599 PAPER - SPECTRA AND PSEUDOSPECTRA

DANIEL REICH

Preface

Analyzing spectra of nonnormal matrices often provides either useless or misleading information. To understand these nonnormal matrices, pseudospectral techniques have been developed and continue to be refined. This paper provides an introduction to the study of pseudospectra. Most of the content closely follows *Spectra and Pseudospectra* by Mark Embree and Lloyd N. Trefethen as well as other works cited.

Contents

1. Introduction

1.1. **Spectrum of a matrix.** Let A be an $N \times N$ matrix with real or complex entries denoted by $A \in \mathbb{C}^{N \times N}$. Let v be a nonzero real or complex column vector of length N denoted by $v \in \mathbb{C}^N$ and let $\lambda \in \mathbb{C}$. Then v is an *eigenvector* of A with its corresponding *eigenvalue* λ, if

$$Av = \lambda v \tag{1}$$

The set of all eigenvalues of A is the *spectrum* of A denoted by $\sigma(A)$. The spectrum can also be defined by the set of points $z \in \mathbb{C}$ where the *resolvent* matrix $(zI - A)^{-1}$ does not exist.

If A has N distinct eigenvalues, then it is guaranteed to have a *complete set of eigenvectors*, a set of N linearly independent eigenvectors $v_1, ..., v_N$ with $Av_j = \lambda_j v_j$. These eigenvectors are unique up to normalization by scalar factors. For any matrix A with a complete set of eigenvectors, let V be the $N \times N$ matrix whose jth column is v_j, then

$$AV = V\Lambda \tag{2}$$

where Λ is the diagonal $N \times N$ matrix whose jth entry is λ_j.

Since the eigenvectors v_j are linearly independent, V is *nonsingular*, and therefore invertible. So we can multiply (2) on the right by V^{-1} to obtain

$$A = V\Lambda V^{-1} \tag{3}$$

known as an *eigenvalue decomposition* or a *diagonalization* of A. A matrix with a complete set of eigenvalues is therefore always *diagonalizable*, also known as *nondefective*.

Diagonalizing a matrix A allows us to more easily see the properties of the matrix A^k. Since $A^k = V\Lambda^k V^{-1}$, we need to understand Λ^k, which is simply a diagonal matrix with entries ${\lambda_j}^k$ $(j = 1, ..., N)$ along the diagonal. Thus we may be able to gain valuable insights into a matrix A^k by finding the eigenvalues of the matrix A. Much of the rest of this paper will be dedicated to understanding the limitations of eigenvalues and alternatives to them.

But before we look at the limitations of eigenvalues, let us note why they are often so useful. Eigenvalues allow us to decouple problems involving matrices or operators into several problems involving scalars, which are often easier to deal with individually. A noteworthy example is Fourier's solution of the heat equation via series expansions. Another use of eigenvalues is for analyzing the phenomenon of resonance. The frequencies at which strings, drums, radio signals, and mechanical systems vibrate correspond to eigenvalues of the linear or linearized operators that govern those systems. A third property of eigenvalues and the one that this paper will investigate most closely is their stability under perturbation, which governs their ability to accurately predict the evolution of a system.

Eigenvalue methods often fail to be useful when problems involve matrices or operators for which the matrix V^{-1}, if it exists, contains very large entries, i.e. $\|V^{-1}\| \gg 1$, where we are assuming V is reasonably scaled. If V is not reasonably scaled then we face the same difficulties when the *condition number* of V is large, i.e.

$$\|V\|\|V^{-1}\| \gg 1. \tag{4}$$

1.2. **Normal Matrix.** A matrix $A \in \mathbb{C}^{N\times N}$ is *normal* iff it has a complete set of orthogonal eigenvectors, that is, if it is unitarily diagonalizable:

$$A = U\Lambda U^* \tag{5}$$

An equivalent and more standard definition is that a matrix A is *normal* iff

$$A^*A = AA^* \tag{6}$$

We already know that eigenvectors are only determined up to a choice of scaling. Since we have just seen that for normal matrices $V^{-1} = V^*$, we know that the condition number of V in the 2-norm, $\kappa(V) = \|V\|_2\|V^{-1}\|_2$, can be made equal to 1 by the right choice of scaling. Therefore, at least in this norm, eigenvalue analysis may fail to be useful for nonnormal matrices, which brings us to our study of Pseudospectra.

2. Pseudospectra

2.1. **Motivation for Pseudospectra.** If A is nearly singular then $\|A^{-1}\|$ is large. If a singular matrix is slightly perturbed, then the norm of its inverse must be large. Consequently, we adopt the convention that $\|A^{-1}\|$ is infinite when A is singular and not invertible.

The eigenvalues of A may be ill-conditioned and not robust to perturbation, but the property that $\|(zI - A)^{-1}\| > \frac{1}{\epsilon}$ is well-conditioned. If z is an eigenvalue, $(zI - A)$ is singular and if z is close to being an eigenvalue of A, then $\|(zI - A)^{-1}\|$ will be large.

We will see in the next section that when a matrix is normal, eigenvalues are not highly sensitive to perturbations, but when a matrix is highly nonnormal the opposite is true. We just stated that if z is close to being an eigenvalue of A, then $\|(zI - A)^{-1}\|$ will be large. But in the case of nonnormal matrices $\|(zI - A)^{-1}\|$ may be large even when z is far from the spectrum. We are now ready to formally introduce Pseudospectra, which we can use to understand these matrices that are far from normal.

Definition 1. Let $A \in \mathbb{C}^{N\times N}$ and $\epsilon > 0$ be arbitrary. The ϵ-*pseudospectrum* $\sigma_\epsilon(A)$ is the set $z \in \mathbb{C}$ such that $\|(zI - A)^{-1}\| > \epsilon^{-1}$.

Definition 2. Let $A \in \mathbb{C}^{N\times N}$ and $\epsilon > 0$ be arbitrary. The ϵ-*pseudospectrum* $\sigma_\epsilon(A)$ is the set $z \in \mathbb{C}$ such that $z \in \sigma(A + E)$ for some $E \in \mathbb{C}^{N\times N}$ with $\|E\| < \epsilon$.

