SPECTRAL ANALYSIS OF NON-HERMITIAN MATRICES

MATHEMATICAL PHYSICS 2010
MATTHEW COUDRON, AMALIA CULIUC, PHILIP VU,
STEPHEN WEBSTER

Abstract. Motivated by work of Contedini-Embree-Trefethen and Goldsheid-Khoruzhenko, we investigate the spectral properties of certain classes of non-Hermitian matrices. We give parametrizations for curves in the plane that contain the spectrum of bi-diagonal matrices with periodic diagonal entries. In the case of period two, we find an asymptotic formula for the spacing between these eigenvalues.

We also study the pseudospectrum $\sigma_\varepsilon(A)$ of a general square matrix A. We generalize the Bauer–Fike Theorem and give lower and upper bounds to show that the asymptotic decay (as $\varepsilon \to 0$) of the diameter of $\sigma_\varepsilon(A)$ near the eigenvalue λ is of order $\varepsilon^{1/k}$, where k is the dimension of the largest Jordan block associated to λ.

1. Introduction

The spectral properties of non-Hermitian operators have been widely studied during the recent years, mainly due to their physical applications [8], [12], [3]. In [9], Hatano and Nelson studied the eigenvalues of a non-Hermitian Hamiltonian to describe the vortex pinning phenomenon in superconductors. This operator has the following discrete form:

$$H_n^g = \begin{pmatrix} v_1 & -e^g & \dots & 0 & -e^{-g} \\ -e^{-g} & v_2 & -e^g & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \vdots & \ddots & \ddots & -e^g \\ -e^g & 0 & \dots & -e^{-g} & v_n \end{pmatrix},$$

where n is the dimension of the matrix, the parameter $g > 0$ is a measure of the strength of the transverse magnetic field in the superconductor, and the values v_i are real numbers representing the potential of the system. A numerical analysis of the matrix H_n^g showed that its spectrum lies along smooth curves in the plane, depending on the value of g. In [7] and [6], Goldsheid and Khoruzhenko proved this result analytically, giving a precise description of the shape of the curves as a function of g. In [2], Contedini, Embree, and Trefethen considered the limit case for H_g^n, a bidiagonal matrix of the form

$$\begin{pmatrix} v_1 & 1 & \dots & 0 & 0 \\ 0 & v_2 & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \vdots & \ddots & \ddots & 1 \\ 1 & 0 & \dots & 0 & v_n \end{pmatrix},$$

where the potential sequence $\{v_i\}_{i=1}^n$ is drawn randomly from a uniform distribution. In our analysis, we begin by considering the spectra of bidiagonal matrices with various potentials. In Section 2 we study periodic potentials of period q defined as

$$v_i = \begin{cases} v & \text{if } i \equiv 1 \mod q \\ 0 & \text{otherwise} \end{cases}$$

We describe the eigenvalue curves corresponding to given periods and for periods 2 and 3, we give parametrizations for these curves. For period $q = 2$, we also derive an expression for the spacing between neighboring eigenvalues. We find that these eigenvalues are evenly spaced and their spacing depends on the size of the matrix. Fixing $v = 1$ and letting $q \to \infty$, we observe that the shape of the eigenvalue curves transitions from an oval to a circle. We present numerical evidence for this conjecture. In Section 3 we investigate the spectra of a class of matrices called Altered-Diagonal Circulant Matrices, each of which is constructed by adding a circulant matrix and a diagonal matrix.

Finally, in Section 4 we study the ϵ-pseudospectrum of general matrices. In the case of normal matrices, as a consequence of the Spectral Theorem, the ϵ-pseudospectra are ϵ-disks around the spectrum. More generally, for diagonalizable matrices, the Bauer-Fike Theorem describes the shape of the ϵ-pseudospectrum and gives upper and lower bounds. Motivated by an observation of Chatelin and Braconnier [1], we extend this result to all square matrices using the Jordan Canonical Form. The new bounds we obtain depend on the condition number of the similarity matrix and the dimension of the largest Jordan block of distinct eigenvalues.

Acknowledgements: We thank the NSF, Williams College and Mount Holyoke College for support.

2. Eigenvalue Curves for Periodic Bidiagonal Matrices

In what follows, we consider $N \times N$ matrices A similar to those studied by Embree, Contedini, and Trefethen. Such matrices have a periodic structure along the diagonal, constants along the super diagonal, and one entry in the bottom left corner. The general form is

$$A_{ij} = \begin{cases} v_i & \text{if } i = j \\ 1 & \text{if } i = j+1 \text{ and for } A_{N1} \\ 0 & \text{otherwise} \end{cases}$$

where $\{v_i\}$ is the periodic sequence $\{v, 0, 0, \cdots 0, v, 0, \cdots, v, 0, \cdots, 0\}$ for $v \in \mathbb{C}$. In Proposition 2.1 we give the general form for the characteristic polynomial of such matrices.

Proposition 2.1. *Let H be an $N \times N$ matrix of the following form:*

$$H = \begin{pmatrix} v_n & \lambda_n & 0 & 0 & \dots & 0 \\ 0 & v_{n-1} & \lambda_{n-1} & 0 & \dots & 0 \\ 0 & 0 & v_{n-2} & \lambda_{n-2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \lambda_1 & 0 & \dots & & & v_1 \end{pmatrix}$$

where $v_j, \lambda_j \in \mathbb{C}$, for all $1 \leq j \leq n$. The the characteristic polynomial of H is

$$p(H) = \prod_{j=1}^{n}(v_j - z) - (-1)^n \prod_{j=1}^{n} \lambda_j$$

Proof. To compute

$$p(H) = \begin{vmatrix} v_n & \lambda_n & 0 & 0 & \dots & 0 \\ 0 & v_{n-1} & \lambda_{n-1} & 0 & \dots & 0 \\ 0 & 0 & v_{n-2} & \lambda_{n-2} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \lambda_1 & 0 & \dots & & & v_1 \end{vmatrix}$$

we expand along the first column:

$$P_n = (v_n - z)\begin{vmatrix} v_{n-1}-z & \lambda_{n-1} & 0 & \dots & 0 \\ 0 & v_{n-2}-z & \lambda_{n-2} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & & & v_1 - z \end{vmatrix} +$$

$$(-1)^{n+1}\lambda_1 \begin{vmatrix} \lambda_n & 0 & & 0 \\ v_{n-1}-z & \lambda_{n-1} & & \\ & \ddots & \ddots & \\ 0 & & v_2 - z & \lambda_2 \end{vmatrix}$$

Then, by the properties of triangular matrices,

$$p(H) = \prod_{j=1}^{n}(v_j - z) - (-1)^n \prod_{j=1}^{n}\lambda_j$$

□

Remarks:

(1) For the particular case $\lambda_1 = \lambda_2 = \dots \lambda_n = \lambda$, we have that

$$p(H) = \prod_{j=1}^{n}(v_j - z) - (-\lambda)^n$$

(2) Due to the commutativity of multiplication, the characteristic polynomial $p(H)$ does not depend on the indexing of the set $\{v_j\}_j$. Furthermore, $p(H)$ does not depend on the values λ_j, but only on their product.

For a fixed period size q, we describe the curves along which the spectrum of A lies. We begin with the case when the period is 2. In order to have an integer number of periods, let N be even. Then the matrix A has the form:

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

Proposition 2.2. *Let A be the $N \times N$ matrix above such that $N = 2m$. Writing v^2 as $r_0 e^{i\phi}$ for some $\phi \in [0, 2\pi)$, $r_0 \in \mathbb{R}^+ \cup \{0\}$, define*

$$r(\theta) = r_0 \cos(\theta - \phi) + \sqrt{16 - r_0^2 \sin^2(\theta - \phi)}$$

for $\theta \in S = \{\theta || \sin(\theta - \phi)| \leq \frac{4}{r_0}, \theta \in [0, 2\pi)\}$ *and*

$$f_{\pm}(\theta) = \frac{1 \pm \sqrt{r(\theta)} e^{\frac{i\theta}{2}}}{2}.$$

Then, for $\theta \in S$,

$$\sigma(A) \subseteq Ran(f_{\pm}(\theta)).$$

Proof. By Proposition 2.1, the characteristic polynomial of A,

$$\Phi_A(z) = z^m (z - v)^m - 1$$

has roots of the form

$$z = \frac{1 \pm \sqrt{v^2 + 4e^{\frac{i2\pi k}{m}}}}{2}$$

where $k = 0, \ldots, m-1$, since solving for the roots of $\Phi_A(z)$ amounts to solving a quadratic equation for each m^{th} root of unity.
We proceed to show that the eigenvalues lie on the two curves described above. For all possible values of k, the points $v^2 + 4e^{\frac{i2\pi k}{m}}$ lie along a circle in the complex plane. Then the eigenvalues of A are on the graph of

$$z(t) = \frac{1 \pm \sqrt{v^2 + 4e^{\frac{it}{m}}}}{2}$$

We now rewrite the function $z(t)$ as a function of the form $r(\theta)e^{i\theta}$. Note that, for all $\theta \in \{\theta || \sin(\theta - \phi)| \leq \frac{a}{r_0}, \theta \in [0, 2\pi)\}$,

$$r(\theta) = r_0 \cos(\theta - \phi) + \sqrt{a^2 - r_0^2 \sin^2(\theta - \phi)}$$

is the function defining a circle of radius a centered at the point (r_0, ϕ), in polar coordinates; since we are shifting a circle by v^2, we choose r_0 to be the length of v^2 considered as a vector on $\mathbb{R}^2$ and ϕ as the argument of v^2. Thus, the spectrum of A lies along the curves

$$\frac{1 \pm \sqrt{r(\theta)} e^{\frac{i\theta}{2}}}{2}$$

in the complex plane.

□

We also calculate the spacing of the eigenvalues, that is, the distance between two adjacent eigenvalues along the curve. As there are in fact two curves due to the $\pm$ sign in $f_{\pm}(x)$, we only just consider adjacent eigenvalues along their corresponding curves. We have already seen that the eigenvalues lie on a circle that is shifted, and then altered by composition with taking a square root. Thus, let us consider the points along the circle in the complex plane and apply the square root function to them. Let $\bar{p} \in \mathbb{Z}$ and consider $p = \bar{p} \mod m$ and $q = \bar{p}+1 \mod m$. Take a branch cut along the ray $2\pi r$ where $r \in \mathbb{I} \cap (0,1)$, that is, r is irrational; this will ensure that none of the eigenvalues will be undefined by our choice of the branch cut. Then label the points $\frac{2\pi k}{m}$ along the circle $r(\theta)$ counterclockwise with the branch cut. Let $u = \frac{2\pi p}{m}$ and $w = \frac{2\pi q}{m}$. Thus, ta! king the square root, scaling, and shifting u and w results in two arbitrary neighboring eigenvalues on either the curve $f_+(\theta)$ or $f_-(\theta)$.

