will be clear from the context and also stated again in the corresponding sections.

[2]See [38] for a detailed description of these quantities.

1.5 Basic properties of eigenfunctions on planar regions

We now state without proof some basic properties properties of the solutions of the eigenvalue problem (1.1) which are useful for the understanding of the following chapters. References to further results and proofs are given in [44].

All eigenvalues λ_k of (1.1) are positive. The first eigenvalue is always simple. We can order the eigenvalues with multiplicity according to

$$0 < \lambda_1 < \lambda_2 \leq \cdots$$

with a limit point at infinity, and the corresponding eigenfunctions can be chosen to form an orthonormal complete set in $L^2(\Omega)$. That is,

$$< u_i, u_j >_\Omega = \delta_{ij},$$

where u_i is the eigenfunction associated with λ_i and δ_{ij} is the Kronecker delta. On some elementary regions the eigenvalues and eigenfunctions are explicitly known. For a rectangle with $0 \leq x \leq a,\ 0 \leq y \leq b$ the eigenfunctions are

$$u_{m,n}(x,y) = \sin\left(\frac{m\pi x}{a}\right)\sin\left(\frac{n\pi y}{b}\right), \qquad m,n = 1,2,\ldots$$

with corresponding eigenvalues

$$\lambda_{m,n} = \pi^2\left[\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2\right].$$

In the case of a disk of radius a the eigenfunctions are given by

$$u_{m,n}(r,\theta) = J_m(\frac{j_{mn}r}{a})[A\cos m\theta + B\sin m\theta], \quad m = 0,1,\ldots,\ n = 1,2,\ldots$$

where j_{mn} is the nth zero of the mth order Bessel function J_m. The eigenvalues are

$$\lambda_{m,n} = \left(\frac{j_{mn}}{a}\right)^2.$$

If for two regions $\Omega_1 \subset \Omega_2$ then for the eigenvalues $\lambda_k^{(1)}$ of Ω_1 and $\lambda_k^{(2)}$ of Ω_2 it follows that

$$\lambda_k^{(1)} \geq \lambda_k^{(2)}.$$

Among all regions with the same area the disk has the smallest eigenvalue λ_1. This is the result of the famous *Faber-Krahn inequality* which states that

$$\lambda_1 \geq \frac{\pi}{|\Omega|} j_{01}^2.$$

But a large region does not necessarily have a small first eigenvalue λ_1. Let ρ be the radius of the largest inscribed disk in a simply connected region Ω. Osserman[3] [56] showed that

$$\lambda_1 \geq \frac{1}{4\rho^2}.$$

In 1994 the value $\frac{1}{4}$ was improved by Bañuelos and Carroll to 0.619 [2]. It still remains an open question what is the largest constant e such that $\lambda_1 \geq \frac{e}{\rho^2}$ for general simply connected regions.

The eigenvalues of (1.1) cannot be arbitrarily distributed. An important result to this effect is *Weyl's law*,

$$\lambda_k \sim \frac{4\pi k}{|\Omega|} \text{ as } n \to \infty.$$

A proof can for example be found in [61].

The nodal lines of u_k are the set of points in Ω where $u_k = 0$. *Courant's nodal line theorem* states that the nodal lines of the kth eigenfunction u_k divide Ω into not more than k subregions [61]. The eigenfunction of the first eigenvalue λ_1 has no nodal lines, and by orthogonality it follows that λ_1 is always simple.

In symmetric regions eigenfunctions can be chosen to have either odd or even symmetry. An odd eigenfunction has a nodal line along the symmetry axis and an even eigenfunction has zero normal derivative along this axis. Further symmetry classes were investigated by Hersch [37]. This can sometimes be used to reduce the eigenvalue problem (1.1) to a problem on a simpler region and was applied by Fox, Henrici and Moler to the eigenvalue problem on the L-shaped region.

Eigenfunctions are real analytic inside Ω. The smoothness on $\partial\Omega$ depends on the region. If a corner of $\partial\Omega$ consists of two straight arcs meeting at an angle π/k, where k is an integer, then any eigenfunction can be continued to an analytic function in the neighborhood of the corner. Otherwise, eigenfunctions can have singularities at

[3] In [3] Bañuelos and Carroll point out that this result even goes back to Makai in 1965.

the corner, which have to be dealt with by the numerical method in order to achieve fast convergence to the eigenfunction. We will say much more about these matters in Chapter 6.

For the eigenvalue problem (1.1) there are different sets of particular solutions. Using separation of variables in polar coordinates for the equation $-\Delta u = \lambda u$ one can derive the solutions

$$J_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta, \quad J_{\alpha k}(\sqrt{\lambda}r)\cos\alpha k\theta \tag{1.4}$$

for $\alpha, \lambda > 0$ and $k \in \mathbb{N}$, where $J_{\alpha k}$ is the Bessel function of the first kind of order αk. We will call the functions in (1.4) Fourier-Bessel sine and Fourier-Bessel cosine functions. If $\alpha k \notin \mathbb{N}$ these functions are not $\mathcal{C}^\infty$ at 0. A similar set of particular solutions is obtained by using Bessel functions of the second kind instead of the first kind in (1.4). We obtain

$$Y_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta, \quad Y_{\alpha k}(\sqrt{\lambda}r)\cos\alpha k\theta.$$

We will only need these functions for the case $\alpha \in \mathbb{N}$. It is important to note that $Y_{\alpha k}(x) \to -\infty$ for $x \to 0$. Therefore, the origin of the polar coordinates has to lie outside the region if we want to use Fourier-Bessel functions of the second kind as particular solutions.

Another class of particular solutions are real plane waves. In cartesian coordinates these are given as

$$\mathrm{Re}\{e^{i\sqrt{\lambda}(x\cos\alpha + y\sin\alpha)}\}, \quad \mathrm{Im}\{e^{i\sqrt{\lambda}(x\cos\alpha + y\sin\alpha)}\},$$

or equivalently in polar coordinates as

$$\mathrm{Re}\{e^{i\sqrt{\lambda}r\cos(\theta-\alpha)}\}, \quad \mathrm{Im}\{e^{i\sqrt{\lambda}r\cos(\theta-\alpha)}\} \tag{1.5}$$

for $-\pi \le \alpha \le \pi$. These are waves oscillating with wavelength $2\pi/\sqrt{\lambda}$ in the direction given by α and constant perpendicular to α. To obtain a set of $2N$ basis functions one usually takes $\alpha = \frac{k\pi}{N}$ for $k = 0, \ldots, N-1$. The following argument shows that this is a sensible choice. It holds that

$$J_n(\sqrt{\lambda}r)e^{in\theta} = \frac{-i^n}{2\pi}\int_0^{2\pi} e^{i\sqrt{\lambda}r\cos(\theta-\tau)}e^{in\tau}d\tau$$

for $n \in \mathbb{N}$. Using the trapezoidal rule we obtain

$$J_n(\sqrt{\lambda}r)e^{in\theta} \approx -\frac{-i^n}{2N}\sum_{k=0}^{2N-1} e^{i\sqrt{\lambda}r\cos(\theta-\frac{\pi k}{N})}e^{in\frac{\pi k}{N}}. \tag{1.6}$$

By combining terms belonging to k and $N+k$ it follows for n even that

$$\begin{aligned} \mathrm{Re}\{J_n(\sqrt{\lambda}r)e^{in\theta}\} &\approx \sum_{k=0}^{N-1} \alpha_k^{(N)} \cos(\sqrt{\lambda}r\cos(\theta - \frac{\pi k}{N})), \\ \mathrm{Im}\{J_n(\sqrt{\lambda}r)e^{in\theta}\} &\approx \sum_{k=0}^{N-1} \beta_k^{(N)} \cos(\sqrt{\lambda}r\cos(\theta - \frac{\pi k}{N})), \end{aligned}$$

for certain real coefficients $\alpha_k^{(N)}$ and $\beta_k^{(N)}$. If n is odd the same formulas are valid with $\sin(\sqrt{\lambda}r\cos(\theta - \frac{\pi k}{N}))$ instead of $\cos(\sqrt{\lambda}r\cos(\theta - \frac{\pi k}{N}))$. Density results and approximation properties of Fourier-Bessel functions and real plane waves are investigated in [68].

A very interesting set of basis functions are evanescent plane waves. These are obtained by choosing a complex shift α in (1.5). Then (1.5) is a wave oscillating with wavelength $2\pi/(\sqrt{\lambda}\cosh \mathrm{Im}\ \alpha)$ along the direction $\mathrm{Re}\ \alpha$ and decaying exponentially in the direction $\mathrm{Re}\ \alpha + \pi/2\ \mathrm{Sign}(\mathrm{Im}\ \alpha)$ [12]. Evanescent plane waves have been applied with great success to obtain accurate eigenvalue approximations on the Bunimovich stadium billiard [6, 81].

Chapter 2

The Method of Particular Solutions (MPS)

In 1967 Fox, Henrici and Moler published a beautiful article “Approximations and bounds for eigenvalues of elliptic operators" [25] describing the Method of Particular Solutions for eigenvalue problems on planar regions. Based on theoretical work of Bergman and Vekua ([10, 80], see also Chapter 6) they approximated solutions of (1.1) by linear combinations of particular solutions that satisfy (1.1a) but not necessarily (1.1b). The boundary conditions were approximated using a collocation method. With this approach they computed the first 10 eigenvalues of the L-shaped region to an accuracy of up to 8 digits. By deriving error estimates they were able to give lower and upper bounds for each eigenvalue.

This simple and elegant method and its application to the L-shaped region led to many related dissertations and articles by Fox's and Mayers' students Donnelly, Mason, Reid and Walsh at Oxford [19, 50, 63] and Moler's students Schryer and Eisenstat at Michigan and Stanford [23, 64].

Unfortunately, the MPS in the form proposed by Fox, Henrici and Moler suffers from problems for more complicated regions, especially regions with several corner singularities. This led to a decline of research in the MPS in the 1970's. Indeed, the best method known as of a year or two ago, developed by Descloux and Tolley in 1983 [18] and improved by Driscoll in 1997 [21], is based on domain decomposition rather than global approximations.

While the MPS got less attention in the numerical analysis community, it was independently rediscovered by physicists working in semiclassical mechanics, quantum chaos and related fields; this literature often speaks of methods of "point matching". One of the originators of this work is Heller, who in the 1980s used a method very similar to the MPS to investigate "scars" in high energy eigenstates of the Bunimovich stadium billiard [34]. It is interesting to note that although Heller's method is now a standard tool in physics, the only indication he gave of it in [34] was the following sentence:

> These are just a few of nearly a dozen types of scars found so far, using a simple algorithm written by the author.

He gave a thorough explanation of his method a few years later in [35]. Heller's approach was generalized and improved by his student Barnett [6].

Another method based on particular solutions is the scaling method of Vergini and Saraceno [82]. The advantage of their method is that it computes good approximations to many high energy eigenstates with just one matrix decomposition, as opposed to the traditional MPS, where several decompositions are needed to get one eigenstate accurately. Investigating this method has led to some interesting theoretical results [7, 8]. Unfortunately, the formulation of Vergini and Saraceno only works for star-shaped regions. But still it is a remarkable method that deserves further investigation.

In this chapter we will first analyze the original MPS of Fox, Henrici and Moler. Then we will discuss in detail the failure of this method for more complicated regions. This failure and understanding it points the way to the more robust methods developed in the later chapters.

2.1 The MPS of Fox, Henrici and Moler

The idea of the MPS as proposed by Fox, Henrici and Moler is to take a set of functions that satisfy (1.1a) and to find a parameter λ for which there exists a linear combination of these functions that is small on the boundary $\partial\Omega$.

Let us consider an infinite wedge with interior angle π/α. The eigenfunctions of this wedge are the functions

$$u(r,\theta) = J_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta \tag{2.1}$$

for arbitrary $\lambda > 0$ and $k \in \mathbb{N}$. The idea of Fox, Henrici and Moler was to approximate

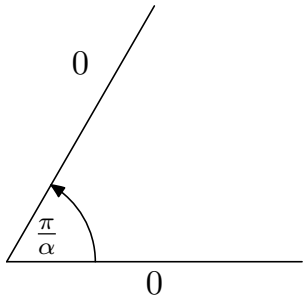


Figure 2.1: An infinite wedge with interior angle π/α. The eigenfunctions (2.1) of this region are known as Fourier-Bessel functions.

eigenfunctions of a polygon containing a corner with interior angle π/α by linear combinations of Fourier-Bessel functions of the form (2.1). Hence, we want to find coefficients $c_k^{(N)}$ and a value for λ such that

$$u(r,\theta) = \sum_{k=1}^{N} c_k^{(N)} J_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta$$

is a good approximation to an eigenfunction of (1.1), i.e. $u(r,\theta)|_{\partial\Omega} \approx 0$. On the arcs adjacent to the corner with interior angle π/α, we automatically have $u(r,\theta) = 0$. The rest of the boundary is discretized with collocation points $z_j = r_j e^{i\theta_j} \in \partial\Omega, j = 1,\ldots,N$. Condition (1.1b) now becomes

$$u(r_j,\theta_j) = \sum_{k=1}^{N} c_k^{(N)} J_{\alpha k}(\sqrt{\lambda}r_j)\sin\alpha k\theta_j = 0, \quad j = 1,\ldots,N.$$

This is equivalent to the system of equations

$$A_B(\lambda)c = 0, \tag{2.2}$$

where $(A_B)_{jk} = J_{\alpha k}(\sqrt{\lambda}r_j)\sin\alpha k\theta_j$. In Chapter 3 we will also introduce a matrix A_I consisting of Fourier-Bessel functions evaluated at interior points of Ω. One can solve (2.2) by looking for the zeros of $\det(A_B(\lambda))$, which was the original approach of Fox, Henrici and Moler.

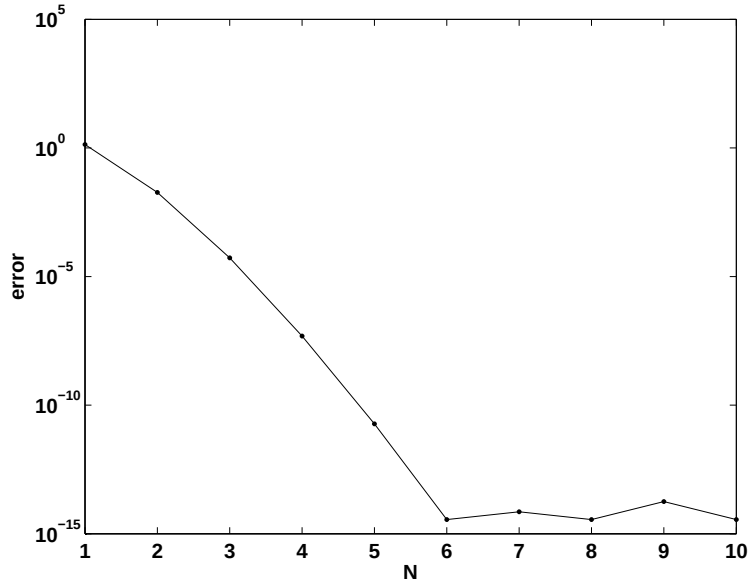


Figure 2.2: The convergence for the first eigenvalue of the unit square. In each step $2N$ basis functions and collocation points are used.

Let us try this method on a simple region. On the unit square $[0,1]^2$ the eigenfunctions are explicitly known as

$$u_{m,n}(x,y) = \sin(m\pi x)\sin(n\pi y), \qquad m,n = 1,2,\ldots$$

with corresponding eigenvalues

$$\lambda_{m,n} = \pi^2\left(m^2+n^2\right).$$

We expand around the corner at $z = 0$. Then the Fourier-Bessel basis functions are automatically zero on the two sides adjacent to $z = 0$. Each of the other two boundary sides is discretized with N collocation points. Therefore, $2N$ basis functions are chosen to obtain a square matrix $A_B(\lambda) \in \mathbb{R}^{2N,2N}$. The convergence behavior of the MPS for the first eigenvalue $2\pi^2$ is shown in Figure 2.2. The figure seems to show at least "spectral" convergence, i.e. convergence at the rate $O(N^{-s})$ for every $s > 0$. Indeed, from the convergence theory developed in Chapter 6 it follows that the rate of convergence is $O(R^{-N})$ for each $R > 1$. This example seems to hint that the MPS might be a powerful method. But the example is still too simple to reveal much. Therefore, let us try a more complicated region.

Figure 2.3 shows the famous L-shaped region. We approximate around the reentrant corner with linear combinations of Fourier-Bessel functions of the form $J_{\frac{2}{3}k}(\sqrt{\lambda}r)\sin\frac{2}{3}k$.

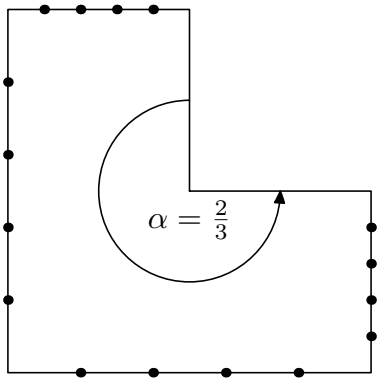


Figure 2.3: Discretization of the L-shaped region for the MPS.

This cancels out the singularity of the eigenfunction at the reentrant corner and theoretically leads again to spectral convergence, as we will show in Chapter 6. The zero boundary conditions are automatically satisfied on the arcs adjacent to the reentrant corner. Each of the other sides is discretized using N collocation points.

To avoid numerical underflow in calculating $\det(A_B(\lambda))$ due to bad scaling of the Fourier-Bessel basis each column of $A_B(\lambda)$ is now scaled to have unit norm. The first eigenvalue of the L-shaped region is $\lambda_1 \approx 9.6397238440219$. The convergence behavior of the MPS to this eigenvalue is shown in Figure 2.4. The MPS does not get more than four digits and breaks down after $N = 15$. This shows that there is a problem with the original MPS as formulated by Fox, Henrici and Moler. They were able to get around this problem and calculate 8 digits by using symmetry properties of the eigenfunctions to reduce the problem size. But as we will show now, such techniques are only able to improve the accuracy of the MPS in a few special cases. On more complex regions the method almost always fails.

2.2 The failure of the original MPS

The MPS tries to find a value $\lambda > 0$ such that there exists a linear combination of Fourier-Bessel basis functions which is small at the boundary collocation points. What happens now if $A_B(\lambda)$ is ill-conditioned for all $\lambda > 0$?

Let $A_B(\lambda) \in \mathbb{R}^{n \times p}$ with $n \geq p$ (we now include the case where there may be more collocation points than basis functions) and assume that the smallest singular value

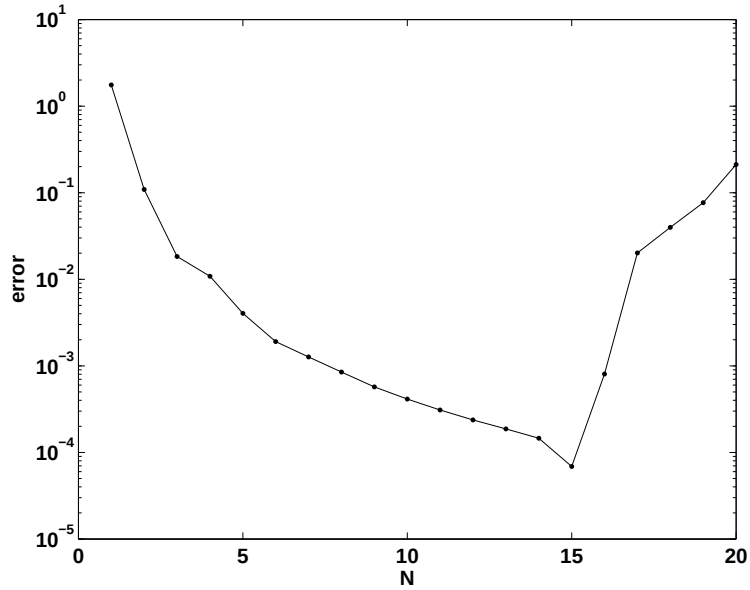


Figure 2.4: Failure of convergence of the original MPS for the first eigenvalue of the L-shaped region.

$\sigma_p(\lambda)$ of $A_B(\lambda)$ satisfies $\sigma_p(\lambda) = O(\epsilon_{mach})$. Then there exists a vector $c \in \mathbb{R}^p$, $\|c\|_2 = 1$ such that $\|A_B(\lambda)c\|_2 = O(\epsilon_{mach})$. If λ is close to an eigenvalue λ_k of (1.1), then

$$\tilde{u}(r,\theta) = \sum_{k=1}^{N} c_k J_{\alpha k}(\sqrt{\lambda} r) \sin \alpha k \theta$$

may be a good approximation of an eigenfunction (in Chapter 5 we will discuss error bounds for the MPS). However, if λ is not close to an eigenvalue of (1.1), then $\tilde{u}(r,\theta)$ satisfies the eigenvalue equation $-\Delta u = \lambda u$ and is numerically zero on the boundary collocation points. The only solution of (1.1), if λ is not an eigenvalue, is $u(r,\theta) = 0$. Therefore, we can expect $\tilde{u}(r,\theta) \approx 0$ in Ω.

The MPS cannot distinguish between functions that are numerically zero in Ω and true eigenfunctions, since it only considers boundary collocation points. But if $A_B(\lambda)$ is ill-conditioned for every $\lambda > 0$, we can always find a linear combination of basis functions that is close to zero at the boundary collocation points, leading to spurious solutions that are close to zero on the whole of Ω if λ is not close to an eigenvalue. In this section we present numerical experiments that demonstrate this behaviour and discuss the matter of when we can expect $A_B(\lambda)$ to be ill-conditioned for all $\lambda > 0$. Let us return to the example of the L-shaped region from Section 2.1. The convergence in Figure 2.4 breaks down after $N = 15$. Figure 2.5 shows the condition number

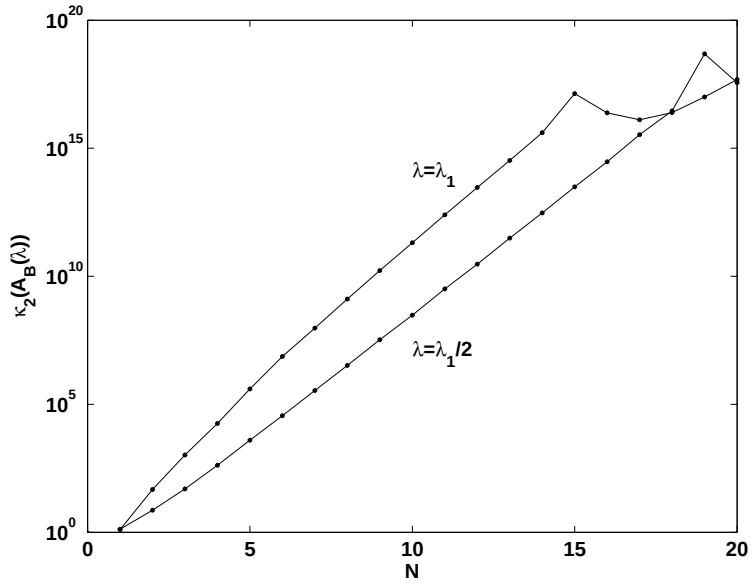


Figure 2.5: The condition number of $A_B(\lambda)$ for $\lambda = \lambda_1$ and the arbitrary value $\lambda = \lambda_1/2$. After $N = 15$ both matrices become numerically singular, making it impossible for the MPS to detect the eigenvalue λ_1.

$\kappa_2(A_B(\lambda))$ measured in the 2−norm for a growing number N of basis functions. For λ we chose two different parameters. The first is $\lambda = \lambda_1$, where $\lambda_1 \approx 9.6397238440219$ is the first eigenvalue on the L-shaped region. The second is the arbitrary choice $\lambda = \lambda_1/2$. The columns of $A_B(\lambda) \in \mathbb{R}^{4N\times 4N}$ (we have N collocation points on each of the 4 sides not adjacent to the reentrant corner) are again scaled to unit norm. Both curves grow exponenentially. After $N = 14$ the results become erroneous due to rounding errors. To detect an eigenvalue of (1.1) the MPS depends on the gap between those two curves, which does not widen much as N increases and is in any case computed incorrectly after $N = 14$.

How can we improve the condition of $A_B(\lambda)$? In Figure 2.5 we already used diagonal scaling of the columns of $A_B(\lambda)$ to improve its condition number. This is crucial here since the scaling of Fourier-Bessel functions becomes exponentially smaller with growing order k, which introduces severe ill-conditioning in $A_B(\lambda)$. Over all possible choices of columnwise scaling a nearly optimal strategy is to scale all columns to unit norm, since for $A \in \mathbb{R}^{m\times n}$ and $\text{rank}(A) = n$,

$$\kappa_2(AD_C) \leq \sqrt{n} \min_{D\in\mathcal{D}_n} \kappa_2(AD).$$

Here, $\mathcal{D}_n \subset \mathbb{R}^{n\times n}$ denotes the set of nonsingular diagonal matrices and $D_C :=$

$\mathrm{diag}(\|A(:,k)\|_2)^{-1}$ is the diagonal matrix that scales all columns of A to unit norm (see [38], p. 125 for a proof).

Scaling alone, although necessary, does not deliver satisfactory results, as Figure 2.5 shows. We could try using different distributions of points on the boundary. Indeed, using points in a Chebyshev distribution on each arc allows us to obtain 8 digits of accuracy before the method breaks down. To make the MPS less dependent on the choice of points it is advisable to use many more points on the boundary than there are expansion terms, as proposed in [54]. But this does not solve the fundamental problem of the MPS that it fails to exclude spurious solutions which are numerically zero everywhere in the region. The following example demonstrates a situation where the MPS fails even to get a few digits of the first eigenvalue. Consider a quadrilateral with four corner singularities defined by the points $0, 1, 1.5 + 1.5i, 1 + 1.5i$. The eigenfunctions have singularities at all four corners (singularities of eigenfunctions are discussed in Chapter 6). Therefore, in order to get fast convergence to the solution, Fourier-Bessel expansions at all corners are needed. The first eigenvalue of (1.1) on this region is $\lambda_1 \approx 24.73768313904717$. Figure 2.6 shows the convergence of the solution for a growing number N of basis terms at each corners. On each side of the boundary 100 points were used. For $N = 3$ the method obtains the first three digits 24.7 correctly, but for larger N it fails completely. The reason is that the four Fourier-Bessel expansions only behave differently very close to the singularities. Otherwise they approximately span the same space of functions on $\overline{\Omega}$. This leads to the matrix $A_B(\lambda)$ being heavily ill-conditioned independently of λ.

Fox, Henrici and Moler were aware of the fact that their method might run into problems for more complicated regions. In [25] they noted:

> In all fairness, it should be reported that results are not always as satisfactory as these examples indicate.... Other methods... are currently being investigated.

In [21] Driscoll wrote about the problems in applying the MPS to a challenging region with several corner singularities:

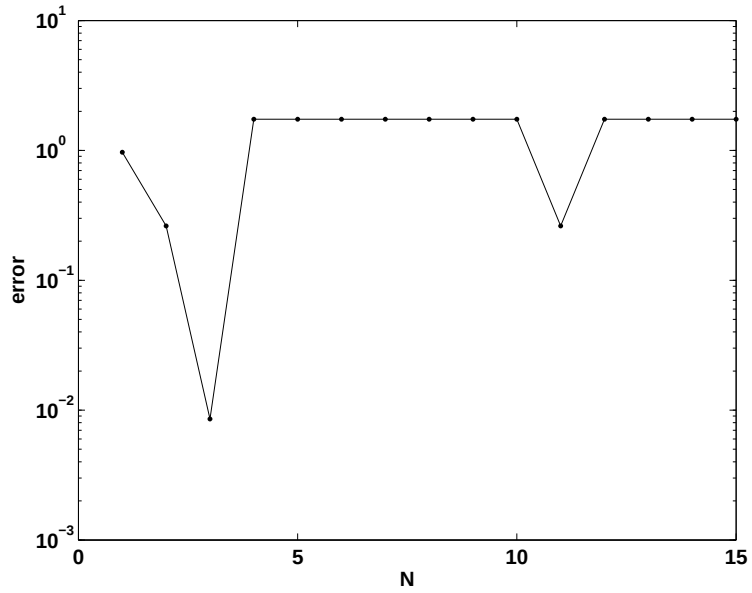


Figure 2.6: The MPS fails completely in the case of a quadrilateral with Fourier-Bessel expansions at all corners.

> As the number of terms in the truncated expansion is increased, the matrix becomes very nearly singular for all values of λ, and detecting the true singularity numerically becomes impossible. In fact, we have been unable to produce more than two or three accurate digits for a few of the smallest eigenvalues with this method.

In Chapter 3 we develop an approach to the MPS that solves these problems and allows highly accurate approximations to eigenvalues and eigenfunctions on planar regions. But before we want to discuss two methods developed by physicists, the PWDM of Heller and its generalization by Barnett. Both methods partially solve the problems of the MPS by introducing a normalization of the trial functions.

2.3 The PWDM of Heller

The idea of Heller's PWDM (Plain Wave Decomposition Method) is very similar to the original MPS of Fox, Henrici and Moler. Two facts about this method are remarkable. The first is that it was developed completely independently of the literature in the Numerical Analysis community about methods based on particular solutions. The

second is that it utilized a simple trick to partially solve the problem of spurious solutions. Heller originally used this method to compute scars in chaotic billiards [34], where he used real plane waves as basis functions. But the method can equally well be applied to other sets of particular solutions like Fourier-Bessel functions or evanescent plane waves.

The idea of the method is the following. Let us pick a basis of N particular solutions. As in the MPS, we could choose N boundary points and obtain the system of equations

$$A(\lambda)c = 0.$$

But as discussed in the last section, this introduces spurious solutions in the search space which destroy the convergence. To avoid such solutions that are numerically zero everywhere in the region, Heller picked one point in the interior of the region and imposed the condition that the trial functions are 0 on $N-1$ boundary points and equal to 1 at the interior point. This leads to the system of equations

$$A(\lambda)c = e_n,$$

where e_N is the Nth unit vector $[0,\dots,0,1]^T \in \mathbb{R}^N$ and the last row of $A(\lambda)$ now consists of the particular solutions evaluated at the interior point. To check the quality of an approximate eigenfunction, it is first normalized in the interior of the region and then evaluated at many boundary points. Heller calls this boundary norm the "tension" of the trial function. If the tension is small, then hopefully the trial function is a good approximation to an eigenfunction of (1.1).

Like the MPS, the method of Heller can also be formulated using a least-squares approach. Let p be the number of basis functions and n the number of boundary points, with $n \geq p$. Furthermore, let $l(\lambda)^T \in \mathbb{R}^{1,p}$ be the row vector of basis functions evaluated at the interior point. Then for a fixed eigenvalue estimate λ, Heller's method can be formulated as

$$\min_{\substack{x\in\mathbb{R}^p \\ l(\lambda)^T x=1}} \|A(\lambda)x\|_2. \tag{2.3}$$

Let $l(\lambda) = QR$ be the full QR decomposition of $l(\lambda)$ and define $\tilde{a} := A(\lambda)Q(:,1)$ and $\tilde{A} := A(\lambda)Q(:,2{:}p)$. Since $l(\lambda)$ is a vector we have $R = [\xi,0,\dots,0]^T \in \mathbb{R}^p$ for one $\xi \in \mathbb{R}$. Equation (2.3) can now be transformed into the standard least-squares problem

$$\min_{z\in\mathbb{R}^{p-1}} \left\| \tilde{A}z + \frac{\tilde{a}}{\xi} \right\|_2.$$

The solution x of (2.3) is obtained as $x = \frac{1}{\xi}Q(:,1) + Q(:,2{:}p)z$ (see [30], Section 12.1.4).

Alternatively, we could attempt to solve the least-squares problem

$$\min_{x\in\mathbb{R}^p} \left\| \begin{bmatrix} A(\lambda) \\ l^T(\lambda) \end{bmatrix} x - e_n \right\|_2 . \tag{2.4}$$

This delivers a trial function that is small at the boundary points but close to one at the interior point, which is often good enough to avoid spurious solutions. Approximations of constrained least squares problems by standard least squares problems are discussed in [30, 78]. In [78] several error bounds are also given.

Heller's method is widely used in the quantum chaos community and related fields. It is easily applicable and often delivers good approximations to eigenmodes. The drawback of the method is the choice of the interior point. If it is close to a nodal line of the exact eigenfunction then even good approximations to the eigenfunction are scaled up by the normalization at the interior point and are discarded as spurious solutions. Hence, this method is only a partial solution to the stability problems of the MPS.

2.4 Barnett's generalization of the PWDM

The PWDM of Heller can have problems if a nodal line is close to the interior point. Barnett's generalization of the PWDM solves this problem [6]. Let

$$\mathcal{A}(\lambda) = \mathrm{span}\{\mathrm{u}^{(1)}, \ldots, \mathrm{u}^{(\mathrm{N})}\}$$

be the space spanned by N particular solutions $u^{(1)}, \ldots, u^{(N)}$ satisfying $-\Delta u^{(k)} = \lambda u^{(k)}$, $k = 1, \ldots, N$ which are twice differentiable in Ω and continuous on $\overline{\Omega}$. For $u, v \in \mathcal{A}(\lambda)$ define the boundary inner product[1]

$$\langle u, v\rangle_{\partial\Omega} := \int_{\partial\Omega} u(s)v(s)\, ds.$$

[1] If λ is an eigenvalue of (1.1), $\langle\cdot,\cdot\rangle_{\partial\Omega}$ is not positive definite and therefore in the strict sense not an inner product.

Furthermore, we need the standard L^2-inner product

$$\langle u, v\rangle_{\partial\Omega} := \int_\Omega u(x,y)v(x,y)\, dxdy.$$

The corresponding norms are defined as $\|u\|_{\partial\Omega} = \langle u, u\rangle_{\partial\Omega}^{1/2}$ and $\|u\|_\Omega = \langle u, u\rangle_\Omega^{1/2}$.

For a function $u \in \mathcal{A}(\lambda)$ we can define the tension

$$t(u) := \frac{\|u\|_{\partial\Omega}}{\|u\|_\Omega}.$$

Furthermore, let us define the minimal tension as

$$t_m(\lambda) := \min_{u\in\mathcal{A}(\lambda)} t(u). \tag{2.5}$$

If $t_m(\lambda) = 0$, then λ is an eigenvalue of (1.1), since then there exists a nonzero function $u \in \mathcal{A}(\lambda)$ satisfying $-\Delta u = \lambda u$ and $\|u\|_{\partial\Omega} = 0$. To compute $t_m(\lambda)$, Barnett proposed the following method. Let $u = \sum_{k=1}^N x_k u^{(k)}$. Then

$$t_m^2(\lambda) = \min_{u\in\mathcal{A}(\lambda)} \frac{\langle u, u\rangle_{\partial\Omega}}{\langle u, u\rangle_\Omega} = \min_{x\in\mathbb{R}^N} \frac{x^T F(\lambda) x}{x^T G(\lambda) x},$$

where F and G are defined by

$$F_{ij}(\lambda) = \langle u_i, u_j\rangle_{\partial\Omega}, \quad G_{ij}(\lambda) = \langle u_i, u_j\rangle_\Omega.$$

Hence, we can represent $t_m^2(\lambda)$ as the minimum of a Rayleigh quotient. The solution is given as the smallest eigenvalue $\mu_1(\lambda)$ of the eigenvalue problem

$$F(\lambda)x(\lambda) = \mu(\lambda)G(\lambda)x(\lambda) \tag{2.6}$$

and we obtain $t_m(\lambda) = \mu_1(\lambda)^{1/2}$.

This method is a true generalization of Heller's method since it guarantees that approximate eigenfunctions are normalized over the whole region Ω instead of being normalized at only one point. But the numerical implementation of Barnett's method has two drawbacks. The first is that almost linearly dependent basis sets lead to a common numerical null space of the matrices $F(\lambda)$ and $G(\lambda)$. One strategy to prevent this is to project the common null space out of the eigenvalue problem, as described in [6]. This issue is also further discussed in Section 4.6.

The other problem is the following. Suppose for the computed smallest eigenvalue $\tilde{\mu}_1(\lambda)$ that

$$\tilde{\mu}_1(\lambda) = \mu_1(\lambda) + f,$$

where $f = O(\epsilon_{mach})$ is a small perturbation in the order of machine accuracy. We then obtain

$$\tilde{t}_m(\lambda) = \sqrt{\mu_1(\lambda) + f}.$$

Therefore, no matter how small $\mu_1(\lambda)$ is, the minimum of the computed value $\tilde{t}_m(\lambda)$ for the tension cannot become lower than $\sqrt{f} = O(\sqrt{\epsilon_{mach}})$, meaning that Barnett's method is limited to an accuracy of $O(\sqrt{\epsilon})$. Since asymptotically the function $t_m(\lambda)$ behaves like $K|\lambda - \lambda_k|$ close to an eigenvalue λ_k for a constant $K > 0$ [7], we can generally not expect to detect eigenvalues to more than 8 digits of accuracy if we work in IEEE double precision. For most applications in physics this restriction to 8 digits of accuracy is usually not harmful. But in this thesis we want to develop a method that is able to detect eigenvalues to an accuracy close to machine precision if the basis of particular solutions admits such accurate approximations. In the next chapter we will develop such a method based on angles between subspaces and in Section 3.5 we show that Barnett's method can be interpreted as a squared version of our new approach.

Chapter 3

Subspace angles and the generalized SVD (GSVD)

In the last chapter we discussed the failure of the MPS in the form put forward by Fox, Henrici and Moler. The reason for the failure is that we do not have a well-conditioned problem since the method does not contain information about the interior of the region. This problem was partially solved by Heller's PWDM and Barnett's generalization of it. But Heller's method only partially solves the problem since it heavily depends on the choice of the interior point, and Barnett's approach cannot find eigenvalues to a higher accuracy than the square root of machine precision. Furthermore, the ill-conditioning of the basis poses stability problems in the formulation as a generalized eigenvalue problem.

In order to reliably find eigenvalues and eigenfunctions of (1.1) we need

1. A well-conditioned problem,
2. A stable algorithm for solving it.

The first of these goals can be achieved by introducing several interior points. To extract approximate eigenfunctions using the information from boundary and interior points we introduce an algorithm that can be formulated either as a problem of finding the angle between certain subspaces or as a generalized singular value problem. The stability of this algorithm will be discussed in detail in Chapter 4.

3.1 The Dirichlet eigenvalue problem and angles between subspaces

Consider the space $\mathcal{A}(\lambda)$ of all solutions of $-\Delta u = \lambda u$ in Ω which are continuous on $\partial\Omega$, i.e.

$$\mathcal{A}(\lambda) := \{u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega) : -\Delta u = \lambda u \text{ in } \Omega\}.$$

Let $\mathcal{D}_0 \subset \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ be the space of functions in this continuity class which are zero on $\partial\Omega$. If for a given $\lambda > 0$ the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ have a nontrivial intersection there exist nonzero functions in $\mathcal{A}(\lambda)$ satisfying the eigenvalue equation and the zero boundary conditions, which are therefore eigenfunctions belonging to the eigenvalue λ. The following lemma is immediately obtained.

Lemma 3.1.1 *The spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ have a nontrivial intersection if and only if $\lambda > 0$ is an eigenvalue of* (1.1).

The principal angle between two subspaces is a useful tool to measure whether they have a nontrivial intersection. Suppose that $\langle \cdot, \cdot \rangle$ is a suitable inner product with induced norm $\| \cdot \|$. Then the principal angle $\theta(\lambda)$ between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ can be defined as

$$\cos\theta(\lambda) := \sup_{\substack{u \in \mathcal{A}(\lambda),\ \|u\|=1 \\ v \in \mathcal{D}_0,\ \|v\|=1}} \langle u, v \rangle. \tag{3.1}$$

What is the right inner product to measure the principal angle between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$? If the standard L^2-inner product $\langle u, v \rangle = \int_\Omega uv dx$ is chosen, then $\theta(\lambda) = 0$ for all $\lambda > 0$ since the eigenfunctions of (1.1) in Ω are in $\mathcal{D}_0$ and form a complete orthonormal set of $L^2(\Omega)$. We need to incorporate the information on the boundary of the region. One way to do this is by introducing a mixed inner product of the form

$$\langle u, v \rangle := \int_\Omega uv dx + \int_{\partial\Omega} uv dx = \langle u, v \rangle_\Omega + \langle u, v \rangle_{\partial\Omega} \tag{3.2}$$

The following theorem shows that this inner product leads to a useful meaning of the angle between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$.

Theorem 3.1.2 *If $\theta(\lambda)$ is defined by the inner product* (3.2)*, then the value $\lambda > 0$ is an eigenvalue of* (1.1) *if and only if $\theta(\lambda) = 0$.*

Proof If λ is an eigenvalue of (1.1), any eigenfunction u associated with λ is an element of $\mathcal{D}_0$ and of $\mathcal{A}(\lambda)$. It follows that $\theta(\lambda) = 0$. Conversely, in Chapter 5 we show

$$\frac{|\lambda - \lambda_k|}{\lambda_k} \leq c \tan\theta(\lambda)$$

for a constant c that only depends on the region Ω, where

$$\frac{|\lambda - \lambda_k|}{\lambda_k} = \min_n \frac{|\lambda_n - \lambda|}{\lambda_n}.$$

The minimum is taken over all eigenvalues λ_n of (1.1). If $\theta(\lambda) = 0$ it follows that $\lambda = \lambda_k$. ∎

3.2 Principal angles in finite dimensional spaces

Before we turn the idea of using principal angles for the MPS into an algorithm we give an introduction to principal angles in finite dimensional spaces and their calculation. The following definition is due to Björck and Golub [16].

Principal angles between subspaces Let $\mathcal{A}$ and $\mathcal{B}$ be subspaces of $\mathbb{R}^m$ with $q = \dim(\mathcal{A}) \geq \dim(\mathcal{B}) = p$. The principal angles $\theta_1 \leq \cdots \leq \theta_p$ are recursively defined as

$$\cos\theta_k := \langle u_k, v_k \rangle = \max_{\substack{u \in \mathcal{A},\ \|u\|_2 = 1 \\ v \in \mathcal{B},\ \|v\|_2 = 1}} \langle u, v \rangle, \quad u \perp u_1, \ldots, u_{k-1}, \quad v \perp v_1, \ldots, v_{k-1}. \tag{3.3}$$

The vectors u_k and v_k are the principal vectors associated with the principal angles θ_k.

Let $Q_A \in \mathbb{R}^{m \times q}$ and $Q_B \in \mathbb{R}^{m \times p}$ be orthogonal bases of $\mathcal{A}$ and $\mathcal{B}$. Björck and Golub showed that the principal angles θ_k and the associated pairs of principal vectors u_k and v_k can be obtained from the singular value decomposition

$$Q_A^T Q_B = \tilde{U} \Sigma \tilde{V}^T, \tag{3.4}$$

where $\tilde{U} \in \mathbb{R}^{q \times q}$ and $\tilde{V} \in \mathbb{R}^{p \times p}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{q \times p}$ is diagonal with $\Sigma = \mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_{\mathrm{p}})$. The v_k are the columns of the matrix $Q_B \tilde{V}$ and the u_k are the first p columns of $Q_A \tilde{U}$. The Björck-Golub algorithm for angles between subspaces therefore consists of two steps:

- Compute orthogonal bases Q_A and Q_B of $\mathcal{A}$ and $\mathcal{B}$.
- Compute the singular values of $Q_A^T Q_B$ to obtain the cosines of the principal angles.

If one is interested in very small angles it is necessary to work with the sines of the principal angles. Consider the case in which $\theta_1 = O(\sqrt{\epsilon_{mach}})$. Then $\cos\theta_1 \approx 1 - \frac{\theta_1^2}{2} = 1 - O(\epsilon_{mach})$. Therefore, the cosine of principal angles can only be determined up to the square root of machine precision. Sines of principal angles do not have this restriction.

The sines of the principal angles between the spaces $\mathcal{A}$ and $\mathcal{B}$ can be elegantly introduced using the CS decomposition.

Theorem 3.2.1 (CS Decomposition) *Let* $Q = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$ *be a matrix with orthonormal columns, where* $Q_1 \in \mathbb{R}^{m_1 \times n}, Q_2 \in \mathbb{R}^{m_2 \times n}$ *and* $m_1 \geq n$. *Then there exist orthogonal matrices* U, W *and* V *such that*

$$\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix} V^T.$$ [1]

The matrices S *and* C *are diagonal with entries*

$$0 = s_1 = \cdots = s_r < s_{r+1} \leq \cdots \leq s_{r+j} < s_{r+j+1} = \cdots = s_n = 1$$

and

$$1 = c_1 = \cdots = c_r > c_{r+1} \geq \cdots \geq c_{r+j} > c_{r+j+1} = \cdots = c_n = 0.$$

Depending on Q, *it is possible that* $r = 0$ *or* $r + j = n$. *Furthermore,* $s_k^2 + c_k^2 = 1,\ k = 1, \ldots, n$.

Proof A proof for the general case involving a row and column partitioning of Q can be found in [59], where the history and many applications of the CS decomposition are also reviewed. Here, we only need the simple case of a row partitioning. The proof given here follows the one in [30].

[1] Often the CS decomposition is written down in the form $\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} C \\ S \end{bmatrix} V^T$. But for our purpose it is more suitable to use the notation given here.

Let USV^T be a singular value decomposition of Q_1, where S contains the singular values $s_1 \leq \cdots \leq s_n$ of Q_1 in ascending order. Since Q is orthogonal, $s_k \leq 1$ for $k = 1, \ldots, n$. Define $\begin{bmatrix} K_1 & K_2 \end{bmatrix} = Q_2 V$, where $K_1 \in \mathbb{R}^{m_2 \times r+j}$ and $K_2 \in \mathbb{R}^{m_2 \times n-r-j}$. Then

$$\begin{bmatrix} U & 0 \\ 0 & I_{m_2 \times m_2} \end{bmatrix}^T \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix} V = \begin{bmatrix} \tilde{S} & 0 \\ 0 & I_{m_1-r-j \times n-r-j} \\ K_1 & K_2 \end{bmatrix},$$

where $\tilde{S} \equiv \mathrm{diag}(\mathrm{s}_1, \ldots, \mathrm{s}_{\mathrm{r+j}}) \in \mathbb{R}^{\mathrm{r+j} \times \mathrm{r+j}}$ contains the singular values of Q_1 which are smaller than 1. Since the columns of the right-hand side matrix have unit norm and are mutually orthogonal, $K_2 = 0$, and the matrix

$$\tilde{K}_1 = K_1 \mathrm{diag}(1/\sqrt{1 - s_1^2}, \ldots, 1/\sqrt{1 - s_{r+j}^2})$$

has orthonormal columns. Define $W = \begin{bmatrix} \tilde{K}_1 & \tilde{K}_1^{\perp} \end{bmatrix}$ with $\tilde{K}_1^{\perp}$ chosen such that W is orthogonal. Then $W^T Q_2 V = C$, which finishes the proof. ∎

The decomposition $Q_2 = WCV^T$ is just the singular value decomposition of Q_2. The remarkable property of the CS decomposition is that the singular value decompositions of Q_1 and Q_2 both have the same right singular vectors, which are the columns of the matrix V.

With the help of the CS decomposition it is easy to formulate a notion of sines of angles between two subspaces. The cosines of the angles between the spaces $\mathcal{A}$ and $\mathcal{B}$ are the singular values of $Q_A^T Q_B$. Define the matrix

$$Q = \begin{bmatrix} (I - Q_A Q_A^T) Q_B \\ Q_A Q_A^T Q_B \end{bmatrix}.$$

Since Q has orthonormal columns, the CS decomposition can be applied, leading to

$$\begin{bmatrix} (I - Q_A Q_A^T) Q_B \\ Q_A Q_A^T Q_B \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix} V^T. \tag{3.5}$$

WCV^T is the singular value decomposition of $Q_A Q_A^T Q_B$. Since premultiplying $Q_A^T Q_B$ with Q_A does not change the singular values the diagonal elements of C are the cosines of the principal angles between $\mathcal{A}$ and $\mathcal{B}$. From Theorem 3.2.1 it follows that $s_k^2 + c_k^2 = 1,\ k = 1, \ldots, n$. Hence, the s_k are the sines of the principle angles and we obtain $s_k = \sin\theta_k$. We do not need the full CS decomposition to compute the sines of the principal angles since they are just the singular values of $(I - Q_A Q_A^T) Q_B$, which can be directly computed.

Up to now we have assumed that $q = \dim(\mathcal{A}) \geq \dim(\mathcal{B}) = p$. Consider now the case in which $q < p$. Then there exist only q principal angles, whose cosines are the singular values of $Q_B^T Q_A \in \mathbb{R}^{p \times q}$. But the singular values of $Q_A^T Q_B$ are identical to those of $Q_B^T Q_A$. However, if we form the CS decomposition of $\begin{bmatrix} (I - Q_A Q_A^T) Q_B \\ Q_A Q_A^T Q_B \end{bmatrix}$ then $c_{q+1} = \cdots = c_p = 0$, and these values are by definition not principal angles between $\mathcal{A}$ and $\mathcal{B}$. But for the ease of notation we will drop the condition $q \geq p$ from now on and define $\theta_{q+1}, \ldots, \theta_p = \pi/2$ whenever $q < p$ while keeping in mind that these are not true principal angles according to Definition 3.2.

