EXPLICIT BOUNDS FOR THE PSEUDOSPECTRA OF VARIOUS CLASSES OF MATRICES AND OPERATORS

FEIXUE GONG[1], OLIVIA MEYERSON[2], JEREMY MEZA[3], ABIGAIL WARD[4]
MIHAI STOICIU[5] (ADVISOR)
SMALL 2014 - MATHEMATICAL PHYSICS GROUP

Abstract. We study the ε-pseudospectra $\sigma_\varepsilon(A)$ of square matrices $A \in \mathbb{C}^{N\times N}$. We give a complete characterization of the ε-pseudospectrum of any 2×2 matrix and describe the asymptotic behavior (as $\varepsilon \to 0$) of $\sigma_\varepsilon(A)$ for any square matrix A. We also present explicit upper and lower bounds for the ε-pseudospectra of bidiagonal and tridiagonal matrices, as well as for finite rank operators.

Contents

Date: October 30, 2014.
[1,2,5]Williams College [3]Carnegie Mellon University [4]University of Chicago.

1. Introduction

The pseudospectra of matrices and operators is an important mathematical object that has found applications in various areas of mathematics: linear algebra, functional analysis, numerical analysis, and differential equations. An overview of the main results on pseudospectra can be found in [14].

In this paper we investigate a few classes of matrices and operators and provide explicit bounds on their ε-pseudospectra. We study 2×2 matrices, bidiagonal and tridiagonal matrices, as well as finite rank operators. We also describe the asymptotic behavior of the ε-pseudospectrum $\sigma_\varepsilon(A)$ of any $n \times n$ matrix A.

The paper is organized as follows: in Section 2 we give the three standard equivalent definitions for the pseudospectrum and we present the "classical" results on ε-pseudospectra of normal and diagonalizable matrices (the Bauer-Fike theorems). Section 3 contains a detailed analysis of the ε-pseudospectrum of 2×2 matrices, including both the non-diagonalizable case (Subsection 3.1) and the diagonalizable case (Subsection 3.2). The asymptotic behavior (as $\varepsilon \to 0$) of the ε-pseudospectrum of any $n \times n$ matrix is described in Section 4, where we show (in Theorem 4.1) that, for any square matrix A, the ε-pseudospectrum converges, as $\varepsilon \to 0$ to a union of disks.

The main result of Section 4 is applied to several classes of matrices: Jordan blocks (Section 5), bidiagonal matrices (Section 6), and tridiagonal matrices (Section 7). In Section 5 we derive explicit lower bounds for $\sigma_\varepsilon(J)$, where J is any Jordan block, while Section 6 is dedicated to the analysis of arbitrary periodic bidiagonal matrices A. We derive explicit formulas (in terms the coefficients of A) for the asymptotic radii of the ε-pseudospectrum of A, as $\varepsilon \to 0$. We continue with a brief investigation of tridiagonal matrices in Section 7, while in the last section (Section 8) we consider finite rank operators and show that the ε-pseudospectrum of an operator of rank m is at most as big as $C\varepsilon^{\frac{1}{m}}$, as $\varepsilon \to 0$.

2. Pseudospectra

2.1. **Motivation and Definitions.** The concept of the spectrum of a matrix $A \in \mathbb{C}^{N \times N}$ provides a fundamental tool for understanding the behavior of A. As is well-known, a complex number $z \in \mathbb{C}$ is in the spectrum of A (denoted $\sigma(A)$) whenever $zI - A$ (which we will denote as $z - A$) is not invertible, i.e., the characteristic polynomial of A has z as a root. As slightly perturbing the coefficients of A will change the roots of the characteristic polynomial, the property of "membership in the set of eigenvalues" is not well-suited for many purposes, especially those in numerical analysis. We thus want to find a characterization of when a complex number is close to an eigenvalue, and we do this by considering the set of complex numbers z such that $\|(z-A)^{-1}\|$ is large, where the norm here is the usual operator norm induced by the Euclidean norm, i.e.

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

The motivation for considering this question comes from the following observation:

Proposition 2.1. *Let $A \in \mathbb{C}^{N\times N}$, λ an eigenvalue of A, and z_n be a sequence of complex numbers converging to λ with $z_n \neq z$ for all n. Then $\|(z_n - A)^{-1}\| \to \infty$ as $n \to \infty$.*

Proof. Let v be the eigenvector corresponding to λ with norm 1, so $Av = \lambda v$. Define

$$v_n = \frac{(z_n - A)v}{\|(z_n - A)v\|}$$

(note that $(z_n - A)v \neq 0$ by assumption). Now, note that

$$\begin{aligned}
\|(z_n - A)^{-1}\| &\geq \|(z_n - A)^{-1} v_n\| \\
&= \left\| \frac{(z_n - A)^{-1}(z_n - A)v}{\|(z_n - A)v\|} \right\| \\
&= \left\| \frac{v}{\|(z_n - A)v\|} \right\| \\
&= \frac{1}{|\lambda - z_n|},
\end{aligned}$$

and since the last quantity gets arbitrarily large for large n, we obtain the result. □

We call the operator $(z - A)^{-1}$ the *resolvent* of A. The observation that the norm of the resolvent is large when z is close to an eigenvalue of A leads us to the first definition of the ε-pseudospectrum of an operator.

Definition 2.1. Let $A \in \mathbb{C}^{N\times N}$, and let $\varepsilon > 0$. The ε-pseudospectrum of A is the set of $z \in \mathbb{C}$ such that

$$\|(z - A)^{-1}\| > 1/\varepsilon$$

Note that the boundary of the ε-pseudospectrum is exactly the $1/\varepsilon$ level curve of the function $z \mapsto \|(z - A)^{-1}\|$. Fig. 2.1 depicts the behavior of this function near the eigenvalues, as described in Prop. 2.1.

The resolvent norm has singularities in the complex plane, and as we approach these points, the resolvent norm grows to infinity. Conversely, if $\|(z - A)^{-1}\|$ approaches infinity, then z must approach some eigenvalue of A [14, Thm 2.4].

(It is also possible to develop a theory of pseudo-spectrum for operators on Banach spaces, and it is important to note that this converse does not necessarily hold for such operators; that is, there are operators [4, 5] such that $\|(z - A)^{-1}\|$ approaches infinity, but z does not approach the spectrum of A.)

The second definition of the ε-pseudospectrum arises from eigenvalue perturbation theory [8].

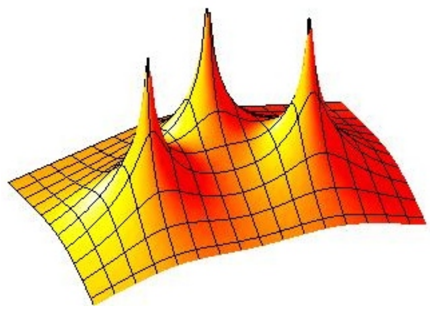

Figure 2.1. Contour Plot of Resolvent Norm

Definition 2.2. Let $A \in \mathbb{C}^{N\times N}$. The ε-pseudospectrum of A is the set of $z \in \mathbb{C}$ such that

$$z \in \sigma(A+E)$$

for some E with $\|E\| < \varepsilon$.

Finally, the third definition of the ε-pseudospectum is the following:

Definition 2.3. Let $A \in \mathbb{C}^{N\times N}$. The ε-pseudospectrum of A is the set of $z \in \mathbb{C}$ such that

$$\|(z-A)v\| < \varepsilon$$

for some unit vector v.

This definition is similar to our first definition in that it quantifies how close z is to an eigenvalue of A. In addition to this, it also gives us the notion of an *ε-pseudoeigenvector*.

Theorem 2.1 (Equivalence of the definitions of pseudospectra)**.** *For any matrix $A \in \mathbb{C}^{N\times N}$, the three definitions above are equivalent.*

The proof of this theorem follows [14, §2].

Proof. Define the following sets

$$M_\varepsilon = \{z \in \mathbb{C} \mid \|(z-A)^{-1}\| > \varepsilon^{-1}\}$$
$$N_\varepsilon = \{z \in \mathbb{C} \mid z \in \sigma(A+E) \text{ for some } E \in \mathbb{C}^{N\times N}, \|E\| < \varepsilon\}$$
$$P_\varepsilon = \{z \in \mathbb{C} \mid \|(z-A)v\| < \varepsilon \text{ for some } v \in \mathbb{C}^N, \|v\| = 1\}$$

To show equivalence, we will show that $M_\varepsilon \subseteq N_\varepsilon \subseteq P_\varepsilon \subseteq M_\varepsilon$.

<u>$N_\varepsilon \subseteq P_\varepsilon$</u>

Let $z \in N_\varepsilon$. Let E be such that $z \in \sigma(A+E)$, with $\|E\| < \varepsilon$. Let v be an eigenvector of norm one for z, that is $\|v\| = 1$ and $(A+E)v = zv$. Then

$$\|(z-A)v\| = \|(z-A-E)v + Ev\| \le \|E\|\|v\| < \varepsilon$$

and hence $z \in P_\varepsilon$.

$\underline{P_\varepsilon \subseteq M_\varepsilon}$

Let $z \in P_\varepsilon$. Let $v \in \mathbb{C}^N$ be such that $\|v\| = 1$ and $\|(z-A)v\| < \varepsilon$. Observe

$$(z-A)v = \|(z-A)v\| \frac{(z-A)v}{\|(z-A)v\|} = su$$

where we define $u = \frac{(z-A)v}{\|(z-A)v\|}$ and $s = \|(z-A)v\|$. Then, we see that $s^{-1}v = (z-A)^{-1}u$, which implies that

$$\|(z-A)^{-1}\| \geq \|(z-A)^{-1}u\| = |s^{-1}|\|v\| = s^{-1} > \varepsilon^{-1}$$

and hence $z \in M_\varepsilon$.

$\underline{M_\varepsilon \subseteq N_\varepsilon}$

Let $z \in M_\varepsilon$. We will construct an E with $\|E\| < \varepsilon$ and $z \in \sigma(A+E)$. Since $\|(z-A)^{-1}\| = \sup_{\|v\|=1} \|(z-A)^{-1}v\| > \varepsilon^{-1}$, there exists some v with $\|v\| = 1$ such that $\|(z-A)^{-1}v\| > \varepsilon^{-1}$. Observe

$$(z-A)^{-1}v = \|(z-A)^{-1}v\| \frac{(z-A)^{-1}v}{\|(z-A)^{-1}v\|} = \|(z-A)^{-1}v\|u$$

where u is the unit vector $\frac{(z-A)^{-1}v}{\|(z-A)^{-1}v\|}$. So $\frac{v}{\|(z-A)^{-1}v\|} = (z-A)u$, from which we have

$$zu = Au + \frac{v}{\|(z-A)^{-1}v\|}$$

Define $E = \frac{v}{\|(z-A)^{-1}v\|}u^*$. Note $Eu = \frac{v}{\|(z-A)^{-1}v\|}$ and $\|E\| < \varepsilon$. Then we have $zu = Au + Eu$, and so $z \in \sigma(A+E)$. □

As all three definitions are now shown to be equivalent, we can unambiguously denote the ε-pseudospectrum of A as $\sigma_\varepsilon(A)$.

Fig. 2.2 depicts an example of ε-pseudospectra for a specific matrix and for various ε. We see that the boundaries of ε-pseudospectra for a matrix are curves in the complex plane around the eigenvalues of the matrix. We are interested in understanding geometric and algebraic properties of these curves.

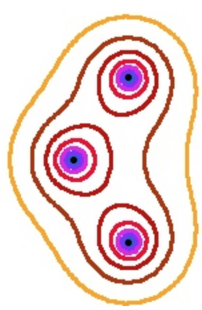

FIGURE 2.2. The curves bounding the ε-pseudospectra of an operator A, for different values of ε.

When a matrix A is the direct sum of smaller matrices, we can look at the pseudospectra of the smaller matrices to understand the ε-pseudospectrum of A. We get the following theorem from [14]:

Theorem 2.2.

$$\sigma_\varepsilon(A_1 \oplus A_2) = \sigma_\varepsilon(A_1) \cup \sigma_\varepsilon(A_2).$$

2.2. **Normal Matrices.** Recall that a matrix A is *normal* if $AA^* = A^*A$, or equivalently if A can be diagonalized with an orthonormal basis of eigenvectors.

