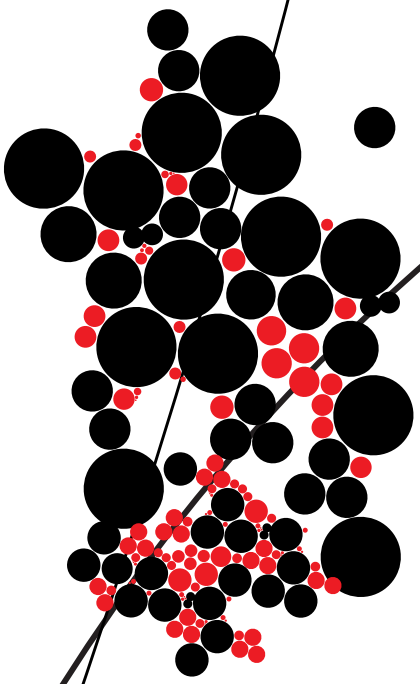

Spectra of finite and infinite Jordan blocks

G.P. van der Beek

26 AUGUST 2015

MASTER THESIS IN APPLIED MATHEMATICS

CHAIR HYBRID SYSTEMS

GRADUATION COMMITTEE

prof. dr. H.J. Zwart

dr. ir. G. Meinsma

dr. G.A.M. Jeurnink

UNIVERSITY OF TWENTE.

Preface

The following document is the result of my final project at the University of Twente. It was hard work and I especially would like to thank prof. dr. H.J. Zwart for his supervision. Without his help this thesis would not be there. I also would like to thank my parents, who kept believing in me although my body was not supportive.

This thesis is a rather theoretical one for a Master in Applied Mathematics. But the plots in this thesis are fascinating and I hope I can show some of that enthusiasm in the following pages. Those plots are what made me start this project and also what kept me motivated to complete it. The mathematics of operator and perturbation theory is hard, but I hope that this report clearly explains what I did the last months.

Gert-Paul van der Beek
Enschede
August 2015

Abstract

The eigenvalues of Jordan blocks are very sensitive to perturbations. This is known for a long time, but why the eigenvalues of a single Jordan block converge to the spectrum of the shift operator when the dimension runs to infinity, is unknown. In this thesis we show why Jordan blocks are so sensitive to perturbations, what has been studied about them in the literature and what the location of the eigenvalues is after perturbation. We also study the shift operator, calculate its spectrum and show that this spectrum is not sensitive to perturbations. Important to note is that the shift operator can be seen as a single Jordan block on an infinite space.

We did not find a definite answer to the relation between the two, but by studying the pseudospectra of both the matrix and the operator we give some clues on why the spectrum of both structures are related.

Contents

Introduction 1

Eigenvalues are useful properties of matrices. When we have a square matrix A, then for non-zero solutions to the equation

$$Av = \lambda v$$

we call λ an eigenvalue and v an eigenvector of A. Together this is an eigenpair of the matrix A.

Eigenvalues are often used to study the stability of systems, but under small perturbations eigenvalues can change significantly. Therefore it is useful to study the sensitivity of the eigenvalues. How sensitive the eigenvalues are is especially visible if we plot the spectrum (the collection of all eigenvalues) in the complex plane.

1.1 Problem description

In this thesis we study the spectrum of random perturbations of the Jordan Canonical Form. We especially look at matrices $A \in \mathbb{C}^{nk \times nk}$ defined as

$$A = \begin{bmatrix} C & D & & \\ & \ddots & \ddots & \\ & & \ddots & D \\ & & & C \end{bmatrix}, \quad C, D \in \mathbb{C}^{k \times k}. \tag{1.1}$$

All other values are zero, a convention we continue in the rest of this thesis. This matrix is perturbed by a random Gaussian matrix with a small variance σ^2:

$$A + \mathcal{N}_{nk \times nk}(0, \sigma^2), \quad \sigma^2 \ll 1.$$

The eigenvalues of this perturbed matrix converge to the eigenvalues of C (for a given n) if $\sigma \to 0$, but this happens slowly. In Figure 1.1 this is illustrated with σ decreasing with each plot. When $\sigma = 0.1$ the spectrum is random. But for $10^{-2} < \sigma < 10^{-16}$ we see that the spectrum of the perturbation resembles the spectrum of the shift operator corresponding to A. This operator is the shift operator. The shift operator A_∞ is defined as $g = A_\infty f$ with

$$(g_k) = (Cf_k + Df_{k+1}), k \in \mathbb{Z}.$$

By inspecting of $Af = g$, for $f \in \mathbb{C}^{(2n+1)k}$, i.e.

$$A \begin{bmatrix} f_{-n} \\ \vdots \\ f_n \end{bmatrix} = \begin{bmatrix} g_{-n} \\ \vdots \\ g_n \end{bmatrix},$$

we see that corresponds to the equations

$$g_k = Cf_k + Df_{k+1}, \ k \in \{-n, \dots, n\}.$$

If we let $n \to \infty$ this becomes the operator

$$g_k = Cf_k + Df_{k+1}.$$

This operator is the above shift operator A_∞. In Chapter 3 we will explain this in more detail.

That the spectrum of the shift operator and the spectrum a single Jordan block are related was noticed in another Master Thesis [Fir12]. This was also visible in the plots of Figure 1.1. In that thesis no explanation could be found. The goal of this thesis is to analyse why these two behaviours are related.

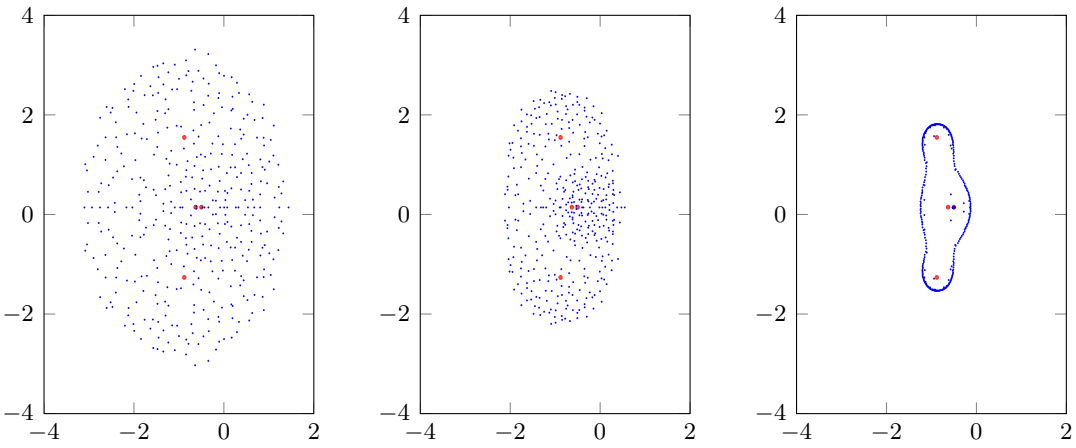

Figure 1.1: Simulations from [Fir12, p.25] with $\sigma = 0.1, 0.05$ and 10^{-3}.

1.2 Structure of this thesis

We start our analysis with a simplified version of the problem we stated in the previous section. In the first chapters of this thesis we only perturb with a deterministic matrix, that means that we perturb by a small value ε or a deterministic matrix with norm $\|\cdot\| < \varepsilon$. We will introduce randomness in Chapter 5. We also start our analysis with the assumption

that C and D are 0 or 1. We then get a single Jordan block like

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

There is a lot known about perturbations of these Jordan blocks. In Chapter 2 we use this to explain what happens if we perturb J in one or in multiple places. We extend this to perturbations of the original problem, the block matrix (1.1).

In Chapter 3 we look at the case when the size n is infinite. We then get an operator that acts on an infinite sequence. We start with the operator that corresponds to Jordan block J. Later in that chapter we show how the operator corresponding to the block matrix (1.1) looks like and what happens after perturbations.

In Chapter 4 we look at pseudospectra, another way to look at the structure of a matrix.

In Chapter 5 we introduce randomness and explain the difference between the deterministic and random case.

1.3 Notation

In this thesis all matrices are written with in upper case, like A, and vectors with in lower case, like v. Just like in matrix (1.1) above, empty values in a matrix are zero. When we write a norm $\|\cdot\|$ we normally mean the Eucledian norm,

$$\|v\|_2 = \sqrt{\sum_{i=1}^{N} |v_i|^2},$$

when not specified otherwise.

Zero matrices of size $n \times k$ are written as $0_{n\times k}$, matrices of size $n \times k$ are written as $[\cdot]_{n\times k}$ and $\text{diag}_k(a)$ is the $k \times k$ matrix with a on its diagonal. Values at position (n, m) inside a matrix A are written like A_{nm}. The unit disk with radius r, $\{z \in \mathbb{C} \mid |z| \leq r\}$, is written as $\mathbb{D}_r$.

The spectrum, the collection of all eigenvalues, is denoted with Λ.

We will most times work in the ℓ^2 space, i.e. the linear space consisting of all sequences v such that $\sum_{i=1}^{\infty} |v_i|^2 < \infty$. The corresponding norm is

$$\|v\|_2^2 = \sum_{i=1}^{\infty} |v_i|^2.$$

In some places we also need the ℓ^1 and ℓ^p spaces, i.e. the linear spaces consisting of all sequences v such that $\sum_{i=1}^{\infty} |v_i| < \infty$ and $\sum_{i=1}^{\infty} |v_i|^p < \infty$, respectively. The corresponding norms are

$$\|v\|_1 = \sum_{i=1}^{\infty} |v_i| \text{ and } \|v\|_p^p = \sum_{i=1}^{\infty} |v_i|^p.$$

We perturb at different places and to make that easy we introduce the matrix Δ with all zeros, except at position (a, b). Formally we define

$$\Delta_{a,b} := \begin{cases} 1 & \text{at position } (a, b) \\ 0 & \text{otherwise} \end{cases} . \tag{1.2}$$

Eigenvalues and perturbations 2

If we want to know why the behaviour of the finite and infinite dimensional case is related we have to understand both in detail. We study both separately and in this chapter we start with the finite case, this means we work with ordinary matrices. The eigenvalues of the matrices we study, like

$$J = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix} \text{ or } A = \begin{bmatrix} C & D & & \\ & \ddots & \ddots & \\ & & \ddots & D \\ & & & C \end{bmatrix} \text{ with } C, D \in \mathbb{C}^{k\times k}, \tag{2.1}$$

are trivial without a perturbation. The eigenvalues of J and A are 0 and the eigenvalues of C, respectively. But the problem becomes more complicated if we perturb our matrix with a small perturbation. In this chapter we look at the eigenvalues of such perturbed matrices. We start with the Jordan block J and perturb at one location and we add perturbations of more locations later in this chapter. We will also see why matrices like J and A are so sensitive to perturbations. In the second half of this chapter we turn to the eigenvalues of perturbations of block matrix A.

2.1 One perturbation

Since we expect that the value in the left bottom of the matrix has the biggest influence on the eigenvalues, we perturb at position $(n, 1)$ and we look at the eigenvalues of $J + \Delta_{n,1}\varepsilon$.

Example 2.1 *To find the eigenvalues of $J + \Delta_{n,1}\varepsilon$, we need to find the roots of the characteristic polynomial of $J + \Delta_{n,1}\varepsilon - \lambda I$. So by applying Cramer's rule twice, first on the*

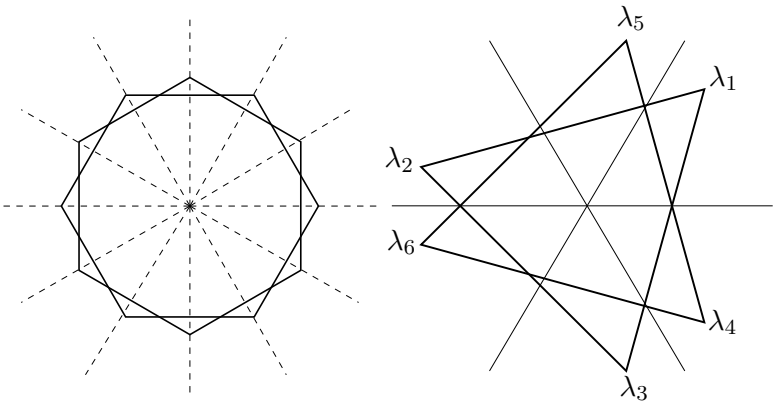

Figure 2.1: [CB94, p. 4] Perturbation of a matrix with one eigenvalue when $n = 6$. Left: one Jordan block, ε both positive and negative. Right: Two Jordan blocks.

first row and afterwards on the first column we get

$$\det \begin{bmatrix} -\lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\lambda & 1 \\ \varepsilon & 0 & 0 & 0 & -\lambda \end{bmatrix} = -(-1)^n \varepsilon \det \begin{bmatrix} 1 & & \\ -\lambda & \ddots & \\ & \ddots & 1 \\ & & -\lambda \end{bmatrix} + (-\lambda)^n$$
$$= (-\lambda)^n + (-1)^{n+1}\varepsilon = 0.$$

Thus all eigenvalues are distributed on a circle with radius $r = |\sqrt[n]{\varepsilon}|$ around the origin. Specifically the solutions are

$$\varepsilon^{1/n} \left(\cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right) \right), \; k = (0, \dots, n-1).$$

This means that all eigenvalues are distributed evenly around the circle with radius r. The difference between the argument of two succeeding eigenvalues is

$$\arg(\lambda_{i+1}) - \arg(\lambda_i) = \frac{2\pi}{n}.$$

For $k = 0$ a real eigenvalue is

$$\lambda(\varepsilon) = \sqrt[n]{\varepsilon}, \; \varepsilon > 0.$$

When $\varepsilon > 0$ and $n \to \infty$ we see that $\lambda(\varepsilon) \to 1$.

That the eigenvalues are distributed evenly around a circle is also visible in the two illustrations of Figure 2.1. There we see that depending on whether ε is positive of negative, the positions of the eigenvalues change slightly. For multiple Jordan Blocks, multiple rings can form.

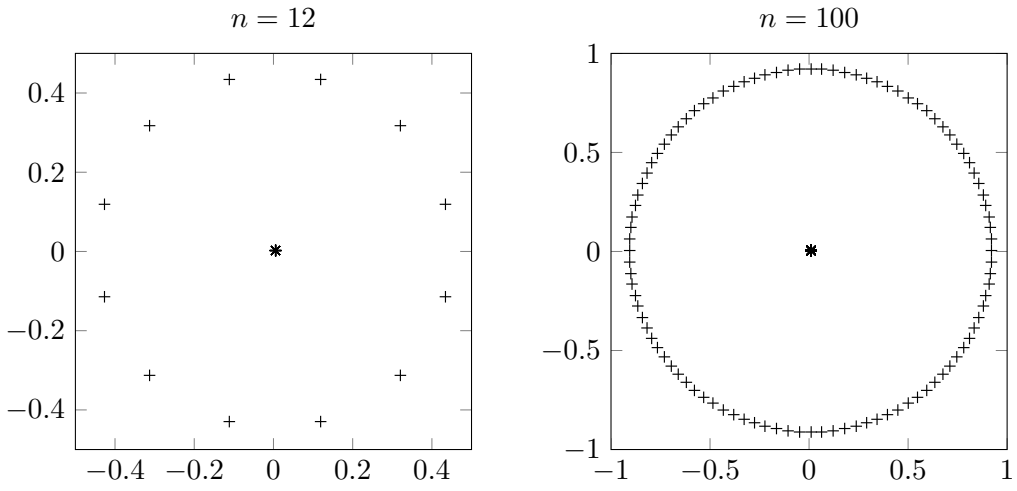

FIGURE 2.2: 1 Random perturbation of $\varepsilon = 10^{-4}$ of a Jordan block with zeros on the diagonal. Visible are the original eigenvalue (*) and the perturbed eigenvalues (+).

Lidskii [Lid66] proved that if we perturb $J + \varepsilon B$, with matrix size n and ε is small enough, the eigenvalues of $J + \varepsilon B$ lie on a circle. To be exact, the eigenvalues of the perturbed system λ_p are related to the eigenvalues λ of J.

$$\lambda_p = \lambda + (\xi)^{1/n}\varepsilon^{1/n} + o(\varepsilon^{1/n}), \xi = yBx, \tag{2.2}$$

with y, x the left and right eigenvector of J. This shows that is of the order $\mathcal{O}(\varepsilon^{1/n})$. In Figure 2.2 the eigenvalues of $J + \Delta_{n,1}\varepsilon$ are plotted for $\varepsilon = 10^{-4}$ and two values of n. There it is visible that it goes to 1.

Why do the eigenvalues of matrix change from all 0 to a circle with radius of almost 1 if we just add one small perturbation? The problem is that the matrices like (1.1) are not normal. In the next section we explain what normal matrices are and why nonnormality is a problem for the stability of eigenvalues.

2.2 Normal matrices

In the previous section we stated that the problem with Example 2.1 was that the matrix was nonnormal.