3.3 A subspace angle algorithm for the MPS

We now return to the question of how to implement a subspace angle algorithm for the Method of Particular Solutions. As with the MPS of Fox, Henrici and Moler, we want to work on a set of discretization points. But instead of working only on boundary points we now add some interior points. Let $z_1, \ldots, z_N \in \partial\Omega$ be the boundary collocation points. In addition we choose a number of interior points $\tilde{z}_1, \ldots, \tilde{z}_M \in \Omega$, which in practice we generally take to be random, though other choices are also possible.

Since it is not possible in a practical algorithm to work with the space of all functions that satisfy the eigenvalue equation $-\Delta u = \lambda u$ in Ω, the space $\mathcal{A}(\lambda)$ now consists only of the span of the basis of particular solutions $u^{(1)}, \ldots, u^{(p)}$ of $-\Delta u = \lambda u$. Also, instead of working with the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ themselves, we work with their representations at the boundary and interior discretization points. Then the bases of these spaces can be written in matrix form. As in the original MPS of Fox, Henrici and Moler, the matrix $A_B(\lambda)$ denotes the basis functions evaluated on the boundary collocation points, while we additionally introduce the matrix $A_I(\lambda)$ of basis functions evaluated on the interior points. Hence, the discretized space $\mathcal{A}(\lambda)$ is the span of the columns of

$$A(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}.$$

Similarly, the columns of the matrix

$$D_0 = \begin{bmatrix} 0 \\ I_{M \times M} \end{bmatrix} \in \mathbb{R}^{(N+M) \times M}$$

provide a basis of the space of functions that are zero at the boundary collocation points.

Theorem 3.3.1 *Let*

$$\begin{bmatrix} Q_B(\lambda) \\ Q_I(\lambda) \end{bmatrix} R(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}$$

be a QR decomposition of $A(\lambda)$ *and let*

$$\begin{bmatrix} Q_B(\lambda) \\ Q_I(\lambda) \end{bmatrix} = \begin{bmatrix} U(\lambda) & 0 \\ 0 & W(\lambda) \end{bmatrix} \begin{bmatrix} S(\lambda) \\ C(\lambda) \end{bmatrix} V(\lambda)^T. \tag{3.6}$$

be the CS decomposition of $Q(\lambda)$. *Then the principal angles* $\phi_k(\lambda)$, $k = 1, \ldots, p$ *between* $\mathcal{A}(\lambda)$ *and* $\mathcal{D}_0$ *are given by*

$$s_k(\lambda) = \sin\theta_k(\lambda), \quad c_k(\lambda) = \cos\theta_k(\lambda).$$

Proof From Section 3.2 it follows that the sines and cosines of the principal angles between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ are obtained from the CS decomposition of

$$\begin{bmatrix} (I - D_0 D_0^T) Q(\lambda) \\ D_0 D_0^T Q(\lambda) \end{bmatrix} = \begin{bmatrix} Q_B(\lambda) \\ 0 \\ 0 \\ Q_I(\lambda) \end{bmatrix}.$$

Using (3.6) we find

$$\begin{bmatrix} (I - D_0 D_0^T) Q(\lambda) \\ D_0 D_0^T Q(\lambda) \end{bmatrix} = \begin{bmatrix} U(\lambda) & 0 & & \\ 0 & I & & \\ & & 0 & I \\ & & W(\lambda) & 0 \end{bmatrix} \begin{bmatrix} S(\lambda) \\ 0 \\ C(\lambda) \\ 0 \end{bmatrix} V(\lambda)^T.$$

Hence, $S(\lambda)$ and $C(\lambda)$ define the sines and cosines of the principal angles between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$. ∎

How does this result relate to the angle $\theta(\lambda)$ between the original non-sampled spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$? In the case of the sampled spaces we have

$$\cos\theta_1(\lambda) = \max_{\substack{x \in \mathbb{R}^p \\ y \in \mathbb{R}^M}} \langle A(\lambda)x, D_0 y \rangle$$

under the condition that $\|A(\lambda)x\|_2 = \|D_0 y\|_2 = 1$ for the angle $\theta_1(\lambda)$ between the sampled spaces. Since

$$\langle A(\lambda)x, D_0 y \rangle = \langle \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix} x, \begin{bmatrix} 0 \\ I \end{bmatrix} y \rangle = \langle A_B(\lambda)x, 0 \rangle + \langle A_I(\lambda)x, y \rangle,$$

this angle is a discrete analogue of the angle $\theta(\lambda)$ for the non sampled spaces in the inner product (3.2). In both cases the boundary part of the inner product is always zero for inner products between elements of $\mathcal{A}(\lambda)$ and of $\mathcal{D}_0$. But nevertheless it is important since the normalization of the elements of $\mathcal{A}(\lambda)$ depends on the boundary and interior part of the inner product, no matter whether we work with the sampled or non-sampled spaces.

From (3.6) it follows that $\|Q_B(\lambda)v_1(\lambda)\|_2 = s_1(\lambda)$ and $\|Q_I(\lambda)v_1(\lambda)\|_2 = c_1(\lambda)$, where $v_1(\lambda)$ is the first column of $V(\lambda)$. Consider the case $s_1(\lambda) \ll 1$. Then also $\|Q_B(\lambda)v_1(\lambda)\|_2 \ll 1$ and

$$\|Q_I(\lambda)v_1(\lambda)\|_2 = c_1(\lambda) = \sqrt{1 - s_1^2(\lambda)} \approx 1.$$

Therefore, if $s_1(\lambda) \ll 1$ there exists a function in $\mathcal{A}(\lambda)$ that is small on the boundary points and bounded away from 0 in the interior of Ω. This function is expected to be a good approximation to an eigenfunction of (1.1). Hence, the subspace angle method automatically excludes the possibility of numerically zero approximate eigenfunctions. The subspace angle method can be written down in four steps.

- Choose N boundary collocation points and M interior discretization points.
- Repeat for every λ
  1. Form the matrices $A_B(\lambda)$ and $A_I(\lambda)$.
  2. Compute the QR factorization $\begin{bmatrix} Q_B(\lambda) \\ Q_I(\lambda) \end{bmatrix} R(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}$.
  3. Compute the smallest singular value $s_1(\lambda)$ of $Q_B(\lambda)$.

The choice of points is done once and for all while the steps 1–3 are repeated for each value of λ. The numerical stability of Step 2 and 3 will be further discussed in Chapter 4. We want to finish this section by applying the subspace angle method to the L-shaped region. In addition to just collocation points on the boundary we now add random interior points as shown in Figure 3.1. Figure 3.2 shows a plot of the sine $s_1(\lambda)$ of the principal angle $\theta_1(\lambda)$, which we call the subspace angle curve. On each boundary side not adjacent to the reentrant corner 100 equally spaced points were chosen. In the interior of the region 50 points were randomly distributed. The approximation

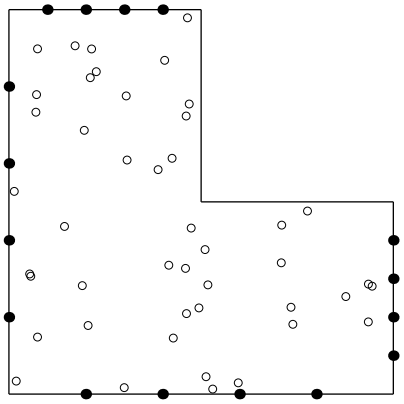


Figure 3.1: In addition to boundary collocation points the subspace angle method utilizes interior points.

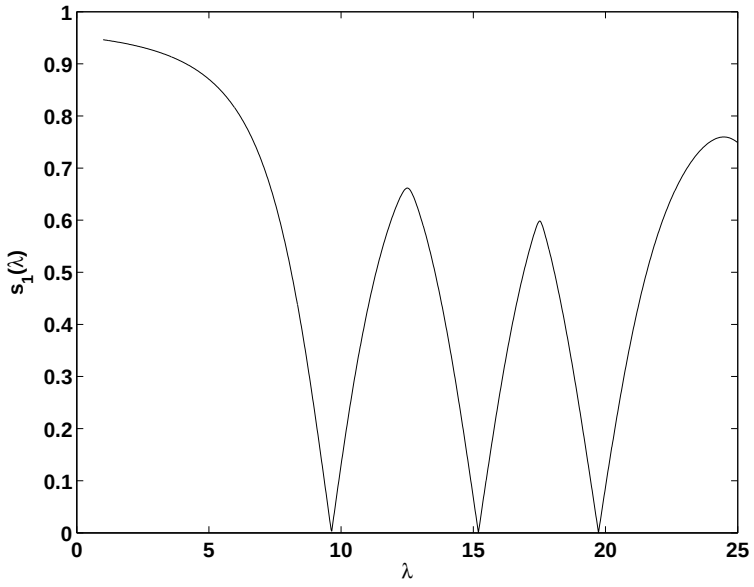


Figure 3.2: The subspace angle curve on the L-shaped region. The first three minima show the positions of the first three eigenvalues on this region.

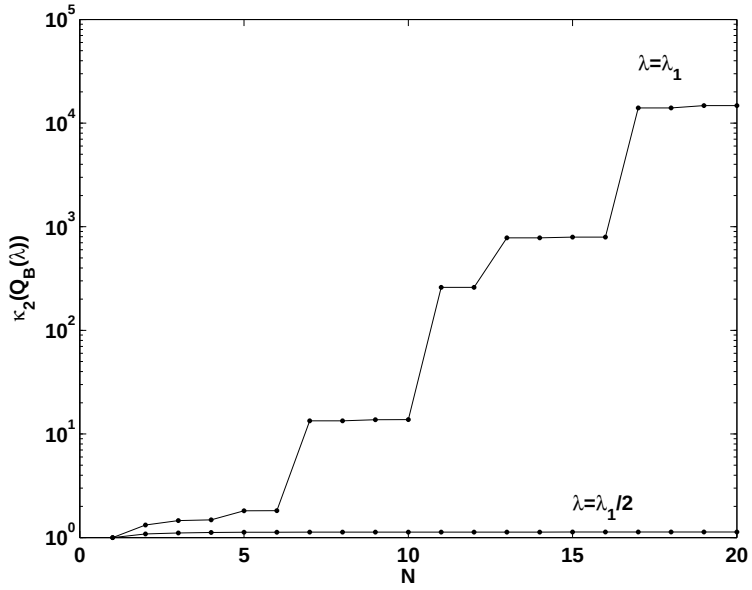


Figure 3.3: The same plot as in Figure 2.5 but now for the matrix $Q_B(\lambda)$. There is a clear gap between the condition numbers for the eigenvalue λ_1 and the arbitrary value $\lambda_1/2$ which widens for a growing number N of basis functions.

basis consists of 20 Fourier-Bessel terms of the form $J_{\frac{2k}{3}}(\sqrt{\lambda}r)\sin\frac{2k\theta}{3}$, $k = 1, \ldots, 20$, with origin at the reentrant corner.

In Figure 2.5 we compared the condition number of $A_B(\lambda)$ in the original MPS for a growing number of basis functions in the two cases $\lambda = \lambda_1$ and $\lambda = \lambda_1/2$. Let us do the same for the matrix $Q_B(\lambda)$. The result is shown in Figure 3.3. In the subspace angle method there is a clear gap in the condition numbers of $Q_B(\lambda_1)$ and $Q_B(\lambda_1/2)$ that widens nicely as the number N of basis terms grows, making it possible to determine the eigenvalue λ_1 to high accuracy.

In Figure 3.4 we show the approximation error $|\lambda - \lambda_1|$ for a growing number N of basis functions. We compared the eigenvalue approximations with the value $\lambda_1 \approx 9.6397238440219$, which we believe to be correct to 14 digits. In Chapter 5 we will show that this value is correct to at leat 13 rounded digits. The minimum of the subspace angle curve was in each step determined with the Matlab function `fminsearch`.

Why is it possible to determine the minimum to such high accuracy using `fminsearch`? Figure 3.5 shows the subspace angle curve close to the value λ_1 for $N = 50$, 60 and 80 basis functions. By increasing the number of basis functions close to λ_1 the subspace

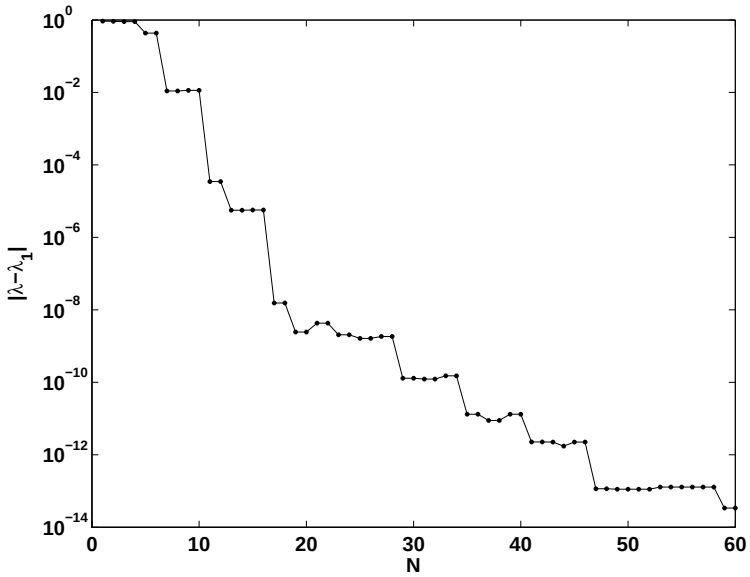


Figure 3.4: The approximation error for the first eigenvalue decreases exponentially on the L-shaped region.

angle curve more and more looks like

$$s_1(\lambda) \approx K|\lambda - \lambda_k|$$

for a value $K > 0$. This asymptotically linear behavior makes it possible to determine the eigenvalue to high accuracy. In Section 3.5 we show that the subspace angle method is closely related to Barnett's method and that for the value $t_m(\lambda)$ defined in (2.5) we have $t_m(\lambda) \approx \tan\theta_1(\lambda)$. Barnett showed [7] that close to an eigenvalue λ_k $t_m^2(\lambda) = C|\lambda - \lambda_k|^2 + O(|\lambda - \lambda_k|^4)$ for a constant $C > 0$ if we approximate from the space of all particular solutions. Since $s_1(\lambda) = \sin\theta_1(\lambda)$ we can expect the subspace angle curve to have a similar asymptotic behavior close to an eigenvalue λ_k if N is high enough.

3.4 The MPS and the generalized singular value decomposition

The original MPS of Fox, Henrici and Moler can be formulated as a singular value decomposition to obtain approximations for eigenvalues and eigenfunctions of (1.1). The approach of Barnett uses generalized eigenvalue problems, and in this chapter we

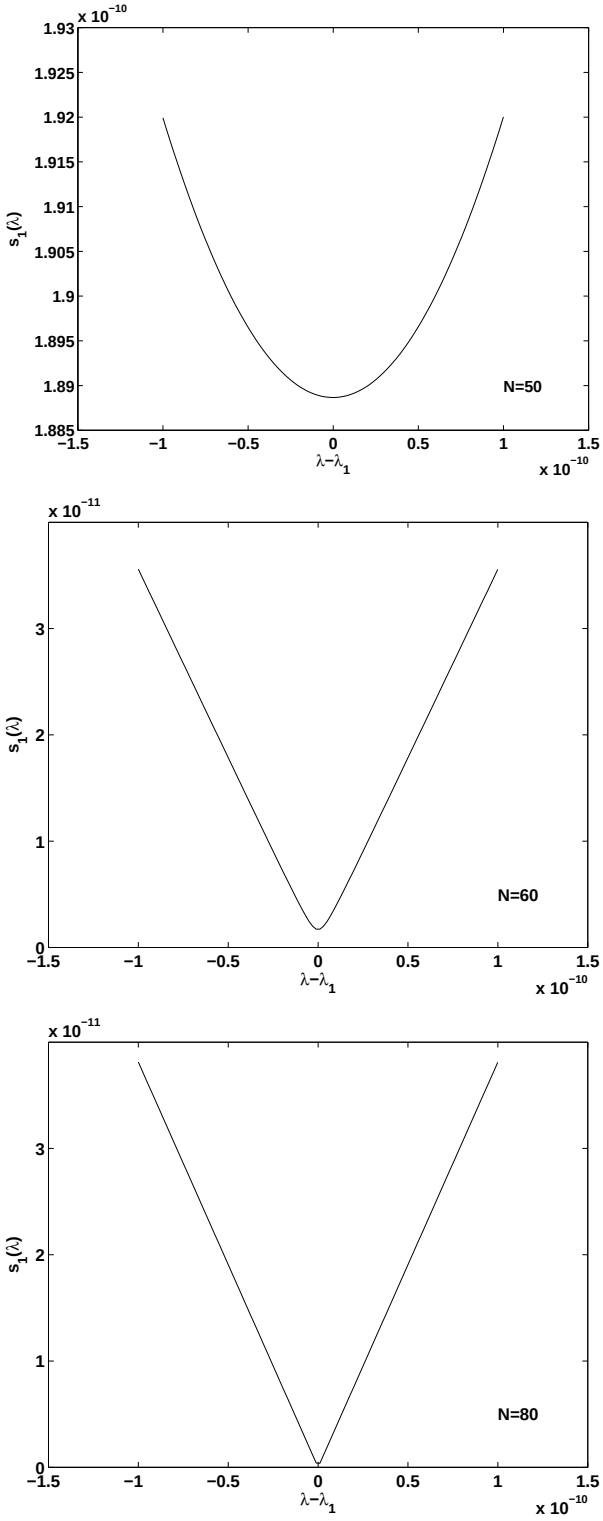


Figure 3.5: The asymptotic behavior of the subspace angle curve close to the eigenvalue λ_1. For a growing number N of basis functions the curve seems to behave linearly close λ_1.

introduced an approach based on principal angles between certain subspaces. We now want to show how this is connected to the Generalized Singular Value Decomposition (GSVD) which will lead us to a natural framework for all methods based on particular solutions discussed so far.

The GSVD was introduced by Van Loan in [77]. He introduced B-singular values and defined them as the elements of the set

$$\mu(A,B) = \{\mu | \mu \geq 0, \det(A^TA - \mu^2 B^TB) = 0\}.$$

This definition also explains why the μ were subsequently called generalized singular values. Ordinary singular values are just the solutions of the equation $\det(A^TA - \mu^2 I) = 0$, while now there is also a matrix B involved. In [58] Paige and Saunders introduced a slightly more general form of the GSVD and also gave a more constructive proof, which will be the basis of the results given here.

Theorem 3.4.1 (Generalized Singular Value Decomposition) *Let $A \in \mathbb{R}^{m_1 \times n}$ with $m_1 \geq n$ and $B \in \mathbb{R}^{m_2 \times n}$. Assume that $Y = \begin{bmatrix} A \\ B \end{bmatrix}$ has linearly independent columns. There exist orthogonal matrices $U \in \mathbb{R}^{m_1 \times m_1}$ and $W \in \mathbb{R}^{m_2 \times m_2}$ and a nonsingular matrix $X \in \mathbb{R}^{n \times n}$ such that*

$$A = USX^{-1}, \quad B = WCX^{-1}, \tag{3.7}$$

where S and C are defined as in Theorem 3.2.1.

Proof Let $QR = Y$ be the QR decomposition of Y and partition Q in the same way as Y is partitioned into A and B, i.e.

$$\begin{bmatrix} Q_A \\ Q_B \end{bmatrix} R = \begin{bmatrix} A \\ B \end{bmatrix}. \tag{3.8}$$

Applying the CS decomposition to Q we obtain

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix} V^T R.$$

Since Y has linearly independent columns, the matrix R^{-1} exists. With $X = R^{-1}V$ the decomposition of A and B in (3.7) follows. ∎

The pairs $(s_k, c_k),\ k = 1, \ldots, n$ are called generalized singular value pairs of the pencil $\{A, B\}$. The generalized singular values are defined as $\sigma_k = s_k/c_k$. If $c_k > 0$ then σ_k is finite. Using the notation from Theorem 3.2.1 there are $r + j$ finite generalized singular values $\sigma_1 \leq \cdots \leq \sigma_{r+j}$ and $n - r - j$ infinite generalized singular values $\sigma_{r+j+1} = \cdots = \sigma_n = \infty$. The kth column x_k of the matrix X is called the right generalized singular vector for the generalized singular value pair (s_k, c_k).

The main part of the proof is the CS decomposition. The GSVD is a simple consequence of this. Paige and Saunders proved the GSVD without the restrictions that $m_1 \geq n$ and $\operatorname{rank}(Y) = n$. But for our purposes this generality is not necessary. If a stable way of computing the CS decomposition is known then this can be directly used to compute the GSVD since the GSVD is just a QR decomposition plus a CS decomposition. This procedure was discussed by Van Loan in [79]. A different approach was taken by Paige in [57], who used an algorithm based on cyclic transformations of A and B. This idea was refined by Bai and Demmel in [4], which forms the basis for the Lapack implementation of the GSVD.

The GSVD has several interesting properties. By combining the equations for A and B in (3.7) we arrive at

$$c_k^2 A^T A x_k = s_k^2 B^T B x_k,\ k = 1, \ldots, n.$$

Therefore, the squares of the finite generalized singular values $\sigma_1, \ldots, \sigma_{r+j}$ are the finite generalized eigenvalues of the generalized eigenvalue problem

$$A^T A x = \sigma^2 B^T B x.$$

The singular values $\sigma_k,\ k = 1, \ldots, \min\{m, n\}$ of a matrix $A \in \mathbb{R}^{m \times n}$ can be characterized as

$$\sigma_k = \max_{\substack{H \subset \mathbb{R}^n \\ \dim(H) = k}} \min_{x \in H \setminus \{0\}} \frac{\|Ax\|_2}{\|x\|_2}. \tag{3.9}$$

Let $m \geq n$. By ordering the singular values in ascending order (i.e. $\sigma_1 \leq \cdots \leq \sigma_n$) an equivalent minimax characterization can be derived:

$$\sigma_k = \min_{\substack{H \subset \mathbb{R}^n \\ \dim(H) = k}} \max_{x \in H \setminus \{0\}} \frac{\|Ax\|_2}{\|x\|_2}. \tag{3.10}$$

For generalized singular values a similar characterization is possible.

Theorem 3.4.2 *The generalized singular values $\sigma_k,\ k = 1, \ldots, r+j$ of $\{A, B\}$ can be characterized as*

$$\sigma_k = \min_{\substack{H \subset \mathbb{R}^n \\ dim(H)=k}} \max_{x \in H \setminus \{0\}} \frac{\|Ax\|_2}{\|Bx\|_2}.$$

Proof The singular values s_k of the matrix Q_A from (3.8) can be characterized as

$$s_k = \min_{\substack{H \subset \mathbb{R}^n \\ dim(H)=k}} \max_{y \in H \setminus \{0\}} \frac{\|Q_A y\|_2}{\|y\|_2}.$$

Since $\sigma_k = s_k/\sqrt{1-s_k^2}$ we obtain

$$\sigma_k = \min_{\substack{H \subset \mathbb{R}^n \\ dim(H)=k}} \max_{\substack{y \in H \\ \|y\|_2=1}} \frac{\|Q_A y\|_2}{\sqrt{1-\|Q_A y\|_2^2}}.$$

The matrix R is nonsingular. Therefore $\dim(H) = \dim(\{x \in \mathbb{R}^n | Rx \in H\})$. Since also

$$\frac{\|Ax\|_2}{\|Bx\|_2} = \frac{\|Q_A y\|_2}{\|Q_B y\|_2} = \frac{\|Q_A y\|_2}{\sqrt{1-\|Q_A y\|_2^2}}$$

for $y = Rx$ and $\|y\|_2 = 1$ the result follows. ∎

Generalized singular values are closely related to principal angles between subspaces.

Theorem 3.4.3 *Let $\theta_1 \leq \cdots \leq \theta_{r+j} < \pi/2$ be the principal angles between the subspaces $\mathcal{A}$ and $\mathcal{B}$ of $\mathbb{R}^n$. Let P_A be the orthogonal projector onto $\mathcal{A}$ and $P_A^\perp$ its orthogonal complement. Let the matrix B be defined such that its columns form a basis of $\mathcal{B}$. Then the finite generalized singular values $\sigma_k,\ k = 1, \ldots, r+j$ of the pencil $\{P_A^\perp B, P_A B\}$ are related to the principal angles θ_k by $\sigma_k = \tan\theta_k$.*

Proof The proof is a simple consequence of the CS decomposition in (3.5). Let $B = Q_B R$ and multiply (3.5) by R to obtain

$$\begin{bmatrix} (I - Q_A Q_A^T)B \\ Q_A Q_A^T B \end{bmatrix} = \begin{bmatrix} U & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} S \\ C \end{bmatrix} V^T R. \tag{3.11}$$

With $P_A = Q_A Q_A^T$ and $P_A^\perp = I - Q_A Q_A^T$ equation (3.11) is just the generalized singular value decomposition of the pencil $\{P_A^\perp B, P_A B\}$. Since the generalized singular value pairs $(s_k, c_k), k = 1, \ldots, n$ are the sines and cosines of the principal angles between $\mathcal{A}$ and $\mathcal{B}$ and $c_k > 0$ for $k = 1, \ldots, r+j$, the result follows. ∎

An interesting corollary of this statement is a minimax characterization for the tangents of principal angles between subspaces.

Corollary 3.4.4 *Let the notation be as in Theorem* 3.4.3. *Then* $\theta_1, \ldots, \theta_{r+j}$ *can be characterized as*

$$\tan\theta_k = \min_{\substack{H\subset\mathbb{R}^n \\ dim(H)=k}} \max_{x\in H\setminus\{0\}} \frac{\|P_A^{\perp} Bx\|_2}{\|P_A Bx\|_2}, \quad k = 1, \ldots, r+j. \tag{3.12}$$

Proof The result follows directly from Theorem 3.4.2 and 3.4.3. ∎

To conclude this section we show how the generalized singular values of the pencil $\{A, B\}$ can be expressed as angles between certain subspaces.

Corollary 3.4.5 *Let* $A \in \mathbb{R}^{m_1\times n}$ *with* $m_1 \geq n$ *and* $B \in \mathbb{R}^{m_2\times n}$. *Define* $Y = \begin{bmatrix} A \\ B \end{bmatrix}$ *and assume that* $rank(Y) = n$. *Let* $\mathcal{Y}$ *be the space spanned by the columns of* Y *and define* $\mathcal{D}_0 \subset \mathbb{R}^{m_1+m_2}$ *as the space of vectors which first* m_1 *entries are zero. Let the finite generalized singular values of the pencil* $\{A, B\}$ *be* $\sigma_1, \ldots, \sigma_{r+j}$. *Then the principal angles* $0 \leq \theta_k < \pi/2$ *between* $\mathcal{Y}$ *and* $\mathcal{D}_0$ *are given as* $\tan\theta_k = \sigma_k$.

Proof Let $P_{\mathcal{D}_0}$ be the projector onto $\mathcal{D}_0$ and $P_{\mathcal{D}_0}^{\perp}$ its orthogonal complement. The generalized singular value pairs of $\{P_{\mathcal{D}_0}^{\perp} Y, P_{\mathcal{D}_0} Y\}$ are identical to those of $\{A, B\}$. The proof therefore follows immediately from Theorem 3.4.3. ∎

A similar result is also proved in [87] by Zha.

3.5 The GSVD as a unified approach for the Method of Particular Solutions

We are now ready to show how to apply the GSVD to the Method of Particular Solutions. Let $\mathcal{A}(\lambda) := \text{span}\{u^{(1)}, \ldots, u^{(n)}\}$ be a given space of particular solutions

satisfying $-\Delta u = \lambda u$ in Ω. As in the approach of Barnett we can attempt to minimize the boundary tension

$$t(u) = \frac{\|u\|_{\partial\Omega}}{\|u\|_{\Omega}},$$

over all $u \in \mathcal{A}(\lambda)$, where $\|u\|_{\partial\Omega}$, $\|u\|_{\Omega}$ and the corresponding inner products are defined as in Section 2.4. Every $u \in \mathcal{A}(\lambda)$ can be written as

$$u = \sum_{k=1}^{n} x_k u^{(k)}.$$

We can rewrite this expression using a matrix-vector product form by defining a semi-infinite matrix $A^{(s)}(\lambda)$ as[2]

$$A^{(s)}(\lambda) = \left[u_1(z), \ldots, u_n(z)\right], \quad z \in \overline{\Omega}.$$

(A Matlab toolbox that can operate with such matrices was recently developed by Battles [9]). The columns of this matrix are not vectors of functions evaluated at discrete points but the functions themselves. Every element $u \in \mathcal{A}(\lambda)$ now has the simple form $u = A^{(s)}(\lambda)x$. By defining the two semi-infinite matrices

$$\begin{aligned} A_B^{(s)}(\lambda) &= A^{(s)}(\lambda), \quad z \in \partial\Omega, \\ A_I^{(s)}(\lambda) &= A^{(s)}(\lambda), \quad z \in \Omega, \end{aligned}$$

the tension $t(u)$ can be reformulated as

$$t(x) = \frac{\|A_B^{(s)}(\lambda)x\|_{\partial\Omega}}{\|A_I^{(s)}(\lambda)x\|_{\Omega}}. \tag{3.13}$$

We are interested in the minimum

$$t_m(\lambda) = \min_{x \in \mathbb{R}^n} \frac{\|A_B^{(s)}(\lambda)x\|_{\partial\Omega}}{\|A_I^{(s)}(\lambda)x\|_{\Omega}}.$$

If the matrices $A_B^{(s)}(\lambda)$ and $A_I^{(s)}(\lambda)$ were discrete, the solution would simply be given by the smallest generalized singular value of the pencil $\{A_B^{(s)}(\lambda), A_I^{(s)}(\lambda)\}$. Since this is not the case one way of computing $t_m(\lambda)$ is to square $t(x)$ and to solve the corresponding eigenvalue problem. In our semi-infinite matrix notation this becomes

$$A_B^{(s)}(\lambda)^T A_B^{(s)}(\lambda)x(\lambda) = \mu(\lambda) A_I^{(s)}(\lambda)^T A_I^{(s)}(\lambda)x(\lambda)$$

[2]In this section we use the index s to distinguish semi-infinite matrices from ordinary matrices

where the matrices $A_B^{(s)}(\lambda)^T A_B^{(s)}(\lambda)$ and $A_I^{(s)}(\lambda)^T A_I^{(s)}(\lambda)$ are defined as

$$(A_B^{(s)}(\lambda)^T A_B^{(s)}(\lambda))_{ij} = \langle u_i, u_j \rangle_{\partial\Omega}, \quad (A_I^{(s)}(\lambda)^T A_I^{(s)}(\lambda))_{ij} = \langle u_i, u_j \rangle_{\Omega}.$$

These are just the matrices $F(\lambda)$ and $G(\lambda)$ from the generalized eigenvalue problem (2.6). But we want to avoid working with a squared formulation. So how can we work with $A_B^{(s)}(\lambda)$ and $A_I^{(s)}(\lambda)$ directly? We could attempt to discretize these two matrices by evaluating the column functions of $A_B^{(s)}(\lambda)$ on a number of boundary collocation points and the column functions of $A_I^{(s)}(\lambda)$ on some interior points. This then leads to the discrete problem

$$t_m(\lambda) \approx \sigma_1(\lambda) = \min_{x \in \mathbb{R}^n} \frac{\|A_B(\lambda)x\|_2}{\|A_I(\lambda)x\|_2},$$

where $\sigma_1(\lambda)$ is the smallest generalized singular value of the discretized pencil $\{A_B(\lambda), A_I(\lambda)\}$[3]. From Corollary 3.4.5 it follows that $\sigma_1(\lambda) = \tan\theta_1(\lambda)$, where $\theta_1(\lambda)$ is the angle between the discretized spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$. Therefore, we have

$$t_m(\lambda) \approx \tan\theta_1(\lambda).$$

So when do we have $F(\lambda) \approx A_B(\lambda)^T A_B(\lambda)$ and $G(\lambda) \approx A_I(\lambda)^T A_I(\lambda)$? Every entry of $F(\lambda)$ and $G(\lambda)$ is an L^2 inner product, which is evaluated by a quadrature rule. Using the trapezium rule and equidistributed points on $\partial\Omega$ we obtain

$$F_{ij}(\lambda) \approx h \sum_{k=1}^{N} u^{(i)}(z_k) u^{(j)}(z_k),$$

where h is the distance between two points on $\partial\Omega$. If the same evaluation points are used as discretization points for $A_B(\lambda)$ we find $F \approx hA_B(\lambda)^T A_B(\lambda)$, where the error $F(\lambda) - hA_B(\lambda)^T A_B(\lambda)$ is determined by the error of the quadrature rule. Similarly, we have $G(\lambda) \approx \tilde{h}^2 A_I(\lambda) A_I(\lambda)$ using a quadrature rule on a regular grid with grid size $\tilde{h}$. By a scaling argument we can assume $h = \tilde{h} = 1$. Hence, the error $t_m(\lambda) - \sigma_1(\lambda)$ depends on the underlying quadrature rule. We do not want to discuss this in greater detail here since for practical purposes it is not necessary to choose the discretization points such that $\sigma_1(\lambda)$ is close to $t_m(\lambda)$. This would lead to a very high number of interior points. Instead we just need enough interior points to guarantee that the

[3] We recently learned that in unpublished work Eisenstat also considered using the GSVD for the Method of Particular Solutions. His starting point was the minimization of error bounds for the MPS.

approximate eigenfunctions stay bounded away from zero. In most experiments this was achieved by choosing just a few dozen interior points.

Let us now summarize the results from this chapter. To repair the MPS of Fox, Henrici and Moler it is necessary to additionally use interior points. Using boundary and interior points we can formulate a subspace angle method that minimizes the smallest principal angle $\theta_1(\lambda)$ between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$, where these two spaces are represented on the boundary and interior discretization points. The subspace angle method computes for each λ the value $s_1 = \sin\theta_1(\lambda)$. An equivalent formulation of the subspace angle method can be derived as the generalized singular value decomposition of the pencil $\{A_B(\lambda), A_I(\lambda)\}$. For the smallest generalized singular value $\sigma_1(\lambda)$ of this pencil it follows from Corollary 3.4.5 that $\sigma_1(\lambda) = \tan\theta_1(\lambda)$. Therefore, the subspace angle method can be seen as a direct generalization of the original MPS by going over from computing the smallest singular value of $A_B(\lambda)$ to computing the smallest generalized singular value of $\{A_B(\lambda), A_I(\lambda)\}$. Barnett's method solves the generalized eigenvalue problem $F(\lambda)x(\lambda) = \mu_1(\lambda)G(\lambda)x(\lambda)$, which can be interpreted as the square of the GSVD of $\{A_B(\lambda), A_I(\lambda)\}$. Squaring the GSVD problem leads to a loss of accuracy as already discussed in Chapter 2. Further implications of the squared generalized eigenvalue formulation in comparison to the GSVD are discussed in Chapter 4. Heller's PWDM uses exactly one point in the interior but does not use a GSVD approach to solve the resulting problem. Using the GSVD we can incorporate an arbitrary number of interior points, thereby solving stability problems which can result from only having one interior point.

At the beginning of this chapter we posed the two goals of formulating a well-conditioned problem and having a stable method to solve this problem. From the first example of the L-shaped region given in this chapter it seems that using interior points together with the subspace angle method achieves these goals. Indeed, in the next chapter we show that the smallest subspace angle $\theta_1(\lambda)$ is well-conditioned if λ is close to an eigenvalue λ_k of (1.1), allowing us to approximate eigenvalues of (1.1) to high accuracy.

Chapter 4

Numerical stability

In the last chapter we derived the subspace angle method and its equivalent formulation as a generalized singular value problem. The first results on the L-shaped region looked promising. But we haven't yet discussed the effect of ill-conditioning in the basis on the reliability of the method. It is well known that singular values are perfectly conditioned. The same is true for eigenvalues of symmetric matrices. However, for generalized singular values the picture looks different. Depending on the pencil $\{A, B\}$, the condition number of generalized singular values can be arbitrarily bad. Since the GSVD underlies the subspace angle method, this raises the question how reliable the subspace angle method is and if we can trust the results that we obtain with it. This section starts with two examples that show how ill-conditioning in the approximation basis can introduce visible numerical errors in the computed generalized singular values. Then we will discuss currently known perturbation results and condition numbers for generalized singular value problems and apply them to the subspace angle method to obtain accurate bounds on the forward error of the method. In the last section of this chapter we compare the condition numbers in the subspace angle method to those of the corresponding formulations as generalized eigenvalue problem in Barnett's method. For all computations in this chapter we use standard Matlab functions.

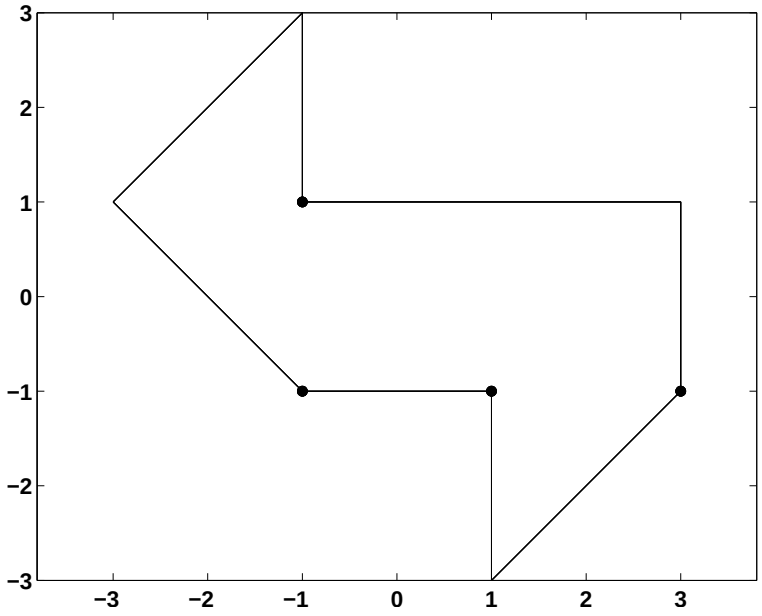


Figure 4.1: The GWW-1 isospectral drum. The singular corners are marked by dots.

4.1 Two examples for highly ill-conditioned problems

Our first example is the first of the two famous Gordon-Webb-Wolpert (GWW) isospectral drums, shown in Figure 4.1. There are four singular corners around which an eigenfunction of (1.1) cannot be analytically continued (more on analytic continuation in Chapter 6). Highly accurate eigenvalue approximations can only be achieved if these singularities are reflected in the approximation basis by using expansions with Fourier-Bessel functions around the singular corners.

To obtain the first eigenvalue to 12 digits of accuracy, an expansion with 60 Fourier-Bessel basis functions around each of the 4 singular corners is necessary. The resulting eigenvalue approximation is $\lambda_1 \approx 2.53794399980$[1]. The tangent of the smallest principal angle computed with the Matlab GSVD function for this shape is plotted in Figure 4.2 (for GSVD computations we use throughout this thesis the Matlab GSVD function, which performs a QR followed by a CS decomposition). Before the curve bends down to the first eigenvalue, it is heavily oscillating, but then it becomes smoother

[1]Beautiful pictures of the first eigenmodes on the isospectral drums and computations of their eigenvalues to 12 digits of accuracy were published by Driscoll in 1997 [21].

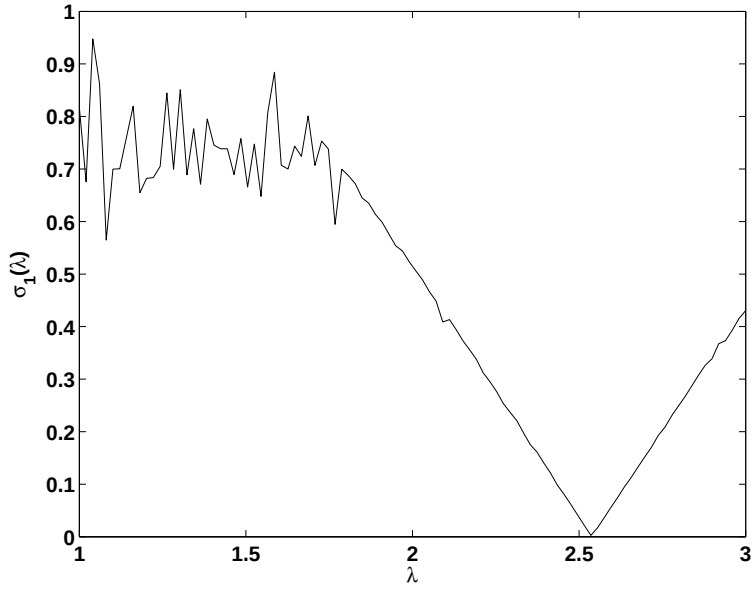


Figure 4.2: The smallest generalized singular value $\sigma_1(\lambda)$ for the GWW-1 isospectral drum. Away from the eigenvalue, the curve shows large oscillations.

close to the eigenvalue. Figure 4.3 shows the condition number of $A(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}$[2] for different values of λ. For all these values of λ the basis is numerically singular. At the eigenvalue λ_1 the condition number is 1.2×10^{16}. But still we are able to approximate λ_1 to an accuracy of 12 digits.

An artificial but more striking example of the possible effects of ill-conditioning is given in Figure 4.4. This shows the curve of the smallest generalized singular value for the MPS on the unit square with 20 Fourier-Bessel basis terms around each of the four corners. An expansion at a single corner would be sufficient to obtain the first eigenvalue $\lambda_1 = 2\pi^2$ up to an accuracy of machine precision; the expansions at the other corners are redundant. Due to the redundant information in the basis the curve shows large oscillations. But these oscillations seem to decrease near λ_1. In this chapter we show that even in the presence of such oscillations, highly accurate approximations to eigenvalues of (1.1) are still possible.

[2]Here and in all other examples we work with bases in which every column is scaled to unit norm to avoid ill-conditioning effects caused by the different scaling of basis functions.

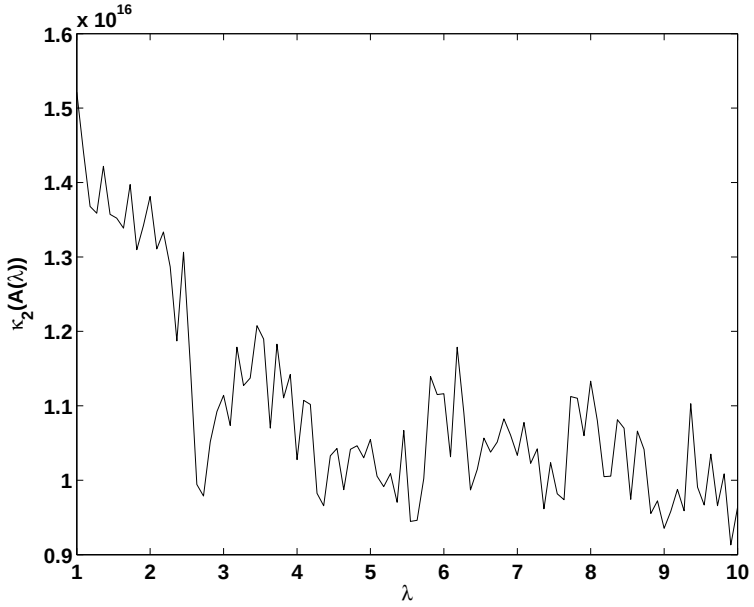


Figure 4.3: The condition number of the Fourier-Bessel basis in the case of the GWW-1 isospectral drum. The basis is numerically singular for all values of λ.

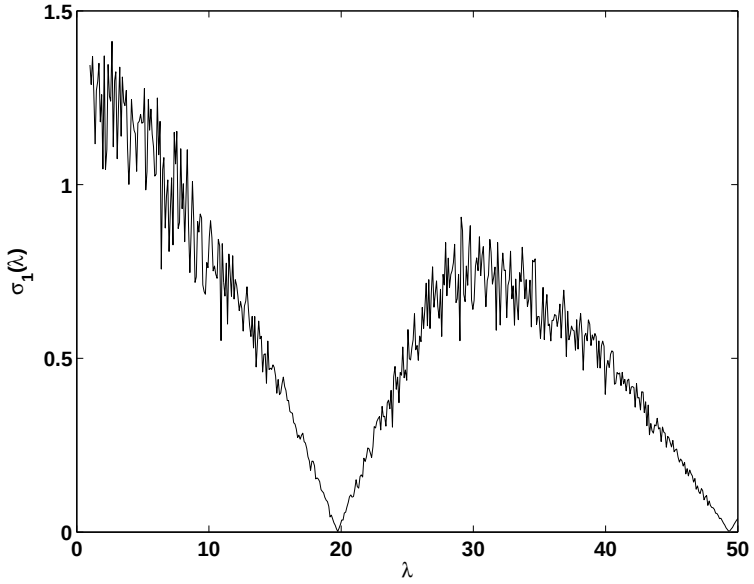


Figure 4.4: On a unit square expansions at all four corners lead to large oscillations in the computed generalized singular values.

4.2 Perturbation results for principal angles between subspaces

The subspace angle algorithm computes the smallest principal angle between the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ represented by the matrices $A(\lambda)$ and D_0. The space $\mathcal{D}_0$ of functions that are zero at the boundary points is exactly represented by the orthonormal columns of D_0. But the matrix $A(\lambda)$ consists of basis functions evaluations and is usually highly ill-conditioned. Therefore, it is feasible to ask what changes in the principal angles are caused by small changes in $A(\lambda)$.

Let $A \in \mathbb{R}^{m\times n}, m \geq n$ be given and denote by $\mathcal{R}(A)$ the space spanned by the columns of A. Let $\delta A \in \mathbb{R}^{m\times n}$ be a small perturbation of A. The first question is how far away is the space $\mathcal{R}(A)$ from the space $\mathcal{R}(A+\delta A)$ spanned by the columns of $A+\delta A$. This question was answered by Wedin in 1983.

Theorem 4.2.1 (Wedin, [86]) *Let $\theta_1, \ldots, \theta_n$ be the principal angles between $\mathcal{R}(A)$ and $\mathcal{R}(A+\delta A)$. Then*

$$\sin\theta_k \leq \|\delta A\|_2 \|A^\dagger\|_2, \quad k = 1, \ldots, n.$$

Proof Let P_A be the projector onto $\mathcal{R}(A)$ and $P_{A+\delta A}$ the projector onto $A+\delta A$. Then $(I - P_{A+\delta A})(A+\delta A) = 0$ and therefore

$$(I - P_{A+\delta A})\delta A = -(I - P_{A+\delta A})A.$$

We have $P_A = AA^\dagger$ and therefore

$$(I - P_{A+\delta A})\delta A A^\dagger = -(I - P_{A+\delta A})P_A.$$

The sines of the principal angles θ_k are the singular values of $(I-P_{A+\delta A})P_A$. Therefore,

$$\sin\theta_k \leq \|(I - P_{A+\delta A})P_A\|_2 \leq \|\delta A\|_2\|A^\dagger\|_2. \quad \blacksquare$$

Hence, if the columns of A form a highly ill-conditioned basis of $\mathcal{R}(A)$, the space $\mathcal{R}(A+\delta A)$ can flutter arbitrarily under small perturbations δA. Let us now ask how the principal angles between the range of the column spaces of the matrices A and

B change under small perturbations. From the last theorem we can expect that the perturbation bounds essentially depend on the condition numbers of A and B. But depending on the direction of the perturbations one can obtain significantly better bounds. This was analyzed by Golub and Zha in 1994.

Theorem 4.2.2 (Golub, Zha, [31]) *Let A and $\tilde{A}$, and B and $\tilde{B}$ have the same rank, i.e.*

$$rank(A) = rank(\tilde{A}) = p \quad rank(B) = rank(\tilde{B}) = q$$

and assume $p \geq q$. For any orthogonally invariant norm $\|\cdot\|$, let the condition numbers of A and B be defined as follows:

$$\kappa(A, \|\cdot\|) = \|A\|\|A^\dagger\|_2, \quad \kappa(B, \|\cdot\|) = \|B\|\|B^\dagger\|_2.$$

Let S be the matrix of sines of the principal angles and C be the matrix of cosines of the principal angles between $\mathcal{R}(A)$ and $\mathcal{R}(B)$. Similarly let $\tilde{S}$ and $\tilde{C}$ be the corresponding matrices for the principal angles between the perturbed spaces $\mathcal{R}(\tilde{A})$ and $\mathcal{R}(\tilde{B})$. Then we have

$$\|C - \tilde{C}\|_2 \leq \sqrt{2} \left\{ \kappa(A, \|\cdot\|) \cos\theta_1 \frac{\|A - \tilde{A}\|}{\|A\|} + \kappa(B, \|\cdot\|) \cos\phi_1 \frac{\|B - \tilde{B}\|}{\|B\|} \right\}$$

and

$$\|S - \tilde{S}\|_2 \leq \sqrt{2} \left\{ \kappa(A, \|\cdot\|) \cos\theta_2 \frac{\|A - \tilde{A}\|}{\|A\|} + \kappa(B, \|\cdot\|) \cos\phi_2 \frac{\|B - \tilde{B}\|}{\|B\|} \right\},$$

with

$$\begin{aligned} \theta_1 &= \theta_{\min}(\mathcal{C}(A, \tilde{A}), \mathcal{R}(B)), \quad \theta_2 = \theta_{\min}(\mathcal{C}(A, \tilde{A}), \mathcal{R}(B)^\perp) \\ \phi_1 &= \theta_{\min}(\mathcal{R}(\tilde{A}), \mathcal{C}(B, \tilde{B})), \quad \phi_2 = \theta_{\min}(\mathcal{R}(\tilde{A})^\perp, \mathcal{C}(B, \tilde{B})), \end{aligned}$$

where $\mathcal{C}(A, \tilde{A})$ is the orthogonal complement of $\mathcal{R}(A) \cap \mathcal{R}(\tilde{A})$ in $\mathcal{R}(A) + \mathcal{R}(\tilde{A})$, and $\mathcal{C}(B, \tilde{B})$ is the orthogonal complement of $\mathcal{R}(B) \cap \mathcal{R}(\tilde{B})$ in $\mathcal{R}(B) + \mathcal{R}(\tilde{B})$. For the spectral norm we can substitute the constant $\sqrt{2}$ with 1.