Definition 3. Let $A \in \mathbb{C}^{N\times N}$ and $\epsilon > 0$ be arbitrary. The *ϵ-pseudospectrum* $\sigma_\epsilon(A)$ is the set $z \in \mathbb{C}$ such that $\|(zI - A)v\| < \epsilon$ for some $v \in \mathbb{C}^N$ with $\|v\| = 1$.

2.2. **Equivalence of definitions.** First we must note that in each definition above, the spectrum of A is clearly contained in the pseudospectrum. Therefore for $z \in \sigma(A)$ the equivalence of these definitions is trivial. So let us assume that $z \notin \sigma(A)$ implying the existence of $(zI - A)^{-1}$.

Definition 1 implies Definition 2. Suppose the conditions of Definition 1, i.e suppose $\|(zI - A)^{-1}\| > \epsilon^{-1}$. Then there must exist some $u \in \mathbb{C}^N$ so that $(zI-A)^{-1}u = (\epsilon')^{-1}v$ for some $v \in \mathbb{C}^N$ with $\|v\| = \|u\| = 1$ and with $\epsilon' < \epsilon$. Consequently, $zv - Av = \epsilon' u$. We need a matrix E such that $Ev = \epsilon' u$ where $\|E\| \le \epsilon'$, for then v will be an eigenvector of $A + E$ with eigenvalue z. The matrix E can be constructed by $E = \epsilon' u w^*$, where $w^* v = 1$ and $\|w\| = 1$. In the 2-norm, we can choose $w = v$. In the case of an arbitrary norm, the existence of w is guaranteed by the Hahn-Banach theorem.

Definition 2 implies Definition 3. Suppose the conditions of Definition 2, i.e. suppose $(A + E)v = zv$ for some $E \in \mathbb{C}^{N\times N}$ with $\|E\| < \epsilon$ and some nonzero $v \in \mathbb{C}^N$, which we may take to be normalized since an eigenvector is unique only up to scalar normalizations, $\|v\| = 1$. Then $(zI - A)v = Ev$ and $\|(zI - A)v\| \le \|E\|\|v\| = \epsilon$.

Definition 3 implies Definition 1. It follows from Definition 3 that if $\|(zI - A)v\| < \epsilon$, there exists some $u \in \mathbb{C}^N$ and ϵ' such that $(zI - A)v = \epsilon' u$ with $\|v\| = \|u\| = 1$ and with $\epsilon' < \epsilon$. We know that $(zI - A)$ is invertible, so we have $(zI - A)^{-1}u = (\epsilon')^{-1}v$ so that $\|(\epsilon')^{-1}v\| = \|(zI - A)^{-1}u\| \le \|(zI - A)^{-1}\|$. Therefore, $\|(zI - A)^{-1}\| \ge (\epsilon')^{-1} > \epsilon^{-1}$.

Definition 1 in the case of the 2-norm. Suppose the conditions of Definition 1 using the 2-norm, i.e suppose $\|(zI - A)^{-1}\|_2 > \epsilon^{-1}$. Since $\|(zI - A)^{-1}\|_2$ is the reciprocal of the smallest singular value of $zI - A$, it follows that the *ϵ-pseudospectrum* $\sigma_\epsilon(A)$ is the set $z \in \mathbb{C}$ such that the smallest singular value of $zI - A$ is less than ϵ.

2.3. **The usefulness of multiple definitions.** The equivalent definitions above provide various methods for calculating and analyzing pseudospectra. Definition 1 allows one to test for a given value of ϵ whether z is in the ϵ-pseudospectrum. This definition is well suited for calculating a contour outlining the boundary of the pseudospectrum. Definition 2 allows for direct computation of the ϵ-pseudospectrum via eigenvalues of a slightly perturbed matrix. These computations provide a cloud of points sampling the pseudospectrum's boundary. Definition 3 allows one to test possible pseudoeigenvector-pseudoeigenvalue pairs.

3. Perturbation Theory

Let us further understand the implications of our second definition of the pseudospectrum by investigating the sensitivity of the eigenvalues of a matrix A to small perturbations of the matrix.

Given a matrix $A \in \mathbb{C}^{N\times N}$, we are interested in the evolution of the eigenvalues and eigenvectors of the perturbed matrix $A(t) = A + tE$, where $t \ll 1$ and $\|E\| = 1$.

A distinct Eigenvalue $\lambda_j \in \sigma(A)$ is a simple root of the characteristic equation $\det(\lambda I - A) = 0$. By the Implicit Function theorem λ must be a C^∞ function of t in some neighborhood of $t = 0$. Accordingly, $\lambda_j(t) \in \sigma(A(t))$ is a simple root of the characteristic equation $\det(\lambda(t)I - A(t)) = 0$. We can write each eigenvalue $\lambda_j(t)$ of $A(t)$ by the following Taylor series:

$$\lambda_j(t) = \lambda_j + \alpha_{1j}t + \alpha_{2j}t^2 + ... \tag{7}$$

Similarly, we can use the eigenvectors of A as a basis of $\mathbb{C}^{N\times N}$ and write the eigenvector $v_j(t)$ of $A(t)$ by the Taylor series

$$v_j(t) = v_j + \sum_{k\neq j}^{N}[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...]v_k \tag{8}$$

where we choose the coefficient of v_j to be 1 since eigenvectors are only unique up to normalization by scalar factors.