Definition 2.1 (Normal Matrix) *A matrix $A \in \mathbb{C}^{n\times n}$ is normal if $A^*A = AA^*$, with A^* the complex conjugate of A.*

If we look to a normal matrix, it is known that the eigenvalues of the perturbed matrix lie close to the original eigenvalues. This is illustrated in Figure 2.3 for normal matrix

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix} + \varepsilon E, \ \|E\| = 1. \tag{2.3}$$

There it is visible that the perturbed eigenvalues lie within ε of the original eigenvalues. So with normal matrices, small perturbations also cause small perturbations of the eigenvalues. Another illustration is in Figure 2.4 where it is visible how random perturbations influence the eigenvalues of a normal matrix with original eigenvalues 1 to 9.

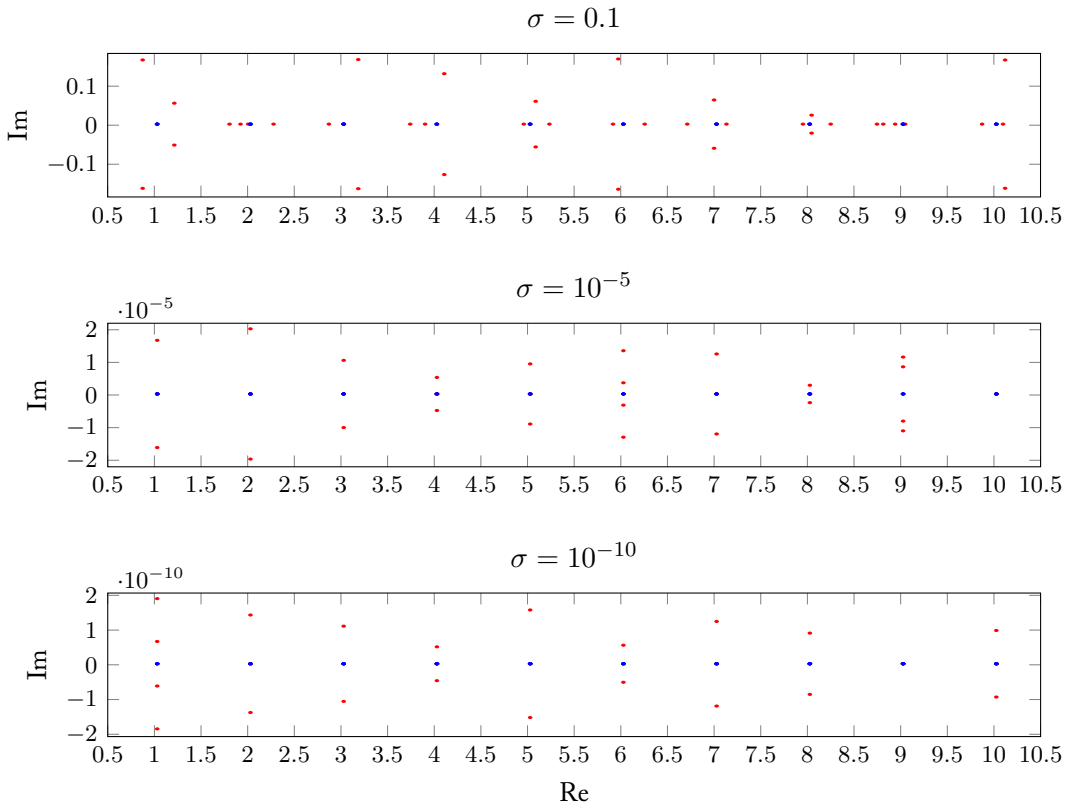

Figure 2.4: Random perturbations of normal matrices

But what makes normal matrices so special? If a matrix is normal, it has a complete set of orthogonal eigenvectors [TE05, p. 9]. Why this is a problem, it visible in example 2.2.

For the diagonal matrix in (2.3) it is clear that it is a normal matrix. To check if a matrix is normal or how "far away" from matrix is from normal can be by the condition number of a matrix that was introduced in Wilkerson's *The algebraic eigenvalue problem* [Wil65]. There the condition number $\kappa(A)$ is defined as

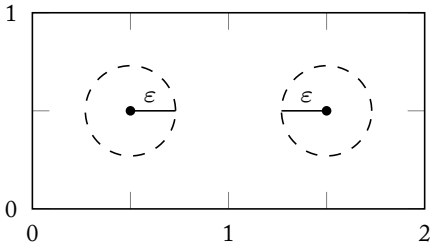

Figure 2.3: Perturbations of a normal matrix with eigenvalues $\frac{1}{2}$ and $\frac{3}{2}$.

$$\kappa(A) = \|A\|\|A\|^{-1}, \text{with } A \text{ non singular.}$$

Example 2.2 *Let's take the most simple Jordan Block*

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

the eigenvalues are 0 *(with multiplicity* 2*). But if we compute the eigenvectors we see that the only eigenvector is*

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Therefore it is impossible to span the complete space.

Wilkerson came to this definition via the analysis of the Jordan Canonical Form. He showed that if λ is a simple[1] eigenvalue of A, with y, x A's left and right eigenvectors, and y^* the complex conjugate of y, for $\lambda(\varepsilon)$ an eigenvalue of $A + \varepsilon E, ||E|| = 1$, with

$$\lambda(\varepsilon) = \lambda + \frac{y^* E x}{y^* x}\varepsilon + \mathcal{O}(\varepsilon^2).$$

So we see that

$$\lambda(\varepsilon) - \lambda \leq \left| \frac{y^* E x}{y^* x} \right| \varepsilon \leq \frac{1}{|y^* x|}\varepsilon.$$

The denominator $s := |y^* x|$ is called the INDIVIDUAL CONDITION NUMBER of an eigenvalue.

Example 2.3 *If we look at the rate of change of* $\lambda(\varepsilon)$*,* $\lambda(\varepsilon)$ *the eigenvalues of* $A + \varepsilon E$*,*

$$\frac{\mathrm{d}\lambda(\varepsilon)}{\mathrm{d}\varepsilon} = \frac{1}{n}\varepsilon^{\frac{1}{n}-1} = \frac{1}{n\sqrt[n]{\varepsilon^{n-1}}},$$

we can see that the rate of change at the origin (when $\varepsilon = 0$*) is infinite and why in the case of multiple eigenvalues Wilkinsons theory above is not valid.*

Wilkinson also uses Gershorins theorem (Theorem 2.1) to show other ways to analyse perturbations of systems with multiple eigenvalues. This is one of the first theorems on bounds on eigenvalues, established in the 1930's.

Theorem 2.1 (Gershgorin) *Let* $A = (a_{ij}) \in \mathbb{C}^{n \times n}$ *and let the* GERSHGORIN DISKS *of* A *be defined by*

$$\mathcal{G}_i := \left\{ \mu : |\mu - a_{ii}| \leq \sum_{j \neq i} |a_{ij}| \right\}.$$

Then

$$\Lambda(A) \subset \bigcup_{i=1}^{n} \mathcal{G}_i.$$

Moreover, if the union of k *of the sets* $\mathcal{G}_i$ *are disjoint from the others, then that union contains exactly* k *eigenvalues of* A*.*

[1] That means all eigenvalues are different

Example 2.4 *If we take $J+\Delta_{n,1}\varepsilon$ from Example 2.1, then we see we have $n-1$ Gershgorin disks $\{\mu : |\mu| \leq 1\}$ and one disk $\{\mu : |\mu| \leq \varepsilon\}$. There is no union of k $(k < n)$ disjoint disks and the spectrum is contained in the union of all Gershgorin disks: $\{\mu : |\mu| \leq 1\}$.*

We already saw in Figure 2.2 that our perturbed matrix stayed within the circle with radius 1. So it seems already a good bound. In the next section we will see if we can make more direct relationship between the perturbation and the resulting eigenvalues.

2.3 Perturbation theory

We notices that the bound we derived in the previous section was already good. However, we only measure here the eigenvalues based on the values on the diagonal. When we extend our problem to (1.1) the diagonal items may say less about the size of our spectrum. So we want to find a relation between the eigenvalues of a matrix and its perturbation. We can use the Schur decomposition [GVL13, Th 7.2.3] to find the distance between the original eigenvalue and the eigenvalue of its perturbation.

Theorem 2.2 *Let $Q^*AQ = D + N$ be a Schur decomposition of $A \in \mathbb{C}^{n\times n}$. This means that Q is unitary, D diagonal and N is an uppertriangular matrix. If $\mu \in \Lambda(A + E)$ and p is the smallest positive integer such that $|N|^p = 0$ then*

$$\min_{\lambda\in\Lambda(A)} |\lambda - \mu| \leq \max\{\theta, \theta^{1/p}\},$$

where

$$\theta = ||E|| \sum_{k=0}^{p-1} ||N||^k.$$

Example 2.4 (continued) *We see that $p = n$, because with every multiplication of J with itself the n^{th} superdiagonal, becomes the $(n + 1)^{th}$ superdiagonal. Therefore after n multiplications we have the zero matrix left. And since $\|N\| = 1$ we get*

$$\theta = n||E||_2 = n|\varepsilon|.$$

So the difference between the original and the perturbed eigenvalues is the maximum of θ and $\theta^{1/p}$. So if $n > 1/\varepsilon$ we have a bound that is bigger than 1, but if ε is smaller we can find a bound within the unit circle. This is also what we would expect, since when n grows the bound grows to the unit circle.

How fast this happens depends on ε and n. In Figure 2.5 it is visible how fast the eigenvalues grow for increasing ε. We cannot show the growth from the origin, because MATLAB is not precise enough. We see although, that for bigger n, our perturbed eigenvalues are big, even if ε small.

Explicit bounds on the eigenvalues are given in a matrix and its perturbation are found in a theory by Elsner. Which shows that especially the situation with A a Jordan block J is causing a big difference in eigenvalues.

Theorem 2.3 ([BEK90]) *Let A be the (possibly multiple) eigenvalues $\lambda_1, \dots, \lambda_n$. Let the eigenvalues of $\tilde{A} = A + E$ be $\tilde{\lambda}_1, \dots, \tilde{\lambda}_n$. Then there is a permutation $j_1, \dots, j_n$ of the integers $1, \dots, n$ such that*

$$|\tilde{\lambda}_{j_i} - \lambda_i| \leq 4(||A||_2 + ||\tilde{A}||_2)^{1-\frac{1}{n}} ||E||_2^{\frac{1}{n}}.$$

So we see that we can also found a bound of the difference between all individual eigenvalues.

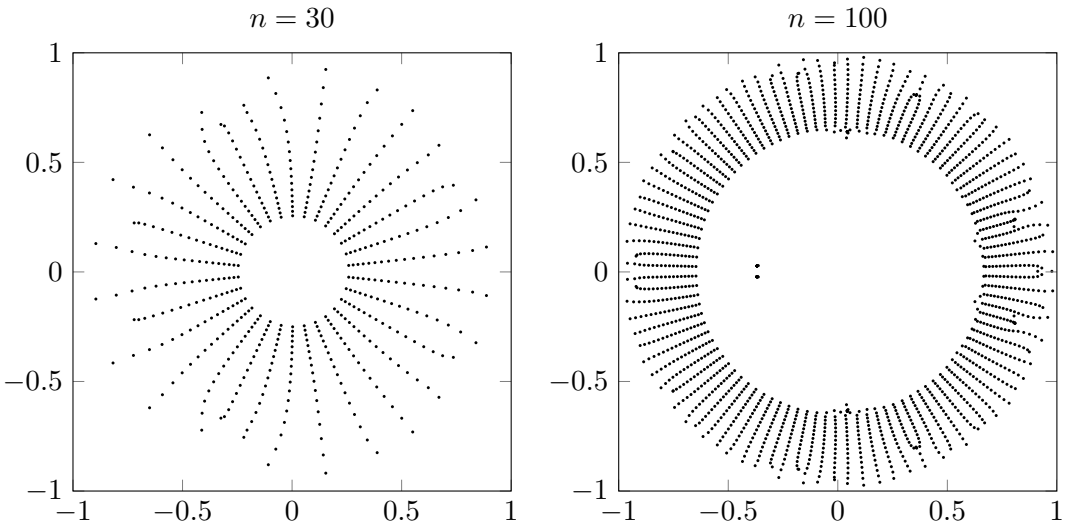

FIGURE 2.5: $A + \varepsilon E$, $||E|| = 1$, $\varepsilon = 10^{-17}, \dots, 10^{-1}$.

2.4 Characteristic polynomial with multiple perturbations

We also want to know what happens when we have multiple perturbations. Just as we did in Example 2.1 we can calculate the characteristic polynomial for more than one perturbation. For two perturbations ε_1, ε_2 in the left corner, we want to calculate the eigenvalues of $J + \varepsilon_1 \Delta_{n,1} + \varepsilon_2 \Delta_{n-1,1}$. These are the solutions of

$$\det \begin{bmatrix} -\lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \varepsilon_2 & \dots & 0 & -\lambda & 1 \\ \varepsilon_1 & 0 & 0 & 0 & -\lambda \end{bmatrix} = -(-1)^{n+1}\varepsilon_2 \det \begin{bmatrix} 1 & & & \\ -\lambda & \ddots & & \\ & \ddots & -\lambda & 1 \\ & & 0 & -\lambda \end{bmatrix}$$

$$- (-1)^n \varepsilon_1 \det \begin{bmatrix} 1 & & \\ -\lambda & \ddots & \\ & \ddots & -\lambda \end{bmatrix} + (-\lambda)^n$$

$$= (-1)^{n+1}\lambda\varepsilon_2 + (-1)^{n+1}\varepsilon_1 + (-\lambda)^n$$

$$= (-\lambda)^n + (-1)^{n+1}(\varepsilon_2\lambda + \varepsilon_1) = 0. \tag{2.4}$$

We arrived at the first equality by applying Cramer's rule, just as we did in Example 2.1. Since we look for solutions for n large, there will be no analytic solution for this equation.

Similarly, when we apply three perturbations $\varepsilon_1, \varepsilon_2, \varepsilon_3$ to J we want to know the eigenvalues of $J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1} + \varepsilon_3\Delta_{n,2}$. Again no analytic solution exists for the zeros of the characteristic polynomial. Our characteristic polynomial in this case becomes

$$\begin{aligned}
\det \begin{bmatrix} -\lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \varepsilon_2 & \dots & 0 & -\lambda & 1 \\ \varepsilon_1 & \varepsilon_3 & 0 & 0 & -\lambda \end{bmatrix} &= -(-1)^{n+1}\varepsilon_2 \det \begin{bmatrix} 1 & & & \\ -\lambda & \ddots & & \\ & \ddots & -\lambda & 1 \\ \varepsilon_3 & & 0 & -\lambda \end{bmatrix} \\
&\quad - (-1)^n\varepsilon_1 \det \begin{bmatrix} 1 & & & \\ -\lambda & \ddots & & \\ & \ddots & \ddots & \\ & & -\lambda & 1 \end{bmatrix} - \lambda(-\lambda^{n-1} + (-1)^n\varepsilon_3) \\
&= (-1)^{n+1}\lambda\varepsilon_2 + (-1)^{n+1}\varepsilon_1 - \lambda(-\lambda^{n-1} + (-1)^n\varepsilon_3) \\
&= (-\lambda)^n + (-1)^{n+1}((\varepsilon_2 + \varepsilon_3)\lambda + \varepsilon_1) = 0. \qquad (2.5)
\end{aligned}$$

It is also possible to construct the characteristic polynomial for many more perturbations. Davies and Hager [DH09, p. 8] found that for

$$J + \delta \begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix}$$

the characteristic polynomial is

$$f(\lambda) := \sum_{i,j=1}^{k} C_{i,j}(R\lambda)^{j-i+k-1}, \tag{2.6}$$

where $\delta = R^N, R \in (0, \infty)$. Using this result it is possible to find the characteristic polynomial without multiple times applying Cramer's rule.

Although there is no analytic solution to the equations (2.4) and (2.5), we can use Rouché's theorem to find a region where the zeros are located.

2.5 Regions of eigenvalues

We can use Rouché's theorem to prove that a function with an analytic solution has an equal number of zeros in a certain region.

Theorem 2.4 (Rouché's theorem [CZ12]) *Let f and g be functions that are holomorphic on the domain $\mathcal{Y}$, and suppose that $\mathcal{Y}$ contains a simple, closed contour Γ. If $|f(s)| > |g(s)|$ for $s \in \Gamma$, then f and $f + g$ have the same number of zeros inside Γ. (A zero of order p counts for p zeros.)*

To find such solutions, we need the following lemma.

Lemma 2.1 *For the absolute value the following statements hold*

1. $g > h$ *and* $g > -h$ *if and only if* $g > |h|$.
2. $|a + b| \leq |a| + |b|$ *(triangle inequality).*
3. $\big|\, |a| - |b| \,\big| \leq |a + b|$.
4. $\big|\, |a| - |b| \,\big| \leq |a - b|$.

Proof *We only prove item 3, since item 4 follows from item 3 with* $b := -b$.

If we set $a := a + b$ *and* $b := -b$ *in item 2 we arrive at* $|a| - |b| \leq |a + b|$. *If we reverse the argument and choose* $b := -a$ *instead, we arrive at* $|b| - |a| \leq |a + b|$. *Because of item 1 we arrive at the result.* ■

We defined in the introduction the region inside the disk with radius r as $\mathbb{D}_r$. We want to find a region $\mathbb{D}_r$ for which Rouché's theorem is valid, by proving the inequalities from Rouché's theorem on the circle $|\lambda| = r$ for a given r. If we know that Rouché's theorem is valid for two functions on a certain region, then we know that they have equal zeros inside this region. So we can find a region and a more simple function and use Rouché to prove that all zeros of our difficult characteristic polynomial lie within this region.