This result does not only depend on the size of the perturbation but also on its direction. The problem with this error bound is that it treats all principal angles together. If the theorem delivers a perturbation bound for the principal angles of, say, 10^{-5}, this is reasonable for the larger angles but catastrophic for very small principal angles. We are especially interested in very small angles. Therefore, we need to treat them separately. Since the subspace angle algorithm can be interpreted as a generalized singular value computation, we can apply condition numbers for generalized singular values to obtain error estimates.

4.3 Condition numbers for generalized singular value problems

It is well known that the singular values of a matrix A are perfectly conditioned under perturbations in A. This follows from the fact that if σ is a singular value of A, then $\pm\sigma$ are eigenvalues of the symmetric matrix $\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$, and eigenvalues of symmetric matrices are perfectly conditioned (see [40] for a proof).

For generalized singular value problems the situation is more complicated. Condition numbers for these problems were defined and analyzed by Sun [71].

Let $A \in \mathbb{R}^{n\times p}$ and $B \in \mathbb{R}^{m\times p}$ and define $Y = \begin{bmatrix} A \\ B \end{bmatrix}$. Furthermore, let $\operatorname{rank}(Y) = p$. We define a perturbed pencil $\{\tilde{A}, \tilde{B}\}$ as $\tilde{A} = A + E$ and $\tilde{B} = B + F$. If (s, c) is a generalized singular value pair of $\{A, B\}$, the corresponding perturbed generalized singular value pair of $\{\tilde{A}, \tilde{B}\}$ is denoted by $(\tilde{s}, \tilde{c})$. Furthermore, let $\sigma = \frac{s}{c}$ and $\tilde{\sigma} = \frac{\tilde{s}}{\tilde{c}}$ be the corresponding generalized singular values. Then a condition number $c(\sigma)$ for σ can be defined in the following way.

Condition number for generalized singular values Let γ_A,γ_B and ξ be positive parameters. Then the condition number $c(\sigma)$ of a generalized singular value σ of the pencil $\{\tilde{A}, \tilde{B}\}$ is defined as

$$c(\sigma) = \lim_{\delta\to 0} \sup_{\left\| \left(\frac{\|E\|_2}{\gamma_A}, \frac{\|F\|_2}{\gamma_B}\right)^T \right\|_\infty \le \delta} \frac{|\tilde{\sigma} - \sigma|}{\xi\delta}.$$

If $\gamma_A = \gamma_B = \xi = 1$ then $c(\sigma)$ is an absolute condition number. Relative condition numbers are obtained for the parameters $\gamma_A = \|A\|_2$,$\gamma_B = \|B\|_2$ and $\xi = \sigma$. By setting $E = 0$ or $F = 0$ one obtains the condition numbers

$$c_A(\sigma) = \lim_{\delta\to 0} \sup_{\frac{\|E\|_2}{\gamma_A} \le \delta, F=0} \frac{|\tilde{\sigma} - \sigma|}{\xi\delta} \quad c_B(\sigma) = \lim_{\delta\to 0} \sup_{\frac{\|F\|_2}{\gamma_A} \le \delta, E=0} \frac{|\tilde{\sigma} - \sigma|}{\xi\delta}.$$

The forward error can be estimated as

$$\frac{|\tilde{\sigma} - \sigma|}{\xi} \le c(\sigma) \left\| \left(\frac{\|E\|_2}{\gamma_A}, \frac{\|F\|_2}{\gamma_B} \right)^T \right\|_\infty + O\left(\left\| \left(\frac{\|E\|_2}{\gamma_A}, \frac{\|F\|_2}{\gamma_B} \right)^T \right\|_\infty^2 \right). \tag{4.1}$$

The following theorem expresses the condition number $c(\sigma)$ in computable quantities.

Theorem 4.3.1 (Sun, [71]) *Let x be the right generalized singular vector associated with σ. Then the condition number $c(\sigma)$ can be expressed as*

$$c(\sigma) = \frac{\|x\|_2(\gamma_B\|Ax\|_2 + \gamma_A\|Bx\|_2)}{\xi\|Bx\|_2^2}. \tag{4.2}$$

If we assume that $\gamma_A = \gamma_B = \xi = 1$ then we obtain for the resulting absolute condition number the expression

$$c(\sigma) = \frac{\|x\|_2}{\|Bx\|_2}(1+\sigma) = \frac{\|x\|_2}{c}(1+\sigma) \tag{4.3}$$

since $\sigma = \frac{s}{c} = \frac{\|Ax\|_2}{\|Bx\|_2}$ and $\|Bx\|_2 = c$. Therefore, if the generalized singular value σ is small, the condition number $c(\sigma)$ mainly depends on $\|x\|_2$. Let $\begin{bmatrix} Q_A \\ Q_B \end{bmatrix} R$ be the QR decomposition of $\begin{bmatrix} A \\ B \end{bmatrix}$ and let τ be the smallest singular value of R. Then a crude upper bound on $\|x\|_2$ is given by τ^{-1} since from Theorem 3.4.1 it follows that there is a vector v with $\|v\|_2 = 1$ such that $x = R^{-1}v$. We find $\|x\|_2 \leq \tau^{-1}$. By using the structure of the GSVD we can often give better bounds on $\|x\|_2$.

Lemma 4.3.2 *Let $\sigma = s/c$ be a generalized singular value of the pencil $\{A, B\}$ and let x be its corresponding right generalized singular vector. Then*

$$\|x\|_2 \leq \min\{\frac{s}{\tau_A}, \frac{c}{\tau_B}, \frac{1}{\tau}\}, \tag{4.4}$$

where τ_A and τ_B are the smallest singular values of A and B and τ is the smallest singular value of $Y = \begin{bmatrix} A \\ B \end{bmatrix}$.

Proof The case $\|x\|_2 \leq \tau^{-1}$ was already discussed. From Theorem 3.4.1 it follows that $\|Ax\|_2 = s$ and $\|Bx\|_2 = c$. With $\hat{x} := x/\|x\|_2$ we obtain $\|x\|_2 = \frac{s}{\|A\hat{x}\|_2}$ and $\|x\| = \frac{c}{\|B\hat{x}\|_2}$. Since $\|A\hat{x}\|_2 \geq \tau_A$ and $\|B\hat{x}\|_2 \geq \tau_B$, the proof follows. ∎

From Lemma 4.3.2 it follows that if one of the matrices A, B or Y is well-conditioned, then $\|x\|_2$ is small. Otherwise $\|x\|_2$ is only small if $s \ll 1$ or $c \ll 1$. The case $c \ll 1$ is not interesting for us since then $s \approx 1$. But the subspace angle method aims to find a λ for which the sine $s_1(\lambda)$ of the smallest principal angle between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ is small.

4.4 Backward stablility of the subspace angle method

To apply the bound on the forward error in (4.1) we need to know the backward error of the subspace angle method. The algorithm consists of two steps. First the QR decomposition

$$\begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix} = \begin{bmatrix} Q_B(\lambda) \\ Q_I(\lambda) \end{bmatrix} R(\lambda)$$

is computed. Then in the second step the smallest singular value $s_1(\lambda)$ of $Q_B(\lambda)$ is computed, which is the sine of the wanted principal angle. A general stability analysis for the Björck-Golub algorithm for computing principal angles between subspaces is given by Drmač in [22]. Due to the special structure of the subspace angle algorithm we can give a simplified analysis here.

Matlab computes the QR factorization of a given matrix A using the Lapack QR factorization, which is based on Householder reflections. An analysis of Householder QR algorithms can be found in [38]. Let $\tilde{\gamma} := \frac{cku}{1-cku}$, where c is a small integer constant and u is the unit round-off. Then for the Householder QR algorithm the following theorem holds, which summarizes Theorem 19.4 of [38] and the discussion afterwards.

Theorem 4.4.1 *Let $\hat{Q} \in \mathbb{R}^{m\times n}$ and $\hat{R} \in \mathbb{R}^{n\times n}$ be the computed QR factors of $A \in \mathbb{R}^{m\times n}$, $(m \geq n)$ obtained via the Householder QR algorithm. Then there exists a matrix $Q \in \mathbb{R}^{m\times n}$ with orthonormal columns such that*

$$A + \Delta A = Q\hat{R},$$

where

$$\|\Delta A(:,j)\|_2 \leq \tilde{\gamma}_{mn}\|A(:,j)\|_2, \quad j = 1:n.$$

For $\hat{Q}$ it holds that

$$\|\hat{Q} - Q\|_F \leq \sqrt{n}\tilde{\gamma}_{mn}.$$

We can now prove a mixed stability result of the subspace angle algorithm. The following theorem is essentially a simplified version of Theorem 2.1 of [22], where the general form of the Björck-Golub algorithm was considered. In the following theorems we ignore the extra treatment of the possible $O(\epsilon_{mach})$ error from the evaluation of the basis functions in $A(\lambda)$.

Theorem 4.4.2 (Mixed stability of the subspace angle algorithm) *Let $\tilde{s}_1(\lambda)$ be the computed sine of the smallest principal angle from the subspace angle algorithm applied to the matrix of sampled basis functions $A(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix} \in \mathbb{R}^{(n+m)\times p}$ with $A_B(\lambda) \in \mathbb{R}^{n\times p}$ and $A_I(\lambda) \in \mathbb{R}^{m\times p}$. Then there exists a value $\bar{s}_1(\lambda)$, which is the exact sine computed from the subspace angle method applied to the matrix $\bar{A}(\lambda) = A(\lambda) + \Delta A(\lambda)$, such that $|\tilde{s}_1(\lambda) - \bar{s}_1(\lambda)| \leq q(n,p)\epsilon_{mach} + \sqrt{p}\tilde{\gamma}_{(n+m)p}$ and $\|\Delta A(:,j)(\lambda)\|_2 \leq \tilde{\gamma}_{(n+m)p}\|A(:,j)(\lambda)\|_2$, where $q(n,p)$ is a modestly growing function of n and p.*

Proof Let $\hat{Q}(\lambda) = \begin{bmatrix} \hat{Q}_B(\lambda) \\ \hat{Q}_I(\lambda) \end{bmatrix}$ and $\hat{R}(\lambda)$ be the computed QR factors of $A(\lambda)$ by the Householder QR algorithm. Then the computed sine $\tilde{s}_1(\lambda)$ is the smallest singular value of $\hat{Q}_B(\lambda) + \Delta\hat{Q}_B(\lambda)$, where $\|\Delta\hat{Q}_B(\lambda)\|_2 \leq q(n,p)\epsilon_{mach}$ and $q(n,p)$ is a modestly growing function of n and p (see [1] for details). Since singular values are perfectly conditioned we have $|\tilde{s}_1(\lambda) - \hat{s}_1(\lambda)| \leq q(n,p)\epsilon_{mach}$, where $\hat{s}_1(\lambda)$ is the exact smallest singular value of $\hat{Q}_B(\lambda)$. From Theorem 4.4.1 it follows that there exists a matrix $\bar{Q}(\lambda)$ with orthonormal columns such that $\|\hat{Q}(\lambda) - \bar{Q}(\lambda)\|_F \leq \sqrt{p}\tilde{\gamma}_{(n+m)p}$ and $A(\lambda) + \Delta A(\lambda) = \bar{Q}(\lambda)\hat{R}(\lambda)$ with $\|\Delta A(:,j)(\lambda)\|_2 \leq \tilde{\gamma}_{(n+m)p}\|A(:,j)(\lambda)\|_2$. Let $\bar{s}_1(\lambda)$ be the exact smallest singular value of $\bar{Q}_B(\lambda)$. Then

$$\begin{aligned} |\tilde{s}_1(\lambda) - \bar{s}_1(\lambda)| &\leq |\tilde{s}_1(\lambda) - \hat{s}_1(\lambda)| + |\hat{s}_1(\lambda) - \bar{s}_1(\lambda)| \\ &\leq q(n,p)\epsilon_{mach} + \sqrt{p}\tilde{\gamma}_{(n+m)p} \end{aligned}$$

since

$$\|\hat{Q}(\lambda) - \bar{Q}(\lambda)\|_2 \leq \|\hat{Q} - \tilde{Q}\|_F \leq \sqrt{p}\tilde{\gamma}_{(n+m)p},$$

and therefore $|\hat{s}_1(\lambda) - \bar{s}_1(\lambda)| \leq \sqrt{p}\tilde{\gamma}_{(n+m)p}$. ∎

Theorem 4.4.2 states that the subspace angle method has a backward error component resulting from the QR factorization, namely from $A(\lambda) + \Delta A(\lambda) = \bar{Q}(\lambda)\hat{R}(\lambda)$, where $\bar{Q}(\lambda)$ is a matrix with orthonormal columns and $\hat{R}$ is the computed upper triangular factor and a forward error component resulting from the subsequent singular value decomposition. However, this forward error component is in the order of machine precision since singular values are perfectly conditioned. The important influence is the error produced by working with $A(\lambda) + \Delta A(\lambda)$ instead of working with $A(\lambda)$. Theorem 4.2.1 states that $\mathcal{R}(A(\lambda)+\Delta A(\lambda))$ can flutter almost arbitrarily under small perturbations $\Delta A(\lambda)$ if $A(\lambda)$ is ill-conditioned.

4.5 The forward error of the subspace angle method

In the last section we derived the backward error of the subspace angle method. By combining this with the condition numbers derived in Theorem 4.3.1 we can now derive bounds for the forward error of the computed subspace angle. Although under small perturbations $\Delta A(\lambda)$ the space $\mathcal{R}(A(\lambda))$ can flutter arbitrarily, small principal angles only suffer from small absolute perturbations, which will still allow us enough accuracy for the subspace angle method to work and will also explain the behavior in the Figures 4.2 and 4.4.

Theorem 4.5.1 *Let* $A(\lambda) = \begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}$ *be the matrix of sampled basis functions with* $A_B(\lambda) \in \mathbb{R}^{n\times p}$ *and* $A_I(\lambda) \in \mathbb{R}^{m\times p}$. *Let* $\tau_B(\lambda)$ *be the smallest singular value of* $A_B(\lambda)$, $\tau_I(\lambda)$ *the smallest singular value of* $A_I(\lambda)$ *and* $\tau(\lambda)$ *the smallest singular value of* $A(\lambda)$. *Let* $s_1(\lambda)$ *be the exact sine of the smallest principal angle delivered by the subspace angle method and let* $\tilde{s}_1(\lambda)$ *be the computed value. Let* $c_1(\lambda) = \sqrt{1 - s_1(\lambda)^2}$ *be the corresponding cosine. With*

$$\nu(\lambda) := \min\left(\frac{s_1(\lambda)}{\tau_B(\lambda)}, \frac{c_1(\lambda)}{\tau_I(\lambda)}, \frac{1}{\tau(\lambda)}\right)$$

the forward error of the subspace angle method is bounded by

$$\begin{aligned} |\tilde{s}_1(\lambda) - s_1(\lambda)| &\leq \nu(\lambda)\left(1 + \frac{s_1(\lambda)}{c_1(\lambda)}\right)\frac{1}{c_1(\lambda)}\tilde{\gamma}_{(n+m)p}\sqrt{p}\|A(\lambda)\|_2 \\ &+ q(n,p)\epsilon_{mach} + \sqrt{p}\tilde{\gamma}_{(n+m)p} + O(p(\tilde{\gamma}_{(n+m)p}\|A(\lambda)\|_2)^2), \quad (4.5) \end{aligned}$$

where $q(n,p)$ *is a modestly growing function of* n *and* p.

Before we give a proof let us have a closer look at this error bound. The important part of (4.5) is the first line. The second line can safely be assumed to be $O(\epsilon_{mach})$. Consider now the case $s_1(\lambda) \ll 1$. Then $c_1(\lambda) \approx 1$ and (4.5) becomes

$$|\tilde{s}_1(\lambda) - s_1(\lambda)| \lessapprox \nu(\lambda)\tilde{\gamma}_{(n+m)p}\sqrt{p}\|A(\lambda)\|_2.$$

If $\tilde{\gamma}_{(n+m)p} = O(\epsilon_{mach})$ and $\|A(\lambda)\|_2 = O(1)$ the forward error is proportional to $\nu(\lambda)\epsilon_{mach}$. But if $A_B(\lambda)$ and $A_I(\lambda)$ have a common numerical nullspace then $\tau(\lambda) \approx$

$\tau_B(\lambda) \approx \tau_I(\lambda) \approx O(\epsilon_{mach})$[3] and therefore $\nu(\lambda) \approx s_1(\lambda)O(\frac{1}{\epsilon_{mach}})$. We obtain

$$|\tilde{s}_1(\lambda) - s_1(\lambda)| \lessapprox K s_1(\lambda) \tag{4.6}$$

for a modest constant $K > 0$. K will usually be larger than 1. Therefore, we cannot expect $\tilde{s}_1(\lambda)$ to have any correct digits. But the absolute error $|\tilde{s}_1(\lambda) - s_1(\lambda)|$ is proportional to $s_1(\lambda)$. Therefore, if say $s_1(\lambda) = 10^{-8}$, then also $\tilde{s}_1(\lambda)$ won't be much larger than 10^{-8}. It is not excluded that $\tilde{s}_1(\lambda)$ might have a much smaller magnitude than $s_1(\lambda)$. But the resulting minima coming from oscillations of $\tilde{s}_1(\lambda)$ are easily distinguishable from true minima of $s_1(\lambda)$ as the Figures 4.2 and 4.4 show. This gives us enough information to determine the eigenvalues of (1.1) to high accuracy, since we are only interested in the minima of the subspace angle curve and if the unperturbed subspace angle curve becomes small, (4.6) guarantees that the computed curve also becomes small. This reflects very well the behavior observed in Figure 4.4. But the plot in Figure 4.2 looks much better close to the eigenvalue than predicted by the error bound. This is discussed after the following proof.

Close to an eigenvalue the behavior in Figure 4.2 looks much better than the predicted error bound. This is discussed after the following proof.

Proof of Theorem 4.5.1 From Theorem 4.4.2 it follows that $|\tilde{s}_1(\lambda) - \bar{s}_1(\lambda)| \leq q(n,p)\epsilon_{mach} + \sqrt{p}\tilde{\gamma}_{(n+m)p}$, where $\bar{s}_1(\lambda)$ is the exact value delivered from the subspace angle method applied to the matrix $A(\lambda)+\Delta A(\lambda)$ with $\|\Delta A(:,j)(\lambda)\|_2 \leq \tilde{\gamma}_{(n+m)p}\|A(:,j)(\lambda)\|_2$. Let $\Delta A(\lambda)$ be partitioned as $A(\lambda)$, i.e. $\Delta A(\lambda) = \begin{bmatrix}\Delta A_B(\lambda)\\ \Delta A_I(\lambda)\end{bmatrix}$. Define $\bar{\theta}_1(\lambda) = \arcsin \bar{s}_1(\lambda)$. From Corollary 3.4.5 it follows that the smallest generalized singular value $\bar{\sigma}_1(\lambda)$ of the pencil $\{A_B(\lambda) + \Delta A_B(\lambda), A_I(\lambda) + \Delta A_I(\lambda)\}$ is the tangent of $\bar{\theta}(\lambda)$, i.e. $\bar{\sigma}_1(\lambda) = \tan\bar{\theta}(\lambda)$. Using (4.1) we find by choosing $\gamma_A = \gamma_B = \xi = 1$ that

$$|\bar{\sigma}_1(\lambda) - \sigma_1(\lambda)| \leq c(\sigma_1(\lambda))\|\Delta A(\lambda)\|_2 + O(\|\Delta A(\lambda)\|^2),$$

where $\sigma_1(\lambda) = s_1(\lambda)/c_1(\lambda)$ is the smallest generalized singular value of $\{A_B(\lambda), A_I(\lambda)\}$. A short calculation shows that from $\|\Delta A(:,j)(\lambda)\|_2 \leq \tilde{\gamma}_{(n+m)p}\|A(:,j)(\lambda)\|_2$ it follows that $\|\Delta A(\lambda)\|_2 \leq \sqrt{p}\tilde{\gamma}_{(n+m)p}\|A(\lambda)\|_2$. Also for $\theta, \bar{\theta} \in [0, \pi/2)$ it holds that $|\sin\theta - \sin\bar{\theta}| \leq |\tan\theta - \tan\bar{\theta}|$. We obtain

$$|\bar{s}_1(\lambda) - s_1(\lambda)| \leq c(\sigma_1(\lambda))\tilde{\gamma}_{(n+m)p}\sqrt{p}\|A(\lambda)\|_2 + O(p(\tilde{\gamma}_{(n+m)p}\|A(\lambda)\|_2)^2). \tag{4.7}$$

[3]This holds since we assume all columns of $A(\lambda)$ to be scaled to unit norm. Otherwise, the computed smallest singular values of $A_B(\lambda)$ and $A_I(\lambda)$ can become arbitrarily small if the columns are badly scaled.

The distance $|\tilde{s}_1(\lambda) - \bar{s}(\lambda)|$ is given by Theorem 4.4.2 as

$$|\tilde{s}_1(\lambda) - \bar{s}_1(\lambda)| \leq q(n,p)\epsilon_{mach} + \sqrt{p}\tilde{\gamma}_{(n+m)p}. \tag{4.8}$$

From Theorem 4.3.1 and Lemma 4.3.2 the condition number $c(\sigma_1(\lambda))$ can be estimated as

$$c(\sigma_1(\lambda)) \leq \nu(\lambda)(1 + \frac{s_1(\lambda)}{c_1(\lambda)})\frac{1}{c_1(\lambda)}. \tag{4.9}$$

Combining (4.7), (4.8) and (4.9) finishes the proof. ∎

How sharp is the estimate in Theorem 4.5.1 ? The amplification factor of the backward error mainly depends on the estimate of $\|x_1(\lambda)\|_2$ from Lemma 4.3.2, where $x_1(\lambda)$ is the right generalized singular vector for the smallest generalized singular value $\sigma_1(\lambda)$. So let us have a look at the estimated value for $\|x(\lambda)\|_2$ from Lemma 4.3.2 and the computed value of $\|x(\lambda)\|_2$ from Matlab's GSVD function. Figure 4.5 shows the estimated value of $\|x_1(\lambda)\|_2$ (dashed line) compared to the computed value of $\|x_1(\lambda)\|_2$ (solid line) around the eigenvalue λ_1 of the GWW-1 isospectral drum. The closer λ is to λ_1 the smaller becomes $\|x_1(\lambda)\|_2$, which is just the predicted behavior by Lemma 4.3.2. But what does it mean for the subspace angle method if $\|x_1(\lambda)\|_2 \approx 10^{15}$ away from λ_1? Since $s_1(\lambda) = \|A_B(\lambda)x_1(\lambda)\|_2 = O(1)$ away from λ_1, the vector $x_1(\lambda)$ must lie close to the nullspace of $A_B(\lambda)$ and is just scaled up such that $s_1(\lambda) = O(1)$. By a similar argument the vector $x_1(\lambda)$ also lies in the nullspace of $A_I(\lambda)$ away from λ_1. Hence, the approximate eigenfunction is meaningless away from λ_1 and just governed by rounding errors. This is the reason for the oscillations at the beginning of the curve in Figure 4.2. When λ approaches λ_1 the vector $x_1(\lambda)$ moves out of the nullspace of $A_I(\lambda)$ but stays in the nullspace of $A_B(\lambda)$. For example, at $\lambda = \lambda_1 - 10^{-5}$ we have $\|x_1(\lambda)\|_2 \approx 10^9$. Therefore, if $x_1(\lambda)$ is in the numerical nullspace of $A_B(\lambda)$ we can expect $\|A_B(\lambda)x_1(\lambda)\|_2 \approx 10^{-7}$ (since then $\|A_B(\lambda)x_1(\lambda)\|_2/\|x_1(\lambda)\|_2 \approx 10^{-16}$) and $\|A_I(\lambda)x_1(\lambda)\|_2/\|x_1(\lambda)\|_2 \approx 10^{-9}$ (since $\|A_I(\lambda)x_1(\lambda)\|_2 \approx 1$). Indeed, we obtain the following values: $\|A_B(\lambda)x_1(\lambda)\|_2 \approx 9.95 \times 10^{-6}$, $\|A_B(\lambda)x_1(\lambda)\|_2/\|x_1(\lambda)\|_2 \approx 7.85 \times 10^{-15}$, $\|A_I(\lambda)x_1(\lambda)\|_2/\|x_1(\lambda)\|_2 \approx 7.89 \times 10^{-10}$. These values differ slightly from the predictions since the smallest singular values of $A_B(\lambda)$ and $A_I(\lambda)$ are not exactly 10^{-16} but in the magnitude of 10^{-15}. But the qualitative behavior corresponds to what we predicted.

Let us summarize these results. In the ill-conditioned case the subspace angle method always seems to choose a function that is associated with a right singular vector x_1

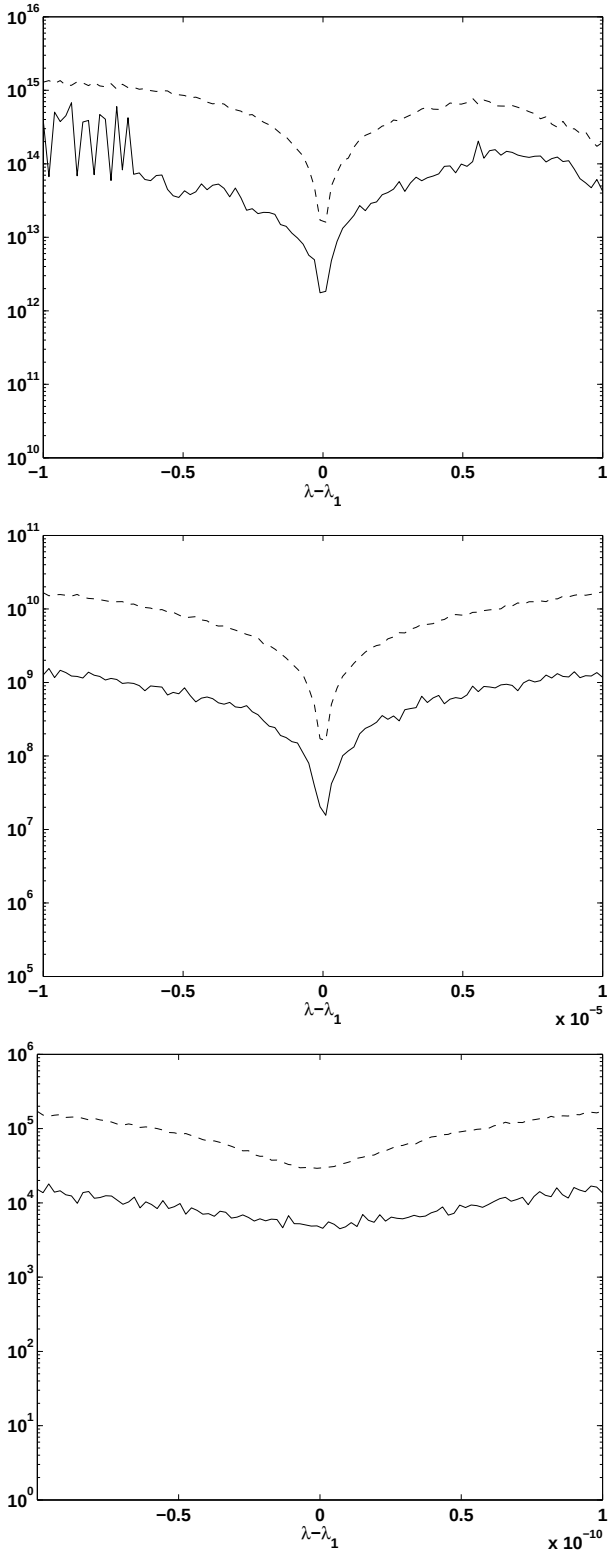


Figure 4.5: The three figures show a comparison between the estimated value of $\|x(\lambda)\|$ (dashed line) and the computed value for $\|x(\lambda)\|_2$ (solid line) around the first eigenvalue λ_1 of the isospectral drum. The closer we get to λ_1 the smaller $\|x(\lambda)\|_2$ is.

which lives in the approximate nullspace of $A_B(\lambda)$. When λ approaches an eigenvalue of (1.1) the vector x_1 moves out of the nullspace of $A_I(\lambda)$ but stays in the nullspace of $A_B(\lambda)$. Therefore, the dependence of the norm estimate in Lemma 4.3.2 on the smallest singular value of $A_B(\lambda)$ reflects this behavior very well.

In Chapter 2 we explained that the original MPS of Fox, Henrici and Moler fails because it cannot distinguish between true eigenfunctions and functions that are zero everywhere on the region. In the ill-conditioned case of the subspace angle method it also happens that far away from an eigenvalue the right generalized singular vector $x_1(\lambda)$ lives in the nullspace of $A(\lambda)$. But the key difference to the MPS of Fox, Henrici and Moler is that such approximate eigenfunctions are scaled up by the high norm of $x_1(\lambda)$ so that they are of unit norm at the boundary and interior points. This scaling guarantees that away from an eigenvalue we cannot obtain an approximate eigenfunction that is close to zero everywhere.

Except in extreme cases like the square region with expansions at all corners or at similar regions, where all Fourier-Bessel expansions approximately span the same space (for example if the region is a small perturbation of the square), the effect of ill-conditioning is much less severe than predicted from the purely algebraic results in Theorem 4.5.1. For example, in Figure 4.2 the oscillations are only visible far away from an eigenvalue. The reason is that the basis functions in this example were chosen to reflect the approximation problem, i.e. all basis functions contain useful information to obtain accurate approximations for the eigenvalues and eigenfunctions on the region. Therefore, although the basis is highly ill-conditioned, it only has redundant information far away from an eigenvalue where the high number of basis functions is useless. Closer to an eigenvalue the structure of the problem leads to a much better behavior than can be predicted by purely looking at the condition number. This is also the reason why the oscillations are so much different than for the square region in Figure 4.4. Here, the basis was artificially chosen to always contain redundant information by introducing expansions around all corners of the region. An expansion around only one corner already delivers enough information to approximate the eigenvalues on the square to high accuracy. Therefore, by artificially introducing redundant information we obtain high oscillations around the subspace angle curve which stay bounded from above relative to the curve. This is exactly the behavior predicted by Theorem 4.5.1. But for most applications this extreme behavior won't

occur. One of the few examples where such a behavior can be observed without artifically introducing redundant information is in the case of perturbations of the square such that we obtain a quadrilateral at which the eigenfunctions have singularities at all four corners. Then approximations at all corners are necessary to obtain highly accurate eigenvalue approximations but we can observe similar oscillations as in Figure 4.4 since except very close to the corners all expansions approximately span the same space in the interior of the region. Then Theorem 4.5.1 tells us that although we have these oscillations, we can obtain approximations to the eigenvalues and eigenfunctions to high accuracy since the oscillations stay bounded from above relative to the subspace angle curve. Hence, we are still able to spot the minima of the subspace angle curve with high accuracy.

4.6 The GSVD and generalized eigenvalue problems

In Section 3.5 we derived the connections between the subspace angle method and the generalized eigenvalue approach of Barnett. Let us now have a closer look at the comparison of the numerical stability of both methods.

A normwise perturbation bound for generalized eigenvalue problems is given in [26]. There, the authors derive condition numbers for the generalized eigenvalue problem

$$Ax = \lambda Bx$$

under normwise perturbation of A and B. Let the distance between the pencils $\{A, B\}$ and $\{\tilde{A}, \tilde{B}\}$ be defined as

$$\delta = \min\{\omega > 0; \|A - \tilde{A}\|_2 \leq \omega\alpha \text{ and } \|B - \tilde{B}\|_2 \leq \omega\beta\},$$

with $\alpha, \beta \geq 0$. Setting $\alpha = \|A\|_2$ and $\beta = \|B\|_2$ leads to a relative normwise distance. Let λ be a semi-simple finite eigenvalue of the pencil $\{A, B\}$. Then the condition number $c_{\text{eig}}(\lambda)$ associated with λ is

$$c_{\text{eig}}(\lambda) = \frac{(\alpha + |\lambda|\beta)\|x\|_2\|y\|_2}{|y^*Bx|}, \tag{4.10}$$

where y is the left eigenvector associated with λ. This condition number is slightly different from the one given by Stewart and Sun in [65]. They can treat infinite

eigenvalues by stating the eigenvalue problem as $\beta Ax = \alpha Bx$. But this leads to the use of more complex metrics since now the effect of perturbations on a two-dimensional parameter space (α, β) has to be considered. This is done in [65] by using the chordal metric

$$\chi(\langle\alpha, \beta\rangle, \langle\gamma, \delta\rangle) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2 + |\beta|^2}\sqrt{|\gamma|^2 + |\delta|^2}},$$

which leads to the condition number

$$c_{\text{chordal}}((\alpha, \beta)) = \frac{\|x\|_2\|y\|_2}{\sqrt{|\alpha|^2 + |\beta|^2}},$$

where (α, β) is normalized such that $\alpha = y^*Ax$ and $\beta = y^*Bx$. Here we are only interested in small eigenvalues λ. Therefore, the use of the chordal condition number is not necessary and we use the condition number $c_{eig}(\lambda)$ defined in (4.10). Furthermore, we will set $\alpha = \beta = 1$ in (4.10).

In Section 3.5 we showed that Barnett's method can be interpreted as minimizing the smallest eigenvalue $\mu_1(\lambda)$ of

$$A_B(\lambda)^T A_B(\lambda)x(\lambda) = \mu(\lambda)A_I(\lambda)^T A_I(\lambda)x(\lambda), \tag{4.11}$$

which is equivalent to finding the smallest generalized singular value $\sigma_1(\lambda)$ of the pencil $\{A_B(\lambda), A_I(\lambda)\}$ since $\mu_1(\lambda) = \sigma_1^2(\lambda)$. But the condition numbers of the two problems differ significantly. The condition number of the smallest eigenvalue $\mu_1(\lambda)$ of 4.11 is given as

$$c_{\text{eig}}(\mu(\lambda)) = \frac{(1 + \mu_1(\lambda))\|x(\lambda)\|_2^2}{\|A_I(\lambda)x(\lambda)\|_2^2}.$$

This is approximately the square of the condition number

$$c(\sigma_1(\lambda)) = \frac{(1 + \sigma_1(\lambda))\|x(\lambda)\|_2}{\|A_I(\lambda)x(\lambda)\|_2}.$$

for the corresponding generalized singular value $\sigma_1(\lambda)$. Therefore, in terms of numerical stability it is always advisable to use the formulation as a GSVD problem instead of a generalized eigenvalue problem. In Figure 4.6 we show the curve of $\mu_1(\lambda)$ for the GWW-1 isospectral drum computed by using the generalized eigenvalue formulation (4.11). Without rounding errors it should be equivalent to the square of the curve in Figure 4.2. But the curve in 4.6 seems to be completely garbled. Many of the values are negative, although the generalized eigenvalue problem only admits nonnegative

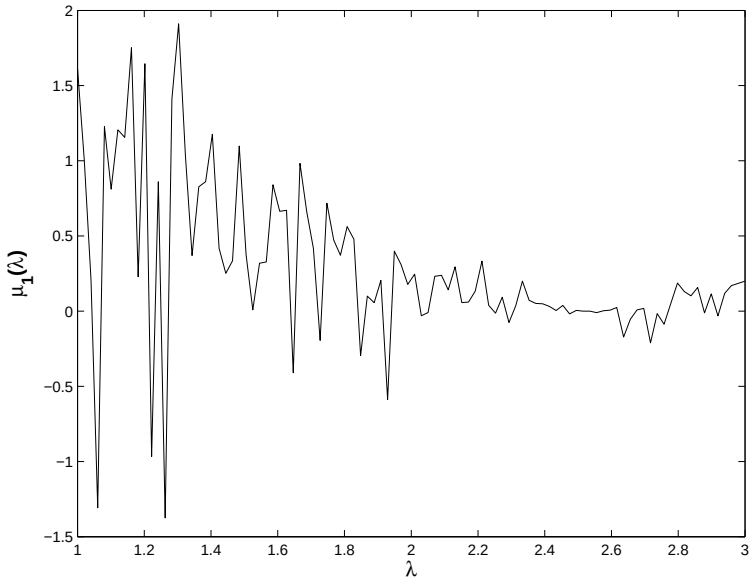


Figure 4.6: The generalized eigenvalue curve for the GWW-1 isospectral drum

eigenvalues. Some of the values returned by Matlab were even complex. These are artefacts of the ill-conditioning present in the eigenvalue formulation.

In [6] Barnett overcomes these ill-conditioning issues by projecting out the nullspace which causes this ill-conditioning. Using our notation this can be done in the following way. Let

$$A_I(\lambda) = U(\lambda)\Sigma(\lambda)V(\lambda)^T$$

be the singular value decomposition of $A_B(\lambda)$. Now define a threshold $\tilde{\epsilon}$ and let $\sigma_1(\lambda) \geq \cdots \geq \sigma_k(\lambda) > \tilde{\epsilon}$, $k \geq 1$ be the singular values of $A_I(\lambda)$ that are larger than $\tilde{\epsilon}$. Partition $V(\lambda)$ as $V(\lambda) = [V_1(\lambda)\ V_2(\lambda)]$ with $V_1(\lambda) = V(:, 1{:}k)$ and $V_2(\lambda) = V(:, k+1{:}\text{end})$. Then the regularized generalized eigenvalue problem is defined as

$$V_1(\lambda)^T A_B(\lambda)^T A_B(\lambda) V_1(\lambda)\tilde{x}(\lambda) = \tilde{\mu}(\lambda) V_1^T A_I(\lambda)^T A_I(\lambda) V_1(\lambda)\tilde{x}(\lambda).$$

A similar strategy was proposed and analyzed by Fix and Heiberger in [24]. The right-hand side matrix now has the singular values $\sigma_1^2 \geq \cdots \geq \sigma_k^2 > \tilde{\epsilon}^2$. Therefore, to remove all numerically zero singular values of $A_I(\lambda)^T A_I(\lambda)$ we need to chose $\tilde{\epsilon} > \sqrt{\epsilon_{mach}}$. In [6] Barnett uses a threshold of $\tilde{\epsilon}^2 = 10^{-14}$, i.e. $\tilde{\epsilon} = 10^{-7}$.

We can apply the same strategy to the GSVD formulation. Then, instead of finding the smallest generalized singular value $\sigma_1(\lambda)$ of the pencil $\{A_B(\lambda), A_I(\lambda)\}$ we find the

smallest generalized singular value $\tilde{\sigma}_1(\lambda)$ of $\{A_B(\lambda)V_1(\lambda), A_I(\lambda)V_1(\lambda)\}$. However, the following strategy to obtain a regularization matrix $V_1(\lambda)$ is more suitable. Let

$$\begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix} = \begin{bmatrix} Q_B(\lambda) \\ Q_I(\lambda) \end{bmatrix} R(\lambda)$$

be the initial QR decomposition which has to be formed in the subspace angle method. Now let

$$R(\lambda) = U_R(\lambda)\Sigma_R(\lambda)V_R(\lambda)^T \tag{4.12}$$

be the SVD of $R(\lambda)$. The regularization matrix $V_1(\lambda)$ is defined as the first k columns of $V_R(\lambda)$ associated with those singular values of $R(\lambda)$, which are above the threshold $\tilde{\epsilon}$. The generalized singular values of $\{A_B(\lambda)V_1(\lambda), A_I(\lambda)V_1(\lambda)\}$ are now obtained from the CS decomposition of the pencil $\{Q_B(\lambda)U_R(:, 1{:}k), Q_I(\lambda)U_R(:, 1{:}k)\}$. The smallest generalized singular value of $\{A_B(\lambda), A_I(\lambda)\}$ is only modestly changed with this strategy if it is not too ill-conditioned. This is shown in the following theorem.

Theorem 4.6.1 *Let $\sigma_1 = s_1/c_1$ be the smallest generalized singular value and x_1 its corresponding right generalized singular vector of the pencil $\{A, B\}$ with $A \in \mathbb{R}^{n\times p}$ and $B \in \mathbb{R}^{m\times p}$. Let the regularization matrix $V_1 \in \mathbb{R}^{p\times k}$ be obtained by the strategy described above and denote by $\tilde{\sigma}_j$ $j = 1, \ldots, k$ the generalized singular values of the pencil $\{AV_1, BV_1\}$. Then*

a) For all generalized singular values $\tilde{\sigma}_j$ of the pencil $\{AV_1, BV_1\}$,

$$\sigma_j \leq \tilde{\sigma}_j$$

b) If $\tilde{\epsilon}\|x_1\|_2 < c_1$, then

$$\sigma_1 \leq \tilde{\sigma}_1 \leq \frac{s_1 + \tilde{\epsilon}\|x_1\|_2}{c_1 - \tilde{\epsilon}\|x_1\|_2}.$$

Proof Let V_2 be the orthogonal complement of V_1, i.e. $V = [V_1\ V_2]$ is an orthogonal matrix. Then $\|AV_2y\|_2 \leq \tilde{\epsilon}\|y\|_2$ and $\|BV_2y\|_2 \leq \tilde{\epsilon}\|y\|_2$ for all $y \in \mathbb{R}^p$ since $\| \begin{bmatrix} AV_2 \\ BV_2 \end{bmatrix} \|_2 \leq \tilde{\epsilon}$. Let $x_1 = V_1y_1 + V_2y_2$. We have

$$\|AV_1y_1\|_2 - \|Ax_1\|_2 \leq \|AV_1y_1 - Ax_1\|_2 = \|AV_2y_2\|_2 \leq \tilde{\epsilon}\|y_2\|_2$$

and

$$\|Bx_1\|_2 - \|BV_1y_1\|_2 \leq \|Bx_1 - BV_1y_1\|_2 = \|BV_2y_2\|_2 \leq \tilde{\epsilon}\|y_2\|_2$$

It follows that

$$\frac{\|AV_1y_1\|_2}{\|BV_1y_1\|_2} \leq \frac{\|Ax_1\|_2 + \tilde{\epsilon}\|y_2\|_2}{\|Bx_1\|_2 - \tilde{\epsilon}\|y_2\|_2}.$$

if $\|Bx_1\|_2 - \tilde{\epsilon}\|y_2\|_2 > 0$. But this holds since $\|Bx_1\|_2 = c_1$, $\|y_2\|_2 \leq \|x_1\|_2$ and therefore

$$\|Bx_1\|_2 - \tilde{\epsilon}\|y_2\|_2 = c_1 - \tilde{\epsilon}\|y_2\|_2 \geq c_1 - \tilde{\epsilon}\|x_1\|_2 > 0.$$

Together with $\|Ax_1\|_2 = s_1$ we find

$$\frac{\|AV_1y_1\|_2}{\|BV_1y_1\|_2} \leq \frac{s_1 + \tilde{\epsilon}\|x_1\|_2}{c_1 - \tilde{\epsilon}\|x_1\|_2}.$$

From the minimax characterization in Theorem 3.4.2 it follows that

$$\tilde{\sigma}_1 \leq \frac{s_1 + \tilde{\epsilon}\|x_1\|_2}{c_1 - \tilde{\epsilon}\|x_1\|_2}.$$

The fact that $\sigma_j \leq \tilde{\sigma}_j$ $j = 1, \ldots, k$ follows immediately from Theorem 3.4.2 since restricting the pencil $\{A, B\}$ to $\{AV_1, BV_1\}$ corresponds to minimizing only over a subset of all possible spaces of dimension j in Theorem 3.4.2. ∎

A similar result for the regularization of the ill-conditioned eigenvalue problem was proved in [24]. If $\sigma_1(\lambda_k) \ll 1$ at an eigenvalue λ_k of (1.1) we obtain for the perturbed generalized singular value $\tilde{\sigma}_1(\lambda_k)$:

$$\tilde{\sigma}_1(\lambda_k) \lessapprox \frac{\sigma_1(\lambda_k) + \tilde{\epsilon}\|x_1(\lambda_k)\|_2}{1 - \tilde{\epsilon}\|x_1(\lambda_k)\|_2} = \sigma_1(\lambda_k) + (1 + \sigma_1(\lambda_k))\tilde{\epsilon}\|x_1(\lambda_k)\|_2 + O((\tilde{\epsilon}\|x(\lambda_k)\|_2)^2),$$

since $s_1(\lambda_k) \approx \sigma(\lambda_k)$ and $c_1(\lambda_k) \approx 1$. The magnification factor $(1+\sigma_1(\lambda_k))\|x(\lambda_k)\|_2$ is essentially the condition number $c(\sigma_1(\lambda_k))$. This can be expected since we ask for the change of a singular value under a small perturbation in the pencil $\{A_B(\lambda_k), A_I(\lambda_k)\}$.

In Figure 4.7 we plot the curves for the sine of the smallest generalized singuar value $\sigma_1(\lambda)$ and for the smallest generalized eigenvalue $\mu_1(\lambda)$ on the GWW-1 isospectral drum in the regularized case. For both plots we use the same regularization matrix V_1 obtained from the QR decomposition of $\begin{bmatrix} A_B(\lambda) \\ A_I(\lambda) \end{bmatrix}$ followed by the SVD of the R factor with a threshold of $\tilde{\epsilon} = 10^{-14}$. The oscillations at the beginning of the subspace angle curve now fully disappear and the generalized eigenvalue curve, although still garbled, now shows a much better behavior than the non-pivoted curve in Figure 4.6. The reason for the better behavior of the generalized singular value curve is

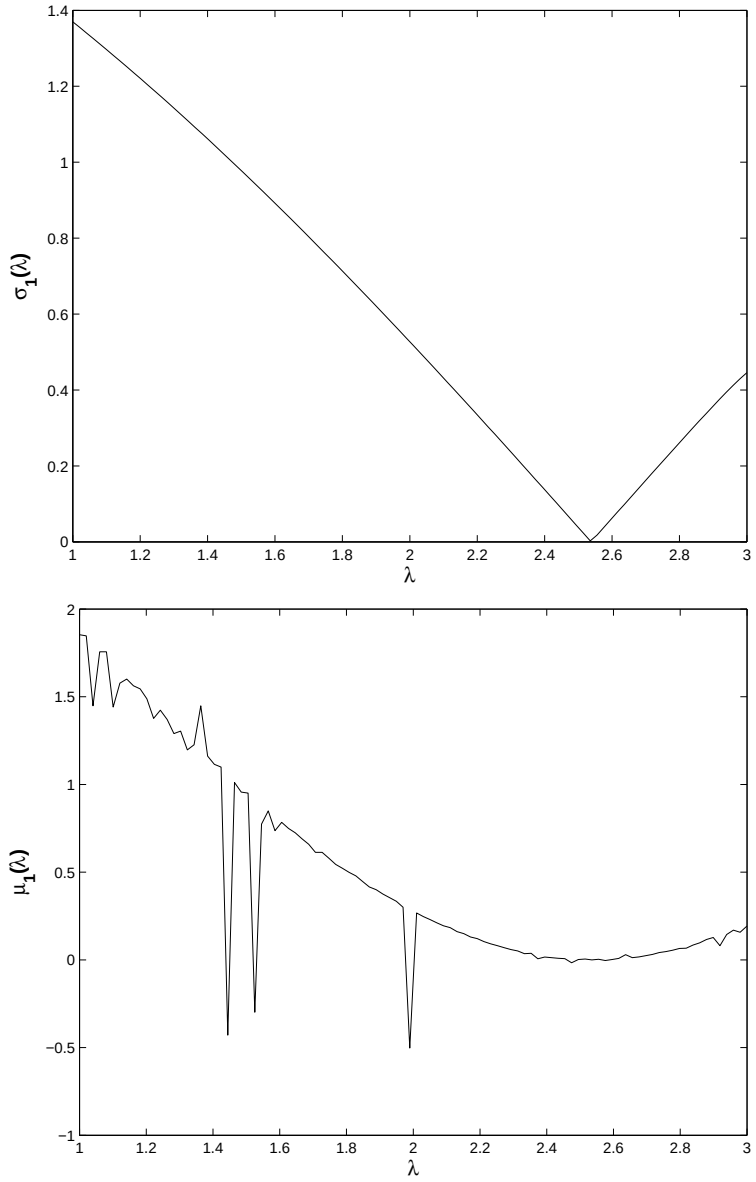


Figure 4.7: The pivoted generalized singular value and generalized eigenvalue curve on the GWW-1 isospectral drum.

that due to the pivoting also the condition number of $A_B(\lambda)V_1(\lambda)$ is now in the order of magnitude of 10^{14} for all $\lambda > 0$ instead of 10^{16} in the non-pivoted case. This gives enough accuracy to remove the oscillations of the subspace angle curve away from the first eigenvalue λ_1 of (1.1) on this region (compare to Figure 4.2). Also the generalized eigenvalue curve improves. But the threshold $\tilde{\epsilon}$ is too small to remove the whole common numerical nullspace of $A_I(\lambda)^T A_I(\lambda)$ and $A_B(\lambda)^T A_B(\lambda)$. To achieve this we would need that $\tilde{\epsilon} > 10^{-8}$. In this example there is hardly any penalty due to the pivoting. The value $\sigma_1(\lambda)$ at the eigenvalue λ_1 grows from 1.7×10^{-11} to 2.1×10^{-11}. The reason is that $\|x(\lambda_1)\|_2 \approx 4.9 \times 10^3$, and therefore the error bound in Theorem 4.6.1 is approximately $\tilde{\epsilon}\|x(\lambda_1)\|_2 \approx 4.9 \times 10^{-11}$.

