The University of Reading

Computing the Spectra and Pseudospectra of Non-Self-Adjoint Random Operators Arising in Mathematical Physics

BY

Ratchanikorn Chonchaiya

Thesis Submitted for the degree of
Doctor of Philosophy
Department of Mathematics and Statistics

NOVEMBER 2010

Abstract

In this thesis, we derive three different classes of spectral inclusion sets, meaning sets that enclose the spectrum or pseudospectra, of an infinite tridiagonal (self-adjoint or non-self-adjoint) matrix A understood as a bounded linear operator on $\ell^2(\mathbb{Z})$. The first inclusion set is the union of certain pseudospectra of $n \times n$ principal submatrices of A (we call that "method 1"). The second version is a very similar construction but with slightly modified circulant-type $n \times n$ submatrices (we call that "method 1*"). In the third version, we work with lower bounds on $n \times \infty$ and $\infty \times n$ submatrices of $A - \lambda I$, which effectively leads to the study of related $n \times n$ matrices (we call that "method 2"). Our third set not only yields an upper bound but also sequences of approximations of the spectra and pseudospectrum of A that are convergent as $n \to \infty$.

In chapter 5 we study the particular tridiagonal operator $A^b v_i = b_i v_{i-1} + v_{i+1}$ where (b_i) is a bounded sequence with $b_i = 1$ or $b_i = -1$ randomly. Our motivation is that this non-self-adjoint operator, and the corresponding non-self-adjoint finite random matrices, have been studied extensively in the mathematical physics literature, starting with work by Feinberg and Zee [16] who studied a model of a particle hopping asymmetrically on a one dimensional lattice. We show that the spectrum of A^b is symmetric about the axes and under $90^\circ-$rotation, the closed unit disk is contained in the spectrum, the numerical range of A^b is the square with $2, -2, 2i$ and $-2i$ as its corners. Further, we apply method 1 to A^b and show that for this operator, the sequences of approximations that it generates to the spectra and pseudospectra are convergent as $n \to \infty$. We finish with some conjecture about the spectrum of A^b.

Acknowledgement

Firstly, I would like to thank my supervisors Prof. Simon Chandler-Wilde and Dr. Marko Lindner for their help, guidance, support and patience, without which I could not have completed this research. Their supervisions have been inspirational. I am also grateful to Dr. Beatrice Pelloni, Dr. Titus Hilberdink and Prof. Les Bunce for their suggestions and comments on my work.

To all my friends in the department of Mathematics, thank you for sharing the good memories and the enjoyable times whilst I was carrying out my research at Reading, especially Keith Pham who has been my best friend for the whole of my PhD life. Thanks a lot for your spaghetti, kebabs and pizzas which are really fantastic.

To all my friends outside the world of maths, thank you for your encouragement, support and your kindness in teaching me so many new things and activities. I have learnt so many useful Chinese words and Chinese recipes from all the food parties, coffee breaks, card and dice games. This has made my time spent in the university thoroughly enjoyable. I should also like to thank Steven Liang for tolerating a maths guy like me since the very start of my studying. These friendships I hope will last.

I also wish to express my appreciation for the financial support received from the Higher Education Strategic Scholarship for Frontier Research granted by the Thai Commission on Higher Education.

I would like to thank my mother and father, my brother Vivach and my sisters Sa and Mai for their unconditional love and family support during my time at the university.

Finally, last but not least, I would like to thank Ros Young for her encouragement and support. I have been introduced to many new words to improve my English and my confidence using spoken English has improved a great deal since I came here from our many conversations.

I thank you all from the bottom of my heart.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Ratchanikorn Chonchaiya

Contents

Chapter 1

Introduction

1.1 A Partial History of Spectral Theory

This thesis is concerned with the study of the spectral theory of infinite tridiagonal matrices. Eigenvalues have been one of the most powerful tools in applied mathematics. They are used in many scientific research fields, for example, fluid mechanics, quantum mechanics, economics, functional analysis and acoustics.

The eigenvalue problem was first introduced by Augustin Louis Cauchy in the study of quadratic forms and also in the study of extreme values by Lagrange multiplier methods. The terms “eigenvalue” and “eigenvector” are sometimes called “characteristic value” and “characteristic vector”, respectively. Those words were coined by Cauchy (see [27]). In 1846, Carl Gustav Jacob Jacobi proposed a numerical iterative method for the calculation of the eigenvalues and eigenvectors of a real Hermitian matrix. Jean d’ Alembert, a French mathematician, was the pioneer who studied differential equations using eigenvalues.

At the beginning of the twentieth century, the eigenvalues of integral operators were studied by David Hilbert. He considered the operator as infinite dimensional matrix. In 1904, the German word “eigen” was used by

Hilbert to denote eigenvalues and eigenvectors.

The origin of the spectral theory of matrices is the concept of an eigenvalue. Much research work on the eigenvalue problem has been developed. In 1829, Cauchy published his result, which is a combination between previously developed results and his own ideas. He showed that real symmetric matrices have real eigenvalues and that the corresponding quadratic form can be diagonalized using a linear transformation. In the 1870's that paper provided important results which gave motivation to develop many theorems on the solid spectral theory of matrices.

The important solution of eigenvalue problems for second-order differential equations was developed by Charl-François Sturm, a Swiss mathematician, in 1836 and Joseph Liouville, a French mathematician, in 1838. The so-called Sturm-Liouville theory was an important milestone in the spectral theory of ordinary differential operators. In 1848, the paper of Cauchy had influenced Jacobi to determine that the eigenvalues of the quadratic forms given by

$$\sum_{k=1}^{n} a_k x_k^2 - 2\sum_{k=1}^{n-1} b_{k+1} x_k x_{k+1}$$

are the roots of the denominator of a limited continued fraction. Nowadays, we call infinite self-adjoint tridiagonal matrices, Jacobi operators. For some more details of the history of spectral theory see [30].

The research in this thesis is a contribution, in particular to the part of spectral theory which deals with the computation of the spectra of infinite (particularly non-self-adjoint) tridiagonal matrices. This part of spectral theory has attracted significant interest recently.

Beginning in 1996, motivated by the studies of statistical mechanics of the magnetic flux lines in superconductors with columnar defect [22, 24], Hatano and Nelson initiated a study of a non-self-adjoint Anderson model which has become to be known as the Hatano-Nelson model. They were studying the eigenvalues of an operator H defined by

$$(Hf)_m = e^{-g} f_{m-1} + v_m f_m + e^g f_{m+1}, \qquad m \in \mathbb{Z},$$

where g is a fixed real parameter (without loss of generality we may assume that g is positive), and $(v_m)_{m\in\mathbb{Z}}$ is a sequence of i.i.d. random variables taking values in some compact subset of the real line, under the constraint that the eigenfunction f is periodic, i.e.

$$f_{m+n} = f_m, \qquad m \in \mathbb{Z},$$

for some $n \in \mathbb{N}$, in which case the eigenvalues of H are the eigenvalues of the finite random non-normal matrix

$$A_n = \begin{pmatrix} v_1 & e^g & & & e^{-g} \\ e^{-g} & v_2 & e^g & & \\ & \ddots & \ddots & \ddots & \\ & & e^{-g} & v_{n-1} & e^g \\ e^g & & & e^{-g} & v_n \end{pmatrix}_{n\times n}.$$

In the paper [23] they also considered the multidimensional version of H. The case when the corner entries in A_n are replaced by zeros is not interesting because it is similar (via a diagonal transform) to a self-adjoint matrix and therefore it has purely real spectrum.

In 1998, Goldsheid and Khoruzhenko [19] developed a theory to study the distribution of eigenvalues in the non-self-adjoint Anderson model. In this paper they found that the eigenvalues of A_n are arranged along a curve in the complex plane and they also derived an equation for that curve. The curve is sometimes called a 'bubble with wings' (see [22, 47, 48]).

In 2001, Davies [11] found that for randomly generated large matrices, which are far from self-adjoint, their spectra depend sensitively on the entries of the matrix, which is the difficulty in studying the behaviour of non-self-adjoint operators. For the non-self-adjoint Anderson model of Hatano-Nelson, the behaviour of the spectrum of a finite $n \times n$ matrix when $n \to \infty$ does not explain the behaviour of the spectrum of the infinite matrix. The reason is there are many approximate eigenvalues which are not close to true eigenvalues of the infinite dimensional matrix. In [12] Davies studied the spectrum of H as an operator on $\ell^2(\mathbb{Z})$ in the case when $(v_m)_{m\in\mathbb{Z}}$ is pseudo-ergodic which holds almost surely if $(v_m)_{m\in\mathbb{Z}}$ is random. He then found the condition which makes $0 \in \operatorname{spec} H$. Martinez [37, 38] obtained a sharp bound on the spectra of many operators and also estimated the size of the hole in the spectrum of the non-self-adjoint Anderson operator.

Trefethen, Contedini and Embree [48] studied the spectra and pseudospectra of random bidiagonal matrices of the form

$$A_n = \begin{pmatrix} x_1 & 1 & & & \\ & x_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & x_{n-1} & 1 \\ & & & & x_{n-1} \end{pmatrix}_{n\times n},$$

where each x_i is a random variable taking values independently in a compact subset of $\mathbb{C}$, from some distribution X. They studied the spectral properties of the "one-way-model" by Brezin, Feinberg and Zee [5, 16, 17]. When the entries on the main diagonal generate from $\{\pm 1\}$, then they obtained the following matrix

$$A_n = \begin{pmatrix} \pm 1 & 1 & & & \\ & \pm 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \pm 1 & 1 \\ & & & & \pm 1 \end{pmatrix}_{n\times n}.$$

The spectral behaviour of the bidiagonal case, A_n, is basically the same as the non-periodic Hatano-Nelson matrix. However, we can see that it is much easier to study the bidiagonal case because its resolvent can be computed immediately. Trefethen et al. also studied the spectrum of the corresponding infinite matrix,

$$A = \begin{pmatrix} \ddots & \ddots & & & & \\ & \pm 1 & 1 & & & \\ & & \pm 1 & 1 & & \\ & & & \pm 1 & 1 & \\ & & & & \pm 1 & \ddots \\ & & & & & \ddots \end{pmatrix}.$$

The main result for a random bidiagonal doubly infinite matrix case is, with probability 1, that the spectrum of the matrix A is the union of the two closed unit disks centred at 1 and -1, respectively.

In 2008, Lindner [35] generalizes the above result from the case of one random and one constant diagonal to the case of two random diagonals, so that

$$A = \begin{pmatrix} \ddots & \ddots & & & & \\ & \sigma_{-1} & \tau_{-1} & & & \\ & & \sigma_0 & \tau_0 & & \\ & & & \sigma_1 & \tau_1 & \\ & & & & \sigma_2 & \ddots \\ & & & & & \ddots \end{pmatrix},$$

where $\sigma_k \in \Sigma$ and $\tau_k \in \mathcal{T}$ are taken independently from random distributions on Σ and $\mathcal{T}$, which are arbitrary compact subsets of $\mathbb{C}$, respectively, under the condition for every $\varepsilon > 0, \sigma \in \Sigma$ and $\tau \in \mathcal{T}$, that $Pr\left(|\sigma_k - \sigma| < \varepsilon\right)$ and $Pr\left(|\tau_k - \tau| < \varepsilon\right)$ are both non-zero. This is a proper generalization of [48] because the set $\mathcal{T}$ may contain zero. In order to say something about his result, we need the following definition. For $\varepsilon > 0$, let

$$\Sigma_{\cup}^{\varepsilon} := \bigcup_{\sigma \in \Sigma} \overline{U}_{\varepsilon}(\sigma) \text{ and } \Sigma_{\cap}^{\varepsilon} := \bigcap_{\sigma \in \Sigma} U_{\varepsilon}(\sigma)$$

with $U_{\varepsilon}(\sigma) = \{\lambda \in \mathbb{C} : |\lambda - \sigma| < \varepsilon\}$ and $\overline{U}_{\varepsilon}(\sigma) = \{\lambda \in \mathbb{C} : |\lambda - \sigma| \leq \varepsilon\}$. Then his main result is:

Theorem 1.1. *If A is the above bi-infinite bi-diagonal random matrix then, with probability 1,*

$$\operatorname{spec} A = \operatorname{spec}_{\mathrm{ess}} A = \Sigma_{\cup}^{T} \setminus \Sigma_{\cap}^{t},$$

where $T = \max\{|\tau| : \tau \in \mathcal{T}\}$ *and* $t = \min\{|\tau| : \tau \in \mathcal{T}\}$.

In 1999, Joshua Feinberg and Anthony Zee started studying a model describing the propagation of a particle hopping on a 1-dimensional lattice. Feinberg and Zee [16] studied the equation

$$v_{k+1} + b_{k-1} v_{k-1} = \lambda v_k \tag{1.1}$$

where the real numbers b_k are generated from some random distribution and λ is the spectral parameter.

Zee and Feinberg studied the distribution of the eigenvalues of the $n \times n$ matrix A_n^b defined by

$$A_n^b = \begin{pmatrix} 0 & 1 & & & & \\ b_1 & 0 & 1 & & & \\ & b_2 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & b_{n-2} & 0 & 1 \\ & & & & b_{n-1} & 0 \end{pmatrix}, \tag{1.2}$$

which is obtained from (1.1) when $v_0 = v_{n+1} = 0$ and each b_k is ± 1, randomly. They noticed that when n was large, the spectrum has a complicated fractal-like form.

In 2002, Holz, Orland and Zee [28] studied the spectrum of the infinite matrix

$$A^b = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & 1 & & & & \\ & b_{-1} & 0 & 1 & & & \\ & & b_0 & 0 & 1 & & \\ & & & b_1 & 0 & 1 & \\ & & & & b_2 & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}, \qquad (1.3)$$

where $b_i = \pm 1$, for all 6 possible cases with a 4-periodic sequence, i.e. with $b_{i+4} = b_i, i \in \mathbb{Z}$. They found that the spectrum for each of these 6 patterns corresponds to a certain curve. They also stated some open questions on the spectrum of the infinite random matrix e.g. does the spectrum contain a hole in the complex plane or not? Is the spectrum of the operator localized or delocalized?

In 2010, Chien and Nakazato [10] studied the numerical range of tridiagonal operators A defined by $Ae_j = e_{j-1} + r^j e_{j+1}$, where $r \in \mathbb{R}, j \in \mathbb{N}$ and $\{e_1, e_2, \ldots\}$ is the standard orthonormal basis for $\ell^2(\mathbb{N})$. In the third section of this paper, they emphasised on the case $r = -1$ and they showed that

$$\begin{aligned} W(A) = \{z \in \mathbb{C} : -1 \le \mathrm{Re}(z) \le 1, -1 \le \mathrm{Im}(z) \le 1\} \\ \setminus \{1+i, 1-i, -1+i, -1-i\}. \end{aligned}$$

A large part of this thesis is concerned with deriving new inclusion sets for the spectra of tridiagonal bounded linear operators, so that it is appropriate to briefly review other methods for computing inclusion sets. Firstly, we let A be any bounded linear operator.

1. A **Trivial upper bound** on the spectrum of A is $\{\lambda : |\lambda| \le \|A\|\}$ since $|\lambda| \le \|A\|$ if $\lambda \in \operatorname{spec} A$. The bound improves as one takes powers: From $\lambda \in \operatorname{spec} A$ we get $\lambda^n \in \operatorname{spec}(A^n)$ for $n \in \mathbb{N}$ and hence $|\lambda| = |\lambda^n|^{1/n} \le \|A^n\|^{1/n}$, leading to the sharper bound $\operatorname{spec} A \subseteq \left\{\lambda : |\lambda| \le \|A^n\|^{1/n}\right\}$.

2. **Gershgorin Circle Theorem** is a method to compute the upper bound of the spectrum of an operator. It is introduced, in the matrix case, by Semyon Aranovich Gershgorin in 1931 [18].

   Let $A = (a_{ij})$ be an $n \times n$ matrix and $d_i = \sum_{j \neq i} |a_{ij}|$ for $i = 1, \dots, n$. Then the set

   $$D_i = \{\lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq d_i\} \qquad i = 1, \dots, n$$

   is called the ith Gershgorin disc of the matrix A. Gershgorin's circle Theorem (see Theorem 2.49 below) says that every eigenvalue of A lies within at least one of the Gershgorin discs. In Theorem 2.50, we generalise the Gershgorin circle theorem (see [18, 50]) to infinite matrices.

3. The **Numerical range** of A is defined by

   $$W(A) := \{(Ax, x) : \|x\| = 1\}$$

   with $(\cdot, \cdot)$ denoting the inner product. The closure of $W(A)$ is an upper bound on $\operatorname{spec} A$. Also this bound improves as one takes powers (see 5. below). Hausdorff [26] and Toeplitz [46] proved that $W(A)$ is convex. Moreover, $W(A)$ is invariant under unitary transformation, i.e., $W(U^*AU) = W(A)$.

4. The **Pseudospectrum** has been introduced independently at least 6 times: by J. M. Varah (1967), H. Landau (1975), S. K. Godunov (1982), L. N. Trefethen (1990), D. Hinrichsen and A. J. Pritchard (1992), and E. B. Davies (1997). It is a tool to understand the behaviour of the spectrum of non-normal operators.

   Definition 1.2. *Let A be a bounded linear operator on a Banach space X. For $\varepsilon > 0$, the ε-**pseudospectrum** is*

   $$\operatorname{spec}_\varepsilon A = \{\lambda \in \mathbb{C} : \|(\lambda I - A)^{-1}\| > \frac{1}{\varepsilon}\},$$

   with the convention that $\|(\lambda I - A)^{-1}\| := \infty$ if $\lambda \in \operatorname{spec} A$.

Sometimes, the pseudospectrum is defined as $\{\lambda \in \mathbb{C} : \|(\lambda I - A)^{-1}\| \geq \frac{1}{\varepsilon}\}$, Chaitin-Chaitelin and Harrabi [7] proved that

$$\overline{\mathrm{spec}_\varepsilon A} = \left\{\lambda \in \mathbb{C} : \left\|(A - \lambda I)^{-1}\right\| \geq \varepsilon^{-1}\right\}$$

is true when A is an operator on Hilbert space and it is not true in general when A is an operator on any Banach space. Moreover, in [43], Shargorodsky has given two examples to show that the pseudospectra of closed densely defined operators on a Hilbert space can jump with respect to ε.

5. **Higher order numerical range** is introduced as a better tool than the numerical range since $W(A)$ is always convex and so does not give much information about the spectrum even though it is easier to compute $W(A)$. Martinez [38] used this technique to determine the spectrum of non-self-adjoint operators completely in many cases.

   Following [14, 15], if p is a polynomial then $\mathrm{Hull}_p(A) := \{z \in \mathbb{C} : |p(z)| \leq \|p(A)\|\} \supseteq \mathrm{spec}\, A$. The intersection over all polynomials p (of order at most $n \in \mathbb{N}$) of $\mathrm{Hull}_p(A)$ is the *higher order hull of* A, denoted by $\mathrm{Hull}_\infty(A)$ (resp. $\mathrm{Hull}_n(A)$). For each polynomial p, put $\mathrm{Num}_p(A) := \{z \in \mathbb{C} : p(z) \in \overline{W(p(A))}\} \supseteq \mathrm{spec}\, A$. The intersection over all polynomials p (of order at most $n \in \mathbb{N}$) of $\mathrm{Num}_p(A)$ is the *higher order numerical range of* A, denoted $\mathrm{Num}_\infty(A)$ (resp. $\mathrm{Num}_n(A)$). Higher order hulls and numerical ranges are related: Clearly $\mathrm{Num}_n(A) \subseteq \mathrm{Hull}_n(A)$ for all $n \in \mathbb{N}$. In fact, equality holds for all n! The set $\mathrm{Num}_\infty(A) = \mathrm{Hull}_\infty(A) \supseteq \mathrm{spec}\, A$ coincides with the so-called polynomial convex hull of $\mathrm{spec}\, A$, which is the complement of the unbounded component of $\mathbb{C} \setminus \mathrm{spec}\, A$ (i.e. it is $\mathrm{spec}\, A$ plus everything enclosed by it).

1.2 Overview of Thesis

In this thesis we study the spectrum of infinite tridiagonal matrices with a bounded set of entries understood as bounded linear operators on $\ell^2(\mathbb{Z})$.

One particular example of this type to be studied in Chapter 5 is a class of matrices of the form (1.3). We start in Chapter 3 where we introduce two methods to approximate the spectrum for arbitrary infinite tridiagonal matrices of the form

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & \beta_{-2} & \gamma_{-1} & & & & \\ & \alpha_{-2} & \beta_{-1} & \gamma_0 & & & \\ & & \alpha_{-1} & \boxed{\beta_0} & \gamma_1 & & \\ & & & \alpha_0 & \beta_1 & \gamma_1 & \\ & & & & \alpha_1 & \beta_2 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}, \tag{1.4}$$

where the box marks the matrix entry at $(0,0)$. Here, (α_i), (β_i), and (γ_i) are bounded sequences of complex numbers. We think of A as a linear operator acting via matrix-vector multiplication on $\ell^p(\mathbb{Z})$. We develop inclusion sets for $\operatorname{spec} A$ and $\operatorname{spec}_\varepsilon A$ in terms of the spectra of finite section operators $A_{n,k} : X_{n,k} \to X_{n,k}$ defined by $A_{n,k} := P_{n,k} A|_{X_{n,k}}$, with matrix representation

$$A_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix},$$

obtaining results reminiscent of the Gershgorin theorem and its generalisations [50].

Defining $\Sigma_\varepsilon^n(A) := \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_\varepsilon A_{n,k}$, we show that, for $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\operatorname{spec} A \subseteq \overline{\Sigma_{\varepsilon_n}^n(A)}, \quad \operatorname{spec}_\varepsilon A \subseteq \Sigma_{\varepsilon+\varepsilon_n}^n(A),$$

where ε_n is given explicitly as the solution of a nonlinear equation, with $\varepsilon_n < \eta_n := 2(\|\alpha\|_\infty + \|\gamma\|_\infty) \sin(\pi/(2n+2))$. In general $\overline{\Sigma_{\epsilon_n}^n(A)}$ may be much larger than $\operatorname{spec} A$ and does not converge to $\operatorname{spec} A$ as $n \to \infty$, but in some cases $\overline{\Sigma_{\epsilon_n}^n(A)} = \operatorname{spec} A$ for all n.

Then we modify our first method to get sharper inclusion sets in certain cases for the spectrum and pseudospectra of A, in Section 3.3, in terms of the quasi-circulant matrices

$$\hat{A}_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \alpha_k \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ \gamma_{k+n+1} & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}.$$

We obtain the following results: for $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\operatorname{spec} A \subseteq \overline{\Pi^n_{\varepsilon_n}(A)}, \quad \operatorname{spec}_\varepsilon A \subseteq \Pi^n_{\varepsilon+\varepsilon_n}(A),$$

where $\Pi^n_\varepsilon(A) := \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_\varepsilon \hat{A}_{n,k}$ and ε_n is given explicitly as the solution of a nonlinear equation, with $\varepsilon_n < \eta_n := 2(\|\alpha\|_\infty + \|\gamma\|_\infty)\sin(\pi/2n)$.

In Chapter 4, we let $P_{n,k} : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ be the projection operator given by

$$(P_{n,k}x)_j = \begin{cases} x_j, & j = k+1, k+2, \dots, k+n, \\ 0, & \text{otherwise,} \end{cases}$$

and let $X_{n,k} := P_{n,k}(\ell^2(\mathbb{Z}))$ be the n-dimensional range of $P_{n,k}$. The method we propose in Chapter 4 modifies the above methods, with something of the flavours of [13, 47]. For $n \in \mathbb{N}$, $\eta > 0$, and $\lambda \in \mathbb{C}$ let

$$B^+_{n,k}(\lambda) := P_{n,k}(A-\lambda I)^*(A-\lambda I)\big|_{X_{n,k}}, \quad B^-_{n,k}(\lambda) := P_{n,k}(A-\lambda I)(A-\lambda I)^*\big|_{X_{n,k}},$$

and let

$$\Gamma^n_\eta(A) := \bigcup_{k\in\mathbb{Z}} \left\{\lambda \in \mathbb{C} : \min\left(\min \operatorname{spec} B^+_{n,k}(\lambda), \min \operatorname{spec} B^-_{n,k}(\lambda)\right) < \eta^2\right\}.$$

Then we show that, for $n \in \mathbb{N}$ and $\varepsilon > 0$,

$$\operatorname{spec} A \subseteq \overline{\Gamma^n_{\eta_n}(A)}, \quad \Gamma^n_\varepsilon(A) \subseteq \operatorname{spec}_\varepsilon A \subseteq \Gamma^n_{\varepsilon+\eta_n}(A),$$

so that $\overline{\Gamma^n_{\eta_n}(A)} \to \operatorname{spec} A$ in the Hausdorff metric as $n \to \infty$.

In Chapter 5, we focus on a study of the spectrum and pseudospectra of the operator A^b defined by (1.3), with b a random sequence of ± 1's. We investigate the symmetries of the spectrum of A^b and compute the numerical range of A^b. Besides that exploration, our main result is that the spectra and pseudospectra of the finite section matrices A_n^b, defined by (1.2), are contained in those of the two-sided infinite matrix A^b. Combining this result with our results on inclusion sets for tri-diagonal operators (as discussed in more detail in Chapter 3), we show that, with probability one, $\overline{\Sigma_{\varepsilon_n}^n(A^b)} \to \operatorname{spec} A^b$ in the Hausdorff metric as $n \to \infty$. Note that for $n \in \mathbb{N}$, $\overline{\Sigma_{\varepsilon_n}^n(A^b)}$ is the closure of the union of the ε_n-pseudospectra of all 2^{n-1} distinct tridiagonal submatrices of the random operator A^b (2^{n-1} is the number of different sequences of ± 1's that can be chosen as the first subdiagonal of A_n^b) with $\varepsilon_n = 4\sin\left(\dfrac{\pi}{(2n+2)}\right)$. Further, we quantify the rate of convergence by showing that, with probability one,

$$\operatorname{spec} A^b \subseteq \overline{\Sigma_{\varepsilon_n}^n(A^b)} \subseteq \operatorname{spec}_{\varepsilon_n} A^b$$

for $n \in \mathbb{N}$. As a consequence of this and related results, we derive also convergence of $\operatorname{spec}_{\varepsilon_n} A_n^b$ to $\operatorname{spec} A^b$ (with probability one) in the Hausdorff metric as $n \to \infty$.

Chapter 2

Preliminaries

In this chapter we introduce the background material needed for this thesis. Most of the material covered in this chapter is standard and therefore, we do not include the proofs of all the results. We have included proofs that we believe contribute to the better understanding of the material that is to be presented and have given references to the proofs of all other results. The aim was to make the thesis as self-contained as possible. We begin by setting forth the basic notation we will use throughout this thesis.

2.1 General Notation and Standard Results in Linear Operator Theory

$\mathbb{Z}$ denotes the set of all integers and $\mathbb{N}, \mathbb{R}$ and $\mathbb{C}$ denote the natural, real and complex numbers, respectively. Moreover, $\mathbb{C}^n = \underbrace{\mathbb{C} \times \cdots \times \mathbb{C}}_{n\text{terms}}$. For any $z \in \mathbb{C}$ we will write $z = x + iy$ where $x = \text{Re}\ (z)$ is the real part of z and $y = \text{Im}\ (z)$ is the imaginary part. We will write $|z|$ for the modulus of z. We will denote by $\overline{z}$ the complex conjugate of z if z is a complex number and by $\overline{A}$ the closure of the set A when A is a subset of a metric space.

The polynomial convex hull of a compact subset K of $\mathbb{C}$ is defined to be the complement of the unbounded component of $\mathbb{C} \setminus K$, i.e. K together with all open regions enclosed by this set.

Let X, Y be Banach spaces, i.e. complete normed vector spaces. We define $B(X, Y)$ as the set of bounded linear operators from X to Y, and we also define $B(X) := B(X, X)$ as the set of all bounded linear operators on X. A linear space with an inner product $(\cdot, \cdot)$ that is complete with respect to the norm defined by $\|x\| = \sqrt{(x, x)}$ for every $x \in X$ is called a Hilbert space.

Let $\ell^p(\mathbb{Z}, U)$ denote the standard space of U-valued sequences $x = (x_j)_{j\in\mathbb{Z}}$ which has p-norm, $\|\cdot\|_p$ defined by $\|x\|_p = (|x_1|^p + |x_2|^p + \dots)^{1/p}$. We omit the second parameter U in this notation if $U = \mathbb{C}$, i.e. $\ell^p(\mathbb{Z}) := \ell^p(\mathbb{Z}, \mathbb{C})$. In this thesis, if we do not specify otherwise, we use $X = \ell^2(\mathbb{Z})$,i.e., $p = 2, U = \mathbb{C}$, therefore X is a Hilbert space with the usual inner product.

Definition 2.1. *If X and Y are Banach spaces, we will say that $A \in B(X, Y)$ is invertible if there exists an operator $B \in B(Y, X)$ such that*

$$AB = I_Y, BA = I_X,$$

where I_X and I_Y denote the identity operators on X and Y, respectively. We call B the inverse of A and denote it by A^{-1}.

Theorem 2.2. *Suppose X is a Banach space and $A \in B(X)$ is invertible* *with inverse A^{-1}. Then, for any $E \in B(X)$ with $\|E\| < \dfrac{1}{\|A^{-1}\|}$, $A+E$ is invertible and*

$$\left\|(A+E)^{-1}\right\| \leq \frac{\|A^{-1}\|}{1-\|E\|\,\|A^{-1}\|}.$$

Conversely, for any $\mu > \dfrac{1}{\|A^{-1}\|}$, there exists $E \in B(X)$ with $\|E\| < \mu$ such that $A+E$ is not invertible (in fact, not injective).

Proof. See [49, Theorem 4.1] ■

Theorem 2.3. *Let X and Y be Hilbert spaces and let $A \in B(X,Y)$. Then there exists one and only one $A^* \in B(Y,X)$ with the property*

$$(Ax, y) = (x, A^*y) \tag{2.1}$$

*for all $x \in X, y \in Y$. This operator A^* is called the **adjoint operator** to A.*

Proof. See [31]. ■

Theorem 2.4. *Let X and Y be Hilbert spaces and let $A \in B(X,Y)$ and A^* be the adjoint of A. Then $A^* \in B(Y,X)$ and $\|A\| = \|A^*\|$. Further, A^* is invertible iff A is invertible, and if they are both invertible then $(A^*)^{-1} = (A^{-1})^*$ and so $\|A^{-1}\| = \|(A^*)^{-1}\|$.*

Definition 2.5. *Let X be a Hilbert space and $A \in B(X)$. A is said to be **normal** if A commutes with its adjoint A^*, i.e, $AA^* = A^*A$.*

Definition 2.6. *Let X be a Hilbert space and $A \in B(X)$. A is said to be **unitary** if the adjoint of A is its inverse, i.e, $AA^* = A^*A = I$, the identity operator.*

Definition 2.7. *Let X be a Hilbert space and $A \in B(X)$. A is said to be **self-adjoint (or Hermitian)** if $A^* = A$.*

Theorem 2.8. *Let $A \in B(X,Y)$.*

1. *$(A^*)^* = A$.*

2. $(A+B)^* = A^* + B^*$.

3. $(AB)^* = B^*A^*$.

4. $(cA)^* = \overline{c}A^*$ *for any* $c \in \mathbb{C}$.

Note that unitary operators and self-adjoint operators are examples of normal operators.

Lemma 2.9. *Let S be a linear subspace of a Hilbert space X. Then S is dense in X if and only if*

$$x \in X, (x,z) = 0, \forall z \in S \quad \Rightarrow \quad x = 0. \tag{2.2}$$

Proof. ($\Rightarrow$) If S is dense in X, then for all $x \in X, \varepsilon > 0$, there exists a $y \in S$ such that $\|x - y\| < \varepsilon$. So if $(x,z) = 0, \forall z \in S$ then $(x,y) = 0 \Leftrightarrow (x,x) = (x, x-y)$. Thus, by Cauchy-Schwarz inequality,

$$\|x\|^2 \leq \|x\| \, \|x-y\| < \|x\| \, \varepsilon$$

which implies that $\|x\| \leq \varepsilon$, since $\|x\| \geq 0$.

($\Leftarrow$) Next, we show that if (2.2) holds then S is dense in X. Note that $\overline{S}$ is a closed linear subspace of X. If S is not dense in X then $\overline{S} \neq X$. Let $x^* \in X \setminus \overline{S}$. Then, by the projection theorem, $x^* = s + x$ where $s \in \overline{S}$ and $x \in \overline{S}^{\perp}$, i.e. $(x,z) = 0$ for all $z \in \overline{S}$, and, since $x^* \notin \overline{S}, x \neq 0$.

■

Theorem 2.10. *Let X and Y be Hilbert spaces and $A \in B(X,Y)$. Then $A^* \in B(Y,X)$ is injective if and only if $A(X)$ is dense in Y.*

Proof. First note that A^* is injective if and only if

$$y \in Y, A^*y = 0 \Rightarrow y = 0.$$

This is equivalent to

$$\begin{aligned} & y \in Y, (A^*y, x) = 0 \text{ for all } x \in X \Rightarrow y = 0 \\ \Leftrightarrow & y \in Y, (y, Ax) = 0 \text{ for all } x \in X \Rightarrow y = 0 \\ \Leftrightarrow & y \in Y, (y, z) = 0 \text{ for all } z \in A(X) \Rightarrow y = 0. \end{aligned}$$

The required result follows from Lemma 2.9. ■

Theorem 2.11. *Let X and Y be Banach spaces. Then $A \in B(X,Y)$ is invertible iff A is bijective.*

Proof. See [42, Corollary 2.12]. ■

The main statement of Theorem 2.11 is that the inverse map $B : Ax \longmapsto x$ from Y to X is automatically bounded and linear if A is bounded, linear and bijective. This fact requires X and Y to be Banach spaces.

Corollary 2.12. *Let X and Y be Hilbert spaces and $A \in B(X,Y)$. If A is not invertible then*

(a) A is not injective,

(b) A^* is not injective,

or (c) $A(X)$ is not closed.

Proof. Suppose that A and A^* are injective and $A(X)$ is closed. By Theorem 2.10, $A(X)$ is dense in X. Hence, A is surjective. Since A is both injective and surjective, it follows that A is invertible by Theorem 2.11. ■

Lemma 2.13. ***(Lax-Milgram)*** *Let X be a Hilbert space. If an operator $A \in B(X)$ has a constant $c > 0$ such that $Re(Ax,x) \geq c\|x\|^2$ for all $x \in X$ then A is invertible and $\|A^{-1}\| \leq c^{-1}$.*

Proof. See [31]. ■

Definition 2.14. *Let X be a Banach space and $A \in B(X)$. Then*

$$\nu(A) := \inf_{x \in X \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \inf_{\|x\|=1} \|Ax\|$$

*is referred to as the **lower norm** of A. We will say that A is bounded below if $\nu(A) > 0$.*

Proposition 2.15. *Let X be a Banach space. An operator $A \in B(X)$ is bounded below if and only if A is injective and has a closed range.*

Proof. See [33, Lemma 2.32]. ■

Note that, if A has closed range then $Y = A(X)$ is a Banach space and, provided also that A is injective, then $A : X \to Y$ is bijective, and so invertible. Therefore, this proposition is a corollary of Theorem 2.11.

Theorem 2.16. *If X is a Hilbert space then $A \in B(X)$ is invertible iff $\nu(A) > 0$ and $\nu(A^*) > 0$. Furthermore, if A is invertible, then A^* is also invertible, and*

$$\nu(A) = \nu(A^*) = \frac{1}{\|A^{-1}\|} = \frac{1}{\|(A^*)^{-1}\|}$$

Proof. See [33, Lemma 2.35], and Theorem 2.4. ∎

2.2 Spectral Theory

Definition 2.17. *Let X be a Banach space and $A \in B(X)$. Let*

$$\operatorname{spec} A = \{\lambda \in \mathbb{C} : A - \lambda I \textit{ is not an invertible operator on } X\}$$

be the (invertibility) spectrum of A,

$$\operatorname{spec}_{\text{point}} A = \{\lambda \in \mathbb{C} : (A - \lambda I) \textit{ is not an injective operator on } X\}$$

be the point spectrum of A and

$$\operatorname{spec}_{\text{ess}} A = \{\lambda \in \mathbb{C} : (A - \lambda I) \textit{ is not Fredholm on } X\}$$

be the essential spectrum of A.

Note that Definition 2.45 below explains what we mean by Fredholmness. In the special case $X = \ell^p(\mathbb{Z})$, we write $\operatorname{spec}^p A, \operatorname{spec}^p_{\text{point}} A$ and $\operatorname{spec}^p_{\text{ess}} A$ for $\operatorname{spec} A, \operatorname{spec}_{\text{point}} A$ and $\operatorname{spec}_{\text{ess}} A$, respectively, to underline the dependence on p. However, for any banded operator A (see Definition 2.46), $\operatorname{spec}^p A$ and $\operatorname{spec}^p_{\text{ess}} A$ do not depend on the choice of $p \in [1, \infty]$ (see [32, 34]), which is why we will simply write $\operatorname{spec} A$ and $\operatorname{spec}_{\text{ess}} A$ in that case.

Definition 2.18. *Let A be a bounded linear operator on a Banach space X. For $\varepsilon > 0$, the ε-**pseudospectrum** is*

$$\operatorname{spec}_{\varepsilon} A = \{\lambda \in \mathbb{C} : \left\|(\lambda I - A)^{-1}\right\| > \frac{1}{\varepsilon}\},$$

with the convention that $\|(\lambda I - A)^{-1}\| := \infty$ if $\lambda \in \operatorname{spec} A$.

From the definition, it follows that the ε-pseudospectra associated with various ε are nested sets,

$$\operatorname{spec}_{\varepsilon_1} A \subseteq \operatorname{spec}_{\varepsilon_2} A, \qquad 0 < \varepsilon_1 \leq \varepsilon_2.$$

Theorem 2.19. *Let X be a Banach space and $A \in B(X)$ and $\varepsilon > 0$ be arbitrary. The ε-pseudospectrum $\operatorname{spec}_{\varepsilon}(A)$ of A is the set of $\lambda \in \mathbb{C}$ defined equivalently by any of the conditions*

$$a.)\ \left\|(A - \lambda I)^{-1}\right\| > \varepsilon^{-1}. \tag{2.3}$$

$$b.)\lambda \in \operatorname{spec}(A + E) \text{ for some } E \in B(X) \text{ with } \|E\| < \varepsilon. \tag{2.4}$$

$$c.)\lambda \in \operatorname{spec}(A) \text{ or } \|(A - \lambda I)x\| < \varepsilon \text{ for some } x \in X \text{ with } \|x\| = 1. \tag{2.5}$$

If $\|(A - \lambda I)x\| < \varepsilon$ as in (2.5), then λ is said to be an ε-pseudoeigenvalue of A and x the corresponding ε-pseudoeigenvector (or pseudoeigenfunction or pseudomode).

Proof. See [47, page 16] ∎

Theorem 2.20. *For any matrix A,*

$$\operatorname{spec}_{\varepsilon} A \supseteq \operatorname{spec} A + B_{\varepsilon}(0) \qquad \forall \varepsilon > 0,$$

and if A is normal and $\|\cdot\| = \|\cdot\|_2$, then

$$\operatorname{spec}_{\varepsilon} A = \operatorname{spec} A + B_{\varepsilon}(0) \qquad \forall \varepsilon > 0.$$

Proof. See [47, page 19]. ∎

Proposition 2.21. *Let X be a Hilbert space. For any $A \in B(X)$, $\lambda \in \mathbb{C}$, and $\varepsilon > 0$, we have* $\left\|\left(A^* - \overline{\lambda} I\right)^{-1}\right\| = \left\|(A - \lambda I)^{-1}\right\|$. *Therefore,* $\operatorname{spec} A^* = \left\{\overline{\lambda} : \lambda \in \operatorname{spec} A\right\}$ *and* $\operatorname{spec}_{\varepsilon} A^* = \left\{\overline{\lambda} : \lambda \in \operatorname{spec}_{\varepsilon} A\right\}$.

Proof. See Theorem 2.16. ∎

Theorem 2.22. *For any $A \in B(X)$, where X is a Hilbert space, if A is self-adjoint, then* $\operatorname{spec} A \subseteq \mathbb{R}$.

Theorem 2.23. *Let X, Y be Hilbert spaces and $A \in B(X, Y)$. Then, A^*A is Hermitian and*

$$\|A^*A\| = \|A\|^2 .$$

Proof. Since $(A^*A)^* = A^*A$, it follows that A^*A is Hermitian. Consider

$$\begin{aligned} \|A\|^2 &= \sup\{(Ax, Ax) : x \in X, \|x\| = 1\} \\ &= \sup\{(A^*Ax, x) : x \in X, \|x\| = 1\} \\ &\leq \sup\{\|A^*Ax\| : x \in X, \|x\| = 1\} \\ &= \|A^*A\| , \end{aligned}$$

and we can see that $\|A^*A\| \leq \|A^*\| \|A\| = \|A\|^2$. ∎

Theorem 2.24. *For any $A \in B(X, Y)$, where X and Y are Hilbert spaces,*

$$\nu(A^*A) = \min \operatorname{spec}(A^*A) = \{\nu(A)\}^2 = \inf_{\substack{\phi \in X \\ \|\phi\|=1}} (A^*A\phi, \phi).$$

Proof. We firstly show that $\nu(A^*A) = \{\nu(A)\}^2$ by considering 3 cases, $\nu(A) \neq 0$ and $\nu(A^*) \neq 0$, $\nu(A) = 0$, and $\nu(A) \neq 0$ but $\nu(A^*) = 0$.

Case 1 : $\nu(A) \neq 0$ and $\nu(A^*) \neq 0$. In this case, A, A^* and A^*A are all invertible and, by Theorem 2.16 and Theorem 2.23, $\nu(A^*A) = \|(A^*A)^{-1}\| = \|A^{-1}(A^{-1})^*\| = \|A^{-1}\|^2 = (\nu(A))^2$.

Case 2 : $\nu(A) = 0$. We can see that $\nu(A^*A) \leq \|A^*\| \nu(A) = 0$ and so $\nu(A^*A) = \{\nu(A)\}^2$.

Case 3 : $\nu(A) \neq 0$ and $\nu(A^*) = 0$. In this case define $\widetilde{A} : X \to A(X)$ by $\widetilde{A}x = Ax$. We know that $\nu(\widetilde{A}) = \nu(A) > 0$. From Proposition 2.15, $\widetilde{A}$

is injective and $\widetilde{A}(X)$ is closed. It is obvious that $\widetilde{A}$ is surjective. Hence $\widetilde{A}$ is a bijection. From Theorem 2.11, $\widetilde{A}$ is invertible. Thus, from Proposition 2.16, $\left\|\widetilde{A}^{-1}\right\| = \dfrac{1}{\left\|\nu(\widetilde{A})\right\|} < \infty$, i.e. $\widetilde{A}$ has bounded inverse. Further, for $x \in X, y \in A(X)$,

$$(Ax, y) = (x, A^*y),$$

so $\widetilde{A}^*y = A^*y$, $y \in A(X)$, and so $\widetilde{A}^*\widetilde{A}x = A^*Ax$, $x \in X$, so that

$$\nu(\widetilde{A}^*\widetilde{A}) = \nu(A^*A).$$

By Theorem 2.23, we have then that

$$\begin{aligned} \left\|\widetilde{A}^{-1}\right\|^2 &= \left\|\widetilde{A}^{-1}\left(\widetilde{A}^{-1}\right)^*\right\| \\ &= \left\|\widetilde{A}^{-1}\left(\widetilde{A}^*\right)^{-1}\right\| \\ &= \left\|\left(\widetilde{A}^*\widetilde{A}\right)^{-1}\right\|. \end{aligned}$$

Thus, and by Theorem 2.16,

$$\left(\nu\left(\widetilde{A}\right)^2\right) = \nu\left(\widetilde{A}^*\widetilde{A}\right).$$

Therefore,

$$\left(\nu\left(A\right)^2\right) = \nu\left(A^*A\right)$$

Now we are showing that $\inf \operatorname{spec}(A^*A) = \nu(A^*A)$. If $\mu < \nu(A^*A)$ then

$$\begin{aligned} ((A^*A - \mu I)x, x) &= (A^*Ax, x) - \mu(x, x) \\ &= (Ax, Ax) - \mu(x, x) \\ &= \|Ax\|^2 - \mu\|x\|^2 \\ &\geq \left((\nu(A))^2 - \mu\right)\|x\|^2 \\ &= (\nu(A^*A) - \mu)\|x\|^2, \end{aligned}$$

so $A^*A - \mu I$ is invertible by Lemma 2.13. So $\inf \operatorname{spec}(A^*A) \geq \nu(A^*A)$. If $\mu = \nu(A^*A)$ then, for every $\varepsilon > 0$, there exists $x \in X$ with $\|x\| = 1$ such that

$$\|A^*Ax\| < \nu(A^*A) + \varepsilon,$$

and then

$$\begin{aligned}\|(A^*A-\mu)x\|^2 &= ((A^*A-\mu)\,x,(A^*A-\mu)\,x)\\ &= (A^*Ax,A^*Ax)-\mu\,(A^*A,x)-\mu\,(x,A^*Ax)+\mu^2(x,x)\\ &= \|A^*Ax\|^2-2\mu\,\|Ax\|^2+\mu^2\,\|x\|^2\\ &= \|A^*Ax\|^2-\mu\,\|Ax\|^2+\mu^2.\end{aligned}$$

Now

$$\|Ax\|\geq\nu(A)\,\|x\|=\nu(A),$$

so

$$\begin{aligned}\|(A^*A-\mu)\,x\|^2 &< (\nu(A^*A)+\varepsilon)^2-2\mu\,(\nu(A))^2+\mu^2\\ &= \left(2\varepsilon\nu(A^*A)+\varepsilon^2\right)+(\nu(A^*A)-\mu)^2\\ &= 2\varepsilon\nu(A^*A)+\varepsilon^2.\end{aligned}$$

Since $\varepsilon>0$ is arbitrarily small, we have shown that $\nu(A^*A-\mu I)=0$, so that $\mu=\nu(A^*A)\in\operatorname{spec}(A^*A)$.

Thus $\inf\operatorname{spec}(A^*A)=\nu(A^*A)$. ■

Corollary 2.25. *For any $A\in B(X)$, where X is a Hilbert space,* $\operatorname{spec}(A^*A)\subseteq[0,\infty)$.

Proof. It is obvious that A^*A is self-adjoint. From Theorem 2.22, we can see that $\operatorname{spec}(A^*A)\subseteq\mathbb{R}$. By Theorem 2.24,

$$\min\operatorname{spec}(A^*A)=\{\nu(A)\}^2\geq 0.$$

Therefore, if $\lambda\in\operatorname{spec}(A^*A)$ then $\lambda\geq 0$. ■

By Theorem 2.16 we have then that

Theorem 2.26. *If X is a Hilbert Space and $A\in B(X)$ then*

$$\operatorname{spec}A=\{\lambda:\nu(\lambda I-A)=0\ \textit{or}\ \nu\left((\lambda I-A)^*\right)=0\}$$

and, for $\varepsilon > 0$*,*

$$\begin{aligned}
\operatorname{spec}_{\varepsilon} A &= \operatorname{spec} A \bigcup \{\lambda : \nu(\lambda I - A) < \varepsilon\} \\
&= \operatorname{spec} A \bigcup \{\lambda : \nu\left((\lambda I - A)^*\right) < \varepsilon\} \\
&= \{\lambda : \nu(\lambda I - A) < \varepsilon\} \cup \{\lambda : \nu\left((\lambda I - A)^*\right) < \varepsilon\} \\
&= \left\{\lambda : \nu\left((\lambda I - A)^*(\lambda I - A)\right) < \varepsilon^2\right\} \\
&\cup \left\{\lambda : \nu\left((\lambda I - A)(\lambda I - A)^*\right) < \varepsilon^2\right\} \\
&= \left\{\lambda : \min \operatorname{spec}\left((\lambda I - A)^*(\lambda I - A)\right) < \varepsilon^2\right\} \\
&\cup \left\{\lambda : \min \operatorname{spec}\left((\lambda I - A)(\lambda I - A)^*\right) < \varepsilon^2\right\} \\
&= \left\{\lambda : \inf_{\phi \in X} \left((\lambda I - A)^* (\lambda I - A)\phi, \phi\right) < \varepsilon^2\right\} \\
&\cup \left\{\lambda : \inf_{\phi \in X} \left((\lambda I - A)(\lambda I - A)^* \phi, \phi\right) < \varepsilon^2\right\}.
\end{aligned}$$

Definition 2.27. *For a bounded operator A on a Hilbert space X,* ***the numerical range of*** *A is*

$$W(A) = \{(Ax, x) : x \in X, \|x\| = 1\}$$

Theorem 2.28. ***(Toeplitz-Hausdorff)*** *Let X be a Hilbert space and $A \in B(X)$. The numerical range of the operator A is a convex set, and*

$$\operatorname{spec} A \subseteq \overline{W(A)} \subseteq \{\lambda : |\lambda| \leq \|A\|\}.$$

Proof. See [13]. ■

Definition 2.29. *Let X be a topological space, and f is a function from X into the extended real number $\mathbb{R}^*$; $f : X \to \mathbb{R}^*$. Then f is said to be upper semicontinuous if $f^{-1}([-\infty, \alpha)) = \{x \in X : f(x) < \alpha\}$ is an open set in X for all $\alpha \in \mathbb{R}$.*

Definition 2.30. *Let $U \subseteq \mathbb{R}^n$. $f : U \to \mathbb{R}$ is said to be harmonic if it satisfies Laplace's equation, i.e.*

$$\frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + \cdots + \frac{\partial^2 f}{\partial x_n^2} = 0$$

everywhere on U.