2.5.1 An outer region

We want to find a region containing the eigenvalues of $J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1}$, as we did in the previous section. If we assume that

1. $r = 1 + \varepsilon, \varepsilon > 0$
2. $|\varepsilon_1| + |\varepsilon_2| < 1$

and we define

$$\begin{aligned} f(\lambda) &= (-\lambda)^n + (-1)^{n+1}\varepsilon_1, \\ h(\lambda) &= (-1)^{n+1}\varepsilon_2\lambda, \\ g(\lambda) &= f(\lambda) + g(\lambda) := \chi(\lambda), \end{aligned}$$

then it follows from assumption 2 that $|\varepsilon_1| + |\varepsilon_2| + |\varepsilon_2|\varepsilon < 1 + \varepsilon$ (since then also $|\varepsilon_2| < 1$) and because $\varepsilon > 0$ also that $|\varepsilon_1| + |\varepsilon_2| + |\varepsilon_2|\varepsilon < (1 + \varepsilon)^n$. We use this to prove that for λ with $|\lambda| = r$ inequality (2.4) holds. For $|\lambda| = r$ we have

$$\begin{aligned} |h(\lambda)| = |\varepsilon_2|(1 + \varepsilon) &< (1 + \varepsilon)^n - |\varepsilon_1| && (2.7) \\ &= |\lambda|^n - |\varepsilon_1| \\ &\leq |(-\lambda)^n + (-1)^{n+1}\varepsilon_1| = |f(\lambda)|. && (2.8) \end{aligned}$$

By Rouché's theorem, $g = f+h$ has the same number of zeros inside $\mathbb{D}_{1+\varepsilon}$ as f. Since f has n zeros inside $\mathbb{D}_{1+\varepsilon}$, so has g. Thus we know that all eigenvalues of $J+\varepsilon_1\Delta_{n,1}+\varepsilon_2\Delta_{n-1,1}$ lie within $\mathbb{D}_{1+\varepsilon}$.

Because the only condition is that $\varepsilon > 0$, we can choose ε as small as we would like. As long as assumption 2 holds we have that

$$\Lambda(J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1}) \subseteq \mathbb{D}_1$$

has n zeros inside $\mathbb{D}_{1+\varepsilon}$. So we can let the region which contains the spectrum of $J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1}$ shrink to the unit circle if we keep decreasing our ε.

Note that when n is large we don't need assumption 2. If assumption 2 does not hold, we can choose every $\varepsilon, \varepsilon_1, \varepsilon_2$ we want. Given we fix $\varepsilon, \varepsilon_1, \varepsilon_2$, $|\varepsilon_2|(1+\varepsilon) + |\varepsilon_1|$ is also a fixed value, while $(1+\varepsilon)^n$ goes to infinity when $n \to \infty$ (since $1+\varepsilon > 1$). If we define $v := |\varepsilon_2|(1+\varepsilon) + |\varepsilon_1|$ we see there always exists a n such that $(1+\varepsilon)^n > v$. This can be seen, because $(1+\varepsilon)^n \to \infty$, while v does not depend on n. For this n and our chosen values of $\varepsilon, \varepsilon_1, \varepsilon_2$, inequality (2.7) is valid. So $|h| < |f|$. Therefore we have the same number of zeros inside f and g. Thus $f(\lambda)$ has n zeros inside $\varepsilon, \varepsilon_1, \varepsilon_2$. So we can always choose a n such that all our zeros lie inside $\mathbb{D}_{1+\varepsilon}$.

2.5.2 An region dependent on n

We can also find a radius r for an outer region dependent on n that is smaller than 1 if $\varepsilon_1, \varepsilon_2$ are small. We assume that

1. $r = \sqrt[n]{|\varepsilon_1|} + \beta$, $\beta > \frac{|\varepsilon_2|}{|\varepsilon_1|n}$,
2. $|\varepsilon_1| < 1$,
3. $n \geq 3$

and we define

$$\begin{aligned} f(\lambda) &= (-\lambda)^n + (-1)^{n+1}\varepsilon_1, \\ h(\lambda) &= (-1)^{n+1}\varepsilon_2\lambda, \\ g(\lambda) &= f(\lambda) + g(\lambda) := \chi(\lambda). \end{aligned}$$

From assumption 2 it follows that $\sqrt[n]{|\varepsilon_1|} < 1$. From this and assumption 1 we see that

$$|\varepsilon_2|(\sqrt[n]{|\varepsilon_1|} + \beta) < (1+\beta)|\varepsilon_2| < (1+\beta)n\beta|\varepsilon_1|.$$

By assumption 2, we also know that $|\varepsilon_1| < |\varepsilon_1|^m$ when $m < 1$ and by assumption 3 we know that $\frac{1}{2}n(n-1) > n$. We use both results to prove that for λ with $|\lambda| = r$ inequality (2.4)

holds. For $|\lambda| = r$ we have

$$\begin{aligned}
|h(\lambda)| = |\varepsilon_2| r = |\varepsilon_2|(\sqrt[n]{|\varepsilon_1|} + \beta) &< (1+\beta) n\beta|\varepsilon_1| \\
&= n\beta|\varepsilon_1| + n\beta^2|\varepsilon_1| \\
&< n\beta|\varepsilon_1|^{\frac{n-1}{n}} + \frac{n(n-1)}{2}\beta^2|\varepsilon_1|^{\frac{n-2}{n}} + \sum_{k=3}^{n}\binom{n}{k}\beta^k|\varepsilon_1|^{\frac{n-k}{n}} \\
&= \left(\sqrt[n]{|\varepsilon_1|} + \beta\right)^n - |\varepsilon_1| = |\lambda|^n - |\varepsilon_1| \\
&= |(-\lambda)^n| - |(-1)^{n+1}\varepsilon_1| \leq |f(\lambda)|. \quad \text{(By Lemma 2.1.4.)}
\end{aligned}$$

By Rouché's theorem, $g = f + h$ has the same number of zeros inside $\mathbb{D}_{\sqrt[n]{|\varepsilon_1|}+\beta}$ as f. Since f has n zeros inside this disk, so has g. Thus the eigenvalues of $J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1}$ lie within $\mathbb{D}_{\sqrt[n]{|\varepsilon_1|}+\beta}$.

2.5.3 Summary

In the previous subsections we found a number of regions where our eigenvalues are located. How these are located is visible in Figure 2.6 and Figure 2.7. In Figure 2.6 we show how a random perturbation of a Jordan block behaves compared to the perturbation with two random variables, as in section 2.4 was calculated. In Figure 2.6 we show how the region we computed in section 2.5.2 compared to the eigenvalues of these perturbed matrices.

In section 5.2 we will calculate the probabilities that the constraints on ε are when our ε are random.

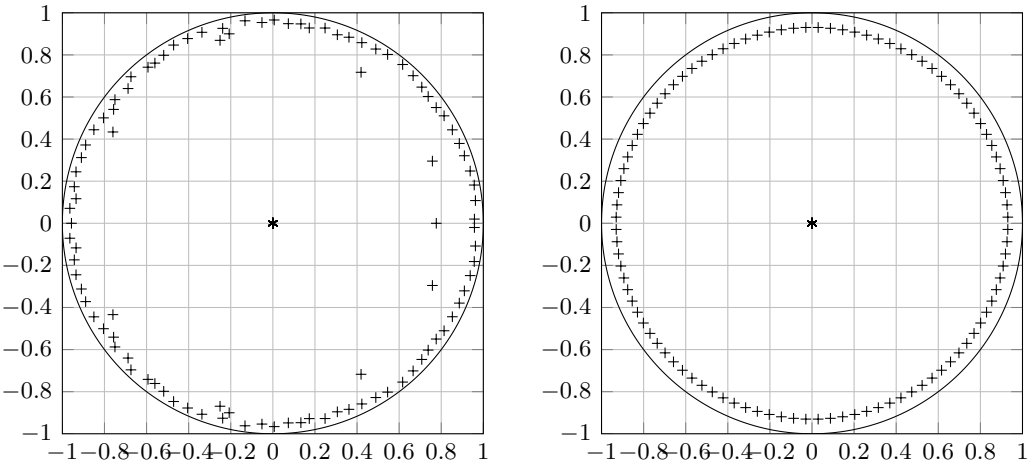

FIGURE 2.6: Two perturbations to a Jordan block J. Values used are $n = 100, \varepsilon = 2 \cdot 10^{-3}$. The solid circle is the circle with radius 1.

2.6 Block matrices

We know the spectrum matrix A in equation (2.1) when it consists of scalar values and there is one perturbation. When the elements of A are matrices we can do the same. We

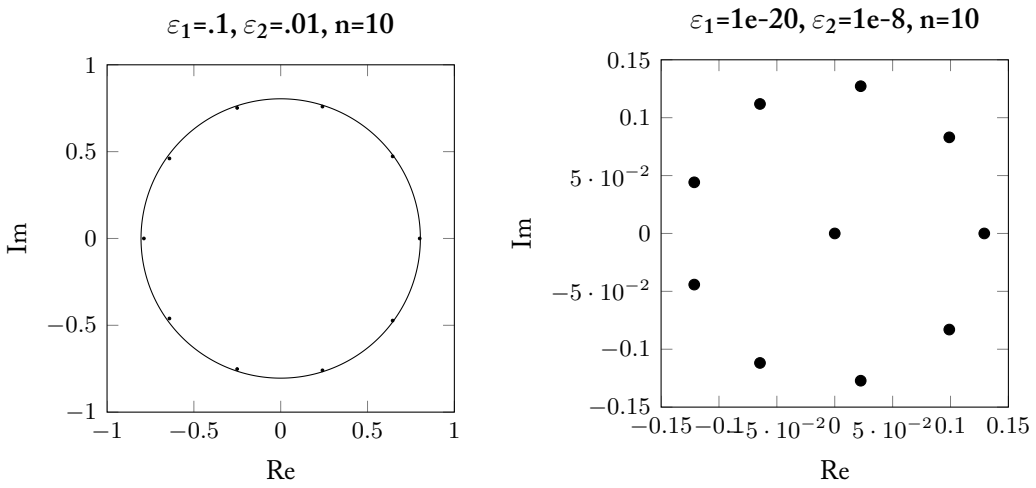

Figure 2.7: Left: the region of section 2.5.2 compared to the actual eigenvalues for $n = 10, \varepsilon_1 = 0.1, \varepsilon_2 = 0.01$. Solid line is r, the dotted line is the circle with radius 1. Right: actual eigenvalues for $n = 10, \varepsilon_1 = 10^{-20}, \varepsilon_2 = 10^{-8}$. $r > 1$ so is not shown.

call matrices that consist of other matrices block matrices. When there is no perturbation, we know that the spectrum Λ of an uppertriangular matrix with matrix C on the diagonal consists of the eigenvalues of C.

When there is a perturbation E in the left bottom of our matrix the situation is more complex. In section 2.6.1 we first work out some theories which help us to find the determinant of block matrices. In subsection 2.6.2 we use these theories to calculate the characteristic polynomial in case there is a perturbation E.

2.6.1 Theories

We begin by summarising some results for block matrices. We will put them all below each other.

Lemma 2.2 ([Lie02])

$$\det \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} = \det(A_{11}) \det(A_{22}).$$

Theorem 2.5 *Upper triangular block matrix*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

is invertible if and only if submatrices A_{11} and A_{22} are invertible.

Proof *A is invertible if and only if* $\det(A) \neq 0$. *We know by Lemma 2.2 that* $\det(A) = \det(A_{11}) \det(A_{22})$. *That means that* $\det(A_{11}) \neq 0$ *and* $\det(A_{22}) \neq 0$ *and thus that A_{11} and A_{22} are invertible.*

The other implication is very similar. If A_{11} and A_{22} are invertible, $\det(A_{11}) \neq 0$ *and* $\det(A_{22}) \neq 0$. *Since* $\det(A) = \det(A_{11})\det(A_{22})$, *it directly follows that* $\det(A) \neq 0$ *and thus that A is invertible.* ■

Theorem 2.6 *The inverse of block matrix $A \in \mathbb{C}^{n\times n}$, when A_{11}, A_{12}, A_{22} are invertible, is*

$$A^{-1} := \begin{bmatrix} A_{11} & A_{12} \\ 0_{n-k,k} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0_{n-k,k} & A_{22}^{-1} \end{bmatrix}.$$

Proof *We know that $AA^{-1} = I$ and we can easily verify that our A^{-1} is correct by*

$$\begin{bmatrix} A_{11} & A_{12} \\ 0_{n-k,k} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0_{n-k,k} & A_{22}^{-1} \end{bmatrix} = \begin{bmatrix} I & -A_{12}A_{22}^{-1} + A_{12}A_{22}^{-1} \\ 0_{n-k,k} & I \end{bmatrix} = I.$$ ■

Theorem 2.7 *Let $C, D \in \mathbb{C}^{k\times k}$, with C invertible. Let $A \in \mathbb{C}^{nk\times nk}$ be defined by*

$$A = \begin{bmatrix} C & D & & \\ & \ddots & \ddots & \\ & & \ddots & D \\ & & & C \end{bmatrix}. \tag{2.9}$$

Then all values on the mth superdiagonal of A^{-1} are

$$(-1)^m (C^{-1}D)^m C^{-1}, m = 0, \ldots, n-1.$$

Proof *We define $B := A^{-1}$ and partition A as*

$$A_n := \left[\begin{array}{c|c} C & G \\ \hline 0_{nk-k,k} & A_{n-1} \end{array}\right] = \left[\begin{array}{c|ccc} C & D & 0 \cdots & 0 \\ \hline 0 & C & D & \\ \vdots & & \ddots & \ddots \\ & & & D \\ 0 & & & C \end{array}\right].$$

If we partition B the same as A we can see by Theorem 2.6 that (since $B_{n2} = -C^{-1}GA_{n-1}^{-1}$)

$$A_n^{-1} = B_n = \left[\begin{array}{c|c} B_{n1} & B_{n2} \\ \hline B_{n3} & B_{n4} \end{array}\right] = \left[\begin{array}{c|c} C^{-1} & -C^{-1}\begin{bmatrix} D & 0 & \cdots & 0 \end{bmatrix} A_{n-1}^{-1} \\ \hline 0_{nk-k,k} & A_{n-1}^{-1} \end{array}\right].$$

Since A_{n-1} has the same structure as A it follows in the same way that if $A_{n-2} \in \mathbb{C}^{k(n-2)\times(n-2)k}$

$$A_{n-1}^{-1} = \left[\begin{array}{c|c} C^{-1} & -C^{-1}\begin{bmatrix} D & 0 & \cdots & 0 \end{bmatrix} A_{n-2}^{-1} \\ \hline 0_{(n-2)k,k} & A_{n-2}^{-1} \end{array}\right].$$

we can generalise this to for all $m = n, \dots, 2$ *and find*

$$A_m^{-1} = \left[\begin{array}{c|c} B_{m1} & B_{m2} \\ \hline B_{m3} & B_{m4} \end{array}\right] = \left[\begin{array}{c|c} C^{-1} & -C^{-1}\begin{bmatrix} D & 0 & \cdots & 0 \end{bmatrix} A_{m-1}^{-1} \\ \hline 0_{(m-1)k,k} & A_{m-1}^{-1} \end{array}\right].$$

If we continue this till we keep a $2k \times 2k$ *matrix we get*

$$A_2^{-1} = \begin{bmatrix} C & D \\ 0 & C \end{bmatrix}^{-1} = \begin{bmatrix} C^{-1} & -C^{-1}DC^{-1} \\ 0 & C^{-1} \end{bmatrix}.$$

Hence B_{32} *becomes*

$$\begin{aligned} B_{32} &= -C^{-1}\begin{bmatrix} D & 0 \end{bmatrix} A_2^{-1} \\ &= -C^{-1}\begin{bmatrix} D & 0 \end{bmatrix} \begin{bmatrix} C^{-1} & -C^{-1}DC^{-1} \\ 0 & C^{-1} \end{bmatrix} \\ &= -C^{-1}D\begin{bmatrix} C^{-1} & -C^{-1}DC^{-1} \end{bmatrix} \\ &= -C^{-1}D\begin{bmatrix} C^{-1} & B_{22} \end{bmatrix}. \end{aligned}$$

So we can generalise this and we can see this becomes a recursion for B_{m2}*:*

$$\begin{aligned} B_{m2} &= -C^{-1}\begin{bmatrix} D & 0 & \cdots & 0 \end{bmatrix} A_{m-1}^{-1} \\ &= -C^{-1}D\begin{bmatrix} C^{-1} & B_{(m-1)2} \end{bmatrix} \\ &= \dots \\ &= -C^{-1}D\begin{bmatrix} C^{-1} & -C^{-1}DC^{-1} & \cdots & (-1)^{m+1}(C^{-1}D)^{n-m-1}C^{-1} \end{bmatrix}. \end{aligned}$$ ■

