

involve

a journal of mathematics

Explicit bounds for the pseudospectra of
various classes of matrices and operators

Feixue Gong, Olivia Meyerson, Jeremy Meza,
Mihai Stoiciu and Abigail Ward



Explicit bounds for the pseudospectra of various classes of matrices and operators

Feixue Gong, Olivia Meyerson, Jeremy Meza,
Mihai Stoiciu and Abigail Ward

(Communicated by Stephan Garcia)

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 -pseudospectra of 2×2 matrices and describe the asymptotic behavior (as $\varepsilon \rightarrow 0$) of $\sigma_\varepsilon(A)$ for every square matrix A . We also present explicit upper and lower bounds for the ε -pseudospectra of bidiagonal matrices, as well as for finite-rank operators.

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 [Trefethen and Embree 2005].

In this paper we describe the asymptotic behavior of the ε -pseudospectrum of all $n \times n$ matrices. We apply this asymptotic bound to several classes of matrices and operators, including 2×2 matrices, bidiagonal matrices, and finite-rank operators, and additionally provide explicit bounds on their ε -pseudospectra.

The paper is organized as follows: in Section 2, we give the three standard equivalent definitions for the pseudospectrum and 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 nondiagonalizable and the diagonalizable cases. The asymptotic behavior (as $\varepsilon \rightarrow 0$) of the ε -pseudospectrum of each $n \times n$ matrix is described in Section 4, where we show (in Theorem 4.2) that, for every square matrix, the ε -pseudospectrum converges, as $\varepsilon \rightarrow 0$, to a union of disks. We apply the main result of Section 4 to several classes of matrices: matrices with a simple eigenvalue, matrices with an eigenvalue with geometric multiplicity 1, 2×2 matrices, and Jordan blocks.

MSC2010: 15A18, 15A60, 47A10, 65F15.

Keywords: spectrum, pseudospectrum, bidiagonal matrices, perturbation of eigenvalues.

Section 5 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, given by Theorem 4.2, of the ε -pseudospectrum of A as $\varepsilon \rightarrow 0$. In the last section (Section 6), we consider finite-rank operators and show that the ε -pseudospectrum of an operator of rank m is at most as big as $C\varepsilon^{1/m}$ as $\varepsilon \rightarrow 0$.

2. Pseudospectra

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 observation that if z_n is a sequence of complex numbers converging to an eigenvalue λ of A , then $\|(z_n - A)^{-1}\| \rightarrow \infty$ as $n \rightarrow \infty$. 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 $z \mapsto \|(z - A)^{-1}\|$. Figure 1 depicts the behavior of this function near the eigenvalues.

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 [Trefethen and Embree 2005, Theorem 2.4].

(It is also possible to develop a theory of pseudospectrum 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 [Davies 1999a; 1999b] such that $\|(z - A)^{-1}\|$ approaches infinity, but z does not approach the spectrum of A .)

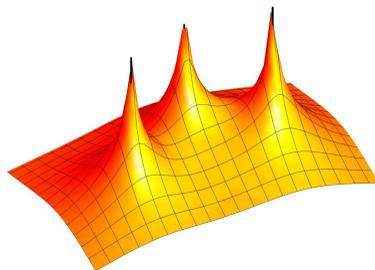


Figure 1. Contour plot of the resolvent norm.

The second and third definitions of the ε -pseudospectrum arise from eigenvalue perturbation theory [Kato 1995].

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 $\|E\| < \varepsilon$.

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 .

The third 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 2.4 (equivalence of the definitions of pseudospectra). *For every matrix $A \in \mathbb{C}^{N \times N}$, the three definitions above are equivalent.*

The proof of this theorem is given in [Trefethen and Embree 2005, Section 2]. As all three definitions are equivalent, we can unambiguously denote the ε -pseudospectrum of A as $\sigma_\varepsilon(A)$.

While the investigation of the set $\sigma_\varepsilon(A)$ can be motivated by questions in numerical analysis, the main impetus for the in-depth study of the ε -pseudospectrum is the study of the size and behavior of the norms $\|e^{tA}\|$ (with $t \in [0, \infty)$) and $\|A^k\|$ (with $k \in \{0, 1, 2, \dots\}$), where A is a matrix or an operator that defines the differential equation $x' = Ax$ or the difference equation $x_{k+1} = Ax_k$.

As explained in [Trefethen and Embree 2005, Part IV] and in [Böttcher 2006], the ε -pseudospectra of A can be used to define and compute the Kreiss constant $\mathcal{K}(A)$, which in turn can be used, via the Kreiss matrix theorem [Trefethen and Embree 2005, Section 18], to find upper and lower bounds for $\sup_{t \geq 0} \|e^{tA}\|$ and $\sup_{k \geq 0} \|A^k\|$. Thus, while the spectrum and the numerical range of A only provide information on $\|e^{tA}\|$ and $\|A^k\|$ in the limits $t \rightarrow \infty$ and $t \rightarrow 0$, as well as for $k \rightarrow \infty$, the

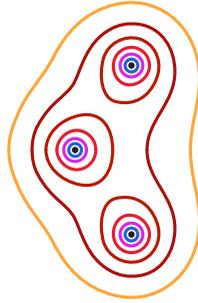


Figure 2. The curves bounding the ε -pseudospectra of a matrix A for different values of ε .

pseudospectrum of A gives information on the size of these norms (and therefore on the size of the solutions of the corresponding linear differential/difference equations) for all values of t and k .

Figure 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.