Proposition 2.3. *The distance between two neighboring eigenvalues on either the curve $f_+(\theta)$ or $f_-(\theta)$ is*

$$\left| \frac{2\pi}{m\sqrt{v+4e^{iu}}} + O\left(\frac{1}{m^2}\right) \right|.$$

Proof. The distance between the points labeled u and w on the circle $4e^{i\theta}$ is

$$4\tan\left(\frac{i2\pi}{m}\right) = \frac{8\pi}{m} + O\left(\frac{1}{m^3}\right).$$

We then apply $g(x) = \frac{1\pm\sqrt{x+v^2}}{2}$ for $\bar{u} = 4e^u$ and $\bar{w} = 4e^w$ to obtain our neighboring eigenvalues. We see that $g(x)$ is an analytic function, and hence, the Taylor series expansion gives

$$\begin{aligned} g(\bar{u}) - g(\bar{w}) &= (u-w)\left(g'(\bar{u}) + O\left(\frac{1}{m}\right)\right) \\ &= \left(\frac{8\pi}{m} + O\left(\frac{1}{m^3}\right)\right)\left(g'(\bar{u}) + O\left(\frac{1}{m}\right)\right) \\ &= \left(\frac{8\pi}{m} + O\left(\frac{1}{m^3}\right)\right)\left(\pm\frac{1}{4\sqrt{4e^{ui}+v}} + O\left(\frac{1}{m}\right)\right) \\ &= \frac{\pm 2\pi}{m\sqrt{v+4e^{iu}}} + O\left(\frac{1}{m^2}\right) \end{aligned}$$

Taking the absolute value of the expression completes the proof. □

We include some plots of the spectrum of a matrix with period 2 and the sequence $\{1,0,1,0,\cdots,0\}$ on the diagonal. Figure 1 plots the eigenvalues of a 10×10 such matrix. The numerics show that the spectrum lies on an ellipse-like curve. In Figure 2, we plot the spectrum

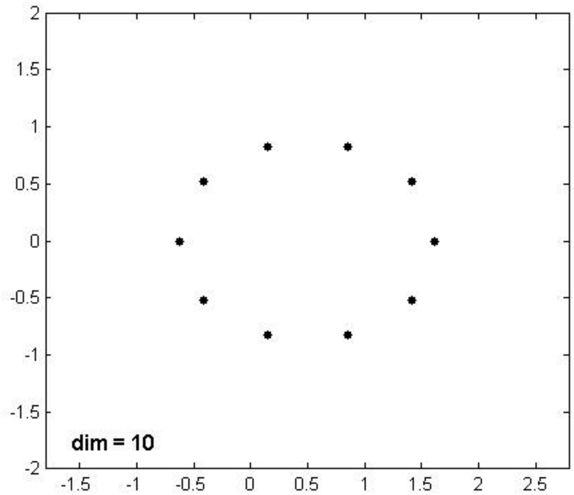

FIGURE 1. *The spectrum of a* 10 *by* 10 *matrix with period 2 diagonal.*

for $N = 500$ and the same periodic structure on the diagonal.

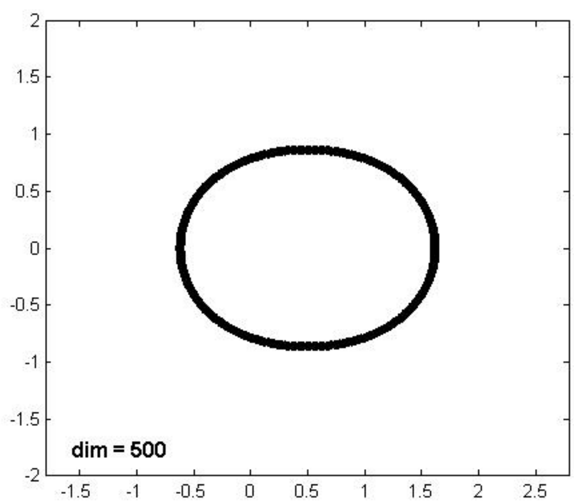

FIGURE 2. *The spectrum of a* 500 *by* 500 *matrix.*

In Figure 3 we plot the proposed curve, $f_{\pm}(\theta)$, along with the eigenvalues of the $N = 10$ matrix, verifying that the eigenvalues are indeed contained in the range of $f_{\pm}(\theta)$.

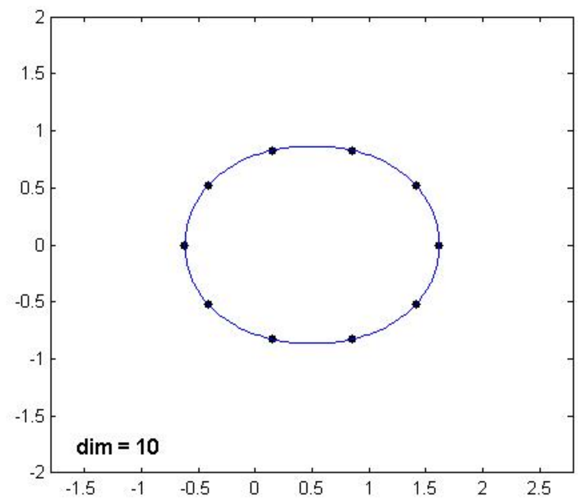

FIGURE 3. *The spectrum of a* 10 *by* 10 *matrix plotted with the function* $f_{\pm}(\theta)$.

We proceed by exploring the case of higher periods for the periodic structure $\{1, 0, 0, \cdots 0, 1, 0, \cdots, 1, 0, \cdots, 0\}$. As before, in order to obtain an integer number of periods, we choose N to be a multiple of the period size. The next case to consider is that of period 3.

For $N = 3m$, let A be a $N \times N$ bidiagonal matrix defined as follows:

$$A_{ij} = \begin{cases} 1 & \text{if } i = j+1 \text{ or } (i,j) = (N,1) \\ v_i & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

where v_i is 1 if $i \equiv 1 \bmod 3$ and 0 in all other cases.

The matrix A has the following form:

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

By Proposition 2.1, the characteristic polynomial of the matrix A is

$$\Phi_A(z) = z^{2m}(z-1)^m - 1,$$

so its eigenvalues are roots of the cubic equation

$$z^2(z-1) = e^{\frac{2k\pi i}{m}}$$

with $0 \le k \le n-1$. A straightforward computation using the cubic formula leads to the following expressions for z:

$$z_1^k = \frac{1}{3} + \frac{2^{\frac{1}{3}}}{3p_k} + \frac{p_k}{3 \cdot 2^{\frac{1}{3}}}$$

$$z_2^k = \frac{1}{3} - \frac{1+\sqrt{3}i}{3 \cdot 4^{\frac{1}{3}} p_k} - \frac{1-\sqrt{3}i \cdot p_k}{6 \cdot 2^{\frac{1}{3}}}$$

$$z_3^k = \frac{1}{3} + \frac{1+\sqrt{3}i}{3 \cdot 4^{\frac{1}{3}} p_k} - \frac{1+\sqrt{3}i \cdot p_k}{6 \cdot 2^{\frac{1}{3}}}$$

where

$$p_k = \left(2 + 27 \cdot e^{\frac{2k\pi i}{m}} + 3\sqrt{3} \cdot \sqrt{4e^{\frac{2k\pi i}{m}} + 27e^{\frac{4k\pi i}{m}}}\right)^{\frac{1}{3}}$$

Each formula generates m eigenvalues, corresponding to the m possible roots of unity. In Figure 4, the differently colored points correspond to the eigenvalues resulting from the three expressions for z. Numerically, we observe that the spectrum of the matrix A appears to lie along a smooth curve. However, the expressions for the eigenvalues z make it significantly more complicated to describe the curve than it was in the period 2 case. In an attempt to parametrize this curve, we consider a circulant matrix C whose spectrum is equal to the spectrum of A. If C is of the form

$$C = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \dots & a_N \\ a_N & a_0 & a_1 & a_2 & \dots & a_{N-1} \\ a_{N-1} & a_N & a_0 & a_1 & \dots & a_{N-2} \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ a_1 & a_2 & \dots & & & a_0 \end{pmatrix},$$

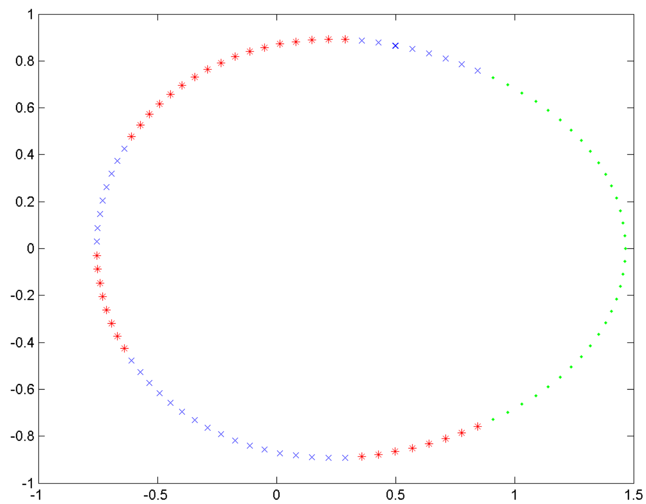

FIGURE 4. *The spectrum of a* 99 *by* 99 *matrix.*

then the spectrum of C lies along the curve

$$f(\theta) = a_0 + a_1 e^{i\theta} + a_2 e^{2i\theta} + \cdots + a_{N-1} e^{i(N-1)\theta}.$$

Therefore, if for an appropriate choice of coefficients a_k the matrices C and A have equal spectra, then $f(\theta)$ will be the desired parametrization. To find the expression for the a_k coefficients, we identify the coefficients of the two characteristic polynomials. For the case $N = 3$, we obtain the following system of equations:

$$\begin{cases} a_0 = & \frac{1}{3} \\ a_1 a_2 = & \frac{1}{9} \\ a_1 + a_2 = & \frac{2}{27} + 1 \end{cases}$$

The numerical solution consists of six triplets (a_0, a_1, a_2). The curves that these coefficients describe coincide for the first three solutions, as well as for the last three, so we obtain two different figures. In the figures below (Figure 5, Figure 6) we plot the two curves, together with the eigenvalues of A.

For higher matrix dimensions, finding and solving the system of equations becomes increasingly difficult. Since the characteristic equation of A is

$$z^{2m}(z-1)^m = 1,$$

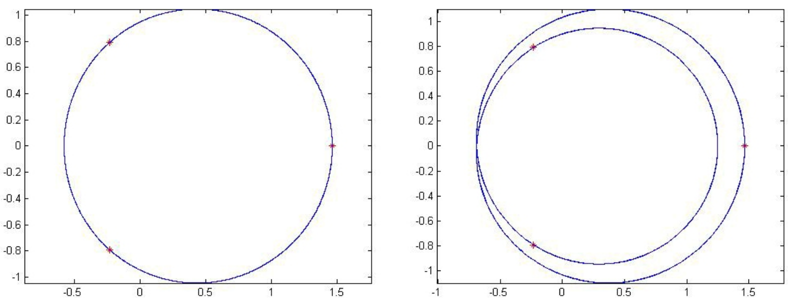

FIGURE 5. *Resulting curves for the* 3×3 *case.*

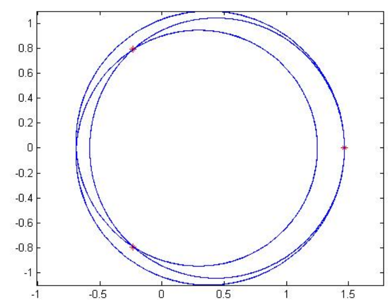

FIGURE 6. *The resulting curves for the* 3×3 *case plotted on the same figure.*

the eigenvalues z will be roots of the polynomial

$$z^2(z-1) - e^{\frac{2k\pi i}{m}}.$$

Therefore, for any fixed k between 0 and $n-1$, we attempt to reduce the calculations to the case of a 3×3 circulant matrix. The system of equations becomes:

$$\begin{cases} a_0 = & \frac{1}{3} \\ a_1 a_2 = & \frac{1}{9} \\ a_1 + a_2 = & e^{\frac{2k\pi i}{m}} + 1 \end{cases}$$

Solving for (a_0, a_1, a_2), we obtain 6 parametric equations for any chosen kth root of unity. The curves corresponding to these equations will contain the three eigenvalues that are the roots of the cubic polynomial in z. If by repeating the process for all k we find that the equations describe the same curve, then all eigenvalues will lie on that curve.