Theorem 2.8

$$\begin{aligned} \det\begin{bmatrix} P & Q \\ R & S \end{bmatrix} &= \det(P)\det(S - RP^{-1}Q) \quad (\textit{if } P \textit{ is invertible}). \\ &= \det(S)\det(P - QS^{-1}R) \quad (\textit{if } S \textit{ is invertible}). \end{aligned}$$

Proof *We can write*

$$\begin{aligned} \begin{bmatrix} P & Q \\ R & S \end{bmatrix} &= \begin{bmatrix} P & 0 \\ R & I \end{bmatrix} \begin{bmatrix} I & P^{-1}Q \\ 0 & S - RP^{-1}Q \end{bmatrix} \\ &= \begin{bmatrix} I & Q \\ 0 & S \end{bmatrix} \begin{bmatrix} P - QS^{-1}R & 0 \\ S^{-1}R & I \end{bmatrix} \end{aligned}$$

and the result follows from Lemma 2.2. ■

2.6.2 Eigenvalues

To calculate the eigenvalues of our matrix (2.9) with some disturbance E,

$$\begin{bmatrix} C & D & & \\ & \ddots & \ddots & \\ & & & D \\ E & & & C \end{bmatrix}, \tag{2.10}$$

we first compute its characteristic polynomial

$$\chi(\lambda) = \det \left[\begin{array}{c|ccccc} C-\Omega & D & 0 & \cdots & 0 \\ \hline 0 & C-\Omega & D & & \\ \vdots & & \ddots & \ddots & \\ 0 & & & & D \\ E & & & & C-\Omega \end{array}\right], \ \Omega = \mathrm{diag}_k(\lambda), \ E \in \mathbb{C}^{k\times k}. \tag{2.11}$$

We simplify this by writing $\bar{C} = C - \Omega$ and use Theorem 2.8 to get

$$\det \left[\begin{array}{c|c} P & Q \\ \hline R & S \end{array}\right] = \det(S)\det(P - QS^{-1}R). \tag{2.12}$$

To calculate the determinant of (2.12) we need to know S^{-1}. Since

$$S^{-1} = \begin{bmatrix} C^{-1} & -C^{-1}DC^{-1} & \cdots & (-1)^{n-1}(C^{-1}D)^{n-1}C^{-1} \\ & \ddots & \ddots & \end{bmatrix}$$

by Theorem 2.7, we know that $S^{-1}_{1(n-1)} = (-1)^{n-1}(\bar{C}^{-1}D)^{n-2}\bar{C}^{-1}$ and thus that

$$\begin{aligned} QS^{-1}R &= \begin{bmatrix} D & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} S^{-1}_{11} & \cdots & S^{-1}_{1(n-1)} \\ \vdots & \ddots & \vdots \\ S^{-1}_{(n-1)1} & \cdots & S^{-1}_{(n-1)(n-1)} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ E \end{bmatrix} = DS^{-1}_{1(n-1)}E \\ &= (-1)^{n-1}D(\bar{C}^{-1}D)^{n-2}\bar{C}^{-1}E. \end{aligned} \tag{2.13}$$

After filling $\det(S) = \det(\bar{C})^{n-1}$ and (2.13) into (2.12) we see that the eigenvalues of matrix (2.9) with some disturbance E are the zeros of

$$\chi(\lambda) = \det(\bar{C})^{n-1}\det(\bar{C} + (-1)^n D(\bar{C}^{-1}D)^{n-2}\bar{C}^{-1}E) \tag{2.14}$$

$$= \det(\bar{C})^{n-1}\det(\bar{C} + (-1)^n(\bar{C}^{-1}D)^{n-1}E). \tag{2.15}$$

The number of zeros depends on the structure of E. It is possible all eigenvalues are different than the eigenvalues of A. But for some E, the perturbed and unperturbed matrix share some eigenvalues. The eigenvalue that C are the perturbed matrix have in common is the most positive real one. What happens is visible in Figure 2.8. Here it is visible that the eigenvalue in (0,0) is shared.

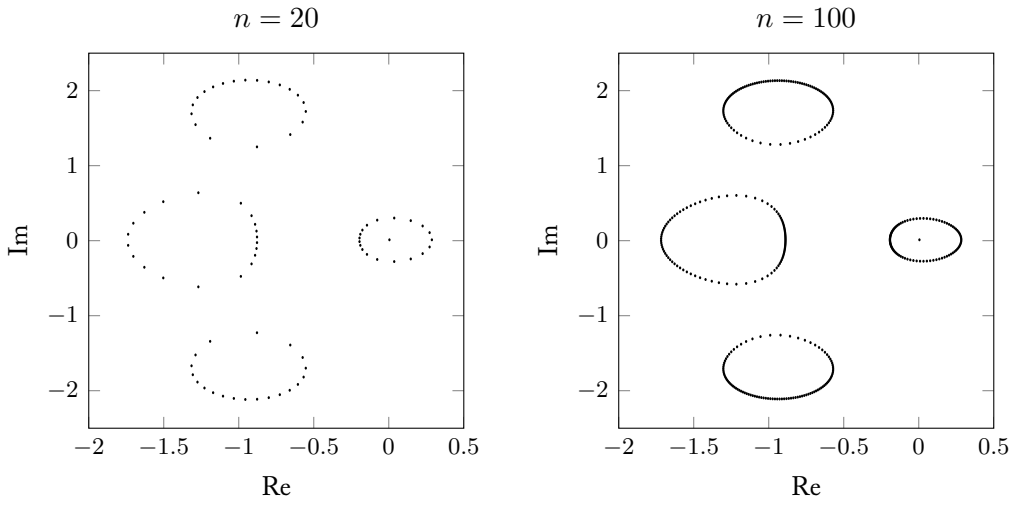

Figure 2.8: Solutions of the characteristic polynomial for block matrices (2.15) for different sizes. With matrices from [Fir12][(3.15)] with $k_p = 1, k_d = 2, k_{dd} = 0, h = 2, \tau = 0.38$.

Spectrum of an operator 3

In Chapter 2 we inspected the eigenvalues of our Jordan matrix when the size was finite. To know the relation between the finite and infinite dimensional case, we inspect the infinite dimensional case in this chapter.

First we need to know what happens when the size of our matrix grows to infinity. We start again with a single Jordan block with zeros on its diagonal. So we have the matrix $T_n \in \mathbb{C}^{(2n+1)\times(2n+1)}$, with $n \in \mathbb{N}$, that is defined as

$$T_n = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & 0 \end{bmatrix}_{1+2n\times 2n+1} . \tag{3.1}$$

The eigenvalues of this matrix are obviously 0, but what is T_n when $n \to \infty$? To show this, we first look at $T_n f$ for $n \to \infty$. We start with the sequence of equations $g = T_n f$, with $f, g \in \mathbb{C}^{2n+1}$. We choose to represent f as

$$f = (f_{-n}, \ldots, f_{-1}, f_0, f_1, \ldots, f_n)^T$$

and we see that

$$g_k = \begin{cases} f_{k+1} & \text{if } k = -n, \ldots, n-1 \\ 0 & \text{if } k = n \end{cases} . \tag{3.2}$$

The Eucledian norm of f is $||f||^2 = \sum_{k=-n}^{n} |f_k|^2$ and if we let $n \to \infty$, we see that $||f||$ converges to the ℓ^2-norm of f. Note that the same claim holds for g. Letting $n \to \infty$ in (3.2) we arrive at the equation

$$g_k = f_{k+1}, k \in \mathbb{Z}. \tag{3.3}$$

So we see that when $n \to \infty$ the matrix (3.1) becomes the operator on ℓ^2 defined by (3.3). However, we cannot compute the eigenvalues like we did in the finite dimensional case. In the next section we explain what the spectrum of an operator is and how to calculate it.

3.1 Spectrum of an operator

In the finite dimensional case we checked for which λ the determinant of the matrix $A-\lambda I$ was zero. Then we knew when $A-\lambda I$ was not invertible. In the infinite dimensional case we do the same. The spectrum of the operator $\mathbb{S}$ is defined as the set

$$\Lambda(\mathbb{S}) := \{z \in \mathbb{C} \mid (zI - \mathbb{S}) \text{ not invertible}\}. \tag{3.4}$$

For $f, g \in \ell^2$ the operator $\lambda I - \mathbb{S}$ is invertible if for all g there exists a unique f such that $(\lambda I - \mathbb{S})f = g$. If the operator $\mathbb{S}$ is not invertible then λ is part of the spectrum.

Definition 3.1 (Regular value) *All values that are not contained in the spectrum are called the regular values of an operator. We write them as* $\rho(\mathbb{S})$.

3.2 Spectrum of the bilateral shift operator

In the introduction of this chapter we derived the operator $g_k = f_{k+1}$. This is called bilateral left shift operator $\mathbb{S}$. So if we have a sequence $f = (\ldots, f_{-1}, f_0, f_1, \ldots)$, then $\mathbb{S}(f) = (\ldots, f_0, f_1, f_2, \ldots)$. Hence $\mathbb{S}$ moves all items one places to the left.

Given the discussion in the previous section and since $\mathbb{S}$ is the bilateral shift operator, to compute the spectrum of $\mathbb{S}$ we have to show that for a given λ there exists a unique f for every g such that

$$\lambda f_n - f_{n+1} = g_n,\ n \in \mathbb{Z}. \tag{3.5}$$

Then we have the set of all regular values, $\rho(\mathbb{S})$, and the spectrum is the remaining set $\mathbb{C}\backslash\rho(\mathbb{S}) = \Lambda(\mathbb{S})$.

We will show that the f_n solving (3.5) is given by a convolution and therefore we will use Theorem 3.1, and especially inequality (3.9) in our proof. The p-norm of f is defined as

$$||f||_p = \left(\sum_{k=-\infty}^{\infty} |f_k|^p \right)^{\frac{1}{p}},\ 1 \leq p < \infty. \tag{3.6}$$

Theorem 3.1 *Suppose* $f \in \ell^1, g \in \ell^p$. *Then*

$$\sum_{k=-\infty}^{\infty} |f_{n-k} g_k|^p < \infty \tag{3.7}$$

for all $n \in \mathbb{Z}$. *Define the convolution of* f *and* g *(notation* $h = f * g$*) as*

$$h_n = \sum_{k=-\infty}^{\infty} f_{n-k} g_k,\ n \in \mathbb{Z}. \tag{3.8}$$

Then $h \in \ell^p(\mathbb{Z})$ *and*

$$||h||_p \leq ||f||_1 ||g||_p. \tag{3.9}$$

Proof *See Appendix C.1.* ■

3.2.1 Regular values

We first look for all regular values. So we need to find λ such that there exists a unique f for every g in the equation (3.5). We start to prove that $|\lambda| < 1$ and $|\lambda| > 1$ are regular values of our operator.

Let's assume that $|\lambda| < 1$. We then can work out the recurrence of $(\lambda I - \mathbb{S})f = g$ as

$$\begin{aligned} f_n &= \lambda f_{n-1} - g_{n-1} = \lambda(\lambda f_{n-2} - g_{n-2}) - g_{n-1} = \dots \\ &= -\sum_{k=1}^{N} \lambda^{k-1} g_{n-k} + \lambda^N f_{n-N}. \end{aligned} \tag{3.10}$$

This motivates us to choose h_k as

$$h_k = \begin{cases} \lambda^{k-1} & \text{for } k \geq 1 \\ 0 & \text{for } k < 1 \end{cases}, \tag{3.11}$$

and then define a candidate solution $f = (f_n)$ for $(\lambda I - \mathbb{S})f = g$ as

$$f_n = -\sum_{k=1}^{\infty} \lambda^{k-1} g_{n-k} = -\sum_{k=-\infty}^{\infty} h_k g_{n-k} = -\sum_{l=-\infty}^{\infty} h_{n-l} g_l. \quad (l = n-k) \tag{3.12}$$

Since we assumed that $|\lambda| < 1$ we have $||h||_1 = \sum_{k=-\infty}^{\infty} |h_k| = \sum_{k=1}^{\infty} |\lambda^k| < \infty$. Since $g \in \ell^2$, we then can use Theorem 3.1 to show that

$$||f||_2^2 = \sum_{n=-\infty}^{\infty} |f_n|^2 = \sum_{n=-\infty}^{\infty} \left| \sum_{l=-\infty}^{\infty} h_{n-l} g_l \right|^2 \leq ||h||_1 ||g||_2 < \infty.$$

We therefore see that for $|\lambda| < 1$ and $g \in \ell^2$, (3.12) is a solution of $(\lambda I - \mathbb{S})f = g$. We still need to prove that is unique, but we will do that in the next section.

If $|\lambda| > 1$, then we cannot say anything from (3.12) since the sum of λ^k diverges. But we can solve the equation $(\lambda I - \mathbb{S})f = g$ forward. Therefore we introduce

$$h_k := \begin{cases} \frac{1}{\lambda^{-k+1}} & \text{for } k \leq 0 \\ 0 & \text{for } k > 0 \end{cases}, \tag{3.13}$$

and see that

$$\begin{aligned} f_n &= \frac{g_n + f_{n+1}}{\lambda} = \frac{g_n}{\lambda} + \frac{1}{\lambda}\left(\frac{g_{n+1} + f_{n+2}}{\lambda}\right) = \dots \\ &= \sum_{k=0}^{N} \frac{g_{n+k}}{\lambda^{k+1}} + \frac{f_{n+N+1}}{\lambda^{N+1}}. \end{aligned} \tag{3.14}$$

This motivates us to choose the infinite sum

$$f_n = \sum_{k=0}^{\infty} \frac{g_{n+k}}{\lambda^{k+1}} = \sum_{k=-\infty}^{\infty} h_{-k} g_{n+k} = \sum_{m=-\infty}^{\infty} h_{n-m} g_m \quad (m = n+k) \tag{3.15}$$

as a candidate solution for $(\lambda I - \mathbb{S})f = g$

Since $|\lambda| > 1$ we have that $||h||_1 = \sum_{k=0}^{\infty} |\lambda^{-(k+1)}| < \infty$. Since $g \in \ell^2$, we can use Theorem 3.1 to prove that

$$||f||_2^2 = \sum_{n=-\infty}^{\infty} |f_n|^2 = \sum_{n=-\infty}^{\infty} \left| \sum_{m=-\infty}^{\infty} h_{n-m} g_m \right|^2 \leq ||h||_1 ||g||_2 < \infty. \tag{3.16}$$

We therefore see that for $|\lambda| > 1$ and $g \in \ell^2$, (3.15) is a solution of $(\lambda I - \mathbb{S})f = g$. We still need to prove that is unique, but we will do that in the next section.

3.2.1.1 Uniqueness

We now have proved that for $|\lambda| > 1$ and $|\lambda| < 1$ there exists a $f \in \ell^2$ for every $g \in \ell^2$. But for every g this f also needs to be unique for $\lambda \in \rho(\mathbb{S})$.

To check the uniqueness of (3.12) we have to show that from

$$\lambda f_n - f_{n+1} = 0 \tag{3.17}$$

it follows that $f = 0$. We assume the opposite, that $f \neq 0$. We therefore first assume that $f_0 \neq 0$ and from (3.17) we get the recursion

$$f_1 = \lambda f_0,\ f_2 = \lambda f_1 = \lambda^2 f_0,\ \ldots,\ f_n = \lambda^n f_0.$$

If $|\lambda| \geq 1$, then it follows that $f_n \neq 0$ for all n and

$$\begin{aligned} \|f\|^2 &\geq \sum_{n=0}^{\infty} |f_n|^2 \\ &= \sum_{n=0}^{\infty} |\lambda|^{2n} |f_0|^2 = \infty. \end{aligned}$$

But since $f_0 \neq 0$ we find $f \notin \ell^2$. Therefore we see our assumption is false and we see that $g = 0$ implies $f = 0$ in (3.12).

But if $f_0 = 0$ we see without loss of generality that it directly follows that all f_n with $n \geq 1$ are zero.

To show that also for $|\lambda| < 1$ the solution is unique we have to show that from equation (3.17) it follows that $f = 0$. This time we solve it in the other direction and we then get

$$f_{-1} = \frac{f_0}{\lambda},\ f_{-2} = \frac{f_{-1}}{\lambda} = \frac{f_0}{\lambda^2},\ \ldots,\ f_{-n} = \frac{f_0}{\lambda^n}.$$

Let's again assume the opposite, that $f = 0$. And that we therefore again assume that $f_0 \neq 0$.

If $|\lambda| \leq 1$ then it follows that $f_n \neq 0$ for all n, since

$$\begin{aligned} \|f\|^2 &\geq \sum_{n=0}^{\infty} |f_{-n}|^2 \\ &= \sum_{n=0}^{\infty} \frac{|f_0|^2}{|\lambda|^{2n}} = \infty. \end{aligned}$$

But since $f_0 \neq 0$ we find $f \notin \ell^2$. Therefore we see our assumption is false and we see that $g = 0$ implies $f = 0$ in (3.15).

But if $f_0 = 0$ we see without loss of generality that it directly follows that all f_n with $n < 0$ are zero.

3.2.2 Spectrum

We proved that $|\lambda| > 1$ and $|\lambda| < 1$ are not part of the spectrum, so we expect that $|\lambda| = 1$ is the complete spectrum. The last step is to prove that this is indeed the case.