3.1. **Distinct Eigenvalues.**

In a neighborhood of a matrix with distinct eigenvalues, the eigenvalues are continuous differentiable functions of the matrix entries. When considering pseudospectra, it is crucial to know how small perturbations in matrix coefficients affect the eigenvalues. For every eigenvalue of A, we have associated left and right eigenvectors such that

$$Av_j = \lambda_j v_j \tag{9}$$

and

$$u_j{}^*A = \lambda_j u_j{}^* \tag{10}$$

It follows directly that for any two distinct eigenvalues λ_i and λ_j, $u_j{}^*v_i = 0$, which can be seen by the following calculation:

$$\begin{aligned} &u_j{}^*Av_i = u_j{}^*\lambda_j v_i \qquad \text{and} \qquad u_j{}^*Av_i = u_j{}^*\lambda_i v_i \\ \Rightarrow &\lambda_j u_j{}^*v_i = \lambda_i u_j{}^*v_i \\ \Rightarrow &u_j{}^*v_i = 0 \end{aligned}$$

We can find an exact expression for α_{1j} in (7) as follows:

$$\begin{aligned}
A(t)v_j(t) &= Av_j(t) + tEv_j(t) \\
&= A\left(v_j + \sum_{k\neq j}^{N}[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...]v_k\right) + tEv_j(t) \\
&= Av_j + A\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k + tEv_j(t) \\
&= \lambda_j v_j + A\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k + tEv_j(t)
\end{aligned}$$

and since $A(t)v_j(t) = \lambda_j(t)v_j(t)$, it follows from (10) that

$$\begin{aligned}
{u_j}^*\lambda_j(t)v_j(t) &= {u_j}^*\lambda_j v_j + {u_j}^*A\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k + {u_j}^*tEv_j(t) \\
&= {u_j}^*\lambda_j v_j + \lambda_j {u_j}^*\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k + {u_j}^*tEv_j(t) \\
&= \lambda_j {u_j}^*v_j + {u_j}^*tEv_j(t)
\end{aligned}$$

so,

$$\begin{aligned}
\lambda_j(t) - \lambda_j &= \frac{{u_j}^*tEv_j(t)}{{u_j}^*v_j} \\
&= t\frac{{u_j}^*Ev_j}{{u_j}^*v_j} + \mathcal{O}(t^2)
\end{aligned}$$

and therefore

$$\alpha_{1j} = \frac{{u_j}^*Ev_j}{{u_j}^*v_j} \tag{11}$$

α_{1j} is always bounded by the condition number of λ_j, which is defined as $\kappa(\lambda_j) = \frac{\|u_j\|\|v_j\|}{|{u_j}^*v_j|}$. By the Cauchy-Schwarz inequality $|{u_j}^*v_j| \leq \|u_j\|\|v_j\|$ so that $\kappa(\lambda_j) \geq 1$. When $u_j = v_j, \kappa(\lambda_j) = 1$ and this is the case when A is normal. But when A is nonnormal, $\kappa(\lambda_j)$ may be arbitrarily large. In fact, by choosing $E = \frac{u_j {v_j}^*}{\|u_j\|\|v_j\|}$, we can achieve these arbitrarily large values of $\alpha_{1j} = \frac{1}{{u_j}^*v_j}$ and see that they occur when u_j and v_j are nearly orthogonal, which corresponds to A being highly nonnormal.

3.2. **Eigenvectors.** To understand how a small perturbation of A affects the eigenvectors, we need to find an expression for $\beta_{kj}^{(1)}$ in (8). We can accomplish this as follows:

$$
\begin{aligned}
0 &= A(t)v_j(t) - \lambda_j(t)v_j(t) \\
&= {u_i}^* A(t)v_j(t) - {u_i}^*\lambda_j(t)v_j(t) \\
&= {u_i}^* A(t)v_j(t) - {u_i}^*\lambda_j(t)v_j - {u_i}^*\lambda_j(t)\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k \\
&= {u_i}^* A(t)v_j(t) - {u_i}^*\lambda_j(t)\left[\beta_{ij}^{(1)}t + \beta_{ij}^{(2)}t^2 + ...\right]v_i \\
&= {u_i}^* A(t)v_j(t) - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= {u_i}^*\left[A + tE\right]v_j(t) - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= {u_i}^* Av_j(t) + {u_i}^* tEv_j(t) - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= {u_i}^* Av_j(t) + {u_i}^* tEv_j - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= \lambda_i {u_i}^* v_j(t) + {u_i}^* tEv_j - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= \lambda_i {u_i}^* v_j + \lambda_i {u_i}^*\sum_{k\neq j}^{N}\left[\beta_{kj}^{(1)}t + \beta_{kj}^{(2)}t^2 + ...\right]v_k + {u_i}^* tEv_j - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2) \\
&= \lambda_i {u_i}^*\beta_{ij}^{(1)}tv_i + {u_i}^* tEv_j - {u_i}^*\lambda_j\beta_{ij}^{(1)}tv_i + \mathcal{O}(t^2)
\end{aligned}
$$

so,

$$\lambda_j {u_i}^* v_i t\beta_{ij}^{(1)} - \lambda_i {u_i}^* v_i t\beta_{ij}^{(1)} = {u_i}^* tEv_j + \mathcal{O}(t^2)$$

and therefore

$$\beta_{ij}^{(1)} = \frac{{u_i}^* Ev_j}{{u_i}^* v_i(\lambda_j - \lambda_i)} \qquad \text{where } i \neq j \tag{12}$$

By choosing our normalization so that ${u_i}^* v_i = 1$, and choosing $E = \frac{u_i {v_j}^*}{\|u_i\|\|v_j\|}$, we can achieve very large values of $\beta_{ij}^{(1)} = \frac{\|u_i\|\|v_j\|}{\lambda_j - \lambda_i}$ when $\|u_i\|\|v_j\| \gg 1$, which is the case for nonnormal matrices. For any normal or nonnormal matrix, we can also achieve large values of $\beta_{ij}^{(1)}$ for eigenvalues that are close to one another.

4. Toeplitz Operators

Our perturbation analysis has demonstrated the weakness of spectral methods in analyzing nonnormal matrices and confirmed the need for pseudospectral methods. Now let us begin to understand the importance of these pseudospectra by looking at one of the best understood families of nonnormal matrices: non-Hermitian Toeplitz matrices. An $N \times N$ *Toeplitz matrix* is a matrix whose entries are constant along diagonals:

$$A = \begin{bmatrix} a_o & a_{-1} & & \cdots & a_{1-N} \\ a_1 & a_0 & & & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & & a_0 & a_{-1} \\ a_{N-1} & \cdots & & a_1 & a_0 \end{bmatrix}$$

A semi-infinite matrix of the same form is a *Toeplitz operator* and a doubly-infinite matrix of this form is a *Laurent operator*. A *circulant matrix* is the finite-dimensional analogue of a Laurent operator in which the entries wrap around periodically, i.e. $a_j = a_{j-N}$ for $1 \leq j \leq N-1$.