A numerical approximation choosing for instance $m = 2$ (N=6) shows, however, that the curves not coincide (Figure 7). In fact, when

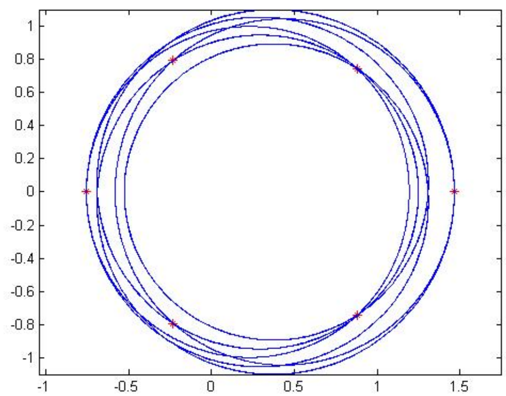

Figure 7. *Curves for the* 6×6 *case, all plotted in the same figure.*

the number of periods, m, is large, the curves completely fill up a space resembling an annulus, which contains the eigenvalues of A (Figure 8). We conjecture that the shape below is an annulus of center a_0.

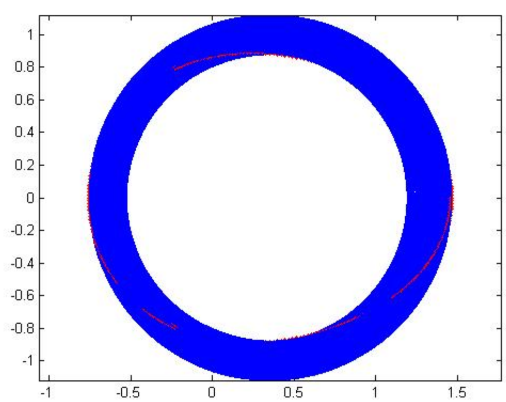

Figure 8. *All curves plotted together for the case* $m = 100$ *(a 300* $\times$ *300 matrix).*

For higher period diagonals, we observe numerically that the spectrum of A lies on a smooth curve. Fixing the number of periods, m, and letting the size of the period, q, increase to infinity (the diagonal sequence becomes $\{1, 0, 0, \cdots 0, 1, 0, \cdots, 0, 1, \cdots, 0\}$), we state the following conjecture:

Conjecture 2.4. As $q \to \infty$, the curve containing the eigenvalues of A approaches the shape of a circle.

In the following figure, we fix $m = 30$ and plot the spectrum of A for values of q between 4 and 100.

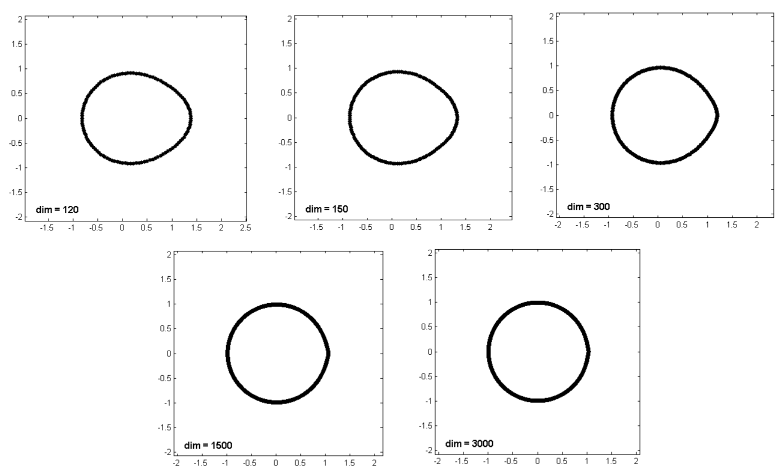

FIGURE 9. *The spectrum of A for $q = 4, q = 5, q = 10, q = 50$ and $q = 100$ respectively, when A is a $30q \times 30q$ matrix. As q increases, the shape of the eigenvalues curve begins to resemble a circle.*

3. Altered-diagonal circulant matrices

Consider a circulant matrix

$$C_n = \begin{pmatrix} a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2 & \vdots & \vdots & \ddots & a_1 \\ a_1 & a_2 & \dots & a_{n-1} & a_0 \end{pmatrix}$$

and a diagonal matrix

$$V_n = \begin{pmatrix} v_0 & 0 & \dots & 0 & 0 \\ 0 & v_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 & v_{n-1} \end{pmatrix}$$

In this section we will study matrices of the form

$$M_n = C_n + V_n = \begin{pmatrix} a_0+v_0 & a_1 & \dots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0+v_1 & a_1 & \dots & a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2 & \vdots & \vdots & \ddots & a_1 \\ a_1 & a_2 & \dots & a_{n-1} & a_0+v_{n-1} \end{pmatrix}$$

Where $a_i \in \mathbb{C}$, $v_i \in \mathbb{C}$. These will be referred to as Altered-Diagonal Circulant Matrices, or ADCM.

3.1. **The Finite Fourier Transform.** It is well known that an $n \times n$ circulant matrix may be diagonalized via conjugation by the $n \times n$ unitary discrete Fourier transform (UDFT). This suggests that the UDFT may provide a favorable basis in which to study altered-diagonal circulant matrices. To further investigate this we define the $n \times n$ unitary discrete Fourier transform.

$$(\mathcal{F}_n)_{ij} = \frac{1}{\sqrt{n}} \cdot \omega_n^{(i-1)(j-1)}$$

Here $\omega_n = e^{\frac{2*\pi*i}{n}}$ is the first n^{th} root of unity. So,

$$\mathcal{F}_n = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & \omega_n^1 & \omega_n^2 & \dots & \omega_n^{n-1} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \vdots & \vdots & \ddots & \omega_n^{(n-2)(n-1)} \\ 1 & \omega_n^{n-1} & \dots & \omega_n^{(n-2)(n-1)} & \omega_n^{(n-1)(n-1)} \end{pmatrix}$$

It is easy to check that $\mathcal{F}_n^{-1} = \mathcal{F}_n^*$, so $(\mathcal{F}_n^{-1})_{ij} = \omega_n^{-(i-1)(j-1)}$

$$\mathcal{F}_n^{-1} = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & \omega_n^{-1} & \omega_n^{-2} & \dots & \omega_n^{-(n-1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \vdots & \vdots & \ddots & \omega_n^{-(n-2)(n-1)} \\ 1 & \omega_n^{-(n-1)} & \dots & \omega_n^{-(n-2)(n-1)} & \omega_n^{-(n-1)(n-1)} \end{pmatrix}$$

Now let us see how conjugation by the UDFT affects an altered diagonal circulant matrix. Define an $n \times n$ ADCM

$$(M_n)_{ij} = \begin{cases} a_0 + v_{i-1} & \text{if } j = i \\ a_{((j-i) \mod n)} & \text{if } j \neq i \end{cases}$$

In the above definition, and throughout the rest of this paper $(j-i)$ mod n will denote the smallest positive integer equal to n modulo $(j-i)$ (so it denotes an integer rather than an equivalence class). Now

$$(C_n)_{ij} = a_{((j-i) \mod n)}$$

and we define the *symbol* f of C_n by $f(z) = \sum_{k=0}^{n-1} a_k \cdot z^k$. Since C_n is circulant it is well known that

$$(\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} = \begin{cases} f(\omega_n^{i-1}) & \text{if } j = i \\ 0 & \text{if } j \neq i. \end{cases}$$

For the following computation it will be beneficial to write V_n in the following form:

$$(V_n)_{ij} = \begin{cases} v_{i-1} & \text{if } j = i \\ 0 & \text{if } j \neq i. \end{cases}$$

So, by standard matrix multiplication,

$$\begin{aligned}
&(\mathcal{F}_n \cdot M_n \cdot \mathcal{F}_n^{-1})_{ij} = (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + (\mathcal{F}_n \cdot V_n \mathcal{F}_n^{-1})_{ij} \\
&= (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + \sum_{k=1}^{n} (\mathcal{F}_n \cdot V_n)_{ik} \cdot (\mathcal{F}_n^{-1})_{kj} \\
&= (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + \sum_{k=1}^{n} \left(\sum_{l=1}^{n} (\mathcal{F}_n)_{il} \cdot (V_n)_{lk} \right) \cdot (\mathcal{F}_n^{-1})_{kj} \\
&= (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + \sum_{k=1}^{n} \left((\mathcal{F}_n)_{ik} \cdot (V_n)_{kk} \right) \cdot \left(\mathcal{F}_n^{-1} \right)_{kj} \\
&= \left(\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1} \right)_{ij} + \sum_{k=1}^{n} v_{k-1} \cdot \frac{1}{\sqrt{n}} \omega_n^{(i-1)(k-1)} \cdot \frac{1}{\sqrt{n}} \omega_n^{-(k-1)(j-1)} \\
&= (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + \frac{1}{n} \sum_{k=1}^{n} v_{k-1} \cdot \omega_n^{(k-1)(i-j)} \\
&= (\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1})_{ij} + \frac{1}{n} \sum_{k=0}^{n-1} v_k \cdot \omega_n^{k(i-j)}
\end{aligned}$$

$$= \begin{cases} f(\omega_n^{i-1}) + \frac{1}{n} \displaystyle\sum_{k=0}^{n-1} v_k & \text{if } j = i \\ \frac{1}{n} \displaystyle\sum_{k=0}^{n-1} v_k \cdot \omega_n^{k(i-j)} & \text{if } j \neq i \end{cases}$$

$$= \begin{cases} p_{i-1} + b_0 & \text{if } j = i \\ b_{((j-i) \mod n)} & \text{if } j \neq i \end{cases}$$

Where we define $p_i = f(\omega_n^i)$ and $b_i = \frac{1}{n}\sum_{k=0}^{n-1} v_k \cdot \omega_n^{-k\cdot i}$. Thus,

$$\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1} = \begin{pmatrix} p_0 & 0 & \dots & 0 & 0 \\ 0 & p_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 & p_{n-1} \end{pmatrix}$$

Where the p_i lie on the curve $f(\partial\mathbb{D})$,

$$\mathcal{F}_n \cdot V_n \cdot \mathcal{F}_n^{-1} = \begin{pmatrix} b_0 & b_1 & \dots & b_{n-2} & b_{n-1} \\ b_{n-1} & b_0 & b_1 & \dots & b_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_2 & \vdots & \vdots & \ddots & b_1 \\ b_1 & b_2 & \dots & b_{n-1} & b_0 \end{pmatrix}$$

$$\mathcal{F}_n \cdot M_n \cdot \mathcal{F}_n^{-1} = \begin{pmatrix} p_0 + b_0 & b_1 & \dots & b_{n-2} & b_{n-1} \\ b_{n-1} & p_1 + b_0 & b_1 & \dots & b_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_2 & \vdots & \vdots & \ddots & b_1 \\ b_1 & b_2 & \dots & b_{n-1} & p_{n-1} + b_0 \end{pmatrix}$$

Thus, the Fourier transform effectively swaps the roles of C_n and V_n. That is to say, C_n is circulant and V_n is diagonal while $\mathcal{F}_n \cdot V_n \cdot \mathcal{F}_n^{-1}$ is circulant, and $\mathcal{F}_n \cdot C_n \cdot \mathcal{F}_n^{-1}$ is diagonal. This may seem to be a meaningless exchange, but it allows us to view M_n as either V_n perturbed by C_n (which is nice to study in the original basis where V_n is diagonal), or as C_n perturbed by V_n (which is nice to study in the Fourier basis where C_n is diagonal). In either case the object of interest is a diagonal matrix which has been perturbed by a circulant matrix. Analysis of such perturbations will be the subject of the next subsection.