If we look again at (3.10) and choose $n = 0$, we get

$$|f_0 - \lambda^N f_{-N}| = -\sum_{k=1}^{N} \lambda^{k-1} g_{-k}. \tag{3.18}$$

If we write $\lambda = |\lambda| e^{i\theta} = e^{i\theta}$ and choose

$$g_k = \frac{e^{i\theta(k+1)}}{|k|+1}, k \in \mathbb{Z},$$

we see that $g_k \in \ell^2$, since $|e^{i\theta}| = 1$. It follows that

$$|f_0 - e^{i\theta N} f_{-N}| = \left| -\sum_{k=1}^{N} \lambda^{k-1} g_{-k} \right| = \left| \sum_{k=1}^{N} \frac{e^{i\theta(k-1)}}{|k|+1} (e^{i\theta})^{-k+1} \right| = \sum_{k=1}^{N} \frac{1}{|k|+1}.$$

We know that this sum diverges if we take the limit $N \to \infty$. But if $f \in \ell^2$, then f_0 and f_{-N} must be finite. Therefore we cannot find a $f \in \ell^2$ such that (3.18) is valid. So $|\lambda| = 1$ is not a regular value and thus part of the spectrum of $\mathbb{S}$.

Concluding, the spectrum of $\mathbb{S}$ equals $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$.

3.3 Spectrum of Operator corresponding to a single Jordan block

Like we did in the introduction of this chapter, we will will in a similar way look at a single Jordan block with ones on its diagonal when the size becomes infinite. So we have the matrix $V_n \in \mathbb{C}^{(2n+1)\times(2n+1)}$, with $n \in \mathbb{N}$, that is defined as

$$V_n = \begin{bmatrix} 1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}_{2n+1 \times 2n+1}. \tag{3.19}$$

The eigenvalues of this matrix are obviously 1, but what happens when $n \to \infty$? We see that $V_n = T_n + I$, with T_n from (3.1). So

$$\mathbb{S}_{new} = I + \mathbb{S}.$$

Then we arrive at the equation

$$g_k = f_k + f_{k+1}, k \in \mathbb{Z}. \tag{3.20}$$

In section 3.2 we already computed the spectrum for a similar mapping and we can use this to inspect

$$\lambda f_k - (f_k + f_{k+1}) = (\lambda - 1)f_k - f_{k+1} = g_k,\ k \in \mathbb{Z}.$$

If we follow the same steps as in the previous section, but with $\lambda - 1$ instead of λ, we see our spectrum is $|\lambda - 1| = 1$. This means that – compared to the spectrum of $\mathbb{S}$ – the spectrum is shifted 1 to the right. Therefore the spectrum of (3.20) is a circle in the complex plane with center 1 and radius 1.

3.4 Spectrum of the z-transformation

Till now we computed the spectrum of two operators. In the introduction in Chapter 1 we already derived the operator corresponding to the block matrix A, see (1.1). This operator

$$g_k = Cf_k + Df_{k+1}, k \in \mathbb{Z}. \tag{3.21}$$

If we apply the Fourier Transform (see Appendix B) we get the transformation

$$\check{z}(\theta) = (\underbrace{C + De^{-i\theta}}_{\check{A}(\theta)})\check{x}(\theta), \text{where } \theta \in (0, 2\pi) \tag{3.22}$$

and $\check{x}$ is the Fourier-transform of x. The mathematics is explained in Appendix B. To compute the spectrum of this operator we need a lemma from [CIZ09].

Definition 3.2 $L_\infty(\mathbb{D}_1; \mathbb{C}^{k\times q}) = \{F : \mathbb{D}_1 \to \mathbb{C}^{k\times q} \mid F \textit{ is measurable and } ||F||_\infty = \operatorname{ess\,sup}_{0<\theta<2\pi} ||F(e^{i\theta})|| < \infty\}$, *where* $\mathbb{D}_1$ *is the unit circle* $\{z \in \mathbb{C} \mid |z| = 1\}$.

Lemma 3.1 ([CIZ09, A. 3]) $L_\infty(\mathbb{D}_1; \mathbb{C}^{k\times k})$ *is a Banach algebra and* $F \in L_\infty(\mathbb{D}_1; \mathbb{C}^{k\times k})$ *is boundedly invertible if and only if there exists a* $\gamma > 0$ *such that* $\{\theta : |\det(F(e^{i\theta}))| < \gamma\}$ *has measure zero. If* F *is continuous, then* $L_\infty(\mathbb{D}_1; \mathbb{C}^{k\times k})$ *is boundedly invertible if and only if* $\det(F(e^{i\theta})) \neq 0$ *for all* $\theta \in [0, 2\pi]$.

This implies that when $\check{A}$ in (3.22) is continuous its spectrum can be written like [CIZ09]

$$\Lambda(\check{A}) = \bigcup_{\theta\in[0,2\pi]} \Lambda\left(\check{A}(e^{i\theta})\right),$$

where, for a fixed θ, $\lambda \in \Lambda(\check{A}(e^{i\theta}))$ if and only if $\det\left(\lambda I - \check{A}(e^{i\theta})\right) = 0$.

This means that for every θ we get a ordinary matrix for which we can compute the eigenvalues. The spectrum of A is the union of all these eigenvalues.

If we choose $C = 0, D = 1$ in (3.21), we get the shift operator from section 3.2. We computed there that the spectrum is the circle with radius 1 and the origin as center. If we inspect (3.22) with $C = 0, D = 1$, we see $\check{A}(\theta) = e^{i\theta}$. In general, it is true that the spectrum of the shift operator stays the same after we apply the z-transformation:

$$\Lambda(\check{A}) = \Lambda(A).$$

3.5 Other kinds of spectra

In the finite dimensional case the spectrum was just the collection of all eigenvalues. In Section 3.1 we defined the spectrum as the set where the operator is not invertible.

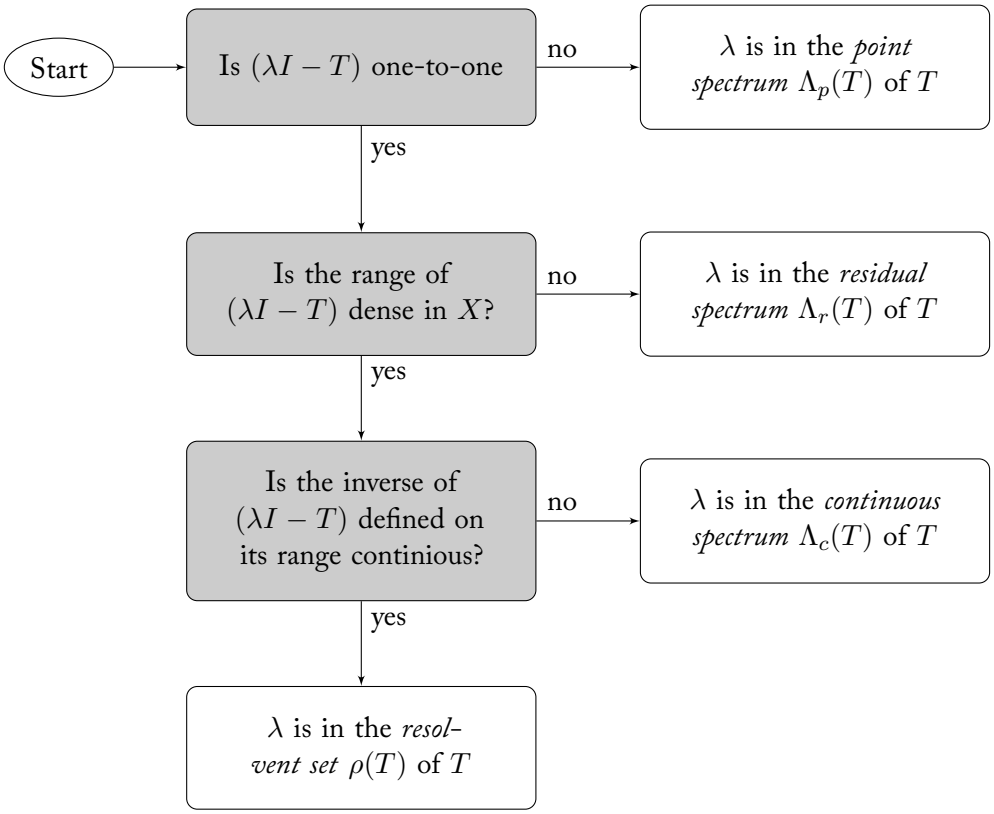

FIGURE 3.1: Flow chart of the spectrum of an operator T (from [NS71, p. 413])

In the finite case, only the point spectrum Λ_p was non-empty. That was the set of eigenvalues. Now we also have the continuous spectrum Λ_c and the residual spectrum Λ_r. The spectrum $\Lambda(T)$ of an operator is the union of those three and can then be written as

$$\Lambda(T) = \Lambda_p(T) \cup \Lambda_c(T) \cup \Lambda_r(T).$$

See e.g. [NS71, p. 412]. The point spectrum $\Lambda_p(T)$ consists of all the eigenvalues of T, the continuous spectrum contains the scalars that are not eigenvalues but make the range of $T - \lambda$ a proper dense subset of the space and the residual spectrum consists of all other scalars in the spectrum. This is also illustrated in Figure 3.1.

If we look at the spectrum of $\check{A}$ of Section 3.4 then in general the point spectrum is empty. Thus $\Lambda(\check{A})$ contains no eigenvalues in most situations.

We also proved already in section 3.2 that if λ is part of the spectrum of the shift operator and there exists a solution, it is also unique. But it is also possible that we find a solution outside of the space we are working in, so not for all $\check{y} \in \ell^2$ there exists a $\check{x} \in \ell^2$.

Another way of looking the problem is that if $\lambda \in \Lambda(\check{A})$ then $(\lambda I - \check{A})^{-1}$ is unbounded.

Definition 3.3 (Approximate eigenvalue) *λ is an approximate eigenvalue if there exists a sequence x_n with $\|x_n\| = 1$ such that*

$$\lim_{n\to\infty} \|\mathbb{S}x_n - \lambda x_n\| = 0.$$

We know that our bilateral shift operator $\mathbb{S}$ has no eigenvalues, but its spectrum equals $|\lambda| = 1$. We can show that $|\lambda| = 1$ is also an approximate eigenvalue.

If $\lambda < 1$, then

$$\big\|\mathbb{S}f - \lambda f\big\| \geq \big|\, \|\mathbb{S}f\| - |\lambda|\, \|f\| \,\big| = \big|1 - |\lambda|\,\big|\, \|f\|$$

and thus bounded from below and thus not part of the approximate point spectrum. Since $\|\mathbb{S}\| = 1$, we know that the spectrum (and approximate point spectrum) is contained within the ball of radius 1, therefore we do not need to check $\lambda > 1$. In general we know that the an approximate eigenvalue is a subset of the spectrum.

Let $|\lambda| = 1$ and let x_n be the vector

$$\frac{1}{\sqrt{n}}(\dots, 0, 1, \lambda^{-1}, \lambda^{-2}, \dots, \lambda^{1-n}, 0, \dots)$$

then $||x_n|| = 1$ and

$$\|\mathbb{S}x_n - \lambda x_n\| = \sqrt{\frac{2}{n}}.$$

This holds for all n, thus the approximate point spectrum is $|\lambda| = 1$.

3.6 Spectrum Operator with scalar perturbation

We want to compare the spectra corresponding to

$$J = \begin{bmatrix} \ddots & \ddots & & & \\ & 0 & 1 & & \\ & & \ddots & 1 & \\ & & & 0 & \ddots \\ & & & & \ddots \end{bmatrix} \text{ and } J_p = \begin{bmatrix} \ddots & \ddots & & & \\ & 0 & 1 & & \\ & & \ddots & 1 & \\ \varepsilon & & & 0 & \ddots \\ & & & & \ddots \end{bmatrix}.$$

We showed in Chapter 2 that from the relations

$$f_n = \lambda f_{n-1} - g_{n-1},\ f_n = \frac{g_n + f_{n+1}}{\lambda}$$

if followed that $|\lambda| < 1$ and $|\lambda| > 1$ are not part of the spectrum of J. Finally we showed that for $|\lambda| = 1$ we could not find a bounded sequence in our space for which the equations are valid. Now there is an extra ε in J_p. Therefore we have when $m = n$ an extra equation and the spectrum is more complex:

$$f_m = \begin{cases} \lambda f_{m-1} - g_{m-1} & m = 1, \dots, n-1 \\ \lambda f_{m-1} + \varepsilon f_1 & m = n \end{cases}. \tag{3.23}$$

We can write our problem as $\mathbb{S} + \varepsilon\Delta_{1,m}$ with $\Delta_{a,b}$ defined as in section 1.3. If we again let $|\lambda| = 1$ and let x_n be the vector

$$\frac{1}{\sqrt{n}}(\dots, 0, 1, \lambda^{-1}, \lambda^{-2}, \dots, \lambda^{1-n}, 0, \dots)$$

then $||x_n|| = 1$ and

$$\begin{aligned} \|(\mathbb{S} + \varepsilon\Delta_{1,m})x_n - \lambda x_n\| &= \|\mathbb{S}x_n - \lambda x_n + \varepsilon\Delta_{1,m}x_n\| && (3.24)\\ &\leq \|\mathbb{S}x_n - \lambda x_n\| + \|\varepsilon\Delta_{1,m}x_n\| && (3.25)\\ &= \sqrt{\frac{2}{n}} + \varepsilon. && (3.26) \end{aligned}$$

If we let $n \to \infty$, this goes to ε. So we need more to show what the approximate point spectrum is.

If we look at $(\mathbb{S} + \Delta_{1,m}\varepsilon)x_{\text{ext}}$, where the values in red are extended, we see that

$$(\mathbb{S} + \Delta_{1,m}\varepsilon)x_{\text{ext}} = \begin{bmatrix} -\lambda & 1 & & & & \\ & -\lambda & 1 & & & \\ & & \ddots & \ddots & & \\ & & & & 1 & \\ & \varepsilon & & & -\lambda & 1 \\ & & & & & -\lambda \end{bmatrix} \begin{bmatrix} 0 \\ x_1 \\ \vdots \\ \vdots \\ x_n \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ \vdots \\ 0 \\ -\lambda x_n \end{bmatrix}. \qquad (3.27)$$

Since $\|x\| = 1$, we need

$$\begin{aligned} x_1 &= \frac{1}{\sum_{k=1}^{n} |\lambda|^{-k}} = \frac{1}{\frac{|\lambda|^n - 1}{|\lambda|^n(|\lambda|-1)}} = \frac{|\lambda|^n(|\lambda| - 1)}{|\lambda|^n - 1}, \\ x_n &= \lambda^{1-n}. \end{aligned}$$

We then get

$$\begin{aligned} \|(\mathbb{S} + \varepsilon\Delta_{1,m})x_{\text{ext}}\|^2 &= |x_1|^2 + |\lambda x_n|^2 = |x_1|^2 + |\lambda|^{-2n} \\ &= \frac{|\lambda|^n(|\lambda| - 1)}{|\lambda|^n - 1} + |\lambda|^{-2n} \\ &= \frac{|\lambda| - 1}{1 - 1/|\lambda|^n} + |\lambda|^{-2n}. \end{aligned}$$

We know from Example 2.1 that $\lambda = \sqrt[n]{\varepsilon}$. In the same example we can see that if n becomes big, the angle between the eigenvalues is small. We can choose a sequence of λ (depend on n) such that $\lambda \to \mu, |\mu| = 1$, when $n \to \infty$. So we see that we can choose the unit circle as an approximate point spectrum of our operator, since then $\|(\mathbb{S} + \varepsilon\Delta_{1,m})x_{\text{ext}}\| \to 0$.

3.7 Spectrum of operator with matrix perturbation

We can also prove that perturbations do not influence the spectrum of operators corresponding to the matrices

$$A = \begin{bmatrix} \ddots & \ddots & & & \\ & C & D & & \\ & & \ddots & D & \\ & & & C & \ddots \\ & & & & \ddots \end{bmatrix} \text{ and } F = \begin{bmatrix} \ddots & \ddots & & & \\ & C & D & & \\ & & \ddots & D & \\ & E & & C & \ddots \\ & & & & \ddots \end{bmatrix},$$

where F is perturbed with E at position $F_{1,n}$. We see that the characteristic polynomial of $A + E$ equals

$$(\lambda I - A - E) = (\lambda - A)\left[I + (\lambda I - A)^{-1}E\right].$$

So we see that the characteristic polynomial is a factor $(I + (\lambda I - A)^{-1}E)$ bigger than the characteristic polynomial of A.

Pseudospectra 4

In Chapter 2 we found that the Jordan block is very sensitive to perturbations. Now instead of looking at the eigenvalues, we look at complex numbers that are almost the eigenvalues. We look where the inverse is not infinite, but bigger then ε^{-1}. This is made formal in the following definition.

Definition 4.1 (ε-pseudospectrum [TE05]) *Let $A \in \mathbb{C}^{N\times N}$ and $\varepsilon > 0$ be arbitrary. The ε-pseudospectrum $\Lambda_\varepsilon(A)$ of A is the set of $z \in \mathbb{C}$ such that $||(zI - A)^{-1}|| > \varepsilon^{-1}$.*

In the definition of the pseudospectrum above, we also introduced the resolvent.