The pseudospectra of these matrices is particularly well-behaved: Thm. 2.3 shows that the ε-pseudospectrum of a normal matrix is exactly a disk of radius ε around each eigenvalue, as in shown in Fig. 2.3. This is clear for diagonal matrices; it follows for normal matrices since as we shall see, the ε-pseudospectrum of a matrix is invariant under a unitary change of basis.

FIGURE 2.3. The ε-pseudospectrum of a normal matrix. Note that each boundary is a perfect disk around an eigenvalue.

Theorem 2.3. *Let $A \in \mathbb{C}^{N\times N}$. Then,*

$$\sigma(A) + B(0,\varepsilon) \subseteq \sigma_\varepsilon(A) \qquad \textit{for all } \varepsilon > 0 \tag{2.1}$$

and if A is normal and the operator norm is the 2-norm, then

$$\sigma_\varepsilon(A) = \sigma(A) + B(0,\varepsilon) \qquad \textit{for all } \varepsilon > 0 \tag{2.2}$$

Conversely, (2.2) *implies that A is normal.*

Proof. Let $z \in \sigma(A) + B(0,\varepsilon)$. Write $z = \lambda + w$ for some $\lambda \in \sigma(A)$ and $w \in \mathbb{C}$ with $|w| < \varepsilon$. Let v be an eigenvector for A with eigenvalue λ. Then, v is an eigenvector for the perturbed system $A + wI$ with eigenvalue z. Thus, $z \in \sigma(A + wI)$, and hence $z \in \sigma_\varepsilon(A)$. This shows (2.1).

To show (2.2), we now suppose A is a normal matrix. It suffices to show the backwards inclusion to (2.1). Write $A = U\Lambda U^*$ for a unitary U and diagonal Λ. We have

$$\|(z-A)^{-1}\| = \|(UzU^* - U\Lambda U^*)^{-1}\| = \|U(z-\Lambda)^{-1}U^*\| = \|(z-\Lambda)^{-1}\| = \frac{1}{\operatorname{dist}(z,\sigma(A))}$$

which implies that if $z \in \sigma_\varepsilon(A)$, then $\mathrm{dist}(z, \sigma(A)) < \varepsilon$. Rewriting this, we get $\sigma_\varepsilon(A) \subseteq \sigma(A) + B(0, \varepsilon)$.

The converse of this proof follows from the discussion below of the condition number and the full proof is postponed for now; see Corollary 2.1.

□

2.3. **Non-normal Diagonalizable Matrices.** Now suppose A is diagonalizable but not normal, i.e. we cannot diagonalize A by an isometry of $\mathbb{C}^N$. In this case we do not expect to get an exact characterization of the ε-pseudospectra as we did previously. The following example shows that there exist matrices with pseudospectra that are larger than the disk of radius ε:

Example 2.1. For the non-normal diagonalizable matrix $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$,

$$B(2, C\varepsilon) \subseteq \sigma_\varepsilon(A)$$

for any $C < \sqrt{2}$ as $\varepsilon \to 0$.

Proof. Let $z \in B(2, C\varepsilon)$, where $C < \sqrt{2}$. We will construct E with $\|E\| < \varepsilon$ so that z is an eigenvalue of $A + E$.

Let $E = \begin{pmatrix} 0 & 0 \\ -\alpha\varepsilon & \alpha\varepsilon \end{pmatrix}$ where $|\alpha| < \frac{1}{\sqrt{2}}$. Direct calculation gives $\|E\| = \sqrt{2}|\alpha|\varepsilon < \varepsilon$. We find the eigenvalues of $A + E$:

$$\det(A + E - zI) = (1 - z)(2 + \alpha\varepsilon - z) - \alpha\varepsilon = z^2 - z(3 + \alpha\varepsilon) + 2 = 0$$

The roots of this equation are given by

$$z = \frac{3 + \alpha\varepsilon \pm \sqrt{1 + 6\alpha\varepsilon + \alpha^2\varepsilon^2}}{2}.$$

Recall for x small, the Taylor expansion $(1 + x)^p = 1 + px + O(x^2)$. This gives that the root in which we add the discriminant has the expansion

$$z = \frac{3 + \alpha\varepsilon + (1 + 3\alpha\varepsilon + O(\varepsilon^2))}{2} \implies z = 2 + 2\alpha\varepsilon + O(\varepsilon^2)$$

Since we may take any α with $|\alpha| < \frac{1}{\sqrt{2}}$, we obtain the result.

□

We begin here the characterization of the behavior of non-normal, diagonalizable matrices. Write $A = VDV^{-1}$. Define the *condition number* of V as

$$\kappa(V) = \|V\|\|V^{-1}\| = \frac{s_{\max}(V)}{s_{\min}(V)}$$

where $s_{\max}(V)$ and $s_{\min}(V)$ are the maximum and minimum singular values of V, respectively. Note that there is some ambiguity when we define $\kappa(V)$, as V is not uniquely determined. If the eigenvalues are distinct, then $\kappa(V)$ becomes unique if the eigenvectors are normalized by $\|v_j\| = 1$.While this choice may not necessarily be the one that minimizes $\kappa(V)$, [13] shows that this choice exceeds the minimal value by at most a factor of $\sqrt{N}$.

Note that $\kappa(V)$ is necessarily greater than or equal to 1, and we claim $\kappa(V) = 1$ if and only if A is normal. The backward direction follows from our unitarily equivalent norm: if A is normal, then V is unitary, which implies that $\|V\| = \|V^{-1}\| = 1$ and hence $\kappa(V) = 1$. To show the forward direction, suppose $\kappa(V) = \frac{s_{\max}(V)}{s_{\min}(V)} = 1$. Thus, in the singular value decomposition of V, the singular values are all the same, which implies that $V = sUW^*$ for $s > 0$ and unitary matrices U, W. Define $V' = V/s$, and note that V' is a unitary matrix. Also note

$$V'DV'^{-1} = (V/s)D(sV^{-1}) = VDV^{-1} = A$$

and thus A is unitarily diagonalizable and hence normal.

The condition number gives an upper bound on the size of the pseudo spectrum of a non-normal diagonalizable matrix:

Theorem 2.4 (Bauer-Fike)**.** *Let $A \in \mathbb{C}^{N\times N}$ and let A be diagonalizable, $A = VDV^{-1}$. Then for each $\varepsilon > 0$,*

$$\sigma(A) + B(0,\varepsilon) \subseteq \sigma_\varepsilon(A) \subseteq \sigma(A) + B(0, \varepsilon\kappa(V))$$

where

$$\kappa(V) = \|V\|\|V^{-1}\| = \frac{s_{max}(V)}{s_{min}(V)}.$$

Corollary 2.1. *If $\sigma_\varepsilon(A) = \sigma(A) + B(0,\varepsilon)$, then A is normal.*

Proof (Thm. 2.4). The first inclusion was established in the previous theorem. For the second inclusion, let $z \in \sigma_\varepsilon(A)$. Then

$$(z - A)^{-1} = (z - VDV^{-1})^{-1} = [V(z - D)V^{-1}]^{-1} = V(z - D)^{-1}V^{-1}$$

which implies that

$$\varepsilon^{-1} < \|(z - A)^{-1}\| \le \kappa(V)\|(z - D)^{-1}\| = \frac{\kappa(V)}{\mathrm{dist}(z, \sigma(A))},$$

whence

$$\mathrm{dist}(z, \sigma(A)) < \varepsilon\kappa(V).$$

□

2.4. **Non-diagonalizable Matrices.** So far we have considered normal matrices, and more generally diagonalizable matrices. We now relax our constraint that our matrix be diagonalizable, and provide similar bounds on the pseudospectra. While not every matrix is diagonalizable, every matrix can be put in Jordan normal form. Below we give a brief review of the Jordan form.

Let $A \in \mathbb{C}^{N\times N}$ and suppose A has only one eigenvalue, λ with geometric multiplicity one. Writing A in Jordan form, there exists a matrix V such that $AV = VJ$, where J is a single Jordan block of size N. Write

$$V = \begin{pmatrix} v_1, & v_2, & \dots, & v_n \end{pmatrix}$$

Then,

$$AV = \begin{pmatrix} Av_1, & Av_2, & \dots, & Av_n \end{pmatrix} = \begin{pmatrix} \lambda v_1, & v_1 + \lambda v_2, & \dots, & v_{n-1} + \lambda v_n \end{pmatrix} = VJ,$$

where v_1 is the right eigenvector associated with λ and $v_2, ..., v_n$ are *generalized right eigenvectors*, that is right eigenvectors for $(A - \lambda I)^k$ for $k > 1$.

We also know that there exists a matrix $U^* = V^{-1}$ such that $U^* A = JU^*$.

$$U = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}.$$

Then,

$$U^* A = \begin{pmatrix} u_1 A \\ u_2 A \\ \vdots \\ u_n A \end{pmatrix} = \begin{pmatrix} \lambda u_1 + u_2 \\ \lambda u_2 + u_3 \\ \vdots \\ \lambda u_{n-1} + u_n \\ \lambda u_n \end{pmatrix} = JU^*.$$

Thus, u_n is the left eigenvector associated with λ and $u_1, ..., u_{n-1}$ are *generalized left eigenvectors.*

Similar to how $\kappa(V)$ quantifies the normality of a matrix, we can also quantify the normality of an eigenvalue:

Definition 2.4. For any simple eigenvalue λ_j of a matrix A, the *condition number of λ_j* is defined as

$$\kappa(\lambda_j) = \frac{\|u_j\| \|v_j\|}{|u_j^* v_j|},$$

where v_j and u_j^* are the right and left eigenvectors associated with λ_j, respectively.

Note: The Cauchy-Schwarz inequality implies that $|u_j^* v_j| \leq \|u_j\| \|v_j\|$, so $\kappa(\lambda_j) \geq 1$, with equality when u_j and v_j are collinear. An eigenvalue for which $\kappa(\lambda_j) = 1$ is called a normal eigenvalue; a matrix A with all eigenvalues simple is normal if and only if $\kappa(\lambda_j) = 1$ for all eigenvalues.

With this definition, we can find finer bounds for the pseudospectrum of a matrix; in particular, we can find bounds for the components of the pseduospectrum centered around each eigenvalue.

The following theorem can be found for example in [3].

Theorem 2.5 (Asymptotic pseudospectra inclusion regions)**.** *Suppose $A \in \mathbb{C}^{N \times N}$ has N distinct eigenvalues. Then, as $\varepsilon \to 0$,*

$$\sigma_\varepsilon(A) \subseteq \bigcup_{j=1}^{N} B\left(\lambda_j, \varepsilon\kappa(\lambda_j) + \mathcal{O}\left(\varepsilon^2\right)\right).$$

We can drop the $\mathcal{O}(\varepsilon^2)$ term, for which we get an increase in the radius of our inclusion disks by a factor of N [2, Thm. 4].

Theorem 2.6 (Bauer-Fike theorem based on $\kappa(\lambda_j)$)**.** *Suppose $A \in \mathbb{C}^{N\times N}$ has N distinct eigenvalues. Then $\forall \varepsilon > 0$,*

$$\sigma_\varepsilon(A) \subseteq \bigcup_{j=1}^{N} B\left(\lambda_j, \varepsilon N \kappa(\lambda_j)\right).$$

The following has a proof in [14].

Theorem 2.7 (Asymptotic formula for the resolvent norm)**.** *Let $\lambda_j \in \sigma(A)$ be an eigenvalue of index k_j. For any $z \in \sigma_\varepsilon(A)$, for small enough ε,*

$$|z - \lambda_j| \le (C_j \varepsilon)^{\frac{1}{k_j}},$$

where $C_j = \|V_j T_j^{k_j-1} U_j^*\|$ and $T = J - \lambda I$.