Several properties of pseudospectra are proven in [Trefethen and Embree 2005, Section 2]. One of which is that if $A \in \mathbb{C}^{N \times N}$, then $\sigma_\varepsilon(A)$ is nonempty, open, and bounded, with at most N connected components, each containing one or more eigenvalues of A . This leads us to the following notation:

Notation. For $\lambda \in \sigma(A)$, we write $\sigma_\varepsilon(A) \upharpoonright \lambda$ to be the connected component of $\sigma_\varepsilon(A)$ that contains λ .

Another property, which follows straight from the definitions of pseudospectra, is that $\bigcap_{\varepsilon > 0} \sigma_\varepsilon(A) = \sigma(A)$. From these properties, it follows that there is ε small enough so that $\sigma_\varepsilon(A)$ consists of exactly $|\sigma(A)|$ connected components, each an open set around a distinct eigenvalue. In particular, there is ε small enough so that $\sigma(A) \cap \sigma_\varepsilon(A) \upharpoonright \lambda = \{\lambda\}$.

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 [Trefethen and Embree 2005]:

Theorem 2.5.
$$\sigma_\varepsilon(A_1 \oplus A_2) = \sigma_\varepsilon(A_1) \cup \sigma_\varepsilon(A_2).$$

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 are particularly well-behaved: Theorem 2.6 shows that the ε -pseudospectrum of a normal matrix is exactly the union of disks of radius ε around each eigenvalue, as in shown in Figure 3. This is clear for diagonal

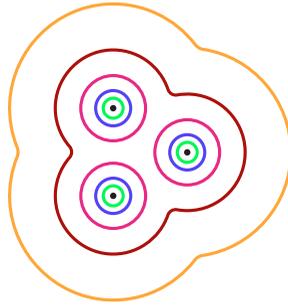


Figure 3. The curves bounding the ε -pseudospectra of a normal matrix for different values of ε . Note that the boundaries are unions of perfect circles around each eigenvalue.

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.

Theorem 2.6. *Let $A \in \mathbb{C}^{N \times N}$. Then,*

$$\sigma(A) + B(0, \varepsilon) \subseteq \sigma_\varepsilon(A) \quad \text{for all } \varepsilon > 0. \tag{2-1}$$

Furthermore, A is a normal matrix if and only if

$$\sigma_\varepsilon(A) = \sigma(A) + B(0, \varepsilon) \quad \text{for all } \varepsilon > 0. \tag{2-2}$$

The proof of this theorem can be found in [Trefethen and Embree 2005, Section 2].

Nonnormal 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. That is, there exist matrices with pseudospectra larger than the union of disks of radius ε . Regardless, we can still characterize the behavior of nonnormal, diagonalizable matrices.

Theorem 2.7 (Bauer–Fike). *Let $A \in \mathbb{C}^{N \times N}$ be diagonalizable with $A = VDV^{-1}$. Then for each $\varepsilon > 0$,*

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

where

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

and $s_{\max}(V), s_{\min}(V)$ are the maximum and minimum singular values of V .

Here, $\kappa(V)$ is known as the condition number of V . Note that $\kappa(V) \geq 1$, with equality attained if and only if A is normal. Thus, $\kappa(V)$ can be thought of as a measure of the normality of a matrix. However, 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 $\|v_j\| = 1$.

Nondiagonalizable 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 1. 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

$$V = (v_1 \ v_2 \ \dots \ v_n).$$

Then,

$$AV = (Av_1 \ Av_2 \ \dots \ Av_n) = (\lambda v_1 \ v_1 + \lambda v_2 \ \dots \ v_{n-1} + \lambda v_n) = VJ,$$

and hence v_1 is a right eigenvector associated with λ and v_2, \dots, v_n are *generalized right eigenvectors*, that is, right eigenvectors for $(A - \lambda I)^k$ for $k > 1$. Similarly, there exists a matrix U such that $U^*A = JU^*$, where now the rows of U^* are left generalized eigenvectors associated with λ .

We can also quantify the normality of an eigenvalue in the same way $\kappa(V)$ quantifies the normality of a matrix.

Definition 2.8. For each 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 $|u_j^* v_j| \leq \|u_j\| \|v_j\|$, 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 simple eigenvalues 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 pseudospectrum centered around each eigenvalue. The following theorem can be found, for example, in [Baumgärtel 1985].

Theorem 2.9 (asymptotic pseudospectra inclusion regions). *Suppose $A \in \mathbb{C}^{N \times N}$ has N distinct eigenvalues. Then, as $\varepsilon \rightarrow 0$,*

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

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 [Bauer and Fike 1960, Theorem 4].

Theorem 2.10 (Bauer–Fike theorem based on $\kappa(\lambda_j)$). *Suppose $A \in \mathbb{C}^{N \times N}$ has N distinct eigenvalues. Then for all $\varepsilon > 0$,*

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

The above two theorems give us upper bounds on the pseudospectra of A only when A has N distinct eigenvalues. These results can be generalized for matrices that do not have distinct eigenvalues. The following is proven in [Trefethen and Embree 2005, Section 52].

Theorem 2.11 (asymptotic formula for the resolvent norm). *Let $\lambda_j \in \sigma(A)$ be an eigenvalue of A with k_j the size of the largest Jordan block associated to λ_j . For every $z \in \sigma_\varepsilon(A)$, for small enough ε ,*

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

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

We extend these results by providing lower bounds for arbitrary matrices, as well as explicit formulas for the ε -pseudospectra of 2×2 matrices.

3. Pseudospectra of 2×2 matrices

The following section presents a complete characterization of the ε -pseudospectra of all 2×2 matrices. We classify matrices by whether they are diagonalizable or nondiagonalizable and determine the ε -pseudospectra for each class. We begin with an explicit formula for computing the norm of a 2×2 matrix.

Let A be a 2×2 matrix with complex coefficients and let s_{\max} denote the largest singular value of A .

