Theory and Numerics of Spectral Value Sets

von

Eduardo Gallestey Alvarez

Dissertation
zur Erlangung des Grades eines Doktors der Naturwissenschaften
-Dr.rer.nat.-

Vorgelegt im Fachbereich 3 (Mathematik & Informatik)
der Universität Bremen
im August 1998

Die Dissertation lag dem Promotionsausschuß des Fachbereichs Mathematik/Informatik der Universität Bremen vor. An dem Promotionsverfahren haben Prof. Dr. Diederich Hinrichsen (Bremen) und Prof. Dr. Anthony Pritchard (Warwick) als Gutachter mitgewirkt. Das Kolloquium fand am 2. Oktober 1998 statt.

"Gedruckt mit Unterstützung des Deutschen Akademischen Austauschdienstes".

A mis padres, por su amor incondicional.

Acknowledgments

There are many people to thank for helping me during my stay at the University of Bremen. First, I want to thank the "Deutscher Akademischer Austauschdienst" for its (not only financial) support. Further, I am grateful to Dr. F. Wirth and, specially, to Dr. W. Manthey for their help with various parts of this thesis. Moreover, I would like to extend these thanks to my friends at the Institute for Dynamical Systems of the Bremen University for creating a stimulating atmosphere of nice and pleasant work. My gratitude also goes to Prof. Anthony Pritchard of the University of Warwick for useful discussions concerning my work and for accepting to be reviewer of this thesis. Furthermore, my supervisor Prof. Dr. Diederich "Didi" Hinrichsen is especially acknowledged for trusting in me, for the time and patience he has invested in teaching me how to convert my imprecise ideas into mathematics. Finally, I am extremely indebted to Sandra Leive, in the meantime my spouse: this thesis would have not been completed without her dedication and love.

Contents

1 Introduction **1**
1.1 Why spectral value sets? . . . 1
1.2 Brief historical remarks . . . 2
1.3 Concerns and structure of the thesis . . . 3

2 Spectral Value Sets in Finite Dimensional Spaces **5**
2.1 Preliminaries . . . 5
2.1.1 Reduced spectral value sets $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ . . . 6
2.1.2 Example: companion matrix . . . 7
2.2 Properties of $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$ . . . 8
2.2.1 Monotonicity . . . 9
2.2.2 Openness and boundedness . . . 10
2.2.3 Upper semicontinuity . . . 10
2.2.4 Asymptotic properties . . . 13
2.2.5 Smoothness of spectral contours . . . 16
2.3 Summary . . . 23

3 Calculation and Visualisation of Spectral Value Sets **25**
3.1 Existing methods . . . 25
3.1.1 Grid method . . . 25
3.1.2 Random perturbations . . . 26
3.1.3 Path following . . . 26
3.2 The SH algorithm . . . 27
3.2.1 Outline . . . 27
3.2.2 Discussion . . . 31
3.2.3 Performance . . . 32
3.3 Examples . . . 33
3.3.1 Matrix Grcar . . . 34
3.3.2 Matrix Fish . . . 34
3.3.3 Wilkinson matrix . . . 36
3.4 Summary . . . 38

4 Spectral Value Sets in Infinite Dimensional Spaces **39**
4.1 Preliminaries . . . 39
4.1.1 Some notations . . . 39

4.1.2 Main definitions and assumptions 40
4.1.3 The transfer function 41
4.1.4 Auxiliary lemma 41
4.2 Characterisation of spectral value sets 42
4.2.1 Closedness radius 43
4.2.2 Spectral value sets 46
4.2.3 C_g-stability radius 49
4.3 Summary 50

5 Abstract results on convergences of operators 51

5.1 Preliminaries 51
5.1.1 Notations 51
5.1.2 Standard convergence theorems 53
5.1.3 Normwise convergence of a product 54
5.1.4 Uniform convergence of operator valued functions 56
5.2 On convergence of discrete operators 58
5.2.1 Discrete operators 58
5.2.2 Normwise convergence 58
5.2.3 Strong convergence 59
5.3 Discrete-regular convergence 60
5.3.1 Preliminaries 60
5.3.2 Conditions for regular convergence in $\mathcal{L}(V)$. 63
5.3.3 Conditions for regular convergence in $\mathcal{L}(W, V)$. 66
5.3.4 Regular convergence in $\mathcal{L}(W, V_a)$ 69
5.3.5 Convergence of eigenelements 72
5.4 Summary 74

6 Approximation of Spectral Value Sets 75

6.1 Preliminaries 75
6.1.1 Norm of discrete operators 75
6.1.2 The role of normwise convergence 76
6.1.3 Compactness of the transfer function 77
6.2 On the approximation of transfer functions 79
6.2.1 Main result 79
6.2.2 Some corollaries 81
6.3 On approximation of the resolvent of Riesz operators 83
6.3.1 Preliminaries 84
6.3.2 Approximation of the resolvent operator 85
6.4 Projection Schemes 89
6.4.1 Preliminaries 89
6.4.2 Projections and discrete operators 91
6.4.3 Projection framework 92
6.4.4 Remarks on the computation of finite dimensional operators 99
6.5 Summary 100

7 Applications **101**
7.1 Delay operators . . . 101
7.1.1 Preliminaries . . . 101
7.1.2 Approximation in the case of delay system . . . 104
7.1.3 Example: Damped oscillator . . . 109
7.2 The Orr-Sommerfeld operator . . . 111
7.2.1 Preliminaries . . . 111
7.2.2 Spectral properties . . . 114
7.2.3 Unstructured perturbations . . . 117
7.2.4 Approximation scheme . . . 118
7.2.5 Structured perturbations . . . 123

A Convergence of level curves **127**

Bibliography **131**

Chapter 1

Introduction

1.1 Why spectral value sets?

Spectral analysis plays a basic role in many areas of applied mathematics and, in particular, in control theory. On the other hand, often mathematical models do not represent *exactly* the dynamics of a physical plant. Consequently, since the spectrum of an operator may be highly sensitive to perturbations, whenever knowledge of the spectra is a vital requirement, one should not only examine the spectrum of the model operator but also the spectrum of nearby ones.
In perturbation theory and numerical analysis there are a number of results available which yield bounds for the variation of the spectrum of a perturbed matrix or operator

$$A \rightsquigarrow A + \Delta$$

in terms of the perturbation. Typical examples of this kind of theorems are the Gershgorin Circles Theorem [27, Theorem 7.2-1] and the Bauer-Fike Theorem [27, Theorem 7.2-2].
"Bauer-Fike like" theorems are especially interesting for applications in control theory and numerical analysis. The reason is that in the bounds are given in terms of the *maximal size* of the perturbation. Their drawback is that these bounds may be very conservative, specially in the case of highly nonnormal matrices and/or operators. Moreover, they have another limitation: if, as it is common in applications, some of the entries of the matrix A are *fixed* a priori, there is no easy way to incorporate this information in the bounds.
In order to study such problems Hinrichsen and Pritchard suggested the use of *structured perturbations* [44]

$$A \rightsquigarrow A + D\Delta E. \tag{1.1}$$

By introducing perturbation structures through the matrices D and E, it is often possible to perturb *just* those entries of A which are uncertain or, for example, to introduce some kind of scaling in the perturbation.
We want to illustrate the flexibility and convenience of the perturbation structures (1.1) with one example. Consider a linear time invariant system

$$\begin{aligned} \dot{x} &= A_0 x + Du \\ y &= Ex \end{aligned}$$

where A_0, D, E are matrices of appropriated dimensions. A usual task is the design of a feedback matrix F_0 such that the closed loop system

$$\dot{x} = Ax, \quad A := A_0 + DF_0E$$

has poles in a predetermined domain C_g of the complex plane. We note immediately that, even if we are able to find such a F_0, small errors in the implementation of the control can destroy its good properties. Thus, it is natural to ask how sensitive is the spectrum of A to small errors Δ on the matrix F_0. This question can be tackled by investigating whether the spectra of matrices of the form

$$A_0 + D(F_0 + \Delta)E = A_0 + DF_0E + D\Delta E = A + D\Delta E$$

remains in the desired region C_g for reasonably small perturbation sizes.
These ideas motivate the following mathematical approach [42].

Definition 1.1.1 Let

$$(A, D, E) \in L_{n,l,q}(\mathbb{K}) := \mathbb{K}^{n\times n} \times \mathbb{K}^{n\times l} \times \mathbb{K}^{q\times n},$$

where $\mathbb{K} = \mathbb{C}, \mathbb{R}$, and denote by $\sigma(A)$ the spectrum of A. Let $\rho > 0$. The associated *spectral value set* of level ρ, $\sigma_{\mathbb{K}}(A, D, E; \rho)$, is the set

$$\sigma_{\mathbb{K}}(A, D, E; \rho) = \bigcup_{\Delta\in\mathbb{K}^{l\times q},\, \|\Delta\|<\rho} \sigma(A + D\Delta E). \tag{1.2}$$

In the sequel, we shall say that the spectral value sets are *unstructured* if $l = q = n$ and the matrices D and E are the identity matrix.

1.2 Brief historical remarks

Before we go to mathematics, let us make some comments on the history of spectral value sets. The unstructured real case $\sigma_{\mathbb{R}}(A, I, I; \rho)$ was studied by Hinrichsen and Pritchard as early as in [45]. In that paper some bounds were found and the behavior of the sets under similarity transformations of A was investigated. In [42] Hinrichsen and Kelb proved a theorem which permitted the characterisation and calculation of $\sigma_{\mathbb{R}}(A, D, E; \rho)$ when $l = 1$. Later, in [43], the same authors analysed the general case $\sigma_{\mathbb{R}}(A, D, E; \rho)$ in the 2-norm using the formula developed by Qiu and et al. [63] for the real stability radius: it turns out that the main role in the analysis of *real* spectral value sets in the 2-norm is played by the function

$$f : \mathbb{C} \setminus \sigma(A) \to \mathbb{R}_+, \quad s \mapsto \inf_{\gamma\in(0,1]} \sigma_2 \begin{pmatrix} \operatorname{Re} G(s) & -\gamma \operatorname{Im} G(s) \\ \gamma^{-1} \operatorname{Im} G(s) & \operatorname{Re} G(s) \end{pmatrix}, \quad G(s) := E(sI - A)^{-1}D,$$

where $\sigma_2(P)$ denotes the second singular value of any matrix P. This function, which is not even continuous, is the subject of current investigations [52], [4]. Its evaluation is difficult.
On the other hand, *complex* spectral value sets $\sigma_{\mathbb{C}}(A, D, E; \rho)$ were analysed in [42] and this dissertation should be understood as a natural extension of the results obtained in this paper. Finally, we stress that complex spectral value sets are important wherever the notion of

spectrum of *non-normal* matrices or operators are used, in particular, in numerical analysis. *Unstructured* complex spectral values sets $\sigma_{\mathbb{C}}(A, I, I; \rho)$ have been analysed independently by Godunov [25], Trefethen [79] and others, under the names of spectral portraits or pseudospectra. The pseudospectra has found application in the stability analysis of the method of lines [68], in the stability of spectral methods [67], in hydrodynamical stability [66] and in the study of the transitory behavior of the solution of linear differential equations [68].

1.3 Concerns and structure of the thesis

In this thesis we consider complex spectral value sets ($\mathbb{K} = \mathbb{C}$). We shall address the following topics:

1. In Chapter 2 a characterisation of the function

$$\psi : \mathbb{C} \setminus \sigma(A) \to \mathbb{R}_+, \quad s \mapsto \|E(sI - A)^{-1}D\|,$$

   is given and some of its properties are investigated. With the help of these results, we characterise $\sigma_{\mathbb{C}}(A, D, E; \rho)$ and its boundary.

2. Chapter 3 deals with the calculation of spectral value sets. First the existing algorithms are briefly reviewed, then we present a new algorithm for calculating spectral value sets under complex perturbations. Examples which show the effectiveness of the proposed method are also presented.

3. Spectral value sets can also be defined when A, D, and E are linear operators acting on Banach spaces, so Chapter 4 deals with spectral value sets of infinite dimensional systems. We study some of their properties and give a characterisation of these sets in terms of the norm of the associated transfer function $s \mapsto G(s)$, $G(s) = ER(s, A)D$, where $R(s, A)$ denotes the resolvent operator of A. This characterisation is an infinite dimensional analogue of the results obtained in [42] for the matrix case. Furthermore, in this chapter we study two related objects: closedness radius and C_g-stability radius.

4. One of the difficulties in the infinite dimensional case is the evaluation the quantity $\|G(s)\|$. The natural approach in solving this problem is to approximate the operators by finite dimensional ones. Thus, we investigate conditions which must be imposed on A, D, E and on the approximation methods which guarantee uniform approximations of the map $s \mapsto \|G(s)\|$ in given compact sets of $\mathbb{C}$. With this aim in mind, Chapter 5 presents some abstract approximation results useful in the study of the approximation schemes to be presented in Chapter 6.

5. The aim of Chapter 7 is to show applications of the theory developed in the thesis. These investigations imply solving difficult numerical problems, thus the examples will be also a test for the quality and performance of our numerical algorithm.

   In the first section of this chapter we analyse robustness issues of delay operators under certain structured perturbations. We shall see that our theory can be applied to this case in a straightforward manner.

The second part of this chapter is more ambitious and deals with spectral value sets of the Orr-Sommerfeld operator [19]. The Orr-Sommerfeld operator plays a central role in hydrodynamical stability theory and the investigation of stability and robustness issues of this highly non-normal operator is a classical problem of applied mathematic. Recently, the topic has received new momentum with the introduction in this field (Reddy et.al. [66]) of "pseudospectra" ideas. Up until now, only the effect of unstructured bounded perturbations has been investigated. As an application of our theoretical results we shall study the robustness of the Orr-Sommerfeld operator to certain structured perturbations which take into account neglected nonlinearities. We shall see that new interesting results can be achieved with the help of spectral value sets.

Chapter 2

Spectral Value Sets in Finite Dimensional Spaces

The aim of this chapter is to characterise spectral value sets $\sigma_{\mathbb{C}}(A, D, E; \rho)$ in the matrix case. We begin with a known result due to Hinrichsen and Kelb [42] which relates spectral value sets with the map obtained by considering the *operator norm* of a certain *transfer function*. Further, we investigate the properties of this map and relate them with those of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ and its boundary $\partial\sigma_{\mathbb{C}}(A, D, E; \rho)$.
For completeness we recall the definition of operator (or induced) norms: if the vector spaces $\mathbb{C}^l$ and $\mathbb{C}^q$ are provided with norms $\|\cdot\|_{\mathbb{C}^l}$ and $\|\cdot\|_{\mathbb{C}^q}$, respectively, the *operator norm* of a matrix $\Delta \in \mathbb{C}^{l\times q}$ with respect to $\|\cdot\|_{\mathbb{C}^l}$ and $\|\cdot\|_{\mathbb{C}^q}$ is defined as

$$\|\Delta\| = \sup_{\|y\|_{\mathbb{C}^q}=1} \|\Delta y\|_{\mathbb{C}^l}. \tag{2.1}$$

2.1 Preliminaries

The foundations for investigations related to the sets $\sigma_{\mathbb{C}}(A, D, E; \rho)$ were built by Hinrichsen and Kelb in [42]. As a main result they proved a theorem which is the backbone of our work. This theorem, presented here in its original form, makes clear the close relationship between spectral value sets and the operator norm of the transfer function

$$G : \rho(A) \to \mathbb{C}^{q\times l}, \quad s \mapsto E(sI - A)^{-1}D, \tag{2.2}$$

where

$$\rho(A) := \mathbb{C} \setminus \sigma(A)$$

denotes the *resolvent set* of A. In the sequel $\mathrm{cl}\,(\Omega)$ means *closure* of a subset Ω in $\mathbb{C}$.

Theorem 2.1.1 *Let $(A, D, E) \in L_{n,l,q}(\mathbb{C})$ and $G(s) = E(sI - A)^{-1}D$ be the associated transfer function. Additionally, let $\rho > 0$ and $\|\cdot\|$ be any operator norm on $\mathbb{C}^{q\times l}$. Then*

1. *$\sigma_{\mathbb{C}}(A, D, E; \rho) \setminus \sigma(A)$ is a bounded open subset of $\mathbb{C}$ and*

$$\sigma_{\mathbb{C}}(A, D, E; \rho) = \sigma(A) \cup \{s \in \rho(A); \|G(s)\| > \rho^{-1}\}. \tag{2.3}$$

2. *The boundary of the set* $\sigma_{\mathbb{C}}(A, D, E; \rho) \setminus \sigma(A)$ *in* $\rho(A)$ *is given by*

$$\partial(\sigma_{\mathbb{C}}(A, D, E; \rho) \setminus \sigma(A)) = \{s \in \rho(A); \|G(s)\| = \rho^{-1}\} =: C_\rho. \tag{2.4}$$

3. $\sigma_{\mathbb{C}}(A, D, E; \rho)$ *is the union of* $\sigma(A)$ *and of those connected components of* $\mathbb{C} \setminus \mathrm{cl}\,(C_\rho)$ *which contain at least one pole of* $G(s)$.

Remark 2.1.2 We do not prove Theorem 2.1.1. One reason is that a more general version of it will be proved in Chapter 4.

Remark 2.1.3 Theorem 2.1.1 is extremely important: it gives a *computable formula* for the calculation of spectral value sets.

Remark 2.1.4 The set C_ρ defined in (2.4) is called the *spectral contour of level* ρ. Point 3 in Theorem 2.1.1 yields a rule for determining $\sigma_{\mathbb{C}}(A, D, E; \rho)$ from knowledge of C_ρ.

Remark 2.1.5 The reader should be careful with intuitive ideas on the sets C_ρ. Although they are called "contours", they might have holes at certain points of $\sigma(A)$!

2.1.1 Reduced spectral value sets $\vartheta_{\mathbb{C}}(A, D, E; \rho)$

It is well known [16, Theorem 17.5] that not every eigenvalue of A is automatically a pole of the transfer function $G(s) = E(sI - A)^{-1}D$. This is the reason for the somewhat involved results "modulo $\sigma(A)$" of Theorem 2.1.1. We could obtain more compact statements by assuming controllability and observability of the triplet (A, D, E). Unfortunately, this assumption is not natural in this context and we must take this fact into account in our results.
Nevertheless, a detailed analysis shows that considering noncontrollable and/or unobservable modes of A is not useful. Indeed, if an eigenvalue λ of A is a noncontrollable and/or unobservable mode of the triplet (A, D, E) then λ cannot be "moved" by means of the structured perturbations. In other words, λ is invariant to such perturbations. This is a consequence of well known results on static linear feedback design [51, p. 205].
In view of these arguments we propose an alternative approach. Let

$$(A_{\min}, D_{\min}, E_{\min}) \in L_{k,l,q}(\mathbb{C}), \quad k \leq n, \tag{2.5}$$

be a *minimal realisation* of G [16, Definition 21.12]. The following objects are important for our discussion:

$$P(G) := \sigma(A_{\min}), \quad \rho(A_{\min}) := \mathbb{C} \setminus P(G). \tag{2.6}$$

The set $P(G)$ is well defined because minimal realisations are similar [51, Theorem 6.2-4]. Note also that $P(G) \subset \sigma(A)$.

Definition 2.1.6 Let $(A, D, E) \in L_{n,l,q}(\mathbb{C})$ and $G(s) = E(sI - A)^{-1}D$ be the associated transfer function. Additionally, let $\rho > 0$ and $(A_{\min}, D_{\min}, E_{\min}) \in L_{k,l,q}(\mathbb{C})$, $k \leq n$, be a minimal realisation of G. The *reduced spectral value set* of level ρ, denoted $\vartheta_{\mathbb{C}}(A, D, E; \rho)$, is the set

$$\vartheta_{\mathbb{C}}(A, D, E; \rho) := \{s \in \mathbb{C} \;:\; \psi(s) > \rho^{-1}\}, \tag{2.7}$$

where

$$\psi : \mathbb{C} \to \mathbb{R}_+ \cup \{+\infty\}, \quad s \mapsto \begin{cases} \|E_{\min}(sI - A_{\min})^{-1} D_{\min}\| & \text{if} \quad s \in \rho(A_{\min}) \\ +\infty & \text{if} \quad s \in P(G) \end{cases} . \tag{2.8}$$

Note that $G \equiv 0$ implies $\vartheta_{\mathbb{C}}(A, D, E; \rho) = \emptyset$ for every $\rho > 0$. Furthermore, it follows from Definition 2.1.6 and Theorem 2.1.1 that

$$P(G) \subset \vartheta_{\mathbb{C}}(A, D, E; \rho), \quad \vartheta_{\mathbb{C}}(A, D, E; \rho) \subset \sigma_{\mathbb{C}}(A, D, E; \rho).$$

Moreover,

$$\sigma_{\mathbb{C}}(A, D, E; \rho) = [\sigma(A) \setminus P(G)] \cup \vartheta_{\mathbb{C}}(A, D, E; \rho) \tag{2.9}$$

and in the case where $P(G) = \sigma(A)$ we have $\sigma_{\mathbb{C}}(A, D, E; \rho) = \vartheta_{\mathbb{C}}(A, D, E; \rho)$.
Also we *redefine* the notion of "spectral contour" making it more appealing. The set

$$C_\rho := \{s \in \mathbb{C} \; : \; \psi(s) = \rho^{-1}\} \tag{2.10}$$

is called *spectral contour of level* ρ. It can be shown that $\partial\vartheta_{\mathbb{C}}(A, D, E; \rho) = C_\rho$.
In our opinion it helps to make the things clearer if through this section we consider *reduced* spectral value sets instead of spectral value sets in the sense of the original definition. We stress that $\sigma_{\mathbb{C}}(A, D, E; \rho)$ and $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ are *identical* modulo $\sigma(A) \setminus P(G)$. Moreover, for calculations the expression of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ given in (2.3) is easier to apply than (2.9). Thus, the use of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ will be restricted to the theoretical considerations given below. In concrete examples, if $\sigma_{\mathbb{C}}(A, D, E; \rho) = \vartheta_{\mathbb{C}}(A, D, E; \rho)$, we shall prefer the first notation.

2.1.2 Example: companion matrix

Before we begin with a formal study of the spectral value sets, let us develop some intuition about $\sigma_{\mathbb{C}}(A, D, E; \rho)$ with the help of an example. Consider the matrices

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 5040 & -13068 & 13132 & -6769 & 1960 & -322 & 28 \end{pmatrix} \tag{2.11}$$

$$D = I_7, \qquad E = I_7. \tag{2.12}$$

Here, and in the sequel, I_n denotes the $n \times n$ identity matrix.
The matrix (2.11) is well known in numerical analysis [83]: it is the companion matrix of the monic polynomial with roots $1, \ldots, 7$. Clearly, the transfer function corresponding to (A, I_7, I_7) is

$$G(s) = (sI_7 - A)^{-1}, \quad s \in \rho(A),$$

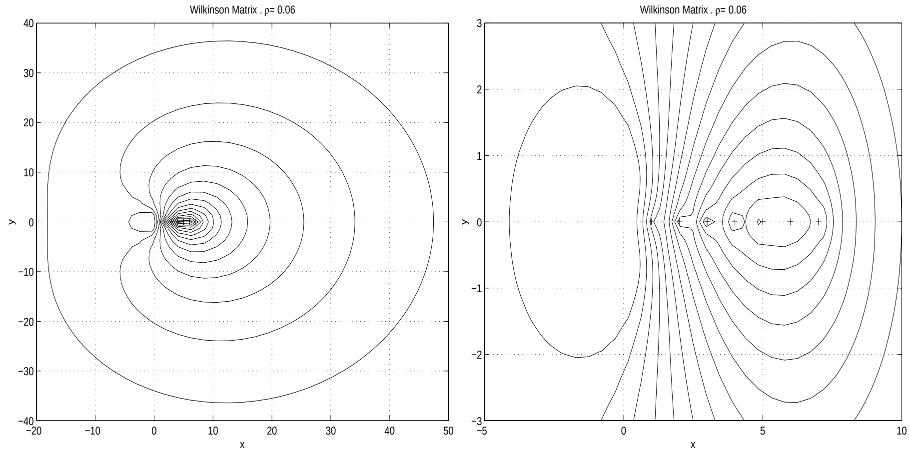

Figure 2.1: Spectral value sets of the matrices (2.11) and (2.12).

and $P(G) = \sigma(A)$.

In Figure 2.1 we depict $\sigma_{\mathbb{C}}(A, D, E; \rho) = \vartheta_{\mathbb{C}}(A, D, E; \rho)$ for some values of ρ and the matrix norm given by the largest singular value. For clarity in the pictures, the concrete values of ρ are omitted. Two pictures are presented, the second being a zoom of the first one. The spectrum of A is represented by crosses. A detailed description of the method used in these calculations is given in Chapter 3.

In Figure 2.1 one observes that

1. As ρ increases from zero to infinity the corresponding (reduced) spectral value sets grow and take different forms: for small values of ρ the sets are a sort of oval regions around the eigenvalues, while with increasing ρ, these regions merge until the set $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ becomes a connected open bounded subset of $\mathbb{C}$.

2. The sets $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ are open and bounded.

3. The interior of the oval region situated at the left in the second picture, does not contain poles of G. Thus, according to Point 3 of Theorem 2.1.1, we conclude that this set does not belong to $\vartheta_{\mathbb{C}}(A, D, E; \rho)$. It follows that $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ is not necessarily simply connected even for large values of ρ.

4. The spectral contours C_ρ possess some kind of smoothness properties.

Our aim in the sequel is to show that these "experimental results" are, in fact, representative and largely correct.

2.2 Properties of $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$

The key point in the investigation of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ is to understand the behavior of the function displayed in (2.8). The function (2.8) corresponding to the matrices (2.11), (2.12) is depicted in Figure 2.2 in form of a three dimensional plot in logarithm scale. Again, we make some observations based on our previous experiment. The aim is to attract the attention of the reader to the most important features.

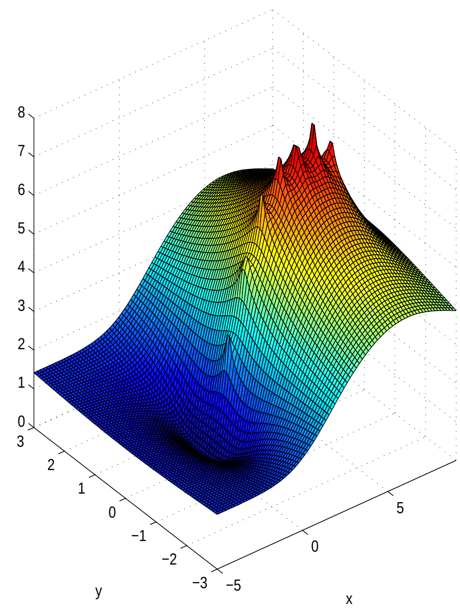

Figure 2.2: Function ψ in the rectangle $[-5, 9] \times [-3, 3]$.

1. ψ is a continuous function. The set $P(G)$, in this case equal to $\sigma(A)$, can be easily recognized as peaks in the graph. In a neighborhood of $\lambda \in P(G)$, $\psi(s)$ becomes large, while it tends to zero as $|s| \to \infty$.

2. ψ may have minima in $\mathbb{C}$ but has no maxima away of $P(G)$, where it is equal to $+\infty$. Further, ψ seems to enjoy some differentiability properties.

In this section we show that these experimental observations are correct and relate them to the properties of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ listed in the preliminaries. We proceed in each subsection as follows: first, we *prove* a property of ψ and thereafter we interpret this property with regard to the corresponding (reduced) spectral value sets.

2.2.1 Monotonicity

In Proposition 2.2.1 we state a simple property of the map $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$.

Proposition 2.2.1 *The set valued map $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$ is monotonic, i.e.,*

$$\rho_1 < \rho_2 \implies \vartheta_{\mathbb{C}}(A, D, E; \rho_1) \subset \vartheta_{\mathbb{C}}(A, D, E; \rho_2).$$

Proof: By Definition 2.1.6, if $s \in \vartheta_{\mathbb{C}}(A, D, E; \rho_1)$ then $\psi(s) > \rho_1^{-1}$. Thus

$$\psi(s) > \rho_2^{-1}, \text{ for any } \rho_2 > \rho_1$$

and we conclude that $s \in \vartheta_{\mathbb{C}}(A, D, E; \rho_2)$ as well. □

Remark 2.2.2 This result holds also for $\rho \mapsto \sigma_{\mathbb{C}}(A, D, E; \rho)$.

2.2.2 Openness and boundedness

The fact that for every $\rho > 0$ the set $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ is open and bounded has been stated implicitly in Theorem 2.1.1. Nevertheless, we state and prove it here for completeness. For the proof, we shall need a lemma which is a direct consequence of the fact that the *rational* transfer function G has poles at $P(G)$ and is strictly proper.

Lemma 2.2.3 *The function ψ defined in (2.8) is continuous. Moreover,*

$$\lim_{|s| \to \infty} \psi(s) = 0. \tag{2.13}$$

With the help of this lemma we can show

Proposition 2.2.4 *For every $\rho > 0$, the set $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ is an open bounded subset of $\mathbb{C}$.*

Proof: Let $s_0 \in \vartheta_{\mathbb{C}}(A, D, E; \rho)$. Then, by definition, $\psi(s_0) > \rho^{-1}$. Since ψ is continuous, there exists a neighborhood $O(s_0) \subset \mathbb{C}$ such that $\psi(s) > \rho^{-1}$ for every $s \in O(s_0)$. Thus, $O(s_0) \subset \vartheta_{\mathbb{C}}(A, D, E; \rho)$ and the openness of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ follows. The boundedness of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ is a consequence of (2.13). □

Remark 2.2.5 Definition 2.1.6 implies that the set of zeros of $G(\cdot)$, i.e.,

$$Z(G) := \{s \in \mathbb{C} \,:\, G(s) = 0\} \tag{2.14}$$

does not belong to $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ for any value of $\rho > 0$:

$$Z(G) \cap \vartheta_{\mathbb{C}}(A, D, E; \rho) = \emptyset, \quad \forall\, \rho > 0.$$

We conclude that the connected components of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ are, in general, not simply connected. However, since G is a *rational* matrix function, $Z(G)$ is a finite set of isolated points whenever $G \not\equiv 0$:

$$Z(G) = \{s_i \in \mathbb{C},\ i = 1, \ldots, k_{Z(G)}\}, \quad k_{Z(G)} < \infty.$$

Finally, we note that the set $Z(G)$ differs from the usual definition of zero of a transfer function used in control theory. The usual definition accounts, roughly speaking, just for rank losses of G [51, p. 448].

2.2.3 Upper semicontinuity

The results of this paragraph are based on the *subharmonicity* of ψ on open subsets of $\rho(A_{\min})$ [10, Definition 5.9].

Definition 2.2.6 Let Ω be an open domain in $\mathbb{C}$. $f : \bar{\Omega} \to \mathbb{R}$ is called *subharmonic in* Ω if

1. f is continuous on $\bar{\Omega}$.

2. For any $s_0 \in \Omega$ and any $\rho > 0$ such that the open disc $|s - s_0| < \rho$ is contained in Ω, the inequality

$$f(s_0) \leq \frac{1}{2\pi} \int_0^{2\pi} f(s_0 + re^{i\theta}) d\theta$$

holds for every $0 < r \leq \rho$.

We need the following theorem [41, Theorem 3.13.1].

Theorem 2.2.7 *Let X be a Banach space and Ω a open bounded subset of $\mathbb{C}$. Consider a function $f : \bar{\Omega} \to X$ holomorphic in Ω and continuous on $\bar{\Omega}$. Then $\|f\|$ is a subharmonic function in Ω.*

Theorem 2.2.7 can be applied to the function defined in (2.8).

Proposition 2.2.8 *Let $\Omega \subset \mathbb{C}$ be an open bounded set such that $\bar{\Omega} \subset \rho(A_{\min})$. Then the function ψ (2.8) is subharmonic in Ω.*

Proof: G is holomorphic on $\rho(A_{\min})$ and $\mathbb{C}^{q \times l}$ endowed with an operator norm is a Banach space. Thus, Theorem 2.2.7 implies the subharmonicity of ψ in Ω. □

Subharmonic functions are well known in Complex Analysis. As the name suggests, these are functions which have harmonic majorants. In this work we are interested only in one elementary property of the subharmonic functions: the fact that, in analogy to harmonic functions, a Maximum Principle holds [10, Corollary 5.9].

Theorem 2.2.9 *Let Ω be a bounded domain in $\mathbb{C}$ and $\partial\Omega$ be its boundary. Suppose f is subharmonic in Ω. Then*

$$\sup_{s \in \Omega} f(s) = \max_{\zeta \in \partial\Omega} f(\zeta).$$

So we have

Corollary 2.2.10 *Under the hypothesis of Proposition 2.2.8 the following equality holds*

$$\sup_{s \in \Omega} \psi(s) = \max_{\zeta \in \partial\Omega} \psi(\zeta).$$

Corollary 2.2.10 is of crucial importance for the next chapter. Moreover, it also implies

Corollary 2.2.11 *ψ has no local maxima unless it is constant. The last can occur only if $G \equiv 0$.*

Proof: The first statement is, clearly, a consequence of Corollary 2.2.10. On the other hand, the second assertion can be proved as follows. We must show that ψ can not be a constant *different* from zero through some open $\Omega \subset \rho(A_{\min})$. Indeed, let us suppose that this is the case. We choose some $s_0 \in \Omega$ and vectors u and y of suitable dimensions and of norm equal to one such that

$$\psi(s_0) = \|G(s_0)\| = |y^* G(s_0) u| > 0, \quad s_0 \in \Omega, \tag{2.15}$$

and see that the scalar function

$$g : \rho(A_{\min}) \to \mathbb{C}, \quad s \mapsto y^* G(s) u,$$

is holomorphic. Moreover, since ψ is constant, $s \mapsto |y^* G(s) u|$ of g has a local maximum at s_0. Using classical complex analysis we see that g *must* be constant through $\rho(A_{\min})$. Finally, due to (2.13), we conclude that g is identically equal to zero. This contradicts (2.15). □

Remark 2.2.12 We shall see in Chapter 4 that this property does not necessarily hold for transfer functions of infinite dimensional systems.

With regard to spectral value sets Corollary 2.2.11 implies

Proposition 2.2.13 *Let $G \not\equiv 0$. Then the map $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$ is upper semicontinuous in the sense of Hausdorff [3, p. 25], i.e., for every $\epsilon > 0$ there exists a $\delta > 0$ such that*

$$\vartheta_{\mathbb{C}}(A, D, E; \hat{\rho}) \subset O_\epsilon(\vartheta_{\mathbb{C}}(A, D, E; \rho)), \quad \forall\, \hat{\rho} \in (\rho - \delta, \rho + \delta), \tag{2.16}$$

where "O_ϵ" denotes an ϵ-neighborhood of a set.

Proof: Let $\epsilon > 0$ be fixed and consider an ϵ-neighborhood $O_\epsilon(\vartheta_{\mathbb{C}}(A, D, E; \rho))$ of the reduced spectral value set $\vartheta_{\mathbb{C}}(A, D, E; \rho)$. By Proposition 2.2.1, the map $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$ is monotonic. Thus, for every $\delta > 0$, $\rho - \delta > 0$, we have the relationship

$$\vartheta_{\mathbb{C}}(A, D, E; \rho - \delta) \subset \vartheta_{\mathbb{C}}(A, D, E; \rho) \subset O_\epsilon(\vartheta_{\mathbb{C}}(A, D, E; \rho)).$$

It follows that only increments of ρ are of interest.
Now, let us suppose that there exist $\rho > 0$ and $\epsilon > 0$ such that for every $\delta > 0$

$$\vartheta_{\mathbb{C}}(A, D, E; \rho + \delta) \not\subset O_\epsilon(\vartheta_{\mathbb{C}}(A, D, E; \rho)). \tag{2.17}$$

Simple considerations using the continuity of ψ show that (2.17) implies the existence of a point $s_0 \in \mathbb{C}$ such that

$$\psi(s_0) = \rho^{-1}, \text{ but } s_0 \notin O_\epsilon(\vartheta_{\mathbb{C}}(A, D, E; \rho)).$$

It follows that we can find a neighborhood $O(s_0)$ of the point s_0 which does not intersect $\vartheta_{\mathbb{C}}(A, D, E; \rho)$:

$$O(s_0) \cap \vartheta_{\mathbb{C}}(A, D, E; \rho) = \emptyset.$$

Thus, by definition,

$$\psi(s) \le \rho^{-1}, \quad \forall\, s \in O(s_0)$$

and we conclude that s_0 is a local maximum of ψ. This is a contradiction to Corollary 2.2.11.
□

Remark 2.2.14 Proposition 2.2.13 is the formal statement of the following (intuitively) obvious fact: the sets $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ do not "expand" suddenly for small increments of ρ. Furthermore, the possible existence of local minima of ψ implies that the connected components of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ are not necessarily simply connected.

2.2.4 Asymptotic properties

Let us study now the asymptotic properties of the map $\rho \mapsto \vartheta_{\mathbb{C}}(A, D, E; \rho)$, that is, its behavior for small and large values of ρ. The expected result is

1. For small sizes of the perturbations, the reduced spectral value sets are small regions around the eigenvalues of $A_{\min}$.

2. For large ρ we expect $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ to occupy large regions of the complex plane.

Useful formulae

In order to study the function ψ in a small neighborhood of the poles $P(G)$ and at infinity, we shall make use of results on the resolvent operator of $A_{\min}$:

$$R(s) := (sI - A_{\min})^{-1}, \quad s \in \rho(A_{\min}).$$

Let $\lambda_i \in P(G) = \sigma(A_{\min})$ and let Γ_i be a (positively oriented) Jordan contour which surrounds λ_i and contains no other points of $\sigma(A_{\min})$ in its interior region. Then the matrix

$$P_i := \frac{-1}{2i\pi} \int_{\Gamma_i} R(s) ds \tag{2.18}$$

is called the *spectral projector* associated to λ_i. Further,

$$D_i := (A_{\min} - \lambda_i I) P_i,$$

is called the *nilpotent matrix* associated to λ_i. These matrices do not depend on the choice of Γ_i.

Let d be the number of *different* eigenvalues of $A_{\min}$. Then there are d matrices P_i and D_i, $i = 1, \dots, d$ and it holds that $R(s)$ is given by [53, I.5.3]

$$R(s) = \sum_{i=1}^{d} \left(\frac{P_i}{s - \lambda_i} - \sum_{k=1}^{l_i - 1} \frac{D_i^k}{(s - \lambda_i)^{k+1}} \right), \quad \forall\, s \in \rho(A_{\min}), \tag{2.19}$$

where l_i is the maximal size of a Jordan block associated with λ_i.

Finally we write down the series expansion of $R(s)$ in a neighborhood of infinity [53, I.5.2]:

$$R(s) = \sum_{k=0}^{\infty} \frac{A_{\min}^k}{s^{k+1}}. \tag{2.20}$$

This series converges absolutely for every $s \in \mathbb{C}$ such that

$$|s| > \gamma(A_{\min}) := \max_{s \in \sigma(A_{\min})} |s|.$$

The expressions (2.19), (2.20) give us useful representations of $G(s)$.

$\vartheta_{\mathbb{C}}(A,D,E;\rho)$ in a neighborhood of a pole

Let us suppose, for simplicity, that the spectrum of $A_{\min} \in \mathbb{C}^{k\times k}$ consists of simple eigenvalues. Then, by (2.19), we have

$$G(s) = \sum_{i=1}^{k} \frac{EP_iD}{s-\lambda_i}, \quad s \in \rho(A_{\min}). \tag{2.21}$$

Since we have chosen $A_{\min}$ to be the system matrix of a minimal realisation of G, all the EP_iD, $i=1,\ldots,k$, are different from zero. Then, for each $i=1,\ldots,k$, the dominant term in the sum (2.21) in a small enough neighborhood $O(\lambda_i) \subset \mathbb{C}$ of the eigenvalue λ_i is the ith summand. Thus, we conclude that the ρ^{-1} level curves of ψ which lie in $O(\lambda_i)$ can be *approximated* by a circle with center in λ_i and radius $\rho\|EP_iD\|$. The proposition below is the formal statement of this (intuitively obvious!) fact.

Proposition 2.2.15 *Suppose that the eigenvalues of $A_{\min} \in \mathbb{C}^{k\times k}$ are simple. Further, for a given $\rho > 0$, define*

$$\vartheta^0_{\mathbb{C}}(A,D,E;\rho) := \cup_{i=1}^{k}\{s \in \mathbb{C} \;:\; |s-\lambda_i| < \rho\|EP_iD\|\}.$$

Then the set valued map $\rho \mapsto \vartheta_{\mathbb{C}}(A,D,E;\rho)$ is such that

$$\lim_{\rho\to 0} \frac{1}{\rho}\, d_H(\mathrm{cl}\,(\vartheta_{\mathbb{C}}(A,D,E;\rho)), \mathrm{cl}\,(\vartheta^0_{\mathbb{C}}(A,D,E;\rho))) = 0.$$

where $d_H(\cdot,\cdot)$ denotes the Hausdorff distance (Definition A.0.9, Appendix A).

The proof of this result is omitted because it would be very technical and, at least in our opinion, of less interest. However, its main idea is quite instructive and is discussed below. For a fixed $i \in \{1,2,\ldots,k\}$ consider an open bounded set $\Omega \subset \mathbb{C}$ such that

$$\lambda_i \in \Omega \text{ and } \bar{\Omega} \subset \mathbb{C} \setminus (\sigma(A_{\min}) \setminus \{\lambda_i\}).$$

Then there exists $M < \infty$ with the property (see (2.21))

$$\max_{s\in\bar{\Omega}} \| \sum_{\substack{j=1\\ j\neq i}}^{k} \frac{EP_jD}{s-\lambda_j} \| < M < \infty.$$

Now, one multiplies (2.21) by $\rho(s-\lambda_i)$, where $\rho \in (0, 1/M)$ and $s \in C_\rho \cap \Omega$. Note that $C_\rho \cap \Omega \neq \emptyset$ for ρ small enough. Then, after some straightforward transformations, one obtains the following inequalities:

$$\frac{\rho\|EP_iD\|}{1+\rho M} \leq |s-\lambda_i| \leq \frac{\rho\|EP_iD\|}{1-\rho M}.$$

These relationships can be written as

$$\rho\|EP_iD\| - o(\rho) \leq |s-\lambda_i| \leq \rho\|EP_iD\| + o(\rho), \tag{2.22}$$

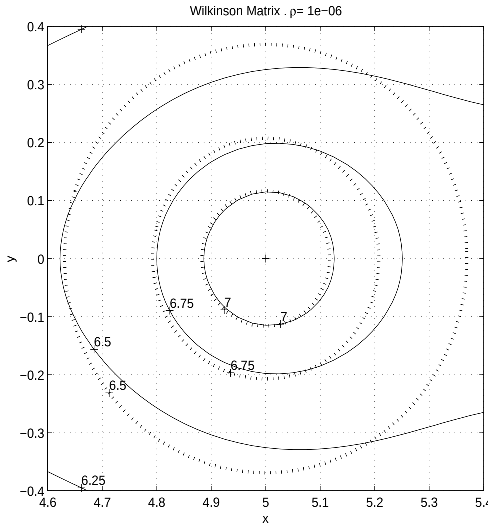

Figure 2.3: Behavior of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ for small perturbations.

where

$$o : \mathbb{R}_+ \to \mathbb{R}_+, \quad \lim_{\rho \to 0} \frac{o(\rho)}{\rho} = 0.$$

The inequalities (2.22) "prove" Proposition 2.2.15.
Figure 2.3 illustrates this result for the matrices (2.11), (2.12) in a neighborhood of the eigenvalue $\lambda_5 = 5$. The spectral contours C_ρ are represented by continuous lines while the circles of interest are depicted using dotted lines. The spectral projector P_5 has been calculated using Matlab code written by the author. It has been found that $\|P_5\| \approx 1.1664 \times 10^6$. One observes that for $\rho \approx 10^{-7}$, the spectral contours C_ρ match the circles $|s - 5| = \rho \|P_5\|$.

$\vartheta_{\mathbb{C}}(A, D, E; \rho)$ in a neighborhood of infinity

In this case the situation is more complicated and we have not been able to state sharp results (like Proposition 2.2.15) of a reasonable degree of generality. Nevertheless, we give here an analysis which should help to understand the general situation. Our approach is essentially the same used in the former subsection: to find suitable series development of G. The difference is that we must work with the set of *zeros* of G: infinite and $Z(G)$.
We begin with the investigation of spectral value sets of large sizes. First, we apply (2.20) in order to obtain a series development for G at infinity. The result is that, for any $s \in \mathbb{C}$ such that $|s| > \gamma(A_{\min})$, the series

$$G(s) = E \left(\sum_{k=0}^{\infty} \frac{A_{\min}^k}{s^{k+1}} \right) D \tag{2.23}$$

converges absolutely.
Let us suppose, for simplicity, that $ED \neq 0$. Further, let us introduce the function

$$G_r : \rho(A_{\min}) \to \mathbb{C}^{q \times l}, \quad s \mapsto E A_{\min} (sI - A_{\min})^{-1} D. \tag{2.24}$$

Then it follows from (2.23) and (2.20) that

$$sG(s) = ED + E\left(\sum_{k=1}^{\infty} \frac{A_{\min}^k}{s^k}\right) D = ED + EA_{\min}\left(\sum_{j=0}^{\infty} \frac{A_{\min}^j}{s^{j+1}}\right) D = ED + G_r(s).$$

In other words, for a given $\rho > 0$ and $s \in C_\rho \cap \{z \in \mathbb{C}, |z| > \gamma(A_{\min})\}$ we have

$$\rho\|ED\| - \rho\|G_r(s)\| \leq |s| \leq \rho\|ED\| + \rho\|G_r(s)\|. \tag{2.25}$$

Note that $C_\rho \cap \{z \in \mathbb{C}, |z| > \gamma(A_{\min})\}$ is not empty for ρ large enough. Moreover, let us assume that the matrix E is invertible. Then, by (2.24),

$$\|G_r(s)\| \leq \delta\|G(s)\|, \text{ where } \delta := \|EA_{\min}E^{-1}\|,$$

and it follows from (2.25) and $\rho\|G(s)\| = 1$ that

$$\rho\|ED\| - \delta \leq |s| \leq \rho\|ED\| + \delta. \tag{2.26}$$

The inequalities (2.26) show that for ρ large enough a connected component of C_ρ lies in the region $\{s \in \mathbb{C},\ \rho\|ED\| - \delta \leq |s| \leq \rho\|ED\| + \delta\}$.
Let us consider now spectral value sets near the zeros of G. This is important because, by Remark 2.2.5, for any $\rho > 0$ some small neighborhood of $Z(G)$ remains in $\mathbb{C} \setminus \vartheta_{\mathbb{C}}(A, D, E; \rho)$.
We proceed by the same scheme: we write down Taylor series of G around the elements of $Z(G)$:

$$G(s) = G_{i0}(s - s_i) + G_{i1}(s - s_i)^2 + \ldots, \quad \forall\, s_i \in Z(G). \tag{2.27}$$

Clearly, in a small neighborhood $O(s_i) \subset \mathbb{C}$ of each $s_i \in Z(G)$, the function G can be approximated by the first nonzero summand of the series (2.27). We conclude that the corresponding spectral contours are approximated by certain circles. We omit the details.
Figure 2.4 depicts the situation for the matrices (2.11), (2.12) when the spectral norm is used. The continuous lines represent level curves of ψ and the dotted ones those of $\frac{ED}{|s|}$. Note that in the case of these matrices $\|EA_{\min}E^{-1}\| = \|A\| = 20455$. Furthermore, $Z(G) = \emptyset$.