3. Pseudospectra of 2 × 2 Matrices

The following section presents a complete characterization of the ε-pseudospectrum of any 2×2 matrix. We classify 2×2 matrices by whether they are diagonalizable or non-diagonalizable and then determine the ε-pseudospectra for each class. We begin with an explicit formula for computing the norm of a 2×2 matrix, and then analyze a 2×2 Jordan block.

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, with $a, b, c, d \in \mathbb{C}$. Let $s_{\max}$ denote the largest singular value of A.

Then,

$$\begin{aligned} \|A\|^2 &= s_{\max} \\ &= \text{ largest eigenvalue of } A^*A \\ &= \frac{\operatorname{Tr}(A^*A) + \sqrt{\operatorname{Tr}(A^*A)^2 - 4\det(A^*A)}}{2}. \end{aligned}$$

Letting $\rho = |a|^2 + |b|^2 + |c|^2 + |d|^2$, we can expand the norm as

$$\|A\|^2 = \frac{1}{2}\left(\sqrt{\rho^2 - 4\left(|a|^2|d|^2 + |b|^2|c|^2 - (ad\overline{bc} + \overline{ad}bc)\right)} + \rho\right).$$

This formula for the norm of the matrix allows us to compute the ε-pseudospectra of 2×2 matrices.

3.1. Non-diagonalizable 2×2 Matrices.

Proposition 3.1. *Let A be any non-diagonalizable 2×2 matrix, and let λ denote the eigenvalue of A. Write $A = VJV^{-1}$ where*

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad a, b, c, d \in \mathbb{C}, \qquad \text{and} \qquad J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}.$$

Then $\sigma_\varepsilon(A)$ is exactly a disk. More specifically, given any ε,

$$\sigma_\varepsilon(A) = B(\lambda, |k|)$$

where

$$|k| = \sqrt{C\varepsilon + \varepsilon^2} \qquad \textit{and} \qquad C = \frac{|a|^2 + |c|^2}{|ad - bc|}. \tag{3.1}$$

Proof. Let $z = \lambda + k$ where $k \in \mathbb{C}$. Then,

$$(z - A)^{-1} = (z - VJV^{-1})^{-1} = V(z - J)^{-1}V^{-1}.$$

Let $k = z - \lambda$. We obtain that

$$\|(z - A)^{-1}\| = \frac{1}{|k^2(ad - bc)|} \left\| \begin{pmatrix} adk - ac - bck & a^2 \\ -c^2 & -bck + ac + adk \end{pmatrix} \right\|.$$

Now, let

$$M = \begin{pmatrix} adk - ac - bck & a^2 \\ -c^2 & -bck + ac + adk \end{pmatrix}.$$

A simple calculation gives

$$\begin{aligned} \operatorname{Tr}(M^*M) &= (|a|^2 + |c|^2)^2 + 2|k|^2|bc - ad|^2 \\ \det(M^*M) &= |k|^4|bc - ad|^4 \end{aligned}$$

From the formula for the norm, we obtain that

$$\begin{aligned} \|(z - A)^{-1}\| &= \|V(z - J)^{-1}V^{-1}\| \\ &= \frac{\sqrt{\operatorname{Tr}(M^*M) + \sqrt{\operatorname{Tr}(M^*M)^2 - 4\det(M^*M)}}}{|k|^2|ad - bc|\sqrt{2}}. \end{aligned}$$

Note that this function depends only on $|k| = |z - \lambda|$; thus for any ε, $\sigma_\varepsilon(A)$ will be a disk. Solving for k in the above equation to find the curve bounding the pseudospectrum, we obtain

$$|k| = \sqrt{\varepsilon \frac{|a|^2 + |c|^2}{|ad - bc|} + \varepsilon^2}.$$

□

Example 3.1. Perhaps the simplest example of such a matrix is the 2×2 Jordan block.

Let $J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ where $\lambda \in \mathbb{C}$. In Jordan form, J decomposes as

$$J = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Plugging into Equation 3.1 with $a = 1, b = 0, c = 0, d = 1$, we get that

$$|k| = \frac{\sqrt{\varepsilon^2(1 - 0)(1 - 0) + \varepsilon|1 - 0|(1 + 0)}}{|1 - 0|} = \sqrt{\varepsilon + \varepsilon^2}.$$

Hence,

$$\sigma_\varepsilon(J) = B\left(\lambda, \sqrt{\varepsilon + \varepsilon^2}\right).$$

We can check this result by direct application of the first definition of pseudospectra.

Let $z \in \sigma_\varepsilon(J)$. Note

$$(z - J)^{-1} = \begin{pmatrix} z-\lambda & -1 \\ 0 & z-\lambda \end{pmatrix}^{-1} = \frac{1}{(z-\lambda)^2}\begin{pmatrix} z-\lambda & 1 \\ 0 & z-\lambda \end{pmatrix}.$$

Using our equation for the norm of a 2×2 matrix, we get

$$\frac{1}{\varepsilon} < \|(z-J)^{-1}\| = \frac{1}{|z-\lambda|^2}\sqrt{\frac{1}{2}\left(2|z-\lambda|^2 + 1 + \sqrt{4|z-\lambda|^2+1}\right)}.$$

Let $w = |z-\lambda|$, and square both sides to get $\frac{2w^4}{\varepsilon^2} < 2w^2+1+\sqrt{4w^2+1}$. This simplifies to $w < \sqrt{\varepsilon+\varepsilon^2}$. Thus, $z \in \sigma_\varepsilon(J) \iff |z-\lambda| < \sqrt{\varepsilon+\varepsilon^2} \iff z \in B\left(\lambda, \sqrt{\varepsilon+\varepsilon^2}\right)$.

3.2. **Diagonalizable 2×2 Matrices.**

Proposition 3.2. *Let A be any diagonalizable 2×2 matrix and let λ_1, λ_2 be the eigenvalues of A and v_1, v_2 be the eigenvectors associated with the eigenvalues. Then $\sigma_\varepsilon(A)$ is the set of points z that satisfy the equation*

$$(\varepsilon^2 - |z-\lambda_1|^2)(\varepsilon^2 - |z-\lambda_2|^2) - \varepsilon^2|\lambda_1-\lambda_2|^2\cot^2(\theta) < 0,$$

where θ is the angle between the two eigenvectors.

Proof. Since A is diagonalizable, we know that $A = VDV^{-1}$ where

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Note that V is an invertible matrix with columns corresponding to the eigenvectors of A.

Without loss of generality, we can write $z = \lambda_1 + k$.

Let $\gamma = \lambda_1 - \lambda_2$ and $r = ad - bc$. Then,

$$\begin{aligned}(z-A)^{-1} &= (z - VDV^{-1})^{-1} = V(z-D)^{-1}V^{-1} \\ &= \frac{1}{r}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \frac{1}{k} & 0 \\ 0 & \frac{1}{\gamma+k} \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ &= \left(\frac{1}{rk(\gamma+k)}\right)\begin{pmatrix} ad\gamma + rk & -ab\gamma \\ cd\gamma & -bc\gamma + rk \end{pmatrix}.\end{aligned}$$

Let $M = \begin{pmatrix} ad\gamma + rk & -ab\gamma \\ cd\gamma & -bc\gamma + rk \end{pmatrix}$.

Then, $\|(z-A)^{-1}\| = \frac{1}{|rk(\gamma+k)|}\|M\|$.

Calculating $\mathrm{Tr}(M^*M)$ and $\det(M^*M)$ gives

$$\begin{aligned}\mathrm{Tr}(M^*M) &= |cd\gamma|^2 + |kr + ad\gamma|^2 + |ab\gamma|^2 + |kr - bc\gamma|^2 \\ &= 2|kr|^2 + 2|r|^2 \,\mathrm{Re}[k\overline{\gamma}] + |\gamma|^2(|a|^2+|c|^2)(|b|^2+|d|^2) \\ &= |r|^2\left(2|k|^2 + 2\,\mathrm{Re}[k\overline{\gamma}] + |\gamma|^2 + |\gamma|^2\left(\frac{(|a|^2+|c|^2)(|b|^2+|d|^2)}{|r|^2} - 1\right)\right).\end{aligned}$$

Note that the first part of this expression simplifies to

$$2|k|^2 + 2\,\mathrm{Re}[k\overline{\gamma}] + |\gamma|^2 = |k+\gamma|^2 + |k|^2.$$

For the second part, we have

$$\begin{aligned}\frac{(|a|^2+|c|^2)(|b|^2+|d|^2)}{|r|^2} - 1 &= \frac{1}{|r|^2}\left((|a|^2+|c|^2)(|b|^2+|d|^2) - |r|^2\right) \\ &= \frac{1}{|r|^2}\left(|ab|^2 + |cd|^2 + bc\overline{ad} + \overline{bc}ad\right).\end{aligned}$$

Now, note that our eigenvectors v_1, v_2 are exactly the columns of V. So, we have

$$|v_1 \cdot v_2|^2 = |\bar{a}b + \bar{c}d|^2 = |ab|^2 + |cd|^2 + \overline{ad}bc + ad\overline{bc}.$$

Thus, the second term in our trace simplifies to

$$\frac{|v_1\cdot v_2|^2}{|r|^2} = \frac{|v_1|^2|v_2|^2\cos^2\theta}{|v_1|^2|v_2|^2\sin^2\theta} = \cot^2\theta,$$

where the denominator comes from the fact that $r = \det V$ is the area of the parallelogram spanned by v_1 and v_2. Thus, we get that

$$\mathrm{Tr}(M^*M) = |r|^2\left(|\gamma+k|^2 + |k|^2 + |\gamma|^2\cot^2\theta\right).$$

For the determinant, we have

$$\begin{aligned}\det(M^*M) &= (|cd\gamma|^2 + |kr+ad\gamma|^2)(|ab\gamma|^2 + |kr - bc\gamma|^2) \\ &\quad - ((kr-bc\gamma)\overline{cd\gamma} - ab\gamma(\overline{kr+ad\gamma}))(-(kr+ad\gamma)\overline{ab\gamma} + cd\gamma(\overline{kr-bc\gamma})) \\ &= (k^2r^2 + kr^2\gamma)(\overline{k^2r^2 + kr^2\gamma}) \\ &= |r|^4|k|^2|k+\gamma|^2\end{aligned}$$

We then plug these in to the equation for the norm of a 2×2 matrix and get

$$\begin{aligned}&\varepsilon^{-1} < \|(z-A)^{-1}\| = \sqrt{\frac{\mathrm{Tr}(M^*M) + \sqrt{\mathrm{Tr}(M^*M)^2 - 4\det(M^*M)}}{2|r|^2|k(\gamma+k)|^2}} \\ \iff &\frac{2|k(\gamma+k)|^2}{\varepsilon^2} < |\gamma+k|^2 + |k|^2 + |\gamma|^2\cot^2\theta + \sqrt{(|\gamma+k|^2+|k|^2+|\gamma|^2\cot^2\theta)^2 - 4|k(\gamma+k)|^2} \\ \iff &\left(\frac{2|k(\gamma+k)|^2}{\varepsilon^2} - (|\gamma+k|^2+|k|^2+|\gamma|^2\cot^2\theta)\right)^2 < (|\gamma+k|^2+|k|^2+|\gamma|^2\cot^2\theta)^2 - 4|k(\gamma+k)|^2 \\ \iff &\frac{4|k(\gamma+k)|^4}{\varepsilon^4} - \frac{4|k(\gamma+k)|^2}{\varepsilon^2}(|\gamma+k|^2+|k|^2+|\gamma|^2\cot^2\theta) + 4|k(\gamma+k)|^2 < 0 \\ \iff &4|k(\gamma+k)|^2 + \varepsilon^4 - (|\gamma+k|^2+|k|^2+|\gamma|^2\cot^2\theta)\varepsilon^2 < 0 \\ \iff &(\varepsilon^2 - |k|^2)(\varepsilon^2 - |k+\gamma|^2) - \varepsilon^2|\gamma|^2\cot^2\theta < 0.\end{aligned}$$

□

Note that for normal matrices, the eigenvectors are orthogonal, so $\theta = \pi/2$ and $\cot\theta = 0$. Therefore the equation above reduces to

$$(\varepsilon^2 - |k|^2)(\varepsilon^2 - |k+\gamma|^2) < 0$$

which describes two disks of radius ε centered around λ_1, λ_2, as we expect.