Then,

$$\|A\|^2 = s_{\max} = \frac{\text{Tr}(A^*A) + \sqrt{\text{Tr}(A^*A)^2 - 4 \det(A^*A)}}{2}. \tag{3-1}$$

Nondiagonalizable 2×2 matrices. Every nondiagonalizable 2×2 matrix must have exactly one eigenvalue of geometric multiplicity 1. In this case, we can Jordan-decompose the matrix and use the first definition of pseudospectra to show that $\sigma_\varepsilon(A)$ must be a perfect disk.

Proposition 3.1. *Let A be a nondiagonalizable 2×2 matrix, and let λ denote the eigenvalue of A . Write $A = V J V^{-1}$, where*

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}. \tag{3-2}$$

Given $\varepsilon > 0$,

$$\sigma_\varepsilon(A) = B(\lambda, |k|), \tag{3-3}$$

where

$$|k| = \sqrt{C\varepsilon + \varepsilon^2} \quad \text{and} \quad C = \frac{|a|^2 + |c|^2}{|ad - bc|}. \tag{3-4}$$

Proof. Let $z = \lambda + k$, where $k \in \mathbb{C}$. Then we have $(z - A)^{-1} = V(z - J)^{-1}V^{-1}$.

Taking the norm, this yields

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

From (3-1), we obtain that

$$\varepsilon^{-1} < \|(z - A)^{-1}\| = \frac{\sqrt{\text{Tr}(M^*M)} + \sqrt{\text{Tr}(M^*M)^2 - 4 \det(M^*M)}}{|k|^2 |ad - bc| \sqrt{2}}.$$

Note that this function depends only on $|k| = |z - \lambda|$; thus for every $\varepsilon > 0$, we have that $\sigma_\varepsilon(A)$ will be a disk. Solving for k to find the curve bounding the pseudospectrum, we obtain

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

Diagonalizable 2×2 matrices. Diagonalizable 2×2 matrices must have two distinct eigenvalues or be a multiple of the identity matrix. In either case, the pseudospectra can be described by the following proposition.

Proposition 3.2. *Let A be a 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 the boundary of $\sigma_\varepsilon(A)$ is the set of points z that satisfy the equation*

$$(\varepsilon^2 - |z - \lambda_1|^2)(\varepsilon^2 - |z - \lambda_2|^2) - \varepsilon^2 |\lambda_1 - \lambda_2|^2 \cot^2(\theta) = 0, \tag{3-5}$$

where θ is the angle between the two eigenvectors.

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

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{3-6}$$

Without loss of generality, let $z = \lambda_1 + k$.

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

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

where

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

Calculating $\text{Tr}(M^*M)$, we obtain

$$\text{Tr}(M^*M) = |r|^2(|\gamma + k|^2 + |k|^2 + |\gamma|^2 \cot^2 \theta), \tag{3-7}$$

where θ is the angle between the two eigenvectors, which are exactly the columns of V . For the determinant, we have

$$\det(M^*M) = |r|^4 |k|^2 |k + \gamma|^2. \tag{3-8}$$

Plugging the above into (3-1), we get

$$\varepsilon^{-1} = \|(z - A)^{-1}\| = \sqrt{\frac{\text{Tr}(M^*M) + \sqrt{\text{Tr}(M^*M)^2 - 4 \det(M^*M)}}{2|r|^2 |k(\gamma + k)|^2}}.$$

Rewriting and simplifying, we obtain the curve describing the boundary of the pseudospectrum:

$$(\varepsilon^2 - |k|^2)(\varepsilon^2 - |k + \gamma|^2) - \varepsilon^2 |\gamma|^2 \cot^2 \theta = 0. \quad \square$$

Note that for normal matrices, the eigenvectors are orthogonal. Therefore the equation above reduces to

$$(\varepsilon^2 - |k|^2)(\varepsilon^2 - |k + \gamma|^2) = 0, \tag{3-9}$$

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), we obtain

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

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

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

Proposition 3.3. *Let A be a diagonalizable 2×2 matrix with two distinct eigenvalues, λ_1, λ_2 . Then, $\sigma_\varepsilon(A) \upharpoonright_{\lambda_i}$ asymptotically tends toward a disk. In particular,*

$$\frac{r_{\max}(\lambda_i)}{r_{\min}(\lambda_i)} = 1 + \mathcal{O}(\varepsilon),$$

where $r_{\max}(\lambda_i), r_{\min}(\lambda_i)$ are the maximum and minimum distances from λ_i to $\partial\sigma_\varepsilon(A) \upharpoonright_{\lambda_i}$. Moreover, for A diagonalizable but not normal, $\sigma_\varepsilon(A) \upharpoonright_{\lambda_i}$ is never a perfect disk.

Proof. Let ε be small enough so that the ε -pseudospectrum is disconnected. Without loss of generality, we will consider $\sigma_\varepsilon(A) \upharpoonright_{\lambda_1}$.

