PSEUDO-EIGENVALUES AND SOME PROPERTIES OF TOEPLITZ MATRICES

by

SELDA TOSUN ÇİFTÇİ

THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

OF

THE ABANT İZZET BAYSAL UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF MATHEMATICS

JULY 2008

Approval of the Graduate School of Natural and Applied Sciences.

Prof. Dr. Nihat ÇELEBİ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Zafer Ercan
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality as a thesis for the degree of Master of Science.

Assist. Prof. Dr. Yusuf Cesur
Supervisor

Examining Committee Members

Associate Prof. Dr. Soner Durmuş

Assistant Prof. Dr. Erol Yılmaz

Assistant Prof. Dr. Yusuf Cesur

ABSTRACT

PSEUDO-EIGENVALUES AND SOME PROPERTIES OF TOEPLITZ MATRICES

Tosun Çiftçi, Selda

M.Sc., Department of Mathematics

Supervisor: Assistant Prof. Dr. Yusuf Cesur

July 2008, 45 pages

In this thesis we investigate the eigenvalues and pseudospectra of a nonhermitian Toeplitz matrix A. These are usually highly sensitive to perturbations having condition numbers that increase exponentially with the dimension N. An equivalent statement is that the resolvent $R(z) = (zI - A)^{-1}$ of a Toeplitz matrix may be much larger in norm than the eigenvalue alone would be-exponentially large as a function of N, even when z is near to the spectrum. Because of these facts, the meaningfulness of the eigenvalues of nonhermitian Toeplitz matrices for any purposes should be considered suspect. In many applications it is more meaningful to investigate the ε-pseudoeigenvalues: the complex numbers z with $\left\|(zI - A)^{-1}\right\| \geqslant \varepsilon^{-1}$.

In the second part of the thesis we investigate the pseudospectra of Triangular Toeplitz Matrices and analyzes the pseudospectra of lower and upper triangular and in particular relates them to the symbols of the matrices and theory to the spectra of the Triangular Toeplitz matrices. Our results are reasonably complete in the triangular case, and preliminary in the cases of nontriangular Toeplitz matrices, with smoothly varying coefficients.

This study presents computed examples of pseudospectrum for Triangular Toeplitz Matrices with using Mathematica and Matlab, and applications in numerical analysis.

Keywords: spectrum, pseudospectrum, eigenvalue, pseudoeigenvalue, Toeplitz Matrix

ÖZET

ÖTELENMİŞ ÖZDEĞERLER VE TOEPLİTZ MATRİSLERİN ÖZELLİKLERİ

Tosun Çiftçi, Selda

Master Tezi, Matematik Bölümü

Tez Yöneticisi: Yrd. Doç. Dr. Yusuf Cesur

Temmuz 2008, 45 sayfa

Bu tezde, hermitian olmayan bir A Toeplitz matrisinin özdeğerlerini araştırıyoruz. Genellikle bunlar küçük ötelemelere karşı aşırı duyarlıdırlar ve koşul sayıları N boyutuna bağlı olarak üstel olarak artmaktadır. Buna eşdeğer bir ifade ise bir Toeplitz matrisine ait özdeğerleri $R(z) = (zI - A)^{-1}$ rezolvent fonksiyonu ile incelenebilmesidir. Rezolventin, spektruma ait z noktalarına çok yakın noktalarda N'e bağlı olarak üstel olarak hızla norm değerleri artar. Bu nedenle hermitian olmayan Toeplitz matrislerine ait birçok teorik çalışmada rezolventten yararlanılarak çalışılır. Bu bakış altında birçok uygulamada ε-ötelenmiş özdeğerleri araştırma işlemlerinde rezolventten yararlanılır. Kısaca bir z kompleks sayısında $\|(zI - A)^{-1}\| \geq \varepsilon^{-1}$ yardımıyla ε-ötelenmiş özdeğerleri incelenebilmektedir.

Çalışmanın ikinci kısmında ise üçgen Toeplitz matrislerin pseudospektrumu araştırılmaktadır. Burada ayrıca alt ve üst Toeplitz matrislerin spektrumlarına ait özellikleri semboller yardımıyla da açıklanmaktadır. Yaptığımız çalışmada elde ettiğimiz sonuçlar, üçgen ve üçgen olmayan Toeplitz matrislerinde ve değişken katsayılı durumlarda da elde edilmiştir.

Bu çalışmada üçgen Toeplitz matrislerinin pseudospektrumu ile ilgili örneklerin araştırılmasında Mathematica ve Matlab programlarından ve nümerik analiz tekniklerinden yararlanılmıştır.

Anahtar Kelimeler: spektrum, pseudospektrum, özdeğer, ötelenmiş özdeğer, Toeplitz Matris

ACKNOWLEDGEMENTS

I would like to acknowledge my indebtedness to my supervisor Assistant Prof. Dr. Yusuf Cesur who has guided, encouraged and believed in me throughout this thesis. I am grateful to him for a critical reading of my studies and valuable suggestions in correcting my errors. I also like to thank to the other members of my Committee, Associate Prof. Dr. Soner Durmuş and Assistant Prof. Dr. Erol Yılmaz for their assistance, and also to B.A.P. for their financing help of our thesis.

It is also a pleasure to acknowledge a special dept of my gratitude to all the professors who gave me help in my master studies.

Finally, I would like to thank to my family for their un-vanishing support and encouragement.

To My Family

TABLE OF CONTENTS

ABSTRACT iii

ÖZET v

ACKNOWLEDGEMENTS vii

TABLE OF CONTENTS ix

1 INTRODUCTION 1

1.1 History of pseudospectra 1

1.2 Introduction 4

1.3 Basic Definitions 5

2 EIGENVALUES AND PSEUDOSPECTRA OF MATRICES 10

2.1 Eigenvalues 10

2.2 Pseudospectra of Matrices 11

2.3 A Matrix Example 18

3 TOEPLITZ MATRICES 27

3.1 Triangular Toeplitz Matrices 27

REFERENCES 39

CHAPTER 1

INTRODUCTION

1.1 History of pseudospectra

In this part of study we attempt to survey the history of the subject of pseudospectra.

In the cases of Hermitian and near-Hermitian systems, mathematical physicists have for years spoken of *quasimodes*, which are the same as what we would call *pseudomodes* or *pseudoeigenvectors*.

In the nonnormal context, the earliest definition of pseudospectra we have encountered is given in J.M. Varah's 1967 The Computation of Bounds for the Invariant Subspaces of a General Matrix Operator [2]. Varah Introduced the notion of an ε-pseudoeigenvalue under the name $r-$ *approximate eigenvalue*. Motivated by his analysis of the accuracy of eigenpairs produced by computer implementations of the inverse iteration algorithm, Varah incorporated a parameter, η_1, in his definition to describe floating-point precision.

Definition 1.1 . *$\{\lambda \quad y\}$ is an $r-$ approximate $\{$eigenvalue eigenvector$\}$ of A if there exists a matrix E with $\|E\|_2 = r.\eta_1$ such that $\{\lambda \quad y\}$ is an exact $\{$eigenvalue eigenvector$\}$ of $A+E$.*

From Varah's unpublished 1967 thesis [2]

Landau applied this consept to the theory of Toeplitz matrices and associated integral operators.

Definition 1.2 *λ is in ε-approximate eigenvalue of A, if there exists $\varphi \in P(rQ)$, with $\|\varphi\| = 1$, such that $\|A_r\varphi - \lambda\varphi\| \leqslant \varepsilon$. We call φ an ε-approximate eigenfunction corresponding to λ.*

definition of pseudospectra from Landau, 1975[3].

In Novasibirsk, S.K. Godunov and his colleagues conducted research related to pseudospectra throught the 1980s. Those involved included A.G. Antonov, A.Y. Bulgakov (later Haydar Bulgak), O.P. Kirilyuk, V.I. Kostin, A.N. Malyshev, and S.I. Razzakov. Godunov, together with Ryabenkii and others, had made significant contributions in the 1960s to the study of how nonnormality affects the numerical stability of discretized differential equations. In fact, in their 1962 monograph, Godunov and Ryabenkii introduce the spectrum of the family of operators R_h, indexed by a mesh parameter h [5]. A point $z \in \mathbb{C}$ is in this set if, for every $\varepsilon > 0$, z is an ε-pseudoeigenvalue of R_h for all sufficiently small h. First published sketch of pseudospectra? From Kostin and Razzakov, 1985[7].

The Novasibirsk group defined the ε-spectrum by relative rather than absolute perturbations: $\sigma_\varepsilon = \left\{z \in \mathbb{C} : \left\|(z - A)^{-1}\right\| \geqslant (\varepsilon \|A\|)^{-1}\right\}$. In [4] and [6], computed plots of pseudospectra were presented and called 'spectral portraits of matrices' containing various 'patches and spectrum', have been invented at least five times: computations based on both contour plotting and curve tracing were discussed. A sketch of two pseudospectra, had appeared earlier in a 1985 paper by Kostin and Razzakov [7].

Pseudospectra were investigated in several papers by J.W. Demmel in the mid-1980s, appearing with the labels $S(A, \varepsilon)$ in [8] and $\sigma(\varepsilon, A)$ in [9]. The former paper contains the first published computer plot.

Beginning in the mid-1980s, D. Hinrichsen and A.J. Pritchard wrote a number of papers on the stability radius of a matrix, i.e., the distance to the set of unstable matrices. In a 1992 paper [11] they introduced the term spectral value set and notation $\sigma(A, \rho)$ to denote the real structured ε-pseudospectra of a nonnormal matrix. Another paper by Hinrichsen and Kelb in 1993 extended these ideas to complex perturbations and also to more general structured perturbations [10].

Trefethen's first publications that mentioned pseudospectra were [12] and [16] in 1990 (the first uses the term ε-approximate eigenvalues). These were an outgrowth of a 1987 paper by Trefethen and Trummer, who found eigenvalues that were extraordinarily sensitive to perturbations but failed fully to appreciate their significance [17]. A crucial collaborator in this work was Trefethen's student Satish Reddy, who began to work with pseudospectra in 1988 and made many contributions after that date. Reddy's early on these topics was summarized in his 1991 thesis at MIT [15].

In 1992 Trefethen's paper 'Pseudospectra of Matrices' appeared, which presented the idea of pseudospectra and exhibited thirteen examples [13]. It was after this point that the idea began to be widely known. A later companion paper, 'Pseudospectra of Linear Opertors', presented ten examples involving oprators on infinite-dimensional spaces [14].