3.2. **Circulant Perturbations.** In this section we will derive the formal power series in r for the evolution of eigenvectors and eigenvalues of a diagonal matrix $D_n + r * C_n$ where D_n is diagonal, and C_n is circulant (C_n here is not necessarily the same as in the previous section, although it is circulant in both cases). Note that this power series is only "formal" because, for the moment, there is no guarantee that it will have a non-zero radius of convergence. The goal is to first derive the formal power series in general and then find explicit radii of convergence is certain special cases. To begin, define

$$D_n = \begin{pmatrix} p_0 & 0 & \dots & 0 & 0 \\ 0 & p_1 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 & p_{n-1} \end{pmatrix}$$

and

$$C_n = \begin{pmatrix} b_0 & b_1 & \dots & b_{n-2} & b_{n-1} \\ b_{n-1} & b_0 & b_1 & \dots & b_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_2 & \vdots & \vdots & \ddots & b_1 \\ b_1 & b_2 & \dots & b_{n-1} & b_0 \end{pmatrix}$$

where $b_i, p_i \in \mathbb{C}$.

When $r = 0$, we need only find the eigenvectors and eigenvalues of D_n. Since D_n is diagonal, these are quite simple, indeed this is one advantage of using a diagonal matrix as the starting point of a perturbation. We will denote the q^{th} eigenvalue of D_n by λ^q, and the corresponding eigenvector (an $n \times 1$ matrix) by V^q. Thus $\lambda^q = p_q$ and $(V^q)_{i1} = \delta(i-q)$ (here δ denotes the dirac delta function).

Since the power series must be written in n variables standard multi-index notation will be used. Thus b will be the $n \times 1$ vector defined by $b_{i1} = b_i$ (where b_i is defined as above), and α will be an $n \times 1$ vector of positive integers, where α_{i1} is the exponent of b_i in the product b^α.

$$(V^q(r))_{i1} = \delta(i-1-q) + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha$$

$$\lambda^q(r) = p_q + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha$$

For a given degree d the $c^q_{i,\alpha}$ and k^q_α with $|\alpha| \le d$ must be chosen such that

$$(D_n + r \cdot C_n - \lambda^q(r) \cdot I_n) \cdot V^q(r) \cong O(r^{(d+1)})$$

of course the values of $c^q_{i,\alpha}$ and k^q_α with $|\alpha| > d$ cannot affect this condition (I_n in the above denotes the $n \times n$ identity matrix; the condition $O(r^d)$ applied to a vector simply means that every component of the vector, or equivalently the norm of the vector, must be $O(r^d)$). Additionally, from now on the $\mod n$ in the subscript of $b_{i \mod n}$ will be omitted for notational simplicity, but these subscripts are still to be interpreted as described earlier. By an easy computation, we have

$$(D_n + r\cdot C_n - \lambda^q(r)\cdot I_n)_{ij} = \begin{cases} p_{i-1} - p_q - \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha + r \cdot b_0 & \text{if } j = i \\ r \cdot b_{(j-i)} & \text{if } j \neq i \end{cases}$$

Therefore, by standard matrix multiplication

$$((D_n + r \cdot C_n - \lambda^q(r) \cdot I_n) \cdot V^q(r))_{i1} = \sum_{j=1}^{n} (D_n + r \cdot C_n - \lambda^q(r) \cdot I_n)_{ij} \cdot V^q(r)_{j1}$$

$$= \left(p_{i-1} - p_q - \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha + r \cdot b_0 \right) \cdot \left(\delta\,(i-1-q) + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha \right)$$

$$+ \sum_{j\neq i} r \cdot b_{(j-i)} \cdot \left(\delta(j-1-q) + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{j-1,\alpha} \cdot b^\alpha \right)$$

$$= (p_{i-1} - p_q) \cdot \delta(i-1-q) + (p_{i-1} - p_q) \cdot \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha$$

$$- \delta(i-1-q) \cdot \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha - \Big(\sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha\Big) \cdot \Big(\sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha\Big)$$

$$+ r \cdot b_0 \cdot \delta(i-1-q) + r \cdot b_0 \cdot \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha$$

$$+ \sum_{j\neq i} r \cdot b_{(j-i)} \cdot \delta(j-1-q) + \sum_{d=1}^{\infty} r^{d+1} \sum_{\alpha:|\alpha|=d} \sum_{j\neq i} c^q_{j-1,\alpha} \cdot b^{\alpha+\delta(j-i)})$$

$$= (p_{i-1} - p_q) \cdot \sum_{d=2}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha$$

$$- \delta(i-1-q) \cdot \sum_{d=2}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha - \sum_{d=2}^{\infty} r^d \sum_{\alpha:|\alpha|=d} \sum_{\gamma+\beta=\alpha} c^q_{i-1,\gamma} \cdot k^q_\beta b^\alpha$$

$$+ r \cdot ((p_{i-1} - p_q) \cdot \sum_{\alpha:|\alpha|=1} c^q_{i-1,\alpha} \cdot b^\alpha - \delta(i-1-q) \cdot \sum_{\alpha:|\alpha|=1} k^q_\alpha \cdot b^\alpha + \sum_{j=1}^{n} \cdot b_{(j-i)} \cdot \delta(j-1-q))$$

$$+ \sum_{d=2}^{\infty} r^d \sum_{\alpha:|\alpha|=d} \sum_{j=1}^{n} c^q_{j-1,\alpha-\delta(j-i)} \cdot b^\alpha$$

$$= r \cdot (\sum_{\alpha:|\alpha|=1} ((p_{i-1} - p_q) \cdot c^q_{i-1,\alpha} - \delta(i-1-q) \cdot k^q_\alpha) \cdot b^\alpha + b_{q+1-i})$$

$$+ \sum_{d=2}^{\infty} r^d \sum_{\alpha:|\alpha|=d} ((p_{i-1} - p_q) \cdot c^q_{i-1,\alpha} - -\delta(i-1-q) \cdot k^q_\alpha - \sum_{\gamma+\beta=\alpha} c^q_{i-1,\gamma} \cdot k^q_\beta +$$

$$\sum_{j=1}^{n} c^q_{j-1,\alpha-\delta(j-i)}) \cdot b^\alpha$$

Note that for γ and β in the above equations we require $|\gamma|, |\beta| > 0$ (this restriction was omitted for space considerations). Equivalently we may omit this condition (allow $|\gamma| = 0$ and also omit the $-k^q_\alpha$ term in the above (and define $c^q_{i,\gamma} = 1$ when $|\gamma| = 0$), but for clarity we will use the first convention in the future. Now to satisfy the requirement that

$$(D_n + r \cdot C_n - \lambda^q(r) \cdot I_n) \cdot V^q(r) \cong O(r^{(d+1)})$$

for all d, and for general $b_i \in \mathbb{C}$ we must set each of the coefficients of b^α in the above expression equal to zero. Thus, we obtain

$$(p_{i-1} - p_q) \cdot c^q_{i-1,\delta(w)} - \delta(i-1-q) \cdot k^q_{\delta(w)} + \delta(w-q-1+i) = 0$$

$$(p_{i-1} - p_q) \cdot c^q_{i-1,\alpha} - \delta(i-1-q) \cdot k^q_\alpha - \sum_{\gamma+\beta=\alpha} c^q_{i-1,\gamma} \cdot k^q_\beta + \sum_{j=1}^{n} c^q_{j-1,\alpha-\delta(j-i)} = 0$$

These equations produce the following recursion for the power series coefficients

$$c^q_{i,\delta(w)} = \frac{\delta(w+i-q)}{p_q - p_i}$$

for $i \neq q$, and

$$k^q_{\delta(w)} = \delta(w)$$

$$c^q_{i,\alpha} = \frac{1}{p_q - p_i} \cdot (\sum_{j=1}^{n} c^q_{j,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} c^q_{i,\gamma} \cdot k^q_\beta)$$

for $i \neq q$, and

$$k^q_\alpha = \sum_{j=1}^{n} c^q_{j,\alpha-\delta(j-q)} - \sum_{\gamma+\beta=\alpha} c^q_{q,\gamma} \cdot k^q_\beta$$

Note that the above recursion does not determine the values of $c^q_{q,\alpha}$. Intuitively speaking, this is because eigenvectors are only determined up to a constant multiple. Thus we are free to choose the values of $c^q_{q,\alpha}$ since these coefficients simply determine how much the eigenvector is "stretched" as r grows. To make this notion precise we introduce a new recursion which corresponds to the case in which all of the $c^q_{q,\alpha}$ are set to zero:

$$t^q_{i,\delta(w)} = \frac{\delta(w+i-q)}{p_q - p_i}$$

for $i \neq q$,

$$t^q_{q,\delta(w)} = 0$$

and

$$s^q_{\delta(w)} = \delta(w)$$

$$t^q_{i,\alpha} = \frac{1}{p_q - p_i} \cdot (\sum_{j=1}^{n} t^q_{j,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} t^q_{i,\gamma} \cdot s^q_\beta)$$

for $i \neq q$

$$t^q_{q,\alpha} = 0$$

and

$$s^q_\alpha = \sum_{j=1}^{n} t^q_{j,\alpha-\delta(j-q)} - \sum_{\gamma+\beta=\alpha} t^q_{q,\gamma} \cdot s^q_\beta$$

So s^q_α here plays the role of the eigenvalue coefficients k^q_α in the earlier recursion. Similarly, $t^q_{i,\alpha}$ plays the role of the eigenvector coefficients $c^q_{i,\alpha}$ in the earlier recursion.