Chapter 5

A posteriori accuracy bounds

In this chapter we want to answer the following question. Given the subspace angle $\theta(\lambda)$ between the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ how can we bound the relative distance of λ to the next eigenvalue of (1.1). To answer this question we cannot work in spaces sampled at boundary and interior points. Therefore, by $\mathcal{A}(\lambda)$ we always denote the space consisting of all particular solutions in $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ satisfying (1.1a) but not necessarily (1.1*b*), and by $\mathcal{D}_0$ we denote the space of all functions in $\mathcal{C}^2(\Omega) \cap \mathcal{C}(\overline{\Omega})$ which are zero on $\partial\Omega$.

Error bounds for the MPS were derived by Fox, Henrici and Moler in 1967 [25] and these results were simplified and extended by Moler and Payne in 1968 [55]. An excellent overview of error bounds for elliptic eigenvalue problems was written by Still in 1988 [66]. The bounds by Moler and Payne are included as special cases in that paper. In this chapter we first review the bounds by Moler and Payne and extend them to the subspace angle method. Then we apply these bounds to obtain a highly accurate inclusion for the first eigenvalue of the L-shaped region computed with the subspace angle method. In this thesis we are mostly concerned with the eigenvalue problem (1.1), but most of the bounds given in this chapter can be extended to the following more general setting. Let $\mathcal{H}$ be a separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Let T be an operator with domain $D(T)$, such that $D(T)$ is dense in $\mathcal{H}$. Furthermore, let T by symmetric, i.e.

$$\langle u, Tv \rangle = \langle Tu, v \rangle, \quad u, v \in D(T)$$

and let T have a pure point spectrum $\{\lambda_i\}$ and corresponding orthonormal eigenvectors u_i, which are complete in $\mathcal{H}$. By $\hat{T}$ we denote an extension of T to a domain $D(\hat{T})$ such that $\hat{T}u = Tu$ for all $u \in D(T)$ and $D(T) \subset D(\hat{T}) \subset \mathcal{H}$.

For the eigenvalue problem (1.1) the Hilbert space $\mathcal{H}$ is the space of square integrable functions $L^2(\Omega)$ with inner product

$$\langle u, v \rangle_\Omega = \int_\Omega u(x,y)v(x,y)dxdy.$$

For the operator T we have $T = -\Delta$ and

$$D(-\Delta) = \{u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega) : u|_{\partial\Omega} = 0\},$$

which is just the space $\mathcal{D}_0$ used in previous chapters. The extension $D(\hat{T})$ in this case is the space $\mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$

5.1 Accuracy bounds and the subspace angle method

The first error bound for the MPS of which we are aware was proved by Fox, Henrici and Moler[1] [25].

Theorem 5.1.1 (Fox, Henrici, Moler [25]) *Let λ and u be an approximate eigenvalue and eigenfunction of* (1.1) *normalized to $\|u\|_\Omega = 1$ which satisfy the eigenvalue equation* (1.1a) *but not necessarily the zero boundary conditions* (1.1b). *Let*

$$\epsilon = \max_{x \in \partial\Omega} |u(x)|$$

and assume $\epsilon < 1$. Then there exists an eigenvalue λ_k of (1.1) *satisfying*

$$\frac{|\lambda_k - \lambda|}{\lambda} \leq \frac{\sqrt{2}\epsilon + \epsilon^2}{1 - \epsilon^2}. \tag{5.1}$$

The advantage of this theorem is that all information needed for the upper bound in (5.1) can be obtained from computed data. This enabled Fox, Henrici and Moler to give upper and lower bounds for the computed eigenvalues.

[1] They proved the error bound not only for the eigenvalue problem (1.1) but also for eigenvalues of slightly more general elliptic operators.

We do not give the original proof of this theorem here but instead present a proof by Moler and Payne for a very similar bound.

Theorem 5.1.2 (Moler, Payne [55]) *Let T be a symmetric operator with domain $D(T)$ and extension $D(\hat{T})$ as defined above. Let λ be an approximate eigenvalue and $u \in D(\hat{T})$ the corresponding approximate eigenvector of T satisfying $\hat{T}u = \lambda u$. Assume there exists $w \in D(\hat{T})$ with*

$$\hat{T}w = 0 \tag{5.2}$$

and

$$u - w \in D(T). \tag{5.3}$$

Let

$$\epsilon = \frac{\|w\|}{\|u\|},$$

and assume $\epsilon < 1$. Then there exists an eigenvalue λ_k of T satisfying

$$\frac{|\lambda|}{1+\epsilon} \leq |\lambda_k| \leq \frac{|\lambda|}{1-\epsilon}.$$

Proof Let

$$a_n = \langle u, u_n \rangle, \quad b_n = \langle w, u_n \rangle,$$

where u_n is the normalized eigenfunction associated with the eigenvalue λ_n. We have

$$\langle u - w, Tu_n \rangle = \langle T(u-w), u_n \rangle = \langle \hat{T}u, u_n \rangle - \langle \hat{T}w, u_n \rangle.$$

and therefore

$$\lambda_n(a_n - b_n) = \lambda a_n$$

or equivalently

$$\frac{\lambda_n - \lambda}{\lambda_n} a_n = b_n.$$

Choose λ_k such that

$$\frac{|\lambda_k - \lambda|}{|\lambda_k|} = \min_n \frac{|\lambda_n - \lambda|}{|\lambda_n|}.$$

For this k it holds that

$$\frac{|\lambda_k - \lambda|}{|\lambda_k|} |a_n| \leq |b_n|$$

for all n. We obtain

$$\frac{|\lambda_k - \lambda|^2}{|\lambda_k|^2} \sum_{n=1}^{\infty} a_n^2 \leq \sum_{n=1}^{\infty} b_n^2.$$

Since

$$\epsilon^2 = \frac{\sum_{n=1}^{\infty} b_n^2}{\sum_{n=1}^{\infty} a_n^2}$$

the proof follows. ∎

To evaluate the error bound in Theorem 5.1.2 the norm of w has to be determined. In the case $T = -\Delta$ this means that a boundary value problem has to be solved. But since the boundary values of w are known $\|w\|_\Omega$ can be easily estimated. From the maximum principle for harmonic functions it follows that

$$\|w\|_\Omega \leq |\Omega| \max_{x \in \overline{\Omega}} |w(x)| = |\Omega| \max_{x \in \partial\Omega} |w(x)| = |\Omega| \max_{x \in \partial\Omega} |u(x)|,$$

where $|\Omega| = \int_\Omega 1 dxdy$ is the area of Ω. We obtain

$$\frac{|\lambda_k - \lambda|}{\lambda_k} \leq \frac{|\Omega| \max_{x \in \partial\Omega} |u(x)|}{\|u\|_\Omega}. \tag{5.4}$$

Another possibility is to bound $\|w\|_\Omega$ using eigenvalues of a Stekloff eigenvalue problem [66, 43], which is defined as

$$\begin{aligned} \Delta^2 u &= 0 \text{ in } \Omega \\ u &= \Delta u - q \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega. \end{aligned} \tag{5.5}$$

It can be shown (see for example [42]) that the smallest eigenvalue q_1 of (5.5) is characterized by

$$q_1 = \min_{\Delta h = 0 \text{ in } \Omega} \frac{\int_{\partial\Omega} h^2 ds}{\int_\Omega h^2 dx}.$$

It follows that

$$\|h\|_\Omega \leq q_1^{-\frac{1}{2}} \|h\|_{\partial\Omega}$$

for all functions h satisfying $\Delta h = 0$ in Ω. This immediately leads to the following bound of ϵ in Theorem 5.1.2 if $T = -\Delta$:

$$\frac{|\lambda_k - \lambda|}{\lambda_k} \leq \epsilon \leq q_1^{-\frac{1}{2}} \frac{\|u\|_{\partial\Omega}}{\|u\|_\Omega} \tag{5.6}$$

As a corollary we obtain a bound on the relative error in terms of the principal angle $\theta(\lambda)$ between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$.

Corollary 5.1.3 *For the angle $\theta(\lambda)$ between the spaces $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ it holds that*

$$\frac{|\lambda_k - \lambda|}{\lambda_k} \leq q_1^{-1/2} \tan\theta(\lambda).$$

Proof Denote by $|||u||| = \sqrt{\|u\|_\Omega^2 + \|u\|_{\partial\Omega}^2}$ the mixed norm induced by the inner product $\langle \cdot, \cdot \rangle$ defined in (3.2). For every $u \in \mathcal{A}(\lambda)$ we have

$$\sup_{\substack{v \in \mathcal{D}_0 \\ |||v|||=1}} \langle u, v \rangle = \|u\|_\Omega, \tag{5.7}$$

since from $\langle u, v \rangle = \int_\Omega u(x,y)v(x,y)dxdy$ for every $v \in \mathcal{D}_0$ and the Cauchy-Schwarz inequality it follows that $\langle u, v \rangle \leq \|u\|_\Omega$ for every $v \in \mathcal{D}_0$ with $|||v||| = 1$. Equality in (5.7) follows from the fact that u can be expanded in Ω in terms of the eigenfunctions $u_k \in \mathcal{D}_0$ of (1.1). Combining (3.1) and (5.7), we get

$$\cos\theta(\lambda) = \sup_{\substack{u \in \mathcal{A}(\lambda) \\ |||u|||=1}} \|u\|_\Omega. \tag{5.8}$$

It follows that

$$\tan\theta(\lambda) = \inf_{\substack{u \in \mathcal{A}(\lambda) \\ |||u|||=1}} \frac{\sqrt{1 - \|u\|_\Omega^2}}{\|u\|_\Omega} = \inf_{u \in \mathcal{A}(\lambda)} \frac{\|u\|_{\partial\Omega}}{\|u\|_\Omega}. \tag{5.9}$$

Since for every $u \in \mathcal{A}(\lambda)$ (5.6) holds it follows from (5.9) that

$$\frac{|\lambda_k - \lambda^*|}{\lambda_k} \leq q_1^{-\frac{1}{2}} \tan\theta(\lambda). \quad \blacksquare$$

Hence, the subspace angle is a measure for the optimal error bound which is possible by approximating from $\mathcal{A}(\lambda)$, while Theorem 5.1.2 only uses one element $u \in \mathcal{A}(\lambda)$ to obtain a bound on the relative eigenvalue error. In applications $\mathcal{A}(\lambda)$ is not the space of all particular solutions but the span of a finite number of particular solutions. Then Corollary 5.1.3 is still valid. But usually at an eigenvalue λ_k we will have $\tan\theta(\lambda_k) > 0$, i.e., the bound on the relative error can be larger than zero at an eigenvalue. Similar bounds for the eigenvectors are also possible. Moler and Payne established the following theorem.

Theorem 5.1.4 (Moler, Payne [55]) *Using the hypotheses and notation of Theorem 5.1.2, assume in addition that* $\|u\|_\Omega = 1$. *Let*

$$\alpha = \min_{\lambda_n \neq \lambda_k} \frac{|\lambda_n - \lambda|}{|\lambda_n|},$$

and let u_k *be the normalized projection of* u *onto the eigenspace of* λ_k. *Then*

$$\|u - u_k\|_\Omega \leq \frac{\epsilon}{\alpha}(1 + \frac{\epsilon^2}{\alpha^2})^{\frac{1}{2}}.$$

If we choose for ϵ the tangent of $\theta(\lambda)$ we obtain the following corollary.

Corollary 5.1.5 *We use the notation of Corollary 5.1.3 and Theorem 5.1.4. Then for every* $\delta > 0$ *and* $\tilde{\epsilon} = q_1^{-\frac{1}{2}} \tan\theta(\lambda)$ *there exists a function* $u \in \mathcal{A}(\lambda)$, $\|u\|_\Omega = 1$ *such that*

$$\|u - u_k\|_\Omega \leq \frac{\tilde{\epsilon}}{\alpha}(1 + \frac{\tilde{\epsilon}^2}{\alpha^2})^{\frac{1}{2}} + \delta.$$

Proof The proof follows by choosing a function $u \in \mathcal{A}(\lambda)$ that comes sufficiently close to the infimum in (5.9). ∎

Further results for the case in which the approximate eigenfunction u satisfies neither $\hat{T}u = \lambda u$ nor $u \in D(T)$ are given in [43] and [66]. We finish this section with a very interesting result by Still [66]. The idea is the following. If an approximate eigenfunction $\tilde{u}$ satisfies the zero boundary conditions (1.1b) but not necessarily the eigenvalue equation (1.1a) the Rayleigh quotient

$$\rho(\tilde{u}) := \frac{\langle \tilde{u}, -\Delta\tilde{u}\rangle_\Omega}{\langle \tilde{u}, \tilde{u}\rangle_\Omega}$$

is a quadratically good approximation to an eigenvalue in the sense that if the distance of $\tilde{u}$ to a normalized eigenvector u_k is $O(\epsilon)$, then the distance of $\rho(\tilde{u})$ to λ_k is $O(\epsilon^2)$. Unfortunately, the functions $u \in \mathcal{A}(\lambda)$ do not necessarily satisfy the zero boundary conditions. But if we define w as in Theorem 5.1.2 as the harmonic function with the same boundary data as $u \in \mathcal{A}(\lambda)$, we can apply the Rayleigh quotient to $u - w$ and obtain

$$\rho(u - w) = \frac{\langle u - w, -\Delta(u - w)\rangle_\Omega}{\langle u - w, u - w\rangle_\Omega} = \lambda\left(1 + \frac{\langle w, u - w\rangle_\Omega}{\langle u - w, u - w\rangle_\Omega}\right).$$

From the properties of the Rayleigh quotient we can hope that

$$\lambda\left(1+\frac{\langle w, u-w\rangle_\Omega}{\langle u-w, u-w\rangle_\Omega}\right)$$

is a quadratically good approximation to an eigenvalue. This is made precise in the following theorem by Still, which also includes the case that the approximate eigenfunction u does not necessarily satisfy $-\Delta u = \lambda u$.

Theorem 5.1.6 (Still [66]) *Given $u \in D(\hat{T})\backslash D(T)$, $\|u\| = 1$, $\lambda \in \mathbb{R}$, define the function r by*

$$r = \hat{T}u - \lambda u.$$

Let $d_+(\rho)$ and $d_-(\rho)$ be defined as

$$d_+(\rho) = \min_{\lambda_\nu > \lambda_k} |\lambda_\nu - \rho|, \quad d_-(\rho) = \min_{\lambda_\nu < \lambda_k} |\lambda_\nu - \rho|,$$

for a given eigenvalue λ_k of T.

a) For a solution $w \in D(\hat{T})$ of

$$\hat{T}w = 0, \quad u - w \in D(T),$$

let

$$v_1 := u - w, \quad \epsilon_1 := \frac{\|\lambda w + r\|_\Omega}{\|u - w\|_\Omega}.$$

Then with the Rayleigh quotient $\rho_1 = \rho(v_1)$ given by

$$\rho_1 = \lambda + \frac{\langle \lambda w + r, v_1\rangle}{\langle v_1, v_1\rangle},$$

the inequality

$$-\frac{(2\epsilon_1)^2}{d_+(\rho_1)} \leq \lambda_k - \rho_1 \leq \frac{(2\epsilon_1)^2}{d_-(\rho_1)}$$

holds for some eigenvalue λ_k of T.

b) For a solution $R \in D(\hat{T})$ of

$$\hat{T}R = \hat{T}u - \lambda u, \quad u - R \in D(T)$$

let

$$v_2 := u - R, \quad \epsilon_2 := |\lambda| \frac{\|R\|}{\|u - R\|}.$$

Then with $\rho_2 = \rho(v_2)$ given by

$$\rho_2 = \lambda + \frac{\langle \lambda R, v_2 \rangle}{\langle v_2, v_2 \rangle},$$

the inequality

$$-\frac{(2\epsilon_2)^2}{d_+(\rho_2)} \leq \lambda_k - \rho_2 \leq \frac{(2\epsilon_2)^2}{d_-(\rho_2)}$$

is valid.

Since for the MPS it holds that $\hat{T}u = \lambda u$ and therefore $r = 0$ part b) of Theorem 5.1.6 gives no more information than part a). In practice, the correction term

$$\frac{\langle \lambda w + r, v_1 \rangle}{\langle v_1, v_1 \rangle}$$

usually cannot be easily computed. But nevertheless, the result is interesting, since it shows that with a small correction to λ a quadratic accuracy is possible.

5.2 Verifying 13 digits of the first eigenvalue on the L-shaped region

In this section we use the bound by Moler and Payne to verify the first eigenvalue of the L-shaped region to 13 rounded digits of accuracy and compare it to approximate bounds obtained from the computed subspace angle. The starting point is the subspace angle method. We discretize the boundary of the L-shaped region with 500 Chebyshev distributed points on each side not adjacent to the reentrant corner. In addition 50 interior points are randomly chosen. Using a Chebyshev distribution on the boundary has the effect that near the corners of the region the absolute value of the approximate eigenfunction stays smaller than with equally distributed points. The eigenfunction is approximated with a basis of $N = 80$ Fourier-Bessel terms around the reentrant corner. The matrix $A(\lambda)$ of particular solutions evaluated at boundary and interior points is normalized such that $\|A(\lambda)(:,k)\|_2 = 1$ for $k = 1, \ldots, N$. The subspace angle method performs a QR factorization of $A(\lambda)$ and computes the smallest singular value of the first part $Q_B(\lambda)$ of $Q(\lambda)$ corresponding to the boundary points. We denote the corresponding singular vector of Q_B by v. As eigenvalue estimate

we use the value $\lambda = 9.6397238440219$ which was obtained from the computation underlying Figure 3.4.

The resulting approximate eigenfunction evaluated at all points is given as $u = Q(\lambda)v$. But we need the coefficient vector of u in the original basis $A(\lambda)$. Therefore, we have to solve

$$A(\lambda)c = Q(\lambda)v.$$

Since $A(\lambda) = Q(\lambda)R(\lambda)$ this is equivalent to the system of equations $R(\lambda)c = v$. Due to ill-conditioning the error between the computed vector $\tilde{c}$ and the true vector c might be large. Nevertheless, the residual $\|R(\lambda)\tilde{c} - v\|_2$ will be small. The reason is that by ignoring that the computed Q is not exactly orthogonal the vector $\tilde{c}$ is the exact right generalized singular vector of a small perturbation of the pencil $\{A_B(\lambda), A_I(\lambda)\}$. By Lemma 4.3.2 and the discussion after Theorem 4.5.1 we can then expect $\|\tilde{c}\|_2 = O(1)$ if λ is close to an eigenvalue of (1.1). Since

$$\|R(\lambda)\tilde{c} - v\|_2 = \|\Delta R\tilde{c}\|_2 \leq \|\Delta R\|_2\|\tilde{c}\|_2,$$

where ΔR is the backward error of solving $R(\lambda)c = v$ it follows that $\|R(\lambda)\tilde{c} - v\|_2$ is small.

Indeed, in our case we have

$$\|R(\lambda)\tilde{c} - v\|_2 \approx 1.11 \times 10^{-16}$$

and

$$\|A(\lambda)\tilde{c} - Q(\lambda)v\|_2 \approx 2.19 \times 10^{-15}.$$

Therefore, the coefficient vector $\tilde{c}$ defines an approximate eigenfunction that is small at the boundary points ($\|A_B(\lambda)\tilde{c}\|_2 \approx 2.48 \times 10^{-14}$) and large at the interior points ($\|A_I(\lambda)\tilde{c}\|_2 \approx 1$). To apply the error bound from (5.4) we need to estimate the expressions $\max_{x\in\partial\Omega}|u(x)|$ and $\|u\|_\Omega$, where u is now the approximate eigenfunction defined by the coefficient vector $\tilde{c}$. The L^2-norm $\|u\|_\Omega$ can easily be estimated with a trick already used in [25]. Let G be the circular sector around the reentrant corner

with radius 1 and angle $\frac{3}{2}\pi$ that fits into the L-shaped region. Then

$$\begin{aligned}
\|u\|_\Omega^2 &\geq \|u\|_G^2 \\
&= \int_G (u(x,y)^2) d(x,y) \\
&= \sum_{k,j=1}^{N} \hat{c}_k \hat{c}_j \int_0^1 J_{\frac{2}{3}k}(\sqrt{\lambda}r) J_{\frac{2}{3}j}(\sqrt{\lambda}r) r dr \int_0^{\frac{3}{2}\pi} \sin\frac{2}{3}k\theta \sin\frac{2}{3}j\theta d\theta \\
&= \frac{3}{4}\pi \sum_{j=1}^{N} \hat{c}_j^2 \int_0^1 J_{\frac{2}{3}k}(\sqrt{\lambda}r)^2 r dr, \qquad (5.10)
\end{aligned}$$

where $\hat{c}_k$ is the kth coefficient of u in the non-scaled Fourier-Bessel basis (for the actual computations we use a scaled basis). The last integral has an analytic expression in terms of Fourier-Bessel functions. But for convenience we just evaluate it to high accuracy using the Matlab `quadl` function. In addition we use only the integral of the lowest order Bessel term in the above sum. We find

$$\|u\|_\Omega \geq |\hat{c}_1| \sqrt{\frac{3}{4}\pi \int_0^1 J_{\frac{2}{3}k}(\sqrt{\lambda}r)^2 r dr}. \qquad (5.11)$$

Figure 5.1 shows $|u(x)|$ on the boundary collocation points. The plot is scaled with the area of the L-Shaped region and the lower bound for $\|u\|_\Omega$ from (5.11). Hence, the maximum of the curve is an upper bound for the error in (5.4).

The computed curve shows oscillations around the true function values due to rounding error effects. Since these oscillations also lead to values that are larger than the true values we still obtain a good upper bound. With the upper bound of the scaled $|u(x)|$ of 1.5×10^{-14} we obtain the inclusion from (5.4) that

$$9.639723844021754 \leq \lambda_1 \leq 9.639723844022043$$

for the true first eigenvalue λ_1 of the L-shaped region. This gives 13 rounded digits correctly. Indeed, from Figure 3.4 we believe that the true value is 9.6397238440219 to 14 digits of accuracy.

How does this bound compare to the subspace angle estimate in Theorem 5.1.3 ? We know neither the constant q_1 nor the exact subspace angle $\theta(\lambda)$. However, if we assume that $\tan\theta(\lambda) \approx \sigma_1(\lambda)$, where $\sigma_1(\lambda)$ is the smallest generalized singular value of the pencil $\{A_B(\lambda), A_I(\lambda)\}$ we can use (5.1.3) for an approximate estimate if q_1 is not too small. In our case we obtain $\sigma_1(\lambda) \approx 2.48 \times 10^{-14}$ leading to a similar error

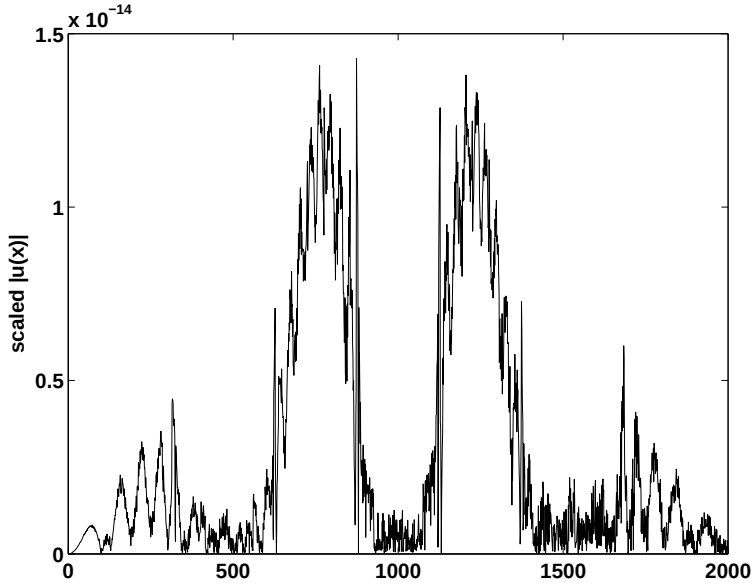


Figure 5.1: Numerically computed values of $|u(x)|$ on the boundary collocation points after scaling by the square root of the area and the estimated value of $\|u\|_\Omega$. Rounding errors lead to oscillations around the true function values.

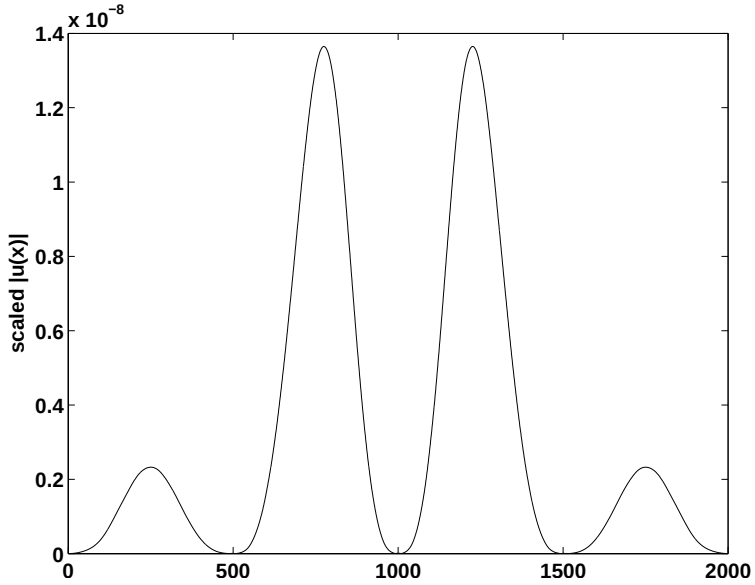


Figure 5.2: The same plot as in Figure 5.1 but now for the value $\lambda = 9.6397238$, which is the first 8 digits of the first eigenvalue of the L-shaped region.

estimate as in (5.4). If we use the approximate eigenvalue $\lambda = 9.6397238$, which is correct to 8 digits we obtain $\sigma_1(\lambda) \approx 2.8 \times 10^{-8}$. Indeed, the plot in Figure 5.2 shows that the relative error is at most 1.5×10^{-8}. This demonstrates that usually it is not necessary to apply the error bound in (5.4) to obtain a good error estimate. A good rule of thumb is that the relative error of the eigenvalue approximation is smaller than or equal to the computed smallest generalized singular value $\sigma_1(\lambda)$.

Chapter 6

Convergence rates via complex approximation theory

While the previous chapters were concerned with the stable computation of eigenvalues and the derivation of accuracy bounds we will now discuss what the approximation properties of the basis are. In the 1940's Vekua discovered close relationships between solutions of elliptic partial differential equations and holomorphic functions in the complex plane [80]. This work was one of the motivations for the Method of Particular Solutions by Fox, Henrici and Moler. A very good survey of this theory was written by Henrici in 1957 [36] and we will review some of the results of his paper to give a short introduction to Vekua's theory. Closely related is the question of analytic continuation of solutions of elliptic PDE's. Classical papers on this subject were written in the 1950's by Garabedian [28] and Lewy [48] and we review Garabedian's results in the special case of the eigenvalue problem (1.1). We will then show how to connect Vekua's results with classical results from complex approximation to obtain convergence estimates for the Method of Particular Solutions. Using analytic continuation and conformal mappings we derive bounds for the exponential convergence of the MPS on regions with at most one corner singularity. For regions with multiple corner singularities we review algebraic convergence rates which go back to Eisenstat in 1974 [23] and were later improved by Still in the 1980's [67, 69]. Based on these results we show how to obtain increasing algebraic convergence rates for regions with multiple singularities.

6.1 An introduction to Vekua's theory

In this section we will closely follow Henrici's beautiful presentation of Vekua's theory in [36]. Before we start let us briefly review what is meant by a real analytic function and by a complex analytic (holomorphic) function and how these two are related[1].

Let u be a real function of the two variables x and y in a region Ω. Throughout this chapter we assume that Ω is bounded and simply connected. Furthermore, the boundary $\partial\Omega$ is assumed to be a piecewise analytic Jordan curve. The function u is called real analytic at a point $(x_0, y_0) \in \Omega$ if in a neighborhood of this point it can be represented as a Taylor series of the form

$$u(x,y) = \sum_{n,m\geq 0} c_{n,m}(x-x_0)^n(y-y_0)^m,$$

where

$$c_{nm} = \frac{1}{n!m!}\frac{\partial^{n+m}u(x_0,y_0)}{\partial x^n \partial y^m}.$$

A complex function ϕ of the complex variable $z = x + iy$ is holomorphic at z_0 if it is complex differentiable at z_0. This is equivalent to having a Taylor series expansion of the form

$$\phi(z) = \sum_{k=0}^{\infty} c_k(z-z_0)^k.$$

in a neighborhood of z_0. The coefficients c_k are given as

$$c_k = \frac{1}{k!}\frac{d}{dz}\phi(z_0).$$

If ϕ depends on several complex variables $z_1, \ldots, z_n \in \mathbb{C}^n$ it is called holomorphic if it is holomorphic in each of its complex variables. This is equivalent to the property that ϕ has a convergent Taylor series in the N complex variables $z_1, \ldots, z_n$ (see for example Chapter 2 of [72]).

Let u be real analytic at a point (x_0, y_0). Then u can be extended to a holomorphic function in two complex variables by allowing x and y to take complex values in a small neighborhood $S \subset \mathbb{C}^2$ around (x_0, y_0). This follows from the absolute convergence of the Taylor series of u [49]. We can write the holomorphic continuation in the following

[1] Often the term analytic is used for holomorphic functions. To distinguish between real analytic and complex analytic functions we will always use the term holomorphic instead of complex analytic.

way. Let $z = x + iy$ and $z^* = x - iy$. We have $\overline{z} = z^*$ if and only if x and y are real. Define

$$U(z, z^*) = u\left(\frac{z + z^*}{2}, \frac{z - z^*}{2i}\right).$$

It follows that $U(z, \overline{z}) = u(x, y)$. If we let z and z^* vary independently around $z_0 = x_0 + iy_0$ and $\overline{z_0}$, there exists a neighborhood $S(z_0) \subset \mathbb{C}$ of z_0 such that U is holomorphic in the region $[S(z_0), S^*(z_0)] \subset \mathbb{C}^2$, where $S^*(z_0) = \{\overline{z} : z \in S(z_0)\}$ is the complex conjugate region of $S(z_0)$. For an arbitrary real analytic function this holomorphic continuation is only possible in a small neighborhood $S(z_0)$ around z_0. Vekua's theory asserts that for solutions of certain elliptic PDEs this continuation is not only possible in the small but in the large, i.e. if u is analytic in Ω then U is holomorphic in $[\Omega, \Omega^*]$. To state the results of Vekua Henrici defines three classes of functions:

I This class contains all functions which are twice continuously differentiable in Ω.

II This class contains all functions f which are real analytic in Ω.

III This class consists of all functions u of Class II which possess a holomorphic extension U into the region $[\Omega, \Omega^*] \in \mathbb{C}^2$. Hence, for every point $z_0 \in \Omega$ there exists a neighborhood $S(z_0)$ such that $\Omega \subset S(z_0)$ and U is holomorphic in $[S(z_0), S^*(z_0)]$.

A simple example given by Henrici is the class of harmonic functions. Let u be harmonic in Ω. Then u is real analytic, i.e. in Class II and it is well known that there exists a function ϕ holomorphic in Ω such that

$$u(x, y) = \mathrm{Re}\{\phi(z)\}, \quad z \in \Omega.$$

Define the function $\bar{\phi}(z) := \overline{\phi(\overline{z})}$, which is holomorphic for $z \in \Omega^*$. Now let

$$U(z, z^*) = \frac{1}{2}[\phi(z) + \overline{\phi(\overline{z^*})}].$$

Then U is holomorphic in $[\Omega, \Omega^*]$ and it holds that

$$U(z, \overline{z}) = \frac{1}{2}[\phi(z) + \overline{\phi(z)}] = u(x, y).$$

Therefore, U is the unique holomorphic extension of u into $[\Omega, \Omega^*]$. Let now the linear elliptic partial differential equation $\mathcal{L}u$ be defined as

$$\mathcal{L}u(x,y) = \Delta u(x,y) + a(x,y)\frac{\partial u}{\partial x}(x,y) + b(x,y)\frac{\partial u}{\partial y}(x,y) + c(x,y)u(x,y) = 0. \quad (6.1)$$

The following theorem is a classical result of the theory of elliptic PDEs [29].

Theorem 6.1.1 *If the coefficient functions a, b, c are in Class II then every solution u of $\mathcal{L}u = 0$ in Class I is also in Class II.*

Vekua proved the following even stronger result.

Theorem 6.1.2 (Vekua) *If the coefficient functions a, b, c are in Class III, then every solution u of $\mathcal{L}u = 0$ in Class I is also in Class III.*

For Theorem 6.1.2 it is essential that Ω is simply connected. Consider the following example. The harmonic function $u(x,y) = \log(x^2+y^2)$ is harmonic in every annulus A surrounding the origin, but the holomorphic extension $U(z,z^*) = \log z + \log z^*$ is not holomorphic in $[A, A^*]$.

If $\mathcal{L} = \Delta$ then any holomorphic function $\phi(z)$ defines a solution of $\mathcal{L}u = 0$ via $u(x,y) = \mathrm{Re}\{\phi(z)\}$. Vekua showed that this is just a special case of a more general relationship between holomorphic functions and solutions of elliptic PDEs with coefficient functions belonging to Class III. Define

$$\begin{aligned} A(z,z^*) &= \frac{1}{4}\left\{a\left(\frac{z+z^*}{2}, \frac{z-z^*}{2i}\right) + ib\left(\frac{z+z^*}{2}, \frac{z-z^*}{2i}\right)\right\}, \\ B(z,z^*) &= \frac{1}{4}\left\{a\left(\frac{z+z^*}{2}, \frac{z-z^*}{2i}\right) - ib\left(\frac{z+z^*}{2}, \frac{z-z^*}{2i}\right)\right\}, \\ C(z,z^*) &= \frac{1}{4}c\left(\frac{z+z^*}{2}, \frac{z-z^*}{2i}\right). \end{aligned}$$

Let ϕ be any holomorphic function in Ω and fix $z_0 \in \Omega$. Define the integral operator

$$\begin{aligned} I[\phi; z_0](z,z^*) &= \frac{1}{2}\left\{G(z,\overline{z_0},z,z^*)\phi(z) + \int_{z_0}^{z} \phi(t)H(t,\overline{z_0},z,z^*)dt + \right. \\ &\qquad \left. G(z_0,z^*,z,z^*)\bar{\phi}(z^*) + \int_{\overline{z_0}}^{z^*} \bar{\phi}(t^*)H^*(z_0,t^*,z,z^*)dt^*\right\} \quad (6.2) \end{aligned}$$

with

$$\begin{aligned} H(t,t^*,z,z^*) &= B(t,t^*)G(t,t^*,z,z^*) - \frac{\partial G}{\partial t}(t,t^*,z,z^*), \\ H^*(t,t^*z,z^*) &= A(t,t^*)G(t,t^*,z,z^*) - \frac{\partial G}{\partial t^*}(t,t^*,z,z^*). \end{aligned}$$

For $z^* = \overline{z}$ these equations simplify to

$$\begin{aligned} I[\phi;z_0](z,\overline{z}) &= \mathrm{Re}\left\{G(z,\overline{z_0},z,\overline{z})\phi(z) + \int_{z_0}^{z} \phi(t)H(t,\overline{z_0},z,\bar{z})dt\right\} \\ &:= \mathrm{Re}\{V[\phi;z_0](z,\overline{z})\}. \end{aligned}$$

The function $G(t,t^*,z,z^*)$ is the complex Riemann function for $\mathcal{L}$. We will not go into further detail about its definition here. A detailed description is given in [36]. For some equations this function is explicitly known. If $\mathcal{L} = \Delta$, then $G(t,t^*,z,z^*) = 1$, and for the Helmholtz operator $\mathcal{L} = \Delta + \lambda$ it is given as

$$G(t,t^*,z,z^*) = J_0(\sqrt{\lambda}\sqrt{(z-t)(z^*-t^*)}).$$

We can now establish a $1-1$ relationship between holomorphic functions and solutions of elliptic PDEs with coefficient functions in Class III.

Theorem 6.1.3 (Vekua) *Fix $z_0 \in \Omega$. Then there exists a unique function ϕ holomorphic in Ω with $\phi(z_0)$ real such that*

$$\begin{aligned} u(x,y) &= \mathrm{Re}\{\mathrm{V}[\phi;\mathrm{z_0}]\}(\mathrm{z},\overline{\mathrm{z}}), \quad \mathrm{z} = \mathrm{x} + \mathrm{iy} \in \Omega \\ U(z,z^*) &= I[\phi;z_0](z,z^*), \quad (z,z^*) \in [\Omega,\Omega^*]. \end{aligned}$$

Moreover,

$$\phi(z) = 2U(z,\overline{z_0}) - U(z_0,\overline{z_0})G(z_0,\overline{z_0},z,\overline{z_0}). \tag{6.3}$$

An equivalent integral representation, which does not depend on the complex Riemann function but can be approximated directly from the coefficients of the elliptic equation, was developed by Bergman [11]. In [64] Schryer uses it to construct particular solutions for elliptic PDEs with polynomial coefficient functions.

From now on we will say that ϕ is associated with u or u is associated with ϕ if $u = \mathrm{Re}\{V[\phi;z_0]\}$. Let us apply this theorem to two examples. First let $\mathcal{L} = \Delta$. Then V reduces to the identity operator and we obtain

$$u(x,y) = \mathrm{Re}\{\phi(z)\}.$$

Furthermore, from (6.2) we find

$$U(z,z^*) = I[\phi; z_0](z,z^*) = \frac{1}{2}\left\{\phi(z) + \bar{\phi}(z^*)\right\}.$$

These are just the results we derived earlier for $\mathcal{L} = \Delta$. Now let $\mathcal{L} = \Delta + \lambda$ and let Ω be the wedge with interior angle π/α (see Figure 2.1). An eigenfunction of this region is given as

$$u(r,\theta) = J_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta, \quad k \in \mathbb{N},\ \lambda > 0.$$

We want to derive the holomorphic function ϕ associated with u. Let $z = re^{i\theta}$. Then

$$u(r,\theta) = J_{\alpha k}(\sqrt{\lambda}\sqrt{z\overline{z}})\left(\left(\sqrt{\frac{z}{\overline{z}}}\right)^{\alpha k} - \left(\sqrt{\frac{\overline{z}}{z}}\right)^{\alpha k}\right)/(2i).$$

From Theorem 6.1.2 it follows that the unique analytic continuation of $u(x,y)$ into the region $[\Omega, \Omega^*] \subset \mathbb{C}^2$ is given by

$$U(z,z^*) = J_{\alpha k}(\sqrt{\lambda}\sqrt{zz^*})\left(\left(\sqrt{\frac{z}{z^*}}\right)^{\alpha k} - \left(\sqrt{\frac{z^*}{z}}\right)^{\alpha k}\right)/(2i) \tag{6.4}$$

since $U(z,z^*)$ is holomorphic in $[\Omega,\Omega^*]$ and $u(x,y) = U(z,\overline{z})$. Bessel functions can be expressed by power series as

$$J_\nu(x) = x^\nu \sum_{n=0}^{\infty} a_n(-\frac{x^2}{4})^n$$

for certain parameters a_n.[2] Therefore, (6.4) becomes

$$U(z,z^*) = \frac{1}{2i}\left(z^{\alpha k} - (z^*)^{\alpha k}\right)\sqrt{\lambda}^{\alpha k}\sum_{n=0}^{\infty} a_n(-\frac{1}{4})^n(\lambda z z^*)^n. \tag{6.5}$$

From Theorem 6.1.3 it follows that

$$\phi(z) = 2U(z,\overline{z_0}) - U(z_0,\overline{z_0})G(z_0,\overline{z_0},z,\overline{z_0}).$$

By a continuity argument we can fix $z_0 = 0$ and obtain together with (6.5)

$$\phi(z) = \frac{1}{i}\sqrt{\lambda}^{\alpha k} a_0 z^{\alpha k}.$$

For $u(x,y) = J_{\alpha k}(\sqrt{\lambda}r)\cos\alpha k\theta$ we obtain similarly $\phi(z) = \sqrt{\lambda}^{\alpha k} a_0 z^{\alpha k}$.

Let us summarize these results.

[2] The a_n are defined by the recurrence relation $a_0 = \frac{1}{2^\nu\Gamma(\nu+1)}$, $a_n = a_{n-1}/(n(n+\nu))$.

Lemma 6.1.4 *Let Ω be a wedge with interior angle π/α and let $\mathcal{L} = \Delta + \lambda$. Then*

$$\begin{aligned} -\frac{1}{a_0\sqrt{\lambda}^{\alpha k}} J_{\alpha k}(\sqrt{\lambda}r)\sin\alpha k\theta &= \mathrm{Re}\{\mathrm{V}[\mathrm{iz}^{\alpha\mathrm{k}},0]\}(\mathrm{z},\overline{\mathrm{z}}), \\ \frac{1}{a_0\sqrt{\lambda}^{\alpha k}} J_{\alpha k}(\sqrt{\lambda}r)\cos\alpha k\theta &= \mathrm{Re}\{\mathrm{V}[\mathrm{z}^{\alpha\mathrm{k}},0]\}(\mathrm{z},\overline{\mathrm{z}}). \end{aligned}$$

The functions defined by $\mathrm{Re}\{V[iz^{\alpha k},0]\}$ and $\mathrm{Re}\{V[z^{\alpha k},0]\}(z,\overline{z})$ are sometimes called generalized harmonic polynomials (see for example [52]) since for $\mathcal{L} = \Delta$ they lead to the harmonic polynomials $r^{\alpha k}\sin\alpha k\theta$ and $r^{\alpha k}\cos\alpha k\theta$.

To establish rates of convergence for the Method of Particular Solutions the smoothness of the holomorphic function ϕ associated with a solution u of $\mathcal{L}u = 0$ is important. This was analyzed by Eisenstat in [23]. He showed that u and its associated function ϕ have the same smoothness behavior on the boundary. To state his theorem we need the following definition.

Let f be defined on a closed subset S of the complex plane. Then f is Hölder continuous with exponent $0 < \gamma \leq 1$ if there exists $K > 0$ such that

$$|f(z_1) - f(z_2)| \leq K|z_1 - z_2|^{\gamma}, \quad \text{for all } z_1, z_2 \in S.$$

Define $C^{p,\gamma}(\overline{\Omega})$ as the space of functions that are p times continuously differentiable in Ω and whose pth derivative is Hölder continuous with exponent γ in $\overline{\Omega}$. Eisenstat proved the following extension of Theorem 6.1.3.

Theorem 6.1.5 (Eisenstat [23]) *Let Ω have no interior or exterior cusps, i.e. for the angle π/α_q at each corner q it holds that $0 < \pi/\alpha_q < 2\pi$. Fix $z_0 \in \overline{\Omega}$.*

1. *Let $\Phi \in \mathcal{C}^{p,\gamma}(\overline{\Omega})$ be holomorphic in Ω and define*

$$u(x,y) := Re\{V[\Phi(z), z_0]\}(z,\overline{z}), \quad z = x + iy \in \overline{\Omega}.$$

   Then $u \in \mathcal{C}^{p,\gamma}(\overline{\Omega})$ and satisfies $\mathcal{L}u = 0$ in Ω.

2. *If $u \in \mathcal{C}^{p,\gamma}(\overline{\Omega})$ satisfies $\mathcal{L}u = 0$ in Ω, then there exists a unique holomorphic function $\Phi \in \mathcal{C}^{p,\gamma}(\overline{\Omega})$ with $\Phi(z_0)$ real such that*

$$\begin{aligned} u(x,y) &= Re\{V[\Phi(z), z_0]\}(z,\overline{z}), \quad z = x + iy \in \overline{\Omega} \\ U(z,z^*) &= I[\phi; z_0](z,z^*), \quad [z,z^*] \in [\Omega,\Omega^*]. \end{aligned} \tag{6.6}$$

Moreover, $\Phi(z) = 2U(z,\overline{z_0}) - U(z_0,\overline{z_0})G(z_0,\overline{z_0},z,\overline{z_0})$.

6.2 Analytic continuation of eigenfunctions via reflection

An important application of Vekua's theory are reflection principles for solutions of elliptic PDE's. Let the boundary $\partial\Omega$ contain a segment σ of the y–axis. Courant and Hilbert [61] described how to analytically continue solutions of the eigenvalue problem (1.1) across σ. The analytic continuation is simply given as $\tilde{u}(x,y) = -u(-x,y)$ for (x,y) in the mirror region Ω' defined by reflecting Ω at the y-axis. In 1954 Garabedian [28] established reflection principles for the equation

$$\mathcal{L}u(x,y) = \Delta u(x,y) + c(x,y)u(x,y) = 0$$

with zero boundary conditions on an arbitrary analytic arc. This was generalized to elliptic PDEs of the form (6.1) and more general boundary conditions by Lewy [48]. In the book by Garabedian [29] the extension of these results to some nonlinear elliptic PDEs is described. Here, we just discuss the case $\mathcal{L} = \Delta + \lambda$ and use the technique described in [28]. Let σ be an analytic segment of the boundary which is parameterized in the form $\overline{w} = R(w)$ for $w \in \sigma$, where R is a holomorphic function in a neighborhood of σ. For example, if σ is a part of the x–axis then $\overline{w} = w$ on σ and therefore $R(w) := w$. If σ is a circle with radius r then $\overline{w}w = r^2$ on σ and $R(w) := r^2/w$. This example shows that in general R is not an entire function. Assume that $z_0 \in \sigma$. Then the function $u(x,y) = U(z,\overline{z})$ can be expressed by Theorem 6.1.3 as

$$U(z,\overline{z}) = \mathrm{Re}\left\{G(z,\overline{z_0},z,\overline{z})\phi(z) - \int_{z_0}^{z} \phi(t)\frac{\partial G}{\partial t}(t,\overline{z_0},z,\overline{z})dt\right\}.$$

By partial integration and using (6.3) this equation becomes

$$U(z,\overline{z}) = U(z_0,\overline{z_0})G(z_0,\overline{z_0},z,\overline{z}) + 2\mathrm{Re}\left\{\int_{z_0}^{z} \frac{\partial U}{\partial t}(t,\overline{z_0})G(t,\overline{z_0},z,\overline{z})dt\right\}, \qquad (6.7)$$

since $U(z_0,\overline{z_0})G(z_0,\overline{z_0},z,\overline{z})$ is real and $G(z,\overline{z_0},z,\overline{z}) = G(z_0,\overline{z_0},z,\overline{z_0}) = 1$ for $\mathcal{L} = \Delta+\lambda$. If $u(x,y)$ satisfies zero Dirichlet boundary conditions on σ we have $U(z_0,\overline{z_0}) = 0$ and therefore

$$U(z,\overline{z}) = 2\mathrm{Re}\left\{\int_{z_0}^{z} \frac{\partial U}{\partial t}(t,\overline{z_0})G(t,\overline{z_0},z,\overline{z})dt\right\}. \qquad (6.8)$$

Define

$$T(z) = \int_{z_0}^{z} \frac{\partial U}{\partial t}(t, \overline{z_0}) G(t, \overline{z}_0, z, R(z)) dt. \tag{6.9}$$

From (6.8) it follows that

$$0 = U(z, \overline{z}) = 2\mathrm{Re}\{T(z)\} = T(z) + \overline{T(z)}$$

for $z \in \sigma$ since $u(x, y) = U(z, \overline{z})$ satisfies the zero boundary conditions there and $\overline{z} = R(z)$ on σ. T is holomorphic in Ω close to σ and has zero real part on σ. Therefore, we can reflect it across σ as

$$\tilde{T}(z) = -\overline{T(\overline{R(z)})} \tag{6.10}$$

for z in the reflection of Ω close to σ. On σ we have

$$\tilde{T}(z) = -\overline{T(\overline{R(z)})} = T(z).$$

Hence, $\tilde{T}$ defines a holomorphic continuation of T if $\overline{R(z)}$ maps complex numbers close to σ on the other side of the boundary line σ. But this is always the case as can be seen by linearizing $R(z)$ close to z_0. To obtain the reflected function $\tilde{u}(x, y)$ the following two steps are necessary. From (6.9) it follows that

$$\overline{T(\overline{R(z)})} + \int_{z_0}^{z} \frac{\partial U}{\partial t}(t, \overline{z}_0) G(t, \overline{z}_0, z, R(z)) dt = 0 \tag{6.11}$$

is a Volterra integral equation[3] defining the holomorphic continuation of $U(z, \overline{z}_0)$ outside Ω. From (6.7) the analytic continuation of $u(x, y)$ across σ can then be obtained. Computing the analytic continuation of an eigenfunction u is complicated since it involves the application of Vekua operators and the solution of an integral equation. But usually we are only interested in the existence of the analytic continuation into a certain region and this only depends on $\partial\Omega$. Consider for example solutions of the eigenvalue problem (1.1) on the unit disk. Then $R(w) := 1/w$ and we can analytically continue any eigenfunction to the whole of the complex plane.

Another simple consequence of reflection principles for eigenfunctions is the analytic continuation in the neighborhood of certain corners of regions. Let u be an eigenfunction in a wedge with interior angle π/k, $k \in \mathbb{N}$. Then by continued reflection u can be analytically continued to an eigenfunction in a wedge with interior angle π. One

[3] A Volterra integral equation has the form $w(z) - \int_{z_0}^{z} K(z, t) w(t) dt = f(z)$. If f is holomorphic in Ω and the kernel $K(z, t)$ is holomorphic in $[\Omega, \Omega] \subset \mathbb{C}^2$ the solution $w(z)$ is holomorphic in Ω [36]. After partial integration (6.11) is of this form in the region of analyticity of $\tilde{T}$.