The papers by Demmel and Wilkinson mentioned above cite each other, and they both cite the paper of Varah [2]. None of the papers discussed here published before 1990 cite any of the others. These data suggest that pseudospectra have been invented at least five times:

J.M. Varah 1967 σ-approximate eigenvalues, 1979 ε-spectrum

H.J. Landau 1975 ε-approximate eigenvalues

S.K. Godunov et al. 1982 spectral portrait

L.N. Terefethen 1990 ε-pseudospectrum

D. Hinrichsen and A.J. Pritchard 1992 spectral value set

1.2 Introduction

Many of the nonhermitian matrices in applications have eigenvalues that are highly sensitive to perturbations. Indeed, for a family of matrices with variable dimension N, the sensitivity of the eigenvalues often increases exponentially as $N \to \infty$. In such circumstances eigenvalue analysis may lead to misleading conclusions, and the reasons lie deeper than the possibility of rounding errors on a computer. Highly sensitive eigenvalues are a reflection of the fact that the basis of eigenvectors is highly ill conditioned, and when this is the case, it is unlikely that there is good reason for working with that basis. Yet any statement about eigenvalues depends ultimately on the properties of the associated eigenvectors, whether or not they are mentioned explicitly.

We believe that in problems like this, a fruitful alternative is the analysis of pseudoeigenvalues. In particular, the purpose of this study is to investigate the pseudospectra of various kinds of Toeplitz matrices, a special family of matrices with broad applications to integral equations, finite-difference equations, matrix iterations, spline approximation, signal processing, and other problems. Let A be a real or complex square matrix of dimension N, and let $\|.\|$ denote the 2-norm. The definition of pseudoeigenvalues is as follows:

Definition. Given $\varepsilon > 0$, the number $z \in \mathbb{C}$ is an ε-pseudoeigenvalue of A if any of the following equivalent conditions is satisfied:

(i) z is an eigenvalue of $A + E$ for some $E \in \mathbb{C}^{N \times N}$ with $\|E\| \leqslant \varepsilon$;

(ii) $\exists u \in \mathbb{C}^n$, such that $\|(A - zI)u\| \leqslant \varepsilon$;

(iii) $\|(zI - A)^{-1}\| \geqslant \varepsilon^{-1}$;

(iv) $\sigma_N(zI - A) \leqslant \varepsilon$.

The set of all ε-pseudoeigenvalues of A, the ε-pseudospectrum, is denoted by $\sigma_\varepsilon(A)$ or simply σ_ε.

A pseudoeigenvalue, in other words, need to be near to any exact eigenvalue, but it is an exact eigenvalue of some nearby matrix. The vector u in (ii) is a (normalized) ε-pseudoeigenvector. The matrix $(zI-A)^{-1}$ in (iii) is the resolvent of A at the point z, and indeed, any statement about pseudoeigenvalues is equivalent to a statement about norms of the resolvent. In condition (iv), $\sigma_N(zI-A)$ denotes the smallest singular value of $zI - A$.

We shall use the following notation: D is the open unit disk in the complex plane, S is the unit circle, and $\Delta = D \cup S$ is the closed unit disk. The corresponding disks and circles of arbitrary radius r are denoted by D_r, S_r and Δ_r.

1.3 Basic Definitions

Definition 1.3 *(Linear operator). A linear operator T is an operator such that*

(i) the domain $D(T)$ of T is a vector space and the range $R(T)$ lies in a vector space over the same field.

(ii) for all $x, y \in D(T)$ and scalars α,

$$T(x+y) = Tx + Ty$$

$$T(\alpha x) = \alpha Tx$$

Definition 1.4 *(Bounded linear operator). Let X and Y be normed space and $T : D(T) \to Y$ a linear operator, where $D(T) \subset X$. The operator T is said to be bounded if there is a real number c such that for all $x \in D(T)$,*

$$\|Tx\| \leqslant c\,\|x\|$$

Definition 1.5 *(Norm). A norm is a function $\|.\| : \mathbb{C}^m \to \mathbb{R}$ and which has the following properties,* *(N1)* $\|x\| \geqslant 0$

(N2) $\|x\| = 0 \Leftrightarrow x = 0$

(N3) $\|\alpha x\| = |\alpha| . \|x\|$

(N4) $\|x + y\| \leqslant \|x\| + \|y\|$

here x and y are arbitrary vectors and $\alpha \in \mathbb{C}$ scalar.

Definition 1.6 *(Normed space, Banach space). A normed space X is a vector space with a norm defined on $\langle x, \|x\| \rangle$. A Banach space is complete normed space.*

Definition 1.7 *(Two Norm of Matrix). For a matrix A,*

$$\|A\| = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2)^{1/2}$$

Definition 1.8 *(Two Norm of Vector). For a vector x,*

$$\|x\| = (\sum_{i=1}^{n} x_i^2)^{1/2}$$

Definition 1.9 *(Eigenvalues, eigenvectors, eigenspaces, spectrum, resolvent set of matrix).*

Let X be a finite dimensional normed space and A be a matrix operator.

An eigenvalue of a square matrix A=(α_{ij}) is a number λ such that $Av = \lambda v$ has a solition $v \neq 0$.

This v is called an eigenvector of A corresponding to that eigenvalue λ.

The eigenvectors corresponding to that eigenvalue λ and the zero vector form a vector subspace of X which is called the eigenspace of A corresponding to that eigenvalue λ.

The set $\sigma(A)$ of all eigenvalues of A is called the spectrum of A.

Its complement $\rho(A) = \mathbb{C} - \sigma(A)$ in the complex plane is called the resolvent set of A [1].

Definition 1.10 *(Regular value, resolvent set, spectrum).*

Let $X \neq \{0\}$ be a complex normed space and $A : D(A) \to X$ a linear operator with domain $D(A) \subset X$. A regular value λ of A is a complex number and $R_\lambda(A) = (A - \lambda I)^{-1}$ such that,

(R1) $R_\lambda(A)$ exists,

(R2) $R_\lambda(A)$ is bounded,

(R3) $R_\lambda(A)$ is defined on a set which is dense in X. The resolvent set $\rho(A)$ of T is the set of all regular values λ of A.

Its complement $\sigma(A) = \mathbb{C} - \rho(A)$ in the complex plane $\mathbb{C}$ is called the spectrum of A, and a $\lambda \in \sigma(A)$ is called a spectral value of A.

Furthermore, the spectrum $\sigma(A)$ is partitioned into three disjoint sets as follows.

(i) The point spectrum or discreet spectrum $\sigma_p(A)$ is the set such that $R_\lambda(A)$ does not exists. A $\lambda \in \sigma_p(A)$ is called an eigenvalue of A.

(ii) The continuous spectrum $\sigma_c(A)$ is the set such that $R_\lambda(A)$ exists and satisfies (R3) but not (R2), that is, $R_\lambda(A)$ is bounded.

(iii) the residual spectrum $\sigma_r(A)$ is the set such that $R_\lambda(A)$ exists and may be bounded (or not) but does not satisfy (R3), that is the domain of $R_\lambda(A)$ is not dense on X.

For instance $\sigma_c(A) = \sigma_r(A) = \emptyset$ in the finite dimensional case.

Definition 1.11 *(Condition number of λ_j).*

The condition number of λ_j is,

$$K(\lambda_j) = \frac{\|u_j\| \, \|v_j\|}{\left\|u_j^* v_j\right\|}, \qquad K(\lambda_j) \geqslant 1 .$$

Definition 1.12 *(Toeplitz matrix). Given* $2n-1$ *numbers* a_k*, where* $k=-n+1,...,-1,0,1,...,n-1$*, a Toeplitz matrix is a matrix which has constant values along negative-sloping diagonals, i.e., a matrix of the form*

$$\begin{bmatrix} a_0 & a_{-1} & a_{-2} & \cdots & a_{-n+1} \\ a_1 & a_0 & a_{-1} & \ddots & \vdots \\ a_2 & a_1 & a_0 & \ddots & a_{-2} \\ \vdots & \ddots & \ddots & \ddots & a_{-1} \\ a_{n-1} & \cdots & a_2 & a_1 & a_0 \end{bmatrix}.$$

Definition 1.13 *(Symbol of A). The symbol of a Toeplitz matrix is the function*

$$f(z) = \sum_k a_k z^k \ ,$$

this is a finite sum or infinite series depending on the context.

Definition 1.14 *(Analytic Function). Let* $z = x+iy$ *and* $f(z) = u(x,y)+iv(x,y)$ *on some region* G *containing the point* $z_0 \in \mathbb{C}$*. If* $f(z)$ *satisfies the Cauchy-Riemann equations and has continuous first partial derivatives in the neighborhood of* z_0*, then* $f'(z_0)$ *exists and is given by*

$$f'(z_0) = \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} \ ,$$

and the function is said to be analytic or complex differentiable .

Definition 1.15 *(Laurent operators). A Laurent operator* A *is a bounded linear operator on* $\ell_2(Z)$ *with the property that the matrix of* A *with respect to the* *standard orthonormal basis* $\{e_j\}_{j=-\infty}^{\infty}$ *of* $\ell_2(Z)$ *is of the form*

$$\begin{bmatrix} \ddots & \ddots & & & \\ \ddots & a_0 & a_{-1} & a_{-2} & \\ & a_1 & a_0 & a_{-1} & \\ & a_2 & a_1 & a_0 & \ddots \\ & & & \ddots & \ddots \end{bmatrix}$$

Here a_0 *denotes the entry* a_0 *located in the zero row zero column position. In other words, a bounded linear operator* A *on* $\ell_2(Z)$ *is a Lourent operator if and only if* $\langle Ae_k, e_j \rangle$ *depends on the difference* $j - k$ *only.*

Definition 1.16 *(Fourier transform).*

The Fourier transform is a generalization of the complex Fourier series in the limit as $L \to \infty$*. Replace the discrete* A_n *with the continuous* $F(k)dk$ *while letting* $n/L \to k$*. Then change the sum to an integral, and the equations become*

$$f(x) = \int_{-\infty}^{\infty} F(k) e^{2\pi i k x} dk$$

$$F(k) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx.$$

CHAPTER 2

EIGENVALUES AND PSEUDOSPECTRA OF MATRICES

2.1 Eigenvalues

In the simplest context of matrices, the definitions are as follows. Let A *be an* $N \times N$ *matrix with real or complex coefficients; we write* $A \in C^{N \times N}$*. Let* v *be a nonzero real or complex column vector of length* N*, and let* λ *be a real or complex scalar; we write* $v \in C^N$ *and* $\lambda \in C$*. Then* v *is an eigenvector of* A*, and* $\lambda \in C$ *is its corresponding eigenvalue, if*

$$Av = \lambda v \tag{2.1}$$

(Even if A *is real, its eigenvalues are in general complex unless* A *is self-adjoint.) The set of all the eigenvalues of* A *is the spectrum of* A*, a nonempty subset of the complex plane* C *that we denote by* $\sigma(A)$*.*

The spectrum can also be defined as the set of points $z \in C$ *where the resolvent matrix,*

$$(z - A)^{-1} \tag{2.2}$$

does not exist. Where $(z - A)$ *is shorthand for* $zI - A$*, where* I *is the identity matrix.*

Eigenvalues are useful for so many problems, more than one of which may be important in a particular application.