Definition 2.31. *Let $G \subseteq \mathbb{R}^n$ and let $\varphi : G \to \mathbb{R} \cup \{-\infty\}$ be an upper semicontinuous function. Then φ is said to be subharmonic if for every $x \in G$ and $r > 0$ such that $\overline{B_r(x)} \subseteq G$, and every real-valued continuous function h on $\overline{B_r(x)}$ that is harmonic in $B_r(x)$ and satisfies $\varphi(x) \leq h(x)$ for all x in the boundary of $B_r(x)$, it holds that $\varphi(x) \leq h(x)$ for all $x \in B_r(x)$.*

Note that the resolvent norm $R_A : \lambda \longmapsto \left\|(A - \lambda I)^{-1}\right\|$ is a subharmonic function on $\mathbb{C} \setminus \operatorname{spec} A$ (see [2] and [47, Theorem 4.2]) subject to the following maximum principle due to Daniluk (see [21, Theorem 3.32] or [3, Theorems 7.5, 7.6] but also [43, 44]):

Theorem 2.32. *If $U \subseteq \mathbb{C} \setminus \operatorname{spec} A$ is open and $R_A(\lambda) \leq M$ for all $\lambda \in U$ then $R_A(\lambda) < M$ for all $\lambda \in U$.*

Definition 2.33. *Let S and T be two non-empty subsets of $\mathbb{C}$. We define their Hausdorff distance $d_H(S,T)$ by*

$$\begin{aligned} d_H(S,T) &= \max\left\{\sup_{s\in S}\inf_{t\in T}|s-t|\,, \sup_{t\in T}\inf_{s\in S}|s-t|\right\}, \\ &= \max\left\{\sup_{s\in S} d(s,T), \sup_{t\in T} d(t,S)\right\}, \end{aligned}$$

where, for $a \in \mathbb{C}$ and a non-empty set $C \subseteq \mathbb{C}$, $d(a,C) := \inf_{c\in C}|a-c|$.

Theorem 2.32 implies the following corollary.

Corollary 2.34. *Suppose $A \in B(X)$ where X is a Hilbert space. Then, for every $\varepsilon > 0$,*

$$\overline{\operatorname{spec}_\varepsilon A} = \left\{\lambda \in \mathbb{C} : \left\|(A-\lambda I)^{-1}\right\| \geq \varepsilon^{-1}\right\}. \tag{2.6}$$

For $\varepsilon \geq 0$

$$d_H\left(\operatorname{spec}_\varepsilon A, \operatorname{spec}_\eta A\right) \to 0 \tag{2.7}$$

as $\eta \to \varepsilon^+$, and

$$d_H\left(\operatorname{spec} A, \operatorname{spec}_\varepsilon A\right) \to 0 \tag{2.8}$$

as $\varepsilon \to 0$.

Proof. See [7] for the proof of (2.6). In order to prove (2.7), let $\varepsilon > 0$. We will show that for every $r > 0$ there exists a $\delta > 0$ such that $\operatorname{spec}_{\varepsilon+\delta}A \subseteq B_r(\operatorname{spec}_\varepsilon A)$,where $B_r(S)$ for a set S is the union of all $B_r(s)$ with s in S(i.e. the r-neighbourhood of S).

Suppose the converse is true. So there is a $r > 0$ such that for every $\delta > 0$ there is a $\lambda \in \operatorname{spec}_{\varepsilon+\delta}A$ with $d(\lambda, \operatorname{spec}_\varepsilon A) > r$. In particular, putting $\delta = 1/n$ with $n = 1, 2, ...$, there is a sequence $\lambda_1, \lambda_2, ...$ such that $R_A(\lambda_n) > 1/(\varepsilon+1/n)$ and $U_r(\lambda_n)$ is disjoint from $\operatorname{spec}_\varepsilon A$. Since $R_A(\lambda) \to 0$ as $|\lambda|$ goes to infinity, this sequence must be bounded and therefore have a convergent subsequence. Call its limit λ_0. It follows that $R_A(\lambda_0) \geq 1/\varepsilon =: M$ and that the disk $U := B_r(\lambda_0)$ is disjoint from $\operatorname{spec}_\varepsilon A$, so that $R_A(\lambda) \leq 1/\varepsilon = M$ for all $\lambda \in U$. By Theorem 2.32 it follows that $R_A(\lambda) < M = 1/\varepsilon$ for all $\lambda \in U = B_r(\lambda_0)$ but this contradicts $R_A(\lambda_0) \geq M$.

(2.8) can be shown by a similar, slightly simpler argument (putting $\varepsilon = 0$). ■

2.3 Finite and Infinite Matrices

2.3.1 Classical Matrix Algebra

Let X and Y be finite dimensional vector spaces (over the field $\mathbb{C}$) , i.e. $X \cong \mathbb{C}^n$ and $Y \cong \mathbb{C}^m$ for some $m, n \in \mathbb{N}$. Let $T \in B(X, Y)$ and $(e_1, \ldots, e_n)$ and $(f_1, \ldots, f_m)$ be ordered bases of X and Y, respectively. We can write

$$T(e_j) = \sum_{i=1}^{m} a_{ij} f_i$$

for every $j = 1, \ldots, n$ and $a_{ij} \in \mathbb{C}$. From elementary linear algebra, there exists an associated $m \times n$ matrix A, defined by $A = (a_{ij})$, such that $Tx = Ax$ for every $x \in X$. Note that the matrix representation depends on the bases of the vector space and also the order of elements in bases.

Note that if $X = \mathbb{C}^n$ and $Y = \mathbb{C}^m$ are equipped with the 2-norm, where $m, n \in \mathbb{N}$ then $\mathbb{C}^n$ and $\mathbb{C}^m$ are Hilbert spaces and the associated linear mapping $T : \mathbb{C}^n \to \mathbb{C}^m$ is a bounded linear operator. Moreover, by Theorem 2.3 there exists an adjoint operator $T^* \in B(\mathbb{C}^m, \mathbb{C}^n)$ and an adjoint matrix associated with T^* can be defined as the following.

Definition 2.35. *Let $T : \mathbb{C}^n \to \mathbb{C}^m$ be a bounded linear operator. The **adjoint** of the $n \times m$ matrix $A = (a_{ij})$ is an $m \times n$ matrix $A^H = (a^H_{ij})$ defined by $a^H_{ij} = \overline{a_{ji}}$, i.e. A^H is the conjugate transpose of A. It holds that*

$$T^*x = A^H x$$

for every $x \in \mathbb{C}^m$. If A is a matrix with real entries, A^H is equal to A^T, the transpose of the matrix A.

Adjoints satisfy the following identities:

(a) $\left(A^H\right)^H = A$.

(b) $(A + B)^H = A^H + B^H$.

(c) $(AB)^H = B^H A^H$

(d) $(cA)^H = \bar{c}A^H$ for any $c \in \mathbb{C}$.

Definition 2.36. *A matrix A is **self-adjoint (or Hermitian)** if it equals its complex conjugate transpose A^H.*

Definition 2.37. *A matrix A is **unitary** if A has its inverse equal to its complex conjugate transpose A^H.*

Definition 2.38. *A matrix A is **normal** if it commutes with its adjoint, i.e. $AA^H = A^H A$.*

An equivalent characterization is that A is normal if it has a complete set of orthogonal eigenvectors, that is, if it is unitarily diagonalizable:

$$A = UDU^H,$$

where U is a unitary matrix and D is a diagonal matrix. Since A and A^H have the same eigenvalues; they are simultaneously diagonalizable (see [48]).

Proposition 2.39. *The eigenvalues of a Hermitian matrix are real.*

Definition 2.40. *A real number λ is said to be a **singular value** of an operator A if it is the square root of one of the eigenvalues of A^*A.*

Definition 2.41. *Suppose A is an $m \times n$ matrix whose entries come from the set of complex numbers $\mathbb{C}$. Then there exists a factorization of the form*

$$A = U\Sigma V^*,$$

*where U is an $m \times m$ unitary matrix over $\mathbb{C}$, the matrix Σ is an $m \times n$ diagonal matrix with nonnegative real numbers on the diagonal, and V is an $n \times n$ unitary matrix over $\mathbb{C}$. This factorization is said to be a **singular value decomposition** of A.*

2.3.2 Matrix Representation of Operators and Sequence Space

Let $p \in [1, \infty]$, $X = \ell^p(\mathbb{Z})$ and $A \in B(X)$. For $k \in \mathbb{Z}$ let $E_k : \mathbb{C} \to X$ and $R_k : X \to \mathbb{C}$ be extension and restriction operators, defined by $E_k x = (\ldots, 0, x, 0, \ldots)$ for $x \in \mathbb{C}$, with the x standing at the kth place in the sequence, and by $R_k x = x_k$, for $x = (x_j)_{j\in\mathbb{Z}} \in X$. Then, the matrix entries of $[A]$ are defined as

$$a_{ij} := R_i A E_j \in B(\mathbb{C}) \cong \mathbb{C}, \qquad i, j \in \mathbb{Z}, \tag{2.9}$$

and $[A]$ is called the **matrix representation** of A.

Conversely, given a matrix $M = [m_{ij}]_{i,j\in\mathbb{Z}}$ with entries in $\mathbb{C}$, we will say that M induces the operator

$$(Bx)_i = \sum_{j=-\infty}^{\infty} m_{ij} x_j, \qquad i \in \mathbb{Z}, \tag{2.10}$$

if the sum converges in $\mathbb{C}$ for every $i \in \mathbb{Z}$ and every $x = (x_j)_{j\in\mathbb{Z}} \in \ell^p(\mathbb{Z})$ and if the resulting operator B is a bounded mapping on $\ell^p(\mathbb{Z})$

It is not hard to see that if M is an infinite matrix and B is induced, via (2.10), by M then the matrix representation $[B]$ from (2.9) is equal to M. It does not work quite like that the other way round: For $p = \infty$, there are operators $A \in B(\ell^p(\mathbb{Z}))$(e.g. see Example 1.26 c in [33]) for which the matrix representation $M := [A]$ induces an operator B that is different from A. However, for every $A \in B(\ell^p(\mathbb{Z}))$ with $p \in [1, \infty)$, the matrix $M := [A]$ with entries (2.9) induces the operator $B = A$.

Definition 2.42. *If $b = (b_i)_{i\in\mathbb{Z}}$ is a bounded sequence of $b_i \in \mathbb{C}$, then by M_b we will denote the **multiplication operator**, acting on every $x \in \ell^p(\mathbb{Z})$ by*

$$(M_b x)_i = b_i x_i \qquad \forall i \in \mathbb{Z}.$$

Definition 2.43. *For every $k \in \mathbb{Z}$, a **shift operator** is defined by*

$$(V_k x)_i = x_{i-k} \qquad i \in \mathbb{Z},$$

for every $x \in \ell^p(\mathbb{Z})$.

Lemma 2.44. *For every $k \in \mathbb{Z}$ and $b \in \ell^\infty(\mathbb{Z})$*

$$V_k M_b = M_{V_k b} V_k.$$

Proof. Let $x \in \ell^p(\mathbb{Z})$ be given.

$$(V_k M_b x)_i = (M_b x)_{i-k} = b_{i-k} x_{i-k} = (V_k b)_i (V_k x)_i = (M_{V_k b} V_k x)_i .$$

■

Definition 2.45. *Let X be any Banach space and $A \in B(X)$. Define $\alpha :=$ dim ker$A$ and $\beta :=$ dim coker A, where coker $A := X/$ im A. A is called a Fredholm operator if both numbers α and β are finite, in which case its image is closed.*

Definition 2.46. *An operator A on $\ell^p(\mathbb{Z})$ is called a **band operator of band-width** w if it can be written in the form*

$$A = \sum_{|\gamma|\le w} M_{b^{(\gamma)}} V_\gamma$$

with $b^{(\gamma)} \in \ell^\infty(\mathbb{Z})$ for every shift operator $\gamma \in \mathbb{Z}$ involved in the summation.

The matrix representation $[A]$ of a band operator A consequently is as the following:

Definition 2.47. *We call $[a_{ij}]_{i,j\in\mathbb{Z}}$ a band matrix of band-width w if all entries a_{ij} with $|i-j| > w$ vanish, or, what is equivalent, if the matrix is only supported on the $k-th$ diagonals with $|k| \leq w$.*

Throughout this thesis, we use the notation $\|\cdot\|$ as the 2-norm, i.e.,

$$\|A\| = \|A\|_2 = \sup_{\|v\|_2=1} \|Av\|_2 .$$

Moreover, when mentioning "the norm" we mean the 2-norm also. The norm of a finite dimensional matrix A is its largest singular value and the norm of the inverse is the inverse of the smallest singular value. This is most easily seen by looking at the singular value decomposition with unitary matrices U, V and $S = \text{diag}\,(s_1, s_2, \ldots, s_n)$ of A. Indeed, since unitary matrices are isometries, $\|A\| = \|USV^*\| = \|S\| = \max |s_k|$ and

$$\begin{aligned}\left\|A^{-1}\right\| &= \left\|(V^*)^{-1}S^{-1}U^{-1}\right\|, \\ &= \left\|S^{-1}\right\| \\ &= \left\|\text{diag}(\frac{1}{s_1}, \frac{1}{s_2}, \ldots, \frac{1}{s_n})\right\| \\ &= \frac{1}{\min |s_k|}.\end{aligned}$$

Hence $\|(A-\lambda I)^{-1}\| = (s_{\min}(A-\lambda))^{-1}$ where $s_{min}(A-\lambda I)$ denotes the smallest singular value of $A-\lambda I$, suggesting a fourth definition of the ε-pseudospectrum in the case of a finite matrix A:

$$\text{spec}_{\varepsilon}(A) = \{\lambda \in \mathbb{C} : s_{\min}(A-\lambda I) < \varepsilon\}. \tag{2.11}$$

Proposition 2.48. *Let A be the matrix representation of a bounded linear operator on a finite-dimensional Hilbert space X and λ_{min} denote the smallest eigenvalue of the Hermitian matrix A^HA. Then*

$$\nu(A) = \sqrt{\lambda_{min}(A^HA)}.$$

Proof. Note that, if λ is an eigenvalue of $A^H A$ and x is a corresponding eigenvector, i.e., $A^H Ax = \lambda x$, then

$$\lambda = \frac{(A^H Ax, x)}{(x, x)} = \frac{(Ax, Ax)}{(x, x)} = \frac{\|Ax\|^2}{\|x\|^2}.$$

We can observe that,

$$\lambda_{min}(A^H A) = \inf_{x\neq 0} \frac{\|Ax\|_2^2}{\|x\|_2^2} = \left(\inf_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2}\right)^2 = (\nu(A))^2 .$$

■

Let $A = (a_{ij})$ be an $n \times n$ matrix and $d_i = \sum_{j\neq i} |a_{ij}|$ for $i = 1, \ldots, n$. Then the set

$$D_i = \{\lambda \in \mathbb{C} : |\lambda - a_{ii}| \leq d_i\} \qquad i = 1, \ldots, n$$

is called the ith Gershgorin disc of the matrix A.

Theorem 2.49. ***(Gershgorin Circle Theorem)*** *Every eigenvalue of A lies within at least one of the Gershgorin discs.*

Proof. See [50]. ■

Gershgorin's original proof can be adapted to infinite matrices hence yielding a spectral inclusion set that we can use as a benchmark for comparison with the results of our methods 1, 1* and 2 in Chapter 3 and Chapter 4.

Theorem 2.50. *Let A be the bounded linear operator which operates on all spaces $\ell^p(\mathbb{Z})$, $p \in [1, \infty]$, via multiplication by (1.4). Putting*

$$r_i \ := \ \max\left(|\alpha_{i-1}| + |\gamma_{i+1}| \, , \, |\alpha_i| + |\gamma_i| \right) \tag{2.12}$$

for every $i \in \mathbb{Z}$, it holds that

$$\operatorname{spec} A \ \subset \ \overline{\bigcup_{i\in\mathbb{Z}} (\beta_i + r_i\mathbb{D})}, \tag{2.13}$$

where $\beta_i + r_i\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z - \beta_i| \leq r_i\}$ is the closed disk (Gershgorin circle) with radius $r_i \geq 0$ around $\beta_i \in \mathbb{C}$.

Proof. Suppose $\lambda \in \mathbb{C}$ is not contained in the right-hand side of (2.13), i.e. $|\beta_i - \lambda| > r_i$ for all $i \in \mathbb{Z}$ and

$$\delta \;:=\; \inf_{i\in\mathbb{Z}} \Big(|\beta_i - \lambda| \;-\; r_i \Big) \;>\; 0.$$

We show that $A - \lambda I$ is invertible on all spaces $\ell^p(\mathbb{Z})$, i.e. $\lambda \notin \operatorname{spec} A$. We do this in three steps:

1. Let $x \in \ell^\infty(\mathbb{Z})$ nonzero and $\varepsilon > 0$ be arbitrary and put $y := (A - \lambda I)x$. Write $x = (x_i)_{i\in\mathbb{Z}}$ and $y = (y_i)_{i\in\mathbb{Z}}$, and fix $i \in \mathbb{Z}$ s.t. $|x_i| > \|x\|_\infty - \varepsilon$. From $(\beta_i - \lambda)x_i = y_i - \alpha_{i-1}x_{i-1} - \gamma_{i+1}x_{i-1}$ we get

$$|\beta_i - \lambda||x_i| \;\le\; |y_i| \;+\; (|\alpha_{i-1}| + |\gamma_{i+1}|)\|x\|_\infty \;\le\; |y_i| \;+\; r_i\|x\|_\infty$$

by (2.12) and hence

$$\begin{aligned} \|y\|_\infty \;&\ge\; |y_i| \;\ge\; |\beta_i - \lambda||x_i| - r_i\|x\|_\infty \\ &\ge\; (|\beta_i - \lambda| \;-\; r_i)\|x\|_\infty \;-\; |\beta_i - \lambda|\varepsilon \\ &\ge\; \delta\|x\|_\infty \;-\; |\beta_i - \lambda|\varepsilon. \end{aligned}$$

Since this inequality holds for every $\varepsilon > 0$ (with i dependent on ε but $|\beta_i - \lambda|$ bounded) it follows that $\|y\|_\infty \;\ge\; \delta\|x\|_\infty$, so that $A - \lambda I$ is bounded below by $\delta > 0$ and hence is injective with closed range as a mapping $\ell^\infty(\mathbb{Z}) \to \ell^\infty(\mathbb{Z})$. Consequently, $A - \lambda I$ is also injective on $\ell^1(\mathbb{Z}) \subset \ell^\infty(\mathbb{Z})$.

2. Now let B be the operator that acts on $\ell^p(\mathbb{Z})$, $p \in [1, \infty]$, via multiplication by the transpose matrix $A^\top = (a_{ji})$ of (1.4)$.Putting$y:=(B-$\lambda I)x$ with $x \in \ell^\infty$ and arguing as in **1.** (note that now $|\alpha_{i-1}| + |\gamma_{i+1}| \le r_i$ is replaced by $|\alpha_i| + |\gamma_i| \le r_i$), one gets that also $B - \lambda I$ is bounded below by δ as a mapping $\ell^\infty \to \ell^\infty(\mathbb{Z})$. So again, $B - \lambda I$ is injective with closed range on $\ell^\infty(\mathbb{Z})$ and hence injective on $\ell^1(\mathbb{Z})$.

3. Via the duality $(u, v) := \sum_{i\in\mathbb{Z}} u_i v_i$ between $\ell^1(\mathbb{Z})$ and $\ell^\infty(\mathbb{Z})$, $A - \lambda I : \ell^\infty(\mathbb{Z}) \to \ell^\infty(\mathbb{Z})$ is the adjoint operator of $B - \lambda I : \ell^1(\mathbb{Z}) \to \ell^1(\mathbb{Z})$, and $B - \lambda I : \ell^\infty(\mathbb{Z}) \to \ell^\infty(\mathbb{Z})$ is the adjoint operator of $A - \lambda I : \ell^1(\mathbb{Z}) \to$

$\ell^1(\mathbb{Z})$. By **1.** and **2.** we conclude that both $A - \lambda I$ and $B - \lambda I$ are invertible (injective with closed and dense range) on $\ell^\infty(\mathbb{Z})$ and their inverses are bounded above by $1/\delta$. Consequently, their pre-adjoints $B - \lambda I$ and $A - \lambda I$ are invertible on $\ell^1(\mathbb{Z})$ with their inverses bounded above by $1/\delta$. By Riesz-Thorin interpolation (e.g. [32, section 1.5.11]), it follows that $A - \lambda I$ is invertible, with the inverse bounded by $1/\delta$, on all spaces $\ell^p(\mathbb{Z})$ with $p \in [1, \infty]$.

From $\|(A-\lambda I)^{-1}\|_{\ell^p(\mathbb{Z})\to\ell^p(\mathbb{Z})} \le 1/\delta$ for all $p \in [1, \infty]$, we imply $\operatorname{dist}(\lambda, \operatorname{spec} A) \ge \delta$. Recall that δ is, by its definition, the distance of λ from the right-hand side of (2.13). □

Theorem 2.50 can be generalised in different directions:

Firstly, it is clear that one can pass from two-sided infinite matrices $(a_{ij})_{i,j\in\mathbb{Z}}$ to one-sided infinite matrices $(a_{ij})_{i,j\in\mathbb{N}}$ by replacing $\mathbb{Z}$ with $\mathbb{N}$ in (2.13), where $r_2, r_3, \ldots$ are as defined in (2.12) but $r_1 := \max(|\alpha_1|, |\gamma_2|)$ (i.e. (2.12) with $i = 1$, $\alpha_0 := 0$ and $\gamma_1 := 0$). Of course one can also pass to finite matrices $(a_{ij})_{i,j=1}^n$ and thereby recover Gershgorin's theorem (in a somewhat weaker form than usual since, instead of (2.12), $r_i := |\alpha_{i-1}| + |\gamma_{i+1}|$ and $r_i := |\alpha_i| + |\gamma_i|$ are both already enough for (2.13) in the finite matrix case, see [50]).

Secondly, our proof shows that one can go away from the tridiagonal case (1.4) to infinite matrices with more than three (even infinitely many) nonzero diagonals as long as the diagonal suprema are summable. So if A is given by a matrix $(a_{ij})_{i,j\in\mathbb{I}}$ (with $\mathbb{I} = \mathbb{Z}$ or $\mathbb{N}$) such that

$$\sum_{k\in\mathbb{Z}} d_k \;<\; \infty \qquad \text{with} \qquad d_k \;:=\; \sup_{\substack{i,j\,\in\,\mathbb{I}\\ i-j\,=\,k}} |a_{ij}|, \quad k \in \mathbb{Z},$$

in which case we say that A belongs to the Wiener algebra, then (2.13) holds with

$$r_i \;:=\; \max\left(\sum_{j\in\mathbb{I}\setminus\{i\}} |a_{ij}|\,,\; \sum_{j\in\mathbb{I}\setminus\{i\}} |a_{ji}| \right) \;\le\; \sum_{k\in\mathbb{Z}\setminus\{0\}} d_k, \qquad i \in \mathbb{I}$$

instead of (2.12).

2.4 Limit Operator and Pseudoergodicity

Throughout this thesis, we need the concept of limit operators. For $p \in [1, \infty]$, let $A \in B(\ell^p(\mathbb{Z}))$ be a band operator, i.e., A is induced by a band matrix, say M. From the boundedness of A we get that every diagonal d_k of M is a bounded sequence of elements in $\ell^p(\mathbb{Z})$. We then put

$$\|A\|_{\mathcal{W}} := \sum_{k=-\infty}^{+\infty} \|d_k\|_\infty = \sum_{k=-\infty}^{+\infty} \sup_{j \in \mathbb{Z}} |a_{j+k,j}|$$

and denote by $\mathcal{W}$ the closure of the set of all band operators on $\ell^p(\mathbb{Z})$ in the norm $\|\cdot\|_{\mathcal{W}}$. The set $\mathcal{W}$, equipped with the norm $\|\cdot\|_{\mathcal{W}}$ turns out to be a Banach algebra and is called **the Wiener algebra**.

Definition 2.51. *Let $A \in \mathcal{W}$. We will call B a **limit operator** of A with respect to the sequence $h = (h_m)_{m \in \mathbb{N}} \subseteq \mathbb{Z}$ with $|h_m| \to \infty$ if, entrywise,*

$$[V_{-h_m} A V_{h_m}] \to [B] \qquad \text{as } m \to \infty.$$

The set of all limit operators of $A \in \mathcal{W}$ is denoted by $\sigma^{op}(A)$.

Proposition 2.52. *Let $A \in \mathcal{W}$. If A is a band operator then each limit operator of A is also a band operator.*

Proof. See Proposition 3.6 [33]. ■

The following proposition is a very useful statement.

Proposition 2.53. *Suppose $A \in \mathcal{W}$. Then the following statements are equivalent:*

(FC) All limit operators of A are injective on $\ell^\infty(\mathbb{Z})$.

(i) All limit operators of A are invertible on one of the spaces $\ell^p(\mathbb{Z})$ where $p \in [1, \infty]$.

(ii) All limit operators of A are invertible on the space $\ell^p(\mathbb{Z})$ for all $p \in [1, \infty]$ and

$$\sup_{p\in[1,\infty]} \sup_{B\in\sigma^{\mathsf{op}}(A)} \left\|B^{-1}\right\|_{B(\ell^p(\mathbb{Z}))} < \infty.$$

Further, on every $\ell^p(\mathbb{Z})$ space it holds that

$$\operatorname{spec}_{\text{ess}}(A) = \bigcup_{B\in\sigma^{\mathsf{op}}(A)} \operatorname{spec}(B) = \bigcup_{B\in\sigma^{\mathsf{op}}(A)} \operatorname{spec}^{\infty}_{\text{point}}(B).$$

Proof. See [8]. ∎

Definition 2.54. *Let D be a closed subset of $\mathbb{C}$. A bounded sequence $b = (b_i) \subseteq \ell^p(\mathbb{Z})$ is said to be* ***pseudo-ergodic with respect to*** *D if, for every finite set $S \subseteq \mathbb{Z}$, every function $c : S \to D$ and every $\varepsilon > 0$, there is a $\gamma \in \mathbb{Z}$ such that*

$$\sup_{\alpha\in S} |b_{\gamma+\alpha} - c_\alpha| < \varepsilon.$$

Moreover, we call the sequence b ***pseudo-ergodic*** *if it is pseudo-ergodic with respect to $D = \overline{\{b_i : i \in \mathbb{Z}\}}$.*

For example, a sequence $(b_i) \in \{\pm 1\}^{\mathbb{Z}}$ is pseudoergodic iff every finite pattern of ± 1's can be found somewhere in the sequence b. There is some connection between pseudoergodicity and limit operators, which is one of the key ingredients for studying spectral theory of random operators.

Proposition 2.55. *Suppose $b = (b_i) \in \ell^\infty(\mathbb{Z})$. Then b is* ***pseudo-ergodic*** *if and only if*

$$\sigma^{\mathsf{op}}(M_b) = \left\{M_c : c = (c_i) \subseteq \overline{\{b_i : i \in \mathbb{Z}\}}\right\}$$

Proof. see [33]. ∎

Chapter 3

Spectral and Pseudospectral Inclusion Sets for Infinite Tridiagonal Matrices

In this chapter, we aim to compute optimal upper bounds (inclusion sets) for the spectrum and pseudospectrum of the operator corresponding to the infinite tri-diagonal matrix

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & \beta_{-2} & \gamma_{-1} & & & & \\ & \alpha_{-2} & \beta_{-1} & \gamma_0 & & & \\ & & \alpha_{-1} & \boxed{\beta_0} & \gamma_1 & & \\ & & & \alpha_0 & \beta_1 & \gamma_1 & \\ & & & & \alpha_1 & \beta_2 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}, \tag{3.1}$$

where the box marks the matrix entry at $(0,0)$. Here, (α_i), (β_i), and (γ_i) are bounded sequences of complex numbers, and the operator acts by multiplication by the matrix A, i.e., if $x = (x_j)_{j\in\mathbb{Z}}$ then $y = Ax$ has ith entry given by

$$y_i = \alpha_{i-1}x_{i-1} + \beta_i x_i + \gamma_{i+1}x_{i+1}, \quad i \in \mathbb{Z}.$$

The operator A is a bounded linear operator on many Banach spaces, in particular on $\ell^p(\mathbb{Z})$ for $1 \le p \le \infty$.

There are two inclusion sets to be discussed in this chapter. For $n \in \mathbb{N}$ and $k \in \mathbb{Z}$, let $A_{n,k}$ and $\hat{A}_{n,k}$ denote the order n tri-diagonal matrices

$$A_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}$$

and

$$\hat{A}_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \alpha_k \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ \gamma_{k+n+1} & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}.$$

Define

$$\Sigma^n_\varepsilon(A) := \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_\varepsilon A_{n,k} \text{ and } \Pi^n_\varepsilon(A) := \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_\varepsilon \hat{A}_{n,k}.$$

We will compute upper bounds for the spectrum and pseudospectrum of A using the ordinary finite submatrices, $A_{n,k}$ and the periodised submatrices, $\hat{A}_{n,k}$, calling these method 1 and method 1*, respectively. For method 1 a main result that we will show is that, for $\varepsilon > 0$, $n \in \mathbb{N}$,

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)}, \tag{3.2}$$

where

$$f(n) = 2\sin\left(\frac{\theta}{2}\right)\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right),$$

and θ is the unique solution in the range $\left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$ of the equation

$$2\sin\left(\frac{t}{2}\right)\cos\left(\left(n+\frac{1}{2}\right)t\right) + \frac{\|\alpha\|\,\|\gamma\|}{(\|\alpha\|+\|\gamma\|)^2}\sin\left((n-1)\,t\right) = 0.$$

In particular, if $\|\alpha\|_\infty = 0$ or $\|\gamma\|_\infty = 0$ then (3.35) holds with

$$f(n) = 2\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)\sin\left(\frac{\pi}{4n+2}\right).$$

The corresponding result for method 1* is that, for $\varepsilon > 0$, $n \in \mathbb{N}$,

$$\operatorname{spec}_\epsilon A \subseteq \Pi^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Pi^n_{f(n)}(A)}, \tag{3.3}$$

where

$$f(n) = 2\sin\left(\frac{\pi}{2n}\right)\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right).$$

To prove the theorems in this chapter, we need the following inequality:

Lemma 3.1. *For $a, b \in \mathbb{R}$ and $\theta > 0$, we have the following inequality*

$$(a+b)^2 \leq a^2(1+\theta) + b^2(1+\theta^{-1}),$$

where equality holds iff $a\theta = b$.

Proof.

$$\begin{aligned}(a^2+b^2) &= (a^2+b^2) + (a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}})^2 - (a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}})^2 \\ &\leq (a^2+b^2) + (a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}})^2 \\ &= a^2(1+\theta) + b^2(1+\theta^{-1}) - 2ab.\end{aligned} \tag{3.4}$$

It follows that

$$(a+b)^2 \leq a^2(1+\theta) + b^2(1+\theta^{-1}).$$

Obviously, from (3.4), equality holds if and only if $(a\theta^{\frac{1}{2}} - b\theta^{-\frac{1}{2}}) = 0$ if and only if $a\theta = b$. ■

For $k \in \mathbb{Z}$ define $\chi^{(n,k)} \in \ell^\infty(\mathbb{Z})$ by

$$\chi_i^{(n,k)} = \begin{cases} 1 & \text{if } i = k+1, k+2, \ldots, k+n \\ 0 & \text{otherwise.}\end{cases}$$

3.1 Inclusion sets in terms of finite section matrices

In this section we will compute inclusion sets in terms of the finite section matrices

$$A_{n,k} := \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}, \tag{3.5}$$

for $n \in \mathbb{N}, k \in \mathbb{Z}$.

Theorem 3.2. *If $\varepsilon > 0$, $n \in \mathbb{N}$, $w_j > 0$, for $j = 1, \ldots, n$, and $w_0 = w_{n+1} = 0$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A),$$

where

$$f(n) = \|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}} + \|\gamma\|_\infty \sqrt{\frac{T_n^+}{S_n}},$$

$$S_n = \sum_{i=1}^n w_i^2,\ T_n^- = \sum_{i=1}^n (w_{i-1} - w_i)^2, \ \ \textit{and}\ \ T_n^+ = \sum_{i=1}^n (w_{i+1} - w_i)^2.$$

Proof. Let $\lambda \in \operatorname{spec}_\varepsilon A$. Then either there exists $x \in \ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A - \lambda I)x\| < \varepsilon$, or the same holds with A replaced by its adjoint. In the first case, let $y = (A - \lambda I)x$, so $\|y\| < \varepsilon$. For $k \in \mathbb{Z}$, define $e^{(k)} \in \ell^\infty(\mathbb{Z})$ by

$$e_i^{(k)} = \begin{cases} w_{i-k} & \text{if } i = k+1, k+2, \ldots, k+n, \\ 0 & \text{otherwise,} \end{cases}$$

and let

$$E_{i,k}^+ := \begin{cases} \left| e_{i+1}^{(k)} - e_i^{(k)} \right| & \text{if } i = k+1, \ldots, k+n, \\ 0 & \text{otherwise,} \end{cases} \tag{3.6}$$

$$E^-_{i,k} := \begin{cases} \left| e^{(k)}_{i-1} - e^{(k)}_i \right| & \text{if } i = k+1, \dots, k+n, \\ \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

Then, for $k \in \mathbb{Z}$,

$$\sum_{k \in \mathbb{Z}} (e^{(k)}_i)^2 = S_n. \tag{3.8}$$

Further, for $k \in \mathbb{Z}$,

$$\sum_{k \in \mathbb{Z}} (E^-_{i,k})^2 = \sum_{k=i-n}^{i-1} (e^{(k)}_{i-1} - e^{(k)}_i)^2 = T^-_n \tag{3.9}$$

and

$$\sum_{k \in \mathbb{Z}} (E^+_{i,k})^2 = \sum_{k=i-n}^{i-1} (e^{(k)}_{i+1} - e^{(k)}_i)^2 = T^+_n. \tag{3.10}$$

For $k \in \mathbb{Z}$, let

$$P_k := \left\| M_{\chi^{(n,k)}} (A - \lambda I) M_{\chi^{(n,k)}} M_{e^{(k)}} x \right\| = \left\| (A_{n,k} - \lambda I_n) \widetilde{x}_{n,k} \right\|,$$

where $\widetilde{x}_{n,k} := (w_1 x_{k+1}, w_2 x_{k+2}, \dots, w_n x_{k+n})^T$, and let $Q_k := \| M_{e^{(n,k)}} x \| = \| \widetilde{x}_{n,k} \|$. We will prove that $P_k < (\varepsilon + f(n)) Q_k$ for some $k \in \mathbb{Z}$, which will show that $\lambda \in \operatorname{spec}_{\varepsilon + f(n)} A_{n,k}$.

Note first that, using (3.6) and (3.7),

$$\begin{aligned} P_k^2 &= \sum_{i=k+1}^{k+n} \left| y_i e^{(k)}_i + \alpha_{i-1} (e^{(k)}_{i-1} - e^{(k)}_i) x_{i-1} + \gamma_{i+1} (e^{(k)}_{i+1} - e^{(k)}_i) x_{i+1} \right|^2 \\ &\leq \sum_{i=k+1}^{k+n} \left(|y_i| \, e^{(k)}_i + E^-_{i,k} \, |\alpha_{i-1} x_{i-1}| + E^+_{i,k} \, |\gamma_{i+1} x_{i+1}| \right)^2. \end{aligned}$$

So, for all $\theta > 0$ and $\phi > 0$, by Lemma 3.1,

$$
\begin{aligned}
P_k^2 &\le \sum_{i=k+1}^{k+n} \left[\left(|y_i| \, e_i^{(k)} \right)^2 (1+\theta) + (1+\theta^{-1}) \left(E_{i,k}^- \, |\alpha_{i-1} x_{i-1}| + E_{i,k}^+ \, |\gamma_{i+1} x_{i+1}| \right)^2 \right] \\
&\le \sum_{i=k+1}^{k+n} \left[\left(|y_i| \, e_i^{(k)} \right)^2 (1+\theta) \right. \\
&\quad \left. + (1+\theta^{-1}) \left((1+\phi) \left(E_{i,k}^- \right)^2 |\alpha_{i-1} x_{i-1}|^2 + (1+\phi^{-1}) \left(E_{i,k}^+ \right)^2 |\gamma_{i+1} x_{i+1}|^2 \right) \right] \\
&= \sum_{i \in \mathbb{Z}} \left[\left(|y_i| \, e_i^{(k)} \right)^2 (1+\theta) \right. \\
&\quad \left. + (1+\theta^{-1}) \left((1+\phi) \left(E_{i,k}^- \right)^2 |\alpha_{i-1} x_{i-1}|^2 + (1+\phi^{-1}) \left(E_{i,k}^+ \right)^2 |\gamma_{i+1} x_{i+1}|^2 \right) \right].
\end{aligned}
$$

Thus, and using (3.8), (3.9) and (3.10),

$$
\begin{aligned}
\sum_{k \in \mathbb{Z}} P_k^2 &\le (1+\theta) \sum_{i \in \mathbb{Z}} |y_i|^2 \sum_{k \in \mathbb{Z}} (e_i^{(k)})^2 + (1+\theta^{-1}) \left[(1+\phi) \, \|\alpha\|_\infty^2 \sum_{i \in \mathbb{Z}} |x_{i-1}|^2 \sum_{k \in \mathbb{Z}} (E_{i,k}^-)^2 \right. \\
&\qquad \left. + (1+\phi^{-1}) \, \|\gamma\|_\infty^2 \sum_{i \in \mathbb{Z}} |x_{i+1}|^2 \sum_{k \in \mathbb{Z}} (E_{i,k}^+)^2 \right] \\
&\qquad = (1+\theta) \, \|y\|^2 \, S_n + (1+\theta^{-1}) \left((1+\phi) \, \|\alpha\|_\infty^2 \, T_n^- + (1+\phi^{-1}) \, \|\gamma\|_\infty^2 \, T_n^+ \right).
\end{aligned}
$$

Similarly, $\sum_{k \in \mathbb{Z}} Q_k^2 = S_n \, \|x\|^2 = S_n$. Now, by Lemma 3.1,

$$
\inf_{\phi > 0} \left((1+\phi) \, \|\alpha\|_\infty^2 \, T_n^- + (1+\phi^{-1}) \, \|\gamma\|_\infty^2 \, T_n^+ \right) = \left(\|\alpha\|_\infty \sqrt{T_n^-} + \|\gamma\|_\infty \sqrt{T_n^+} \right)^2.
$$

Thus,

$$
\begin{aligned}
\sum_{k \in \mathbb{Z}} P_k^2 &\le \left[(1+\theta) \, \|y\|^2 + \frac{1}{S_n} (1+\theta^{-1}) \left(\|\alpha\|_\infty \sqrt{T_n^-} + \|\gamma\|_\infty \sqrt{T_n^+} \right)^2 \right] S_n \\
&= \left[(1+\theta) \, \|y\|^2 + (1+\theta^{-1}) \, [f(n)]^2 \right] \sum_{k \in \mathbb{Z}} Q_k^2.
\end{aligned}
$$

Applying Lemma 3.1 again, we see that

$$
\inf_{\theta > 0} \left[(1+\theta) \, \|y\|^2 + (1+\theta^{-1}) [f(n)]^2 \right] = (\|y\| + f(n))^2,
$$

so that

$$\sum_{k\in\mathbb{Z}} P_k^2 \le (\|y\| + f(n))^2 \sum_{k\in\mathbb{Z}} Q_k^2 < (\varepsilon + f(n))^2 \sum_{k\in\mathbb{Z}} Q_k^2.$$

Thus, for some $k \in \mathbb{Z}$,

$$P_k < (\varepsilon + f(n))\, Q_k,$$

so that $\lambda \in \operatorname{spec}_{\varepsilon+f(n)} A_{n,k}$.

In the case when there exists $x \in \ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A - \lambda I)^* x\| < \varepsilon$, the same argument shows that $\overline{\lambda} \in \operatorname{spec}_{\varepsilon+f(n)} A^*_{n,k}$, for some $k \in \mathbb{Z}$, so that $\lambda \in \operatorname{spec}_{\varepsilon+f(n)} A_{n,k}$. □

Corollary 3.3. $\operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)}$.

Proof. We can see that, if

$$\lambda \in \operatorname{spec} A = \bigcap_{\varepsilon>0} \operatorname{spec}_\varepsilon A \subseteq \bigcap_{\varepsilon>0} \left(\bigcup_{k\in\mathbb{Z}} \operatorname{spec}_{\varepsilon+f(n)} A_{n,k} \right),$$

so that for all $\varepsilon > 0$ there is a $k \in \mathbb{Z}$ with $\left\|(A_{n,k} - \lambda I_n)^{-1}\right\| > \dfrac{1}{\varepsilon + f(n)}$, then $s := \sup_{k\in\mathbb{Z}} \left\|(A_{n,k} - \lambda I_n)^{-1}\right\| \ge \dfrac{1}{f(n)}$. If $s > \dfrac{1}{f(n)}$ then there exists a $k \in \mathbb{Z}$ with $\lambda \in \operatorname{spec}_{f(n)} A_{n,k} \subseteq \Sigma^n_{f(n)}(A)$. If $s = \dfrac{1}{f(n)}$ then put $D :=$Diag$\{A_{n,k} : k \in \mathbb{Z}\}$, so that

$$\left\|(D - \lambda I)^{-1}\right\| = \sup_{k\in\mathbb{Z}} \left\|(A_{n,k} - \lambda I_n)^{-1}\right\| = s = \frac{1}{f(n)}.$$

Take $r > 0$ small enough that $\lambda + r\mathbb{D} \subseteq \rho(D) := \mathbb{C} \setminus \operatorname{spec} D$. By Theorem 2.32, there are $\mu_1, \mu_2, \ldots \in \rho(D)$ with $|\mu_m - \lambda| < \dfrac{r}{m}$ and $\left\|(D - \mu_m I)^{-1}\right\| > \left\|(D - \lambda I)^{-1}\right\| = \dfrac{1}{f(n)}$ for $m = 1, 2, \ldots$. Hence, λ is in the closure of $\operatorname{spec}_{f(n)} D = \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_{f(n)} A_{n,k} = \Sigma^n_{f(n)}(A)$. ■

Take $w_k = 1$, $k = 1, \ldots, n$. Then it is obvious that $S_n = n$ and $T_n^+ = T_n^- = 1$ in Theorem 3.2 and Theorem 3.3, giving the following corollary.

Corollary 3.4. *If* $\varepsilon > 0$, $n \in \mathbb{N}$, *then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = (\|\alpha\|_{\alpha} + \|\gamma\|_{\infty})\sqrt{\frac{1}{n}}.$$

Note that, with this "cut-off" truncation, our function $f(n) = O(n^{-\frac{1}{2}})$ as $n \to \infty$. Taking $w_k = 1 - \dfrac{2|\frac{n+1}{2} - k|}{n+1}$, $k = 1, \ldots, n$, a hat-function, we obtain the following corollary:

Corollary 3.5. *If* $\varepsilon > 0$, $n \in \mathbb{N}$, *then*

$$\operatorname{spec}_{\varepsilon} A \subseteq \Sigma^n_{\varepsilon + f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = \begin{cases} \dfrac{2\sqrt{3}(\|\alpha\|_{\infty} + \|\gamma\|_{\infty})}{\sqrt{(n+1)(n+2)}}, & \text{if } n \text{ is even,} \\[2ex] \dfrac{2\sqrt{3}(\|\alpha\|_{\infty} + \|\gamma\|_{\infty})}{\sqrt{n^2+2n+3}}, & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Straightforward computations yield

$$S_n = \begin{cases} \dfrac{n(n+2)}{3(n+1)} & \text{if } n \text{ is even,} \\ \dfrac{n^2+2n+3}{3(n+1)} & \text{if } n \text{ is odd,} \end{cases}$$

and

$$T_n^+ = T_n^- = \begin{cases} \dfrac{4n}{(n+1)^2} & \text{if } n \text{ is even,} \\ \dfrac{4}{(n+1)} & \text{if } n \text{ is odd.} \end{cases}$$

Thus the conclusion follows by Theorem 3.2. ■

We see, by comparing Corollary 3.4 and Corollary 3.5, that careful choice of the weights w_n can reduce the value of $f(n)$ significantly. This suggests as a new challenging problem how to choose $w_1, w_2, \ldots, w_n$ to minimise $f(n)$ in Theorem 3.2. In the following subsections we will solve this problem, computing the minimum $f(n)$ for:

- an operator A which has $\|\gamma\|_{\infty} = 0$

- an operator A which has $\|\gamma\|_\infty = \|\alpha\|_\infty$
- an arbitrary tridiagonal operator A.

3.1.1 The Optimal Bound For the Bi-diagonal Case

Let A be a matrix of the form (3.1) which has $\|\gamma\|_\infty = 0$. Then, in Theorem 3.2,

$$f(n) = \|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}},$$

and, from the definitions of S_n and T_n^-, we see that, where $w = (w_1, \ldots, w_n)^T$,

$$S_n = \|w\|_2^2$$

and

$$T_n^- = \|B_n w\|_2^2$$

where

$$B_n = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}.$$

Thus

$$B_n^T B_n = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & 1 & -1 & \\ & & & 1 & -1 \\ & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}.$$

We can notice that $B_n^T B_n$ is symmetric. Moreover, from Corollary 2.25 we know that all eigenvalues of $B_n^T B_n$ are non-negative real numbers. From

Proposition 2.48, we can see that

$$\lambda_{min}(B_n^T B_n) = (\nu(B_n))^2 = \left(\inf_{x\neq 0} \frac{\|B_n x\|_2}{\|x\|_2}\right)^2 = \left(\inf \sqrt{\frac{T_n^-}{S_n}}\right)^2 = \inf \frac{T_n^-}{S_n}. \tag{3.11}$$

where $\lambda_{min}(B_n^T B_n)$ denotes the smallest eigenvalue of $B_n^T B_n$.

Suppose that $w = (w_1, w_2, \ldots, w_n)^T \neq 0$, $\lambda \in \mathbb{C}$ and $B_n^T B_n w = \lambda w$. Thus, we have

$$-w_{j-1} + (2-\lambda)w_j - w_{j+1} = 0 \qquad \text{for } j = 1, \ldots, n-1, \tag{3.12}$$

and

$$-w_{n-1} + w_n = 0,$$

where $w_0 = 0$. Now, (3.12) holds iff

$$w_j = Dm_1^j + Em_2^j, \quad j = 0, 1, \ldots, n,$$

where m_1, m_2 are the roots of $m^2 - (2-\lambda)m + 1 = 0$, and D and E are constants. Note that, $m_1 m_2 = 1$ and $m_1 + m_2 = 2 - \lambda$.

Put $m_1 = e^{i\theta}$, for some $\theta \in \mathbb{C}$, giving

$$\lambda = 2 - (e^{i\theta} + e^{-i\theta}) = 2 - 2\cos\theta = 4\sin^2\left(\frac{\theta}{2}\right).$$

Since

$$w_j = D(e^{i\theta})^j + E(e^{-i\theta})^j,$$

it follows that

$$w_j = B\cos(j\theta) + C\sin(j\theta), j = 1, 2, \ldots, n,$$

for some constant B and C. Since $w_0 = 0$, we have $B = 0$ so that $w_j = C\sin(j\theta)$, with $C \neq 0$. We can see that θ is not a multiple of 2π because this would make $w_j = 0$ for all $j = 1, 2, \ldots, n$. Since $-w_{n-1} + (1-\lambda)w_n = 0$, we obtain

$$\cos\left(\left(n + \frac{1}{2}\right)\theta\right)\sin(\frac{\theta}{2}) = 0,$$

so that $\theta = \dfrac{(2r-1)\pi}{2n+1}$ where $r = 1, \ldots, n$. That means the smallest eigenvalue of $B_n^T B_n$ is $4\sin^2(\dfrac{\pi}{4n+2})$ with the corresponding eigenvector being

$$w_j = \sin\left(\frac{j\pi}{2n+1}\right), j = 1, \ldots, n.$$

Thus, from (3.11), we see that we have shown that

$$\inf_{w \in \mathbb{R}^n \neq 0} \sqrt{\frac{T_n^-}{S_n}} = \sqrt{\lambda_{min}(B_n^T B_n)} = 2\sin\left(\frac{\pi}{4n+2}\right).$$

Thus, in the bidiagonal case, Theorem 3.2 has the following corollary.