Let $z_{\max} \in \partial\sigma_\varepsilon(A)$ such that $|z_{\max} - \lambda_1|$ is a maximum. Set $r_{\max}(\lambda) = |z_{\max} - \lambda|$. Consider the line joining λ_1 and λ_2 . Suppose for contradiction that z_{\max} does 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

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

Thus, from Proposition 3.2, we have that $z' \in \sigma_\varepsilon(A)$ 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 this line, where $z_{\min} \in \partial\sigma_\varepsilon(A)$ such that $r_{\min} = |z_{\min} - \lambda_1|$ 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 $y = |\lambda_2 - \lambda_1|$. The equation describing $r_{\max}(\lambda_1)$ becomes

$$(r_{\max}(\lambda_1)^2 - \varepsilon^2)((y - r_{\max}(\lambda_1))^2 - \varepsilon^2) = \varepsilon^2 y^2 \cot^2 \theta.$$

Similarly, we can obtain the equation for $r_{\min}(\lambda_1)$. Solving for $r_{\max}(\lambda_1)$ and $r_{\min}(\lambda_1)$, we get

$$r_{\max}(\lambda_1) = \frac{1}{2}(y - \sqrt{y^2 + 4\varepsilon^2 - 4y\varepsilon \csc \theta}), \quad (3-10)$$

$$r_{\min}(\lambda_1) = \frac{1}{2}(\sqrt{y^2 + 4\varepsilon^2 + 4y\varepsilon \csc \theta} - y). \quad (3-11)$$

For ε small, we can use the approximation $(1 + \varepsilon)^p = 1 + p\varepsilon + \mathcal{O}(\varepsilon^2)$. Then,

$$\begin{aligned} \frac{r_{\max}(\lambda_1)}{r_{\min}(\lambda_1)} &= \frac{1 - \sqrt{1 + 4(\varepsilon/y)^2 - 4(\varepsilon/y) \csc \theta}}{\sqrt{1 + 4(\varepsilon/y)^2 + 4(\varepsilon/y) \csc \theta} - 1} \\ &= \frac{1 + \eta\varepsilon + \mathcal{O}(\varepsilon^2)}{1 - \eta\varepsilon + \mathcal{O}(\varepsilon^2)}, \end{aligned} \quad (3-12)$$

where $\eta = (\cos \theta \cot \theta)/y$. Using the geometric series approximation $1/(1 - x) = 1 + x + \mathcal{O}(x^2)$, we find that

$$\frac{r_{\max}}{r_{\min}} = 1 + \frac{(2 \cos \theta \cot \theta)\varepsilon}{y} + \mathcal{O}(\varepsilon^2). \quad (3-13)$$

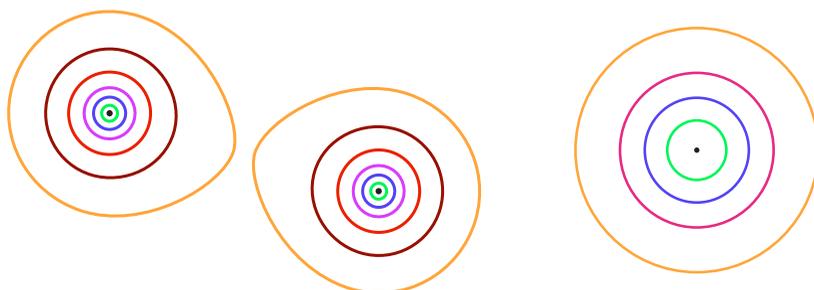


Figure 4. The ε -pseudospectra of a diagonalizable 2×2 matrix.

Thus, $\sigma_\varepsilon(A)$ 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 therefore $\cos \theta \cot \theta \neq 0$ and $(r_{\max}(\lambda))/r_{\min}(\lambda) \neq 1$. \square

This result can be observed by looking at plots of the pseudospectra of diagonalizable 2×2 matrices.

The image on the left in Figure 4 shows the pseudospectra of a particular 2×2 matrix. One can see that for large enough values of ε , the pseudospectra around either eigenvalue are not perfect disks. The image on the right is the pseudospectra of the same matrix (restricted to one eigenvalue), with smaller values of epsilon. Here, the pseudospectra appear to converge to disks. We find that this result holds in general for every $N \times N$ matrix and this is proven in the following section.

4. Asymptotic union of disks theorem

In Propositions 3.1 and 3.2, we showed that the ε -pseudospectra 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 ε -pseudospectra exhibit pathological properties for large ε ; for example, the nondiagonalizable matrix given in Figure 5 has, for larger ε , an ε -pseudospectrum that is not convex and not simply connected.

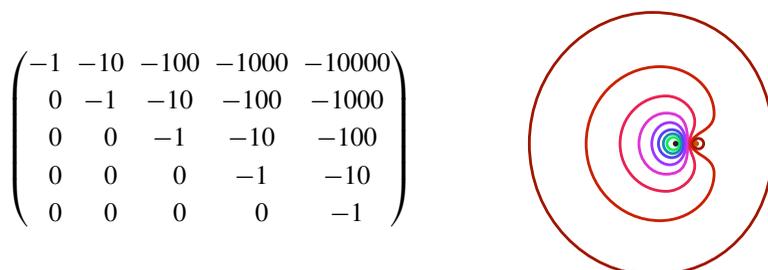


Figure 5. A Toeplitz matrix and its pseudospectra.

Thus, pseudospectra may behave poorly for large enough ε ; however, in the limit as $\varepsilon \rightarrow 0$, these properties disappear and the pseudospectra behave as disks centered around the eigenvalues with well-understood radii. In order to understand this asymptotic behavior, we use the following set-up (which follows [Moro et al. 1997]).

Let $A \in \mathbb{C}^{N \times N}$ and fix $\lambda \in \sigma(A)$. Write the Jordan decomposition of A as

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

where J consists of Jordan blocks J_1, \dots, J_m corresponding to the eigenvalue λ , and \hat{J} consists of Jordan blocks corresponding to 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 corresponding to λ of size $n \times n$. Arrange the Jordan blocks in J in weakly decreasing order, according to size. That is,

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

where J_1, \dots, J_ℓ are $n \times n$.

Further partition P ,

$$P = (P_1 \ \dots \ P_\ell \ \dots \ P_m),$$

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

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

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

The following theorem is presented by Moro, Burke, and Overton [Moro et al. 1997] and due to Lidskiĭ [1966].