Lemma 3.1. *With the above notation, we have*

$$s_\alpha^q = k_\alpha^q$$

and

$$\sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c_{i,\alpha}^q \cdot b^\alpha = (1 + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c_{q,\alpha}^q \cdot b^\alpha) \cdot (\sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} t_{i,\alpha}^q \cdot b^\alpha)$$

Proof. We have

$$\begin{aligned}
&\left(1 + \sum_{d=1}^{\infty} r^d \cdot \sum_{\alpha:|\alpha|=d} c_{q,\alpha}^q b^\alpha\right) \left(\sum_{d=1}^{\infty} r^d \cdot \sum_{\alpha:|\alpha|=d} t_{i,\alpha}^q b^\alpha\right) \\
&= \sum_{d=1}^{\infty} r^d \cdot \sum_{\alpha:|\alpha|=d} t_{i,\alpha}^q b^\alpha + \left(\sum_{d=1}^{\infty} r^d \cdot \sum_{\alpha:|\alpha|=d} c_{q,\alpha}^q b^\alpha\right) \left(\sum_{d=1}^{\infty} r^d \cdot \sum_{\alpha:|\alpha|=d} t_{i,\alpha}^q b^\alpha\right) \\
&= \sum_{d=1}^{\infty} r^d \left(\sum_{\alpha:|\alpha|=d} t_{i,\alpha}^q + \sum_{\gamma+\beta=\alpha} \left(t_{i,\gamma}^q c_{q,\beta}^q\right) b^\alpha\right) \sum_{d=1}^{\alpha} r^d \left(\sum_{\alpha:|\alpha|=d} c_{i,\alpha}^q b^\alpha\right)
\end{aligned}$$

so the lemma is true if and only if

$$c_{i,\alpha}^q = t_{i,\alpha}^q + \sum_{\gamma+\beta=\alpha} t_{i,\gamma}^q \cdot c_{q,\beta}^q$$

we attempt to prove this new condition by induction. That is, we assume this is true when $|\alpha| < d$ (it is easy to verify in the base case), and now when $|\alpha| = d$

$$\begin{aligned}
c_{i,\alpha}^q &= \frac{1}{p_q - p_i} \left(\sum_{j=1}^{n} c_{j,\alpha-\delta(j-i)}^q - \sum_{\gamma+\beta=\alpha} c_{i,\gamma}^q k_\beta^q\right) \\
&= \frac{1}{p_q - p_i} \left(\sum_{j=1}^{n} \left(t_{i,\alpha-\delta(j-i)}^q + \sum_{\varepsilon_1+\varepsilon_2=\alpha-\delta(j-i)} t_{i,\varepsilon_1}^a c_{q,\varepsilon_2}^q\right)\right. \\
&\quad \left. - \sum_{\gamma+\beta=\alpha} \left(t_{i,\gamma}^q + \sum_{\varepsilon_{\gamma 1}+\varepsilon_{\gamma 2}=\gamma} t_{i,\varepsilon_{\gamma 1}}^q c_{q,\varepsilon_{\gamma 2}}^q\right) k_\beta^q\right)
\end{aligned}$$

$$= \frac{1}{p_q - p_i}\left(\sum_{j=1}^{n} t^q_{i,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} t^q_{i,\gamma} k^q_\beta\right) + \frac{1}{p_q - p_i}$$

$$\left(\sum_{j=1}^{n}\sum_{\varepsilon_1+\varepsilon_2=\alpha-\delta(j-i)} t^q_{i,\varepsilon_1} c^q_{q,\varepsilon_2} - \sum_{\gamma+\beta=\alpha}\sum_{\varepsilon_{\gamma 1}+\varepsilon_{\gamma 2}=\gamma} t^q_{i,\varepsilon_{\gamma 1}} c^q_{q,\varepsilon_{\gamma 2}} k^q_\beta\right).$$

Note that by the earlier recursion

$$t^q_{i,\alpha} = \left(\sum_{j=1}^{n} t^q_{i,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} t^q_{i,\gamma} k^q_\beta\right),$$

thus obtaining,

$$c^q_{i,\alpha} = t^q_{i,\alpha} + \frac{1}{p_q - p_i}\left(\sum_{j=1}^{n}\sum_{\varepsilon_1+\varepsilon_2=\alpha+\delta(j-i)} t^q_{i,\varepsilon_1}\cdot c^q_{q,\varepsilon_2} - \sum_{\gamma+\beta+\varepsilon=\alpha} t^q_{i,\gamma}\cdot k^q_\beta \cdot c^q_{q,\varepsilon}\right)$$

$$= t^q_{i,\alpha} + \frac{1}{p_q - p_i}\left(\sum_{\gamma+\beta=\alpha}\sum_{j=1}^{n} t^q_{i,\gamma-\delta(j-i)}\cdot c^q_{q,\beta} - \sum_{\gamma+\beta=\alpha}\sum_{\varepsilon_{\gamma 1}+\varepsilon_{\gamma 2}=\gamma} t^q_{i,\varepsilon_{\gamma 1}}\cdot k^q_{\varepsilon_{\gamma 2}}\cdot c^q_{q\beta}\right)$$

$$= t^q_{i,\alpha} + \frac{1}{p_q - p_i}\left(\sum_{\gamma+\beta=\alpha}\sum_{j=1}^{n} t^q_{i,\gamma-\delta(j-i)}\cdot c^q_{q,\beta} - \sum_{\gamma+\beta=\alpha}\sum_{\varepsilon_{\gamma 1}+\varepsilon_{\gamma 2}=\gamma} t^q_{i,\varepsilon_{\gamma 1}}\cdot k^q_{\varepsilon_{\gamma 2}}\cdot c^q_{q\beta}\right)$$

$$= t^q_{i,\alpha} + \frac{1}{p_q - p_i}\left(\sum_{j=1}^{n}\sum_{\varepsilon_1+\varepsilon_2=\alpha-\delta(j-i)} t^q_{i,\varepsilon_1} c^q_{q,\varepsilon_2} - \sum_{\gamma+\beta+\varepsilon=\alpha} t^q_{i,\gamma} c^q_{q,\varepsilon} k^q_\beta\right)$$

$$= t^q_{i,\alpha} + \sum_{\gamma+\beta=\alpha}\frac{1}{p_q - p_i}\left(\sum_{j=1}^{n} t^q_{i,\gamma-\delta(i-j)} - \sum_{\varepsilon_{\gamma 1}+\varepsilon_{\gamma 2}=\gamma} t^q_{i,\varepsilon_{\gamma 1}} k^q_{\varepsilon_{\gamma 2}}\right) c^q_{q,\beta}$$

$$= t^q_{i,\alpha} + \sum_{\gamma+\beta=\alpha} t^q_{i,\gamma}\cdot c^q_{q,\beta}$$

The last line follows from the recursion defined for $t^q_{i,\gamma}$ earlier.

The second equivalence above is taken to be formal power series equivalence. □

To represent this relationship more clearly we define an "unscaled" eigenvector

$$(V^q_u(r))_{i1} = \delta(i-1-q) + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} t^q_{i-1,\alpha}\cdot b^\alpha$$

Corollary 3.2. *With the above notation, we have*

$$V^q(r) = \left(1 + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{q,\alpha} \cdot b^\alpha\right) \cdot V^q_u(r)$$

Proof. This follows directly by applying Lemma 3.1 to each element of the vector on the right hand side. □

The notation used above was necessary to prove the given statements. However, we can simplify our notation by removing the t's and s's, which we will not need for the rest of our work. This is accomplished by defining "scaling coefficients" $s^q_\alpha = c^q_{q,\alpha}$ which may be chosen freely (we have relieved the symbol s of its earlier role in which $s^q_\alpha = k^q_\alpha$ and from now on we will use only k^q_α as the eigenvalue coefficient). Having done this we then redefine the c's as $c^q_{i,\alpha} = t^q_{i,\alpha}$, allowing them to take the place of the t's so that they are free of the arbitrary choice of scaling coefficients. To conclude we rewrite our results in our new notation.

$$(V^q_u(r))_{i1} = \delta(i-1-q) + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} c^q_{i-1,\alpha} \cdot b^\alpha$$

$$V^q(r) = (1 + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} s^q_\alpha \cdot b^\alpha) \cdot V^q_u(r)$$

$$\lambda^q(r) = p_q + \sum_{d=1}^{\infty} r^d \sum_{\alpha:|\alpha|=d} k^q_\alpha \cdot b^\alpha$$

and the recursion corresponding to our new notation is

$$c^q_{i,\delta(w)} = \frac{\delta(w+i-q)}{p_q - p_i}$$

for $i \neq q$,

$$c^q_{q,\delta(w)} = 0$$

and

$$k^q_{\delta(w)} = \delta(w)$$

$$c^q_{i,\alpha} = \frac{1}{p_q - p_i} \cdot (\sum_{j=1}^{n} c^q_{j,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} c^q_{i,\gamma} \cdot k^q_\beta) = \frac{1}{p_q - p_i} \cdot (\sum_{j \neq q} c^q_{j,\alpha-\delta(j-i)} - \sum_{\gamma+\beta=\alpha} c^q_{i,\gamma} \cdot k^q_\beta)$$

for $i \neq q$, and

$$k_\alpha^q = \sum_{j=1}^{n} c^q_{j,\alpha-\delta(j-q)} - \sum_{\gamma+\beta=\alpha} c^q_{q,\gamma} \cdot k^q_\beta = \sum_{j\neq q} c^q_{j,\alpha-\delta(j-q)}$$

4. Restricting pseudospectra using Jordan decomposition

In the preceding sections, we studied the spectral properties of matrices utilizing only the set of eigenvalues. We now proceed to introduce the ϵ pseudospectrum as an extension of the definition of the spectrum.

For an $n \times n$ complex matrix A and $z \in \mathbb{C}$, let $\sigma(A)$ denote the spectrum of A. Then the following statements are equivalent:

(1) $z \in \sigma(A)$
(2) $(z - A)v = 0$ for some vector $v \neq 0$
(3) $(z - A)$ is not invertible

The following three characterizations of the pseudospectrum relax these properties of the spectrum, creating a more inclusive set.

Definition 4.1. Let $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ and $\epsilon > 0$. Let $\|\cdot\|$ be an operator norm. The **ϵ-pseudospectrum** $\sigma_\epsilon(A)$ of A is the set of $z \in \mathbb{C}$ such that:

(1) $z \in \sigma(A + E)$ for a matrix $E \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ with norm $\|E\| < \epsilon$
(2) $\|(z - A)v\| < \epsilon$ for some vector v such that $\|v\| = 1$ (in which case we call z an **ϵ-pseudoeigenvalue** with **ϵ-pseudoeigenvector** v)
(3) $\|(z - A)^{-1}\| > \frac{1}{\epsilon}$ (with the convention that if $z \in \sigma(A)$, then $\|(z - A)^{-1}\| = \infty$).

Theorem 4.2. *([14]) Conditions 1, 2 and 3 are equivalent.*

Proof. We follow the proof outlined in [14].
In the trivial case that $z \in \sigma(A)$, z satisfies conditions

(1) $E = 0$, $z \in \sigma(A + E)$ and $\|E\| = 0 < \epsilon$
(2) $\|(z - A)v\| = 0 < \epsilon$ for some $v \neq 0$),
(3) $\|(z - A)^{-1}\| = \infty > \frac{1}{\epsilon}$.

If $z \notin \sigma(A)$, then $(z - A)^{-1}$ exists, and we prove (1)⇒(2), (2)⇒(3), and (3)⇒(1).
First, we show (1)⇒(2). Suppose $\exists\, E \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ with $\|E\| < \epsilon$ such that $z \in \sigma(A + E)$, that is, there is a vector $v \in \mathbb{C}^n$ (which we may choose so that $\|v\| = 1$) satisfying $(A+E)v = zv$. Then $Ev = (z-A)v$ and

$$\|(z - A)v\| = \|Ev\| \leq \|E\|\|v\| = \|E\| < \epsilon.$$

Next, we show (2)$\Rightarrow$(3). Suppose $\exists\ v \in \mathbb{C}^n$ with $\|v\| = 1$ so that $\|(z-A)v\| < \epsilon$. Then $(z-A)v = su$ for some $u \in \mathbb{C}^n$, $\|u\| = 1$ and some nonnegative $s < \epsilon$. So $(z-A)^{-1}(su) = v$, $\|(z-A)^{-1}u\| = \frac{\|v\|}{s}$, and $\|(z-A)^{-1}\| \geq \frac{1}{s} > \frac{1}{\epsilon}$.