Definition 4.2 (Resolvent) *The resolvent of matrix A is $(zI - A)^{-1}$.*

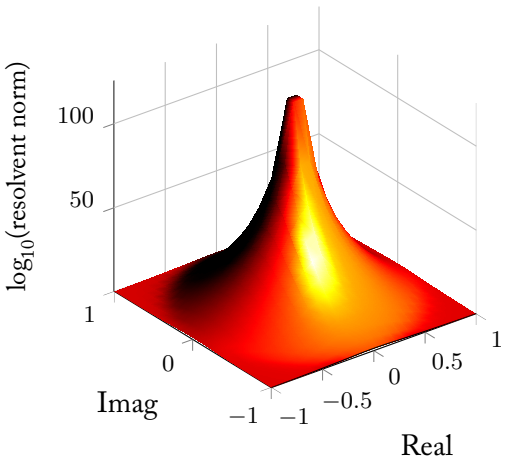

FIGURE 4.1: Plot of the resolvent norm $\|z - J\|^{-1}$. Where J is the Jordan block in (2.1) with $n = 100$.

The norm of the resolvent is plotted in Figure 4.1. For different ε we can plot level curves where $||(zI - A)^{-1}|| = \varepsilon$. If we combine different level curves in one plot, for different values of ε, we get a typical representation of pseudospectra. Figure 4.2 is an example of that.

The best source on pseudospectra is Trefethen and Embree's book *Spectra and Pseudospectra* [TE05]. In this book they describe the history of pseudospectra and a lot of different applications. Trefethen has researched pseudospectra since 1991. In that year he wrote a book chapter [Tre92] in which he analysed the pseudospectra of different matrices. He starts with the Jordan block , because this is the non-normal matrix most sensitive to perturbations. In the same chapter he published pseudospectra (Figure 4.2) of this matrix.

There are three equivalent definitions of the pseudospectrum [TE05, Ch. 2]:

1. $\Lambda_\varepsilon(A) = \{z \in \mathbb{C} \mid \|(zI - A)^{-1}\| \geq \varepsilon^{-1}\}$
2. $\Lambda_\varepsilon(A) = \{z \in \mathbb{C} \mid z \in \Lambda(A + E) \text{ for some } E \text{ with } \|E\| \leq \varepsilon\}$

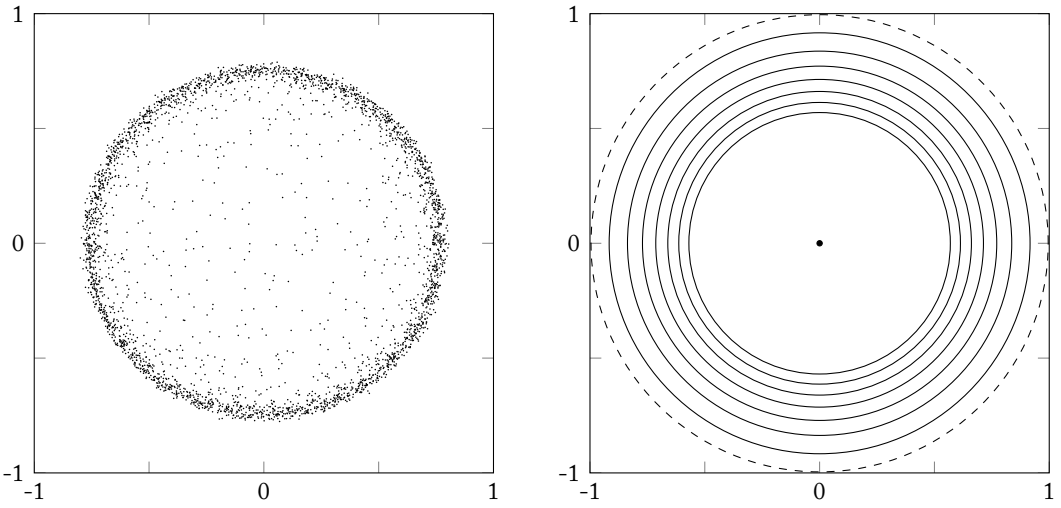

Figure 4.2: [Tre92, p.254] Disturbances from the Jordan canonical form J . Left the eigenvalues of 100 random complex perturbations of J with E ($||E|| = 10^{-3}$), right the pseudospectra of J with $\varepsilon = 10^{-2}, 10^{-3}, \ldots, 10^{-8}$. The dotted line is the numerical range.

3. $\Lambda_\varepsilon(A) = \{z \in \mathbb{C} \mid \text{there exists a } v \in \mathbb{C}^n \text{ with } ||v|| = 1 \text{ such that } \|(A - zI)v\| \leq \varepsilon\}$

Definition 1 is the most intuitive, but definition 2 helps us a lot more to connect the finite to the infinite dimensional case that we described in Chapter 3. We are looking what happens to the eigenvalues when we perturb a matrix since we perturb our matrix with a random matrix E and look at its spectrum.

The complete proof that all definitions are equivalent can be found in [TE05, Ch. 2], but if we know why item 2 implies item 1, it will help us understand what happens when we perturb a matrix with a small matrix E.

If $z \in \Lambda(A+E)$, that means that there exists a v, with $\|v\| = 1$, such that $(A+E)v = zv$. If we move Av to the right and invert $zI - A$, we arrive at $v = (zI - A)^{-1}Ev$. We can use this equation to prove that when $z \in \rho(A)$

$$1 = \|v\| = \|(zI - A)^{-1}Ev\| \leq \|(zI - A)^{-1}\|\|Ev\| \leq \|(zI - A)^{-1}\|\varepsilon.$$

So we see that definition 2 of a pseudospectrum implies also that definition 1 is true. This means that we can perturb a matrix A and learn something about its behaviour.

4.1 Poor man's pseudospectra

We can use definition 2 to build pseudospectra. We then simulate the ε-pseudospectrum by perturbing a matrix E, with $\|E\| < \varepsilon$, and compute the eigenvalues of this perturbed matrix. This gives results that are not as accurate as the algorithms used today, but perturbing the matrices and looking at its eigenvalues was a cheap way of creating pseudospectra plots. Therefore there are called "poor man's pseudospectra". This difference is also visible in Figure 4.2. The contour line for $\varepsilon = 10^{-2}$ is the solid line at the outside of the right

plot. When we perturb J with $\|E\| < 10^{-3}$, we see the perturbed eigenvalues lie inside the 10^{-3}-pseudospectrum.

Trefethen [Tre99] gives three problems with "poor man's pseudospectra". These spectra are of course not as accurate as computing them using definition 1, since we get a cloud of points and not a circle. Secondly, pseudospectra are not essentially about perturbations and by representing them as perturbations they can be easily misinterpreted. Finally, random perturbation to an operator do not make sense. Therefore these days there more accurate algorithms for computing pseudospectra. Modern way's to compute pseudospectra can be found in [Tre99].

4.2 Pseudospectra of operators

We can also define pseudospectra for operators. Trefethen and Embree [TE05][p. 31] define them. They use that for X a Banach-space, $\mathcal{C}(X)$ is the set of all closed operators, $\mathcal{B}(X)$ the set of all bounded operators on X and $\mathcal{D}(x) \subset X$ the domain of X.

Definition 4.3 (ε-pseudospectra for operators [TE05, p. 31]) *Let $A \in \mathcal{C}(X)$ and $\varepsilon > 0$. The ε-pseudospectra $\Lambda_\varepsilon(A)$ of A is the set of $z \in \mathbb{C}$ defined equivalently by any of the conditions*

1. $\|(zI - A)^{-1}\| > \varepsilon^{-1}$
2. $z \in \Lambda(A + E)$ *for some* $E \in \mathcal{B}(X)$ *with* $\|E\| < \varepsilon$
3. $z \in \Lambda(A)$ *or* $\|(zI - A)u\| < \varepsilon$ *for some* $u \in \mathcal{D}(A)$ *with* $\|u\| = 1$.

4.3 Numerical range

In Figure 4.2 we already saw the numerical range as a dotted line. The numerical range is a closed convex hull of the spectrum $\Lambda(A)$. The numerical range of a matrix A is defined as

$$W(A) = \left\{ \frac{x^*Ax}{x^*x} \mid x \in \mathbb{C}^n;\ x \neq 0 \right\} = \{\, x^*Ax \mid x \in \mathbb{C}^n;\ \|x\| = 1\} \tag{4.1}$$

Example 4.1 *Let's take the most simple Jordan Block and a vector x with $\|x\| = 1$, e.g.*

$$J = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, x = \begin{bmatrix} \frac{1}{2}\sqrt{2} \\ \frac{1}{2}\sqrt{2} \end{bmatrix}.$$

Then we see that $\Lambda(J) = \{0\}$, while $x^*Jx = \frac{1}{2}$ and is a part of the set $W(J)$.

So we see in example 4.1 that the numerical range can be much bigger than the spectrum.

We can also look at Jordan block of size n.

Example 4.2 *Let $J \in \mathbb{C}^{n\times n}$. Then we can choose $v \in \mathbb{C}^n$, e.g.*

$$y = \begin{bmatrix} \frac{1}{n}\sqrt{n} \\ \vdots \\ \frac{1}{n}\sqrt{n} \end{bmatrix},$$

such that $\|y\| = 1$ then

$$y^* J y = \frac{n-1}{n}$$

is part of the set $W(J)$.

Till now we only have a theory how to compute the numerical range for matrices. We can also define them for operators, just as we defined pseudospectra for operators.

Definition 4.4 (Numerical range of an operator [Jen, p.12]) *Let $A \in \mathcal{B}(X)$ with X a Hilbert space. The numerical range of an operator A is the set*

$$W(A) = \{\langle u, Au\rangle \mid \|u\| = 1\}.$$

In the operator case we can see something similar to example 4.1. The spectrum of our shift operator on ℓ^2 is the unit circle, but its numerical range is the unit disk [TE05, p. 174]. So we see that the numerical range can be much bigger then the spectrum.

The numerical range can be used to study the behaviour of matrix exponentials, but according to Trefethen and Embree [TE05, p. 166], the numerical range also is determined by the limit when $\varepsilon \to \infty$ of behaviour of the pseudospectrum of $\sigma_\varepsilon(A)$.

In the beginning of Chapter 3 we defined T_n. If we just as in Example 4.2 compute the numerical range, we see that $2n/(2n+1)$ is part of the set $W(T_n)$. So we see that as $n \to \infty,\ 2n/(2n+1) \to 1$.So the set $W(T_n)$ contains for sure a value close to 1 when n goes to infinity, which is also what we would expect. Since we proved that when $n \to \infty$, T_n becomes the shift operator when $n \to \infty$ and the numerical range $W(\mathbb{S})$ of the shift operator also is 1.

4.4 Structured pseudospectra

In Figure 4.2 we perturbed with a complex matrix E. But it is interesting to see that when we perturb with a real matrix, the eigenvalues of our perturbed matrix lie in a different pattern compared to the eigenvalues of our perturbed matrix by complex perturbations. This is visible in Figure 4.3.

When we restrict which ε we can use in our pseudospectra (for instance $\varepsilon \in \mathbb{R}$ instead of $\mathbb{C}$), we have structured pseudospectra. There are interesting applications of structured pseudospectra. We can for instance take the equations

$$\dot{x} = Ax + Bu, y = Cx + Du.$$

After applying the Laplace transformation this leads to the operator-valued function

$$D + C(\lambda I - A)^{-1}B.$$

If A, B, C are matrices, we get the equations [BL]

$$\Lambda(A) \cup \{\lambda \in \rho(A) : \|C(\lambda I - A)^{-1}B\| > 1/\varepsilon\} = \bigcup_{\|E\|<\varepsilon} \Lambda(A + BEC).$$

These structured pseudospectra are also called spectral value sets and are useful for analysis of and control of uncertain linear systems.

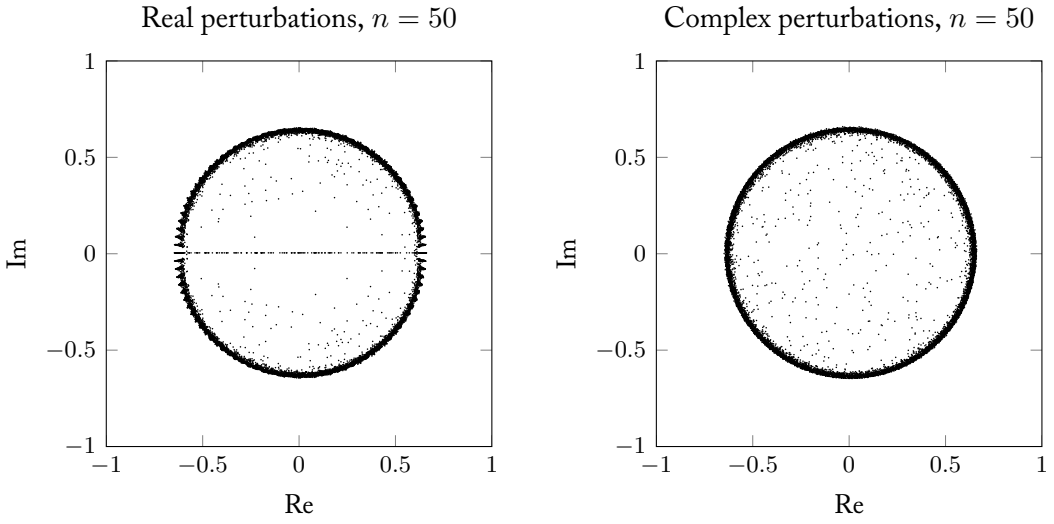

FIGURE 4.3: Structured pseudospectra. Left: real perturbation. Right: complex perturbations. We did 250 random computations to Jordan Block from size 50, $J + E, \|E\| = 10^{-10}, n = 50$.

4.5 Conclusion

In this chapter we showed why pseudospectra are useful. Especially useful is the idea that with random perturbation we can simulate the pseudospectra of a matrix, although this is not as accurate as real algorithms. This helps us to understand what we see when we perturb a matrix with a random perturbation.

Random matrices 5

In Chapter 1 we introduced our problem as the perturbation with a random matrix. In the previous chapters we started with a deterministic perturbation E with norm $\|E\| < \varepsilon$.

Now we return to the original question. We wanted to know the eigenvalues of

$$J + \mathcal{N}_{nk\times nk}(0, \sigma^2).$$

In section 2.5 we introduced regions of eigenvalues when there were too many perturbations to find an explicit solution. Before we answer the question stated above, we assume that all our ε that we introduced are random variables. We look at the probability that the eigenvalues lie within the regions we found. Later we will look at complete random matrices, not just random perturbations at limited places.

We need probability distributions of functions of random variables for the calculations in this chapter. These distributions can be found in Appendix A.

5.1 One eigenvalue

In Example 2.1 we show that when there is one perturbation ε, all eigenvalues lie in a circle with radius $\sqrt[n]{\varepsilon}$. In Appendix A.2 we found that the distribution of $\sqrt[n]{\varepsilon}$, with $\varepsilon \sim \mathcal{N}(0,1)$ is distributed with density function

$$f_Z(x) = \frac{2nx^{n-1}}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}.$$

Theorem 5.1 *$Z = \sqrt[n]{|\alpha|}$ is distributed with*

$$f_Z(x) = \frac{2nx^{n-1}}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}$$

with expected value

$$\mu_Z = \frac{2^{\frac{1}{2n}}\sigma^{\frac{1}{n}}\Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}} \text{ and variance } \sigma_Z^2 = \frac{2^{\frac{1}{n}}\sigma^{\frac{2}{n}}\Gamma\left(\frac{1}{2}+\frac{1}{n}\right)}{\sqrt{\pi}} - \left(\frac{2^{\frac{1}{2n}}\sigma^{\frac{1}{n}}\Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}}\right)^2.$$

As $n \to \infty$, $\mu_Z \to 1$ and $\sigma_Z^2 \to 0$.

Proof *See Appendix A.* ■

We already saw in Figure 1.1 that when our σ^2 became smaller, the eigenvalues of our perturbed matrix are going to a fixed contour. Theorem 5.1 illustrates this clearly. It states that the variance of the eigenvalues of our perturbed matrix are going to zero when $n \to \infty$.

In Figure 5.1 we perturbed with three different ε to see the effect on the spectrum, just as we did in Figure 1.1. We see the same thing happening for the Jordan block with zeros on the diagonal, for a smaller ε the random perturbations are going to the unit circle.

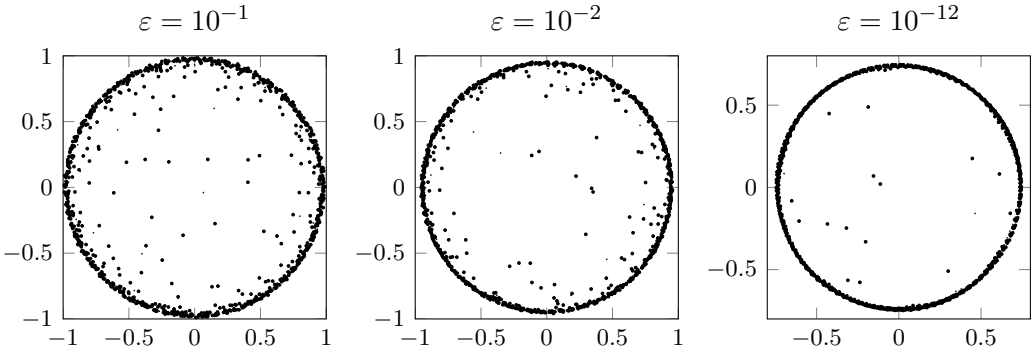

Figure 5.1: Random simulations of $J + \varepsilon E$, $\|E\| = 1$, $n = 100$. Plot are combined plots of 10 experiments.