Corollary 3.6. *Let A be a matrix of the form (3.1) with $\|\gamma\|_\infty = 0$. If $\varepsilon > 0$, $n \in \mathbb{N}$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = 2\sin\left(\frac{\pi}{4n+2}\right)\|\alpha\|_\infty.$$

3.1.2 The Optimal Bound For the case $\|\alpha\|_\infty = \|\gamma\|_\infty$

Let A be a matrix of the form (3.1) which has $\|\alpha\|_\infty = \|\gamma\|_\infty$. From Theorem 3.2 we have that, for $n \in \mathbb{N}$,

$$f(n) = \|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}} + \|\gamma\|_\infty \sqrt{\frac{T_n^+}{S_n}},$$

where $S_n = \sum_{i=1}^n w_i^2$, $T_n^- = \sum_{i=1}^n (w_{i-1} - w_i)^2$, and $T_n^+ = \sum_{i=1}^n (w_{i+1} - w_i)^2$, where $w_0 = w_{n+1} = 0$. We will minimise $f(n)$ under the constraint, which seems appropriate for symmetry reasons when $\|\alpha\|_\infty = \|\gamma\|_\infty$, that $w_j = w_{n+1-j}$ for $j = 1 \ldots, n$, which implies that

$$T_n^+ = T_n^-. \tag{3.13}$$

Since $\|\alpha\|_\infty = \|\gamma\|_\infty$ and (3.13) holds, it follows that our $f(n)$ can be written of the form

$$f(n) = 2\,\|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}}.$$

We will consider the case when n is even and odd separately. Firstly, if $n = 2k$ for some $k \in \mathbb{N}$, it follows that

$$\begin{aligned}
\frac{T_n^-}{S_n} &= \frac{\displaystyle\sum_{i=1}^{n} (w_{i-1} - w_i)^2}{\displaystyle\sum_{i=1}^{n} w_i^2} \\
&= \frac{\frac{1}{2}w_1^2 + (w_2 - w_1)^2 + \cdots + (w_k - w_{k-1})^2}{w_1^2 + \cdots + w_k^2} \\
&= \frac{\|C_k w\|_2^2}{\|w\|_2^2}
\end{aligned}$$

where $w = (w_1, \ldots, w_k)^T$ and

$$C = \begin{pmatrix} \frac{1}{\sqrt{2}} & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{pmatrix}_{k\times k},$$

so that

$$CC^T = \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & & & \\ -\frac{1}{\sqrt{2}} & 2 & -1 & & \\ & -1 & 2 & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}_{k\times k}.$$

From Proposition 2.48, we have

$$\inf \sqrt{\frac{T_n^-}{S_n}} = \nu(C) = s_{min}(C) = \sqrt{\lambda_{min}(CC^T)}.$$

We know that λ is an eigenvalue of CC^T with eigenvector $v = \begin{pmatrix} v_k \\ \vdots \\ v_1 \end{pmatrix}$ iff

$$\begin{pmatrix} \frac{1}{2}-\lambda & -\frac{1}{\sqrt{2}} & & & \\ -\frac{1}{\sqrt{2}} & 2-\lambda & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2-\lambda & -1 \\ & & & -1 & 2-\lambda \end{pmatrix} \begin{pmatrix} v_k \\ v_{k-1} \\ \vdots \\ v_2 \\ v_1 \end{pmatrix} = 0. \tag{3.14}$$

From equation (3.14) we have that $(CC^T - \lambda I)v = 0$ iff

$$(\frac{1}{2}-\lambda)v_k - \frac{1}{\sqrt{2}}v_{k-1} = 0, \tag{3.15}$$

$$-\frac{1}{\sqrt{2}}v_k + (2-\lambda)v_{k-1} - v_{k-2} = 0, \tag{3.16}$$

$$-v_{j+1} + (2-\lambda)v_j - v_{j-1} = 0 \text{ for } j = 1, 2, \ldots, k-2, \tag{3.17}$$

where

$$v_0 := 0. \tag{3.18}$$

Equation (3.17) has general solutions

$$v_j = A\cos(j\theta) + B\sin(j\theta) \text{ for } j = 1, 2, \ldots, k-1$$

where A and B are constants and

$$\lambda = 2(1-\cos\theta) = 4\sin^2\left(\frac{\theta}{2}\right).$$

Since $v_0 = 0$, it follows that $A = 0$. Thus, taking $B = 1$,

$$v_j = \sin(j\theta),$$

for $j = 1, \ldots, k-1$. From equation (3.16), we can see that

$$\begin{aligned} & -\frac{1}{\sqrt{2}}v_k + (2-\lambda)v_{k-1} - v_{k-2} = 0 \\ \Leftrightarrow & -\frac{1}{\sqrt{2}}v_k + 2\cos\theta\sin[(k-1)\theta] - \sin[(k-2)\theta] = 0 \\ \Leftrightarrow & -\frac{1}{\sqrt{2}}v_k + \sin(k\theta) = 0. \end{aligned}$$

Thus $v_k = \sqrt{2}\sin(k\theta)$. Note that θ is not a multiple of 2π since the eigenvector v is not the zero vector. Using equation (3.15) we see that λ is an eigenvalue iff

$$
\begin{aligned}
&(\frac{1}{2}-\lambda)(\sqrt{2}\sin(k\theta)) - \frac{1}{\sqrt{2}}(\sin(k-1)\theta) = 0 \\
\Leftrightarrow\ &(1-2\lambda)\sin(k\theta) - \sin((k-1)\theta) = 0 \\
\Leftrightarrow\ &(4\cos\theta - 3)\sin(k\theta) - \sin((k-1)\theta) = 0 \\
\Leftrightarrow\ &3(\cos\theta - 1)\sin(k\theta) + \cos(k\theta)\sin\theta = 0 \\
\Leftrightarrow\ &-6\sin^2(\frac{\theta}{2})\sin(k\theta) + 2\cos(k\theta)\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) = 0.
\end{aligned}
$$

Since θ is not a multiple of 2π, this implies that

$$
\begin{aligned}
&-3\sin(\frac{\theta}{2})\sin(k\theta) + \cos(k\theta)\cos(\frac{\theta}{2}) = 0 \\
\Leftrightarrow\ &2\cos((k+\frac{1}{2})\theta) - \cos((k-\frac{1}{2})\theta) = 0.
\end{aligned}
$$

It follows that

$$\lambda_{min}(C_k C_k^T) = 4\sin^2(\frac{\theta^*}{2}),$$

where θ^* is the smallest positive solution of $F(\theta) = 0$, and

$$F(\theta) = 2\cos((k+\frac{1}{2})\theta) - \cos((k-\frac{1}{2})\theta).$$

To locate the solution θ^* of equation $F(\theta) = 0$, it is helpful to show that there exists $a \in \mathbb{R}^+$ such that $\theta^* \in (0, a)$ and F is also a monotonic function on $(0, a)$. Note that

$$F(0) = 1 > 0 \text{ and } F\left(\frac{\pi}{2(k+\frac{1}{2})}\right) = -\cos\left(\frac{(k-\frac{1}{2})}{(k+\frac{1}{2})}\frac{\pi}{2}\right) = -\sin\left(\frac{\pi}{(2k+1)}\right) < 0. \tag{3.19}$$

Further,

$$
\begin{aligned}
F'(\theta) &= -\left[(2k+1)\sin((k+\frac{1}{2})\theta) - (k-\frac{1}{2}\sin((k-\frac{1}{2})\theta)\right] \\
&= -\left[(k+\frac{3}{2})\sin((k+\frac{1}{2})\theta) + (k-\frac{1}{2})\left(\sin((k+\frac{1}{2})\theta) - \sin((k-\frac{1}{2})\theta)\right)\right].
\end{aligned}
$$

From the fact that the function $\sin x$ is an increasing function for $x \in (0, \dfrac{\pi}{2})$, it follows that $\sin((k+\frac{\pi}{2})\theta) - \sin((k-\frac{1}{2})\theta) > 0$. Therefore $F'(\theta) < 0$, i.e., F is a monotonic function, on the interval $\left(0, \dfrac{\pi}{2(k+\frac{1}{2})}\right)$. Therefore, θ^* is the unique solution of $F(\theta) = 0$ in the interval $\left(0, \dfrac{\pi}{2k+1}\right)$. Moreover, we can also notice that

$$\begin{aligned} F\left(\frac{\pi}{2k+3}\right) &= 2\cos\left(\frac{\pi}{2} - \frac{\pi}{2k+3}\right) - \cos\left(\frac{\pi}{2} - \frac{2\pi}{2k+3}\right) \\ &= 2\sin\left(\frac{\pi}{2k+3}\right) - \sin\left(\frac{2\pi}{2k+3}\right) \\ &= 2\sin\left(\frac{\pi}{2k+3}\right) - 2\sin\left(\frac{\pi}{2k+3}\right)\cos\left(\frac{\pi}{2k+3}\right) > 0. \end{aligned}$$

Hence and from (3.19), we know that $\theta^* \in \left(\dfrac{\pi}{2k+3}, \dfrac{\pi}{2k+1}\right)$.

Now we are now considering the second case when $n = 2k+1$ for some $k \in \mathbb{N}$. We have then that

$$\begin{aligned} \frac{T_n^-}{S_n} &= \frac{\frac{1}{2}w_1^2 + (w_1 - w_2)^2 + \cdots + (w_k - w_{k+1})^2}{w_1^2 + \cdots + w_k^2 + \frac{1}{2}w_{k+1}^2} \\ &= \frac{\|Dw\|_2^2}{\|w\|_2^2} \end{aligned}$$

where $w = (w_1, \ldots, w_k, \frac{1}{\sqrt{2}}w_{k+1})^T$ and

$$D = \begin{pmatrix} \frac{1}{\sqrt{2}} & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & \sqrt{2} \end{pmatrix}_{(k+1)\times(k+1)},$$

so that

$$DD^T = \begin{pmatrix} \frac{1}{2} & -\frac{1}{\sqrt{2}} & & & \\ -\frac{1}{\sqrt{2}} & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 3 \end{pmatrix}.$$

We know that λ is an eigenvalue of DD^T with eigenvector $v = \begin{pmatrix} v_{k+1} \\ \vdots \\ v_1 \end{pmatrix}$ iff

$$\begin{pmatrix} \frac{1}{2} - \lambda & -\frac{1}{\sqrt{2}} & & & \\ -\frac{1}{\sqrt{2}} & 2-\lambda & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2-\lambda & -1 \\ & & & -1 & 3-\lambda \end{pmatrix} \begin{pmatrix} v_{k+1} \\ v_k \\ \vdots \\ v_2 \\ v_1 \end{pmatrix} = 0. \tag{3.20}$$

From equation (3.14), we have that $(DD^T - \lambda I)v = 0$ iff

$$(\frac{1}{2} - \lambda)v_{k+1} - \frac{1}{\sqrt{2}} v_k = 0, \tag{3.21}$$

$$-\frac{1}{\sqrt{2}} v_{k+1} - (2-\lambda)v_k - v_{k-1} = 0, \tag{3.22}$$

$$-v_{j+1} + (2-\lambda)v_j - v_{j-1} = 0 \text{ for } j = 2, \ldots, k-1, \tag{3.23}$$

and

$$-v_2 + (3-\lambda)v_1 = 0. \tag{3.24}$$

Equation (3.23) has the general solution

$$v_j = A\cos((j-1)\theta) + B\sin((j-1)\theta), \qquad j = 2, \ldots, k,$$

where A and B are constants and

$$\lambda = 2(1-\cos\theta) = 4\sin^2\left(\frac{\theta}{2}\right).$$

Obviously, we have $v_1 = A$. From equation (3.24), we have

$$\begin{aligned} & -v_2 + (3-\lambda)v_1 = 0 \\ & \Leftrightarrow -[A\cos\theta + B\sin\theta] + [1 + 2\cos\theta]A = 0 \\ & \Leftrightarrow B = \frac{(1+\cos\theta)\,A}{\sin\theta}. \end{aligned}$$

Thus, we can see that the constant B depends on A. We choose $A = 1$ then $B = \dfrac{1+\cos\theta}{\sin\theta}$. We have then that

$$\begin{aligned} v_j &= \cos((j-1)\theta) + \left(\frac{1+\cos\theta}{\sin\theta}\right)\sin((j-1)\theta) \\ &= \frac{\cos((j-1)\theta)\sin\theta}{\sin\theta} + \frac{\sin((j-1)\theta)}{\sin\theta} + \frac{\cos(\theta)\sin((j-1)\theta)}{\sin\theta} \\ &= \frac{\sin\theta\cos((j-1)\theta) + \cos\theta\sin((j-1)\theta)}{\sin\theta} + \frac{\sin((j-1)\theta)}{\sin\theta} \\ &= \frac{\sin(j\theta) + \sin((j-1)\theta)}{\sin\theta} \\ &= \frac{2\sin((j-\frac{1}{2})\theta)\cos(\frac{\theta}{2})}{2\sin(\frac{\theta}{2})\cos(\frac{\theta}{2})}, \end{aligned}$$

for $j = 1, 2, \ldots, k$ From equation (3.21), it follows that

$$v_j = \frac{\sin((j-\frac{1}{2})\theta)}{\sin(\frac{\theta}{2})}, \qquad j = 1, \ldots, k,$$

and

$$v_{k+1} = \frac{\sqrt{2}\sin((k+\frac{1}{2})\theta)}{\sin(\frac{\theta}{2})}.$$

Note that θ is not a multiple of 2π since the eigenvector v is not the zero vector. Using equation (3.21),

$$\begin{aligned} &(\frac{1}{2} - \lambda)v_{k+1} - \frac{1}{\sqrt{2}}v_k = 0 \\ \Leftrightarrow\ &(\frac{1}{2} - 2 + 2\cos\theta)\frac{\sqrt{2}\sin((k+\frac{1}{2})\theta)}{\sin(\frac{\theta}{2})} - \frac{\sin((k-\frac{1}{2})\theta)}{\sqrt{2}\sin(\frac{\theta}{2})} = 0. \end{aligned}$$

Multiplying both sides by $\sqrt{2}\sin(\frac{\theta}{2})$, we see that

$$\begin{aligned} &\Leftrightarrow (4\cos\theta - 3)\sin((k+\frac{1}{2})\theta) - \sin((k-\frac{1}{2})\theta) = 0 \\ &\Leftrightarrow 3(\cos\theta - 1)\sin((k+\frac{1}{2})\theta) + \cos((k+\frac{1}{2})\theta)\sin\theta = 0 \\ &\Leftrightarrow -6\sin^2(\frac{\theta}{2})\sin((k+\frac{1}{2})\theta) + 2\cos((k+\frac{1}{2})\theta)\sin(\frac{\theta}{2})\cos(\frac{\theta}{2}) = 0. \end{aligned}$$

Since θ is not a multiple of 2π, this implies that

$$\begin{aligned}&-3\sin(\frac{\theta}{2})\sin((k+\frac{1}{2})\theta)+\cos((k+\frac{1}{2})\theta)\cos(\frac{\theta}{2})=0\\&\Leftrightarrow 2\cos((k+1)\theta)-\cos(k\theta)=0.\end{aligned}$$

Let

$$F(\theta)=2\cos((k+1)\theta)-\cos(k\theta).$$

Similarly to the first case, we will show that there exists a unique solution $\theta^* \in \left(0, \frac{\pi}{2(k+1)}\right)$ to the equation $F(\theta)=0$. We can see that

$$F(0)=1>0 \;\text{ and } F\left(\frac{\pi}{2(k+1)}\right)=-\cos\left(\frac{k\pi}{2(k+1)}\right)=-\sin\left(\frac{\pi}{2(k+1)}\right)<0. \tag{3.25}$$

Further,

$$\begin{aligned}F'(\theta)&=-\left[2(k+1)\sin((k+1)\theta)-(k\sin(k\theta)\right]\\&=-\left[(k+1)\sin((k+1)\theta)+k(\sin((k+1)\theta)-\sin(k\theta))\right].\end{aligned}$$

From the fact that the function $\sin x$ is an increasing function when $x \in (0, \frac{\pi}{2})$, it follows that $\sin((k+1)\theta)-\sin(k\theta)>0$ for every θ on the interval $\left(0, \frac{\pi}{2(k+1)}\right)$. Therefore $F'(\theta)<0$, i.e. F is a monotonic function on the interval $\left(0, \frac{\pi}{2(k+1)}\right)$.

Therefore θ^* is the smallest positive solution of $2\cos((k+1)\theta)-\cos(k\theta)=0$ which is the unique solution in the interval $(0, \frac{\pi}{2(k+1)})$. Moreover, we can also see that

$$\begin{aligned}F(\frac{\pi}{2(k+2)})&=2\cos\left(\left(\frac{(k+1)\pi}{2k+4}\right)\right)-\cos\left(\frac{k\pi}{2k+4}\right)\\&=2\cos\left(\left(\frac{k+1}{k+2}\cdot\frac{\pi}{2}\right)\right)-\cos\left(\frac{k}{k+2}\cdot\frac{\pi}{2}\right)\\&=2\cos\left(\frac{\pi}{2}-\frac{\pi}{k+2}\right)-\cos\left(\frac{\pi}{2}-\frac{2\pi}{k+2}\right)\\&=2\sin\left(\frac{\pi}{k+2}\right)-\sin\left(\frac{2\pi}{k+2}\right)\\&>0.\end{aligned}$$

Hence, from equation (3.25) and $F(\frac{\pi}{2(k+2)}) > 0$, we know that $\theta^* \in \left(\frac{\pi}{2(k+2)}, \frac{\pi}{2(k+1)}\right)$. Hence, $\lambda_{\min}(DD^T) = 4\sin^2\left(\frac{\theta^*}{2}\right)$.

As a consequence, we can see that, from Theorem 3.2, for $\varepsilon > 0$, $n \in \mathbb{N}$,

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = \begin{cases} 2(\|\alpha\|_\infty + \|\gamma\|_\infty)\sin(\frac{\theta_1}{2}) & \text{if } n = 2m+1 \\ \\ 2(\|\alpha\|_\infty + \|\gamma\|_\infty)\sin(\frac{\theta_2}{2}) & \text{if } n = 2m \end{cases}$$

where θ_1 is the unique solution, in the interval $\left(\frac{\pi}{2(m+2)}, \frac{\pi}{2(m+1)}\right)$, of the equation

$$2\cos((m+1)\theta) - \cos(m\theta) = 0,$$

and θ_2 is the unique solution, in the interval $\left(\frac{\pi}{2m+3}, \frac{\pi}{2m+1}\right)$, of the equation

$$2\cos((m+\frac{1}{2})\theta) - \cos((m-\frac{1}{2})\theta) = 0.$$

We can rewrite these results more neatly as the following corollary:

Corollary 3.7. *If $\varepsilon > 0$, $n \in \mathbb{N}$ and $\|\alpha\|_\infty = \|\gamma\|_\infty$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = 2(\|\alpha\|_\infty + \|\gamma\|_\infty)\sin(\frac{\theta}{2})$$

where θ is the unique solution, in the interval $\left(\frac{\pi}{n+3}, \frac{\pi}{n+1}\right)$, of the equation

$$2\cos((\frac{n+1}{2})\theta) - \cos((\frac{n-1}{2})\theta) = 0.$$

3.1.3 The optimal bound for an arbitrary Tridiagonal matrix.

Let A be a matrix of the form (3.1). From Theorem 3.2 we have that, for $n \in \mathbb{N}$,

$$f(n) = \|\alpha\|_\infty \sqrt{\frac{T_n^-}{S_n}} + \|\gamma\|_\infty \sqrt{\frac{T_n^+}{S_n}},$$

where $S_n = \sum_{i=1}^{n} w_i^2$, $T_n^- = \sum_{i=1}^{n} (w_{i-1} - w_i)^2$, and $T_n^+ = \sum_{i=1}^{n} (w_{i+1} - w_i)^2$. Let $r = \|\alpha\|_\infty$ and $s = \|\gamma\|_\infty$. By lemma 3.1 we have then that,

$$(f(n))^2 = \min_{\theta > 0} \left[r^2(1+\theta)\frac{T_n^-}{S_n} + s^2(1+\theta^{-1})\frac{T_n^+}{S_n} \right].$$

Let

$$B_n = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \end{pmatrix}.$$

Note that $B_n^* = B_n^T$, since B_n is real. We have

$$T_n^- = \|B_n w\|^2 = (B_n w)^* B_n w = w^* B_n^* B_n w$$

while

$$T_n^+ = \left\|B_n^T w\right\|^2 = w^* B_n B_n^* w.$$

So

$$r^2(1+\theta)\frac{T_n^-}{S_n} + s^2(1+\theta^{-1})\frac{T_n^+}{S_n} = \frac{w^* D_n w}{w^* w} \tag{3.26}$$

where $D_n = r^2(1+\theta)B_n^* B_n + s^2(1+\theta^{-1})B_n B_n^*$. It is easy to see that D_n is symmetric, real and positive definite. In fact,

$$w^* D_n w \geq \left[r^2(1+\theta) + s^2(1+\theta^{-1})\right] \lambda_1$$

where λ_1 is the smallest eigenvalue of $B_n^* B_n$ and $B_n B_n^*$.

For $0 \leq \phi \leq 1$, let

$$E_n(\phi) = \phi B_n^* B_n + (1-\phi) B_n B_n^*$$

so $E_n(\phi)$ is the matrix

$$E_n(\phi) = \begin{pmatrix} 1+\phi & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2-\phi \end{pmatrix}.$$

Let $\mu(\phi) = \lambda_{min}(E_n(\phi))$, the smallest eigenvalue of a self-adjoint matrix $E_n(\phi)$. From equation (3.26), we have

$$\begin{aligned} D_n &= \left(r^2(1+\theta) + s^2(1+\theta^{-1})\right) \left(\frac{r^2(1+\theta)}{r^2(1+\theta)+s^2(1+\theta^{-1})} B_n^* B_n \right. \\ &\quad \left. + \frac{s^2(1+\theta^{-1})}{r^2(1+\theta)+s^2(1+\theta^{-1})} B_n B_n^* \right) \\ &= \left(r^2(1+\theta) + s^2(1+\theta^{-1})\right) E_n \left(\frac{r^2(1+\theta)}{r^2(1+\theta)+s^2(1+\theta^{-1})} \right). \end{aligned} \tag{3.27}$$

We know that λ is an eigenvalue of $E_n(\phi)$ with an eigenvector $v = (v_n, \ldots, v_1)^T$ iff

$$\begin{pmatrix} 1+\phi & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2-\phi \end{pmatrix} \begin{pmatrix} v_n \\ v_{n-1} \\ \vdots \\ v_2 \\ v_1 \end{pmatrix} = 0. \tag{3.28}$$

From equation (3.28),we have that $(E_n(\phi) - \lambda)v = 0$ iff

$$[(1+\phi) - \lambda]\, v_n - v_{n-1} = 0 \tag{3.29}$$

$$-v_{j+1} + (2-\lambda)v_j - v_{j-1} = 0, \text{ for } j = 2, \ldots, n-1, \tag{3.30}$$

and

$$-v_2 + [(2-\phi) - \lambda]\, v_1 = 0. \tag{3.31}$$

Equation (3.30) implies that

$$v_j = A\cos((j-1)\theta) + B\sin((j-1)\theta), j = 1, \ldots, n, \tag{3.32}$$

for some constants A and B where $\lambda = 2 - 2\cos\theta = 4\sin^2\left(\frac{\theta}{2}\right)$. In equation (3.32), arbitrarily set $A = \sin\theta$, so that $v_1 = \sin\theta$. Then equation (3.31) is satisfied if and only if

$$-\sin\theta\cos\theta - B\sin\theta + (2\cos\theta - \phi)\sin\theta = 0$$

i.e. if and only if

$$B = \cos\theta - \phi.$$

Therefore if v is an eigenvector of $E_n(\phi)$, then, to within multiplication by a constant,

$$\begin{aligned} v_j &= \sin\theta\cos[(j-1)\theta] + (\cos\theta - \phi)\sin[(j-1)\theta] \\ &= \sin(j\theta) - \phi\sin((j-1)\theta), j = 1, \ldots, n. \end{aligned} \tag{3.33}$$

Substituting v_j in equation (3.33) into equation (3.29), we see that λ is an eigenvalue of $E_n(\phi)$ iff $\lambda = 4\sin^2\left(\frac{\theta}{2}\right)$ with $\theta \neq 0$ and

$$F(\theta) = 0,$$

where $F(t) = (2\cos t - 1)\sin(nt) - \sin[(n-1)t] + \phi(1-\phi)\sin[(n-1)t], t \geq 0$. Note that

$$\begin{aligned} F(t) &= 2\cos t\sin(nt) - 2\cos\left(\frac{t}{2}\right)\sin\left(\left(n - \frac{1}{2}\right)t\right) + \phi(1-\phi)\sin((n-1)t) \\ &= \cos t\sin(nt) - \sin(nt) + \sin t\cos(nt) + \phi(1-\phi)\sin((n-1)t) \\ &= \sin((n+1)t) - \sin(nt) + \phi(1-\phi)\sin((n-1)t) \\ &= 2\sin\left(\frac{t}{2}\right)\cos\left(\left(n + \frac{1}{2}\right)t\right) + \phi(1-\phi)\sin((n-1)t) \end{aligned} \tag{3.34}$$

Clearly,

$$\mu(\phi) = 4\sin^2\left(\frac{\theta}{2}\right)$$

where θ is the smallest positive solution of the equation $F(t) = 0$. It is obvious to see that $\mu(\phi) = \mu(1-\phi)$. Now

$$\begin{aligned}
F\left(\frac{\pi}{n+1}\right) &= 2\sin\left(\frac{\pi}{2n+2}\right)\cos\left(\frac{(2n+1)\pi}{2n+2}\right) + \phi(1-\phi)\sin\left(\frac{(n-1)\pi}{n+1}\right)\\
&= 2\sin\left(\frac{\pi}{2n+2}\right)\cos\left(\pi - \frac{\pi}{2n+2}\right) + \phi(1-\phi)\sin\left(\pi - \frac{2\pi}{n+1}\right)\\
&= -2\sin\left(\frac{\pi}{2n+2}\right)\cos\left(\frac{\pi}{2n+2}\right) + \phi(1-\phi)\sin\left(\frac{2\pi}{n+1}\right)\\
&= -\sin\left(\frac{\pi}{n+1}\right) + \phi(1-\phi)\sin\left(\frac{2\pi}{n+1}\right)\\
&< -\sin\left(\frac{\pi}{n+1}\right) + 2\phi(1-\phi)\sin\left(\frac{\pi}{n+1}\right)\\
&= \sin\left(\frac{\pi}{n+1}\right)(2\phi(1-\phi)-1) < 0
\end{aligned}$$

and

$$\begin{aligned}
F\left(\frac{\pi}{2n+1}\right) &= 2\sin\left(\frac{\pi}{2(2n+1)}\right)\cos\left(\frac{(2n+1)\pi}{2(2n+1)}\right) + \phi(1-\phi)\sin\left(\frac{(n-1)\pi}{2n+1}\right)\\
&= 2\sin\left(\frac{\pi}{2(2n+1)}\right)\cos\left(\frac{\pi}{2}\right) + \phi(1-\phi)\sin\left(\frac{(n-1)\pi}{2n+1}\right) > 0
\end{aligned}$$

Also, from (3.34),

$$\begin{aligned}
F'(t) &= -(2n+1)\sin\left(\frac{t}{2}\right)\sin\left[\left(n+\frac{1}{2}\right)t\right] + \cos\left(\frac{t}{2}\right)\cos\left[\left(n+\frac{1}{2}\right)t\right]\\
&\quad + \phi(1-\phi)(n-1)\cos((n-1)t)\\
&= -(2n+1)\sin\left(\frac{t}{2}\right)\sin\left[\left(n+\frac{1}{2}\right)t\right] + \cos\left(\frac{t}{2}\right)\cos\left[\left(n+\frac{1}{2}\right)t\right]\\
&\quad + \phi(1-\phi)(n-1)\left\{\cos\left[\left(n+\frac{1}{2}\right)t\right]\cos\left(\frac{3t}{2}\right) + \sin\left[\left(n+\frac{1}{2}\right)t\right]\sin\left(\frac{3t}{2}\right)\right\}\\
&= A(t)\cos\left[\left(n+\frac{1}{2}\right)t\right] + B(t)\sin\left[\left(n+\frac{1}{2}\right)t\right]
\end{aligned}$$

where

$$A(t) = \cos\left(\frac{t}{2}\right) + \phi(1-\phi)(n-1)\cos\left(\frac{3t}{2}\right),$$

and

$$B(t) = \phi(1-\phi)(n-1)\sin\left(\frac{3t}{2}\right) - (2n+1)\sin\left(\frac{t}{2}\right).$$

Since $t \in \left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$, it follows that

$$\frac{\pi}{2} \le \frac{(2n+1)t}{2} \le \frac{\pi}{2}\left(\frac{2n+1}{n+1}\right) < \pi.$$

Therefore,

$$\cos\left[\left(n+\frac{1}{2}\right)t\right] < 0 \text{ and } \sin\left[\left(n+\frac{1}{2}\right)t\right] > 0.$$

In order to show that $F'(t) < 0$, , it suffices to prove that $A(t) > 0$ and $B(t) < 0$. Note that if $t \in \left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$ then $\frac{3t}{2} \in \left[\frac{3\pi}{2(2n+1)}, \frac{3\pi}{2(n+1)}\right)$. Thus, $\cos\left(\frac{3t}{2}\right) > 0$ and, hence, $A(t) > 0$. We can see that

$$\begin{aligned}
B(t) &= \phi(1-\phi)(n-1)\left[\sin t\cos\left(\frac{t}{2}\right) + \sin\left(\frac{t}{2}\right)\cos t\right] - (2n+1)\sin\left(\frac{t}{2}\right) \\
&= \sin\left(\frac{t}{2}\right)\left[\phi(1-\phi)(n-1)\cos t - (2n+1)\right] + \phi(1-\phi)(n-1)2\sin\left(\frac{t}{2}\right)\cos^2\left(\frac{t}{2}\right) \\
&\le \sin\left(\frac{t}{2}\right)\left\{\frac{1}{4}(n-1)\cos t - (2n+1) + \frac{1}{2}\cos^2\left(\frac{t}{2}\right)\right\} < 0,
\end{aligned}$$

since $\cos t \le 1$. Thus, $F'(t) < 0$ for $t \in \left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$. So θ, defined initially as the smallest positive solution of $F(t) = 0$ in $\left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$, is the unique solution in this interval.

To show that $\theta \in \left[\frac{\pi}{2n+1}, \theta^*\right]$ where $\theta^* < \frac{\pi}{n+1}$ is the smallest positive solution of $F(t) = 0$ in the case $\phi = \frac{1}{2}$, it is equivalent to prove (since $\frac{\pi}{2n+1}$ is the solution of $F(t) = 0$ in the case $\phi = 0$) that

$$\mu(0) \le \mu(\phi) \le \mu\left(\frac{1}{2}\right).$$

Since $E_n(\phi) = \phi E_n(1) + (1-\phi)E_n(0)$, it follows that

$$\begin{aligned}\mu(\phi) &= \min_{\|w\|=1} \left(\phi w^* B_n^* B_n w + (1-\phi)\phi w^* B_n B_n^* w\right) \\ &\geq \phi\mu(1) + (1-\phi)\mu(0) = \mu(0) \text{ since } \mu(1) = \mu(0).\end{aligned}$$

Also $2E_n(\frac{1}{2}) = E_n(\phi) + E_n(1-\phi)$, so that

$$\begin{aligned}\mu(\frac{1}{2}) &= \frac{1}{2}\left[\min_{\|w\|=1} \left(\phi w^* E_n(\phi) w + w^* E_n(1-\phi) w\right)\right] \\ &\geq \frac{1}{2}\left[\mu(\phi) + \mu((1-\phi)) = \mu(\phi)\right] \text{ since } \mu(\phi) = \mu(1-\phi).\end{aligned}$$

From equation (3.27), we have

$$\begin{aligned}(f(n))^2 &= \min_{\tau>0} \lambda_{min}(D_n) \\ &= \min_{\tau>0} \left(r^2(1+\tau) + s^2(1+\tau^{-1})\right) \mu\left(\frac{r^2(1+\tau)}{r^2(1+\tau) + s^2(1+\tau^{-1})}\right).\end{aligned}$$

Let $\phi = \dfrac{r^2(1+\tau)}{r^2(1+\tau) + s^2(1+\tau^{-1})}$, then $\tau = \dfrac{s^2\phi}{r^2(1-\phi)}$, so that

$$(f(n))^2 = \inf_{0<\phi<1} \left\{\left(\frac{r^2}{\phi} + \frac{s^2}{1-\phi}\right)\mu(\phi)\right\}.$$

Theorem 3.8. *For $\varepsilon > 0$, $n \in \mathbb{N}$,*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)}, \tag{3.35}$$

where

$$f(n) = 2 \inf_{0<\phi<1} \left(\sqrt{\frac{\|\alpha\|^2}{\phi} + \frac{\|\gamma\|^2}{1-\phi}} \sin\left(\frac{\theta}{2}\right)\right), \tag{3.36}$$

and θ is the unique solution in the range $\left(\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$ of the equation

$$2\sin\left(\frac{t}{2}\right)\cos\left(\left(n+\frac{1}{2}\right)t\right) + \phi(1-\phi)\sin\left((n-1)t\right) = 0.$$

Taking $\phi = \dfrac{r}{r+s}$ in equation (3.36), so that $\tau := \dfrac{s^2\phi}{r^2(1-\phi)} = \dfrac{r}{s}$, we get

Corollary 3.9. *If $\varepsilon > 0$, $n \in \mathbb{N}$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = 2\sin\left(\frac{\theta}{2}\right)\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right),$$

and θ is the unique solution in the range $\left[\dfrac{\pi}{2n+1}, \dfrac{\pi}{n+1}\right)$ of the equation

$$2\sin\left(\frac{t}{2}\right)\cos\left(\left(n+\frac{1}{2}\right)t\right) + \frac{\|\alpha\|\,\|\gamma\|}{(\|\alpha\|+\|\gamma\|)^2}\sin\left((n-1)\,t\right) = 0.$$

In particular, if $\|\alpha\|_\infty = 0$ or $\|\gamma\|_\infty = 0$ then (3.35) holds with

$$f(n) = 2\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)\sin\left(\frac{\pi}{4n+2}\right).$$

Taking $\phi = \dfrac{1}{2}$ in equation (3.36), then $\tau = \dfrac{s^2\phi}{r^2(1-\phi)} = \dfrac{s^2}{r^2}$, and we get

Corollary 3.10. *If $\varepsilon > 0$, $n \in \mathbb{N}$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

where

$$f(n) = 2\sqrt{2}\sin\left(\frac{\theta}{2}\right)\sqrt{\|\alpha\|^2_\infty + \|\gamma\|^2_\infty},$$

and θ is the unique solution in the range $\left(\dfrac{\pi}{n+1} - \dfrac{2\pi}{(n+1)(n+3)}, \dfrac{\pi}{n+1}\right)$ of the equation

$$2\cos\left(\left(\frac{n+1}{2}\right)t\right) - \cos\left(\left(\frac{n-1}{2}\right)t\right) = 0. \tag{3.37}$$

Proof. We need to show that

$$2\sin\left(\frac{t}{2}\right)\cos\left(\left(n+\frac{1}{2}\right)t\right) + \frac{1}{4}\sin\left((n-1)\,t\right) = 0. \tag{3.38}$$

and (3.37) are equivalent. Since

$$2\cos\left(\left(\frac{n+1}{2}\right)t\right)-\cos\left(\left(\frac{n-1}{2}\right)t\right)=0,$$

multiplying both sides by $2\sin\frac{(n+1)t}{2}-\sin\frac{(n-1)t}{2}$, we have

$$\sin\left((n+1)t\right)-\sin(nt)+\frac{1}{4}\sin\left((n-1)t\right)=0,$$

which is equivalent to (3.38).

□

3.2 Numerical Examples for Method 1

3.2.1 Shift Operator

Corollary 3.3 and Corollary 3.6 tell us that in the bidiagonal case, $\|\gamma\|_\infty = 0$,

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)}. \tag{3.39}$$

for all $\varepsilon > 0$ and $n \in \mathbb{N}$, where

$$f(n) = 2\sin\left(\frac{\pi}{4n+2}\right)\|\alpha\|_\infty. \tag{3.40}$$

We will look at the simplest example of this type, the so-called right shift operator

$$A = V_1 := \begin{pmatrix} \ddots & \ddots & & & \\ \ddots & 0 & 0 & & \\ & 1 & 0 & 0 & \\ & & 1 & 0 & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \tag{3.41}$$

and check how sharp (or how generous) the inclusions (3.39) are in this case. In other words, we ask by what function $f_*(n)$ one could replace $f(n)$ in (3.39) so that the inclusions still hold for the case of the shift operator (3.41). Note that every principal submatrix of size $n \times n$, $A_{n,k}$, is the same, i.e.,

$$A_{n,k} = A_{n,1} = A_n := \begin{pmatrix} 0 & 0 & & & \\ 1 & 0 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 0 \\ & & & 1 & 0 \end{pmatrix}_{n\times n}. \tag{3.42}$$

Obviously, the spectrum (set of eigenvalues) of A_n is $\{0\}$. We will show that in this case

$$f_*(n) = f(n) = 2\sin\left(\frac{\pi}{4n+2}\right)$$

i.e. (3.39) with $f(n)$ given by (3.40) is already the sharpest approximation for the inclusion sets. In order to prove this, we will show that

$$\overline{\text{spec}_{f(n)} A_n} = \overline{B_1(0)}. \tag{3.43}$$

so that spec $A = \mathbb{T}$ would not be covered by $\overline{\text{spec}_{f(n)} A_n}$ if we chose $f(n)$ any smaller.

Proposition 3.11. *Let*

$$g(\lambda) := \nu\,(\lambda - A_n)^2 = \min \text{spec}\,\left[(\lambda - A_n)^*(\lambda - A_n)\right].$$

Then the following hold:

1. $g(1) = [f(n)]^2$
2. $g(\lambda) < g(\mu)$ *for* $0 \le \lambda < \mu$
3. $g(re^{i\theta}) = g(r)$ *for* $r \ge 0, \theta \in \mathbb{R}$

where $f(n) = 2\sin\left(\dfrac{\pi}{4n+2}\right)$.

As a consequence of Proposition 3.11, we get (3.43) and hence the sharpness of (3.39). In order to prove the above proposition, we need to prove the following useful lemma.

Lemma 3.12. *If* $x = (x_1, \ldots, x_n)^T \in \mathbb{C}^n$*, with* $\|x\|_2 = 1$ *and* $\widetilde{x} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)^T \in \mathbb{C}^n$ *defined by* $\widetilde{x}_j = e^{-ij\theta}x_j, j = 1, \ldots, n,$ *then* $\|\widetilde{x}\|_2 = 1$ *and for* $r \geq 0$*,*

$$\|(r - A_n)x\|_2 = \left\|(re^{i\theta} - A_n)\widetilde{x}\right\|_2.$$

Thus, $\nu(r - A_n) = \nu(re^{i\theta} - A_n)$*.*

Proof.

$$\begin{aligned}\|\widetilde{x}\|_2 &= \sqrt{|x_1 e^{-i\theta}|^2 + |x_2 e^{-2i\theta}|^2 + \ldots + |x_n e^{-ni\theta}|^2} \\ &= \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2} \\ &= \|x\|_2 = 1.\end{aligned}$$

For $r \geq 0$ we can see that

$$(re^{i\theta} - A_n)\widetilde{x} = \begin{pmatrix} rx_1 \\ (-x_1 + rx_2)e^{-i\theta} \\ \vdots \\ (-x_{n-1} + rx_n)e^{-(n-1)i\theta} \end{pmatrix}.$$

So,

$$\begin{aligned}\left\|\left(re^{i\theta} - A_n\right)\widetilde{x}\right\|_2 &= \sqrt{|rx_1|^2 + |(-x_1 + rx_2)e^{-i\theta}|^2 + \ldots + |(-x_{n-1} + rx_n)e^{-(n-1)i\theta}|^2} \\ &= \sqrt{|rx_1|^2 + |(-x_1 + rx_2)|^2 + \ldots + |(-x_{n-1} + rx_n)|^2} \\ &= \|(r - V_n)x\|_2.\end{aligned}$$

Therefore,

$$\nu(r - A_n) = \min_{\|x\|=1} \|(r - A_n)x\|_2 = \min_{\|\widetilde{x}\|=1} \left\|(re^{i\theta} - A_n)\widetilde{x}\right\|_2 = \nu(re^{i\theta} - A_n).$$

∎

We are now ready to prove Proposition 3.11.

Proof. [1.] From Theorem 2.16 and Proposition 2.48,

$$g(1) = [\nu(1 - A_n)]^2 = \left[\inf \sqrt{\frac{T_n^-}{S_n}}\right]^2 = [f(n)]^2.$$

[2.] Note that

$$(\lambda - A_n)^*(\lambda - A_n) = \begin{pmatrix} \lambda & -1 & & & \\ & \lambda & -1 & & \\ & & \lambda & \ddots & \\ & & & \ddots & -1 \\ & & & & \lambda \end{pmatrix}_{n\times n} \begin{pmatrix} \lambda & & & & \\ -1 & \lambda & & & \\ & -1 & \ddots & & \\ & & \ddots & \lambda & \\ & & & -1 & \lambda \end{pmatrix}_{n\times n}$$

i.e.

$$(\lambda - A_n)^*(\lambda - A_n) = \begin{pmatrix} \lambda^2 & -\lambda & & & \\ -\lambda & \lambda^2+1 & -\lambda & & \\ & \ddots & \ddots & \ddots & \\ & & -\lambda & \lambda^2+1 & -\lambda \\ & & & -\lambda & \lambda^2+1 \end{pmatrix}_{n\times n}$$

Note that $\widehat{\lambda}$ is an eigenvalue of $(\lambda - A_n)^*(\lambda - A_n)$ with an eigenvector $w = (w_n, w_{n-1}, \ldots, w_1)^T$ iff

$$(\lambda - A_n)^*(\lambda - A_n)w = \widehat{\lambda}w,$$

which can be written as follows

$$\left[\lambda^2 - \widehat{\lambda}\right] w_n - \lambda w_{n-1} = 0 \tag{3.44}$$

$$-\lambda w_{j+1} + \left[(\lambda^2+1) - \widehat{\lambda}\right] w_j - \lambda w_{j-1} = 0 \qquad j = 1, \ldots, n-1 \tag{3.45}$$

$$w_0 = 0. \tag{3.46}$$

Equation (3.45) implies that

$$w_j = Dm_1^j + Em_2^j \qquad j = 1, \ldots, n$$

where m_1 and m_2 are roots of $\lambda m^2 - (1 + \lambda^2 - \widehat{\lambda})m + \lambda = 0$ i.e. $m_1 m_2 = 1$

and $m_1 + m_2 = -\dfrac{1}{\lambda}(1 + \lambda^2 - \widehat{\lambda})$. Thus, put $m_1 = e^{i\theta}$

$$\begin{aligned}\widehat{\lambda} &= (\lambda^2 + 1) + \lambda(m_1 + m_2)\\ &= (\lambda^2 + 1) + \lambda(e^{i\theta} + e^{-i\theta})\\ &= (\lambda^2 + 1) + 2\lambda\cos\theta. \end{aligned} \tag{3.47}$$

Next, we want to show that if $\lambda_1 < \lambda_2$ then $\widehat{\lambda_1} < \widehat{\lambda_2}$. Since $w_j = B\cos(j\theta) + C\sin(j\theta)$ and $w_0 = 0$, it follows that $B = 0$. Hence, from equation (3.44) and (3.47), we obtain

$$\begin{aligned}-\lambda\sin[(n-1)\theta] + \left(\lambda^2 - \widehat{\lambda}\right)\sin(n\theta) &= 0\\ -\lambda\sin[(n-1)\theta] + (1 + 2\lambda\cos\theta)\sin(n\theta) &= 0\\ \lambda\sin[(n+1)\theta] + \sin(n\theta) &= 0.\end{aligned}$$

Therefore,

$$\lambda = -\frac{\sin(n\theta)}{\sin[(n+1)\theta]}.$$

Note that

$$\begin{aligned}\lambda_1 < \lambda_2 &\Leftrightarrow -\frac{\sin(n\theta_1)}{\sin[(n+1)\theta_1]} < -\frac{\sin(n\theta_2)}{\sin[(n+1)\theta_2]}\\ &\Leftrightarrow \frac{\sin(n\theta_1)}{\sin[(n+1)\theta_1]} > \frac{\sin(n\theta_2)}{\sin[(n+1)\theta_2]}.\end{aligned}$$

We can show by mathematical induction that $\sin k\theta < k\sin\theta$ for $k \in \mathbb{N}$ and note that

$$\begin{aligned}f'(\theta) &= \frac{n\sin[(n+1)\theta]\cos(n\theta) - (n+1)\sin(n\theta)\cos[(n+1)\theta]}{\sin^2[(n+1)\theta]}\\ &= \frac{n\sin\theta - \sin(n\theta)\cos[(n+1)\theta]}{\sin^2[(n+1)\theta]}\\ &\geq \frac{n\sin\theta - \sin(n\theta)}{\sin^2[(n+1)\theta]}\\ &> 0,\end{aligned}$$

where

$$f(\theta) = \frac{\sin(n\theta)}{\sin[(n+1)\theta]}.$$

Hence, $f(\theta)$ is an increasing function. Hence, $\lambda_1 < \lambda_2 \Leftrightarrow \theta_1 > \theta_2$. Therefore, from equation (3.47),

$$\lambda_1 < \lambda_2 \Rightarrow \theta_1 > \theta_2 \Rightarrow \cos\theta_1 < \cos\theta_2 \Rightarrow \widehat{\lambda}_1 < \widehat{\lambda}_2,$$

i.e. $g(\lambda) < g(\mu)$ for $0 \leq \lambda < \mu$.

[3.] From Lemma 3.12, we have

$$g(re^{i\theta}) = [\nu(re^{i\theta} - A_n)]^2 = [\nu(r - A_n)]^2 = g(r).$$

■

3.2.2 1 Dimensional Schrödinger Operator

We are now considering the Laurent operator $A := V_1 + V_{-1}$, where V_n is defined as in Definition 2.43, which is a discrete 1D Schrödinger operator with potential zero. The corresponding spectrum is just the curve

$$\operatorname{spec} A = \{t^{-1} + t : t \in \mathbb{T}\} = \{e^{-i\theta} + e^{i\theta} : \theta \in \mathbb{R}\} = \{2\cos\theta : \theta \in \mathbb{R}\} = [-2, 2].$$

Again, the finite submatrices $A_{n,k}$ do not depend on k:

$$A_{n,k} = A_n := \begin{pmatrix} 0 & 1 & & & \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ & & & 1 & 0 \end{pmatrix}_{n\times n}. \tag{3.48}$$

We can see that A_n is a self-adjoint matrix, so $\operatorname{spec} A_n \subseteq \mathbb{R}$. Moreover, λ is an eigenvalue of A_n iff $A_n v = \lambda v$ where v is the corresponding eigenvector, i.e.

$$\begin{pmatrix} -\lambda & 1 & & & \\ 1 & -\lambda & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -\lambda & 1 \\ & & & 1 & -\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n-1} \\ v_n \end{pmatrix} = 0. \tag{3.49}$$

From equation (3.49), we have that $(A_n - \lambda I)v = 0$ iff

$$v_{j+1} - \lambda v_j + v_{j-1} = 0 \text{ for } j = 1, \ldots, n, \tag{3.50}$$

where we put $v_0 := 0$ and $v_{n+1} := 0$. The characteristic equation is

$$t^2 - \lambda t + 1 = 0.$$

Denote the two solutions by $t_1, t_2 \in \mathbb{C}$. Since $x \in \mathbb{R}$, we have $t_1 = \overline{t_2}$. We have $1 = t_1 t_2 = t_1 \overline{t_1} = |t_1|^2$. So, put $t_1 = e^{i\theta}$ and $t_2 = e^{-i\theta}$ for some $\theta \in [0, \pi]$. Then

$$v_j = A\cos(j\theta) + B\sin(j\theta), \quad j = 0, 1, 2, \ldots,$$

where A and B are arbitrary constants. Since $v_0 = 0$ and $v_{n+1} = 0$, it follows that $A = 0$ and hence

$$\sin\left((n+1)\theta\right) = 0 \Rightarrow \theta = \frac{r\pi}{n+1}, \qquad r \in \mathbb{Z}.$$

Under the condition $\theta \in [0, \pi]$ and v is not a zero vector, we know that $\theta \in (0, \pi)$. Therefore,

$$\theta = \frac{r\pi}{n+1}, \quad r = 1, \ldots, n.$$

So we can see that $\operatorname{spec} A_n$ is the set of n points, $\lambda_r = 2\cos\left(\frac{r\pi}{n+1}\right)$ where $r = 1, \ldots, n$. As in Section 3.2.1, we are again asking by what function $f_*(n)$ one could replace $f(n)$ such that (3.2) still holds. Or in other words, for every $n \in \mathbb{N}$, what is the smallest number ε_n for which it holds that

$$\operatorname{spec} A \subseteq \overline{\bigcup_{k \in \mathbb{Z}} \operatorname{spec}_{\varepsilon_n} A_{n,k}}. \tag{3.51}$$

Since we know that $A_{n,k}$ is normal, by Theorem 2.20, (3.51) holds iff

$$[-2, 2] = \operatorname{spec} A \subseteq \overline{\bigcup_{k \in \mathbb{Z}} \operatorname{spec}_{\varepsilon_n} A_{n,k}} = \left\{2\cos\left(\frac{r\pi}{n+1}\right) : r = 1, \ldots, n\right\} + \overline{B_{\varepsilon_n}(0)}. \tag{3.52}$$

Equation (3.52) holds iff the interval $[-2, 2]$ is covered by the union of n closed balls of radius ε_n centred at the eigenvalues of A_n. The smallest ε_n for which this is true is

$$\varepsilon_n = \max\left\{|2-\lambda_1|, \varepsilon_n^*, |2+\lambda_n|\right\},$$

where

$$\begin{aligned}\varepsilon_n^* &= \max\left\{\left|\frac{1}{2}|\lambda_2-\lambda_1|, \frac{1}{2}|\lambda_3-\lambda_2|, \ldots, \frac{1}{2}|\lambda_n-\lambda_{n-1}|\right\}\right. \\ &= \max_{r=1,\ldots,n}\left\{\left|\cos\left(\frac{(r-1)\pi}{n+1}\right)-\cos\left(\frac{r\pi}{n+1}\right)\right|\right\}. \end{aligned} \tag{3.53}$$

Next, we will show that, $\lambda_{r-1}-\lambda_r < \lambda_r - \lambda_{r+1}$ for $r = 1, \ldots, \left\lceil \frac{n-1}{2} \right\rceil$. It suffices to show that $\lambda_{r-1}+\lambda_{r+1} < 2\lambda_r$, which holds as

$$\begin{aligned}\lambda_{r-1}+\lambda_{r+1} &= 2\cos\left(\frac{(r-1)\pi}{n+1}\right)+2\cos\left(\frac{(r+1)\pi}{n+1}\right) \\ &= 2\left(2\cos\left(\frac{r\pi}{n+1}\right)\cos\left(\frac{\pi}{n+1}\right)\right) \\ &< 2\left(2\cos\left(\frac{r\pi}{n+1}\right)\right) = 2\lambda_r.\end{aligned}$$

Since we know that $\lambda_{r-1}-\lambda_r > 0$, we can conclude that $|\lambda_{r-1}-\lambda_r| < |\lambda_r - \lambda_{r+1}|$, for $r = 1, \ldots, \left\lceil \frac{n-1}{2} \right\rceil$. Further, given that $\lambda_0 = 2$ and $\lambda_{n+1} = -2$

$$\begin{aligned}|\lambda_{n-r}-\lambda_{n-r+1}| &= \left|2\cos\left(\frac{(n-r)\pi}{n+1}\right)-2\cos\left(\frac{(n-r+1)\pi}{n+1}\right)\right| \\ &= \left|2\cos\left(\pi-\frac{(r+1)\pi}{n+1}\right)-2\cos\left(\pi-\frac{r\pi}{n+1}\right)\right| \\ &= \left|2\cos\left(\pi-\frac{(r+1)\pi}{n+1}\right)-2\cos\left(\pi-\frac{r\pi}{n+1}\right)\right| \\ &= \left|2\cos\left(\frac{(r+1)\pi}{n+1}\right)-2\cos\left(\frac{r\pi}{n+1}\right)\right| \\ &= |\lambda_{r+1}-\lambda_r|,\end{aligned} \tag{3.54}$$

for $r = 0, \ldots, n$. Thus

$$\varepsilon_n^* = \begin{cases} \frac{1}{2}\left|\lambda_{\frac{n}{2}} - \lambda_{\frac{n+2}{2}}\right| & \text{if } n \text{ is even,} \\ \frac{1}{2}\left|\lambda_{\frac{n-1}{2}} - \lambda_{\frac{n+1}{2}}\right| & \text{if } n \text{ is odd} \end{cases}$$

$$= \begin{cases} 2\sin\left(\dfrac{\pi}{2(n+1)}\right) & \text{if } n \text{ is even,} \\ 2\sin\left(\dfrac{\pi}{2(n+1)}\right)\cos\left(\dfrac{\pi}{2(n+1)}\right) & \text{if } n \text{ is odd} \end{cases}.$$

In order to compute ε_n for each n, we will now firstly consider the case $n = 1$. We can see that the set of eigenvalue of A_n when $n = 1$ is $\{0\}$. Thus, the smallest $\varepsilon_1 = 2 > f(1)$. From (3.54) and definition of ε_n^*, we have that for every $n \geq 2$

$$|2 - \lambda_1| = |2 + \lambda_n| = \left|2 - 2\cos\left(\frac{\pi}{n+1}\right)\right| = 4\sin^2\left(\frac{\pi}{2(n+1)}\right) < \varepsilon_n^*.$$

Therefore, $\varepsilon_n = \varepsilon_n^*$. Where $f(n)$ is defined in Corollary 3.7, we know that

$$4\sin\left(\frac{\pi}{2(n+3)}\right) < f(n) < 4\sin\left(\frac{\pi}{2(n+1)}\right).$$

We can see that $\dfrac{f(n)}{\varepsilon_n} \to 2$ as $n \to \infty$, so that, for this example, $f(n)$ overestimates the smallest ε_n for which (3.51) holds by a factor of about 2 for larger value of n.