1. Diagonalization and separation of variables: use of the eigenfunctions as a basis.

2. Resonance: heightened response to selected inputs.

3. Asymptotics and stability: dominant response to general inputs.

4. They give a matrix a personality.

2 norm, defined by $\|x\|_2 = \left(\sum |x_j|^2\right)^{1/2}$ *for a vector* x *and then by*

$$\|A\|_2 = \max_x \frac{\|Ax\|_2}{\|x\|_2} \tag{2.3}$$

for a matrix A*.*

2.2 Pseudospectra of Matrices

In this section we introduce pseudospectra of finite-dimensional matrices, and we generalize to linear operators in Banach space. In most of the computed examples we take $\|.\|$ *to be the 2-norm* $\|.\|_2$ *defined by the Euclidean inner product as in (2.3). We shall generally let* A *denote a matrix in* $C^{N\times N}$ *or a linear operator on an infinite-dimensional space. Throughout applied mathematics, the question 'Is* A *singular?' is not robust, for an arbitrarily small perturbation can change the answer from yes to no. For applied purposes, a better question is 'Is* $\|A^{-1}\|$ *large?' Now, the condition defining eigenvalues is a condition of matrix singularity. To ask, 'Is* z *eigenvalue of* A *?' is the same as to ask, is*

$$z - A$$

singular ?

Therefor, the property of being an eigenvalue of a matrix is also not robust.

A better question may be, is

$$\left\|(z-A)^{-1}\right\|$$

large ?

This pattern of thinking leads gives to our first definition of pseudospectra:

Definition 2.1 *Let $A \in C^{N\times N}$ end $\varepsilon > 0$ be arbitrary. The $\varepsilon-$pseudospectrum $\sigma_\varepsilon(A)$ of A is the set of $z \in C$ such that*

$$\left\|(z-A)^{-1}\right\| > \varepsilon^{-1} \tag{2.4}$$

The matrix $(z-A)^{-1}$ is known as the resolvent of A at z. In (2.4) we employ the convention that

$$\left\|(z-A)^{-1}\right\| = \infty \quad \text{for} \quad z \in \sigma(A) \tag{2.5}$$

where $\sigma(A)$ is the spectrum (set of eigenvalues) of A, so that in particular, the spectrum is contained in the ε-pseudospectrum for every $\varepsilon > 0$. In words, the ε-pseudospectrum is the open subset of the complex plane bounded by the ε^{-1} level curve of the norm of the resolvent.

Another definition of pseudospectra is based on the connection between the resolvent norm and eigenvalue perturbation theory [18]

Definition 2.2 *$\sigma_\varepsilon(A)$ is the set of $z \in C$ such that*

$$z \in \sigma(A+E) \tag{2.6}$$

for some $E \in C^{N\times N}$ with $\|E\| < \varepsilon$.

From either of these definitions, it follows that the pseudospectra associated *with various ε are nested sets,*

$$\sigma_{\varepsilon_1}(A) \subseteq \sigma_{\varepsilon_2}(A), \qquad 0 < \varepsilon_1 \leqslant \varepsilon_2 \tag{2.7}$$

and that the intersection of all the pseudospectra is the spectrum,

$$\bigcap_{\varepsilon>0} \sigma_\varepsilon(A) = \sigma(A). \tag{2.8}$$

Definition 2.3 *$\sigma_\varepsilon(A)$ is the set of $z \in C$ such that*

$$\|(z - A)v\| < \varepsilon \tag{2.9}$$

for some $v \in C^N$ with $\|v\| = 1$.

The number z in (2.9) (or equivalently in any of our definitions) is an $\varepsilon-$ pseudoeigenvalue of A, and v is a corresponding $\varepsilon-$pseudoeigenvector. (Synonymous terms include pseudoeigenfunction, pseudoeigenmode, and pseudomode.) In words, the $\varepsilon-$ pseudospectrum is the set of $\varepsilon-$ pseudoeigenvalues.

Theorem 2.4 *(Equivalence of the definitions of pseudospectra).*

For any matrix $A \in C^{N\times N}$, the three definitions above are equivalent.

Proof.

For $z \in \sigma_\varepsilon(A)$ the equivalence is trivial, so assume $z \notin \sigma(A)$, implying the existence of $(z - A)^{-1}$. To prove (2.6) $\Rightarrow$ (2.9), suppose that $(A + E)v = zv$ for some $E \in C^{N\times N}$ with $\|E\| < \varepsilon$ and some nonzero $v \in C^N$, which we may take to be normalized, $\|v\| = 1$. Then $\|(z - A)v\| = \|Ev\| < \varepsilon$, as required. To prove (2.9) $\Rightarrow$ (2.4), suppose $(z - A)v = su$ for some $v, u \in C^N$ with $\|v\| = \|u\| = 1$ and $s < \varepsilon$. Then $(z - A)^{-1}u = s^{-1}v$, so $\|(z - A)^{-1}\| \geq s^{-1} > \varepsilon^{-1}$. Finally, to prove (2.4) $\Rightarrow$ (2.6), suppose $\|(z - A)^{-1}\| > \varepsilon^{-1}$. Then $(z - A)^{-1}u = s^{-1}v$ and *consequently $zv - Av = su$ for some $v, u \in C^N$ with $\|v\| = \|u\| = 1$ and $s < \varepsilon$. To establish (2.6), it is enough to show that there exist a matrix $E \in C^{N\times N}$ with $\|E\| = s$ and $Ev = su$, for then v will be an eigenvector of $A+E$ with eigenvalue z. In fact, E can be taken to be a rank-1 matrix of the form $E = suw^*$ for some $w \in C^N$ with $w^*v = 1$. If $\|.\|$ is the 2-norm, this is evident simply by taking $w = v$. In the case of an arbitrary norm $\|.\|$, the existence of a vector v satisfying the required conditions can be interpreted as the existence of a linear functional L on C^N with $\|Lv\| = 1$ and $\|L\| = 1$, which is guaranteed by the Hahn-Banach theorem. A less highbrow version of the same proof can be carried out by the method of dual norms; see [19] and [20].* □

In this section we shall often restrict attention to the case in which C^N is the standard inner product

$$(u, v) = v^*u \tag{2.10}$$

and $\|.\|$ is the corresponding 2-norm,

$$\|v\| = \|v\|_2 = \sqrt{v^*v}\ . \tag{2.11}$$

If $\|.\| = \|.\|_2$, the norm of a matrix is its largest singular value and the norm of the inverse is the inverse of the smallest singular value. In particular,

$$\left\|(z - A)^{-1}\right\|_2 = [s_{\min}(z - A)]^{-1}\ , \tag{2.12}$$

where $s_{\min}(z - A)$ denotes the smallest singular value of $z - A$.

Definition 2.5 *For $\|.\| = \|.\|_2$, $\sigma_\varepsilon(A)$ is the set of $z \in C$ such that*

$$s_{\min}(z - A) < \varepsilon \tag{2.13}$$

Definition 2.6 *(Normal matrix).*

A matrix $A \in C^{N\times N}$ is normal if it has a complete set of orthogonal eigenvectors, that is, if it is unitarily diagonalizable:

$$A = U\Lambda U^* \tag{2.14}$$

(Here U is unitary and Λ is a diagonal matrix of eigenvalues.) For a normal matrix, the ε-pseudospectrum is just the union of the open ε-balls about the points of the spectrum. The resolvent norm satisfies

$$\left\|(z-A)^{-1}\right\|_2 = \frac{1}{dist(z,\sigma(A))} \tag{2.15}$$

where $dist(z,\sigma(A))$ denotes the usual distance of a point to a set in the complex plane. The next theorem expresses these facts with the aid of the following notation for an open ε-ball:

$$\Delta_\varepsilon = \{z \in \mathbb{C} : |z| < \varepsilon\} \tag{2.16}$$

In this theorem a sum of sets has the usual meaning:

$$\sigma(A) + \Delta_\varepsilon = \{z : z = z_1 + z_2\ ,\ z_1 \in \sigma(A)\ ,\ z_2 \in \Delta_\varepsilon\}$$

which is equal to $z : dist(z,\sigma(A)) < \varepsilon$.

Theorem 2.7 *(Pseudospectra of a normal matrix).*

For any $A \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$,

$$\sigma_\varepsilon(A) \supseteq \sigma(A) + \Delta_\varepsilon \qquad \forall\, \varepsilon > 0, \tag{2.17}$$

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

$$\sigma_\varepsilon(A) = \sigma(A) + \Delta_\varepsilon \qquad \forall\, \varepsilon > 0. \tag{2.18}$$

Conversely, if $\|.\| = \|.\|_2$, then (2.18) implies that A is normal.

Proof.

If z is an eigenvalue of A, then $z+\delta$ is an eigenvalue of $A+\delta$ for any $\delta \in \mathbb{C}$; since $\|\delta \mathrm{I}\| = |\delta|$, this establishes (2.17). For (2.18) we note that if A is normal, it can be assumed without loss of generality to be diagonal without any effect on norms if $\|.\| = \|.\|_2$, with diagonal elements a_{jj} equal to the eigenvalues λ_j. In this case the resolvent is also diagonal, which implies that it satisfies (2.15), and as noted above, (2.4) implies that this is equivalent to (2.18). Finally, for the converse, here is a sketch of a proof that can be made precise with the aid of results from [21]. Equation (2.18) implies that each eigenvalue of A has condition number 1. By a standard formula for eigenvalue condition numbers [21], if $\|.\| = \|.\|_2$, it follows that each right eigenvector of A is also a left eigenvector, i.e., that A and A^* are simultaneously diagonalizable. Therefore A is normal.