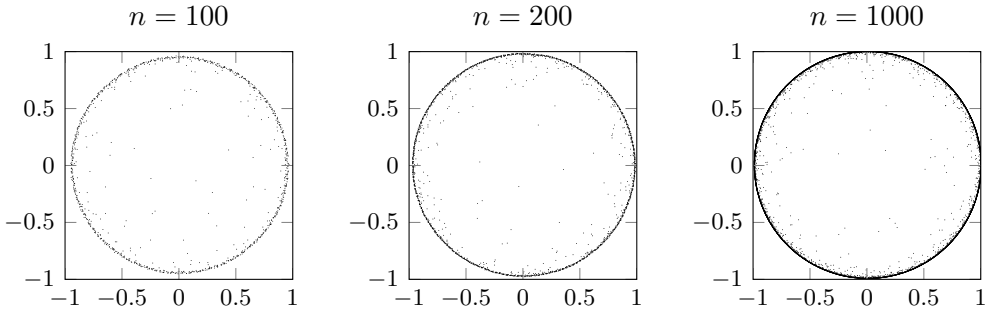

Figure 5.2: Random simulations of $J + \varepsilon E$, $\|E\| = 1$, $\varepsilon = 10^{-2}$. Plot are combined plots of 10 experiments.

5.2 Region of eigenvalues

In this section we assume that all perturbations are normally distributed with mean 0 and variance σ^2:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2), i = \{1, 2, 3\}.$$

For each of the subsections we will use the probability distributions to know the change that the eigenvalues of a randomly perturbed Jordan block lie within the regions found in section 2.5.

Theorem 5.2 *$|\alpha|$ is distributed by*

$$f_Y(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2},$$

and has expected value $\sqrt{\frac{2}{\pi}}\sigma$ and variance $\left(1 - \frac{2}{\pi}\right)\sigma^2$.

Proof *See Appendix A.* ■

In section 2.5.1 we found all our eigenvalues of $J + \varepsilon_1\Delta_{n,1} + \varepsilon_2\Delta_{n-1,1}$ lie within $r = 1 + \varepsilon$, $\varepsilon > 0$ provided

$$|\varepsilon_1| + |\varepsilon_2| < 1.$$

In Theorem A.3 we proved that the sum of two absolute random variables is distributed like one absolute random variable, but with twice the mean and variance. So if $\varepsilon_{1,2} \sim \mathcal{N}(0, \sigma^2)$, then $|\varepsilon_1| + |\varepsilon_2|$ is distributed with density function

$$f_Y(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-x^2/4\sigma^2}.$$

Therefore when $\sigma^2 = 10^{-1}$,

$$P(|\varepsilon_1| + |\varepsilon_2| \geq 1) \approx .11$$

If $\sigma^2 = 10^{-2}$, then this probability is approximately $5 \cdot 10^{-7}$. When $\sigma^2 = 10^{-5}$,

$$P(|\varepsilon_1| + |\varepsilon_2| \geq 1) \approx 0.$$

So we proved that

$$\Lambda\left(\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \varepsilon_2 & & & 0 \\ \varepsilon_1 & & & \end{bmatrix}\right)$$

lies within the unit circle when $\varepsilon_1, \varepsilon_2$ are normally distributed with mean 0 and variance 10^{-5}.

5.2.1 An region dependent on n

In section 2.5.2 we got from these conditions that were dependent on our random variable. We wanted to find a smaller region and found that our eigenvalues lie within the disk

$$r = \sqrt[n]{|\varepsilon_1|} + \beta, \beta > \frac{|\varepsilon_2|}{|\varepsilon_1|n} \tag{5.1}$$

if $|\varepsilon_1| < 1$. By the probability density function of $|\varepsilon_1|$ in (A.2) we see that if $\sigma^2 = 10^{-1}$ that

$$P(|\varepsilon_1| \geq 1) = 1 \cdot 10^{-3}.$$

When $\sigma^2 = 10^{-2}$

$$P(|\varepsilon_1| \geq 1) \approx 0.$$

So we proved that

$$\Lambda\left(\begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ \varepsilon_2 & & & 0 \\ \varepsilon_1 & & & \end{bmatrix}\right)$$

lies within the disk with radius r from (5.1) when $\varepsilon_1, \varepsilon_2$ are normally distributed with mean 0 and variance 10^{-2}.

5.3 Conclusion

We saw that the assumptions we did in section 2.5 are good, since the probability that random perturbations fall inside this region is big for small ε. We also saw an explanation why Figure 1.1 converges when σ becomes small. We did not look at complete random perturbations, that stays a major point of research.

Conclusions and recommendations 6

We started this thesis with the question what the relationship is between the finite perturbation of the Jordan block and the spectrum of the z-transformation. We can conclude that

1. The scalar and the matrix case have a lot in common.
2. The spectra of operators are unaffected by perturbations.
3. There is clear relation between computing pseudospectra and our random perturbations.

6.1 Relation between scalar and matrix case

We found in section 2.1 that one small perturbation of a Jordan block led to a Pusieux series with our eigenvalues distributed around a circle with no eigenvalues in common with the unperturbed eigenvalues. We see that in the matrix case almost the the same happens. In section 2.6 we found that the characteristic polynomial was

$$\chi(\lambda) = \det(\bar{C})^{n-1} \det(\bar{C} + (-1)^n (\bar{C}^{-1} D)^{n-1} E)$$

At the end of section 2.6 we discussed how the characteristic polynomial above looks like. Depending on the structure of E there are a lot of zeros in the long expression of the determinant.

In the finite dimensional case all eigenvalues would be different, but in the infinite dimensional case there can be some common eigenvalues.

6.2 The spectra of operators are unchanged after perturbations

In Chapter 3 we saw that we could perturb an operator, but the spectrum does not change.

6.3 Pseudospectra

We found that the spectrum of a Jordan block is extremely sensitive to perturbations. As a solution pseudospectra were invented to overcome this.

6.4 Random matrices

We calculated that the spectrum of our perturbed matrix lies almost always in the regions we found in Chapter 2. We also noticed that the variance the perturbed matrix is going to zero, what explains why we see the behaviour of Figure 1.1. ø

6.5 Summary

We computed a lot of different spectra, for ordinary matrices and for operators. I created Table 6.1 with an overview of all these different computations to make it visible how they are related.

6.6 Recommendations for further research

While we started this thesis with the question how a randomly perturbed matrix is connected to the spectrum of the z-transform we only looked at random perturbations are certain places. Since random matrix perturbations are a vivid research field, there can be done more research in this field. For instance Bordenave and Capitaine [BC14] discussed random matrix perturbations. It would be interesting to see whether their article could also bring this research a step forward.

Also the research in the last sections of Chapter 3 is not complete. When the exact behaviour of the spectrum of z-transformation can be understood, maybe the complete research question can be answered.

Table 6.1: Eigenvalues (finite) or spectra (infinite) of different matrices/operators.

	Finite	**Infinite**
$\begin{bmatrix} 0 & 1 & \\ & \ddots & 1 \\ & & 0 \end{bmatrix}$	0	$\lvert\lambda\rvert = 1$
$\begin{bmatrix} 1 & 1 & \\ & \ddots & 1 \\ & & 1 \end{bmatrix}$	1	$\lvert\lambda - 1\rvert = 1$
$\begin{bmatrix} 0 & 1 & \\ & \ddots & 1 \\ \varepsilon & & 0 \end{bmatrix}$	$\lvert\lambda\rvert = \sqrt[n-1]{\lvert\varepsilon\rvert}$	$\lvert\lambda\rvert = 1$
$\begin{bmatrix} C & D & \\ & \ddots & D \\ & & C \end{bmatrix}$	$\operatorname{eig}(C)$	$\bigcup_{\theta\in[0,2\pi]} \Lambda\left(\breve{A}(e^{i\theta})\right)$
$\begin{bmatrix} C & D & \\ & \ddots & D \\ E & & C \end{bmatrix}$	$\det\left(\bar{C} + (-1)^n D(\bar{C}^{-1}D)^{n-2}\bar{C}^{-1}E\right) \cdot \det(\bar{C})^{n-1} = 0$	$\bigcup_{\theta\in[0,2\pi]} \Lambda\left(\breve{A}(e^{i\theta})\right)$

$$\bar{C} = C - \lambda I$$

Bibliography

[BC14] Charles Bordenave and Mireille Capitaine, *Outlier eigenvalues for deformed iid random matrices*, arXiv preprint arXiv:1403.6001 (2014).

[BEK90] Rajendra Bhatia, Ludwig Elsner, and Gerd Krause, *Bounds for the variation of the roots of a polynomial and the eigenvalues of a matrix*, Linear Algebra and its Applications **142** (1990), 195–209.

[BL] Albrecht Böttcher and Marko Lindner, *Pseudospectrum*, `http://www.scholarpedia.org/article/Pseudospectrum/`, Accessed: 2015-06-10.

[CB94] Françoise Chatelin and Thierry Braconnier, *About the qualitative computation of Jordan forms*, ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik **74** (1994), no. 2, 105–113.

[CIZ09] Ruth Curtain, Orest V. Iftime, and Hans Zwart, *System theoretic properties of a class of spatially invariant systems*, Automatica **45** (2009), no. 7, 1619–1627.

[CZ12] Ruth F Curtain and Hans Zwart, *An introduction to infinite-dimensional linear systems theory*, vol. 21, Springer Science & Business Media, 2012.

[DH09] E. Brian Davies and Mildred Hager, *Perturbations of Jordan matrices*, Journal of Approximation Theory **156** (2009), no. 1, 82–94.

[ES08] Bennett Eisenberg and Rosemary Sullivan, *Why is the sum of independent normal random variables normal?*, Mathematics Magazine **81** (2008), no. 5, 362–366.

[Fir12] Amir Firooznia, *Controller design for string stability of infinite vehicle strings*, Master's thesis, Eindhoven University of Technology, the Netherlands, 2012.

[GVL13] Gene H. Golub and Charles F. Van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore, 2013.

[Jen] Arne Jensen, *Lectures notes on spectra and pseudospectra of matrices and operators*, `http://people.math.aau.dk/~matarne/11-kaleidoscope2/notes2.pdf`, Accessed: 2014-12-19.

[Lid66] Viktor B. Lidskii, *Perturbation theory of non-conjugate operators*, USSR Computational Mathematics and Mathematical Physics **6** (1966), no. 1, 73–85.

[Lie02] Robert A Liebler, *Basic matrix algebra with algorithms and applications*, CRC Press, 2002.

[NS71] Arch W. Naylor and George R. Sell, *Linear operator theory in engineering and science*, Holt, Rinehart and Winston, Inc., 1971.

[TE05] Lloyd N. Trefethen and Mark Embree, *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton University Press, 2005.

[Tre92] Lloyd N. Trefethen, *Pseudospectra of matrices*, Numerical analysis 1991 (David F. Griffiths and G. Alistair Watson, eds.), Longman Scientific & Technical, Harlow, Essex, 1992, pp. 234–265.

[Tre99] ———, *Computation of pseudospectra*, Acta numerica **8** (1999), 247–295.

[Wil65] James H. Wilkinson, *The algebraic eigenvalue problem*, vol. 87, Clarendon Press Oxford, 1965.

Appendices

Appendix A

Probability distributions

In this thesis we multiple times use distributions that are a function of a Gaussian distributed random variable. In this appendix we work out their distributions, expectation and variance. We assume in all sections that $\alpha \sim \mathcal{N}(0, \sigma^2)$.

A.1 Distribution of $|\alpha|$

Theorem A.1 *$|\alpha|$ is distributed by*

$$f_Y(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2},$$

and has expected value $\sqrt{\frac{2}{\pi}}\sigma$ and variance $\left(1 - \frac{2}{\pi}\right)\sigma^2$.

Proof *Since $Y = P(|\alpha| \leq x) = P(\alpha \leq x) + P(\alpha \geq -x)$. We know if we define the error function* $\operatorname{erf}(y)$ *as*

$$\operatorname{erf}(y) = \frac{2}{\sqrt{\pi}} \int_0^y e^{-u^2} \mathrm{d}u,$$

that (by substituting $u = x(2\sigma^2)^{-1/2}$)

$$f_Y(x) := \begin{cases} \frac{2}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} & x > 0 \\ 0 & x \leq 0 \end{cases}, \tag{A.1}$$

$$F_Y(y) := P(|\alpha| \leq y) = \int_0^y \frac{2}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \,\mathrm{d}x = \operatorname{erf}\left(\frac{y-\mu}{\sqrt{2}\sigma}\right). \tag{A.2}$$

We use this to compute the expectation of $|\alpha|$

$$\begin{aligned} \mathbb{E}(|\alpha|) &= \int_0^\infty \frac{2x}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \,\mathrm{d}x \\ &= \frac{2}{\sigma\sqrt{2\pi}} \int_0^\infty x e^{-x^2/2\sigma^2} \,\mathrm{d}x \\ &= -\sqrt{\frac{2}{\pi}}\sigma \int_0^\infty e^u \,\mathrm{d}u \quad \left(u = -\frac{x^2}{2\sigma^2}\right) \end{aligned}$$

$$= \left[-\sqrt{\frac{2}{\pi}} \sigma e^{u} \right]_{0}^{\infty}$$
$$= \sqrt{\frac{2}{\pi}} \sigma.$$

We use the calculation of

$$\begin{aligned}
\int_0^\infty x^2 e^{-x^2/c^2} \, \mathrm{d}x &= \int_0^\infty \frac{c^2 x}{2} d(e^{-x^2/c^2}) \, \mathrm{d}x \\
&= -\frac{c^2 x}{2} e^{-x^2/c^2} \Big]_0^\infty + \frac{c^2}{2} \int_0^\infty e^{-x^2/c^2} \, \mathrm{d}x \\
&= \frac{c}{2} \int_0^\infty e^{-x^2/c^2} \, \mathrm{d}x \\
&= \frac{c^2 |c|}{2} \int_0^\infty e^{-y^2} \, \mathrm{d}y \quad \left(y = \frac{x}{|c|} \right) \\
&= \frac{c^2 |c|}{4} \sqrt{\pi},
\end{aligned}$$

(where we used that the Gaussian Integral equals $\sqrt{\pi}$) to compute

$$\begin{aligned}
\mathbb{E}(|\alpha|^2) &= \int_0^\infty \frac{2x^2}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \, \mathrm{d}x \\
&= \frac{2}{\sigma\sqrt{2\pi}} \int_0^\infty x^2 e^{-x^2/2\sigma^2} \, \mathrm{d}x \\
&= \frac{2\sigma^3}{\sigma\sqrt{2\pi}} \sqrt{\frac{\pi}{2}} \\
&= \sigma^2.
\end{aligned}$$

Therefore we know that

$$\sigma_{|\alpha|}^2 = \mathbb{E}\left(|\alpha|^2\right) - \mathbb{E}\left(|\alpha|\right)^2 = \left(1 - \frac{2}{\pi}\right) \sigma^2. \tag{A.3}$$

■

A.2 Distribution of $\sqrt[n]{|\alpha|}$

Theorem A.2 *$Z = \sqrt[n]{|\alpha|}$ is distributed with*

$$f_Z(x) = \frac{2nx^{n-1}}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}$$

with expected value

$$\mu_Z = \frac{2^{\frac{1}{2n}} \sigma^{\frac{1}{n}} \Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}} \text{ and variance } \sigma_Z^2 = \frac{2^{\frac{1}{n}} \sigma^{\frac{2}{n}} \Gamma\left(\frac{1}{2} + \frac{1}{n}\right)}{\sqrt{\pi}} - \left(\frac{2^{\frac{1}{2n}} \sigma^{\frac{1}{n}} \Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}} \right)^2.$$

As $n \to \infty$, $\mu_Z \to 1$ and $\sigma_Z^2 \to 0$.