For this example, plots of the inclusion sets in Corollary 3.7 for $n = 4, 8, 16$ and 64, are shown in Figure 3.1.

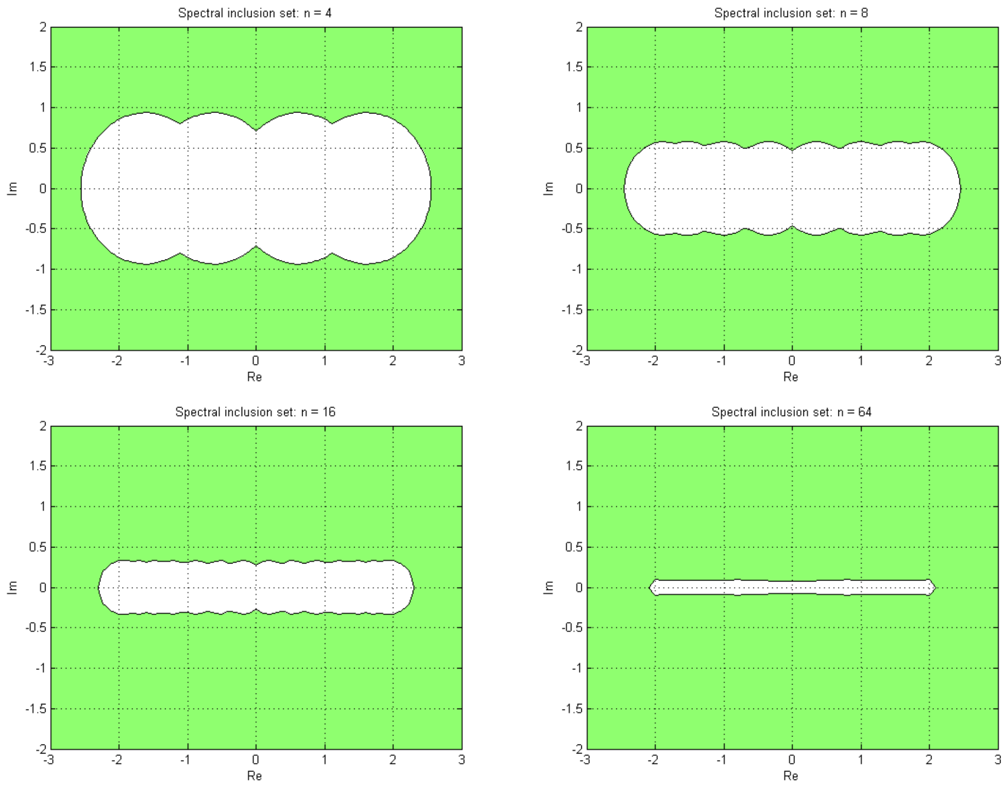

Figure 3.1: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}$, which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 3.7 and $A_{n,k}$ is the ordinary finite submatrix given by (3.5) of the 1-dimensional Schrödinger operator, $A = V_1 + V_{-1}$, with $\operatorname{spec} A = [-2, 2]$. Shown are the inclusion sets when $n = 4, 8, 16$ and 64.

3.2.3 3-periodic Bi-diagonal Operator

In this example, we are considering the bidiagonal operator

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & -1.5 & 1 & & & & \\ & 0 & 1 & 2 & & & \\ & & 0 & 1 & 1 & & \\ & & & 0 & -1.5 & 1 & \\ & & & & 0 & & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}.$$

Since A is not normal, we cannot apply Theorem 2.20 to this operator. However, Corollary 3.3 and Corollary 3.6 tell us that in the bidiagonal case, $\|\gamma\|_\infty = 0$,

$$\operatorname{spec}_\varepsilon A \subseteq \Sigma^n_{\varepsilon+f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)}. \tag{3.55}$$

where $f(n)$ is defined by (3.40) and the finite submatrices, $A_{n,k}$, of A must be one of the following $n \times n$ matrices:

$$\begin{pmatrix} -1.5 & 1 & & & \\ & 1 & 2 & & \\ & & 1 & 1 & \\ & & & -1.5 & \ddots \\ & & & & \ddots \end{pmatrix}, \begin{pmatrix} 1 & 2 & & & \\ & 1 & 1 & & \\ & & -1.5 & 1 & \\ & & & 1 & \ddots \\ & & & & \ddots \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 & & & \\ & -1.5 & 1 & & \\ & & 1 & 2 & \\ & & & 1 & \ddots \\ & & & & \ddots \end{pmatrix}.$$

Therefore, in order to compute our inclusion sets, we do not have to compute the pseudospectra of infinitely many submatrices but only 3 submatrices of size n. In Figure 3.3 we plot the inclusion sets in (3.55) for $\operatorname{spec} A$ for $n = 4, 8, 16, 32$ and 64, and we also plot $\operatorname{spec} A$, computed using the formulae in Theorem 4.4.9 of Davies [13] (cf. Theorem 5.12 in Chapter 5).

These inclusion sets seem to be converging to $\widehat{\operatorname{spec} A}$ where $\widehat{\operatorname{spec} A}$, defined as in Section 2.1, is the complement of the unbounded component of $\mathbb{C} \setminus \operatorname{spec} A$.

3.3 Inclusion sets in terms of quasi-circulant modification matrices

In this chapter so far we have computed the inclusion sets for $\operatorname{spec} A$ using ordinary finite submatrices, $A_{n,k}$. In this section we will prove corresponding results when, instead, we replace $A_{n,k}$ by the "periodised" submatrices, $\hat{A}_{n,k}$, defined in (3.56). We will show numerical examples where this method produce much sharper inclusion sets.

$$\hat{A}_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \alpha_k \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ \gamma_{k+n+1} & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}. \tag{3.56}$$

Theorem 3.13. *If $\varepsilon > 0$, $n \in \mathbb{N}$, $w_j > 0$, for $j = 1, \ldots, n$ and $w_0 = w_{n+1} = 0$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Pi^n_{\varepsilon+f(n)}(A),$$

where

$$f(n) = (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\frac{T_n}{S_n}},$$

$$S_n = \sum_{i=1}^{n} w_i^2 \text{ and } T_n = (w_1 + w_n)^2 + \sum_{i=1}^{n-1} (w_{i+1} - w_i)^2.$$

Proof. Let $\lambda \in \operatorname{spec}_\varepsilon A$. Then either there exists $x \in \ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A - \lambda I)x\| < \varepsilon$, or the same holds with A replaced by its adjoint. In the first case, let $y = (A - \lambda I)x$, so $\|y\| < \varepsilon$.

For $i, k \in \mathbb{Z}$, define $e_i^{(k)}$, $E^+_{i,k}$ and $E^-_{i,k}$ as in the proof of Theorem 3.2. For $k \in \mathbb{Z}$, let

$$P_k := \|(A_{n,k} - \lambda I_n)\widetilde{x}_{n,k}\|,$$

where $\tilde{x}_{n,k} = (w_1 x_{k+1}, w_2 x_{k+2}, \ldots, w_n x_{k+n})^T$. Let

$$A_{i,k} = \begin{cases} |x_{k+n}|\, w_n & \text{if } i = k+1 \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.57}$$

and

$$B_{i,k} = \begin{cases} |x_{k+1}|\, w_1 & \text{if } i = k+n \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.58}$$

We will prove that $P_k < (\varepsilon + f(n)) \|\tilde{x}_{n,k}\|$ for some $k \in \mathbb{Z}$, which will show that $\lambda \in \operatorname{spec}_{\varepsilon+f(n)} A_{n,k}$.

Note first that, using (3.6), (3.7), (3.57) and (3.58)

$$\begin{aligned} P_k^2 = &\left| y_{k+1} e_{k+1}^{(k)} + \alpha_k (x_{k+n} e_{k+n}^{(k)} - x_k e_{k+1}^{(k)}) + \gamma_{k+2} (e_{k+2}^{(k)} - e_{k+1}^{(k)}) x_{k+2} \right|^2 \\ &+ \sum_{i=k+2}^{k+n-1} \left| y_i e_i^{(k)} + \alpha_{i-1} (e_{i-1}^{(k)} - e_i^{(k)}) x_{i-1} + \gamma_{i+1} (e_{i-1}^{(k)} - e_i^{(k)}) x_{i+1} \right|^2 \\ &+ \left| y_{k+n} e_{k+n}^{(k)} + \alpha_{k+n-1} (e_{k+n-1}^{(k)} - e_{k+n}^{(k)}) x_{k+n-1} + \gamma_{k+n+1} (x_{k+1} e_{k+1}^{(k)} - x_{k+n} e_{n+k}^{(k)}) \right|^2 \\ &\le \sum_{i=k+1}^{k+n} \left(|y_i|\, e_{i,k}^{(k)} + |\alpha_{i-1}| \left(A_{i,k} + E_{i,k}^{-} |x_{i-1}| \right) + |\gamma_{i+1}| \left(B_{i,k} + E_{i,k}^{+} |x_{i+1}| \right) \right)^2 . \end{aligned}$$

So, for all $\theta > 0$ and $\phi > 0$, by Lemma 3.1,

$$\begin{aligned} P_k^2 \le &\sum_{i=k+1}^{k+n} \left[(1+\theta) \left(|y_i|\, e_i^{(k)} \right)^2 + (1+\theta^{-1}) \left(|\alpha_{i-1}| \left(A_{i,k} + E_{i-k}^{-} |x_{i-1}| \right) \right. \right. \\ &\left. \left. + \, |\gamma_{i+1}| \left(B_{i,k} + E_{i-k}^{+} |x_{i+1}| \right) \right)^2 \right] \\ &\le \sum_{i=k+1}^{k+n} \left[(1+\theta) \left(|y_i|\, e_i^{(k)} \right)^2 + (1+\theta^{-1}) \left((1+\phi) |\alpha_{i-1}|^2 \left(A_{i,k} + E_{i,k}^{-} |x_{i-1}| \right)^2 \right. \right. \\ &\left. \left. + (1+\phi^{-1}) |\gamma_{i+1}|^2 \left(B_{i,k} + E_{i,k}^{+} |x_{i+1}| \right)^2 \right) \right] . \end{aligned}$$

Note that,

$$\sum_{k \in \mathbb{Z}} x_{k+i}^2 e_{k+i}^{(k)\,2} = \left(e_{i+k}^{(k)} \right)^2 \|x\|^2$$

for all i, and we have that

$$\sum_{k\in\mathbb{Z}}(x_{k+n}e^{(k)}_{k+n} - x_k e^{(k)}_{i+k})^2 \leq \left(e^{(k)}_{k+n}\right)^2 \|x\|^2 + 2e^{(k)}_{k+1}e^{(k)}_{k+n}\sum_{k\in\mathbb{Z}}|x_{k+n}x_k| + \left(e^{(k)}_{k+1}\right)^2 \|x\|^2$$
$$\leq (e^{(k)}_{k+1} + e^{(k)}_{k+n})^2 \|x\|^2 .$$

Thus

$$\begin{aligned}\sum_{k\in\mathbb{Z}} P_k^2 &\leq (1+\theta)\sum_{i\in\mathbb{Z}}|y_i|^2\sum_{k\in\mathbb{Z}}\left(e^{(k)}_i\right)^2 \\ &+ (1+\theta^{-1})\Big[(1+\phi)\|\alpha\|^2_\infty\left\{(e^{(k)}_{k+n}+e^{(k)}_{k+1})^2 + (E^-_{k+2,k})^2 + \cdots + (E^-_{k+n,k})^2\right\} \\ &+ (1+\phi^{-1})\|\gamma\|^2_\infty\{(E^+_{k+1,k})^2 + (E^+_{k+2,k})^2 + \cdots + (E^+_{k+n-1,k})^2 + (e^{(k)}_{k+1}+e^{(k)}_{k+n})^2\}\Big] \\ &\leq (1+\theta)\|y\|^2_2 S_n + (1+\theta^{-1})\left((\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{T_n}\right)^2 \\ &< (1+\theta)\varepsilon^2 S_n + (1+\theta^{-1})\left((\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{T_n}\right)^2\end{aligned}$$

Thus,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} P_k^2 &< \left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})\left((\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{T_n}\right)^2\frac{1}{S_n}\right] S_n \\ &\leq \left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})\,[f(n)]^2\right]\sum_{k\in\mathbb{Z}}\|\tilde{x}_{n,k}\|^2 .\end{aligned}$$

Applying Lemma 3.1 again, we see that

$$\inf_{\theta>0}\left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})[f(n)]^2\right] = (\varepsilon + f(n))^2,$$

so that

$$\sum_{k\in\mathbb{Z}} P_k^2 < (\varepsilon + f(n))^2\sum_{k\in\mathbb{Z}}\|\tilde{x}_{n,k}\|^2 .$$

Thus, for some $k\in\mathbb{Z}$,

$$P_k < (\varepsilon + f(n))\,\|\tilde{x}_{n,k}\| ,$$

so that $\lambda \in \operatorname{spec}_{\varepsilon+f(n)} A_{n,k}$.

In the case when there exists $x\in\ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A-\lambda I)^* x\| < \varepsilon$, the same argument shows that $\overline{\lambda}\in\operatorname{spec}_{\varepsilon+f(n)} A^*_{n,k}$, for some $k\in\mathbb{Z}$, so that $\lambda\in\operatorname{spec}_{\varepsilon+f(n)} A_{n,k}$. ∎

Corollary 3.14. $\operatorname{spec} A \subseteq \overline{\Pi^n_{f(n)}(A)}$.

Proof. We can see that, if

$$\lambda \in \operatorname{spec} A = \bigcap_{\varepsilon>0} \operatorname{spec}_{\varepsilon} A \subseteq \bigcap_{\varepsilon>0} \left(\bigcup_{k\in\mathbb{Z}} \operatorname{spec}_{\varepsilon+f(n)} \hat{A}_{n,k} \right),$$

so that for all $\varepsilon > 0$ there is a $k \in \mathbb{Z}$ with $\left\| \left(\hat{A}_{n,k} - \lambda I_n \right)^{-1} \right\| > \dfrac{1}{\varepsilon + f(n)}$, then $s := \sup_{k\in\mathbb{Z}} \left\| (\hat{A}_{n,k} - \lambda I_n)^{-1} \right\| \geq \dfrac{1}{f(n)}$. If $s > \dfrac{1}{f(n)}$ then there exists a $k \in \mathbb{Z}$ with $\lambda \in \operatorname{spec}_{f(n)} \hat{A}_{n,k} \subseteq \Pi^n_{f(n)}(A)$. If $s = \dfrac{1}{f(n)}$ then put $D :=\operatorname{Diag}\left\{ \hat{A}_{n,k} : k \in \mathbb{Z} \right\}$, so that

$$\left\| (D - \lambda I)^{-1} \right\| = \sup_{k\in\mathbb{Z}} \left\| \left(\hat{A}_{n,k} - \lambda I_n \right)^{-1} \right\| = s = \frac{1}{f(n)}.$$

Take $r > 0$ small enough that $\lambda + r\mathbb{D} \subseteq \rho(D) := \mathbb{C} \setminus \operatorname{spec} D$. By Theorem 2.32, there are $\mu_1, \mu_2, \ldots \in \rho(D)$ with $|\mu_m - \lambda| < \dfrac{r}{m}$ and $\left\| (D - \mu_m I)^{-1} \right\| > \left\| (D - \lambda I)^{-1} \right\| = \dfrac{1}{f(n)}$ for $m = 1, 2, \ldots$. Hence, λ is in the closure of $\operatorname{spec}_{f(n)} D = \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_{f(n)} \hat{A}_{n,k} = \Pi^n_{f(n)}(A)$. ■

From the definitions of S_n and T_n in Theorem 3.13, we know that

$$S_n = \|w\|_2^2$$

and

$$T_n = \|B_n w\|_2^2$$

where

$$B_n = \begin{pmatrix} 1 & -1 & & & \\ & 1 & -1 & & \\ & & \ddots & \ddots & \\ & & & 1 & -1 \\ -1 & & & & 1 \end{pmatrix}.$$

From Theorem 2.16 and Proposition 2.48 we have that

$$\inf_{w\neq 0} \sqrt{\frac{T_n}{S_n}} = \inf_{w\neq 0} \frac{\|B_n w\|_2}{\|w\|_2} = \nu(B_n) = \sqrt{\lambda_{min}(B_n B_n^T)} = s_{\min}(B_n).$$

Further

$$B_n B_n^T = \begin{pmatrix} 2 & -1 & & & 1 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 1 & & & -1 & 2 \end{pmatrix}.$$

We know that λ is an eigenvalue of $B_n B_n^T$ with eigenvector $v = \begin{pmatrix} v_n \\ \vdots \\ v_1 \end{pmatrix}$ iff

$$\begin{pmatrix} 2-\lambda & -1 & & & 1 \\ -1 & 2-\lambda & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2-\lambda & -1 \\ 1 & & & -1 & 2-\lambda \end{pmatrix} \begin{pmatrix} v_n \\ v_{n-1} \\ \vdots \\ v_2 \\ v_1 \end{pmatrix} = 0. \tag{3.59}$$

From equation (3.59), we have that $(B_n B_n^T - \lambda I)v = 0$ iff

$$v_1 + (2-\lambda)v_n - v_{n-1} = 0, \tag{3.60}$$

$$-v_{i+1} + (2-\lambda)v_i - v_{i-1} = 0 \text{ for } i = 2, \ldots, n-1, \tag{3.61}$$

and

$$-v_2 + (2-\lambda)v_1 + v_n = 0. \tag{3.62}$$

Equation (3.61) has general solution

$$v_j = A\cos((j-1)\theta) + B\sin((j-1)\theta), \qquad j = 1, 2, \ldots, n-1,$$

where $\lambda = 2(1-\cos\theta) = 4\sin^2(\frac{\theta}{2})$. From equation (3.60) we have then that,

$$v_1 = A \text{ and } v_2 = A\cos\theta + B\sin\theta.$$

Clearly, for a three-term-recurrence relation, we need to know the first two terms as initial conditions. In the other word, for every $(v_1, v_2) \in \mathbb{C}^2$, there exists exactly one sequence $(v_1, v_2, \ldots, v_n) \in \mathbb{C}^n$ which satisfies the equation (3.60). Provided $\sin\theta \neq 0$, this solution is given by

$$v_j = v_1 cos\left((j-1)\theta\right) + \frac{(v_2 - v_1\cos\theta)}{\sin\theta}\sin\left((j-1)\theta\right)$$

$j = 1, 2, \ldots, n$. Setting $v_1 = v_2 = 1$, then substituting v_1, v_2 and v_n in the equation (3.62) yields

$$\sin\left((n-1)\theta\right) - \sin\theta - \sin\left((n-2)\theta\right) + \sin(2\theta) = 0,$$

i.e.

$$4\cos\left(\frac{n\theta}{2}\right)\cos\left(\frac{n-3}{2}\theta\right)\sin\left(\frac{\theta}{2}\right) = 0.$$

That means the smallest θ which satisfies the above equation is $\theta = \dfrac{\pi}{n}$.

Corollary 3.15. *If* $\epsilon > 0$, $n \in \mathbb{N}$ *, then*

$$\operatorname{spec}_{\varepsilon} A \subseteq \Pi^n_{\varepsilon + f(n)}(A) \text{ and } \operatorname{spec} A \subseteq \overline{\Pi^n_{f(n)}(A)},$$

where

$$f(n) = 2\sin\left(\frac{\pi}{2n}\right)\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right).$$

3.4 Numerical Examples for Method 1*

3.4.1 Shift Operator

Note that every finite $n \times n$ periodised submatrix of the shift operator V_1 is of the form,

$$\widehat{A}_n := \begin{pmatrix} 0 & 0 & & & 1 \\ 1 & 0 & 0 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 0 \\ 0 & & & 1 & 0 \end{pmatrix},$$

i.e., $\widehat{A}_{n,k} = \widehat{A}_n$ for every $k \in \mathbb{Z}$. From the normality property of V_1 and Theorem 2.20, we can compute that $\operatorname{spec}_\varepsilon V_1 = \mathbb{T} \cup \{\lambda \in \mathbb{C} : d(\lambda, \mathbb{T}) < \varepsilon\}$ and

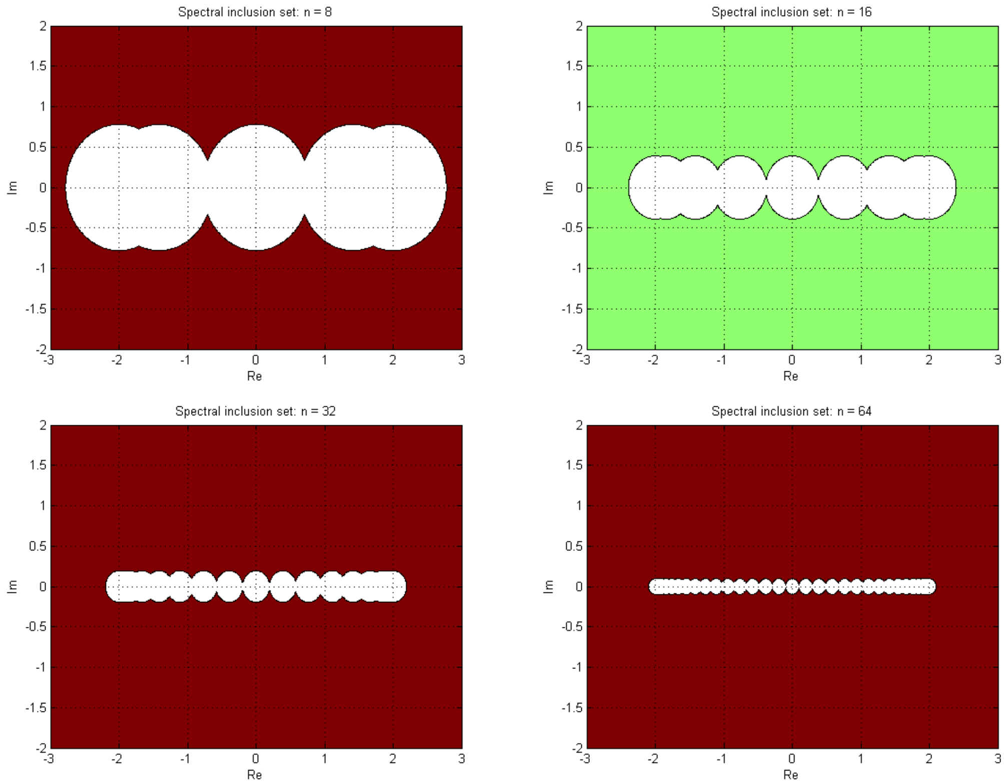

Figure 3.2: Plots of the sets $\overline{\Pi^n_{f(n)}(A)}$, which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 3.15, where A is the 1-dimensional Schrödinger operator, $A = V_1 + V_{-1}$, with $\operatorname{spec} A = [-2, 2]$. Shown are the inclusion sets when $n = 8, 16, 32$ and 64.

$\operatorname{spec}_{\varepsilon+f(n)}\hat{A}_n = \mathbb{T}_n + B_{\varepsilon+f(n)}(0)$, where $\mathbb{T}$ and $\mathbb{T}_n$ are the unit circle and the set of the n-th roots of unity, respectively. Thus Corollary 3.15, in the case $A = V_1$ reduces to

$$\mathbb{T} + B_\varepsilon(0) = \operatorname{spec}_\varepsilon V_1 \subseteq \Pi^n_{\varepsilon+f(n)}(A) = \operatorname{spec}_{\varepsilon+f(n)}\hat{A}_n = \mathbb{T}_n + B_{\varepsilon+f(n)}(0),$$

and

$$\mathbb{T} = \operatorname{spec} V_1 \subseteq \overline{\Pi^n_{\varepsilon+f(n)}(A)} = \overline{\operatorname{spec}_{f(n)}\hat{A}_n} = \mathbb{T}_n + \overline{B_{f(n)}(0)},$$

where $f(n) = 2\sin\left(\frac{\pi}{2n}\right)$.

Let us now consider the sharpness of this value for $f(n)$. Clearly, in this example, for $\varepsilon > 0$,

$$\Pi^n_{\varepsilon+f(n)}(A) \to \operatorname{spec}_\varepsilon(A)$$

and

$$\overline{\Pi^n_{f(n)}(A)} \to \operatorname{spec}(A)$$

as $n \to \infty$ in the Haussdorff metric. Let us now compute the smallest possible ε_n which satisfies

$$\operatorname{spec}_\varepsilon V_1 \subseteq \Pi^n_{\varepsilon+\varepsilon_n}(V_1) = \operatorname{spec}_{\varepsilon+\varepsilon_n}\hat{A}_n,$$

and

$$\mathbb{T} = \operatorname{spec} A \subseteq \overline{\Pi^n_{\varepsilon_n}(V_1)} = \overline{\operatorname{spec}_{\varepsilon_n}\hat{A}_n} = \mathbb{T}_n + \overline{B_\varepsilon(0)}. \tag{3.63}$$

We can compute that $\operatorname{spec}\hat{A}_n = \mathbb{T}_n$. Since we know that the distances between any 2 consecutive roots of unity are equal, therefore, the smallest possible number ε_n is the half of the distance between each pair of the consecutive roots. The arc between 2 consecutive roots on the circumferrence subtends the angle $\frac{2\pi}{n}$. Therefore, the smallest value of ε_n for which (3.63) holds is

$$\varepsilon_n = 2\sin\left(\frac{\pi}{2n}\right) = f(n).$$

3.4.2 1D - Schrödinger Operator

As an example for Corollary 3.15, we will apply this method to the operator $A = V_{-1} + V_1$. Note that every $n \times n$ periodised submatrix of the operator A is of the form

$$\widehat{A}_n = \begin{pmatrix} 0 & 1 & & & 1 \\ 1 & 0 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & 0 & 1 \\ 1 & & & 1 & 0 \end{pmatrix},$$

which means our inclusion sets of the spectrum can be computed using just the periodised matrix $\widehat{A}_n$. Obviously, Corollary 3.15 can be rewritten as

$$\operatorname{spec}_{\epsilon} A \subseteq \bigcup_{k\in\mathbb{Z}} \operatorname{spec}_{\epsilon+f(n)} \hat{A}_{n,k} = \operatorname{spec}_{\epsilon+f(n)} \hat{A}_n.$$

We can see that $\widehat{A}_n$ is a symmetric matrix, so $\operatorname{spec} \widehat{A}_n \subseteq \mathbb{R}$. $(\widehat{A}_n - \lambda I)v = 0$ iff

$$v_1 - \lambda v_n + v_{n-1} = 0 \tag{3.64}$$

$$v_{j+1} - \lambda v_j + v_{j-1} = 0 \text{ for } j = 2, \dots, n-1. \tag{3.65}$$

$$v_2 - \lambda v_1 + v_n = 0. \tag{3.66}$$

The equation (3.65) has general solution

$$v_j = A\cos((j-1)\theta) + B\sin((j-1)\theta)$$

where $j = 1, \dots, n$ and

$$\lambda = 2\cos\theta.$$

From the equation (3.66), we have $v_1 = A$ and then

$$\begin{aligned} v_2 &= A\cos\theta + B\sin\theta \\ &= v_1\cos\theta + B\sin\theta \\ B &= \frac{v_2 - v_1\cos\theta}{\sin\theta} \end{aligned}$$

where $\lambda = 2\cos\theta$. Setting $v_1 = v_2 = 1$, then $A = 1$ and $B = \dfrac{1-\cos\theta}{\sin\theta}$. From equation (3.66),

$$\begin{aligned} & v_n - \lambda v_1 + v_2 = 0 \\ \Leftrightarrow\ & \cos((n-1)\theta) + \left(\frac{1-\cos\theta}{\sin\theta}\right)\sin((n-1)\theta) - 2\cos\theta + 1 = 0 \\ \Leftrightarrow\ & -4\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{n\theta}{2}\right)\sin\left(\frac{(n-3)\theta}{2}\right) = 0. \end{aligned}$$

Therefore, $\theta = \dfrac{2\pi}{n}$ is the smallest possible $\theta > 0$. So we can see that the spec $\hat{A}_n$ is the set of n points, $\lambda_r = 2\cos\left(\dfrac{2r\pi}{n}\right)$, where $r = 1, \ldots, n$. As in Section 3.2.1, we are again asking by what function $f_*(n)$ one could replace $f(n)$ such that (3.2) still holds. Or in other words, for every $n \in \mathbb{N}$, what is the smallest number ε_n for which it holds that

$$\operatorname{spec} A \subseteq \overline{\Pi^n_{\varepsilon_n}(A)}. \tag{3.67}$$

Since $A_{n,k}$ is normal, by Theorem 2.20, (3.67) holds iff

$$[-2,2] = \operatorname{spec} A \subseteq \overline{\Pi^n_{\varepsilon_n}} = \left\{\lambda_r = 2\cos\left(\frac{r\pi}{n+1}\right) : r = 1, \ldots, n\right\} + \overline{B_{\varepsilon_n}(0)}. \tag{3.68}$$

Equation (3.68) holds iff the interval $[-2,2]$ is covered by the union of n closed balls of radius ε_n centred at the eigenvalues of $\hat{A}_n$. The smallest ε_n for which this is true is

$$\varepsilon_n = \max\{|2-\lambda_1|, \varepsilon_n^*, |2+\lambda_n|\}$$

where

$$\begin{aligned} \varepsilon_n^* &= \max\left\{\frac{1}{2}|\lambda_2-\lambda_1|, \frac{1}{2}|\lambda_3-\lambda_2|, \ldots, \frac{1}{2}|\lambda_n-\lambda_{n-1}|\right\} \\ &= \max_{r=1,\ldots,n}\left\{\left|\cos\left(\frac{2(r-1)\pi}{n}\right) - \cos\left(\frac{2r\pi}{n}\right)\right|\right\}. \end{aligned} \tag{3.69}$$

Next, we will show that, $\lambda_{r-1} - \lambda_r < \lambda_r - \lambda_{r+1}$ for $1 \le r \le \dfrac{n}{2}$. It suffices to show that $\lambda_{r-1} + \lambda_{r+1} < 2\lambda_r$, which holds as

$$\begin{aligned}\lambda_{r-1} + \lambda_{r+1} &= 2\cos\left(\frac{2(r-1)\pi}{n}\right) + 2\cos\left(\frac{2(r+1)\pi}{n}\right)\\ &= 2\left(2\cos\left(\frac{2r\pi}{n}\right)\cos\left(\frac{2\pi}{n}\right)\right)\\ &< 2\left(2\cos\left(\frac{2r\pi}{n}\right)\right) = 2\lambda_r.\end{aligned}$$

Since we know that $\lambda_{r-1} - \lambda_r > 0$ for $1 \le r \le \frac{n}{2}$, we can conclude that $|\lambda_{r-1} - \lambda_r| < |\lambda_r - \lambda_{r+1}|$. Further, given that $\lambda_0 = 2$ and $\lambda_{n+1} = -2$

$$\begin{aligned}\left|\lambda_{\frac{n}{2}-r+1} - \lambda_{\frac{n}{2}-r}\right| &= \left|2\cos\left(\frac{2(\frac{n}{2}-r+1)\pi}{n}\right) - 2\cos\left(\frac{2(\frac{n}{2}-r)\pi}{n}\right)\right|\\ &= \left|2\cos\left(2\pi - \frac{2(r-1)\pi}{n}\right) - 2\cos\left(2\pi - \frac{2r\pi}{n}\right)\right|\\ &= \left|2\cos\left(\frac{2(r-1)\pi}{n}\right) - 2\cos\left(\frac{2r\pi}{n}\right)\right|\\ &= \left|\lambda_{r-1} - \lambda_r\right|,\end{aligned}$$

for $r = 0, \ldots, n$. Thus,

$$\begin{aligned}\varepsilon_n^* &= \begin{cases}\frac{1}{2}\left|\lambda_{\frac{n-2}{4}} - \lambda_{\frac{n+2}{4}}\right| & \text{if } \frac{n}{2} \text{ is odd,}\\ \frac{1}{2}\left|\lambda_{\frac{n+4}{4}} - \lambda_{\frac{n}{4}}\right| & \text{if } \frac{n}{2} \text{ is even}\end{cases}\\ &= \begin{cases}2\sin\left(\frac{\pi}{n}\right) & \text{if } \frac{n}{2} \text{ is odd,}\\ \sin\left(\frac{2\pi}{n}\right) & \text{if } \frac{n}{2} \text{ is even}\end{cases}.\end{aligned}$$

In order to compute ε_n^*. we will now firstly consider the case $n = 1$. We can see that the set of eigenvalue of $\hat{A}_1$ is $\{0\}$. Thus, the smallest $\varepsilon_1 = 2 > f(1)$. Where $f(n)$ is defined in Corollary 3.15, we know that $f(n) = 4\sin\left(\frac{\pi}{2n}\right)$. We can see that

$$\varepsilon_n = 2\sin\left(\frac{\pi}{n}\right) < 4\sin\left(\frac{\pi}{2n}\right)\cos\left(\frac{\pi}{2n}\right) \le 4\sin\left(\frac{\pi}{2n}\right) = f(n).$$

Moreover, we also can see that the proportion $\frac{f(n)}{\varepsilon_n} \to 1$ as $n \to \infty$. This shows how sharp the number $f(n)$ is when we approximate the spectrum using large matrices.

For this example, plots of the inclusion sets in Corollary 3.15 for $n = 8, 16, 32$ and 64, are shown in Fig. 3.2.

3.4.3 3-periodic Bi-diagonal Operator

In this example, we are considering the bidiagonal operator

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & -1.5 & 1 & & & & \\ & 0 & 1 & 2 & & & \\ & & 0 & 1 & 1 & & \\ & & & 0 & -1.5 & 1 & \\ & & & & 0 & 1 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}.$$

Since A is not normal, we can not apply Theorem 2.20 to this operator. However, Theorem 3.15 tells us that

$$\mathrm{spec}_{\varepsilon} A \subseteq \bigcup_{k\in\mathbb{Z}} \mathrm{spec}_{\varepsilon+f(n)} \hat{A}_{n,k}.$$

From the numerical results, Figure 3.4, it seems like the inclusion sets are converging to the spectrum of A as $n \to \infty$.

3.5 The conjecture on the convergence of our inclusion sets we got from method 1 and 1*?

In the examples we have considered, we can notice that when we apply method 1 to the infinite operator A, the inclusion sets $\overline{\Sigma^n_{f(n)}(A)}$ are simply connected sets. Note that, if a sequence of the inclusion sets of the infinite tridiagonal operator A converges to the spectrum as n tends to infinity, that means the spectrum has also to be a simply connected set. Therefore, the inclusion sets for the spectra of infinite dimensional operators using method 1 do not converge to the spectrum in general.

In section 3.2.1 and 3.4.1, we have shown analytically that method 1 converges to the polynomial convex hull of spec A, $\widehat{\mathrm{spec}\, A}$ (see Section 2.1), and method 1* converges to $\operatorname{spec} A$ when A is the right shift operator. Moreover, we can see that the convergence of method 1 to $\widehat{\operatorname{spec} A}$ is faster than the convergence of method 1* to $\operatorname{spec} A$.

These numerical results suggest some conjectures. The first two relate to the convergence of the inclusion sets for the spectra of bi-infinite tri-diagonal operators as follows:

1. The inclusion sets for $\operatorname{spec} A$ using method 1 converges to the polynomial convex hull of $\operatorname{spec} A$ and the inclusion sets for $\operatorname{spec} A$ using method 1* converge to $\operatorname{spec} A$, precisely

   $$d_H\left(\overline{\Sigma^n_{f(n)}(A)}, \widehat{\operatorname{spec} A}\right) \to 0 \text{ and } d_H\left(\overline{\Pi^n_{f(n)}(A)}, \operatorname{spec} A\right) \to 0 \text{ as } n \to \infty$$

2. The Haussdorff distance between the inclusion sets of A using method 1 and the polynomial convex hull of $\operatorname{spec} A$ is less than the Haussdorff distance between the inclusion sets of A using method 1* and the spectrum of A, i.e. $d_H\left(\overline{\Sigma^n_{f(n)}(A)}, \widehat{\operatorname{spec} A}\right) \leq d_H\left(\overline{\Pi^n_{f(n)}(A)}, \operatorname{spec} A\right)$.

Although we believe that method 1* converges to $\operatorname{spec} A$, in some case method 1 converge to $\operatorname{spec} A$ even faster than method 1*. Comparing the plots in Figure 3.5 and Figure 3.6 when we apply each method to the operator

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & 1 & & & & \\ & -1 & 0 & 1 & & & \\ & & 1 & 0 & 1 & & \\ & & & 1 & 0 & 1 & \\ & & & & -1 & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix},$$

we see that the inclusion sets $\overline{\Sigma^n_{f(n)}(A)}$ converge to $\operatorname{spec} A$ faster than $\overline{\Pi^n_{f(n)}(A)}$.

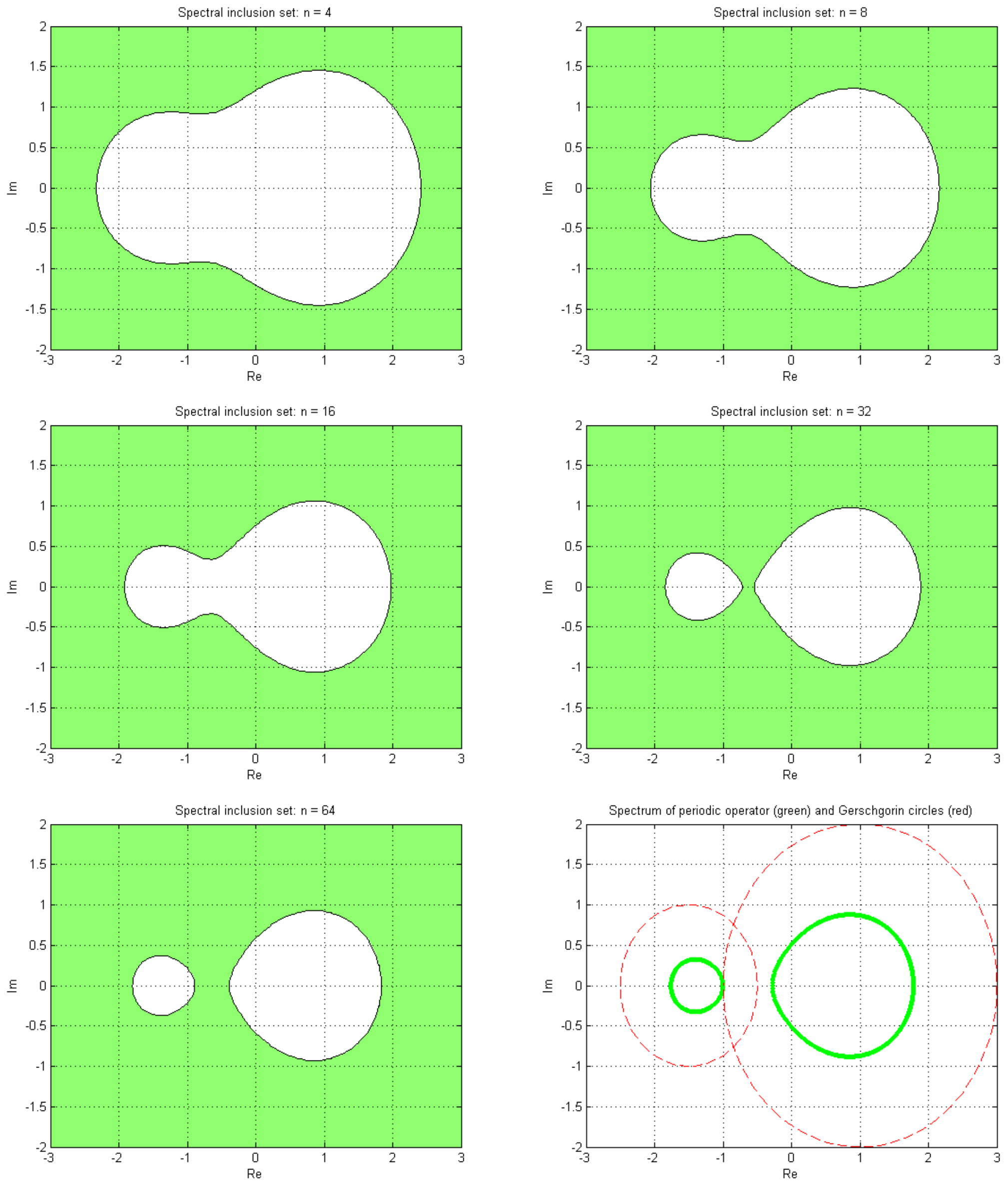

Figure 3.3: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}$, which are inclusion sets for spec A, where $f(n)$ is given as in Corollary 3.6, where A is the periodic bi-diagonal operator which has $\gamma_i = 1$ for all $i \in \mathbb{Z}$ and $(\beta_i) = (\ldots, 1, -1.5, 1, 1, -1.5, 1, \ldots)$. Shown are the inclusion sets when $n = 4, 8, 16, 32$ and 64. The last picture shows the spectrum of the operator A and also the Gershgorin circles (dashed line), defined in Theorem 2.50.

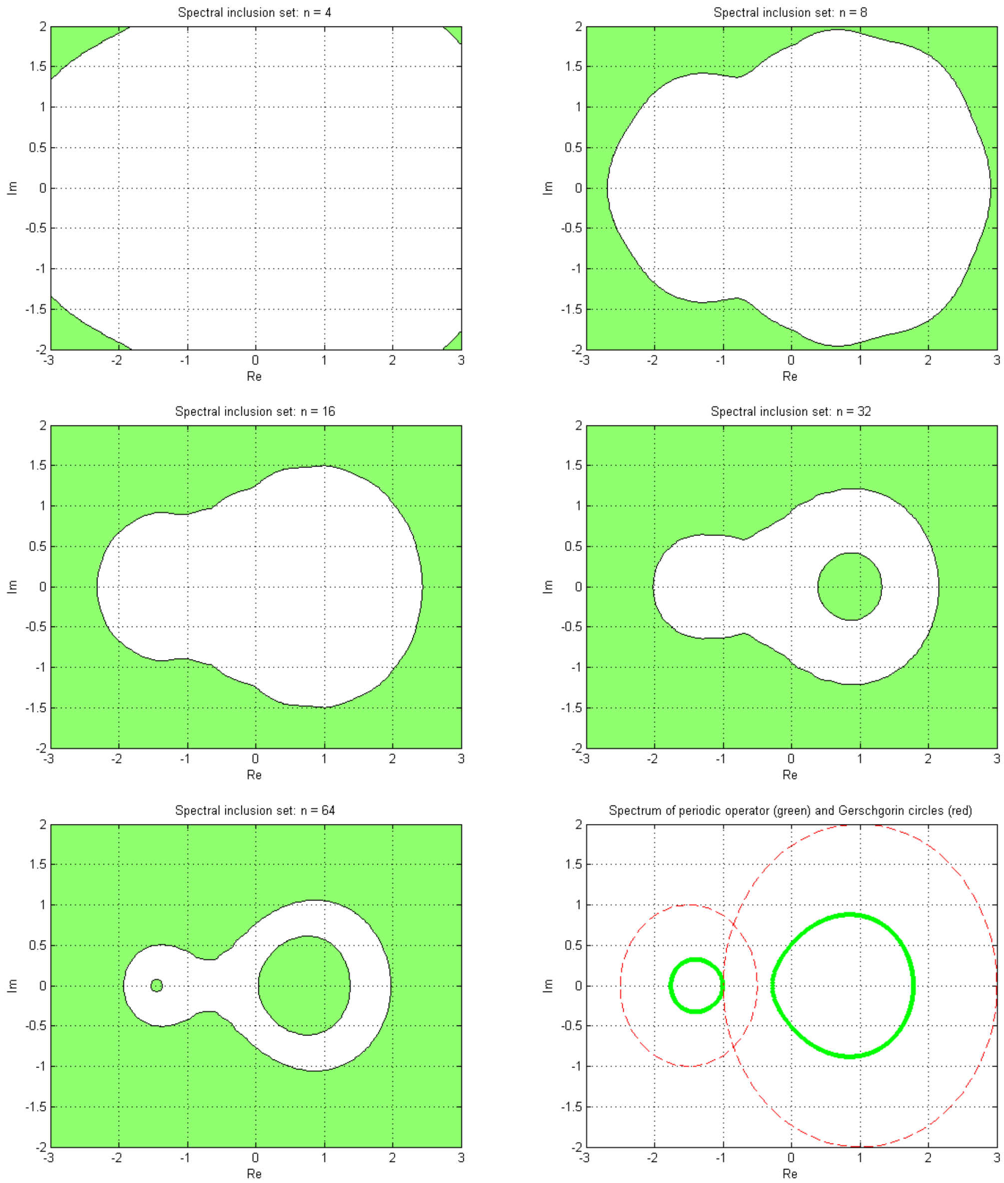

Figure 3.4: Plots of the sets $\overline{\Pi^n_{f(n)}(A)}$, which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 3.15, where A is the periodic bi-diagonal operator which has $\gamma_i = 1$ for all $i \in \mathbb{Z}$ and $(\beta_i) = (\ldots, 1, -1.5, 1, 1, 2, 1, \ldots)$. Shown are the inclusion sets when $n = 4, 8, 16, 32$ and 64. The last picture shows the spectrum of the operator A and also the Gershgorin circles (dashed line), defined in Theorem 2.50.