Let us consider the *family* of Toeplitz operators A of varying dimensions and denote this by $\{A_N\}$. As $N \to \infty$, we in some sense approach the Toeplitz and Laurent operators, but in what sense?

To help us understand this, let us look at a specific and useful family of Toeplitz matrices, the shift operators. Let us begin by understanding the $N \times N$ left shift operators with 1's on the superdiagonal and 0's elsewhere:

$$A_N = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ & \ddots & \ddots & \ddots & \\ \vdots & & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

We know that these finite-dimensional matrices have no nonzero eigenvalues. Recall that for infinite-dimensional matrices, the spectrum is defined as the set of numbers $z \in \mathbb{C}$ for which the resolvent $(zI - A)^{-1}$ does not exist as a bounded operator. When we consider the semi-infinite-dimensional analogue, we can see that the spectrum includes all points in the unit circle by the following computation $Au = z^{-1}u$:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \ddots \\ 0 & 0 & 0 & 0 & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ z^{-1} \\ z^{-2} \\ z^{-3} \\ \vdots \end{bmatrix} = z^{-1} \begin{bmatrix} 1 \\ z^{-1} \\ z^{-2} \\ z^{-3} \\ \vdots \end{bmatrix}$$

Since we want $u \in l^2$, we have that each z^{-1} satisfying $|z| > 1$ is an eigenvalue of the left shift Toeplitz operator with eigenvector u.

Let us analyze the right shift matrix. As with the left shift matrix, there are no nonzero eigenvalues for the finite-dimensional matrices. However, the difference is evident in that the semi-infinite right shift Toeplitz operator has no nonzero eigenvalues; the first entry of any vector applied to the operator is zeroed out. We can instead look at the doubly-infinite analogue, the right shift Laurent operator, where $Av = zv$:

$$\begin{bmatrix} \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & 0 & 0 & 0 & \ddots \\ \ddots & 1 & 0 & 0 & \ddots \\ \ddots & 0 & 1 & 0 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ z^1 \\ 1 \\ z^{-1} \\ \vdots \end{bmatrix} = z \begin{bmatrix} \vdots \\ z^1 \\ 1 \\ z^{-1} \\ \vdots \end{bmatrix}$$

We can see that z includes no eigenvalues of the right shift Laurent operator corresponding to eigenvectors in l^2. At best, eigenvalues z with $|z| = 1$ correspond to eigenvectors $v \in l^\infty$. In the l^2 case, these z are in the spectrum even if not as eigenvalues.

Extending our results for the left shift operator, we can see that any Toeplitz operator with 1's on the nth superdiagonal and 0's elsewhere will have $Au = z^{-n}u$, where $u \in l^2$ for eigenvalues $z : |z| > 1$. Similarly, any Laurent operator with 1's on the nth subdiagonal and 0's elsewhere will have $Av = z^n v$ with $v \in l^\infty$ for eigenvalues $z : |z| = 1$.

Now, it follows naturally to define the *symbol* of any Toeplitz matrix, Toeplitz operator or Laurent operator as the function

$$f(z) = \sum_k a_k z^k \tag{13}$$

where a_k denotes the entries of A as we defined earlier in this section. Additionally, let us define the *symbol curve* as the image of the symbol on the unit circle, $f(\mathbb{T})$ with $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$. We have just seen that the symbol provides information about the spectra of both the left shift Toeplitz and right shift Laurent operators via the equations $Au = f(z)u$ and $Av = f(z)v$, respectively. The following theorem generalizes the relationship between the symbol and the spectra to any Toeplitz or Laurent operator.

Theorem 4.1 (Spectra of Toeplitz and Laurent operators)**.** *Let A be a Laurent or Toeplitz operator with continuous symbol f.*

(i) *If A is a Laurent operator, then $\sigma(A) = f(\mathbb{T})$.*
(ii) *If A is a Toeplitz operator, then $\sigma(A) = f(\mathbb{T})$ together with all the points enclosed by this curve with nonzero winding number.*

The winding number of a point is defined as the number of times the curve passes around that point, where positive numbers correspond to the counter-clockwise direction.

Now let us connect this theorem back to our family of finite Toeplitz matrices and our discussion of Pseudospectra with the following theorem:

Theorem 4.2 (Behavior of pseudospectra as $N \to \infty$)**.** *Let A be a Toeplitz operator with continuous symbol f and let $\{A_N\}$ be the associated family of*

Toeplitz matrices. Then for any $\epsilon > 0$*,*

$$\lim_{N\to\infty} \sigma_\epsilon A_N = \sigma_\epsilon(A), \tag{14}$$

and thus

$$\lim_{\epsilon\to 0} \lim_{N\to\infty} \sigma_\epsilon(A_N) = \sigma(A) \tag{15}$$

So we can see that in the limiting case, the pseudospectra of finite-dimensional Toeplitz matrices are bounded by the symbol curves $f(\mathbb{T})$.