When the matrix only has one eigenvalue and is still diagonalizable (i.e. when it is a multiple of the identity), then we get

$$(\varepsilon^2 - |k|^2)(\varepsilon^2 - |k+0|^2) < 0,$$

which is a disk of radius ε centered around the eigenvalue.

One consequence to note of Proposition 3.2 is that the shape of $\sigma_\varepsilon(A)$ is dependent on both the eigenvalues *and* the eigenvectors of A. Another less obvious consequence is that the pseudospectrum of a 2×2 matrix approaches a union of disks as ε tends to 0. This is proven in the following proposition.

Proposition 3.3. *Let A be a diagonalizable 2×2 matrix. Then, $\sigma_\varepsilon(A)$ asymptotically tends toward a disk. In particular, $\frac{r_{max}}{r_{min}} = 1 + \mathcal{O}(\varepsilon)$, where r_{max}, r_{min} are the maximum and minimum distances from an eigenvalue to the boundary of the closer connected component of $\partial\sigma_\varepsilon(A)$. Moreover, for A diagonalizable but not normal, $\sigma_\varepsilon(A)$ is never a perfect disk.*

Proof. Let ε be small enough so that the ε-pseudospectrum is disconnected. Without loss of generality, we will show that the component around λ_1 tends towards a disk by showing that the ratio of the maximum distance from z to λ_1 to the minimum distance from z to λ_1 is $1 + o(\varepsilon)$.

Let $z_{\max} \in \partial\sigma_\varepsilon(A)$ such that $|z_{\max} - \lambda_1|$ is a maximum. Consider the line joining λ_1 and λ_2. Suppose for contradiction that $z_{\max}$ did not lie on this line. Then, rotate $z_{\max}$ in the direction of λ_2 so that it is on this line, and call this new point z'. Note that $|z' - \lambda_2| < |z_{\max} - \lambda_2|$, but $|z' - \lambda_1| = |z_{\max} - \lambda_1|$. As such, we get that

$$(|z'-\lambda_1|^2-\varepsilon^2)(|z'-\lambda_2|^2-\varepsilon^2) < (|z_{\max}-\lambda_1|^2-\varepsilon^2)(|z_{\max}-\lambda_2|^2-\varepsilon^2) = \varepsilon^2|\lambda_1-\lambda_2|^2\cot^2\theta$$

Thus, from Proposition 3.2 $z' \in \sigma_\varepsilon(A)$ but z' is not on the boundary of $\sigma_\varepsilon(A)$. Starting from z' and traversing the line joining λ_1 and λ_2, we can find $z'' \in \partial\sigma_\varepsilon(A)$ such that $|z'' - \lambda_1| > |z' - \lambda_1| = |z_{\max} - \lambda_1|$. This contradicts our choice of $z_{\max}$ and so $z_{\max}$ must be on the line joining λ_1 and λ_2. A similar argument shows that $z_{\min}$ must also be on the line joining λ_1 and λ_2, where $z_{\min} \in \partial\sigma_\varepsilon(A)$ such that $|z_{\min} - \lambda_1|$ is a minimum.

Since $z_{\max}$ is on the line joining λ_1 and λ_2, we have the exact equality $|z_{\max} - \lambda_2| = |z_{\max} - \lambda_1| + |\lambda_2 - \lambda_1|$. Let $r_{\max} = |z_{\max} - \lambda_1|$ and let $y = |\lambda_2 - \lambda_1|$. The equation describing $r_{\max}$ becomes

$$\left(r_{\max}^2 - \varepsilon^2\right)\left((y - r_{\max})^2 - \varepsilon^2\right) = \varepsilon^2 y^2 \cot^2\theta$$

Solving this equation for $r_{\max}$, we get

$$r_{\max} = \frac{1}{2}\left(y - \sqrt{y^2 + 4\varepsilon^2 - 4y\varepsilon\csc\theta}\right)$$

Similarly, we get

$$r_{\min} = \frac{1}{2}\left(\sqrt{y^2 + 4\varepsilon^2 + 4y\varepsilon\csc\theta} - y\right)$$

For ε small, we can use the approximation $(1+\varepsilon)^p = 1 + p\varepsilon + \frac{p(p-1)}{2}\varepsilon^2 + o(\varepsilon^3)$. Using this, we see

$$\begin{aligned}
\frac{r_{\max}}{r_{\min}} &= \frac{1 - \sqrt{1 + 4(\varepsilon/y)^2 - 4(\varepsilon/y)\csc\theta}}{\sqrt{1 + 4(\varepsilon/y)^2 + 4(\varepsilon/y)\csc\theta} - 1} \\
&= \frac{2(\varepsilon/y)\csc\theta - 2(\varepsilon/y)^2 + 2(\varepsilon/y)^2\csc^2\theta + \mathcal{O}(\varepsilon^3)}{2(\varepsilon/y)\csc\theta + 2(\varepsilon/y)^2 - 2(\varepsilon/y)^2\csc^2\theta + \mathcal{O}(\varepsilon^3)} \\
&= \frac{\csc\theta + (\varepsilon/y)(\csc^2\theta - 1) + \mathcal{O}(\varepsilon^2)}{\csc\theta - (\varepsilon/y)(\csc^2\theta - 1) + \mathcal{O}(\varepsilon^2)} \\
&= \frac{1 + \eta\varepsilon + \mathcal{O}(\varepsilon^2)}{1 - \eta\varepsilon + \mathcal{O}(\varepsilon^2)}
\end{aligned}$$

where $\eta = \cos\theta\cot\theta$. Recall the geometric series approximation $\frac{1}{1-x} = 1 + x + \mathcal{O}(x^2)$. Using this, we get our desired result

$$\begin{aligned}
\frac{r_{\max}}{r_{\min}} &= (1 + \eta\varepsilon + \mathcal{O}(\varepsilon^2))(1 + \eta\varepsilon + \mathcal{O}(\varepsilon^2)) \\
&= 1 + 2\eta\varepsilon + \mathcal{O}(\varepsilon^2) \\
&= 1 + (2\cos\theta\cot\theta)\varepsilon + \mathcal{O}(\varepsilon^2)
\end{aligned}$$

Thus, as $\varepsilon \to 0$, the farthest point on the boundary of $\sigma_\varepsilon(A)$ from λ_1 tends towards the closest point on the boundary of $\sigma_\varepsilon(A)$ from λ_1. Hence, $\sigma_\varepsilon(A)$ tends towards a disk.

Moreover, if A is diagonalizable but not normal, then the eigenvectors are linearly independent but not orthogonal, so θ is not a multiple of $\pi/2$ or π, and hence $\cos\theta\cot\theta \neq 0$ and $\frac{r_{\max}}{r_{\min}} \neq 1$.

□

4. Asymptotic Union of Disks Theorem

In the previous section, we showed that the ε-pseudospectrum for all 2×2 matrices are disks or asymptotically converge to a union of disks. We now explore whether this behavior holds in the general case. It is possible to find matrices whose $\varepsilon-$pseudospectra exhibit pathological properties for large ε; for example, the following non-diagonalizable matrix has, for larger ε, an ε-pseudospectrum that is not convex and not simply connected.

Thus, pseudospectra may behave poorly for large enough ε; however, in the limit as $\varepsilon \to 0$, these properties disappear and the pseudospectra behave as disks centered around the eigenvalues with well-understood radii. The following theorem allows us to understand this asymptotic behavior.

We will use the following set-up (which follows [10]) for our theorem.

$$\begin{pmatrix} -1 & -10 & -100 & -1000 & -10000 \\ 0 & -1 & -10 & -100 & -1000 \\ 0 & 0 & -1 & -10 & -100 \\ 0 & 0 & 0 & -1 & -10 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

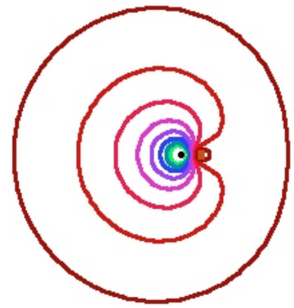

FIGURE 4.1. Pseudospectra of a Toeplitz matrix

For $A \in \mathbb{C}^{N\times N}$, find matrices that Jordanize A. That is,

$$\begin{pmatrix} J & \\ & \hat{J} \end{pmatrix} = \begin{pmatrix} Q \\ \hat{Q} \end{pmatrix} A \begin{pmatrix} P & \hat{P} \end{pmatrix}, \qquad \begin{pmatrix} Q \\ \hat{Q} \end{pmatrix} \begin{pmatrix} P & \hat{P} \end{pmatrix} = I$$

and J consists of Jordan blocks $J_1, \dots, J_m$ corresponding to an eigenvalue λ. $\hat{J}$ is the part of the Jordan form containing the other eigenvalues of A.

Let n be the size of the largest Jordan block corresponding to λ, and suppose there are ℓ Jordan blocks of size $n \times n$. Arrange $J_1, \dots, J_m$ in J so that

$$\dim(J_1) = \cdots = \dim(J_\ell) > \dim(J_{\ell+1}) \geq \cdots \geq \dim(J_m)$$

where $J_1, \dots, J_\ell$ are $n \times n$.

Further partition

$$P = \begin{pmatrix} P_1, & \dots, & P_\ell, & \dots, & P_m \end{pmatrix}$$

in a way that agrees with the above partition of J, so that the first column, x_j, of each P_j is a right eigenvector of A associated with λ. We also partition Q likewise

$$Q = \begin{pmatrix} Q_1 \\ \vdots \\ Q_\ell \\ \vdots \\ Q_m \end{pmatrix}.$$

The last row, y_j, of each Q_j is a left eigenvector of A corresponding to λ.
We now build the matrices

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_\ell \end{pmatrix}, \qquad X = \begin{pmatrix} x_1, & x_2, & \dots, & x_\ell \end{pmatrix},$$

where X and Y are the matrices of right and left eigenvectors, respectively, corresponding to the Jordan blocks of maximal size for λ.

With these matrices defined, we can understand the radii of the ε-pseudospectrum of A as $\varepsilon \to 0$.

Theorem 4.1. *Let $A \in \mathbb{C}^{N\times N}$. Let $\varepsilon > 0$. Then given $\lambda \in \sigma(A)$, for ε small enough, there exists a connected component $U \subseteq \sigma_\varepsilon(A)$ such that $U \cap \sigma(A) = \lambda$; denote this component of the ε-pseudospectrum $\sigma_\varepsilon(A) \restriction_\lambda$.*

Then, as $\varepsilon \to 0$,

$$B(\lambda, (C\varepsilon)^{1/n} + o(\varepsilon^{1/n})) \subseteq \sigma_\varepsilon(A) \restriction_\lambda \subseteq B(\lambda, (C\varepsilon)^{1/n} + o(\varepsilon^{1/n}))$$

where $C = \|XY\|$, with X, Y defined above.

Proof.

Lower Bound: We begin by proving the first containment,

$$B(\lambda, (C\varepsilon)^{1/n} + o(\varepsilon^{1/n})) \subseteq \sigma_\varepsilon(A) \restriction_\lambda .$$

Let $E \in \mathbb{C}^{N\times N}$. From [10, Theorem 2.1], we know there is an eigenvalue of the perturbed matrix $A + \varepsilon E$ admitting a first order expansion

$$\lambda(\varepsilon) = \lambda + (\gamma\varepsilon)^{1/n} + o(\varepsilon^{1/n}),$$

where γ is any eigenvalue of YEX. [10] later shows in Theorem 4.2 that

$$\alpha := \max_{\|E\|\le 1} \gamma_{\max} = \|XY\|,$$

where $\gamma_{\max}$ is the largest eigenvalue of YEX. In fact, the E that maximizes γ is given by $E = vu$ where v and u are the right and left singular vectors of the largest singular value of XY, normalized so $\|v\| = \|u\| = 1$.

We claim that $B(\lambda, (|\alpha|\varepsilon)^{1/n} + o(\varepsilon^{1/n})) \subseteq \sigma_\varepsilon(A)$.