Theorem 4.1 [Lidskiĭ 1966]. *Given ℓ, n as defined above corresponding to the matrix A , there are ℓn eigenvalues of the perturbed matrix $A + \varepsilon E$ admitting a first-order expansion*

$$\lambda_{j,k}(\varepsilon) = \lambda + (\gamma_j \varepsilon)^{1/n} + o(\varepsilon^{1/n})$$

for $j = 1, \dots, \ell$, and $k = 1, \dots, n$, where γ_j are the eigenvalues of YEX and the different values of $\lambda_{j,k}(\varepsilon)$ for $k = 1, \dots, n$ are defined by taking the distinct n -th roots of γ_j .

Lidskii’s result can be interpreted in terms of the ε -pseudospectrum of a matrix A in order to understand the radii of $\sigma_\varepsilon(A)$ as $\varepsilon \rightarrow 0$.

Theorem 4.2. *Let $A \in \mathbb{C}^{N \times N}$. Let $\varepsilon > 0$. 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 by $\sigma_\varepsilon(A) \upharpoonright_\lambda$.*

Then, as $\varepsilon \rightarrow 0$,

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

where $C = \|XY\|$, with X, Y defined above, and n is the size of the largest Jordan block corresponding to λ .

Proof. Lower bound: Give $E \in \mathbb{C}^{N \times N}$, and let $\gamma_{\max}(E)$ be the largest eigenvalue of YEX . It is shown [Moro et al. 1997, Theorem 4.2] that

$$\alpha := \max_{\|E\| \leq 1} \gamma_{\max}(E) = \|XY\|.$$

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

Fix $E = vu$, with v, u defined above, fix $\theta \in [0, 2n\pi]$, and define $\tilde{E} = e^{i\theta} E$. Note that γ is an eigenvalue of YEX if and only if $e^{i\theta}\gamma$ is an eigenvalue of $Y\tilde{E}X$. Since α is an eigenvalue of E , we know that $e^{i\theta}\alpha$ is an eigenvalue of $Y\tilde{E}X$. Considering the perturbed matrix $A + \varepsilon\tilde{E}$, Theorem 4.1 implies that there is a perturbed eigenvalue $\lambda(\varepsilon)$ of the form

$$\lambda(\varepsilon) = \lambda + (e^{i\theta}\alpha\varepsilon)^{1/n} + o(\varepsilon^{1/n}),$$

and thus $\lambda(\varepsilon) \in \overline{\sigma_\varepsilon(A)} \upharpoonright_\lambda$. Ranging θ from 0 to $2n\pi$, we get the desired result.

Upper bound: Using the proof of [Trefethen and Embree 2005, Theorem 52.3], we know that asymptotically

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

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 an $n \times n$ matrix with a 1 in the top right entry and zeros elsewhere. We find

$$PD^{n-1} = (\Omega_1 \ \Omega_2 \ \dots \ \Omega_\ell \ 0),$$

where Ω_j is a matrix whose last column is x_j with zeros elsewhere. This then gives

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

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

We present special cases of matrices to explore the consequences of Theorem 4.2.

Special cases.

(1) λ is simple: 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 4.2 becomes

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

which matches with Theorem 2.9.

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

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

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

First, assume A is nondiagonalizable. In this case, A only has one eigenvalue, λ . Writing $A = VJV^{-1}$, where V and J are as defined in (3-2), we have that,

$$X = (a \ c)^T, \quad Y = \frac{1}{ad - bc} (-c \ a).$$

From Theorem 4.2, we then have that as $\varepsilon \rightarrow 0$,

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

This agrees asymptotically with (3-4); however, (3-4) gives an explicit formula for $\sigma_\varepsilon(A)$.

In the case where A is diagonalizable, A has two eigenvalues, λ_1 and λ_2 . Again, we write $A = DVD^{-1}$, where V and D are as defined in (3-6). From this, we have

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

Thus, as $\varepsilon \rightarrow 0$, we have from Theorem 4.2 that

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

So,

$$\frac{r_{\max}}{r_{\min}} = \frac{(\csc \theta)\varepsilon + o(\varepsilon)}{(\csc \theta)\varepsilon + o(\varepsilon)} = 1 + o(1).$$

This agrees with the ratio we obtain from the explicit formula for diagonalizable 2×2 matrices; however, (3-13) gives us more information on the $o(1)$ term.

(4) A is a Jordan block: From [Trefethen and Embree 2005, p. 470], we know that the ε -pseudospectrum of the Jordan block is exactly a disk about the eigenvalue of J of some radius. An explicit formula for the radius remains unknown; however, we can use Theorem 4.2 to find the asymptotic behavior.

Proposition 4.3 (asymptotic Bound). *Let J be an $N \times N$ Jordan block. Then*

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

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 4.2, we find $C = \|XY\| = \|v_j u_j\| = 1$. Thus,

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

By a simple computation, we can also get a better explicit lower bound on the ε -pseudospectra of an $N \times N$ Jordan block that agrees with our asymptotic bound.

Proposition 4.4. *Let J be an $N \times N$ Jordan block. Then,*

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

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 $J + E$:

$$\begin{aligned} 0 &= \det(J + E - zI) \\ &= \det \begin{pmatrix} \lambda - z & k + 1 & & & \\ & \lambda - z & k + 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & k + 1 \\ k & & & & \lambda - z \end{pmatrix} \\ &= (\lambda - z)^N + (-1)^{N-1} k(1 + k)^{N-1} \\ &= (-1)^{N-1} ((z - \lambda)^N + k(1 + k)^{N-1}), \end{aligned}$$

the right and left eigenvectors, respectively. We will give the explicit expressions for v_j and u_j .