Let us now understand these theorems in the context of the left shift operator. We can see that even for this lower dimension $N = 50$, as $\epsilon \to 0$, the

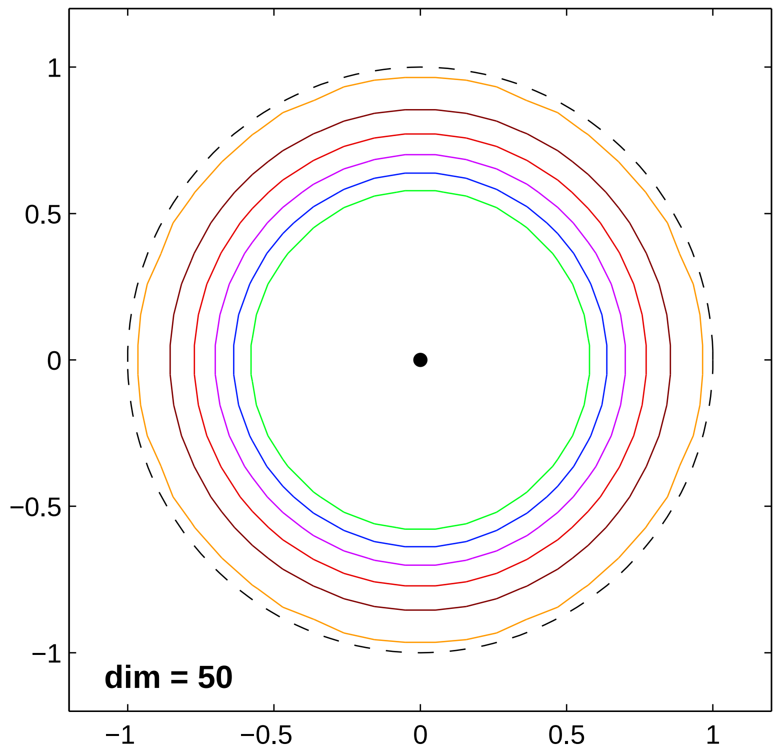

FIGURE 1. The solid line contours are the boundaries of the ϵ-pseudospectra for the left-shift matrix with $N = 50$ and $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}$ as calculated by EigTool. The outer dashed line is the symbol curve.

boundaries of the pseudospectra are converging to the symbol curve. We will discuss the algorithm and program used for computing these pseudospectra in the next section.

Lets look at another Toeplitz matrix A_N, defined by its symbol

$$f_{butterfly}(z) = z^2 - iz + iz^{-1} - z^{-2}. \tag{16}$$

In Figure 2, the spectrum of $\{A_N\}$ is not approaching the spectrum of the

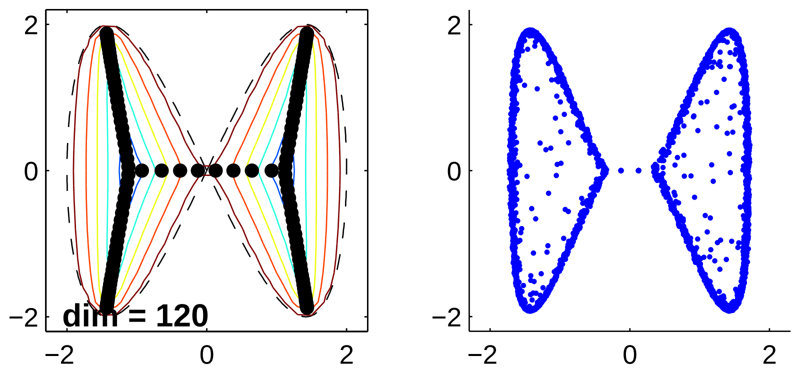

FIGURE 2. On the left, the solid line contours are the boundaries of the ϵ-pseudospectra for the butterfly matrix with $N = 120$ and $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}$ as calculated by EigTool. The outer dashed line is the symbol curve and the bold dots are the eigenvalues. On the right, eigenvalues of 20 random perturbations of norm 10^{-3}.

Toeplitz operator A. However, as the theorem asserted, the pseudospectra are converging, even for $N = 120$. In the figure on the left, the pseudospectra were computed by EigTool. Whereas in the figure on the right, all eigenvalues of 20 perturbed matrices were computed according to definition 2 of the pseudospectrum. As we can see, the brute force method of calculating all eigenvalues agrees with the results obtained from EigTool. Notice that the sparsity pattern of the butterfly matrix allowed for these brute force computations to be done at a significantly lower cost than would have been the case for a dense matrix of the same dimension.

5. Computing Pseudospectra

Now that we have seen Figures 1 & 2, let us understand both the construction of and motivation for the EigTool algorithm for computing pseudospectra.

5.1. **Computational Limitations.** Our second definition of pseudospectra is used to compute the boundary of an ϵ-pseudospectrum, thereby identifying the entire region included in that pseudospectrum. This definition requires direct computation of the full spectra of a perturbed matrix. For large matrices this $\mathcal{O}(N^3)$ cost makes the calculation impossible due to memory and time limitations. To work around this issue, we need an algorithm that computes only the part of the spectrum containing the needed and most useful information. In Matlab, the Eigs routine is used for this purpose.

5.2. **Iterative Methods.** The $\mathcal{O}(N^3)$ cost of calculating full spectra is likely unavoidable for random matrix problems. Fortunately, in practice most matrices have some kind of structure and techniques have been found to exploit many of these structures. For nonsymmetric matrices, Eigs uses the Arnoldi iteration and for symmetric matrices, Eigs uses the Lanczos iteration. Both methods are based on the idea of projecting a matrix of higher dimension onto a lower-dimensional Krylov subspace.

Given a matrix A and a vector b, the set of vectors $b, Ab, A^2b, A^3b, ...$, is a *Krylov sequence*. These sequences can be computed in the form b, Ab, $A(Ab)$, $A(A(Ab)), ...$, by the *black box*:

$$b \to BLACKBOX \to Ab \tag{17}$$

whose cost is often significantly less than $\mathcal{O}(N^2)$. Successively larger groups of these vectors span the Krylov subspaces. Projection into these subspaces reduces the original matrix problem to a sequence of matrix problems of dimensions $1, 2, 3, ...$

The structure of the matrix determines the black box, which in turn determines the work per step. One example is sparse matrices, where all the zero calculations can be avoided, and another is Toeplitz matrices, where the calculations correspond to convolutions in Fourier space. Most linear algebra algorithms, such as Gaussian elimination and QR factorization, that are applied to dense matrices require $\mathcal{O}(N)$ steps, each requiring $\mathcal{O}(N^2)$ work, for a total cost of $\mathcal{O}(N^3)$. For iterative methods, this is the worst-case cost. The ideal iterative method reduces the number of steps from $\mathcal{O}(N)$ to $\mathcal{O}(1)$ and the work per step from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, reducing the cost to $\mathcal{O}(N)$. More typically, one can hope to reduce the cost from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2)$.