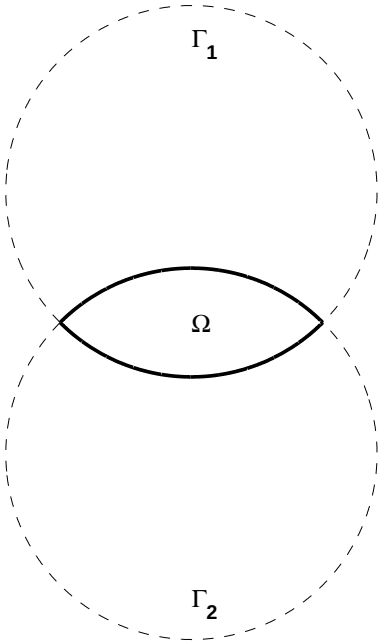


Figure 6.1: A region defined by the intersection of two circles

final reflection then yields the analytic continuation of u in a whole neighborhood of the corner. Hence, any eigenfunction can be analytically continued around a corner that locally consists of two straight lines meeting at an interior angle π/k, where k is an integer. If k is not an integer this is generally not possible. Around a corner with interior angle π/α any eigenfunction u of (1.1) can be expanded into the series,

$$u(r,\theta) = \sum_{k=1}^{\infty} c_k J_{\alpha k}(\sqrt{\lambda} r) \sin \alpha k \theta. \tag{6.12}$$

which is absolutely convergent to u in a neighborhood of the corner [67]. Asymptotically, we obtain

$$u(r,\theta) = \gamma r^{\alpha} \sin \alpha\theta + o(r^{\alpha}), \tag{6.13}$$

for a constant γ depending on the normalization of u. This was proved by Lehmann in 1957 [47] for more general elliptic PDEs. Therefore, if α is not an integer u can in general not be analytically continued around the corner. If the corner does not consist of two straight lines then apart from some special cases u cannot be analytically continued around the corner either. As an example consider the region Ω defined by the intersection of two circles Γ_1 and Γ_2 with radius 1 such that the angles at the two corners are $\pi/2$. Figure 6.1 shows this region.

Theorem 6.2.1 *No eigenfunction on the region Ω of Figure 6.1 can be analytically continued across the two corners of $\partial\Omega$.*

Proof Assume that there is an eigenfunction u of (1.1) on Ω that can be analytically continued around the two corners of Ω. Then we can reflect u across the upper boundary segment of Ω to obtain a function that is analytic inside the disk defined by Γ_1 and also in a neighborhood of Γ_1. By analytic continuation of the lower zero boundary line we obtain $u|_{\Gamma_1} = 0$. Therefore, the analytic continuation of u defines an eigenfunction of (1.1) on the disk enclosed by Γ_1. A nodal line of this eigenfunction is the upper boundary segment of Ω. But nodal lines of eigenfunctions on a disk can only be concentric lines around the center of the disc or straight lines emerging from the center of the disc, a contradiction. ∎

Let us summarize our results in the following definition and theorem.

Definition A corner consisting of two straight lines meeting at an interior angle π/k, where $k \in \mathbb{N}$, is called *regular*. Otherwise, it is called *singular*.

Theorem 6.2.2 *If a corner is regular any eigenfunction can be analytically continued into a neighborhood of it.*

In the following we will use the analytic continuation results from this section to establish exponential convergence of the MPS on certain regions.

6.3 Convergence estimate for regions with no singular corners

In this section we establish convergence estimates of the subspace angle method for regions without singular corners[4]. Before we review some results of complex approximation let us first establish the connection between the subspace angle method and

[4]Exponential convergence of the MPS on regions whose boundary is an analytic Jordan curve was also theoretically established by Still in [67] but without giving exact asymptotic exponential convergence rates.

complex approximation. For a set $S \subset \mathbb{C}$ define the supremum norm $\|\phi\|_{\infty,S}$ as

$$\|\phi\|_{\infty,S} := \sup_{z\in S} |\phi(z)|.$$

Since we assume that Ω is bounded the Vekua operator $\mathrm{Re}\{V[\phi; z_0]\}$ is bounded in $\|\cdot\|_{\infty,\overline{\Omega}}$ by

$$\|\mathrm{Re}\{V[\phi; z_0]\}\|_{\infty,\overline{\Omega}} \leq \|G\|_{\infty,\overline{\Omega}}\|\phi\|_{\infty,\overline{\Omega}} + \int_{z_0}^{z} \|H\|_{\infty,\overline{\Omega}}\|\phi\|_{\infty,\overline{\Omega}} d|t| \leq K_V \|\Phi\|_{\infty,\overline{\Omega}}$$

(see [23]).

For the equation $-\Delta + \lambda u = 0$ the constant K_V depends in addition to the region Ω also on the parameter λ. Let us denote by $\mathcal{V}_A$ the space of all holomorphic functions ϕ associated with functions $u \in \mathcal{A}(\lambda)$ for a fixed $z_0 \in \overline{\Omega}$.

Lemma 6.3.1 *Let (λ_k, u_k) be an eigenpair of (1.1). Fix $z_0 \in \overline{\Omega}$ and let ϕ_k be the holomorphic function associated with u_k. Let $u \in \mathcal{A}(\lambda_k)$ and denote by $\phi \in \mathcal{V}_A$ its associated holomorphic function. Denote by $\theta(\lambda_k)$ the principal angle between $\mathcal{A}(\lambda_k)$ and $\mathcal{D}_0$. Then*

$$\tan\theta(\lambda_k)\|u\|_\Omega \leq C\|\phi - \phi_k\|_{\infty,\overline{\Omega}}$$

for a constant $C > 0$ that depends only on λ_k and Ω.

Proof We have

$$\|u\|_{\partial\Omega} = \|u - u_k\|_{\partial\Omega} \leq C_1\|u - u_k\|_{\infty,\partial\Omega} = C_1\|u - u_k\|_{\infty,\overline{\Omega}} \leq C_1 K_V \|\phi - \phi_k\|_{\infty,\overline{\Omega}}$$

for a constant $C_1 > 0$ that depends on Ω. Therefore, with $C = C_1 K_V$

$$\|u\|_{\partial\Omega} \leq C\|\phi - \phi_k\|_{\infty,\Omega}. \tag{6.14}$$

Since

$$\tan\theta(\lambda_k) = \inf_{u\in\mathcal{A}(\lambda)} \frac{\|u\|_{\partial\Omega}}{\|u\|_\Omega}$$

we obtain

$$\|u\|_\Omega \tan\theta(\lambda_k) \leq \|u\|_{\partial\Omega}.$$

Together with (6.14) the result follows. ∎

Since

$$|\|u\|_\Omega - \|u_k\|_\Omega| \leq \|u - u_k\|_\Omega \leq C\|\phi - \phi_k\|_{\infty,\bar{\Omega}}$$

for a constant $C > 0$ depending on Ω and λ_k the factor $\|u\|_\Omega$ is close to 1 if ϕ is a good approximation of ϕ_k and $\|u_k\|_\Omega = 1$. Hence, the factor $\|u\|_\Omega$ has little influence. The important consequence of this lemma is that $\tan\theta(\lambda_k)$ can be bounded by the error of approximating ϕ_k with functions $\phi \in \mathcal{V}_A$, giving a link between subspace angles and complex approximation. In a similar form this Lemma was already proved in [23] but without using the notion of subspace angles.

After establishing the link between the subspace angle method and the approximation of holomorphic functions let us review some results of complex approximation theory which we will need to establish convergence rates. Let the error $E_{N,K}(\phi)$ of approximating a holomorphic function ϕ in a compact set $K \subset \mathbb{C}$ with polynomials of maximal degree N be defined as

$$E_{N,K}(\phi) = \min_{p\in\Pi_N} \max_{z\in K} |\phi(z) - p(z)|,$$

where Π_N is the space of polynomials of maximal degree N. The first question is whether it is possible at all that $E_{N,K}(\phi) \to 0$ for $N \to \infty$, i.e. if the function ϕ can be arbitrarily well approximated on K by polynomials. If $\mathbb{C}\backslash K$ is connected and ϕ is holomorphic on K this was shown by Runge in 1885. His result is often referred to as the beginning of complex approximation theory (see [27, 53] for an overview of the history). The original theorem of Runge cannot be applied if ϕ is not holomorphic on ∂K. A more general result was proved by Mergelyan in 1951 [53]. He showed that $E_{N,K}(\phi) \to 0$ if $\mathbb{C}\backslash K$ is connected and ϕ is holomorphic in the interior of K and continuous on ∂K. This is a great improvement on Runge's theorem since ϕ needs not be holomorphic on ∂K any longer. The result includes several previous results as special cases (for example, the Weierstrass approximation theorem for approximation on an interval $[a, b]$).

The next question is the speed of convergence of $E_{N,K}(\phi)$, i.e. how fast does $E_{N,K}(\phi)$ go to zero? For simplicity we assume that K is a simply connected compact set bounded by a piecewise analytic Jordan curve. Let $C = \partial K$. Then the equipotential curves C_ρ are defined in the following way.

Equipotential curves Let $w = \Phi(z)$ be the conformal map of the exterior of K to the exterior of the unit disc $\{w \in \mathbb{C} : |w| > 1\}$ normalized in the standard way at

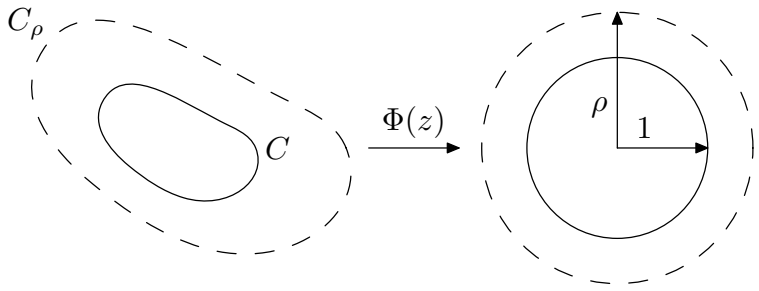


Figure 6.2: The definition of equipotential curves

infinity[5]. The equipotential curve C_ρ of radius $\rho > 1$ is defined as $C_\rho = \{z \in \mathbb{C} : |\Phi(z)| = \rho\}$ (see Figure 6.2).

Furthermore, by the conformal distance of a point z to a region Ω we denote the value ρ such that $z \in C_\rho$.

For $\rho > 1$ the curve C_ρ is always an analytic Jordan curve [84]. The term equipotential curve comes from potential theory. The function $g_K(z) = \log|\Phi(z)|$ is the Green's function for K, which is up to an additional constant the equilibrium potential of K (see [62] for a beautiful introduction). With this definition we can now give the first statement of the rate of convergence of $E_{N,K}(\phi)$.

Theorem 6.3.2 *Suppose $\rho > 1$ is the largest number such that ϕ is analytic inside C_ρ. Then*

$$E_{N,K}(\phi) = O(R^{-N})$$

holds for every $R < \rho$, but for no $R > \rho$.

Proof The classical reference for a proof of this theorem and its implications is Walsh [84]. A more accessible proof is given in [27]. ∎

We are now ready to prove exponential convergence rates for the subspace angle method on regions without singular corners.

[5] $\Phi(z)$ is of the form $\Phi(z) = cz + c_0 + \frac{c_1}{z} + \ldots$ with $c > 0$.

Theorem 6.3.3 *Suppose* (λ_k, u_k) *is an eigenpair of* (1.1) *on* Ω *with boundary* $C = \partial\Omega$. *Let* $\|u_k\|_\Omega = 1$ *and assume that there exists* $R > 1$ *such that* u_k *is analytic inside and on* C_R. *Fix* $z_0 \in \Omega$ *and let*

$$\mathcal{A}_N(\lambda_k) := \left\{ \sum_{j=0}^{N} J_j(\sqrt{\lambda_k} r)(a_j \sin j\theta + b_j \cos j\theta) : a_j, b_j \in \mathbb{R} \right\}$$

be the space of Fourier-Bessel functions of maximum order N *expanded around* z_0. *Then*

$$\tan \theta_N(\lambda_k) = O(R^{-N})$$

as $N \to \infty$, *where* $\theta_N(\lambda_k)$ *is the subspace angle between* $\mathcal{A}_N(\lambda_k)$ *and* $\mathcal{D}_0$.

Proof Although the hypothesis is different, the proof is very similar to the proof of Theorem 8.3 in [23]. From Lemma 6.1.4 it follows with $\alpha = 1$ that $\mathcal{V}_A = \Pi_N$. Let ϕ_k be the holomorphic function associated with u_k. Then from Theorem 6.3.2 it follows that

$$\min_{p_N \in \Pi_N} \|p_N - \phi_k\|_{\infty,\overline{\Omega}} = O(R^{-N}). \tag{6.15}$$

as $N \to \infty$. Let $\tilde{p}_N$ be the best approximation of ϕ_k from Π_N. Together with Lemma 6.3.1 we find

$$\|\mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_\Omega \tan \theta_N(\lambda_k) = O(R^{-N}).$$

We can estimate $\|\mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_\Omega$ as

$$\begin{aligned} \|\mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_\Omega &\geq \|u_k\| - \|u_k - \mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_\Omega \\ &\geq 1 - C\|u_k - \mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_{\infty,\overline{\Omega}} \\ &\geq 1 - CK_V\|\tilde{p}_N - \phi_k\|_{\infty,\overline{\Omega}} \\ &= 1 - O(R^{-N}) \end{aligned}$$

for a constant $C > 0$ that depends on Ω. Therefore, $\|\mathrm{Re}\{V[\tilde{p}_N; z_0]\}\|_{\overline{\Omega}}$ approaches 1 and it follows that

$$\tan \theta_N(\lambda_k) = O(R^{-N}). \quad \blacksquare$$

To find the maximal convergence rate we need to know how far across $\overline{\Omega}$ an eigenfunction u_k of (1.1) can be analytically continued. The singularity z_s of the analytic continuation of u_k with the smallest conformal distance to the region then gives the

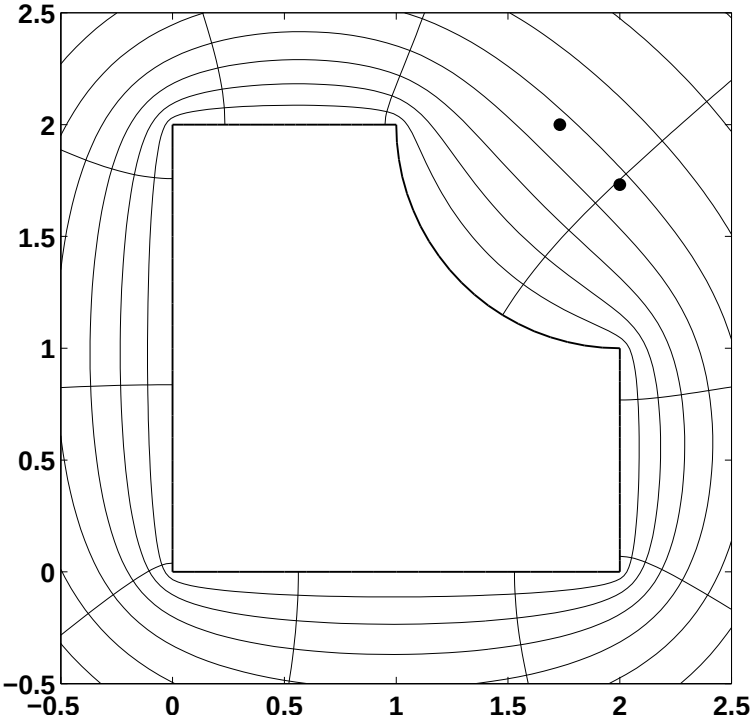


Figure 6.3: The circular L region and some of its equipotential curves. The dots are singularities of the analytic continuation of an eigenfunction of the region.

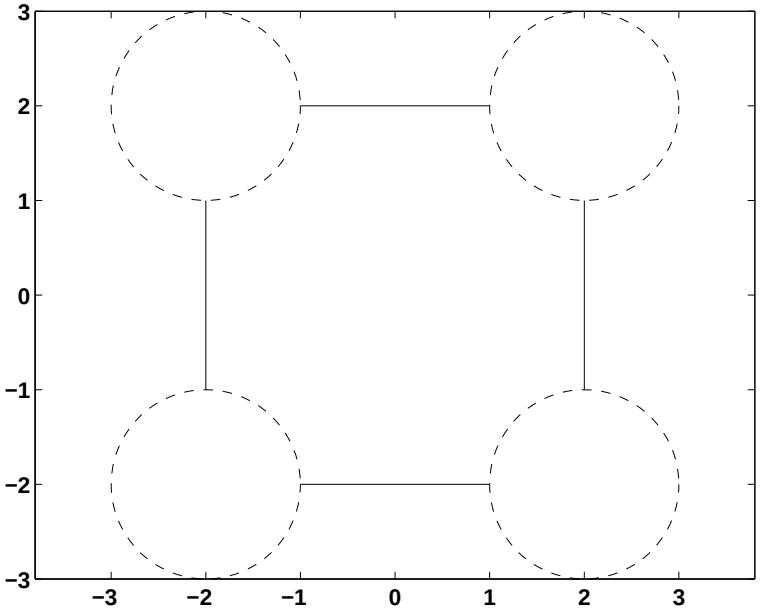


Figure 6.4: Any eigenfunction on the circular L region can be reflected to fill the whole complex plane except for an infinite number of disks of radius 1 positioned on a regular grid.

.

convergence rate of the subspace angle method. Let us demonstrate this with an example. Figure 6.3 shows a region Ω, which is similar to the L-shaped region but with a quarter circle of radius 1 instead of the reentrant corner. The figure also shows some of the equipotential lines C_ρ computed with Driscoll's Schwarz-Christoffel Toolbox [20]. The two dots mark singularities of the analytic continuation of an eigenfunction. Their positions are obtained in the following way. We can reflect any eigenfunction u across each of the straight lines of Ω. The resulting eigenfunction lives in the region shown in Figure 6.4. We can further reflect it to cover the whole complex plane apart from an infinite number of disks of radius 1 ordered on a regular grid. But how far can we reflect an eigenfunction into the disk? Consider for example the two upper disks in Figure 6.4. Let us take a point z_0 lying on the straight line between these disks. We can reflect it across the right circle and obtain a point z_1 lying on the same line but inside the upper right disk. The position of z_1 is obtained from the equation

$$(z_0 - (2+2i))\overline{(z_1 - (2+2i))} = 1.$$

Now we can reflect z_1 again in the upper left circle to obtain a point z_2 in that disk. Then we can reflect again in the right circle and so on. The two limit points of this iteration are not the centers of the disks, but are determined by the condition that the reflection of the left limit point in the right disk is exactly the left limit point and vice versa. Due to the symmetry of the region their distances to the centers of the two disks must be equal. This leads to the two limit points $z_L = -\sqrt{3} + 2i$ in the left disk and to $z_R = \sqrt{3} + 2i$ in the right disk. We cannot analytically continue u further into the two disks on the line $x + 2i$, $x \in \mathbb{R}$. Therefore, the points z_L and z_R are singularities of the analytic continuation of u. By symmetry we obtain many more singularities such as the point $z_{R'} = 2 + i\sqrt{3}$. The points z_R and $z_{R'}$ are plotted in Figure 6.3. There might be other singularities with a smaller conformal distance to Ω. But if such singularities exist their conformal distance will not be significantly smaller than that of z_R and $z_{R'}$ since the upper line in Figure 6.4 is the shortest connection between two disks and therefore leads to the smallest penetration into the two disks by reflection. For example, if we reflect between the lower left circle and the upper right circle we obtain two singularities whose conformal distance to the region is larger than that of z_R. With the Schwarz-Christoffel toolbox we obtain the value $\rho \approx 1.476$ for the conformal distance of z_R to Ω. Approximating with Fourier-Bessel sine and cosine functions we obtain from Theorem 6.3.3

$$\tan\theta_N(\lambda_k) = O(1.476^{-N})$$

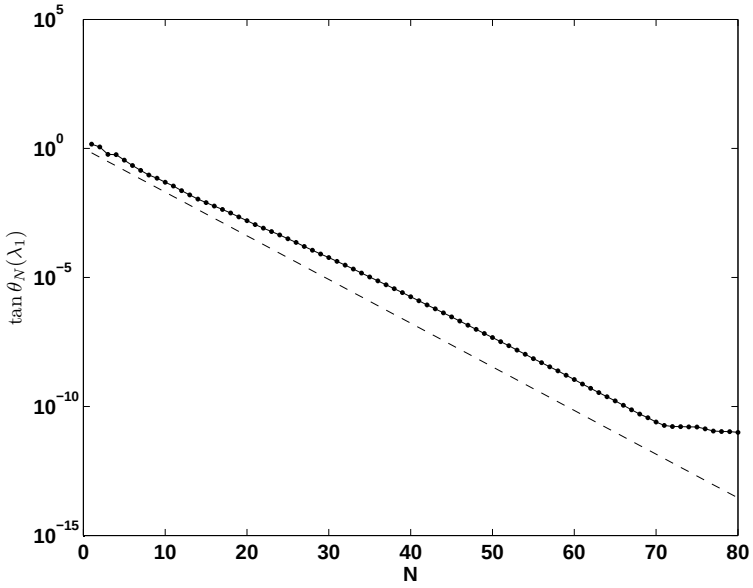


Figure 6.5: The theoretical convergence rate for $N \to \infty$ (dashed) and the measured convergence (solid) on the L circle region. The measured convergence fits well with the predicted rate.

for an eigenvalue λ_k of (1.1) on the circular L region, assuming z_R is the singularity with the smallest conformal distance[6]. In Figure 6.5 we compare the computed value of $\tan\theta_N(\lambda_1)$ with the estimated maximal convergence rate. For λ_1 we used the approximation $\lambda_1 \approx 7.02025391131$. The convergence stops at $N = 70$ since λ_1 is only known to about 12 digits of accuracy. The measured convergence curve first seems to move away from the estimated straight line. But then the slope of the measured convergence slowly approaches the estimated value again and we obtain a good match between the estimated slope and the measured slope. We always have to keep in mind that the estimated convergence rate is an asymptotic rate for $N \to \infty$. The transient behavior of the measured curve can differ from this.

Another interesting example is the half annulus with radii $r_1 = 1$ and $r_2 = 2$ as shown in Figure 6.6, where also some equipotential curves are plotted. The closest singularity is at 0 leading to a theoretical convergence rate of $O(1.16^{-N})$. By increasing r_2 we could make the asymptotic rate of convergence arbitrarily close to 1. The reason for this slow convergence is that the approximation basis is not a very good one. All eigenvalues on the half annulus are also eigenvalues on the full annulus. But

[6]In a strict sense, we can only say that $\tan\theta_N(\lambda_k) = O(R^{-N})$ for $R < 1.476$ if 1.476 is the exact maximum radius of analyticity. But since 1.476 is just a numerically estimated value we will omit this and just say that the rate is $O(1.476^{-N})$.

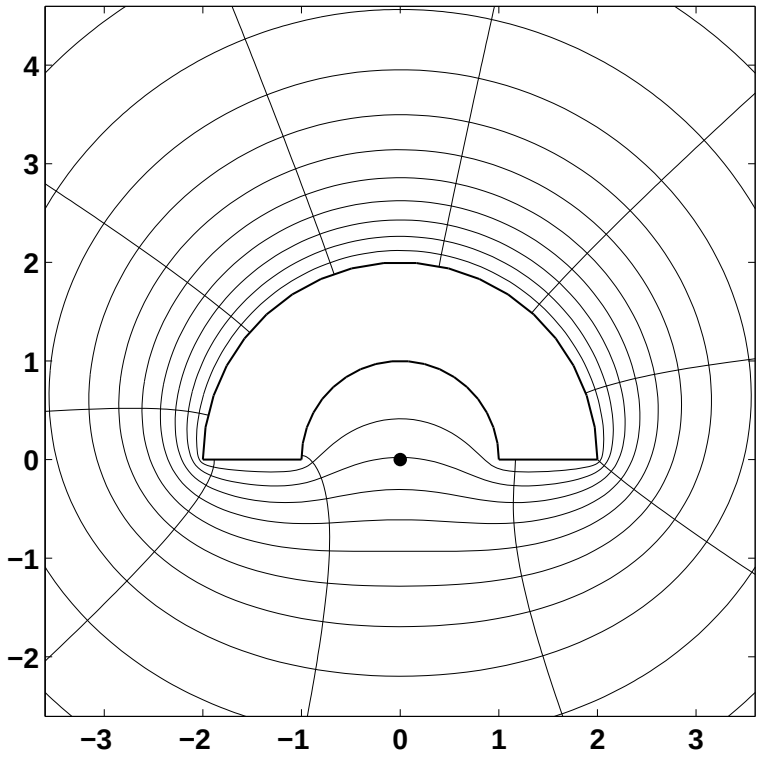


Figure 6.6: A half annulus region and its mapping function. The dot marks the closest singularity of an eigenfunction of the region.

the corresponding eigenfunctions are given as linear combinations of Fourier-Bessel functions of the first and of the second kind by

$$u_{m,n} = \left[Y_m(k_{m,n}) J_m\left(\frac{k_{m,n} r}{a}\right) - J_m(k_{m,n}) Y_m\left(\frac{k_{m,n} r}{a}\right) \right] \left[A \cos m\theta + B \sin m\theta \right],$$

where $k_{m,n}$ is the nth root of

$$Y_m(k) J_m\left(\frac{kb}{a}\right) - J_m(k) Y_m\left(\frac{kb}{a}\right) = 0$$

and Y_m is the mth Bessel function of the second kind [44]. Therefore, although by approximating with Fourier-Bessel sine and cosine functions inside the half annulus we are guaranteed exponential convergence, it is not a good basis since the eigenfunctions also involve Bessel functions of the second kind, which have a singularity at 0. The convergence curve for $\tan\theta_N(\lambda_1)$ on this region is shown in Figure 6.7. It agrees very well with the predicted value of $O(1.16^{-N})$.

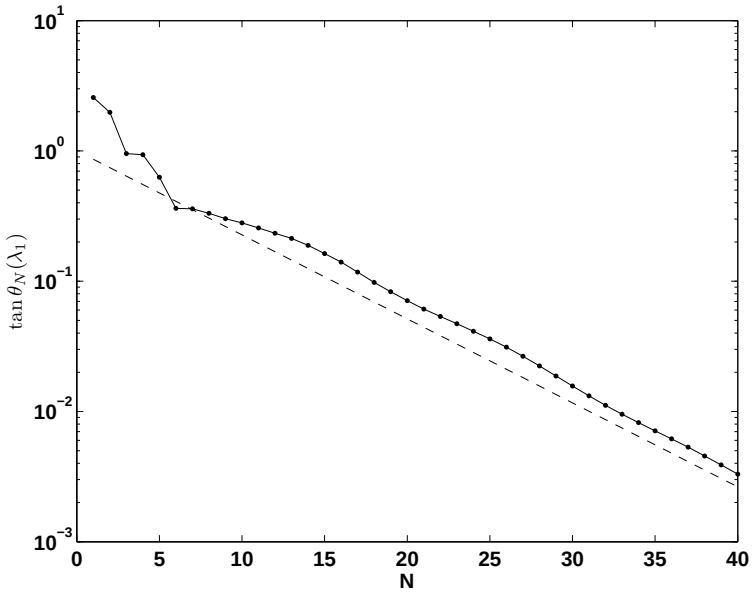


Figure 6.7: Convergence of the subspace angle method on the half annulus region (solid). The dashed line shows the theoretical asymptotic rate of convergence.

6.4 Exponential convergence on regions with one singular corner

Only in the case in which $\partial\Omega$ has no singular corner can we apply Theorem 6.3.3 to obtain an exponential convergence estimate. Let us now extend this result to the case in which $\partial\Omega$ has exactly one singular corner whose adjacent arcs are straight lines. In such a case any eigenfunction on Ω can be analytically continued across $\overline{\Omega}$ except close to the singular corner. This is for example the case for the L-shaped region. We need the following lemma.

Lemma 6.4.1 *Let* 0 *be a corner of* $\partial\Omega$ *with interior angle* π/α*, which is formed by two straight arcs from which the right one is part of the real axis as shown in Figure* 2.1*. Let* u *be an eigenfunction of* (1.1) *on* Ω*. Then the unique holomorphic function* ϕ *associated with* u *by the Vekua operator such that* $u = Re\{V[\phi; 0]\}$ *is purely imaginary on the arcs adjacent to* 0 *and has the absolutely convergent expansion*

$$\phi(z) = \sum_{k=1}^{\infty} i\tilde{c}_k z^{\alpha k}, \quad \tilde{c}_k \in \mathbb{R},\ z \in \{z : |z| < R\} \cap \overline{\Omega}$$

for some $R > 0$.

Proof There exists $R > 0$ such that

$$u(r,\theta) = \sum_{j=1}^{\infty} c_j J_{\alpha j}(\sqrt{\lambda_k} r) \sin \alpha j \theta \tag{6.16}$$

is an absolutely convergent series for $r < R$ and $0 \leq \theta \leq \pi/\alpha$. Together with Lemma 6.1.4 it follows that there exists real coefficients

$$\tilde{c}_j = -\frac{c_j \sqrt{\lambda_k}^{\alpha j}}{2^{\alpha j}\Gamma(\alpha j + 1)},$$

such that for $|z| < R$ we have

$$\phi(z) = \sum_{j=0}^{\infty} i\tilde{c}_j z^{\alpha j}. \tag{6.17}$$

The absolute convergence of this series follows from the fact that

$$\frac{|z|^\nu}{2^\nu \Gamma(\nu+1)}(1-\epsilon) \leq |J_\nu(z)| \leq \frac{|z|^\nu}{2^\nu \Gamma(\nu+1)}, \quad |z| \leq \tau$$

for every $\epsilon > 0$ and $\nu > \nu_0(\tau, \epsilon)$ sufficiently large (see for example [67]). Hence, the terms in (6.17) can be bounded by the terms in (6.16).

From (6.17) it follows that ϕ is purely imaginary on $\{z : |z| < R\} \cap \partial\Omega$. By analytic continuation along the arcs ϕ is also purely imaginary on the whole of both arcs. ∎

Using this lemma we can show that the singularity can be removed by a conformal map of the region.

Lemma 6.4.2 *In the notation of Lemma 6.4.1, define the region Ω^α as*

$$\Omega^\alpha = \{z^\alpha : z \in \Omega\}.$$

Then the function $\tilde{\phi}(w) := \phi(z)$ for $w = z^\alpha$ is analytic on $\overline{\Omega^\alpha}$.

This lemma states that by the transformation $w = z^\alpha$ we obtain a function that is analytic on the whole of $\overline{\Omega^\alpha}$ and therefore in particular at $w = z = 0$. Hence, by the transformation $w = z^\alpha$ we have removed the singularity of ϕ at 0.

Proof From Lemma 6.4.1 it follows that

$$\phi(z) = \sum_{j=0}^{\infty} i\tilde{c}_j z^{\alpha j} \tag{6.18}$$

in $\{z : |z| < R\} \cap \overline{\Omega}$. Now let $w = z^\alpha$. Then

$$\tilde{\phi}(w) = \sum_{j=0}^{\infty} i\tilde{c}_j w^j.$$

This is a power series in w. Its absolute convergence for $|w| < R^\alpha$ follows from the absolute convergence of (6.18). Hence, $\tilde{\phi}$ possesses a power series expansion around $w = 0$ and is therefore holomorphic in a neighborhood of 0. Since by assumption $\tilde{\phi}$ possesses no other singularities on $\overline{\Omega^\alpha}$ it is holomorphic there. ∎

Let us now use these lemmas to determine the rate of convergence of the MPS in a region with one singular corner. Assume that Ω satisfies the hypotheses of Lemma 6.4.1 and let (λ_k, u_k) be an eigenpair of (1.1) on Ω with $\|u_k\|_\Omega = 1$. Denote by ϕ_k the holomorphic function associated with u_k for $z_0 = 0$ and let $\tilde{\phi}_k$ be the conformal transplant of ϕ_k to the region Ω^α. Hence, $\tilde{\phi}_k(w) = \phi_k(z)$ for $w = z^\alpha$. From Lemma 6.4.2 it follows that $\tilde{\phi}_k$ is holomorphic on $\overline{\Omega^\alpha}$. Therefore, there exists $R > 1$ such that

$$\min_{p_N \in \Pi_N} \|\tilde{\phi}_k - p_N\|_{\infty, \overline{\Omega^\alpha}} = O(R^{-N}).$$

This estimate holds for all $R < \rho$, where ρ is the conformal distance of the closest singularity of $\tilde{\phi}_k$ to the region $\overline{\Omega^\alpha}$ (see Theorem 6.3.2). Let

$$\tilde{p}_N(w) = \sum_{j=0}^{N} c_k w^j, \quad c_j \in \mathbb{C}$$

be the best approximating polynomial in Π_N of $\tilde{\phi}_k$. Then

$$\tilde{p}_N(z) = \sum_{k=0}^{N} c_k z^{\alpha k}$$

and

$$\|\phi_k - \tilde{p}_N\|_{\infty, \overline{\Omega}} = O(R^{-N})$$

in the z–domain. The Vekua transform $u = \mathrm{Re}\{V[\tilde{p}_N; 0]\}$ of $\tilde{p}_N$ in the z–domain has the form

$$u(r, \theta) = \sum_{j=0}^{N} J_{\alpha j}(\sqrt{\lambda_k} r)(a_j \sin \alpha j\theta + b_j \cos \alpha j\theta)$$

for real coefficients a_j and b_j depending on the coefficients c_j of $\tilde{p}_N$. Now define

$$\mathcal{A}_N(\lambda) := \{\sum_{j=0}^{N} J_{\alpha j}(\sqrt{\lambda_k} r)(a_k \sin \alpha j\theta + b_j \cos \alpha j\theta) : a_j, b_j \in \mathbb{R}\}.$$

We obtain as in the proof of Theorem 6.3.3

$$\tan \theta_N(\lambda_k) = O(R^{-N})$$

for the angle between $\mathcal{A}_N(\lambda_k)$ and $\mathcal{D}_0$. Let us summarize this result as a theorem.

Theorem 6.4.3 *Let Ω be a region with one singular corner with interior angle π/α as defined in Lemma* 6.4.1 *and let (λ_k, u_k) be an eigenpair of* (1.1) *on this region. Let ϕ_k be the holomorphic function associated with u_k and define $\tilde{\phi}_k$ by $\tilde{\phi}_k(w) = \phi_k(z)$ for $w = z^\alpha$. Let*

$$\mathcal{A}_N(\lambda) := \{\sum_{j=0}^{N} J_{\alpha j}(\sqrt{\lambda_k} r)(a_j \sin \alpha j\theta + b_j \cos \alpha j\theta) : a_j, b_j \in \mathbb{R}\}.$$

Then there exists $R > 1$ such that for the angle $\theta_N(\lambda_k)$ between $\mathcal{A}_N(\lambda_k)$ and $\mathcal{D}_0$ it holds that

$$\tan \theta_N(\lambda_k) = O(R^{-N}).$$

This estimate holds for all $R < \rho$ where ρ is the smallest conformal distance of a singularity of $\tilde{\phi}$ to the region $\overline{\Omega^\alpha}$.

Hence, by adapting the space of particular solutions to the singularity we obtain exponential convergence for regions with one corner singularity. For the L-shaped region this was also investigated by Still in [67]. But he did not use conformal mapping techniques to compute the exact asymptotic rate of convergence but rather gave bounds on the convergence rate by directly estimating the Fourier-Bessel series.

Let us use this result to determine the rate of convergence if Ω is the L-shaped region. By reflection we can determine all singularities around Ω and remove the singularity at the reentrant corner with the map $w = z^{2/3}$. This is shown in Figure 6.8, where $\Omega^{2/3}$ and some equipotential curves are plotted[7]. The dots mark singularities of the analytic continuation of eigenfunctions on Ω under the map to the w–domain.

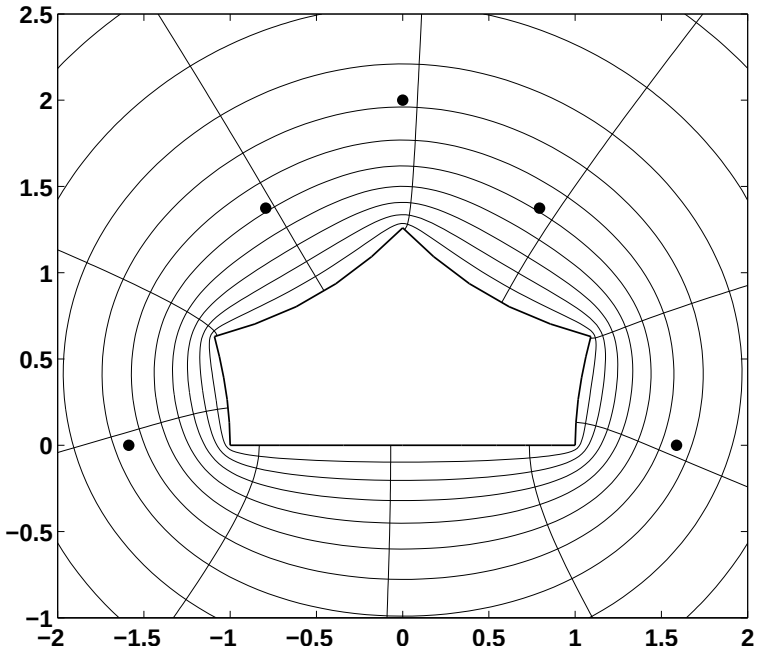


Figure 6.8: Equipotential curves of the L-shaped region after canceling out the reentrant corner at $z = 0$. The dots mark singularities of the analytic continuation of an eigenfunction u.

Computing the minimum conformal distance to $\Omega^{2/3}$ of the singularities leads to $\rho \approx 1.54$. A comparison between the measured convergence and this estimated rate is shown in Figure 6.9. The estimated rate fits well with the measured rate. Again the measured convergence is not fully identical to the estimated behavior for $N \to \infty$ since we only observe the curve up to $N = 60$. But for large N the measured slope will eventually approach the estimated rate.

The convergence plot in Figure 6.9 was computed using Fourier-Bessel sine and cosine functions of the form

$$J_{\frac{2}{3}j}(\sqrt{\lambda_1}r)\sin\frac{2}{3}j\theta$$

and

$$J_{\frac{2}{3}j}(\sqrt{\lambda_1}r)\cos\frac{2}{3}j\theta$$

since this corresponds to polynomial approximation on the region $\Omega^{2/3}$. But the Fourier-Bessel cosine functions do not satisfy the zero boundary conditions on the

[7] To compute the equipotential curves and the conformal distance of the singularities we used Driscoll's Schwarz-Christoffel Toolbox after discretizing $\Omega^{2/3}$ to obtain a polygonal region. The computations can also be done on a slicker way without this discretization, but we will not go into that here.

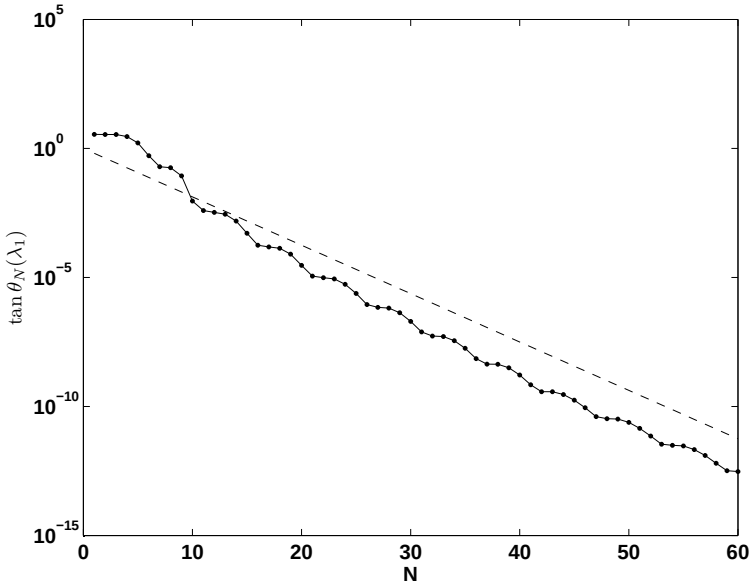


Figure 6.9: Measured convergence (solid) of $\tan\theta_N(\lambda_1)$ compared to the estimated convergence rate (dashed) for the L-shaped region using Fourier-Bessel sine and cosine expansions.

arcs adjacent to the reentrant corner and it seems natural to only use Fourier-Bessel sine functions. How does the convergence rate change in this case?

Approximating only with Fourier-Bessel sine functions corresponds to approximating $\tilde{\phi}(w)$ with polynomials that have purely imaginary coefficients. The best approximating polynomial in this class can be determined with the following Lemma.

Lemma 6.4.4 *Let Ω be the L-shaped region and denote by $\tilde{\Omega}^{2/3}$ the reflection of $\Omega^{2/3}$ at the real line as shown in Figure* 6.10. *Let ϕ be a function holomorphic in $\Omega^{2/3}$ with no singularites in $\tilde{\Omega}^{2/3}$ which is purely imaginary on the part of $\partial\Omega^{2/3}$ intersecting with the real line. Denote by $\tilde{p}_N$ the best approximating polynomial of degree N for the function*

$$\tilde{\phi}(z) = \begin{cases} \phi(z); & z \in \Omega^{2/3} \\ -\overline{\phi(\overline{z})}; & \overline{z} \in \Omega^{2/3} \end{cases}$$

in $\overline{\tilde{\Omega}^{2/3}}$. Then $\tilde{p}_N$ has purely imaginary coefficients and is the best approximating polynomial for ϕ in $\overline{\Omega^{2/3}}$ from the space of polynomials of maximum degree N with purely imaginary coefficients.

Proof Let $\overline{\tilde{\phi}(\overline{z})}$ be the reflection of $\tilde{\phi}(z)$ on the real axis. Then the best approximating

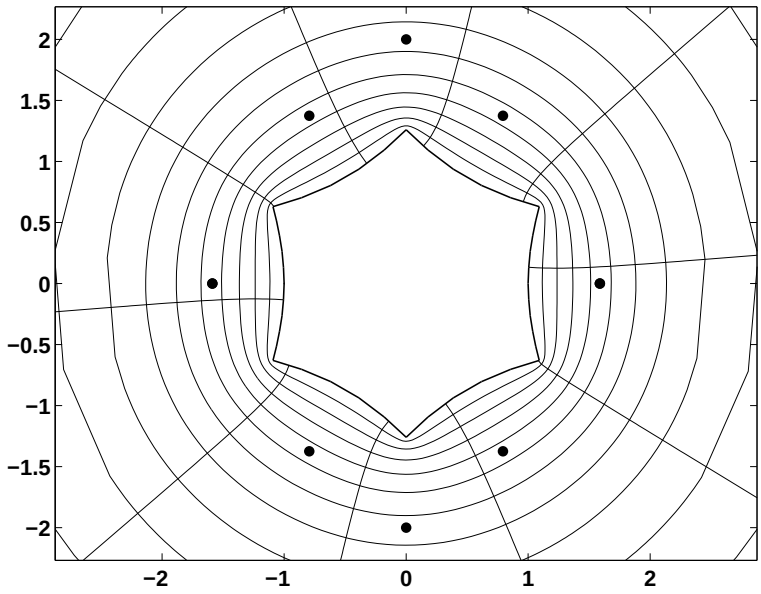


Figure 6.10: The region from Figure 6.8 but now doubled in size by an additional reflection. The dots are again singularities of the analytic continuation of an eigenfunction u_k.

.

polynomial for this reflection is $\overline{\tilde{p}_N(\overline{z})}$. But it can be easily seen that $\overline{\tilde{\phi}(\overline{z})} = -\tilde{\phi}(z)$. Therefore, $\overline{\tilde{p}_N(\overline{z})} = -p_N(z)$. With

$$\tilde{p}_N(z) = \sum_{j=0}^{N} c_k z^j$$

for certain coefficients $c_j \in \mathbb{C}$ we find

$$0 = \overline{\tilde{p}_N(\overline{z})} + \tilde{p}_N(z) = \sum_{j=0}^{N} (c_j + \overline{c_j}) z^k = \sum_{j=0}^{N} 2\mathrm{Re}\{c_j\} z^j$$

and therefore $\mathrm{Re}\{c_j\} = 0$ for $j = 1, \ldots, N$. Now let $\hat{p}_N$ be the best approximating polynomial for ϕ in $\overline{\Omega^{2/3}}$ from the space of polynomials of maximum degree N with purely imaginary coefficients. Since the best approximating polynomial $\tilde{p}_N$ for $\tilde{\phi}$ on $\overline{\tilde{\Omega}^{2/3}}$ also has purely imaginary coefficients and $\tilde{\phi}$ is symmetric around the real axis it follows that $\hat{p}_N = \tilde{p}_N$. ∎

For the MPS on the L-shaped region with Fourier-Bessel functions this result means that the rate of convergence is determined by a conformal map of the region shown in Figure 6.10. We obtain an asymptotic convergence rate of $O(1.44^{-N})$. Figure 6.11

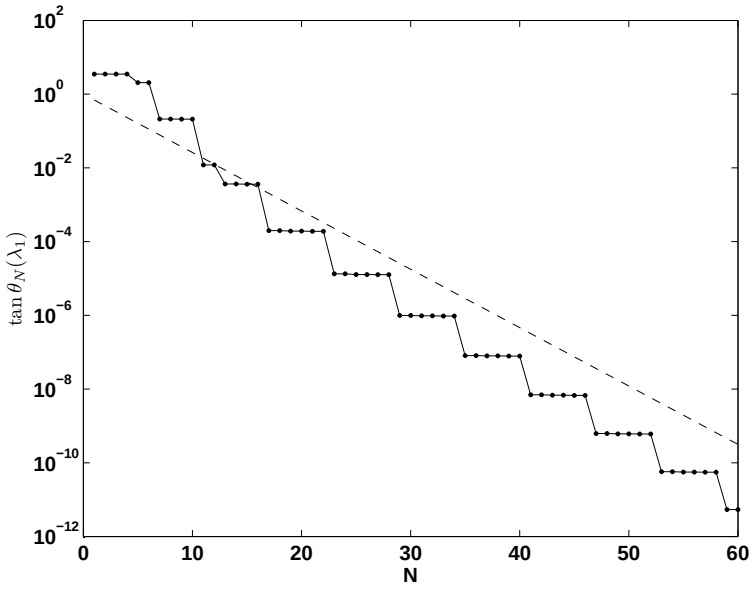


Figure 6.11: Measured and estimated convergence of the MPS on the L-shaped region using only Fourier-Bessel sine functions.

shows the measured and the estimated convergence behavior on the L-shaped region if only Fourier-Bessel sine functions are used. It is interesting to note that the observed transient rate of convergence from Figure 6.11 seems to be about $O(1.51^{-N})$, which is much closer to the asymptotic rate of convergence for the case that Fourier-Bessel sine and cosine functions are used. This shows that the effect of omitting Fourier-Bessel cosine functions from the basis is low and only becomes significant for $N \to \infty$. It is also noticeable that the curve in Figure 6.11 bends up slightly as in the beginning it seems to converge faster than in later steps. Eventually it will settle at a rate of $O(1.44^{-N})$ for $N \to \infty$.