Proof *If we have a function of a random variable $Y = g(X)$, when X is distributed with probability density function f_X, then*

$$f_Z(x) = f_X(g(x)) \times |g'(x)|. \tag{A.4}$$

Since $\sqrt[n]{|\alpha|} = x$, $g(x) = x^n \sim |\alpha|$. Therefore $g(x)$ is distributed as (A.1)*. Thus with f_X as in* (A.1), (A.4) *becomes*

$$f_Z(x) = \frac{2}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2} \left|nx^{n-1}\right| \tag{A.5}$$

$$= \frac{2nx^{n-1}}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}. \tag{A.6}$$

We can use the probability density to compute the expectation.

$$\begin{aligned}
\mathbb{E}(Z) &= \int_0^\infty x f_Z(x)\,\mathrm{d}x \\
&= \int_0^\infty \frac{2nx^n}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}\,\mathrm{d}x \quad \left(u = \frac{x^{2n}}{2\sigma^2}\right) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty \frac{2nx^n}{2nx^{-1}u} \left(2^{1/2n}\sigma^{1/n}u^{1/2n}\right)^{n+1} e^{-u}\,\mathrm{d}u \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty \left(2^{1/2n}\sigma^{1/n}u^{1/2n}\right)^{n+1} u^{-1}e^{-u}\,\mathrm{d}u \\
&= \frac{\sigma^{\frac{n+1}{n}} 2^{\frac{n+1}{2n}}}{\sigma\sqrt{2\pi}} \int_0^\infty u^{\frac{n+1}{2n}-1} e^{-u}\,\mathrm{d}u \\
&= \frac{2^{\frac{1}{2n}}\sigma^{\frac{1}{n}}\Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}}.
\end{aligned}$$

Furthermore in a similar way we can calculate

$$\begin{aligned}
\mathbb{E}\left(Z^2\right) &= \int_0^\infty x^2 f_Z(x)\,\mathrm{d}x \\
&= \int_0^\infty \frac{2nx^{n+1}}{\sigma\sqrt{2\pi}} e^{-x^{2n}/2\sigma^2}\,\mathrm{d}x \quad \left(u = \frac{x^{2n}}{2\sigma^2}\right) \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty \left(2^{1/2n}\sigma^{1/n}u^{1/2n}\right)^{n+2} u^{-1}e^{-u}\,\mathrm{d}u \\
&= \frac{\sigma^{\frac{n+2}{n}} 2^{\frac{n+2}{2n}}}{\sigma\sqrt{2\pi}} \int_0^\infty u^{\frac{n+2}{2n}-1} e^{-u}\,\mathrm{d}u \\
&= \frac{2^{\frac{1}{n}}\sigma^{\frac{2}{n}}\Gamma\left(\frac{1}{2}+\frac{1}{n}\right)}{\sqrt{\pi}}.
\end{aligned}$$

Therefore we know that

$$\sigma_Z^2 = \mathbb{E}\left(\left(\sqrt[n]{|\alpha|}\right)^2\right) - \mathbb{E}\left(\sqrt[n]{|\alpha|}\right)^2 \tag{A.7}$$

$$= \frac{2^{\frac{1}{n}}\sigma^{\frac{2}{n}}\Gamma\left(\frac{1}{2}+\frac{1}{n}\right)}{\sqrt{\pi}} - \left(\frac{2^{\frac{1}{2n}}\sigma^{\frac{1}{n}}\Gamma\left(\frac{n+1}{2n}\right)}{\sqrt{\pi}}\right)^2. \tag{A.8}$$

To compute the asymptotic behavior of the expectation and the variance, it follows from

$$\lim_{n\to\infty}\Gamma\left(\frac{n+1}{2n}\right)=\lim_{n\to\infty}\Gamma\left(\frac{n+2}{2n}\right)=\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}\ \textit{and}\ \lim_{n\to\infty}\sqrt[n]{z}=1,\ z>0.$$

that

$$\mu_{asymp}=\lim_{n\to\infty}\mathbb{E}\left(Z\right)=\frac{\sqrt{\pi}}{\sqrt{\pi}}=1$$
$$\sigma^2_{asymp}=\lim_{n\to\infty}\mathbb{E}\left(Z^2\right)-\left(\lim_{n\to\infty}\mathbb{E}\left(Z\right)\right)^2=1-1^2=0.$$ ■

In Figure A.1 the expected value and variance is plotted against n.

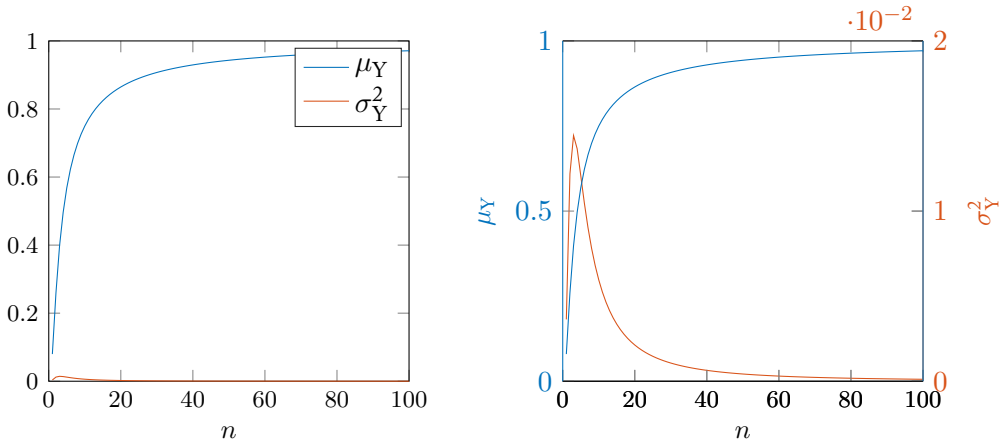

Figure A.1: Expected value and variance of $\sqrt[n-1]{|\alpha|}$, $\alpha\sim\mathcal{N}(0,\sigma^2)$, $\sigma^2=0.01$

A.3 Sum of two random variables

Theorem A.3 *If X, Y are distributed with distribution function* (A.1) *with mean μ_X, μ_Y and variance σ^2_X, σ^2_Y, then $X+Y$ is distributed with* (A.1) *with mean $\mu_X+\mu_Y$ and variance $\sigma^2_X+\sigma^2_Y$.*

Proof *From [ES08]. Let Z_1 and Z_2 be two independent standard normal random variables with joint density function*

$$f(z_1,z_2)=\frac{\exp\left(-\frac{1}{2}\left(z_1^2+z_2^2\right)\right)}{\pi}.$$

This is function is is rotation invariant, which means that this function has the same value for all points with the same distance from the origin. Thus, $f(T(z_1,z_2))=f(z_1,z_2)$ where T is any rotation around the origin.

It follows that for any set A in the plane $P((Z_1,Z_2)\in A)=P((Z_1,Z_2)\in TA)$ where T is a rotation of the plane. Now if X_1 is normal with mean 0 and variance σ_1^2 and X_2 is normal

with mean 0 and variance σ_2^2, then $|X_1| + |X_2|$ has the same distribution as $|\sigma_1 Z_1| + |\sigma_2 Z_2|$. Hence

$$P(X_1 + X_2 \leq t) = P(\sigma_1 Z_1 + \sigma_2 Z_2 \leq t) = P((Z_1, Z_2) \in A)$$

where A is the half plane $\{(z_1, z_2) = \sigma_1 z_1 + \sigma_2 z_2 \leq t\}$. The boundary line $\sigma_1 z_1 + \sigma_2 z_2 = t$ lies at a distance $d = |t|/\sqrt{\sigma_1^2 + \sigma_2^2}$ from the origin. It follows that the set A can be rotated into the set

$$TA = \left\{ (z_1, z_2) \mid z_1 \leq t/\sqrt{\sigma_1^2 + \sigma_2^2} \right\}.$$

Thus $P(X_1 + X_2 < t) = P\left(\sqrt{\sigma_1^2 + \sigma_2^2} Z_1 < t\right)$. It follows that $X_1 + X_2$ is normal with mean 0 and variance $\sigma_1^2 + \sigma_2^2$. This completes the proof. ■

Note The proof is only above the normal case, but I like it a lot more than the convolution proof. I still need to adapt it to the absolute case.

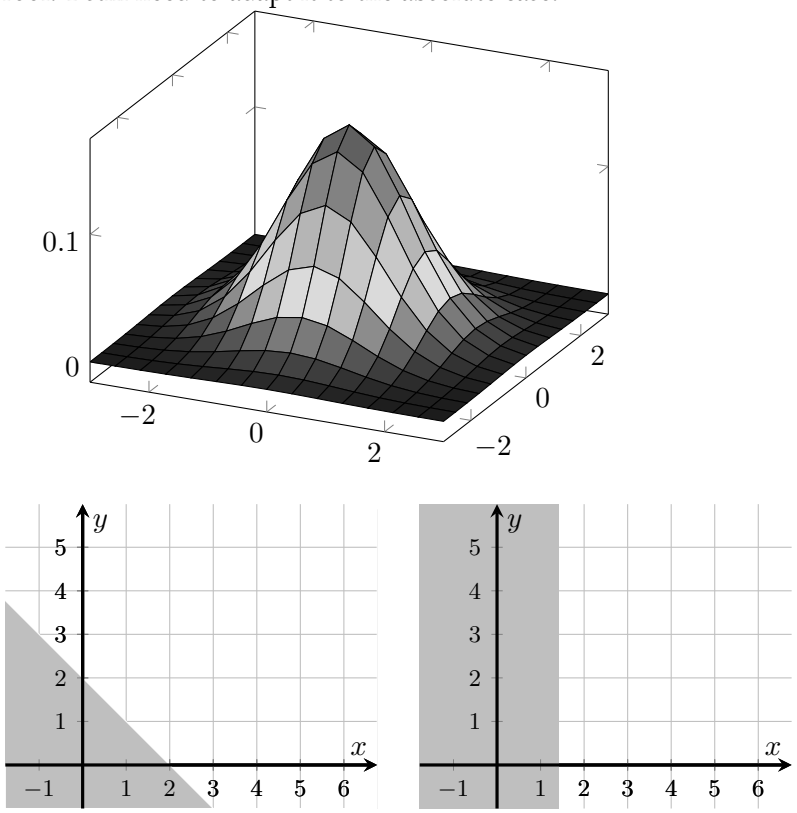

FIGURE A.2: Rotation in set z_1, z_2

A.4 Sum of two absolute variables

We know that

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \right).$$

The sum of two independent normally distributed variables $x \sim \mathcal{N}(\mu_x, \sigma_x^2)$, $y \sim \mathcal{N}(\mu_y, \sigma_y^2)$ is again normally distributed with mean $\mu_x + \mu_y$ and variance $\sigma_x^2 + \sigma_y^2$. We know from section A.1 that

$$\sigma_{|\alpha|}^2 = \mathbb{E}\left(|\alpha|^2\right) - \mathbb{E}(|\alpha|)^2 = \sigma^2 - \frac{2}{\pi}\sigma^2. \tag{A.9}$$

Since $|\alpha_1|$ is distributed with cumulative distribution function $F_X(y)$ with mean $\sigma(2/\pi)^{1/2}$ and variance $(1 - 2/\pi)\sigma^2$, $|\alpha_1| + |\alpha_2|$ is distributed with cumulative distribution function $F_X(y)$ with mean $2\sigma(2/\pi)^{1/2}$ and variance $(2 - \pi)\sigma^2$ (Theorem A.3). It then follows that

$$k(y, \sigma) = P(|\alpha_1| + |\alpha_2| \leq y) = \text{erf}\left(\frac{y - 2\sigma(2/\pi)^{1/2}}{\sqrt{2(2-\pi)}\sigma}\right).$$

Appendix B

System theory in the z-domain

B.1 Norms and spaces

Corresponding to the definition in [Fir12, C.2] we define the $\mathcal{L}_2$-space for $\check{x}\left(z = e^{i\theta}\right)$ as as mapping from $\mathbb{D}_1$ to $\mathbb{C}^n$ for which (B.1) is finite. In Definition B.1 this is made formal.

Definition B.1 ($\mathcal{L}_2$-space)

$$\mathcal{L}_2(\mathbb{D}_1; \mathbb{C}^n) = \{\check{x} : \mathbb{D}_1 \to \mathbb{C}^n \mid \|\check{x}\|_{\mathcal{L}_2} < \infty\},$$

where $\|\cdot\|_{\mathcal{L}_2}$ is called the $\mathcal{L}_2$-norm and is defined as

$$\|\check{x}\|_{\mathcal{L}_2} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} \|\check{x}(e^{i\theta})\|^2_{\mathbb{C}^n} \mathrm{d}\theta}. \tag{B.1}$$

Here $\mathbb{D}_1 = \{z \in \mathbb{C} \mid |z| = 1\}$ is the unit circle.

B.2 Bilateral z-transformation

The bilateral z-transformation of an element $x = (x_k)_{k=-\infty}^{\infty} \in \ell^2(\mathbb{C}^n)$ is defined as

$$\mathcal{Z}(x) = \check{x}(z) = \sum_{k=-\infty}^{\infty} x_k z^{-k} = \sum_{k=-\infty}^{\infty} x_k e^{-ki\theta},$$

where z is a complex variable that acts as a shift operator and can be written as $z = e^{i\theta}$ with $\theta \in [0, 2\pi]$. This is precisely the Fourier series representation of an element $\check{x}(z) \in \mathcal{L}_2(\delta\mathbb{D}; \mathbb{C}^n)$ with the Fourier coefficients

$$x_k = \mathcal{Z}^{-1}(\check{x}(z)) = \frac{1}{2\pi i}\oint \check{x}(z) z^{k-1} \mathrm{d}z$$

with the unit circle as the path of integration. Thus, by substituting $z = e^{i\theta}$

$$x_k = \mathcal{Z}^{-1}(\check{x}(z)) = \frac{1}{2\pi i}\int_0^{2\pi} \check{x}(e^{i\theta}) e^{i\theta(k-1)} (ie^{i\theta}\mathrm{d}\theta) = \frac{1}{2\pi}\int_0^{2\pi} \check{x}(e^{i\theta}) e^{i\theta k}\mathrm{d}\theta$$

Note that an element $\check{x}(z) \in \mathcal{L}_2(\delta\mathbb{D}; \mathbb{C}^n)$ has an inverse bilateral z-transform $x = (x_k)_{k=-\infty}^{\infty} \in \ell^2(\mathbb{C}^n)$, where x_k's are the Fourier coefficients of $\check{x}(z)$.

Appendix C

Proofs

Theorem C.1 (Theorem 3.1) *Suppose $f \in \ell^1(\mathbb{Z}), g \in \ell^p(\mathbb{Z})$. Then*

$$\sum_{k=-\infty}^{\infty} |f_{n-k} g_k|^p < \infty \tag{C.1}$$

for all $n \in \mathbb{Z}$. Define the convolution of f and g (notation $h = f * g$) as

$$h_n = \sum_{k=-\infty}^{\infty} f_{n-k} g_k, \; n \in \mathbb{Z}. \tag{C.2}$$

Then $h \in \ell^p(\mathbb{Z})$ and

$$||h||_p \leq ||f||_1 ||g||_p. \tag{C.3}$$

Proof *If $\tilde{p} = p(p-1)^{-1}$ thus $p^{-1} + \tilde{p}^{-1} = p^{-1} + (p-1)p^{-1} = 1$, we see that*

$$\begin{aligned}
|h_n|^p &= \left| \sum_{k=-\infty}^{\infty} f_{n-k} g_k \right|^p \leq \left(\sum_{k=-\infty}^{\infty} |f_{n-k} g_k| \right)^p = \left(\sum_{k=-\infty}^{\infty} \left| f_{n-k}^{1/\tilde{p}} f_{n-k}^{1/p} g_k \right| \right)^p \\
&\leq \left(\left(\sum_{k=-\infty}^{\infty} |f_{n-k}|^{\frac{\tilde{p}}{\tilde{p}}} \right)^{\frac{1}{\tilde{p}}} \left(\sum_{k=-\infty}^{\infty} \left| f_{n-k}^{1/p} g_k \right|^p \right)^{\frac{1}{p}} \right)^p \quad (\textit{by Hölder: } \frac{1}{p} + \frac{1}{\tilde{p}} = 1) \\
&= \left(\sum_{k=-\infty}^{\infty} |f_{n-k}| \right)^{\frac{p}{\tilde{p}}} \left(\sum_{k=-\infty}^{\infty} \left| f_{n-k}^{1/p} \right|^p |g_k|^p \right) \\
&= \left(\sum_{k=-\infty}^{\infty} |f_{n-k}| \right)^{p-1} \sum_{k=-\infty}^{\infty} |f_{n-k}| \, |g_k|^p \, .
\end{aligned}$$

Therefore

$$
\begin{aligned}
||h||_p^p = \sum_{n=-\infty}^{\infty} |h_n|^p &\leq \sum_{n=-\infty}^{\infty} \left(\left(\sum_{k=-\infty}^{\infty} |f_{n-k}| \right)^{p-1} \sum_{k=-\infty}^{\infty} |f_{n-k}| \, |g_k|^p \right) \\
&= ||f||_1^{p-1} \sum_{k=-\infty}^{\infty} \left(\sum_{n=-\infty}^{\infty} |f_{n-k}| \, |g_k|^p \right) && \textit{(Tonelli-Fubini)} \\
&= ||f||_1^{p-1} \sum_{k=-\infty}^{\infty} \left(\sum_{m=-\infty}^{\infty} |f_m| \, |g_k|^p \right) && (n-k=m) \\
&= ||f||_1^{p-1} \sum_{m=-\infty}^{\infty} |f_m| \sum_{k=-\infty}^{\infty} |g_k|^p \\
&= ||f||_1^{p-1} \, ||f||_1 \, ||g||_p^p = ||f||_1^p \, ||g||_p^p.
\end{aligned}
$$

Thus $||h||_p \leq ||f||_1 ||g||_p$. ■