2.2.5 Smoothness of spectral contours

Our aim in this section is to investigate the analyticity (or lack of it) of the spectral contours

$$C_\rho = \{s \in \mathbb{C} \ :\ \psi(s) = \rho^{-1}\}.$$

Our approach is to analyse the smoothness of ψ with respect to the variables x and y, $x + iy \in \Omega$, where Ω is some open subset of $\rho(A_{\min})$.
Let us begin with some initial considerations on general norms. For each $s \in \rho(A_{\min})$, the function ψ is given by

$$\psi(s) := \max_{\|u\|=1} \|G(s)u\|. \tag{2.28}$$

This is a *parametric optimisation problem with constraints* [29] and C_ρ is difficult to characterise in this general setting. For example, usually norms are non smooth functions of the entries of

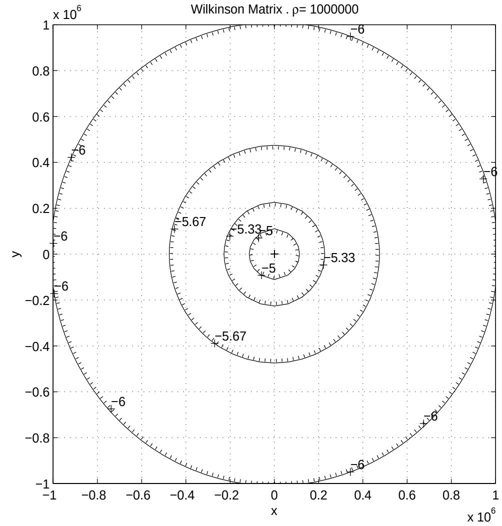

Figure 2.4: Behavior of $\vartheta_{\mathbb{C}}(A, D, E; \rho)$ for large perturbations.

the matrix and one would have to consider subdifferentials, dual sets, etc. [22]. Moreover, in order to obtain relatively "weak" results like local differentiability, one is obliged to assume *non degeneracy* of the *critical points* of Problem (2.28). A characterisation of these non degenerate critical points is not possible without further assumptions. The interested reader may consult [50] for a treatment of these topics.

In view of these difficulties and in order to be able to make substantial statements, we shall restrict our analysis to the case where the norms in the vector spaces are Euclidean, i.e.,

$$\|w\|_2 = \left(\sum_{i=1}^{m} |w_i|^2 \right)^{1/2}, \quad w = (w_1, \ldots, w_m)^T \in \mathbb{C}^m. \tag{2.29}$$

It turns out that for this norm one can give a fairly complete characterisation of the smoothness of ψ.

Preliminaries

We begin with some initial considerations. Let us consider the following action defined for every matrix $M \in \mathbb{C}^{q \times l}$:

$$M_{\mathbb{R}} = \begin{pmatrix} \operatorname{Re} M & -\operatorname{Im} M \\ \operatorname{Im} M & \operatorname{Re} M \end{pmatrix} \in \mathbb{R}^{2q \times 2l}. \tag{2.30}$$

This operation, usually called *realification* [61, p 55], has some remarkable properties.

Lemma 2.2.16 *Let $H \in \mathbb{C}^{p \times q}$ and $M \in \mathbb{C}^{q \times l}$. Then*

1. $(HM)_{\mathbb{R}} = H_{\mathbb{R}} M_{\mathbb{R}}$.

2. $(M_{\mathbb{R}})^T = (M^*)_{\mathbb{R}}$, *where* $(\,\cdot\,)^*$ *denotes conjugate transpose while* $(\,\cdot\,)^T$ *transpose.*

3. If M is unitary then $M_{\mathbb{R}}$ is orthogonal.

The proof of the first two assertions is straightforward and is omitted. The third assertion is given in [27, Exercise P1.4-5]. With the help of these properties one can prove

Proposition 2.2.17 *Suppose that $M \in \mathbb{C}^{q\times l}$ has singular value decomposition [73, Theorem I.4.1]*

$$M = U^*\Sigma V, \tag{2.31}$$

where Σ has size $q \times l$. Then $M_{\mathbb{R}}$ has (real) singular value decomposition

$$M_{\mathbb{R}} = (U_{\mathbb{R}})^T \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} V_{\mathbb{R}}. \tag{2.32}$$

Proof: Let us apply Point 1 of Lemma 2.2.16 to (2.31). The result is

$$M_{\mathbb{R}} = (U^*)_{\mathbb{R}}\Sigma_{\mathbb{R}}V_{\mathbb{R}} = (U^*)_{\mathbb{R}} \begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix} V_{\mathbb{R}}.$$

Further, Point 2 of Lemma 2.2.16 implies

$$(U^*)_{\mathbb{R}} = (U_{\mathbb{R}})^T.$$

Hence, Equation (2.32) holds.
It remains to show that both $U_{\mathbb{R}}$ and $V_{\mathbb{R}}$ are orthogonal matrices. But this is exactly the statement of Point 3 of Lemma 2.2.16. □

The next observation can be found in [73, Theorem I.4.2].

Theorem 2.2.18 *Suppose that $M \in \mathbb{R}^{q\times l}$, $q \geq l$, has singular values $\sigma_1, \ldots, \sigma_l$. Then the symmetric matrix*

$$H(M) := \begin{pmatrix} 0 & M \\ M^T & 0 \end{pmatrix} \in \mathbb{R}^{(q+l)\times(q+l)} \tag{2.33}$$

has spectrum

$$\sigma(H(M)) = \{\pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_l, \underbrace{0, \ldots, 0}_{q-l}\}.$$

The theorem below belongs to the well developed perturbation theory of real symmetric matrices [59].

Theorem 2.2.19 *Let Ω be an open connected subset of $\mathbb{R}^k$ and S^q be the set of real symmetric matrices of size $q \times q$. Suppose that*

$$R : \Omega \to S^q$$

is a differentiable mapping, that is, the components r_{ij} of R are of class C^∞ on Ω. Let $\lambda : \Omega \to \mathbb{R}$ be a continuous function such that, for every $\omega \in \Omega$, $\lambda(\omega)$ is an eigenvalue of $R(\omega)$ with constant *multiplicity. Then $\lambda(\cdot)$ is a differentiable function.*

With the help of the machinery developed above we can investigate the smoothness of ψ in the case when the vector norms under consideration are Euclidean. Indeed, it is well known [73, Theorem 2.10] that with this choice of the vector norms the operator norm on $\mathbb{C}^{q \times l}$ is given by

$$\|M\| = \sigma_1(M),$$

where $\sigma_1(M)$ denotes the *largest singular value* of $M \in \mathbb{C}^{q \times l}$. Thus, essentially, we must investigate the analyticity of the largest singular value of an holomorphic matrix function. In the sequel, we identify $\mathbb{C}$ with $\mathbb{R}^2$. Further, we shall write, with certain abuse of notation, $\psi(x, y)$ and/or $\psi(s)$, $s = x + iy$ depending on the context. The same convention holds for other functions.

Smoothness of C_ρ

We begin by noting that the transfer matrix

$$G(s) = E_{\min}(sI - A_{\min})^{-1} D_{\min}, \quad \forall\, s \in \rho(A_{\min}),$$

where $(A_{\min}, D_{\min}, E_{\min}) \in L_{k,l,q}(\mathbb{C})$, $(k \leq n)$, is a minimal realisation (2.5), can be written as [13, p. 91]

$$G(s) = \frac{1}{p(s)} Q(s), \quad s \in \rho(A_{\min}), \tag{2.34}$$

where $p(s)$ is the *characteristic polynomial* of $A_{\min}$ and $Q(\cdot)$ is certain matrix polynomial of degree $\leq k - 1$. It follows that

$$\|G(s)\| = \frac{1}{|p(s)|}\, \phi(s), \quad \forall\, s = x + iy \in \rho(A_{\min}),$$

with ϕ defined by

$$\phi : \mathbb{R}^2 \to \mathbb{R}_+, \quad (x, y) \mapsto \|Q(x, y)\|. \tag{2.35}$$

The function $(x, y) \mapsto |p(x + iy)|$ is different from zero and differentiable with respect to x and y in $\rho(A_{\min})$. Thus, we conclude that it is enough to characterise the smoothness of ϕ. This shall be our inmediate goal.
For simplicity in the notation let us set

$$M : \mathbb{R}^2 \to S^{2(l+q)}, \quad (x, y) \mapsto H(Q_{\mathbb{R}}(x, y)), \tag{2.36}$$

where $H(\cdot)$ is the matrix function defined in (2.33) and the subscript "$_{\mathbb{R}}$" denotes realification, see (2.30).

Proposition 2.2.20 *Let $\Omega \subset \mathbb{R}^2$ be an open connected set such that the largest eigenvalue of M (2.36) has* constant *multiplicity in Ω. Then ϕ is differentiable in Ω and, consequently, so is ψ in $\Omega \setminus P(G)$.*

Proof: Indeed, it follows from Proposition 2.2.17 that

$$\phi(x,y) = \|Q(x,y)\| = \sigma_1(Q(x,y)) = \sigma_1(Q_{\mathbb{R}}(x,y)).$$

Moreover, by Theorem 2.2.18,

$$\sigma_1(Q_{\mathbb{R}}(x,y)) = \lambda_{\max}(H(Q_{\mathbb{R}}(x,y))) = \lambda_{\max}(M(x,y)),$$

where λ_{max} denotes the largest eigenvalue. Now, we observe that $M(x,y)$ is *real* and *symmetric* for every par (x,y). Moreover, its entries are *polynomials* in the real variables x and y. Thus, our statement is a direct implication of Theorem 2.2.19. □

Remark 2.2.21 The fact that *simple* singular values of rational matrices are smooth functions of x and y is known, see [55] and [49]. Our proof, however, is different and includes the case of multiple identical singular values.

We discuss now the implications of Proposition 2.2.20 with regard to the analyticity of the spectral contours C_ρ. For this we introduce the following notations: Ω_S denotes the set of points where ψ is smooth, while Ω_{NS} denotes the complement of Ω_S in $\rho(A_{\min})$. Recall that the spectral contours are given by

$$C_\rho = \{(x,y) \in \mathbb{R}^2 \ : \ \psi(x,y) = \rho^{-1}\}.$$

First, we consider the case where $C_\rho \subset \Omega_S$. The simplest situation is when ρ^{-1} is a *regular value* [2, p. 4] of ψ, i.e.,

$$\operatorname{grad} \psi(x,y) \neq 0, \quad \forall\, (x,y) \in C_\rho.$$

In this case ψ is smooth in a neighborhood of C_ρ and each connected component of C_ρ is a smooth curve [2, p. 4]. Moreover, since C_ρ is bounded we have that these components are diffeomorphic to a circle [2, Lemma 2.4]. Figure 2.5 depicts this situation for the matrices (2.11), (2.12) in a neighborhood of the eigenvalues $\lambda_5 = 5$ and $\lambda_6 = 6$.

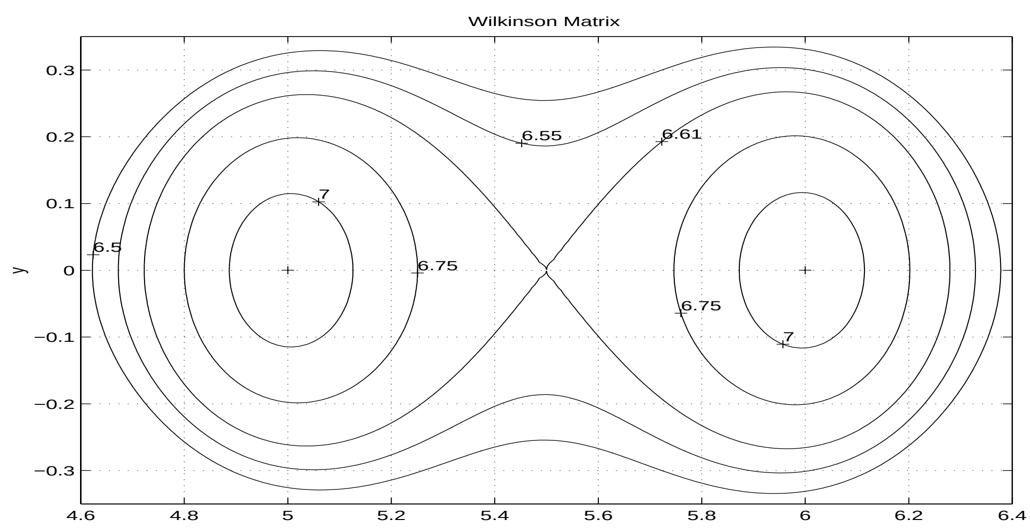

Figure 2.5: Smooth spectral contours

The behavior of C_ρ at points where the gradient of ψ is equal to zero also deserves attention. At these points the spectral contours may have *bifurcations* [2, Definition 4.4] and several smooth curves intersect each other [2, Theorem 4.6], [14]. This constellation can be observed in Figure 2.5 at neighborhood of the point $(5.5, 0)$.
Another structure is observed when the slopes of the level curves *coincide* at the point where the gradient of ψ is zero. Now, the spectral contours do not intersect itself and one obtains a "cusp". The spectral contour labeled with $''0''$ in Figure 2.6 is an example of this situation. The data used there is

$$A = \mathrm{diag}(-1, 1), \quad D = E = I_2. \tag{2.37}$$

Finally, we note that a point (x, y) where the gradient is equal to zero may also be a local minimum of ψ. In this case, there exists a neighborhood $O(x, y) \subset \mathbb{R}^2$ such that

$$C_\rho \cap O(x, y) = \{(x, y)\}.$$

Non smoothness of C_ρ

We finish our analysis with a characterisation of Ω_{NS}. This set is important because at the points where C_ρ intersects Ω_{NS}, ψ is non differentiable and the *normal* to the curves in C_ρ can be discontinuous or even not exist. For simplicity, we shall assume that

$$\text{the singular values of } Q(\cdot) \text{ are pairwise non identical functions} \tag{2.38}$$

for the rest of the section.
Proposition 2.2.20 suggests that in order to know where ψ is *non smooth*, it is necessary to characterise the points where the multiplicity of $\sigma_1(Q(\cdot)) = \lambda_{\max}(M(\cdot))$ might change. In other words, due to (2.38), we must find out where the singular values of Q are not simple.
We have that the singular values of $Q(x, y)$ at a point $(x, y) \in \mathbb{R}^2$ are the positive solutions of the equation

$$P(\zeta, x, y) := \det(\zeta^2 I - Q^*(x + iy)Q(x + iy)) = 0. \tag{2.39}$$

Clearly, $P(\zeta, x, y)$ is a polynomial function in three real variables. Now, if at some point $(x, y) \in \mathbb{R}^2$ the multiplicity of a singular value is bigger than one, we are in presence of a *multiple* root ζ_0 of Equation (2.39). It follows that ζ_0 is also root of the derivative of P with respect to ζ:

$$\begin{aligned} P(\zeta_0, x, y) &= 0, \\ \frac{\partial P}{\partial \zeta}(\zeta_0, x, y) &= 0. \end{aligned} \tag{2.40}$$

We see that (2.40) is a system of two *polynomial* equations in two real variables which, again because of (2.38), have no common factors. We conclude that the solutions of (2.40) define algebraic curves in $\mathbb{R}^2$. The set Ω_{NS} is a subset of these curves.
Let us consider now the question on the number of intersections of a spectral contour C_ρ with the set Ω_{NS}. We shall see that, at least in general, for a *fixed* $\rho > 0$ the corresponding spectral

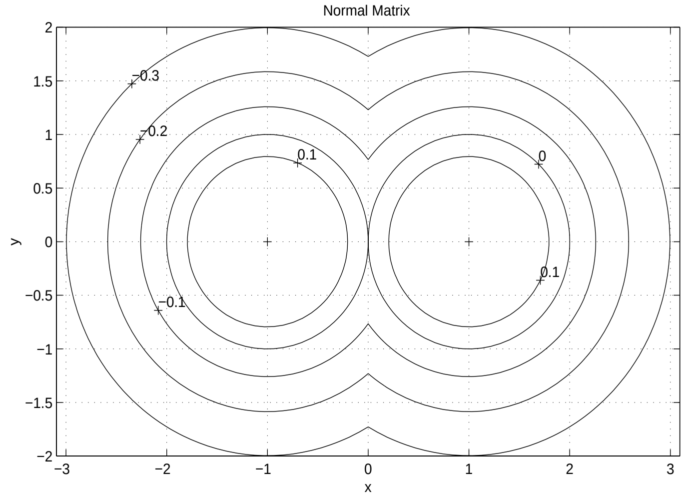

Figure 2.6: Non smooth spectral contours

contour C_ρ may intersect Ω_{NS} only at finite number of points. In fact (in the 2-norm) C_ρ is given by

$$C_\rho = \{s \in \mathbb{C} \ : \ \rho^{-1}|p(s)| = \sigma_1(Q(s))\}. \tag{2.41}$$

Thus, if C_ρ intersects Ω_{NS} the following equations must be satisfied simultaneously (see (2.40) and (2.41))

$$\begin{aligned} p(x-iy)p(x+iy) - \rho^2\zeta_0^2 &= 0, \\ P(\zeta_0, x, y) &= 0, \\ \frac{\partial P}{\partial \zeta}(\zeta_0, x, y) &= 0, \end{aligned} \tag{2.42}$$

This is a system of three polynomial equations in three real variables, which "generically" has only finitely many solutions. If the polynomials (2.42) have common factors, then $C_\rho \subset \Omega_{NS}$. Figure 2.6, which depicts spectral contours corresponding to the matrices (2.37), illustrates our result. Simple calculations show that for these matrices

$$\psi(x,y) = \begin{cases} \max\{(x^2 - 2x + 1 + y^2)^{-\frac{1}{2}}, (x^2 + 2x + 1 + y^2)^{-\frac{1}{2}}\} & \text{if} \quad (x,y) \in \rho(A) \\ +\infty & \text{if} \quad (x,y) \in \sigma(A) \end{cases}$$

and, consequently, that

$$\Omega_{NS} = \{(0,y) \in \mathbb{R}^2, \quad y \in \mathbb{R}\},$$

i.e., the imaginary axis. Moreover, the set Ω_{NS} is intersected at most twice by each of the spectral contours.

2.3 Summary

In this chapter the properties of spectral value sets $\sigma_{\mathbb{C}}(A, D, E; \rho)$ have been investigated. It has been shown that the sets are open and bounded for every $\rho > 0$ and monotonic and upper semicontinuous as set-valued functions of ρ. Furthermore, we studied their behavior for small and large sizes of the perturbations and the smoothness of the spectral contours C_ρ in the 2-norm.

Chapter 3

Calculation and Visualisation of Spectral Value Sets

Essentially, spectral value sets are a *graphical* tool for robustness analysis. Thus, there is need for reliable and efficient tools for their calculation and visualisation. Our aim in this chapter is to present a new numerical method for this purposes: the SH algorithm. We begin with a short "state of the art" review of this kind of calculations. Thereafter, we shall discuss our approach.

3.1 Existing methods

Most of the work in this area has been done for the unstructured complex case ($E = D = I$, $\mathbb{K} = \mathbb{C}$) and using the 2-norm in the spaces $\mathbb{C}^l$ and $\mathbb{C}^q$. In this section only this norm will be considered. Let us now discuss the different methods which are being used for the calculation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$.

3.1.1 Grid method

This is the method used for the generation of the graphics of Chapter 2. The Grid method is the most accurate and robust algorithm, but also the one with the highest computational cost. The quantity $\|G(s)\| = \sigma_{\max}(G(s))$ is calculated on a grid covering the set of interest in $\mathbb{C}$ and C_ρ is visualised by feeding the data into a contour plotter [42].
In the unstructured case (pseudospectra) [79]

$$\sigma_{\mathbb{C}}(A, D, E; \rho) = \sigma_{\mathbb{C}}(A, I, I; \rho)$$

one may simplify the calculations by observing that

$$\|G(s)\| = \|(sI - A)^{-1}\| = \sigma_{\min}^{-1}(sI - A), \tag{3.1}$$

where $\sigma_{\min}$ denotes the smallest singular value. This method has been the widely used by Trefethen and his coauthors in their investigations. It is reliable and robust but still computationally expensive. See [79] for references.
Much work has been done in order to reduce the number of "flops" required in these calculations. Marques and Toumazou [57], [56] have successfully applied the Lanczos method in the

calculation of $\sigma_{\min}(sI - A)$ for large matrices. Lui [54] used inverse iteration and a technique known as continuation which usually gives better initial vectors for the Lanczos algorithm. Toh and Trefethen [76] have used approximations of A obtained by Arnoldi iterations in order to reduce the order of the matrices which have to be dealt with. Recently Braconnier and Higham [8] combined Lanczos Method, Chebyshev acceleration and continuation in one algorithm and improved the performance in calculating $\sigma_{\min}(sI - A)$. Fortran versions of this algorithm have been developed by Braconnier [7]. The same author has also written codes in PVM [23] which can be used in parallel computations of the pseudospectra.

3.1.2 Random perturbations

The method described in this paragraph is straightforward, but inaccurate. Nevertheless, it has the advantage that it approximates $\sigma_{\mathbb{C}}(A, D, E; \rho)$ at a relatively low computational cost. It is applicable to both $\mathbb{K} = \mathbb{C}, \mathbb{R}$ and any operator norms.
The method starts by generating a sequence of random $\Delta_i \in \mathbb{K}^{l \times q}$, $i = 1, .., N$, such that $\|\Delta_i\| = \rho$, for every i. Then for large N an approximation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is given by the set

$$\sigma_{rand}^{N}(A, D, E; \rho; N) := \cup_{i=1}^{N} \sigma(A + D\Delta_i E).$$

It is clear that

$$\sigma_{rand}(A, D, E; \rho; N) \subset \sigma_{\mathbb{C}}(A, D, E; \rho),$$

but, usually, these approximations are far from being tight.
Figure 3.1 is a typical example of this kind of calculation. It represents $\sigma_{rand}(A, I_{32}, I_{32}; 1; 60)$ and the exact $\sigma_{\mathbb{C}}(A, I_{32}, I_{32}; 1)$, where

$$A := \texttt{pentoep}(32, 0, 1/2, 0, 0, 1) \in \mathbb{R}^{32 \times 32} \tag{3.2}$$

is taken from the `Test matrix toolbox for Matlab` [40]. The graphics were generated by means of the functions `ps` and `pscont` of the same toolbox. One sees that the approximation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is rather inexact. Nevertheless, since these graphics give an idea of the *mobility* of $\sigma(A)$ under "typical" perturbations of A, the analysis of these graphics can be useful for practical purposes. Graphics obtained with this method have appeared already in [45], [42].
Another variant of this method is the following. Again one generates random matrices Δ_i, $i = 1, \ldots, N$, but now the matrices have entries $\pm 1 \pm i$. The matrices should then be normalised in order to obtain the desired size of the perturbations. In this case one often obtains better approximations than when just "pure" random matrices are used. Finding a satisfactory explanation to this experimental fact would be interesting. Figure 3.2 represents the result of this method applied to the matrix (3.2). Note that in this example no improvements are observed. This approach has been used in [79].

3.1.3 Path following

The approach suggested by Bruehl in [9] deserves special mention. He determines C_ρ by finding the lines $\sigma_{\min}(sI - A) = \rho$. The goal is reached by using a predictor-corrector algorithm [2]. In this way there is no need for a large number of evaluations of $\sigma_{\min}$. The drawback of the method is that in presence of a topologically complicated structure of C_ρ, for example when $\sigma_{\mathbb{C}}(A, I, I; \rho)$ is not simply connected, the algorithm is likely to fail.

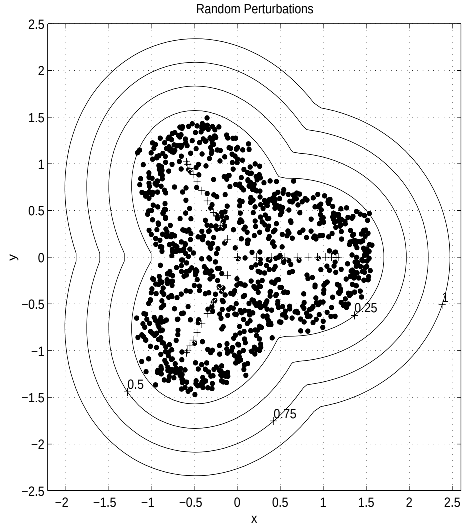

Figure 3.1: First variant of the random perturbations method: $\sigma_{rand}(A, I, I; 1; 60)$ and $\sigma_{\mathbb{C}}(A, I, I; 1)$

3.2 The SH algorithm

The calculation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ remains a difficult task and further investigations leading to more efficient algorithms are required. Previously, with the exception of the Path Following Method, all efforts of the numerical community in the pseudospectra area have been dedicated to diminishing the cost of calculation of $\sigma_{\min}(sI - A)$. We propose the use of the subharmonicity of the norm of $G(s)$, in particular Corollary 2.2.10, in order to reduce *the global number of evaluations* of $\|G(s)\|$.
Corollary 2.2.10 can be applied to the calculation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ in the following way. Let $\Omega \subset \rho(A)$ be an open bounded set such that $\bar{\Omega} \subset \rho(A)$ and let

$$M := \max_{\zeta \in \partial\Omega} \|G(\zeta)\|.$$

Then from Theorem 2.1.1 it follows that

$$\text{if } M < \rho^{-1} \implies \Omega \cap \sigma_{\mathbb{C}}(A, D, E; \rho) = \emptyset. \tag{3.3}$$

We propose to use this fact as the basis of a new algorithm. Given an initial set $R_0 \subset \mathbb{C}$, grids are generated by an iterative procedure and one decides, using (3.3), whether or not a given set remains of interest for the next iteration. The method will have the additional advantage that it does not rely on the properties of a specific norm, as do methods based in the calculation of the smallest singular value. In the sequel we call this method *SH algorithm* (signifying subharmonic).

3.2.1 Outline

Let $(A, D, E) \in L_{n,l,q}(\mathbb{C})$, $\rho > 0$ and $\epsilon > 0$, where ϵ is a small parameter which corresponds to the desired final accuracy of the calculation. As before, $G(s) = E(sI - A)^{-1}D$ denotes the associated transfer function.

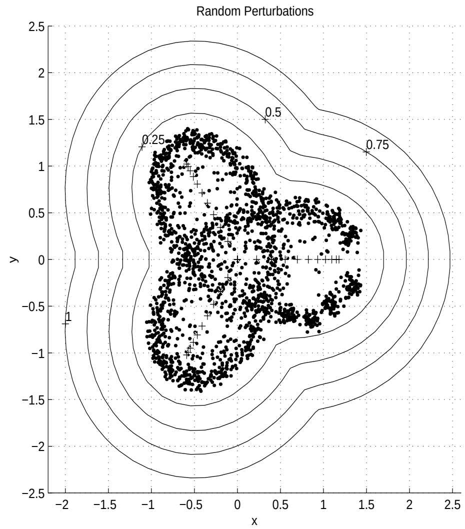

Figure 3.2: Second variant of the random perturbations method: $\sigma_{rand}(A, I, I; 1; 60)$ and $\sigma_{\mathbb{C}}(A, I, I; 1)$

In an iterative procedure, we shall search for a union of rectangles in the complex plane which approximates $\sigma_{\mathbb{C}}(A, D, E; \rho)$. Let $R_i = \cup_{j=1}^{n_i} R_{i,j}$ be the ith approximation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$, where each $R_{i,j}$ is a *rectangle* in $\mathbb{C}$. The general idea of the SH method is the following.

Point 3.2.1 *Given R_i, we find R_{i+1} by subdividing each $R_{i,j}$, $j = 1, \ldots, n_i$ into l congruent subrectangles $R_{i+1,j}$, $j = 1, \ldots, l$, and discarding some of the $R_{i+1,j}$ for the next iteration.*

The elimination is made according to a rule which applies Corollary 2.2.10.

Point 3.2.2 *If $R_{i+1,j}$ does not contain an eigenvalue of A then it is a set of subharmonicity of $\|G\|$. Hence, Corollary 2.2.10 can be applied: if*

$$Q_{i+1,j} := \max_{s \in \partial R_{i+1,j}} \|G(s)\| < \rho^{-1}$$

then

$$R_{i+1,j} \cap \sigma_{\mathbb{C}}(A, D, E; \rho) = \emptyset$$

and the rectangle $R_{i+1,j}$ can be discarded for the next iteration.

Since the use of a minimisation algorithm for the calculation of $Q_{i+1,j}$ can be very expensive, we only calculate a lower bound.

Point 3.2.3 *Let $V_{i+1,j}$ be the set of vertices of $R_{i+1,j}$. Then we approximate $Q_{i+1,j}$ by*

$$M_{i+1,j} := \max_{s \in V_{i+1,j}} \|G(s)\|.$$

We need some stopping criteria for the iterations. We have selected the size of the rectangles $R_{i,j}$. Suppose that $H(R_{i,j})$ denotes the area of $R_{i,j}$. Since by construction

$$H(R_{i+1,j}) = \frac{H(R_{i,k})}{l}, \qquad j = 1, \ldots, n_{i+1},\ k = 1, \ldots, n_i,$$

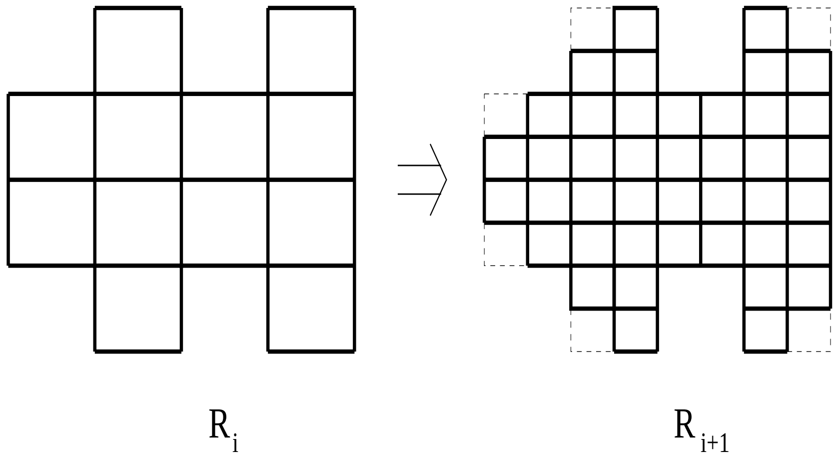

Figure 3.3: Predictor step. The rectangles of R_i are subdivided and some of the small rectangles are discarded (drawn with dashed lines).

we use

Point 3.2.4 *If $H(R_{i+1,j}) < \epsilon$, then stop the iterations.*

We call the actions above the *predictor step.* Figure 3.3 demonstrates how the predictor step could work. Still some work is needed. Due to the error introduced by substituting $Q_{i+1,j}$ with $M_{i+1,j}$, some “useful” $R_{i+1,j}$ may be discarded. Thus a *corrector step* is needed.

Point 3.2.5 *Suppose that after N iterations the predictor step is done. Then we inspect the boundary of R_N and generate the congruent neighboring rectangles $R'_{N,j}$. In R_N are incorporated those $R'_{N,j}$ which are* not discarded *by the rule described in Point 3.2.2. The boundary of the resulting R_N is again analysed until no new $R'_{N,j}$ is accepted.*

Figure 3.4 illustrates the functioning of the corrector step.
At the end of these operations, $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is contained in R_N up to grid precision. R_N and the data about the values of $\|G(s)\|$ are fed in a contour plotter in order to visualise $\sigma_{\mathbb{C}}(A, D, E; \rho)$. Below a *pseudocode* of the algorithm appears. A *free boundary of $R_{i,j}$* is a side of $R_{i,j}$ which is part of the boundary of R_i.

Pseudocode of the SH algorithm.

1. Initialisation.

1.1. Given $(A, D, E) \in L_{n,l,q}$, $\rho > 0$, $\epsilon > 0$, $R_0 = \cup_{j=1}^{n_0} R_{0,j}$.

1.2. Calculate $\sigma(A)$ (or part of it), set $i = 0$.

2. Predictor step.

2.1. Set $R_{i+1} = \emptyset$, $n_{i+1} = 0$.

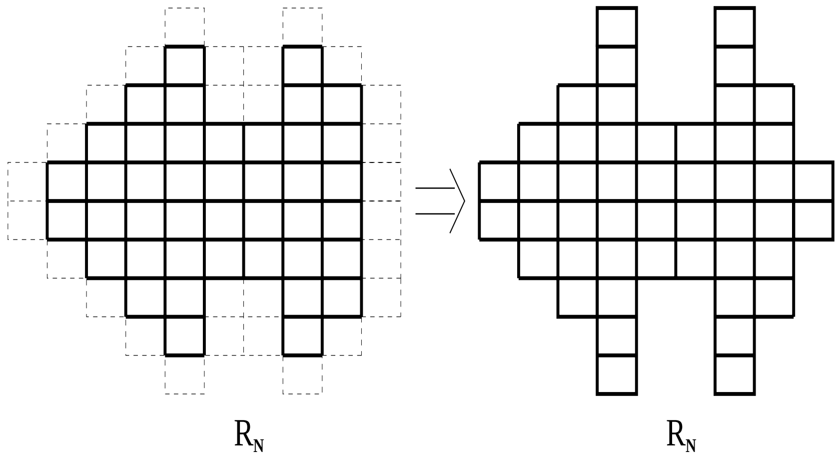

Figure 3.4: Corrector step. The neighboring rectangles have been drawn with dashed lines. Note that some of them are included in the "second" R_N.

2.2. for $j = 1, n_i$

2.2.1. Subdivide $R_{i,j}$ in l subrectangles $R_{i+1,j}$, $j = 1, .., l$.

2.2.2. for $j = 1, l$

2.2.2.1. Process $R_{i+1,j}$

2.3. if $(H(R_{i+1,j}) > \epsilon) \Rightarrow \begin{cases} i = i+1 \\ \textit{goto } 2.1 \end{cases}$

2.4 $N = i + 1$

3. Corrector step.

3.1. for $j = 1, n_N$

3.1.1. Using $N_{N,j}$, $E_{N,j}$, $W_{N,j}$, $S_{N,j}$ (see Function **Process**), detect the free boundaries of $R_{N,j}$

3.1.2. For every free boundary F, **Process** R_F, its neighboring rectangle.

4. Visualisation.

4.1 Feed the data into a contour plotter.

5. Stop.

The function **Process** used above "selects" the useful rectangles according to our selection rules. Its pseudocode looks as follows.

Pseudocode of the function Process.

1. Find $M_{i,j} = \max_{s \in V_{i,j}} \|G(s)\|$.

2. Calculate the boolean variable $U_{i,j}$.

$$U_{i,j} = \begin{cases} 1 & if & M_{i,j} & > & \rho^{-1}; \\ 1 & if & \sigma(A) \cap R_{i,j} & \neq & \emptyset\,; \\ 0 & otherwise\,; \end{cases}$$

3. if $(U_{i,j} = 1) \Rightarrow \begin{cases} R_i & = & R_i \cup R_{i,j} \\ n_i & = & n_i + 1 \end{cases}$

4. Find the free boundaries of $R_{i,j}$. The information is stored in the boolean variables $N_{i,j}$, $E_{i,j}$, $W_{i,j}$, $S_{i,j}$ respectively.

5. Return.

Adjacent rectangles have common vertices. Therefore, in order to avoid repeated calculations of $\|G(s)\|$, we store the vertices $V_{i,j}$ which have already been calculated. For the sake of clarity we have omitted these and other technicalities in the explanation.

3.2.2 Discussion

At this point it is worth making some comments. They should contribute to a better understanding of the algorithm.
First, we say some words about the corrector step. One can wonder why would the corrector step give different conclusions to the predictor step. The reason is the following. As mentioned before, the substitution of $Q_{i+1,j}$ with $M_{i+1,j}$ may result in *actually useful rectangles* $R_{i+1,j}$ be lost for the next iterations. The corrector step solves this defect of the predictor step by investigating the boundary of R_N. It incorporates in R_N useful rectangles $R'_{N,j}$ which *were not contained* in R_{N-1} during the predictor step. This reasoning leads us to a conclusion which deserves special mention.

Remark 3.2.6 By virtue of the corrector step, the SH method is able to enlarge R_0, so that when $\sigma_{\mathbb{C}}(A, D, E; \rho)$ extends outside of R_0, $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is still determined.

We shall illustrate this fact with one example in Section 3.3.
A good question to ask is whether or not a corrector step should be made for each R_i or just for R_N. The answer is that our experiments show that the inclusion of a corrector step at *each* iteration does not generally result in an acceleration of the algorithm. In our implementation we have decided to make a corrector step only for R_N.
On the other hand, another important remark is the following. All the rectangles $R_{i,j} \in R_i$ are subdivided. At a first glance, since we are trying to identify a "boundary" C_ρ, one might think that it is enough *just* to subdivide those rectangles of R_i which lie in the boundary of R_i. But this is not the case. In fact, *minima* of $\|G(s)\|$ are likely to exist and, therefore, new "internal" components of C_ρ can appear in later iterations. See [42] for one example. As a byproduct we obtain an important feature of the SH method.

Remark 3.2.7 The quantity $\|G(s)\|$ is calculated in *all* grid nodes which lie in $\sigma_{\mathbb{C}}(A, D, E; \rho)$. Since

$$\sigma_{\mathbb{C}}(A, D, E; \rho_1) \subset \sigma_{\mathbb{C}}(A, D, E; \rho), \quad \forall\, \rho_1 < \rho,$$

data generated for a given value of ρ can be used, without changes and just by means of the function `contour` of Matlab, for the visualisation of $\sigma_{\mathbb{C}}(A, D, E; \rho_1)$ for any $\rho_1 \in (0, \rho)$.

Finally, we say some words about the use of the spectrum of A in the SH algorithm. It might be thought that our method becomes useless whenever $\sigma(A)$ is difficult to calculate. But this is not the case. Indeed, let ϵ_m denote the machine precision used. Then, when one applies a backward stable numerical eigenvalue algorithm, QR for example, the answer consists of complex numbers which lie in the region of the complex plane in which $\|G(s)\| \geq O(\epsilon_m^{-1})$ [24]. It is easy to see, that these "pseudoeigenvalues" are accurate enough for our algorithm to succeed.

3.2.3 Performance

The calculation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is computationally very intensive. For example, in the unstructured case ($E = D = I$) the operation cost of the Grid method is about $O(mn^3)$, where m denotes the number of nodes in the grid used and n is the order of A. Lui [54] reports operations costs of $O(n^3 + mn^2)$ when continuation techniques are used. In the structured case the costs are even larger, since in this case $G(s)$ must be calculated at each point. In general, a straightforward grid method has a cost

$$P_{grid} = mS(n),$$

where $S(n) = O(n^k)$ is the estimate of the cost of a simple evaluation of $\|G(s)\|$. Here k takes some value in the interval $[2, 3]$, depending on the method used.
The operation cost P of the SH method can be decomposed into two components $P = P_1 + P_2$. P_1 is associated with the calculation of $\|G(s)\|$ in the required points, while P_2 is related to the logical operations implied in the generation and further manipulation of the sets R_i. Let us estimate P through the costs of the predictor step.
Let R_0 be the initial set of interest and suppose that $\sigma_{\mathbb{C}}(A, D, E; \rho) \subset R_0$. Introduce the notation

$$\tau := \frac{H(\sigma_{\mathbb{C}}(A, D, E; \rho))}{H(R_0)},$$

where $H(\sigma_{\mathbb{C}}(A, D, E; \rho))$ and $H(R_0)$ represent the areas of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ and R_0 respectively. Using this notation an estimate of P_1 is given by

$$P_1 = \tau m S(n).$$

In this formula m is the number of nodes contained in a uniform grid in R_0 of resolution ϵ, where ϵ is the parameter introduced in Section 3.2.1.
In contrast to P_1, P_2 *does not* depend on n. It is easy to see that ϵ and R_0 define the maximal possible number of iterations, which we denote by N. The next observation is that $P_2 = O(n_T)$, where n_T is the total number of rectangles n_T manipulated during the iterations. Further, using the notation of Section 3.2.1, we obtain that

$$n_T = \sum_{i=0}^{N} n_i \leq (N+1)n_N \approx (N+1)\tau m.$$

Here we have used the fact that the last iteration $i = N$ is the one which has the largest n_i. It is easy to see that n_N is closely related to the number of nodes of the regular grid, which are *inside* $\sigma_{\mathbb{C}}(A, D, E; \rho)$. This fact allows us to conclude that $n_T \approx \tau m$.

We thus have arrived at the following estimate for P

$$P = \tau m(S(n) + O(N+1)). \tag{3.4}$$

The relative performance P_{rel} of the SH method with respect to a standard grid method is then given by

$$P_{rel} = \frac{P}{P_{grid}} = \tau(1 + \frac{O(N+1)}{S(n)}).$$

The last equality shows that whenever τ is small and n is large the proposed algorithm should perform better than the standard ones. In Section 3.3.3 we shall show, through one example, that the constant implied in $O(N+1)$ is not very large.

The corrector step can involve, at least theoretically, relatively large number of operations. This can occur if the length of C_ρ is large with respect to the area of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ and is an unavoidable cost if the initial rectangle R_0 was a bad guess. Moreover, note that since R_N is corrected with the smallest rectangles, the total operation cost depends, to a certain extent, on the resolution of the grid chosen, i.e., on the parameter ϵ, or equivalently, on the maximal number of iterations N allowed. Experiments show that the use of large N ($N > 7$) is not advisable whenever no information about the actual size of $\sigma_{\mathbb{C}}(A, D, E; \rho)$ is available. However, for reasonable choices of R_0 and N, we have not found examples where the execution time of the corrector step is larger than 10% of the time used by the predictor step.

3.3 Examples

We investigate the speed and reliability of the SH algorithm, comparing it with the usual Grid method. Our implementation is a C++ code which uses the NAG Library [58] for most of the matrix operations. In our implementation $l = 4$ (see Point 3.2.1) and simple lists have been used to form and store the R_i. All computations were performed on a Sun SPARCserver-1000 and the unit roundoff is $2^{-53} \approx 1.11 \times 10^{-16}$. The pictures are generated with the aid of the function `contour` from Matlab 5.2. Equation (3.1) has been used whenever pseudospectra should be calculated.

3.3.1 Matrix Grcar

Let us begin with a well known example, the Grcar matrix [28]. Consider a matrix A of the form

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & & & & \\ -1 & 1 & 1 & 1 & 1 & & & \\ & -1 & 1 & 1 & 1 & \ddots & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & & -1 & 1 & 1 & 1 & 1 \\ & & & & -1 & 1 & 1 & 1 \\ & & & & & -1 & 1 & 1 \\ & & & & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{64\times 64}.$$

We calculated the unstructured spectral value set $\sigma_{\mathbb{C}}(A, I_{64}, I_{64}; 0.002)$, with R_0 the rectangle defined by the corners $(-0.5, 0), (2.5, 3)$. In the upper part $(\operatorname{Im} s \geq 0)$ of Figure 3.5 the results of the calculation of our algorithm (with $N = 6$) are depicted. The lower part is the output of a Matlab implementation of the usual Grid method. The points in the upper part are the vertices of the rectangles $R_{N,j}$, $j = 1, \ldots, n_N$.

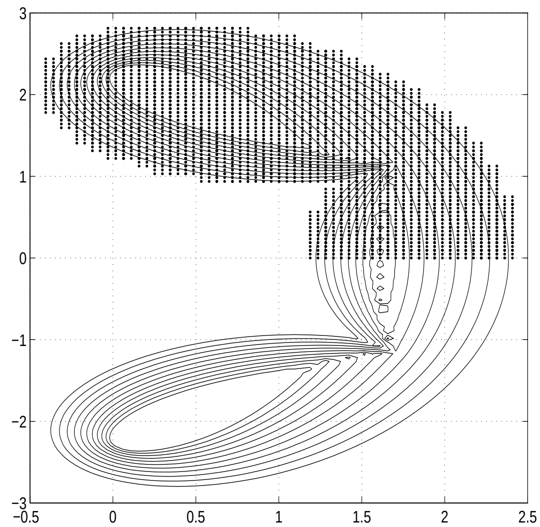

Figure 3.5: $\sigma_{\mathbb{C}}(A, I_{64}, I_{64}; 0.002)$ of Grcar matrix.

In Table 3.1 it can be seen that the SH algorithm made only 2725 evaluations of $\|G(s)\|$ versus 4225 by the Grid method. This represents a saving of about 35%. The times used by both algorithms are in the same relation. The behavior and performance of the SH algorithm observed in this example is typical for cases where the initial rectangle R_0 is larger than the actual $\sigma_{\mathbb{C}}(A, D, E; \rho)$ of interest.

3.3.2 Matrix Fish

The aim of this example is to show the ability of our algorithm to *enlarge* the initial set R_0 when R_0 does not contain the whole $\sigma_{\mathbb{C}}(A, D, E; \rho)$ of interest. With A given by the *pentadiagonal*

Grcar Matrix ($N = 6$)		
	Grid Method	SH Method
Num. Eval	4225	2725
Time	1280 sec	806 sec

Table 3.1: Performances for the Grcar Matrix ($n = 64$).

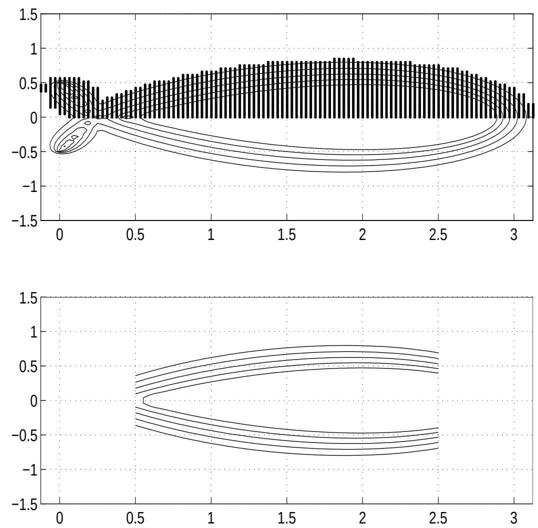

Figure 3.6: $\sigma_{\mathbb{C}}(A, I_{32}, I_{32}; 0.001)$ of the Fish matrix.

Toeplitz matrix [40]

$$A = \texttt{pentoep}(32, 0, 1/2, 1, 1, 1) \in \mathbb{R}^{32\times 32},$$

the set $\sigma_{\mathbb{C}}(A, I_{32}, I_{32}; 0.001)$ has been calculated by the standard Grid method and the SH algorithm. In both cases the same initial R_0 was chosen: a rectangle given by the corners $(0.5, 0)(2.5, 1.5)$.

Fish Matrix ($N = 6$)		
	Grid Method	SH Method
Num. Eval	4225	3187
Time	235 sec	164 sec

Table 3.2: Performances for the Fish Matrix ($n = 32$).

In the lower part of Figure 3.6 the result of applying the Grid method is depicted. The output of our algorithm (with $N = 6$) is represented in the upper part. As can be seen, the "tail" and the "head" of the "fish" have been perfectly recovered.

Table 3.2 depicts the performances of both methods. It can be seen that, although a larger region of the complex plane has been investigated, the SH algorithm has better performance statistics than the standard Grid algorithm.

3.3.3 Wilkinson matrix

With the example of this paragraph we want to illustrate

- the remarkable effect of using structured perturbations on the spectral value sets, and therefore, in the correct estimation of the robustness of the matrix A under study.
- that the performance of the SH algorithm is still good for small matrices and relatively large values of τ, (see Equation (3.4)).

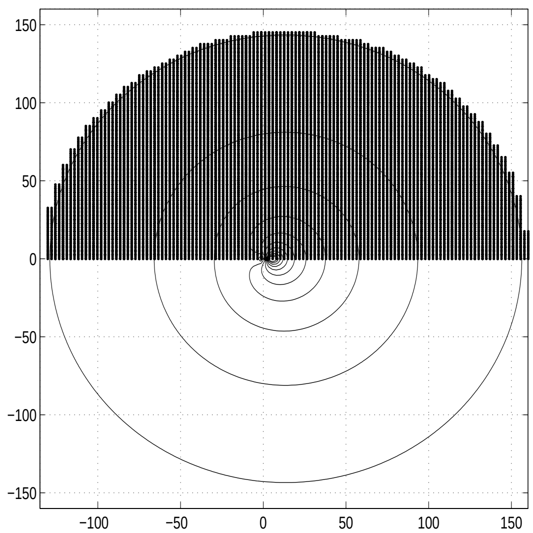

Figure 3.7: $\sigma_{\mathbb{C}}(A, I, I; 1)$ of the Wilkinson matrix ($N = 7$).