Fix $\theta \in [0, 2\pi]$, and define $\tilde{E} = e^{i\theta}E$. Considering the perturbed matrix $A + \varepsilon\tilde{E}$, we can write

$$\lambda(\varepsilon) = \lambda + (\gamma\varepsilon)^{1/n} + o(\varepsilon^{1/n})$$

where γ is now any eigenvalue of $Y\tilde{E}X$. Solving for γ, we obtain:

$$0 = \det(e^{i\theta}YEX - \gamma I) = \det\left(e^{i\theta}(YEX - e^{-i\theta}\gamma I)\right),$$

which implies $0 = \det(YEX - e^{-i\theta}\gamma I)$, and thus $e^{-i\theta}\gamma$ is any eigenvalue of YEX. Since α is an eigenvalue of YEX, we can take $\gamma = e^{i\theta}\alpha$. Thus, we obtain

$$\lambda + (e^{i\theta}\alpha\varepsilon)^{1/n} + o(\varepsilon^{1/n}) \in \sigma_\varepsilon(A).$$

Ranging θ from 0 to $2n\pi$, we get

$$B(\lambda, (|\alpha|\varepsilon)^{1/n} + o(\varepsilon^{1/n})) \subseteq \sigma_\varepsilon(A).$$

Upper Bound: We now prove the second containment,

$$\sigma_\varepsilon(A) \restriction_\lambda \subseteq B(\lambda, (C\varepsilon)^{1/n} + o(\varepsilon^{1/n})).$$

From [14, §52, Theorem 52.3], we have that asymptotically

$$\sigma_\varepsilon(A) \restriction_\lambda \subseteq B(\lambda, (\beta\varepsilon)^{1/n} + o(\varepsilon^{1/n})),$$

where $\beta = \|PD^{n-1}Q\|$ and $J = \lambda I + D$. We claim $\beta = \|XY\| = \alpha$.

Note that $D^{n-1} = \text{diag}[\Gamma_1, \dots \Gamma_\ell, 0]$ where Γ_k is a $n \times n$ matrix with a 1 in the top right entry and zeros elsewhere. We find

$$PD^{n-1} = \begin{pmatrix} & & & & & \\ & x_1 & x_2 & \cdots & x_\ell & \\ & & & & & \end{pmatrix}.$$

This then gives

$$PD^{n-1}Q = \begin{pmatrix} XY & 0 \\ 0 & 0 \end{pmatrix}.$$

Thus $\beta = \|PD^{n-1}Q\| = \|XY\| = \alpha$.

Combining this with our lower bound result, we see that as $\varepsilon \to 0$,

$$\sigma_\varepsilon(A) \restriction_\lambda = B(\lambda, (\alpha\varepsilon)^{1/n} + o(\varepsilon^{1/n})).$$

□

We present special cases of matrices to explore the consequences of Theorem 4.1.

Special Cases:

(1) λ is simple. That is, λ has algebraic multiplicity 1.
Then, $n = 1$ and X and Y become the right and left eigenvectors x and y^* for λ, respectively. Hence, $C = \|XY\| = \|xy^*\| = \|x\|\|y\| = \kappa(\lambda)$ where we normalize so that $|y^*x| = 1$. Then, Theorem 4.1 becomes

$$\sigma_\varepsilon(A) \restriction_\lambda \approx B(\lambda, \kappa(\lambda)\varepsilon)$$

which matches with Theorem 2.5.

(2) λ is semisimple. That is, the algebraic multiplicity of λ equals the geometric multiplicity.
Then, $n = 1$, and we obtain $X = P$ and $Y = Q$, and C becomes the norm of the spectral projector for λ; see [14].

(3) λ has geometric multiplicity 1
In this case we obtain the same result for when λ is simple, except n may not equal 1. In other words,

$$\sigma_\varepsilon(A) \restriction_\lambda \approx B(\lambda, (\kappa(\lambda)\varepsilon)^{1/n}).$$

(4) $A \in \mathbb{C}^{2\times 2}$.
There are two cases, as in Section 3: Pseudospectra of 2×2 Matrices.

First, assume A is non-diagonalizable. In this case, A only has one eigenvalue, λ. Let

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad V^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

where $A = VJV^{-1}$.

Then, we have that

$$X = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad Y = \frac{1}{ad - bc}\begin{pmatrix} -c & a \end{pmatrix}.$$

From Theorem 4.1, we then have that as $\varepsilon \to 0$,

$$\sigma_\varepsilon(A) \subseteq B\left(\lambda, \left(\frac{|a|^2 + |c|^2}{|ad - bc|}\varepsilon\right)^{1/2} + o\left(\varepsilon^{1/2}\right)\right).$$

Equation 3.1 gives the explicit formula for $\sigma_\varepsilon(A)$ as $\varepsilon \to 0$, namely

$$\begin{aligned}\sigma_\varepsilon(A) &= B\left(\lambda, \left(\frac{|a|^2 + |c|^2}{|ad - bc|}\varepsilon + \varepsilon^2\right)^{1/2}\right) \\ &= B\left(\lambda, \left(\frac{|a|^2 + |c|^2}{|ad - bc|}\varepsilon\right)^{1/2} + \mathcal{O}(\varepsilon^{3/2})\right),\end{aligned}$$

which agrees with the above theorem.

In the case where A is diagonalizable, A has two eigenvalues, λ_1 and λ_2. Without loss of generality, we will look at $\sigma_\varepsilon(A)$ around λ_1.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\left[\frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right].$$

From this, we have

$$X = \begin{pmatrix} a \\ c \end{pmatrix}, \qquad Y = \frac{1}{ad - bc}\begin{pmatrix} d & -b \end{pmatrix}.$$

$$\|XY\| = \frac{\left(|a|^2 + |c|^2\right)\left(|b|^2 + |d|^2\right)}{|ad - bc|} = \csc\theta.$$

Thus, as $\varepsilon \to 0$, we have from Theorem 4.1:

$$B\left(\lambda, (\csc\theta)\,\varepsilon + o\left(\varepsilon\right)\right) \subseteq \sigma_\varepsilon(A) \subseteq B\left(\lambda, (\csc\theta)\,\varepsilon + o\left(\varepsilon\right)\right).$$

So,

$$\frac{r_{\max}}{r_{\min}} = \frac{(\csc\theta)\,\varepsilon + o\left(\varepsilon\right)}{(\csc\theta)\,\varepsilon + o\left(\varepsilon\right)}.$$

Note as $\varepsilon \to 0$, that this tends to 1 faster than ε.

This agrees with the ratio we obtain from the explicit formula for diagonalizable 2×2 matrices. However, the formula for the 2×2 matrix gives us slightly more information on the $o(\varepsilon)$ term.

$$r_{\max} = \frac{1}{2}\left(y - \sqrt{y^2 + 4\varepsilon^2 - 4y\varepsilon \csc\theta}\right)$$
$$r_{\min} = \frac{1}{2}\left(\sqrt{y + 4\varepsilon^2 + 4y\varepsilon \csc\theta} - y\right)$$

$$\begin{aligned}\frac{r_{\max}}{r_{\min}} &= \frac{1 - \sqrt{1 + 4(\varepsilon/y)^2 - 4(\varepsilon/y)\csc\theta}}{\sqrt{1 + 4(\varepsilon/y)^2 + 4(\varepsilon/y)\csc\theta} - 1} \\ &= \frac{2(\varepsilon/y)\csc\theta - 2(\varepsilon/y)^2 + 2(\varepsilon/y)^2\csc^2\theta + o(\varepsilon^3)}{2(\varepsilon/y)\csc\theta + 2(\varepsilon/y)^2 - 2(\varepsilon/y)^2\csc^2\theta + o(\varepsilon^3)} \\ &= \frac{\csc\theta + (\varepsilon/y)[\csc^2\theta - 1] + o(\varepsilon^2)}{\csc\theta - (\varepsilon/y)[\csc^2\theta - 1] + o(\varepsilon^2)}.\end{aligned}$$

5. Pseudospectra of $N \times N$ Jordan Block

As the Jordan blocks are the fundamental structural elements of finite dimensional linear operators, it is especially important to understand their pseudospectra.

Proposition 5.1 (Lower Bound)**.** *Let J be an $N \times N$ Jordan block where*

$$J = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}.$$

Then,

$$B(\lambda, \sqrt[N]{\varepsilon}) \subseteq \sigma_\varepsilon(J).$$

Proof. By induction. This is clear for $N = 1$; for $N = 2$, we will show that

$$B(\lambda, \sqrt{\varepsilon}) \subseteq \sigma_\varepsilon(J).$$

Let $z \in B(\lambda, \sqrt{\varepsilon})$. Write $z = \lambda + \omega$, where $|\omega| < \sqrt{\varepsilon}$. Define

$$E = \begin{pmatrix} 0 & 0 \\ \omega^2 & 0 \end{pmatrix}.$$

Note $||E|| = |\omega|^2 < \varepsilon$. Observe

$$\det(J + E - zI) = \det\begin{pmatrix} \lambda - z & 1 \\ \omega^2 & \lambda - z \end{pmatrix} = \det\begin{pmatrix} -\omega & 1 \\ \omega^2 & -\omega \end{pmatrix} = 0.$$

Thus, $z \in \sigma(J + E)$, and hence $z \in \sigma_\varepsilon(J)$.

Now, assume that the proposition is true for all $N \times N$ Jordan matrices. Let J be an $(N+1) \times (N+1)$ Jordan matrix, let $z \in B\left(\lambda, \varepsilon^{N+1}\right)$, and let $z = \lambda + \omega$, where

$|\omega| < \varepsilon^{N+1}$. Let E be a matrix with ω^{N+1} in the bottom-left corner entry and zeros elsewhere. Note that $\|E\| = |\omega|^{N+1} < \varepsilon$. Then,

$$\det(J + E - zI) = (-\omega)^{N+1} + (-1)^N \omega^{N+1} = \omega^{N+1} \left((-1)^{N+1} + (-1)^N\right) = 0.$$

Thus, $z \in \sigma(J + E)$, and so $z \in \sigma_\varepsilon(J)$. □

From the second definition of the pseudospectrum, using a perturbation matrix, we can get a better explicit lower bound on the ε-pseudospectra of an $N \times N$ Jordan block.

Proposition 5.2 (Better Lower Bound)**.** *Let J be an $N \times N$ Jordan block. Then,*

$$B\left(\lambda, \sqrt[N]{\varepsilon(1+\varepsilon)^{N-1}}\right) \subseteq \sigma_\varepsilon(J).$$

Proof. We use the second definition for $\sigma_\varepsilon(J)$. Let

$$E = \begin{pmatrix} 0 & k & & & \\ & 0 & k & & \\ & & \ddots & \ddots & \\ & & & \ddots & k \\ k & & & & 0 \end{pmatrix}$$

where $|k| < \varepsilon$, and note that $\|E\| < \varepsilon$. We take $\det(J + E - zI)$ and set it equal to zero to find the eigenvalues of $J + E$.

$$\begin{aligned} 0 &= \det(J + E - zI) \\ &= \det \begin{pmatrix} \lambda - z & k+1 & & & \\ & \lambda - z & k+1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & k+1 \\ k & & & & \lambda - z \end{pmatrix} \\ &= (\lambda - z)^N + (-1)^{N-1} k(1+k)^{N-1} \\ &= (-1)^{N-1}((z-\lambda)^N + k(1+k)^{N-1}); \\ \iff (z-\lambda)^N &= k(1+k)^{N-1} \\ z - \lambda &= \sqrt[N]{k(1+k)^{N-1}}. \end{aligned}$$

So, $B\left(\lambda, \sqrt[N]{\varepsilon(1+\varepsilon)^{N-1}}\right) \subseteq \sigma_\varepsilon(J)$. □

From [14, pg. 470], we know that the ε-pseudospectrum of the Jordan block is a perfect disk about the eigenvalue of J of some radius. Although we do not find the explicit radius, we can use Theorem 4.1 to find the asymptotic behavior of $\sigma_\varepsilon(J)$,

Proposition 5.3 (Asymptotic Bound)**.** *Let J be an $N \times N$ Jordan block. Then*

$$\sigma_\varepsilon(J) \approx B(\lambda, \varepsilon^{1/N} + o(\varepsilon^{1/N})).$$

Proof. The $N \times N$ Jordan block has left and right eigenvectors u_j and v_j where $\|u_j\| = 1$ and $\|v_j\| = 1$. So, from Theorem 4.1, we find $C = \|XY\| = \|v_j u_j\| = 1$. Thus,

$$\sigma_\varepsilon(J) \approx B(\lambda, \varepsilon^{1/N} + o(\varepsilon^{1/N})).$$

□

However, we can find the perturbation matrix that gives us the boundary of $\sigma_\varepsilon(J)$.