Now suppose A is diagonalizable but not necessarily normal, and let $V \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$ be a matrix of eigenvectors of A. With $\|.\| = \|.\|_2$, the condition number of this basis of eigenvectors, is

$$\kappa(V) \equiv \|V\|_2 \left\|V^{-1}\right\|_2 = \frac{s_{\max}(V)}{s_{\min}(V)} \tag{2.19}$$

where $s_{\max}(V)$ and $s_{\min}(V)$ are the largest and smallest singular values of V. In general, $\kappa(V)$ may be any number in the range $1 < \kappa(V) < \infty$, and the value $\kappa(V) = 1$ is possible if and only if A is normal. $\square$

The condition number of V provides an upper bound for the condition numbers *of the individual eigenvalues of A. This fact is known as the Bauer-Fike theorem.*

Theorem 2.8 *(Bauer-Fike theorem).*

Suppose $A \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$ is diagonalizable, $A = V\Lambda V^{-1}$. Then for each $\varepsilon > 0$, with $\|.\| = \|.\|_2$,

$$\sigma(A) + \Delta_\varepsilon \subseteq \sigma_\varepsilon(A) \subseteq \sigma(A) + \Delta_{\varepsilon\kappa(V)} \tag{2.20}$$

.

Proof.

(cf. [22, 23, 24]). The first inclusion was established in (2.17). For the second we calculate

$$(z-A)^{-1} = (z-V\Lambda V^{-1})^{-1} = [V(z-\Lambda)V^{-1}]^{-1} = (V^{-1})^{-1}(z-\Lambda\ )^{-1}V^{-1} = V(z-\Lambda)^{-1}V^{-1}\ ,$$

which implies

$$\left\|(z-A)^{-1}\right\|_2 \leqslant \kappa(V)\left\|(z-\Lambda)^{-1}\right\|_2 = \frac{\kappa(V)}{dist(z,\sigma(A))}\ ,$$

and the definition (2.4) completes the proof. □

The following theorem collects some basic properties of pseudospectra,

Theorem 2.9 *(Properties of pseudospectra).*

Let $A \in \mathbb{C}^{\mathbb{N}\times\mathbb{N}}$ and $\varepsilon > 0$ be arbitrary.

(i) $\sigma_\varepsilon(A)$ is nonempty, open, and bounded, with at most N connected components, each containing one or more eigenvalues of A.

(ii) If $\|.\| = \|.\|_2$, then $\sigma_\varepsilon(A^*) = \overline{\sigma_\varepsilon(A)}$.

(iii) If $\|.\| = \|.\|_2$, then $\sigma_\varepsilon(A_1 \oplus A_2) = \sigma_\varepsilon(A_1) \cup \sigma_\varepsilon\,(A_2)$.

(iv) For any $c \in \mathbb{C}$, $\sigma_\varepsilon(A + c) = c + \sigma_\varepsilon(A)$.

(v) For any nonzero $c \in \mathbb{C}$, $\sigma_{|c|\varepsilon}(cA) = c\sigma_\varepsilon\,(A)$.

In part (iii), $A_1 \oplus A_2$ denotes the direct sum of two square matrices A_1 and A_2, whose dimensions need not be equal; in other words it is the block diagonal matrix

$$A_1 \oplus A_2 = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

.

2.3 A Matrix Example

Consider the tridiagonal Toeplitz matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 5 & 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 5 & 0 & 1 \\ 0 & \cdots & 0 & 5 & 0 \end{pmatrix} \in \mathbb{C}^{N\times N} \tag{2.21}$$

This matrix is nonsymetric, but it can be symetrized by the diagonal similarity transformation

$$DAD^{-1} = S \tag{2.22}$$

with

$$D = diag(\frac{1}{\sqrt{5}}, \left(\frac{1}{\sqrt{5}}\right)^2, ..., \left(\frac{1}{\sqrt{5}}\right)^N)$$

and

$$S = \begin{pmatrix} 0 & \sqrt{5} & & & \\ \sqrt{5} & 0 & \sqrt{5} & & \\ & \ddots & \ddots & \ddots & \\ & & \sqrt{5} & 0 & \sqrt{5} \\ & & & \sqrt{5} & 0 \end{pmatrix} \in \mathbb{C}^{N\times N} \tag{2.23}$$

it follows that the eigenvalues of A are the same as those of S, namely

$$\lambda_k(A) = \lambda_k(S) = 2\sqrt{5}\cos\frac{k\pi}{N+1}\,, \qquad 1 \leqslant k \leqslant N \tag{2.24}$$

Thus the spectrum of A consist of N distinct real numbers in the interval (-5,5).

The pseudospectra of A, however, lie far from the real axis.For $N = 25$, Figure 2.1 plots the boundaries of $\sigma_\varepsilon(A)$ for $\varepsilon = 10^{-2}, 10^{-3}, ..., 10^{-7}$ with $\|.\| = \|.\|_2$, revealing wide oval-shaped regions in the complex plane. In fact, the ε-pseudospectrum of A is approximately equal to the region bounded by the ellipse that is the image of the circle $|z| = \varepsilon^{1/N}$ under the mapping

$$f(z) = z^{-1} + 5z \tag{2.25}$$

which is known as the symbol of A. For each z inside the ellipse $f(\mathrm{T})$, where $\mathrm{T} = \{z \in \mathbb{C} : |z| = 1\}$, the resolvent norm $\|(z-A)^{-1}\|$ grows exponentially as $N \to \infty$. On the other hand, for each z outside $f(\mathrm{T})$, $\|(z-A)^{-1}\|$ is bounded uniformly with respect to N.

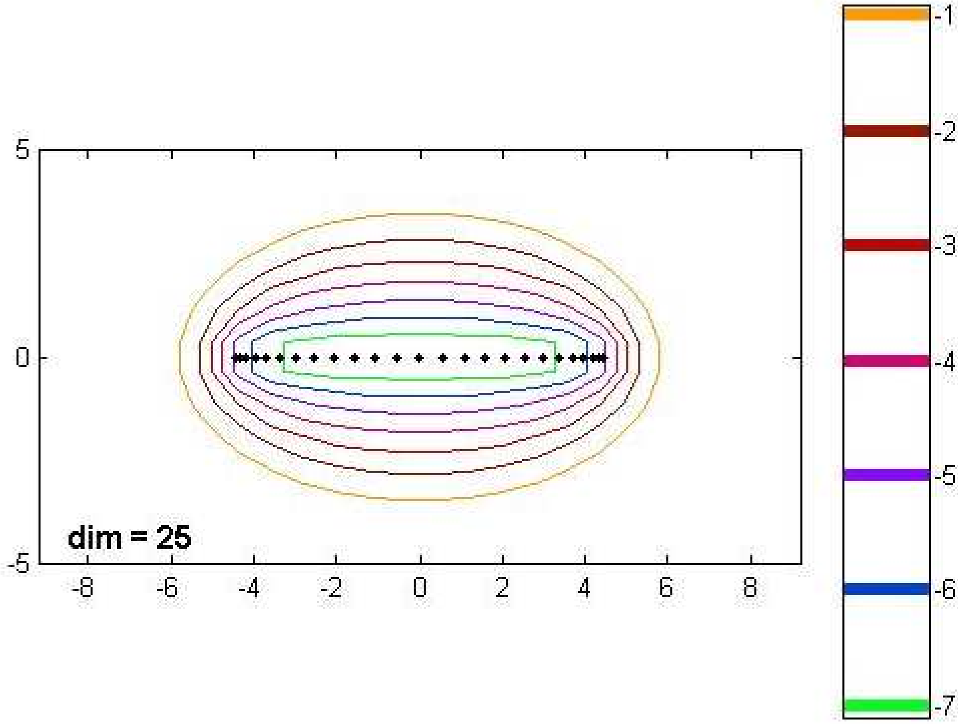

Figure 2.1: Boundaries of pseudospectra $\sigma_\varepsilon(A)$, $\varepsilon = 10^{-2}, 10^{-3}, ..., 10^{-7}$, for the matrix (2.21) of dimension $N = 25$. The eigenvalues are marked by solid dots.

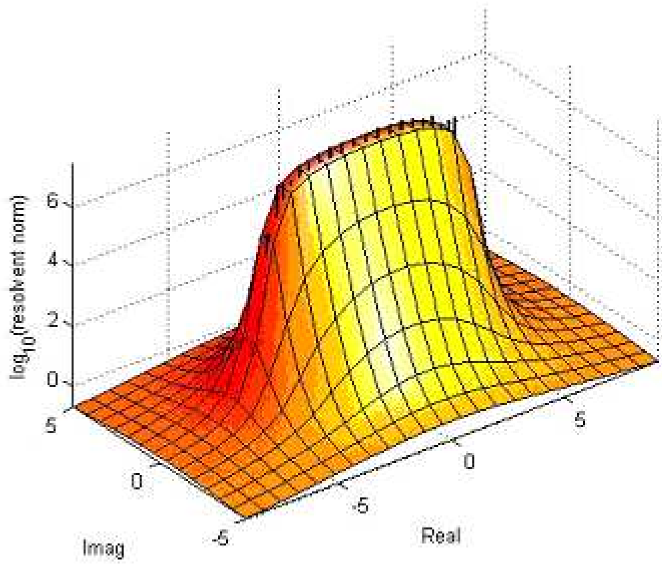

Figure 2.2: The resolvent norm as a function of $z \in \mathbb{C}$ for the matrix of figure 2.1. The spikes occur at eigenvalues of A; in principle they extend to infinity.

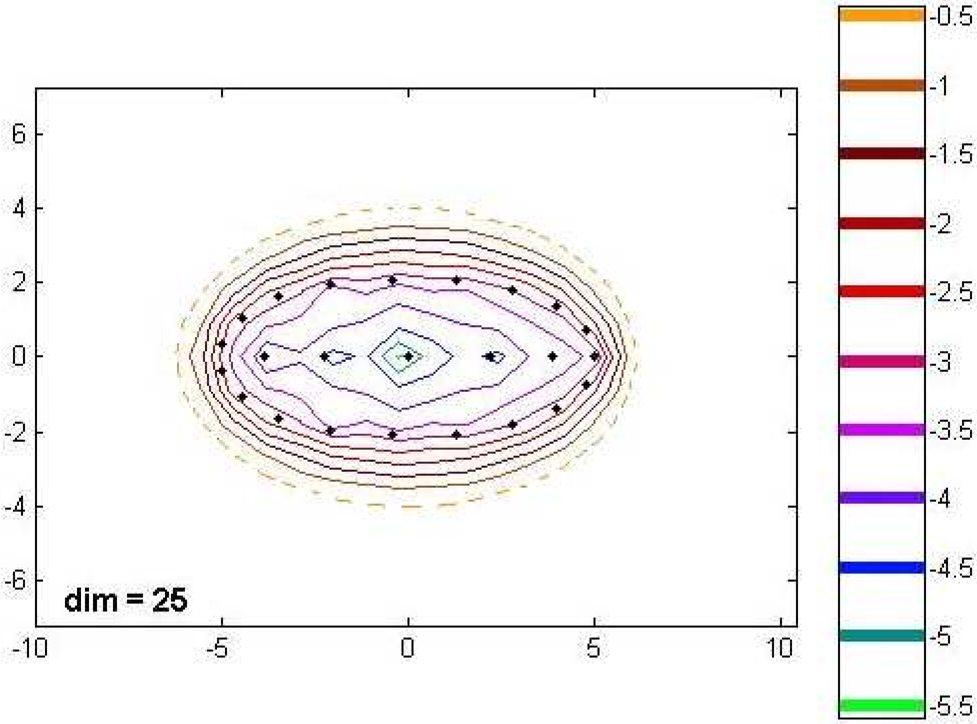

Figure 2.3: Superposition of eigenvalues of $A + E$, where A is the tridiagonal Toeplitz matrix (3.1) of dimension $N = 25$ and E is random matrix with $\|E\| = 10^{-2}$. The eigenvalues of A are real, but the perturbation introduced by E moves them far into the complex plane,close the ellipse defined by (2.25) with $|z| = (0.01)^{\frac{1}{25}}$ (solid curve).The dashed ellipse corresponds to $N \to \infty$ and $|z| = 1$.