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

The cases will be presented as follows:

- **Case 1:** Let A be a $kn \times kn$ matrix with a_1, \dots, a_k distinct.
- **Case 2:** Let A be an $N \times N$ matrix with a_1, \dots, a_k distinct.
- **General case:** Let A be an $N \times N$ matrix with a_1, \dots, a_k not distinct.

To shorten notation for the rest of this section, we define

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

Case 1: The size of A is $kn \times kn$ and the a_i are distinct.

We write the elements of the superdiagonal as b_1, b_2, \dots, b_{N-1} , and we let $p = k(n - 1) + j$.

We have that

$$v_j = \begin{pmatrix} \frac{b_1 \cdots b_{j-1}}{f(a_j - a_1) \cdots f(a_j - a_{j-1})} \\ \frac{b_2 \cdots b_{j-1}}{f(a_j - a_2) \cdots f(a_j - a_{j-1})} \\ \vdots \\ \frac{b_{j-1}}{f(a_j - a_{j-1})} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad u_j^* = \frac{1}{(f(a_1 - a_j) \cdots f(a_k - a_j))^{n-1}} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \frac{b_j \cdots b_p}{f(a_{j+1} - a_j)} \\ \vdots \\ \frac{b_j \cdots b_{N-2}}{f(a_{j+1} - a_j) \cdots f(a_{k-1} - a_j)} \\ \frac{b_j \cdots b_{N-1}}{f(a_{j+1} - a_j) \cdots f(a_k - a_j)} \end{pmatrix}^T.$$

Direct computation will show that these are indeed left and right eigenvectors associated with each eigenvalue a_j .

Case 2: The size of A is $N \times N$, and the a_i are distinct.

We relax 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 $0 < r \leq k$. 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

$$v_j = \left(\frac{b_1 \cdots b_{j-1}}{f(a_j - a_1) \cdots f(a_j - a_{j-1})} \quad \frac{b_2 \cdots b_{j-1}}{f(a_j - a_2) \cdots f(a_j - a_{j-1})} \quad \cdots \quad \frac{b_{j-1}}{f(a_j - a_{j-1})} \quad 1 \quad 0 \quad \cdots \quad 0 \right)^T.$$

We split up the formula for the left eigenvectors into two cases: (i) $1 \leq j \leq r$ and (ii) $r < j \leq k$.

(i) $1 \leq j \leq r$. 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 . When $1 \leq j \leq r$, we have that a_j is in this last partial block. In this case, let $p = kn + j$.

We have that

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

where

$$\mu_j = \frac{f(a_1 - a_j) \cdots f(a_{j-1} - a_j)}{(f(a_1 - a_j) \cdots f(a_k - a_j))^n f(a_1 - a_j) \cdots f(a_r - a_j)}.$$

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

We have that

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

where

$$\mu_j = \frac{f(a_1 - a_j) \cdots f(a_{j-1} - a_j)}{(f(a_1 - a_j) \cdots f(a_k - a_j))^n f(a_1 - a_j) \cdots f(a_r - a_j)}.$$

Case 3: General case. The size of A is $N \times N$ and the a_i are not distinct for $1 \leq i \leq k$.

Let A be an $N \times N$ periodic bidiagonal matrix with period k on the main diagonal. Let n, r be such that $N = kn + r$, where $0 < r \leq k$. Write a_1, \dots, a_k for the entries on the main diagonal (the a_i 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 each 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 2. That is,

$$v_\ell = \left(\frac{b_1 \cdots b_{\ell-1}}{f(a_\ell - a_1) \cdots f(a_\ell - a_{\ell-1})} \quad \frac{b_2 \cdots b_{\ell-1}}{f(a_\ell - a_2) \cdots f(a_\ell - a_{\ell-1})} \quad \cdots \quad \frac{b_{\ell-1}}{f(a_\ell - a_{\ell-1})} \quad 1 \quad 0 \quad \cdots \quad 0 \right)^T.$$

The corresponding left eigenvector for α depends on the first and last positions of α . Let $k(n - 1) = \ell q + s$ and set $q \equiv m \pmod k$. We split up the formula for the left eigenvector of α into two cases, which again mirror the formulas given in Case 2:

(i) $1 \leq \ell \leq r$ and (ii) $r < \ell \leq k$.

For both of these two cases, we define

$$g(b_i) = \begin{cases} b_i & i \geq p, \\ 1 & i < p. \end{cases}$$

(i) $1 \leq \ell \leq r$. In this case, α appears in the partial block. Let $p = kn + \ell$. We have

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

where

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

(ii) $r < \ell \leq k$. In this case, α is in the last complete block. Here, we let $p = k(n - 1) + \ell$. Now, we have

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

where

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

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

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

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

Note. 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, \dots, A_n . To find the ε -pseudospectrum of A , apply the same analysis to these smaller matrices, and from Theorem 2.5, we have that

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

6. 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 an 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 of ε -pseudospectra are bounded by powers of epsilon. The following theorem makes this precise.

Theorem 6.1. *Let V be a Hilbert space and $A : V \rightarrow V$ a finite-rank operator on V . Then there exists C such that for sufficiently small ε ,*

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

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

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

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 every $\varepsilon > 0$,

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

The ε -pseudospectrum of the zero operator is well-understood since this operator is normal; for all ε , it is precisely the ball of radius ε . It thus suffices to consider the ε -pseudospectrum of the finite-rank operator $A' : U \rightarrow U$, 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

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

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 Theorem 4.2, we then know that $\sigma_\varepsilon(A)$ is contained, for small enough ε , in the set $\sigma(A) + C\varepsilon^{1/(m+1)}$.