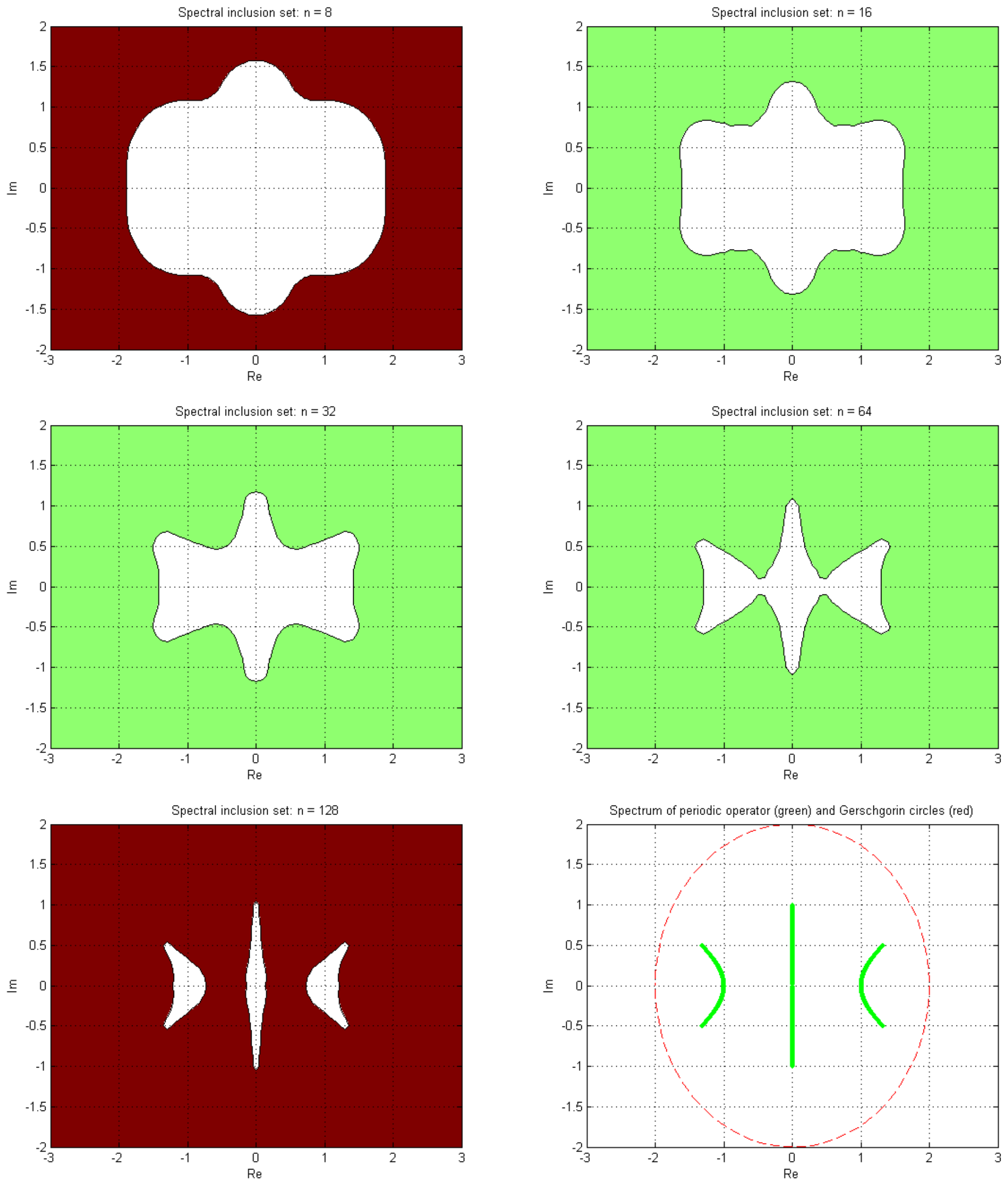

Figure 3.5: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}$, which are inclusion sets for spec A, where $f(n)$ is given as in Corollary 3.6, where A is the periodic tri-diagonal operator which has $\gamma_i = 1$ and $\beta_i = 0$ for all i and $(\alpha_i) = (\ldots, 1, -1, 1, 1, -1, 1, \ldots)$. Shown are the inclusion sets when $n = 4, 8, 16, 32$ and 64. The last picture shows the spectrum of the operator A and also the Gershgorin circles (dashed line), defined in Theorem 2.50.

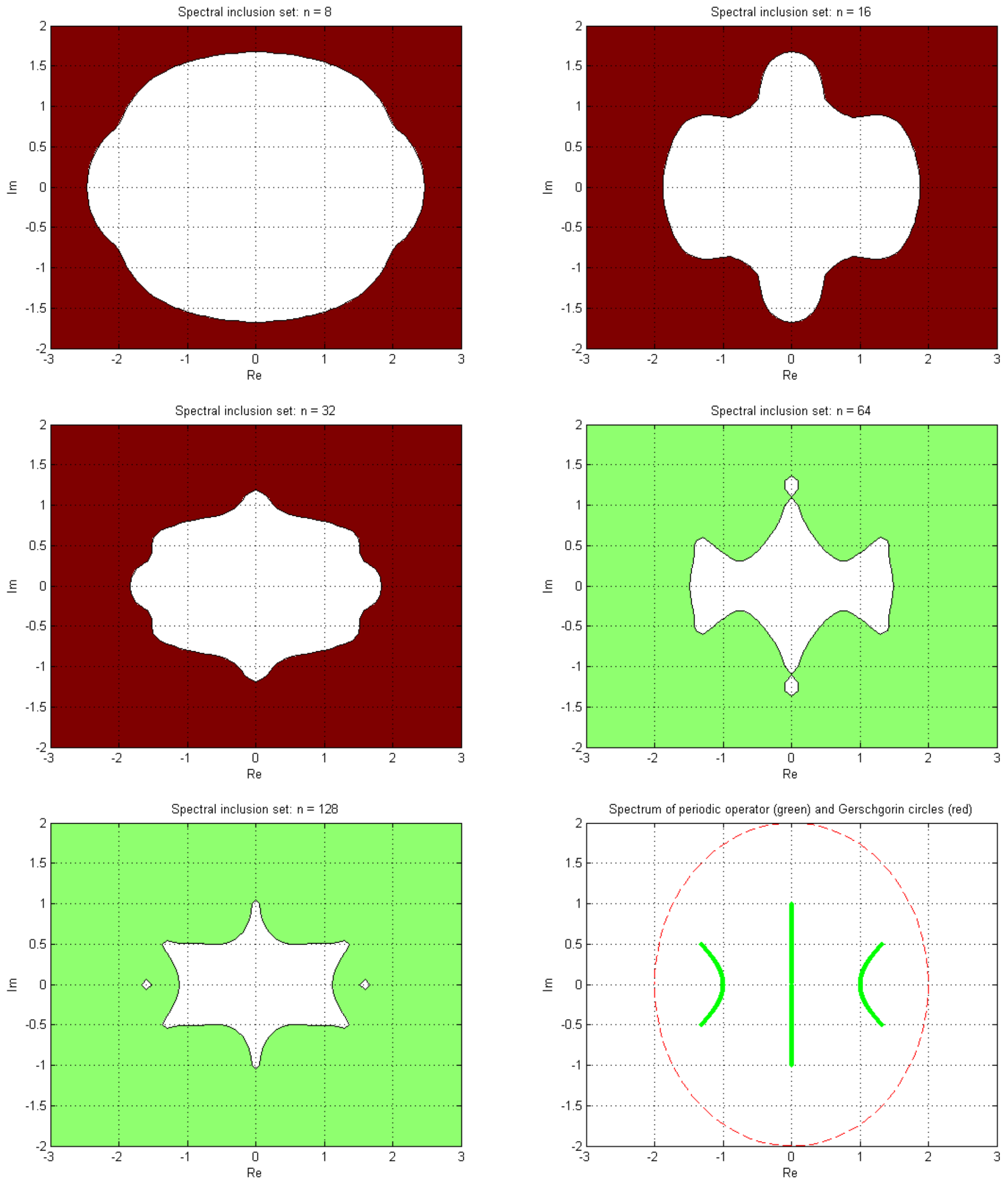

Figure 3.6: Plots of the sets $\overline{\Pi^n_{f(n)}(A)}$, which are inclusion sets for spec A, where $f(n)$ is given as in Corollary 3.15 when $\hat{A}_{n,k}$ is the periodised finite submatrix given by (3.56) of the periodic tri-diagonal operator which has $\gamma_i = 1$ and $\beta_i = 0$ for all i and $(\alpha_i) = (\ldots, 1, -1, 1, 1, -1, 1, \ldots)$. Shown are the inclusion sets when $n = 4, 8, 16, 32$ and 64. The last picture shows the spectrum of the operator A and also the Gershgorin circles (dashed line), defined in Theorem 2.50.

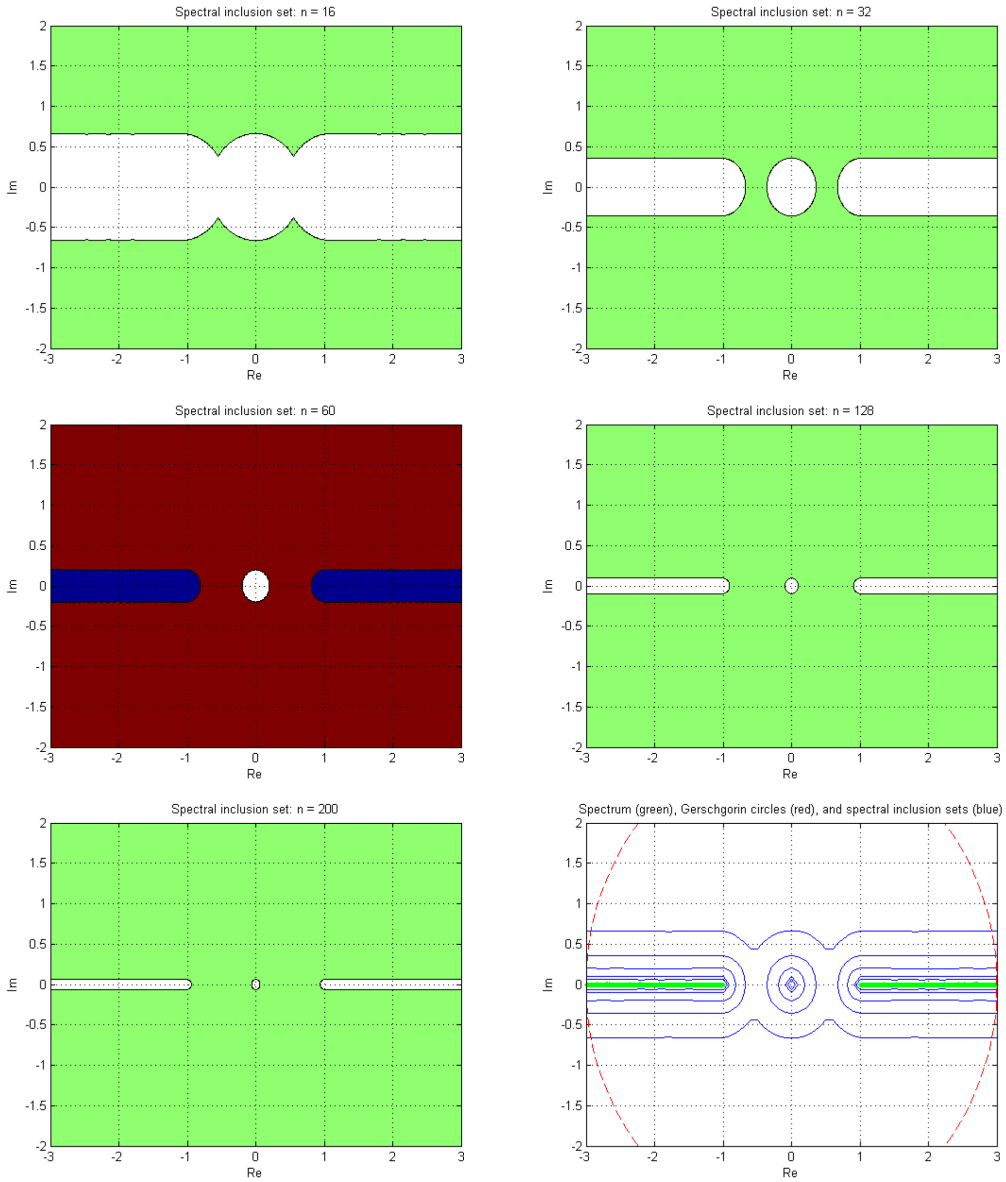

Figure 3.7: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}$, which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 3.15, where A is the periodic tridiagonal operator which has $\gamma_i = 1$ and $\beta_i = 0$ for all i and $(\alpha_i) = (\ldots, 1, -1, 1, 1, -1, 1, \ldots)$, with $\operatorname{spec} A = [-3, -1] \cup [1, 3]$. Shown are the inclusion sets when $n = 4, 8, 16, 32$ and 64. The last picture shows the spectrum of the operator A and also the Gershgorin circles (dashed line), defined in Theorem 2.50.

Chapter 4

A One-sided truncation Method for Approximating the Spectrum and Pseudospectrum of Infinite Tridiagonal Matrices

In this chapter, we will try to improve our upper bounds of the spectrum of tri-diagonal matrices. While the "finite section matrix" method involves the smallest singular value of the two-sided truncation $P_{n,k}\,(A - \lambda I)\,P_{n,k}$, where $P_{n,k}$ is the operator of multiplication by $\chi^{(n,k)}$, the method to be discussed now involves the smallest of all singular values of the two one-sided truncations $(A - \lambda I)\,P_{n,k}$ and $P_{n,k}\,(A - \lambda I)$, where, because of the tridiagonal structure of A, these can be identified with $(n+2) \times n$ instead of $\infty \times n$ matrices. The fact that truncation is only performed from one side should give a more accurate picture of the operator and its spectrum.

4.1 Inclusion sets in terms of one-sided truncation matrices

In Chapter 3, we have studied inclusion sets for the spectra and pseudospectra of the matrix

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & \beta_{-2} & \gamma_{-1} & & & & \\ & \alpha_{-2} & \beta_{-1} & \gamma_0 & & & \\ & & \alpha_{-1} & \boxed{\beta_0} & \gamma_1 & & \\ & & & \alpha_0 & \beta_1 & \gamma_1 & \\ & & & & \alpha_1 & \beta_2 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix},$$

using two different methods. The first method is a naive way to approximate the spectra and pseudospectra of any infinite matrix using the standard $n \times n$ finite sections,

$$A_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}.$$

(See Figure 4.1.) We have shown that

$$\operatorname{spec}_{\varepsilon} A \subseteq \Sigma^n_{\varepsilon+f(n)}(A)$$

and

$$\operatorname{spec} A \subseteq \overline{\Sigma^n_{f(n)}(A)},$$

for every $\varepsilon > 0$ and every $n \in \mathbb{N}$, where $f(n) < \frac{\pi}{n}\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)$.

There are some disadvantages of using the principal finite sections, so the second method has been introduced. It is a method to compute the

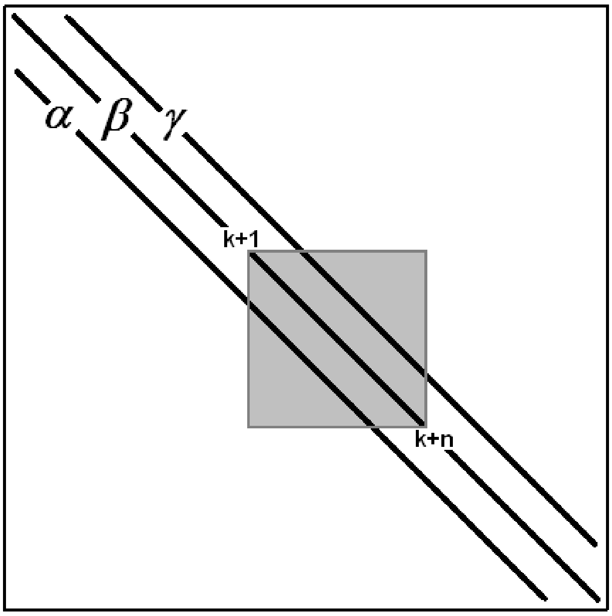

Figure 4.1: The idea of the first method.

spectra and pseudospectra of an infinite tridiagonal matrix using periodised submatrices,

$$\hat{A}_{n,k} = \begin{pmatrix} \beta_{k+1} & \gamma_{k+2} & & & \alpha_k \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & \\ & \ddots & \ddots & \ddots & \\ & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ \gamma_{k+n+1} & & & \alpha_{k+n-1} & \beta_{k+n} \end{pmatrix}.$$

This method gives the following results,

$$\operatorname{spec}_{\varepsilon} A \subseteq \Pi^n_{\varepsilon+f(n)}(A)$$

and

$$\operatorname{spec} A \subseteq \overline{\Pi^n_{f(n)}(A)},$$

where $f(n) < \frac{\pi}{n}(\|\alpha\|_\infty + \|\gamma\|_\infty)$.

In this chapter, we are going to introduce a new idea to compute the inclusion sets for the spectra and pseudospectra of A using one-sided finite truncations. Let $(X_{n,k})$ denote the range of $M_{\chi^{(n,k)}}$, which is an n-dimensional

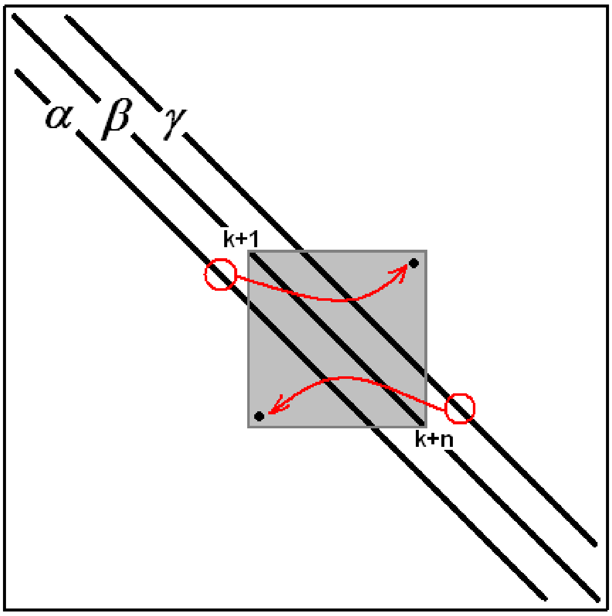

Figure 4.2: The idea of the second method.

subspace of $\ell^2(\mathbb{Z})$. Now, recalling that A is the tridiagonal operator given by (3.1), let

$$B^+_{n,k}, B^-_{n,k} : X_{n,k} \to X_{n,k}$$

be defined by

$$B^+_{n,k} := M_{\chi^{(n,k)}}(A-\lambda I)^*(A-\lambda I)M_{\chi^{(n,k)}}\big|_{X_{n,k}}$$

and

$$B^-_{n,k} := M_{\chi^{(n,k)}}(A-\lambda I)(A-\lambda I)^*M_{\chi^{(n,k)}}\big|_{X_{n,k}},$$

respectively, and let

$$C_{n,k} : X_{n,k} \to \ell^2(\mathbb{Z})$$

be defined by

$$C_{n,k} = (A-\lambda I)M_{\chi^{(n,k)}}\big|_{X_{n,k}} = (A-\lambda I)\big|_{X_{n,k}}.$$

Then

$$C^*_{n,k} = M_{\chi^{(n,k)}}(A-\lambda I)^*,$$

and

$$B^{+}_{n,k} = C^{*}_{n,k} C_{n,k}$$

and, from Theorem 2.24,

$$\begin{aligned}\min \operatorname{spec}(B^{+}_{n,k}) = \min \operatorname{spec}(C^{*}_{n,k} C_{n,k}) &= \inf_{\substack{\|\phi\|=1 \\ \phi\in X_{n,k}}} \left(C^{*}_{n,k} C_{n,k}\phi, \phi\right) \\ &= \inf_{\substack{\|\phi\|=1 \\ \phi\in X_{n,k}}} (C_{n,k}\phi, C_{n,k}\phi) = \inf_{\substack{\|\phi\|=1 \\ \phi\in X_{n,k}}} \|C_{n,k}\phi\|^2 = \nu\left(C_{n,k}\right)^2 .\end{aligned}$$

Therefore

$$\min \operatorname{spec}(B^{+}_{n,k}) = \nu\left((A-\lambda I)\big|_{X_{n,k}}\right)^2 \geq \nu\left(A-\lambda I\right)^2 . \tag{4.1}$$

Similary,

$$\min \operatorname{spec}(B^{-}_{n,k}) = \nu\left((A-\lambda I)^{*}\big|_{X_{n,k}}\right)^2 \geq \nu\left((A-\lambda I)^{*}\right)^2 . \tag{4.2}$$

Recall that, for the lower norm

$$\nu(B) = \inf_{\|x\|=1} \|Bx\| ,$$

of an operator B on a Hilbert space X., it holds, by Proposition 2.15 and Theorem 2.16,

$$\begin{aligned}\nu(B) > 0 &\Leftrightarrow B \text{ is injective and the image of } B \text{ is closed,} \\ \nu(B) > 0 \text{ and } \nu(B^{*}) > 0 &\Leftrightarrow B \text{ is invertible.}\end{aligned}$$

The latter motivates us to define

$$\begin{aligned}\xi(B) &:= \min(\nu(B), \nu(B^{*})) \\ &= \min(\nu(B^{*}B), \nu(BB^{*}))^{\frac{1}{2}} \\ &= \min(\operatorname{spec}(B^{*}B) \cup \operatorname{spec}(BB^{*}))^{\frac{1}{2}} \\ &= \min(s_{\min}(B), s_{\min}(B^{*})),\end{aligned}$$

so that one has

$$\operatorname{spec}(B) = \{\lambda \in \mathbb{C} : \xi(B-\lambda I) = 0\}, \tag{4.3}$$

and

$$\operatorname{spec}_{\varepsilon}(B) = \{\lambda \in \mathbb{C} : \xi(B - \lambda I) < \varepsilon\}, \tag{4.4}$$

by Theorem 2.26.

We now put, for $k \in \mathbb{Z}, n \in \mathbb{N}$ and $B \in B(\ell^2(\mathbb{Z}))$,

$$\begin{aligned}\xi_{n,k}(B) &:= \min\left(\nu\left(B|_{X_{n,k}}\right), \nu\left(B^*|_{X_{n,k}}\right)\right)\\ &= \min\left(\nu\left(BM_{\chi^{(n,k)}}|_{X_{n,k}}\right), \nu\left(B^*M_{\chi^{(n,k)}}|_{X_{n,k}}\right)\right)\\ &= \min\left(\operatorname{spec}\left(M_{\chi^{(n,k)}}B^*BM_{\chi^{(n,k)}}|_{X_{n,k}}\right) \cup \operatorname{spec}\left(M_{\chi^{(n,k)}}BB^*M_{\chi^{(n,k)}}|_{X_{n,k}}\right)\right)^{\frac{1}{2}}\end{aligned}$$

and

$$\xi_n(B) := \inf_{k\in\mathbb{Z}} \xi_{n,k}(B).$$

From $\nu(B) \leq \nu\left(B|_{X_{n,k}}\right)$ and $\nu(B^*) \leq \nu\left(B^*|_{X_{n,k}}\right)$ it follows that,

$$\xi(B) \leq \xi_{n,k}(B) \text{ for all } n, k, \text{ so that } \xi(B) \leq \xi_n(B). \tag{4.5}$$

Recalling (4.3) and (4.4), we introduce the following sets:

Definition 4.1. *For $\varepsilon > 0, k \in \mathbb{Z}, n \in \mathbb{N}$, and $A \in B(\ell^2(\mathbb{Z}))$, we put*

$$\gamma_{\varepsilon}^{n,k}(A) := \{\lambda \in \mathbb{C} : \xi_{n,k}(A - \lambda I) < \varepsilon\},$$

and

$$\Gamma_{\varepsilon}^{n}(A) := \{\lambda \in \mathbb{C} : \xi_{n}(A - \lambda I) < \varepsilon\}.$$

Explicitly, this means that

$$\begin{aligned}\gamma_{\varepsilon}^{n,k}(A) &= \{\lambda \in \mathbb{C} : \min\left(\nu\left((A-\lambda I)|_{X_{n,k}}\right), \nu\left((A-\lambda I)^*|_{X_{n,k}}\right)\right) < \varepsilon\}\\ &= \Big\{\lambda \in \mathbb{C} : \min\left(\operatorname{spec}\left(M_{\chi^{(n,k)}}(A-\lambda I)^*(A-\lambda I)M_{\chi^{(n,k)}}|_{X_{n,k}}\right)\right.\\ &\qquad \left.\cup \operatorname{spec}\left(M_{\chi^{(n,k)}}(A-\lambda)(A-\lambda)^*M_{\chi^{(n,k)}}|_{X_{n,k}}\right)\right) < \varepsilon^2\Big\}\end{aligned} \tag{4.6}$$

and $\Gamma_{\varepsilon}^{n}(A) = \bigcup_{k\in\mathbb{Z}} \gamma_{\varepsilon}^{n,k}(A)$.

It is interesting to compare (4.6) to the corresponding sets (pseudospectrum of two-sided truncations) used in our first method. These are

$$\begin{aligned}\operatorname{spec}_{\varepsilon} A_{n,k} &= \{\lambda \in \mathbb{C} : \min\operatorname{spec}\left(M_{\chi^{(n,k)}}(A-\lambda I)^*M_{\chi^{(n,k)}}(A-\lambda I)M_{\chi^{(n,k)}}\right) < \varepsilon^2\}\\ &= \{\lambda \in \mathbb{C} : \min\operatorname{spec}\left(M_{\chi^{(n,k)}}(A-\lambda I)M_{\chi^{(n,k)}}(A-\lambda I)^*M_{\chi^{(n,k)}}\right) < \varepsilon^2\}.\end{aligned}$$

From (4.4), the definition of $\Gamma^n_\varepsilon(A)$ and (4.5), applied to $B = A - \lambda I$, we immediately get the following lower bound on $\operatorname{spec}_\varepsilon A$:

Theorem 4.2. *If $\varepsilon > 0, n \in \mathbb{N}$ and $A \in B(\ell^2(\mathbb{Z}))$ then*

$$\Gamma^n_\varepsilon(A) \subseteq \operatorname{spec}_\varepsilon A. \tag{4.7}$$

With very much similarity to Theorem 3.2 and Theorem 3.13, we can accompany the lower bound (4.7) on $\operatorname{spec}_\varepsilon A$ by the following upper bound.

Theorem 4.3. *If $\varepsilon > 0$, $n \in \mathbb{N}$, $w_j > 0$, for $j = 1, \dots, n$ and $w_0 = w_{n+1} = 0$, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Gamma^n_{\varepsilon+f(n)}(A),$$

where

$$f(n) = (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\frac{T_n}{S_n}}$$

with $S_n = \sum_{i=1}^{n} w_i^2$ and $T_n = w_1^2 + w_n^2 + \sum_{i=1}^{n-1} (w_{i+1} - w_i)^2$.

Proof. Let $\lambda \in \operatorname{spec}_\varepsilon(A)$. Then either there exists $x \in \ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A - \lambda I)x\| < \varepsilon$, or the same holds with $A - \lambda I$ replaced by its adjoint. In the first case, let $y = (A - \lambda I)x$, so $\|y\| < \varepsilon$. For $i, k \in \mathbb{Z}$, define $e_i^{(k)}, E^+_{i,k}$ and $E^-_{i,k}$ as in the proof of Theorem 3.2.

For $k \in \mathbb{Z}$, let

$$P_k := \|(A - \lambda I)M_{e^{(k)}}x\|.$$

Let $Q_k := \|M_{e^{(k)}}x\|$. We will prove that $P_k < (\varepsilon + f(n))Q_k$, for some $k \in \mathbb{Z}$, which will show that $\nu\left((A - \lambda I)\big|_{X_{n,k}}\right) < \varepsilon + f(n)$, so that $\lambda \in \gamma^{n,k}_{\varepsilon+f(n)}(A) \subseteq \Gamma^n_{\varepsilon+f(n)}(A)$.

Note first that, using (3.6) and (3.7),

$$\begin{aligned}
P_k^2 &= \left|\gamma_{k+1}e_{k+1}^{(k)}x_{k+1}\right|^2 + \left|y_{k+1}e_{k+1}^{(k)} - \alpha_k x_k e_{k+1}^{(k)} + \gamma_{k+2}(e_{k+1}^{(k)} - e_{k+1}^{(k)})x_{k+2}\right|^2 \\
&+ \sum_{i=k+2}^{k+n-1}\left|y_i e_i^{(k)} + \alpha_{i-1}(e_{i-1}^{(k)} - e_i^{(k)})x_{i-1} + \gamma_{i+1}(e_{i+1}^{(k)} - e_i^{(k)})x_{i+1}\right|^2 \\
&+ \left|y_{k+n}w_n + \alpha_{k+n-1}(e_{k+n-1}^{(k)} - e_{k+n}^{(k)})x_{k+n-1} - \gamma_{k+n+1}x_{k+n+1}e_{k+n}^{(k)}\right|^2 + \left|\alpha_{k+n}e_{k+n}^{(k)}x_{k+n}\right|^2 \\
&\le \sum_{i=k+1}^{k+n}\left(|y_i|\, e_i^{(k)} + |\alpha_{i-1}|\, E_{i,k}^{-}\, |x_{i-1}| + |\gamma_{i+1}|\, E_{i,k}^{+}\, |x_{i+1}|\right)^2 \\
&+ |\gamma_{k+1}|^2 \left(E_{k,k}^{+}\right)^2 |x_{k+1}|^2 + |\alpha_{k+n}|^2 \left(E_{k+n+1,k}^{-}\right)^2 |x_{k+n}|^2 .
\end{aligned}$$

So, for all $\theta > 0$ and $\phi > 0$, by Lemma 3.1,

$$\begin{aligned}
P_k^2 &\le \sum_{i=k+1}^{k+n}\left[(1+\theta)\left(|y_i|\, e_i^{(k)}\right)^2 + (1+\theta^{-1})\left(|\alpha_{i-1}|\, E_{i,k}^{-}\, |x_{i-1}| + |\gamma_{i+1}|\, E_{i,k}^{+}\, |x_{i+1}|\right)^2\right] \\
&+ |\gamma_{k+1}|^2 \left(E_{k,k}^{+}\right)^2 |x_{k+1}|^2 + |\alpha_{k+n}|^2 \left(E_{k+n+1,k}^{-}\right)^2 |x_{k+n}|^2 \\
&\le \sum_{i=k}^{k+n+1}\left[(1+\theta)\left(|y_i|\, e_i^{(k)}\right)^2 + (1+\theta^{-1})\left((1+\phi)\, |\alpha_{i-1}|^2 \left(E_{i,k}^{-}\right)^2 |x_{i-1}|^2\right.\right. \\
&\quad \left.\left. + \ (1+\phi^{-1})\, |\gamma_{i+1}|^2 \left(E_{i,k}^{+}\right)^2 |x_{i+1}|^2\right)\right] + |\gamma_{k+1}|^2 \left(E_{k,k}^{+}\right)^2 |x_{k+1}|^2 + |\alpha_{k+n}|^2 \left(E_{k+n+1,k}^{-}\right)^2 |x_{k+n}|^2 .
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} P_k^2 &\le (1+\theta)\sum_{i\in\mathbb{Z}}|y_i|^2 \sum_{k\in\mathbb{Z}}\left(e_i^{(k)}\right)^2 \\
&\quad + (1+\theta^{-1})\left[(1+\phi)\,\|\alpha\|_\infty^2\left\{\left(e_{k+1}^{(k)}\right)^2 + \left(E_{k+2,k}^{-}\right)^2 + \cdots + \left(E_{k+n+1,k}^{-}\right)^2 + \left(e_{k+n}^{(k)}\right)^2\right\}\right. \\
&\quad \left. + \ (1+\phi^{-1})\,\|\gamma\|_\infty^2\,\{\left(e_{k+1}^{(k)}\right)^2 + \left(E_{k+1,k}^{+}\right)^2 + \left(E_{k+2,k}^{+}\right)^2 + \cdots + \left(E_{k+n-1,k}^{+}\right)^2 + \left(e_{k+n}^{(k)}\right)^2\}\right] \\
&\quad \le (1+\theta)\,\|y\|_2^2\, S_n + (1+\theta^{-1})\left(\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)\sqrt{T_n}\right)^2 .
\end{aligned}$$

Moreover, $\sum_{k\in\mathbb{Z}} Q_k^2 = \sum_{k\in\mathbb{Z}}\sum_{i=1}^{n} w_i^2 \left|x_{i+k}\right|^2 = S_n$. Thus, and since $\|y\| < \varepsilon$,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} P_k^2 &< \left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})\left(\left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)\sqrt{T_n}\right)^2 \frac{1}{S_n}\right] S_n \\ &\le \left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})\left[f(n)\right]^2\right]\sum_{k\in\mathbb{Z}} Q_k^2.\end{aligned}$$

Applying Lemma 3.1 again, we see that

$$\inf_{\theta>0}\left[(1+\theta)\varepsilon^2 + (1+\theta^{-1})[f(n)]^2\right] = (\varepsilon + f(n))^2,$$

so that

$$\sum_{k\in\mathbb{Z}} P_k^2 < (\varepsilon + f(n))^2 \sum_{k\in\mathbb{Z}} Q_k^2.$$

Thus, for some $k \in \mathbb{Z}$,

$$P_k < (\varepsilon + f(n))\, Q_k.$$

In the case that there exists $x \in \ell^2(\mathbb{Z})$ with $\|x\| = 1$ and $\|(A-\lambda I)^* x\| < \varepsilon$, we can show similarly that, for some $k \in \mathbb{Z}$, $\nu\left((A-\lambda I)^*\big|_{X_{n,k}}\right) < \varepsilon + f(n)$, so that $\lambda \in \gamma^{n,k}_{\varepsilon+f(n)}(A) \subseteq \Gamma^n_{\varepsilon+f(n)}(A)$. Thus the result is proved. ∎

In order to minimise $f(n) = \left(\|\alpha\|_\infty + \|\gamma\|_\infty\right)\sqrt{\frac{T_n}{S_n}}$, we need to minimise $\frac{T_n}{S_n}$ where

$$\begin{aligned} T_n &= w_1^2 + (w_2 - w_1)^2 + \cdots + (w_n - w_{n-1})^2 + w_n^2, \\ S_n &= w_1^2 + w_2^2 + \cdots + w_n^2, \end{aligned}$$

i.e. we need to minimise

$$\sqrt{\frac{T_n}{S_n}} = \frac{\|Bw\|}{\|w\|}$$

where

$$B = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & -1 & 1 & \\ & & & -1 & 1 \\ & & & & -1 \end{pmatrix}_{(n+1)\times n},$$

and $w = (w_1, w_2, \ldots, w_n)^T$.

Note that

$$\inf_{\|w\|\neq 0} \frac{T_n}{S_n} = \inf_{\|w\|\neq 0} \frac{\|Bw\|^2}{\|w\|^2} = \inf_{\|w\|\neq 0} \frac{(Bw, Bw)_{\mathbb{R}^{n+1}}}{(w, w)_{\mathbb{R}^n}}$$
$$= \inf_{\|w\|\neq 0} \frac{(B^* Bw, w)_{\mathbb{R}^{n+1}}}{(w, w)_{\mathbb{R}^n}} = \lambda_{min}\left(B^T B\right),$$

where

$$B^T B = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}_{n\times n}.$$

We know that λ is an eigenvalue of $B^T B$ with eigenvector $v = (v_1, v_2, \ldots, v_n)^T$ iff $B^T Bv = \lambda v$. i.e.

$$(2-\lambda)v_1 - v_2 = 0, \tag{4.8}$$
$$-v_{i-1} + (2-\lambda)v_i - v_{i+1} = 0 \qquad i = 2, \ldots, n-1, \tag{4.9}$$
$$-v_{n-1} + (2-\lambda)v_n = 0. \tag{4.10}$$

Set $v_0 = v_{n+1} = 0$, then the above equations can be written as

$$-v_{i-1} + (2-\lambda)v_i - v_{i+1} = 0,$$

for $i = 1, \ldots, n$. As in the discussion before Corollary 3.6, we see that the general solution is

$$v_j = C\cos(j\theta) + D\sin(j\theta), \quad j = 1, \ldots, n+1,$$

with $\theta \in (0, \pi)$ and $2\cos\theta = 2 - \lambda$. Let $D = 1$. Since $v_0 = 0$, it follows that $C = 0$, i.e.

$$v_j = \sin(j\theta), \quad j = 1, \ldots, n+1.$$

Since $v_{n+1} = 0$ it follows that $(n+1)\theta = k\pi$ for some $k \in \mathbb{N}$, so that $\theta_k = \dfrac{k\pi}{n+1}$, $\quad k = 1, \ldots, n$. From $2 - \lambda_k = 2\cos\theta_k$ we conclude that

$$\lambda_k = 2(1-\cos\theta_k) = 4\sin^2\frac{\theta_k}{2} = 4\sin^2\frac{k\pi}{2(n+1)}, \quad k = 1, \ldots, n.$$

It follows that $\lambda_{\min} = \lambda_1$ and hence the minimal value for $f(n)$ is

$$f(n) = (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\frac{T_n}{S_n}} = (\|\alpha\|_\infty + \|\gamma\|_\infty)\sqrt{\lambda_1}$$
$$= (\|\alpha\|_\infty + \|\gamma\|_\infty)\, 2\sin\frac{\pi}{2(n+1)}.$$

This minimum is realised by the choice $w = \left(\sin\frac{j\pi}{n+1}\right)_{j=1}^{n}$ for the weight vector w in Theorem 4.3.

Corollary 4.4. *If* $\varepsilon > 0$, $n \in \mathbb{N}$ *, then*

$$\operatorname{spec}_\varepsilon A \subseteq \Gamma^n_{\varepsilon+f(n)}(A),$$

and

$$\operatorname{spec} A \subseteq \overline{\Gamma^n_{f(n)}(A)}.$$

where

$$f(n) = 2\sin\left(\frac{\pi}{2(n+1)}\right)(\|\alpha\|_\infty + \|\gamma\|_\infty). \qquad (4.11)$$

4.2 How to implement a program to approximate the spectrum of an infinite matrix A?

We are considering in this section how to compute $\Gamma^n_\varepsilon(A)$. Let $\left(e^{(j)}\right)_{j\in\mathbb{Z}}$ be the canonical basis for $\ell^2(\mathbb{Z})$, i.e.,

$$e^{(j)}_i = \delta_{ij}, \qquad i \in \mathbb{Z}$$

Note that, for $x \in \ell^2(\mathbb{Z})$.

$$(Ax)_m = \sum_{j\in\mathbb{Z}} a_{mj}x_j,$$

where $a_{mj} := \alpha_{j-1}\delta_{m,j-1}+\beta_j\delta_{m,j}+\gamma_{j+1}\delta_{m,j+1}$, $m, j \in \mathbb{Z}$. For $\phi = \sum_{j=k+1}^{k+n} \phi_j e^{(j)} \in X_{n,k}$, the action of $A\left|_{X_{n,k}} : X_{n,k} \to \ell^2(\mathbb{Z})\right.$ is given by

$$\left(A\left|_{X_{n,k}}\phi\right.\right)_m = \sum_{j=k+1}^{k+n} a_{mj}\phi_j, \qquad m \in \mathbb{Z},$$

and, for $\phi \in X_{n,k}$,

$$\left((A-\lambda I)\left|_{X_{n,k}}\ \phi\right.\right)_m = \sum_{j=k+1}^{k+n} (a_{mj} - \lambda\delta_{mj})\,\phi_j, \qquad m \in \mathbb{Z}.$$

Further, for $\psi \in \ell^2(\mathbb{Z})$,

$$\left(M_{\chi^{(n,k)}}(A-\lambda I)^*\psi\right)_m = \sum_{j\in\mathbb{Z}} \left(\overline{a_{jm}} - \overline{\lambda}\delta_{jm}\right)\psi_j, \quad m = k+1,\ldots,k+n.$$

Thus, for $\phi \in X_{n,k}$, and $m = k+1,\ldots,k+n$,

$$\begin{aligned}
(B_{n,k}^{+}\phi)_m &= \left(M_{\chi^{(n,k)}}(A-\lambda I)^*(A-\lambda I)M_{\chi^{(n,k)}}\phi\right)_m \\
&= \sum_{i\in\mathbb{Z}}\left(\overline{a_{im}} - \overline{\lambda}\delta_{im}\right)\left(\sum_{j=k+1}^{k+n}(a_{ij}-\lambda\delta_{ij})\,\phi_j\right) \\
&= \sum_{i\in\mathbb{Z}}\sum_{j=k+1}^{k+n}\left(\overline{a_{im}} - \overline{\lambda}\delta_{im}\right)\left((a_{ij}-\lambda\delta_{ij})\,\phi_j\right) \\
&= \sum_{i\in\mathbb{Z}}\sum_{j=k+1}^{k+n}\left(|\lambda|^2\delta_{im}\delta_{ij} - \overline{\lambda}\delta_{im}a_{ij} - \lambda\overline{a_{jm}}\delta_{ij} + \overline{a_{im}}a_{ij}\right)\phi_j \\
&= \sum_{j=k+1}^{k+n}\phi_j\sum_{i\in\mathbb{Z}}\left(|\lambda|^2\delta_{im}\delta_{ij} - \overline{\lambda}\delta_{im}a_{ij} - \lambda\overline{a_{im}}\delta_{ij} + \overline{a_{im}}a_{ij}\right) \\
&= \sum_{j=k+1}^{k+n}\left(|\lambda|^2\delta_{mj} - \overline{\lambda}a_{mj} - \lambda\overline{a_{jm}} + \sum_{i=m-1}^{m+1}\overline{a_{im}}a_{ij}\right)\phi_j. \qquad (4.12)
\end{aligned}$$

Let

$$\psi_j^{(1)} = \left(M_{\chi^{(n,k)}}(A-\lambda I)^*(A-\lambda I)M_{\chi^{(n,k)}}\phi\right)_{k+j}, \quad j = 1,\ldots,n,$$

and let $\psi^{(1)} = \left(\psi^{(1)}_{k+1}, \psi^{(1)}_{k+2}, \ldots, \psi^{(1)}_{k+n}\right)^T$ and $\phi = (\phi_{k+1}, \phi_{k+2}, \ldots, \phi_{k+n})^T$. Then (4.12) implies that

$$\psi^{(1)} = \left[|\lambda|^2 I_n - \overline{\lambda} A_{n,k} - \lambda A^*_{n,k} + \widetilde{A}^*_{n,k} \widetilde{A}_{n,k}\right] \phi,$$

where $A_{n,k}$ is given by (3.5) and

$$\widetilde{A}_{n,k} = \begin{pmatrix} \gamma_{k+1} & & & & & & \\ \beta_{k+1} & \gamma_{k+2} & & & & & \\ \alpha_{k+1} & \beta_{k+2} & \gamma_{k+3} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \alpha_{k+n-3} & \beta_{k+n-2} & \gamma_{k+n-1} & & \\ & & & \alpha_{k+n-2} & \beta_{k+n-1} & \gamma_{k+n} \\ & & & & \alpha_{k+n-1} & \beta_{k+n} \\ & & & & & \alpha_{k+n} \end{pmatrix}_{(n+2)\times n}.$$

Similarly, we can show that if we let

$$\psi^{(2)}_j = \left(M_{\chi^{(n,k)}} (A - \lambda I) (A - \lambda I)^* M_{\chi^{(n,k)}} \phi\right)_{k+j}, \quad j = 1, \ldots, n,$$

and $\psi^{(2)} = \left(\psi^{(2)}_{k+1}, \psi^{(2)}_{k+2}, \ldots, \psi^{(2)}_{k+n}\right)^T$ then

$$\widetilde{\psi}^{(2)} = \left[|\lambda|^2 I_n - \overline{\lambda} A_{n,k} - \lambda A^*_{n,k} + \mathring{A}^*_{n,k} \mathring{A}_{n,k}\right] \phi$$

where

$$\mathring{A}_{n,k} = \begin{pmatrix} \alpha_k & & & & & & \\ \beta_{k+1} & \alpha_{k+1} & & & & & \\ \gamma_{k+2} & \beta_{k+2} & \alpha_{k+2} & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & \gamma_{k+n-4} & \beta_{k+n-3} & \alpha_{k+n-2} & & \\ & & & \gamma_{k+n-3} & \beta_{k+n-2} & \alpha_{k+n-1} \\ & & & & \gamma_{k+n-1} & \beta_{k+n} \\ & & & & & \gamma_{k+n+1} \end{pmatrix}_{(n+2)\times n}.$$

Define the $n \times n$ matrices $C^{\pm}_{n,k}$ by

$$C^{+}_{n,k} = \widetilde{A}^{H}_{n,k} \widetilde{A}_{n,k}$$

and

$$C^{-}_{n,k} = \mathring{A}^{H}_{n,k}\mathring{A}_{n,k}.$$

Setting

$$D^{\pm}_{n,k} = |\lambda|^2 I_n - \overline{\lambda}A_{n,k} - \lambda A^{H}_{n,k} + C^{\pm}_{n,k}, \tag{4.13}$$

clearly

$$\begin{aligned}\min\operatorname{spec} D^{+}_{n,k} &= \min\operatorname{spec}\left(M_{\chi^{(n,k)}}\left(A-\lambda I\right)^{*}\left(A-\lambda I\right)M_{\chi^{(n,k)}}\big|_{X_{n,k}}\right)\\ &= \min\operatorname{spec} B^{+}_{n,k},\end{aligned} \tag{4.14}$$

and

$$\begin{aligned}\min\operatorname{spec} D^{-}_{n,k} &= \min\operatorname{spec}\left(M_{\chi^{(n,k)}}\left(A-\lambda I\right)\left(A-\lambda I\right)^{*}M_{\chi^{(n,k)}}\big|_{X_{n,k}}\right)\\ &= \min\operatorname{spec} B^{-}_{n,k}.\end{aligned} \tag{4.15}$$

Recalling Definition 4.1, we see that, for $\varepsilon > 0$ and $n \in \mathbb{N}$,

$$\Gamma^{n}_{\varepsilon}(A) = \{\lambda \in \mathbb{C} : \xi_n(A-\lambda I) < \varepsilon\} \tag{4.16}$$

where $\xi_n(A-\lambda I) = \left\{\inf_{k\in\mathbb{Z}}\left(\min\operatorname{spec} D^{+}_{n,k}, \min\operatorname{spec} D^{-}_{n,k}\right)\right\}^{\frac{1}{2}}$.

In the numerical results that we show in section 4.4 we compute the inclusion sets $\overline{\Gamma^{n}_{f(n)}(A)}$ for spec A using equation (4.16)

4.3 Convergence of One-Sided Truncation Method

From Theorem 4.2 and Theorem 4.3 we get

$$\operatorname{spec}_{\varepsilon} A \subseteq \Gamma^{n}_{\varepsilon+f(n)}(A) \subseteq \operatorname{spec}_{\varepsilon+f(n)} A \tag{4.17}$$

for all $\varepsilon > 0, n \in \mathbb{N}$ and A as in (3.1).

From Theorem 2.32, it follows that the set $\operatorname{spec}_{\varepsilon} A$ depends continuously (in the Hausdorff metric) on $\varepsilon > 0$, i.e.

$$\varepsilon_n \to \varepsilon > 0 \quad \Rightarrow \quad d_H\left(\operatorname{spec}_{\varepsilon_n} A, \operatorname{spec}_{\varepsilon} A\right) \to 0.$$

By (4.11) it follows that $f(n) \to 0$ as $n \to \infty$, so that

$$d_H\left(\operatorname{spec}_{\varepsilon+f(n)} A, \operatorname{spec}_\varepsilon A\right) \to 0, \quad n \to \infty.$$

But from (4.17) it follows that

$$d_H\left(\Gamma^n_{\varepsilon+f(n)}(A), \operatorname{spec}_\varepsilon A\right) \le d_H\left(\operatorname{spec}_{\varepsilon+f(n)} A, \operatorname{spec}_\varepsilon A\right) \to 0,$$

as $n \to \infty$, so that

$$\Gamma^n_{\varepsilon+f(n)}(A) \to \operatorname{spec}_\varepsilon A \text{ and } \Gamma^n_{f(n)}(A) \to \operatorname{spec} A$$

in the Hausdorff metric as $n \to \infty$.