Finally, we show (3)$\Rightarrow$(1). Suppose $\|(z-A)^{-1}\| > \frac{1}{\epsilon}$, and let u be a unit vector, that is, $\|u\| = 1$. Since $\|(z-A)^{-1}u\| > \frac{1}{\epsilon}$, we may decompose $(z-A)^{-1}u$ into a vector $\|v\| = 1$ and a nonnegative real $s < \epsilon$:

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

Thus,

$$(z-A)v = su.$$

Define a functional l from Span(v), a subspace of $\mathbb{C}^n$, into $\mathbb{C}$ by

$$l(\alpha v) = \alpha l(v).$$

$l(v) = 1$, so $\|l\| = \|l(v)\| = 1$. By a corollary to the Hahn-Banach theorem (see 6.10 in [13]), there exists an extension $\tilde{l} : \mathbb{C}^n \longrightarrow \mathbb{C}$ so that, as with l, $\tilde{l}v = 1$ and $\|\tilde{l}\| = 1$. Define a rank one map E from $\mathbb{C}^n$ to Span$(u) \subseteq \mathbb{C}^n$ by

$$E(x) = s\tilde{l}(x)u \text{ for any } x \in \mathbb{C}^n.$$

Because $\|E\|$ is a linear map from $\mathbb{C}^n$ to $\mathbb{C}^n$, $E \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ with norm

$$\begin{aligned}
\|E\| &= \|su\tilde{l}\| \\
&= s \sup_{\|x\|=1} \|u\tilde{l}x\| \\
&\leq s \sup_{\|x\|=1} \|u\|\|\tilde{l}x\| \\
&= s\|u\|\|\tilde{l}\| \\
&= s \\
&< \epsilon.
\end{aligned}$$

Also, $Ev = s(\tilde{l}v)u = su$. Therefore, $(z-A)v = Ev$ and $zv = (A+E)v$, meaning z is an eigenvalue of $(A+E)$. So $z \in \sigma(A+E)$, and z satisfies condition (1). □

The equivalence of these definitions holds for any operator norm. From here on, however, we denote vector and matrix norms as follows.

Let A be an $n \times n$ matrix with entries $a_{ij} \in \mathbb{C}$. We define the infinity norm to be

$$||A||_{\max} = \max_{i,j} |a_{ij}|$$

where a_{ij} are the entries of A. Let $a = (a_i)_{i=1}^n \in \mathbb{C}^n$. The vector 2-norm and max norm are denoted in the usual manner:

$$||a||_2 = \sqrt{\sum_{i=1}^n |a_i|^2}$$
$$||a||_\infty = \max_i |a_i|.$$

In particular, note that from here on, we use $\|A\|$ to denote the operator 2-norm:

$$||A|| = \sup_{||x||_2=1} ||Ax||_2 .$$

The Frobenius 2-norm:

$$||A||_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}.$$

will also be useful.

It is clear that $||A||_{\max}$, $||A||$, and $||A||_2$ are all norms. It will be clear from context whether we use $||\cdot||_2$ to mean a vector norm or a matrix norm.

Lemma 4.3. *Let A be an $n \times n$ matrix with entries $a_{ij} \in \mathbb{C}$. Then*

$$||A|| \leq n\, ||A||_{\max}$$

Proof. We follow the method outlined in [10].

It suffices to show that

$$||A|| \leq ||A||_2 \leq n\, ||A||_{\max} .$$

A^* is the adjoint of A, or the complex conjugate transpose of A, and A^*A is a compact symmetric operator. By results from operator analysis, therefore,

$$||A||^2 = ||A^*A|| = \rho(A^*A)$$

where $\rho(A^*A)$ is the spectral radius of A^*A. Let $\lambda_{i}{}_{i=1}^n$ be the eigenvalues (with multiplicity) of A^*A. A^*A has real, nonnegative eigenvalues, therefore

$$\rho(A^*A) \leq \sum_i \lambda_i(A^*A).$$

A^*A is similar to its Jordan canonical form (see the next section), and hence

$$\sum_i \lambda_i(A^*A) = tr(A^*A).$$

Now let b_{ij} be the entries of A^*A. Then notice

$$b_{ij} = \langle \bar{a}_i, a_j \rangle,$$

where a_j is the j^{th} column vector of A and $\bar{a}_i$ is the complex conjugate a_i, as $\bar{a}_i{}^T$ is the i^{th} row vector of A^*. Thus we have

$$\begin{aligned} tr(A^*A) &= \sum_{i=0}^{n} |b_{ii}| \\ &= \sum_{i=1}^{n} \langle \bar{a}_i, a_i \rangle \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \\ &= ||A||_2^2 , \end{aligned}$$

proving that $||A|| \leq ||A||_2$.

Now for the other inequality,

$$\begin{aligned} ||A||_2^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \\ &\leq \sum_{i=1}^{n} \sum_{j=1}^{n} \max_{i,j} |a_{ij}|^2 \\ &= n^2 \max_{i,j} |a_{ij}|^2 \\ &= n^2 \, ||A||_{\max}^2 , \end{aligned}$$

thus proving that $||A||_2 \leq n \, ||A||_{\max}$.

□

Remark 4.4. Define $Q \in \mathrm{Mat}_{n\times n}\mathbb{C}$ with all entries $q_{ij} = 1$, that is,

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

The matrix J attains equality in Lemma 4.3.

Proof. We continue to follow [10]. Note that n is an eigenvalues of Q with eigenvector $x = (1, 1, \cdots, 1)^T$. As Q is a positive matrix and x is a positive eigenvector, by the Perron-Frobenius Theorem, n is the largest eigenvalue for Q. Alternatively, we can also note Q has rank one, so n is its only eigenvalue. As Q is a self-adjoint matrix, we have

$$\begin{aligned} ||Q|| &= \sqrt{\rho(Q^*Q)} \\ &= \sqrt{\rho(Q^2)} \\ &= \sqrt{\lambda_{max}(Q^2)} \\ &= \sqrt{n^2} \\ &= n \\ &= n\,||Q||_{\max}. \end{aligned}$$

□

Definition 4.5. A matrix $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ is **normal** if it is unitarily diagonalizable, that is, $\exists$ a unitary matrix U and a diagonal matrix D such that

$$A = UDU^{-1}.$$

Equivalently, A has a complete set of orthogonal eigenvectors.

Theorem 4.6. *([14]) For any $A \in Mat_{n\times n}(\mathbb{C})$, $\epsilon > 0$,*

$$\sigma_\epsilon(A) \supseteq \sigma(A) + D_\epsilon(0)$$

and if A is normal (using the operator norm)

$$\sigma_\epsilon(A) = \sigma(A) + D_\epsilon(0)$$

Previously, the Bauer-Fike theorem provided bounds for the pseudospectrum for diagonalizable matrices.

Definition 4.7. A matrix S has **condition number**

$$\kappa(S) = \|S\|\|S^{-1}\|.$$

Theorem 4.8. *(Bauer-Fike) Let $A \in Mat_{n\times n}(\mathbb{C})$ be diagonalizable, that is, there exist invertible matrix S and diagonal matrix D such that $A = SDS^{-1}$. Let $\epsilon > 0$. Then*

$$\sigma(A) + D_\epsilon 0 \subset \sigma_\epsilon(A) \subset \sigma(A) + D_{\epsilon\kappa(S)}0$$

4.1. **Jordan Canonical Form.** For $\lambda \in \mathbb{C}$ and $n \in \mathbb{N}$, let $J_n^\lambda \in \text{Mat}_{n\times n}(\mathbb{C})$ with entries

$$a_{ij} = \begin{cases} \lambda & \text{for } i = j \\ 1 & \text{for } i = j+1 \\ 0 & \text{else} \end{cases}$$

For example,

$$J_3^2 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}.$$

These J_n^λ are called **Jordan blocks**. A matrix $J \in \text{Mat}_{N\times N}\mathbb{C}$ is in **Jordan canonical form** if

$$J = J_{n_1}^{\lambda_1} \oplus J_{n_2}^{\lambda_2} \oplus \cdots \oplus J_{n_m}^{\lambda_m},$$

where $n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{C}$ for each i, $1 \le i \le m \le N$, and

$$\sum_{i=1}^{m} n_i = N.$$

A matrix of this form is called a **block diagonal matrix** because the Jordan blocks $J_{n_i}^{\lambda_i}$ are arranged along the main diagonal of the matrix.

The following theorem states that every matrix is similar to at least one matrix that is in Jordan form.

Theorem 4.9. *(Jordan Decomposition Theorem)*

Let A be any square matrix. Then there exists an invertible matrix S and a Jordan form matrix J (which shares its eigenvalues with A) such that

$$A = SBS^{-1}.$$

A proof may be found in [4].

The generality of this theorem is crucial to our proof.

4.2. **More important facts about block diagonal matrices.** A few well-established facts about block diagonal matrices will prove useful.

Lemma 4.10. *Let* $A \in Mat_{N\times N}(\mathbb{C})$*, let* $A = \bigoplus_{i=1}^{m} A_i$ *where* A_i *is an* $n_i \times n_i$ *matrix with complex entries. Then*

$$\det(A) = \prod_{i=1}^{m} \det(A_i).$$

Proof. Clearly, if $m = 1$, then $\det(A) = \det(A_1)$. It will suffice to prove the lemma for two blocks ($m = 2$), since we can always group blocks A_1 to A_{n-1} as one block and rewrite the matrix as the direct sum of two block matrices. We must show

$$A = A_1 \oplus A_2 \Rightarrow \det(A) = \det(A_1)\det(A_2) \tag{1}$$

We will proceed by induction on n_2, the dimension of the second block. Suppose A_2 is a 1×1 matrix, that is, $A_2 = \{a_{NN}\}$. Then expanding along the last row of A, we get

$$\begin{aligned} \det(A) &= a_{NN}\det(A_1) \\ &= \det(A_2)\det(A_1). \end{aligned}$$

So Equation 1 holds when $n_2 = 1$. Now assume that Equation 1 holds when $n_2 = k$. Suppose A_2 is a $(k+1) \times (k+1)$ matrix. Let B_i be the $k \times k$ matrix obtained from A_2 by deleting the i^{th} column and the first row. Let $C_i = A_1 \oplus B_i$. Then taking the determinant of A by expanding along row $n_1 + 1$, we obtain

$$\det(A) = \sum_{i=1}^{k+1} (-1)^{i+1} a_{(1+n_1)(i+n_1)} \det(C_i).$$

C_i is the direct sum of one block, A_1, with a $k \times k$ block, B_i, so by our induction hypothesis, $\det(C_i) = \det(A_1)\det(B_i)$. Thus we get

$$\begin{aligned}\det(A) &= \sum_{i=1}^{k+1}(-1)^{i+1}a_{(1+n_1)(i+n_1)}\det(A_1)\det(B_i)\\ &= \left(\sum_{i=1}^{k+1}(-1)^{i+1}a_{(1+n_1)(i+n_1)}\det(B_i)\right)\det(A_1)\\ &= \det(A_1)\det(A_2),\end{aligned}$$

proving (1).