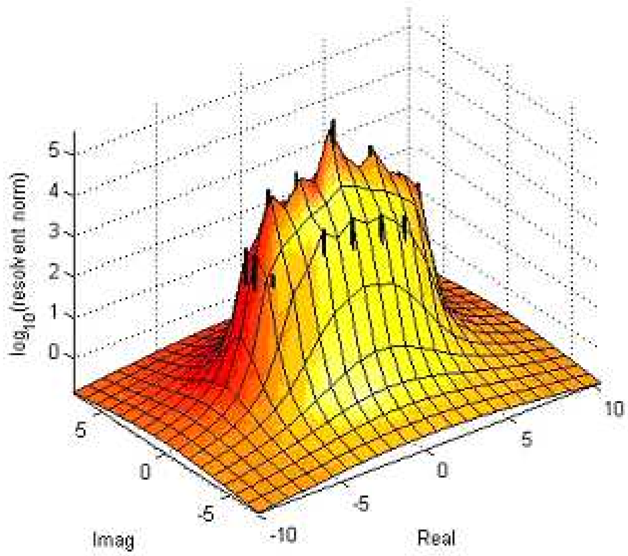

Figure 2.4: The resolvent norm as a function of $z \in \mathbb{C}$ for the matrix of figure 2.3. The spikes occur at eigenvalues of A; in principle they extend to infinity.

These pictures change quantitatively but not qualitatively if N is varied.To illustrate this figure (2.5) and (2.6) shows twelve sets of dots corresponding to same experiment as in figure (2.3), but for $N = 25,\ 75,\ 300$ and 500, and $\|E\| = 10^{-2}, 10^{-3}, 10^{-4}$.

To appreciate how such small perturbations to A can move eigenvalues so dramatically, consider a modification to the $(N,1)$ entry of A:

$$A + E = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 5 & 0 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 5 & 0 & 1 \\ \varepsilon & \cdots & 0 & 5 & 0 \end{pmatrix}, \qquad \varepsilon = 10^{-2}, 10^{-3}, 10^{-4}$$

where

$$E = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ & \ddots & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & \vdots \\ \varepsilon & 0 & \cdots & \cdots & 0 \end{pmatrix}$$

Now apply the similarity transformation that symmetrized A to obtain

$$D\,(A + E)\,D^{-1} = \begin{pmatrix} 0 & \sqrt{5} & & \\ \sqrt{5} & 0 & \ddots & \\ & \ddots & \ddots & \sqrt{5} \\ \left({}^{1}\!/\!\sqrt{5}\right)^{N-1} \varepsilon & & \sqrt{5} & 0 \end{pmatrix}$$

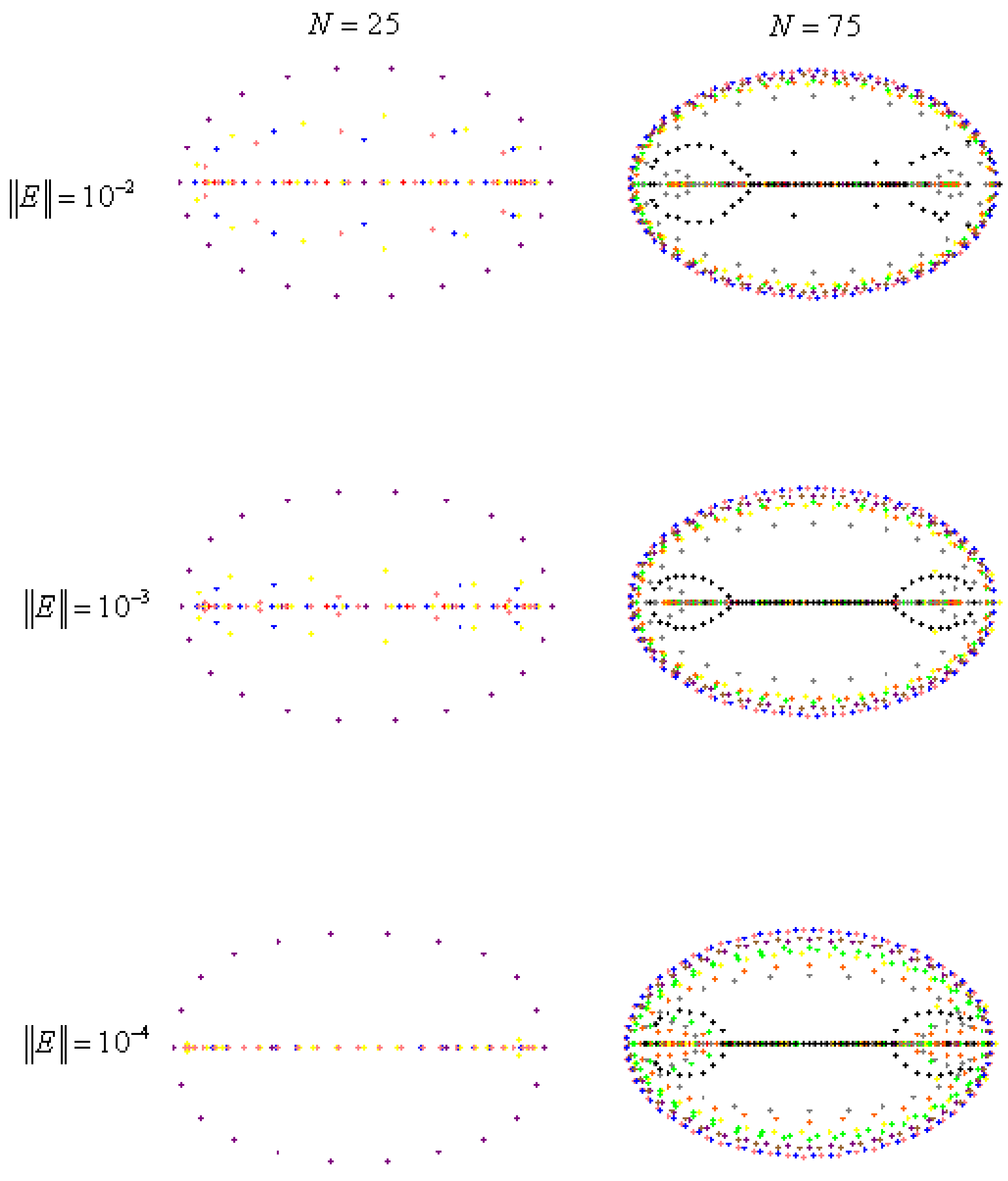

Figure 2.5: Six plots as in figure 2.3 corresponding to N=25, 75 and ten random complex perturbations E, such that $\|E\| = 10^{-2}, 10^{-3}, 10^{-4}$.