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

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

whose pseudospectrum will contain the ball of radius $\varepsilon^{1/(m+1)}$ by Proposition 4.4. \square

Open Questions. The natural question to ask now is whether we can extend this result to more arbitrary operators on Hilbert spaces. In particular, for a bounded operator A , we would like to establish if there exists a continuous function $r_A(\varepsilon)$ such that for sufficiently small ε ,

$$\sigma_\varepsilon(A) \subseteq \sigma(A) + B(0, r_A(\varepsilon)).$$

For a matrix A , we proved in Theorem 4.2 that $r_A(\varepsilon) = C\varepsilon^{1/n}$, where n is the size of the largest Jordan block associated to A , and C is a constant that depends on

the left and right eigenvectors associated to a certain eigenvalue. For a finite-rank operator A , we proved in Theorem 6.1 that $r_A(\varepsilon) = C\varepsilon^{1/(m+1)}$, where m is the rank of the operator and C is as above.

For closed but not necessarily bounded operators, the picture is more complex, as the spectrum need not be bounded or even nonempty. For example, the operator $A : u \mapsto u'$ in $L^2[0, 1]$ with domain $D(A)$ being the set of absolutely continuous functions on $[0, 1]$ satisfying $u(1) = 0$ has empty spectrum. When $D(A)$ is the entire space, the spectrum of A is the entire complex plane. Davies [1999a] also provides an example of an unbounded operator with unbounded pseudospectrum.

Given these examples, we can see that Theorem 6.1 will not generalize to unbounded operators, as the pseudospectrum of an unbounded operator may be unbounded for all ε .

Nonetheless, we do still have a certain convergence of the ε -pseudospectrum to the spectrum [Trefethen and Embree 2005, Section 4], namely $\bigcap_{\varepsilon>0} \sigma_\varepsilon(A) = \sigma(A)$. Also, while the ε -pseudospectrum may be unbounded, each bounded component of it necessarily contains a component of the spectrum. These results imply that the bounded components of the ε -pseudospectrum must converge to the spectrum. Therefore, if we restrict our attention to these bounded components, we can attempt to generalize Theorems 4.2 and 6.1 by asking whether the *bounded* components of $\sigma_\varepsilon(A)$ converge to the spectrum as a union of disks.

Acknowledgements

Support for this project was provided by the National Science Foundation REU Grant DMS-0850577 and DMS-1347804, the Clare Boothe Luce Program of the Henry Luce Foundation, and the SMALL REU at Williams College.

References

- [Bauer and Fike 1960] F. L. Bauer and C. T. Fike, “Norms and exclusion theorems”, *Numer. Math.* **2** (1960), 137–141. MR 0118729 Zbl 0101.25503
- [Baumgärtel 1985] H. Baumgärtel, *Analytic perturbation theory for matrices and operators*, Operator Theory: Advances and Applications **15**, Birkhäuser, Basel, 1985. MR 878974 Zbl 0591.47013
- [Böttcher 2006] A. Böttcher, “Review of “Spectra and pseudospectra: The behavior of nonnormal matrices and operators” by L. N. Trefethen and M. Embree”, *Linear Algebra Appl.* **416**:2-3 (2006), 1098–1101.
- [Davies 1999a] E. B. Davies, “Pseudo-spectra, the harmonic oscillator and complex resonances”, *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.* **455**:1982 (1999), 585–599. MR 1700903 Zbl 0931.70016
- [Davies 1999b] E. B. Davies, “Semi-classical states for non-self-adjoint Schrödinger operators”, *Comm. Math. Phys.* **200**:1 (1999), 35–41. MR 1671904 Zbl 0921.47060
- [Kato 1995] T. Kato, *Perturbation theory for linear operators*, Springer, Berlin, 1995. MR 1335452 Zbl 0836.47009

[Lidskii 1966] V. B. Lidskii, “On the theory of perturbations of nonselfadjoint operators”, *Ž. Vyčisl. Mat. i Mat. Fiz.* **6**:1 (1966), 52–60. In Russian; translated in *{USSR} Computational Mathematics and Mathematical Physics* **6**:1 (1966), 73–85. MR 0196930 Zbl 0166.40501

[Moro et al. 1997] J. Moro, J. V. Burke, and M. L. Overton, “On the Lidskii–Vishik–Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure”, *SIAM J. Matrix Anal. Appl.* **18**:4 (1997), 793–817. MR 1471994 Zbl 0889.15016

[Trefethen and Embree 2005] L. N. Trefethen and M. Embree, *Spectra and pseudospectra: The behavior of nonnormal matrices and operators*, Princeton Univ. Press, 2005. MR 2155029 Zbl 1085.15009

Received: 2015-05-25

Revised: 2015-06-22

Accepted: 2015-06-23

fgong@mit.edu

*Department of Mathematics and Statistics, Williams College,
Williamstown, MA 01267, United States*

olivia.s.meyerson@gmail.com

*Department of Mathematics and Statistics, Williams College,
Williamstown, MA 01267, United States*

jdmeza@alumni.cmu.edu

*Department of Mathematics, Carnegie Mellon University,
Pittsburgh, PA 15213, United States*

mstoiciu@williams.edu

*Department of Mathematics and Statistics, Williams College,
Williamstown, MA 01267, United States*

abigailward@uchicago.edu

*Department of Mathematics, The University of Chicago,
Chicago, IL 60637, United States*

involve

msp.org/involve

INVOLVE YOUR STUDENTS IN RESEARCH

Involve showcases and encourages high-quality mathematical research involving students from all academic levels. The editorial board consists of mathematical scientists committed to nurturing student participation in research. Bridging the gap between the extremes of purely undergraduate research journals and mainstream research journals, *Involve* provides a venue to mathematicians wishing to encourage the creative involvement of students.