Note: Let J be an $n \times n$ Jordan block and E be the perturbation matrix

$$E = \begin{pmatrix} & & & & \varepsilon \\ & & & \iddots & \\ & & \iddots & & \\ & \iddots & & & \\ \varepsilon & & & & \end{pmatrix}.$$

We know that the ε-pseudospectrum of any Jordan block is a perfect disk around the eigenvalue. We observe that as we range θ from 0 to 2π, $\sigma(A + e^{i\theta}E)$ numerically traces the boundary of $\sigma_\varepsilon(J)$ for $\varepsilon > 0$.

For example, suppose J is a 5×5 Jordan block with 0 as an eigenvalue. Letting $\varepsilon = .025$, we get through eigtool that the pseudospectral radius (defined as the farthest point on the boundary of pseudo spectrum from 0) is 0.50765. We now use Mathematica to compute the eigenvalue of $(J + E)$ and obtain 0.50765 as well.

6. Pseudospectra of bidiagonal matrices

The Jordan block is a simple example of a bidiagonal matrix. We now extend our application of Theorem 4.1 to all bidiagonal matrices.

6.1. **Asymptotic Bounds for Periodic Bidiagonal Matrices.**

We investigate various classes of bidiagonal matrices. We describe the asymptotic behavior of the ε-pseudospectrum for the general periodic bidiagonal case, namely A is an $N \times N$ matrix with period k on the main diagonal and nonzero superdiagonal entries

$$A = \begin{pmatrix} a_1 & b_1 & & & & & & \\ & \ddots & \ddots & & & & & \\ & & a_k & \ddots & & & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & \ddots & & \\ & & & & & a_1 & \ddots & \\ & & & & & & \ddots & b_{N-1} \\ & & & & & & & a_r \end{pmatrix}$$

We have from Theorem 4.1, that

$$\sigma_\varepsilon(A) \approx \bigcup_{j=1}^{k} B\left(a_j, (C_j\varepsilon)^{\frac{1}{n_j}}\right),$$

where n_j is the size of the Jordan block corresponding to a_j and $C_j = \|X_j Y_j\|$. For matrices that are of the bidiagonal form, the geometric multiplicity of each eigenvalue is 1, meaning that all the eigenvalues fall in Case 3 of Theorem 4.1. Thus, the constant C_j that multiplies any eigenvalue a_j is simply $C_j = \|v_j\|\|u_j\|$, where v_j and u_j denote the right and left eigenvectors, respectively.

We will begin by introducing ε-pseudospectrum for simple special cases which lead to the most general case.

The cases will be presented as follows:

(1) Let A be a $kn \times kn$ matrix with $a_1, \dots, a_k$ distinct and $b_i = 1$ for all i.
(2) Let A be a $kn \times kn$ matrix with $a_1, \dots, a_k$ distinct.
(3) Let A be an $N \times N$ matrix with $a_1, \dots, a_k$ distinct.
(4) **General Case:** Let A be an $N \times N$ matrix with $a_1, \dots, a_k$ not distinct.

We define:

$$f(x) = \begin{cases} x, & x \neq 0 \\ 1, & x = 1 \end{cases} .$$

Case 1:

- The size of A is $kn \times kn$.
- The a_i's are distinct.
- $b_i = 1$ for all i.

So, A is of the form:

$$A = \begin{pmatrix} a_1 & 1 & & & & & & \\ & \ddots & \ddots & & & & & \\ & & a_k & \ddots & & & & \\ & & & \ddots & \ddots & & & \\ & & & & \ddots & \ddots & & \\ & & & & & a_1 & \ddots & \\ & & & & & & \ddots & 1 \\ & & & & & & & a_k \end{pmatrix}.$$

To find the asymptotic bound for the resolvent norm of A around each eigenvalue, we must determine the right and left eigenvectors associated with each eigenvalue. We claim:

$$v_j = \begin{pmatrix} \frac{1}{f(a_j - a_{j-1}) \cdots f(a_j - a_1)} \\ \frac{1}{f(a_j - a_{j-1}) \cdots f(a_j - a_2)} \\ \vdots \\ \frac{1}{f(a_j - a_{j-1})} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$u_j = \frac{f(a_1 - a_j)\cdots f(a_{j-1} - a_j)}{[f(a_1 - a_j)\cdots f(a_{j-1} - a_j)f(a_{j+1} - a_j)\cdots f(a_k - a_j)]^n} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(a_{j+1} - a_j)\cdots f(a_k - a_j) \\ f(a_{j+2} - a_j)\cdots f(a_k - a_j) \\ \vdots \\ f(a_k - a_j) \\ 1 \end{pmatrix},$$

which can be verified by direction computation of the left and right eigenvectors for each eigenvalue $a_1, \dots, a_k$.

We will check that v_j is a right eigenvector for eigenvalue a_j by cases.

First observe that for $i > j$, we have $(v_j)_i = 0$. Therefore,

$$(Av_j)_i = A_{ii}(v_j)_i + A_{i(i+1)}(v_j)_{i+1} = 0 = a_j(v_j)_i.$$

Now suppose $i \leq j$. Then, we have

$$(Av_j)_i = A_{ii}(v_j)_i + A_{i(i+1)}(v_j)_{i+1} = a_i(v_j)_i + (v_j)_{i+1}.$$

For $1 \leq i \leq j - 2$, we get

$$\begin{aligned} (Av_j)_i &= a_i \cdot \frac{1}{(a_j - a_{j-1})\cdots(a_j - a_i)} + \frac{1}{(a_j - a_{j-1})\cdots(a_j - a_{i+1})} \\ &= \frac{a_i + (a_j - a_i)}{(a_j - a_{j-1})\cdots(a_j - a_i)} \\ &= a_j(v_j)_i. \end{aligned}$$

For $i = j - 1$, we get

$$(Av_j)_{j-1} = a_{j-1} \cdot \frac{1}{a_j - a_{j-1}} + 1 = a_j(v_j)_{j-1}.$$

For $i = j$ we get

$$(Av_j)_j = a_j(v_j)_j + 0.$$

Thus, in all cases we see that $Av_j = a_j v_j$, and so v_j is indeed a right eigenvector for the eigenvalue a_j. The proof that u_j^* is a left eigenvector for eigenvalue a_j is similar.

Case 2:

- The size of A is $kn \times kn$.
- The a_i's are distinct.

Since the superdiagonal of A is not constant, we now write the elements of the superdiagonal as $b_1, b_2, \dots, b_{N-1}$.

We claim:

$$v_j = \begin{pmatrix} \frac{b_1 \cdots b_{j-1}}{f(a_j - a_1)\cdots f(a_j - a_{j-1})} \\ \frac{b_2 \cdots b_{j-1}}{f(a_j - a_2)\cdots f(a_j - a_{j-1})} \\ \vdots \\ \frac{b_{j-1}}{f(a_j - a_{j-1})} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$u_j^* = \left(0, \quad \cdots \quad 0, \quad y, \quad \frac{b_p}{f(a_{j+1}-a_j)} y, \quad \cdots \quad \frac{b_p \cdots b_{N-2}}{f(a_{j+1}-a_j)\cdots f(a_{k-1}-a_j)} y, \quad \frac{b_p \cdots b_{N-1}}{f(a_{j+1}-a_j)\cdots f(a_k-a_j)} y\right),$$

where $p = k(n-1) + j$, and

$$y = \frac{b_j \cdots b_{p-1}}{((a_1 - a_j)\cdots(a_{j-1} - a_j)(a_{j+1} - a_j)\cdots(a_k - a_j))^{n-1}}.$$

Again, direct computation will show that these are indeed left and right eigenvectors associated with any eigenvalue a_j.

Case 3:

- The size of A is $N \times N$.
- The a_i's are distinct.

We now drop our assumption that the size of our matrix is $kn \times kn$, for period k on the diagonal. Let n, r be such that $N = kn + r$, where $0 < r \le k$. In other words, a_r is the last entry on the main diagonal, so the period does not necessarily complete.

For a_j, the right eigenvector is given by

$$v_j = \begin{pmatrix} \frac{b_1 \cdots b_{j-1}}{f(a_j - a_1)\cdots f(a_j - a_{j-1})} \\ \frac{b_2 \cdots b_{j-1}}{f(a_j - a_2)\cdots f(a_j - a_{j-1})} \\ \vdots \\ \frac{b_{j-1}}{f(a_j - a_{j-1})} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We split up the formula for the left eigenvectors into two cases:

(1) $1 \le j \le r$
(2) $r < j \le k$.

(1) $1 \le j \le r$
On the main diagonal, there are n *complete* blocks with entries $a_1, \dots, a_k$, and one *partial* block at the end with entries $a_1, \dots a_r$. In the first case, when $1 \le j \le r$, then a_j is in this last partial block. In this case then, let $p = kn + j$.

We claim

$$u_j = \mu_j \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(a_{j+1} - a_j) \cdots f(a_r - a_j) \\ b_p f(a_{j+2} - a_j) \cdots f(a_r - a_j) \\ \vdots \\ b_p \cdots b_{N-2} f(a_r - a_j) \\ b_p \cdots b_{N-1} \end{pmatrix},$$

where $\mu_j = \frac{b_j \cdots b_{p-1} f(a_1 - a_j) \cdots f(a_{j-1} - a_j)}{[f(a_1 - a_j) \cdots f(a_{j-1} - a_j) f(a_{j+1} - a_j) \cdots f(a_k - a_j)]^n f(a_1 - a_j) \cdots f(a_r - a_j)}$.

(2) $r < j \le k$
In this case, a_j is in the last *complete* block. Now, let $p = k(n-1) + j$.

Now, we claim

$$u_j = \mu_j \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(a_1 - a_j) \cdots f(a_r - a_j) f(a_{j+1} - a_j) \cdots f(a_k - a_j) \\ b_p f(a_1 - a_j) \cdots f(a_r - a_j) f(a_{j+2} - a_j) \cdots f(a_k - a_j) \\ \vdots \\ b_p \cdots b_{p+k-j-1} f(a_1 - a_j) \cdots f(a_r - a_j) \\ b_p \cdots b_{p+k-j} f(a_2 - a_j) \cdots f(a_r - a_j) \\ \vdots \\ b_p \cdots b_{N-2} f(a_r - a_j) \\ b_p \cdots b_{N-1} \end{pmatrix},$$

again where $\mu_j = \frac{b_j \cdots b_{p-1} f(a_1 - a_j) \cdots f(a_{j-1} - a_j)}{[f(a_1 - a_j) \cdots f(a_{j-1} - a_j) f(a_{j+1} - a_j) \cdots f(a_k - a_j)]^n f(a_1 - a_j) \cdots f(a_r - a_j)}$

Case 4: General Case.

- The size of A is $N \times N$.
- The a_i's are not distinct for $1 \le i \le k$

Let A be a $N \times N$ periodic bidiagonal matrix with period k on the main diagonal. Let n, r be such that $N = kn + r$, where $0 < r \le k$. Write $a_1, \dots, a_k$ for the entries on the main diagonal (a_i's not distinct) and $b_1, \dots, b_{N-1}$ for the entries on the superdiagonal. Let a_r be the last entry on the main diagonal.