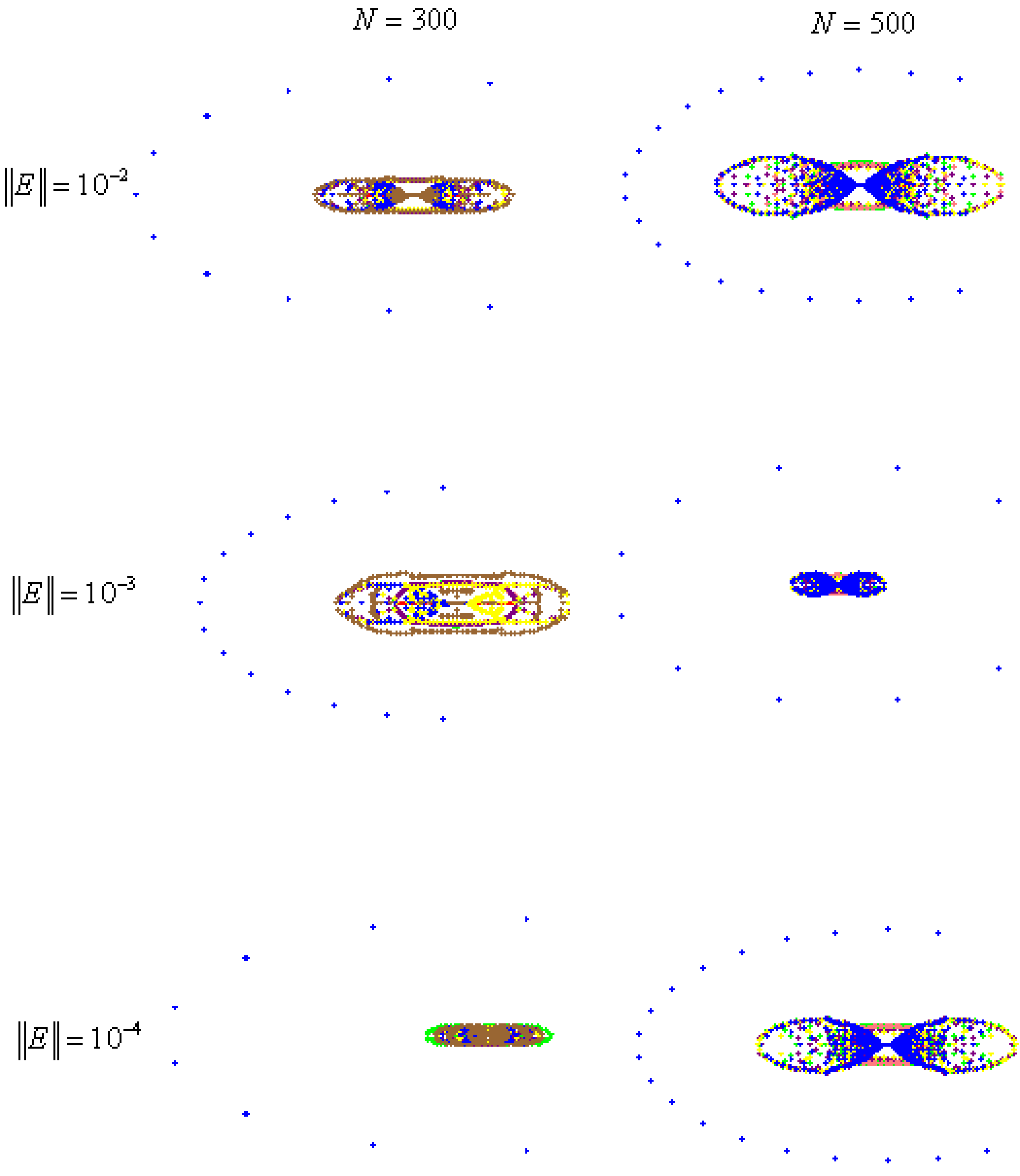

Figure 2.6: Six plots as in figure 2.3 corresponding to N=300, 500 and ten random complex perturbations E, such that $\|E\| = 10^{-2}, 10^{-3}, 10^{-4}$

```
n = 25;
A = DiagonalMatrix[Table[0, {n}]];
For[i = 1, i ≤ n - 1, i++, A[[i+1, i]] = 5 ;];
For[i = 1, i ≤ n - 1, i++, A[[i, i+1]] = 1 ;];
Print[MatrixForm[A]];
In = IdentityMatrix[n];
M = A - λ In;
p[λ_] = Det[A - λ In];
solset = Solve[p[λ] == 0, λ];
For[i = 1, i ≤ n, i++,
  λi = solset[[i, 1, 2]];] ;
Print[      A = , MatrixForm[A]];
Print[A - λ, In,  = , MatrixForm[A],  - λ, MatrixForm[In]];
Print[A - λ, In,  = , MatrixForm[M]];
Print[The characteristic polynomial is];
Print[p[λ] = |A-λ In|];
Print[p[λ] = , p[λ]];
q[λ_] = Factor[p[λ]];
If[Not[p[λ] === q[λ]], Print[p[λ] = , q[λ]]];
Print[To find the eigenvalues of the matrix A];
Print[Solve  , p[λ] == 0];
Print[Get];
For[i = 1, i ≤ n, i++,
  Print[  , λi,  = , λi,  = , Chop[N[λi]]] ];
Print[{(xk,yk)} = , nokta = Table[{ComplexExpand[Re[Chop[N[λi]]]],
    ComplexExpand[Im[Chop[N[λi]]] ]}, {i, n}]]
Needs[Graphics`Colors`];
dots = ListPlot[nokta, PlotStyle → { Red , PointSize[0.015]},
   DisplayFunction → Identity];
Show[dots, PlotRange → All, Axes → False, AxesLabel → {X, Y},
  DisplayFunction → $DisplayFunction];
```

Figure 2.7: Program for plotting $\sigma(A)$ and $\sigma(A + E)$ with Mathematica

A and S have the same eigenvalues, so the spectrum of $A+E$ matches that of $D(A+E)D^{-1}$. Thus the eigenvalues of $A+E$ correspond to those of an exponentially large perturbation of a symmetric matrix. In this light, it might seem remarkable that the eigenvalues of A move so little!

CHAPTER 3

TOEPLITZ MATRICES

3.1 Triangular Toeplitz Matrices

Let A be an upper triangular Laurent operator, Toeplitz operator, or Toeplitz matrix defined by coefficients $a_k \in \mathbb{C}$, $0 \leqslant k \leqslant N-1 \leqslant \infty$,

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{N-1} \\ & a_0 & a_1 & \ddots & \vdots \\ & & a_0 & \ddots & a_2 \\ & & & \ddots & a_1 \\ & & & & a_0 \end{pmatrix} \qquad (N \times N), \tag{3.1}$$

and let $f(z)$ be the symbol of this operator or matrix,

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \tag{3.2}$$

Since we are not concerned with sharp regularity conditions, we assume merely that $f(z)$ belongs to the Wiener class of functions with absolutely convergent Taylor coefficients,

$$\sum_{k=0}^{\infty} |a_k| < \infty \tag{3.3}$$

so that in particular, $f(z)$ is analytic in D and continuous in Δ where D is the open unit disk in the complex plane, S is the unit circle, and $\Delta = D \cup S$ is the closed unit disk. In the case of a Toeplitz matrix this condition is vacuous, and we *shall sometimes write the symbol as $f_N(z)$ to emphasize the finite dimensionality.*

The spectrum $\sigma(A)$ is well understood, and the results are summarized in the following theorem. For a highly readable presentation of the mathematics related to this theorem.

Theorem 3.1 *Let A and f be as described above.*

(i) If A is a Laurent operator, $\sigma(A) = f(S)$.

(ii) If A is a Toeplitz operator, $\sigma(A) = f(\Delta)$.

(iii) If A is a Toeplitz matrix, $\sigma(A) = f(\{0\}) = \{a_0\}$.

Proof. *(i): The result for Laurent operators was first proved by Toeplitz in 1911 [25]. If A is a Laurent operator, then Au is a convolution a^*u, where a is the sequence of values $\{a_{-k}\}$. Since $u \in l^2$ and $a \in l^1 \subseteq l^2$ by assumption, the convolution is equivalent to an operation of pointwise multiplication in the Fourier domain $L^2[-\pi, \pi]$. To be precise, Au has semidiscrete Fourier transform $Au(\theta) = \hat{a}(\theta)\hat{u}(\theta)$, and the vector $(\lambda I - A)^{-1} u$ has Fourier transform*

$$\frac{\hat{u}(\theta)}{\lambda - \hat{a}(\theta)} = \frac{\hat{u}(\theta)}{\lambda - f(e^{i\theta})}$$

Since $\hat{a}(\theta) = f\left(e^{i\theta}\right)$ is a continuous function on a compact domain, the resolvent operator is evidently well defined and bounded in norm if and only if the denominator is never zero, i.e., $\lambda \notin f(S)$ as claimed.

(ii): The result for Toeplitz operators was first proved by Wintner in 1929 [26]. For any $z \in D$, the vector $(1, z, z^2, \ldots)^T$ belongs to l^2 and is obviously an eigenvector of A with eigenvalue $f(z)$. This proves that σ includes $f(D)$, and since σ must be compact, it includes $f(\Delta)$. Conversely, let $\lambda \in \mathbb{C} \quad f(\Delta)$ be arbitrary. Then $[\lambda - f(z)]^{-1}$ is bounded analytic function of the Wiener class in the unit disk, whose Taylor coefficient are easily seen to provide the matrix entries

of an inverse (also Toeplitz) of the operator $\lambda I - A$. In other words, the resolvent $(\lambda I - A)^{-1}$ exists as a bounded operator, so λ is not in the spectrum.

(iii): Since A is triangular, the result for Toeplitz matrices is trivial.

Note that the fact that A is triangular was not used in the proof of (i) above, and thus the same statement carries over to nontriangular Toeplitz matrices and operators. An alternative proof of (ii) can be obtained as a corollary of Theorem 3.2 below.

So much for the standard results. Now, what about pseudospectra? A comparison of statements (ii) and (iii) of Theorem 3.1 shows that the limit $N \to \infty$ is a discontinuous one as far as exact spectra are concerned. The pseudospectra, however, behave continuously. The fundemental observation is that if ε is small and N is large, then σ_ε looks not like $f(0)$ but instead more like

$$\sigma_\varepsilon \approx f\left(\Delta_r\right) \tag{3.4}$$

with

$$r = \left(\frac{\varepsilon}{c_N}\right)^{1/N} = 1 + O\left(N^{-1}\right), \quad c_N = \sum_{k=1}^{N-1} |a_k| \tag{3.5}$$

(The constant c_N is not exactly right, as we shall see below, but its precise value matters little because of the Nth root.) Thus as $N \to \infty$ we have $\sigma_\varepsilon \approx f\left(\Delta_r\right)$, which is the same as the spectrum of the associated Toeplitz operator.

Before stating theorems to this effect, let us consider some examples. Here are three further matrices of interset:

$$A = \begin{pmatrix} 0 & 1 & 1 & & \\ & 0 & 1 & 1 & \\ & & 0 & 1 & 1 \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 2/3 & 2/3^2 & 2/3^3 & 2/3^4 \\ & 1 & 2/3 & 2/3^2 & 2/3^3 \\ & & 1 & 2/3 & 2/3^2 \\ & & & 1 & 2/3 \\ & & & & 1 \end{pmatrix} \tag{3.6}$$

$$C = \begin{pmatrix} 1 & 3 & 3 & 3 & 3 \\ & 1 & 3 & 3 & 3 \\ & & 1 & 3 & 3 \\ & & & 1 & 3 \\ & & & & 1 \end{pmatrix}$$

with spectra

$$\sigma(A) = \{0\}, \qquad \sigma(B) = \sigma(C) = \{1\}$$

Despite appearances, the dimension in each case is N. [Note that A is banded, the entries of B decay geometrically, and the entries of C do not decay at all, which means that it violates our assumption (3.3).] The symbols of these matrices in the limit $N \to \infty$ are

$$f_A(z) = z + z^2,$$

$$f_B(z) = 1 + \frac{2}{3}\left(z + \frac{1}{3}z^2 + \frac{1}{3}z^3 + \cdots\right) = \frac{1 + z/3}{1 - z/3},$$

$$f_C(z) = 1 + 3(z + z^2 + z^3 + \cdots) = \frac{1 + 2z}{1 - z},$$

and the corresponding regions $f(\Delta)$ *are*

$f_A(\Delta)$*: a region bounded by a limaçon,*

$f_B(\Delta)$*: the closed disk ,*

$f_C(\Delta)$*: the closed right half plane.*

Figures 3.1-3.6 show computed ε*-pseudoeigenvalues for* A*,* B*,* C *for* $N = 50,\ 200,\ 300$ *and* 500*. In each plot, the eigenvalues of four random complex perturbations* $A + E$ *with* $\|E\| = 10^{-2}$ *are superimposed, and on top of these are drawn the curves* $f(S_r)$ *with* $r = (\varepsilon/c_N)^{1/N}$ *(solid). In each case most of the computed eigenvalues lie close to the solid curve, in keeping with (3.4).*

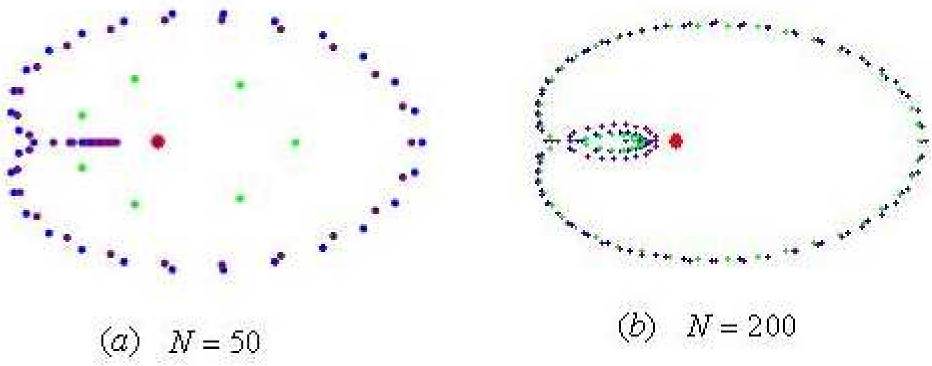

Figure 3.1: Eigenvalues of A+E (3.6), $\|E\| = 10^{-2}$; the results from four random complex perturbations E are superimposed. (a) and (b) shows eigenvalues of four randomly perturbed matrices A+E and this eigenvalues are marked by colour points plotted by Mathematica. The eigenvalues of A are all 0 (red point).