An interesting application of structured perturbations, and therefore of the SH algorithm, is the study of the root sets of polynomials [42]. This goal can be reached by investigating the spectral value sets of the corresponding companion matrices using structures given by suitable matrices D and E. With regard to this problem, the approach used in [75] is worth of being mentioned. Roughly speaking, the authors studied the pseudospectra ($D = E = I$) of a matrix obtained by certain preconditioning of the companion matrix. Nevertheless, we find that using structured spectral value sets is more natural and appealing.
We choose again (Chapter 2, Equation (2.11)) as study object the Wilkinson polynomial of degree 7, i.e., the monic polynomial with roots $1 = 1, \ldots, 7$. This problem has matrices of small dimensions and thus, is also well suited to illustrate the second target of this paragraph. Figure 3.7 depicts $\sigma_{\mathbb{C}}(A, I_7, I_7; 1)$ and Table 3.3 represents the performances of both the SH algorithm and the Grid method. Note that although the matrix is small and τ is relatively large, our algorithm still performs better than the standard one. This result shows that the constant implied in the $O(N + 1)$ of Formula (3.4) is not very large.

Wilkinson Matrix. Experiment 1 ($N = 7$)		
	Grid Method	SH Method
Num. Eval	16641	12403
Time	33 sec	29 sec

Table 3.3: Performances in Experiment 1.

Let us calculate now $\sigma_{\mathbb{C}}(A, D, E; 1)$ using the structures

$$D = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}^T, \quad E = I_7.$$

Note that now only perturbations of the *bottom row* are now allowed, which corresponds to perturbations of *only* the coefficients of the underlying polynomial. This is, perhaps, the ideal setup for studying the root sets of polynomials. Figure 3.8 depicts $\sigma_{\mathbb{C}}(A, D, E; 1)$. One can

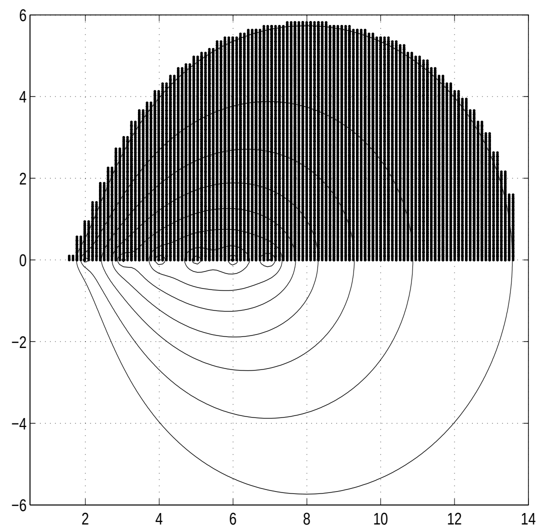

Figure 3.8: $\sigma_{\mathbb{C}}(A, D, E; 1)$ of the Wilkinson matrix ($N = 7$).

observe the striking difference in size and form of $\sigma_{\mathbb{C}}(A, I, I; 1)$ with respect to $\sigma_{\mathbb{C}}(A, D, E; 1)$. This example illustrates the importance of the use of perturbation structures. Again, the SH algorithm performs better that the Grid method.

Wilkinson Matrix. Experiment 2 ($N = 7$)		
	Grid Method	SH Method
Num. Eval	16641	11128
Time	20 sec	17 sec

Table 3.4: Performances in Experiment 2.

3.4 Summary

This chapter has dealt with the calculation of spectral value sets $\sigma_{\mathbb{C}}(A, D, E; \rho)$ in the matrix case. After a brief review of the existing method for the calculation of $\sigma_{\mathbb{C}}(A, D, E; \rho)$, A new algorithm, the SH algorithm, has been presented. It exploits the subharmonicity of $\|G\|$ proved in Proposition 2.2.8. The method draws its efficiency from the fact that, using the Maximum Principle for subharmonic functions, subsets of Ω are discarded which are not part of $\sigma_{\mathbb{C}}(A, D, E; \rho)$, saving calculations of $\|G(s)\|$. The idea has been implemented using rectangles as basic subsets. The SH algorithm is also able to enlarge the initial set Ω when it is smaller than $\sigma_{\mathbb{C}}(A, D, E; \rho)$. Finally, numerical examples have been discussed which illustrate the advantages of the new method. Last, but not least, one of the examples has also shown the importance of the use of structured perturbations.

Chapter 4

Spectral Value Sets in Infinite Dimensional Spaces

Our aim now is to extend the notion of spectral value sets to an infinite-dimensional setting. An immediate difficulty is that for unbounded operators on infinite-dimensional spaces the spectrum is a much more complicated object than the spectrum of a matrix. It is usual to restrict considerations to closed operators. We will also require that the model and perturbed operators are closed and so we will have to investigate under what conditions the property of closedness is preserved under perturbations. Note that in other contexts it may be more appropriate to place further restrictions on the operators. For example, if one were considering dynamical systems one might require that the model and perturbed operators are generators of strongly continuous semigroup. Nevertheless, we do not consider these questions here and our assumptions are just those needed for a meaningful treatment of spectral value sets.
In this chapter we shall define spectral value sets and a closedness radius. We shall see that a characterisation of the sets and a lower bound for the radius can be given in terms of the norm of a transfer function as it was done in [42] in the finite dimensional case.
We remark that our work on spectral value sets is closely related to the results of numerical analysts on pseudospectra of closed operators. Trefethen's paper [77] and more recently Harrabi's [32] and [33] are good examples of the attention which these ideas are receiving by the numerical community.

4.1 Preliminaries

In this short section we define the framework for the sequel: we shall introduce some notations, definitions and objects that are going to play a central role in the rest of the dissertation.

4.1.1 Some notations

Our aim here is to introduce some standard notations, definitions and results which are going to be used through this work.

1. X, Y, U, V, W, $\ldots$ denote Banach spaces unless otherwise stated.

2. $\|\cdot\|_X$ denotes the norm in the Banach space X.

3. X^* denotes the *dual* space of X, i.e., the set of all bounded linear forms in X.

4. A, D, E, R, S, T, ... denote linear operators acting in Banach spaces unless otherwise stated.

5. $\mathrm{Ker}(T)$ denotes the null space of a linear operator $T : X \to Y$, i.e., $\mathrm{Ker}(T) := \{x \in X \ : \ Tx = 0_Y\}$.

6. $\mathrm{Rg}(T)$ denotes the range of a linear operator $T : X \to Y$, i.e., $\mathrm{Rg}(T) := \{y \in Y :: \ y = Tx, \ x \in X\}$.

7. $L(X, Y)$ denotes the set of all linear operators on X to Y with domain $\mathcal{D}(T) \subset X$.

8. $\mathcal{L}(X, Y)$ denotes the set of all *bounded* operators in $L(X, Y)$. We write $\mathcal{L}(X)$ for $\mathcal{L}(X, X)$.

9. $\mathcal{K}(X, Y)$ denotes the set of all compact operators in $\mathcal{L}(X, Y)$, where we shall write $\mathcal{K}(X)$ for $\mathcal{K}(X, X)$.

10. $\mathcal{C}(X)$ denotes the sets of all closed linear operators in X.

11. $\rho(T) \subset \mathbb{C}$ denotes the resolvent set of $T \in \mathcal{C}(X)$.

12. $\sigma(T) := \mathbb{C} \setminus \rho(T)$ denotes the spectrum of $T \in \mathcal{C}(X)$.

13. $R(s, T)$, $s \in \rho(T)$, denotes the resolvent operator of $T \in \mathcal{C}(X)$. If there is no possibility of confusion, we write $R(s)$ for $R(s, T)$.

4.1.2 Main definitions and assumptions

We assume that $\underline{X}$, X, U, Y are complex Banach spaces, A is a closed linear operator in X:

$$A \in \mathcal{C}(X)$$

with domain

$$\mathcal{D}(A) \subset X, \quad \overline{\mathcal{D}(A)} = X,$$

while $D : U \to X$ and $E : \underline{X} \to Y$ are bounded operators:

$$D \in \mathcal{L}(U, X), \quad E \in \mathcal{L}(\underline{X}, Y).$$

We provide $\mathcal{D}(A)$ with the corresponding *graph*-norm

$$\|x\|_{\mathcal{D}(A)} = (\|x\|_X^2 + \|Ax\|_X^2)^{\frac{1}{2}}. \tag{4.1}$$

Since A is closed in X, its domain $\mathcal{D}(A)$ endowed with this norm is complete and hence a Banach space. Our fundamental assumption is

$$\mathcal{D}(A) \subset \underline{X} \subset X \text{ with continuous dense injections.} \tag{4.2}$$

4.1.3 The transfer function

For any $s \in \mathbb{C}$ in the resolvent set $\rho(A)$, the linear operator $sI - A : \mathcal{D}(A) \to X$ is bounded and surjective as an operator from the Banach space $\mathcal{D}(A)$ into the Banach space X and hence, the inverse $(sI - A)^{-1} : X \to \mathcal{D}(A)$ is bounded by the open mapping theorem:

$$R(s) = (sI - A)^{-1} \in \mathcal{L}(X, \mathcal{D}(A)),$$

($\mathcal{D}(A)$ being endowed with the graph norm). So the transfer function

$$G(s) := E(sI - A)^{-1}D \in \mathcal{L}(U, Y), \quad s \in \rho(A). \tag{4.3}$$

is well defined. Actually we have

Lemma 4.1.1 *$G : \rho(A) \to \mathcal{L}(U, Y)$ is analytic on $\rho(A)$.*

Proof: We must show that for every $s_0 \in \rho(A)$, there exists an open neighborhood $O(s_0) \subset \rho(A)$, such that, for every $s \in O(s_0)$, $G(s)$ is given by an absolutely convergent series of the form

$$G(s) = \sum_{k=0}^{\infty} (s - s_0)^k G_k(s_0), \quad G_k(s_0) \in \mathcal{L}(U, Y), \quad k = 1, 2, \ldots.$$

In fact, for any $s_0 \in \rho(A)$ and any $s \in \mathbb{C}$ satisfying $|s - s_0| \, \|R(s_0)\|_{\mathcal{L}(X)} < 1$ the following series converges absolutely [15, Equation A.4.5]

$$R(s) = R(s_0) \left[I + (s_0 - s)R(s_0)\right]^{-1} = R(s_0) \left[I + \sum_{k=1}^{\infty} (s_0 - s)^k R(s_0)^k \right].$$

Choosing $s \in \mathbb{C}$ such that even

$$|s - s_0| \, \|R(s_0)\|_{\mathcal{L}(X, \mathcal{D}(A))} < 1$$

the series is absolutely convergent in $\mathcal{L}(X, \mathcal{D}(A))$. Since the restriction of $E : \underline{X} \to Y$ to $\mathcal{D}(A)$ is continuous we see that

$$E(sI - A)^{-1}D = ER(s_0)D + \sum_{k=1}^{\infty} (s_0 - s)^k ER(s_0)^{k+1} D$$

is absolutely convergent in $\mathcal{L}(U, Y)$. Thus $G : \rho(A) \to \mathcal{L}(U, Y)$ is analytic. □

4.1.4 Auxiliary lemma

We shall need the following lemma. Although fairly well known we prove it for completeness.

Lemma 4.1.2 *Suppose Y, Z are Banach spaces and*

$$M \in \mathcal{L}(Y, Z), \quad N \in \mathcal{L}(Z, Y),$$

then $(I_Z + MN)$ has bounded inverse $(I_Z + MN)^{-1}$ if and only if $(I_Y + NM)$ has. In this case

$$(I_Z + MN)^{-1} M = M(I_Y + NM)^{-1}. \tag{4.4}$$

Proof: Assume $(I_Y + NM)^{-1} \in \mathcal{L}(Y)$. We have

$$M(I_Y + NM) = (I_Z + MN)M.$$

Hence,

$$M = (I_Z + MN)M(I_Y + NM)^{-1}. \tag{4.5}$$

Let us introduce the auxiliary operator

$$L := I_Z - M(I_Y + NM)^{-1}N \in \mathcal{L}(Z).$$

Then, using (4.5), we see that

$$(I_Z + MN)L = (I_Z + MN) - (I_Z + MN)M(I_Y + NM)^{-1}N = (I_Z + MN) - MN = I_Z.$$

So $(I_Z + MN) : Z \to Z$ is surjective. Now suppose $(I_Z + MN)z = 0$, then

$$(I_Y + NM)Nz = N(I_Z + MN)z = 0.$$

But since $(I_Y + NM) : Y \to Y$ is invertible this implies $Nz = 0$ and this together with $(I_Z + MN)z = 0$ implies $z = 0$. It follows $(I_Z + MN) : Z \to Z$ is injective. The invertibility of $(I_Z + MN)Z \to Z$ (as a bounded linear operator on Z) now follows from the open mapping theorem. Finally (4.4) follows now from (4.5). □

4.2 Characterisation of spectral value sets

Our aim is to study the variations of the spectrum $\sigma(A)$ under structured perturbations

$$A \rightsquigarrow A_\Delta = A + D\Delta E, \quad \Delta \in \mathcal{L}(Y, U).$$

with $\mathcal{D}(A_\Delta) := \mathcal{D}(A)$ by definition. The operators D, E are fixed and describe both the structure and unboundedness of the perturbations, whilst Δ is arbitrary.

Definition 4.2.1 Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$. Suppose that (4.2) holds and that a number $\rho > 0$ is given. Then the associated *spectral value set* of level ρ, denoted $\sigma(A, D, E; \rho)$, is defined to be

$$\sigma(A, D, E; \rho) \;=\; \bigcup_{\Delta \in \mathcal{L}(Y,U),\, \|\Delta\| < \rho,\, A_\Delta \in \mathcal{C}(X)} \sigma(A_\Delta).$$

The *point spectral value set* of level ρ, denoted $\sigma_P\,(A, D, E; \rho)$, is given by

$$\sigma_P(A, D, E; \rho) \;=\; \bigcup_{\Delta \in \mathcal{L}(Y,U),\, \|\Delta\| < \rho,\, A_\Delta \in \mathcal{C}(X)} \sigma_P(A_\Delta).$$

where $\sigma_P(A_\Delta)$ denotes the point spectrum of A_Δ.

4.2.1 Closedness radius

We pursue a characterisation of spectral value sets, but note that in the definition there is a requirement that A_Δ be closed. We, therefore, analyse for which $\rho > 0$ the condition $\|\Delta\|_{\mathcal{L}(Y,U)} < \rho$ implies that the perturbed operator A_Δ is closed. The supremal value of ρ for which this implication is valid is called *closedness radius*.

Definition 4.2.2 The closedness radius $r(A,\, D,\, E)$ is given by

$$\begin{aligned} r(A,\, D,\, E) \;&=\; \inf\{\|\Delta\|_{\mathcal{L}(Y,U)},\, \Delta \in \mathcal{L}(Y,U) \\ &\quad \text{such that } A_\Delta \text{ with domain } \mathcal{D}(A) \text{ is not closed}\}. \end{aligned}$$

We set $r(A,\, D,\, E) = \infty$ if A_Δ is closed for all $\Delta \in \mathcal{L}(Y,U)$.
The "stability of closedness" problem has been treated by Kato in [53, Chapter IV] using the concept of *relative-boundedness*. Let T and S be linear operators with the same domain space X and such that $\mathcal{D}(T) \subset \mathcal{D}(S)$. Suppose that the following inequality holds

$$\|Sx\|_Y \le a\|x\|_X + b\|Tx\|_Y, \quad \forall\, x \in \mathcal{D}(T), \tag{4.6}$$

where a and b are real nonnegative constants. If this is the case one says that S is *T-bounded.* The infimum b_0 of the numbers b (with a suitable a depending upon b) for which (4.6) holds is called *T-bound* of S. Note that the T-bound of S is zero if $S \in \mathcal{L}(X,Y)$. The following theorem holds [53, Theorem IV.1.1].

Theorem 4.2.3 *Let S, $T \in L(X,Y)$, where X and Y are Banach spaces, and let S be T-bounded with T-bound smaller than one. Then $R := T + S$ is closed if and only if T is closed.*

An immediate corollary of this theorem is

Corollary 4.2.4 *Suppose $\underline{X} = X$. Then $r(A,\, D,\, E) = \infty$.*

The reason for this is that, under these conditions, $D\Delta E \in \mathcal{L}(X)$ and thus, its A-bound is zero. In this section we obtain a rather different result. It takes into account the presence of structure in the perturbations.

Proposition 4.2.5 *Suppose $s \in \rho(A)$. If*

$$\|\Delta\|_{\mathcal{L}(Y,U)} < \|G(s)\|^{-1}_{\mathcal{L}(U,Y)}, \tag{4.7}$$

then A_Δ is closed and $s \in \rho(A_\Delta)$.

Proof: Suppose that (4.7) holds. Then $\|G(s)\Delta\|_{\mathcal{L}(Y)} < 1$ and so $I_Y - G(s)\Delta$ has bounded inverse. But

$$I_Y - G(s)\Delta = I_Y - E(sI_X - A)^{-1}D\Delta$$

and by (4.2), we have

$$E|_{\mathcal{D}(A)} \in \mathcal{L}(\mathcal{D}(A), Y); \quad (sI_X - A)^{-1}D\Delta \in \mathcal{L}(Y, \mathcal{D}(A)).$$

Hence applying Lemma 4.1.2 with $Z = \mathcal{D}(A)$, $M = (sI_X - A)^{-1}D\Delta$ and $N = E|_{\mathcal{D}(A)}$ we have that

$$I_{\mathcal{D}(A)} - (sI_X - A)^{-1}D\Delta(E|_{\mathcal{D}(A)})$$

has bounded inverse as well. But since $s \in \rho(A)$, this implies that

$$sI_X|_{\mathcal{D}(A)} - A - D\Delta(E|_{\mathcal{D}(A)}) = sI_X|_{\mathcal{D}(A)} - A_\Delta|_{\mathcal{D}(A)} : \mathcal{D}(A) \to X$$

is boundedly invertible and so

$$(sI_X - A_\Delta)^{-1} \in \mathcal{L}(X, \mathcal{D}(A))$$

and thus

$$s \in \rho(A_\Delta).$$

Moreover, an operator and its inverse are closed simultaneously [53, III.5.2]. Thus, we have that $sI_X - A_\Delta \in \mathcal{C}(X)$ and as a consequence A_Δ must also be closed. This completes the proof. □

As an immediate consequence we have the following estimate.

Theorem 4.2.6 *Suppose that $\rho(A) \neq \emptyset$, then*

$$r(A,\, D,\, E) \geq \sup_{s\in\rho(A)} \|G(s)\|^{-1}_{\mathcal{L}(U,Y)} = \left[\inf_{s\in\rho(A)} \|G(s)\|_{\mathcal{L}(U,Y)}\right]^{-1} =: \bar{\rho}. \tag{4.8}$$

Remark 4.2.7 The number $\bar{\rho}$ above will be found in most of the statements of this chapter. In most applications (in particular, in the finite dimensional case) the infimum of $\|G(s)\|_{\mathcal{L}(U,Y)}$ on $\rho(A)$ will be zero. Thus, $\bar{\rho} = \infty$ and we have $r(A,\, D,\, E) = \infty$. Note, however, that $\bar{\rho}$ *may* be finite.

We investigate now the relationship between our results on "stability of closedness" and Theorem 4.2.3. We would like to deduce this theorem from our theory. However, we shall see that only a weaker result can be obtained.
Only the case where A is unbounded will be considered. The reason is that if $A \in \mathcal{L}(X)$, then $\mathcal{D}(A) = X$ and consequently $\underline{X} = X$ as well. Thus, by Corollary 4.2.4, we are always in the scope of Theorem 4.2.3. So, let $A \in \mathcal{C}(X)$ be such that

$$\sup_{\substack{x\in\mathcal{D}(A)\\ \|x\|_X=1}} \|Ax\|_X = \infty. \tag{4.9}$$

We define $U = X$, $\underline{X} = Y = \mathcal{D}(A)$, $D = I_X$, $E = I_{\mathcal{D}(A)}$. By definition, $\Delta \in \mathcal{L}(Y,U)$ iff

$$\|\Delta\|_{\mathcal{L}(Y,U)} = \sup_{x\in\mathcal{D}(A)} \frac{\|\Delta x\|_X}{[\|x\|_X^2 + \|Ax\|_X^2]^{\frac{1}{2}}} < \infty.$$

Thus, saying $\Delta \in \mathcal{L}(Y,U)$ is equivalent to state that Δ is A-bounded with A-bound equal to $\|\Delta\|_{\mathcal{L}(Y,U)}$. We apply Proposition 4.2.5 now. In this case, the transfer function $G(s) \in \mathcal{L}(U,Y)$ is just $(sI - A)^{-1} : X \to \mathcal{D}(A)$, $s \in \rho(A)$ and one obtains the following bound for its norm:

$$\begin{aligned}
\|(sI-A)^{-1}\|_{\mathcal{L}(X,\mathcal{D}(A))} &= \sup_{\substack{y\in X\\ y\neq 0}} \frac{\|(sI-A)^{-1}y\|_{\mathcal{D}(A)}}{\|y\|_X} = \sup_{\substack{x\in\mathcal{D}(A)\\ x\neq 0}} \frac{\|x\|_{\mathcal{D}(A)}}{\|(sI-A)x\|_X} \\
&= \sup_{\substack{x\in\mathcal{D}(A)\\ x\neq 0}} \frac{(\|x\|_X^2+\|Ax\|_X^2)^{\frac{1}{2}}}{\|(sI-A)x\|_X} \geq \sup_{\substack{x\in\mathcal{D}(A)\\ x\neq 0}} \frac{(\|x\|_X^2+\|Ax\|_X^2)^{\frac{1}{2}}}{|s|\|x\|_X+\|Ax\|_X} \qquad (4.10)\\
&= \sup_{\substack{x\in\mathcal{D}(A)\\ \|x\|_X=1}} \frac{(1+\|Ax\|_X^2)^{\frac{1}{2}}}{|s|+\|Ax\|_X} \geq \inf_{z\in\rho(A)} \sup_{\chi\geq 0} \frac{(1+\chi^2)^{\frac{1}{2}}}{|z|+\chi} =: \delta.
\end{aligned}$$

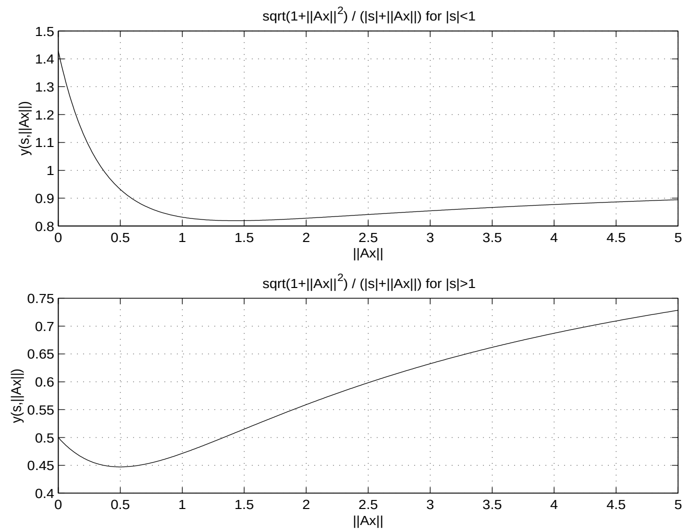

Figure 4.1: Graphics of $y(z,\chi) = \frac{(1+\chi^2)^{\frac{1}{2}}}{|z|+\chi}$ for $|z|<1$ and $|z|>1$

Simple considerations (see Figure 4.1) using (4.9) convince us that

$$\sup_{\chi\geq 0} \frac{(1+\chi^2)^{\frac{1}{2}}}{|z|+\chi} \geq 1.$$

Thus, $\delta \geq 1$ and it follows, by (4.10), that

$$\inf_{s\in\rho(A)} \|(sI-A)^{-1}\|_{\mathcal{L}(X,\mathcal{D}(A))} \geq 1.$$

Hence,

$$\bar{\rho} := \left[\inf_{s\in\rho(A)} \|(sI-A)^{-1}\|_{\mathcal{L}(U,Y)}\right]^{-1} \leq 1. \qquad (4.11)$$

Now, by Theorem 4.2.6, it follows that $A_\Delta = A+\Delta$ is closed whenever

$$\|\Delta\|_{\mathcal{L}(Y,U)} < \bar{\rho}.$$

Remark 4.2.8 It is readily seen that there is a gap between this result and Theorem 4.2.3: our theory does not give information about the closedness of A_Δ if the A-bound of Δ belongs to the interval $[\bar{\rho}, 1)$.

Remark 4.2.9 The bound (4.8) may be conservative even in the unstructured case $U = Y = \underline{X} = X$, $E = D = I_X$. Indeed, since $A + \Delta$ is always closed for $A \in \mathcal{C}(X)$ and $\Delta \in \mathcal{L}(X)$, the closedness radius $r(A, I_X, I_X)$ is infinite. On the other hand, one can construct examples where

$$\inf_{s \in \rho(A)} \|R(s, A)\|_{\mathcal{L}(X)} > 0. \tag{4.12}$$

In that case, $\bar{\rho}$ would be finite. The reader should note, however, that (4.12) might hold only in "exotic" situations. For example, an important operator class for which (4.12) is forbidden is the class of $A \in \mathcal{C}(X)$ which are infinitesimal generators of C_0-semigroups. This follows from the Hille-Yosida Theorem [15, Theorem 2.2.12].

4.2.2 Spectral value sets

Let us return to our main topic: spectral value sets. The next proposition is another step towards its characterisation. It contains a sort of converse statement to Proposition 4.2.5.

Proposition 4.2.10 *Suppose $s \in \rho(A)$ and $\|G(s)\|^{-1}_{\mathcal{L}(U,Y)} < \bar{\rho}$. Then there exists, for every $\varepsilon > 0$, a disturbance $\Delta \in \mathcal{L}(Y, U)$ with*

$$\|G(s)\|^{-1}_{\mathcal{L}(U,Y)} \leq \|\Delta\|_{\mathcal{L}(Y,U)} < \|G(s)\|^{-1}_{\mathcal{L}(U,Y)} + \varepsilon$$

such that A_Δ is closed and $s \in \sigma_P(A_\Delta)$.

Proof: Let $u \in U$ with $\|u\|_U = 1$ be such that

$$\|G(s)u\|^{-1}_Y < \min\{\|G(s)\|^{-1}_{\mathcal{L}(U,Y)} + \varepsilon, \bar{\rho}\}.$$

Then there exist (by the Hahn-Banach Theorem) a linear form $y^* \in Y^*$ of dual norm $\|y^*\|_{Y^*} = 1$ such that

$$y^* G(s)u = \|G(s)u\|_Y.$$

Define $\Delta \in \mathcal{L}(Y, U)$ by

$$\Delta = \|G(s)u\|^{-1}_Y u y^*.$$

Then, it is easy to check that $\Delta\, G(s)u = u$ and $\|\Delta\|_{\mathcal{L}(Y,U)} = \|G(s)u\|^{-1}_Y$, while by Corollary 4.2.6 we can ensure that A_Δ is closed.
Let $x = (sI - A)^{-1}Du$, then $x \in \mathcal{D}(A)$ and $x \neq 0$ since otherwise $u = \Delta E(sI - A)^{-1}Du = \Delta Ex = 0$. But $D\Delta Ex = Du = (sI - A)x$ and so $A_\Delta x = sx$. The proof is complete. □

As a consequence of the previous propositions we obtain a characterisation of the spectral value sets in terms of the superlevel sets of $\|G(\cdot)\|_{\mathcal{L}(U,Y)}$.

Theorem 4.2.11 *Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$. Suppose also that (4.2) holds. If*

$$0 < \rho < \bar{\rho},$$

then

$$\sigma(A, D, E; \rho) = \sigma(A) \cup \{s \in \rho(A) : \|G(s)\|_{\mathcal{L}(U,Y)} > \rho^{-1}\} \tag{4.13}$$

Moreover,

$$\sigma(A, D, E; \rho) \cap \rho(A) = \{s \in \rho(A); \|G(s)\|_{\mathcal{L}(U,Y)} > \rho^{-1}\} = \sigma_P(A, D, E; \rho) \cap \rho(A). \tag{4.14}$$

Proof: "$\subset$ in (4.13)": If $s \in \sigma(A, D, E; \rho) \cap \rho(A)$ then $s \in \sigma(A_\Delta)$ for some $\Delta \in \mathcal{L}(Y, U)$ with $0 < \|\Delta\|_{\mathcal{L}(Y,U)} < \rho$. It follows from Proposition 4.2.5 that

$$\|\Delta\|_{\mathcal{L}(Y,U)} \geq \|G(s)\|^{-1}_{\mathcal{L}(U,Y)}$$

and hence

$$\|G(s)\|_{\mathcal{L}(U,Y)} \geq \|\Delta\|^{-1}_{\mathcal{L}(Y,U)} > \rho^{-1}.$$

"$\supset$ in (4.13)": Now assume $s \in \rho(A)$ and $\|G(s)\|_{\mathcal{L}(U,Y)} > \rho^{-1}$. By Proposition 4.2.10 there exists $\Delta \in \mathcal{L}(Y, U)$ such that A_Δ is closed, $s \in \sigma_P(A_\Delta)$ and $\|\Delta\|_{\mathcal{L}(Y,U)} < \rho$, hence $s \in \sigma_P(A, D, E; \rho)$. This proves the remaining part and hence equality in (4.13).
The first equation in (4.14) follows from (4.13). Using this equality the previous argument shows the inclusion $\subset$ between the second and third sets in (4.14). The converse inclusion $\supset$ is trivial. □

As a consequence of the theorem we obtain that $\sigma(A, D, E; \rho)$ is the disjoint union of the closed set $\sigma(A)$ and the open set $\sigma(A, D, E; \rho) \cap \rho(A)$. Moreover, we obtain the following corollary.

Corollary 4.2.12 *Suppose $0 < \rho < \bar{\rho}$. Then*

$$\overline{\sigma(A, D, E; \rho)} \subset \sigma(A) \cup \{s \in \rho(A) : \|G(s)\|_{\mathcal{L}(U,Y)} \geq \rho^{-1}\} = \bigcap_{\rho < \tilde{\rho} < \bar{\rho}} \sigma(A, D, E; \tilde{\rho}). \tag{4.15}$$

Furthermore, the union of all spectral value sets for the values $0 < \rho < \bar{\rho}$ consists of all points of the complex plane where $\|G(s)\|_{\mathcal{L}(U,Y)}$ does not take its minimum:

$$\bigcup_{0 < \rho < \bar{\rho}} \sigma(A, D, E; \rho) = \mathbb{C} \setminus \left\{ \mu \in \rho(A); \|G(\mu)\|_{\mathcal{L}(U,Y)} = \inf_{s \in \rho(A)} \|G(s)\|_{\mathcal{L}(U,Y)} \right\} \tag{4.16}$$

In particular, if $\|G(\cdot)\|_{\mathcal{L}(U,Y)}$ does not have a global minimum on $\rho(A)$ then the union of all the above spectral value sets is the whole complex plane.

Proof: By Theorem 4.2.11 and by the continuity of $G(\cdot)$ on $\rho(A)$ we have

$$\begin{aligned}\overline{\sigma(A,D,E;\rho)} &= \sigma(A) \cup \overline{\{s \in \rho(A); \|G(s)\|_{\mathcal{L}(U,Y)} > \rho^{-1}\}} \\ &\subset \sigma(A) \cup \{s \in \rho(A); \|G(s)\|_{\mathcal{L}(U,Y)} \geq \rho^{-1}\}.\end{aligned}$$

This proves the first inclusion in (4.15). Further, if $s_0 \in \{s \in \rho(A) : \|G(s)\| \geq \rho^{-1}\}$ then, for any $\tilde{\rho} \in (\rho, \bar{\rho})$,

$$s_0 \in \{s \in \rho(A); \|G(s)\|_{\mathcal{L}(U,Y)} > \tilde{\rho}^{-1}\} \subset \sigma(A,D,E;\tilde{\rho}).$$

This proves the inclusion "$\subset$" for the second equation in (4.15). Finally, let

$$s_0 \in \bigcap_{\rho<\tilde{\rho}<\bar{\rho}} \sigma(A,D,E;\tilde{\rho}).$$

Then either $s_0 \in \sigma(A)$ or, by the previous theorem,

$$s_0 \in \{s \in \rho(A) : \|G(s)\| > \tilde{\rho}^{-1}\} \text{ for all } \tilde{\rho} \in (\rho, \bar{\rho}).$$

Hence $\|G(s_0)\|_{\mathcal{L}(U,Y)} \geq \rho^{-1}$ and this concludes the proof of (4.15).
It follows directly from the previous theorem that a global minimum $\mu \in \rho(A)$ of $\|G(\cdot)\|_{\mathcal{L}(U,Y)}$ will not be contained in any spectral value set $\sigma(A,D,E;\rho)$ with $\rho < \bar{\rho}$. On the other hand, if $\|G(s_0)\|_{\mathcal{L}(U,Y)} > \bar{\rho}^{-1}$ then there exists $\rho < \bar{\rho}$ such that $\|G(s_0)\|_{\mathcal{L}(U,Y)} > \rho^{-1}$ and so $s_0 \in \sigma(A,D,E;\rho)$ by (4.13). This proves (4.16) and the final statement of the corollary. □

Remark 4.2.13 Corollary 4.2.12 makes clear an important difference between the finite dimensional case and the infinite dimensional one: we *have not* been able to prove that the closure of $\sigma(A,D,E;\rho)$ is equal to

$$\sigma(A) \cup \{s \in \rho(A); \|G(s)\|_{\mathcal{L}(U,Y)} \geq \rho^{-1}\}, \tag{4.17}$$

but just that it is a *subset* of (4.17). With regard to this observation the following facts are worth mentioning. For the unstructured case we have, see [33], that if $s \mapsto \|G(s)\|$ is not *locally constant* (i.e., $s \mapsto \|G(s)\|_{\mathcal{L}(U,Y)}$ is not constant in any open subset Ω of $\rho(A)$), then equality in (4.15) holds. In this case, the map $\rho \mapsto \sigma(A,D,E;\rho)$ is upper semicontinuous, as it was in the matrix case (Proposition 2.2.13). Further, in [6, Theorem 5.1] the authors proved that the norm of the resolvent of bounded operators in L_p spaces *can not* be locally constant. It is unclear whether this property holds in general.

Remark 4.2.14 Again we make a comment with respect to the closedness radius in the unstructured case. In Remark 4.2.9 we have pointed out that (4.8) can be conservative. However, Equation (4.16) tells us that this bound is enough for purposes related to spectral value sets. Indeed, even if $\bar{\rho}$ is finite, for $\rho > 0$ such that $\rho < \bar{\rho}$, the corresponding spectral value sets $\sigma(A,D,E;\rho)$ will cover the whole plane minus the set where the infimum of $\|G\|$ is achieved, see Equation (4.16).

4.2.3 C_g-stability radius

Spectral value sets are closely related to the following problem. In certain applications, for example in control design by linear feedback, many desirable properties can be expressed by the requirement that A has its spectrum in a prescribed open subset C_g of the complex plane. Hence, it is of interest to determine the robustness of this property of $\sigma(A)$, i.e., $\sigma(A) \subset C_g$, with respect to perturbations of the form $D\Delta E$, where D, Δ and E are defined as in the sections above. A natural measure for this robustness is the following

Definition 4.2.15 Let C_g be an open subset of the complex plane and suppose that $\sigma(A) \subset C_g$. The number

$$r(A, D, E, C_g) = \inf\{\|\Delta\|_{\mathcal{L}(Y,U)} \,|\, \sigma(A + D\Delta E) \not\subset C_g\}$$

is called C_g*-stability radius.*

Figure 4.2 illustrates the situation. A useful characterisation of $r(A, D, E; C_g)$ can be easily

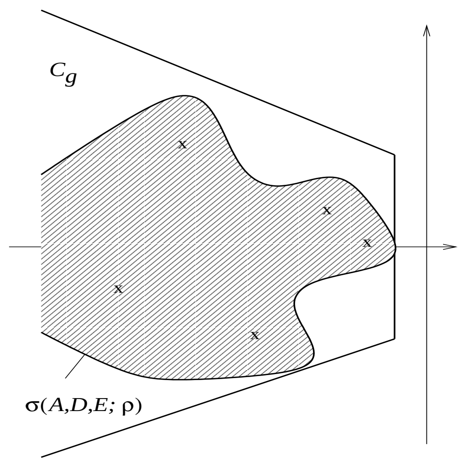

Figure 4.2: Set C_g and the spectral value set $\sigma(A, D, E; \rho)$ for $\rho = r(A, D, E, C_g)$.

obtained by means of Theorem 4.2.11.

Theorem 4.2.16 *Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$ and (4.2). Further, let us assume that the limit $\lim_{|s|\to\infty} \|G(s)\|_{\mathcal{L}(U,Y)}$ exists and is finite. Finally, let $C_g \subset \mathbb{C}$ be an open set such that $\sigma(A) \subset C_g$. Then the following formula holds*

$$r(A, D, E, C_g) = \left[\sup_{s \in \partial C_g} \|G(s)\|_{\mathcal{L}(U,Y)} \right]^{-1}, \tag{4.18}$$

where ∂C_g denotes the boundary of C_g.

Proof: If $G \equiv 0$ the assertion is trivial. Thus, let us suppose that $G \not\equiv 0$. Then it follows from Proposition 4.2.5 that

$$r(A, D, E, C_g) \geq \inf_{s \in \mathbb{C} \setminus C_g} \|G(s)\|^{-1}_{\mathcal{L}(U,Y)}.$$

On the other hand, the proof of Proposition 4.2.10 implies that for every $s \in \mathbb{C} \setminus C_g (\subset \rho(A))$ and every $\varepsilon > 0$, there exists a disturbance $\Delta \in \mathcal{L}(Y, U)$ such that

$$\mathbb{C} \setminus C_g \ni s \in \sigma(A + D\Delta E)$$

and

$$\|G(s)\|^{-1} \leq \|\Delta\|_{\mathcal{L}(Y,U)} < \|G(s)\|^{-1} + \varepsilon.$$

We conclude that

$$r(A, D, E, C_g) = \inf_{s \in \mathbb{C} \setminus C_g} \|G(s)\|^{-1}_{\mathcal{L}(U,Y)} = \left[\sup_{s \in \mathbb{C} \setminus C_g} \|G(s)\|_{\mathcal{L}(U,Y)} \right]^{-1}. \tag{4.19}$$

Since $\sigma(A) \subset C_g$, $G(s)$ is analytic in $\mathbb{C} \setminus \bar{C}_g$. Now, applying the Maximum Principle to $\mathbb{C} \setminus \bar{C}_g$ [41, p. 100], we obtain (4.18) from (4.19). □

We stress that the C_g-stability radius is not (at least directly) related with usual concepts of "stability" in dynamical systems theory like growth bounds for the corresponding semigroups $e^{(A+D\Delta E)t}$. Here, the terminology "stability" means just the robustness of the property "$\sigma(A) \subset C_g$" with respect to perturbations of the nominal operator A.

4.3 Summary

In this chapter we have defined and characterised *spectral value sets* for infinite dimensional systems which satisfy certain mild assumptions. Moreover, related objects like *closedness radius* and C_g-*stability radius* have also been studied. Essentially, we have proved that the key quantity for the analysis is, as in the finite dimensional case, the function $s \mapsto \|ER(s, A)D\|_{\mathcal{L}(U,Y)}$.

Chapter 5

Abstract results on convergences of operators

In the previous chapter we have shown that accessing the norm of the transfer function $G(s) = ER(s,A)D$ is the key for the calculation of closedness radius, spectral value sets and C_g-stability radii. However, in contrast to the matrix case, for most operators acting in infinite dimensional spaces, it is impossible to get formulae for these norms or even for the transfer function. The natural way is then to generate finite dimensional approximations of the transfer function. With this aim in mind we present in this chapter some abstract convergence results related to projection methods and/or discrete operators. These results build the foundation of the rest of this dissertation. Our approach is based on the works [12], [81], [82] and [1].

5.1 Preliminaries

Before we may go to the main statements we must introduce some notations and make some "tuning" comments.

5.1.1 Notations

Approximation shall be the central topic in the sequel. Thus, we need special notation for sequences and convergences: with certain abuse of notation, we denote by "$(q_N)_{N\in\mathbb{N}} \in \mathcal{Q}$" a sequence with elements q_N in the set $\mathcal{Q}$ for every $N \in \mathbb{N}$.
Let us recall now some standard definitions [53, III.3.1].

Definition 5.1.1 Suppose that T, $(T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W,V)$. Then

1. T_N converges *strongly* to T, to be denoted

$$T_N \xrightarrow[V]{s} T,$$

if

$$\lim_{N\to\infty} \|(T_N - T)w\|_V = 0; \quad \text{for all } w \in W.$$

2. T_N converges *in norm* to T, to be denoted

$$T_N \xrightarrow[\mathcal{L}(W,V)]{n} T,$$

if

$$\lim_{N\to\infty} \|T_N - T\|_{\mathcal{L}(W,V)} = 0.$$

In the sequel we shall make use of the following objects.

- For any $x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$, $|x| := \left(\sum_{j=1}^n |x_j|^2\right)^{\frac{1}{2}}$.
- For given numbers a and b such that $-\infty < a < b < \infty$, $\Omega := (a, b)$.
- $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$.
- For any $\alpha \in \mathbb{N}_0$, $D^\alpha := \frac{d^\alpha}{d\omega^\alpha}$.
- $C^k(\Omega; \mathbb{C}^n)$: Vector space of all continuous $\mathbb{C}^n$-valued functions x on Ω such that for all $\alpha \in \mathbb{N}_0$ with $\alpha \le k$, the function $D^\alpha x$ exists and is continuous on Ω.
- $C^k(\bar{\Omega}; \mathbb{C}^n)$: Vector space of all bounded functions $x \in C^k(\Omega; \mathbb{C}^n)$ such that for all $\alpha \in \mathbb{N}_0$ with $\alpha \le k$, the functions $D^\alpha x$ can be extended so as to be bounded and continuous on $\bar{\Omega}$. $C^k(\bar{\Omega}; \mathbb{C}^n)$ is a Banach space equipped with the norm

$$\|x\|_{k,\infty} := \max_{\alpha \le k} \sup_{\omega \in \Omega} |D^\alpha x(\omega)|.$$

- $L_r^2(\Omega; \mathbb{C}^n)$: Vector space of all (in Ω) Lebesgue measurable $\mathbb{C}^n$-valued functions with respect to the weight function

$$r : \Omega \to \mathbb{R}, \quad \int_\Omega r(\omega) d\omega < \infty.$$

  $L_r^2(\Omega; \mathbb{C}^n)$ is a Banach space endowed with the norm

$$\|x\|_{L_r^2(\Omega;\mathbb{C}^n)} := \left(\int_\Omega r(\omega)|x(\omega)|^2 d\omega\right)^{\frac{1}{2}} < \infty.$$

- $L^2(\Omega; \mathbb{C}^n)$: Vector space of all (in Ω) Lebesgue measurable $\mathbb{C}^n$-valued functions such that

$$\|x\|_{L_2} := \left(\int_\Omega |x(\omega)|^2 d\omega\right)^{\frac{1}{2}} < \infty. \tag{5.1}$$

- $H^k(\Omega; \mathbb{C}^n)$: Vector space of all $x \in L^2(\Omega; \mathbb{C}^n)$ such that the functions $D^\alpha x$, $\alpha < k$, are absolutely continuous and $D^k x$ is in $L^2(\Omega; \mathbb{C}^n)$. $H^k(\Omega; \mathbb{C}^n)$ is a Banach space equipped with the norm (see (5.1))

$$\|x\|_{k,L_2} := \left(\sum_{\alpha \le k} \|D^\alpha x\|_{L_2}^2\right)^{\frac{1}{2}}. \tag{5.2}$$

- $H^k(\bar{\Omega};\mathbb{C}^n)$: Vector space of all $x \in H^k(\Omega;\mathbb{C}^n)$ such that the functions $D^\alpha x$, $\alpha < k$, can be extended so as to be absolutely continuous on $\bar{\Omega}$. $H^k(\bar{\Omega};\mathbb{C}^n)$ is a Banach space equipped with the norm (5.2).
- $H_0^k(\bar{\Omega};\mathbb{C}^n)$: Closure of the subspace of $H^k(\bar{\Omega};)$ formed by the $x \in H^k(\bar{\Omega};\mathbb{C}^n)$ such that $D^\alpha x(a) = D^\alpha x(b) = 0$, $\alpha < k$.

Remark 5.1.2 $H^k(\Omega;\mathbb{C}^n)$ is the completion of $C^k(\bar{\Omega};\mathbb{C}^n)$ with respect to the norm (5.2) [15, p. 578].

In this work, unless otherwise stated, by "subspace" (of a Banach space) is meant "closed linear subspace".

5.1.2 Standard convergence theorems

The following theorem will be frequently used in the sequel [12, Banach-Steinhaus Theorem]

Theorem 5.1.3 *Let T, $(T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W,V)$. Then*

$$T_N \xrightarrow[V]{s} T,$$

iff

1. *The sequence $(T_N)_{N\in\mathbb{N}}$ is uniformly bounded, i.e.,*

$$\sup_{N\in\mathbb{N}} \|T_N\|_{\mathcal{L}(W,V)} \leq M.$$

2. $\lim_{N\to\infty} \|(T_N - T)w\|_V = 0$ *for all w in some set W_0 dense in W.*

We shall also use the following lemma [53, Lemma III.3.8]

Lemma 5.1.4 *Suppose that*

$$T,\ (T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W,V), \quad S,\ (S_N)_{N\in\mathbb{N}} \in \mathcal{L}(V,U).$$

Then, if

$$T_N \xrightarrow[V]{s} T \text{ and } S_N \xrightarrow[U]{s} S,$$

it follows that

$$S_N T_N \xrightarrow[U]{s} ST.$$

Finally, we give a simple but useful result [53, Problem III.3.10].

Lemma 5.1.5 *Let $(T_N)_{N\in\mathbb{N}}$, $T \in \mathcal{L}(W,V)$ be such that*

$$T_N \xrightarrow[V]{s} T.$$

Further, let $(w_N)_{N\in\mathbb{N}} \in W$ be a sequence with the property

$$\lim_{N\to\infty} \|w_N - w\|_W = 0$$

for some $w \in W$. Then

$$\lim_{N\to\infty} \|T_N w_N - Tw\|_V = 0.$$

5.1.3 Normwise convergence of a product

We continue this preliminary considerations with an important lemma which will become our main tool in the analysis to be carried out in the next chapter.