We can explicitly find the left and right eigenvectors for any eigenvalue, α. Suppose α first appears in position ℓ of the period k. Then the corresponding right eigenvector for α is the same form as v_ℓ in Case 4. That is,

$$v_\ell = \begin{pmatrix} \frac{b_1 \cdots b_{\ell-1}}{f(a_\ell - a_1)\cdots f(a_\ell - a_{\ell-1})} \\ \frac{b_2 \cdots b_{\ell-1}}{f(a_\ell - a_2)\cdots f(a_\ell - a_{\ell-1})} \\ \vdots \\ \frac{b_{\ell-1}}{f(a_\ell - a_{\ell-1})} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The corresponding left eigenvector for α depends on the first and last positions of α. Let $q \equiv m \pmod{k}$. We split up the formula for the left eigenvector of α into two cases, which again mirror the formulas given in Case 4:

(1) $1 \le \ell \le r$
(2) $r < \ell \le k$.

For both of these two cases, we define

$$f(b_i) = \begin{cases} b_i, & \text{i} \ge p \\ 1, & \text{i} < p \end{cases}.$$

(1) $1 \le \ell \le r$
In this case then, α appears in the partial block. Let $p = kn + \ell$.

We claim

$$u_\ell = \mu_\ell \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(b_{p+m-\ell-1}) f(a_{m+1} - a_\ell) \cdots f(a_r - a_\ell) \\ f(b_{p+m-\ell-1}) f(b_{p+m-l}) f(a_{m+2} - a_\ell) \cdots f(a_r - a_\ell) \\ \vdots \\ f(b_{p+m-\ell-1}) \cdots f(b_{N-2}) \cdots b_{N-2} f(a_r - a_\ell) \\ f(b_{p+m-\ell-1}) \cdots f(b_{N-1}) \end{pmatrix},$$

where $\mu_\ell = \frac{b_\ell \cdots b_{p-1} f(a_1 - a_\ell) \cdots f(a_{\ell-1} - a_\ell)}{[f(a_1 - a_\ell) \cdots f(a_{\ell-1} - a_j) f(a_{\ell+1} - a_\ell) \cdots f(a_k - a_\ell)]^n f(a_1 - a_\ell) \cdots f(a_r - a_\ell)}$.

(2) $r < j \le k$
In this case, α is in the last *complete* block. Here, we let $p = k(n-1) + \ell$

Now, we claim

$$u_\ell = \mu_\ell \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(b_{p+m-\ell-1})f(a_1 - a_\ell)\cdots(a_r - a_\ell)f(a_{m+1} - a_\ell)\cdots f(a_k - a_\ell) \\ f(b_{p+m-\ell-1})f(b_{p+m-l})f(a_1 - a_\ell)\cdots f(a_r - a_\ell)f(a_{m+2} - a_\ell)\cdots f(a_k - a_\ell) \\ \vdots \\ f(b_{p+m-\ell-1})\cdots f(b_{p+k-\ell-1})f(a_1 - a_\ell)\cdots f(a_r - a_\ell) \\ f(b_{p+m-\ell-1})\cdots f(b_{p+k-\ell})f(a_2 - a_\ell)\cdots f(a_r - a_\ell) \\ \vdots \\ f(b_{p+m-\ell-1})\cdots f(b_{N-2})f(a_r - a_\ell) \\ f(b_{p+m-\ell-1})\cdots f(b_{N-1}) \end{pmatrix},$$

where

$$\mu_\ell = \frac{b_\ell \cdots b_{p-1} f(a_1 - a_\ell)\cdots f(a_{\ell-1} - a_\ell)}{[f(a_1 - a_\ell)\cdots f(a_{\ell-1} - a_\ell)f(a_{\ell+1} - a_j)\cdots f(a_k - a_\ell)]^n f(a_1 - a_\ell)\cdots f(a_r - a_\ell)}.$$

From these formulas, we can find the eigenvectors, and hence the asymptotic behavior of the ε-pseudospectrum for any bidiagonal matrix,

$$\sigma_\varepsilon(A) \approx \bigcup_{j=1}^{k} B\left(a_j, (C_j\varepsilon)^{\frac{1}{n_j}}\right)$$

where $C_j = \|v_j\|\|u_j\|$ and n_j is the size of the Jordan block corresponding to a_j.

Note 1: Let A be a periodic, bidiagonal matrix and suppose $b_i = 0$ for some i. Then the matrix decouples into the direct sum of smaller matrices, call them $A_1, \ldots, A_n$. To find the ε-pseudospectrum of A, apply the same analysis to these smaller matrices, and from Theorem 2.2, we have that

$$\sigma_\varepsilon(A) = \bigcup_{i=1}^{n} \sigma_\varepsilon(A_i).$$

Note 2: It is also interesting that given the results in Case 4, we can find different matrices with the same asymptotic ε-pseudospectra. Based on the formulas for the right and left eigenvectors, we observe that we can change the arrangement of the period without changing the left and right eigenvectors for any eigenvalue if the following hold:

(1) The algebraic multiplicity of each a_i remains the same.
(2) The first position of each distinct eigenvalue remains the same. Additionally, suppose α first appears in position ℓ. We can permute the previous entries in positions $1, 2, \ldots, \ell - 1$ as long as we fix the number of times each distinct entry appears.
(3) The last position of each distinct eigenvalue remains the same in a period. Additionally, suppose α last appears in position q on the main diagonal. Let $q \equiv m \pmod{k}$. We can permute the other entries in positions $m+1, \ldots, k$ as long as we fix the number of times each distinct entry appears.

Example 6.1. Let matrix A have period 13 on the main diagonal, with distinct eigenvalues a, b, c. Let A have period

$$\{a, b, a, b, c, a, b, c, a, b, c, b, c\}.$$

Note that the first entries of a, b, c are fixed in positions $1, 2, 5$ respectively and the last entries are fixed in positions $9, 12, 13$ respectively.

We can permute entries $3, 4$ since permuting these entries does not change the number of a or b entries before the first entry of c. We can also permute entries $6, 7, 8$, as well as the entries $10, 11$ for similar reasons. We now give possible permutations of said entries:

- A_1 has period $\{a, b, b, a, c, a, b, c, a, b, c, b, c\}$, permuting entries 3 and 4,
- A_2 has period $\{a, b, a, b, c, b, c, a, a, b, c, b, c\}$, permuting entries 6, 7, 8,
- A_3 has period $\{a, b, b, a, c, a, b, c, a, c, b, b, c\}$, permuting entries 10 and 11.

With these possible permutations,

$$\sigma_\varepsilon(A) \approx \sigma_\varepsilon(A_1) \approx \sigma_\varepsilon(A_2) \approx \sigma_\varepsilon(A_3)$$

We will now present examples of bidiagonal matrices with explicit and asymptotic bounds.

6.2. **Examples of Bidiagonal Matrices.**

Example 6.2. Let A be an $n \times n$ bidiagonal matrix,

$$A = \begin{pmatrix} a & b_1 & 0 & 0 & \dots \\ 0 & a & b_2 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \dots \\ \vdots & \vdots & \vdots & \ddots & b_{n-1} \\ 0 & 0 & 0 & 0 & a \end{pmatrix}$$

where $b_i \neq 0$. The following propositions give explicit lower and upper bounds for the ε-pseudospectrum of A.

Proposition 6.1 (Lower Bound)**.**

$$B\left(a, (\varepsilon|b_1 b_2 \cdots b_{n-1}|)^{\frac{1}{n}}\right) \subseteq \sigma_\varepsilon(A).$$

Proof. Let $z \in B\left(a, (\varepsilon|b_1 b_2 \cdots b_{n-1}|)^{\frac{1}{n}}\right)$. We will show that $z \in \sigma_\varepsilon(A)$. Let E be an $n \times n$ matrix with ε in the lower-left corner and 0's elsewhere. Then $\|E\| \leq \varepsilon$. We take the determinant of $A + E - z$ and set it equal to 0 to get the eigenvalues of A+E.

$$\begin{aligned} 0 = |A + E - z| &= (\lambda - z)^n + (-1)^{n-1} \varepsilon b_1 b_2 \cdots b_{n-1} \\ &\implies (\lambda - z)^n = (-1)^n \varepsilon b_1 b_2 \cdots b_{n-1} \\ &\implies z = \lambda + (\varepsilon b_1 b_2 \cdots b_{n-1})^{\frac{1}{n}}. \end{aligned}$$

Thus, $z \in \sigma_\varepsilon(A)$. □

Proposition 6.2 (Upper Bound)**.** *The Jordan decomposition of A is*

$$A = \begin{pmatrix} 1 & & & & \\ & \frac{1}{b_1} & & & \\ & & \frac{1}{b_1 b_2} & & \\ & & & \ddots & \\ & & & & \frac{1}{b_1 \cdots b_{n-1}} \end{pmatrix} \begin{pmatrix} a & 1 & & & \\ & a & 1 & & \\ & & a & \ddots & \\ & & & \ddots & 1 \\ & & & & a \end{pmatrix} \begin{pmatrix} 1 & & & & \\ & b_1 & & & \\ & & b_1 b_2 & & \\ & & & \ddots & \\ & & & & b_1 \cdots b_{n-1} \end{pmatrix}.$$

Thus $\sigma_\varepsilon(A) \subseteq B\left(a, (|b_1 \cdots b_{n-1}|\varepsilon)^{\frac{1}{n}} + o\left(\varepsilon^{1/n}\right)\right)$.

Combining this upper bound with the lower bound from Prop. 6.1, we get that asymptotically, $\sigma_\varepsilon(A) \approx B\left(a, (|b_1 \cdots b_{n-1}|\varepsilon)^{\frac{1}{n}}\right)$.

Note that if $b_i = 0$ for some i, then the matrix decouples, such that its Jordan decomposition has multiple blocks and $\sigma_\varepsilon(A) \subseteq B\left(a, (|b_k \cdots b_j|\varepsilon)^{\frac{1}{n}} + o\left(\varepsilon^{1/n}\right)\right)$ where $b_k, b_{k+1}, \ldots b_j$ correspond to the largest Jordan block.

We now find bounds for a bidiagonal matrix of period 2 on the main diagonal.

Example 6.3. Restricting to period 2 on the main diagonal and constants on the superdiagonal, we can get explicit lower bounds for our pseudospectra.

Let A be a $2N \times 2N$ bidiagonal matrix with period 2 on the main diagonal, such that

$$A = \begin{pmatrix} a_1 & 1 & & & \\ & a_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & a_1 & 1 \\ & & & & a_2 \end{pmatrix}.$$

Proposition 6.3 (Lower Bound)**.**

$$\bigcup_{j=1}^{2} B\left(a_j, \frac{1}{|a_1 - a_2|}\left(\varepsilon^{\frac{1}{N}}\right)\right) \subseteq \sigma_\varepsilon(A).$$

Proof. Fix $\theta \in [0, 2N\pi)$, and let E be the perturbation matrix with $\varepsilon e^{i\theta}$ on the lower left corner.

Then, $\det(A + E - z) = 0 \Rightarrow (a_1 - z)^N (a_2 - z)^N - \varepsilon = 0$. Solving for z, we find that

$$z = \frac{a_1 + a_2 \pm \sqrt{(a_1 - a_2)^2 + 4(\varepsilon e^{i\theta})^{1/N}}}{2}.$$

Around the eigenvalue λ_1 as $\varepsilon \to 0$, we have

$$\begin{aligned} z &= \frac{a_1 + a_2 + \sqrt{(a_1 - a_2)^2 + 4(\varepsilon e^{i\theta})^{1/N}}}{2} \\ &= \frac{a_1 + a_2}{2} + \frac{a_1 - a_2}{2}\left[1 + \frac{4\left(\varepsilon e^{i\theta}\right)^{\frac{1}{N}}}{(a_1 - a_2)^2}\right]^{1/2} \\ &\approx \frac{a_1 + a_2}{2} + \frac{a_1 - a_2}{2}\left[1 + \frac{2\left(\varepsilon e^{i\theta}\right)^{\frac{1}{N}}}{(a_1 - a_2)^2}\right] \\ &= a_1 + \frac{\left(\varepsilon e^{i\theta}\right)^{1/N}}{a_1 - a_2}. \end{aligned}$$

So, $|z - a_1| \approx \frac{\varepsilon^{\frac{1}{N}}}{|a_1 - a_2|} e^{i\theta/N}$.