4.4 Numerical Examples for Method 2

4.4.1 Shift Operator

As an example of Corollary 4.4, we will apply method 2 to the right shift operator (3.41) and compare the result to the previous methods. Note that for $n \in \mathbb{N}, k \in \mathbb{Z}$, the matrices $\mathring{A}_{n,k}$ and $\widetilde{A}_{n,k}$ from Section 4.2 are in this case

$$\mathring{A}_{n,k} = \begin{pmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & 0 & 1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 0 & 0 & 1 & \\ & & & 0 & 0 & \\ & & & & 0 & \end{pmatrix} \text{ and } \widetilde{A}_{n,k} = \begin{pmatrix} 0 & & & & & \\ 0 & 0 & & & & \\ 1 & 0 & 0 & & & \\ & \ddots & \ddots & \ddots & & \\ & & 1 & 0 & 0 & \\ & & & 1 & 0 & \\ & & & & 1 & \end{pmatrix}. \tag{4.18}$$

Both matrices are of size $(n+2)\times n$ and are independent of k. From (4.13) we get that

$$D^+_{n,k} = D^-_{n,k} = D_n := \begin{pmatrix} |\lambda|^2+1 & -\lambda & & & & \\ -\overline{\lambda} & |\lambda|^2+1 & -\lambda & & & \\ & -\overline{\lambda} & |\lambda|^2+1 & -\lambda & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\overline{\lambda} & |\lambda|^2+1 & -\lambda \\ & & & & -\overline{\lambda} & |\lambda|^2+1 \end{pmatrix}_{n\times n}$$

Let $\omega := \sqrt{\lambda/\overline{\lambda}}$. We can show that

$$G^{-1}\left(D_n - (|\lambda|^2+1)I_n\right)G = -\begin{pmatrix} 0 & |\lambda| & & & & \\ |\lambda| & 0 & |\lambda| & & & \\ & |\lambda| & 0 & |\lambda| & & \\ & & \ddots & \ddots & \ddots & \\ & & & |\lambda| & 0 & |\lambda| \\ & & & & |\lambda| & 0 \end{pmatrix}_{n\times n}, \tag{4.19}$$

where

$$G = \begin{pmatrix} \omega & 0 & & & & \\ 0 & \omega^2 & 0 & & & \\ & 0 & \omega^3 & 0 & & \\ & & \ddots & \ddots & \ddots & \\ & & & 0 & \omega^{n-1} & 0 \\ & & & & 0 & \omega^n \end{pmatrix}_{n\times n}.$$

We have already shown in Chapter 3 that the spectrum (eigenvalues) of the matrix A_n of the form (3.48) is the set $\left\{2\cos\left(\frac{m\pi}{n+1}\right) : m = 1,\ldots,n\right\}$. From (4.19), we have that $\left(D_n - (|\lambda|^2+1)I_n\right)$ is similar to the matrix $-|\lambda|A_n$. It follows that

$$\begin{aligned} \operatorname{spec}\left(D_n - (|\lambda|^2+1)I_n\right) &= \operatorname{spec}\left(-|\lambda|A_n\right) \\ &= \left\{-2|\lambda|\cos\left(\frac{m\pi}{n+1}\right) \;:\; m = 1,\ldots,n\right\}. \end{aligned}$$

Therefore

$$\operatorname{spec} D_n = \left\{ |\lambda|^2 + 1 - 2|\lambda| \cos\left(\frac{m\pi}{n+1}\right) \;:\; m = 1, \dots, n \right\}.$$

$$\min \operatorname{spec} D_n = |\lambda|^2 + 1 - 2|\lambda|\, c_n$$

and

$$\xi_n(A - \lambda I) = \xi_{n,k}(A - \lambda I) := \left(|\lambda|^2 + 1 - 2|\lambda|\, c_n\right)^{\frac{1}{2}},$$

for all $k \in \mathbb{Z}$ where $c_n := \cos\left(\frac{\pi}{n+1}\right)$. Consequently, by equation (4.16),

$$\begin{aligned} \Gamma_\varepsilon^n(A) &= \{\lambda \in \mathbb{C} : \xi_n(A - \lambda I) < \varepsilon\} \\ &= \left\{\lambda \in \mathbb{C} : \left(|\lambda|^2 + 1 - 2|\lambda|\, c_n\right) < \varepsilon^2\right\} \\ &= \left\{\lambda \in \mathbb{C} : (|\lambda| - c_n)^2 < \varepsilon^2 + c_n^2 - 1\right\} \end{aligned}$$

From Corollary 3.15 we get that $\operatorname{spec} A \subseteq \overline{\Gamma_{f(n)}^n(A)}$. So if we put $\varepsilon = f(n)$ in the above computations, we get from (4.11) that

$$\begin{aligned} \overline{\Gamma_{f(n)}^n(A)} &= \left\{\lambda \in \mathbb{C} : (|\lambda| - c_n)^2 \le 4\sin^2\left(\frac{\pi}{2(n+1)}\right) + c_n^2 - 1\right\} \\ &= \left\{\lambda \in \mathbb{C} : (|\lambda| - c_n)^2 \le 2(1 - c_n) + c_n^2 - 1\right\} \\ &= \left\{\lambda \in \mathbb{C} : (|\lambda| - c_n)^2 \le (c_n - 1)^2\right\} \\ &= \{\lambda \in \mathbb{C} : ||\lambda| - c_n| \le |c_n - 1| = 1 - c_n\} \\ &= \{\lambda \in \mathbb{C} : |\lambda| \in [2c_n - 1, 1]\} \end{aligned}$$

Thus, we have shown that

$$\overline{\Gamma_{f(n)}^n(A)} = \{\lambda \in \mathbb{C} : |\lambda| \in [2c_n - 1, 1]\},$$

so that our inclusion set for $\operatorname{spec} A$ is an annulus with center 0, outer radius 1 and thickness

$$2(1 - c_n) = 4\sin^2\left(\frac{\pi}{2(n+1)}\right) < \frac{\pi^2}{(n+1)^2}.$$

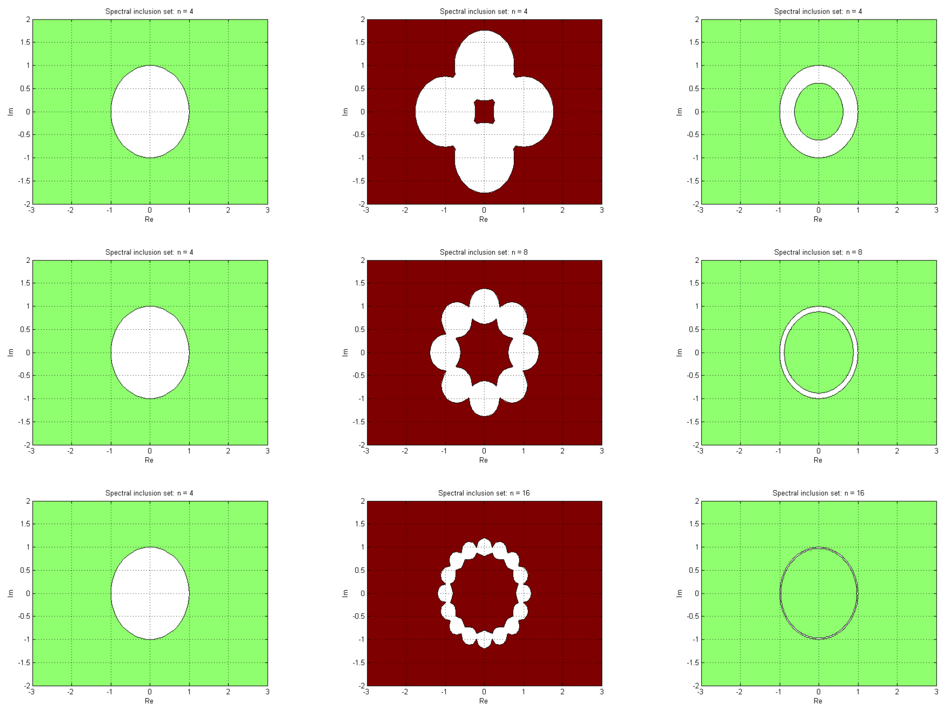

Figure 4.3: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}, \overline{\Pi^n_{f(n)}(A)}$ and $\overline{\Gamma^n_{f(n)}(A)}$ which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 4.4, where A is the shift operator, V_1, with $\operatorname{spec} A = \mathbb{T}$. Shown are the inclusion sets computed from method1 (column 1), method 1* (column 2), method 2 (column 3) when $n = 4, 8$ and 16.

Thus

$$d_H\left(\overline{\Gamma^n_{f(n)}(A)}, \operatorname{spec} A\right) = \max\left(4\sin^2\left(\frac{\pi}{2(n+1)}\right), 0\right)$$
$$= 4\sin^2\left(\frac{\pi}{2(n+1)}\right) < \frac{\pi^2}{(n+1)^2} \to 0$$

as $n \to \infty$. Therefore, $\overline{\Gamma^n_{f(n)}(A)} \to \operatorname{spec} A$ as $n \to \infty$ in Hausdorff metric, as shown already in section 4.3. Now we have a rigorous analytical description for all of the images in Figure 4.3.

4.4.2 3-periodic Bi-diagonal Operator

As the second example, Let us consider the case when when A is the bi-diagonal matrix,

$$A = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 1 & 1 & & & & \\ & 0 & -1 & 1 & & & \\ & & 0 & 1 & 2 & & \\ & & & 0 & 1 & 1 & \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}, \tag{4.20}$$

with only the main diagonal and the first superdiagonal non-zero, and both of these periodic with period 3.

We can compute explicitly that

$$\operatorname{spec}(A) = \left\{ \lambda \in \mathbb{C} : \lambda^3 - \lambda^2 + \lambda + 1 = 2e^{i\theta}, \theta \in [-\pi, \pi] \right\}.$$

From Figure 4.4, we can see that, when n is getting larger, the inclusion set $\overline{\Gamma^n_{f(n)}(A)}$ is converging to spec A faster than $\overline{\Pi^n_{f(n)}(A)}$ of method 1* which uses the periodised submatrices. We would not compare method 1 to method 2 because we believe that they converge to different sets. However, from the numerical results, this seems to convince us that method 2 produces the inclusion sets which give us the fastest convergence to spec A.

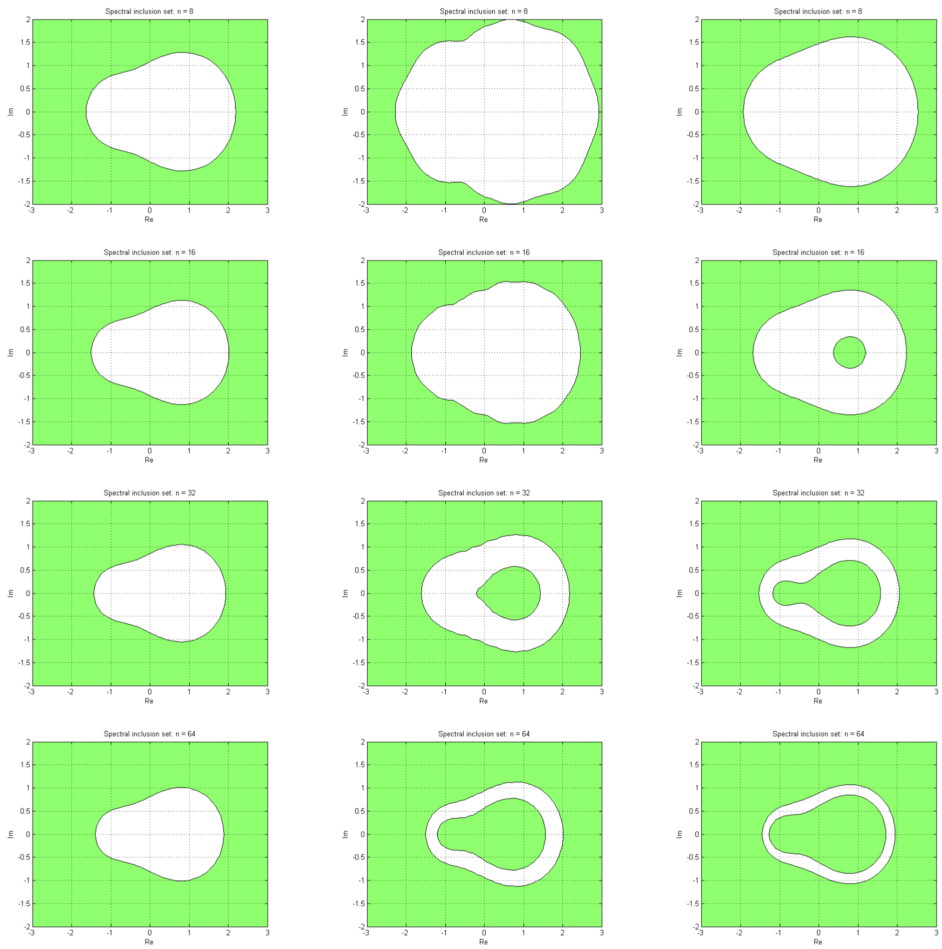

Figure 4.4: Plots of the sets $\overline{\Sigma^n_{f(n)}(A)}, \overline{\Pi^n_{f(n)}(A)}$ and $\overline{\Gamma^n_{f(n)}(A)}$ which are inclusion sets for $\operatorname{spec} A$, where $f(n)$ is given as in Corollary 4.4, where A is the periodic bi-diagonal operator which has $\alpha_i = 0$ for all i, $(\beta_i) = (\ldots, 1, -1, 1, 1, -1, 1, \ldots)$ and $(\alpha_i) = (\ldots, 1, 1, 2, 1, 1, 2, \ldots)$. Shown are the inclusion sets computed from method1 (column 1), method 1* (column 2), method 2 (column 3) when $n = 8, 16, 32$ and 64.

Chapter 5

Spectral Properties of a Random Tridiagonal Matrix Arising in Non-Self-Adjoint Quantum Mechanics

In this chapter, we will investigate the spectral properties of our infinite tridiagonal matrix of the form

$$A^b = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & 1 & & & & \\ & b_{-1} & 0 & 1 & & & \\ & & b_0 & 0 & 1 & & \\ & & & b_1 & 0 & 1 & \\ & & & & b_2 & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix},$$

where $b = (b_i) \in \{\pm 1\}^{\mathbb{Z}}$, which is a non-self-adjoint tri-diagonal matrix. We will be particularly interested in the case where the entries b_i are chosen randomly. In the first section, we will discuss some related research work which gives us some inspiration.

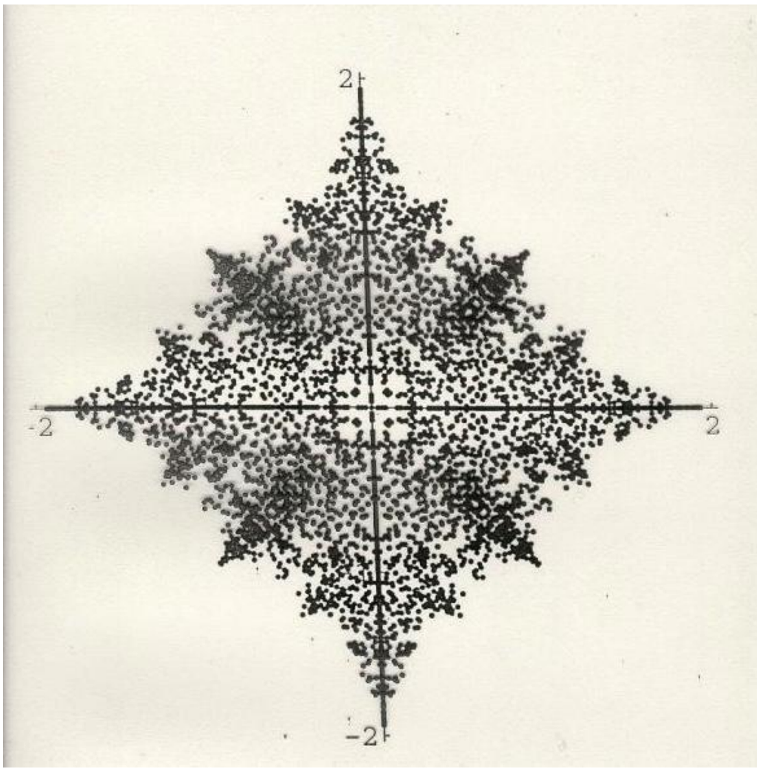

Figure 5.1: A picture of spectral plot from a talk of A. Zee at the Math. Sci. Research Institute, Berkeley in 1999.

5.1 Background and Motivation

Recently, there has been a great deal of work on the spectrum and pseudospectrum of doubly- (and singly-) infinite random tridiagonal matrices and their finite sections (e.g. [1, 5, 6, 11, 16, 17, 19, 20, 23, 24, 25, 29, 37, 39, 40, 45]). In 2001, Trefethen, Contedini and Embree [48] studied the spectra and pseudospectra of random bidiagonal matrices of the form

$$A_n = \begin{pmatrix} x_1 & 1 & & & \\ & x_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & x_{n-1} & 1 \\ & & & & x_n \end{pmatrix}_{n\times n}$$

where each x_i is a random variable taking values independently in a compact subset of $\mathbb{C}$, from some distribution X. If the entries on the main diagonal generate from $X = [-2, 2]$ with uniform probability, this random matrix A is associated with the "one-way-model" by Brezin, Feinberg and Zee [5, 16, 17], when the entries on the main diagonal generate from $\{\pm 1\}$, then they obtained the following matrix

$$A_n = \begin{pmatrix} \pm 1 & 1 & & & \\ & \pm 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \pm 1 & 1 \\ & & & & \pm 1 \end{pmatrix}_{n\times n}$$

In order to prove the main result, they started characterising the spectra, pseudospectra and numerical range of the finite matrix A_n. The main idea of the proof is to show that for every $\lambda \in \mathbb{C}$ they have conditions on the norm of resolvent $(A_n - \lambda I)^{-1}$ which guaranteed the exponential growth, guaranteed almost sure exponential growth, guaranteed almost sure subexponential growth and guaranteed boundedness as $n \to \infty$. It follows that the set of complex numbers has been divided into 4 regions according to the behaviour of $(A_n - \lambda I)^{-1}$, i.e. delocalized (surely), delocalized almost surely, localized almost surely and localized (surely). Note that $(A_n - \lambda I)^{-1}$ is said to be localized if the entries of the colums in $(A_n - \lambda I)^{-1}$ decay exponentially with distance from the diagonal and delocalized if they do not.

Then they proved results for the case of infinite bidiagonal matrices, finite periodic matrices and bi-infinite bidiagonal matrices A of the form

$$A = \begin{pmatrix} \ddots & \ddots & & & & \\ & \pm 1 & 1 & & & \\ & & \pm 1 & 1 & & \\ & & & \pm 1 & 1 & \\ & & & & \pm 1 & \ddots \\ & & & & & \ddots \end{pmatrix},$$

respectively. The main result for a random bidiagonal doubly infinite matrix case is, with probability 1, $\operatorname{spec} A$ is the union of the regions which are delocalized almost surely and localized almost surely. Precisely, $\operatorname{spec} A$ is the union of the two closed unit disks centred at 1 and -1, respectively.

In 2008, Lindner [35] generalizes the above result for the case of one random and one constant diagonals to the case of two random diagonals, so that

$$A = \begin{pmatrix} \ddots & \ddots & & & & \\ & \sigma_{-1} & \tau_{-1} & & & \\ & & \sigma_0 & \tau_0 & & \\ & & & \sigma_1 & \tau_1 & \\ & & & & \sigma_2 & \ddots \\ & & & & & \ddots \end{pmatrix}, \tag{5.1}$$

where $\sigma_k \in \Sigma$ and $\tau_k \in \mathcal{T}$ are taken independently from a random distribution on Σ and $\mathcal{T}$, which are arbitrary compact subsets of $\mathbb{C}$, respectively under the condition for every $\varepsilon > 0, \sigma \in \Sigma$ and $\tau \in \mathcal{T}$, that $Pr\left(|\sigma_k - \sigma| < \varepsilon\right)$ and $Pr\left(|\tau_k - \tau| < \varepsilon\right)$ are both non-zero. This is a proper generalization of [48] because the set $\mathcal{T}$ may contain zero. For every $\varepsilon > 0$, define

$$\Sigma_{\cup}^{\varepsilon} := \bigcup_{\sigma\in\Sigma} \overline{U}_{\varepsilon}(\sigma) \text{ and } \Sigma_{\cap}^{\varepsilon} := \bigcap_{\sigma\in\Sigma} U_{\varepsilon}(\sigma)$$

where $U_\varepsilon(\sigma) = \{\lambda \in \mathbb{C} : |\lambda - \sigma| < \varepsilon\}$. He then proved the following nice result:

Theorem 5.1. *If A is the random matrix shown in (5.1) then, almost surely,*

$$\operatorname{spec} A = \operatorname{spec}_{\text{ess}} A = \Sigma_{\cup}^{T} \setminus \Sigma_{\cap}^{t}$$

where $T = \max\{|\tau| : \tau \in \mathcal{T}\}$ *and* $t = \min\{|\tau| : \tau \in \mathcal{T}\}$.

If all entries on the main diagonal and on the upper diagonal take values independently from the sets $\Sigma = \{\pm 1\}$ and $\mathcal{T} = \{1\}$, respectively, then it is obvious from Theorem 5.1 that the spectrum of that infinite bi-diagonal matrix is the union of the two closed unit disks centred at 1 and -1, respectively.

Moreover, Lindner computed the spectrum of doubly infinite random matrices of the form (5.1) where (σ_i) and (τ_i) are pseudo-ergodic ± 1-sequences, by looking at the union of spectrum of finitely many limit operators. Since the set $\sigma^{\text{op}}(A)$ is very large and very difficult to compute, he computed the spectrum of bi-infinite random ± 1-sequence matrix by considering only periodic limit operators. He approximated the spectrum of the infinite matrix by the union of all eigenvalues of every possibility n-periodic limit operators,

$$\operatorname{spec}_{\text{per}} A := \bigcup_{n=1}^{\infty} \left(\bigcup_{B\in\mathcal{P}_n(A)} \operatorname{spec}_{\text{point}}^{\infty} B \right),$$

where $\mathcal{P}_n(A) \subseteq \sigma^{\mathsf{op}}(A)$ denotes the set of all limit operators of A with $n-$periodic diagonals. Then he proved that $\operatorname{spec}_{\text{per}} A$ is dense in $\operatorname{spec} A$.

There are some difficulties to study the spectrum of A^b and to find the explicit form of eigenvectors. First of all, the eigenvalues of the finite section matrices of the infinite bidiagonal random matrices can be found more easily than those of the finite section matrices of the infinite tridiagonal random matrices since the eigenvalues of the finite section matrices of the infinite bidiagonal random matrices are already the entries on the main diagonal of those finite submatrices.

Moreover, computing the spectra and pseudospectra of infinite bidiagonal random matrices is a lot easier than computing the spectra and pseudospectra of infinite tridiagonal random matrices. The reason is, to find the eigenvectors of bidiagonal matrices we need to solve 2-term recurrence equations which is less complicated than finding the eigenvector of infinite tridiagonal random matrices which satisfy a 3-term recurrence relation.

In 1999, to the best of my knowledge, leading quantum physicists including Anthony Zee, Joshua Feinberg started studying to describe the propagation of a particle hopping on a 1-dimensional lattice. Feinberg and Zee [16] studied the equation

$$v_{k+1} + r_{k-1}v_{k-1} = \lambda v_k \tag{5.2}$$

where the real numbers r_k are generated from some random distribution and λ is the spectral parameter. Zee and Feinberg studied the distribution of the eigenvalues of the $n \times n$ matrix A_n^b defined by

$$A_n^b = \begin{pmatrix} 0 & 1 & & & & \\ b_1 & 0 & 1 & & & \\ & b_2 & 0 & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & b_{n-2} & 0 & 1 \\ & & & & b_{n-1} & 0 \end{pmatrix},$$

which is obtained from (5.2) when $v_0 = v_{n+1} = 0$ and each b_k is ± 1, randomly. They noticed that when n was large, the spectrum has a complicated fractal-like form. The matrix A^b is obtained from one of the simple models suggested in [28]. They studied the eigenvalues of the finite section matrix A_n^b and rewrite the equation (5.2) into the transfer matrix form

$$\begin{pmatrix} x_{k+1} \\ x_k \end{pmatrix} = T_{k-1} \begin{pmatrix} x_k \\ x_{k-1} \end{pmatrix}$$

where T_k defined by

$$T_k = \begin{pmatrix} \lambda & -b_k \\ 1 & 0 \end{pmatrix}$$

Let $P = [p_{11} \quad p_{12}]$ be a vector defined by $P := T_{n-1} \cdots T_2 T_1$. Then they determined the eigenvalues once they solved the equation $\lambda p_{11} + p_{12} = 0$. They had also noticed that when they considered a large finite subsection, the spectrum has a complicated fractal-like form.

In 2002, Holz, Orland and Zee [28] studied the spectrum of the infinite random matrix A^b for all 6 possible cases with a 4-periodic sequence. They found that the spectrum for each of these 6 patterns corresponds to a certain curve. In this paper, they stated some open questions on the spectrum of the infinite random matrix e.g. does the spectrum contain a hole in the complex plane or not? Is the spectrum of the operator localized or delocalized?

In 2010, Chien and Nakazato [10] studied the numerical range of tri-diagonal operators A defined by $Ae_j = e_{j-1} + r^j e_{j+1}$, where $r \in \mathbb{R}, j \in \mathbb{N}$ and $\{e_1, e_2, \ldots\}$ is the standard orthonormal basis for $\ell^2(\mathbb{N})$. In the third section of this paper, they emphasised the case $r = -1$ and they showed that

$$\begin{aligned} W(A) = \{z \in \mathbb{C} : -1 \leq \text{Re}(z) \leq 1, -1 \leq \text{Im}(z) \leq 1\} \\ \setminus \{1+i, 1-i, -1+i, -1-i\}. \end{aligned}$$

5.2 The Initial Investigations

Consider the $n \times n$ finite section matrix of the form

$$A_n^b = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ b_1 & 0 & 1 & \ldots & 0 \\ \vdots & b_2 & \ddots & \ddots & \vdots \\ 0 & \ldots & \ddots & 0 & 1 \\ 0 & \ldots & \ldots & b_{n-1} & 0 \end{pmatrix}. \tag{5.3}$$

All eigenvalues λ and eigenvectors (or eigenfunctions) x of A_n^b satisfy

$$A_n^b x = \lambda x.$$

Taking norms on both sides, we have

$$|\lambda|\,\|x\| = \|\lambda x\| = \left\|A_n^b x\right\| \le \left\|A_n^b\right\|\,\|x\|$$

i.e.

$$|\lambda| \le \left\|A_n^b\right\|.$$

In particular

$$\left\|A_n^b\right\|_\infty = \max_m \sum_n |a_{mn}| = 2.$$

Also for the infinite case, $\left\|A_n^b\right\| = 2$, so $|\lambda| \le 2$. Further, from the basic property of Toeplitz operators (see [4]), we know that

$$A := \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & 1 & & & & \\ & c & 0 & 1 & & & \\ & & c & 0 & 1 & & \\ & & & c & 0 & 1 & \\ & & & & c & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix},$$

has symbol $a(t) = t + ct^{-1}$ where $t \in \mathbb{T} = \{t : |t| = 1\}$. If $c = 1$, it follows that

$$\operatorname{spec} A = \{e^{i\theta} + e^{-i\theta} : \theta \in \mathbb{R}\} = \{2\cos\theta : \theta \in \mathbb{R}\} = [-2, 2] \tag{5.4}$$

and if $c = -1$,

$$\operatorname{spec} A = \{e^{i\theta} - e^{-i\theta} : \theta \in \mathbb{R}\} = \{2i\sin\theta : \theta \in \mathbb{R}\} = i[-2, 2]. \tag{5.5}$$

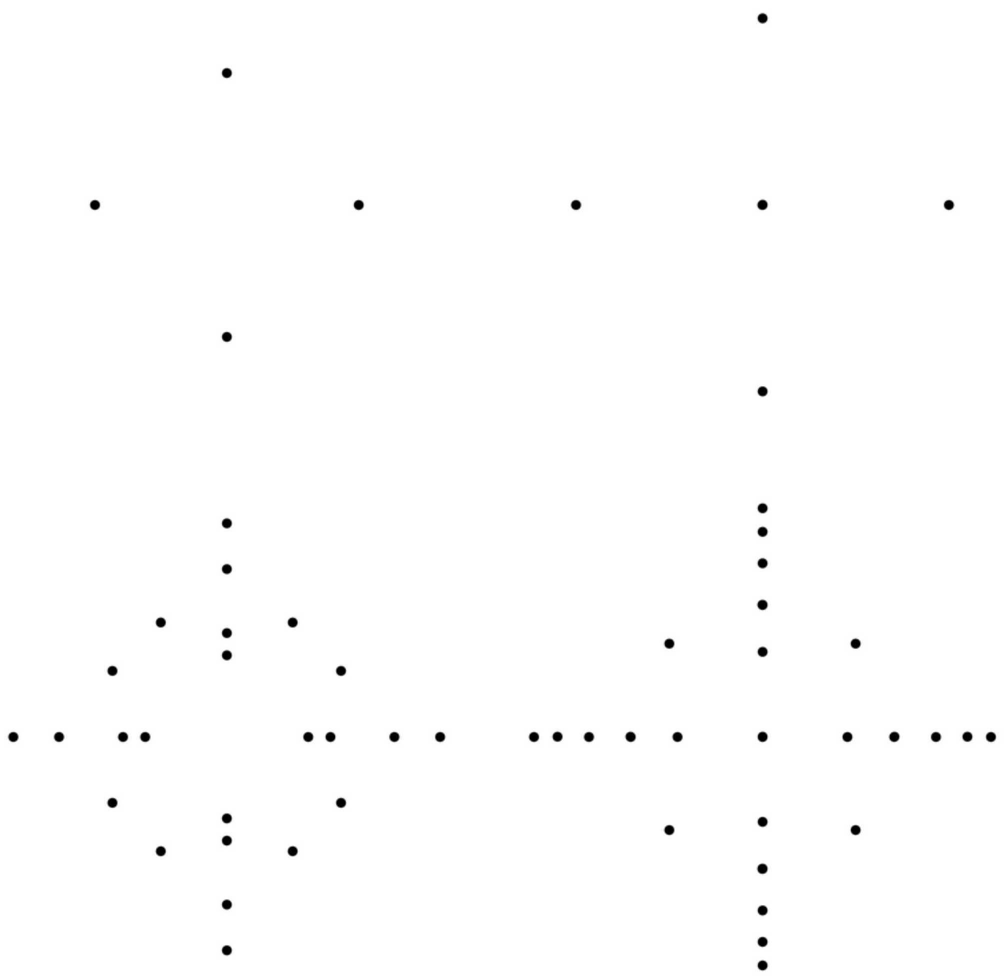

Figure 5.2: The plot of all eigenvalues of all possible finite matrices of order $n = 2, 3, 4$ and 5.

5.2.1 The Symmetry of $\operatorname{spec} A_n^b$

From Figure 5.2, we can see the symmetry of $\operatorname{spec} A_n^b$. This suggests us to prove rigorously the symmetry properties of $\operatorname{spec} A_n^b$ and $\operatorname{spec} A^b$, respectively. To study the symmetry of the spectrum ofA_n^b, we need to prove the following proposition:

Proposition 5.2. *Let $n \geq 3$. The characteristic polynomial of A_n^b is an odd function if n is odd and it is an even function if n is even. Moreover, for $n \geq 3$,*

$$D_n := det(A_n^b - \lambda I) = -\lambda D_{n-1} - b_{n-1} D_{n-2}.$$

Proof. We are proving this statement by using mathematical induction on the size of the finite section, n.
Basis Step : If $n = 3$ we consider the matrix

$$A_3^b = \begin{pmatrix} 0 & 1 & 0 \\ b_1 & 0 & 1 \\ 0 & b_2 & 0 \end{pmatrix}.$$

The characteristic polynomial for A_3 is $\det(A_3 - \lambda I) = -\lambda^3 + (b_1 + b_2)\lambda$ which is an odd function no matter what b_1, b_2 are.

If $n = 4$ we consider the matrix

$$A_4^b = \begin{pmatrix} 0 & 1 & 0 & 0 \\ b_1 & 0 & 1 & 0 \\ 0 & b_2 & 0 & 1 \\ 0 & 0 & b_3 & 0 \end{pmatrix}.$$

The characteristic polynomial D_4 is

$$D_4 = \det(A_4^b - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 & 0 \\ b_1 & -\lambda & 1 & 0 \\ 0 & b_2 & -\lambda & 1 \\ 0 & 0 & b_3 & -\lambda \end{vmatrix}$$

$$= (-\lambda)\begin{vmatrix} -\lambda & 1 & 0 \\ b_1 & -\lambda & 1 \\ 0 & b_2 & -\lambda \end{vmatrix} - \begin{vmatrix} -\lambda & 1 & 0 \\ b_1 & -\lambda & 1 \\ 0 & 0 & b_3 \end{vmatrix}$$

$$= (-\lambda)\begin{vmatrix} -\lambda & 1 & 0 \\ b_1 & -\lambda & 1 \\ 0 & b_2 & -\lambda \end{vmatrix} - b_3\begin{vmatrix} -\lambda & 1 \\ b_1 & -\lambda \end{vmatrix}.$$

Since D_3 is an odd function, it follows that D_4 is an even function. In addition,

$$D_4 = (-\lambda)D_3 - (b_3)D_2.$$

Inductive Step : Let k be any integer such that for every $l \leq k$, then

$$\begin{vmatrix} -\lambda & 1 & & & & \\ b_1 & -\lambda & 1 & & & \\ & b_2 & -\lambda & 1 & & \\ & & b_3 & -\lambda & \ddots & \\ & & & \ddots & \ddots & 1 \\ & & & & b_{l-1} & -\lambda \end{vmatrix}$$

is an odd function if l is odd and is an even function if l is even. Consider

$$\det(A_{k+1} - \lambda I) = \begin{vmatrix} -\lambda & 1 & & & & & & \\ b_1 & -\lambda & 1 & & & & & \\ & b_2 & -\lambda & 1 & & & & \\ & & b_3 & -\lambda & \ddots & & & \\ & & & \ddots & \ddots & 1 & & \\ & & & & b_{k-2} & -\lambda & 1 & \\ & & & & & b_{k-1} & -\lambda & 1 \\ & & & & & & b_k & -\lambda \end{vmatrix} \tag{5.6}$$

Using the $(k+1)th$ column to compute the determinant, we see that

$$\det(A^b_{k+1} - \lambda I) = (-\lambda) \cdot \begin{vmatrix} -\lambda & 1 & & & & & \\ b_1 & -\lambda & 1 & & & & \\ & b_2 & -\lambda & 1 & & & \\ & & b_3 & -\lambda & \ddots & & \\ & & & \ddots & \ddots & 1 & \\ & & & & b_{k-2} & -\lambda & 1 \\ & & & & & b_{k-1} & -\lambda \end{vmatrix}$$

$$- (1) \cdot \begin{vmatrix} -\lambda & 1 & & & & & \\ b_1 & -\lambda & 1 & & & & \\ & b_2 & -\lambda & 1 & & & \\ & & b_3 & -\lambda & \ddots & & \\ & & & \ddots & \ddots & 1 & \\ & & & & b_{k-2} & -\lambda & 1 \\ & & & & & 0 & b_k \end{vmatrix}$$

$$= (-\lambda) \cdot \begin{vmatrix} -\lambda & 1 & & & & & \\ b_1 & -\lambda & 1 & & & & \\ & b_2 & -\lambda & 1 & & & \\ & & b_3 & -\lambda & \ddots & & \\ & & & \ddots & \ddots & 1 & \\ & & & & b_{k-2} & -\lambda & 1 \\ & & & & & b_{k-1} & -\lambda \end{vmatrix}$$

$$- (b_k) \cdot \begin{vmatrix} -\lambda & 1 & & & & \\ b_1 & -\lambda & 1 & & & \\ & b_2 & -\lambda & 1 & & \\ & & b_3 & -\lambda & \ddots & \\ & & & \ddots & \ddots & 1 \\ & & & & b_{k-2} & -\lambda \end{vmatrix}$$

$$= (-\lambda)D_k - b_k D_{k-1}.$$

By the induction hypothesis, we can see that, if k is odd then D_k is odd and D_{k-1} is even i.e. D_{k+1} is even and if k is even then D_k is even and D_{k-1} is odd i.e. D_{k+1} is even as desired. ■

Note that every odd function satisfies $f(0) = 0$. As a result, if n is odd then $\det(A_n^b - \lambda I)$ can be factorised as $\det(A_n^b - \lambda I) = \lambda \cdot m(\lambda)$, where $m(\lambda)$ is an even function. As a consequence, if λ is a root of the characteristic equation, so is $-\lambda$.

Because all coeficients of D_n are real, all its complex roots come in conjugated pairs, so that with λ also $\overline{\lambda}$ is a root of D_n.

In the next subsection we are proving the symmetry property of $\operatorname{spec} A^b$.

5.2.2 The Symmetry of $\operatorname{spec} A^b$

From the results in the subsection 5.2.1, we have already shown that the spectra of the finite submatrices of A^b are symmetric about the $x-$axis, $y-$axis and $90°$ rotation around the origin. This suggests us to show that if $\lambda \in \operatorname{spec} A^b$ then $-\lambda$ and $\overline{\lambda}$ also belong to $\operatorname{spec} A^b$. We need to recall the definition of pseudo-ergodic in our case. A sequence $(b_i) \in \{\pm 1\}^{\mathbb{Z}}$ is said to be pseudo-ergodic iff every finite pattern of ± 1's can be found somewhere in the sequence b.

Proposition 5.3. *If b is pseudo-ergodic then,*

$$\sigma^{op}(A^b) = \left\{ A^c : c \in \{\pm 1\}^{\mathbb{Z}} \right\}.$$

Proof. This follows from $A^b = M_b V_1 + V_{-1}$ and Proposition 2.55. ■

Proposition 5.4. *If $b \in \{\pm 1\}^{\mathbb{Z}}$ is pseudo-ergodic (which holds almost surely if b is random) then*

1. *It holds that*

$$\operatorname{spec} A^b = \operatorname{spec}_{\text{ess}} A^b = \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec} A^c = \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec}^{\infty}_{\text{point}} A^c, \tag{5.7}$$

   so in particular,

$$\operatorname{spec}_{\text{per}} A^b := \bigcup_{n \in \mathbb{N}} \operatorname{spec}^n_{\text{per}} A^b \subseteq \operatorname{spec} A^b \tag{5.8}$$

   (see Fig.5.3, Fig.5.4 and Fig.5.5), where

$$\operatorname{spec}^n_{\text{per}} A^b := \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}, n-periodic} \operatorname{spec}^{\infty}_{\text{point}} A^c. \tag{5.9}$$

2. $\operatorname{spec} A^b$ *is invariant under reflection about either axis as well as under a 90° rotation around the origin.*

Proof.

1. By Proposition 2.53, i.e. for any $A \in \mathcal{W}$,

$$\operatorname{spec}_{\text{ess}} A = \bigcup_{B \in \sigma^{\mathsf{op}}(A)} \operatorname{spec} B = \bigcup_{B \in \sigma^{\mathsf{op}}(A)} \operatorname{spec}^{\infty}_{\text{point}} B,$$

and Proposition 5.3, i.e. $\sigma^{\mathsf{op}}(A^b) = \left\{A^c : c \in \{\pm 1\}^{\mathbb{Z}}\right\}$, we will get that

$$\operatorname{spec}_{\text{ess}} A^b = \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec} A^c = \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec}^{\infty}_{\text{point}} A^c.$$

Moreover, we can see that $\operatorname{spec} A^b \subseteq \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec} A^c = \operatorname{spec}_{\text{ess}} A^b$. On the other hand, of course $\operatorname{spec}_{\text{ess}} A^b \subseteq \operatorname{spec} A^b$. Therefore we obtain the desired result as in the proposition.

2. we let $\lambda \in \operatorname{spec} A^b = \bigcup_{c \in \{\pm 1\}^{\mathbb{Z}}} \operatorname{spec}^{\infty}_{\text{point}} A^c$, i.e. $\lambda \in \operatorname{spec}^{\infty}_{\text{point}} A^d$ for some $d \in \{\pm 1\}^{\mathbb{Z}}$. Hence we have that $A^d v = \lambda v$. Now we will show that $-\lambda \in \operatorname{spec} A^b$. Choose a vector $\hat{v}$ where $\hat{v}_k = (-1)^k v_k$, then

$$\begin{aligned} A^d v = \lambda v &\Leftrightarrow d_{k-1} v_{k-1} - \lambda v_k + v_{k+1} = 0 \\ &\Leftrightarrow d_{k-1}\left((-1)^{k-1} v_{k-1}\right) - \lambda\left((-1)^{k-1} v_k\right) + (-1)^{k-1} v_{k+1} = 0 \\ &\Leftrightarrow d_{k-1}\left((-1)^{k-1} v_{k-1}\right) + \lambda\left((-1)^{k} v_k\right) + (-1)^{k+1} v_{k+1} = 0 \\ &\Leftrightarrow A^d \hat{v} = -\lambda \hat{v} \\ &\Leftrightarrow -\lambda \in \operatorname{spec}^{\infty}_{\text{point}} A^d. \end{aligned}$$

We are now showing that if $\lambda \in \operatorname{spec} A^b$ then $\lambda i \in \operatorname{spec} A^b$. Choose a vector v' such that $v'_k = i^k v_k$.

$$\begin{aligned} A^d v = \lambda v &\Leftrightarrow d_{k-1} v_{k-1} - \lambda v_k + v_{k+1} = 0 \\ &\Leftrightarrow d_{k-1}\left(i^{k+1} v_{k-1}\right) - \lambda\left(i^{k+1} v_k\right) + i^{k+1} v_{k+1} = 0 \\ &\Leftrightarrow -d_{k-1}\left(i^{k-1} v_{k-1}\right) - \lambda i\left(i^{k} v_k\right) + i^{k+1} v_{k+1} = 0 \\ &\Leftrightarrow e_{k-1} v'_{k-1} - \lambda i v'_k + v'_{k+1} = 0 \\ &\Leftrightarrow A^e v' = \lambda i v' \\ &\Leftrightarrow \lambda i \in \operatorname{spec}^{\infty}_{\text{point}} A^e, \end{aligned}$$

where $e_k = -d_k$ for all $k \in \mathbb{Z}$, so that also $e = (e_k) \in \{\pm 1\}^{\mathbb{Z}}$ and hence $\lambda i \in \operatorname{spec} A^b$. Moreover, we know

$$\left\|\left(A^b - \overline{\lambda} I\right)^{-1}\right\| = \left\|\left(\left(A^b - \lambda I\right)^*\right)^{-1}\right\| = \left\|\left(\left(A^b - \lambda I\right)^{-1}\right)^*\right\| = \left\|\left(A^b - \lambda I\right)^{-1}\right\|,$$

we can conclude that $\overline{\lambda} \in \operatorname{spec} A^b$.

Hence, $-\lambda, \overline{\lambda}, \lambda i \in \bigcup_{c\in\{\pm 1\}^{\mathbb{Z}}} \operatorname{spec}^{\infty}_{\text{point}} A^c = \operatorname{spec} A^b$. Therefore, $\operatorname{spec} A^b$ is invariant under reflection about either axis as well as under a 90° rotation around the origin. ■

5.2.3 Numerical Range of A^b

In this section, we will show that the numerical range of A^b is equal to $D = \{x + iy : x, y \in \mathbb{R}, |x| + |y| < 2\}$.

Proposition 5.5. *If b is pseudo-ergodic, then*

$$\overline{D} \subseteq \overline{W(A^b)},$$

and

$$D \subseteq W(A^b).$$

Proof. Let $\lambda \in \overline{D} = \{x+iy : x, y \in \mathbb{R}, |x|+|y| \leq 2\}$, which is the square with $2, -2, 2i$ and $-2i$ as its corners. From (5.4), (5.5) and Proposition 5.4, we know that $[-2, 2] \cup [-2i, 2i] \in \bigcup_{c\in\{\pm 1\}^{\mathbb{Z}}} \operatorname{spec} A^c = \operatorname{spec} A^b$. Note that $W(A^b)$ is convex and $\operatorname{spec} A^b \subseteq \overline{W(A^b)}$ by Theorem 2.28, it follows that the line segments $[-2, 2], [-2i, 2i]$ and $\{(x + yi) \in \mathbb{C} : |x| + |y| = 2\}$ are in $\overline{W(A^b)}$. From the convexity, it is implied that $\lambda \in \overline{W(A^b)}$.

Now, we are proving that $D \subseteq W(A^b)$. Let $\mu \in D$. Since D is an open set, there exists $\varepsilon_1 > 0$ such that $B_{\varepsilon_1}(\mu) \subseteq D$. We claim that μ does not belong to the boundary of $W(A^b)$. Suppose not, it also means μ belongs to the boundary of $\overline{W(A^b)}$, i.e. for every $\varepsilon > 0$, $B_\varepsilon(\mu) \cap \overline{W(A^b)} \neq \phi$ and $B_\varepsilon(\mu) \cap \left(\overline{W(A^b)}\right)^C \neq \phi$, where A^C is the complement of a set A. Since $\overline{D} \subseteq \overline{W(A^b)}$, there exists $\zeta \in B_{\varepsilon_1}(\mu)$ such that $\zeta \notin \overline{D}$ which contradicts the fact that $B_{\varepsilon_1}(\mu) \subseteq D$. That means the supposition is not true. Therefore, $D \subseteq W(A^b)$. ■

Lemma 5.6. $\forall x, y \in \ell^2(\mathbb{Z}), \qquad \overline{(x, y)} = (\bar{x}, \bar{y}).$

Proof. Let $x, y \in \ell^2(\mathbb{Z})$. Then

$$\begin{aligned}\overline{(x,y)} &= \overline{\sum_{i\in\mathbb{Z}} x\bar{y}} \\ &= \sum_{i\in\mathbb{Z}} \bar{x}y \\ &= (\bar{x},\bar{y}).\end{aligned}$$

■

Lemma 5.7. *If* $\lambda \in W(A^b)$ *then* $\overline{\lambda} \in W(A^b)$ *and* $-\lambda \in W(A^b)$.

Proof. Let $\lambda \in W(A^b)$. So, there exists a unit vector $v \in \ell^2(\mathbb{Z})$ such that $\lambda = (A^b v, v)$.
Claim $\overline{\lambda} \in W(A^b)$.

$$\bar{\lambda} = \overline{(A^b v, v)}$$

since $\overline{(x,y)} = (\bar{x},\bar{y})$,

$$= (\overline{A^b v}, \bar{v})$$

since every entry of A^b is real,

$$= (A^b \bar{v}, \bar{v}).$$

Since $\|\bar{v}\| = \|v\| = 1$, it follows that $\bar{\lambda} \in W(A^b)$.

Claim$-\lambda \in W(A^b)$.
Since

$$-\lambda = -(A^b v, v) = (A^b vi, vi),$$

and $\|vi\| = \|v\| = 1$, it follows that $-\lambda \in W(A^b)$. ■

From Lemma 5.7, we know that if $a + bi \in W(A^b)$ then $a - bi \in W(A^b)$. Therefore, $W(A^b)$ is symmetric about the real axis. Since we know that $a-bi$ belongs to $W(A^b)$ if $a + bi \in W(A^b)$ and $-(a - bi)$ also belongs to $W(A^b)$. Thus, $-a + bi \in W(A^b)$ if $a + bi \in W(A^b)$. As a consequence, $W(A^b)$ is symmetric about the imaginary axis

For the next lemma, we will show that $W(A^b) \subseteq D$. We need to show some facts about the real part of the numerical range as following :

$$\begin{aligned} Re(Av, v) &= \frac{1}{2}[(Av, v) + \overline{(Av, v)}] \\ &= \frac{1}{2}[(Av, v) + \overline{(v, (A)^* v)}] \\ &= \frac{1}{2}[(Av, v) + (A^* v, v)] \\ &= (Bv, v) \end{aligned}$$

where $A \in B(\ell^2(\mathbb{Z}))$ and $B = \frac{1}{2}(A + (A^*)$ is self-adjoint. This is very helpful.

Lemma 5.8. *$Re(e^{\frac{i\pi}{4}}(A^b v, v)) < \sqrt{2}$, for all $v \in \ell^2(\mathbb{Z})$ with $\|v\| \leq 1$.*

Proof. Since $A^b = M_b V_1 + V_{-1}$ and $(A^b)^* = V_{-1} M_b + V_1$, it follows that $e^{\frac{i\pi}{4}} A^b = M_{e^{\frac{i\pi}{4}} b} V_1 + e^{\frac{i\pi}{4}} V_{-1}$ and $(e^{\frac{i\pi}{4}} A^b)^* = V_{-1} M_{e^{-\frac{i\pi}{4}} b} + e^{-\frac{i\pi}{4}} V_1$, and let

$$B_{\frac{\pi}{4}} := \frac{1}{2}(e^{\frac{i\pi}{4}} A^b + (e^{\frac{i\pi}{4}} A^b)^*) = (M_c V_1 + V_{-1} M_{\bar{c}}),$$

where $c = \frac{1}{2}(e^{\frac{i\pi}{4}} b + e^{-\frac{i\pi}{4}})$, i.e., $\|c\|_\infty = \frac{1}{\sqrt{2}}$. From the previous statement before this lemma, we know that

$$\begin{aligned} \text{Re } e^{\frac{i\pi}{4}}(A^b v, v) &= \text{Re } (e^{\frac{i\pi}{4}} A^b v, v) \\ &= (B_{\frac{\pi}{4}} v, v) \\ &= (M_c V_1 v, v) + (V_{-1} M_{\bar{c}} v, v) \\ &\leq \|c\|_\infty ((V_1 w, w) + (V_{-1} w, w)) \qquad \text{where } w_i = |v_i|. \end{aligned}$$

Claim 1. $V_1 w \neq mw$ for every constant m. Suppose $V_1 w = mw$ for some constant m. We have $\|V_1 w\| = \|mw\|$. Since V_1 is an isometry, $1 = \|V_1\| = |m|$. It follows that $|w_i| = |w_{i+1}|$ for every i. That means, the sequence $w = (w_i) \notin \ell^2(\mathbb{Z})$, which is a contradiction. Therefore the supposition is not true.

Claim 2.$|(V_1 w, w)| < 1$ and $|(V_{-1} w, w)| < 1$.

By Cauchy-Schwarz Inequality and Claim 1. we know that the equality doesn't hold in any case, i.e.

$$|(V_1 w, w)| < \|V_1 w\| \|w\| = 1.$$

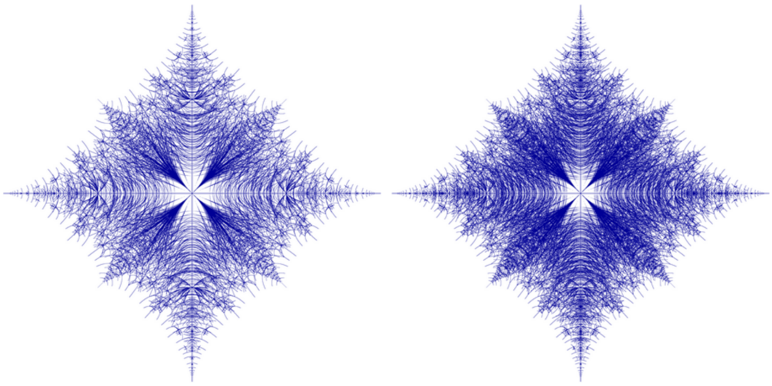

Figure 5.3: This figure shows the sets $\bigcup_{n=1}^{12} \mathrm{spec}_{\mathrm{per}}^{n} A^b$ and $\bigcup_{n=1}^{13} \mathrm{spec}_{\mathrm{per}}^{n} A^b$, as defined in (5.9).