Lemma 5.1.6 *Let $T \in \mathcal{K}(W, V)$ and $S \in \mathcal{L}(V, U)$. Suppose that sequences*

$$\begin{aligned} &(T_N)_{N\in\mathbb{N}}, \quad T_N : W \to V, \quad N \in \mathbb{N} \\ &(S_N)_{N\in\mathbb{N}}, \quad S_N : V \to U, \quad N \in \mathbb{N} \end{aligned}$$

are given such that

$$T_N \xrightarrow[\mathcal{L}(W,V)]{n} T \text{ and } S_N \xrightarrow[U]{s} S.$$

Then

$$S_N T_N \xrightarrow[\mathcal{L}(W,U)]{n} ST.$$

Proof: First, we note that by the Banach Steinhaus Theorem $(S_N)_{N\in\mathbb{N}}$ is uniformly bounded, i.e.,

$$\|S_N\|_{\mathcal{L}(V,U)} < M < \infty, \quad \forall\, N \in \mathbb{N}. \tag{5.3}$$

Without loss of generality we may consider that $\|S\|_{\mathcal{L}(V,U)} < M$. We must show that for every $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that

$$\|STw - S_N T_N w\|_U < \epsilon$$

for every $N > N_\epsilon$ and every $w \in O_W$, where

$$O_W := \{r \in W; \quad \|r\| = 1\}.$$

In fact, let $w \in O_W$. Then we have

$$\begin{aligned} \|STw - S_N T_N w\|_U &= \|(S - S_N)Tw + S_N(T - T_N)w\|_U \\ &\le \|(S - S_N)Tw\|_U + \|S_N(T - T_N)w\|_U. \end{aligned} \tag{5.4}$$

From $T_N \xrightarrow[\mathcal{L}(W,V)]{n} T$ and (5.3), it follows that there exists an $N_{1\epsilon}$ such that

$$\|S_N(T - T_N)w\|_U \le M\|(T - T_N)w\|_V < \epsilon/3 \text{ for } N > N_{1\epsilon}.$$

Furthermore, since $T \in \mathcal{K}(W, V)$, the set $T(O_W)$ is relatively compact. Thus, there exists a finite number $N_{2\epsilon}$ of $v_i \in V$, $i = 1, \ldots, N_{2\epsilon}$, such that for *every* $w \in O_W$, there is some v_i with

$$\|Tw - v_i\|_V \le \frac{\epsilon}{6M}.$$

Thus,

$$\begin{aligned}\|(S-S_N)Tw\|_U &\le \|(S-S_N)v_i\|_U + \|(S-S_N)(Tw-v_i)\|_U \\ &\le \max_{j=1,\dots,N_{2\epsilon}} \|(S-S_N)v_j\|_U + 2M\|Tw-v_i\|_V \\ &\le \max_{j=1,\dots,N_{2\epsilon}} \|(S-S_N)v_j\|_U + \epsilon/3.\end{aligned}$$

It remains only to see that $S_N \xrightarrow[U]{s} S$ implies that there exists a $N_{3\epsilon} < \infty$ such that

$$\max_{j=1,\dots,N_{2\epsilon}} \|(S-S_N)v_j\|_U \le \epsilon/3 \text{ for } N > N_{3\epsilon}.$$

The proof is completed by choosing

$$N_\epsilon = \max_{1,2,3} N_{i\epsilon}.$$

□

A similar statement, that for bounded T and compact S,

$$T_N \xrightarrow[V]{s} T \text{ and } S_N \xrightarrow[\mathcal{L}(V,U)]{n} S \quad \Rightarrow \quad S_N T_N \xrightarrow[\mathcal{L}(W,U)]{n} ST$$

is false as the following counter-example shows. Consider the Banach space $W := \ell^2$, i.e., the set of sequences of real numbers $w := (w_i)_{i\in\mathbb{N}}$ such that

$$\|w\| := \left(\sum_{i=1}^{\infty} |w_i|^2\right)^{1/2} < \infty,$$

and define the sequences of operators

$$T_N w := (w_N, w_{N+1}, w_{N+2}, \dots), \quad S_N w := (w_1, \frac{w_2}{2}, \frac{w_3}{3}, \dots) =: Sw.$$

Then, $T_N \xrightarrow[W]{s} T := 0$, S is compact [15, Example A.3.23] and $S_N \xrightarrow[\mathcal{L}(W)]{n} S$. Nevertheless, the sequence

$$S_N T_N w = (w_N, \frac{w_{N+1}}{2}, \frac{w_{N+2}}{3}, \dots)$$

does not converge in norm to the null operator $ST = 0$. Note also that for the sequence

$$T_N S_N w = (\frac{w_N}{N}, \frac{w_{N+1}}{N+1}, \frac{w_{N+2}}{N+2}, \dots)$$

we have

$$T_N S_N \xrightarrow[\mathcal{L}(W)]{n} 0.$$

The following trivial example shows that the assumption on the compactness of T is essential. Suppose that for some $(S_N)_{N\in\mathbb{N}}$, $S \in \mathcal{L}(W)$ it holds that

$$S_N \xrightarrow[W]{s} S$$

but *not*

$$S_N \xrightarrow[\mathcal{L}(W)]{n} S$$

and let $T = I_W$, $T_N = I_W$, $N \in \mathbb{N}$, where I_W is the identity operator in W. One sees immediately that

$$T_N \xrightarrow[\mathcal{L}(W)]{n} T.$$

However, the statement

$$S_N T_N = S_N \xrightarrow[\mathcal{L}(W)]{n} S = ST$$

does not hold. The reason for the theorem to "fail" is, of course, that the identity operator I_W is not compact.

5.1.4 Uniform convergence of operator valued functions

Lemma 5.1.7 *Let $(S_N(s))_{N\in\mathbb{N}}$, $S(s) \in \mathcal{L}(W,V)$ for each $s \in K$, where $K \subset \mathbb{C}$ is a compact set. Furthermore, let us assume that there exists $M > 0$ such that*

$$\begin{aligned} \|S(s_1) - S(s_2)\|_{\mathcal{L}(W,V)} &\le M|s_1 - s_2|, \quad \forall s_1, s_2 \in K, \\ \|S_N(s_1) - S_N(s_2)\|_{\mathcal{L}(W,V)} &\le M|s_1 - s_2|, \quad \forall s_1, s_2 \in K. \end{aligned} \tag{5.5}$$

Finally, suppose that

$$S_N(s) \xrightarrow[\mathcal{L}(W,V)]{n} S(s), \quad \forall\, s \in K. \tag{5.6}$$

Then, this convergence is uniform in $s \in K$.

Proof: We must show that for every $\epsilon > 0$, there exists $N(\epsilon) \in \mathbb{N}$ such that for every $N > N(\epsilon)$ the following inequality holds

$$\|S(s) - S_N(s)\|_{\mathcal{L}(W,V)} < \epsilon, \quad \forall\, s \in K.$$

Indeed, since K is compact, for every $\epsilon > 0$, it has a finite cover of diameter $\delta := \epsilon/(3M)$, i.e., one can find n_δ elements $s_j \in K$, $j = 1, \ldots, n_\delta$, such that for every $s \in K$, there exists an $i \in \{1, \ldots, n_\delta\}$ for which $|s - s_i| < \delta$.
Let $s \in K$ and choose s_i such that $|s - s_i| < \delta$. It is easy to see that

$$\begin{aligned} \|S(s) - S_N(s)\|_{\mathcal{L}(W,V)} \le \|S(s) - S(s_i)\|_{\mathcal{L}(W,V)} + \|S(s_i) - S_N(s_i)\|_{\mathcal{L}(W,V)} + \\ \|S_N(s_i) - S_N(s)\|_{\mathcal{L}(W,V)}. \end{aligned} \tag{5.7}$$

Now, using (5.5), we obtain bounds for the first and third summands of (5.7):

$$\|S(s) - S(s_i)\|_{\mathcal{L}(W,V)} \le \epsilon/3, \quad \|S_N(s) - S_N(s_i)\|_{\mathcal{L}(W,V)} \le \epsilon/3. \tag{5.8}$$

Finally, by (5.6), there exists a number $N_1(\epsilon) \in \mathbb{N}$ such that

$$\|S(s_j) - S_N(s_j)\|_{\mathcal{L}(W,V)} < \epsilon/3, \quad \forall\, j \in \{1, \ldots, n_\delta\}\}, \quad N > N(\epsilon). \tag{5.9}$$

Now, it is easy to see that for $N > N(\epsilon)$, the inequalities displayed in (5.7), (5.8) and (5.9) give us

$$\|S(s) - S_N(s)\|_{\mathcal{L}(W,V)} < \epsilon, \quad \forall\, N > N(\epsilon).$$

Since $s \in K$ was arbitrarily chosen, the proof is complete. □

The following proposition will be useful in the next chapter.

Proposition 5.1.8 *Let $T \in \mathcal{K}(U,W)$ and $(T_N)_{N\in\mathbb{N}}$ be such that*

$$T_N \xrightarrow[\mathcal{L}(U,W)]{n} T.$$

Furthermore, let $K \subset \mathbb{C}$ be a compact set and suppose that we are given operator valued functions

$$(S_N(s))_{N\in\mathbb{N}},\ S(s) :\in \mathcal{L}(W,V), \quad \forall\, s \in K,$$

such that (5.5) holds and

$$S_N(s) \xrightarrow[V]{s} S(s), \quad \forall\, s \in K.$$

Consider the operator valued functions

$$Q_N(s) := S_N(s)T_N, \quad Q(s) := S(s)T, \quad \forall\, N \in \mathbb{N}, \quad \forall\, s \in K.$$

Then

$$Q_N(s) \xrightarrow[\mathcal{L}(U,V)]{n} Q(s), \quad \forall\, s \in K, \tag{5.10}$$

Moreover, this convergence is uniform in $s \in K$.

Proof: Equation (5.10) follows by a straightforward use of Lemma 5.1.6. It remains to show the uniformity of this convergence. By Lemma 5.1.7, we must show that there exists a constant $M_0 > 0$ for which

$$\begin{aligned}
\|Q(s_1) - Q(s_2)\|_{\mathcal{L}(U,V)} &\le M_0|s_1 - s_2|, \quad \forall s_1,\, s_2 \in K,\\
\|Q_N(s_1) - Q_N(s_2)\|_{\mathcal{L}(U,V)} &\le M_0|s_1 - s_2|, \quad \forall s_1,\, s_2 \in K.
\end{aligned}$$

But this is rather obvious. First, we observe that since $T_N \xrightarrow[\mathcal{L}(U,W)]{n} T$, the sequence $(T_N)_{N\in\mathbb{N}}$ is uniformly bounded. Further, since we have assumed (5.5) it holds that

$$\begin{aligned}
\|Q(s_1) - Q(s_2)\|_{\mathcal{L}(U,V)} &\le M|s_1 - s_2|\|T\|_{\mathcal{L}(U,W)}, \quad \forall\, s_1,\, s_2 \in K,\\
\|Q_N(s_1) - Q_N(s_2)\|_{\mathcal{L}(U,V)} &\le M|s_1 - s_2|\|T_N\|_{\mathcal{L}(U,W)}, \quad \forall\, s_1,\, s_2 \in K.
\end{aligned}$$

Making

$$M_0 := M \max\{\|T\|_{\mathcal{L}(U,W)},\ \sup_{N\in\mathbb{N}} \|T_N\|_{\mathcal{L}(U,W)}\}$$

ends the proof. □

5.2 On convergence of discrete operators

The aim of this section is to introduce the concept of *discrete* operators. We shall also show how to obtain normwise and strong convergence using discrete operators.

5.2.1 Discrete operators

The concept of *discrete operator* [12] can be introduced as follows. Let W and V be Banach spaces. Suppose that two sequences $(W_N)_{N\in\mathbb{N}}$ and $(V_N)_{N\in\mathbb{N}}$ of *finite dimensional* subspaces in W and V, respectively, are given and, with them, sequences of projections

$$(\pi_N^W)_{N\in\mathbb{N}}, \quad \pi_N^W : W \to W_N, \quad N \in \mathbb{N} \tag{5.11}$$
$$(\pi_N^V)_{N\in\mathbb{N}}, \quad \pi_N^W : V \to V_N, \quad N \in \mathbb{N}. \tag{5.12}$$

Further, let us denote by

$$\iota_N^W : W_N \to W, \quad \iota_N^V : V_N \to V, \quad N \in \mathbb{N} \tag{5.13}$$

the corresponding natural embeddings. Finally, let us suppose that we are given a sequence of linear operators

$$\mathcal{T}_N : W_N \to V_N, \quad N \in \mathbb{N}. \tag{5.14}$$

Note that for each $N \in \mathbb{N}$, $\mathcal{T}_N$ has matrix representation with respect to given bases in W_N and V_N.

Definition 5.2.1 Suppose that a framework (5.11), (5.12), (5.13), (5.14) is given. The operator sequence, denoted $(T_N)_{N\in\mathbb{N}}$,

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \forall\, N \in \mathbb{N}, \tag{5.15}$$

is called *discrete sequence.* Furthermore, its elements T_N, defined by a relationship (5.15) are called *discrete operators.*

Remark 5.2.2 Note that we use here the notion of *projection* in a nonstandard sense. Indeed, a projection π (of W onto $\hat{W}$), is *usually* an operator $\pi \in \mathcal{L}(W)$ such that $\pi^2 = \pi$ (idempotent) and $\mathrm{Rg}(\pi) = \hat{W}$, where $\hat{W}$ is a (closed linear) subspace of W [53, Section 1.3.4]. On the other hand, for us a *projection* is a closed linear operator $\pi : W \to \hat{W}$, such that the operator $\iota\pi : W \to W$ is a projection of W into W in the usual sense with $\hat{W} = \pi(W)$. Here $\iota : \hat{W} \to W$ denotes the natural embedding of the subspace $\hat{W}$ into W. The reason for this unusual terminology is that it fits more easily into the usual framework of regular convergence to be introduced later.

5.2.2 Normwise convergence

The following proposition shows how to obtain convergence in norm with the help of discrete sequences.

Proposition 5.2.3 *Let W and V be Banach spaces and consider some $T \in \mathcal{K}(W,V)$. Suppose that sequences of projections and embeddings*

$$(\pi_N^W)_{N\in\mathbb{N}}, \quad \pi_N^W : W \to W_N, \quad \iota_N^W : W_N \to W$$
$$(\pi_N^V)_{N\in\mathbb{N}}, \quad \pi_N^V : V \to V_N, \quad \iota_N^V : V_N \to V$$

are given such that

$$\pi_N^{W^*}\iota_N^{W^*} \xrightarrow[W^*]{s} I_{W^*}, \qquad \iota_N^V \pi_N^V \xrightarrow[V]{s} I_V,$$

where $\pi_N^{W^*} := (\pi_N^W)^*$, is the adjoint operator to π_N^W and $\iota_N^{W^*} : W^* \to W_N^*$, which is a projection of W^*, the dual space of W, onto W_N^*, the dual space of W_N^*. Then the discrete sequence $(T_N)_{N\in\mathbb{N}}$ given by

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \mathcal{T}_N := \pi_N^V T \iota_N^W, \quad N \in \mathbb{N} \tag{5.16}$$

is such that

$$T_N \xrightarrow[\mathcal{L}(W,V)]{n} T.$$

Proof: By the Banach Steinhaus Theorem

$$\|\iota_N^V \pi_N^V\|_{\mathcal{L}(V)} < M < \infty, \quad \forall\, N \in \mathbb{N}.$$

Thus, we have

$$\begin{aligned}
\|T - T_N\|_{\mathcal{L}(W,V)} &= \|T - \iota_N^V \mathcal{T}_N \pi_N^W\|_{\mathcal{L}(W,V)} \\
&\le \|(I_V - \iota_N^V \pi_N^V)T\|_{\mathcal{L}(W,V)} + \|\iota_N^V \pi_N^V\|_{\mathcal{L}(V)} \|T(I_W - \iota_N^W \pi_N^W)\|_{\mathcal{L}(W,V)} \\
&= \|(I_V - \iota_N^V \pi_N^V)T\|_{\mathcal{L}(W,V)} + M\|(I_{W^*} - \pi_N^{W^*}\iota_N^{W^*})T^*\|_{\mathcal{L}(V^*,W^*)},
\end{aligned}$$

where the equality of the norms of a bounded operator and its adjoint [53, III.3.3] has been used. Since $T \in \mathcal{K}(W,V)$ so is $T^* \in \mathcal{K}(V^*,W^*)$ [53, Theorem III.4.10]. Using Lemma 5.1.6 we see now that both summands converge to zero. □

Remark 5.2.4 Conditions of the type

$$\pi_N^{W^*}\iota_N^{W^*} \xrightarrow[W^*]{s} I_{W^*}$$

hold often in applications. For example, whenever $(\pi_N^W)_{N\in N}$ is a sequence of orthogonal projections in a Hilbert space W and $\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W$.

5.2.3 Strong convergence

We continue with a simple proposition. It is similar to Proposition 5.2.3, but addresses strong convergence.

Proposition 5.2.5 *Let W and V be Banach spaces and consider some $T \in \mathcal{L}(W,V)$. Suppose that sequences of projections and embeddings*

$$\begin{array}{lll} (\pi_N^W)_{N\in\mathbb{N}}, & \pi_N^W : W \to W_N, & \iota_N^W : W_N \to W \\ (\pi_N^V)_{N\in\mathbb{N}}, & \pi_N^V : V \to V_N, & \iota_N^V : V_N \to V \end{array}$$

are given such that

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W, \quad \iota_N^V \pi_N^V \xrightarrow[V]{s} I_V.$$

Then the sequence $(T_N)_{N\in\mathbb{N}}$,

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \mathcal{T}_N := \pi_N^V T \iota_N^W, \quad N \in \mathbb{N}$$

is such that

$$T_N \xrightarrow[V]{s} T.$$

Proof: First, we note that by the Banach Steinhaus Theorem

$$\| \iota_N^V \pi_N^V \|_{\mathcal{L}(W)} \leq M < \infty, \quad \forall N \in \mathbb{N}.$$

We may suppose , without loss of generality, that $\|T\|_{\mathcal{L}(W,V)} < M$ as well. Let $w \in W$. Then

$$\begin{aligned} \|(T - T_N)w\|_V &\leq \|(I_V - \iota_N^V \pi_N^V)Tw\|_V + \| \iota_N^V \pi_N^V T(I_W - \pi_N^W)w\|_V \\ &\leq \|(I_V - \iota_N^V \pi_N^V)Tw\|_V + M^2 \|(I_W - \iota_N^W \pi_N^W)w\|_W. \end{aligned}$$

The proof is completed by using the strong convergence of both projections to the identity. □

5.3 Discrete-regular convergence

Our aim in this section is to make clear how strong convergence of *resolvent operators* can be achieved. For this we study here the concept of (discrete-) regular convergence for operators in $\mathcal{L}(W,V)$. We also show how regular convergence can be achieved with the help of projection methods. Clearly, these results may be applied to operators in $\mathcal{L}(X)$, while for general closed operators $A \in \mathcal{C}(X)$, one converts the corresponding domains $\mathcal{D}(A)$ into Banach spaces by means of the graph norm. Note that we do not try to give an extensive treatment of the topic, our definitions and lemmas are just adapted to our situation.

5.3.1 Preliminaries

We begin by recalling some standard results. For example, the following definitions [12, III.5.3.b], [81, Definition 2.1.1].

Definition 5.3.1 A sequence $(v_N)_{N\in\mathbb{N}} \in V$ is called *relatively compact* in V whenever every subsequence $(v_N)_{N\in\mathbb{N}_1}$, $\mathbb{N}_1 \subset \mathbb{N}$, contains a convergent (in V) subsequence $(v_N)_{N\in\mathbb{N}_2}$, $\mathbb{N}_2 \subset \mathbb{N}_1$.

We use below the notion of *discrete operator* introduced in Definition 5.2.1.

Definition 5.3.2 Let $T \in \mathcal{L}(W,V)$, where W and V are Banach spaces. We say that a sequence of discrete operators $(T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W,V)$ converges *regularly* to T, to be denoted

$$T_N \xrightarrow[\mathcal{L}(W,V)]{r} T,$$

if

1. $T_N \xrightarrow[V]{s} T$.
2. The following regularity condition is satisfied: each sequence $(w_N)_{N\in\mathbb{N}}$ in W such that
   (a) $(w_N)_{N\in\mathbb{N}}$ is bounded and $w_N \in W_N$, $\forall\, N \in \mathbb{N}$, and
   (b) $(T_N w_N)_{N\in\mathbb{N}}$ is relatively compact in V,

   is *itself* relatively compact in W.

The following lemma is useful in checking regular convergence.

Lemma 5.3.3 *Let $T \in \mathcal{L}(W,V)$, where W and V are Banach spaces and $(T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W,V)$. Suppose that $(T_N)_{N\in\mathbb{N}}$ satisfies the assumptions:*

i. $T_N \xrightarrow[V]{s} T$.

ii. Every bounded sequence $(w_N)_{N\in\mathbb{N}}$ in W such that $w_N \in W_N$, $\forall\, N \in \mathbb{N}$, has the following property: if $N_1 \subset \mathbb{N}$ is such that $(T_N w_N)_{N\in\mathbb{N}1}$ is convergent, then there exists $\mathbb{N}_2 \subset \mathbb{N}_1$ such that $(w_N)_{N\in\mathbb{N}_2}$ converges.

Then

$$T_N \xrightarrow[\mathcal{L}(W,V)]{r} T.$$

Proof: Only Point 2 of Definition 5.3.2 must be proved. Indeed, let $(w_N)_{N\in\mathbb{N}} \in W$ be such that (a) and (b) in Definition 5.3.2 above holds. Consider any subsequence $(w_N)_{N\in\mathbb{N}_1}$, $\mathbb{N}_1 \subset \mathbb{N}$. Then, due to 2.$(b)$, there exists $\mathbb{N}_2 \subset \mathbb{N}_1$ such that $(T_N w_N)_{N\in\mathbb{N}_2}$ converges. Further, by (ii), there exists $\mathbb{N}_3 \subset \mathbb{N}_2$ for which $(w_N)_{N\in\mathbb{N}_3}$ is convergent. It follows that $(w_N)_{N\in\mathbb{N}}$ is relatively compact. □

Regular convergence in $\mathcal{L}(W,V)$ can be used for convergence studies related to the class of bounded Fredholm operators.

Definition 5.3.4 An operator $T \in \mathcal{L}(W,V)$ is called (bounded) *Fredholm operator*, to be denoted $T \in \mathcal{F}(W,V)$, if its range $\mathrm{Rg}(T)$ is closed in V, $\operatorname{codim} \mathrm{Rg}(T) < \infty$ and $\dim \mathrm{Ker}(T) < \infty$.

Definition 5.3.5 Let $T \in \mathcal{F}(W,V)$. The *index* $\mathrm{ind}(T)$ of T is the (finite) number

$$\mathrm{ind}(T) := \dim \mathrm{Ker}(T) - \operatorname{codim} \mathrm{Rg}(T).$$

The set of all operators $T \in \mathcal{F}(W,V)$ such that $\operatorname{ind}(T) = 0$ shall be denoted by $\mathcal{F}_0(W,V)$. The set $\mathcal{F}_0(W,V)$ has a nice characterisation [81, pp. 653].

Lemma 5.3.6 *An operator $T \in \mathcal{F}_0(W,V)$ iff it is representable as a sum $T := R - S$, where $R \in \mathcal{L}(W,V)$ has bounded inverse $R^{-1} \in \mathcal{L}(V,W)$ and $S \in \mathcal{K}(W,V)$.*

The statement below can be found in [81, Theorem 1, pp. 655].

Lemma 5.3.7 *Let $T \in \mathcal{F}_0(W,V)$ be such that $\operatorname{Ker}(T) = \{0\}$. Then $\operatorname{Rg}(T) = V$ and $T^{-1} \in \mathcal{L}(V,W)$.*

The following result ([81, Theorem 1, pp 655]) show why regular convergence and operators in $\mathcal{F}_0(W,V)$ are important for us.

Proposition 5.3.8 *Let $T \in \mathcal{F}_0(W,V)$ be such that $\operatorname{Ker}(T) = \{0\}$. Furthermore, suppose that projections and embeddings are given which satisfy*

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W, \quad \iota_N^V \pi_N^V \xrightarrow[V]{s} I_V.$$

Consider the subspaces of W and V, respectively, given by

$$W_N = \pi_N^W W, \quad V_N = \pi_N^V V, \quad \forall N \in \mathbb{N},$$

and, with them, a sequence

$$\mathcal{T}_N : W_N \to V_N, \quad \forall N \in \mathbb{N},$$

such that $\mathcal{T}_N \in \mathcal{F}_0(W_N, V_N)$ for N large enough. Finally, suppose that

$$\iota_N^V \mathcal{T}_N \pi_N^W \xrightarrow[\mathcal{L}(W,V)]{r} T.$$

Then, also for N large enough, the inverses $\mathcal{T}_N^{-1}$ exist in $\mathcal{L}(V_N, W_N)$ and

$$\iota_N^W \mathcal{T}_N^{-1} \pi_N^V \xrightarrow[\mathcal{L}(V,W)]{r} T^{-1}. \tag{5.17}$$

Remark 5.3.9 By definition, the convergence in (5.17) implies

$$\iota_N^W \mathcal{T}_N^{-1} \pi_N^V \xrightarrow[W]{s} T^{-1}.$$

The following proposition will play an important role in the next chapter. It is a direct consequence of [80, Theorem 4.17, pp 69].

Proposition 5.3.10 *Let W and V be Banach spaces and suppose that projections and embeddings are given such that*

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W, \quad \iota_N^V \pi_N^V \xrightarrow[V]{s} I_V.$$

These projections define subspaces

$$W_N = \pi_N^W W, \quad V_N = \pi_N^V V, \quad \forall N \in \mathbb{N}.$$

Let $\Omega \subset \mathbb{C}$ be open and bounded and suppose that the operator valued functions

$$T : \Omega \to \mathcal{F}_0(W, V), \quad \mathcal{T}_N : \Omega \to \mathcal{F}_0(W_N, V_N), \quad N \in \mathbb{N}$$

are analytic and such that $T(s)$ is boundedly invertible for every $s \in \Omega$. Moreover, we assume that

$$\iota_N^V \mathcal{T}_N(s) \pi_N^W \xrightarrow[\mathcal{L}(W,V)]{r} T(s), \quad \forall\, s \in \Omega.$$

Finally, let us suppose that for each compact set $K \subset \Omega$ the following inequality holds

$$\sup_{N \in \mathbb{N}} \max_{s \in K} \|\mathcal{T}_N(s)\|_{\mathcal{L}(W_N, V_N)} < \infty.$$

Then, again for each compact set $K \subset \Omega$, there exists a $N_0 = N_0(K) \in \mathbb{N}$ such that $\mathcal{T}_N(s)$ is boundedly invertible for every $s \in K$ and every $N > N_0$. Moreover,

$$\sup_{N > N_0} \max_{s \in K} \|(\mathcal{T}_N(s))^{-1}\|_{\mathcal{L}(V_N, W_N)} < \infty. \tag{5.18}$$

5.3.2 Conditions for regular convergence in $\mathcal{L}(V)$.

We are specially interested in conditions under which projection methods yield regular convergence. We shall give a result which guarantees regular convergence of bounded Fredholm operators. For this, we introduce a "device" which allows a rather general treatment of the topic: *nice sequences.*

Let W_N, $V_N \subset V$ be finite dimensional subspaces of a Banach space V for each $N \in \mathbb{N}$. Suppose that sequences of projections and embeddings

$$\begin{aligned} (\pi_N^W)_{N \in \mathbb{N}}, \quad \pi_N^W : V \to W_N, \quad \iota_N^W : W_N \to V \\ (\pi_N^V)_{N \in \mathbb{N}}, \quad \pi_N^V : V \to V_N, \quad \iota_N^V : V_N \to V \end{aligned}$$

are given such that

$$\iota_N^W \pi_N^W \xrightarrow[V]{s} I_V, \quad \iota_N^V \pi_N^V \xrightarrow[V]{s} I_V. \tag{5.19}$$

Definition 5.3.11 The sequence $(\mathcal{P}_N)_{N \in \mathbb{N}}$ given by

$$\mathcal{P}_N : W_N \to V_N, \quad \mathcal{P}_N := \pi_N^V|_{W_N}, \quad N \in \mathbb{N}, \tag{5.20}$$

is called *nice*, to be denoted $(\mathcal{P}_N)_{N \in \mathbb{N}} \in \mathcal{N}$, if for sufficiently large $N \in \mathbb{N}$ the inverses

$$\mathcal{P}_N^{-1} : V_N \to W_N,$$

exist and the sequence formed by $\|\mathcal{P}_N^{-1}\|_{\mathcal{L}(W_N, V_N)}$, N sufficiently large, is uniformly bounded.

Nice sequences are important due to the following fact [12, Exercise 4.11].

Lemma 5.3.12 *Let $(\mathcal{P}_N)_{N\in\mathbb{N}} \in \mathcal{N}$. Then, the operators*

$$\mathcal{Q}_N : V \to W_N, \quad v \mapsto \mathcal{P}_N^{-1}\pi_N^V v, \tag{5.21}$$

are well defined for N large enough and are projections of V into W_N. Moreover, if (5.19) *holds then*

$$\iota_N^W \mathcal{Q}_N \xrightarrow[V]{s} I_V. \tag{5.22}$$

Proof: Obviously, $\iota_N^W \mathcal{Q}_N \in \mathcal{L}(V)$ for every $N \in \mathbb{N}$. Further, the identities

$$\iota_N^W \mathcal{Q}_N \, \iota_N^W \mathcal{Q}_N = \iota_N^W \mathcal{P}_N^{-1} \underbrace{\pi_N^V \iota_N^W \mathcal{P}_N^{-1}}_{\mathcal{P}_N\mathcal{P}_N^{-1}=I_{V_N}} \pi_N^V = \iota_N^W \mathcal{P}_N^{-1}\pi_N^V = \iota_N^W \mathcal{Q}_N.$$

show that $\iota_N^W \mathcal{Q}_N$ is idempotent and thus, that $(\mathcal{Q}_N)_{N\in\mathbb{N}}$ is a sequence of projections (in our sense!, see Remark 5.2.2). It remains to show (5.22). Indeed, it follows from (5.19) that $(\pi_N^V)_{N\in\mathbb{N}}$ is uniformly bounded. Furthermore, by assumptions, $(\mathcal{P}_N^{-1})_{N\in\mathbb{N}}$, N sufficiently large, is also uniformly bounded. Thus, see (5.21),

$$\| \iota_N^W \mathcal{Q}_N \|_{\mathcal{L}(V)} < M < \infty, \quad N \text{ sufficiently large.}$$

Since

$$\iota_N^W \pi_N^W = \iota_N^W \underbrace{\mathcal{Q}_N \, \iota_N^W}_{\mathcal{P}_N^{-1}\pi_N^V|_{W_N}=I_{W_N}} \pi_N^W = \iota_N^W \mathcal{Q}_N \, \iota_N^W \pi_N^W,$$

we have the identity

$$I_V - \iota_N^W \mathcal{Q}_N = (I_V - \iota_N^W \mathcal{Q}_N)(I_V - \iota_N^W \pi_N^W).$$

Thus, for any $v \in V$

$$\|(I_V - \iota_N^W \mathcal{Q}_N)v\|_V \leq (1+M)\|v - \iota_N^W \pi_N^W v\|_V.$$

By (5.19), the second factor above tends to zero when $N \to \infty$ and we obtain (5.22). □

Remark 5.3.13 Note that if $(\pi_N^V)_{N\in\mathbb{N}} = (\pi_N^W)_{N\in\mathbb{N}}$, they define a nice sequence $(\mathcal{P}_N)_{N\in\mathbb{N}} \in \mathbb{N}$.

The next proposition is based on [1, Theorem IV.15.3]. The concept of *nice* sequences introduced in Definition 5.3.11 plays a fundamental role in it.

Proposition 5.3.14 *Let V be a Banach space and $T \in \mathcal{L}(V)$ be of the form*

$$T := I_V - S, \tag{5.23}$$

where $S \in \mathcal{K}(V)$. Furthermore, let $(W_N)_{N\in\mathbb{N}}$, $(V_N)_{N\in\mathbb{N}}$ be sequences of finite dimensional subspaces of a Banach space V. Suppose that sequences of projections and embeddings

$$(\pi_N^W)_{N\in\mathbb{N}}, \quad \pi_N^W : V \to W_N, \quad \iota_N^W : W_N \to V$$
$$(\pi_N^V)_{N\in\mathbb{N}}, \quad \pi_N^V : V \to V_N, \quad \iota_N^V : V_N \to V$$

are given such that (5.19) *hold. Finally, suppose that the sequence of operators* $(\mathcal{P}_N)_{N\in\mathbb{N}}$ *given by* (5.20) *belongs to* $\mathcal{N}$. *Under these conditions, the sequence* $(T_N)_{N\in\mathbb{N}}$ *defined by*

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \mathcal{T}_N := \pi_N^V T \iota_N^W, \quad N \in \mathbb{N}, \tag{5.24}$$

is such that

$$T_N \xrightarrow[\mathcal{L}(V)]{r} T.$$

Proof: The first condition for regularity follows by Proposition 5.2.5 and $T \in \mathcal{L}(V)$.
The second condition can be verified as follows, see Lemma 5.3.3. Let $(w_N)_{N\in\mathbb{N}}$, with $w_N \in W_N$, be a bounded sequence in V such that

$$\lim_{N\to\infty} v_N = v \in V, \tag{5.25}$$

where

$$v_N := T_N w_N = \pi_N^V (I_V - S) w_N \in V_N, \quad N \in \mathbb{N}. \tag{5.26}$$

Then, by (5.26), we have that

$$\mathcal{P}_N w_N = \pi_N^V w_N = v_N + \pi_N^V S w_N = \pi_N^V (v_N + S w_N), \quad N \in \mathbb{N}. \tag{5.27}$$

Using the notation (see (5.21))

$$\mathcal{Q}_N := \mathcal{P}_N^{-1} \pi_N^V, \quad N \text{ sufficiently large},$$

and applying $\mathcal{P}_N^{-1}$ to (5.27), we obtain

$$w_N = \mathcal{Q}_N (v_N + S w_N), \quad N \text{ sufficiently large}.$$

We have that $S \in \mathcal{K}(V)$ and $(w_N)_{N\in\mathbb{N}}$ is bounded. Thus, $(Sw_N)_{N\in\mathbb{N}}$ has a convergent subsequence $(Sw_N)_{N\in\mathbb{N}_1}$, $\mathbb{N}_1 \subset \mathbb{N}$. We denote its limit by v_S. Furthermore, by (5.25), $(v_N)_{N\in\mathbb{N}_1}$ converges to v. It follows that

$$\lim_{\substack{N\to\infty \\ N\subset\mathbb{N}_1}} \iota_N^W w_N = \lim_{\substack{N\to\infty \\ N\subset\mathbb{N}_1}} \iota_N^W \mathcal{Q}_N (v_N + S w_N) = v + v_S.$$

In establishing the last equality we have used (Lemma 5.3.12)

$$\iota_N^W \mathcal{Q}_N \xrightarrow[V]{s} I_V$$

and Lemma 5.1.5. We have shown that $(w_N)_{N\in\mathbb{N}_1}$ converges to $w := (v + v_S) \in V$ which completes the proof. □

Remark 5.3.15 For bounded operators in Hilbert spaces and with $(\pi_N^V)_{N\in\mathbb{N}} = (\pi_N^W)_{N\in\mathbb{N}}$ being orthogonal projections, a convergence result somewhat similar to Proposition 5.3.14 has been known under the name of *quasitriangular convergence* [31].

Remark 5.3.16 A result similar to Proposition 5.3.14 can be obtained for *unbounded operators* of the form

$$T := (I_{V_0} - S)R^{-1}, \tag{5.28}$$

where

$$S,\ R \in \mathcal{K}(V_0)$$

with R^{-1} *unbounded* on V_0. In order to see this, one notes that considering the equation

$$Tz = x, \quad z \in \mathcal{D}(T), \quad x \in V_0,$$

is equivalent to solve for

$$u \in V := V_0 \times V_0$$

the equation

$$Fu = \begin{pmatrix} 0 \\ x \end{pmatrix},$$

where

$$F := \begin{pmatrix} I_{V_0} & -R \\ 0 & I_{V_0} - S \end{pmatrix} = I_V - \begin{pmatrix} 0 & -R \\ 0 & -S \end{pmatrix}.$$

One sees immediately that F has the form (5.23) required in Proposition 5.3.14.

5.3.3 Conditions for regular convergence in $\mathcal{L}(W, V)$.

Let us return now to the general situation. The proposition below is the analog to Proposition 5.3.14. It gives sufficient conditions for regular convergence in $\mathcal{L}(W, V)$.

Proposition 5.3.17 *Let $T \in \mathcal{F}_0(W, V)$. Furthermore, suppose that projections and embeddings are given such that that*

$$\iota_N^V \pi_N^V \xrightarrow[V]{s} I_V. \tag{5.29}$$

Finally, let $W_N := R^{-1}V_N$, $N \in \mathbb{N}$, and define the sequence of projections

$$\pi_N^W : W \to W_N, \quad w \mapsto R^{-1}\,\iota_N^V \pi_N^V R w, \quad N \in \mathbb{N}. \tag{5.30}$$

where R is derived from a representation of T in accordance with Lemma 5.3.6. Under these conditions, the sequence $(T_N)_{N\in\mathbb{N}} \in \mathcal{L}(W, V)$ defined by

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \mathcal{T}_N := \pi_N^V T\, \iota_N^W, \quad N \in \mathbb{N}, \tag{5.31}$$

is such that

$$T_N \xrightarrow[\mathcal{L}(W,V)]{r} T. \tag{5.32}$$

Proof: We begin by proving that

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W. \tag{5.33}$$

Indeed, consider some $w \in W$. Then

$$\| \iota_N^W \pi_N^W w - w\|_W \le \|R^{-1}\|_{\mathcal{L}(V,W)} \| \iota_N^V \pi_N^V Rw - Rw\|_V = \|R^{-1}\|_{\mathcal{L}(V,W)} \|(\iota_N^V \pi_N^V - I_V)Rw\|_V$$

and the second factor above tends to zero due to (5.29).
Having proved that, we see that the first condition for regularity in $\mathcal{L}(W,V)$, i.e., $T_N \xrightarrow[V]{s} T$, follows by Proposition 5.2.5.
In order to prove the second condition of regularity we need some initial considerations. Recall that for T we have fixed a representation $T := R - S$, where R and S enjoy the properties given in Lemma 5.3.6. Consider the operator

$$Q := TR^{-1} = I_V - SR^{-1} \in \mathcal{L}(V).$$

Using (5.30), it is easy to see that

$$T_N = \iota_N^V \pi_N^V Q R \, \iota_N^W \pi_N^W = \iota_N^V \pi_N^V Q \, \iota_N^V \pi_N^V R = Q_N R, , \quad N \in \mathbb{N},$$

where

$$Q_N := \iota_N^V \mathcal{Q}_N \pi_N^V, \quad \mathcal{Q}_N := \pi_N^V Q \, \iota_N^V \quad N \in \mathbb{N}.$$

Let us show that $Q_N \xrightarrow[\mathcal{L}(V)]{r} Q$. For this, we shall use Proposition 5.3.14. Indeed, $SR^{-1} \in \mathcal{K}(V)$. Furthermore, by assumptions

$$\iota_N^V \pi_N^V \xrightarrow[V]{s} I_V.$$

Thus, by Proposition 5.3.14 and Remark 5.3.13, it follows that

$$Q_N \xrightarrow[\mathcal{L}(V)]{r} Q. \tag{5.34}$$

Now, let us return to the proof of the second condition for regularity. For this, Lemma 5.3.3 will be used. Let $(w_N)_{N\in\mathbb{N}}$ be a bounded sequence in W with $w_N \in W_N$, $N \in \mathbb{N}$, and such that

$$\lim_{N\to\infty} T_N w_N = \lim_{N\to\infty} Q_N R w_N = v \in V. \tag{5.35}$$

Since R is bounded, $(Rw_N)_{N\in\mathbb{N}}$ is also bounded (in V) and, by the definition of $(V_N)_{N\in\mathbb{N}}$, $Rw_N \in V_N$ for all $N \in \mathbb{N}$. Further,

$$\lim_{N\to\infty} Q_N(Rw_N) = v \in V. \tag{5.36}$$

Since $Q_N \xrightarrow[\mathcal{L}(V)]{r} Q$, Equation (5.36) implies that there exists a subsequence $\mathbb{N}_1 \subset \mathbb{N}$ such that

$$\lim_{\substack{N\to\infty \\ N\in\mathbb{N}_1}} Rw_N = v_R \in V.$$

Finally, we use the bounded invertibility of R in order to obtain

$$\lim_{\substack{N\to\infty \\ N\in\mathbb{N}_1}} w_N = R^{-1} v_R \in W.$$

The proof is complete. □

Remark 5.3.18 In Proposition 5.3.17, the assumptions on the sequences of projections $(\pi_N^W)_{N\in\mathbb{N}}$ and $(\pi_N^V)_{N\in\mathbb{N}}$ are *symmetric* in the following sense: if

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W \tag{5.37}$$

and

$$\pi_N^V : V \to V_N, \quad \pi_N^V := R\,\iota_N^W \pi_N^W R^{-1}, \quad N \in \mathbb{N}, \tag{5.38}$$

then (5.32) holds for the sequence defined by (5.31). This fact is easily seen by observing that, for the sequences of projections (5.37) and (5.38), the relationships displayed in (5.29) and (5.30) also apply.

Remark 5.3.19 Proposition 5.3.17 can be strengthened if W and V are *Hilbert* spaces. Indeed, it is well known [26, Chapter XIV.2] that the condition imposed on T, i.e., $T \in \mathcal{F}_0(W,V)$, hold for its *adjoint* $T^* \in \mathcal{L}(V,W)$ as well. Thus, if we assume that $(\pi_N^V)_{N\in\mathbb{N}}$ is a sequence of *orthogonal* projections such that (5.29) holds, we obtain, eventually using the "symmetry" mentioned in Remark 5.3.18, regular convergence of the corresponding *adjoint* operators.

Remark 5.3.20 In Proposition 5.3.8 the sequence $(\mathcal{T}_N)_{N\in\mathbb{N}}$ was required to be such that $\mathcal{T}_N \in \mathcal{F}_0(W_N, V_N)$ for N large enough. The following simple arguments show that this condition is satisfied *automatically* by the sequence $(\mathcal{T}_N)_{N\in\mathbb{N}}$ defined in (5.31). Indeed, $T = R - S \in \mathcal{F}_0(W,V)$, where R and S enjoy the properties of Lemma 5.3.6. Thus, we see that

$$\mathcal{T}_N := \pi_N^V T\,\iota_N^W = \mathcal{R}_N - \mathcal{S}_N, \quad N \in \mathbb{N},$$

where

$$\mathcal{R}_N := \pi_N^V R\,\iota_N^W, \quad \mathcal{S}_N := \pi_N^V S\,\iota_N^W, \quad N \in \mathbb{N}.$$

Clearly,

$$\mathcal{R}_N \in \mathcal{L}(W_N, V_N), \quad \mathcal{R}_N^{-1} \in \mathcal{L}(V_N, W_N), \quad \mathcal{S}_N \in \mathcal{K}(W_N, V_N), \quad N \in \mathbb{N}.$$

Thus, $\mathcal{T}_N \in \mathcal{F}_0(W_N, V_N)$ for every $N \in \mathbb{N}$. It follows that (5.17) holds under the conditions of Proposition 5.3.17.

Remark 5.3.21 The following simple result is worth mentioning. It is a direct consequence of [82, Proposition 3.5, Theorem 4.1, Remark 3.1]: if

$$R \in \mathcal{L}(W,V), \quad R^{-1} \in \mathcal{L}(V,W), \quad (R_N)_{N\in\mathbb{N}} \in \mathcal{F}(W_N, V_N), \quad S \in \mathcal{K}(W,V)$$

then

$$R_N \xrightarrow[\mathcal{L}(W,V)]{r} R, \quad S_N \xrightarrow[\mathcal{L}(W,N)]{n} S \quad \Rightarrow \quad R_N - S_N \xrightarrow[\mathcal{L}(W,V)]{r} R - S.$$

Note that the convergence $S_N \xrightarrow[\mathcal{L}(W,V)]{n} S$ can be achieved if the projections $(\pi_N^W)_{N\in\mathbb{N}}$ are orthogonal, see Proposition 5.2.3.

5.3.4 Regular convergence in $\mathcal{L}(W, V_a)$

The results given above make use of the existence of a sequence $(W_N)_{N\in\mathbb{N}}$ of subspaces of W and projections onto them with the property

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W.$$

In some sense this is a natural and simple condition. However, in applications, the Banach space W is often the domain $\mathcal{D}(A)$ of some $A \in \mathcal{C}(X)$ endowed with the graph norm. In this situation, finding (or handling) the sequences $(\pi_N^W)_{N\in\mathbb{N}}$ can be difficult, for instance, if boundary conditions must be satisfied. We present here an approach which avoids this difficulty. For the sake of simplicity, and because our description covers most of the interesting cases, we formulate these results for the class of *ordinary differential operators*. For certain generalisations see [81] and [82].

Let us introduce the spaces

$$W := H^m([a,b];\mathbb{C}), \quad V := L^2([a,b];\mathbb{C}), \quad Z := W \cap \operatorname{Ker}(L). \tag{5.39}$$

where

$$\begin{aligned} &L : W \to \mathbb{C}^m, \quad w \mapsto \begin{pmatrix} L_1 w \\ \vdots \\ L_m w \end{pmatrix}, \\ &L_k w = \sum_{i=0}^{m-1} \left[\alpha_{ik} w^{(i)}(a) + \beta_{ik} w^{(i)}(b)\right], \quad k = 1, \dots, m. \end{aligned} \tag{5.40}$$

where both the α-s and the β-s are real or complex constants. Now, consider the operator

$$\begin{aligned} &T : Z \to V, \quad w \mapsto w^{(m)} - Cw, \\ &Cw := \sum_{i=0}^{m-1} u_i(t) w^{(i)}, \quad a < t < b, \end{aligned} \tag{5.41}$$

where the u_i, $i = 0, \dots, m-1$, are continuous functions in $[a, b]$. Since

$$C \in \mathcal{L}(H^m((a,b);\mathbb{C}), H^1((a,b);\mathbb{C}))$$

and the embedding

$$\iota : H^1((a,b);\mathbb{C}) \to L^2((a,b);\mathbb{C})$$

is compact [20, Theorem V.4.17], one sees easily that $C \in \mathcal{K}(W, V)$. Thus, *if* the operator

$$T_m : Z \to V, \quad w \mapsto w^{(m)} \tag{5.42}$$

is boundedly invertible it follows, by Lemma 5.3.6, that

$$T \in \mathcal{F}_0(Z, V).$$

Further, we consider the *extension* T_0 [53, Example III.2.7] of the operator T defined by:

$$T_0 : W \to V, \quad w \mapsto w^{(m)} - Cw,$$

and introduce the "augmented" operator:

$$T_a := \begin{pmatrix} T_0 \\ L \end{pmatrix} : W \to V_a,$$

where L is the operator (5.40) which defines the boundary conditions and

$$V_a := V \times \mathbb{C}^m. \tag{5.43}$$

It holds [82, pp 690]

Lemma 5.3.22 *Suppose that $T_m^{-1} \in \mathcal{L}(V, Z)$. Then*

$$T_a \in \mathcal{F}_0(W, V_a) \text{ and } \mathrm{Ker}(T_a) = \{0\}.$$

As a consequence $T_a^{-1} \in \mathcal{L}(V_a, W)$.

Due to Lemma 5.3.22, and by means of suitable sequences of projections, we may apply Proposition 5.3.17 in order to obtain *regularly* convergent approximations of T_a. The following simple results make possible the application of this approach in the approximation of T^{-1}.

Lemma 5.3.23 *Let T_m^{-1}, $T^{-1} \in \mathcal{L}(V, Z)$ and $I_a : V \to V_a$ be defined by*

$$I_a v := \begin{pmatrix} v \\ 0^{m\times 1} \end{pmatrix}, \quad \forall\, v \in V. \tag{5.44}$$

Then T^{-1} and $T_a^{-1} I_a$ represent the same operator in $\mathcal{L}(V, Z)$.

Proof: Both T^{-1} and $T_a^{-1} I_a$ are defined on the whole V. Further, consider some $v \in V$. Then there exists a unique $w := T^{-1} v \in Z$. It is enough to show that for any $v \in V$, the solution w_a of the equation

$$T_a w_a = I_a v \tag{5.45}$$

is unique and such that $w_a = w$. Indeed, by Lemma 5.3.22, $T_a^{-1} \in \mathcal{L}(V_a, W)$. Thus, the existence and uniqueness of w_a in W is clear. Further, the "first row" of Equation (5.45) states that $T_0 w_a = v$, while the "last" one states $w_a \in \mathrm{Ker}(L)$. But, this is exactly the solution of $Tw = v$ for $v \in V$. □

The preparation above allows us to state the following proposition.