Performing the same calculation for the eigenvalue a_2, we find that

$$|z - a_2| \approx \frac{\varepsilon^{\frac{1}{N}}}{|a_1 - a_2|} e^{i\theta/N}.$$

Ranging θ over $[0, 2N\pi)$, we get that the disks of radius $\frac{\varepsilon^{\frac{1}{N}}}{|a_1 - a_2|}$, centered around a_1 and a_2 are contained in our ε-pseudospectrum. In other words,

$$\bigcup_{j=1}^{2} B\left(a_j, \frac{1}{|a_1 - a_2|}\left(\varepsilon^{\frac{1}{N}}\right)\right) \subseteq \sigma_\varepsilon(A).$$

□

We also have an asymptotic upper bound for $\sigma_\varepsilon(A)$:

$$|z - a_j| \le \frac{\varepsilon^{\frac{1}{N}}}{|a_1 - a_2|}\left(1 + |a_1 - a_2|^2\right)^{\frac{1}{2N}}$$

and the ratio of the upper and lower bounds is just $\left(1 + |a_1 - a_2|^2\right)^{\frac{1}{2N}}$.

Proposition 6.4 (Asymptotic Lower Bound)**.**

$$\bigcup_{j=1}^{2} B\left(a_j, \frac{1}{|a_1 - a_2|}\left(\sqrt{1 + |a_1 - a_2|^2}\varepsilon\right)^{\frac{1}{N}} + o(\varepsilon^{\frac{1}{N}})\right) \subseteq \sigma_\varepsilon(A).$$

Proof. From Theorem 4.1, we can find an asymptotic lower bound for the bidiagonal matrix of period 2.

We Jordan decompose matrix A such that $QAP = J$ where $QP = I$. Then, we multiply the first column of P, call it x_1, by the N^{th} row of Q, call it y_1 to get the constant that multiplies a_1. We find that

$$\|x_1 y_1\| = \frac{1}{|a_1 - a_2|^N}\sqrt{1 + |a_1 - a_2|^2}.$$

For a_2, we find $x_2 y_2$ by multiplying the N^{th} column of P with the last row of Q, and get that

$$\|x_2 y_2\| = \|x_1 y_1\|.$$

So, as $\varepsilon \to 0$

$$\bigcup_{j=1}^{2} B\left(a_j, \frac{1}{|a_1 - a_2|}\left(\sqrt{1 + |a_1 - a_2|^2 \varepsilon}\right)^{\frac{1}{N}} + o(\varepsilon^{\frac{1}{N}})\right) \subseteq \sigma_\varepsilon(A)$$

□

This lower bound differs from the lower bound given in Proposition 6.3 by a constant, $\alpha = (\sqrt{1 + |a_1 - a_2|^2})^{\frac{1}{N}}$ and some $o(\varepsilon^{\frac{1}{N}})$.

7. Tridiagonal Matrices

The natural extension to bidiagonal matrices is to consider tridiagonal matrices. The study of tridiagonal matrices arises from a quantum mechanics random matrix problem known as the Anderson model, as introduced by Anderson [1]. Hatano, Nelson, and Shnerb [**?**] studied the non-Hermitian analog of the Anderson model, now known as the Hatano-Nelson model, which takes the form

$$A = \begin{pmatrix} a_1 & e^g & & & e^{-g} \\ e^{-g} & a_2 & e^g & & \\ & \ddots & \ddots & \ddots & \\ & & e^{-g} & a_{N-1} & e^g \\ e^g & & & e^{-g} & a_N \end{pmatrix}$$

where g is a fixed real parameter and $a_1, \dots a_N$ are iid random variables. We are interested in the nonperiodic Hatano-Nelson matrix

$$A = \begin{pmatrix} a_1 & e^g & & & \\ e^{-g} & a_2 & e^g & & \\ & \ddots & \ddots & \ddots & \\ & & e^{-g} & a_{N-1} & e^g \\ & & & e^{-g} & a_N \end{pmatrix}$$

In this case, we can use a diagonal similarity transformation of $D = \operatorname{diag}(1, e^g, e^{2g}, \dots, e^{(N-1)g})$ to get that the nonperiodic Hatano-Nelson matrix is similar to the matrix

$$DAD^{-1} = \begin{pmatrix} a_1 & 1 & & & \\ 1 & a_2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & a_{N-1} & 1 \\ & & & 1 & a_N \end{pmatrix}$$

This matrix will only have one Jordan block per eigenvalue [7]. In fact, since this matrix is symmetric and tridiagonal with nonzero off-diagonal entries, then we know it has N distinct eigenvalues [12, 6].

For the case when $a = a_1 = a_2 = \ldots = a_N$, then A is Toeplitz, and it is known that the eigenvalues are of the form $\lambda_k = a + 2\cos\left(\frac{k\pi}{N+1}\right)$ and eigenvectors take the form $u_j^{(k)} = \sin\frac{kj\pi}{n+1}$ [6, 15]. The Toeplitz case has been studied [11] and the sensitivity of the eigenvalues investigated; however by finding the eigenvectors, we can give asymptotic results on the pseudospectrum.

For the case when the main diagonal is periodic with period 2, i.e.

$$a_j = \begin{cases} a_1 & \text{if } j \text{ odd} \\ a_2 & \text{if } j \text{ even} \end{cases}$$

then we can find an explicit formula for the eigenvalues and eigenvectors [9, 6]. With these formulae, we can then find an asymptotic upper bound on $\sigma_\varepsilon(A)$, as we did for bidiagonal matrices.

Further research would attempt to generalize these results to period k on the main diagonal, and then ultimately to arbitrary diagonal entries.

8. Finite Rank operators

The majority of this paper has focused on both explicit and asymptotic characterizations of ε-pseudospectra for various classes of finite dimensional linear operators. A natural next step is to consider finite rank operators on infinite dimensional space.

In section 2 we defined ε-pseudospectra for matrices, although our definitions are exactly the same in the infinite dimensional case. For our purposes, the only noteworthy difference between matrices and operators is that the spectrum of an operator is no longer defined as the collection of eigenvalues, but rather

$$\sigma(A) = \{\lambda \mid \lambda I - A \text{ does not have a bounded inverse}\}$$

As a result, we do not get the same properties for pseudospectra as we did previously; in particular, $\sigma_\varepsilon(A)$ is not necessarily bounded.

That being said, the following theorem shows that finite rank operators behave similarly to matrices, in that asymptotically the radii ε-pseudospectra are bounded by roots of epsilon. The following theorem makes this precise.

Theorem 8.1. *Let V be a Hilbert space and $A : V \to V$ a finite rank operator on H. Then there exists C such that for sufficiently small ε,*

$$\sigma_\varepsilon(A) \subseteq \sigma(A) + B(0, C\varepsilon^{\frac{1}{m+1}}).$$

where m is the rank of A. Furthermore, this bound is sharp in the sense that there exists a rank-m operator A and a constant c such that

$$\sigma_\varepsilon(A) \supseteq \sigma(A) + B(0, c\varepsilon^{\frac{1}{m+1}})$$

for sufficiently small ε.

Proof. Since A has finite rank, there exists a finite dimensional subspace U such that $V = U \oplus W$ and $A(U) \subseteq U$, and $A(W) = \{0\}$. Choosing an orthonormal basis for A which respects this decomposition we can write $A = A' \oplus 0$. Then the spectrum of A is $\sigma(A') \cup \{0\}$, and we know that for any ε,

$$\sigma_\varepsilon(A) = \sigma_\varepsilon(A') \cup \sigma_\varepsilon(0).$$

The ε-pseudospectrum of the zero operator is well-understood since this operator is normal; for any ε, it is precisely the ball of radius ε. It thus suffices to consider the ε-pseudospectrum of the finite rank operator $A' : U \to U$, where U is finite dimensional. The ε-pseudospectrum of this operator goes like $\varepsilon^{1/j}$, where j is the dimension of the largest Jordan block; we will prove that $j \leq m+1$. Note that the rank of the $n \times n$ Jordan block given by

$$A = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots \\ 0 & \lambda & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \dots \\ \vdots & \vdots & \vdots & \ddots & b_{n-1} \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

is n if $\lambda \neq 0$, and $n-1$ if $\lambda = 0$. Since we know that the rank of A is larger than or equal to the rank of the largest Jordan block, we have an upper bound on the dimension of the largest Jordan block: it is of size $m+1$, with equality attained when $\lambda = 0$. By previous theorems, we then know that $\sigma_\varepsilon(A)$ is contained, for small enough ε, in the set $\sigma(A) + C\varepsilon^{\frac{1}{m+1}}$.

Note that this bound is sharp; we can see this by taking V to be $\mathbb{R}^{m+1}$ and considering the rank-m operator

$$A_m = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \ddots & \dots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

the pseudospectrum of which will contain the ball of radius $\varepsilon^{1/m+1}$ by proposition 5.1. ☐

The natural question to ask now is whether we can extend this result to more arbitrary operators on Hilbert spaces. For operators on Banach spaces, we do still have a certain convergence of the ε-pseudospectrum to the spectrum [14, §4], namely $\cap_{\varepsilon>0} \sigma_\varepsilon(A) = \sigma(A)$. Also, while the ε-pseudospectrum may be unbounded, any bounded component of it necessarily contains a component of the spectrum.

From these results, we see that we can no longer get a complete asymptotic result for $\sigma_\varepsilon(A)$, as there could be an unbounded component of $\sigma_\varepsilon(A)$ that has no intersection with the spectrum. For example, there exist operators with bounded spectrum and unbounded ε-pseudospectra for all $\varepsilon > 0$.

However, the above result states that the bounded components of the ε-pseudospectrum must converge to the spectrum. If we restrict our attention to these bounded components, we can generalize Thms. 4.1 and 8.1 by asking whether the bounded components of $\sigma_\varepsilon(A)$ converge to the spectrum as a union of disks.

References

[1] P. W. Anderson. Absence of diffusion in certain random lattices. *Phys. Rev.*, 109:1492–1505, Mar 1958.

[2] Friedrich L Bauer and CT Fike. Norms and exclusion theorems. *Numerische Mathematik*, 2(1):137–141, 1960.

[3] Hellmut Baumgärtel. *Analytic perturbation theory for matrices and operators*, volume 17. Birkhäuser Basel, 1985.

[4] EB Davies. Pseudo–spectra, the harmonic oscillator and complex resonances. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 455(1982):585–599, 1999.

[5] EB Davies. Semi-classical states for non-self-adjoint schrödinger operators. *Communications in mathematical physics*, 200(1):35–41, 1999.

[6] Joseph Frederick Elliott. The characteristic roots of certain real symmetric matrices. Master's thesis, University of Tennessee, 1953.

[7] Carla Ferreira, Beresford Parlett, and Froilán M Dopico. Sensitivity of eigenvalues of an unsymmetric tridiagonal matrix. *Numerische Mathematik*, 122(3):527–555, 2012.

[8] Tosio Kato. *Perturbation theory for linear operators*, volume 132. springer, 1995.

[9] S Kouachi. Eigenvalues and eigenvectors of some tridiagonal matrices with non-constant diagonal entries. *Applicationes mathematicae*, 35(1):107, 2008.

[10] Julio Moro, James V Burke, and Michael L Overton. On the lidskii–vishik–lyusternik perturbation theory for eigenvalues of matrices with arbitrary jordan structure. *SIAM Journal on Matrix Analysis and Applications*, 18(4):793–817, 1997.

[11] Silvia Noschese, Lionello Pasquini, and Lothar Reichel. Tridiagonal toeplitz matrices: properties and novel applications. *Numerical Linear Algebra with Applications*, 20(2):302–326, 2013.

[12] Beresford N Parlett. *The symmetric eigenvalue problem*, volume 7. SIAM, 1980.

[13] AVD Sluis. Condition numbers and equilibration of matrices. *Numerische Mathematik*, 14(1):14–23, 1969.

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

[15] W. C. Yueh. Eigenvalues of Several Tridiagonal Matrices. *Appl. Math. E-Notes*, 5:66–74, 2005.