6.5 Convergence on regions with multiple singularities

Let us now apply the subspace angle method to the eigenvalue problem (1.1) on a region Ω with more than one singular corner. For regions with one corner singularity we were able show exponential convergence by canceling out the singularity with a conformal map if the corner is bounded by the intersection of two straight arcs. So the first question is if we can just use the same strategy for regions with multiple

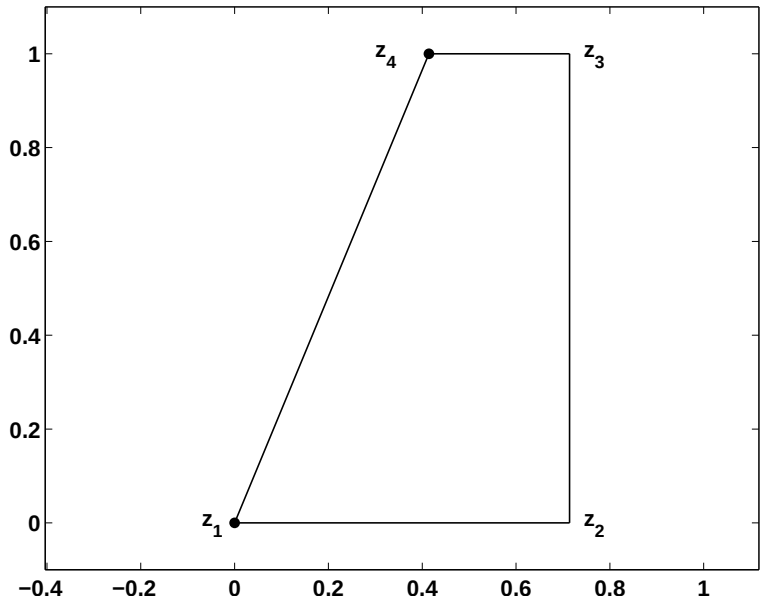


Figure 6.12: A region with two singular corners at z_1 and z_4.

singularities. We will demonstrate all results in this section at the example region shown in Figure 6.12. The four points of the quadrilateral are $z_1 = 0$, $z_2 = .3 + 1/\tan(\frac{3\pi}{8})$, $z_3 = 0.3 + 1/\tan(\frac{3\pi}{8}) + 1i$, $z_4 = 1/\tan(\frac{3\pi}{8}) + 1i$. The corresponding interior angles π/α_k, $k = 1, \ldots, 4$ are defined by $\alpha_1 = \frac{8}{3}$, $\alpha_2 = 2$, $\alpha_3 = 2$, $\alpha_4 = \frac{8}{5}$. Hence, the corners at z_1 and z_4 are singular. Let u_k be an eigenfunction of (1.1) on the region Ω shown in Figure 6.12 and ϕ_k its associated holomorphic function around z_1, i.e. $u_k = \mathrm{Re}\{V[\phi_k; z_1]\}$. Then from Lemma 6.4.1 it follows that

$$\phi_k(z) = \sum_{j=1}^{\infty} i c_j z^{\alpha_1 j}, \quad c_j \in \mathbb{R}$$

close to z_1. But close to z_4 we cannot expect ϕ_k to have this form. Close to that corner the eigenfunction can be expressed as a convergent series of the form

$$u(r, \theta) = \sum_{j=1}^{\infty} a_j J_{\alpha_4 j}(\sqrt{\lambda_k} r) \sin \alpha_4 j\theta \tag{6.19}$$

with origin of the polar coordinates at z_4. Together with (6.5) it follows that asymptotically

$$\phi_k(z) \sim \sum_{j,l \geq 0} c_{j,l} (z - z_4)^{j + l\alpha_4}, \quad c_{j,l} \in \mathbb{C} \tag{6.20}$$

as $z \to z_4$. Generally, this needs not be a convergent series. One can show that if α is irrational ϕ_k has an asymptotic expansion close to a corner π/α in the terms $z^{j+\alpha l}$

$j, l \in \mathbb{N}$. If $\alpha = p/q$ is rational with (p, q) relatively prime the asymptotic expansion can also have terms of the form $(z^p \log z)^m z^{j+\alpha l}$, $j, l, m \in \mathbb{N}$ [85]. Let us now assume we introduce the conformal map $w = (z - z_4)^{\alpha_4}$ to straighten out the corner at z_4. Then from (6.20) it follows that

$$\phi_k(w) \sim \sum_{j,l \geq 0} c_{j,l} w^{\frac{j}{\alpha_4}+l}$$

close to $w = 0$. But this asymptotic series still has singular terms, which are now of the form w^{j/α_4+l}. Hence, although we have straightened out the corner at z_4 we haven't canceled out the corresponding singularity of ϕ_k. It follows that we can always only cancel out one singularity of ϕ_k, namely the singularity z_s such that $u_k = \mathrm{Re}\{V[\phi_k; z_s]\}$. Due to the behavior of ϕ_k at the other singularities it is not possible to cancel them out with the strategy used in the previous section. But as we will see later, by a suitable choice of basis functions we can reduce the order of the corner singularities and still obtain fast convergence.

Algebraic convergence estimates for the MPS were first analyzed by Eisenstat in 1974 [23]. This was further developed by Still [66, 67, 69] in the 1980's. In 1999 Melenk [52] published algebraic convergence results for approximation in Sobolev spaces. Here, we will mostly use the estimates in the form given by Still.

Let Ω have corners at $z_1, \ldots, z_n$. Denote by $\omega_1 = \pi/\alpha_1, \ldots, \omega_n = \pi/\alpha_n$ the corresponding interior angles. Let $\bar{\omega}$ be the largest interior angle and define $\mu := \min\{1, 2 - \bar{\omega}/\pi\}$. If Ω has a reentrant corner then $\mu < 1$ and $\mu\pi$ is the exterior angle at the reentrant corner. Furthermore, we assume that Ω has no interior or exterior cusps, i.e. $0 < \omega_k < 2\pi$ for $k = 1, \ldots, n$. Let $z_0 \in \Omega$ and define

$$\mathcal{A}_N(\lambda_k) := \left\{ \sum_{j=0}^{N} J_j(\sqrt{\lambda_k} r)(a_j \sin j\theta + b_j \cos j\theta) : a_j, b_j \in \mathbb{R} \right\},$$

where the polar coordinates are around z_0. Still proved the following theorem, which we present for the special case of the eigenvalue problem (1.1).

Theorem 6.5.1 (Still [69]) *Let (λ_k, u_k) be an eigenpair of (1.1). Let $p \in \mathbb{N}$, $0 < \gamma \leq 1$, be defined by $p + \gamma = \frac{\pi}{\bar{\omega}}$. Then $u \in \mathcal{C}^{p,\gamma}(\overline{\Omega})$ and for any $\epsilon > 0$ there exists a constant $c(\epsilon)$ such that*

$$\min_{u_N \in \mathcal{A}_N(\lambda)} \|u_k - u_N\|_{\infty,\overline{\Omega}} \leq \frac{c(\epsilon)}{N^{\mu(p+\gamma)-\epsilon}}.$$

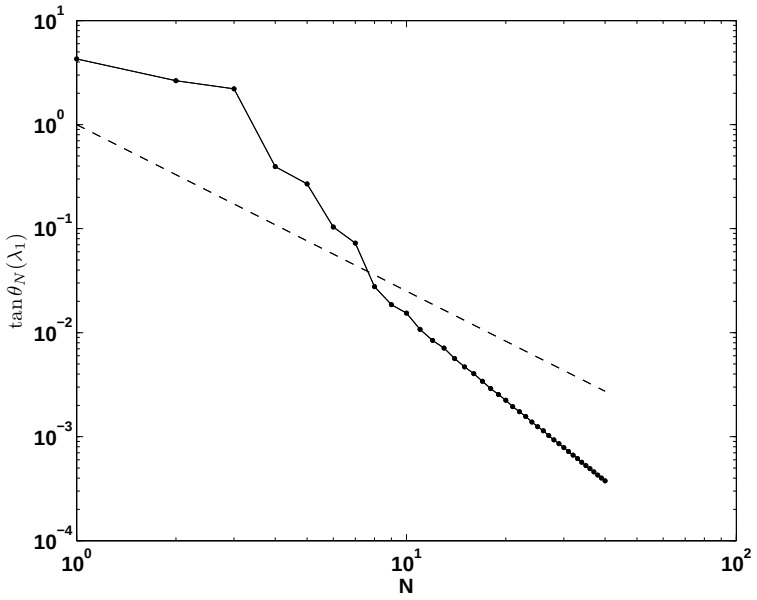


Figure 6.13: The plot shows the convergence behavior on a region with two singular corners (solid line). The dashed line is the theoretical rate from Theorem 6.5.1.

This result was also proved by Eisenstat in a more general setting in [23]. Since eigenfunctions are analytic around nonsingular corners we only need to consider singular corners for Theorem 6.5.1.

Figure 6.13 shows the convergence of the subspace angle method using the approximation space $\mathcal{A}(\lambda_1)$ as defined above. For λ_1 we use the approximation $\lambda_1 \approx 48.4161682676614$, which is believed to be correct to all given digits. The dashed line in Figure 6.13 shows the convergence bound from Theorem 6.5.1. The corner at z_4 is the singular corner with the largest interior angle. From Theorem 6.5.1 the convergence estimate

$$\tan\theta_N(\lambda_1) \leq \frac{c(\epsilon)}{N^{\frac{8}{5}-\epsilon}}$$

for all $\epsilon > 0$ follows. The observed convergence in Figure 6.13 is even faster than predicted by algebraic convergence estimates in the supremum norm. Indeed, Melenk showed that bounding the approximation error in L^2 can lead to improved algebraic convergence rates [52].[8] Therefore, although the error in the supremum norm is an upper bound for the tangent of the subspace angle, more suitable function space settings might give sharper bounds in this case. All algebraic convergence estimates

[8]For the approximation of functions that are holomorphic in the neighborhood of the region as described by Theorem 6.3.2 the asymptotic exponential rates are the same in the supremum and L^2 norm.

depend on the smoothness of the eigenfunctions at the singular corners. If we can improve the smoothness at the corners, faster convergence rates are possible. This possibility was investigated by Eisenstat [23]. Consider the corner z_4. As already stated in (6.19), close to z_4 u_k has the series representation

$$u_k(r,\theta) = \sum_{j=1}^{\infty} a_j J_{\alpha_4 j}(\sqrt{\lambda_k} r) \sin \alpha_4 j\theta.$$

By canceling out lower order terms of this series we can improve the smoothness of u_k at z_4. If we enrich the approximation space by linear combinations of $J_{\alpha_4 j}(\sqrt{\lambda_k} r) \sin \alpha_4 j\theta$, $j = 1, \ldots, n$ the problem of approximating u_k can be interpreted as the problem of approximating

$$\tilde{u}_k(r,\theta) := u_k(r,\theta) - \sum_{j=1}^{n} a_j J_{\alpha_4 j}(\sqrt{\lambda_k} r) \sin \alpha_4 j\theta = \sum_{j=n+1}^{\infty} a_j J_{\alpha_4 j}(\sqrt{\lambda_k} r) \sin \alpha_4 j\theta$$

close to z_4, which has a much weaker singularity at z_4 than the function u_k. If we do the same at z_1 we can make the function $\tilde{u}_k$ as smooth at the corners as we wish. It cannot become analytic close to the corners since no matter how many singular terms we use the remaining terms in the series expansions around z_1 and z_4 will stay singular. But still we can obtain high algebraic convergence rates. This is shown in Figure 6.14, where we compare the convergence of the subspace angle method if the approximation space is enlarged by the first two singular Fourier-Bessel terms around the singularities at z_1 and z_4 to the case of approximating just with Fourier-Bessel terms in the interior of Ω as done in Figure 6.13.

Just by adding four singular terms we drastically increase the rate of convergence while the additional computational effort is negligible since the number of basis functions only grows from 201 to 205 at the step $N = 100$. We can even improve the rate of convergence more by choosing in each step the same number of terms at the two singularities as in the interior of the region. Figure 6.15 shows a double-logarithmic and a semi-logarithmic plot of the resulting convergence curve. It decreases faster than linearly in the double-logarithmic plot indicating super-algebraic convergence, but decreases slower than linearly in the semi-logarithmic plot, which indicates a rate slower than exponential convergence. We achieve an accuracy close to machine precision after $N = 26$, which corresponds to 105 basis terms (26 Fourier-Bessel sine terms around each of the singular corners, 52 Fourier-Bessel sine and cosine terms in the interior and 1 Bessel term of order zero in the interior). In the case of using

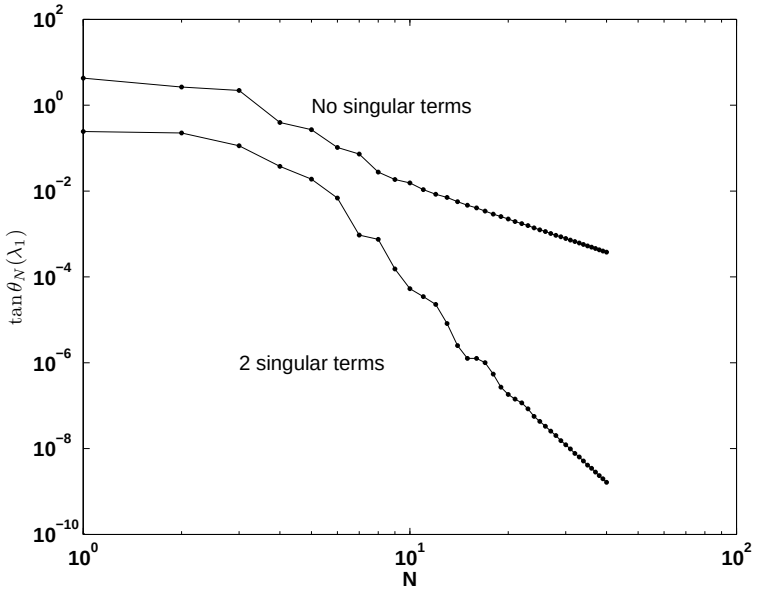


Figure 6.14: Comparison of the subspace angle method on Ω by only using Fourier-Bessel terms in the interior of the region and by adding two singular terms around each of the two singular corners z_1 and z_4.

only Fourier-Bessel terms in the interior of the region shown in Figure 6.13 we needed 201 terms to bring the subspace angle down to just about 10^{-4}. This shows how essential it is to capture the corner singularities correctly. We can even reduce the number of basis functions more. From Theorem 6.5.1 it follows that close to z_1 the rate of convergence is $O(N^{-\frac{8}{3}})$, while close to z_4 it is $O(N^{-\frac{8}{5}})$. Hence, close to z_1 the convergence is about twice as fast as close to z_4. Therefore, it makes sense to choose twice as many singular terms around z_4 as around z_1 to make up for this difference in the convergence rate. This rule of thumb was pointed out by Descloux and Tolley [18]. As a result we obtain the convergence curve shown in Figure 6.16. At about $N = 14$ the subspace angle is already close to machine precision. This corresponds to 28 Fourier-Bessel terms around z_4, 14 Fourier-Bessel terms around z_1 and 29 Fourier-Bessel terms in the interior of the region, i.e. overall 71 terms, which saves 34 basis terms compared to the case that we approximate with the same number of terms around both singularities as shown in Figure 6.15.

Let us summarize the results of this chapter. For regions with zero or one singular corner we proved exponential convergence (Theorem 6.3.3 and 6.4.3). Moreover, by conformal mapping techniques we were able to determine the asymptotic exponential rate of convergence. For regions with multiple singularities the situation is different.

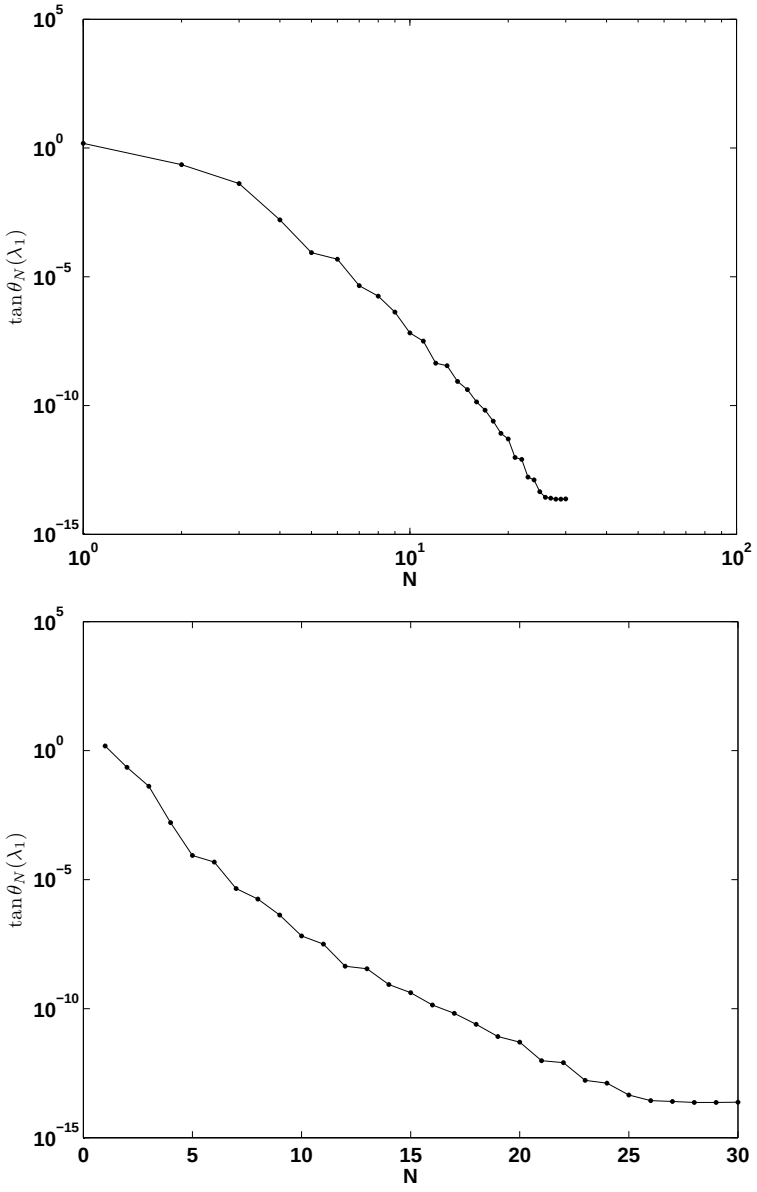


Figure 6.15: Convergence of the subspace angle method if in each step the number of terms at each singularity is also N. The upper plot shows the convergence behavior on a double-logarithmic scale while the lower plot uses a semi-logarithmic scale. The convergence appears to be faster than algebraic but slower than exponential.

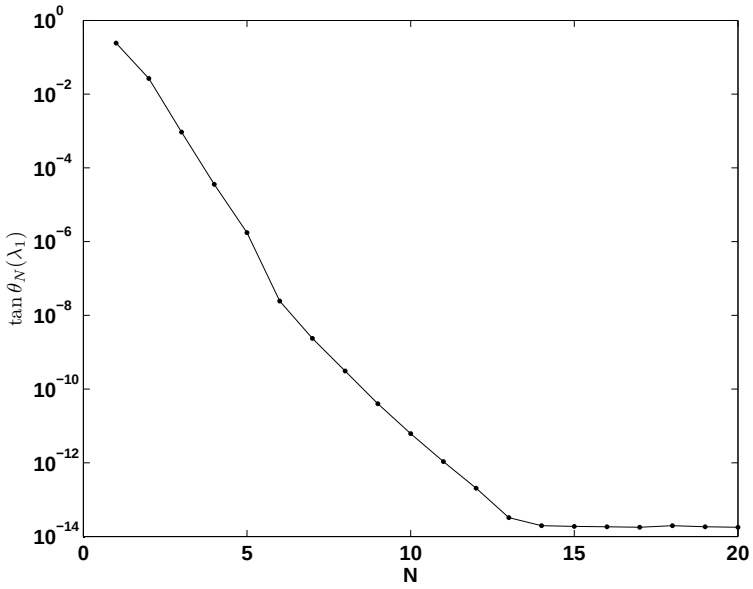


Figure 6.16: The rate of convergence if twice as many singular terms are used around z_4 as z_1.

Algebraic convergence rates for this case were previously analyzed by Eisenstat and Still and in the setting of Sobolev spaces by Melenk. By a suitable adjustment of the approximation spaces high algebraic convergence rates can be achieved. We demonstrated numerically that this can be further improved by approximating not only with an increasing number of Fourier-Bessel terms in the interior of the region but also with an increasing number of Fourier-Bessel terms at the singularities. The numerical convergence then seems better than algebraic but not yet exponential. This is plausible, since we increase the smoothness of the function we wish to approximate in each step by adding more singular terms. Therefore, the rate of algebraic convergence grows with increasing N. But since the singularities do not fully disappear we cannot expect exponential convergence. Nevertheless, the convergence can be made very fast even in the case of regions with multiple singular corners. The only restriction is that if a singular corner is not formed by two straight lines but by arbitrary analytic curves logarithmic terms can appear in the expansion of ϕ_k close to that corner. Just by using Fourier-Bessel functions we cannot capture those logarithmic terms and they can lead to a slow down of convergence.

6.6 A note on the convergence of eigenvalues

We conclude this chapter with a note on the rate of convergence of an eigenvalue approximation λ to an eigenvalue λ_k of (1.1) in a region Ω. From Chapter 5 we know that

$$\frac{|\lambda - \lambda_k|}{\lambda_k} \leq c \tan\theta(\lambda)$$

for a constant $c > 0$ that depends on Ω. Therefore, if λ is the minimum of the subspace angle curve, then

$$\frac{|\lambda - \lambda_k|}{\lambda_k} \leq c \tan\theta(\lambda) \leq c \tan\theta(\lambda_k)$$

and for a growing number of basis functions λ converges at least as fast as $\tan\theta(\lambda_k)$. But can the rate of convergence be faster than that of $\tan\theta(\lambda_k)$? If $\tilde{u}$ is an approximate eigenfunction from $\mathcal{D}_0$ instead of $\mathcal{A}(\lambda)$ then a good eigenvalue estimate is given by the Rayleigh quotient

$$\rho(\tilde{u}) = \frac{\langle \tilde{u}, -\Delta\tilde{u}\rangle_\Omega}{\langle \tilde{u}, \tilde{u}\rangle_\Omega}.$$

If the distance of $\tilde{u}$ to an eigenvector u_k is $O(\epsilon)$, then the distance of $\rho(\tilde{u})$ to λ_k is $O(\epsilon^2)$ leading to a squared convergence behavior for the eigenvalue approximations. The question is if such a "squared convergence" behavior also exists for the Method of Particular Solutions. The Rayleigh quotient does not give us any new information when we approximate from $\mathcal{A}(\lambda)$ since if $-\Delta u = \lambda u$, then $\rho(u) = \lambda$. However, there are several examples where we numerically observe faster convergence for the eigenvalue approximation than for the corresponding subspace angle. Figure 6.17 shows the convergence of the minimum of $\tan\theta_N(\lambda)$ for a growing number N of basis functions on the L-shaped region. The dotted curve shows the distance of the corresponding values λ to the first eigenvalue λ_1. Until about $N = 17$ the eigenvalue approximation λ converges faster than $\tan\theta_N(\lambda)$. This changes when $\tan\theta_N(\lambda) \approx 10^{-8}$ at $N = 17$. Then both curves seem to decrease with the same rate. Another striking example is given in Figure 6.18. It shows the convergence of the smallest subspace angle on the half annulus region from Figure 6.6. While the subspace angle converges smoothly the corresponding eigenvalue approximations first converge with a faster rate but then start oscillating and stagnate. We cannot yet explain the speedup of the eigenvalue convergence in these two examples. But the conclusion is that if we are only interested in a certain accuracy of an eigenvalue approximation then often we need fewer basis terms than predicted from our convergence theory for the subspace

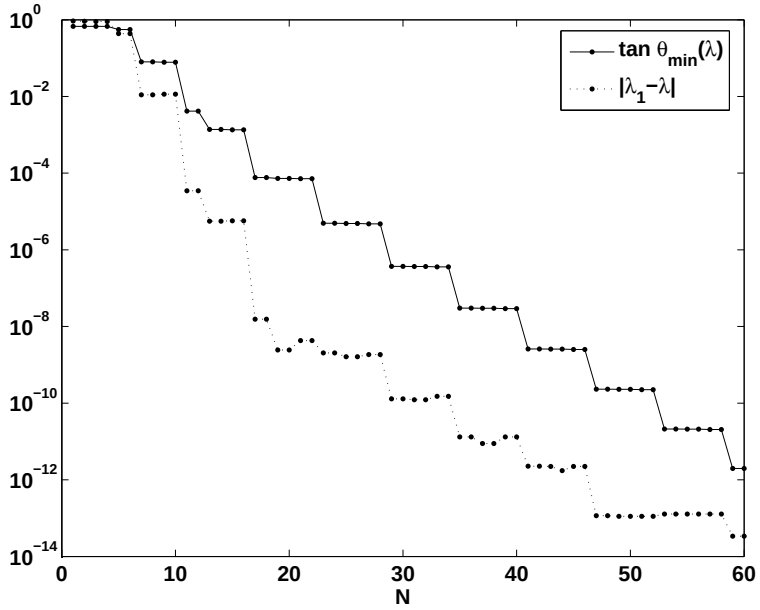


Figure 6.17: Comparison of the eigenvalue and subspace angle convergence on the L-shaped region

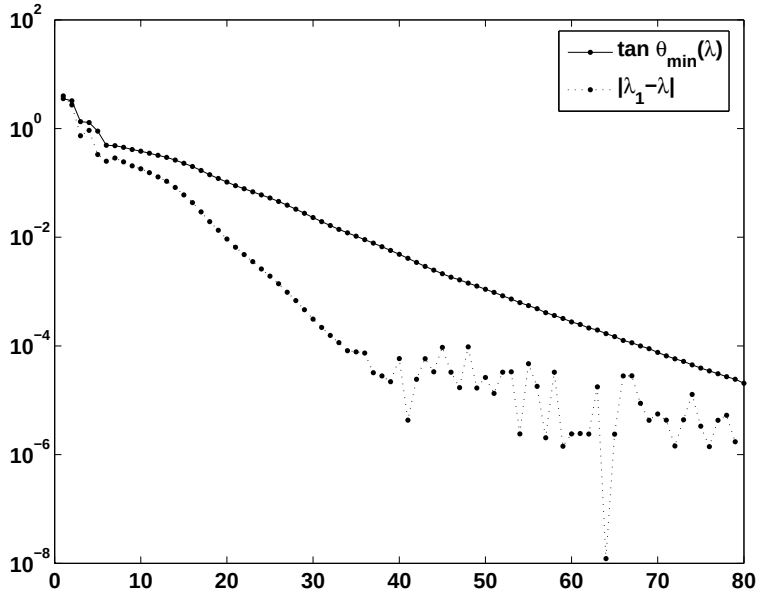


Figure 6.18: Comparison of the eigenvalue and subspace angle convergence on the half annulus region.

angle. Therefore, it is useful to look not only at the value of the subspace angle when observing the convergence of the subspace angle method for growing N but also to check the number of stable digits in the corresponding eigenvalue approximations, which in some cases can be significantly higher than predicted by the value of the subspace angle.

Chapter 7

A domain decomposition method based on the GSVD

The Method of Particular Solutions is a global approximation method in the sense that the basis functions in $\mathcal{A}(\lambda)$ live in the whole region Ω. With the right tools from linear algebra this idea can be turned into a stable and rapidly converging method, as we have seen in the previous chapters. We can prove exponential convergence for regions with zero or one corner singularity and also obtain arbitrarily fast algebraic rates for regions with more than one corner singularity.

An alternative to global approximation methods are domain decomposition methods that use particular solutions in each subdomain. This idea was introduced by Descloux and Tolley in 1983 [18]. Their method converges exponentially in arbitrary polygonal regions and also overcomes the ill-conditioning of the basis functions. In its original form, the accuracy of their method was limited to the square root of machine precision. Also, domain decomposition methods are more complicated to implement than global approximation methods. The problem of the limited accuracy was solved by Driscoll in 1997 [21], who computed the first 25 eigenmodes of the two isospectral drums to 12 digits of accuracy with this method. Instead of minimizing the smallest eigenvalue of a certain parameter-dependent eigenvalue problem Driscoll's improvement computes zeros of the derivative of an eigenvalue.

In this chapter we present a modification of the method of Descloux and Tolley based on the minimization of a generalized singular value. This approach provides another

solution of the square root of machine precision problem and has the additional advantages that it avoids computing L^2–inner products and evaluating derivatives of possibly ill-conditioned eigenvalues. Like that of Descloux and Tolley, our method is easily adapted to planar regions other than polygons.

Descloux and Tolley used Taylor series estimates to show the exponential convergence of their method. This has the disadvantage that it only works if an eigenfunction can be expanded into a convergent Taylor series on each subdomain. Also, Taylor series estimates only give optimal convergence rates on circles. We overcome these problems by using Vekua's theory and analytic continuation to establish exponential convergence rates. These estimates are asymptotically optimal and can be applied to a larger class of regions than Taylor series estimates.

7.1 The method of Descloux and Tolley and its reformulation as a GSVD problem

Let Ω be a polygonal region. Assume that Ω is partitioned into subregions $\Omega_1, \ldots, \Omega_p$ such that $\Omega_j \cap \Omega_l = \emptyset$ for $j \neq l$ and $\partial\Omega_j \cap \partial\Omega$ either contains no corner or consists of two straight arcs meeting at a corner of $\partial\Omega$. An example of such a decomposition is given in Figure 7.1 for the region described in Section 6.5. The internal boundary $\partial\Omega_j \cap \partial\Omega_l$ between two subdomains Ω_j and Ω_l is denoted by Γ_{jl}. If Γ_{jl} consists of only a finite number of points (as is for example the case for Γ_{13} in Figure 7.1) we set $\Gamma_{jl} := \emptyset$. The number of nonempty internal boundary segments Γ_{jl} is denoted by n. In Figure 7.1 we have $n = 4$. Let π/α_j be the interior angle of the corner of $\partial\Omega$ intersecting with $\partial\Omega_j$. If $\partial\Omega_j$ has no such corner we define a corner on $\partial\Omega \cap \partial\Omega_j$ with interior angle π. Therefore, we can assume from now on that every boundary segment $\partial\Omega_j$ contains a corner of $\partial\Omega$. Let z_j be the position of this corner in $\partial\Omega_j$. Around each corner z_j we can approximate an eigenfunction u_k of (1.1) with a Fourier-Bessel series of the form

$$f_{j,N_j}(r,\theta) = \sum_{i=1}^{N_j} a_i^{(j)} f_{j,N_j,i}, \tag{7.1}$$

with

$$f_{j,N_j,i}(r,\theta) = J_{\alpha_j i}(\sqrt{\lambda} r) \sin \alpha_j i\theta,$$

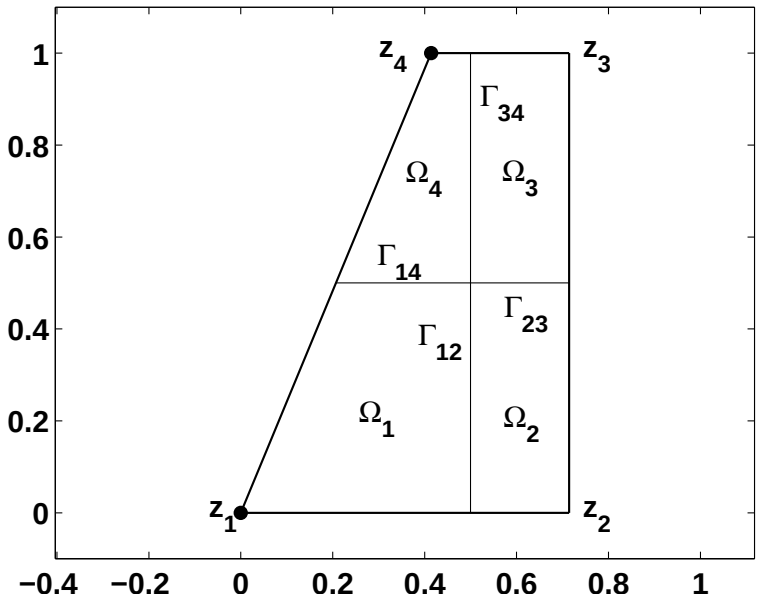


Figure 7.1: A domain decomposition for the method of Descloux and Tolley.

where the origin of the polar coordinates is z_j. Outside $\overline{\Omega_j}$ we define $f_{j,N_j,i}(r,\theta) := 0$, $i = 1,\ldots,N_j$ to restrict the support of f_{j,N_j} to $\overline{\Omega_j}$. On the boundary segments of Ω adjacent to z_k this expansion automatically satisfies the zero boundary conditions. But we need to satisfy compatibility conditions on the internal boundary segements Γ_{jl} between Ω_j and Ω_l. Let Γ_{jl} be a nonempty boundary segment. Then we need to satisfy

$$f_{j,N_j}(x,y) = f_{l,N_l}(x,y), \quad \nabla f_{j,N_j}(x,y) = \nabla f_{l,N_l}(x,y)$$

for $(x,y) \in \Gamma_{jl}$[1]. Let $a = (a_1^1,\ldots,a_{N_1}^{(1)},\ldots,a_1^{(p)},\ldots,a_{N_p}^{(p)})$ be the vector of all coefficients from the Fourier-Bessel expansions (7.1) with length $N := N_1 + \cdots + N_p$. Each such vector can be assigned a unique basis function from the space

$$\tilde{\mathcal{A}}_N(\lambda) := \left\{ f \in \mathcal{C}^2\left(\bigcup_{j=1}^{p} \Omega_j\right) : f|_{\Omega_j} = f_{j,N_j} \right\},$$

[1]It would be sufficient to demand that $\frac{\partial}{\partial n_j} f_{j,N_j}(x,y) + \frac{\partial}{\partial n_j} f_{j,N_j}(x,y) = 0$ for $(x,y) \in \Gamma_{jl}$, where $\frac{\partial}{\partial n_j}$ is the outward normal derivative on Ω_j. But since the effort of computing the normal derivative is essentially the same as that of computing the full derivative we will work with the full derivative as Descloux and Tolley did.

of all particular solutions of the method of Descloux and Tolley (on the internal boundary lines Γ_{jl} we let f undefined). We define the two quadratic forms

$$\begin{aligned}\mathcal{I}(\lambda,a) &= \sum_{j<l}\int_{\Gamma_{jl}} |f_{j,N_j}-f_{l,N_l}|^2+|\nabla f_{j,N_j}-\nabla f_{l,N_l}|^2 ds\\ \mathcal{K}(\lambda,a) &= \sum_{j=1}^{p}\int_{\Omega_j} |f_{j,N_j}|^2 d(x,y),\end{aligned}$$

where the dependence on λ comes from the Fourier-Bessel expansions f_{j,N_j} on λ. Both quadratic forms can be written as $\mathcal{I}(\lambda,a)=a^T I(\lambda)a$ and $\mathcal{K}(\lambda,a)=a^T K(\lambda)a$, where $I(\lambda)\in\mathbb{R}^{N\times N}$ is symmetric positive semi-definite and $K(\lambda)\in\mathbb{R}^{N\times N}$ is symmetric positive definite. Therefore, the minimum of $\mathcal{I}(\lambda,a)/\mathcal{K}(\lambda,a)$ over all $a\in\mathbb{R}^N$ is the smallest eigenvalue $\mu_1(\lambda)$ of the generalized eigenvalue problem

$$I(\lambda)x(\lambda)=\mu(\lambda)K(\lambda)x(\lambda). \tag{7.2}$$

Descloux and Tolley did not use the formulation as a generalized eigenvalue problem. By only evaluating $\mathcal{K}(\lambda,a)$ in a sector contained in each subdomain one can use the same trick as in (5.10) and obtain a diagonal right-hand side matrix $K(\lambda)$ whose diagonal elements are explicitly known. Therefore, it is easy to reduce (7.2) to the standard eigenvalue problem $K^{-1/2}IK^{-1/2}y(\lambda)=\mu_1(\lambda)y(\lambda)$. While the formulation as a generalized eigenvalue problem has the disadvantage that the two matrices $I(\lambda)$ and $K(\lambda)$ can have a common numerical null-space caused by linear dependencies of the basis functions on each subdomain, this problem is avoided in the formulation as a standard eigenvalue problem.

As in the method of Barnett, the problem with (7.2) is that a squaring is involved, which leads to a loss of accuracy. This effect was analyzed by Driscoll in [21]. Following an idea of Vavasis he repaired the method of Descloux and Tolley by finding the zero of the derivative $\mu_1'(\lambda)$ instead of minimizing the locally quadratic function $\mu_1(\lambda)$. By differentiating (7.2) with respect to λ and multiplying on the left by $x(\lambda)$ one obtains

$$\mu'(\lambda)=\frac{x(\lambda)(I'(\lambda)-\mu(\lambda)K'(\lambda))x(\lambda)}{x(\lambda)^T K(\lambda)x(\lambda)}.$$

Since $\mu'(\lambda)$ behaves linearly around a zero of $\mu(\lambda)$, the accuracy of solving $\mu'(\lambda)=0$ is comparable to the accuracy to which the values $\mu'(\lambda)$ can be determined. With this modified algorithm Driscoll computed the first 25 eigenvalues of the GWW-isospectral drums to 12 digits of accuracy. The disadvantage of this approach is that in addition

to the value $\mu(\lambda)$ we have to compute its derivative $\mu'(\lambda)$. Furthermore, we are still working with eigenvalue formulations, which can be ill-conditioned.

We have found that a solution of this problem problem is to reformulate (7.2) as a generalized singular value problem. We need the semi-norm

$$\|f\|_\Gamma := \left(\sum_{j<l} \int_{\Gamma_{j,l}} |f_{j,N_j}(s) - f_{l,N_l}(s)|^2 + |\nabla f_{j,N_j}(s) - \nabla f_{l,N_l}(s)|^2 ds \right)^{\frac{1}{2}}$$

and the norm

$$\|f\|_{\tilde{\Omega}} := \left(\sum_{j=1}^{p} \int_{\Omega_j} |f_{j,N_j}(x,y)|^2 dxdy \right)^{\frac{1}{2}} = \left(\sum_{j=1}^{p} \|f_{j,N_j}\|^2_{\Omega_j} \right)^{\frac{1}{2}}$$

for $f \in \tilde{\mathcal{A}}_N(\lambda)$. These are just the square roots of the quadratic forms $\mathcal{I}$ and $\mathcal{K}$. The semi-norm $\|\cdot\|_\Gamma$ is well defined since although f is not defined on Γ_{jl} the restriction f_{j,N_j} of f to the subdomain Ω_j is defined on Γ_{jl}. Now let

$$\sigma(\lambda) := \min_{u \in \tilde{\mathcal{A}}(\lambda) \setminus \{0\}} \frac{\|u\|_\Gamma}{\|u\|_{\tilde{\Omega}}}. \tag{7.3}$$

As in Section 3.5 we can discretize Ω and Γ_{jl} to turn (7.3) into a generalized singular value problem. Only the matrices will be slightly more complicated because of the structure of the norms used here. We discretize each Γ_{jl} by points $z_k^{(jl)} \in \Gamma_{jl}$, $k = 1, \ldots, M_{jl}$ and each subdomain Ω_j by points $\tilde{z}_t^{(j)} \in \Omega_j$, $t = 1, \ldots, L_j$. Therefore, we have $M = \sum_{j<l} M_{jl}$ discretrization points on the interior boundaries between the subdomains and $L = \sum_{j=1}^{p} L_j$ discretization points in the union of all subdomains. We now define the matrix $A_1(\lambda) \in \mathbb{R}^{3M,N}$ as

$$A_1(\lambda) = \begin{bmatrix} F(\lambda) \\ \nabla F(\lambda) \end{bmatrix},$$

where each row of $F(\lambda)$ is associated with one internal boundary collocation point z_k^{jl} on Γ_{jl} and defined as

$$[f_{1,N_1,1}(z_k^{jl}), \ldots, f_{1,N_1,N_1}(z_k^{jl}), \ldots, f_{j,N_j,1}(z_k^{jl}), \ldots, f_{j,N_j,N_j}(z_k^{jl}), \\ \ldots, -f_{l,N_l,1}(z_k^{jl}), \ldots, -f_{l,N_l,N_l}(z_k^{jl}), \ldots]$$

Since we restricted the support of a Fourier-Bessel basis function $f_{l,N_l,k}$ to $\overline{\Omega_l}$ all elements of the row which are not associated with basis functions in Ω_j or Ω_l are zero. The matrix ∇F is defined as

$$\nabla F = \begin{bmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \end{bmatrix}.$$

Hence, each boundary collocation point z_k^{jl} is assigned to two rows in ∇F. One row consists of the partial derivatives in the x-direction of the basis functions evaluated at z_k^{jl} and the other row consists of the corresponding derivatives in the y-direction evaluated at z_k^{jl}, where the derivatives belonging to basis functions in Ω_l are as in the definition of F multiplied by -1. Again, we set all elements in the rows associated with z_k^{jl} that do not belong to basis functions in Ω_j or Ω_l to zero. Furthermore, we define a matrix $A_2(\lambda)$ as

$$A_2(\lambda) = \begin{bmatrix} G_1 & & \\ & \ddots & \\ & & G_p \end{bmatrix},$$

where the ith column of G_j is the basis function $f_{j,N_j,i}$ evaluated at the points $\tilde{z}_t^{(j)} \in \Omega_j$, $t = 1, \ldots, L_j$. The definition of $\sigma(\lambda)$ in (7.3) now becomes

$$\hat{\sigma}(\lambda) = \min_{x \in \mathbb{R}^N \setminus \{0\}} \frac{\|A_1(\lambda)x\|_2}{\|A_2(\lambda)x\|_2}$$

and $\hat{\sigma}(\lambda)$ is just the smallest generalized singular value of the pencil $\{A_1(\lambda), A_2(\lambda)\}$.

As an example, let us do this for the region Ω shown in Figure 7.1 with the decomposition into four subdomains given there. Figure 7.2 shows the value $\hat{\sigma}(\lambda)$ for various values of λ. In each subdomain we used 10 Fourier-Bessel basis functions. All nonempty interior boundary lines Γ_{jl} were discretized with 50 equally spaced points and each subdomain Ω_i was discretized with 20 random interior points. The curve has two minima pointing to the first two eigenvalues.

In this example we used a division into four subdomains. But Ω only has two singular corners at z_1 and z_4. Therefore, we could attempt to only divide Ω into two subdomains such that each of the subdomains has one singular corner. For example, let Ω be subdivided by the straight line formed by Γ_{23} and Γ_{14}. Then we need to additionally impose the condition that the norm of the approximate eigenfunctions is be minimized on the line from z_2 to z_3. But this is easily accomplished similarly to the subspace angle method by discretizing the boundary line from z_2 to z_3 with collocation points and including in the matrix $A_1(\lambda)$ two blocks which consist of the basis functions around z_1 and z_4 evaluated at these additional boundary collocation points. Figure 7.3 shows the value $\hat{\sigma}(\lambda_1)$ for a growing number N of Fourier-Bessel basis functions in each subdomain, where $\lambda_1 \approx 48.4161682676614$ is the first eigenvalue of (1.1) on Ω. In the case of two subdomains we have an accuracy close to

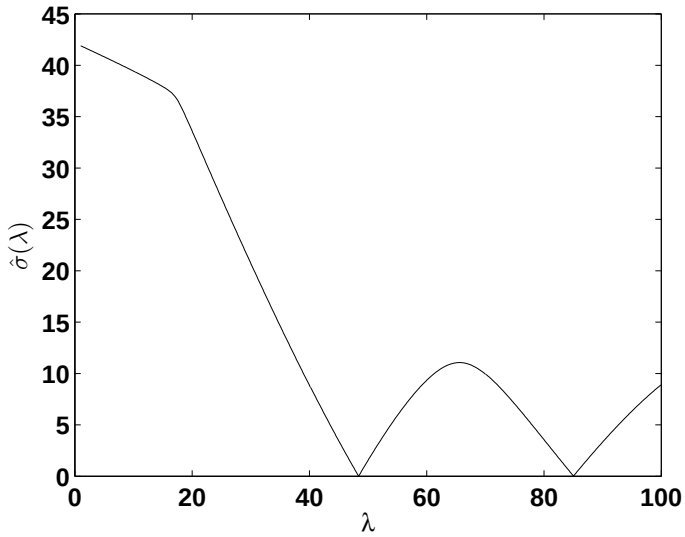


Figure 7.2: The domain decomposition GSVD method on a quadrilateral with 2 singular corners.

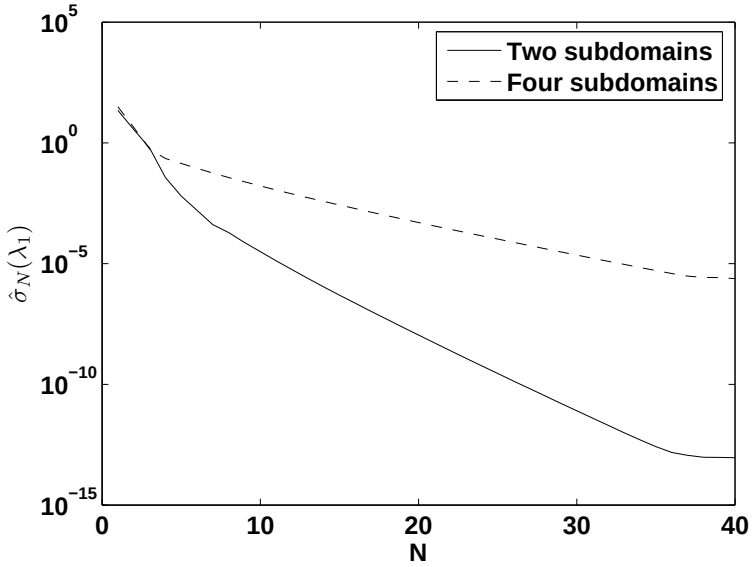


Figure 7.3: Comparison of the convergence of $\hat{\sigma}(\lambda_1)$ for a growing number N of basis functions in each subdomain in the cases of division into two subdomains and four subdomains.

machine precision after $N = 40$ steps, while in the case of four subdomains we have an accuracy of only 10^{-5} at $N = 40$. The reason for this slow convergence in the case of four subdomains is that the singularity of the eigenfunction at z_4 is very close Ω_3 slowing down the convergence there. Hence, in the case of four subdomains not only do we need more basis functions but the convergence rate is slower.

A striking feature of both curves is that in contrast to global approximations whose convergence on this region was discussed in Section 6.5, we seem to observe exponential convergence. In the next section we prove that the domain decomposition method indeed converges exponentially and compute the asymptotic rate for the solid curve in Figure 7.3.

7.2 Exponential convergence of the domain decomposition method

Based on Taylor series estimates Descloux and Tolley proved exponential convergence of their domain decomposition method. However, their estimates have two (related) disadvantages. First, we need to guarantee that the subdomains are chosen such that the eigenfunction has a convergent Taylor series on each subdomain. Second, Taylor series estimates only deliver the true asymptotic convergence on circles. For other regions Taylor polynomials are not optimal and therefore do not deliver good bounds for the asymptotic convergence rate. The results in this section are illustrated for the region Ω from Figure 7.1 using a subdivision into two subdomains with internal boundary $\Gamma := \Gamma_{12} \cup \Gamma_{14}$. Let us denote by $\hat{\Omega}_1$ the subdomain below Γ and by $\hat{\Omega}_2$ the subdomain above Γ. Since with only two subdomains we also have to minimize the error of the approximate eigenfunctions on the right boundary segment from z_2 to z_3, we introduce a slightly different quadratic form $\tilde{\mathcal{I}}(\lambda, a)$ defined as

$$\begin{aligned}\tilde{\mathcal{I}}(\lambda, a) &= \sum_{j<l} \int_{\Gamma_{jl}} |f_{j,N_j}(s) - f_{l,N_l}(s)|^2 + |\nabla f_{j,N_j}(s) - \nabla f_{l,N_l}(s)|^2 ds \\ &+ \sum_{j=1}^{p} \int_{\partial\Omega \cap \partial\Omega_j} |f_{j,N_j}(s)|^2 ds.\end{aligned}$$

If the domain decomposition of Descloux and Tolley is used, the last sum of $\tilde{\mathcal{I}}(\lambda, a)$ is always zero and we have $\tilde{\mathcal{I}}(\lambda, a) = \mathcal{I}(\lambda, a)$. Therefore, this slightly more general

approach includes the original method of Descloux and Tolley as a special case. For each λ the method computes in the non-sampled case

$$\sigma(\lambda) = \min_{a \in \mathbb{R}^N \{0\}} \frac{\tilde{\mathcal{I}}^{1/2}(\lambda, a)}{\mathcal{K}^{1/2}(\lambda, a)}.$$

Let (λ_k, u_k) be an eigenpair of (1.1) on Ω. We can estimate $\tilde{\mathcal{I}}(\lambda, a)$ as

$$\begin{aligned}
\tilde{\mathcal{I}}(\lambda_k, a) &= \sum_{j<l} \int_{\Gamma_{jl}} |f_{j,N_j}(s) - f_{l,N_l}(s)|^2 + |\nabla f_{j,N_j}(s) - \nabla f_{l,N_l}(s)|^2 ds \\
&+ \sum_{j=1}^{p} \int_{\partial\Omega \cap \partial\Omega_j} |f_{j,N_j}(s)|^2 ds \\
&\leq C_1 \sum_{j<l} \Big[\|f_{j,N_j} - u_k\|^2_{\infty,\Gamma_{jl}} + \|u_k - f_{l,N_l}\|^2_{\infty,\Gamma_{jl}} + \|\nabla f_{j,N_j} - \nabla u_k\|^2_{\infty,\Gamma_{jl}} \\
&+ \|\nabla u_k - \nabla f_{l,N_l}\|^2_{\infty,\Gamma_{jl}} \Big] + C_2 \sum_{j=1}^{p} \|f_{j,N_j} - u_k\|^2_{\infty,\partial\Omega\cap\partial\Omega_j} \\
&\leq C_1 \sum_{j<l} \|\nabla f_{j,N_j} - \nabla u_k\|^2_{\infty,\Gamma_{jl}} + \|\nabla f_{l,N_l} - \nabla u_k\|^2_{\infty,\Gamma_{jl}} \\
&+ C_3 \sum_{j=1}^{p} \|u_k - f_{j,N_j}\|^2_{\infty,\overline{\Omega_j}}, \qquad (7.4)
\end{aligned}$$

where $C_1, C_2, C_3 > 0$ are constants which depend on Ω. Hence, we need to estimate the rate of convergence of the functions f_{j,N_j} to u_k restricted to $\overline{\Omega_j}$ and the rate of convergence of the derivatives of f_{j,N_j} to the derivatives of u_k on the internal boundary lines Γ_{jl}. Take for example the subdomain $\hat{\Omega}_1$ from the region in Figure 7.2. The only singularity of u_k in $\overline{\hat{\Omega}_1}$ is at $z_1 = 0$. Let ϕ_k be the holomorphic function associated with u_k such that $u_k = \mathrm{Re}\{V[\phi_k; z_1]\}$. From Lemma 6.4.1 it follows that

$$\phi_k(z) = \sum_{k=1}^{\infty} i c_k z^{\frac{8k}{3}}, \quad c_k \in \mathbb{R}$$

close to z_1, and using Lemma 6.4.2, we know that the function $\tilde{\phi}_k(w) := \phi_k(z)$ for $w = z^\alpha$ is holomorphic in a neighborhood around z_1. Therefore, $\tilde{\phi}_k$ is holomorphic on $\overline{\hat{\Omega}_1^{8/3}}$. We can now proceed exactly as in the example of the L-shaped region in Section 6.4. The two closest singularities to $\hat{\Omega}_1$ in the conformal sense are z_4 and the point $z_1' = \frac{2}{\tan\frac{3}{8}\pi} + 0.6$. The singularity at z_1' is obtained by reflection of the eigenfunction u_k at the boundary line connecting z_2 and z_3. Since we only use Fourier-Bessel sine functions we have to reflect the region $\overline{\hat{\Omega}_1^{8/3}}$ across the real line before computing the

mapping to the exterior of the unit disk. The conformal distances of $z_4^{8/3}$ and $z_1'^{8/3}$ to this reflected region are then approximately given as 2.82 and 5.34. Therefore, we obtain

$$\min_{u \in \tilde{\mathcal{A}}_N(\lambda_k)} \|u_k - u\|_{\infty,\overline{\hat{\Omega}_1}} = O(2.82^{-N_1}),$$

if N_1 is the number of Fourier-Bessel basis terms in $\hat{\Omega}_1$. Similarly, for approximating with Fourier-Bessel functions around z_4 in the region $\hat{\Omega}_2$ we obtain

$$\min_{u \in \tilde{\mathcal{A}}_N(\lambda_k)} \|u_k - u\|_{\infty,\overline{\hat{\Omega}_2}} = O(1.86^{-N_2}),$$

where the conformally closest singularity is $z_4' = \frac{1}{\tan\frac{8}{3}\pi} + 0.6$, which is the image of the singularity at z_4 under reflection of u_k at the boundary line from z_2 to z_3. Combining these results we obtain

$$\min_{u \in \tilde{\mathcal{A}}_N(\lambda_1)} \|u_k - u\|_{\infty,\overline{\Omega}} = O(1.86^{-N})$$

if $N := N_1 = N_2$. On arbitrary polygonal regions we obtain similarly the following result.