Proposition 5.3.24 *Let W, V and V_a be given by (5.39) and (5.43). Suppose that T_m^{-1}, $T^{-1} \in \mathcal{L}(V, Z)$. Furthermore, let us assume that $(W_N)_{N\in\mathbb{N}}$ is a sequence of subspaces of W and that projections and embeddings*

$$\pi_N^W : W \to W_N, \quad \iota_N^W : W_N \to W, \quad N \in \mathbb{N},$$

are given such that

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W.$$

Finally, let

$$V_N := TW_N, \quad V_{aN} := T_a W_N, \quad N \in \mathbb{N},$$

and

$$\pi_N^V := T\,\iota_N^W \pi_N^W T^{-1}, \quad \pi_N^{V_a} := T_a\,\iota_N^W \pi_N^W T_a^{-1}, \quad N \in \mathbb{N}. \tag{5.46}$$

Then, the sequence $(T_{aN})_{N\in\mathbb{N}}$*, defined by*

$$T_{aN} := \iota_N^{Va} \mathcal{T}_{aN} \pi_N^W, \quad \mathcal{T}_{aN} := \pi_N^{Va} T_a\,\iota_N^W, \quad N \in \mathbb{N},$$

is such that $\mathcal{T}_{aN} \in \mathcal{F}_0(W_N, V_{aN})$ *for* N *large enough, and*

$$\iota_N^W \mathcal{T}_{aN}^{-1} I_a \pi_N^V \xrightarrow[\mathcal{L}(V,Z)]{r} T^{-1}.$$

Proof: By Lemma 5.3.22, $T_a \in \mathcal{F}_0(W, V_a)$ with $\mathrm{Ker}(T_a) = \{0\}$ and $T_a^{-1} \in \mathcal{L}(V_a, W)$. On the other hand, Proposition 5.3.17 guarantees the convergence

$$T_{aN} \xrightarrow[\mathcal{L}(W,V_a)]{r} T_a.$$

It follows (Proposition 5.3.8, Remark 5.3.20) that $\mathcal{T}_{aN} \in \mathcal{F}_0(W_N, V_{aN})$ for $N \in \mathbb{N}$ and

$$\iota_N^W \mathcal{T}_{aN}^{-1} \pi_N^{V_a} \xrightarrow[\mathcal{L}(V_a,W)]{r} T_a^{-1}.$$

Thus,

$$\iota_N^W \mathcal{T}_{aN}^{-1} \pi_N^{V_a} I_a \xrightarrow[\mathcal{L}(V,W)]{r} T_a^{-1} I_a. \tag{5.47}$$

On the other hand, we have the identity

$$\pi_N^{V_a} I_a = T_a\,\iota_N^W \pi_N^W \underbrace{T_a^{-1} I_a}_{T^{-1}} = T_a\,\iota_N^W \pi_N^W T^{-1} = T_a T^{-1} T\,\iota_N^W \pi_N^W T^{-1} = T_a T^{-1} \pi_N^V = I_a \pi_N^V,$$

which, together with Lemma 5.3.23 and (5.47), proves the relationship

$$\iota_N^W \mathcal{T}_{aN}^{-1} I_a \pi_N^V \xrightarrow[\mathcal{L}(V,Z)]{r} T^{-1}.$$

□

Remark 5.3.25 The advantage of this approach over the former one is, clearly, that the assumptions on the sequence $(W_N)_{N\in\mathbb{N}}$ are not so restrictive: the subspaces of $(W_N)_{N\in\mathbb{N}}$ are *not* assumed to lie in $\mathrm{Ker}(L)$.

5.3.5 Convergence of eigenelements

We finish this section with results on the approximate solution of eigenvalue problems.
Let $A \in \mathcal{C}(X)$ with resolvent set $\rho(A)$ and let $\lambda \in \sigma_P(A)$ be an isolated eigenvalue with finite *algebraic* multiplicity m. Suppose that a bounded domain $\Omega \subset \mathbb{C}$, $\Omega \cap \sigma(A) = \lambda$ is given such that $\partial\Omega \subset \rho(A)$ is a (positively oriented) rectifiable closed simple curve. The spectral projector associated to λ is defined by

$$P_\Omega = \frac{1}{2i\pi} \int_{\partial\Omega} R(s)ds. \tag{5.48}$$

Now, let $(A_N)_{N\in\mathbb{N}}$ be a sequence in $\mathcal{C}(X)$ such that $\partial\Omega \subset \rho(A_N)$ for N large enough. Then, a sequence of spectral projectors can be also defined

$$P_{N\Omega} = \frac{1}{2i\pi} \int_{\partial\Omega} R(s, A_N)ds. \tag{5.49}$$

For convenience we use the notations P_Ω and $P_{N\Omega}$ for the spectral projections. However, the reader should be aware of the fact that these projectors do not depend on Ω, but on the part of the spectrum which lies in it. We have used already the notion of spectral projections in Chapter 2, Equation (2.18).
With these notations the following proposition can be stated.

Proposition 5.3.26 *Let $A \in \mathcal{C}(X)$ be such that the natural embedding $\iota : \mathcal{D}(A) \to X$ is compact. Suppose that a bounded domain $\Omega \subset \mathbb{C}$, whose boundary $\partial\Omega$ is a rectifiable closed simple curve, is given with the property:*

$$\sigma(A) \cap \Omega = \{\lambda\}, \tag{5.50}$$

where $\lambda \in \sigma_P(A)(= \sigma(A))$ has algebraic multiplicity m. Finally, let us suppose that a discrete sequence $(A_N)_{N\in\mathbb{N}}$ is given such that

$$A_N \underset{\mathcal{L}(\mathcal{D}(A),X)}{\overset{r}{\longrightarrow}} A.$$

Then,

1. $s\,\iota - A_N \underset{\mathcal{L}(\mathcal{D}(A),X)}{\overset{r}{\longrightarrow}} s\,\iota - A, \quad \forall\, s \in \rho(A).$

2. *For N large enough, $\sigma(\mathcal{A}_N) \cap \Omega = \{\lambda_1^{(N)}, \ldots, \lambda_m^{(N)}\}$, counting their multiplicities, and*

$$\lim_{N\to\infty} |\lambda_i^{(N)} - \lambda| = 0, \quad i = 1, \ldots, m.$$

3. *The subspaces $P_{N\Omega}X$ and $P_\Omega X$ converge in* gap *[53, Chapter IV.2.1].*

Proposition 5.3.26 can be found in the literature [12, Table 5.1] and [12, Summary p. 238]. Similar results have been stated in [80, Theorem 55 and 62]. See also [82, Theorem 6.3].

Remark 5.3.27 If λ is a simple eigenvalue ($m = 1$) and X is a Hilbert space, the eigenprojections P_Ω take the form [15, Lemma 2.5.7.f]

$$P_\Omega = \langle \cdot , \psi \rangle \phi,$$

where ϕ and ψ are eigenvectors associated to λ of A and $\bar{\lambda}$ of its adjoint A^*, respectively. These vectors are related through the equalities [15, Lemma 2.5.7.f]

$$\langle \phi , \psi \rangle = 1$$

and are sometimes called *biorthogonal.* Moreover, Item 3 of Proposition 5.3.26 reduces to

$$\lim_{N\to\infty} \|\phi^{(N)} - \phi\|_X = 0,$$

and

$$\psi^{(N)} \xrightarrow[X]{s} \psi, \tag{5.51}$$

where $\phi^{(N)}$ and $\psi^{(N)}$ are, again, the eigenvectors associated to λ_N of A_N and $\bar{\lambda}_N$ of A_N^*, respectively.
Finally, if there is regular convergence of the corresponding adjoint operators (see Remark 5.3.19), (5.51) becomes

$$\lim_{N\to\infty} \|\psi^{(N)} - \psi\|_X = 0.$$

For generalised eigenvalue problems, i.e.,

$$Rx = \lambda Sx, \quad R,\, S \in \mathcal{L}(W, V), \tag{5.52}$$

we have a similar result [81, pp 683]. Below we give a reformulation of [82, Theorem 6.3]. In its statement we will use the notation $\sigma(R, S)$ for the spectrum of Problem (5.52), i.e., the set of $\lambda \in \mathbb{C}$ such that $R - \lambda S$ has no bounded inverse defined on the whole V.

Proposition 5.3.28 *Let* $R \in \mathcal{L}(W, V)$, $R^{-1} \in \mathcal{L}(V, W)$, $S \in \mathcal{K}(W, V)$, *and* $\mathrm{Ker}(R - \mu_0 S) = \{0\}$ *for some* $\mu_0 \in \Omega \subset \mathbb{C}$. *Here* Ω *is a bounded domain for which*

$$\sigma(S, R) \cap \Omega = \lambda$$

holds and whose boundary $\partial\Omega$ *is a rectifiable closed simple curve, where* λ *is an element of* $\sigma(R, S)$ *with algebraic multiplicity* m. *Let us suppose that sequences of projections* $(\pi_N^W)_{N\in\mathbb{N}}$, $(\pi_N^V)_{N\in\mathbb{N}}$ *and embeddings* $(\iota_N^W)_{N\in\mathbb{N}}$, $(\pi_N^V)_{N\in\mathbb{N}}$ *are given and, using them, define the following sequences of discrete operators*

$$\begin{aligned} R_N &:= \iota_N^V \mathcal{R}_N \pi_N^W, \quad \mathcal{R}_N := \pi_N^V R\, \iota_N^W, \quad N \in \mathbb{N}, \\ S_N &:= \iota_N^V \mathcal{S}_N \pi_N^W, \quad \mathcal{S}_N := \pi_N^V S\, \iota_N^W, \quad N \in \mathbb{N}. \end{aligned}$$

Finally, suppose that

$$R_N - \mu S_N \xrightarrow[\mathcal{L}(W,V)]{r} R - \mu S, \quad \forall\, \mu \in \Omega.$$

Then, the eigenelements of the generalised eigenvalue problems

$$\begin{aligned} Rx &= \lambda Sx \\ \mathcal{R}_N x_N &= \lambda_N \mathcal{S}_N x_N. \end{aligned}$$

are such that the statements 2 and 3 of Proposition 5.3.26 and Remark 5.3.27 hold with $\sigma(\mathcal{R}_N, \mathcal{S}_N)$ *for* $\sigma(\mathcal{A}_N)$.

5.4 Summary

The aim of this chapter has been to build the foundation for the approximation of spectral value sets of infinite dimensional operator by spectral value sets of finite dimensional ones. Essentially, the chapter has dealt with the following approximation concepts: strong, normwise and discrete-regular convergence. We began by discussing elementary properties of these convergences and proving certain useful results. Thereafter, we have shown how these types of convergences can be achieved by means of discrete operators. Finally, we have discussed some implications of regular convergence for eigenvalue problems.

Chapter 6

Approximation of Spectral Value Sets

We have mentioned already that, while the results of Section 4.2 are theoretically satisfactory, they are useful in applications only if the values of the norm of the transfer function associated with (A, D, E) can be obtained. In effect, the existence of the transfer function on the resolvent set is guaranteed under our assumptions. But, since the transfer function is usually not explicitly known, we cannot apply the formula in Theorem 4.2.11 to determine the spectral value sets. In Chapter 3 we developed efficient algorithms for calculating spectral value sets in the matrix case. Further, in Chapter 5 we have discussed several abstract approximation results on convergence of discrete operators. How could we use all this machinery for the calculations of $\sigma(A, D, E; \rho)$ in the infinite-dimensional setting? The answer to this question is given in the following.

6.1 Preliminaries

Essentially, our aim in this section is to motivate the use of *discrete* operators and/or sequences for the calculation of spectral value sets of infinite dimensional systems.

6.1.1 Norm of discrete operators

Let us suppose that we are given a framework as in (5.11), (5.12), (5.13), (5.14):

$$\begin{array}{llll} W_N \in W, & \pi_N^W : W \to W_N, & \iota_N^W : W_N \to W, & N \in \mathbb{N}, \\ V_N \in V, & \pi_N^V : V \to V_N, & \iota_N^V : V_N \to V, & N \in \mathbb{N}, \end{array}$$

where W and V are given Banach spaces. Further, let

$$T_N := \iota_N^V \mathcal{T}_N \pi_N^W, \quad \mathcal{T}_N : W_N \to V_N, \quad N \in \mathbb{N},$$

be a sequence of discrete operators, see Definition 5.2.1. We recall that each $\mathcal{T}_N$, $N \in \mathbb{N}$, allows a matrix representation and, thus, its norm $\|\mathcal{T}_N\|_{\mathcal{L}(W_N,V_N)}$ can be computed with standard matrix methods. Further, one observes that if W_N and V_N are provided with the norms induced by W and V respectively, then the norm of T_N is given by

$$\|T_N\|_{\mathcal{L}(W,V)} = \|\mathcal{T}_N\|_{\mathcal{L}(W_N,V_N)}, \quad N \in \mathbb{N}.$$

This fact makes discrete operators remarkable in our context. Indeed, by Theorem 4.2.11, the sets $\sigma(A,D,E;\rho)$ can be studied by looking at the norm of the corresponding transfer function G. Since the norm of discrete operators can be computed in a straightforward manner, we would substitute the calculation of $\|G(s)\|_{\mathcal{L}(U,Y)}$ by the "easy" calculation of $\|G_N(s)\|_{\mathcal{L}(U,Y)}$, where $(G_N(s))_{N\in\mathbb{N}}$ is, for each s of interest, a sequence of discrete operators which approximates *in some sense* $G(s)$.

6.1.2 The role of normwise convergence

In view of Corollary 4.2.12 and Equation (4.15), we should examine whether the upper level sets of $s \mapsto \|G_N(s)\|_{\mathcal{L}(U,Y)}$ approximate (in the sense of the Hausdorff metric) those of $s \mapsto \|G(s)\|_{\mathcal{L}(U,Y)}$.

It is easily seen that strong convergence of the transfer functions is not a condition which could, in general, guarantee this. In fact, from the Banach Steinhaus Theorem it follows that if

$$T_N \xrightarrow[V]{s} T,$$

then

$$\liminf_{N\to\infty} \|T_N\|_{\mathcal{L}(W,V)} \geq \|T\|_{\mathcal{L}(W,V)}.$$

Thus, if the approximations are such that

$$G_N(s) \xrightarrow[Y]{s} G(s), \quad \forall\, s \in \rho(A),$$

then the upper level sets of $\|G_N(\cdot)\|_{\mathcal{L}(U,Y)}$ *may* converge to sets which are just *upper bounds* of $\sigma(A,D,E;\rho)$.

On the other hand, Corollary A.0.15, Appendix A, shows that, given a locally connected compact set $K \subset \mathbb{C}$, in order to approximate upper level sets of $s \mapsto \|G(s)\|$ associated to regular values of $\|G\|$ (Definition A.0.13) and contained in K, it suffices to have uniform convergence in K:

$$\lim_{N\to\infty} \max_{s\in K} \big|\, \|G_N(s)\|_{\mathcal{L}(U,Y)} - \|G(s)\|_{\mathcal{L}(U,Y)} \big| = 0. \tag{6.1}$$

Since

$$\big|\|G_N(s)\|_{\mathcal{L}(U,Y)} - \|G(s)\|_{\mathcal{L}(U,Y)}\big| \leq \|G_N(s) - G(s)\|_{\mathcal{L}(U,Y)},$$

one has that

$$\lim_{N\to\infty} \|G_N(s) - G(s)\|_{\mathcal{L}(U,Y)} = 0 \text{ uniformly in } s \in K$$

implies (6.1). This leads us to the consideration of the following problem.

Problem 6.1.1 *Find conditions guaranteeing that a sequence of discrete operators* $(G_N(\cdot))_{N\in\mathbb{N}}$ *be such that*

$$\lim_{N\to\infty} \|G_N(s) - G(s)\|_{\mathcal{L}(U,Y)} = 0 \tag{6.2}$$

uniformly *on compacts subsets* $K \subset \rho(A)$ *.*

Remark 6.1.2 Note that the solution of Problem 6.1.1 gives only *sufficient* conditions for the desired convergence. As an illustration we mention the following fact. Consider Toeplitz operators

$$T = T(a) : X \to X, \quad X = l^2(\mathbb{N}, \mathbb{C})$$

with piecewise continuous symbol

$$a : \Gamma \to \mathbb{C}, \quad \Gamma = \{s \in \mathbb{C};\ |s| = 1\},$$

the associated infinite matrix and its finite dimensional truncations:

$$T = (t_{i-j})_{i,j \in \mathbb{N}}, \quad \mathcal{T}_N = (t_{i-j})_{i,j \in \underline{N}}, \quad \underline{N} = \{1, \ldots, N\}.$$

In this case, one has only strong convergence of the corresponding discrete operators T_N to T. However, for any $\rho > 0$, the *unstructured spectral value sets* $\sigma(\mathcal{T}_N, I_N, I_N; \rho)$ converge towards the corresponding unstructured spectral value set $\sigma(T, I, I; \rho)$. On the other hand, one should say that this convergence is in some sense "anomalous": the *spectra* of the $\mathcal{T}_N$ do *not*, in general, approximate the spectrum of T, see [5].

6.1.3 Compactness of the transfer function

Our is aim is to find conditions leading to the solution of Problem 6.1.1, and in particular, such that (6.2) holds. In this case we must be aware of the following fact [53, Theorem III.4.7].

Lemma 6.1.3 *$\mathcal{K}(W, V)$ is closed in $\mathcal{L}(W, V)$ with respect to the norm topology.*

This lemma has the following implications. Since for each $s \in \rho(A)$ the transfer functions $G_N(s)$ are discrete operators, and hence compact, in order to have convergence in norm to the transfer function $G(s)$, a necessary condition is the compactness of $G(s)$ *itself*. Thus, we can solve Problem 6.1.1 only under conditions which guarantee the compactness of $G(s)$ for every $s \in \rho(A)$.

Lemma 6.1.4 *Suppose that $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$ and that (4.2) holds, i.e.,*

$$\mathcal{D}(A) \subset \underline{X} \subset X \text{ with continuous dense injections.}$$

Let us assume that the embedding

$$\iota : \mathcal{D}(A) \to \underline{X} \text{ is compact,} \tag{6.3}$$

then $G(s) \in \mathcal{K}(U, Y)$ for each $s \in \rho(A)$.

Proof: First, note that the product of a compact operator with a bounded one is again a compact operator [53, Theorem III.4.8]. Thus, the proof follows from $G(s) = E\, \iota\, R(s) D$, i.e.,

$$U \underset{D}{\to} X \underset{R(s)}{\to} \mathcal{D}(A) \underset{\iota}{\to} \underline{X} \underset{E}{\to} Y,$$

$\iota \in \mathcal{K}(\mathcal{D}(A), \underline{X})$ and the boundedness of the rest of the operators. □

As we shall now see, assuming (6.3) has important implications for the resolvent $R(s)$ and the spectrum $\sigma(A)$ of A.

Lemma 6.1.5 *The compactness of*

$$\iota : \mathcal{D}(A) \to \underline{X}$$

is a necessary and sufficient condition for

$$R(s) \in \mathcal{K}(X, \underline{X}), \quad \forall\, s \in \rho(A). \tag{6.4}$$

Moreover, $\iota \in \mathcal{K}(\mathcal{D}(A), \underline{X})$ *implies that the spectrum* $\sigma(A)$ *is a set of isolated eigenvalues with finite algebraic multiplicities.*

Proof: First, we note the following. For any $s, s_0 \in \rho(A)$, we have the identity [53, I.(5.5), III.6.1]

$$R(s) = R(s_0)[I_X - (s - s_0)R(s)].$$

$R(s)$ is in $\mathcal{L}(X)$. Thus, this equality shows that if for some $s_0 \in \rho(A)$, $R(s_0) \in \mathcal{K}(X, \underline{X})$, then $R(s) \in \mathcal{K}(X, \underline{X})$ for every $s \in \rho(A)$.
Suppose $R(s_0) \in \mathcal{K}(X, \underline{X})$ for some $s_0 \in \rho(A)$. Then, from

$$(s_0 I - A) \in \mathcal{L}(\mathcal{D}(A), X) \text{ and } \iota := R(s_0)(s_0 I - A),$$

it follows that

$$\iota \in \mathcal{K}(\mathcal{D}(A), \underline{X}).$$

Conversely, suppose that $\iota \in \mathcal{K}(\mathcal{D}(A), \underline{X})$ and consider some $s_0 \in \rho(A)$. Then

$$R(s_0) \in \mathcal{L}(X, \mathcal{D}(A))$$

and we conclude, by [53, Theorem III.4.8], that

$$\iota \circ R(s_0) \in \mathcal{K}(X, \underline{X}) \implies R(s_0) \in \mathcal{K}(X, \underline{X}).$$

Now, proving the assertion made on $\sigma(A)$ is simple. Consider the injection

$$\iota_{\underline{X}} : \underline{X} \to X.$$

The operator $\iota_{\underline{X}}$ is continuous by (4.2) and thus bounded. Further, we have just shown that $R(s) \in \mathcal{K}(X, \underline{X})$ for all $s \in \rho(A)$. Thus, $\iota_{\underline{X}} \circ R(s) \in \mathcal{K}(X)$. We may now obtain the mentioned properties of $\sigma(A)$ from [53, III.6.8 and Theorem III.6.29]. □

Remark 6.1.6 Condition 6.3 is often satisfied in applications. For example, the following embeddings are compact.

1. $\iota : C^{k+1}(\bar{\Omega}; \mathbb{C}^n) \to C^k(\bar{\Omega}; \mathbb{C}^n)$, $k \in \mathbb{N}_0$ [20, Theorem V.1.1].
2. $\iota : H^1(\Omega; \mathbb{C}^n) \to L^2(\Omega; \mathbb{C}^n)$ [20, Theorem V.4.17].
3. $\iota : H^k(\Omega; \mathbb{C}^n) \to H^{k_1}(\Omega; \mathbb{C}^n)$ for $k_1 < k$; k, $k_1 \in \mathbb{N}_0$ [20, Theorem V.4.18].

The notations used above were introduced in Section 5.1.1.

6.2 On the approximation of transfer functions

We go now towards the statement of the most important approximation result. As before, let us assume that $\underline{X}$, X, U, Y are complex Banach spaces, $A \in \mathcal{C}(X)$ with domain $\mathcal{D}(A)$ and $\mathcal{D}(A) \subset \underline{X} \subset X$ with continuous dense injections, while $D \in \mathcal{L}(U, X)$ and $E \in \mathcal{L}(\underline{X}, Y)$.
The following notations are going to be useful:

1. $F := R(s_0)D \in \mathcal{L}(U, \underline{X})$ for some fixed $s_0 \in \rho(A)$.

2. $H(s) := ER(s)|_{\underline{X}} \in \mathcal{L}(\underline{X}, Y)$ for every $s \in \rho(A)$.

Note that using those notations and the resolvent identity [53, III.6.1, I.5.2]

$$R(s) = [I_X - (s - s_0)R(s)]R(s_0) \tag{6.5}$$

the transfer function $G(s) := ER(s, A)D$, for $s \in \rho(A)$, can be written as

$$G(s) = G(s_0) - (s - s_0)H(s)F = EF - (s - s_0)H(s)F, \quad s \in \rho(A).$$

Let $K \subset \rho(A)$ be a compact set. Essentially, we shall deal with three operator sequences:

$$\begin{aligned}&(F_N)_{N\in\mathbb{N}} \in \mathcal{L}(U, \underline{X}), \quad (E_N)_{N\in\mathbb{N}} \in \mathcal{L}(\underline{X}, Y),\\ &(H_N(s))_{N\in\mathbb{N}} \in \mathcal{L}(\underline{X}, Y), \quad s \in K.\end{aligned}$$

Note that under these conditions the "transfer functions"

$$G_N(s) = E_N F_N - (s - s_0)H_N(s)F_N, \quad s \in K, \quad N \in \mathbb{N},$$

are well defined in $\mathcal{L}(U, Y)$.

6.2.1 Main result

We are now ready for the formulation of the main result of this section. It gives sufficient conditions for the solution of Problem 6.1.1.

Theorem 6.2.1 *Suppose that $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$ and that (4.2) holds. Furthermore, let $K \subset \rho(A)$ be a given compact set and $s_0 \in K$ be fixed. Let*

$$F := R(s_0)D \text{ and } H(s) := ER(s)|_{\underline{X}}, \quad s \in K.$$

and consider the corresponding transfer function

$$G(s) := ER(s)D = EF - (s - s_0)H(s)F, \quad s \in \rho(A).$$

Further, let us suppose that we are given operator sequences

$$\begin{aligned}&(F_N)_{N\in\mathbb{N}} \in \mathcal{L}(U, \underline{X}), \quad (E_N)_{N\in\mathbb{N}} \in \mathcal{L}(\underline{X}, Y),\\ &(H_N(s))_{N\in\mathbb{N}} \in \mathcal{L}(\underline{X}, Y), \quad s \in K,\end{aligned}$$

which, under these conditions, define $\mathcal{L}(U, Y)$-valued functions

$$G_N(s) := E_N F_N - (s - s_0)H_N(s)F_N, \quad N \in \mathbb{N}, \quad s \in K.$$

Finally, let us suppose that

1. $F \in \mathcal{K}(U, \underline{X})$.
2. $F_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} F$.
3. $H_N(s) \xrightarrow[Y]{s} H(s), \forall\, s \in K$.
4. $E_N \xrightarrow[Y]{s} E$.

Then

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K. \tag{6.6}$$

Moreover, if, additionally, there exists $M > 0$ *independent of* N *such that*

$$\|H_N(s_1) - H_N(s_2)\|_{\mathcal{L}(\underline{X},Y)} < M|s_1 - s_2|, \quad \forall\, s_1,\, s_2 \in K, \tag{6.7}$$

then the convergence in (6.6) *is uniform in* $s \in K$.

Proof: We have

$$G_N(s_0) = E_N F_N \xrightarrow[\mathcal{L}(U,Y)]{n} EF = G(s_0), \tag{6.8}$$

$$H_N(s) F_N \xrightarrow[\mathcal{L}(U,Y)]{n} H(s)F, \quad \forall\, s \in K, \tag{6.9}$$

where both statements follow by a straightforward use of Lemma 5.1.6. Further, we see that

$$\|G(s) - G_N(s)\|_{\mathcal{L}(U,Y)} \leq \|G(s_0) - G_N(s_0)\|_{\mathcal{L}(U,Y)} + |s - s_0| \|H(s)F - H_N(s)F_N\|_{\mathcal{L}(U,Y)}, \quad \forall\, s \in K. \tag{6.10}$$

Since $|s - s_0|$ is uniformly bounded in K, (6.6) follows using (6.8), (6.9) and (6.10).
The last statement of the theorem, i.e. the uniform convergence, can be proved using Proposition 5.1.8. Indeed, it is easily seen that, for every $s_1,\, s_2 \in K$,

$$\begin{aligned} \|H(s_1) - H(s_2)\|_{\mathcal{L}(\underline{X},Y)} &\leq \|E\| \|R(s_1) - R(s_2)\|_{\mathcal{L}(\underline{X})} \\ &\leq |s_1 - s_2| \|E\| \|R(s_1)\|_{\mathcal{L}(\underline{X})} \|R(s_2)\|_{\mathcal{L}(\underline{X})} \\ &\leq M_1 |s_1 - s_2|, \end{aligned}$$

where

$$M_1 := \|E\| \max_{s \in K} \|R(s)\|^2_{\mathcal{L}(\underline{X})}.$$

Since we have assumed (6.7), we may apply Proposition 5.1.8 to the convergence (6.9) and obtain

$$H_N(s) F_N \xrightarrow[\mathcal{L}(U,Y)]{n} H(s)F, \quad \forall\, s \in K, \text{ uniformly in } s \in K \tag{6.11}$$

Now, using (6.11) and (6.10), we see that the convergence in (6.6) is also *uniform* in $s \in K$. $\square$

Remark 6.2.2 Using the resolvent identity (6.5), it can be easily shown that if $R(s_0)D \in \mathcal{K}(U, \underline{X})$ for some $s_0 \in \rho(A)$, then $R(s)D \in \mathcal{K}(U, \underline{X})$ for every $s \in \rho(A)$. This fact is important for applications, since it means that one has freedom in choosing the point $s_0 \in K$ for which the compactness of $R(s_0)D$ is to be proved.

Remark 6.2.3 The operator $D : U \to X$ is bounded. Thus, by Lemma 6.1.5, the compactness of the embedding $\iota : \mathcal{D}(A) \to \underline{X}$ is a sufficient condition for the compactness of $F := R(s)D : U \to \underline{X}$ for all $s \in \rho(A)$. We have seen in Remark 6.1.6 that the compactness of these embeddings is not rare in applications. Moreover, if the operator D is compact, this also implies that $F \in \mathcal{K}(U, \underline{X})$. We conclude that Condition 1 in our theorem is not as restrictive as one would think at first sight.

Remark 6.2.4 We have already seen (Proposition 5.2.3) that the norm of a bounded operator and of its adjoint coincide. In other words,

$$\|G(s) - G_N(s)\|_{\mathcal{L}(U,Y)} = \|G^*(s) - G_N^*(s)\|_{\mathcal{L}(Y^*,U^*)}.$$

That means that considering the adjoint operators can be advantageous in some applications. For example, if $E \in \mathcal{K}(\underline{X}, Y)$ whereas $D \notin \mathcal{K}(U, X)$. Of course, in that case the conditions of Theorem 6.2.1 must be formulated for the adjoints operators rather than for the original A, D and E. We omit the details.

6.2.2 Some corollaries

We shall consider below some useful corollaries of Theorem 6.2.1. In their statements we use the following notation: for a sequence $(A_N)_{N\in\mathbb{N}} \in \mathcal{C}(X)$ such that $\mathcal{D}(A) \subset \mathcal{D}(A_N)$ for $N \in \mathbb{N}$, we shall denote *strong convergence* of resolvents in a subset Ω of the resolvent set $\rho(A)$ with the symbol

$$R(s, A_N) \xrightarrow[X]{s} R(s, A), \quad \forall\, s \in \Omega, \tag{6.12}$$

i.e., for each $s \in \Omega$ we have

1. $s \in \rho(A_N)$ for N large enough (depending upon s) and
2. $\lim_{N\to\infty} \|R(s, A_N)x - R(s, A)x\|_{\mathcal{L}(X)} = 0$, for all $x \in X$.

Lemma 6.2.5 *Let K be a* compact *subset of $\rho(A)$. If*

$$R(s, A_N) \xrightarrow[X]{s} R(s, A), \quad \forall\, s \in K, \tag{6.13}$$

then, there exists a finite $N(K) \in \mathbb{N}$ such that

$$K \subset \rho(A_N), \quad \forall\, N > N(K).$$

Proof: Indeed, consider some $s_0 \in K$ and let N be large enough so that $s_0 \in \rho(A_N)$. Then

$$\{s \in \mathbb{C} \;:\; |s - s_0| < \|R(s_0, A_N)\|^{-1}_{\mathcal{L}(X)}\} \subset \rho(A_N).$$

The assertion follows immediately using the compacity of K and the fact that, due to (6.13), the sequence $(R(s, A_N))_{N\in\mathbb{N}}$ is *uniformly bounded* in N and s in the compact set K ([53, Theorem VIII.1.1]). □

Remark 6.2.6 In other words, if the set Ω in (6.12) is a compact subset of $\rho(A)$, then we may assume, without loss of generality, that $K \subset \rho(A_N)$ for all $N \in \mathbb{N}$. In the sequel we shall be making this assumption.

Corollary 6.2.7 *Suppose that $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$ and that (4.2) holds. Furthermore, let $K \subset \rho(A)$ be a given compact set and $s_0 \in K$ be fixed. Let*

$$F := R(s_0)D \text{ and } H(s) := ER(s)|_{\underline{X}}, \quad s \in K.$$

and consider the corresponding transfer function

$$G(s) := ER(s)D = EF - (s - s_0)H(s)F, \quad s \in \rho(A).$$

Further, let us suppose that we are given operator sequences

$$(A_N)_{N\in\mathbb{N}} \in \mathcal{C}(X), \quad \mathcal{D}(A) \subset \mathcal{D}(A_N), \quad N \in \mathbb{N},$$
$$(D_N)_{N\in\mathbb{N}} \in \mathcal{L}(U, X), \quad (E_N)_{N\in\mathbb{N}} \in \mathcal{L}(\underline{X}, Y)$$

which define $\mathcal{L}(U, Y)$-valued functions

$$G_N(s) := E_N R(s, A_N) D_N = E_N F_N - (s - s_0) H_N(s) F_N, \quad N \in \mathbb{N},$$

for $s \in K \cap \rho(A_N)$, where

$$F_N := R(s_0, A_N) D_N, \quad H_N(s) := E_N R(s, A_N)|_{\underline{X}}, \quad N \in \mathbb{N}.$$

Finally, let us suppose that

1. *$F \in \mathcal{K}(U, \underline{X})$.*
2. *$F_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} F$.*
3. *$R(s, A_N)|_{\underline{X}} \xrightarrow[\underline{X}]{s} R(s)|_{\underline{X}}$, $\forall\, s \in K$.*
4. *$E_N \xrightarrow[Y]{s} E$.*

Then

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall s \in K. \tag{6.14}$$

Moreover, this convergence is uniform in $s \in K$.

Proof: In order to use Theorem 6.2.1 only Condition 3 and Equation (6.7) remain to be proved. But this is easy. By Lemma 6.2.5, there exists $N(K) \in \mathbb{N}$ such that $K \subset \rho(A_N)$ for $N > N(K)$. Thus, the operators $H_N(s)$ are well defined in the whole K for $N > N(K)$. Further, using Lemma 5.1.4 we obtain that

$$H_N(s) = E_N R(s, A_N)|_{\underline{X}} \xrightarrow[Y]{s} ER(s)|_{\underline{X}} = H(s), \quad \forall\, s \in K,$$

i.e., Condition 3. For the proof of (6.7) we observe that

$$\begin{aligned}\|H_N(s_1) - H_N(s_2)\|_{\mathcal{L}(\underline{X},Y)} &= \|E_N[R(s_1, A_N) - R(s_2, A_N)]\|_{\mathcal{L}(\underline{X},Y)} \\ &\leq |s_1 - s_2| \|E_N\|_{\mathcal{L}(\underline{X},Y)} \|R(s_1, A_N)\|_{\mathcal{L}(\underline{X})} \|R(s_2, A_N)\|_{\mathcal{L}(\underline{X})},\end{aligned}$$

where again the resolvent identity (6.5) has been used. Further, since by assumption $E_N \xrightarrow[Y]{s} E$, the sequence $(E_N)_{N\in\mathbb{N}}$ is uniformly bounded. Thus, the proof is complete if we show that there exists $M > 0$ with the property

$$\|R(s, A_N)\|_{\mathcal{L}(\underline{X})} \leq M, \quad \forall\, s \in K, \quad \forall\, N > N(K).$$

But this follows from [53, Theorem VIII.1.1]. □

The following simple corollary of Theorem 6.2.7 is important for applications.

Corollary 6.2.8 *Suppose that the assumptions of Corollary 6.2.7 hold but, in place of Conditions 1, 2 and 3, we assume that*

1. $U = X$, $D := D_N := I_X$.
2. $R(s_0) \in \mathcal{K}(X, \underline{X})$.
3. $R(s_0, A_N) \xrightarrow[\mathcal{L}(X,\underline{X})]{n} R(s_0)$.

Then

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall s \in K.$$

Moreover, this convergence is uniform in $s \in K$.

Proof: The only point to prove is Condition 3 of Corollary 6.2.7. But this follows immediately from the fact that if $R(s_0, A_N) \xrightarrow[\mathcal{L}(X,\underline{X})]{n} R(s_0)$, for some $s_0 \in \rho(A)$, then $R(s, A_N) \xrightarrow[\mathcal{L}(X,\underline{X})]{n} R(s)$ for every $s \in \rho(A)$ [53, VIII.1.1]. □

6.3 On approximation of the resolvent of Riesz operators

Corollary 6.2.8 is specially interesting in the case of unstructured perturbations, i.e., $E = D = I_X$, $U = Y = \underline{X} = X$. However, it is readily seen that the condition

$$R(s, A_N) \xrightarrow[\mathcal{L}(X,X)]{n} R(s)$$

is a very strong requirement on the approximation scheme. For instance, one knows [53, Theorem IV.2.25]

Lemma 6.3.1 *Let T, $T_N \in \mathcal{C}(X)$, $N \in \mathbb{N}$, with nonempty resolvent and denote the gap between T and T_N [53, Chapter IV.2.2] by $\hat{\delta}(T, T_N)$. Then*

$$\lim_{N\to\infty} \hat{\delta}(T, T_N) \to 0$$

iff the following conditions hold

1. *Each $s \in \rho(T)$ belongs to $\rho(T_N)$ for N large enough.*
2. *$R(s, T_N) \underset{\mathcal{L}(X,X)}{\overset{n}{\longrightarrow}} R(s, T)$, while it is sufficient that this be true for some $s \in \rho(T)$.*

As regards our problem Lemma 6.3.1 can be interpreted as follows. The desired uniform approximation of the resolvent can be reached only under very restrictive conditions on the approximation method, namely, *convergence in gap*. How restrictive this condition can be, is illustrated by the following fact [53, Theorem IV.2.23.a].

Lemma 6.3.2 *If $T \in \mathcal{L}(X, Y)$, $\lim_{N\to\infty} \hat{\delta}(T, T_N) \to 0$ iff $T_N \in \mathcal{L}(X, Y)$ for N large enough and $T_N \underset{\mathcal{L}(X,Y)}{\overset{n}{\longrightarrow}} T$.*

In spite of this, we shall show in the sequel that for an important class of infinite-dimensional operators, namely *Riesz* operators, the "unstructured" version of Problem 6.1.1 can be solved satisfactorily.

6.3.1 Preliminaries

We begin with some definitions.

Definition 6.3.3 A sequence of vectors $(\phi_i)_{i\in\mathbb{N}}$ in a Hilbert space X is called a *Riesz basis* in X if the following two conditions hold:

(a) $\mathrm{cl}\,(\mathrm{span}_{i\geq 1}\{\phi_i\}) = X$.

(b) There exist positive constants m and M such that for arbitrary $N \in \mathbb{N}$ and arbitrary scalars α_i, $i = 1, \ldots, N$, the following inequalities hold

$$m^2 \sum_{i=1}^{N} |\alpha_i|^2 \leq \| \sum_{i=1}^{N} \alpha_i \phi_i \|^2 \leq M^2 \sum_{i=1}^{N} |\alpha_i|^2.$$

It can be easily shown [15, Lemma 2.3.2.b] that in this case each element $x \in X$ has a unique representation

$$x = \sum_{i=1}^{\infty} \langle\, x \,,\, \psi_i \,\rangle\, \phi_i, \tag{6.15}$$

where $(\psi_i)_{i\in\mathbb{N}}$ is the *biorthogonal sequence* corresponding to $(\phi_i)_{i\in\mathbb{N}}$, i.e.,

$$\langle\, \phi_i \,,\, \psi_j \,\rangle = \delta_{ij}. \tag{6.16}$$

See Remark 5.3.27. Moreover, it holds that [15, Lemma 2.3.2.b]

$$m^2 \sum_{i=1}^{\infty} |\langle x\,,\,\psi_i\rangle|^2 \leq \|x\|^2 \leq M^2 \sum_{i=1}^{\infty} |\langle x\,,\,\psi_i\rangle|^2. \tag{6.17}$$

The following lemma gives a useful characterisation of a Riesz basis [15, Exercise 2.21.a].

Lemma 6.3.4 *A sequence of vectors $(\phi_i)_{i\in\mathbb{N}}$ in a Hilbert space X forms a Riesz basis in X iff it is similar to an orthonormal basis $(e_i)_{i\in\mathbb{N}}$ of X, i.e., there exists a bounded invertible operator $T \in \mathcal{L}(X)$ such that $\phi_i = Te_i$, $i = 1, \ldots, \infty$.*

We may define now the main objects of this section: Riesz operators.

Definition 6.3.5 A closed linear operator A in a Hilbert space X is called *Riesz operator* in X, to be denoted $A \in \mathcal{R}(X)$, if

1. $\sigma(A) = \{\lambda_i,\, i = 1, 2, \ldots\}$, where each λ_i is a *simple* eigenvalue of A. The sequence $(\lambda_i)_{i\in\mathbb{N}}$ has no accumulation points in $\mathbb{C}$.

2. The eigenvectors $(\phi_i)_{i\in\mathbb{N}}$ form a Riesz basis of X.

Remark 6.3.6 Note that this definition differs slightly from Definition 2.3.4 in [15]. Here we do *not* allow the existence in $\mathbb{C}$ of accumulation points of the spectrum.

A Riesz operator A has the following properties [Theorem 2.3.5][15].

1. $\mathcal{D}(A) = \{x \in X; \quad \sum_{i=1}^{\infty} |\lambda_i|^2 |\langle x\,,\,\psi_i\rangle|^2 < \infty\}$.

2. $Ax = \sum_{i=1}^{\infty} \lambda_i \langle x\,,\,\psi_i\rangle\, \phi_i$, $x \in \mathcal{D}(A)$, $(\psi_i)_{i\in\mathbb{N}}$ given as in (6.16).

3. $R(s) : X \to X$ is compact and is given by

$$R(s)f = (sI_X - A)^{-1}f = \sum_{i=1}^{\infty} \frac{1}{s-\lambda_i} \langle f\,,\,\psi_i\rangle\, \phi_i, \quad f \in X. \tag{6.18}$$

As for Riesz basis, it is easily proved that [15, Exercise 2.21.c]

Lemma 6.3.7 *An operator $A \in \mathcal{R}(X)$ iff it is similar to a closed operator Q whose eigenvectors form an orthogonal basis for X, i.e., there exists a bounded invertible operator $T \in \mathcal{L}(X)$, $T^{-1} : \mathcal{D}(A) \to \mathcal{D}(Q)$, such that $A = TQT^{-1}$.*

6.3.2 Approximation of the resolvent operator

Now we shall show how results on spectral approximation of linear operators can be useful for the *uniform* approximation of the resolvent of Riesz operators on compact sets $K \subset \rho(A)$.
For the sake of clarity in the exposition we shall numerate the eigenvalues of A (remember that A is Riesz and hence $\sigma(A)$ is a numerable set of eigenvalues). Suppose that some "initial" compact set $K \subset \mathbb{C}$ is given such that $\sigma(A) \cap K$ is not empty. Since $A \in \mathcal{R}(X)$, we know that

$$\sigma(A) \cap K = \{\lambda_i \in \sigma(A);\ i = 1, \ldots, L_K\},\ \text{where } L_K < \infty. \tag{6.19}$$

We begin the numeration of $\sigma(A)$ with the eigenvalues of $\sigma(A) \cap K$. Furthermore, we consider increasing ϵ-neighborhoods of K and numerate the rest of the eigenvalues consecutively as they "begin" to lie in these neighborhoods. This numerating system, although clearly imperfect, is enough for our purposes.

Lemma 6.3.8 *Suppose that $A \in \mathcal{R}(X)$ and let $K \subset \rho(A)$ be a compact set. Denote by $R_L(s)$ the operator given by*

$$R_L(s)x := R(s)P_L = P_L R(s), \quad x \in X, \tag{6.20}$$

where P_L denotes, for every $L \in \mathbb{N}$, the spectral projection associated to the first L eigenvalues of A (see Equation (5.48)). Then

$$R_L(s) \xrightarrow[\mathcal{L}(X,X)]{n} R(s), \quad \forall\, s \in K,$$

with uniform convergence in $s \in K$.

Proof: Since $A \in \mathcal{R}(X)$, we know that

$$\iota_L P_L \xrightarrow[X]{s} I_X,$$

where ι_L is the corresponding embedding. In view of this, the statement follows easily from $R(s) \in \mathcal{K}(X)$ and Corollary 6.2.8 with "$R_N(s) \equiv R(s)$" and "$(P_L)_{L\in\mathbb{N}}$" playing the role of "$(E_N)_{N\in\mathbb{N}}$". □

Remark 6.3.9 Note that

$$P_L R(s) f = \sum_{i=1}^{L} \frac{1}{s-\lambda_i} \langle\, f\,,\, \psi_i\,\rangle\, \phi_i, \quad f \in X.$$

Lemma 6.3.8 is a nice result because it shows a method leading to convergence in norm of the resolvent operators. In particular, it shows that truncations of the series (6.18) are the "canonical" candidates for approximations of $R(s)$.
In applications, however, the eigenfunctions and eigenvalues of A, or even its resolvent $R(s, A)$, are seldom known. Thus, Lemma 6.3.8 is satisfactory only from the theoretical point of view. We need a more "realistic" approach; for example, the next proposition.

Proposition 6.3.10 *Suppose that $A \in \mathcal{R}(X)$ and that an operator sequence $(A_N)_{N\in\mathbb{N}} \in \mathcal{C}(X)$ is given with the following property: If $\Omega_0 \subset \mathbb{C}$ is a bounded domain such that*

$$\sigma(A) \cap \Omega_0 = \{\lambda_1, \ldots, \lambda_m\}$$

then,

1. *For N large enough,*

$$\sigma(A_N) \cap \Omega_0 = \{\lambda_1^{(N)}, \ldots, \lambda_m^{(N)}\}. \tag{6.21}$$

2. Each of the eigenvalues $\lambda_i^{(N)}$, $i = 1, \dots, m$, is simple and

$$\lim_{N\to\infty} |\lambda_i^{(N)} - \lambda_i| = \lim_{N\to\infty} \|\phi_i^{(N)} - \phi_i\|_X = \lim_{N\to\infty} \|\psi_i^{(N)} - \psi_i\|_X = 0, \; i = 1, \dots, m. \tag{6.22}$$

where, for each $i = 1, \dots, m$, ϕ_i and $\phi_i^{(N)}$ denote eigenvectors corresponding to λ_i and $\lambda_i^{(N)}$, respectively, while ψ_i and $\psi_i^{(N)}$ are biorthogonal with respect to ϕ_i and $\phi_i^{(N)}$, see (6.16).

Let K be a compact subset of $\rho(A)$ and denote by $N(K)$ a number for which

$$K \subset \rho(A_N), \quad \forall\, N > N(K).$$

($N(K)$ is finite due to (6.21), (6.22)).
Then, for every $\epsilon > 0$, one can choose *a bounded domain $\Omega \subset \mathbb{C}$ (depending upon ϵ) with the following properties: its boundary $\partial\Omega$ is a rectifiable closed simple curve, $K \subset \Omega$, and there exists $\hat{N}(K) > N(K)$ such that*

$$\|R(s, A_N)\, \iota_{N\Omega} P_{N\Omega} - R(s)\|_{\mathcal{L}(X)} < \epsilon, \quad \forall\, s \in K, \quad N > \hat{N}(K). \tag{6.23}$$

Here $P_{N\Omega}$ denotes the eigenprojection (see (5.49))

$$P_{N\Omega} = \frac{1}{2i\pi} \int_{\partial\Omega} R(s, A_N) ds.$$

and $\iota_{N\Omega}$ the embedding of the corresponding sum of eigenspaces into X.