5.3. **Arnoldi Iteration.** Arnoldi iteration reduces a matrix $A \in \mathbb{C}^{N\times N}$ to *Hessenberg form*, a matrix H whose entries below the first subdiagonal are all 0. The full reduction is the orthogonal similarity transform $H = Q^*AQ$, where H has the same eigenvalues as A. In practice, computing the full reduction would result in the same $\mathcal{O}(N^3)$ cost that is motivating our use of these iterations in the first place. Instead, we consider the rectangular matrix $Q_n \in \mathbb{C}^{N\times n}$, whose columns are the first n columns of Q. Let $H_n \in \mathbb{C}^{n+1\times n}$ denote the $(n+1)\times n$ upper-left section of H, which is also a Hessenberg matrix, then we have $AQ_n = Q_{n+1}H_n$:

$$A_{N\times N}\left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_n \\ & & \end{array}\right] = \left[\begin{array}{c|c|c} & & \\ q_1 & \cdots & q_{n+1} \\ & & \end{array}\right]\begin{bmatrix} h_{1,1} & \cdots & h_{1,n} \\ h_{2,1} & & \vdots \\ & \ddots & \\ & & h_{n+1,n} \end{bmatrix}$$

where the n^{th} column Aq_n can be written

$$Aq_n = h_{1,n}q_1 + \ldots + h_{n,n}q_n + h_{n+1,n}q_{n+1}. \tag{18}$$

So we can see that the vectors q_{n+1} can be found recursively from the previous n vectors q_n.

Algorithm for Arnoldi Iteration

```
b = arbitrary
q_1 = b/||b||
for n = 1, 2, 3, ...
   v = A q_n
   for j = 1 to n
      h_{j,n} = q_j^* v
      v = v - h_{j,n} q_j
   h_{n+1,n} = ||v||
q_{n+1} = v / h_{n+1,n}
```

Since H_n is not square, let us denote the matrix composed of its first n rows by $\tilde{H}_n$. The eigenvalues of $\tilde{H}_n$, known as *Ritz values*, are computed by standard QR algorithms. The Ritz values are estimates of the eigenvalues of H, and therefore of A as well, and their convergence is a strong indicator that they are indeed eigenvalues of A. Although I will not give a detailed explanation of why this is so, the Arnoldi iteration is an attempt to solve a problem involving a Krylov subspace that is analogous to least-squares. Since $n \ll N$, this method cannot recover all the eigenvalues of A. According to Trefethen & Bau [1], the Arnoldi iterations rapidly converge, often with geometric convergence, to eigenvalues near the edge of the spectrum, which are precisely the ones of greatest interest in most applications.

5.4. **Lanczos Iteration.** The Lanczos iteration is the Arnoldi iteration specialized for the hermitian case $A = A^*$. We know that $h_{i,j} = q_i{}^*Aq_j$, which implies that $h_{i,j} = 0$ for $i > j+1$ since $Aq_j \in \langle q_1, \ldots, q_{j+1}\rangle$ and the Krylov vectors are orthogonal. Taking the conjugate transpose, we have $h^*_{i,j} = q_j{}^*A^*q_i = q_j{}^*Aq_i$, so by the same argument, $h_{i,j} = 0$ for $j > i+1$. We know that H_n is just the $(n+1) \times n$ upper-left section of H, and we know that $H^* = Q^*A^*Q$. So when A is hermitian, $H = H^*$, and therefore our Hessenberg matrices H_n are not only tridiagonal, but also symmetric. This allows us to replace the bounds of the inner loop of the algorithm for the Arnoldi iteration by $n-1$ to n, leading to a three-term recurrence relationship instead of the $(n+1)$-term recurrence relationship in equation (18). The result is, of course, a much less costly iteration.

5.5. **EigTool.** EigTool is Matlab routine that was constructed for the purpose of calculating pseudospectra. The algorithm is as follows:

Algorithm for EigTool
1) Choose max dimension p for subspace - larger p for better pseudospectra.
2) Choose number of eigenvalues k - larger k for better pseudospectra.
3) Run Eigs to obtain H_p, where p denotes the dimension $(p+1) \times p$.
4) Define a grid over the region of $\mathbb{C}$ enclosing converged Ritz values.
5) For each grid point z,
   a) Perform reduced QR factorization of shifted matrix: $zI_p - H_p = Q_p R$
   b) Get $\lambda_{\max}(z)$ from Lanczos iteration on $(R^* R)^{-1}$, random starting vector.
   c) $\sigma_{\min}(z) := 1/\sqrt{\lambda_{\max}(z)}$
6) Start GUI and create contour plot of the $\sigma_{\min}$ values.
7) Allow adjustment of parameters (e.g. grid size, contour levels) in GUI.

We have already understood 1-3. In step 4, defining a grid allows us to avoid computing singular values in uninteresting regions of the complex plane. According to Trefethen [3] and Trefethen & Bau [1], the areas where Ritz values rapidly converge are most likely our areas of interest, whereas in the areas where they do not, the resolvent norm is most likely small and these areas are therefore not of interest. Consequently, calculating pseudospectra becomes a local problem and thus significantly less costly.

Step 5 is to find the singular values of H_p by exploiting the structure of the Hessenberg matrix. The QR factorization of the matrix, after shifting for each grid point, gives an upper-triangular matrix R with the same singular values as the shifted matrix. We can then use the Lanczos iteration to find the largest eigenvalue of $(R^* R)^{-1}$, so that we can obtain $\sigma_{\min}(R) = \sigma_{\min}(zI_p - H_p)$.

Step 6 applies the definition of the pseudospectrum in the case of the 2-norm, creating contours where $\sigma_{\min}(z) < \epsilon$ for various values of ϵ. Finally, step 7 allows one to adjust the parameters that only require redoing steps 5 and 6.