□

Lemma 4.11 is a closely related fact concerning the characteristic polynomial of A, defined $p(A) = \det(zI - A)$:

Lemma 4.11. *Let A and $\{A_i\}_{i=1}^m$ be matrices such that*

$$A = A_1 \oplus A_2 \oplus \cdots \oplus A_m.$$

Then

$$p(A) = \prod_{i=1}^{n} p(A_i).$$

Proof.

$$\begin{aligned}zI - A &= zI - (A_1 \oplus A_2 \oplus \cdots \oplus A_m)\\ &= (zI - A_1) \oplus (zI - A_2) \oplus \cdots \oplus (zI - A_m)\\ \det(zI - A) &= \prod_{i=1}^{m} \det(zI - A_i)\end{aligned}$$

as desired. □

Next comes a convenient fact about the norms of block matrices.

Lemma 4.12. *Let $A = \bigoplus_{i=1}^m A_i$, where A_i are $n_i \times n_i$ matrices. Then*

$$||A|| = \max_i\{||A_i||\}.$$

Proof. As with Lemma 4.10, it suffices to prove Lemma 4.12 for the case $m = 2$. Thus,

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

where A_1 is an $n_1 \times n_1$ matrix and A_2 is an $n_2 \times n_2$ matrix, and we wish to show that

$$||A|| = \max(||A_1||, ||A_2||). \tag{2}$$

Then we see that $A_i : \mathbb{C}^{n_i} \to \mathbb{C}^{n_i}$, and that $\mathbb{C}^{n_i}$ are each Hilbert spaces. Moreover, $\mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2} \cong \mathbb{C}^{n_1+n_2}$. That is to say, the subspace of $\mathbb{C}^{n_1+n_2}$ that acts as the domain and range of the block matrix A_1 is orthogonal to that of A_2 in $\mathbb{C}^{n_1+n_2}$. Thus we have, by norms of direct sums, for $A_1 \oplus A_2 : \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2} \to \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2}$,

$$\begin{aligned}
||A|| &= ||A_1 \oplus A_2|| \\
&= \sup_{||(x,y)||_2=1} ||(A_1x, A_2y)||_2 \\
&= \sup_{||x||_2^2+||y||_2^2=1} \sqrt{||A_1x||_2^2 + ||A_2y||_2^2} \\
&\leq \sup_{||x||_2^2+||y||_2^2=1} \sqrt{||A_1||^2\,||x||_2^2 + ||A_2||^2\,||y||_2^2} \\
&\leq \sup_{||x||_2^2+||y||_2^2=1} \sqrt{||\max(||A_1||, ||A_2||)||^2\,||x||_2^2 + (\max(||A_1||, ||A_2||))^2\,||y||_2^2} \\
&= \sup_{||x||_2^2+||y||_2^2=1} \sqrt{(\max(||A_1||, ||A_2||))^2\left(||x||_2^2 + ||y||_2^2\right)} \\
&= \max(||A_1||, ||A_2||).
\end{aligned}$$

Now we will show inequality the other direction. Without loss of generality, suppose $\max(||A_1||, ||A_2||) = ||A_1||$. Then

$$\begin{aligned}
||A_1 \oplus A_2|| &= \sup_{||(x,y)||_2=1} ||(A_1x, A_2y)||_2 \\
&\geq \sup_{||x||_2=1, ||y||_2=0} \sqrt{||A_1x||_2^2 + ||A_2y||_2^2} \\
&= \sup_{||x||_2=1} ||A_1x||_2 \\
&= ||A_1|| \\
&= \max(||A_1||, ||A_2||),
\end{aligned}$$

and we have proven Equation 2.

□

4.3. **Disk Bounds for Pseudospectra.** For any $r > 0, c \in \mathbb{C}$, denote the disk of radius r about c by

$$D_r(c) = \{z \in \mathbb{C} \mid |c - z| \leq r\}.$$

Theorem 4.13. *Let $J \in Mat_{N\times N}(\mathbb{C})$ be in Jordan canonical form, that is, let*

$$J = J_{n_1}^{\lambda_1} \oplus J_{n_2}^{\lambda_2} \oplus \cdots \oplus J_{n_m}^{\lambda_m},$$

where $n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{C}$ for each i, $1 \leq i \leq m \leq N$. Let $n = \max_i(n_i)$. Then for $0 < \epsilon < \frac{1}{n}$, we have

$$\bigcup_{i=1}^{m} D_{(\epsilon)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(J) \subseteq \bigcup_{i=1}^{m} D_{(n_i\epsilon)^{\frac{1}{n_i}}}(\lambda_i).$$

Proof. Note that for all i, $n_i \leq n$ and $\epsilon < \frac{1}{n} \leq \frac{1}{n_i}$.

For the first inclusion, let $\alpha \in D_{\epsilon^{\frac{1}{n_i}}}(\lambda_i)$ for some specific i. We may write

$$\alpha - \lambda_i = re^{i\theta},$$

where $r, \theta \in \mathbb{R}$ and $r \leq \epsilon^{\frac{1}{n_i}}$. Let $\beta = r^n e^{in\theta}$ so that

$$(\alpha - \lambda_i)_i^n = \beta.$$

Let

$$j_i = \sum_{k=1}^{i-1} n_k,$$

so that the entry (a, b) of $J_{n_i}^{\lambda_i}$ is the entry $(j_i + a, j_i + b)$ of J. Let $E \in \text{Mat}_{N\times N}(C)$ with entries

$$e_{ab} = \begin{cases} \beta & \text{for } (a,b) = (j_i + n_i, j_i + 1) \\ 0 & \text{else.} \end{cases}$$

Clearly

$$\|E\| = |\beta| \leq \epsilon$$

and

$$P(J_{n_i}^{\lambda_i}) = (z - \lambda_i)^{n_i} - \beta,$$

so the zeros of $P(J_{n_i}^{\lambda_i})$ are given by:

$$(z - \lambda_i)^{n_i} = \beta$$

One solution is

$$z - \lambda_i = \alpha - \lambda_i$$
$$z = \alpha$$

So α is a zero of $P(J_{n_i}^{\lambda_i})$. By Lemma 4.11, α is also a zero of $P(J)$, thus $\alpha \in \sigma_\epsilon(J)$. For the second inclusion, let $z \in \sigma_\epsilon(J)$, so that by the definition of pseudospectrum,

$$\|(zI - J)^{-1}\| \geq \frac{1}{\epsilon}.$$

Note that

$$(zI - J)^{-1} = \bigoplus_{i=1}^{m} (zI - J_{n_i}^{\lambda_i})^{-1}.$$

By Lemma 4.12,

$$\|(zI - J)^{-1}\| = \max_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|$$

and by Lemma 4.3,

$$\max_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\| \leq \max_i (n_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|_{\max})$$

Thus,

$$\max_i (n_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|_{\max}) \geq \frac{1}{\epsilon},$$

meaning that for some i,

$$n_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|_{\max} \geq \frac{1}{\epsilon}.$$

It can be easily verified that $(zI - J_{n_i}^{\lambda_i})^{-1}$ has entries

$$d_{ab} = \begin{cases} (z-\lambda)^{a-b-1} & \text{for } a \leq b \\ 0 & \text{for } a > b \end{cases}$$

and therefore that

$$\|(zI - J_{n_i}^{\lambda_i})^{-1}\|_{max} = \begin{cases} \|(z-\lambda_i)^{-n_i}| & \text{if } |z-\lambda_i| \leq 1 \\ |(z-\lambda_i)^{-1}| & \text{if } |z-\lambda_i| > 1 \end{cases}$$

The second case leads to

$$n_i \geq n_i |(z-\lambda_i)^{-1}| \geq \frac{1}{\epsilon},$$

and $\epsilon \geq \frac{1}{n_i}$, a contradiction. Therefore, the first case must be true, from which we find that

$$|(z-\lambda_i)^{-n_i}| = \|(zI - J_{n_i}^{\lambda_i})^{-1}\|_{max} \geq \frac{1}{n_i \epsilon},$$

$$|(z-\lambda_i)^{n_i}| \leq n_i \epsilon,$$

$$|(z-\lambda_i)| \leq (n_i \epsilon)^{\frac{1}{n_i}},$$

and

$$z \in D_{(n_i\epsilon)^{\frac{1}{n_i}}}(\lambda_i) \in \bigcup_{i=1}^{m} D_{(n_i\epsilon)^{\frac{1}{n_i}}}(\lambda_i).$$

□

Remark 4.14. Because $0 < \epsilon \leq 1$, if $n_1 > n_2$, then $(\epsilon)^{\frac{1}{n_1}} \leq (\epsilon)^{\frac{1}{n_2}}$. Thus, an expression such as

$$\bigcup_{i=1}^{m} D_{(\epsilon)^{\frac{1}{n_i}}}(\lambda_i)$$

may be understood in terms of the largest block dimension corresponding to each distinct eigenvalues. That is, if $0 < \epsilon < 1$ and we let

$$L = \{i \mid n_i \geq n_j \forall j : \lambda_j = \lambda_i\},$$

then

$$\bigcup_{i=1}^{m} D_{(\epsilon)^{\frac{1}{n_i}}}(\lambda_i) = \bigcup_{i \in L} D_{(\epsilon)^{\frac{1}{n_i}}}(\lambda_i).$$

In order to generalize this theorem to all square matrices, we define the **condition number** of a matrix.

Definition 4.15. An invertible matrix S has **condition number**

$$\kappa(S) = \|S\|\|S^{-1}\|.$$

The next lemma restates theorem 4ϵ in [5], relating matrix similarity, pseudospectra, and condition numbers.

Lemma 4.16. *Let $A, B \in Mat_{n\times n}(\mathbb{C})$ and suppose $\exists$ an invertible matrix S such that $A = SBS^{-1}$. Then for all $\epsilon \geq 0$,*

$$\sigma_{\frac{\epsilon}{\kappa(S)}}(B) \subseteq \sigma_\epsilon(A) \subseteq \sigma_{\kappa(S)\epsilon}(B).$$

Proof. Begin with the right hand side inclusion. Suppose $z \in \sigma_\epsilon(A)$. Then

$$\begin{aligned}
A &= SBS^{-1} \\
z - A &= SzS^{-1} - SBS^{-1} \\
&= S(z - B)S^{-1} \\
(z - A)^{-1} &= S(z - B)^{-1}S^{-1} \\
\|(z - A)^{-1}\| &\leq \|S\|\|(z - B)^{-1}\|\|S^{-1}\| \\
&= \kappa(S)\|(z - B)^{-1}\| \\
\epsilon^{-1} &\leq \|(z - A)^{-1} \\
&\leq \kappa(S)\|(z - B)^{-1}\| \\
(\epsilon\kappa(S))^{-1} &\leq \|(z - B)^{-1}\|
\end{aligned}$$

This shows the right hand side inclusion, and since $B = S^{-1}AS$, it also implies that for all $\zeta \geq 0$,

$$\sigma_\zeta(B) \subseteq \sigma_{\kappa(S^{-1})\zeta}(A).$$

Note that

$$\kappa(S) = \|S\|\|S^{-1}\| = \kappa(S^{-1}),$$

and let $\zeta = \frac{\epsilon}{\kappa(S)}$ to show

$$\sigma_{\frac{\epsilon}{\kappa(S)}}(B) \subseteq \sigma_\epsilon(A).$$

□

We are now ready to prove our main result.