Proof: Indeed, by Lemma 6.3.8, there exists a number $L(K) \in \mathbb{N}$ such that

$$\|R(s) - R_{L(K)}(s)\|_{\mathcal{L}(X)} < \epsilon/2, \quad \forall\, s \in K,$$

where $R_L(s)$ is given by (6.20). Let $\Omega \subset \mathbb{C}$ be a bounded domain such that $\partial\Omega$ is a rectifiable closed simple curve, $K \subset \Omega$, and

$$\sigma(A) \cap \Omega = \{\lambda_1, \dots, \lambda_{L(K)}\}.$$

See Figure 6.1. For any $s \in \Omega$ and $N > N(K)$ we have then

$$\begin{aligned} \|R(s) - R(s, A_N)\, \iota_{N\Omega} P_{N\Omega}\|_{\mathcal{L}(X)} &\le \|R(s) - R_{L(K)}(s)\|_{\mathcal{L}(X)} + \|R_{L(K)}(s) - R(s, A_N)\, \iota_{N\Omega} P_{N\Omega}\|_{\mathcal{L}(X)} \\ &\le \epsilon/2 + \|R(s)\, \iota_{\Omega} P_{\Omega} - R(s, A_N)\, \iota_{N\Omega} P_{N\Omega}\|_{\mathcal{L}(X)}, \end{aligned}$$

where $P_\Omega := P_{L(K)}$ and $\iota_\Omega := \iota_{L(K)}$, respectively. The proof is done if we are able to prove that for N large enough

$$\|R(s)\, \iota_{\Omega} P_{\Omega} - R(s, A_N)\, \iota_{N\Omega} P_{N\Omega}\|_{\mathcal{L}(X)} < \epsilon/2, \quad \forall\, s \in K. \tag{6.24}$$

By [53, Equation III.6.35], and since the eigenvalues λ_i, $\lambda_i^{(N)}$, $i = 1, \dots, L(K)$, are all simple, one sees, increasing N if necessary, that

$$\begin{aligned} R(s)\, \iota_{\Omega} P_{\Omega} &:= \sum_{i=1}^{L(K)} \frac{1}{s - \lambda_i} P_\Omega^i, \\ R(s, A_N)\, \iota_{N\Omega} P_{N\Omega} &:= \sum_{i=1}^{L(K)} \frac{1}{s - \lambda_i^{(N)}} P_{N\Omega}^i, \end{aligned}$$

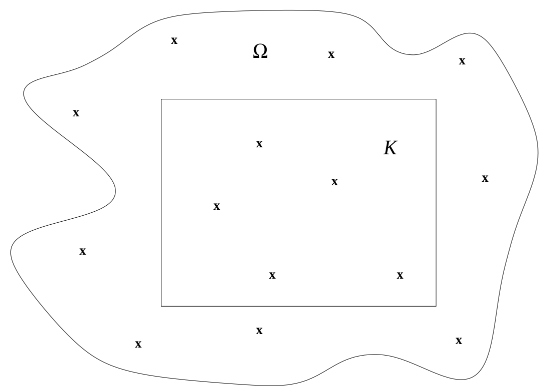

Figure 6.1: Sets K and Ω. The crosses represent $\sigma(A) \cap \Omega$.

where

$$P_\Omega^i := \langle \cdot \,,\, \psi_i \rangle\, \phi_i, \quad P_{N\Omega}^i := \langle \cdot \,,\, \psi_i^{(N)} \rangle\, \phi_i^{(N)}, \quad i = 1, \ldots, L(K),$$

and λ, ϕ_i, ψ_i and $\lambda_i^{(N)}$, $\phi_i^{(N)}$, $\psi_i^{(N)}$, $i = 1, \ldots, L(K)$, are the eigenelements of A and A_N, respectively, corresponding to Ω. The eigenprojections P_Ω^i, $P_{N\Omega}^i$ have been explicitly written down using the corresponding scalar product.
Then with these notations we have

$$\|(R(s)\,\iota_{\,\Omega} P_\Omega - R(s, A_N)\,\iota_{\,N\Omega} P_{N\Omega}) \cdot \|_{\mathcal{L}(X)} \leq \sum_{i=1}^{L(K)} \| \frac{1}{s - \lambda_i} P_\Omega^i - \frac{1}{s - \lambda_i^{(N)}} P_{N\Omega}^i \|_{\mathcal{L}(X)}, \tag{6.25}$$

for all $N > N(K)$ and every $s \in K$. Furthermore, using (6.22), we can prove, for each $i = 1, \ldots, L(K)$, the convergences

$$P_{N\Omega}^i \xrightarrow[\mathcal{L}(X,X)]{n} P_\Omega^i,$$

$$\frac{1}{s - \lambda_i^{(N)}} \xrightarrow[X]{s} \frac{1}{s - \lambda_i}, \quad \forall\, s \in K,$$

and, by Proposition 5.1.8, we obtain

$$\frac{1}{s - \lambda_i^{(N)}} P_{N\Omega}^i \xrightarrow[\mathcal{L}(X,X)]{n} \frac{1}{s - \lambda_i} P_\Omega^i, \quad \forall\, s \in K, \quad i = 1, \ldots, L(K), \tag{6.26}$$

uniformly in $s \in K$. A straightforward use of (6.26) and (6.25) ends the proof. □

Remark 6.3.11 The resolvent $R(s, A)$ of a Riesz operator is compact. Thus, the embedding $\iota : \mathcal{D}(A) \to X$ is also compact [12, Proposition 2.37] and we may find in Proposition 5.3.26 and Remark 5.3.27 sufficient conditions to be imposed on a discrete sequence $(A_N)_{N \in \mathbb{N}}$ in order to satisfy the assumptions of Proposition 6.3.10.

Remark 6.3.12 Essentially, what we have done in Proposition 6.3.10 is to approximate in norm the compact transfer function

$$G(s) := R(s, A) P_\Omega, \quad s \in \rho(A).$$

Doing this makes sense because, since we have chosen Ω as in Lemma 6.3.8, the functions $G(\cdot)$ and $R(\cdot, A)$ are very close on K.

Remark 6.3.13 In applications the determination of Ω may be a complicated and costly "trial-and-error" process.

6.4 Projection Schemes

The aim of this section is to develop approximation schemes able to generate sequences of discrete operators which satisfy the conditions of Theorem 6.2.1 and Corollary 6.2.7. As main tool we shall use the results of Chapter 5.
Indeed, we see that Proposition 5.2.3 gives sufficient conditions on the projections for

$$F_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} F$$

as required in Theorem 6.2.1 and Corollary 6.2.7. Further, Proposition 5.2.5 addresses the strong convergence

$$E_N \xrightarrow[Y]{s} E.$$

On the other hand, the third condition in Theorem 6.2.1 and Corollary 6.2.7, respectively, i.e.,

$$H_N(s) \xrightarrow[Y]{s} H(s), \quad \forall\, s \in K,$$
$$R(s, A_N)|_{\underline{X}} \xrightarrow[\underline{X}]{s} R(s)|_{\underline{X}}, \quad \forall\, s \in K,$$

is more difficult to handle. The case where $R(s)$ and/or $H(s)$ can be *explicitly* obtained is perhaps the only one which is easy to solve. Indeed, if these operators are known, with the help of suitable sequences of projections and Proposition 5.2.5, we may generate discrete sequences with the desired properties. Unfortunately, this is rarely the case. Thus, in order to develop a general approach, we use the results of Chapter 5 related to *discrete-regular convergence*.

6.4.1 Preliminaries

We begin with some words about notation. Here we shall deal with discrete operators and their corresponding finite dimensional versions. So, in order to avoid a "notational overflow", we shall denote them with the symbols: A_N and $\mathcal{A}_N$, D_N and $\mathcal{D}_N$, E_N and $\mathcal{E}_N$.
We recall that our underlying framework is

$$A \in \mathcal{C}(X), \quad D \in \mathcal{L}(U, X), \quad E \in \mathcal{L}(\underline{X}, Y) \text{ and } (4.2).$$

The problem to be solved is: given a compact set $K \subset \rho(A)$, to construct a sequence of discrete operators ("transfer functions") $(G_N(s))_{N\in\mathbb{N}}$ with the property

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K,$$

uniformly in $s \in K$. Moreover, since we plan to use Theorem 6.2.1, we shall assume that

$$F := R(s_0)D \in \mathcal{K}(U, \underline{X}) \text{ for some } s_0 \in K,$$

where by Remark 6.2.2, the point $s_0 \in \rho(A)$ is any point in K. Sometimes, we shall also make use of the $\mathcal{L}(\underline{X}, Y)$-valued function

$$H(s) := ER(s)|_{\underline{X}}, \quad s \in \rho(A).$$

Finally we also introduce the linear operator $\hat{A}$ defined by

$$\hat{A} : \mathcal{D}(\hat{A}) \to \underline{X}, \quad \hat{A}x := Ax, \quad \forall\, x \in \mathcal{D}(\hat{A}), \tag{6.27}$$

where $\mathcal{D}(\hat{A})$ is defined by

$$\mathcal{D}(\hat{A}) := \{x \in \mathcal{D}(A) \text{ such that } Ax \in \underline{X}\}. \tag{6.28}$$

Lemma 6.4.1 *The operator $\hat{A}$ is closed in $\underline{X}$: $\hat{A} \in \mathcal{C}(\underline{X})$. Furthermore, $\rho(\hat{A}) = \rho(A)$.*

Proof: We must show that the graph $\Gamma(\hat{A})$ [53, III.5.2] of $\hat{A}$ is closed in $\underline{X} \times \underline{X}$. Indeed, let us suppose that $(x_N)_{N \in \mathbb{N}}$ is a sequence in $\mathcal{D}(\hat{A})$ such that

$$\lim_{N \to \infty} (x_N, \hat{A}x_N) = (x, y) \in \underline{X} \times \underline{X}.$$

We must prove that $(x, y) \in \Gamma(\hat{A})$ (i.e., $x \in \mathcal{D}(\hat{A})$, $y = \hat{A}x$). The pairs $(x_N, \hat{A}x_N)_{N \in \mathbb{N}}$ are elements of $\Gamma(A)$ for every $N \in \mathbb{N}$, they form a convergent sequence and $A \in \mathcal{C}(X)$. Thus, $(x, y) \in \Gamma(A)$. It follows that $x \in \mathcal{D}(A)$ and $y = Ax$. Using the definition of $\mathcal{D}(\hat{A})$ and $y \in \underline{X}$, see (6.28), we conclude that $x \in \mathcal{D}(\hat{A}) \subset \underline{X}$.
The second statement can be proved as follows:

1. $\rho(A) \subset \rho(\hat{A})$: Let $s \in \rho(A)$. Then, see (6.27), (6.28), the operator

   $$R'(s, \hat{A}) : \underline{X} \to \underline{X}, \quad x \mapsto (sI_{\underline{X}} - \hat{A})^{-1}x$$

   is well defined for every $x \in \underline{X}$. Moreover, since $sI_{\underline{X}} - \hat{A}$ is closed in $\underline{X}$, $R'(s, \hat{A}) \in \mathcal{C}(\underline{X})$ as well. It follows, by the closed graph theorem [53, Theorem III.5.20], that $R'(s, \hat{A})$ is in $\mathcal{L}(\underline{X})$. Thus, $R(s, \hat{A}) = R'(s, \hat{A})$ and $s \in \rho(\hat{A})$.

2. $\rho(\hat{A}) \subset \rho(A)$: Let $s \in \rho(\hat{A})$ and ι denote the (continuous) embedding of $\underline{X}$ into X, see (4.2). Further, consider the set $X' := \iota\,\underline{X} \subset X$. Then the operator

   $$R'(s, A) : X' \to X, \quad x \mapsto \iota\, R(s, \hat{A})x,$$

   is bounded. Moreover, by [53, Extension Principle III.2.2], $R'(s, A)$ can be extended to cl $X' = X$ with conservation of the norm. This extension is clearly $R(s, A)$ and it follows that $R(s, A) \in \mathcal{L}(X)$ and $s \in \rho(A)$.

□

Finally, we recall the resolvent identity (6.5)

$$R(s) = [I_X - (s - s_0)R(s)]\, R(s_0), \quad s,\, s_0 \in \rho(A),$$

which was frequently used in the proof of Theorem 6.2.1.

6.4.2 Projections and discrete operators

Our immediate goal here is to introduce the sequences of subspaces, projections and discrete operators which will be used through this section.
Let us suppose that $\mathcal{D}(\hat{A})$, $\underline{X}$, Y, U are *separable Banach spaces* and consider sequences of finite dimensional subspaces

$$(U_N)_{N\in\mathbb{N}}, \quad (Z_N)_{N\in\mathbb{N}}, \quad (X_N)_{N\in\mathbb{N}}, \quad (Y_N)_{N\in\mathbb{N}}$$

of

$$U, \quad \mathcal{D}(\hat{A}), \quad \underline{X}, \quad Y,$$

respectively. We assume that $\dim \underline{X} = \infty$ and that $\dim Z_N$ and $\dim X_N$ grow as $N \to \infty$; we do not exclude, however, the possibility that U and/or Y are finite dimensional, in which case the sequences of subspaces $(U_N)_{N\in\mathbb{N}}$ and/or $(Y_N)_{N\in\mathbb{N}}$ are taken stationary and equal to U and Y, respectively. The "canonical" choice for $(X_N)_{N\in\mathbb{N}}$ will be $X_N = AZ_N$, but other choices are also possible.
Usually the subspaces U_N, Z_N, X_N and Y_N will be defined via sequences of projections:

$$\begin{aligned} \pi_N^Z &: \mathcal{D}(\hat{A}) \to Z_N, & \pi_N^X &: \underline{X} \to X_N \\ \pi_N^U &: U \to U_N, & \pi_N^Y &: Y \to Y_N. \end{aligned} \tag{6.29}$$

We shall also need the natural embeddings of the finite dimensional subspaces into the corresponding spaces, i.e.,

$$\begin{aligned} \iota_N^Z &: Z_N \to \mathcal{D}(\hat{A}), & \iota_N^X &: X_N \to \underline{X}, \\ \iota_N^U &: U_N \to U, & \iota_N^Y &: Y_N \to Y. \end{aligned} \tag{6.30}$$

Using all these projections and embeddings we may define sequences of discrete operators

$$(F_N)_{N\in\mathbb{N}}, \quad (P_N)_{N\in\mathbb{N}}, \quad (A_N)_{N\in\mathbb{N}}, \quad (H_N(s))_{N\in\mathbb{N}}, \quad (E_N)_{N\in\mathbb{N}}$$

by the formulas

$$\begin{array}{rcccrcll} F_N & : & U & \to & \underline{X}, & F_N u & := & \iota_N^X \mathcal{F}_N \pi_N^U u, & u \in U, \\ P_N & : & \mathcal{D}(\hat{A}) & \to & \underline{X}, & P_N z & := & \iota_N^X \mathcal{P}_N \pi_N^Z x, & z \in \mathcal{D}(\hat{A}), \\ A_N & : & \mathcal{D}(\hat{A}) & \to & \underline{X}, & A_N z & := & \iota_N^X \mathcal{A}_N \pi_N^Z z, & z \in \mathcal{D}(\hat{A}), \\ H_N(s) & : & \underline{X} & \to & Y, & H_N(s) x & := & \iota_N^Y \mathcal{H}_N(s) \pi_N^X x, & x \in \underline{X}, \\ E_N & : & \underline{X} & \to & Y, & E_N x & := & \iota_N^Y \mathcal{E}_N \pi_N^X x, & x \in \underline{X}, \end{array} \tag{6.31}$$

where the finite dimensional operators are just

$$\begin{array}{rcccrcll} \mathcal{F}_N & : & U_N & \to & X_N, & \mathcal{F}_N & := & \pi_N^X F \iota_N^U, & N \in \mathbb{N}, \\ \mathcal{P}_N & : & Z_N & \to & X_N, & \mathcal{P}_N & := & \pi_N^X \hat{\imath} \, \iota_N^Z, & N \in \mathbb{N}, \\ \mathcal{A}_N & : & Z_N & \to & X_N, & \mathcal{A}_N & := & \pi_N^X \hat{A} \, \iota_N^Z, & N \in \mathbb{N}, \\ \mathcal{H}_N(s) & : & X_N & \to & Y_N, & \mathcal{H}_N(s) & := & \pi_N^Y H(s) \, \iota_N^X, & N \in \mathbb{N}, \\ \mathcal{E}_N & : & X_N & \to & Y_N, & \mathcal{E}_N & := & \pi_N^Y E \, \iota_N^X, & N \in \mathbb{N}, \end{array} \tag{6.32}$$

In the formulas above we have used the continuous embedding

$$\hat{\iota} : \mathcal{D}(\hat{A}) \to \underline{X}. \tag{6.33}$$

6.4.3 Projection framework

Our plan now is to show how and under which conditions the discrete operators defined in (6.31) satisfy the conditions of Theorem 6.2.1 and/or Corollary 6.2.7.
In the following we often use the expression "e-computable operator" to mean that the corresponding matrix expressions of the operators (6.32) can be *explicitly* calculated. In general, at least the operators $\hat{A}$, D and E must be of this sort. However, $R(s)$ and $H(s)$ could fail to be e-computable. In such cases, one must find a way to construct the sequence $(H_N(s))_{N\in\mathbb{N}}$ with the information available. Since we shall not be able to treat all the possible cases, we shall study in detail only some of them.

Case 1: $H(\cdot)$ is e-computable.

This is a straightforward case: one only needs to impose some "natural" conditions on the projections.

Proposition 6.4.2 *Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U,X)$, $E \in \mathcal{L}(\underline{X},Y)$ and suppose that (4.2) holds. Further, let $K \subset \rho(A)$ be a given compact set and suppose that $R(s_0)D \in \mathcal{K}(U,\underline{X})$ for some $s_0 \in K$ and that both $F = R(s_0)D$ and $H(s) = ER(s)|_{\underline{X}}$, $s \in K$, are e-computable. Finally, let us assume that sequences of projections and embeddings are given such that*

$$\pi_N^{U^*}\iota_N^{U^*} \xrightarrow[U]{s} I_{U^*}, \quad \iota_N^X\pi_N^X \xrightarrow[\underline{X}]{s} I_{\underline{X}}, \quad \iota_N^Y\pi_N^Y \xrightarrow[Y]{s} I_Y. \tag{6.34}$$

Then, the sequences $(F_N)_{N\in\mathbb{N}}$, $(H_N(s))_{N\in\mathbb{N}}$, $s \in K$, and $(E_N)_{N\in\mathbb{N}}$ defined in (6.31) generate a discrete operator sequence

$$G_N(s) = E_NF_N - (s-s_0)H_N(s)F_N = \iota_N^Y\mathcal{G}_N(s)\pi_N^U, \quad s \in K, \quad N \in \mathbb{N},$$

where

$$\mathcal{G}_N(s)_N := \mathcal{E}_N\mathcal{F}_N - (s-s_0)\mathcal{H}_N(s)\mathcal{F}_N, \quad N \in \mathbb{N}, \tag{6.35}$$

such that

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K,$$

uniformly in $s \in K$.

Proof: We shall use Theorem 6.2.1. For this we must check its Conditions 2, 3, 4 and Equation (6.7). Indeed,

1. $F_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} F$: follows by Proposition 5.2.3.

2. $H_N(s) \xrightarrow[Y]{s} H(s)$, $\forall\, s \in K$: follows by Proposition 5.2.5.

3. $E_N \xrightarrow[Y]{s} E$: follows by Proposition 5.2.5 as well.

In order to prove Equation (6.7) we note that for any s_1, $s_2 \in K$ we have

$$\begin{aligned} H_N(s_1) - H_N(s_2) &= \iota_N^Y \pi_N^Y E[R(s_1)|_{\underline{X}} - R(s_2)|_{\underline{X}}]\, \iota_N^X \pi_N^X \\ &= -(s_1 - s_2)\, \iota_N^Y \pi_N^Y E R(s_1)|_{\underline{X}} R(s_2)|_{\underline{X}}\, \iota_N^X \pi_N^X. \end{aligned}$$

Since, by assumptions, both $(\iota_N^Y \pi_N^Y)_{N\in\mathbb{N}}$, $(\iota_N^X \pi_N^X)_{N\in\mathbb{N}}$ are strongly convergent sequences, they are uniformly bounded. Moreover, $E \in \mathcal{L}(\underline{X}, Y)$ and since $K \subset \rho(A) = \rho(\hat{A})$ is compact, there exists $M_0 < \infty$ such that $\|R(s)\|_{\mathcal{L}(\underline{X})} < M_0$ for all $s \in K$. Thus, if

$$M := \max\{\sup_{N\in N} \|\iota_N^Y \pi_N^Y\|_{\mathcal{L}(Y)},\ \sup_{N\in N} \|\iota_N^X \pi_N^X\|_{\mathcal{L}(\underline{X})},\ M_0,\ \|E\|_{\mathcal{L}(\underline{X},Y)}\}.$$

we obtain

$$\|H_N(s_1) - H_N(s_2)\|_{\mathcal{L}(\underline{X},Y)}\} \leq M^5 |s_1 - s_2|.$$

□

Figure 6.2 illustrates the interrelation between subspaces and operators. Note that the diagram is not commutative.

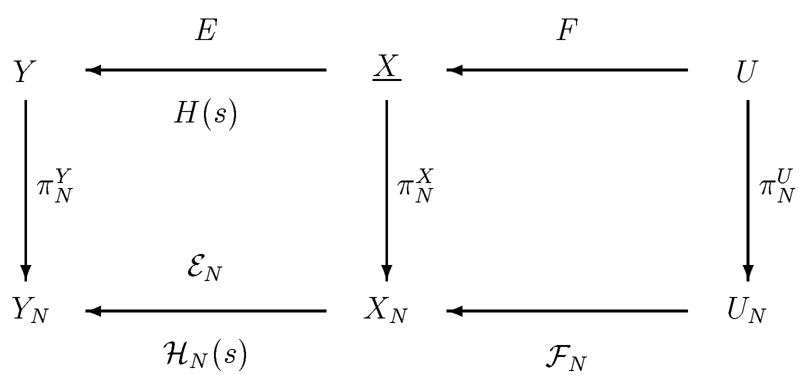

Figure 6.2: Interrelation between subspaces and operators when $H(\cdot)$ is e-computable.

Case 2: $H(\cdot)$ is not e-computable

This a more difficult and interesting case. The general approach is to consider the regular convergence

$$A_N \xrightarrow[\mathcal{L}(\mathcal{D}(\hat{A}),\underline{X})]{r} A$$

in $\mathcal{L}(\mathcal{D}(\hat{A}), \underline{X})$, where $(A_N)_{N\in\mathbb{N}}$ is defined as in (6.31). This makes sense because, since the operator $\hat{A}$ is closed in $\underline{X}$, see Lemma 6.4.1, we may provide its domain $\mathcal{D}(\hat{A})$ with the graph norm

$$\|x\|_{\mathcal{D}(\hat{A})} = \left(\|\hat{A}x\|_{\underline{X}}^2 + \|x\|_{\underline{X}}^2\right)^{1/2}, \quad x \in \mathcal{D}(\hat{A}),$$

and so $\mathcal{D}(\hat{A})$ becomes a Banach space. Note, however, that also other norms may meet the requirement of converting $\mathcal{D}(\hat{A})$ into a Banach space. The (continuous) embedding

$$\hat{\imath} : \mathcal{D}(\hat{A}) \to \underline{X}$$

and the sequence $(P_N)_{N\in\mathbb{N}}$ of discrete operators

$$P_N = \iota_N^X \mathcal{P}_N \pi_N^Z, \quad \mathcal{P}_N := \pi_N^X \, \hat{\imath} \, \iota_N^Z, \quad N \in \mathbb{N},$$

introduced already in (6.31), (6.32), will play an important role in our approach. At last, recall that, by Lemma 6.4.1, $\rho(A) = \rho(\hat{A})$.

Proposition 6.4.3 *Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U,X)$, $E \in \mathcal{L}(\underline{X},Y)$ and suppose that (4.2) holds. Further, let $K \subset \rho(A)$ be a given compact set such that $0 \in K$ and suppose that*

$$F := R(0,A)D \in \mathcal{K}(U,\underline{X}), \quad \hat{\imath} \in \mathcal{K}(\mathcal{D}(\hat{A}),\underline{X}). \tag{6.36}$$

Furthermore, let us assume that

$$\iota_N^X \pi_N^X \xrightarrow[\underline{X}]{s} I_{\underline{X}}, \quad \pi_N^{U^*} \iota_N^{U^*} \xrightarrow[U]{s} I_{U^*}, \quad \iota_N^Y \pi_N^Y \xrightarrow[Y]{s} I_Y,$$

and define

$$\pi_N^Z : \mathcal{D}(\hat{A}) \to Z_N, \quad \pi_N^Z = \hat{A}^{-1} \iota_N^X \pi_N^X \hat{A}, \quad N \in \mathbb{N}.$$

Then, the sequences $(F_N)_{N\in\mathbb{N}}$, $(P_N)_{N\in\mathbb{N}}$, $(A_N)_{N\in\mathbb{N}}$, $(E_N)_{N\in\mathbb{N}}$ defined in (6.31) generate discrete operator sequences

$$H_N(s) := \iota_N^X \mathcal{H}_N(s) \pi_N^X, \quad s \in \rho(A_N), \quad N \in \mathbb{N},$$
$$G_N(s) := \iota_N^Y \mathcal{G}_N(s) \pi_N^U, \quad s \in \rho(A_N), \quad N \in \mathbb{N},$$

where

$$\mathcal{S}_N(s) := (s\mathcal{P}_N - \mathcal{A}_N)^{-1}, \quad N \in \mathbb{N}, \tag{6.37}$$
$$\mathcal{H}_N(s) := \mathcal{E}_N \mathcal{P}_N \mathcal{S}_N(s), \quad N \in \mathbb{N}, \tag{6.38}$$
$$\mathcal{G}_N(s)_N := \mathcal{E}_N \mathcal{F}_N - s\mathcal{H}_N(s)\mathcal{F}_N, \quad N \in \mathbb{N}, \tag{6.39}$$

such that

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K,$$

uniformly in $s \in K$.

Proof: As in the proof of Proposition 6.4.2 we obtain immediately

$$F_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} F, \quad E_N \xrightarrow[Y]{s} E. \tag{6.40}$$

Thus, in order to use Theorem 6.2.1, only the following statements remains to be proved:

$$H_N(s) \xrightarrow[Y]{s} H(s), \quad \forall\, s \in K, \tag{6.41}$$

and the existence of an $M > 0$ such that

$$\|H_N(s_1) - H_N(s_2)\|_{\mathcal{L}(\underline{X},Y)} < M|s_1 - s_2|, \quad \forall\, s_1,\, s_2 \in K, \quad \forall\, N \in \mathbb{N}. \tag{6.42}$$

Let us prove them. We have $0 \in K \subset \rho(A)$. Thus, by Lemma 6.4.1, $0 \in \rho(\hat{A})$. It follows that $\hat{A}^{-1} \in \mathcal{L}(\underline{X}, \mathcal{D}(\hat{A}))$ and, due to (6.36), we conclude that

$$s\,\hat{\imath} - \hat{A} \in \mathcal{F}_0(\mathcal{D}(\hat{A}), \underline{X}), \quad \forall\, s \in K.$$

Now, we apply Propositions 5.2.5 and 5.3.17 and obtain

$$P_N \quad \xrightarrow[\underline{X}]{s} \quad \hat{\imath}\,, \tag{6.43}$$

$$sP_N - A_N \quad \xrightarrow[\mathcal{L}(\mathcal{D}(\hat{A}),\underline{X})]{r} \quad s\,\hat{\imath} - \hat{A}, \quad \forall\, s \in \rho(A). \tag{6.44}$$

As in the proof of Proposition 5.3.17, we can show that

$$\iota_N^Z \pi_N^Z \xrightarrow[\mathcal{D}(\hat{A})]{s} I_{\mathcal{D}(\hat{A})}.$$

Moreover, since $K \subset \rho(A)$ is compact, it is a simple matter to show that the operator functions

$$\begin{aligned} \mathcal{T}_N &: K \to \mathcal{F}_0(Z_N, X_N), \quad s \mapsto s\mathcal{P}_N - \mathcal{A}_N, \quad N \in \mathbb{N}, \\ T &: K \to \mathcal{F}_0(\mathcal{D}(\hat{A}), \underline{X}), \quad\quad s \mapsto s\,\hat{\imath} - \hat{A} \end{aligned}$$

satisfy the conditions of Proposition 5.3.10. Thus, according to this proposition, there exists a $N_0 = N(K)$ such that the operators $s\mathcal{P}_N - \mathcal{A}_N$ are invertible for all $s \in K$ and all $N > N_0$. Finally, by Proposition 5.3.8 and Lemma 5.1.4, we conclude that

$$\iota_N^X \mathcal{P}_N \mathcal{S}_N \pi_N^X \xrightarrow[\underline{X}]{s} \hat{\imath}\,(s\,\hat{\imath} - \hat{A})^{-1} = R(s, \hat{A}), \quad \forall\, s \in \rho(A).$$

The last relationship, together with $E_N \xrightarrow[Y]{s} E$ and (again) Lemma 5.1.4, implies (6.41).
It remains to show Equation (6.42). For this one observes that for $\mathcal{S}_N(s)$ we have an identity similar to (6.5), i.e., for every $s_1,\, s_2 \in K$

$$\mathcal{S}_N(s_2) - \mathcal{S}_N(s_1) = -\mathcal{S}_N(s_1)\mathcal{P}_N\mathcal{S}_N(s_2)(s_2 - s_1), \tag{6.45}$$

which is easily obtained from the equality

$$-\mathcal{P}_N(s_2 - s_1) = (s_1\mathcal{P}_N - \mathcal{A}_N) - (s_2\mathcal{P}_N - \mathcal{A}_N).$$

Thus,

$$\mathcal{H}_N(s_1) - \mathcal{H}_N(s_2) = -\mathcal{E}_N\mathcal{P}_N\mathcal{S}_N(s_1)\mathcal{P}_N\mathcal{S}_N(s_2)(s_1 - s_2), \quad \forall\, N > N_0. \tag{6.46}$$

Clearly, the uniformly boundedness of the operator sequences in the RHS of (6.46) would prove (6.42). Let us show that this is case. Indeed, it follows from (6.40) and (6.43) that the sequences $(E_N)_{N\in\mathbb{N}}$ and $(P_N)_{N\in\mathbb{N}}$ are uniformly bounded. On the other hand, we obtain from Proposition 5.3.10 (see the inequality in (5.18)) that

$$\sup_{N>N_0} \max_{s\in K} \|\mathcal{S}_N(s)\|_{\mathcal{L}(X_N, Z_N)} < \infty.$$

These arguments, together with (6.46), ensure that Equation (6.42) holds. The proof is complete.
□

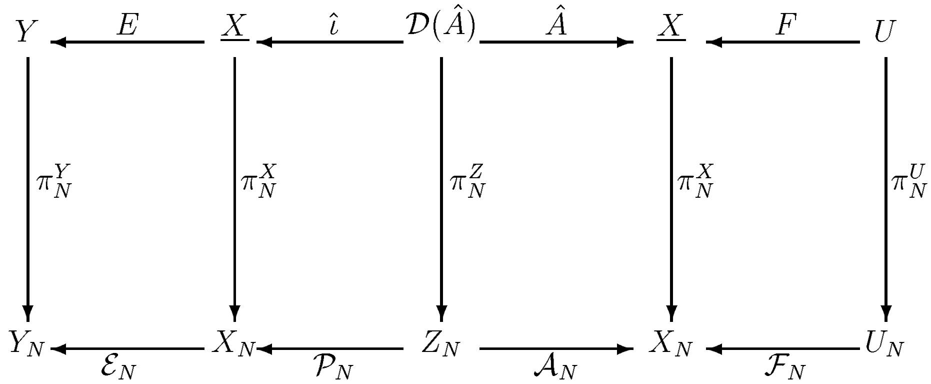

Figure 6.3: Interrelation between subspaces and operators when $H(\cdot)$ is not e-computable.

The situation is depicted in Figure 6.3.

Remark 6.4.4 The calculation of the matrices $(\mathcal{F}_N)_{N\in\mathbb{N}}$ can be simplified if $D \in \mathcal{K}(U, \underline{X})$. In that case one generates the sequence of discrete operators $(D_N)_{N\in\mathbb{N}}$ given by

$$D_N := \iota_N^X \mathcal{D} \pi_N^U, \quad \mathcal{D}_N := \pi_N^X D\, \iota_N^U, \quad N \in \mathbb{N}.$$

Then $D_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} D$ by Proposition 5.2.3 and it is easy to see, using Lemma 5.1.6, that the "new" $(F_N)_{N\in\mathbb{N}}$ is such that

$$F_N := P_N S_N(s_0) D_N \xrightarrow[\mathcal{L}(U,\underline{X})]{n} R(s_0, \hat{A}) D =: F.$$

Moreover, using the identity (6.45), we see that the formulas (6.37) for $(\mathcal{G}_N(\cdot))_{N\in\mathbb{N}}$ take a simpler form, namely,

$$\mathcal{G}_N(s) = \mathcal{E}_N \mathcal{P}_N \mathcal{S}_N(s) \mathcal{D}_N, \quad \forall\, s \in K. \tag{6.47}$$

Remark 6.4.5 Let us suppose, for simplicity, that $D \in \mathcal{K}(U, \underline{X})$. If the sequence $(P_N)_{N\in\mathbb{N}}$, see (6.31), is such that the inverses $\mathcal{P}_N^{-1}$ exist for N large enough then, one has the following identity

$$\mathcal{P}_N \mathcal{S}_N(s) = \mathcal{P}_N \left(s\mathcal{P}_N - \mathcal{A}_N\right)^{-1} = \left(sI_{X_N} - \mathcal{A}_N \mathcal{P}_N^{-1}\right)^{-1}.$$

Let us introduce a new discrete sequence

$$\hat{\mathcal{A}}_N := \mathcal{A}_N \mathcal{P}_N^{-1}, \quad N \in \mathbb{N}.$$

It is obvious now that (see (6.47))

$$\mathcal{G}_N(s) = \mathcal{E}_N \mathcal{R}(s, \hat{\mathcal{A}}_N) \mathcal{D}_N, \quad \forall\, s \in K, \quad \forall\, N \in \mathbb{N}.$$

and, by Proposition 6.4.3, that

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K,$$

uniformly in s. Thus, the triplets $(\hat{\mathcal{A}}_N, \mathcal{D}_N, \mathcal{E}_N)_{N\in\mathbb{N}}$ solve Problem 6.1.1 and consequently, if ρ is a *regular* value of $\|G\|$ (Definition A.0.13), their associated spectral value sets converge to the spectral value sets of the given infinite dimensional operators (A, D, E) :

$$\sigma(\hat{\mathcal{A}}_N, \mathcal{D}_N, \mathcal{E}_N; \rho) \cap K \underset{N\to\infty}{\longrightarrow} \sigma(A, D, E; \rho) \cap K.$$

This is a nice result. Note, however, that the inverses $(\mathcal{P}_N^{-1})_{N\in\mathbb{N}}$ *may not* exist and still the approximation method may work. In the following chapter we will see an example of this sort.

Case 3: Approximation by means of "nice" sequences

The norm in $\mathcal{D}(\hat{A})$ used in the previous section can be restrictive and uncomfortable in applications. Thus, it is interesting to know that, under certain circumstances, the "difficult" regular convergence in $\mathcal{L}(\mathcal{D}(\hat{A}), \underline{X})$ can be substituted by the "easier" regular convergence in $\mathcal{L}(\underline{X})$. Of course, there is a price to pay: one must impose new conditions on A. A description of this possibility is the aim of this section.

The framework of projections and discrete sequences to be used here is essentially the same as before, see (6.29), (6.30), (6.31), (6.32). There is only one but *important* exception which makes the whole difference: the sequences of projections $(\pi_N^Z)_{N\in\mathbb{N}}$ and embeddings $(\iota_N^Z)_{N\in\mathbb{N}}$ are *redefined* to be

$$\pi_N^Z : \underline{X} \to Z_N, \quad \iota_N^Z : Z_N \to \underline{X}, \quad N \in \mathbb{N}. \tag{6.48}$$

The subspaces $(Z_N)_{N\in\mathbb{N}}$ are in $\hat{\imath}\,\mathcal{D}(\hat{A})$, see Equation (6.33) for the definition of $\hat{\imath}$. Compare also with (6.29) and (6.30) and note that the Banach space $\mathcal{D}(\hat{A})$ does *not* appear now. Furthermore, the sequences $(P_N)_{N\in\mathbb{N}}$ and $(A_N)_{N\in\mathbb{N}}$ of (6.31) are not defined. Their role is played by other discrete sequences which will be introduced immediately.

First, by means of (6.48) and (6.29), we define the operators

$$P_N := \iota_N^X \mathcal{P}_N \pi_N^Z, \quad \mathcal{P}_N := \pi_N^X|_{Z_N}. \quad N \in \mathbb{N}, \tag{6.49}$$

Note that, in some sense, they are similar to the operators $(P_N)_{N\in\mathbb{N}}$ of the previous case. Furthermore, in Chapter 5 we showed that if $(\mathcal{P}_N)_{N\in\mathbb{N}} \in \mathcal{N}$ (Definition 5.3.11) then the sequence

$$(\mathcal{Q}_N)_{N\in\mathbb{N}}, \quad \mathcal{Q}_N := \mathcal{P}_N^{-1}\pi_N^X : \underline{X} \to Z_N, \quad N \in \mathbb{N},$$

consists, for each $N \in \mathbb{N}$, of projections of $\underline{X}$ into Z_N. Moreover, if the convergences

$$\iota_N^Z \pi_N^Z \xrightarrow[\underline{X}]{s} I_{\underline{X}}, \quad \iota_N^X \pi_N^X \xrightarrow[\underline{X}]{s} I_{\underline{X}},$$

take place then

$$\iota_N^Z \mathcal{Q}_N \xrightarrow[\underline{X}]{s} I_{\underline{X}}.$$

Note that these convergences are meant in the norm of $\underline{X}$ and not in the norm of the Banach space $\mathcal{D}(\hat{A})$.

Due to the property of "niceness", the inverses of $(\mathcal{P}_N)_{N\in\mathbb{N}}$, defined in (6.49), form a uniformly bounded sequence. Thus, the following sequence is well defined:

$$A_N : \underline{X} \to \underline{X}, \quad A_N x := \iota_N^X \mathcal{A}_N \pi_N^X x, \quad x \in \underline{X},$$

where

$$\mathcal{A}_N : X_N \to X_N, \quad \mathcal{A}_N := \pi_N^X \hat{A} \mathcal{P}_N^{-1} \iota_N^X, N \in \mathbb{N}.$$

Using these notations we can state the following proposition.

Proposition 6.4.6 *Let $A \in \mathcal{C}(X)$, $D \in \mathcal{L}(U, X)$, $E \in \mathcal{L}(\underline{X}, Y)$ and suppose that (4.2) holds. Further, let $K \subset \rho(A)$ be a given compact set and suppose that $F := R(s_0)D \in \mathcal{K}(U, \underline{X})$ for some $s_0 \in K$. Moreover, suppose that $\hat{A}$ (6.27) is of one of the forms*

1. *$\hat{A} = I_{\underline{X}} - S$, $S \in \mathcal{K}(\underline{X})$.*
2. *$\hat{A} = (I_{\underline{X}} - S)R^{-1}$, $S \in \mathcal{K}(\underline{X})$, $1 \in \rho(S)$, $R \in \mathcal{K}(\underline{X})$, R^{-1} exists, but is unbounded.*

Finally, suppose that we are given sequences of projections and embeddings such that

$$\begin{aligned} &\iota_N^Z \pi_N^Z \xrightarrow[\underline{X}]{s} I_{\underline{X}}, \qquad \iota_N^X \pi_N^X \xrightarrow[\underline{X}]{s} I_{\underline{X}}, \\ &\pi_N^{U^*} \iota_N^{U^*} \xrightarrow[U]{s} I_{U^*}, \quad \iota_N^Y \pi_N^Y \xrightarrow[Y]{s} I_Y, \\ &(\mathcal{P}_N)_{N\in N} \in \mathcal{N}. \end{aligned}$$

Then, the sequences of discrete operators $(A_N)_{N\in\mathbb{N}}$, $(F_N)_{N\in\mathbb{N}}$, $(E_N)_{N\in\mathbb{N}}$ defined above generate discrete operator sequences

$$\begin{aligned} H_N(s) &:= \iota_N^X \mathcal{H}_N(s) \pi_N^X, \quad s \in \rho(A_N), \quad N \in \mathbb{N}, \\ G_N(s) &:= \iota_N^Y \mathcal{G}_N(s) \pi_N^U, \quad s \in \rho(A_N), \quad N \in \mathbb{N}, \end{aligned}$$

where

$$\begin{aligned} \mathcal{S}_N(s) &:= (sI_{X_N} - \mathcal{A}_N)^{-1}, && N \in \mathbb{N}, \\ \mathcal{H}_N(s) &:= \mathcal{E}_N \mathcal{S}_N(s), && N \in \mathbb{N}, \\ \mathcal{G}_N(s)_N &:= \mathcal{E}_N \mathcal{F}_N - (s - s_0)\mathcal{H}_N(s)\mathcal{F}_N, && N \in \mathbb{N}, \end{aligned}$$

such that

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in K,$$

uniformly in $s \in K$.

The proof of this proposition is similar to the proof of Proposition 6.4.3. The only difference is that the regular convergence

$$sI_{\underline{X}} - A_N \xrightarrow[\mathcal{L}(\underline{X})]{r} sI_{\underline{X}} - \hat{A}$$

is proved by means of Proposition 5.3.14 or, with certain care, by Remark 5.3.16. We omit the details. The operators and spaces are represented in Figure 6.4.

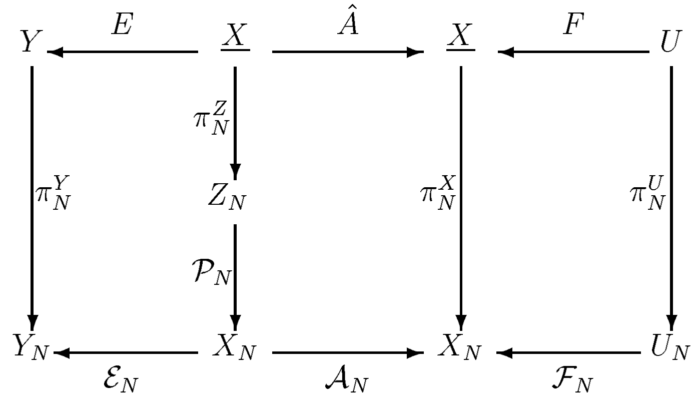

Figure 6.4: Interrelation between subspaces and operators when $\mathcal{P}_N \in \mathcal{N}$.

Remark 6.4.7 Results somewhat similar to Proposition (6.4.6) have been independently obtained in [46] using the concept of quasitriangular convergence mentioned in Remark 5.3.15.

Remark 6.4.8 The nice point in this case is that we always have three operator sequences $(\mathcal{A}_N)_{N\in\mathbb{N}}$, $(\mathcal{F}_N)_{N\in\mathbb{N}}$ and $(\mathcal{E}_N)_{N\in\mathbb{N}}$ which solves Problem 6.1.1 and, consequently, if ρ is regular it holds that

$$\overline{\sigma(\mathcal{A}_N, \mathcal{F}_N, \mathcal{E}_N; \rho)} \cap K \underset{N\to\infty}{\longrightarrow} \overline{\sigma(A, D, E; \rho)} \cap K \tag{6.50}$$

in the sense of the Hausdorff metric.

6.4.4 Remarks on the computation of finite dimensional operators

In applications the subspaces U_N, Z_N, X_N and Y_N will usually be defined via some bases. We shall show a standard method which may help in obtaining the necessary sequences of projections. As illustration, consider, for each $N \in \mathbb{N}$, bases in Z_N and X_N

$$z_i^{(N)} \in Z_N, \quad x_i^{(N)} \in X_N, \quad i = 1, \ldots, N, \quad N \in \mathbb{N}$$

Further, consider the *dual* bases in Z_N^* and X_N^*, i.e., the elements of the dual spaces Z_N^* and X_N^*, respectively, such that

$$\langle\, z_i^{(N)}\,,\, z_j^{(N)*}\,\rangle = \delta_{ij} \quad \langle\, x_i^{(N)}\,,\, x_j^{(N)*}\,\rangle = \delta_{ij}$$
$$i = 1, \ldots, N; \qquad j = 1, \ldots, N.$$

The linear functionals

$$z_j^{(N)*} : Z_N \to \mathbb{C}, \quad x_j^{(N)*} : X_N \to \mathbb{C}, \quad j = 1, \ldots, N$$

can be extended [1, pp 344-345] with conservation of the norm to the whole $\mathcal{D}(\hat{A})$ and $\underline{X}$, respectively. Using these notations, we may define projections

$$\pi_N^Z : \mathcal{D}(\hat{A}) \to Z_N, \quad \pi_N^X : \underline{X} \to X_N$$

by the expressions

$$\pi_N^Z z = \sum_{i=1}^{N} \langle\, z \,,\, z_i^{(N)*} \,\rangle\, z_i^{(N)}; \quad \pi_N^X x = \sum_{i=1}^{N} \langle\, x \,,\, x_i^{(N)*} \,\rangle\, x_i^{(N)}.$$

The same procedure is used for generating the corresponding projections

$$\pi_N^U : U \to U_N, \quad \pi_N^Y : Y \to Y_N.$$

Now, the corresponding matrices of (6.32) can be calculated as follows

$$\begin{aligned}
\mathcal{F}_N &:= (f_{ij}^N) = (\langle\, F u_j^{(N)} \,,\, x_i^{(N)*} \,\rangle),\\
\mathcal{P}_N &:= (p_{ij}^N) = (\langle\, z_j^{(N)} \,,\, x_i^{(N)*} \,\rangle),\\
\mathcal{A}_N &:= (a_{ij}^N) = (\langle\, \hat{A} z_j^{(N)} \,,\, x_i^{(N)*} \,\rangle),\\
\mathcal{H}_N(s) &:= (h_{ij}^N) = (\langle\, H(s) x_j^{(N)} \,,\, y_i^{(N)*} \,\rangle),\\
\mathcal{E}_N &:= (e_{ij}^N) = (\langle\, E x_j^{(N)} \,,\, y_i^{(N)*} \,\rangle).
\end{aligned} \tag{6.51}$$

In words: one takes each element of the basis, applies the desired operator and project the resulting vector into the corresponding finite dimensional space.

Remark 6.4.9 Note that in the formulas (6.51), the dual basis $(z_i^{(N)*})_{N\in\mathbb{N}}$ was not used. This means that, for the calculations, there is not need in constructing them explicitly. The same holds for $(u_i^{(N)*})_{N\in\mathbb{N}}$

6.5 Summary

In this chapter the results obtained in Chapter 5 have been used in order to prove theorems leading to the approximation in norm of transfer functions of infinite dimensional systems. It has also been shown that this convergence is uniform in compact subsets of the resolvent set. Furthermore, the approximation results have been specialised to the case of Riesz operators, where we have shown that the norm of their resolvent operator can be approximated well enough as to permit the calculation of pseudospectrum (unstructured spectral value sets). Finally, we applied the approximation results to the particular case of discrete operators and showed how calculations should be carried out.

Chapter 7

Applications

The aim of this chapter is to illustrate with some examples how the calculation of spectral value sets for operators in Banach spaces can be carried out. For this, we will make use of the main results of the previous chapters: the results on the relationship between level curves of the norm of the transfer function and the spectral value sets, basically Theorem 4.2.11, the approximation schemes of Chapter 5, for example Proposition 6.4.3, and, last but not least, the SH algorithm developed in Chapter 3. Two examples will be considered: delay equations and the Orr-Sommerfeld operator.

7.1 Delay operators

One usual assumption in the mathematical description of a causal physical process is that the behavior of the process depends only on the present state of the system. However, there exist situations where this assumption is not satisfied. In those cases one is forced to take into account information about the *past* of the system as well and this leads to the consideration of delay differential equations, see, for example [17] for an extensive treatment of this topic.
As for the usual situation, robustness and stability are important issues which should be investigated. Spectral value sets can be useful tools in this context and the aim of this section is to develop a suitable framework which enable these investigations. We proceed as follows: first, in the preliminaries, we introduce the necessary mathematical objects, that is, we select a state space description for the delay differential equations. Furthermore, since this state space description is infinite dimensional, we show how the approximation results of the previous chapters can be applied here. Finally, we illustrate the obtained results with one example.

7.1.1 Preliminaries

Let us consider the delay differential equation

$$\dot{z}(t) = Lz_t := \sum_{i=1}^{k_L} L_i z(t-h_i) + \int_{-h}^{0} L_0(\tau) z(t+\tau)d\tau, \quad t > 0, \tag{7.1}$$

$$z_t(\tau) := z(t+\tau), \quad \tau \in [-h, 0], \tag{7.2}$$

where

$$\begin{aligned} &z(0) := x^0 \in \mathbb{C}^n, \\ &z(\tau) := x^1(\tau) \text{ a.e. on } [-h,0), \quad x^1 \in L^2([-h,0);\mathbb{C}^n) \end{aligned} \tag{7.3}$$

and $h > 0$ is fixed, $0 = h_1 < \cdots < h_{k_L} = h$, $L_i \in \mathbb{C}^{n\times n}$, $i = 1,\ldots,k_L$, and $L_0 \in L^2([-h,0];\mathbb{C}^{n\times n})$. It is well known that for every initial condition of the form (7.3) there exists a unique function $z(\cdot)$ on $[0,\infty)$ that is absolutely continuous and satisfies (7.1) almost everywhere [15, Section 2.4], [47].
Suppose now that the operator L is uncertain and that this uncertainty can be adequately modelled by perturbations of (7.1) of the form

$$L \rightsquigarrow L + D_0\Delta M, \tag{7.4}$$

where $D_0 \in \mathbb{C}^{n\times l}$ is a fixed matrix, $\Delta \in \mathbb{C}^{l\times q}$, $\|\Delta\| < \rho$, with $\rho < \infty$ given apriori and M is a linear operator from $C([-h,0];\mathbb{C}^n)$ into $\mathbb{C}^q$ of a form analogous to L, i.e.,

$$Mz_t := \sum_{i}^{k_M} M_i z(t-r_i) + \int_{-h}^{0} M_0(\tau)z(t+\tau)d\tau, \quad z_t \in C([-h,0];\mathbb{C}^n),$$

where $0 = r_1 < \cdots < r_{k_M} = h$, $M_i \in \mathbb{C}^{q\times n}$, $i = 1,\ldots,k_M$, and $M_0 \in L^2([-h,0];\mathbb{C}^{q\times n})$. Note that $M \in \mathcal{L}(C([-h,0];\mathbb{C}^n),\mathbb{C}^q)$ [71]. It is readily seen that these structures permit the application of a wide range of perturbations to the operator L.