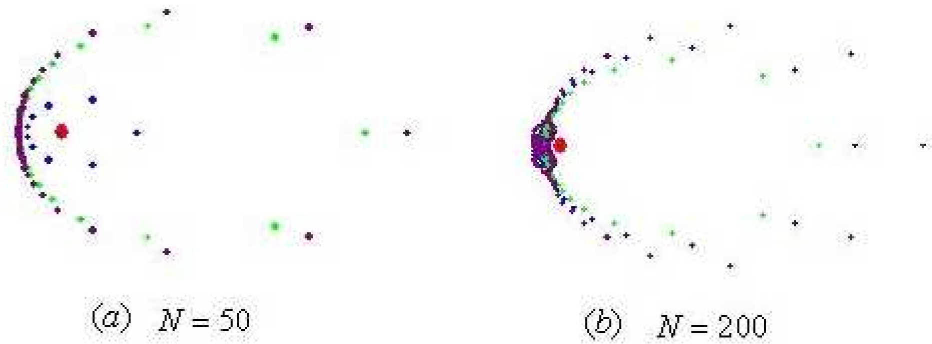

Figure 3.2: Same as Figure 3.1, but for the matrix B of (3.6). The eigenvalues of B are all 1 (red point).

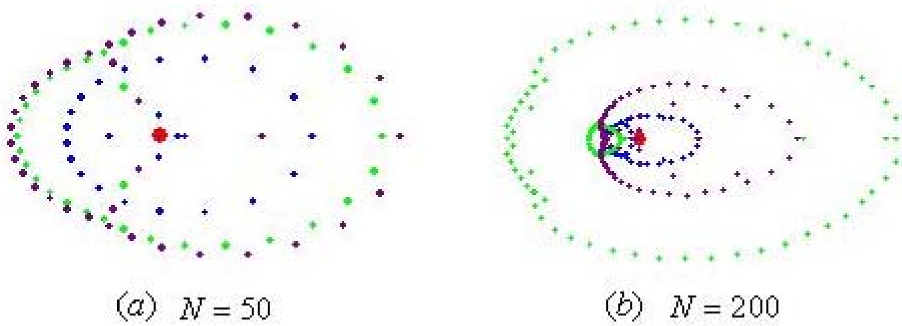

Figure 3.3: Same as Figures 3.1 and 3.2 but for the matrix C of (3.6). The eigenvalues of C are all 1 (red point).

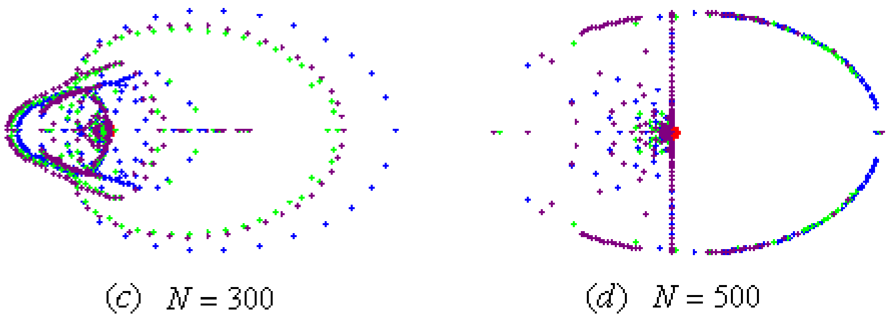

Figure 3.4: Same as Figures 3.1-3.3 but for the matrix A of (3.6). The eigenvalues of A are all 0 (red point). (c) and (d) shows eigenvalues of four randomly perturbed matrices A+E and this eigenvalues are marked by colour points plotted by Mathematica .

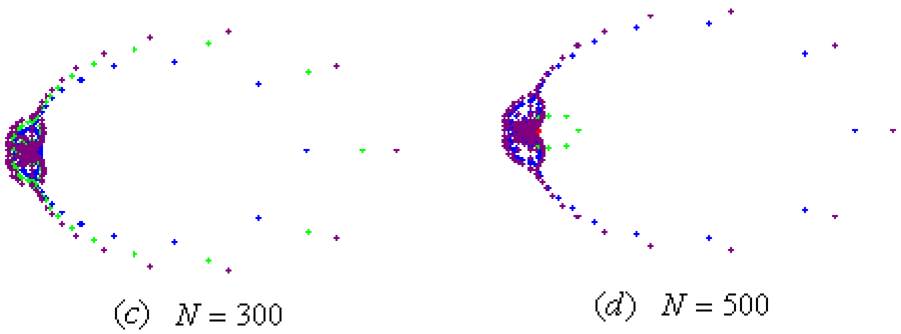

Figure 3.5: Same as Figure 3.1-3.4 but for the matrix B of (3.6). The eigenvalues of B are all 1 (red point).

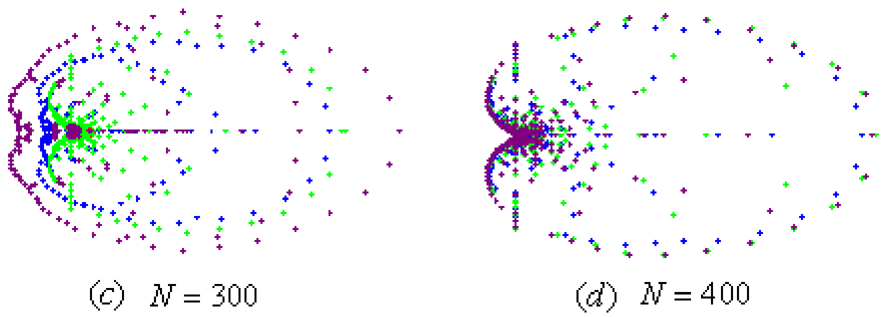

Figure 3.6: Same as Figures 3.1-3.5 but for the matrix C of (3.6). The eigenvalues of C are all 1 (red point). Note that the different scales in (a) and (b). As $N \to \infty$, the pseudospectra grow linearly to fill the right half plane.

The following two theorems make these observations precise. Although the theorems are stated for upper triangular Toeplitz matrices, the same results are valid in the lower triangular case.

Theorem 3.2 *Let A_N be an $N \times N$ (nondiagonal) triangular Toeplitz matrix with entries $\{a_k\}$ and symbol $f_N(z)$. If c_N and r are defined by*

$$c_N = \sum_{k=1}^{N-1} |a_k| > 0, \qquad r = \left(\frac{\varepsilon}{c_N}\right)^{1/N}, \tag{3.7}$$

then for any $\varepsilon \geqslant 0$,

$$f_N(\Delta_r) \subseteq \sigma_\varepsilon(A_N) \subseteq f_N(\Delta) + \Delta_\varepsilon. \tag{3.8}$$

Theorem 3.3 *Let A be a triangular Toeplitz operator with absolutely summable entries $\{a_k\}$ and symbol $f(z)$, and let A_N denote the $N \times N$ triangular Toeplitz matrix defined by $a_0, ..., a_N - 1$. Then*

$$\lim_{N\to\infty} \sigma_\varepsilon(A_N) = f(\Delta) + \Delta_\varepsilon = \sigma_\varepsilon(A) \tag{3.9}$$

for each $\varepsilon > 0$, and therefore

$$\lim_{\varepsilon\to 0} \lim_{N\to\infty} \sigma_\varepsilon(A_N) = f(\Delta) = \sigma(A). \tag{3.10}$$

Before proving these theorems, let us clarify a few points. In Theorem 3.2, the motivation is problems with $\varepsilon \ll c_N$ and thus $r < 1$, but this is not essential; the theorem holds as stated even if $r > 1$. Numerical experiments like those of Figures 3-5 suggest that the first set inclusion in (3.8) is typically much sharper than the second. In Theorem 3.3, the limit of sets is defined by $\lim\limits_{v\to\infty} \sigma_v = \{z \in : z_v \to z \; for \; some \; z_v \in \sigma_v\}$. The assumption of absolute summability of the entries a_k in this theorem can certainly be weakened.