Lemma 7.2.1 *There exist numbers $R_j > 1$, $j = 1, \ldots, p$ and functions $u_N \in \mathcal{A}(\lambda_k)$ such that for their restrictions $u_N|_{\Omega_j} = f_{j,N_j}$ to Ω_j it holds that*

$$\|f_{j,N_j} - u_k\|_{\infty,\overline{\Omega_j}} = O(R_j^{-N_j})$$

as $N_j \to \infty$.

Proof The proof proceeds exactly as in the example given above. It is only essential that u_k has at most one corner singularity in each subdomain $\overline{\Omega_j}$, which is guaranteed by the domain decomposition. ∎

It is now left to estimate the convergence of the derivatives of the Fourier-Bessel basis functions to ∇u_k on the internal boundary line Γ. We need the following technical lemma.

Lemma 7.2.2 *Let $\{f^{(N)}\}$ be a sequence of real analytic functions defined in the interval $I := [a, b]$ and having a holomorphic continuation to the region $\Omega \subset \mathbb{C}$. If $\|f^{(N)}\|_{\infty,I} = O(R^{-N})$ for $N \to \infty$ then $\|f'^{(N)}\|_{\infty,I} = O((R-\delta)^{-N})$ for every $\delta > 0$.*

Proof Fix $\delta > 0$. Then there exists a neighborhood S of I in Ω such that $\|f^{(N)}\|_{\infty,S} = O((R-\delta)^{-N})$ for $N \to \infty$.[2] Now choose $\epsilon > 0$ such that $K_\epsilon := \{z : |z - x_0| = \epsilon,\ z \in \mathbb{C}\} \subset S$ for all $x_0 \in I$. Then

$$f'^{(N)}(x_0) = \frac{1}{2\pi i} \int_{K_\epsilon} \frac{f^{(N)}(\xi)}{(\xi - x_0)^2} d\xi$$

for $x_0 \in I$. It follows that

$$|f'^{(N)}(x_0)| \leq \frac{1}{\epsilon} \|f^{(N)}\|_{\infty,S} = O((R-\delta)^{-N})$$

and therefore

$$\|f'^{(N)}\|_{\infty,I} = O((R-\delta)^{-N}) \tag{7.5}$$

for every $\delta > 0$. ∎

We are now able to estimate the expression $\|\nabla f_{j,N_j} - \nabla u_j\|_{\infty,\Gamma_{jl}}$. By Lemma 7.2.1 there exists a sequence $a^{(N)} \in \mathbb{R}^N$ and associated basis functions f_{j,N_j}, $j = 1, \ldots, p$, $N = \sum_{j=1}^{p} N_j$ such that

$$\|f_{j,N_j} - u_k\|_{\infty,\Gamma_{j,l}} = O(R_j^{-N_j})$$

for all $j = 1, \ldots, p$. Now fix $\delta_1 > 0$. Then there exists a region D such that $\Gamma_{jl} \subset D$ and

$$\|f_{j,N_j} - u_k\|_{\infty,D} = O((R_j - \delta_1)^{-N_j}).$$

Let $(x_0, y_0) \in \Gamma_{jl}$ and define $\hat{u}_{N_j}(x) := f_{j,N_j}(x, y_0) - u_k(x, y_0)$ in a small interval $I := [x_0 - \epsilon, x_0 + \epsilon]$ for an $\epsilon > 0$ such that $I \times y_0 \in D$. From Theorem 6.1.2 it follows that $\hat{u}_{N_j}$ can be continued to a holomorphic function in a neighborhood of I independent of N_j. We can now use Lemma 7.2.2 and find that for every $\delta_2 > 0$

$$\|\hat{u}'_{N_j}\|_{\infty,I} = O((R_j - \delta_1 - \delta_2)^{-N_j}),$$

which implies that

$$\left| \frac{\partial}{\partial x} f_{j,N_j}(x_0, y_0) - \frac{\partial}{\partial x} u_k(x_0, y_0) \right| = O((R_j - \delta)^{-N_j})$$

for $\delta = \delta_1 + \delta_2$. Similarly, we obtain

$$\left| \frac{\partial}{\partial y} f_{j,N_j}(x_0, y_0) - \frac{\partial}{\partial y} u_k(x_0, y_0) \right| = O((R_j - \delta)^{-N_j}).$$

[2] In the case of polynomial approximation this is called overconvergence, the effect that approximations to analytic functions in a region Ω also converge in a neighborhood of Ω if the function is analytic there (see [84], §4.6–4.7).

Since (x_0, y_0) was chosen arbitrarily on Γ_{jl} we find

$$\|\nabla f_{j,N_j} - \nabla u_k\|_{\infty,\Gamma_{jl}} = O((R-\delta)^{-N_j})$$

for every $\delta > 0$. Combining this result with (7.4) yields

$$\tilde{\mathcal{I}}(\lambda_k, a^{(N)}) = \sum_{j=1}^{p} O((R_j - \delta)^{-2N_j}) \tag{7.6}$$

for every $\delta > 0$ and $N_j \to \infty$. R_j is the exponential rate of convergence of approximating u_k on the subdomain Ω_j. If $N_1 = \cdots = N_p =: \tilde{N}$ and $\tilde{R} = \min_j R_j$, then

$$\tilde{\mathcal{I}}(\lambda_k, a^{(p\tilde{N})}) = O((\tilde{R} - \delta)^{-2\tilde{N}}).$$

Let us now estimate $\mathcal{K}(\lambda_k, a^{(N)})$. Assume that $\|u_k\|_\Omega = 1$. Then

$$\begin{aligned}
\mathcal{K}(\lambda_k, a^{(N)}) &= \sum_{j=1}^{p} \|f_{j,N_j}\|_{\Omega_j}^2 \\
&\geq \sum_{j=1}^{p} \left[\|u_k\|_{\Omega_j} - \|f_{j,N_j} - u_k\|_{\Omega_j}\right]^2 \\
&\geq \sum_{j=1}^{p} \left[\|u_k\|_{\Omega_j} - C\|f_{j,N_j} - u_k\|_{\infty,\overline{\Omega_j}}\right]^2 \\
&\geq \sum_{j=1}^{p} \left[\|u_k\|_{\Omega_j} - O(R_j^{-N_j})\right]^2 \to \sum_{j=1}^{p} \|u_k\|_{\Omega_j}^2 = 1
\end{aligned} \tag{7.7}$$

for a constant $C > 0$ which depends on Ω. Combining all results we obtain the following theorem which establishes exponential convergence rates for the domain decomposition method.

Theorem 7.2.3 *Let (λ_k, u_k) be an eigenpair of (1.1) with $\|u_k\|_\Omega = 1$. Then*

$$\min_{a \in \mathbb{R}^N} \frac{\tilde{\mathcal{I}}(\lambda_k, a)}{\mathcal{K}(\lambda_k, a)} = \sum_{j=1}^{p} O((R_j - \delta)^{-2N_j})$$

for every $\delta > 0$ and $N_j \to \infty$, $j = 1, \ldots, p$. The numbers R_j are the exponential convergence rates from Lemma 7.2.1 for approximating u_k on $\overline{\Omega_j}$ with functions in $\tilde{\mathcal{A}}_N(\lambda_k)$.

Proof Let $0 < \epsilon < 1$. From (7.7) it follows that there exists N_0 such that $\mathcal{K}(\lambda, a^{(N)}) \geq 1 - \epsilon$ for $N_j > N_0$, $j = 1, \ldots, p$. Together with (7.6) we find

$$\min_{a \in \mathbb{R}^N} \frac{\tilde{\mathcal{I}}(\lambda_k, a)}{\mathcal{K}(\lambda_k, a)} \leq \frac{\tilde{\mathcal{I}}(\lambda_k, a^{(N)})}{\mathcal{K}(\lambda_k, a^{(N)})} \leq \frac{1}{1-\epsilon} \tilde{\mathcal{I}}(\lambda_k, a^{(N)}) = \sum_{j=1}^{p} O((R_j - \delta)^{-2N_j}).$$

for $N_j \to \infty$, $j = 1, \ldots, p$. ∎

For our example domain Ω of Figure 7.1 we computed the rates of convergence R_1 and R_2 on the two subdomains $\hat{\Omega}_1$ and $\hat{\Omega}_2$ in this section. They were $R_1 \approx 2.82$ and $R_2 \approx 1.86$. If we use the same number N of Fourier-Bessel basis functions on both subdomains it follows from Theorem 7.2.3 that

$$\sigma(\lambda) = \min_{a \in \mathbb{R}^{2N}} \frac{\tilde{\mathcal{I}}^{1/2}(\lambda_k, a)}{\mathcal{K}^{1/2}(\lambda_k, a)} = O((1.86)^{-N}).$$

Figure 7.4 shows the measured convergence of

$$\hat{\sigma}(\lambda) \approx \sigma(\lambda)$$

for a growing number of basis functions. This time we did not use the fixed value $\lambda = \lambda_1$ but the minimum of the curve of $\hat{\sigma}(\lambda)$. The dotted line is the convergence behavior of the position λ of the minimum of the curve to the first eigenvalue λ_1 and the dashed line is the estimated rate 1.86^{-N}. In this plot we used the same number N of basis functions on $\hat{\Omega}_1$ and $\hat{\Omega}_2$. But the rate of convergence on $\hat{\Omega}_1$ is approximately 2.82^{-N} and on $\hat{\Omega}_1$ it is 1.86^{-N}. To balance these different convergence rates we can use different numbers of basis functions on the two subdomains as in Section 6.5. We want to achieve $2.82^{-N_1} = 1.86^{-N_2}$ which results in $\frac{N_2}{N_1} = \frac{\log 2.82}{\log 1.86} \approx 1.67$. Therefore, it is more suitable to use $3N$ basis functions on $\hat{\Omega}_2$ and $2N$ basis functions on $\hat{\Omega}_1$. The resulting convergence curve is plotted in Figure 7.5. The curve for $\hat{\sigma}(\lambda)$ reaches its minimum in Figure 7.5 at $N = 13$, which corresponds to 65 Fourier-Bessel basis functions, while the minimum in Figure 7.4 is reached at $N = 38$ corresponding to 76 basis functions.

Let us compare the convergence rates computed in this chapter with the convergence estimates of Descloux and Tolley. The radius ρ_1 of $\hat{\Omega}_1$ is

$$\rho_1 = \max_{z \in \hat{\Omega}_1} |z - z_1| \approx 0.87,$$

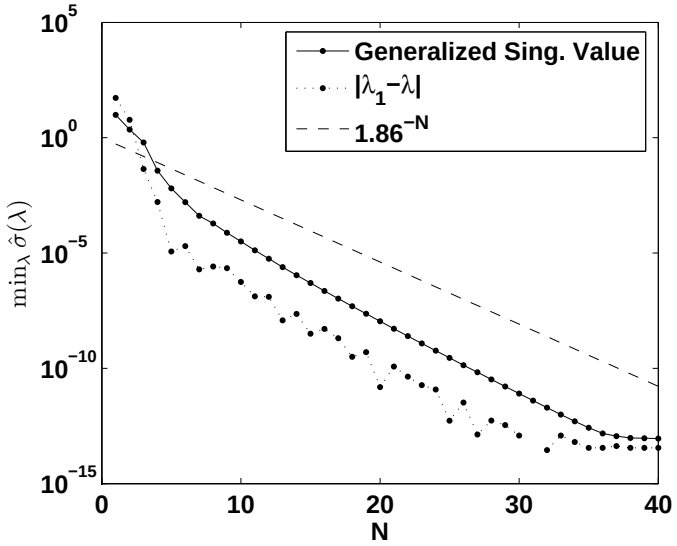


Figure 7.4: Comparison of estimated and measured convergence for the domain decomposition method

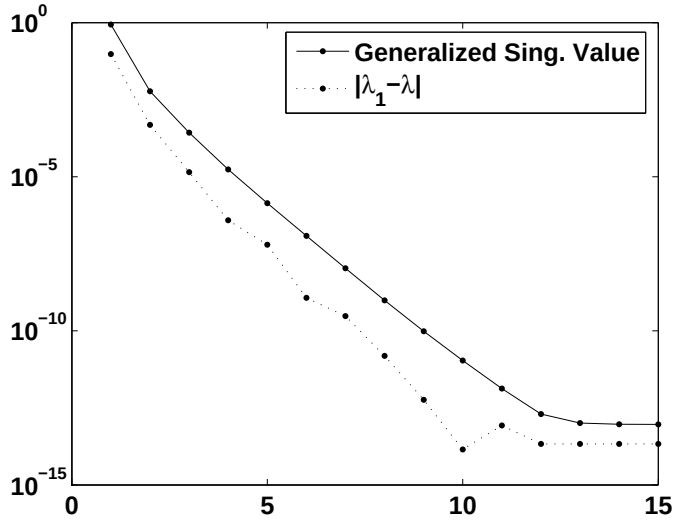


Figure 7.5: Convergence of the domain decomposition method with $2N$ basis functions on $\hat{\Omega}_1$ and $3N$ basis functions on $\hat{\Omega}_2$.

while the closest singularity is z_4 with $|z_4| \approx 1.08$. Using Taylor series estimates as Descloux and Tolley did we find $(\frac{|z_4|}{\rho_1})^{8/3} \approx 1.78$ for the exponential rate of convergence on $\hat{\Omega}_1$, while we computed an exponential rate of $R_1 \approx 2.82$ on this subdomain. Similarly, Taylor series estimates deliver an exponential rate of 1.05 compared to our computed value of 1.86 on $\hat{\Omega}_2$. Hence, by just using Taylor series estimates we obtain a convergence estimate that does not have anything to do with the true convergence behavior, while our value comes close to the slope of the observed curve in Figure 7.4 and is asymptotically correct for $N \to \infty$.

When is it preferable to use a domain decomposition method and when should we use global approximations as in the subspace angle method? The obvious advantage of a method based on global approximations is its lower programming effort. Furthermore, using a bad domain decomposition can considerably slow down the convergence, as shown in Figure 7.3. But finding an optimal domain decomposition is a nontrivial task, if possible at all. In Section 6.5 we needed at least 71 basis functions to obtain a smallest generalized singular value close to machine precision. In the domain decomposition method presented here this was achieved by using 65 basis functions (the case $N = 13$ in Figure 7.5). However, in the domain decomposition method we also have to compute derivatives of the basis functions, resulting in an overall higher computational effort. The theoretical advantage of domain decomposition methods is that they converge exponentially on polygonal regions. But as shown in the example presented here, this does not necessarily mean that the computational effort is lower to obtain an accuracy close to machine precision.

The picture looks different for multiply connected regions. Consider the region Ω shown in Figure 7.6. This region has four singular corners with interior angles $\frac{3\pi}{2}$. Fourier-Bessel functions to capture these singularities are of the form $J_{\frac{2\pi}{3}k}(\sqrt{\lambda}r)\sin\frac{2\pi}{3}k\theta$, resulting in branch lines that always intersect the region. Therefore, we cannot use global basis functions that are adapted to the singularities. However, by using a domain decomposition, we can divide Ω into four simply connected subdomains, on which it is possible to use basis functions that are adapted to the singularities, making accurate eigenvalue computations possible. In Chapter 8 we compute some of the eigenvalues and eigenfunctions of the region in Figure 7.6 using the domain decomposition GSVD method.

Another application where domain decomposition methods are of advantage is if

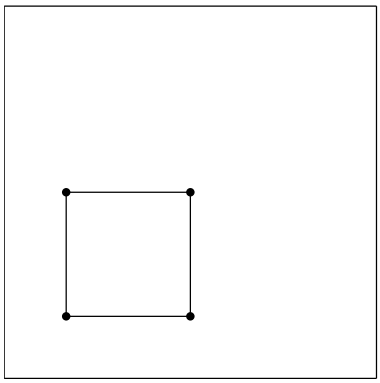


Figure 7.6: A multiply connected region with four singular corners.

we have several subdomains which are only weakly linked and we have additional information about the eigenfunctions in each subdomain. Then this approach can be used to quickly find good approximations for eigenfunctions in the whole region.

Chapter 8

Examples of computed eigenvalues and eigenfunctions

In this chapter we present accurate computations of eigenvectors and eigenfunctions of several different regions. Most of the examples are computed with the subspace angle method as described in Chapter 3. If we use a different approach like the domain decomposition GSVD we state it in the corresponding section. For each eigenfunction we print all digits that we believe to be correct. The plotted eigenfunctions are normalized such that their maximum absolute value is one and we always plot black level curves which go from -0.9 to 0.9 in steps of 0.2. Further examples including unbounded regions can be found in [75]. Section 8.5 was also published in [14].

8.1 The L-shaped region

The L-shaped region was the central example in the paper by Fox, Henrici and Moler in 1967. At the end of the 1970's Moler used it to demonstrate the power of his new computer numerical system MATLAB. Also, every numerical analyst will have seen the MATLAB logo, which is a variant of the first eigenfunction of the L-shaped region. After using this famous region in all chapters of the thesis we finally want to show some of its eigenfunctions computed with the subspace angle method. In Figure 8.1 we show six eigenfunctions of the L-shaped region. The 3rd and the 104th eigenvalue are special. They are also eigenvalues of the unit square and have the exact values

$\lambda_3 = 2\pi^2$ and $\lambda_{104} = 50\pi^2$. The 104th eigenvalue is especially interesting as it is a triple degeneracy, i.e. $\lambda_{103} = \lambda_{104} = \lambda_{105}$. This comes from the fact that

$$50 = 5^2 + 5^2 = 1^2 + 7^2 = 7^2 + 1^2$$

and therefore there exist three linearly independent eigenfunctions $\sin(5\pi x)\sin(5\pi y)$, $\sin(\pi x)\sin(7\pi x)$, and $\sin(7\pi x)\sin(\pi y)$. The numbers of the eigenvalues are obtained by counting the minima of the subspace angle curve. To count multiple eigenvalues correctly we also need the higher subspace angles from Definiton 3.2. The idea is that if we have a double eigenvalue the smallest and the second smallest subspace angle will go to zero since we have a two-dimensional eigenspace and therefore also a two-dimensional intersection between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$. In Figure 8.2 some higher subspace angle curves around the value $\lambda = 50\pi^2$ are plotted. At the triple degeneracy the curves for the smallest three subspace angles go to zero.

8.2 The circular L region

Let us now have a look at a slight variation of the L-shaped region. Instead of the reentrant corner we have a quarter circle of radius one. The asymptotic rate of convergence of the MPS on this region was computed in Section 6.3. Figure 8.3 shows some of the eigenvalues and eigenfunctions of this region. As far as we are aware of there are no degenerate eigenvalues any longer on this region.

8.3 Symmetric and unsymmetric dumbbells

In this section we compare the eigenvalues and eigenfunctions of two dumbbell shapes shown in Figure 8.4. The left dumbbell consists of two squares of side length π which are coupled by a bridge of length and width $\frac{\pi}{4}$. In the unsymmetric dumbbell the side length of the right square is reduced from π to 0.9π. Let us first discuss the symmetric dumbbell. Without the connecting bridge the region would consist of two squares, each with eigenvalues

$$i^2 + j^2, \quad i, j = 1, 2, \ldots.$$

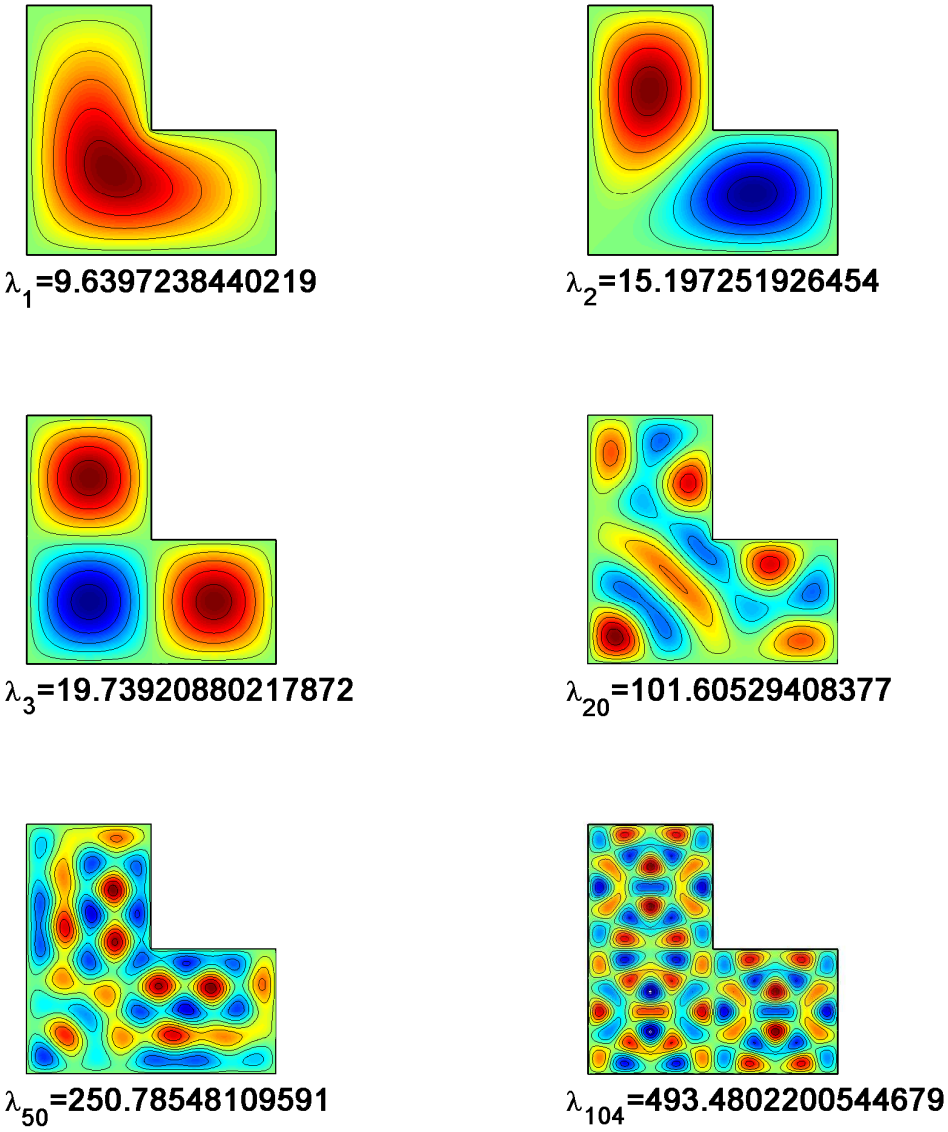


Figure 8.1: Some eigenfunctions of the L-shaped region.

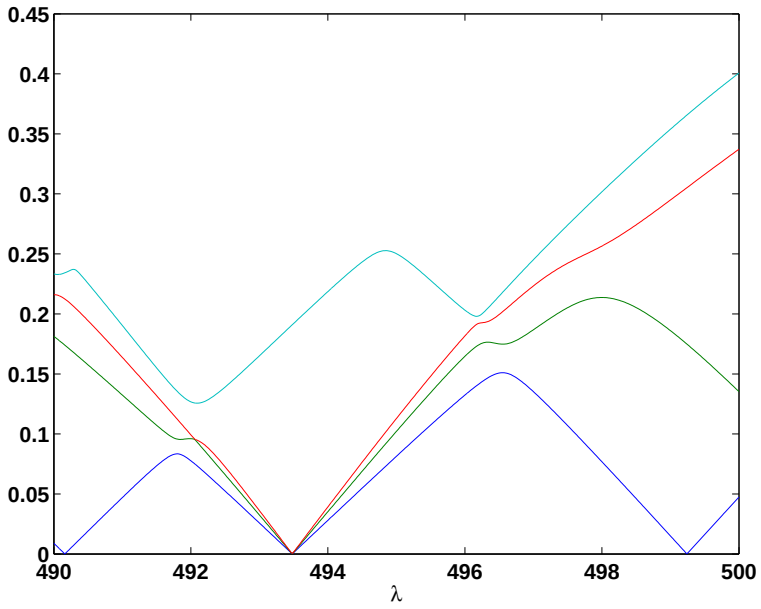


Figure 8.2: Sines $s_1(\lambda), \ldots, s_4(\lambda)$ of the subspace angles $\theta_1(\lambda), \ldots, \theta_4(\lambda)$ between $\mathcal{A}(\lambda)$ and $\mathcal{D}_0$ for the L-shaped region. At $\lambda = 50\pi^2$ there is a triple degeneracy.

Hence, the eigenvalues of the two unconnected squares are

$$2, 2, 5, 5, 5, 5, 8, 8, 10, 10, 10, 10, \ldots.$$

By introducing the connection between the squares we obtain a simply connected region with lower eigenvalues and with broken degeneracies. But we can expect that the first eigenvalues of the connected region will be close to the first eigenvalues of the two unconnected squares. Some of the eigenfunctions of the symmetric dumbbell are plotted in Figure 8.5.

If we break the symmetry the eigenfunctions will change dramatically. Assume that the bridge between the squares at the unsymmetric dumbbell does not exist. Then the eigenvalues of the left square are different from the eigenvalues of the right square. Hence, an eigenfunction of the region consisting of both squares is always zero on one of the squares. If we introduce the connection between the two squares we can expect that for example the first eigenfunction on this region will be small in the right square since without the bridge it would be zero there. Correspondingly the eigenfunction belonging to the second eigenvalue will be small on the left square. Hence, the eigenfunctions belonging to smaller eigenvalues become localized due to the small perturbation that destroys the symmetry of the dumbbell. Only for higher eigenvalues can we expect global eigenfunctions to occur since then the local wavelength of an

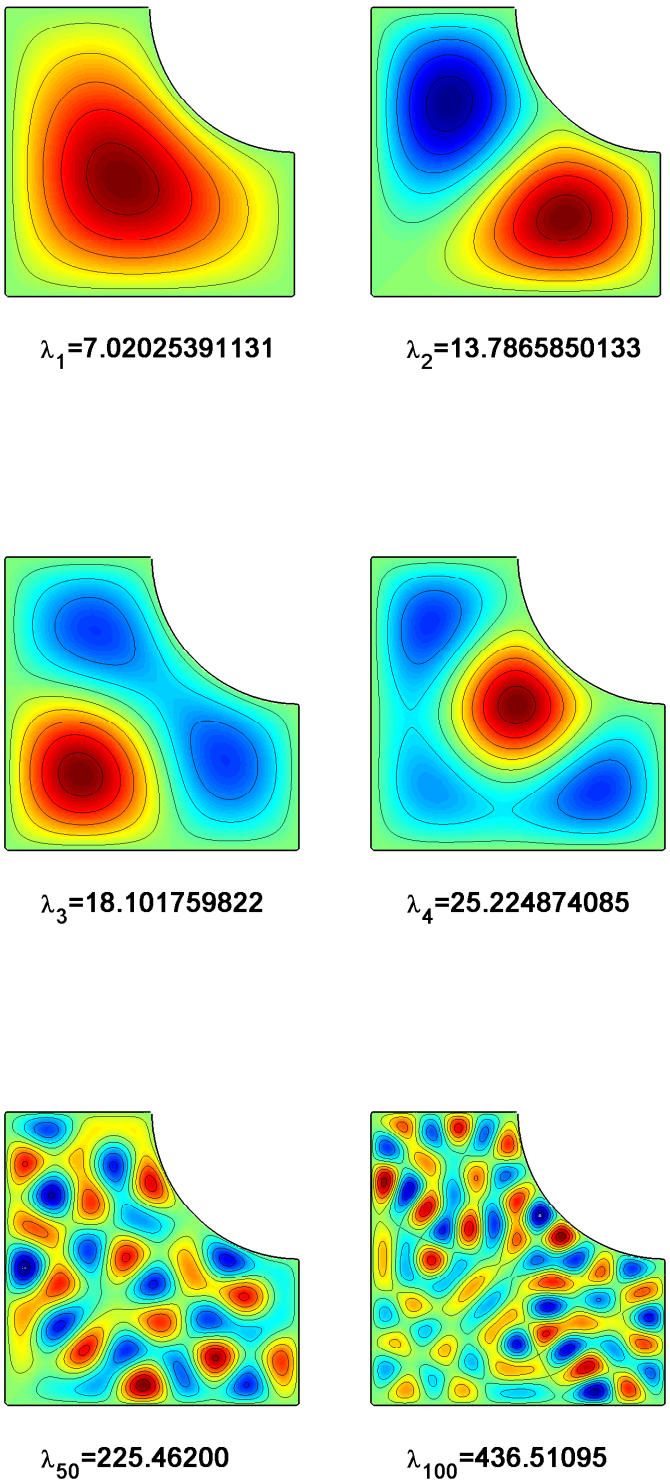


Figure 8.3: Some eigenfunctions of the circular L region.

Figure 8.4: A symmetric and a nonsymmetric dumbbell region.

eigenfunction becomes smaller than the width of the bridge. Some eigenfunctions of the unsymmetric dumbbell are plotted in Figure 8.6.

Figure 8.7 compares the subspace angle curves for the symmetric and the nonsymmetric dumbbell. The blue and the green curve for the smallest and second smallest subspace angle are almost identical since smaller eigenvalues are clustered in pairs. Close to $\lambda = 5$ all four subspace angle curves become small, indicating a cluster of four eigenvalues. In the unsymmetric curve the eigenvalues are more separated and we can observe some interesting features of the subspace angle curves. Consider for example the first two eigenvalues. The blue curve for the smallest subspace angle has minima close to these eigenvalues. But between them the blue curve goes up again and almost crosses the green curve belonging to the second smallest subspace angle. In the figure it seems that the curves even cross. Only by zooming into the graph does it become visible that the blue and the green curve come close between $\lambda = 2$ and $\lambda = 3$ but avoid each other. These avoided crossings were investigated by Barnett in [8]. It is interesting to look at the approximate eigenfunctions belonging to the smallest and the second smallest subspace angle before and after the avoided crossing. This is shown in Figure 8.8 and 8.9. The left plot of Figure 8.8 shows the approximate eigenfunction corresponding to the smallest subspace angle at $\lambda = 2.1$, just before the avoided crossing, while the right plot shows the approximate eigenfunction corresponding to the second smallest subspace angle at this value of λ. Figure 8.9 shows the same functions but now for the value $\lambda = 2.2$, which is after the avoided crossing. Before the avoided crossing occurs the function associated with the smallest subspace angle looks like the first eigenfunction on this region while after the avoided crossing it looks like the second eigenfunction. The second smallest subspace angle shows just the opposite behavior. Thus we see that the most rapid change in the approximate eigenfunctions appears at the avoided crossings and furthermore, the functions associated with higher subspace angles are approximations to the eigenfunctions of neighboring eigenvalues.

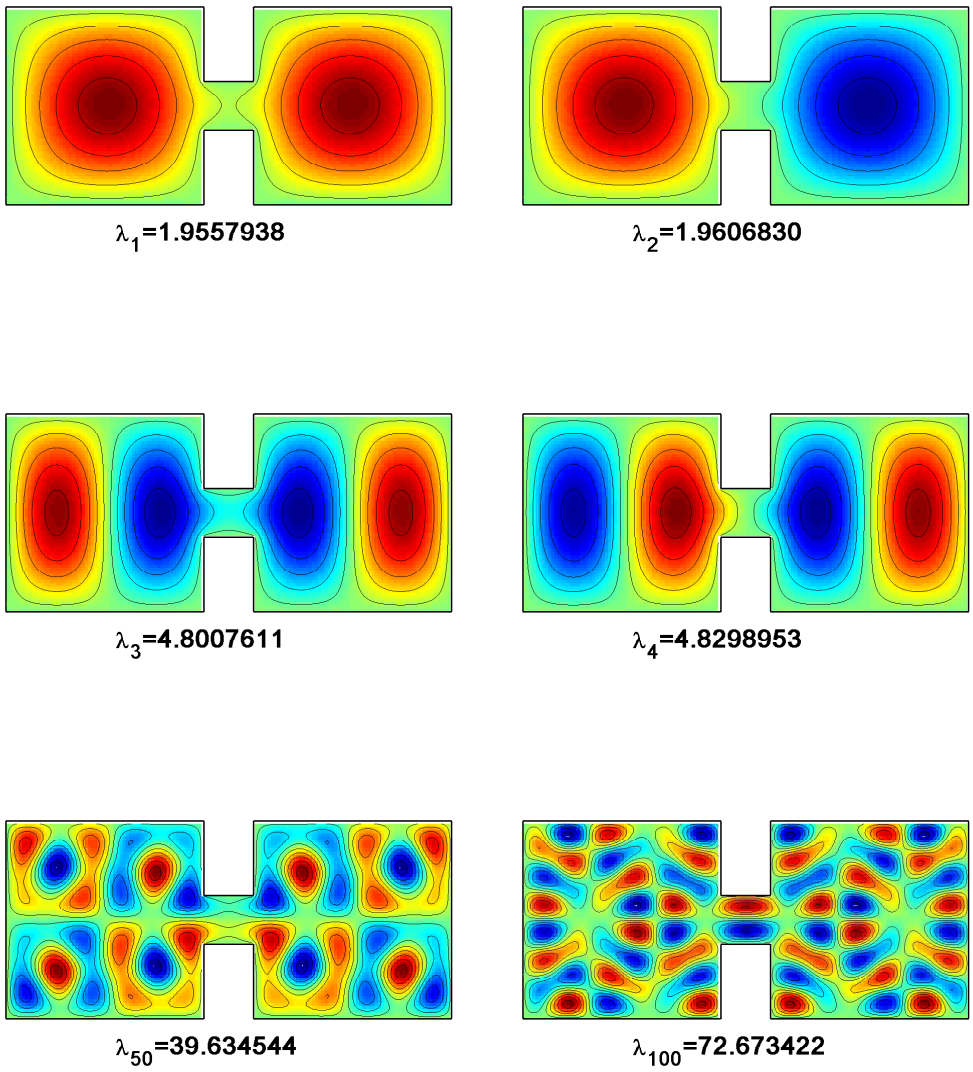


Figure 8.5: Some eigenfunctions of the symmetric dumbbell. The first eigenvalues are close to those of the square with side length π.

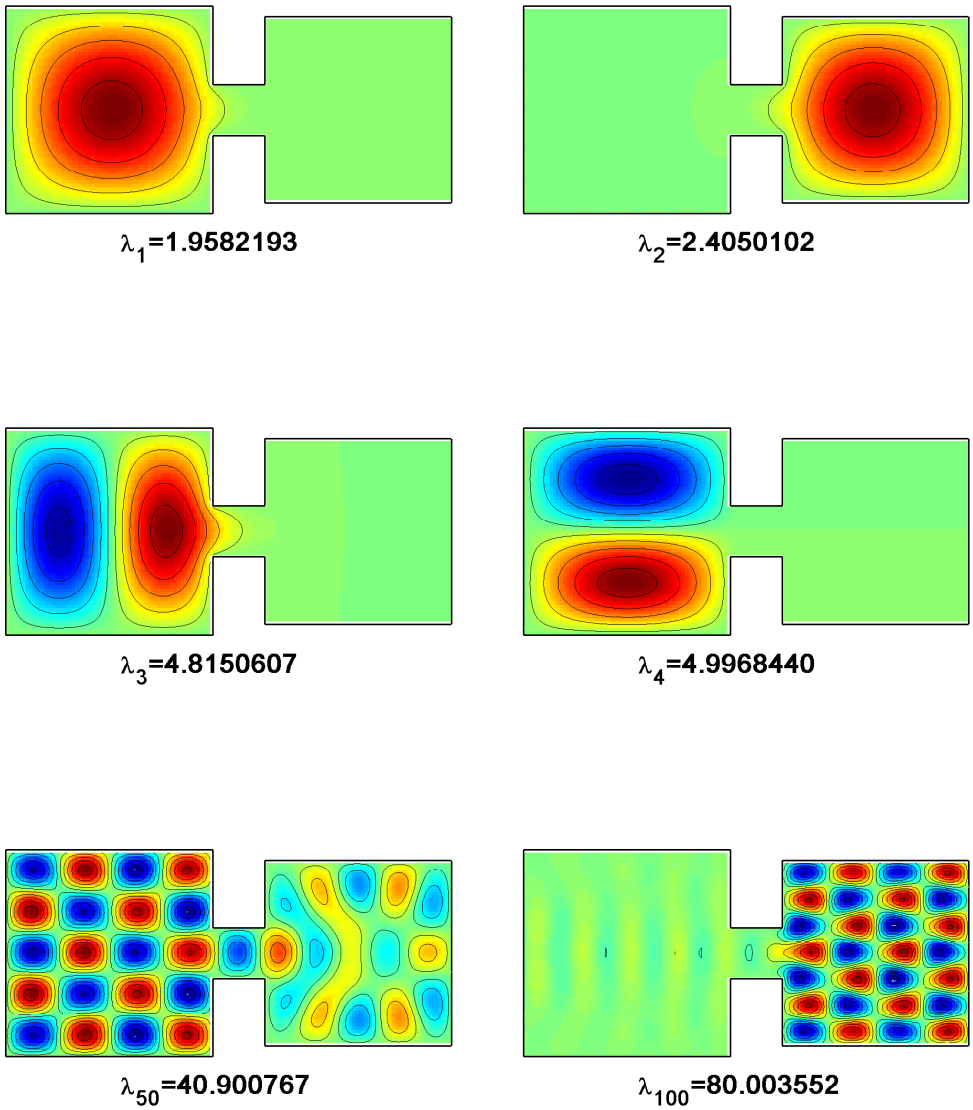


Figure 8.6: Some eigenfunctions of the nonsymmetric dumbbell. The first eigenvalues are localized. But also for some higher eigenvalues localization can occur at the corresponding eigenfunctions.

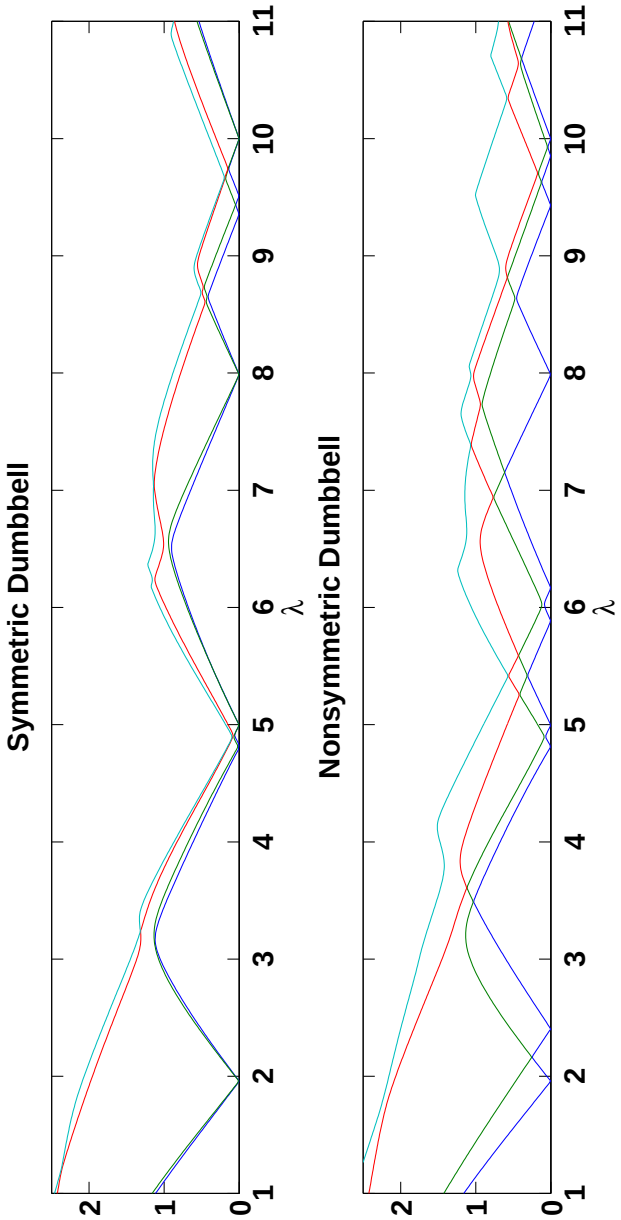


Figure 8.7: Subspace angle curves for the symmetric and the nonsymmetric dumbbell.

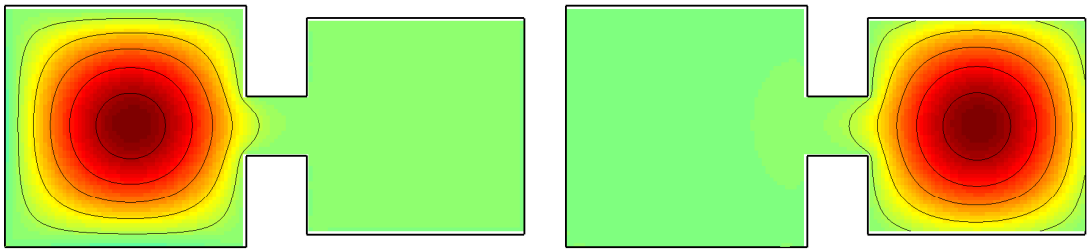


Figure 8.8: Approximate eigenfunctions corresponding to the smallest (left) and second smallest (right) subspace angle at the value $\lambda = 2.1$ before the avoided crossing occurs.

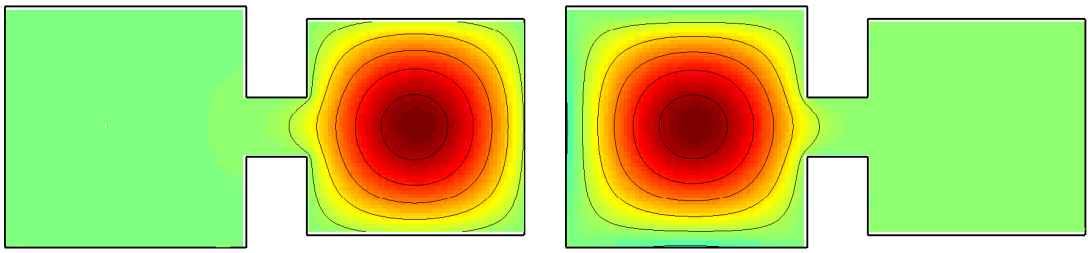


Figure 8.9: The same plot as in Figure 8.9 but now for the case $\lambda = 2.2$, after the avoided crossing.

8.4 The GWW isospectral drums

In 1966 Kac [41] asked the famous question "Can one hear the shape of a drum?". This question asks if there exist two distinct regions which have the same spectrum. A beautiful survey of this question was given by Protter in 1987 [60]. But the answer was first found by Gordon, Webb and Wolpert in 1992 [32] and it is no, one cannot hear the shape of a drum. Two of the simplest isospectral regions which they found are the GWW isospectral drums. However, the proof of isospectrality does not give the eigenvalues. Highly accurate computations of the eigenvalues were first done by Driscoll in 1997 [21]. He used a modification of the domain decomposition method by Descloux and Tolley (see Chapter 7 for an introduction) to compute the first 25 eigenvalues of the GWW isospectral drums to 12 digits of accuracy. In Figure 8.10 we present some of the eigenvalues of the isospectral drums computed with the subspace angle method, showing that our method is at least as accurate as that of Descloux, Tolley and Driscoll.

8.5 Eigenvalue avoidance

The phenomenon of eigenvalue avoidance is linked to the question of how likely it is that a given operator has multiple eigenvalues. In 1929 von Neumann and Wigner [83] showed that the set of real symmetric $N \times N$ matrices with multiple eigenvalues has codimension 2, which means that this set has two degrees of freedom less than the set of all symmetric matrices and is therefore unlikely to be encountered by chance. Let us look at the family of matrices $F(t) := A + tB$, where A and B are real symmetric $N \times N$ matrices and t is a real parameter. If A and B are randomly chosen the eigenvalues $\lambda_k(t), k = 1, \ldots, N$ of $F(t)$ might come very close to each other. But they will probably not intersect since we only have one degree of freedom t but two conditions for a multiple eigenvalue. This eigenvalue avoidance phenomenon is beautifully explained by Peter Lax in his textbook *Linear Algebra* [46], and illustrated by a picture on the cover. Eigenvalue avoidance is not only observed for finite dimensional operators. Uhlenbeck in 1976 [76] and Teytel in 1999 [73] showed that these results can be generalized to certain classes of selfadjoint operators

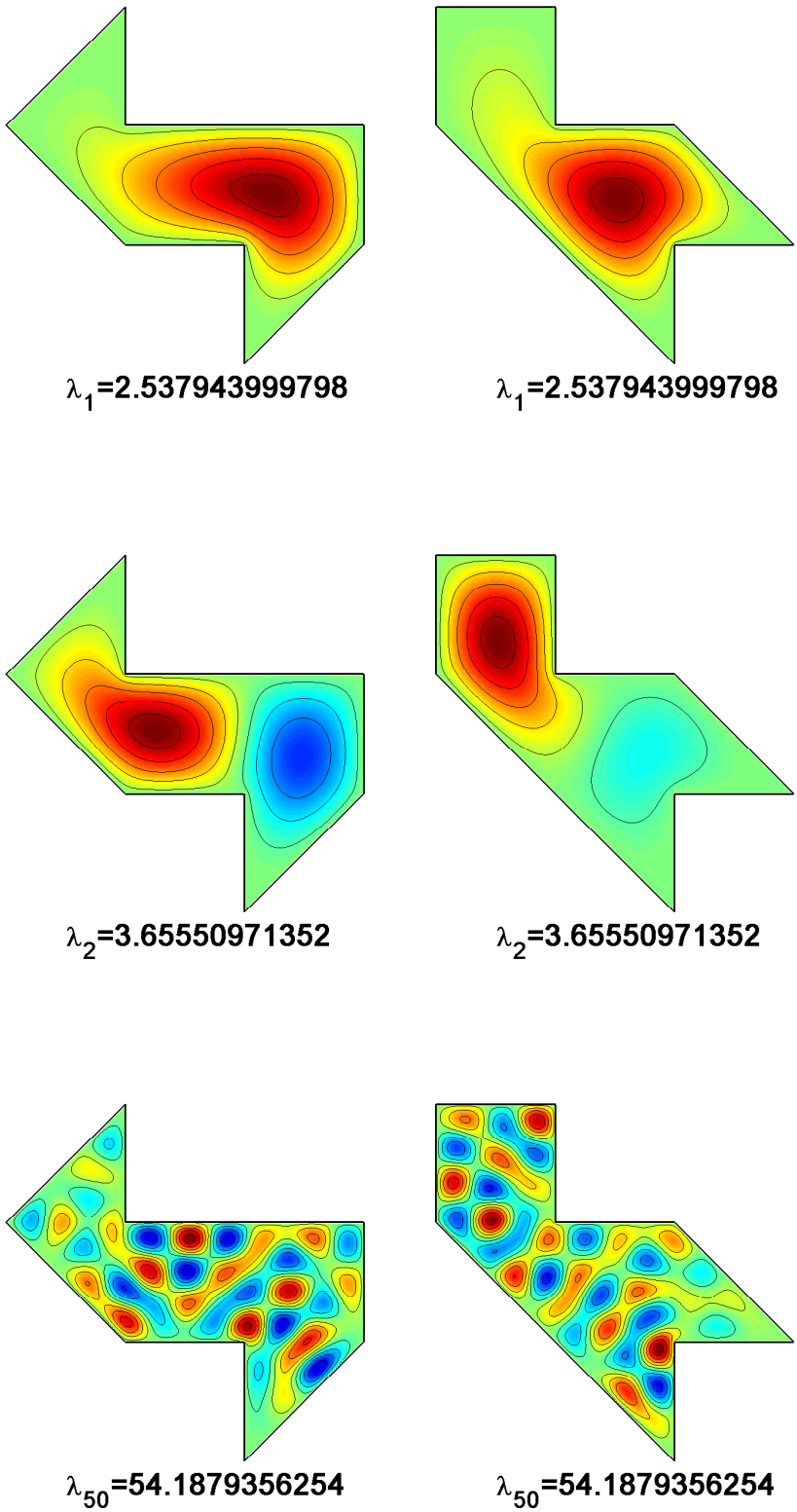


Figure 8.10: Some eigenfunctions of the GWW isospectral drums. Both regions have the same spectrum.