Example Suppose L reduces to a sum of a finite number of points delays, i.e.,

$$Lz_t := \sum_{i=1}^{k} L_i z(t-h_i),$$

If the matrices L_i are uncertain, it is natural to consider perturbations of the form $D_0 = I_n$ and

$$Mz_t := \sum_{i=1}^{k} M_i z(t-h_i),$$

with given matrices $M_i \in \mathbb{C}^{n\times n}$, $i = 1,\ldots,k$. By choosing different M_i one can specify the type and size of the uncertainty which is allowed for each matrix L_i. □

We explain now how this problem can be handled within the framework developed in Chapter 4. First of all, we need a state space description of System (7.1). One approach is to use (see (7.3))

$$x(t) = (z(t), z_t) \in X := \mathbb{C}^n \times L^2([-h,0];\mathbb{C}^n).$$

The evolution in time of $x(t)$ defines the strongly continuous semigroup $S(t)$ of bounded linear operators in X defined by [71]

$$S(t)x = (z(t), z_t) \in X; \quad x \in X; \quad t \geq 0, \tag{7.5}$$

where $z(t)$, $t \geq 0$, is the unique solution of (7.1) with initial condition (7.3). The infinitesimal generator A of $S(t)$ is given by

$$\begin{aligned} Ax &:= \begin{pmatrix} Lx^1 \\ \frac{d}{dt}x^1 \end{pmatrix}, \\ \mathcal{D}(A) &:= \{x := (x^0, x^1) \in X \mid x^1 \in H^1([-h,0];\mathbb{C}^n),\ x^0 = x^1(0)\}. \end{aligned} \tag{7.6}$$

Obviously, the behavior of a solution $z(\cdot)$ to (7.1) is determined by the properties of the operator A in (7.6).
As usual, we shall focus our attention on the spectrum of A. In particular, we are interested in its robustness with respect to perturbations D and E with clear interpretation in terms of the original data (7.1). Some general properties of $\sigma(A)$ are the following: it is discrete, the multiplicity of each eigenvalue is finite and for every $\rho \in \mathbb{R}$, there are only finitely many elements of $\sigma(A)$ in $C_\rho^+ := \{s \in \mathbb{C};\quad \operatorname{Re} s > \rho\}$. Moreover, the spectrum of A has the following representation

$$\sigma(A) = \{\lambda \in \mathbb{C};\quad \det(\mathcal{F}(s)) = 0\},$$

where $\mathcal{F}(s) : \mathbb{C} \to \mathbb{C}^{n\times n}$ denotes the matrix function

$$\mathcal{F}(s) := sI_n - L(e^{s\cdot}I_n), \tag{7.7}$$

and I_n denotes the $n \times n$ identity matrix [47]. Furthermore, the resolvent

$$R(s,A) : X \to X$$

is a compact operator and is given by [47]

$$R(s)(\phi^0,\phi^1) = x$$

where $x = (x^0, x^1)$ is

$$\begin{aligned} x^0 &= F(s)^{-1}\left[\phi^0 + L\left(\int_{\cdot}^{0} e^{s(\cdot-\tau)}\phi^1(\tau)d\tau\right)\right] \\ x^1(\theta) &= e^{s\theta}x^0 + \int_\theta^0 e^{s(\theta-\tau)}\phi^1(\tau)d\tau, \quad \theta \in [-h,0]. \end{aligned}$$

We now introduce perturbation structures by means of suitable operators D and E. We choose them so that they contain the same information as the perturbations given in (7.4). For this, we define $U = \mathbb{C}^l$, $Y = \mathbb{C}^q$, and

$$\underline{X} := \{x = (x_0, x_1) \in \mathbb{C}^n \times C([-h,0];\mathbb{C}^n)\}. \tag{7.8}$$

The operators $D : U \to X$ and $E : \underline{X} \to Y$ are given by

$$\begin{aligned} D : U \to X;\quad Du &:= \begin{pmatrix} D_0 u \\ 0 \end{pmatrix} \in X, \\ E : \underline{X} \to Y;\quad Ex &:= \quad Mx_1 \quad \in Y. \end{aligned} \tag{7.9}$$

Our aim is to study the mobility of $\sigma(A)$ under perturbations with the structures (7.9). Namely, we consider $\sigma(A_\Delta)$ with A_Δ given by

$$A_\Delta := A + D\Delta E. \tag{7.10}$$

Note that A_Δ is the infinitesimal generator of the semigroup corresponding to the perturbed operator (7.4).
Following the "spirit" of the previous chapter we see that the conditions

1. $\mathcal{D}(A) \subset \underline{X} \subset X$ with continuous dense injections,

2. $D : U \to X$ is bounded,

3. $E \in \mathcal{L}(\underline{X}, X)$,

hold. Thus, Theorem 4.2.11 can be applied and we obtain that, in order to calculate the sets $\sigma(A, D, E; \rho)$, we must look for those regions of the complex plane where the corresponding transfer function

$$G(s) := ER(s)D, \quad s \in \rho(A),$$

has the property $\|G(s)\|_{\mathcal{L}(U,Y)} > \rho^{-1}$. In this case the transfer function G takes a simple form which, if only point delays in the operator M are allowed, looks as follows

$$G(s) := \left(\sum_{i=1}^{k_M} M_i e^{s - r_i} \right) \mathcal{F}(s)^{-1} D_0. \quad s \in \rho(A). \tag{7.11}$$

We stress that in this situation considering *unstructured* perturbations of the nominal operator (7.6) makes no sense. In this situation, it is only meaningful to consider perturbations with clear interpretations in terms of System (7.1).

7.1.2 Approximation in the case of delay system

The problem on the approximation of delay operators has been matter of extensive investigations, see for example [71], [47], [48] and references therein. Our results are, essentially, extensions of those obtained by Ito and Kappel in [48]. Nevertheless, there are differences between our approach and the others mentioned. The aim in those papers was to approximate the semigroup $S(t)$ (introduced in (7.5)) uniformly over finite intervals of time. Furthermore, the authors considered, if any, only output operators of finite rank, i.e.,

$$E : X \to \mathbb{C}^p, \quad Ex := E_0 x^0, \quad E_0 \in \mathbb{C}^{q \times n}, \quad x = (x^0, x^1) \in X.$$

Although our investigations are related to those papers, they pursue other goals. First, we have to deal with a more general class of output operators E, namely $E \in \mathcal{L}(\underline{X}, Y)$ of the form (7.9). Secondly, we are not interested in the approximation of the semigroup $S(t)$, but in a scheme able to approximate the transfer function G in norm and uniformly in s in any compact $K \subset \rho(A)$, see Problem 6.1.1.
In [47] the authors used with satisfactory performances an approximation scheme based on projections of $\mathcal{D}(A)$ onto subspaces formed by the span of *linear splines* and projections of X

onto the span of *step functions.* Later, in [48], they showed that schemes based on other piecewise polynomials work as well.

We use a similar approach here but with one difference: following the approach of Subsection 6, we shall *not* approximate the operators A, D and E in X, but in $\underline{X}$. Our approximation scheme will make use of linear splines in $\underline{X}$, while quadratic splines are going to be utilised in $\mathcal{D}(\hat{A})$, where $\mathcal{D}(\hat{A})$ is the subspace of $\mathcal{D}(A)$ defined in (6.28). Note that in this case (see (7.6) and (7.8)) $\mathcal{D}(\hat{A})$ is just

$$\mathcal{D}(\hat{A}) = \{x = (x_0, x_1) \in \mathcal{D}(A) \,|\, x^1 \in C^1([-h, 0]; \mathbb{C}^n)\}.$$

Note also that $D : U \to \underline{X}$ is compact. Thus, we shall apply Proposition 6.4.3 together with Remark 6.4.4 in order to achieve the desired results. We shall use a projection scheme which fits in the framework of Figure 6.4.

Subspaces and projections

The projection scheme to be described here is based on the theory of *B-splines.* The monograph [72] is the main reference.

For a given $N \in \mathbb{N}$, we introduce a uniform mesh [72] with respect to the interval $[-h, 0]$, i.e.,

$$t_k^N := -kr, \quad k = 0, \dots, N, \quad r := \frac{h}{N}. \tag{7.12}$$

Let us consider $N + 1$ linear splines

$$\{e_k^{(N)}\}_{k=1}^{N+1}; \quad e_k^{(N)}(t) := N_2(-\frac{t}{r} - k + 1), \quad t \in [-h, 0],$$

where

$$N_2(\theta) := \begin{cases} \theta & for \quad 0 \le \theta \le 1, \\ 2 - \theta & for \quad 1 \le \theta \le 2, \\ 0 & elsewhere \end{cases}$$

and $N + 2$ quadratic splines given by

$$\{b_k^{(N)}\}_{k=0}^{N+1}; \quad b_k^{(N)}(t) := 0.5 * N_3(-\frac{t}{r} - k + 2), \quad t \in [-h, 0],$$

where

$$N_3(\theta) := \begin{cases} \theta^2 & for \quad 0 \le \theta \le 1 \\ -2\theta^2 + 6\theta - 3 & for \quad 1 \le \theta \le 2 \\ \theta^2 - 6\theta + 9 & for \quad 2 \le \theta \le 3 \\ 0 & elsewhere \end{cases}$$

They are represented in Figure 7.1.2. It is well known that $\{e_k^{(N)}\}_{k=1}^{N+1}$ and $\{b_k^{(N)}\}_{k=0}^{N+1}$ are bases in the spaces of linear and quadratic *splines with simple knots* [72, Example 4.3] corresponding to the uniform mesh (7.12). This can be seen, for example, from the fact that these spaces

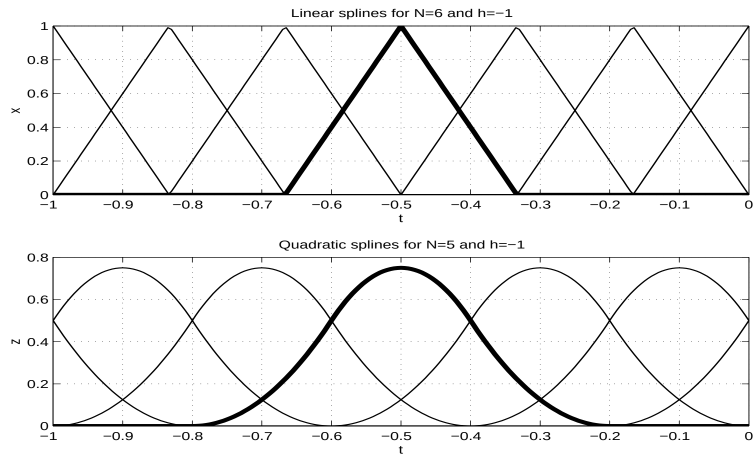

Figure 7.1: Bases of linear and quadratic splines for $N = 6$ and $N = 5$, respectively.

have dimension $N + 1$ and $N + 2$, respectively, [72, Theorem 4.4] and the linear independence of $\{e_k^{(N)}\}_{k=1}^{N+1}$ and $\{b_k^{(N)}\}_{k=0}^{N+1}$.
Now, following [47], one introduces

$$Z_N^1 \subset C^1([-h,0];\mathbb{C}^n), \quad X_N^1 \subset C([-h,0];\mathbb{C}^n),$$

which are the spaces spanned by the *columns* of

$$[b_0^{(N)} I_n,\, b_1^{(N)} I_n,\, \ldots,\, b_{N+1}^{(N)} I_n] \quad \text{and} \quad [e_1^{(N)} I_n,\, e_2^{(N)} I_n,\, \ldots,\, e_{N+1}^{(N)} I_n],$$

respectively. Furthermore, as in [47], we introduce

$$\begin{aligned} Z_N &:= \; Z_N^0 \times Z_N^1 = Q Z_N^1 \; \subset \mathcal{D}(\hat{A}), \\ X_N &:= \; \mathbb{C}^n \times X_N^1 \; \subset \underline{X}, \end{aligned}$$

where Q denotes the operation $f \to (f(0), f)$. It can be easily proved that the *columns* of

$$\begin{aligned} \{B_i^{(N)}\}_{i=0}^{N+1} &:= \left[Q(b_0^{(N)} I_n),\, Q(b_1^{(N)} I_n),\, \ldots,\, Q(b_{N+1}^{(N)} I_n) \right], \\ \{E_i^{(N)}\}_{i=0}^{N+1} &:= \left[\begin{pmatrix} I_n \\ 0 \end{pmatrix}, \begin{pmatrix} 0_n \\ e_1^{(N)} I_n \end{pmatrix}, \ldots, \begin{pmatrix} 0_n \\ e_{N+1}^{(N)} I_n \end{pmatrix} \right] \end{aligned}$$

are bases in Z_N and X_N, respectively. These spaces have dimension $n(N + 2)$. See [71] and [48] for similar formulae.

Remark 7.1.1 The matrix I_n above accounts for the dimension n present in the state x, see (7.5): one uses $N + 2$ splines for the approximation of each of the n "coordinates" of x.

A *dual* basis $\{f_i^{(N)}\}_{i=0}^{N+1}$ with respect to the basis in X_N given by the columns of $\{E_i^{(N)}\}_{i=0}^{N+1}$ is given by the linear "functionals" [72, Example 4.40]

$$f_i^{(N)}x = \begin{cases} (x^0)^T & for \quad i=0, \\ (x^1(t_{i-1}^N))^T & for \quad i=1,\dots,N+1, \end{cases}$$

for every $x=(x^0,x^1)\in\underline{X}$. Note that $f_i^{(N)}x\in\mathbb{C}^{1\times n}$ and that the relationship

$$f_i^{(N)}E_j^{(N)} = \delta_{ij}I_n, \quad i,j=0,\dots,N+1,$$

holds in a trivial way.
Having defined these bases, we may introduce projections

$$\pi_N^X : \underline{X}\to X_N, \quad N\in\mathbb{N},$$

defined by

$$\pi_N^X x := \sum_{i=0}^{N+1} f_i^{(N)}(x)E_i^{(N)}, \quad \forall\, x\in\underline{X}. \tag{7.13}$$

The fact that

$$\iota_N^X\pi_N^X \xrightarrow[\underline{X}]{s} I_{\underline{X}}$$

is implied by [72, Lemma 6.59, Corollary 6.21]. Further, let us assume that $0\in\rho(A)$. Then, by the definitions of the spaces Z_N, X_N and of the operator $\hat{A}$ we observe that

$$X_N = \hat{A}Z_N \subset \underline{X}, \quad \forall\, N\in\mathbb{N},$$

and all this means that

$$\pi_N^Z := \hat{A}^{-1}\,\iota_N^X\pi_N^X\hat{A}, \quad N\in\mathbb{N}, \tag{7.14}$$

is a projection of $\mathcal{D}(\hat{A})$ onto Z_N for every $N\in\mathbb{N}$. From the proof of Theorem 5.3.17, we know that

$$\iota_N^Z\pi_N^Z \xrightarrow[\mathcal{D}(\hat{A})]{s} I_{\mathcal{D}(\hat{A})}.$$

In this case the sequences of spaces $(U_N)_{N\in\mathbb{N}}$ and $(Y_N)_{N\in\mathbb{N}}$ are chosen stationary and equal to U and Y, respectively. Hence, the projections π_N^U, $\pi_N^{U^*}$ and π_N^Y are just the identity matrices in $\mathbb{C}^l$ and $\mathbb{C}^q$, respectively, and the strong convergences

$$\iota_N^U\pi_N^U \xrightarrow[U]{s} I_U, \quad \pi_N^{U^*}\,\iota_N^{U^*} \xrightarrow[U^*]{s} I_{U^*}, \quad \iota_N^Y\pi_N^Y \xrightarrow[Y]{s} I_Y \tag{7.15}$$

hold trivially.

Discrete operators

Above, as in [48], we have assumed that $0 \in \rho(A)$. Further, the system of projections (7.13), (7.14), (7.15) fit into the projection framework of Case 2 in Section 6.4.3. Thus, by simple inspection we see that the conditions of Proposition 6.4.3 hold in this case and we conclude that the projections (7.13), (7.14) and (7.15) generate discrete operators with the necessary properties for the solution of Problem 6.1.1. Furthermore, an additional simplification is given by the fact that $D \in \mathcal{K}(U, \underline{X})$, because we may use the formulas for $(G_N(s))_{N\in\mathbb{N}}$ displayed in Remark 6.4.4, Equation (6.47). The description of the approximation method is complete with

Proposition 7.1.2 *Consider the following sequences of discrete operators, where the notation* $d := n(N+2)$ *will be used.*

1. $P_N : \mathcal{D}(\hat{A}) \to \underline{X}$, $P_N := \iota_N^X \mathcal{P}_N \pi_N^Z$, $\mathcal{P}_N := \pi_N^X\, \hat{\imath}\, \iota_N^Z$, $N \in \mathbb{N}$, *where* $\hat{\imath} : \mathcal{D}(\hat{A}) \to \underline{X}$ *is the corresponding embedding. For each* $N \in \mathbb{N}$, *the sequence* $(\mathcal{P}_N)_{N\in\mathbb{N}}$, $\mathcal{P}_N : Z_N \to X_N$, *has matrix representation with respect to the bases* $\{B_i^{(N)}\}_{i=0}^{N+1}$ *and* $\{E_i^{(N)}\}_{i=0}^{N+1}$ *given by*

$$\mathcal{P}_N = \begin{pmatrix} 0.5 & 0.5 & 0 & \dots & \dots & \dots \\ 0.5 & 0.5 & 0 & \dots & \dots & \dots \\ 0 & 0.5 & 0.5 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0.5 & 0.5 & 0 \\ \dots & \dots & \dots & 0 & 0.5 & 0.5 \end{pmatrix} \otimes I_n \in \mathbb{C}^{d\times d}.$$

2. $A_N : \mathcal{D}(\hat{A}) \to \underline{X}$, $A_N := \iota_N^X \mathcal{A}_N \pi_N^Z$, $\mathcal{A}_N := \pi_N^X \hat{A}\, \iota_N^Z$, $N \in \mathbb{N}$, *where for each* $N \in \mathbb{N}$, *the sequence* $(\mathcal{A}_N)_{N\in\mathbb{N}}$, $\mathcal{A}_N : Z_N \to X_N$, *has matrix representation with respect to the bases* $\{B_i^{(N)}\}_{i=0}^{N+1}$ *and* $\{E_i^{(N)}\}_{i=0}^{N+1}$ *given by*

$$\mathcal{A}_N := \begin{pmatrix} \mathcal{A}_N^0 & \mathcal{A}_N^1 & \dots & \dots & \dots & \mathcal{A}_N^{N+1} \\ r^{-1} & -r^{-1} & 0 & \dots & \dots & \dots \\ 0 & r^{-1} & -r^{-1} & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & r^{-1} & -r^{-1} & 0 \\ \dots & \dots & \dots & 0 & r^{-1} & -r^{-1} \end{pmatrix} \otimes I_n \in \mathbb{C}^{d\times d},$$

$$\mathcal{A}_N^i = L(b_i^N), \quad i = 0, \dots, N+1.$$

3. $D_N : U \to \underline{X}$, $D_N := \iota_N^X \mathcal{D}_N \pi_N^U$, $\mathcal{D} := \pi_N^X D\, \iota_N^U$, $N \in \mathbb{N}$, *where for each* $N \in \mathbb{N}$, *the sequence* $(\mathcal{D}_N)_{N\in\mathbb{N}}$, $\mathcal{D}_N : U_N \to X_N$ *has matrix representation*

$$\mathcal{D}_N := \begin{pmatrix} D_0 \\ 0_n \\ \vdots \\ 0_n \end{pmatrix} \in \mathbb{C}^{d\times l}.$$

4. $E_N : \underline{X} \to Y$, $E_N := \iota_N^Y \mathcal{E}_N \pi_N^X$, $\mathcal{E}_N := \pi_N^Y E \iota_N^X$, $N \in \mathbb{N}$, *where for each* $N \in \mathbb{N}$, *the sequence* $(\mathcal{E}_N)_{N\in\mathbb{N}}$, $\mathcal{E}_N : X_N \to Y_N$ *has matrix representation given by*

$$\begin{aligned} \mathcal{E}_N &:= \begin{pmatrix} \mathcal{E}_N^0, & \dots, & \mathcal{E}_N^{N+1} \end{pmatrix} \in \mathbb{C}^{q\times d} \\ \mathcal{E}_N^i &= M(e_i^N), \quad i = 0, \dots, N+1. \end{aligned}$$

Then, the sequence of discrete operators $(G_N(s))_{N\in\mathbb{N}}$, $s \in \rho(A)$, *given by*

$$\begin{aligned} G_N(s) &:= \qquad \iota_N^Y \mathcal{G}_N(s) \pi_N^U, & \forall\, s \in \rho(A), \quad N \in \mathbb{N}, \\ \mathcal{G}_N(s) &:= \mathcal{E}_N \mathcal{P}_N \left(s\mathcal{P}_N - \mathcal{A}_N\right)^{-1} \mathcal{D}_N, & \forall\, s \in \rho(A), \quad N \in \mathbb{N}, \end{aligned} \tag{7.16}$$

is such that

$$G_N(s) \xrightarrow[\mathcal{L}(U,Y)]{n} G(s), \quad \forall\, s \in \rho(A).$$

Moreover, this convergence is uniform in $s \in K$, *for any compact* $K \subset \rho(A)$.

The matrices above have been written down following the method of Section 6.4.4. See [71] and [48] for other formulas of this type.

7.1.3 Example: Damped oscillator

In this subsection we consider an example in which the transfer function $G(s) = ER(s)D$ can be explicitly calculated. Our aim is to compare the approximated spectral value sets with the exact ones and, in this way, to develop some intuition about the convergence process.
Consider the perturbed delay equation

$$\ddot{\xi}(t) + 2\dot{\xi}(t) + 6\xi(t) + \xi(t-1) + \delta\dot{\xi}(t-\alpha) = 0.$$

Here α is a fixed delay, $0 \le \alpha \le 1$, and the term $\delta\dot{\xi}(t-\alpha)$ is the perturbation with $\delta \in \mathbb{C}$ unknown. Setting $z = [\xi \;\; \dot{\xi}]^T$ this equation can be re-written as

$$\dot{z}(t) = Lz_\tau(t) + D_0 \delta M z(t),$$

where the operator L is given by (compare with Equation (7.1))

$$Lz(t) = A_0 z(t) + A_1 z(t-1)$$

and

$$\begin{aligned} A_0 &= \begin{bmatrix} 0 & 1 \\ -6 & -2 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ -1 & 0 \end{bmatrix}, \\ D_0 &= [0 \;\; 1]^T, \quad Mz(t) = [0 \;\; 1] z(t-\alpha). \end{aligned}$$

It was shown in the preliminaries that the above expressions can be associated with an abstract equation on $X := \mathbb{C}^2 \times L^2([-1,0];\mathbb{C}^2)$, namely

$$\dot{x}(t) = Ax(t) + D\delta Ex(t),$$

where

$$Ax(t) = \left\{ \begin{array}{l} A_0 x^1(t) + A_1 x^1(t-1) \\ \frac{d}{dt} x^1(\cdot) \end{array} \right\}$$
$$x = (x^0, x^1) \in \mathcal{D}(A) \subset \mathbb{C}^2 \times H^1([-1,0]; \mathbb{C}^2),$$

and

$$D = [D_0 \;\; 0]^T, \quad Ex(t) = [0 \;\; 1] x^1(t - \alpha),$$

for $x = (x^0, x^1) \in \underline{X} := \mathbb{C}^2 \times C([-1,0]; \mathbb{C}^2)$. We must approximate the transfer function of the

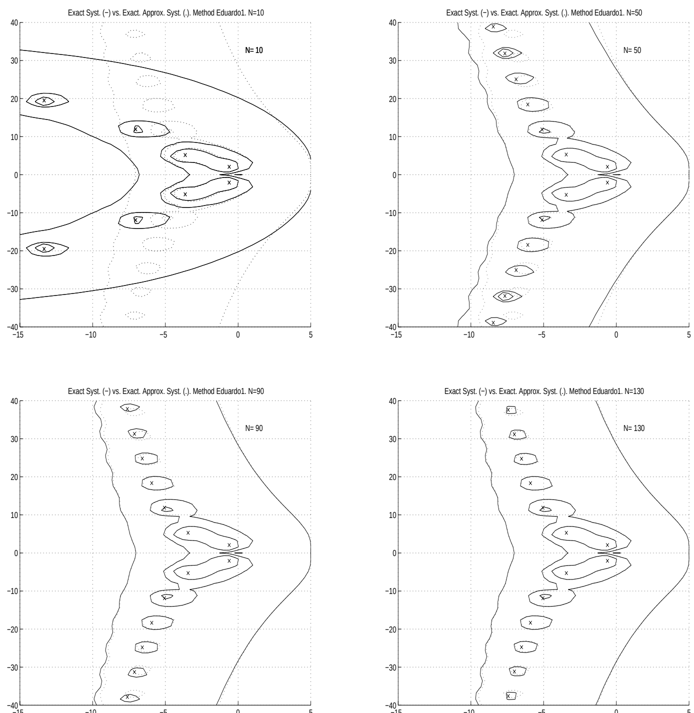

Figure 7.2: Approximation of $\sigma(A, D, E; \rho)$

system

$$\dot{x}(t) = Ax(t) + Du(t),$$
$$y(t) = Ex(t).$$

In order to do this, we shall use the projection method explained in the previous subsection. The exact transfer function of the original system can be calculated, see (7.11), and is given by

$$G(s) = \frac{e^{-\alpha s}\, s}{s^2 + 2s + 6 + e^{-s}}. \tag{7.17}$$

Note that

$$\lim_{\mathrm{Re}\, s \to +\infty} \|G(s)\|_{\mathcal{L}(U,Y)} = 0,$$

so, the stability radius for closedness $r(A, D, E) = \infty$. Moreover, Proposition 7.1.2 ensures that the spectral value sets of the delay system can be approximated arbitrarily closely. For the calculations the SH algorithm of Chapter 3 has been used.
We calculated the level curves of

$$s \mapsto \|\mathcal{G}_N(s)\|_{\mathcal{L}(U_N,Y_N)}, \quad s \in K,$$

where K is the rectangle given by the corners $(-15, -40)$ and $(5, 40)$, for three different values of ρ: $2.4, 0.4, 0.2$. The delay was taken $\alpha = 0.25$. In Figure 7.1.3 the continuous lines are those based on the approximate transfer functions (7.16), while the dotted ones correspond to the exact contours computed by Formula (7.17). In all cases they form a sort of elliptic region of the complex plane. The convergence of the approximant is relatively slow and in order to get the actual sets near the imaginary axis it was necessary to compute approximations of high order ($N > 100$). Nevertheless, the graphics illustrate the result proved before, namely, the approximation of the sets $\sigma(A, D, E; \rho)$.

7.2 The Orr-Sommerfeld operator

The stability analysis of parallel laminar flows has a long history. In particular, the flow of a viscous fluid in a channel formed by two infinite parallel planes, usually called Poiseuille flow, has attracted the attention of mathematicians and physicists for more than a century. The example of this section is related to this topic: we shall study spectral value sets of the associated Orr-Sommerfeld operator.

7.2.1 Preliminaries

The nominal flow, whose stability must be investigated, is described by the following evolution (Navier-Stokes) equations

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial z} + v\frac{\partial u}{\partial y} = -\frac{\partial p}{\partial z} + \frac{1}{R}\Delta u \tag{7.18}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial z} + v\frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + \frac{1}{R}\Delta v \tag{7.19}$$

$$\frac{\partial u}{\partial z} + \frac{\partial v}{\partial y} = 0. \tag{7.20}$$

Here $\Delta := \frac{\partial^2}{\partial z^2} + \frac{\partial^2}{\partial y^2}$ is the Laplace operator, z denotes the direction of the pressure gradient parallel to the planes, y the distance normal to them measured from the channel center, u, v the corresponding velocity components, p the pressure, t the time and R the Reynolds number. The boundary conditions are

$$u(t, z, \pm 1) = v(t, z, \pm 1) = 0.$$

The basic flow, represented in Figure 7.2.1, is given by

$$U(y) = 1 - y^2, \quad V = 0, \quad P = -\frac{2}{R}z. \tag{7.21}$$

A linearisation the Navier-Stokes equations (7.18), (7.19), (7.20) about the exact solution (7.21)

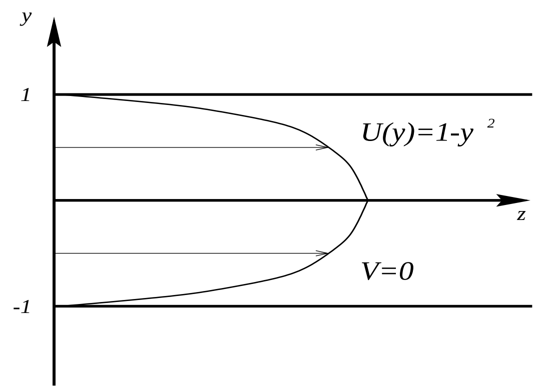

Figure 7.3: Poiseuille flow.

leads to the Orr-Sommerfeld equation. First, one introduces a stream function ψ, i.e., a function such that

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial z}.$$

In terms of ψ, Equation (7.20) is satisfied automatically and eliminating the pressure the Navier-Stokes equations become

$$\frac{\partial \Delta\psi}{\partial t} + \frac{\partial \psi}{\partial y}\frac{\partial \Delta\psi}{\partial z} - \frac{\partial \psi}{\partial z}\frac{\partial \Delta\psi}{\partial y} = \frac{1}{R}\Delta\Delta\psi,$$
$$\frac{\partial \psi}{\partial z}(t, z, \pm 1) = \frac{\partial \psi}{\partial y}(t, z, \pm 1) = 0.$$

The stream function ψ has as remarkable property that it is *constant* along the current lines defined by the solution of (7.18), (7.19), (7.20). In the case of (7.21) the associated stream function is

$$\psi_0(y) := y - y^3/3 + c, \tag{7.22}$$

where c is an arbitrary constant.
In order to analyse the evolution in time of small time varying perturbations to (7.21), one assumes that the perturbation is of sinusoidal form along the direction z. This situation is mathematically described by a decomposition

$$\psi(t, z, y) = \psi_0(y) + \phi(t, y)e^{\imath\alpha z},$$

where α is real and positive, ψ_0 is the stream function of the nominal flow (7.22), and $\phi(t, y)$ is assumed to be small. Note that now complex solutions are allowed. Clearly, ϕ must satisfy the boundary conditions:

$$\phi(t, \pm 1) = \frac{\partial \phi}{\partial y}(t, \pm 1) = 0.$$

Using these expressions, one obtains the following equation for $\phi(t, y)$:

$$\frac{\partial \Delta \phi}{\partial t} = \left[\frac{1}{R}\Delta\Delta - \imath\alpha(U\Delta - \frac{\partial^2 U}{\partial y^2})\right]\phi + \imath\alpha(\frac{\partial \phi}{\partial y} - \phi\frac{\partial}{\partial y})(\Delta\phi)e^{\imath\alpha z}, \tag{7.23}$$

$$\Delta := \frac{\partial^2}{\partial y^2} - \alpha^2, \tag{7.24}$$

$$\phi(t, \pm 1) = \frac{\partial \phi}{\partial y}(t, \pm 1) = 0. \tag{7.25}$$

In the linear stability theory, the last term in (7.23) is neglected and one studies the resulting linear operator (at this stage written just formally)

$$A := \Delta^{-1}\left[\frac{1}{R}\Delta\Delta - \imath\alpha(U\Delta - \frac{\partial^2 U}{\partial y^2})\right]$$

in order to make stability statements. Note that A is parameterised through R and α.
For physical as well as for theoretical reasons it is convenient to study A in the Banach space X defined by [62]

$$X := H_0^1([-1; 1]; \mathbb{C}). \tag{7.26}$$

The norm in X

$$\|x\|_X := \int_{-1}^{1}(|\frac{\partial x}{\partial y}|^2 + \alpha^2|x|^2)dy$$

is usually called *energy norm*. See also [62, p. 223] where this space is denoted by $L_{2,0}^{(1)}$. Integrating by parts one sees that

$$\|x\|_X^2 = -\int_{-1}^{1} x^* M x dy = -\langle Mx, x\rangle_{L^2([-1,1];\mathbb{C})}, \tag{7.27}$$

where

$$M : \mathcal{D}(M) \to L^2([-1,1];\mathbb{C}), \quad x \mapsto \frac{\partial^2 x}{\partial y^2} - \alpha^2 x, \tag{7.28}$$

where

$$\mathcal{D}(M) := H_0^1([-1,1];\mathbb{C}) \cap H^2([-1,1];\mathbb{C}).$$

As an operator in $L^2([-1,1];\mathbb{C})$, M has the following properties [15, Example A.4.26].

1. $\sigma(M) = \{\tilde{\mu}_i := -(\alpha^2 + i^2\pi^2), \quad i \in \mathbb{N}\}$. The eigenvalues are simple.

2. M^{-1} exists, is selfadjoint and compact. Moreover, its eigenfunctions $(\tilde{e}_N)_{N\in\mathbb{N}}$ form an orthonormal basis in $L^2([-1,1];\mathbb{C})$.

Definition 7.2.1 The linear operator

$$\begin{aligned} &A : \mathcal{D}(A) \to \mathcal{D}(M), \quad x \mapsto M^{-1}Sx, \\ &S := \frac{1}{R}M^2 - \imath\alpha(UM - \frac{\partial^2 U}{\partial y^2}), \\ &\mathcal{D}(A) := \{x \in H^4([-1,1];\mathbb{C}), \quad x(\pm 1) = \frac{\partial x}{\partial y}(\pm 1) = 0\}. \end{aligned} \tag{7.29}$$

is called *Orr-Sommerfeld operator.*

Most of the investigations related to the Orr-Sommerfeld operator have had the following pattern: one makes *predictions* about the stability of the flow (7.21) based on the *location* of the spectrum of the Orr-Sommerfeld operator (7.29). One expects the basic flow (7.21) to be stable for low Reynolds numbers R (high viscosity and/or low velocities), while for large R, instability (turbulence) should be observed. Usually one looks for the lowest (critical) Reynolds number R_c for which $\sigma(A)$ "leaves" the left half of the complex plane for some value of α. Accurate calculations give the estimation $R_c = 5772$ for $\alpha = 1.02$ [19].
However, the results of this approach are *not* in good agreement with the experiments. In fact, turbulence has been observed for $R \approx 1000$ [19, pp 452], where as an interesting fact we mention that *laminar* flows can be observed at substantially higher values of R ($R \approx 8000$) than the theoretical value $R_c = 5772$ [19, pp 453]. Explaining these discrepancies has been one of the hardest problems in the theory of hydrodynamical stability [19], [37]. The results of this section are also related to this problem.

7.2.2 Spectral properties

The spectral properties of the Orr-Sommerfeld operator have been matter of extensive investigations. It is known that the spectrum $\sigma(A)$ is purely discrete [19, pp 156] and there are satisfactory methods for their calculation [60] so long as R is not too large. Moreover, the eigenfunctions of A are *complete* in X [62]. Note that the completeness of a function system *does not* mean that it is a basis. See [53, V.2.5] for the definition of these concepts.
It would be interesting to know whether the Orr-Sommerfeld operator is a *Riesz* operator or not, see Definition 6.3.5. We show here that the answer to this question is, at least in some sense, positive. The key is the following theorem. Note that it is simply a *reformulation* of [53, Theorem V.4.15, Remark V.4.16.c].

Theorem 7.2.2 *Let X be a Hilbert space, T be a selfadjoint operator in X, with compact resolvent and simple eigenvalues $\cdots < \mu_2 < \mu_1$. Further, suppose that its eigenvalues have the property*

$$\lim_{i\to\infty} \mu_i - \mu_{i+1} = \infty.$$

Let $(P_i)_{i\in\mathbb{N}}$ be the eigenprojections of T, so that

$$P_i = \langle \cdot\, , e_i \rangle_X \, e_i, \quad i \in \mathbb{N}.$$

Finally, consider some T-bounded operator S. Then, $T+S$ is closed with compact resolvent. Moreover, there exists $L \in \mathcal{L}(X)$ with $L^{-1} \in \mathcal{L}(X)$ such that the eigenprojections $(Q_i)_{i\in\mathbb{N}}$ of $T+S$ can be indexed as $\{Q_{0j}, Q_i\}$, $j=0,\ldots,m<\infty$, $i=n+1,n+2,\ldots$, in such a way that the equalities

$$\sum_{j=0}^{m} Q_{0j} = L\left(\sum_{i=0}^{n} P_i\right)L^{-1}, \quad Q_i = LP_iL^{-1}, \quad i>n, \tag{7.30}$$

hold.

The results obtained in [62] make possible the application of these ideas to the Orr-Sommerfeld case. Essentially, the authors have proved all the facts we need for our purposes. For clarity in the exposition we state some of them in the following lemma.

Lemma 7.2.3 *The operator*

$$T : \mathcal{D}(A) \to \mathcal{D}(M), \quad x \mapsto \frac{1}{R}M^{-1}M^2x,$$

where M is given by (7.28), has selfadjoint extension $\bar{T}$ in X. Further, $\bar{T}$ is selfadjoint, *has (selfadjoint) compact resolvent and its eigenfunctions $(e_i)_{i\in\mathbb{N}}$ form a basis in X. Finally, its spectrum $\sigma(\bar{T})$ consists of* simple *eigenvalues $(\mu_i)_{i\in\mathbb{N}}$ such that*

$$\lim_{i\to\infty} \mu_i - \mu_{i+1} = \infty. \tag{7.31}$$

These statements can be found along the paper [62] as the authors proved that the Orr-Sommerfeld operator satisfies the assumptions of the main theorem [62, p.220]. In particular, it is proved [62, p. 224, Equation (20), $\delta=0$] that $\bar{T}$ is the unbounded operator given by the *inverse* of the selfadjoint compact operator

$$R_1 : X \to X, \quad [R_1x](y) = \frac{1}{R}\int_{-1}^{1}\left[\frac{\partial g_1}{\partial \tau}(y,\tau)\frac{\partial x}{\partial \tau}(\tau)\right]d\tau + \frac{1}{R}\alpha^2\int_{-1}^{1}[g_1(y,\tau)x(\tau)]\,d\tau,$$

where $g_1(y,\tau)$ is certain Green function such that both g_1 and $\frac{\partial g_1}{\partial \tau}$ are continuous; more exactly, it is the solution of [62, p. 223, Equation (15), $\delta=0$]

$$M^2g_1(y,\tau) = \delta(y-\tau), \quad g_1(\pm1,\tau) = \frac{\partial g_1}{\partial \tau}(\pm1,\tau) = 0,$$

with $\delta(y-\tau)$ denoting the usual Dirac function. The assertion (7.31) is in [62, p. 226; Equation (23), p.225], while the rest follows from the standard theory of selfadjoint compact operators, see [53, V.3.8].

Theorem 7.2.4 *The Orr Sommerfeld operator A (7.29) has closed extension $\bar{A}$ in X. Moreover, if the Reynolds number R and the wave number α are such that $\sigma(\bar{A})$ is a set of simple eigenvalues, then*

$$\bar{A} \in \mathcal{R}(X).$$

Proof: By (7.29), A can be expressed as

$$A = T + S_1,$$

where

$$T : \mathcal{D}(A) \to \mathcal{D}(M), \quad x \mapsto \frac{1}{R} M^{-1} M^2 x,$$
$$S_1 : \mathcal{D}(M) \to \mathcal{D}(M), \quad x \mapsto -\imath\alpha M^{-1}(UM - \frac{\partial^2 U}{\partial y^2})x.$$

By Lemma 7.2.3, $\bar{T}$ exists and satisfies the conditions of Theorem 7.2.2. Further, the existence of $\bar{S}_1$ is proved in [62, pp. 226]. It is the operator in $\mathcal{L}(X)$ given by [62, pp. 226, Equation (28)]

$$[\bar{S}_1 x](y) = -\imath\alpha U(y)x(y) + 2\imath\alpha \int_{-1}^{1} \left[g_2(y,\tau) \frac{\partial U}{\partial \tau}(\tau) \frac{\partial x}{\partial \tau}(\tau) \right] d\tau + 2\imath\alpha \int_{-1}^{1} \left[g_2(y,\tau) \frac{\partial^2 U}{\partial \tau^2}(\tau) x(\tau) \right] d\tau,$$

where $g_2(y,\tau)$ is the solution of [62, p. 223, Equation (16), $\delta = 0$]

$$M g_2(y,\tau) = \delta(y-\tau), \quad g_2(\pm 1, \tau) = 0,$$

with $\delta(y-\tau)$ denoting the usual Dirac function. Thus, $\bar{A}$ exists and is given by

$$\bar{A} = \bar{T} + \bar{S}_1.$$

Moreover, by Theorem 7.2.2, $\bar{A}$ has compact resolvent and, consequently, by Lemma 6.1.5, $\sigma(\bar{A})$ is a set of eigenvalues $(\lambda_i)_{i \in \mathbb{N}}$ with no other limit point than infinite. By assumption, these eigenvalues are all simple. In other words, $\bar{A}$ satisfies the first condition for being Riesz operator. It remains to show that the eigenfunctions of $\bar{A}$ form a Riesz basis: we have assumed that $\sigma(\bar{A})$ contains only simple eigenvalues. As a consequence, the eigenprojections $(Q_i)_{i \in \mathbb{N}}$ have rank one and can be written as

$$Q_i := \langle \cdot \, , \, \psi_i \rangle_X \, \phi_i, \quad i \in \mathbb{N},$$

where ϕ_i and ψ_i, $i \in \mathbb{N}$, are the eigenfunctions and their adjoints, respectively. Further on, by Theorem 7.2.2, there exists a $L \in \mathcal{L}(X)$ with $L^{-1} \in \mathcal{L}(X)$ such that the eigenprojections $(Q_i)_{i \in \mathbb{N}}$ of $\bar{A}$ can be indexed in such a way that Equation (7.30) holds. Moreover, since the eigenvalues are all simple, $m = 0$, $n = 0$, and we obtain for every $x \in \mathcal{D}(A)$ the following equations:

$$Q_i x = \langle \, x \, , \, \psi_i \, \rangle_X \, \phi_i = L P_i L^{-1} x = \langle \, L^{-1} x \, , \, e_i \, \rangle_X \, L e_i = \langle \, x \, , \, (L^{-1})^* e_i \, \rangle_X \, L e_i, \quad i \in \mathbb{N},$$

where $(e_i)_{i \in \mathbb{N}}$ denote the eigenfunctions of $\bar{T}$. Hence,

$$\phi_i = L e_i, \quad i \in \mathbb{N}.$$

By Lemma 7.2.3, the eigenfunctions $(e_i)_{i \in \mathbb{N}}$ form an orthogonal basis in X. Thus, $(\phi_i)_{i \in \mathbb{N}}$ is similar to a Riesz basis and, using Lemma 6.3.4, we see that $(\phi_i)_{i \in \mathbb{N}}$ forms a Riesz basis in X. We conclude that $\bar{A} \in \mathcal{R}(X)$. □

Remark 7.2.5 The obvious bound for $\sigma(A) = \sigma(\bar{A})$ given by

$$\lambda_i \in \{s \in \mathbb{C}, \ |s - \mu_i| < \|\bar{S}_1\|_{\mathcal{L}(X)}\}, \quad i \in \mathbb{N},$$

where as before $(\mu_i)_{i\in\mathbb{N}}$ denotes the spectrum of T, shows that the eigenvalues of A are simple at least for R small enough. In our calculations, the eigenvalues of A have been always simple. Moreover, the simplicity of $\sigma(A)$ is usually *assumed* [66] in studies related to the Orr-Sommerfeld operator. However, this property of $\sigma(A)$ seems *not* to hold for every R [36].

Remark 7.2.6 The completeness of the eigenfunctions of Orr-Sommerfeld operator belongs to the well documented facts of the theory of hydrodynamical stability. On the other hand, their property of being *basis* of X has not been so widely known [39]. To the best of our knowledge, this is the first time that this fact appears in the literature.

Remark 7.2.7 In the case of *nearly parallel shear flows* [38], the corresponding "Orr-Sommerfeld" operator has the form $T + S_1$, where T is the same as before and the "new" $\bar{S}_1$ is a T*-compact* operator [53, IV.1.3]. It follows that Theorem 7.2.4 is true also in this case. Note that the completeness of the eigenfunctions have already been established in [38].

7.2.3 Unstructured perturbations

A recent school of thought has brought new life into the problem on the discrepancies between experiment and theory. The reasoning is as follows. The Orr Sommerfeld operator is highly non-normal. Thus, there is a potential for *transient growth* of initially small perturbations even if all the eigenmodes decay exponentially. Interesting references where this point of view is developed are [30], [78], [35], [65] and [21] among others. The most important paper in our context is the paper by Reddy et.al. [66], where pseudospectra ideas are explicitly used.
In fact, in the paper as a measure for the transient behavior, Reddy et.at. [66] studied the pseudospectra (Figure 7.4) and numerical range of A. They gave *lower bounds* for the critical Reynolds number R_c ($R_c \geq 82.2$) which agree with former theoretical and experimental results. As usual, they restricted their analysis to the even modes of A. A detailed discussion of Figure 7.4 is given later.
In order to calculate the pseudospectra, because the norm of $R(s, A)$ can not be computed explicitly, the authors of [66] approximated $s \to \|R(s)\|_{\mathcal{L}(X)}$ using the projection scheme of [37]. This method, which was originally designed for the calculation of eigenvalues and eigenfunctions of A, is based on collocation in Chebyshev points of expansions in Chebyshev polynomials with explicit enforcement of the boundary conditions. In the terminology of Chapter 6, the authors generated discrete operators $(A_N)_{N\in\mathbb{N}}$ and approximate the pseudospectra of A with the pseudospectra of $(A_N)_{N\in\mathbb{N}}$. We give some details on this scheme below.
It is important to know whether these approximations generate correct pictures. In [66], at least in a strict mathematical sense, this convergence was not proved. The authors showed it experimentally: they calculated the contours using two different discretisation schemes and observed that for large N the contours become identical and stationary, see Figure 7.4.
It should be noted that the approximation in this case is much more difficult and involved than in the case of Section 7.1. The reason is that the Orr-Sommerfeld operator is an *integro-differential* operator and thus, projection methods are difficult to apply in a straightforward manner.

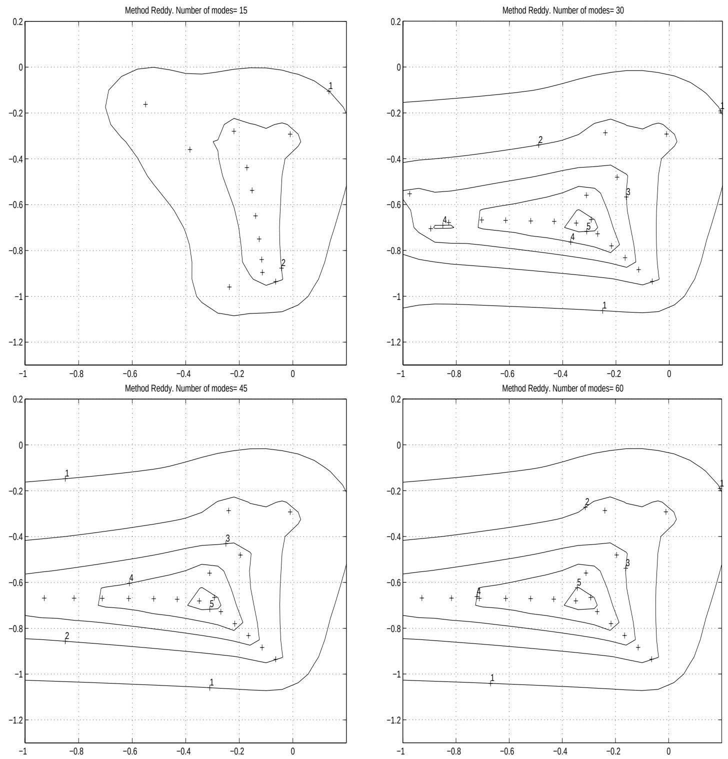

Figure 7.4: Unstructured spectral value sets of the Orr-Sommerfeld operator for $R = 3000$, $\alpha = 1$, $\rho = 10^{-1}, \ldots, 10^{-8}$ and $N = 15, 30, 45, 60$.

7.2.4 Approximation scheme

Our immediate aim is to work out an approximation scheme which can be useful for the approximation of the spectral value sets related to the Orr-Sommerfeld operator. Our approach is to combine the results of Section 5.3.4 with Proposition 6.3.10. The description of the approximation scheme takes place in several steps.