Proof of Theorem 3.2. *To prove the first set of inclusion in (3.8), we construct pseudoeigenvectors as follows. Given* $\varepsilon \geqslant 0$*, let* r *be defined by (3.7). Now, given any* $\lambda \in f_N(\Delta_r)$*, let* $\lambda = f_N(z)$ *for some* $z \in \Delta_r$ *and define* $u = (1, z, z^2, \ldots, z^{N-1})^T$*. Then we readily calculate*

$$(\lambda I - A_N)\, u = z^N \begin{pmatrix} 0 & & & & & \\ a_{N-1} & \cdot & & & & \\ \vdots & \cdot & \cdot & & & \\ a_3 & & \cdot & \cdot & & \\ a_2 & a_3 & & \cdot & \cdot & \\ a_1 & a_2 & a_3 & \cdots & a_{N-1} & 0 \end{pmatrix} u, \tag{3.11}$$

and since the norm of this matrix is bounded by c_N*, this implies*

$$\frac{\|(\lambda I - A_N)\, u\|}{\|u\|} \leqslant |z|^N c_N \leqslant \varepsilon. \tag{3.12}$$

In other words, u *is an* ε*-pseudoeigenvector of* A_N*, and thus* $\lambda \in \sigma_\varepsilon(A_N)$ *by condition (ii) of the definition in the introduction.*

Our proof of the second set inclusion in (3.8) is less elementary; we do not know if it can be simplified. If $\lambda \in \sigma_\varepsilon(A_N)$*, then by condition (iv) in the introduction,* $\sigma_N(\lambda I - A_N) \leqslant \varepsilon$*, or equivalently, with the definition*

$$H_N = \begin{pmatrix} a_{N-1} & \cdots & a_2 & a_1 & a_0 - \lambda \\ \vdots & \iddots & a_1 & a_0 - \lambda & \\ a_2 & \iddots & a_0 - \lambda & & \\ a_1 & \iddots & & & \\ a_0 - \lambda & & & & \end{pmatrix} \tag{3.13}$$

$\sigma_N(H_N) \leqslant \varepsilon$*. This matrix* H_N *is known as a Hankel matrix, and by a result* *in the theory of AAK [27], or Caratheodory-Fejer (CF), approximation that goes back to Takagi in 1924,* $\sigma_N(H_N)$ *is equal to the distance in the supremum norm* $\|\phi\|_\infty = \sup_{z\in S}|\phi(z)|$ *between*

$$a_{N-1}z + \cdots + (a_0 - \lambda)z^N$$

and its best approximation $r(z) + g(z)$*, where* $r(z)$ *is a rational function of order at most N-1 with no poles in* Δ *[27,28]. (The order of a rational function is the maximum of the degrees of its numerator and denominator.) Therefore,*

$$\varepsilon \geqslant \left\|a_{N-1}z + \cdots + (a_0-\lambda)z^N - (r+g)(z)\right\|_\infty$$

$$= \left\|a_{N-1}z^{-1} + \cdots + (a_0-\lambda)z^{-N} - (r+g)(z^{-1})\right\|_\infty$$

$$= \left\|(a_0-\lambda) + \cdots + a_{N-1}z^{N-1} - z^N(r+g)(z^{-1})\right\|_\infty$$

$$= \left\|f_N(z) - \lambda - z^N(r+g)(z^{-1})\right\|_\infty \tag{3.14}$$

Now the function $(r+g)(z)$ *has winding number* $\geqslant N-1$ *about the origin when* z *traverses* S*, and therefore* $z^N(r+g)(z^{-1})$ *has winding number* $\geqslant 1$*. If* $\lambda \notin f_N(\Delta)$*, on the other hand, then* $f_N(z)-\lambda$ *is a function with winding number* 0 *on* S *and minimum modulus there equal to the distance from* λ *to* $f_N(\Delta)$*. As in the proof of Rouche's theorem, these facts are consistent with (3.14) only if that distance is* $\geqslant \varepsilon$*, or in other words,* $\lambda \in f_N(\Delta) + \Delta_\varepsilon$*.*

In Theorem 3.2, the constant c_N *entering into the definition of* r *can be improved by using the ideas of* AAK/CF *approximation in the first part of the proof as well as the second. The number* c_N *appeared only as an upper bound on the norm of the matrix in (3.11), and a sharper bound would be the number* $\tilde{c}_N$ *defined* *by*

$$\tilde{c}_N = \inf_{z_0 \in \mathbb{C}} \|f_N - z_0\|_\infty \tag{3.15}$$

that is, the radius of the smallest disk containing $f_N(D)$. As remarked above, however, the improvement will be insignificant in most applications because of the Nth root.

Proof of Theorem 3.3. *For any fixed $\varepsilon > 0$, the value r of (3.7) converges to 1 as $N \to \infty$, so by (3.8) we have $f(\Delta) \subseteq \lim_{N\to\infty} \sigma_\varepsilon(A_N) \subseteq f(\Delta) + \Delta_\varepsilon$. The second inclusion must be an equality, however, since in particular we can always perturb A by a multiple of the identity. This establishes the first equality of (3.9).*

Equation (3.10) follows from (3.9).

Note that Theorems 3.2 and 3.3 contain a rather intriguing assertion: although triangular Toeplitz matrices and operators have highly sensitive eigenvalues in general, their eigenvalues are completely insensitive outside the region $f(\Delta)$.

We close this section by mentioning that the pseudospectra of our matrices A, B, and C, besides showing something of the variety of behavior that may arise even within the narrow class of triangular Toeplitz matrices, also offer some more direct lessons for numerical analysis. The matrix A, which is the simplest nonhermition example that is not just a Jordan block, illustrates that in general pseudospectra are not just disks. This is an indication that looking at the Jordan cannonical form alone, although sufficient for linear perturbation analysis in theoreticalsense [29], is insufficient for perturbation analysis in practice-because a perturbation that is "small" for practical purposes may still lie far outside the linear range. The matrix B, with pseudospectra consisting of disks eccentrically situated with respect to the spectral point 1, provided an example in [30] to illustrate that the convergence of a matrix iteration may be determined in practice by the pseudospectrum rather than the spectrum; similar phenomena and further Toeplitz examples are discussed in [31]. The matrix C, finally, is adapted from *a well-known example devised by Kahan to prove that QR decomposition with column pivoting is not a failsafe method for determining the rank of a matrix [32]. Kahan's matrix differs slightly from C, having the Toeplitz structure modified by a column scaling, but the essence of it can still be seen in C: though the exact spectrum contains only the point 1, other points in the right half plane-in particular the point $z = 4$-lie wel within the pseudospectrum and indeed are ε-pseudoeigenvalues for values of ε that decrease exponentially as $N \to \infty$.*

REFERENCES

[1] *M. Gasanov, Y.Cesur. "On a nonlinear eigenvalue problem "Integral Equations and Operator Theory, 491-500, Birkhauser Verlag, Basel, 1997. (7)*

[2] *J. M. Varah. The compitation of bounds for the invariants subspaces of a general matrix operators. Technical Report CS 66, Computer Sicience Department, Stanford University, 1967. (1,3)*

[3] *H. J. Landau. On Szegeö's eigenvalue distribution theorem and non-Hermitian kernels. J. d'Analyse Math., 28:335-357, 1975. (2)*

[4] *S. K. Godunov, O. P. Kirilyuk, and V. I. Kostin. Spectral portraits of matrices. Technical Report Preprint 3, Inst. of Math., Sib. Branch of USSR Acad. Sci., 1990. In Russian. (2)*

[5] *S. K. Godunov and V. S. Ryabenkii. Theory of Difference Schemes: An Introduction. North-Holland, Amsterdam, 1964. Russian original published by Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow, 1962. (2)*

[6] *V. I. Kostin. On definition of matrices' spectra. In M. Durand and F. E. Dabaghi, editors, High Performance Computing II. North-Holland, Amsterdam, 1991. (2)*

[7] *V. I. Kostin and S. I. Razzakov. On convergence of the power orthogonal method of spectrum computing. Trans. Inst. Math. Sib. Branch Acad. Sci., 6:55-84, 1985. In Russian. (2)*

[8] *J. W. Demmel. A counterexample for two conjectures about stability. IEEE Trans. Auto. Control, AC-32:340-343, 1987. (2)*

[9] *J. W. Demmel. Nearest defective matrices and the geometry of ill-conditioning. In M. G. Cox. and S. Hammarling, editors, Reliable Numerical Computatiton, pages 35-55. Oxford University Press, Oxford, 1990. (2)*

[10] *D. Hinrichsen and B. Kelb. Spectral value set: A graphical tool for robustness analysis. Sys. Control Lett., 21:127-136, 1993. (3)*

[11] *D. Hinrichsen and A. J. Pritchard. On spectral variations under bounded real matrix perturbations. Numer. Math., 60:500-524, 1992. (2)*

[12] *L. N. Trefethen. Approximation theory and numerical linear algebra. In J. C. Mason and M. G. Cox, editors, Algorithms for Approximation II. Chapman and Hall, London, 1990 (3)*

[13] *L. N. Trefethen. Pseudospectra of matrices. In D. F. Griffiths and G. A. Watson, editors, Numerical Analysis 1991, pages 234-266. Longman Scientific and Technical, Harlow, Essex, UK, 1992. (3)*

[14] *L. N. Trefethen. Pseudospectra of linear operators. SIAM Review, 39:383-406, 1997. (3)*

[15] *S. C. Reddy. Pseudospectra of Operators and Discretization Matrices and an Application to Stability of the Method of Lines. Ph.D. thesis, MIT, Cambridge, MA. 1991. (3)*

[16] *S. C. Reddy and L. N. Trefethen. Lax-stability of fully discrete spectral methods via stability regions and pseudo-eigenvalues. Comp. Methods Appl. Mech. Eng., 80:147-164, 1990. (3)*

[17] *L. N. Trefethen and M. R. Trummer. An instability pheonomenon in spectral methods. SIAM J. Numer. Anal., 24:1008-1023, 1987. (3)*

[18] T. Kato. Perturbation Theory for Linear Opertators. Springer-Verlag, Berlin, corrected second edition, 1980. (12)

[19] R. A. Horn and C. R. Johnson. Matrix Analysis. Cambridge University Press, Cambridge, 1985. (14)

[20] J. H. Wilkinson. Sensitivity of eigenvalues Utilitas Math., 30:243-286, 1986. (14)

[21] L. N. Trefethen and M. Embre. Spectra and Pseudospectra. 2005. (16)

[22] F. L. Bauer and C. T. Fike. Norms and exclusion theorems. Numer. Math., 2:137-141, 1960. (17)

[23] G. W. Stewart and J. Sun. Matrix Perturbation Theory. Academic Press, San Diego, 1990. (17)

[24] J. H. wilkinson. The Algebraic Eigenvalue Problem. Oxford University Press, Oxford, 1965. (17)

[25] O. Toeplitz, Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veranderlichen, Math. Ann. 70:351-376, 1911. (28)

[26] A. Wintner. Zur Theorie der beschrankten Bilinearformen, Math. Z. 30:228-282, 1929. (28)

[27] V. M. Adamjan, D. Z. Arov, and M. G. Krein. Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem, Math. USSR-Sb. 15:31-73, 1971. (36)

[28] L. N. Trefethen. Rational Chebyshev approximation on the unit disk, Numer. Math. 37:297-320, 1981. (36)

[29] T. Kato, Perturbation Theory for Linear Operators. Springer-Verlag, New York, 1976. (37)

[30] L. N. Trefethen. Approximation theory and numerical linear algebra, in Algorithms for Approximation II, (J. C. Mason and M. G. Cox, Eds.), Chapman, London, 1990. (37)

[31] N. M. Nachtigal, L. Reichel, and L. N. Trefethen. A hybrid GMRES algorithm for nonsymmetric linear system, SIAM J. Matrix Anal. Appl., to appear. (37)

[32] W. Kahan. Numerical linear algebra, Canadian Math. Bull. 9:757-801, 1965. (37)