6. Pseudospectra Computations: Twisted Toeplitz

In section 4, we analyzed and understood many spectral and pseudospectral properties of Toeplitz matrices, where the entries are constant along diagonals. We will now numerically explore a specific family of twisted Toeplitz matrices:

Definition. *Let $c_{-n}, ..., c_m$ be 2π-periodic coefficient functions. The associated family of twisted Toeplitz matrices is the set of (m, n)-periodic matrices $\{A_N\}_{N \geq 1}$ with coefficients*

$$a_{jk} = c_{(j-k)(\text{mod } N)}(x_j)$$

where

$$x_j = 2\pi j/N, \quad 1 \leq j \leq N.$$

6.1. **Setup.** Let us consider an $N \times N$ bidiagonal matrix A defined by

$$a_{j,j} = x_j \qquad a_{j,j+1} = \frac{1}{2}x_j \tag{19}$$

The eigenvalues of A are its diagonal entries, which are all real numbers in $(0, 2\pi]$. We constructed the perturbed matrices $A(t) = A+tE$ with normally distributed complex matrices E.

6.2. **Computations.** We will now compare pseudospectra calculations made with EigTool to those made with other algorithms.

The eigenvalues shown in Figure 3, for the case $N = 60$ and $t = 10^{-12}$, were computed by the Matlab routine Condeig. The routine reduces the system to Upper Hessenberg using unitary transformations and then applies iterative QR factorizations with shifts, an essential part of the EigTool algorithm. The error in this calculation was estimated by Condeig so we are confident that the precision of the eigenvalues is high compared with the sizes of perturbations being considered. Not surprisingly, in view of the similarity of the algorithms, this calculation agreed with the pseudospectra for $t = 10^{-12}$ as calculated by EigTool, which is one of the curves seen in Figure 4.

We then tested our first-order perturbation expansions from equations (7) and (8). We calculated the eigenvalues for the same matrix $A(t)$ using our first-order perturbation approximation and found that these eigenvalues were computed with large errors, which is shown in Figure 5. Additionally and not surprisingly, the first-order perturbation approximation of the eigenvectors was increasingly inaccurate for larger and larger values of N. The error in both the first order approximations of the eigenvalues and the eigenvectors is seen in Figure 6.

6.3. **Inaccurate Perturbations.** Clearly, the perturbation approximations were highly inaccurate for all perturbations when $N \geq 30$. The first-order eigenvalue computations were accurate for $N \leq 30$ as were the eigenvector approximations to a lesser extent.

The condition number of the matrix V of eigenvectors of A grew exponentially with N. Consequently, the first-order coefficients α_1 and $\beta^{(1)}$ in equations (11) and (12), respectively, became increasingly large. For the case $N = 60$, they reached the order of 10^{18}, dominating even the smallest t values. The theory that was used to obtain the perturbation expansion required these constants to be on the same order as the eigenvalues and eigenvectors being perturbed. Thus the failure in the expansions.

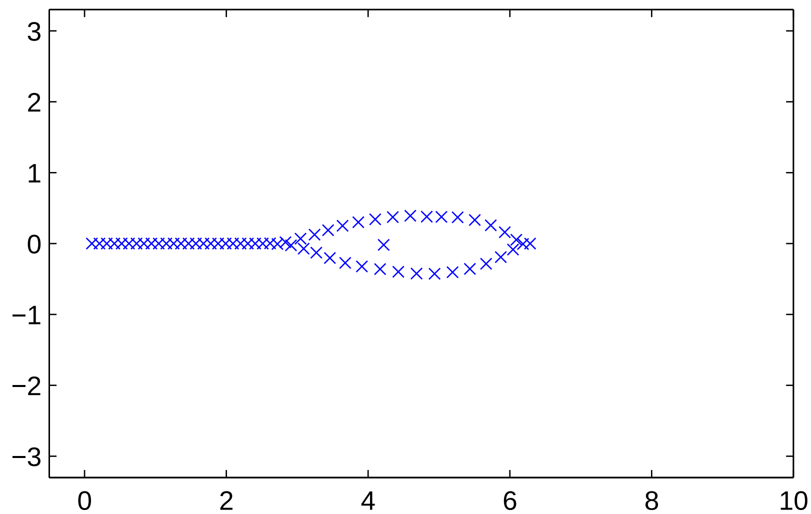

FIGURE 3. Eigenvalues calculated with Matlab's Condeig routine for $N = 60$ and $t = 10^{-12}$.

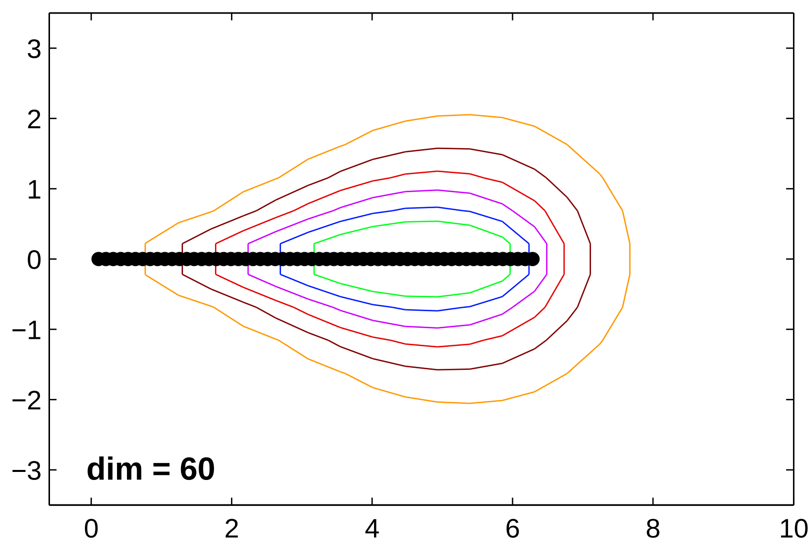

FIGURE 4. The pseudospectra as computed by Eigtool with $N = 60$ and $t = 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}, 10^{-10}, 10^{-12}$. The thick line at $im(\lambda) = 0$ is the plot of the actual eigenvalues of A.