Spaces and "augmented" operators. The first step is to introduce the spaces

$$W := H^4([-1,1];\mathbb{C}), \quad V := L^2([-1,1];\mathbb{C}), \quad V_a := V \times \mathbb{C}^4.$$

As in Section 5.3.4, we consider the operators (see (7.29))

$$M_0 : W \to V, \quad x \mapsto \left(\frac{\partial^2}{\partial y^2} - \alpha^2\right) x,$$

$$S_0 : W \to V, \quad x \mapsto \left[\frac{1}{R} M_0^2 - \imath\alpha \left(U M_0 - \frac{\partial^2 U}{\partial y^2}\right)\right] x,$$

and the (boundary condition) operator

$$L : W \to \mathbb{C}^4, \quad x \mapsto \operatorname{col}\left(x(1),\, \frac{\partial x}{\partial y}(1),\, x(-1),\, \frac{\partial x}{\partial y}(-1)\right). \tag{7.32}$$

We introduce also the “augmented” operators

$$M_a := \begin{pmatrix} M_0 \\ L \end{pmatrix} : W \to V_a, \quad S_a := \begin{pmatrix} S_0 \\ \lambda_a L \end{pmatrix} : W \to V_a, \tag{7.33}$$

where $\lambda_a \in \rho(T)$ is a *nonzero* constant with a role to be made clear later.

Approximating sequence. Our plan is to use Proposition 6.3.10. Thus, we must find an operator sequence $(\mathcal{A}_N)_{N\in\mathbb{N}}$ such that its eigenelements approximate the eigenelements of A, see the assumptions of Proposition 6.3.10. In the solution of this problem we shall use the fact that the eigenvalue problems

$$Sx = \lambda Mx, \quad x \in \mathcal{D}(A)$$

and

$$Ax := M^{-1}Sx = \lambda x, \quad x \in \mathcal{D}(A),$$

actually coincide.
The main point is Lemma 7.2.8. In its statement and in the sequel we shall use the notation $\sigma(R,S)$ for the spectrum of a generalised eigenvalue problem $Rx = \lambda Sx$.

Lemma 7.2.8 $\sigma(S_a, M_a) = \sigma(A) \cup \lambda_a$. *The eigenvectors of* (S_a, M_a) *and* A *corresponding to* $\sigma(A)$ *coincide as well.*

Proof: Similar functional-theoretical arguments to those used for the Orr-Sommerfeld operator [19, pp. 156] show that the spectrum $\sigma(S_a, M_a)$ has only eigenvalues. Moreover, it is clear that $\sigma(A) \subset \sigma(S_a, M_a)$. We must show that (S_a, M_a) has no spectrum other than $\sigma(A) \cup \lambda_a$. Indeed, consider the equation

$$S_a\phi_a = sM_a\phi_a, \quad \phi_a \in W, \quad s \in \rho(A). \tag{7.34}$$

We must find out for which $s \in \rho(A)$, this equation has *unique solution* $\phi_a = 0$. We proceed as follows: Equation (7.34) means

$$S_0\phi_a - sM_0\phi_a = 0, \tag{7.35}$$
$$(\lambda_a - s)L\phi_a = 0. \tag{7.36}$$

1. If $s = \lambda_a$, (7.36) holds for every $\phi_a \in W$ and thus (7.35) has nonzero solutions, because it is a differential equation with no boundary conditions. It follows $\lambda_a \in \sigma(S_a, M_a)$.
2. If $s \neq \lambda_a$, Equation (7.36) means $\phi_a \in \mathcal{D}(A)$. Since $s \in \rho(A)$, $\phi_a = 0$ is the *unique* solution of (7.35). We conclude that $s \in \rho(S_a, M_a)$.

The proof is complete. □

Let us continue with the exposition. Clearly, it holds that

$$S_a \in \mathcal{L}(W, V_a), \quad S_a^{-1} \in \mathcal{L}(V_a, W), \quad M_a \in \mathcal{K}(W, V_a).$$

Thus

$$S_a - sM_a \in \mathcal{F}_0(W, V_a), \quad \forall\, s \in \mathbb{C}.$$

Suppose that we are given a sequence of *orthogonal* projections $(\pi_N^W)_{N\in\mathbb{N}}$ such that

$$\iota_N^W \pi_N^W \xrightarrow[W]{s} I_W. \tag{7.37}$$

Consider also

$$\pi_N^{V_a} := S_a \,\iota_N^W \pi_N^W S_a^{-1}, \quad N \in \mathbb{N}. \tag{7.38}$$

Then, by the proof of Proposition 5.3.24 and Remark 5.3.21, it follows that

$$S_{aN} - sM_{aN} \xrightarrow[\mathcal{L}(W,V_a)]{r} S_a - sM_a, \quad s \in \mathbb{C}, \tag{7.39}$$

where

$$M_{aN} := \iota_N^{V_a} \mathcal{M}_a \pi_N^W, \quad \mathcal{M}_{aN} := \pi_N^{V_a} M_a \,\iota_N^W, \quad N \in \mathbb{N}, \tag{7.40}$$

$$S_{aN} := \iota_N^{V_a} \mathcal{S}_a \pi_N^W, \quad \mathcal{S}_{aN} := \pi_N^{V_a} S_a \,\iota_N^W, \quad N \in \mathbb{N}. \tag{7.41}$$

Moreover, by Remark 5.3.19, the *adjoint* operators converge regularly as well.
Now, using Proposition 5.3.28, we obtain that the eigenelements of the problem

$$\mathcal{S}_{aN} x_{aN} = \lambda_N \mathcal{M}_{aN} x_{aN}$$

approximate those of $S_a x = \lambda M_a x$ and hence, by Lemma 7.2.8, those of $Sx = \lambda Mx$.
Finally, we note that *if* the matrices $(\mathcal{M}_{aN})_{N\in\mathbb{N}}$ are invertible for N sufficiently large, then one can define a sequence $(\mathcal{A}_N)_{N\in\mathbb{N}}$ given by

$$\mathcal{A}_N := \mathcal{M}_{aN}^{-1} \mathcal{S}_{aN}, \quad N \in \mathbb{N}, \tag{7.42}$$

which is the desired operator sequence of Proposition 6.3.10.

Discretisation procedure We have learned in the previous section that all we need here for a satisfactory application of our approximation results is a sequence of orthogonal projections $(\pi_N^W)_{N\in\mathbb{N}}$ such that (7.37) holds and such that the projections (7.38) can be calculated.
However, there is another approximation scheme which, due to its accurate approximations of the *eigenelements* of A at a comparatively low computational costs [60], has become standard in calculations related to the Orr-Sommerfeld operator. Note that approximation of the eigenelements is the main requirement in Proposition 6.3.10. This scheme, which we call "Herbert's method" [37], belongs to the class of *spectral methods* [11], and does not use (7.38).
Herbert's method is based on the following objects. First, as projections $(\pi_N^W)_{N\in\mathbb{N}}$ one takes *orthogonal projections* (with respect to the inner product in W) onto the span of polynomials of degree $\leq N$. Further, let

$$V_N := W_N, \quad N \in \mathbb{N},$$

and let us denote by $(\pi_N^V)_{N\in\mathbb{N}}$ the sequence of operators which map every *continuous* function in $[-1,1]$ onto its *interpolatory* N-th degree polynomial at the Chebyshev points

$$y_j = \cos(\frac{j\pi}{N}), \quad j = 0,\dots,N.$$

The properties of the elements of $(\pi_N^V)_{N\in\mathbb{N}}$ are quite complicated. Although they are idempotent $((\pi_N^V)^2 = \pi_N^V)$, they are not projectors in the usual sense. The reason is that they are not closed (and hence not bounded) neither in $V = L^2([-1,1];\mathbb{C})$ nor in $L_r^2([-1,1];\mathbb{C})$ with weight $r(y) = 1/\sqrt{1-y^2}$ [1, Exercise 15.20]. Actually, one has [74, Theorem 14.3.2]

$$\pi_N^W \underset{L_r^2([-1,1];\mathbb{C})}{\overset{s}{\longrightarrow}} \iota\,,$$

where

$$\iota : C([-1,1];\mathbb{C}) \to L_r^2([-1,1];\mathbb{C})$$

is the canonical (continuous) embedding. Now, we may introduce the sequence

$$\pi_N^{V_a} = \begin{pmatrix} \pi_N^V \\ L \end{pmatrix}, \quad N \in \mathbb{N}. \tag{7.43}$$

Note that

$$\dim \mathrm{Rg}(\pi_N^{V_a}) = N+5, \quad \dim \mathrm{Rg}(\pi_N^W) = N+1, \quad N \in \mathbb{N}. \tag{7.44}$$

Herbert's method, used also in [37] and [66], chose $(\pi_N^W)_{N\in\mathbb{N}}$, $(\pi_N^{V_a})_{N\in\mathbb{N}}$ as main building block of their discretisation procedure. In order to calculate the corresponding finite dimensional operators (see Section 6.4.4) one must fix basis in W_N and V_N; here Chebyshev polynomials are the usual choice. Let us give some details. Consider the expansion

$$x(y) = \sum_{i=0}^{N} x_i T_i(y) \in W_N, \quad N \in \mathbb{N},$$

where $\{x_i\}$ are the expansion coefficients and $\{T_i\}$ are the first $N+1$ Chebyshev polynomials. The main part of the discretisation is to find the family of matrices

$$\{D^{(l)}\} \in \mathbb{C}^{(N+1)\times(N+1)}$$

which convert the expansion coefficients $\{x_i\}$ to the values $\{x^{(l)}(y_j)\}$, $l = 1,\dots,4$. The matrices are defined by

$$D_{ji}^{(l)} = T_i^{(l)}(y_j), \quad j,i = 0,1,\dots,N. \tag{7.45}$$

Having these matrices, the finite dimensional operators $(\mathcal{S}_{aN})_{N\in\mathbb{N}}$, $(\mathcal{M}_{aN})_{N\in\mathbb{N}}$ can be computed using an involved, although straightforward, procedure.
Due to the treatment of the boundary conditions (see also (7.44)) the *matrix representations* of $(\mathcal{S}_{aN})_{N\in\mathbb{N}}$ and $(\mathcal{M}_{aN})_{N\in\mathbb{N}}$ are in $\mathbb{C}^{(N+5)\times(N+1)}$. At first glance this is a problem because we need the inverses of the *operators* $\mathcal{M}_{aN}$. The dilemma is solved as follows. By Proposition 5.3.17, $\mathcal{S}_{aN}$, $\mathcal{M}_{aN} \in \mathcal{F}_0(W_N, V_{aN})$, $N \in \mathbb{N}$. Thus, all these matrices *have rank* $N+1$. In order to obtain square matrices one neglects the rows 0, 1, $N-1$, and N of $\mathcal{S}_{aN}$ and $\mathcal{M}_{aN}$ which, from "heuristic reasons", contain redundant information. Indeed, these lines correspond to collocation points near the boundaries and this information is already represented in the matrices by the operator L (7.32).

Another important point is the constant λ_a given in (7.33) which is set equal to a complex number far away from the compact set K where the pseudospectrum must be calculated. Finally, energy norm calculations are achieved by means of "weighting matrices" B so that

$$\|x\|_X = \|Bx\|_{L^2([-1,1];\mathbb{C})}\,.$$

The corresponding operator norms can also be calculated using this approach. Appendix A of [66] gives a detailed description of the calculation of weight matrices B. See also [37] and [64] for more details on the whole Herbert's method.

The main question now is whether the spectral value sets of the finite dimensional operators associated to Herbert's method converge to the spectral value sets of the Orr-Sommerfeld operator. As we mentioned before, to the best of our knowledge there is not formal proof of this fact. Vainikko [82, Section 8] proved that a slight modification of the projections $(\pi_N^W)_{N\in\mathbb{N}}$, $(\pi_N^{V_a})_{N\in\mathbb{N}}$, namely by projecting W onto W_{N+5} rather than on W_N, is able to provide regular convergence when applied to ordinary differential operators of the type of Section 5.3.4. Note that with this choice the matrix representations of the finite dimensional operators are automatically square. We have not been able to prove the same fact for the Herbert's method, whereas the main difficulty is (7.44) and the elimination of rows to be made later. Nevertheless, since the eigenelements of A are successfully approximated by this scheme, see [60] [37], we believe that regular convergence (7.39) does take place here as well. We stress that we *could* use other sequences of projections so that we would then be in a position to *prove* regular convergence; for example, Vainniko's scheme. However, we shall not do that because former experience with those schemes shows that, *numerically speaking*, those "better" projections do not perform as perfectly as Herbert's method [37] which, nowadays, is the first choice in hydrodynamical stability [70], [18].

Calculation of $\sigma(A, I_X, I_X, \rho)$. The compact K of interest is the region of the complex plane represented in Figure 7.4. As in [66], we take $\lambda_a = -200$. Since we know that the eigenelements

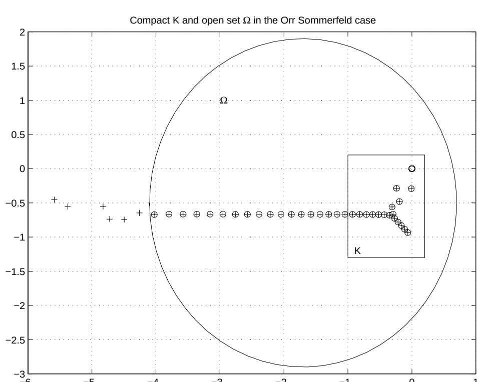

Figure 7.5: Sets K and Ω in the Orr-Sommerfeld case together with $\sigma(\mathcal{A}_N)$ and $\sigma(\mathcal{A}_N\mathcal{P}_{N\Omega})$.

(spectrum, eigenvectors and the corresponding biorthogonal vector) converge, we may carry out the calculation of $\sigma(A, I_X, I_X, \rho)\cap K$ (Figure 7.4) by the scheme of Proposition 6.3.10. In other words, one approximates the norm of $R(s, A)$ by the norm of $R(s, \mathcal{A}_N)\mathcal{P}_{N\Omega}$, where we have used the notations of Section 6.3. As open domain Ω we have chosen

$$\Omega := \{s \in \mathbb{C} \; : \; |s + 1.7 + 0.5i| < 2.4\}, \tag{7.46}$$

see Figure 7.5.
With the help of the `Matlab` code mentioned in Chapter 2 we were able to calculate the eigenprojections $\mathcal{P}_{N\Omega}$ accurately. Indeed, in Figure 7.5 we represented $\sigma(\mathcal{A}_N\mathcal{P}_{N\Omega})$ (symbol 'o') together with $\sigma(\mathcal{A}_N) \cap \Omega$ (symbol '+'). As it was expected, one observes that

$$\sigma(\mathcal{A}_N\mathcal{P}_{N\Omega}) = \sigma(\mathcal{A}_N) \cap \Omega \cup \{0\}.$$

We conclude that the pictures of Figure 7.4 are correct.

7.2.5 Structured perturbations

In order to explain the discrepancies between spectral analysis of A and the experimental results we propose a new approach. It will stress the usefulness of an analysis based on spectral value sets.
Let us illustrate our ideas by means of Figure 7.4. In fact, Reddy's graphs show that the spectrum $\sigma(A)$ is extremely *sensitive* to perturbations and that the sets $\sigma(A, I_X, I_X; \rho)$ cross the imaginary axis (with small ρs) for Reynolds numbers much lower that R_c. This suggests that, due to *small perturbations*, the Poiseuille flow may become unstable "earlier" as predicted by the eigenvalues alone.
Moreover, we believe that these results can be improved. In fact, an adequate representation of the effect of nonlinearities and other uncertainties can be better carried out using *structured* perturbations. As illustration we shall consider the following ones. They take into account the nonlinearities which are present in Equation (7.23). We define (see (7.26))

$$U = X, \quad \underline{X} = \mathcal{D}(M) \text{ equipped with the graph norm;} \quad Y = L^2([-1,1];\mathbb{C}) \tag{7.47}$$

and (see (7.28))

$$D : U \to X, \quad u \mapsto P_\Omega u, \tag{7.48}$$

$$E : \underline{X} \to Y, \quad x \mapsto Mx, \tag{7.49}$$

where Ω is the domain (7.46) and P_Ω is the corresponding spectral projection. One sees immediately that

$$\begin{aligned} &\mathcal{D}(A) \subset \underline{X} \subset X \text{ with continuous dense injections} \\ &D : U \to X \text{ is compact} \\ &E : \underline{X} \to Y \text{ is bounded.} \end{aligned}$$

This is a framework which satisfies the requirements given in Section 4.1.
We can not calculate $\|E(sI - A)^{-1}D\|_{\mathcal{L}(U,Y)}$ explicitly and the use of approximations is obligatory. In the previous section we have already considered approximations to A and D. The new

operator to be approximated is (7.49) and it is clear that we should use the operators $(\mathcal{M}_{aN})_{N\in\mathbb{N}}$ introduced before. Note that we already used these matrices in the construction of $(\mathcal{A}_N)_{N\in\mathbb{N}}$.
We shall not try to *state* convergence results as we did in the previous section, see Proposition 7.1.2. The reason is that the approximation scheme we are using here, in particular, the sequence (7.43) and especially the "elimination of rows" explained afterwards, do not fit in our main convergence results, for example Proposition 6.4.3.

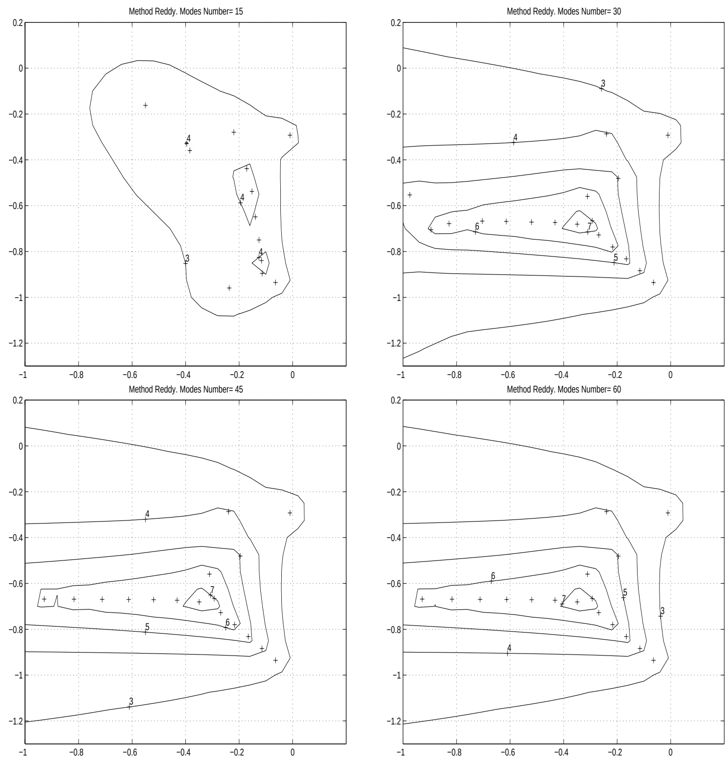

Figure 7.6: Structured spectral value sets of the Orr-Sommerfeld operator for $R = 3000$, $\alpha = 1$, $\rho = 10^{-3}, \dots, 10^{-8}$ and $N = 15, 30, 45, 60$.

In Figure 7.6 we give some computational results for such perturbations. These pictures have been obtained using the SH algorithm of Chapter 3. Note that they become stationary for $N \geq 45$.
Let us compare these new pictures with the ones, corresponding to unstructured perturbations, represented in Figure 7.4. Both figures represent the results for the case $R = 3000$, $\alpha = 1$. The crosses represent points in $\sigma(A)$ and the contours are the boundaries of the spectral value sets for different values of ρ. These are $10^{-1}, \dots, 10^{-8}$. The relevant sets in the two figures are marked with the numbers $1, \dots, 8$ etc. Around some points in $\sigma(A)$ the contours are not visible for the scale used, indicating that they hardly move for the above class of perturbations. As ρ increases the sets associated with a given eigenvalue expand and then merge with other sets associated with other eigenvalues.
The graphs are qualitatively similar and show the high sensitivity of $\sigma(A)$. Nevertheless, the quantitative differences between the two cases is striking. For example, in the first picture of

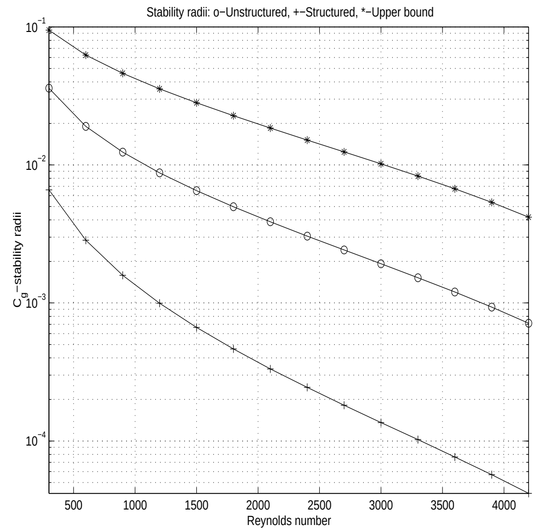

Figure 7.7: Stability radii for different values of R.

Figure 7.4, we see that perturbations at level $\rho \approx 10^{-2}$ may destabilise the system, whereas in the structured case destabilisation occurs at the level $\rho < 10^{-4}$. One sees that as ρ increases the difference between the two sets becomes even larger.

Figure 7.7 gives more insight into the problem. It represents the C_g-stability radii (see Definition 4.2.15) for both unstructured and structured perturbations and different values of R. Here C_g is just the left half of the complex plane. In order to calculate the stability radii we computed the H_∞-norm of the transfer function $R(s, \mathcal{A}_N)\mathcal{P}_{N\Omega}$ and $\mathcal{M}_{aN}R(s, \mathcal{A}_N)\mathcal{P}_{N\Omega}$, respectively. The computations were carried out using the function `normhinf` of the computing package "Matlab" and approximations of order $N = 60$. These curves show that the robustness of $\sigma(A)$ diminishes when R becomes larger and that our structured perturbations have stronger destabilising effects than unstructured perturbations. We have included in this picture the (obvious) upper bound for $r(\mathcal{A}_N, I, I, C_g)$ given by the quantity

$$d^*(\sigma(\mathcal{A}_N), \partial C_g) = \sup_{\lambda \in \sigma(\mathcal{A}_N)} \operatorname{Re} \lambda.$$

Since this bound is exact for *normal* operators, this graph shows that the non-normality of A increases *exponentially* with R.

Moreover, a detailed analysis of Figure 7.7 shows "mysterious" changes in the slopes of the curves at Reynolds numbers $500 \leq R \leq 3000$. Are there important physical reasons behind of this behavior? Future investigations should clarify this question.

We believe that these pictures prove that spectral value sets analysis can shed light into problems of hydrodynamical stability.

Appendix A

Convergence of level curves

Our aim here is to show that convergence of functions in the chordal metric implies convergence of the corresponding level curves in the sense of Hausdorff distance. We begin with some definitions.

Definition A.0.9 Let K_1 and K_2 be nonempty bounded closed subsets of a metric space K. Then the *Hausdorff metric* d_H is defined as

$$d_H(K_1, K_2) := \max\{d^*(K_1, K_2),\ d^*(K_2, K_1)\},$$

where

$$\begin{aligned} d^*(A, B) &= \sup_{x \in A} d(x, B) \\ d(x, B) &= \inf_{y \in B} d(x, y) \end{aligned}$$

and d is the metric on K.

Let $\hat{\mathbb{C}} := \mathbb{C} + \cup\{\infty\}$. Then for two points a, $b \in \hat{\mathbb{C}}$, we denote by $\kappa(a, b)$ the chordal distance between a and b:

$$\begin{aligned} \kappa(a, b) &= \frac{|a - b|}{\sqrt{(1 + |a|^2)(1 + |b|^2)}}, \qquad (a,\ b \text{ both finite}) \\ \kappa(a, \infty) &= \kappa(\infty, a) = \frac{1}{\sqrt{1 + |a|^2}}, \\ \kappa(\infty, \infty) &= 0. \end{aligned}$$

Further, let $\hat{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{+\infty\}$. The notation $C(K; \hat{\mathbb{R}}_+)$ will be used for the set of continuous functions from the metric space K into $\hat{\mathbb{R}}_+$.

Definition A.0.10 Suppose that $f, g \in C(K; \hat{\mathbb{R}}_+)$. The *chordal distance* between f and g is defined by

$$d_\kappa(f, g) = \sup_{s \in K}\{\kappa(f(s), g(s))\}.$$

We say that $(f_N)_{N\in\mathbb{N}}$ converges on K to f in the sense of the chordal metric, denoted by $f_N \underset{K}{\rightarrow} f$, if

$$\lim_{N\to\infty} d_\kappa(f_N, f) = 0.$$

We shall also use the concept of uniform convergence. We say that $(f_N)_{N\in\mathbb{N}}$ converges on K uniformly to f, to be denoted $f_N \rightrightarrows f$, if

$$\lim_{N\to\infty} \max_{s\in K} |f_N(s) - f(s)| = 0.$$

We continue with a simple lemma whose proof appears in [34]. In its enunciation and in the rest of this chapter we shall use the notations f^{-1} and f_N^{-1} for $\frac{1}{f}$ and $\frac{1}{f_N}$, respectively.

Lemma A.0.11 *Let $f, (f_N)_{N\in\mathbb{N}} \in C(K;\hat{\mathbb{R}}_+)$ and $0 < r < \infty$. If $f_N \underset{K}{\rightarrow} f$, then*

1. *$f_N \rightrightarrows f$ on $\{s \in K : f(s) \le r\}$.*
2. *$f_N^{-1} \rightrightarrows f^{-1}$ on $\{s \in K : f(s) \ge r\}$.*

Conversely, if there exist A, $B \subset K$, $A \cap B \ne \emptyset$, such that

$$f_N \rightrightarrows f \text{ on } A \quad \text{and} \quad f_N^{-1} \rightrightarrows f^{-1} \text{ on } B,$$

then

$$f_N \underset{A\cap B}{\rightarrow} f.$$

This lemma has the following useful corollary.

Corollary A.0.12 *Let $K \subset \hat{\mathbb{C}}$ be compact. Then $f_N \underset{K}{\rightarrow} f$ iff for every $s \in K$ there exists an ϵ-neighborhood $V_\epsilon(s)$ of s such that either $(f_N)_{N\in\mathbb{N}}$ or $(f_N^{-1})_{N\in\mathbb{N}}$ converge uniformly on $V_\epsilon(s)$ to f or f^{-1}, respectively.*

Definition A.0.13 Let $K \subset \hat{\mathbb{C}}$, $f \in C(K;\hat{\mathbb{R}}_+)$. We say that $\rho > 0$ is *regular* for f, if for any $s \in K$ such that $f(s) = \rho$, and any ϵ-neighborhood $V_\epsilon(s)$ of s, there exists $s_l, s_b \in V_\epsilon(s) \cap K$ such that $f(s_l) < \rho$ and $f(s_b) > \rho$.

Roughly speaking, ρ is regular if there are no local extrema in the set $\{s \in K, \quad f(s) = \rho\}$. For example, if f is non constant and meromorphic in an open set $K \subset \mathbb{C}$, the numbers $\rho > 0$ such that there exist $s \in K$ with $|f(s)| = \rho$ are regular for $|f|$.
The main result in this chapter is the theorem below. It makes apparent the relationship between convergence in the sense of the chordal metric and convergence of the corresponding level curves in the sense of the Hausdorff metric.

Theorem A.0.14 *Let K be a compact and locally connected subset of a metric space. Suppose that f, $(f_N)_{N\in\mathbb{N}}$ are elements of $C(K;\hat{\mathbb{R}}_+)$ such that*

$$f_N \underset{K}{\rightarrow} f.$$

Furthermore, suppose that $\rho > 0$ is regular for f. Let D_ρ and D_ρ^N, be the ρ-level curve in K of f and f_N, respectively, i.e.,

$$D_\rho = \{s \in K : f(s) = \rho\}, \quad D_\rho^N = \{s \in K : f_N(s) = \rho\}.$$

Then

$$\lim_{N\to\infty} d_H(D_\rho^N, D_\rho) = 0. \tag{A.1}$$

Proof: Introduce the notation

$$S_r = \{s \in K : f(s) \geq r\}, \quad r > 0,$$

where by definition the set $\{s \in K : f(s) = \infty\}$ is contained in S_r. We also note that the regularity condition for ρ implies $\partial S_\rho = D_\rho$.
We will first show that for every $\epsilon > 0$ there exists $N_1 \in \mathbb{N}$ such that the relation $N > N_1$ implies

$$d^*(D_\rho^N, D_\rho) < \epsilon. \tag{A.2}$$

Let V_ϵ be an ϵ-neighborhood of D_ρ:

$$V_\epsilon = \{s \in K;\ d(s, s_0) < \epsilon \text{ for some } s_0 \in D_\rho\}.$$

The set $S_\rho \cup V_\epsilon$ is open. Thus $E := K \setminus (S_\rho \cup V_\epsilon)$ is closed. It follows that $\max_{s\in E} f(s) < \rho$. Since $f_N \rightrightarrows f$ on E (Lemma A.0.11), there exists $N_1 \in \mathbb{N}$ such that for $N > N_1$, $f_N(s) < \rho$ for all $s \in E$. Hence,

$$D_\rho^N \subset S_\rho \cup V_\epsilon, \quad N > N_1. \tag{A.3}$$

We have now two cases:

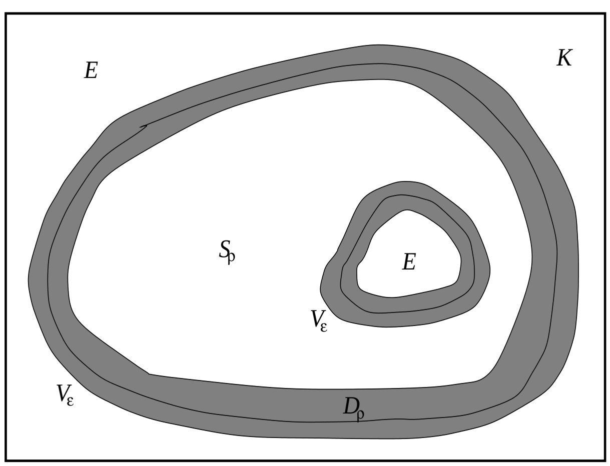

Figure A.1: Sets in the proof of Theorem A.0.14. D_ρ is represented by the thick line and V_ϵ by the shaded zone.

1. First case: $\partial V_\epsilon \cap S_\rho = \emptyset$. First, we prove

$$S_\rho \subset V_\epsilon. \tag{A.4}$$

In fact, suppose (A.4) does not hold. Then there exists a $s_0 \in S_\rho$ with $s_0 \notin V_\epsilon$. Since $\partial V_\epsilon \cap S_\rho = \emptyset$, we have $s_0 \notin \partial V_\epsilon$, i.e., $s_0 \in V_\epsilon^C$, where V_ϵ^C is the (open) complement in K of $\mathrm{cl}\,(V_\epsilon)$. Let us denote by $S_\rho^{s_0}$ the connected component of s_0 in $\{s \in K;\ f(s) > \rho\}$. Note that $S_\rho^{s_0}$ exists because K is locally connected. Moreover, since $\{s \in K;\ f(s) > \rho\}$ is open in K, $S_\rho^{s_0}$ is open in K as well. Now, it is clear that

$$S_\rho^{s_0} \cap V_\epsilon^C \neq \emptyset$$

and is open in K. Moreover, since $\partial S_\rho^{s_0} \subset D_\rho \subset V_\epsilon$, one has that the set

$$S_\rho^{s_0} \cap V_\epsilon \neq \emptyset$$

and is open in K too. One also notes

$$V_\epsilon \cap V_\epsilon^C = \emptyset, \quad K = \partial V_\epsilon \cup V_\epsilon \cup V_\epsilon^C.$$

The last three relationships imply that

$$S_\rho^{s_0} = \left(S_\rho^{s_0} \cap V_\epsilon\right) \cup \left(S_\rho^{s_0} \cap V_\epsilon^C\right)$$

The sets V_ϵ and V_ϵ^C are open sets in K with empty intersection. It follows that $S_\rho^{s_0}$ is not connected. Thus, we have a contradiction and (A.4) holds.

Expression (A.4) together with (A.3) imply that $D_\rho^N \subset V_\epsilon$ and, consequently, that (A.2) holds.

2. Second case: $\partial V_\epsilon \cap S_\rho \neq \emptyset$. Let us define $\tilde{S} := S_\rho \setminus V_\epsilon \neq \emptyset$ and $\tilde{\rho} := \min_{s \in \tilde{S}} f(s)$. Obviously, $\tilde{S} \subset S_{\tilde{\rho}}$. Moreover, since $\partial S_\rho = D_\rho \subset V_\epsilon$, we have that $\tilde{\rho} > \rho$. We apply now Lemma A.0.11. It tells us that $f_N^{-1} \rightrightarrows f^{-1}$ on $\tilde{S}$. Then $f_N^{-1} < \frac{1}{\rho}$ on $\tilde{S}$ for all N sufficiently large. Hence,

$$D_\rho^N \subset V_\epsilon \cup (K \setminus S_\rho)$$

for all N sufficiently large. Combining this with (A.3) we obtain $D_\rho^N \subset V_\epsilon$ for all N sufficiently large and consequently (A.2).

Our proof is complete if we show that $d^*(D_\rho, D_\rho^N) < \epsilon$ for N large enough.
Let $s_0 \in D_\rho$. Since ρ is regular, in each $\epsilon/2$-neighborhood $V_{\epsilon/2}(s_0)$ of s_0 in K, one can find points s_{0l}, s_{0b} in $V_{\epsilon/2}(s_0)$ such that $f(s_{0l}) < \rho$ and $f(s_{0b}) > \rho$. Thus, by Lemma A.0.11, there exists an $N(s_0)$ with the property that for all $N > N(s_0)$ one can find points s_{l_N}, s_{b_N} in $V_{\epsilon/2}(s_0)$ for which $f_N(s_{l_N}) < \rho$ and $f_N(s_{b_N}) > \rho$. By continuity and the connectness of $V_{\epsilon/2}(s_0)$, the equations $f_N(s) = \rho$, $N > N(s_0)$ have a root s_N in $V_{\epsilon/2}(s_0)$. Hence,

$$d^*(\{s_N\}, D_\rho^N) < \epsilon \text{ for all } s \in V_{\epsilon/2}(s_0) \cap D_\rho,\ N > N(s_0).$$

Because D_ρ is compact, this implies

$$d^*(D_\rho, D_\rho^N) < \epsilon$$

for all N sufficiently large. The last inequality together with (A.2) gives

$$d_H(D_\rho, D_\rho^N) \leq \epsilon.$$

The proof is complete. □

By Corollary A.0.12, uniform convergence implies chordal convergence. Thus, we have

Corollary A.0.15 *Let K be a compact and locally connected subset of a metric space. $f, (f_N)_{N\in\mathbb{N}} \in C(K; \hat{\mathbb{R}}_+)$, $f_N \rightrightarrows f$, $\rho > 0$ be regular for f. Let D_ρ and D_ρ^N be the ρ-level curves in K of f and f_N, respectively. Then $D_\rho^N \underset{N\to\infty}{\longrightarrow} D_\rho$ in the Hausdorff metric. Moreover, with some additional effort it can be proved that* (A.1) *remains valid if we replace D_ρ and D_ρ^n by S_ρ and the corresponding S_ρ^n, respectively.*

The proof of Theorem A.0.14 is mainly due to Ribalta [69].

Bibliography

[1] Krasnoselski. M. A. *Näherungsverfahren zur Lösung von Operatorgleichungen.* Akademie Verlag Berlin, 1973.

[2] E. L. Allgower and K. Georg. Continuation and path following. *Acta Numerica*, pages 1–64, 1993.

[3] B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-Linear Parametric Optimization.* Akademie-Verlag. Berlin, 1982.

[4] B. Bernhardsson, A. Rantzer, and L. Qiu. Real perturbation values and real quadratic forms in a complex vector space. *Linear Algebra Appl.*, 278:131–154, 1998.

[5] A. Böttcher. Pseudospectra and singular values of large convolution operators. *J. Integral Equations and Applications*, 6:267–301, 1994.

[6] A. Böttcher, S. Grudsky, and B. Silbermann. Norm of inverses, spectra, and pseudospectra of large truncated wiener-hopf operators and toeplitz matrices. *New York J. Math.*, 3:1–31, 1997.

[7] T. Braconnier. Fvspspack: a Fortran and PVM Package to Compute The Field of Values and Pseudospectra of Large Matrices. Numerical Analysis Report No. 293, Manchester Center for Computational Mathematics, 1996.

[8] T. Braconnier and N. J. Higham. Computing the field of values and pseudospectra using the Lanczos method with continuation. *BIT*, 36(3):422–440, 1996.

[9] M. Bruehl. A curve tracing algorithm for computing the pseudospectrum. *BIT*, 33(3):441–455, 1996.

[10] R. B. Burckel. *An Introduction to Classical Complex Analysis*, volume 64 of *Mathematische Reihe.* Birkhauser, Basel, 1979.

[11] C. Canuto, M. Hussaini, A. Quarteroni, and T. Zang. *Spectral methods in fluid dynamics*, volume xiv of *Springer Series in Computational Physics.* Springer-Verlag, 1988.

[12] F. Chatelin. *Spectral Approximation of Linear Operators.* Academic Press, London, 1983.

[13] C.T. Chen. *Linear System Theory and Design.* Holt, Rinehart & Winston, New York, 1984.

[14] M. G. Crandal and P. H. Rabinowitz. Bifurcation of simple eigenvalues. *J. Funct. Anal.*, 8:321–340, 1971.

[15] R.F. Curtain and H.J. Zwart. *An Introduction to Infinite-Dimensional Linear Systems Theory.* Number 21 in Texts in Applied Mathematics. Springer, New York, 1995.

[16] D. F. Delchamps. *State Space and Input-Output Linear Systems.* Springer Verlag, 1988.

[17] O. Diekman, S. A. van Gils, S. M.Verduyn Lunel, and H.O. Walther. *Delay Equations.* Springer-Verlag, 1995.

[18] J. Dongarra, B. Straughan, and D. Walker. Chebyshev tau-qz algorithm methods for calculating spectra of hydrodynamic stability problems. *Appl. Numer. Math.*, 22(4):399–434, 1996.

[19] P. G. Drazin and W. H. Reid. *Hydrodynamic Stability.* Cambridge University Press, Cambridge, UK, 1981.

[20] D. E. Edmunds and W. D. Evans. *Spectral Theory and Differential Operators.* Oxford University Press, London, 1987.

[21] B. Farrell and P. Ioannou. Optimal excitation of three-dimensional perturbations in viscous constant shear flow. *Phys. Fluids*, A 5(6):1390–1400, 1993.

[22] R. Fletcher. *Practical Methods of Optimization.* John Wiley & Sons, 1987.

[23] A. Geist, A. Beguelin, J. Dongara, W. Jiang, R. Manchek, and V. Sunderam. *PVM: A Users Guide and Tutorial for Networked Parallel Computing.* The MIT Press, Cambridge, Massachusetts, 1994.

[24] S. K. Godunov. Spectral portraits of matrices and criteria of spectrum dichotomy. In *Proceedings of International Conference on Computer Arithmetic, Scientific Computation and Mathematical Modelling. SCAN-91*, Oldenburg, 1991.

[25] S. K. Godunov, O. P. Kiriljuk, and W. I. Kostin. Spectral portraits of matrices. Technical Report 3, Inst. of Math. Acad.Sci.USSR, Soviet Union, 1990.

[26] I. Gohberg, S. Goldberg, and M. A. Kaashoek. *Classes of Linear Operators*, volume 49 of *Operator Theory: Advances and Applications.* Birkhäuser Verlag, 1990.

[27] G. H. Golub and C. F. Van Loan. *Matrix Computation.* John Hopkins University Press, 1988.

[28] J. Grcar. Operator coefficient methods for linear equations. Technical Report Rep. SAND89-8691, Sandia National Lab, November 1989.

[29] J. Guddat, F. Guerra, and H. Jongen. *Parametric Optimization: Singularities, Pathfollowing and Jumps.* John Wiley & Sons, 1990.

[30] L. Gustavsson. Resonant growth of three-dimensional disturbances in plane poiseuille flow. *J. Fluid Mech.*, 112, 1981.

[31] P. Halmos. Quasitriangular operators. *Acta Szeged.*, 29:283–293, 1968.

[32] A. Harrabi. Pseudospectra of a sequence of bounded operators. *Mathematical Modelling and Numerical Analysis*, 1997.

[33] A. Harrabi. Note on pseudospectra of closed operators. Technical Report TR/PA/90/08, CERFACS, 1998.

[34] W. K. Hayman. *Meromorphic Functions.* Oxford Mathematical Monographs. Oxford University Press, London, 1964.

[35] D. Henningson and S. Reddy. On the role of linear mechanisms in transition to turbulence. *Phys. Fluids*, pages 1396–1398, 1994.

[36] D. S. Henningson. and P. J. Schmidt. Vector eigenfunction expansions for plane flows. *Studies in Applied Mathematics*, 87:15–43, 1992.

[37] T. Herbert. Die neutrale Fläche der ebenen Poiseuille-Strömung, 1977. Habilitationsschrift, Univ. Stuttgart. Germany.

[38] I. H. Herron. A completeness theorem for the linear stability problem of nearly parallel flows. *J. Math. Anal. Appl.*, 1982.

[39] I. H. Herron. Property of basisness of the eigenfunctions of the Orr-Sommerfeld operator. Personal comunication, 1997.

[40] N. J. Higham. The Test Matrix Toolbox for Matlab (version 3.0). Numerical Analysis Report No. 279, Manchester Centre for Computational Mathematics, 1995.

[41] E. Hille and R.S. Phillips. *Functional Analysis and Semigroups.* American Mathematical Society, Providence, Rhode Island, 1957.

[42] D. Hinrichsen and B. Kelb. Spectral value sets: a graphical tool for robustness analysis. *Systems & Control Letters*, 21:127–136, 1993.

[43] D. Hinrichsen and B. Kelb. Stability radii and spectral value sets for real matrix perturbations. In U. Helmke and R. Mennicken, editors, *Proc. Conf. MTNS*, Regensburg, 1993.

[44] D. Hinrichsen and A.J. Pritchard. Stability radius for structured perturbations and the algebraic Riccati equation. *Systems & Control Letters*, 8:105–113, 1986.

[45] D. Hinrichsen and A.J. Pritchard. On spectral variations under bounded real matrix perturbations. *Numerische Mathematik*, 60:509–524, 1992.

[46] M. Huhtanen. On a resolvent condition in finite-dimensional controller design. Technical report, Centro Int. Math., Portugal, 1998. To appear.

[47] K. Ito and F. Kappel. A uniformly differentiable approximation scheme for delay systems using splines. *Applied Mathematics and Optimization*, 23:217–262, 1991.

[48] K. Ito and F. Kappel. Two families of approximation schemes for delay systems. *Results in Mathematics*, 21:93–137, 1992.

[49] S. Ji-guang. A note on simple non-zero singular values. *J. Comput. Mathematics.*, 6(3):258–266, July 1988.

[50] H. Jongen, P. Jonker, and F. Twilt. *Nonlinear Optimization in* $\mathbb{R}^n$. Verlag Peter Lang, 1983.

[51] T. Kailath. *Linear systems.* Prentice Hall, Englewood Cliffs, New Jersey, 1980.

[52] M. Karow. Hermitesch-symmetrische ungleichungen und der reelle stabilitätsradius. Master's thesis, University of Bremen, 1997.

[53] T. Kato. *Perturbation Theory for Linear Operators.* Springer-Verlag, Berlin-Heidelberg-New York, 1976.

[54] S. H. Lui. Computation of pseudospectra by continuation. *SIAM J. Sci. Comput.*, 18:565–573, 1997.

[55] A. G. J. MacFarlane and Y. S. Hung. Analytic properties of the singular values of a rational matrix. *Internat. J. Control*, 37(2):221–234, 1983.

[56] O. Marques and V. Toumazou. Spectral portrait computation by Lanczos method (augmented matrix method). Technical Report TR/PA/95/05, CERFACS, 1995.

[57] O. Marques and V. Toumazou. Spectral portrait computation by Lanczos method (normal equation method). Technical Report TR/PA/95/02, CERFACS, 1995.

[58] NAG Ltd, Wilkinson House. Jordan Hill Rode. Oxford. United Kingdom. *NAG Fortran Library Manual, Mark 16*, 1995.

[59] K. Nomizu. Characteristic roots and vectors of a differentiable family of symmetric matrices. *Linear and Multilinear Algebra*, 1:159–162, 1972.

[60] S. A. Orzag. Accurate solution of the Orr-Sommerfeld stability equation. *J. Fluid Mech.*, 50:689–703, 1971.

[61] V. Prasolov. *Problems and Theorems in Linear Algebra.* American Mathematical Society, 1994.

[62] R. C. Prima and G. J. Habetler. A completeness theorem for non-selfadjoint eigenvalue problems in hydrodynamic stability. *Arch. Rat. Mech. Anal.*, 34:218–227, 1969.

[63] Li Qiu, B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young, and J.C. Doyle. A formula for computation of the real stability radius. *Automatica*, 31(6):879–890, 1995.

[64] S. Reddy. A spectral method for computing transient growth. Personal Comunication, 1997.

[65] S. Reddy and D. Henningson. Energy growth in viscous channel flows. *J. Fluid Mech.*, pages 209–238, 1993.

[66] S.C. Reddy, P. Schmidt, and D. Henningson. Pseudospectra of the Orr-Sommerfeld operator. *SIAM J. Appl. Math.*, 53:15–47, 1993.

[67] S.C. Reddy and L.N. Trefethen. Lax-stability of fully discrete spectral methods via stability regions and pseudo-eigenvalues. *Comp. Meth. Appl. Mech. Engr.*, 80:147–164, 1990.

[68] S.C. Reddy and L.N. Trefethen. Stability of the method of lines. *Numerische Mathematik*, 62:236–267, 1992.

[69] A. Ribalta. Approximation of level curves. Personal communication, 1997.

[70] R. Russell. A comparison of collocation and finite differences for two-point boundary value problems. *SIAM J. Numer. Anal.*, 1977.

[71] D. Salamon. Structure and stability of finite dimensional approximations for functional differential equations. *SIAM J. Contr. & Opt.*, 1985.

[72] L. Schumaker. *Spline Functions Basic Theory.* John Wiley & Sons, 1981.

[73] Gilbert W. Stewart and Ji-guang Sun. *Matrix perturbation theory.* Computer Science and Scientific Computing. Academic Press, Boston, 1990.

[74] G. Szego. *Orthogonal Polynomials*, volume XXIII of *Colloquium Publications.* American Mathematical Society, Providence, Rhode Island, 3rd edition, 1967.

[75] K. C. Toh and L. N. Trefethen. Pseudozeros of polynomials and pseudospectra of companion matrices. *Numerische Mathematik*, 68:403–425, 1994.

[76] K. C. Toh and L. N. Trefethen. Calculation of pseudospectra by the Arnoldi iteration. *SIAM J. Sci. Comput.*, 17(1):1–15, 1996.

[77] L. N. Trefethen. Pseudospectra of linear operators. *SIAM Review*, 39(3):383–406, 1997.

[78] L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A.Driscol. Hydrodynamic stability without eigenvalues. *Science*, 261:578–584, 1993.

[79] L.N. Trefethen. Pseudospectra of matrices. In D.F. Griffiths and G.A. Watson, editors, *Numerical Analysis*, volume 91, pages 234–266. Longmann, 1992.

[80] G. Vainikko. *Funktionalanalysis der Diskretisierungs-methoden.* Teubner Verlag, Leipzig, 1976.

[81] G. Vainikko. Approximative methods for nonlinear equations (two approaches to the convergence problem). *Nonlinear Analysis: Theory, Methods and Applications*, 2(6):647–687, 1978.

[82] G. Vainikko. Regular convergence of operators and approximate solution of equations. *J. Sov. Math.*, 15:675–705, 1981.

[83] J. Wilkinson. The perfidious polynomial. In G. Golub, editor, *Studies in Numerical Analysis*, volume 24 of *Studies in Mathematics*, pages 1–28. Mathematical Association of America, 1984.