Similarly, we can show that $|(V_{-1}w, w)| < 1$.
Hence, $\left|(B_{\frac{\pi}{4}}w, w)\right| \leq \|c\|_\infty (|(V_1 w, w)| + |(V_{-1}w, w)|) < (\frac{1}{\sqrt{2}})2 = \sqrt{2}$. It follows that Re $e^{\frac{i\pi}{4}}(A^b v, v) =$ Re $(B_{\frac{\pi}{4}}w, w) < \sqrt{2}$. ∎

Consequently, by proposition 5.5 and lemma 5.8 we have the following result:

Theorem 5.9. *Let* $D = \{x + iy : x, y \in \mathbb{R}, |x| + |y| < 2\}$. *If* b *is pseudo-ergodic then* $D = W(A^b)$.

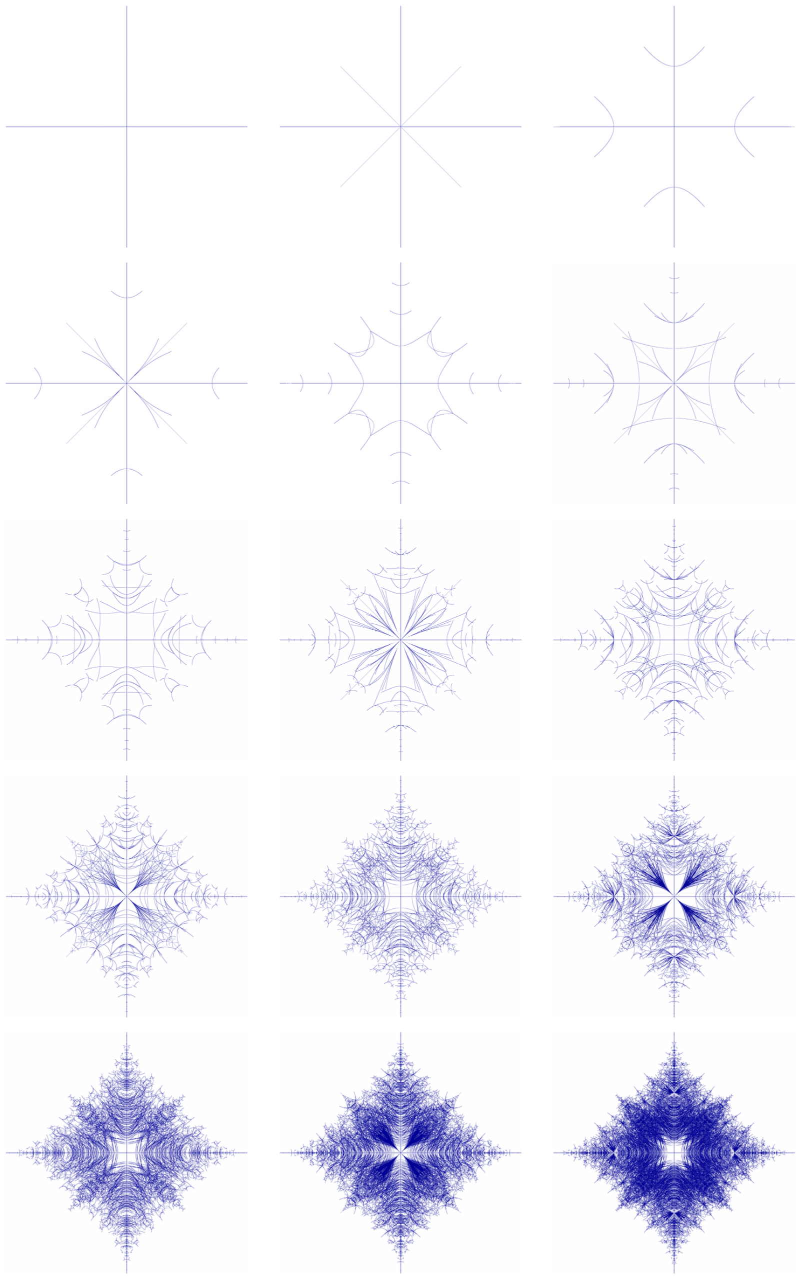

Figure 5.4: This figure shows the sets $\mathrm{spec}^n_{\mathrm{per}} A^b$, as defined in 5.9, for $n = 1, \ldots, 15$. Note that each set $\mathrm{spec}^n_{\mathrm{per}} A^b$ consists of k analytic arcs, where $2^n/n \leq k \leq 2^n$. Recall that Fig.5.3 shows the union of the first twelve pictures and thirteen pictures, respectively, of this figure.

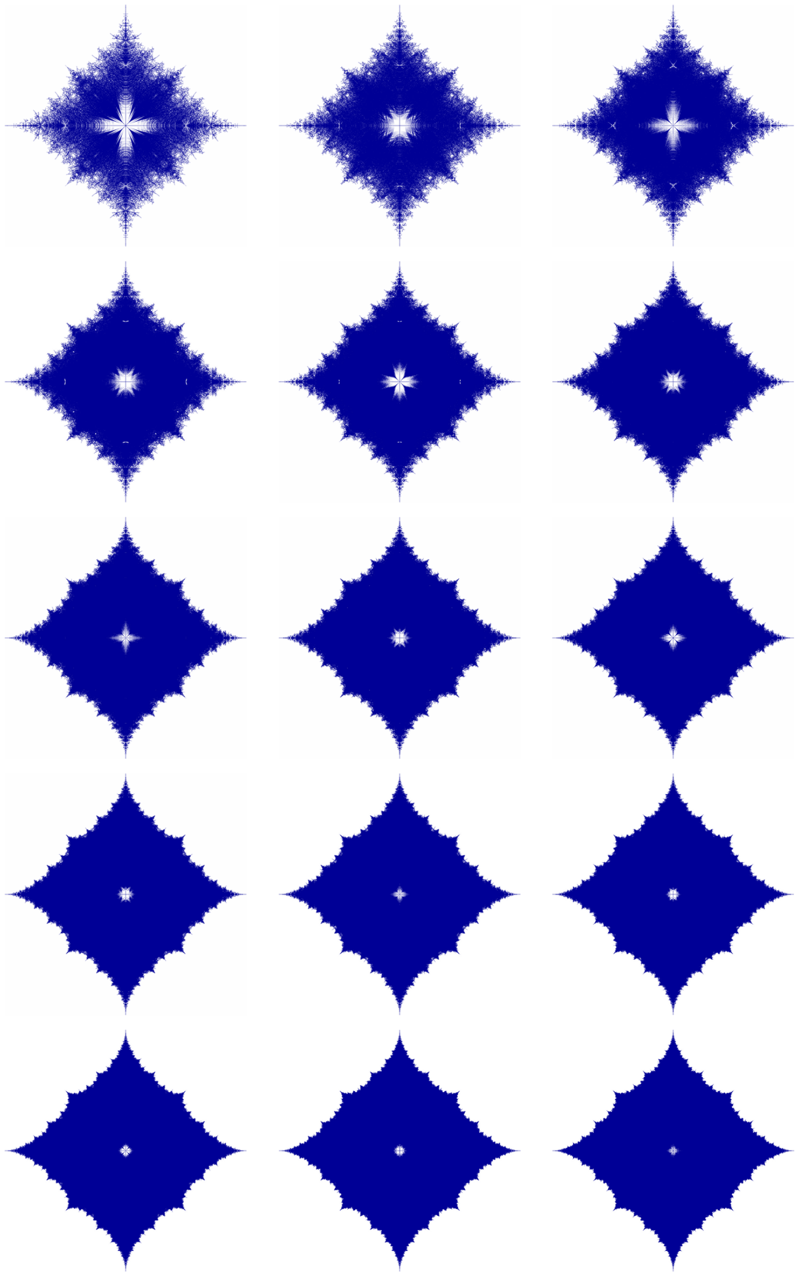

Figure 5.5: This figure shows the sets $\mathrm{spec}^n_{\mathrm{per}} A^b$, as defined in 5.9, for $n = 16, \ldots, 30$. Note that each set $\mathrm{spec}^n_{\mathrm{per}} A^b$ consists of k analytic arcs, where $2^n/n \le k \le 2^n$.

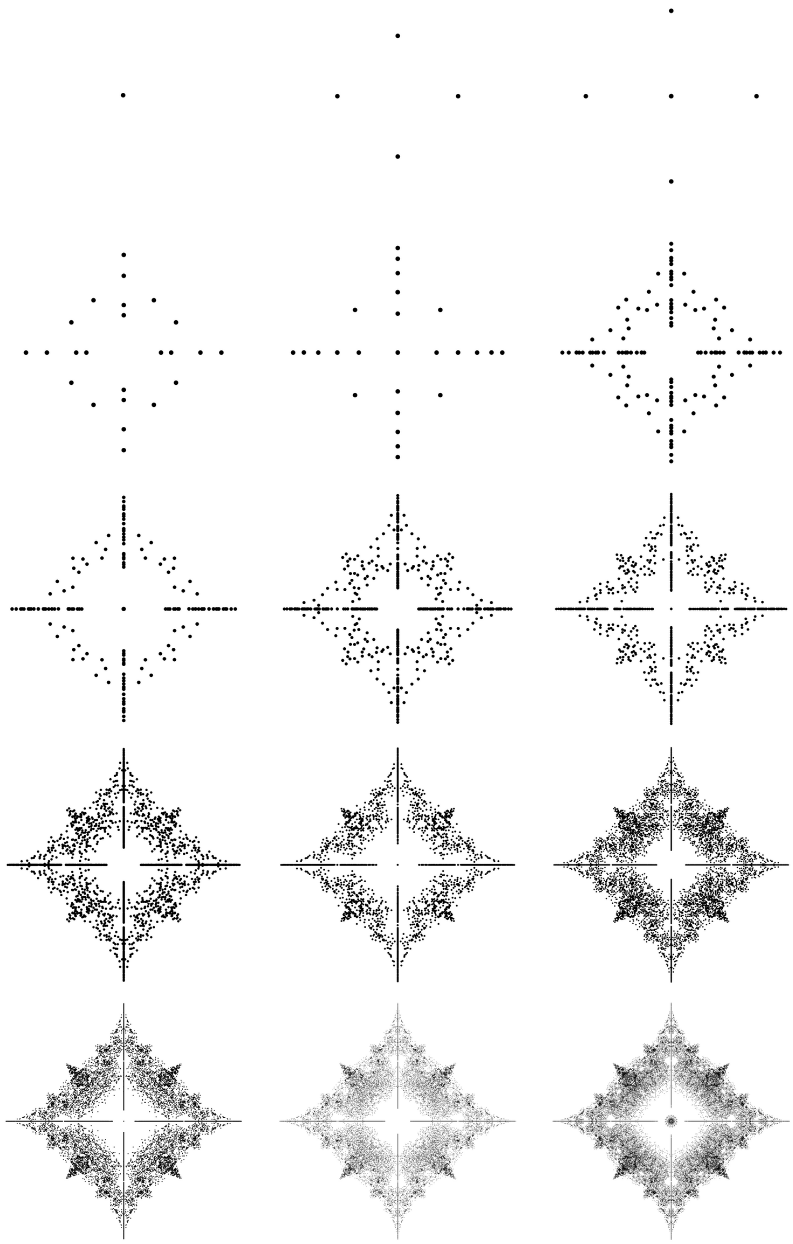

Figure 5.6: This figure shows the union $\bigcup_{c\in\{\pm1\}^{n-1}} \operatorname{spec} A_n^c$ of all $n \times n$ matrix eigenvalues for $n = 1, \ldots, 15$. Note that the first picture we have used heavier pixels for the sake of visibility.

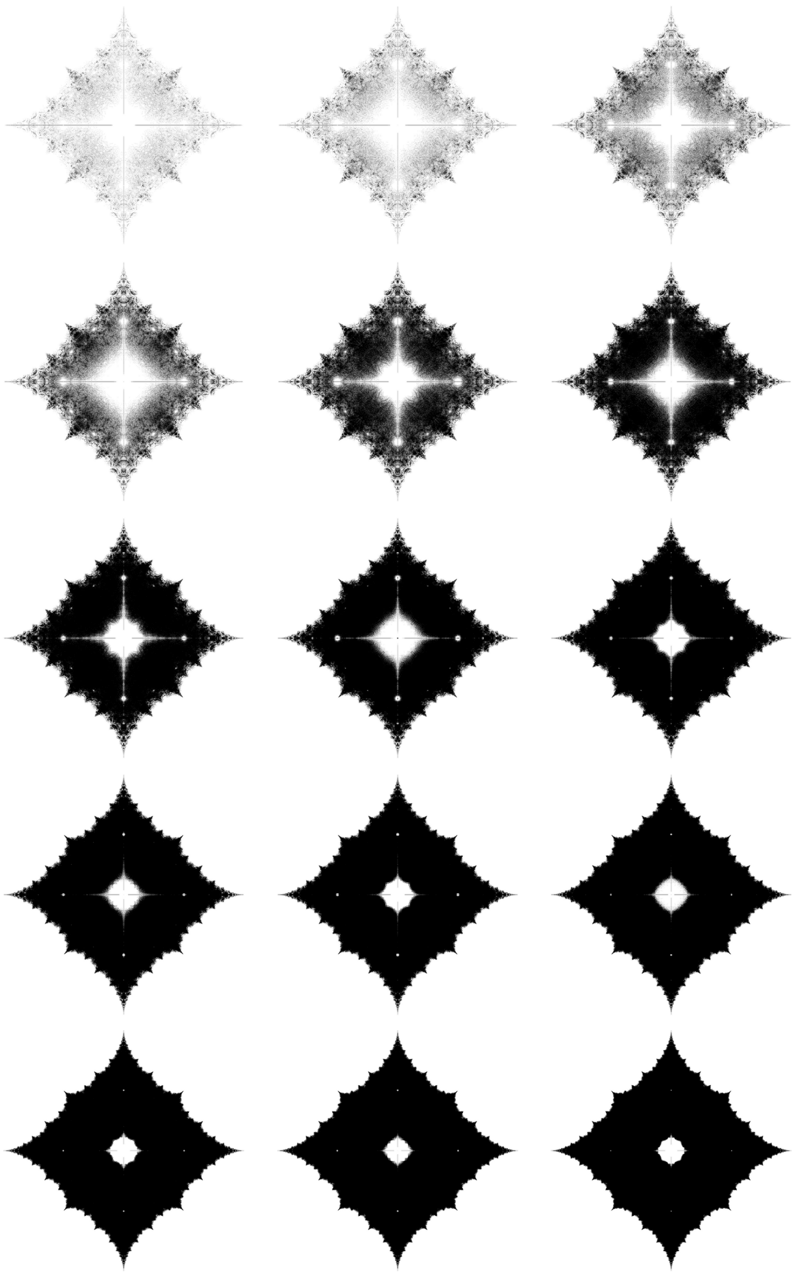

Figure 5.7: This figure shows the union $\bigcup_{c\in\{\pm 1\}^{n-1}} \operatorname{spec} A_n^c$ of all $n \times n$ matrix eigenvalues for $n = 16, \ldots, 30$.

5.3 Spectral Theory of Periodic Matrices

Recall from Proposition 5.4, if b is pseudo-ergodic, we have the following result:

$$\operatorname{spec}(A^b) = \operatorname{spec}_{\text{ess}}(A^b) = \bigcup_{c\in\{\pm1\}^{\mathbb{Z}}} \operatorname{spec}^{\infty}_{\text{point}}(A^c). \tag{5.10}$$

Generally it is difficult to evaluate the rightmost term in (5.10) since it would be a very large computation and the point spectrum of A^c is sometimes difficult to determine. We are presenting the approach which has been used by Davies and co-workers which is to look at a large number of periodic limit operator of A^b. More precisely, one look at the subset $\bigcup_{B\in\sigma^{\text{op}}(A), n-periodic} \operatorname{spec}^{\infty}_{\text{point}} B$ of $\operatorname{spec} A$ for large values of $n \in \mathbb{N}$. We know (see e.g. Lindner [36, Theorem 5.37]) that $\operatorname{spec} B = \operatorname{spec}^{\infty}_{\text{point}} B$ if B is $n-$periodic and its computation reduces to the computation of the spectra of certain finite matrices by treating B as a block Laurent matrix with $n\times n$ block entries (see Section 5.3.2 and 5.3.3 for the known related results).

5.3.1 The Point Spectrum of the Matrix A^b

In this section we will discuss about the point spectrum of A^b when b is N-periodic. Let λ be an eigenvalue of A^b with a corresponding eigenvector $x \in \ell^\infty(\mathbb{Z})$. Then $A^b x = \lambda x, \lambda \neq 0$, i.e.

$$b_j x_{j-1} - \lambda x_j + x_{j+1} = 0, \qquad j \in \mathbb{Z}$$

i.e.

$$\begin{pmatrix} x_j \\ x_{j+1} \end{pmatrix} = M_j \begin{pmatrix} x_{j-1} \\ x_j \end{pmatrix}, \qquad j \in \mathbb{Z}$$

where $M_j = \begin{pmatrix} 0 & 1 \\ -b_j & \lambda \end{pmatrix}$. Note that $M_{j+N} = M_j$ for every $j \in \mathbb{Z}$. Then

$$\begin{pmatrix} x_{mN-1} \\ x_{mN} \end{pmatrix} = \widetilde{M} \begin{pmatrix} x_{(m-1)N-1} \\ x_{(m-1)N} \end{pmatrix}, \qquad m \in \mathbb{Z}$$

where

$$\begin{aligned}\widetilde{M} &= M_{mN-1}M_{mN-2}\dots M_{(m-1)N}\\ &= M_{N-1}M_{N-2}\dots M_0.\end{aligned}$$

Thus,

$$\begin{pmatrix} x_{mN-1} \\ x_{mN} \end{pmatrix} = \widetilde{M}^m \begin{pmatrix} x_{-1} \\ x_0 \end{pmatrix}, \qquad m \in \mathbb{Z}. \tag{5.11}$$

Lemma 5.10. *If C is an invertible 2×2 matrix and v is a non-zero 2×1 vector and $(C^m v)_{m \in \mathbb{Z}}$ is a bounded sequence, then C has an eigenvalue μ with $|\mu| = 1$.*

Proof. Let $C = XDX^{-1}$ be the Jordan normal form of C with

$$\text{either } D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ or } D = \begin{pmatrix} a & 0 \\ 1 & a \end{pmatrix}$$

depending on whether C is diagonalizable or not. If there exists a nonzero vector $v \in \mathbb{C}^2$ such that $C^m v = XD^m X^{-1} v$ remains bounded as $m \to \pm\infty$ then there is also a nonzero vector $w := X^{-1} v \in \mathbb{C}^2$, say $w = \begin{pmatrix} x \\ y \end{pmatrix}$, such that $D^m w = X^{-1} C^m v$ remains bounded as $m \to \pm\infty$.

Case 1. (Diagonalizable case). If D is the first one of the two matrices above then

$$D^m w = \begin{pmatrix} a^m & 0 \\ 0 & b^m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a^m x \\ b^m y \end{pmatrix}.$$

Since this is bounded as $m \to \pm\infty$ we know that both $|a^m x| = |a|^m |x|$ and $|b^m y| = |b|^m |y|$ remain bounded and hence ($|a| = 1$ or $x = 0$) and ($|b| = 1$ or $y = 0$) hold. Since $w \neq 0$, not both x and y are zero and at least one of $|a|$ and $|b|$ has to be equal to 1.

Case 2. (non-diagonalizable case). If D is the first one of the two matrices above then

$$D^m w = \begin{pmatrix} a^m & 0 \\ ma^{m-1} & a^m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a^m x \\ ma^{m-1}x + a^m y \end{pmatrix}.$$

Since this is bounded as $m \to \pm\infty$ we know that the first component $|a^m x| = |a|^m |x|$ is bounded as $m \to \pm\infty$. This implies that $|a| = 1$ or $x = 0$. If $|a| = 1$ we are finished and if $x = 0$ we get from the boundedness of the second component $|ma^{m-1}x + a^m y| = |a^m y| = |a|^m |y|$ that $|a| = 1$ since $y \neq 0$ by $w \neq 0$. ■

From Lemma 5.10, we know that $\widetilde{M}$ has an eigenvalue α with $|\alpha| = 1$. Let $\begin{pmatrix} h \\ k \end{pmatrix}$ be the corresponding eigenvector. Define the sequence $z = (z_i)_{i\in\mathbb{Z}}$ by

$$z_{-1} = h \text{ and } z_0 = k$$

which is satisfied

$$b_j z_{j-1} - \lambda z_j + y_{j+1} = 0, \qquad j \in \mathbb{Z}.$$

From (5.11), for $m \in \mathbb{Z}$,

$$\begin{pmatrix} z_{mN-1} \\ z_{mN} \end{pmatrix} = \tilde{M}^m \begin{pmatrix} z_{-1} \\ z_0 \end{pmatrix} = \alpha^m \begin{pmatrix} z_{-1} \\ z_0 \end{pmatrix}.$$

So, it is easy to see that $z \in \ell^\infty(\mathbb{Z})$. In particular, $\begin{pmatrix} z_{N-1} \\ z_N \end{pmatrix} = \alpha \begin{pmatrix} z_{-1} \\ z_0 \end{pmatrix}$. Thus

$$\begin{pmatrix} -\lambda & 1 & 0 & \dots & b_0\alpha^{-1} \\ b_1 & -\lambda & 1 & \dots & 0 \\ \vdots & b_2 & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & -\lambda & 1 \\ \alpha & 0 & \dots & b_{N-1} & -\lambda \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{N-2} \\ z_{N-1} \end{pmatrix} = 0$$

i.e.

$$(A_N^b + B_{N,\alpha} - \lambda I) \begin{pmatrix} z_0 \\ z_1 \\ \vdots \\ z_{N-2} \\ z_{N-1} \end{pmatrix} = 0$$

where, for $N \geq 2$,

$$A_N^b = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ b_1 & 0 & 1 & \dots & 0 \\ \vdots & b_2 & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & 0 & 1 \\ 0 & 0 & \dots & b_{N-1} & 0 \end{pmatrix}$$

and

$$B_{N,\alpha} = \begin{pmatrix} 0 & 0 & 0 & \dots & b_0\alpha^{-1} \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & 0 & 0 \\ \alpha & 0 & \dots & 0 & 0 \end{pmatrix}$$

$B_{N,\alpha}$ can be written in the form $(B_{N,\alpha})_{ij} = \delta_{i,1}\delta_{j,J}b_0\alpha^{-1} + \delta_{i,N}\delta_{j,1}\alpha$. This is an excellent way to write it because it works even for the periodicity $N = 1$, provided we set $A_1 = 0$. Let

$$A_{N,\alpha}^b := A_N^b + B_N^\alpha.$$

We have shown that, if $\lambda \in \operatorname{spec}_{\text{point}}^\infty (A^b)$ then $\lambda \in \operatorname{spec} A_{N,\alpha}^b$ for some α with $|\alpha| = 1$.

Conversely, let $\lambda \in \operatorname{spec} (A_{N,\alpha}^b)$ for some α with $|\alpha| = 1$. So, there exists a bounded non-zero vector $x = [x_0, x_1, \dots, x_{N-1}]^T$ such that $(A_{N,\alpha}^b - \lambda I)x = 0$. Then, defining $x_N = \alpha x_0$ and $x_{-1} = \alpha^{-1}x_{N-1}$, it follows that

$$b_j x_{j-1} - \lambda x_j + x_{j+1} = 0,$$

for $0 \leq j \leq N-1$. Define the bounded sequence y by $y_k = \alpha^m x_j$ if $k = mN+j$ for some $m \in \mathbb{Z}$ and some j in the range $0 \leq j \leq N-1$. Then, for $k \in \mathbb{Z}$,

$$b_k y_{k-1} - \lambda y_k + y_{k+1} = \alpha^m b_j x_{j-1} - \alpha^m \lambda x_j + \alpha^m x_{j+1} = 0,$$

i.e. there exists $y = (y_k)_{k\in\mathbb{Z}}$ such that $(A^b - \lambda I)y = 0$, i.e. $\lambda \in \operatorname{spec}_{\text{point}}^\infty A^b$.

Then, we have shown the following result.

Theorem 5.11. *If b is N-periodic with $N \in \mathbb{N}$ then*

$$\operatorname{spec} A^b = \operatorname{spec}_{\text{ess}} A^b = \operatorname{spec}_{\text{point}}^\infty A^b = \bigcup_{|\alpha|=1} \operatorname{spec} A_{N,\alpha}^b.$$

Example.

Case 1 the period $N = 1$ (Laurent Operators)

$$\begin{aligned}\operatorname{spec}(A^b_{N,\alpha}) &= \{\lambda : \det(A^b_{N,\alpha} - \lambda I) = 0\} \\ &= \{\lambda : b_0\alpha^{-1} + \alpha - \lambda = 0\}.\end{aligned}$$

Therefore

$$\operatorname{spec}^\infty_{\text{point}}(A^b) = \{b_0\alpha^{-1} + \alpha : |\alpha| = 1\}.$$

In particular, if $b_0 = 1$ then

$$\operatorname{spec}^\infty_{\text{point}}(A^b) = \{\alpha^{-1} + \alpha : |\alpha| = 1\} = \{e^{is} + e^{-is} : s \in \mathbb{R}\} = [-2, 2],$$

and if $b_0 = -1$ then

$$\operatorname{spec}^\infty_{\text{point}}(A^b) = \{\alpha^{-1} - \alpha : |\alpha| = 1\} = \{e^{is} - e^{-is} : s \in \mathbb{R}\} = [-2i, 2i].$$

Case 2 the period $N = 2$.

$$\begin{aligned}\operatorname{spec}(A^b_{N,\alpha}) &= \{\lambda : \begin{vmatrix} -\lambda & b_0\alpha^{-1} + 1 \\ b_1 + \alpha & -\lambda \end{vmatrix} = 0\} \\ &= \{\lambda : \lambda^2 - (b_1 + \alpha)(b_0\alpha^{-1} + 1) = 0\} \\ &= \{\lambda : \lambda^2 - (b_0 b_1 \alpha^{-1} + b_0 + b_1 + \alpha) = 0\} \\ &= \{\lambda : \lambda^2 = \frac{b_0 b_1}{\alpha} + b_0 + b_1 + \alpha\} \\ &= \{\lambda : \lambda^2 = \frac{\alpha^2 + (b_0 + b_1)\alpha + b_0 b_1}{\alpha}\}.\end{aligned}$$

Therefore,

$$\operatorname{spec}^\infty_{\text{point}}(A^b) = \{\pm\sqrt{\frac{(b_0 + \alpha)(b_1 + \alpha)}{\alpha}} : |\alpha| = 1\}.$$

In particular, if $b_0 = 1$ and $b_1 = -1$ then

$$\begin{aligned}
\operatorname{spec}^{\infty}_{\text{point}}(A^b) &= \{\pm\sqrt{\frac{(\alpha+1)(\alpha-1)}{\alpha}} : |\alpha| = 1\} \\
&= \{\pm\sqrt{\frac{\alpha^2-1}{\alpha}} : |\alpha| = 1\} \\
&= \{\pm\sqrt{\alpha-\alpha^{-1}} : |\alpha| = 1\} \\
&= \{\pm\sqrt{2i\sin\theta} : \theta \in \mathbb{R}\} \\
&= \{\pm e^{\frac{i\pi}{4}}\sqrt{p} : -2i \le p \le 2i\} \\
&= \pm e^{\frac{i\pi}{4}}[-\sqrt{2i}, \sqrt{2i}].
\end{aligned}$$

Case 3 the period $N = 3$.

$$\begin{aligned}
\operatorname{spec} A^b_{N,\alpha} &= \{\lambda : \begin{vmatrix} -\lambda & 1 & b_0\alpha^{-1} \\ b_1 & -\lambda & 1 \\ \alpha & b_2 & -\lambda \end{vmatrix} = 0\} \\
&= \{\lambda : -\lambda^3 - \alpha - b_0b_1b_2\alpha^{-1} + b_0\lambda + b_1\lambda + b_2\lambda = 0, \text{ where } |\alpha| = 1\} \\
&= \{\lambda : -\lambda^3 + (b_0+b_1+b_2)\lambda - (\alpha + b_0b_1b_2\alpha^{-1}) = 0 \text{ where } |\alpha| = 1\}.
\end{aligned}$$

If $b_i \in \{1, -1\}$ for every $i \in \mathbb{Z}$, it follows that

$$\operatorname{spec}^{\infty}_{\text{point}}(A^b) = \{\lambda : -\lambda^3 + (b_0+b_1+b_2)\lambda - p = 0, \text{ where } p \in [-2,2] \cup [-2i, 2i]\}.$$

In particular, if $b_0 = 1, b_1 = -1$ and $b_2 = -1$ then

$$\begin{aligned}
\operatorname{spec}^{\infty}_{\text{point}}(A^b) &= \{\lambda : -\lambda^3 - \lambda - p = 0, \text{ where } p \in [-2,2]\} \\
&= \{\lambda : -2 \le -\lambda^3 - \lambda \le 2\} \\
&= \{\lambda : -2 \le \lambda^3 + \lambda \le 2\}.
\end{aligned}$$

In the next section, we will prove the same Theorem 5.11 by a different technique.

5.3.2 Computing the spectra by using Floquet-Bloch Technique.

As we noted at the beginning of section 5.3 the characteristic of the spectra of A^b is Theorem 5.11 is not new, but is known as a standard part of Floquet-Bloch theory. In particular, it can be derived from the following result [Theorem 4.4.9, [13]]

Theorem 5.12. *Let $\mathcal{K}$ be a finite-dimensional inner product space and let $\mathcal{H} = \ell^2(\mathbb{Z}^N, \mathcal{K})$ be the space of all square-summable $\mathcal{K}$ valued sequences on $\mathbb{Z}^N$. Let $A : \mathcal{H} \to \mathcal{H}$ be defined by*

$$(Af)(n) := \sum_{m \in \mathbb{Z}^N} a_{n-m} f_m$$

where $a : \mathbb{Z}^N \to L(\mathcal{K})$ satisfies $\sum_{n \in \mathbb{Z}^N} \|a_n\| < \infty$. Then

$$\mathrm{spec}\,(A) = \mathrm{spec}_{\mathrm{ess}}\,(A) = \bigcup_{\theta \in [-\pi,\pi]^N} \mathrm{spec}\,(b(\theta))$$

where

$$b(\theta) := \sum_{n \in \mathbb{Z}^N} a_n e^{in\cdot\theta} \in L(\mathcal{K})$$

for all $\theta \in [-\pi, \pi]^N$.

Let us show how this theorem implies Theorem 5.11. In fact we shall show a slightly better result.

Let A be a bounded operator on $\ell^2(\mathbb{Z})$ with a tridiagonal matrix $[A] = [a_{mn}]_{m,n\in\mathbb{Z}}$, i.e. one that satisfies $a_{mn} = 0$ if $|m - n| > 1$. Suppose also that A is periodic with period N in the sense that $a_{m+N,n+N} = a_{mn}$ for all $m, n \in \mathbb{Z}$.

Let $\sigma : \mathbb{Z}\times\{0, 1, \dots, N-1\} \to \mathbb{Z}$ be the map $\sigma(m, j) := Nm+j$. It is easy to show that σ is injective, and by using Division Algorithm, we can show that σ is surjective. Therefore σ is bijective. Since $\sigma : \mathbb{Z} \times \{0, 1, \dots, N-1\} \to \mathbb{Z}$ is a bijection, clearly, the mapping $f : \ell^2(\mathbb{Z}, \mathbb{C}^N) \to \ell^2(\mathbb{Z})$ defined by

$$f(x) = \tilde{x},$$

for $x \in \ell^2(\mathbb{Z}, \mathbb{C}^N)$, where $\tilde{x} \in \ell^2(\mathbb{Z})$ is defined by $f(x)(\sigma(m,j)) = (x(m))_j$ or $\tilde{x}(\sigma(m,j)) = (f^{-1}(x)(m))_j$, $m \in \mathbb{Z},\ j \in \{0, 1, \dots N-1\}$ is a bijection and it is also isometric [1], and so an isometric isomorphism. Thus $\ell^2(\mathbb{Z}) \cong \ell^2(\mathbb{Z}, \mathbb{C}^N)$.

For $g \in \ell^2(\mathbb{Z})$,

$$(Ag)(m) := \sum_{n\in\mathbb{Z}} a_{mn} g_n, \quad m \in \mathbb{Z}.$$

Let $\tilde{A} \in L(\ell^2(\mathbb{Z}, \mathbb{C}^N))$ be defined by

$$\tilde{A} := f^{-1} A f.$$

Then, for every $\lambda \in \mathbb{C}$,

$$\lambda - \tilde{A} = f^{-1}(\lambda - A) f$$

so

$$\mathrm{spec}\,(A) = \mathrm{spec}\,(\tilde{A}).$$

Moreover, for $x \in \ell^2(\mathbb{Z}, \mathbb{C}^N)$, $m \in \mathbb{Z}, i \in \{0, 1, \dots, N-1\}$

$$\begin{aligned}
((\tilde{A}x)(m))_i &= (f\tilde{A}x)\sigma(m,i) \\
&= f(f^{-1}Afx)\sigma(m,i) \\
&= A\tilde{x}\sigma(m,i) \\
&= \sum_{n\in\mathbb{Z}} a_{\sigma(m,i),n}\tilde{x}_n \\
&= \sum_{n\in\mathbb{Z}} \sum_{j=0}^{N-1} a_{\sigma(m,i),\sigma(n,j)}\tilde{x}_{\sigma(n,j)} \\
&= \sum_{n\in\mathbb{Z}} \sum_{j=0}^{N-1} a_{\sigma(m,i),\sigma(n,j)}(x(n))_j \\
&= \sum_{n\in\mathbb{Z}} \sum_{j=0}^{N-1} (C_{m-n})_{ij}(x(n))_j \\
&= \sum_{n\in\mathbb{Z}} C_{m-n} x(m)
\end{aligned}$$

[1] Since $\|f(x)\| = \sqrt{\sum_{m\in\mathbb{Z}} \sum_{j=0}^{N-1} |\tilde{x}(\sigma(m,j))|^2} = \sqrt{\sum_{m\in\mathbb{Z}} \sum_{j=0}^{N-1} |(x(m))_j|^2} = \|x\|$

where $C_m \in L(\mathbb{C}^N)$ has matrix representation $(C_m)_{ij} := a_{\sigma(m,i),\sigma(0,j)} = a_{\sigma(m+N,i),\sigma(N,j)}$.
Since $a_{m,n} = 0$ if $|m-n| > 1$, it follows that $\sum_{n\in\mathbb{Z}} \|C_n\| = \|C_{-1}\| + \|C_0\| + \|C_1\| < \infty$.

Then, from Theorem 4.4.9 in [13],

$$\begin{aligned}\operatorname{spec}(A) &= \operatorname{spec}(\tilde{A})\\ &= \bigcup_{\theta\in[-\pi,\pi]} \operatorname{spec}(b(\theta))\\ &= \bigcup_{|\alpha|=1} \operatorname{spec}(C_{-1}\alpha + C_0 + C_1\alpha^{-1})\end{aligned}$$

define $a_i := a_{i,i+1}, b_i := a_{i,i}, c_i := a_{i+1,i}$ and let $C_{J,\alpha} = C_{-1}\alpha + C_0 + C_1\alpha^{-1}$, i.e.

$$\begin{aligned}C_{N,\alpha} &= \begin{pmatrix} 0 & 0 & \dots & 0 & 0\\ 0 & 0 & 0 & \dots & 0\\ \vdots & 0 & \ddots & \ddots & \vdots\\ 0 & \dots & \ddots & 0 & 0\\ a_0 & 0 & \dots & 0 & 0\end{pmatrix}\alpha + \begin{pmatrix} b_0 & a_1 & 0 & \dots & 0\\ c_1 & b_1 & a_2 & \dots & 0\\ \vdots & c_2 & \ddots & \ddots & \vdots\\ 0 & \dots & \ddots & b_{N-2} & a_{N-1}\\ 0 & 0 & \dots & c_{N-1} & b_{N-1}\end{pmatrix}\\ &+ \begin{pmatrix} 0 & 0 & \dots & 0 & c_0\\ 0 & 0 & 0 & \dots & 0\\ \vdots & 0 & \ddots & \ddots & \vdots\\ 0 & \dots & \ddots & 0 & 0\\ 0 & 0 & \dots & 0 & 0\end{pmatrix}\alpha^{-1}\\ &= \begin{pmatrix} b_0 & a_1 & 0 & \dots & c_0\alpha^{-1}\\ c_1 & b_1 & a_2 & \dots & 0\\ \vdots & c_2 & \ddots & \ddots & \vdots\\ 0 & \dots & \ddots & b_{N-2} & a_{N-1}\\ a_0\alpha & 0 & \dots & c_{N-1} & b_{N-1}\end{pmatrix}.\end{aligned}$$

We have shown the following result:

Theorem 5.13.

$$\operatorname{spec}(A) = \bigcup_{|\alpha|=1} \operatorname{spec}(C_{N,\alpha}),$$

where $C_{N,\alpha} = C_N + D_N^{\alpha}$ *such that*

$$C_N = \begin{pmatrix} b_1 & a_1 & 0 & \dots & 0 \\ c_2 & b_2 & a_2 & \dots & 0 \\ \vdots & c_3 & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & b_{N-1} & a_{N-1} \\ 0 & 0 & \dots & c_N & b_N \end{pmatrix}$$

and $(D_N^{\alpha})_{mn} = \delta_{m,1}\delta_{n,N}c_0\alpha^{-1} + \delta_{m,N}\delta_{n,1}a_0\alpha.$

Clearly, Theorem 5.11 follows from the result which implies that

$$\operatorname{spec}(A^b) = \bigcup_{|\alpha|=1} \operatorname{spec}(A^b_{N,\alpha}).$$

5.3.3 An alternative view point

From the view point of S. Prössdorf and B. Silbermann [41], by using the concept of isomorphism to rearrange the entries of the matrix A^b, we can consider our matrix as a finite matrix with each entry is Laurent. Then we can compute the spectrum by using the concept of symbol of Laurent operators.

$$A^b = \begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & 1 & & & & \\ & b_{-1} & 0 & 1 & & & \\ & & b_0 & 0 & 1 & & \\ & & & b_1 & 0 & 1 & \\ & & & & b_2 & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix}$$

with $b = (b_i)_{i\in\mathbb{Z}}$ periodic, i.e. $b_{j+N} = b_j, j \in \mathbb{Z}$. Since the mapping $f : \ell^2(\mathbb{Z}) \to (\ell^2(\mathbb{Z}))^N$, defined by

$$\begin{pmatrix} \vdots \\ x_{-1} \\ x_0 \\ x_1 \\ \vdots \end{pmatrix} \mapsto \left(\begin{pmatrix} \vdots \\ x_1 \\ x_{N+1} \\ \vdots \end{pmatrix}, \begin{pmatrix} \vdots \\ x_2 \\ x_{N+2} \\ \vdots \end{pmatrix}, \dots, \begin{pmatrix} \vdots \\ x_N \\ x_{2N} \\ \vdots \end{pmatrix} \right),$$

is a bijection and also isometric[2], hence $\ell^2(\mathbb{Z}) \cong (\ell^2(\mathbb{Z}))^N$. Let $\tilde{A}^b = fA^bf^{-1}$ be defined by

$$\tilde{A}^b \begin{pmatrix} x(1) \\ x(2) \\ \vdots \\ x(N) \end{pmatrix} = \begin{pmatrix} \tilde{A_{11}} & \tilde{A_{12}} & \cdots & \tilde{A_{1N}} \\ \tilde{A_{21}} & \tilde{A_{22}} & \cdots & \tilde{A_{2N}} \\ \vdots & \ddots & \ddots & \vdots \\ \tilde{A_{N1}} & \tilde{A_{N2}} & \cdots & \tilde{A_{NN}} \end{pmatrix} \begin{pmatrix} x(1) \\ x(2) \\ \vdots \\ x(N) \end{pmatrix},$$

each $\tilde{A_{ij}}$ is Laurent then $\tilde{A}^b : (\ell^2(\mathbb{Z}))^N \to (\ell^2(\mathbb{Z}))^N$.

From [41, section 4.98], we know that, if A is a Fredholm operator, then $(\det \operatorname{smb} A)(t) \neq 0$ for all $t \in \mathbb{T}$

If $N = 2$

$$\tilde{A}^b - \lambda I = \left(\begin{array}{ccccc|ccccc} \ddots & & & & & \ddots & & & & \\ & -\lambda & & & & \ddots & 1 & & & \\ & & -\lambda & & & & -1 & 1 & & \\ & & & -\lambda & & & & -1 & 1 & \\ & & & & \ddots & & & & \ddots & \ddots \\ \hline \ddots & \ddots & & & & \ddots & & & & \\ & 1 & 1 & & & & -\lambda & & & \\ & & 1 & 1 & & & & -\lambda & & \\ & & & 1 & \ddots & & & & -\lambda & \\ & & & & \ddots & & & & & \ddots \end{array} \right)$$

[2] Since $\|f(x)\| = \sqrt{\sum_{j\in\mathbb{Z}} |x_j|^2} = \sqrt{\sum_{i=0}^{N-1} \sum_{m\in\mathbb{Z}} \left|x_{\sigma(m,i)}\right|^2} = \sqrt{\sum_{i=0}^{N-1} \|x(i)\|^2} = \|x\|$

We have then that

$$\text{smb}\,(\tilde{A}^b - \lambda I) = \begin{pmatrix} -\lambda & 1-t \\ 1+\frac{1}{t} & -\lambda \end{pmatrix}.$$

If $N \geq 3$, then we have

$$\text{smb}(\tilde{A}^b - \lambda I) = \begin{pmatrix} -\lambda & 1 & 0 & \dots & \dots & a_0 t \\ a_1 & -\lambda & 1 & \dots & \dots & 0 \\ 0 & a_2 & -\lambda & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & a_{N-2} & -\lambda & 1 \\ \frac{1}{t} & 0 & \dots & \dots & a_{N-1} & -\lambda \end{pmatrix} (= A^b_{N,\alpha} - \lambda I \quad \text{for } \alpha = \frac{1}{t}).$$

Since $\text{spec}\,(\tilde{A}^b) = \text{spec}\,(A^b)$ and from Theorem 5.11, it follows that

$$\begin{aligned} \{\lambda \in \mathbb{C} : \tilde{A}^b - \lambda I \text{ is not invertible }\} &= \text{spec}\,(\tilde{A}^b) \\ &= \text{spec}\,(A^b) \\ &= \bigcup_{|\alpha|=1} \text{spec}\,(A^b_{N,\alpha}) \\ &= \bigcup_{|\alpha|=1} \{\lambda \in \mathbb{C} : \det(A^b_{N,\alpha} - \lambda I) = 0\} \\ &= \{\lambda \in \mathbb{C} : \det(\text{smb}\tilde{A}^b - \lambda I) = 0 \text{ for some } t \in \mathbb{T}\}. \end{aligned}$$

In the next section, we will construct a sequence $b \in \{\pm 1\}^{\mathbb{Z}}$ for which $\text{spec}^{\infty}_{\text{point}}\, A^c$ contains the open unit disk. As a consequence of (5.7) and the closedness of spectra, this shows that $\text{spec}\, A^b$ contains the closed unit disk.

5.4 Eigenvalue Problem meets Sierpinski Triangle

Fix $\lambda \in \mathbb{C}$. In this section we are looking for a sequence $b \in \{\pm 1\}^{\mathbb{Z}}$ such that $\lambda \in \text{spec}^{\infty}_{\text{point}}\, A^b$; that is, there exists a non-zero $v \in \ell^{\infty}(\mathbb{Z})$ with $A^b v = \lambda v$,

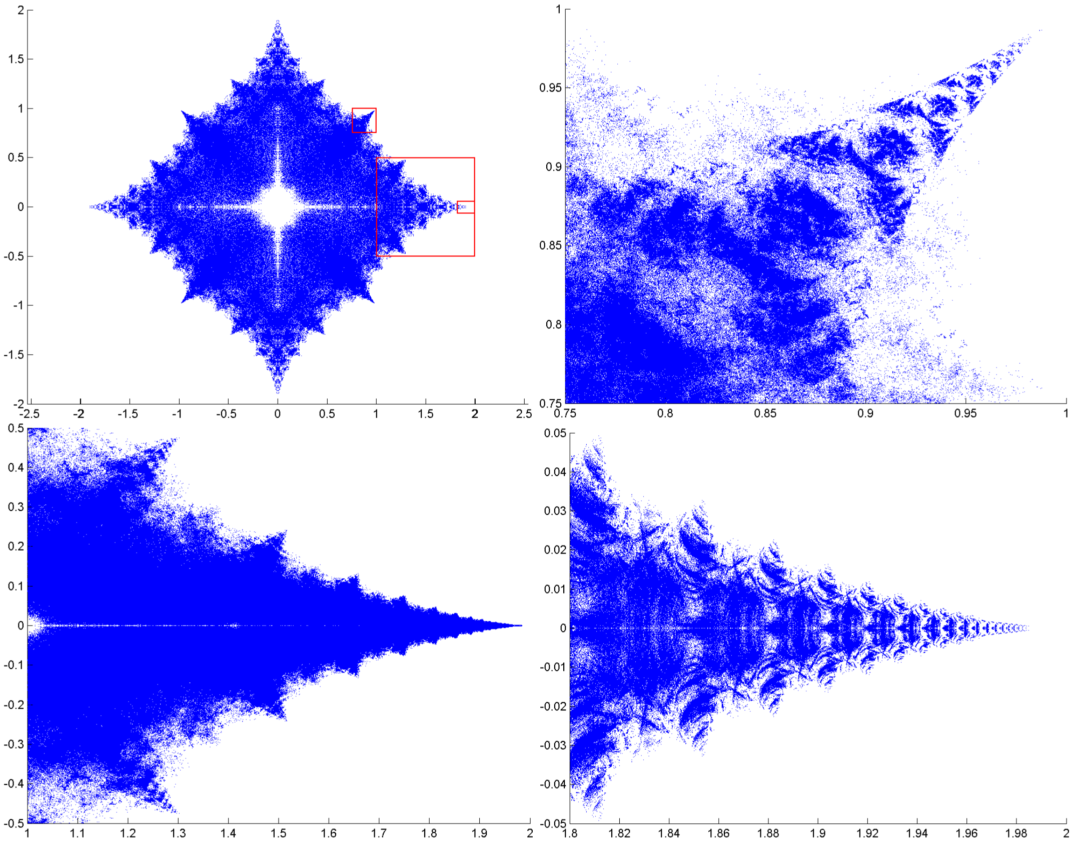

Figure 5.8: This figure illustrates a zoom into the set $\bigcup_{c\in\{\pm 1\}^{24}} \operatorname{spec} A^c_{25}$, i.e. the 10th picture of Fig.5.7

i.e.

$$v(i+1) = \lambda v(i) - b(i)v(i-1) \tag{5.12}$$

for every $i \in \mathbb{Z}$.

Starting from $v(-1) = 1$, $v(0) = 0$ and $v(1) = 1$, we will successively use (5.12) to compute $v(i)$ and $b(i)$ for $i = 2, 3, \ldots$ (an analogous procedure is possible for $i = -2, -3, \ldots$). Doing so we get $b(0) = -1$, $v(2) = \lambda$, $b(1) = 1$, $v(3) = \lambda^2 + 1$, etc. In general, $v(i)$ is a polynomial of degree $i - 1$ in λ. Since we want v to be a bounded sequence, we are trying to keep the coefficients of these polynomials small. So our strategy will be to choose $b(2), b(3), \ldots \in \{\pm 1\}$ such that each $v(i)$ is a polynomial in λ with coefficients in $\{-1, 0, 1\}$. The following table, where we abbreviate -1 by $-$, $+1$ by $+$, and 0 by a space, shows that this seems to be possible.

		j $\rightarrow$						coefficients of λ^{j-1} in the polynomial $v(i)$										
i	$b(i)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	$\cdots$
1	$+$	$+$																
2	$-$		$+$															
3	$+$	$+$		$+$														
4	$-$				$+$													
5	$+$	$+$		$+$		$+$												
6	$+$		$+$				$+$											
7	$-$	$-$				$-$		$+$										
8	$+$								$+$									
9	$-$	$+$				$+$		$-$		$+$								
10	$-$		$+$				$+$				$+$							
11	$+$	$+$		$+$		$+$				$+$		$+$						
12	$+$				$+$								$+$					
13	$-$	$-$		$-$						$-$		$-$		$+$				
14	$+$		$-$								$-$				$+$			
15	$-$	$+$								$+$				$-$		$+$		
16	$+$																$+$	
$\vdots$	$\vdots$	$\vdots$																

(5.13)

For $i, j \in \mathbb{N}$, we denote the coefficient of λ^{j-1} in the polynomial $v(i)$ by $p(i,j)$. Then the right part of table (5.13) shows the values $p(i,j)$ for $i, j = 1, ..., 16$. From (5.12) it follows that

$$p(i+1, j) \;=\; p(i, j-1) \;-\; b(i)p(i-1, j) \tag{5.14}$$

holds for $i = 2, 3, ...$ and $j = 1, 2, ..., i$ with $p(i', j') := 0$ if $j' < 1$ or $j' > i'$.

If, for some i, j, one has that $p(i, j-1) \neq 0$ and $p(i-1, j) \neq 0$ then, by (5.14) and $p(i+1, j) \in \{-1, 0, 1\}$, this implies that

$$b(i) \;=\; p(i, j-1)/p(i-1, j) \;=\; p(i, j-1) \cdot p(i-1, j) \tag{5.15}$$

since otherwise $p(i+1, j) \in \{-2, 2\}$. As an example, look at $p(15, 1) = 1$ and $p(14, 2) = -1$. For $b(15) = 1$, we would get from (5.14) that $p(16, 2) = 2 \notin$

$\{-1,0,1\}$, so it remains to take $b(15)=-1=p(15,1)\cdot p(14,2)$. The same value $b(15)=-1$ is enforced by $p(15,9)$ and $p(14,10)$, as well as by $p(15,13)$ and $p(14,14)$. We will prove that this coincidence, i.e. that the right-hand side of (5.15) is (if non-zero) independent of j, is not a matter of fortune. As a result we get that the table (5.13) continues without end, only using values from $\{-1,0,1\}$ for $p(i,j)$ and from $\{\pm 1\}$ for $b(i)$. To prove this, we employ a particular self-similarity in the triangular pattern of (5.13); more precisely, it can be shown that the pattern of non-zero entries of $p(\cdot,\cdot)$ forms a so-called infinite discrete Sierpinski triangle.