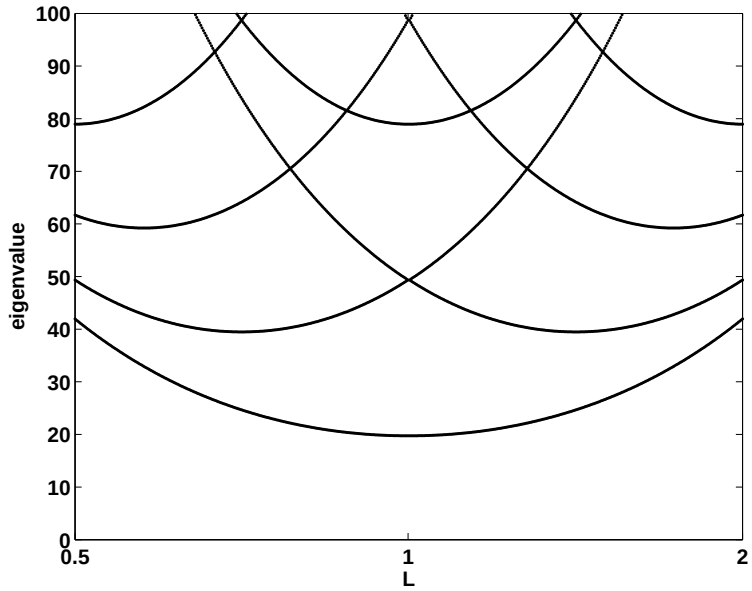


Figure 8.11: Eigenvalue curves for a parameter-dependent rectangle. At $L = 1$ the second and third eigenvalues intersect. Eigenvalue crossings for some higher eigenvalues are also visible.

acting on Hilbert spaces. A beautiful analysis of eigenvalue crossings on triangles was published in 1984 by Berry and Wilkinson [13].

In this section we want to make this phenomenon visible using the subspace angle method. Consider a rectangle with side lengths L and $1/L$. If we let L take values from 0.5 to 2 then for $L = 1$ we obtain a square and the shapes for $L = 0.5$ and $L = 2$ are identical. Hence, for the eigenvalues $\lambda_k(L)$ on this region we have $\lambda_k(0.5) = \lambda_k(2)$. At $L = 1$ the second and third eigenvalue cross since on a square we have $\lambda_2 = \lambda_3$. The eigenvalue curves $\lambda_k(L)$ for the first eigenvalues on this region are shown in Figure 8.11. Now assume that we perturb the shape slightly, i.e. the new shape is defined by the four points 0, L, $L - p + i/L$, i/L, where $p > 0$ is a small perturbation. Then, as in the finite dimensional case, we cannot expect eigenvalue crossings to occur any more. The eigenvalues might still come close but they will not intersect. For the value $p = 0.2$ this is shown in Figure 8.12. All eigenvalue crossings have disappeared. It seems that the eigenvalues avoid each other.

If we decrease p further then $\lambda_2(1)$ and $\lambda_3(1)$ will come closer and eventually be equal for $p = 0$. How small can we make p and still numerically detect that these two eigenvalues are distinct? This is a good test for the accuracy of the subspace

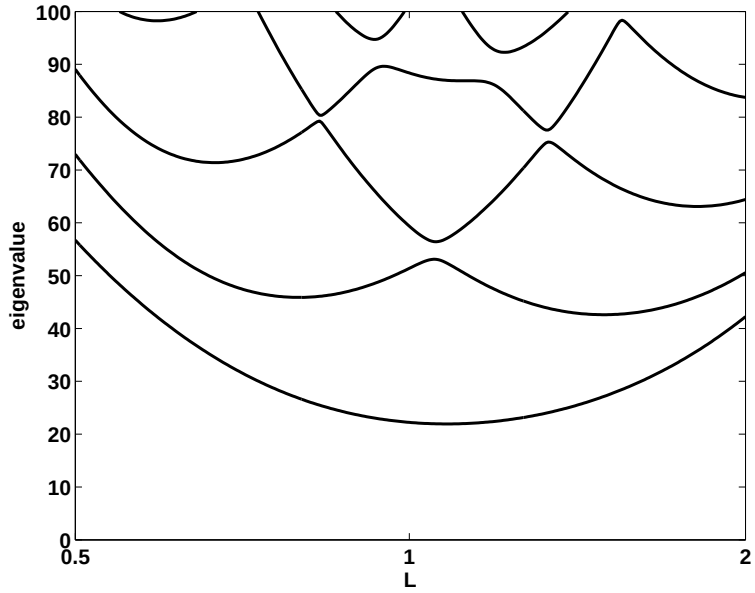


Figure 8.12: Eigenvalue curves for a perturbed parameter-dependent rectangle. All eigenvalue crossings have disappeared. The eigenvalues avoid each other.

angle method. Figure 8.13 shows the subspace angle curve close to $5\pi^2$ (the second eigenvalue on the unperturbed square) for $p = 10^{-13}$ and $L = 1$. The curve has two minima pointing to the two different eigenvalues λ_2 and λ_3 on the perturbed region. Since the subspace angle is of the order of magnitude of 10^{-14}, we can assume that the minima are within a relative error of roughly 10^{-14} the correct eigenvalues. The curve in Figure 8.13 is also another nice example for the perturbation results derived in Chapter 4 for the subspace angle method. Due to ill-conditioning in the basis the curve shows oscillations which become smaller as we approach the minima making it possible to detect the eigenvalues to high accuracy and therefore to distinguish λ_2 and λ_3 even for the small perturbation $p = 10^{-13}$.

8.6 A region with a hole

Until now we have always considered simply connected regions. In this section we want to give our first example of a multiply connected region. It is an annulus, in which the outer and the inner circle have different centers. The inner circle has radius 0.5 and center at 0. The outer circle has radius 1 and center at 0.4. As basis functions we use linear combinations of Fourier-Bessel functions of the first and of the second

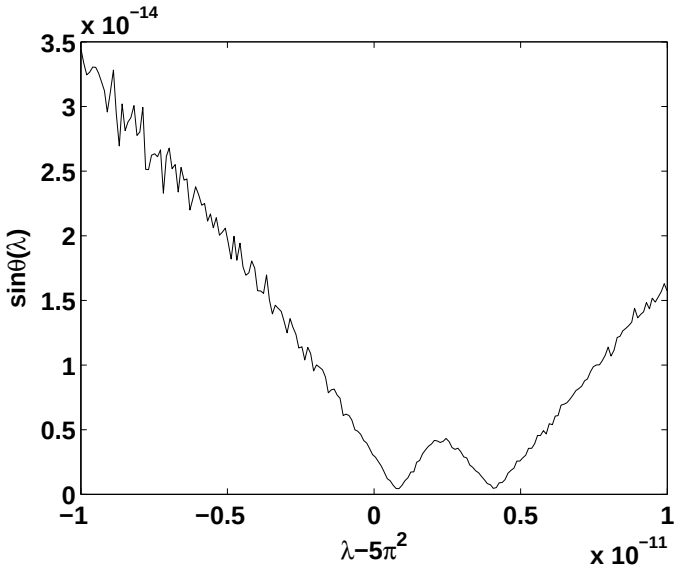


Figure 8.13: The subspace angle curve for a slightly perturbed unit square. Although the perturbation p has the value 10^{-13} two distinct eigenvalues are recognizable.

kind in the form

$$\sum_{k=0}^{N} Y_k(\sqrt{\lambda}r)\left(A_Y^{(N)}\cos k\theta + B_Y^{(N)}\sin k\theta\right) + \sum_{k=0}^{N} J_k(\sqrt{\lambda}r)\left(A_J^{(N)}\cos k\theta + B_J^{(N)}\sin k\theta\right). \tag{8.1}$$

The idea is that this is analogous to approximating a holomorphic function in an annulus with a Laurent series. This approach can also be justified from the fact that every solution of $-\Delta u = \lambda u$ in a circular annulus $R_1 < |z| < R_2$ can be expanded in a series of the form (8.1) with $N \to \infty$ (see [80], §22). Figure 8.14 shows some of the eigenvalues of the annulus region.

The subspace angle method can also do regions with several holes. Such a region is shown in Figure 8.15. The inner circle is of radius 1 and the outer circle is of radius 2. In the upper plot the small holes are of radius 0.4, while in the lower plot the radius of the right small hole is reduced to 0.3. The eigenfunctions are approximated by linear combinations of Fourier-Bessel functions of the first and second kind around the big center hole together with Fourier-Bessel functions of the second kind around the small holes. Similarly to the example of the dumbbell we can see localization effects of the eigenfunctions if the symmetry is broken.

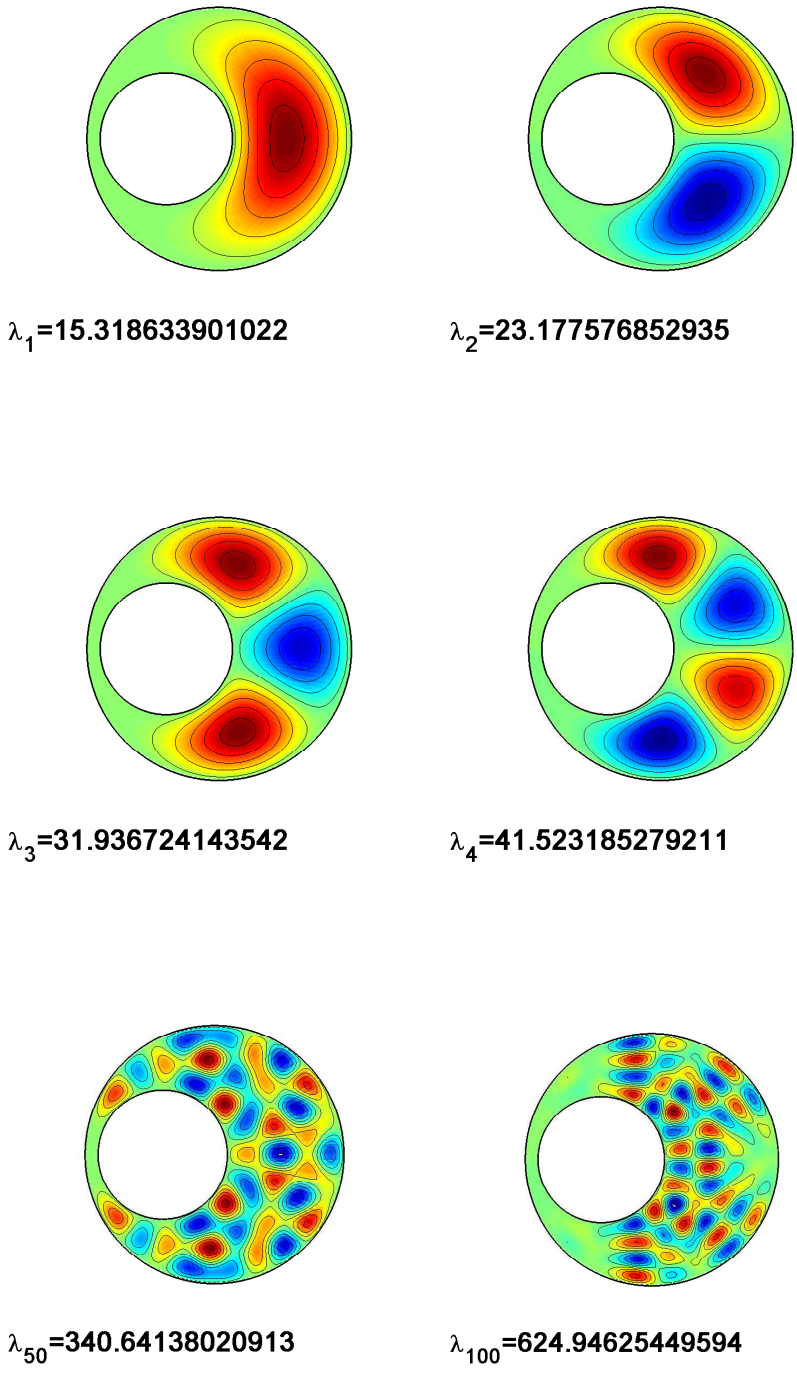


Figure 8.14: Some eigenfunctions of a circle with a hole. An interesting localization effect is visible in the eigenfunction of λ_{100}.

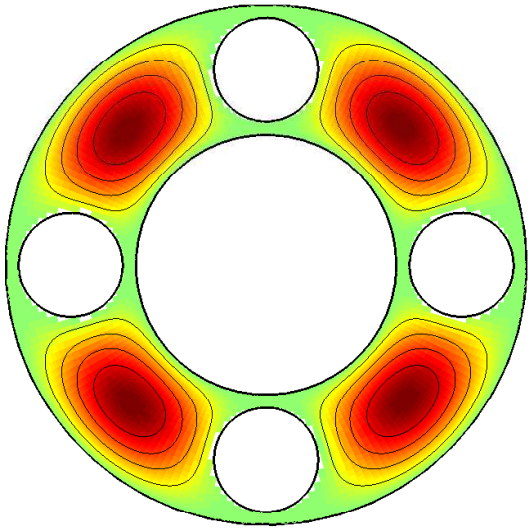


Figure 8.15: Eigenfunctions on a circular region with five holes. As in the example of the dumbbell the breaking of symmetry leads to localization.

8.7 A square with a square shaped hole

While in the previous section the holes in the region did not have singular corners, we now consider a square with a square shaped hole. The inner boundary is a square with side length one and lower left corner at zero while the outer boundary is a square with side length three and lower left corner at $-0.5 - 0.5i$. The four inner corners of the region are singular with interior angle $\frac{3}{2}\pi$. The outer corners are regular. We could compute the eigenvalues and eigenfunctions by cutting the region along its symmetry axis in two halves to obtain two simply connected regions. The odd modes are then obtained by computing the eigenvalues on the half region with zero Dirichlet boundary conditions and the even modes are obtained by computing the eigenvalues of the half region with zero Neumann conditions along the symmetry axis and zero Dirichlet boundary conditions on the other sides. But by slightly changing the position of the inner square the symmetry would be lost and this would not be possible any longer. Therefore, we directly use the domain decomposition GSVD method to compute the eigenvalues and eigenfunctions on the whole region. This can be done by dividing the region into four subdomains, each of which contains one singular corner. Figure 8.16 shows some eigenfunctions of the region computed with the domain decomposition GSVD. To obtain the first 7 digits of the presented eigenvalues around 40 basis functions are needed at each singular corner. The corresponding smallest generalized singular value is of the magnitude 10^{-3} which shows a squared convergence effect for the eigenvalue on this region. The exponential convergence rate is relatively slow since the eigenfunctions have singularities inside the inner square.

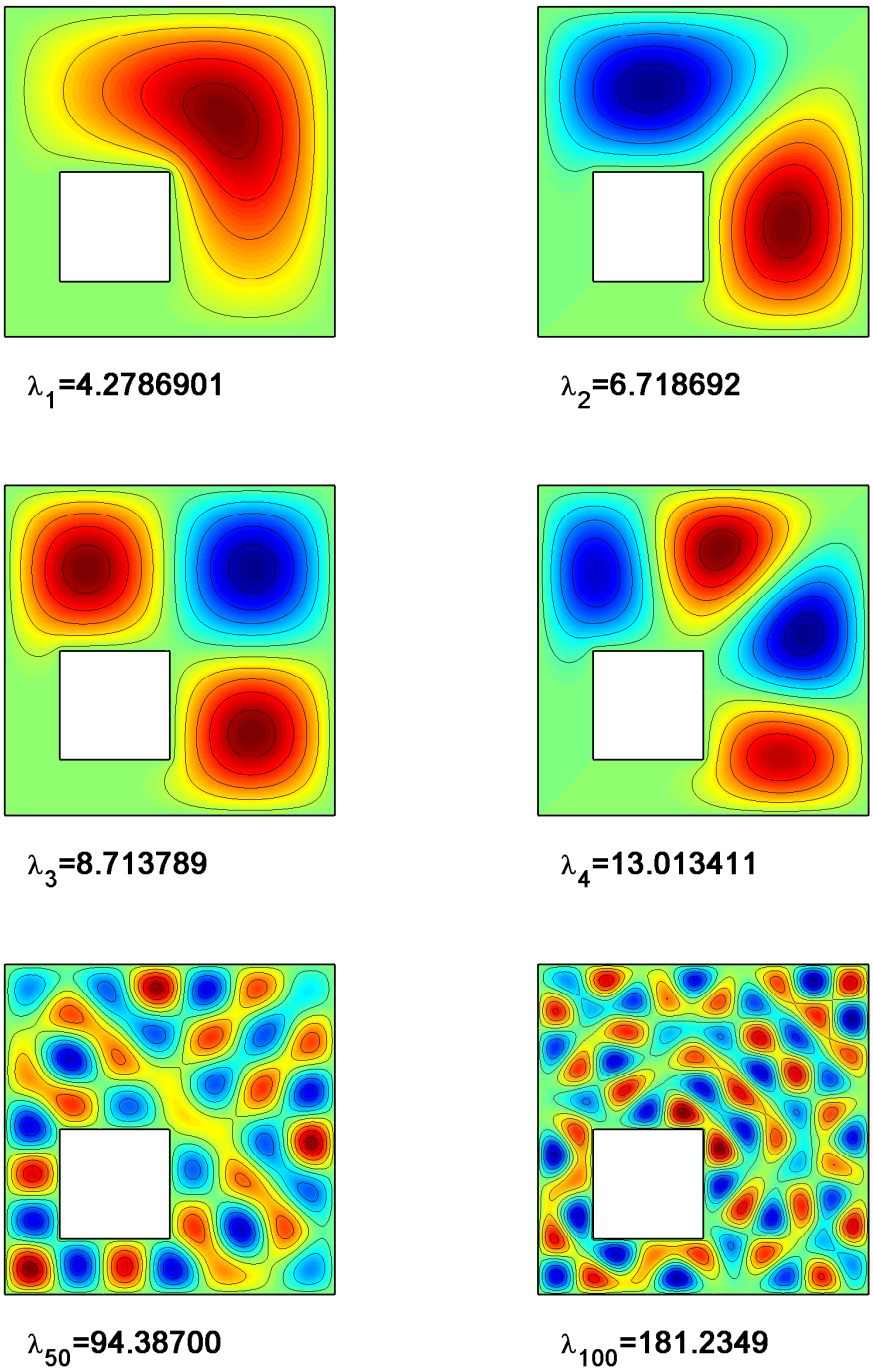


Figure 8.16: Some eigenfunctions on the square with a square-shaped hole. Only for eigenfunctions belonging to higher eigenvalues is the local wavelength small enough to fully penetrate the lower left part of the region.

Chapter 9

Conclusions

This chapter summarizes the contributions of this thesis and gives an outlook to further research questions in numerical linear algebra, approximation theory and for further applications.

9.1 Numerical linear algebra

The first chapters of this thesis were concerned with how tools from linear algebra can be applied to the MPS in order to obtain a stable and accurate algorithm. The original MPS of Fox, Henrici and Moler used the determinant of square matrices containing the basis functions evaluated at boundary collocation points to determine an eigenvalue. In Chapter 2 we showed that this approach generally fails on more complicated regions. The method gets somewhat better behaved if one chooses many more boundary collocation points than there are basis functions [54]. In terms of numerical linear algebra this means going away from determinants of square matrices to singular values of rectangular matrices. But as we showed in Chapter 2 the fundamental problem of spurious solutions remains. We need a method which also uses information about the approximate eigenfunctions in the interior of the region. Two such approaches are the PWDM by Heller and Barnett's method. But while the first one only partially solves the stability problems of the MPS the latter one is only accurate up to $O(\sqrt{\epsilon_{mach}})$ and explicitly has to deal with the ill-conditioning in the basis functions as we showed in Section 4.6.

Based on principal angles between subspaces we developed a stable and accurate method in Chapter 3. The principal idea was to introduce additional interior points and to minimize a certain angle between two subspaces. We then showed that this is equivalent to minimizing the smallest generalized singular value of the pencil $\{A_B(\lambda), A_I(\lambda)\}$. Hence, the MPS becomes stable by going over from singular value computations to generalized singular values.

We also showed that Barnett's approach can be interpreted as solving the generalized eigenvalue problem

$$A_B(\lambda)^T A_B(\lambda) x(\lambda) = \mu(\lambda) A_I(\lambda)^T A_I(\lambda),$$

which is just a squared formulation of our method. In Section 4.6 we compared the GSVD approach with the formulation as generalized eigenvalue problem and showed that the GSVD is a more stable and accurate tool for the MPS. Again it is advantageous to use a tool for rectangular matrices (GSVD) rather than one for square matrices (generalized eigenvalues).

The GSVD and other algorithms that work on rectangular matrices are still less developed than square matrix methods. While there is a variety of methods for large structured eigenvalue and generalized eigenvalue problems, we are only aware of two methods for the GSVD of large and structured problems [39, 87]. Such methods would especially be useful for the domain decomposition GSVD approach proposed in Chapter 7.

Also the stability of the GSVD for matrix pencils $\{A, B\}$ such that $Y = \begin{bmatrix} A \\ B \end{bmatrix}$ is ill-conditioned has not yet been very much explored in the literature. In most articles about the GSVD it is assumed that Y is well-conditioned. In this thesis we showed that the GSVD can also deliver meaningful results for heavily ill-conditioned problems. Since ill-conditioned bases appear in a variety of applications further research of the GSVD in such situations can lead to new robust algorithms.

Another question is that of the resolvent and pseudospectra for the Method of Particular Solutions. Figure 9.1 shows an extension of Figure 3.2 into the complex plane by taking complex values for λ. This raises the question of how subspace angles are connected to pseudospectra and the resolvent norm of the Laplacian. Let

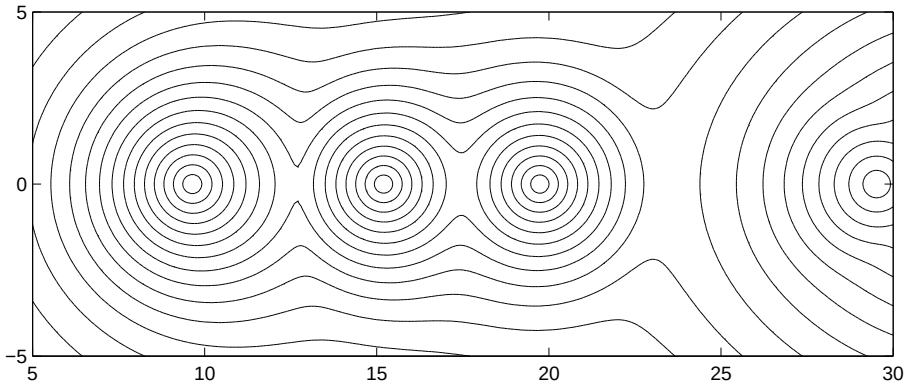


Figure 9.1: Extension of Figure 3.2 to the complex plane. The plot shows the level curves $|s_1(\lambda)| = 0.0.5, 0.1, 0.15, \ldots$ for the L-shaped region.

$r(\lambda) := (\lambda I + \Delta)^{-1}$ be the resolvent function of the Laplacian. The resolvent norm is defined as

$$\|r(\lambda)\| := \sup_{\substack{u \in L^2(\Omega) \\ \|u\|_\Omega = 1}} \|r(\lambda)u\|_\Omega.$$

For the MPS we can define a similar function using the cotangent of the subspace angle between $\mathcal{A}(\lambda)$ (here $\mathcal{A}(\lambda)$ is the space of all possible particular solutions as defined in (1.2)) and $\mathcal{D}_0$. Then

$$\cot\theta(\lambda) = \sup_{\substack{u \in \mathcal{A}(\lambda) \\ \|u\|_{\partial\Omega} = 1}} \|u\|_\Omega.$$

Both functions, the resolvent and $\cot\theta(\lambda)$, have poles at the eigenvalues. Finding a meaningful connection between both would also lead to a meaningful connection between subspace angles and pseudospectra. Although the Laplacian is a selfadjoint operator and therefore its pseudospectra are simply disks around the eigenvalues, a connection to subspace angles could lead to interesting new insights.

An example is the method of Vergini and Saraceno (see [82] or the introduction in the thesis of Barnett [6]). It solves a generalized eigenvalue problem that depends on the parameter λ and computes from the generalized eigenvalues approximate distances to the eigenvalues of (1.1) closest to λ. The advantage is that only one matrix decomposition is necessary to obtain approximations for several eigenvalues of (1.1). Also the resolvent norm is a distance computation since for the Laplacian $\|r(\lambda)\| = \frac{1}{|\lambda - \lambda_k|}$, where λ_k is the eigenvalue of (1.1) closest to λ. But while the resolvent norm is well understood there are still many open questions concerning the method of Vergini and

Saraceno. For example, currently it only works satisfactory on star-shaped regions. But it has many connections to Barnett's method, which can be regarded as a squared version of our subspace angle computations. Therefore, connecting the resolvent and $\cot\theta(\lambda)$ might give new insight into the method of Vergini and Saraceno.

9.2 Approximation theory

Another major aspect of this thesis is the approximation theory for the Method of Particular Solutions. Based on results of Vekua and Garabedian we derived in Chapter 6 exponential convergence estimates for the MPS for regions with zero or one singular corner and computed the asymptotic exponential rates for several regions using conformal mapping techniques. For regions with multiple singular corners we have numerical results indicating faster than algebraic convergence if an increasing number of basis functions at the singular corners is used. In Chapter 7 we extended our results to domain decomposition methods and thereby improved the original estimates of Descloux and Tolley.

But still there are several open questions concerning the approximation theory of the MPS. We do not yet have a theoretical analysis of the convergence rate shown in Figure 6.15 which would give convergence estimates of our method for regions with multiple singularities. Also the convergence of the MPS in the case of multiply connected regions has not yet been investigated. For example, in the region shown in Figure 8.14 we observed rapid convergence to the solution. We think that there is a close connection to rational approximation in the complex plane.

In this thesis we always used Fourier-Bessel basis sets. These are easily adapted to reflect corner singularities of the eigenfunctions and are directly connected to polynomials via the Vekua theory. However, in some applications other basis sets are preferable. For example, in the case of the Bunimovich stadium billiard a combination of real plane waves and evanescent waves leads to very good results [6, 81]. How do Fourier-Bessel basis sets compare with real plane waves? Figure 9.2 compares the convergence behavior of the subspace angle method for the first eigenvalue on the circular L region with Fourier-Bessel functions and real plane waves. Both basis sets should lead to the same convergence behavior as there are no corner singularities on

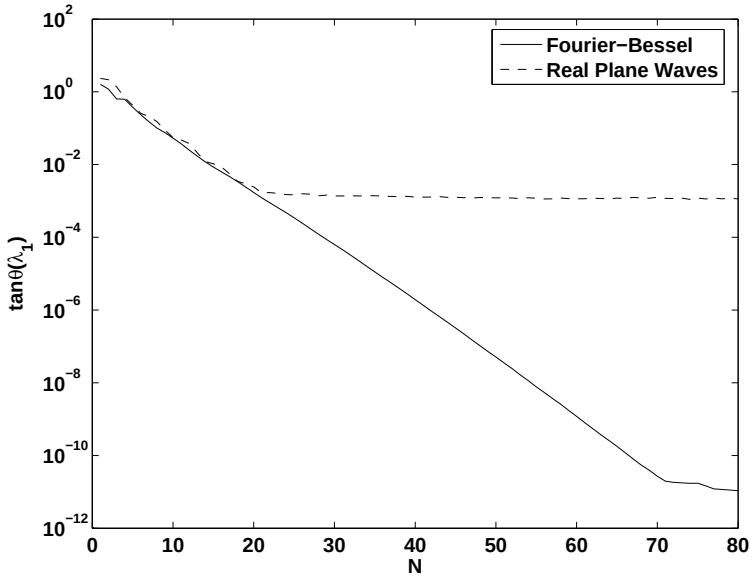


Figure 9.2: Comparison of real plane waves and Fourier-Bessel functions on the circular L region. Although they should both have to the same theoretical asymptotic rate of convergence the numerical behavior is completely different.

this region. But with real plane waves the convergence stops at about 10^{-3} while the convergence of the Fourier-Bessel basis set continues until about 10^{-11} (it cannot converge further since we know λ_1 only to 12 digits of accuracy). Explaining this phenomenon would lead to new insight to the question of when to use which basis set.

A very different basis set is used in the Method of Fundamental Solutions. There, one approximates the solution of (1.1) by linear combinations of singular particular solutions (in our case Fourier-Bessel functions of the second kind), which lie on a curve enclosing $\overline{\Omega}$ (see for example [17, 51]). In [17] it is stated that these basis sets behave favorably compared to Fourier-Bessel functions since all fundamental solutions behave uniformly on the region. This is not true for Fourier-Bessel functions which become exponentially small in Ω for increasing order. But further numerical experiments are necessary to determine if fundamental solutions really lead to a better numerical behavior than Fourier-Bessel basis sets, and it is yet unclear how corner singularities are approximated with such basis sets.

9.3 Further applications

In this thesis we focused on solutions of the Laplace eigenvalue problem (1.1) with Dirichlet boundary conditions. Neumann boundary conditions can also easily be implemented. Instead of the basis functions evaluated at boundary collocation points the matrix $A_B(\lambda)$ then contains the normal derivative of the basis functions evaluated at the boundary collocation points. The algorithms described in this thesis can also be applied to more general elliptic eigenvalue problems if particular solutions are known. The numerical construction of particular solutions for elliptic PDEs with polynomial coefficient functions was discussed by Schryer in [64]. If we want to go over to three dimensional problems we could similarly as in 2d use particular solutions for the three dimensional Laplace eigenvalue problem and find the eigenvalues with the subspace angle method. Difficulties arise if the region has corners. In three dimensions corners can have almost arbitrary shapes and it is a hard problem to find particular solutions in 3d which are adapted to the corner singularities.

Another interesting application in 2d is the extension of the subspace angle method to the biharmonic eigenvalue problem. This would give us a tool to compute Chladni figures to high accuracy using particular solutions[1].

In this thesis we have only treated interior eigenvalue problems. But also of great interest is the solution of Helmholtz problems in the exterior of a region. Adapting the subspace angle method to such problems could lead to many interesting applications of the MPS in scattering theory.

A fascinating topic is eigenvalue problems on fractal drums. Computing eigenvalues and eigenfunctions of the Koch snowflake shown in Figure 9.3 is a beautiful example of this [5, 45]. Since approximations of this fractal have thousands of corners the approach of capturing corner singularities by singular Fourier-Bessel functions does not seem feasible. The basis size would be too big. Currently, the most successful approach seems to be the one proposed in [5] which uses a conformal mapping technique for regions with thousands of corners to transplant the eigenvalue problem on the snowflake into a system of nonlinear equations in a reference region. It would be

[1] Napoleon would be fascinated by this!

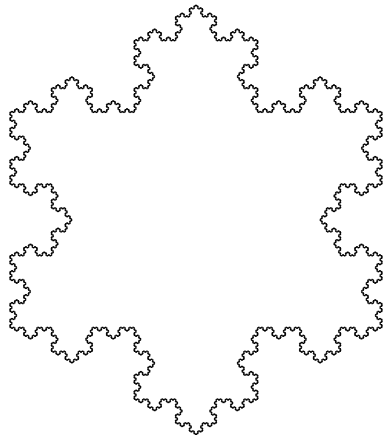


Figure 9.3: The Koch snowflake

fascinating to also find a way of using particular solutions for computing eigenvalues to high accuracy for this region.

9.4 Do we have the best method for computing eigenvalues on planar regions?

It is dangerous to ask such a question since the best method always depends too much on the specific application. If we want general purpose solvers for arbitrary regions then methods like boundary element methods (BEM) or finite element methods (FEM) are probably the best choice. The picture looks different if we focus on regions with piecewise analytic boundary and a small number of singular corners. Then the subspace angle method is easily implementable and at the same time highly accurate. Certainly, we could also tune general purpose methods like FEM to deliver rapid convergence on such regions. But the beauty of the MPS together with the subspace angle approach is that writing a code for a certain region is often just a matter of minutes due to its simple idea. For such problems the subspace angle method is probably the best choice for many applications. It is not only fast and accurate but also easily implementable. In Figure 9.4 we show Matlab code that fits on a single page and computes the first three eigenvalues on the L-shaped region to 10 digits of accuracy in just five seconds on a modern computer[2]. This is what we are striving

[2]This is an example of a"Ten Digit Algorithm" as proposed by Trefethen in [74].

for in numerical analysis, finding simple and beautiful algorithms which deliver fast and accurate solutions to our problems.

```
% Ldrum.m  Compute eigenvalues of Laplacian on L-shaped region
%          T. Betcke and L. N. Trefethen 9/03
%
% The first three eigenvalues are computed by the method of
% particular solutions (Betcke & Trefethen, SIAM Review 2005).

% Compute subspace angles for various values of lambda:
  N = 36; k = 1:N;                        % orders in Bessel expansion
  np = 2*N;                               % no. of bndry & interior pts
  t1 = 1.5*pi*(.5:np-.5)'/np;             % angles of bndry pts
  r1 = 1./max(abs(sin(t1)),abs(cos(t1))); % radii of bndry pts
  t2 = 1.5*pi*rand(np,1);                 % angles of interior pts
  r2 = rand(np,1)./max(...
           abs(sin(t2)),abs(cos(t2)));    % radii of interior pts
  t = [t1;t2]; r = [r1;r2];               % bndry and interior combined
  lamvec = .2:.2:25; S = [];              % trial values of lam
  for lam = lamvec
    A = sin(2*t*k/3).*...
          besselj(2*k/3,sqrt(lam)*r);
    [Q,R] = qr(A,0);
    s = min(svd(Q(1:np,:))); S = [S s];   % subspace angle for this lam
  end

% Convert to signed subspace angles:
  I = 1:length(lamvec);                   % all lam points
  J = I(2:end-1);                         % interior points
  J = J( S(J)<S(J-1) & S(J)<S(J+1) );     % local minima
  J = J + (S(J-1)>S(J+1));                % points where sign changes
  K = 0*I; K(J) = 1;
  S = S.*(-1).^cumsum(K);                 % introduce sign flips
  subplot(3,1,1)
  hold off, plot(lamvec,S), hold on       % plot signed angle function
  plot([0 max(lamvec)],[0 0],'-k')        % plot lam axis

% Find eigenvalues via 9th-order interpolation:
  for j = length(J):-1:1
    I = J(j)-5:J(j)+4;
    lam = polyval(polyfit(S(I)/norm(S(I)),lamvec(I),9),0);
    plot(lam*[1 1],[-1 1],'r')
    text(lam,.6,sprintf('%13.9f',lam),'color','r')
  end

% Plot the first eigenfunction:
  [X,Y] = meshgrid(-1:.05:1,-1:.05:1); Z = X(:)+i*Y(:);
  p = [0 1i -1+1i -1-1i 1-1i 1];
  [in on] = inpolygon(real(Z),imag(Z),real(p),imag(p));
  zB = Z(on); zI = Z(in&~on); z = [zB;zI]; t = mod(angle(z/i),2*pi);
  A = besselj(2*k/3,sqrt(lam)*abs(z)).*sin(2*t*k/3);
  [Q,R] = qr(A,0); [U,S,V] = svd(Q(1:length(zB),:));
  V = V(:,end); Q = Q*V; [t,I] = max(abs(Q)); Q = Q/Q(I);
  F = NaN*zeros(size(Z));
  F(in&~on) = Q(length(zB)+1:end); F(on) = Q(1:length(zB),:);
  F = reshape(F,length(X),length(Y)); subplot(3,1,2:3)
  surf(X,Y,F), view(-150,40), axis off, zlim([0 .7])
```

Figure 9.4: This code computes the first three eigenvalues on the L-shaped region to 10 digits of accuracy. Instead of using fminsearch the subspace angle curve is converted to a curve which has sign changes close to the eigenvalues. These are then determined by a 9th-order polynomial interpolation.

Bibliography

[1] E. Anderson, Z. Bai, C. Bischof, L. S. Blackford, J. Demmel, Jack J. Dongarra, J. Du Croz, S. Hammarling, A. Greenbaum, A. McKenney, and D. Sorensen. *LAPACK Users' guide (third ed.)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1999.

[2] R. Bañuelos and T. Carroll. Brownian motion and the fundamental frequency of a drum. *Duke Math. J.*, 75(3):575–602, 1994.

[3] R. Bañuelos and T. Carroll. Addendum to: "Brownian motion and the fundamental frequency of a drum". *Duke Math. J.*, 82(1):227, 1996.

[4] Z. Bai and J. W. Demmel. Computing the generalized singular value decomposition. *SIAM J. Sci. Comput.*, 14(6):1464–1486, 1993.

[5] L. Banjai. Eigenfrequencies of fractal drums. Preprint Nr. 11/2005, Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, 2005.

[6] A. H. Barnett. *Dissipation in deforming chaotic billiards*. PhD thesis, Department of Physics, Harvard University, Cambridge, Massachusetts, 2000.

[7] A. H. Barnett. Inclusion of Dirichlet eigenvalues in the semiclassical limit via a boundary spetral problem. *In preparation*, 2004.

[8] A. H. Barnett. Quasi-orthogonality on the boundary for Euclidian Laplace eigenfunctions. *Submitted to Comm. Pure Appl. Math.*, 2004.

[9] Z. Battles. *Numerical Linear Algebra for Continuous Functions*. PhD thesis, Oxford University, Computing Laboratory, Oxford, 2005.

[10] S. Bergman. Functions satisfying certain partial differential equations of elliptic type and their representation. *Duke Math. J.*, 14:349–366, 1947.

[11] S. Bergman. *Integral Operators in the Theory of Linear Partial Differential Equations*. Springer, Berlin, 1961.

[12] M. V. Berry. Evanescent and real waves in quantum billiards and Gaussian beams. *J. Phys. A*, 27:L391–L398, 1994.

[13] M. V. Berry and M. Wilkinson. Diabolical points in the spectra of triangles. *Proc. R. Soc. Lond.*, A392:15–43, 1984.

[14] T. Betcke and L. N. Trefethen. Computations of eigenvalue avoidance in planar domains. *Proc. Appl. Math. Mech.*, 4(1):634–635, 2004.

[15] T. Betcke and L. N. Trefethen. Reviving the method of particular solutions. *SIAM Review*, 47(3):469–491, 2005.

[16] Å. Björck and G. H. Golub. Numerical methods for computing angles between linear subspaces. *Math. Comp.*, 27:579–594, 1973.

[17] A. Bogomolny. Fundamental solutions method for elliptic boundary value problems. *SIAM J. Numer. Anal.*, 22(4):644–669, 1985.

[18] J. Descloux and M. Tolley. An accurate algorithm for computing the eigenvalues of a polygonal membrane. *Comput. Methods Appl. Mech. Engrg.*, 39(1):37–53, 1983.

[19] J. D. P. Donnelly. Eigenvalues of membranes with reentrant corners. *SIAM J. Numer. Anal.*, 6:47–61, 1969.

[20] T. A. Driscoll. Schwarz-Christoffel toolbox for MATLAB. Available at http://www.math.udel.edu/ driscoll/software/SC/index.html.

[21] T. A. Driscoll. Eigenmodes of isospectral drums. *SIAM Rev.*, 39(1):1–17, 1997.

[22] Z. Drmač. On principal angles between subspaces of Euclidean space. *SIAM J. Matrix Anal. Appl.*, 22(1):173–194, 2000.

[23] S. C. Eisenstat. On the rate of convergence of the Bergman-Vekua method for the numerical solution of elliptic boundary value problems. *SIAM J. Numer. Anal.*, 11:654–680, 1974.

[24] G. Fix and R. Heiberger. An algorithm for the ill-conditioned generalized eigenvalue problem. *SIAM J. Numer. Anal.*, 9(1):78–88, 1972.

[25] L. Fox, P. Henrici, and C. Moler. Approximations and bounds for eigenvalues of elliptic operators. *SIAM J. Numer. Anal.*, 4:89–102, 1967.

[26] V. Frayssé and V. Toumazou. A note on the normwise perturbation theory for the regular generalized eigenproblem. *Numer. Linear Algebra Appl.*, 5(1):1–10, 1998.

[27] D. Gaier. *Lectures on complex approximation.* Birkhäuser, Boston, 1987.

[28] P. R. Garabedian. Applications of analytic continuation to the solution of boundary value problems. *J. Ratl. Mech. Anal.*, 3:383–393, 1954.

[29] P. R. Garabedian. *Partial differential equations.* John Wiley & Sons, 1964.

[30] G. H. Golub and C. F. van Loan. *Matrix computations - third edition.* Johns Hopkins University Press, Baltimore, 1996.

[31] G. H. Golub and H. Zha. Perturbation analysis of the canonical correlations of matrix pairs. *Linear Algebra Appl.*, 210:3–28, 1994.

[32] C. Gordon, D. Webb, and S. Wolpert. Isospectral plane domains and surfaces via Riemannian orbifolds. *Invent. Math.*, 110:1–22, 1992.

[33] M. Gutzwiller. Quantum chaos. *Scientific American*, 1992.

[34] E. J. Heller. Bound-state eigenfunctions of classically chaotic Hamiltonian systems: Scars of periodic orbits. *Phys. Rev. Lett.*, 53:1515–1518, 1984.

[35] E. J. Heller. Wavepacket dynamics and quantum chaology. In M. J. Giannoni, A. Voros, and J. Zinn-Justin, editors, *Proceedings of the 1989 Les Houches Summer School on "Chaos and Quantum Physics"*, pages 547–663, North-Holland, 1991. Elsevier Science Publishers B. V.

[36] P. Henrici. A survey of I. N. Vekua's theory of elliptic partial differential equations with analytic coefficients. *Z. Angew. Math. Phys.*, 8:169–203, 1957.

[37] J. Hersch. Erweiterte symmetrieeigenschaften von lösungen gewisser linearer rand- und eigenwertprobleme. *J. Reine Angew. Math.*, 218:143–158, 1965.

[38] N. J. Higham. *Accuracy and stability of numerical algorithms - second edition.* SIAM, 2002.

[39] M. E. Hochstenbach. A Jacobi–Davidson type method for the generalized singular value problem. Preprint, Department of Mathematics, Case Western Reserve University, Cleveland, Ohio, USA, September 2004. Submitted.

[40] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1985.

[41] M. Kac. Can one hear the shape of a drum? *Amer. Math. Monthly*, 73(4, Part 2):1–23, 1966.

[42] J. R. Kuttler. Remarks on a Stekloff eigenvalue problem. *SIAM J. Numer. Anal.*, 9, 1972.

[43] J. R. Kuttler and V. G. Sigillito. Bounding eigenvalues of elliptic operators. *SIAM J. Math. Anal.*, 9, 1978.

[44] J. R. Kuttler and V. G. Sigillito. Eigenvalues of the Laplacian in two dimensions. *SIAM Review*, 26(2):163–193, 1984.

[45] M. L. Lapidus, J. W. Neuberger, R. J. Renka, and C. A. Griffith. Snowflake harmonics and computer graphics; numerical computation of spectra on fractal drums. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 6(7):1185–1210, 1996.

[46] P. Lax. *Linear Algebra*. John Wiley & Sons Inc, 1996.

[47] R. S. Lehmann. Development of the mapping function at an analytic corner. *Pacific J. Math.*, 7:1437–1449, 1957.

[48] H. Lewy. On the reflection laws of second order differential equations in two independent variables. *Bull. Amer. Math. Soc.*, 65:37–58, 1959.

[49] Renardy M. and Rogers R. *An introduction to partial differential equations*. Springer, 1993.

[50] J. C. Mason. Chebyshev polynomial approximations for the L-membrane eigenvalue problem. *SIAM J. Appl. Math.*, 15:172–186, 1967.

[51] R. Mathon and R. L. Johnston. The approximate solution of elliptic boundary-value problems by fundamental solutions. *SIAM J. Numer. Anal.*, 14(4):638–650, 1977.

[52] J. M. Melenk. Operator adapted spectral element methods i: harmonic and generalized harmonic polynomials. *Numerische Mathematik*, 84(1):35–69, 1999.

[53] S. N. Mergelyan. Uniform approximations to functions of a complex variable. *Uspehi Mat. Nauk.*, 7(2):31, 122 1952. English transl., Amer. Math. Soc. Transl. (1), No. 101.

[54] C. B. Moler. Accurate bounds for the eigenvalues of the laplacian and applications to rhombical domains. Technical Report CS-TR-69-121, Department of Computer Science, Stanford University, 1969.

[55] C. B. Moler and L. E. Payne. Bounds for eigenvalues and eigenfunctions of symmetric operators. *SIAM J. Numer. Anal.*, 5:64–70, 1968.

[56] R. Osserman. A note on hayman's theorem on the bass note of a drum. *Comm. Math. Helvetici*, 52:545–555, 1977.

[57] C. C. Paige. Computing the generalized singular value decomposition. *SIAM J. Sci. Statist. Comput.*, 7(4):1126–1146, 1986.

[58] C. C. Paige and M. A. Saunders. Towards a generalized singular value decomposition. *SIAM J. Numer. Anal.*, 18(3):398–405, 1981.

[59] C. C. Paige and M. Wei. History and generality of the CS decomposition. *Linear Algebra Appl.*, 208/209:303–326, 1994.

[60] M. H. Protter. Can one hear the shape of a drum? Revisited. *SIAM Review*, 29:185–197, 1987.

[61] Courant R. and Hilbert D. *Methods of mathematical physics*, volume I. Interscience, New York, 1953.

[62] T. Ransford. *Potential theory in the complex plane*, volume 28 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1995.

[63] J. K. Reid and J. E. Walsh. An elliptic eigenvalue problem for a re-entrant region. *SIAM J. Appl. Math.*, 13:837–850, 1965.

[64] N. L. Schryer. Constructive approximation of solutions to linear elliptic boundary value problems. *SIAM J. Numer. Anal.*, 9:546–572, 1972.

[65] G. W. Stewart and J. Sun. *Matrix perturbation theory.* Academic Press, Inc., 1990.

[66] G. Still. Computable bounds for eigenvalues and eigenfunctions of elliptic differential operators. *Numer. Math.*, 54, 1988.

[67] G. Still. *Defektminimierungsmethoden zur Lösung elliptischer Rand- und Eigenwertaufgaben.* Habilitation. University of Trier, 1989.

[68] G. Still. On density and approximation properties of special solutions of the Helmholtz equation. *Z. Angew. Math. Mech.*, 72(7):277–290, 1992.

[69] G. Still. Approximation theory methods for solving elliptic eigenvalue problems. *ZAMM Z. Angew. Math. Mech.*, 83(7):468–478, 2003.

[70] Strutt, J. W., Lord Rayleigh. *Theory of sound 2nd ed.* Dover, New York, 1945.

[71] J. G. Sun. Condition number and backward error for the generalized singular value decomposition. *SIAM J. Matrix Anal. Appl.*, 22(2):323–341 (electronic), 2000.

[72] J. Taylor. *Several complex variables with connections to algebraic geometry and Lie groups.* American Mathematical Society, 2002.

[73] M. Teytel. How rare are multiple eigenvalues? *Comm. Pure Appl. Math.*, 52:917–934, 1999.

[74] L. N. Trefethen. Ten digit algorithms. Numerical Analysis Group Research Reports NA-05/13, Oxford University Computing Laboratory, Oxford, United Kingdom, 2005.

[75] L. N. Trefethen and T. Betcke. Computed eigenmodes of planar regions. *AMS Contemporary Mathematics, to appear*, 2005.

[76] K. Uhlenbeck. Generic properties of eigenfunctions. *Amer. J. Math.*, 98:1059–1078, 1976.

[77] C. F. Van Loan. Generalizing the singular value decomposition. *SIAM J. Numer. Anal.*, 13(1):76–83, 1976.

[78] C. F. van Loan. A generalized SVD analysis of some weighting methods for equality constrained least squares. In *Matrix Pencils*, volume 973 of *Lecture Notes in Math.*, pages 245–262. Springer, New York, 1983.

[79] C. F. Van Loan. Computing the CS and the generalized singular value decompositions. *Numer. Math.*, 46(4):479–491, 1985.

[80] I. N. Vekua. *New Methods for Solving Elliptic Equations*. North-Holland, Amsterdam, 1967.

[81] E. Vergini. *Estudio Cuántico y Semiclásico de Bilares Clásicamente Caóticos*. PhD thesis, Universidad de Buenos Aires, 1995.

[82] E. Vergini and M. Saraceno. Calculation by scaling of highly excited states of billiards. *Phys. Rev. E*, 52(3):2204–2207, 1995.

[83] J. von Neumann and E. Wigner. Über das verhalten von eigenwerten bei adiabatischen prozessen. *Phys. Z.*, 30:467–470, 1929.

[84] J. L. Walsh. *Interpolation and approximation by rational functions in the complex domain, 3rd edition*, volume XX of *American Mathematical Society Colloquium Publications*. American Mathematical Society, 1960.

[85] W. Wasow. Asymptotic development of the solution of dirichlet's problem at analytic corners. *Duke Mathematical Journal*, 27:47–56, 1957.

[86] P. A. Wedin. On angles between subspaces of a finite dimensional inner product space. In *Matrix Pencils*, volume 973 of *Lecture Notes in Math.*, pages 263–285. Springer, New York, 1983.

[87] H. Zha. Computing the generalized singular values/vectors of large sparse or structured matrix pairs. *Numer. Math.*, 72(3):391–417, 1996.