MANAGING EDITOR

Kenneth S. Berenhaut Wake Forest University, USA

BOARD OF EDITORS

Colin Adams	Williams College, USA	Suzanne Lenhart	University of Tennessee, USA
John V. Baxley	Wake Forest University, NC, USA	Chi-Kwong Li	College of William and Mary, USA
Arthur T. Benjamin	Harvey Mudd College, USA	Robert B. Lund	Clemson University, USA
Martin Bohner	Missouri U of Science and Technology, USA	Gaven J. Martin	Massey University, New Zealand
Nigel Boston	University of Wisconsin, USA	Mary Meyer	Colorado State University, USA
Amarjit S. Budhiraja	U of North Carolina, Chapel Hill, USA	Emil Minchev	Ruse, Bulgaria
Pietro Cerone	La Trobe University, Australia	Frank Morgan	Williams College, USA
Scott Chapman	Sam Houston State University, USA	Mohammad Sal Moslehian	Ferdowsi University of Mashhad, Iran
Joshua N. Cooper	University of South Carolina, USA	Zuhair Nashed	University of Central Florida, USA
Jem N. Corcoran	University of Colorado, USA	Ken Ono	Emory University, USA
Toka Diagana	Howard University, USA	Timothy E. O'Brien	Loyola University Chicago, USA
Michael Dorff	Brigham Young University, USA	Joseph O'Rourke	Smith College, USA
Sever S. Dragomir	Victoria University, Australia	Yuval Peres	Microsoft Research, USA
Behrouz Emamizadeh	The Petroleum Institute, UAE	Y.-F. S. Pétermann	Université de Genève, Switzerland
Joel Foisy	SUNY Potsdam, USA	Robert J. Plemmons	Wake Forest University, USA
Errin W. Fulp	Wake Forest University, USA	Carl B. Pomerance	Dartmouth College, USA
Joseph Gallian	University of Minnesota Duluth, USA	Vadim Ponomarenko	San Diego State University, USA
Stephan R. Garcia	Pomona College, USA	Bjorn Poonen	UC Berkeley, USA
Anant Godbole	East Tennessee State University, USA	James Propp	U Mass Lowell, USA
Ron Gould	Emory University, USA	József H. Przytycki	George Washington University, USA
Andrew Granville	Université Montréal, Canada	Richard Rebarber	University of Nebraska, USA
Jerold Griggs	University of South Carolina, USA	Robert W. Robinson	University of Georgia, USA
Sat Gupta	U of North Carolina, Greensboro, USA	Filip Saidak	U of North Carolina, Greensboro, USA
Jim Haglund	University of Pennsylvania, USA	James A. Sellers	Penn State University, USA
Johnny Henderson	Baylor University, USA	Andrew J. Sterge	Honorary Editor
Jim Hoste	Pitzer College, USA	Ann Trenk	Wellesley College, USA
Natalia Hritonenko	Prairie View A&M University, USA	Ravi Vakil	Stanford University, USA
Glenn H. Hurlbert	Arizona State University, USA	Antonia Vecchio	Consiglio Nazionale delle Ricerche, Italy
Charles R. Johnson	College of William and Mary, USA	Ram U. Verma	University of Toledo, USA
K. B. Kulasekera	Clemson University, USA	John C. Wierman	Johns Hopkins University, USA
Gerry Ladas	University of Rhode Island, USA	Michael E. Zieve	University of Michigan, USA

PRODUCTION

Silvio Levy, Scientific Editor

Cover: Alex Scorpan

See inside back cover or msp.org/involve for submission instructions. The subscription price for 2016 is US \$160/year for the electronic version, and \$215/year (+\$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Involve (ISSN 1944-4184 electronic, 1944-4176 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

Involve peer review and production are managed by EditFLOW[®] from Mathematical Sciences Publishers.

PUBLISHED BY

 **mathematical sciences publishers**
nonprofit scientific publishing

<http://msp.org/>

© 2016 Mathematical Sciences Publishers

involve

2016

vol. 9

no. 3

A combinatorial proof of a decomposition property of reduced residue systems	361
YOTSANAN MEEMARK AND THANAKORN PRINYASART	
Strong depth and quasigeodesics in finitely generated groups	367
BRIAN GAPINSKI, MATTHEW HORAK AND TYLER WEBER	
Generalized factorization in $\mathbb{Z}/m\mathbb{Z}$	379
AUSTIN MAHLUM AND CHRISTOPHER PARK MOONEY	
Cocircular relative equilibria of four vortices	395
JONATHAN GOMEZ, ALEXANDER GUTIERREZ, JOHN LITTLE, ROBERTO PELAYO AND JESSE ROBERT	
On weak lattice point visibility	411
NEIL R. NICHOLSON AND REBECCA RACHAN	
Connectivity of the zero-divisor graph for finite rings	415
REZA AKHTAR AND LUCAS LEE	
Enumeration of m -endomorphisms	423
LOUIS RUBIN AND BRIAN RUSHTON	
Quantum Schubert polynomials for the G_2 flag manifold	437
RACHEL E. ELLIOTT, MARK E. LEWERS AND LEONARDO C. MIHALCEA	
The irreducibility of polynomials related to a question of Schur	453
LENNY JONES AND ALICIA LAMARCHE	
Oscillation of solutions to nonlinear first-order delay differential equations	465
JAMES P. DIX AND JULIO G. DIX	
A variational approach to a generalized elastica problem	483
C. ALEX SAFSTEN AND LOGAN C. TATHAM	
When is a subgroup of a ring an ideal?	503
SUNIL K. CHEBOLU AND CHRISTINA L. HENRY	
Explicit bounds for the pseudospectra of various classes of matrices and operators	517
FEIXUE GONG, OLIVIA MEYERSON, JEREMY MEZA, MIHAI STOICIU AND ABIGAIL WARD	