Proposition 5.14. *For every $i\in\mathbb{N}$, there exist $b(i)$, $c(i)\in\{\pm 1\}$ such that*

(i) it holds that

$$\begin{pmatrix} p(2i-1,2j-1) & p(2i-1,2j) \\ p(2i,2j-1) & p(2i,2j) \end{pmatrix} = \begin{cases} p(i,j)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } i+j \text{ is even,} \\ c(i)\,p(i-1,j)\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} & \text{if } i+j \text{ is odd} \end{cases}$$

for every $j=1,\dots,i$, and

(ii) $p(i,j-1)\cdot p(i-1,j)\in\{0,b(i)\}$ for all $j=2,\dots i-1$.

So in particular, by (i), all values $p(i,j)$ are in $\{-1,0,1\}$.

As an immediate consequence we get the following result for which we note that the table (5.13) can be extended to negative values of i in a similar fashion.

Corollary 5.15. *For the sequence $b\in\{\pm 1\}^{\mathbb{Z}}$ from Proposition 5.14, it holds that the closed unit disk $\overline{\mathbb{D}}:=\{z\in\mathbb{C}:|z|\le 1\}$ is contained in $\operatorname{spec}A^b$. Consequently, for every pseudo-ergodic $c\in\{\pm 1\}^{\mathbb{Z}}$, one has $\overline{\mathbb{D}}\subset\operatorname{spec}A^c$.*

Proof. Let $\lambda\in\mathbb{D}:=\{z\in\mathbb{C}:|z|<1\}$ and let b be the sequence from Proposition 5.14. Then, for every $i\in\mathbb{Z}$,

$$|v(i)| = \left|\sum_{j=1}^{|i|} p(i,j)\lambda^{j-1}\right| \le \sum_{j=1}^{|i|} |p(i,j)|\,|\lambda^{j-1}| \le \sum_{j=1}^{\infty} |\lambda|^{j-1} = \frac{1}{1-|\lambda|},$$

showing that $v \in \ell^\infty(\mathbb{Z})$, and, by our construction (5.12), $A^b v = \lambda v$. So $\mathbb{D} \subset \operatorname{spec}^\infty_{\text{point}} A^b \subset \operatorname{spec} A^b$. Since $\operatorname{spec} A^b$ is closed, it holds that $\overline{\mathbb{D}} \subset \operatorname{spec} A^b$. The claim for a pseudo-ergodic c then follows from $\operatorname{spec} A^b \subset \operatorname{spec} A^c$. ■

Proof of Proposition 5.14 Firstly, it is easy to see (by (5.13), (5.14) and induction) that $p(i', j') = 0$ if $i' + j'$ is odd, whence $p(2i-1, 2j) = 0 = p(2i, 2j-1)$ for all i, j.

We will now prove (i) and (ii) by induction over $i \in \mathbb{N}$. Therefore, let (i) be satisfied for $i = 1, ..., k$, and let (ii) be satisfied for $i = 1, ..., 2k$. (The base case is easily verified by looking at table (5.13)). We will then prove (i) for $i = k+1$ and (ii) for $i = 2k+1$ and $2k+2$.

Part (i). We let $i = k+1$ and start with the case when $i+j$ is even. By (5.14), we have that

$$\begin{aligned} p(2i-1, 2j-1) &= p(2i-2, 2j-2) - b(2i-2) \cdot p(2i-3, 2j-1) \\ &= p(2(i-1), 2(j-1)) \\ &\quad -b(2i-2) \cdot p(2(i-1)-1, 2j-1), \end{aligned} \tag{5.16}$$

where, by induction (and since $i-1+j-1$ is even), $p(2(i-1), 2(j-1)) = p(i-1, j-1)$ if $j > 1$ and it is 0 if $j = 1$. Also by induction, $p(2(i-1) - 1, 2j-1) = c(i-1)p(i-2, j)$ since $i-1+j$ is odd. To determine $b(2i-2)$, take $J \in \{1, ..., 2i-4\}$ such that $p(2i-2, J) \neq 0$ (whence $J =: 2j'$ has to be even) and $p(2i-3, J+1) \neq 0$ (if no such J exists then we are free to choose $b(2i-2)$ in which case we will put $b(2i-2) := c(i-1)b(i-1)$). From (i) and $0 \neq p(2i-2, J) = p(2(i-1), 2j')$ it is clear that $i-1+j'$ is even and $i-1+j'+1$ is odd. Now, by (ii) and (i), we have that

$$\begin{aligned} b(2i-2) &= p(2i-2, J)\, p(2i-3, J+1) \\ &= p(2(i-1), 2j')\, p(2(i-1)-1, 2(j'+1)-1) \\ &= p(i-1, j')\, c(i-1)\, p(i-2, j'+1) \;=\; c(i-1)\, b(i-1). \end{aligned}$$

Inserting all these results in (5.16), we get that

$$\begin{aligned} p(2i-1,2j-1) &= \begin{cases} p(i-1,j-1)-c(i-1)b(i-1)c(i-1)p(i-2,j) & \text{if } j>1, \\ 0-c(i-1)b(i-1)c(i-1)p(i-2,j) & \text{if } j=1 \end{cases} \\ &= \left.\begin{cases} p(i-1,j-1)-b(i-1)p(i-2,j) & \text{if } j>1, \\ -b(i-1)p(i-2,j) & \text{if } j=1 \end{cases}\right\} = p(i,j). \end{aligned}$$

We already saw that $p(2i-1,2j)=0=p(2i,2j-1)$ for all i,j; so we are left with

$$p(2i,2j)=p(2i-1,2j-1)-b(2i-1)p(2i-2,2j)=p(i,j)-b(2i-1)0=p(i,j).$$

Now suppose $i+j$ is odd. Then, almost exactly as above,

$$\begin{aligned} p(2i-1,2j-1) &= p(2i-2,2j-2)-b(2i-2)p(2i-3,2j-1) \\ &= 0-c(i-1)b(i-1)p(i-1,j) \;=\; c(i)p(i-1,j) \end{aligned}$$

with $c(i):=-c(i-1)b(i-1)$, and

$$\begin{aligned} p(2i,2j) &= p(2i-1,2j-1)-b(2i-1)p(2i-2,2j) \\ &= c(i)p(i-1,j)-b(2i-1)p(i-1,j) \end{aligned} \tag{5.17}$$

since $i-1+j$ is even. To determine $b(2i-1)$, we again take $J\in\{1,...,2i-3\}$ such that $p(2i-1,J)\neq 0$ (whence $J=:2j'-1$ is odd) and $p(2i-2,J+1)\neq 0$ (if no such J exists then we are free to choose $b(2i-1)$ in which case we will put $b(2i-1):=c(i)$). From (i) and $0\neq p(2i-2,J+1)=p(2(i-1),2j')$ it is clear that $i-1+j'$ is even and $i+j'$ is odd. Now, by (ii) and (i), we have that

$$\begin{aligned} b(2i-1) &= p(2i-1,J)\,p(2i-2,J+1)=p(2i-1,2j'-1)\,p(2(i-1),2j') \\ &= c(i)\,p(i-1,j')\,p(i-1,j')=c(i). \end{aligned}$$

Inserting this into (5.17), we get that

$$p(2i,2j) \;=\; c(i)p(i-1,j)-c(i)p(i-1,j) \;=\; 0.$$

Part (ii). Let $i=2k+1$ and suppose $j\in\{2,...,2k\}$ is such that $p(i,j-1)\neq 0$ (whence $i+j-1$ is even, i.e. $j=:2j'$ is even) and $p(i-1,j)\neq 0$. If no such j exists then the product in (ii) is always zero and there is nothing to show. From $0 \neq p(i-1,j) = p(2k,2j')$ and (i) we get that $k+j'$ is even and $k+1+j'$ is odd. Now we have that

$$\begin{aligned} p(i,j-1)\,p(i-1,j) &= p(2k+1,2j'-1)\,p(2k,2j') \\ &= p(2(k+1)-1,2j'-1)\,p(2k,2j') \\ &= c(k+1)\,p(k,j')\,p(k,j') \;=\; c(k+1) \;=:\; b(i) \end{aligned}$$

is independent of j. Now let $i = 2k+2$ and suppose $j \in \{2,...,2k+1\}$ is such that $p(i,j-1) \neq 0$ (whence $i+j-1$ is even, i.e. $j =: 2j'+1$ is odd) and $p(i-1,j) \neq 0$. (Again, if no such j exists then there is nothing to show.) From $0 \neq p(i,j-1) = p(2k+2,2j') = p(2(k+1),2j')$ and (i) we get that $k+1+j'$ is even and $k+1+j'+1$ is odd. Now we have that

$$\begin{aligned} p(i,j-1)\,p(i-1,j) &= p(2k+2,2j')\,p(2k+1,2j'+1) \\ &= p(2(k+1),2j')\,p(2(k+1)-1,2(j'+1)-1) \\ &= p(k+1,j')\,c(k+1)\,p(k,j'+1) \\ &= c(k+1)\,b(k+1) \;=:\; b(i) \end{aligned}$$

is independent of j.

5.5 The Relationship between the Spectrum of Singly Tridiagonal Random Matrices and the Spectrum of Doubly Tridiagonal Random Matrices.

In this section, we are going to prove that the spectrum of one sided infinite matrices, A^b_+ is equal the spectrum of two sided infinite matrices, A^b. Let us introduce $A^{b,c}$ as a new notation for an infinite tridiagonal matrix of the form

$$\begin{pmatrix} \ddots & \ddots & & & & & \\ \ddots & 0 & c_{-2} & & & & \\ & b_{-1} & 0 & c_{-1} & & & \\ & & b_0 & 0 & c_0 & & \\ & & & b_1 & 0 & c_1 & \\ & & & & b_2 & 0 & \ddots \\ & & & & & \ddots & \ddots \end{pmatrix} \tag{5.18}$$

where $b, c \in \ell^\infty(\mathbb{Z})$.

Theorem 5.16. *If b is pseudo-ergodic then*

$$\operatorname{spec} A^b = \operatorname{spec} A^b_+,$$

where

$$A^b_+ = \begin{pmatrix} 0 & 1 & & & \\ b_1 & 0 & 1 & & \\ & b_2 & 0 & 1 & \\ & & b_3 & 0 & \ddots \\ & & & \ddots & \ddots \end{pmatrix}.$$

Proof. ($\subseteq$) Let us consider this infinite dimensional block matrix

$$\widetilde{A}^b_+ = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A^b_+ \end{array}\right]$$

We have then that $\operatorname{spec} \widetilde{A}^b_+ = \{0\} \cup \operatorname{spec} A^b_+ = \operatorname{spec} A^b_+$. In the next step, we are investigating the behaviour of the essential spectrum of $\widetilde{A}^b_+$. Since b is "+–pseudo-ergodic", from [8], we have

$$\operatorname{spec}_{\text{ess}} A^b_+ = \bigcup_{B \in \sigma^{\mathsf{op}}(A^b_+)} \operatorname{spec} B. \tag{5.19}$$

Since $\sigma^{\mathsf{op}}(A^b_+) = \left\{A^c : c \in \{\pm 1\}^{\mathbb{Z}}\right\}$, it follows that

$$\operatorname{spec}_{\text{ess}} A^b_+ \supseteq \operatorname{spec} A^b.$$

Hence

$$\operatorname{spec} A^b \subseteq \operatorname{spec} A^b_+.$$

($\supseteq$) Let $\lambda \in \operatorname{spec} A^b_+$. If $\lambda \in \operatorname{spec}_{\text{ess}} A^b_+$ then $\lambda \in \operatorname{spec} A^b$ by (5.7) and (5.19). So let $\lambda \in \operatorname{spec} A^b_+ \setminus \operatorname{spec}_{\text{ess}} A^b_+$. By Theorem 2.12, λ is an eigenvalue of A^b_+, i.e., $A^b_+ - \lambda I$ is not injective or $\overline{\lambda}$ is an eigenvalue of $\left(A^b_+\right)^*$, i.e., $\left(A^b_+ - \lambda I\right)^*$ is not injective, since the range of $\left(A^b_+ - \lambda I\right)$ is closed if $\lambda \notin \operatorname{spec}_{\text{ess}} A^b_+$.

If λ is an eigenvalue of A^b_+ with the corresponding eigenvector $(x_0, x_1, x_2, \ldots)^T$, then

$$A^b_+ x - \lambda x = \begin{pmatrix} -\lambda & 1 & & & \\ b_1 & -\lambda & 1 & & \\ & b_2 & -\lambda & 1 & \\ & & b_3 & -\lambda & \ddots \\ & & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \\ \vdots \end{pmatrix}.$$

Then

$$\left(\begin{array}{cccc|c|cccc} \ddots & \ddots & & & & & & & \\ \ddots & -\lambda & b_2 & & & & & & \\ & 1 & -\lambda & b_1 & & & & & \\ & & 1 & -\lambda & \mathbf{1} & & & & \\ \hline & & & \mathbf{-1} & -\lambda & \mathbf{1} & & & \\ \hline & & & & \mathbf{1} & -\lambda & 1 & & \\ & & & & & b_1 & -\lambda & 1 & \\ & & & & & & b_2 & -\lambda & \ddots \\ & & & & & & & \ddots & \ddots \end{array}\right) \begin{pmatrix} \vdots \\ x_2 \\ x_1 \\ x_0 \\ \hline 0 \\ \hline x_0 \\ x_1 \\ x_2 \\ \vdots \end{pmatrix} = 0,$$

i.e.

$$\left(A^{\widetilde{c},\widetilde{d}} - \lambda I\right) \widetilde{x} = 0$$

where

$$\widetilde{c}_i = \begin{cases} 1 & \text{if } i \leq -2 \text{ or } i = 0, \\ -1 & \text{if } i = -1 \\ b_i & \text{if } i \geq 1 \end{cases},$$

$$\widetilde{d}_i = \begin{cases} 1 & \text{if } i \geq 0, \\ b_{-i} & \text{if } i \leq -1 \end{cases},$$

and

$$\widetilde{x}_i = \begin{cases} x_i & \text{if } i \geq 0\ , \\ x_{-i-2} & \text{if } i \leq -2 \\ 0 & \text{if } i = -1 \end{cases}.$$

Then we have $\widetilde{c}, \widetilde{d} \in \{\pm 1\}^{\mathbb{Z}}$, so that $A^{\widetilde{c},\widetilde{d}} \in \sigma^{\mathsf{op}}(A^{c,d})$ if c, d are pseudo-ergodic ± 1 sequences. We have then that $\lambda \in \operatorname{spec}^{\infty}_{\text{point}} A^{\widetilde{c},\widetilde{d}} \subseteq \operatorname{spec} A^{c,d} \subseteq \operatorname{spec} A^b$.

Note that from $\lambda \in \operatorname{spec} A^b$, we obtain, by (5.7) and (5.19), $\lambda \in \operatorname{spec}_{\text{ess}} A^b_+$, which contradicts our assumption, implying that this second case is impossible. Therefore $\operatorname{spec} A^b_+ = \operatorname{spec}_{\text{ess}} A^b_+ = \operatorname{spec} A^b = \operatorname{spec}_{\text{ess}} A^b$ ■

5.6 Lower Bound on $\operatorname{spec}(A^b)$

As an auxiliary step, we will study matrices of the form $A^{b,c}$ with $b, c \in \ell^\infty(\mathbb{Z})$. It can be noticed that, for every $x \in \ell^\infty(\mathbb{Z})$ and for each $i \in \mathbb{Z}$

$$\begin{aligned} \left(A^{b,c}x\right)_i &= b_i x_{i-1} + c_i x_{i+1} \\ &= b_i \left(V_1 x\right)_i + c_i \left(V_{-1} x\right)_i \\ &= \left(M_b V_1 x\right)_i + \left(M_c V_{-1} x\right)_i \end{aligned}$$

So,

$$A^{b,c} = M_b V_1 + M_c V_{-1}.$$

Lemma 5.17. *If* $b, c \in \{\pm 1\}^{\mathbb{Z}}$, *then*

$$\operatorname{spec} A^{b,c} = \operatorname{spec} A^e,$$

where $e = bV_1 c \in \{\pm 1\}^{\mathbb{Z}}$.

Proof. Choose $a \in \{\pm 1\}^{\mathbb{Z}}$ such that

$$a_{i+1} = a_i c_i, \qquad i \in \mathbb{Z}$$

i.e. so that

$$V_{-1}a = ac.$$

(E.g. choose $a_0 = 1$, and set

$$a_{i+1} := a_i c_i \qquad i = 0, 1, 2, \ldots$$

and

$$a_i := a_{i+1}c_i, \qquad i = -1, -2, \ldots).$$

Then, noting that $M_a^{-1} = M_a$,

$$\begin{aligned} M_a A^{b,c} M_a^{-1} &= M_a \left(M_b V_1 + M_c V_{-1} \right) M_a \\ &= M_{ab} V_1 M_a + M_{ac} V_{-1} M_a \\ &= M_{abV_1 a} V_1 + M_{acV_{-1}a} V_{-1} \\ &= M_e V_1 + V_{-1} = A^e, \end{aligned}$$

since $M_{acV_{-1}a} = M_{(ac)^2} = I$, where $e := abV_1 a = bV_1 c$, since $V_1 a V_1 c = a$, it follows that $V_1 a = aV_1 c$. Therefore $\operatorname{spec} A^{b,c} = \operatorname{spec} A^e$. ∎

Corollary 5.18. *If $b, c \in \{\pm 1\}^{\mathbb{Z}}$ and $A^{b,c}$ is symmetric, i.e. $b = V_1 c$, then* $\operatorname{spec} A^{b,c} = [-2, 2]$.

Proof. Note that $A^{b,c}$ is symmetric if

$$b_i = c_{i-1}, \qquad i \in \mathbb{Z}$$

i.e.

$$b = V_1 c.$$

It follows that,

$$(bV_1 c)_i = b_i (V_1 c)_i = b_i c_{i-1} = 1 \qquad \text{for every } i \in \mathbb{Z}.$$

Hence, $\operatorname{spec} A^{b,c} = [-2, 2]$. ∎

We agree to call the pair $(b,c) \in \{\pm 1\}^{\mathbb{Z}} \times \{\pm 1\}^{\mathbb{Z}}$ *pseudo-ergodic* if $\sigma^{\mathsf{op}}(A^{b,c}) = \{A^{d,e} : d,e \in \{\pm 1\}^{\mathbb{Z}}\}$, which implies that both b and c are pseudo-ergodic sequences but also means that they are so, roughly speaking, independently of each other.

We are now proving some results on the resolvent norm of the infinite matrix. Then we will prove the result of the inclusion sets of the spectrum of periodic matrices.

Lemma 5.19. *If b and c are pseudo-ergodic then $\|(A^b - \lambda I)^{-1}\| = \|(A^c - \lambda I)^{-1}\|$ for all $\lambda \in \mathbb{C}$.*

Proof. This follows immediately from Theorem 5.12 (ix) in [9] and the fact that A^b is a limit operator of A^c, and vice versa. ∎

Corollary 5.20. *If b and c are pseudo-ergodic then* $\operatorname{spec} A^b = \operatorname{spec} A^c$ *and* $\operatorname{spec}_\varepsilon A^b = \operatorname{spec}_\varepsilon A^c$ *for all $\varepsilon > 0$.*

Remark 5.21. **a)** Note that both claims also follow immediately from Theorem 6.28 (v) and (vi) in [9].

b) Lemma 5.19 and Corollary 5.20 show that, from the (pseudo)spectral point of view, we can talk about "the pseudo-ergodic operator A^b" without having a particular pseudo-ergodic sequence b in mind since residual norm, spectrum and pseudospectrum of A^b don't depend on b – as long as it is pseudo-ergodic. □

Note that for all $\lambda \in \mathbb{C}$, if (b,c) and d are pseudo-ergodic then by Lemma 5.19,

$$\begin{aligned} \|(A^{b,c} - \lambda I)^{-1}\| &= \|M_a(A^{b,c} - \lambda I)^{-1} M_a^{-1}\| = \|(M_a A^{b,c} M_a^{-1} - \lambda I)^{-1}\| \\ &= \|(A^e - \lambda I)^{-1}\| = \|(A^d - \lambda I)^{-1}\|, \end{aligned} \tag{5.20}$$

where $a_{i+1} = a_i c_i$ and $e = bV_1 c$ as defined in the proof of Lemma 5.17.

Corollary 5.22. *If (b,c) and d are pseudo-ergodic then*

$$\operatorname{spec} A^{b,c} = \operatorname{spec} A^d \qquad \textit{and} \qquad \operatorname{spec}_\varepsilon A^{b,c} = \operatorname{spec}_\varepsilon A^d \qquad \forall \varepsilon > 0.$$

Theorem 5.23. *If $b \in \{\pm 1\}^{\mathbb{Z}}$ is a pseudo-ergodic sequence and $n \in \mathbb{N}$, then*

$$\operatorname{spec} A_n^b \subseteq \operatorname{spec} A^b.$$

Proof. Let $\lambda \in \operatorname{spec} A_n^b$ and $x = (x_0, x_1, \ldots, x_{n-1})^T$ such that $A_n^b x = \lambda x$. So,

$$A_n^b x - \lambda x = \begin{pmatrix} -\lambda & 1 & & & & \\ b_1 & -\lambda & 1 & & & \\ & b_2 & -\lambda & 1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & b_{n-2} & -\lambda & 1 \\ & & & & b_{n-1} & -\lambda \end{pmatrix}_{n\times n} \begin{pmatrix} x_0 \\ x_1 \\ \\ \vdots \\ x_{n-2} \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

But then, by using a simple reflection technique, we get

$$\left(\begin{array}{c|c|cccc|c|cccc|c|cc} \ddots & & & & & & & & & & & & & \\ -\lambda & \mathbf{1} & & & & & & & & & & & & \\ \hline \mathbf{-1} & -\lambda & \mathbf{1} & & & & & & & & & & & \\ \hline & \mathbf{1} & -\lambda & 1 & & & & & & & & & & \\ & & b_1 & -\lambda & \ddots & & & & & & & & & \\ & & & \ddots & \ddots & 1 & & & & & & & & \\ & & & & b_{n-1} & -\lambda & \mathbf{1} & & & & & & & \\ \hline & & & & & \mathbf{-1} & -\lambda & \mathbf{1} & & & & & & \\ \hline & & & & & & \mathbf{1} & -\lambda & b_{n-1} & & & & & \\ & & & & & & & 1 & \ddots & \ddots & & & & \\ & & & & & & & & \ddots & -\lambda & b_1 & & & \\ & & & & & & & & & 1 & -\lambda & \mathbf{1} & & \\ \hline & & & & & & & & & & \mathbf{-1} & -\lambda & \mathbf{1} & \\ \hline & & & & & & & & & & & \mathbf{1} & -\lambda & \\ & & & & & & & & & & & & & \ddots \end{array}\right) \left(\begin{array}{c} \vdots \\ x_0 \\ \hline 0 \\ \hline x_0 \\ x_1 \\ \vdots \\ x_{n-1} \\ \hline 0 \\ \hline x_{n-1} \\ \vdots \\ x_1 \\ x_0 \\ \hline 0 \\ \hline x_0 \\ \vdots \end{array}\right) = 0.$$

If we call this new infinite matrix $A^{\tilde{c},\tilde{d}} - \lambda I$ then we have $\tilde{c}, \tilde{d} \in \{-1, 1\}^{\mathbb{Z}}$, so that $A^{\tilde{c},\tilde{d}} \in \sigma^{\mathsf{op}}(A^{c,d})$ if c, d are pseudo-ergodic ± 1 sequences. But we have just seen that

$$\lambda \in \operatorname{spec}^{\infty}_{\text{point}} A^{\tilde{c},\tilde{d}} \subseteq \operatorname{spec} A^{c,d} = \operatorname{spec} A^b.$$

■

The fact that the RHS of the inclusion $\operatorname{spec} A_n^b \subseteq \operatorname{spec} A^b$ does not depend on the pseudo-ergodic sequence b (see Corollary 5.20), whereas the LHS clearly does, shows that this formula can be easily improved. In fact, we can say something about the RHS without looking at the actual entries of the sequence b (see Remark 5.21 b):

Corollary 5.24. *If b is pseudo-ergodic and $n \in \mathbb{N}$ then*

$$\bigcup_{c_1,\ldots,c_{n-1}\in\{\pm 1\}} \operatorname{spec} A_n^c \subseteq \operatorname{spec} A^b.$$

Proof. This easily follows from Proposition 5.23 and $\operatorname{spec} A_n^c \subset \operatorname{spec} A^c = \operatorname{spec} A^b$ for all pseudo-ergodic $c \in \{\pm 1\}^{\mathbb{Z}}$, by Corollary 5.20. ∎

We now carry on with the analogues of Proposition 5.23 and Corollary 5.24 for pseudospectra. In fact, we have the following inequality for resolvent norms.

Proposition 5.25. *If b is pseudo-ergodic and $n \in \mathbb{N}$ then, for all $\lambda \in \mathbb{C}$,*

$$\|(A_n^b - \lambda I_n)^{-1}\| \;\le\; \|(A^b - \lambda I)^{-1}\|,$$

independent of the parameter $p \in [1,\infty]$ of the underlying space $\ell^p(\mathbb{Z})$.

Proof. Put $M := \|(A_n^b - \lambda I_n)^{-1}\|$, which we will understand as $M = \infty$ if $\lambda \in \operatorname{spec} A_n^b$. From Proposition 5.23 we know that in this case $\lambda \in \operatorname{spec} A^b$, so that also the RHS is ∞ and, in this sense, the inequality holds. So we are left with $M < \infty$, i.e., $\lambda \notin \operatorname{spec} A_n^b$. Then, for every $\delta > 0$, there exists an $x = (x_1, ..., x_n)^\top \in \mathbb{C}^n$ such that $\|x\| = 1$ and $y := (A_n^b - \lambda I_n)x$ has $\|y\| < \dfrac{1}{M-\delta}$. Now put $c, d \in \{\pm 1\}^{\mathbb{Z}}$ such that $A^{c,d}$ is the matrix that we studied in the proof of Proposition 5.23. Here our current proof has to bifurcate depending on $p \in [1,\infty]$.

Case 1: $p = \infty$

Define $\widetilde{x} \in \ell^\infty(\mathbb{Z})$ exactly as in the proof of Proposition 5.23. Then $\widetilde{y} := (A^{b,c} - \lambda I)\widetilde{x}$ is of the form $\widetilde{y} = (\cdots, y^\top, 0, (J_n y)^\top, 0, y^\top, 0, (J_n y)^\top, \cdots)^\top \in \ell^\infty(\mathbb{Z})$ and $\|\widetilde{y}\|_\infty = \|y\|_\infty$, as well as $\|\widetilde{x}\|_\infty = \|x\|_\infty$, so that

$$\left\|(A^{c,d} - \lambda I)^{-1}\right\| \ge \frac{\|\widetilde{x}\|_\infty}{\|\widetilde{y}\|_\infty} = \frac{\|x\|_\infty}{\|y\|_\infty} > M - \delta.$$

Case 2: $p < \infty$

For any $m \in \mathbb{N}$, let $\widetilde{x}^{(m)}$ be the sequence $\widetilde{x}$ from case 1, but with all entries of index outside $\{-m(n+1), \ldots, m(n+1)\}$ put to zero (where we suppose that at index zero there is one of the 0 entries of $\widetilde{x}$). Then $\widetilde{y}^{(m)} := (A^{b,c} - \lambda I)\widetilde{x}^{(m)}$ is the same as $\widetilde{y}$ from Case 1 for entries with index between $-m(n+1)+1$ and $m(n+1)-1$, it is zero outside $\{-m(n+1)-1, \dots, m(n+1)+1\}$ and it has $\widetilde{y}^{(m)}(-m(n+1)) = x_1$ if m is even and x_n if m is odd, $\widetilde{y}^{(m)}(m(n+1)) = -x_1$ if m is even and $-x_n$ if m is odd.

As a result, we have

$$\|\widetilde{x}^{(m)}\|_p^p = 2m\,\|x\|_p^p \qquad \text{and} \qquad \|\widetilde{y}^{(m)}\|_p^p = 2m\,\|y\|_p^p + \begin{cases} 2|x_1|^p & \text{if } m \text{ is even,} \\ 2|x_n|^p & \text{if } m \text{ is odd.} \end{cases}$$

From $\|x\|_p = 1$, $\|y_p\| < \frac{1}{M-\delta}$ and $|x_1|, |x_n| \leq \|x\|_p = 1$ we hence get that $\|\widetilde{x}^{(m)}\| = 2m$ and $\|\widetilde{y}\|_p^p < 2m\frac{1}{(M-\delta)^p} + 2$, so that

$$\|(A^{c,d} - \lambda I)^{-1}\|^p \;\geq\; \frac{\|\widetilde{x}^{(m)}\|_p^p}{\|\widetilde{y}^{(m)}\|_p^p} \;>\; \frac{1}{\frac{2m}{(M-\delta)^p} + 2} \;=\; \frac{1}{\frac{1}{(M-\delta)^p} + \frac{1}{m}}.$$

In either case, since these inequalities hold for all $\delta > 0$ and all $m \in \mathbb{N}$, we get that $\|(A^{c,d} - \lambda I)^{-1}\| \geq M = \|(A_n^b - \lambda I_n)^{-1}\|$. The claim now follows by choosing (e, f) pseudo-ergodic and noting that

$$\left\|(A^b - \lambda I)^{-1}\right\| = \left\|(A^{e,f} - \lambda I)^{-1}\right\| \geq \left\|(A^{c,d} - \lambda I)^{-1}\right\| \geq \left\|(A_n^b - \lambda I_n)^{-1}\right\|$$

by (5.20) and the fact $A^{c,d}$ is a limit operator of $A^{e,f}$. ∎

Similar to the discussion before, the RHS of the inequality in Proposition 5.25 does not depend on the pseudo-ergodic sequence b (by Lemma 5.19), whereas the LHS does. So, we have the following strengthening:

Corollary 5.26. *If b is pseudo-ergodic and $n \in \mathbb{N}$ then, for every $\varepsilon > 0$,*

$$\sup_{c_1, c_2, \dots, c_{n-1} \in \{\pm 1\}} \|(A_n^c - \lambda I)^{-1}\| \;\leq\; \|(A_n^b - \lambda I)^{-1}\|.$$

Proof. For any $n \in \mathbb{N}$ and pseudo-ergodic c, we have

$$\|(A_n^c - \lambda I_n)^{-1}\| \;\leq\; \|(A^c - \lambda I)^{-1}\| \;=\; \|(A^b - \lambda I)\|,$$

by Proposition 5.25 and Lemma 5.19. ∎

Corollary 5.27. *If b is pseudo-ergodic and $n \in \mathbb{N}$ then, for every $\varepsilon > 0$,*

$$\bigcup_{c_1,c_2,\ldots,c_{n-1}\in\{\pm1\}} \mathrm{spec}_\varepsilon A_n^c \;\subseteq\; \mathrm{spec}_\varepsilon A^b,$$

and in particular

$$\mathrm{spec}_\varepsilon A_n^b \;\subseteq\; \mathrm{spec}_\varepsilon A^b.$$

We may regard Corollary 5.27 as a generalisation of Corollary 5.24 from the case $\varepsilon = 0$ to $\varepsilon \geq 0$.

So far in this section we have established lower bounds on spectrum and pseudospectrum of A^b in terms of the same quantities for the finite matrix A_n^b. We will now look at upper bounds on $\mathrm{spec}\, A^b$ and $\mathrm{spec}_\varepsilon A^b$ in terms of certain pseudospectra of the finite matrix.

5.7 Computing the Inclusion sets using Method 1 and their numerical results.

In this section, we are now applying Corollary 3.3 and Corollary 3.9 to the infinite random tri-diagonal matrix A^b, to approximate the spectrum of this operator. Then we have

Corollary 5.28.

$$\mathrm{spec}\, A^b \subseteq \overline{\bigcup_{b\in\{\pm1\}^{\mathbb{Z}}} \mathrm{spec}_{\,f(n)} A_n^b}$$

where

$$f(n) = 4\sin\left(\frac{\theta}{2}\right),$$

and θ is the unique solution in the range $\left[\frac{\pi}{2n+1}, \frac{\pi}{n+1}\right)$ of the equation

$$2\sin\left(\frac{t}{2}\right)\cos\left(\left(n+\frac{1}{2}\right)t\right) + \frac{1}{4}\sin\left((n-1)\,t\right) = 0.$$

Firstly, we have considered the plots of the finite subsection for the every possible ± 1- sequences. The following pictures show us the spectrum (eigenvalues) of the finite $n \times n$ matrices,

$$A_n^b = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ b_1 & 0 & 1 & \dots & 0 \\ \vdots & b_2 & \ddots & \ddots & \vdots \\ 0 & \dots & \ddots & 0 & 1 \\ 0 & \dots & \dots & b_{n-1} & 0 \end{pmatrix},$$

where $n = 1, \dots, 30$.

We are now showing the convergence of the inclusions set for spec A^b using method 1, i.e., the union of pseudospectra of $n \times n$ principal submatrices of A^b. We start off by letting $\Sigma_n = \overline{\bigcup_{c \in \{\pm 1\}^{n-1}} \operatorname{spec}_{f(n)} A_n^c}$. Then we know that

$$\Sigma_n = \overline{\Sigma_{f(n)}^n(A^b)}.$$

From Corollary 3.3 and Corollary 5.27, we can see that

$$\operatorname{spec} A^b \subseteq \Sigma_n \subseteq \overline{\operatorname{spec}_{f(n)} A^b}.$$

Then, by Corollary 2.34, we have the following corollary to see the convergence when we apply method 1 to the operator A^b.

Corollary 5.29. *If $b \in \{\pm 1\}^{\mathbb{Z}}$ is pseudo-ergodic then*

$$\Sigma_n \to \operatorname{spec} A^b,$$

as $n \to \infty$ in the Haussdorff metric.

Fig.5.9 illustrates the spectral inclusion sets Σ_n (the black area) where $n = 2, \dots, 18$. The red area is the closure of the numerical range of the operator A^b and the blue area is the union of all lower bounds on spec A^b. In order to demonstrate that Σ_n does not converge to the numerical range $W(A^b)$, we have some numerical evidence to show that eventually Σ_n invades the numerical range at the point $\lambda = 1.5+0.5i$, starting from $n = 34$, so that $1.5 + 0.5i \notin \Sigma_{34}$. Therefore λ cannot be part of spec A^b but it's numerically too expensive to compute and draw a whole image of Σ_{34}. Then, we have a conjecture on the picture of the spec A^b as in the Fig.5.10.

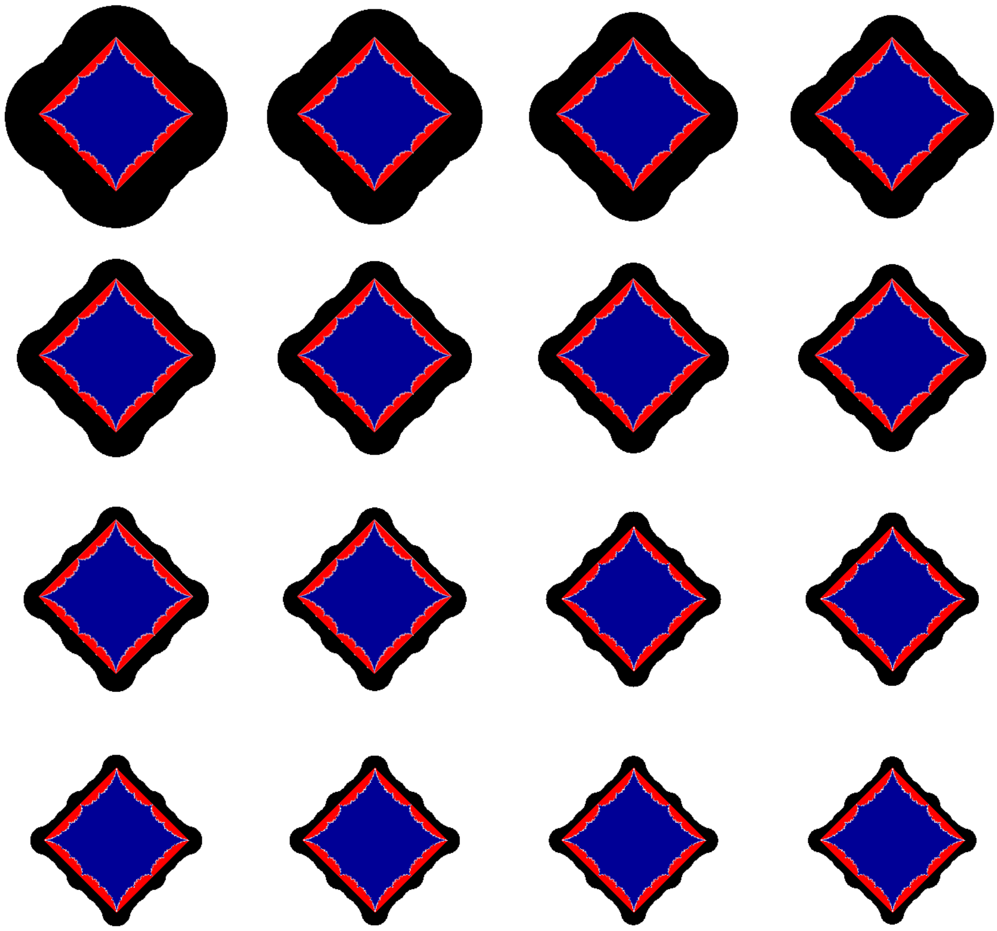

Figure 5.9: Plots of the sets Σ_n, which are inclusion sets for spec A^b, where $f(n)$ is given as in Corollary 5.28. A_n^b is the ordinary finite submatrix given by (5.3) of the operator $A^b = M_bV_1 + V_{-1}$, when $b \in \{\pm 1\}^{\mathbb{Z}}$ is pseudo-ergodic. Shown are the inclusion sets when $n = 3, 4, \ldots, 18$.

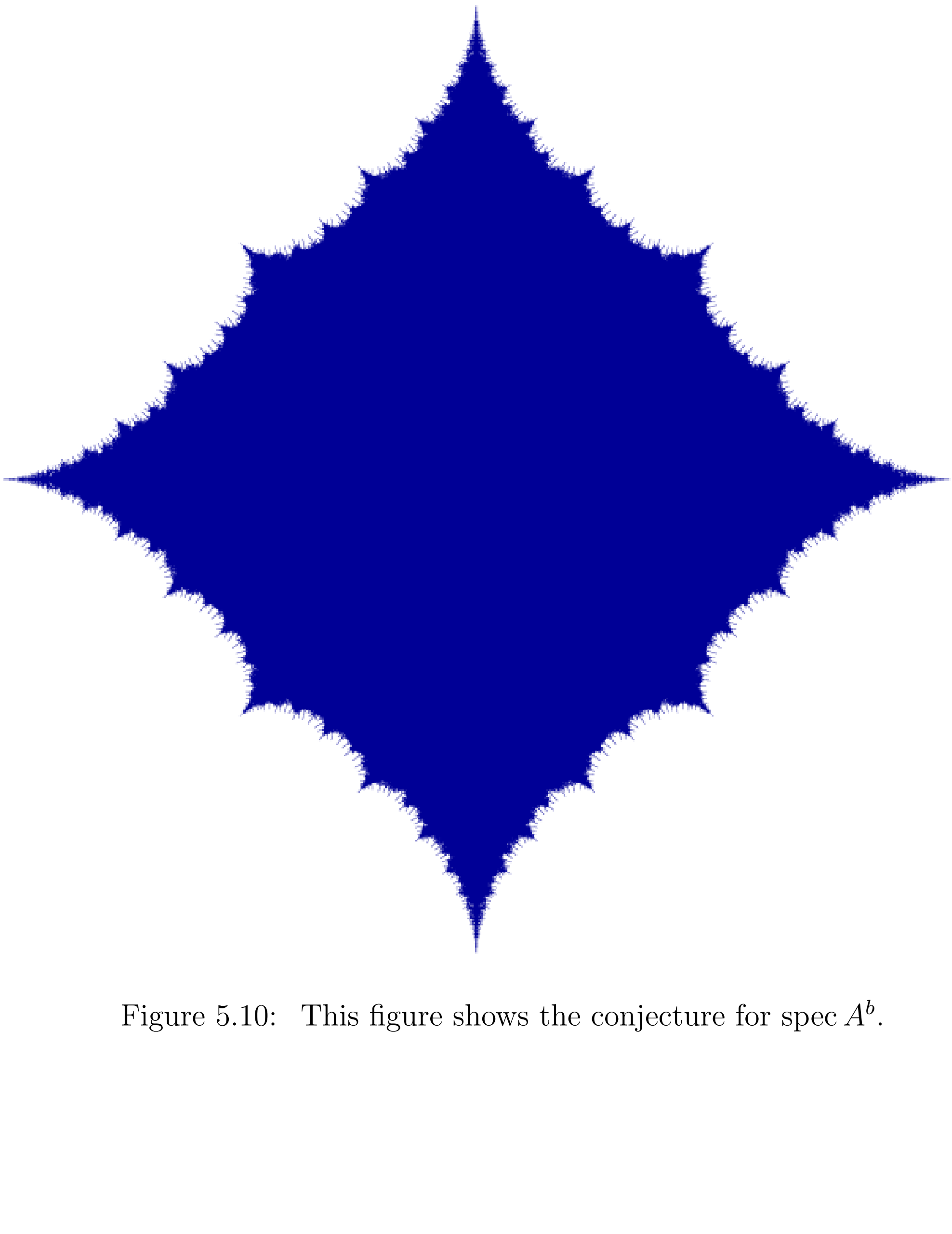

Figure 5.10: This figure shows the conjecture for spec A^b.

Bibliography

[1] P. W. Anderson: Absence of diffusion in certain random lattices. *Phys. Rev.,* 109:1492-1505, 1958.

[2] S. Boyd and C.A. Desoer: Subharmonic functions and performance bounds on linear time-invariant feedback systems, *IMA Journal of Mathematical control and Information*, **2** (1985), 153–170.

[3] A. Böttcher and S. M. Grudsky: *Spectral Properties of Banded Toeplitz Matrices*, SIAM, Philadelphia 2005.

[4] A. Böttcher and B. Silbermann: *Analysis of Toeplitz Operators*, 1st ed. Akademie-Verlag 1989 and Springer 1990; 2nd ed. Springer 2006.

[5] E. Brézin and A. Zee: Non-hermitean delocalization: Multiple scattering and bounds, *Nucl. Phys. B* **509** (1998), 599–614.

[6] P. W. Brouwer, P. G. Silvestrov, and C. W. J. Beenakker: Theory of directed localization in one dimension, *Phys. Rev. B,* **56** (1997), 4333–4335.

[7] F. Chaitin-Chaitelin and A. Harrabi: About definitions of pseudospectra of closed operators in Banach spaces, *Tech. Rep. TR/PA/98/08,* CERFACS.

[8] S. N. Chandler-Wilde and M. Lindner: Sufficiency of Favard's condition for a class of band-dominated operators on the axis, *J. Funct. Anal.* **254** (2008), 1146–1159.

[9] S. N. Chandler-Wilde and M. Lindner: Limit Operators, Collective Compactness, and the Spectral Theory of Infinite Matrices, *Memoirs of the AMS* (also see *Preprint 2008-7, TU Chemnitz* or *Preprint NI08017-HOP of INI Cambridge*).

[10] M. Chien and H. Nakazato: The numerical range of a tridiagonal operator, *J. Math. Anal. Appl.* **373** (2011), 297–304.

[11] E. B. Davies: Spectral properties of random non-self-adjoint matrices and operators, *Proc. Royal Soc. A.* **457** (2001), 191–206.

[12] E. B. Davies: Spectral theory of pseudo-ergodic operators, *Commun. Math. Phys.* **216** (2001), 687–704.

[13] E. B. Davies *Linear Operators and Their Spectra* , 1st Ed., Cambridge University Press, 2007.

[14] E. B. Davies: *Linear Operators and their Spectra*, Cambridge University Press, 2007.

[15] E. B. Davies: Spectral bounds using higher order numerical ranges, *LMS Journal of Computation and Mathematics*, **8** (2005), 17–45.

[16] J. Feinberg and A. Zee: Non-Hermitean Localization and De-Localization, *Phys. Rev. E* **59** (1999), 6433–6443.

[17] J. Feinberg and A. Zee: Spectral Curves of Non-Hermitean Hamiltonians, *Nucl. Phys. B* **552** (1999), 599–623.

[18] S. Gershgorin: Über die Abgrenzung der Eigenwerte einer Matrix, *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk* **7** (1931), 749-754.

[19] I. Y. Goldsheid, and B. A. Khoruzhenko: Distribution of eigenvalues in non-Hermitian Anderson models, *Phys. Rev. Lett.,* **80** (1998), 2898–2900.

[20] I. Y. Goldsheid, and B. A. Khoruzhenko: Eigenvalues curves of asymmetric tri-diagonal matrices, *Elect. J. Prob.,* **5** (2000), 1–28.

[21] R. Hagen, S. Roch and B. Silbermann: *C*-Algebras and Numerical Analysis*, Marcel Dekker, Inc, 2001.

[22] N. Hatano: Localization in non-Hermitian quantum mechanicsand flux-line pinning in superconductors, *Physica A: Statistical and Theoretical Physics,* **254** (1998), 317–331.

[23] N. Hatano and D. R. Nelson: Localization transitions in non-Hermitian quantum mechanics, *Phys. Rev. Lett.,* **77** (1996), 570–573.

[24] N. Hatano and D. R. Nelson: Vortex pinning and non-Hermitian quantum mechanics, *Phys. Rev.,* **56** (1997), 8651–8673.

[25] N. Hatano, T. Wanatabe, and J. Yamasaki: Localization, resonance and non-Hermitian quantum, *Physica. A,* **314** (2002), 170–176.

[26] F. Hausdorff: Der Wertvorrat einer Bilinearform, *Math Z*, **3** (1919), 314–316.

[27] T. Hawkins: Cauchy and the spectral theory of matrices , *Historia Mathematica 2*, **3** (1975), 1-29.

[28] D. E. Holz, H. Orland and A. Zee: On the Remarkable Spectrum of a Non-Hermitian Random Matrix Model, *Journal of Physics A: Mathematical and General* **36** (2003), 3385-3400.

[29] R. A. Janik, M. A. Nowak, G. Papp and I. Zahaed: Localization transitions from free random variables, *Acta. Phys. Polon. B,* **30** (1999), 45–58.

[30] M. Kline *Mathematical Thought from Ancient to Modern Time.*, Oxford university press., New York, 1972.

[31] R. Kress *An Introduction to the Mathematical Theory of Inverse Problems.*, Springer, Berlin, 1996.

[32] V. G. Kurbatov: *Functional Differential Operators and Equations*, Kluwer Academic Publishers, Dordrecht, Boston, London 1999.

[33] M. Lindner: *Infinite Matrices and their Finite Sections: An Introduction to the Limit Operator Method*, Frontiers in Mathematics, Birkhäuser 2006.

[34] M. Lindner: Fredholmness and index of operators in the Wiener algebra are independent of the underlying space, *Operators and Matrices* **2** (2008), 297–306.

[35] M. Lindner: A note on the spectrum of bi-infinite bi-diagonal random matrices, Technische Universitt Chemnitz, Fakultt fr Mathematik (Germany). Preprint 12, 2008

[36] M. Lindner: Fredholm Theory and Stable Approximation of Band Operators and Their Generalisations, *Habilitation thesis, uni, Germany*, **40** (2008), 493–504.

[37] C. Martínez: Spectral Properties of Tridiagonal Operators, *PhD thesis, Kings College, London* 2005.

[38] C. Martínez: Spectral estimates for the one-dimensional non-self-adjoint Anderson model, *J. Operator Theory*, **56** (2006), 59–88.

[39] D. R. Nelson and N. Hatano: Winding numbers, complex currents, and non-Hermitian localization, *Phys. Rev. Lett.,* **80** (1998), 5172–5175.

[40] D. R. Nelson and N. M. Shnerb: Non-Hermitian localization and population biology, *Phys. Rev. E,* **58** (1998), 1383–1403.

[41] S. Prössdorf and B. Silbermann *Numerical Analysis for Integral and Related Operator Equations* , Birkhäuser, 1991.

[42] W. Rudin: *Functional Analysis,* McGraw-Hill, 2nd edition, 1973.

[43] E. Shargorodsky: On the level sets of the resolvent norm of a linear operator, *Bulletin of the LMS*, **40** (2008), 493–504.

[44] E. Shargorodsky and S. Shkarin: The level sets of the resolvent norm and convexity properties of Banach spaces, *Archiv der Mathematik*, **93** (2009), 59–66.

[45] D. J. Thouless: Electrons in disordered systems and the theory of localization, *Phys. Reports,* **13** (2001), 93–142.

[46] O. Toeplitz: Das algebraische Analogon zu einem Satze von Fejer, *Math Z*, **2** (1918), 187–197.

[47] L. N. Trefethen and M. Embree: *Spectra and pseudospectra The behaviour of Nonnormal Matrices and Operators*, Princeton University Press, 2005.

[48] L. N. Trefethen, M. Contedini and M. Embree: Spectra, pseudospectra, and localization for random bidiagonal matrices, *Comm. Pure Appl. Math.* **54** (2001), 595–623.

[49] L. N. Trefethen and M. Embree: *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton University Press, Princeton, NJ, 2005.

[50] R. S. Varga *Gershgorin and his Circles*, Springer, 2004.