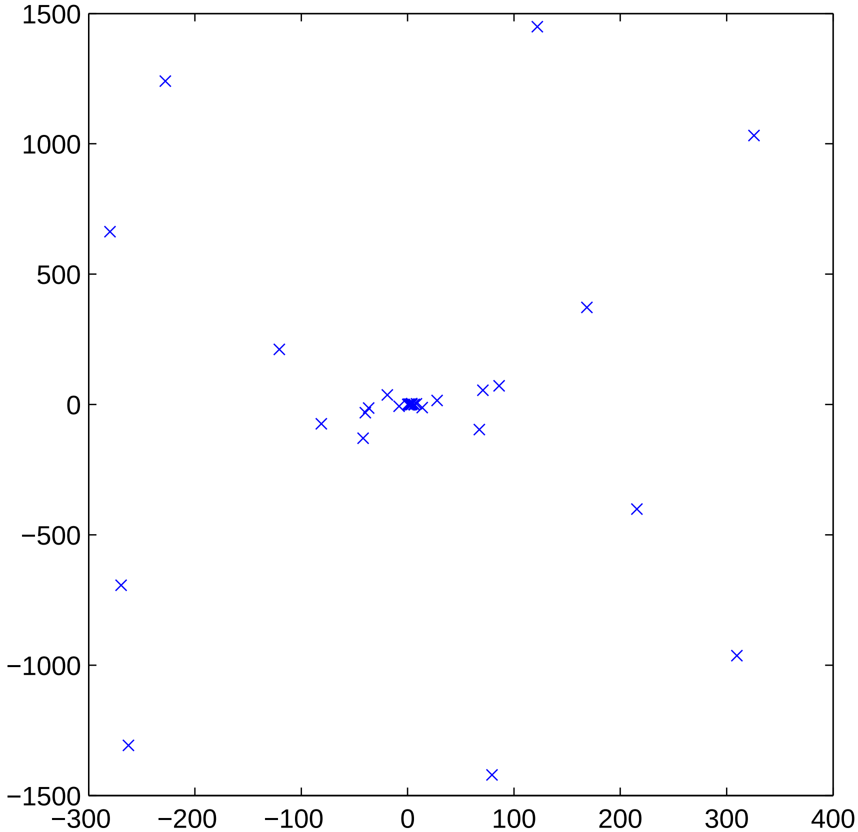

FIGURE 5. Eigenvalues calculated to first-order using a perturbation expansion with $N = 60$ and $t = 10^{-12}$.

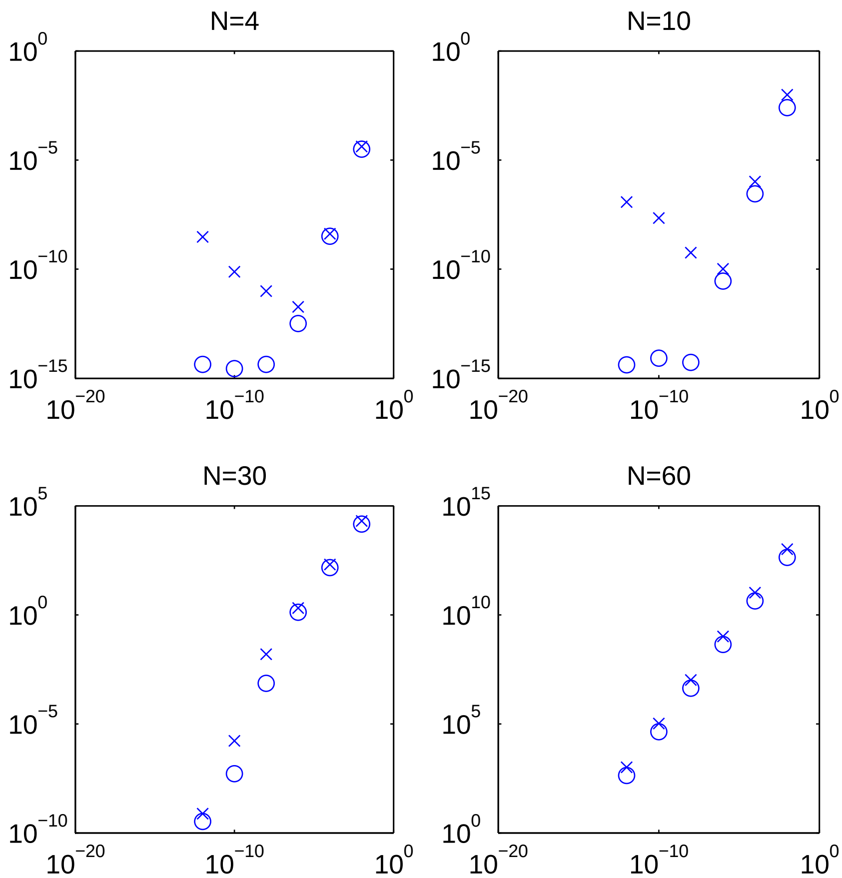

FIGURE 6. Error vs. t, where 'o' represents the norm of the difference between the eigenvalues computed to first-order using a perturbation expansion and the eigenvalues computed using Matlab's Condeig routine. 'x' represents the norm of the difference between the eigenvectors computed to first order using a perturbation expansion and the eigenvectors computed using Matlab's Condeig routine.

References

[1] Bau, David & Trefethen, Lloyd N. (1997). *Numerical Linear Algebra*. Philadelphia: SIAM.

[2] Embree, Mark & Trefethen, Lloyd N. (2005). *Spectra and Pseudospectra: The behavior of Nonnormal Matrices and Operators*. Princeton: Princeton University Press.

[3] Trefethen, Lloyd N. (1999). Computation of Pseudospectra. *Acta Numerica*, 1-49.

[4] Trefethen, Lloyd N., & Wright, Thomas G. (2001). Large-scale computation of pseudospectra using Arpack and Eigs. *SIAM Journal of Scientific Computing*, 23(2), 591-605.

[5] Wright, Thomas G., EigTool, 2002. Software available at **http://www.comlab.ox.ac.uk/pseudospectra/eigtool**.