Theorem 4.17. *Let $A \in Mat_{N\times N}(\mathbb{C})$ be a matrix. Let $S, J \in Mat_{N\times N}(\mathbb{C})$ be a corresponding invertible matrix and Jordan form matrix such that $A = SJS^{-1}$. As before, write $J = \bigoplus_{i=1}^m J_{n_i}^{\lambda_i}$ and $n = \max_i n_i$. Then for $0 < \epsilon < \frac{1}{n}$, we have*

$$\bigcup_{i=1}^m D_{\left(\frac{\epsilon}{\kappa(S)}\right)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(A) \subseteq \bigcup_{i=1}^m D_{(n_i\kappa(S)\epsilon)^{\frac{1}{n_i}}}(\lambda_i).$$

Proof. By Lemma 4.16

$$\sigma_{\frac{\epsilon}{\kappa(S)}}(J) \subseteq \sigma_\epsilon(A) \subseteq \sigma_{\kappa(S)\epsilon}(J).$$

Applying Theorem 4.13,

$$\bigcup_{i=1}^m D_{\left(\frac{\epsilon}{\kappa(S)}\right)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_{\frac{\epsilon}{\kappa(S)}}(J)$$

and

$$\sigma_{\kappa(S)\epsilon}(J) \subseteq \bigcup_{i=1}^{m} D_{(n_i\kappa(S)\epsilon)^{\frac{1}{n_i}}}(\lambda_i).$$

Therefore,

$$\bigcup_{i=1}^{m} D_{\left(\frac{\epsilon}{\kappa(S)}\right)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(A) \subseteq \bigcup_{i=1}^{m} D_{(n_i\kappa(S)\epsilon)^{\frac{1}{n_i}}}(\lambda_i)$$

as desired.

□

4.4. **An optimal upper bound.** Here we use the upper bound in Theorem 4.17 and a tighter restriction of ϵ to prove an optimal upper bound.

Let $n \in \mathbb{N}$. We define a function $f_n : (0, 1) \to \mathbb{R}$ by

$$f_n(x) = \frac{1 - x^n}{x^n - x^{n+1}}.$$

We also define the interval Q_n as the image

$$Q_n = f_n\left(\left(0, \frac{n}{n+1}\right)\right).$$

Lemma 4.18. *$f_n(x)$ is strictly decreasing on the interval $0 < x < \frac{n}{n+1}$.*

Proof. By the product rule,

$$f_n'(x) = \frac{(x^n - x^{n+1})n - (1 - x^n)(nx^{n-1} - (n+1)x^n)}{(x^n - x^{n+1})(x^n - x^{n+1})}.$$

The denominator is positive, and the numerator reduces to

$$-x^{2n} + (n+1)x^n - nx^{n-1}.$$

We compute

$$\begin{aligned} x &< \frac{n}{n+1} \\ (n+1)x &< n \\ (n+1)x^n &< nx^{n-1} \\ -x^{2n} + (n+1)x^n - nx^{n-1} &< 0 \\ f'(x) &< 0. \end{aligned}$$

□

Remark 4.19. The interval on which $f'(x) < 0$ actually extends farther to the right. As a consequence, in the upcoming theorem, the upper bound on ϵ will *not* be the lowest possible.

Corollary 4.20. *Let* $n \in \mathbb{N}$*. Given* $q \in Q_n$*, there exists a unique* $x_q \in \left(0, \frac{n}{n+1}\right)$ *such that* $f_n(x_q) = q$*.*

Proof. Existence follows from the definition of Q_n, and uniqueness from Lemma 4.18.

□

Define the inverse function of f_n on Q_n as:

$$f_n^{-1} : Q_n \to \left(\left(0, \frac{n}{n+1}\right)\right),$$

where for $q \in Q_n$,

$$f_n^{-1}(q) = x_q.$$

By Corollary 4.20, f_n^{-1} is well defined. Also, for $n \in \mathbb{N}$, let

$$\eta_n = \frac{n^n}{n(n+1)^n}$$

For $n \in \mathbb{N}$, we define

$$\xi_n = \frac{\left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1}}{1 - \left(\frac{n}{n+1}\right)^n}$$

Lemma 4.21. *Let* $n \in \mathbb{N}$*. Then*

(1)

$$\eta_n \leq \xi_n$$

(2)

$$\eta_n \leq \frac{1}{n}.$$

Proof. For (1), let $j \in \mathbb{N}, 0 \leq jl \leq n-1$. The following two statements are clear:

$$(n-1-0)(n-1-1)\cdots(n-1-(j-1)) \leq n^j$$

and

$$(n-1-j)! \leq (n-1-j)!j!.$$

Multiply each side to find

$$\begin{aligned}
(n-1)! &\leq (n-1-j)!j!n^j \\
n^j &\geq \frac{(n-1)!}{(n-1-j)!j!} \\
n^j(n^{n-j-1}) &\geq \frac{(n-1)!}{(n-1-j)!j!} n^{n-j-1} \\
n^{n-1} &\geq \frac{(n-1)!}{(n-1-j)!j!} n^{n-j-1} \\
&= \binom{n-1}{j} n^{n-i-j} \\
\sum_{j=0}^{n-1} n^{n-1} &\geq \sum_{j=0}^{n-1} \binom{n-1}{j} n^{n-1-j} 1^j,
\end{aligned}$$

the binomial expansion of $(n+1)^{n-1}$. Continuing,

$$\begin{aligned}
n^n &\geq (n+1)^{n-1} \\
\frac{n^n}{(n+1)^{n-1}} &\leq 1 \\
n^2 - n - n^2 &\geq n + 1 - \frac{n^n}{(n+1)^{n-1}} \\
n - \frac{n^2}{n+1} &\geq 1 - \left(\frac{n}{n+1}\right)^n \\
\frac{1 - \frac{n}{n+1}}{1 - \left(\frac{n}{n+1}\right)^n} &\geq \frac{1}{n} \\
\frac{\left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1}}{1 - \left(\frac{n}{n+1}\right)^n} &\geq \frac{n^n}{n(n+1)^n} \\
\xi_n &\geq \eta_n.
\end{aligned}$$

For (2), see that

$$\begin{aligned}
\frac{n^n}{n(n+1)^n} &= \frac{1}{1+n} \left(\frac{n}{n+1}\right)^{n-1} \\
&< \frac{1}{1+n} \\
&< \frac{1}{n}
\end{aligned}$$

$\square$

Lemma 4.22. *If* $0 < \epsilon < \xi_n$, *then* $\frac{1}{\epsilon} \in Q_n$.

Proof. The only possible discontinuities in the rational function $f_n(x)$ are at $x = 0$ and $x = 1$, so $f_n(x)$ is continuous on the interval $\left(0, \frac{n}{n+1}\right)$. Observe that

$$\lim_{x \to 0^+} (f_n) = \infty$$

and

$$\begin{aligned} f_n(\frac{n}{n+1}) &= \frac{1 - \left(\frac{n}{n+1}\right)^n}{\left(\frac{n}{n+1}\right)^n - \left(\frac{n}{n+1}\right)^{n+1}} \\ &= \frac{1}{\xi_n} \\ &< \frac{1}{\epsilon}. \end{aligned}$$

Thus

$$f_n(\frac{n}{n+1}) < \frac{1}{\epsilon} < \lim_{x \to 0^+} f_n(x)$$

By the Intermediate Value Theorem, there is some $x_q \in \left(0, \frac{n}{n+1}\right)$ such that $f_n(x_q) = \frac{1}{\epsilon}$. Thus, $\frac{1}{\epsilon} \in Q_n$. □

Lemma 4.23. *Let* $n \in \mathbb{N}$, $0 < \epsilon < \xi_n$. *If* $q > \frac{1}{\epsilon}$, *then* $q \in Q_n$ *and*

$$f_n^{-1}(q) < f_n^{-1}(\frac{1}{\epsilon}).$$

Proof.

$$\frac{1}{\epsilon} < q < \lim_{x \to 0^+} f_n(x).$$

By the Intermediate Value Theorem, $q \in Q_n$, and by Lemma 4.18, $f_n^{-1} q < f_n^{-1}(\frac{1}{\epsilon})$. □

Theorem 4.24. *Let* $J \in Mat_{N \times N}(\mathbb{C})$ *be in Jordan Canonical Form, that is, let*

$$J = J_{n_1}^{\lambda_1} \oplus J_{n_2}^{\lambda_2} \oplus \cdots \oplus J_{n_m}^{\lambda_m},$$

where $n_i \in \mathbb{N}$ *and* $\lambda_i \in \mathbb{C}$ *for each* i, $1 \leq i \leq m \leq N$. *Let* $\eta = \min_i \eta_{n_i}$. *Then for* $0 < \epsilon < \eta$, *we have*

$$\bigcup_{i=1}^{m} D_{(\epsilon)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(J) \subseteq \bigcup_{i=1}^{m} D_{f_{n_i}^{-1}(\frac{1}{\epsilon})}(\lambda_i).$$

Proof. By Lemma 4.21

$$0 < \epsilon < \eta \leq \eta_{n_i} \leq \xi_{n_i},$$

as required by Lemmas 4.22 and 4.23, and $\epsilon < \frac{1}{n}$ as required by Theorem 4.17. The first inclusion of this theorem is identical to that of Theorem 4.13.

For the second inclusion, let $z \in \sigma_\epsilon(J)$, so that by the definition of the ϵ-pseudospectrum,

$$\|(zI - J)^{-1}\| \geq \frac{1}{\epsilon}.$$

Note that

$$(zI - J)^{-1} = \bigoplus_{i=1}^{m} (zI - J_{n_i}^{\lambda_i})^{-1}.$$

By Lemma 4.12,

$$\|(zI - J)^{-1}\| = \max_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|$$

So for some i,

$$\frac{1}{\epsilon} \leq \|(zI - J_{n_i}^{\lambda_i})^{-1}\|. \tag{3}$$

Equation 3 leads to $|z - \lambda_i| \leq (n_i\epsilon)^{\frac{1}{n_i}}$ as in Theorem 4.13. Let $r = |z - \lambda_i|$. Substituting, we find

$$\begin{aligned} r &\leq (n_i\epsilon)^{\frac{1}{n_i}} \\ &\leq (n_i\eta_{n_i})^{\frac{1}{n_i}} \\ &\leq (n_i\frac{n_i^{n_i}}{n_i(n_i+1_i)^{n_i}})^{\frac{1}{n_i}} \\ &= \frac{n_i}{n_i+1} \end{aligned}$$

So either $r = \lambda_i$, in which case

$$z \in D_{f_{n_i}^{-1}(\epsilon)}(\lambda_i)$$

is trivial, or $r \neq \lambda_i$, so that

$$0 < r < \frac{n_i}{n_i+1}. \tag{4}$$

By Equation 4 and Corollary 4.20, r is the unique $x \in (0, \frac{n_i}{n_i+1})$ such that $f_{n_i}^{-1}(x) = f_{n_i}^{-1}(r)$. We may write

$$r = f_{n_i}^{-1}(f_{n_i}(r)) \tag{5}$$

Equation 3 also yields

$$\begin{aligned}
\frac{1}{\epsilon} &\le \max_{\|v\|=1} \|(zI - J_{n_i}^{\lambda_i})^{-1} v\| \\
&= \max_{\|v\|=1} \left\| \begin{pmatrix} (z-\lambda_i)^{-1} & (z-\lambda_i)^{-2} & \cdots & (z-\lambda_i)^{-n_i} \\ 0 & (z-\lambda_i)^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & (z-\lambda_i)^{-2} \\ 0 & \cdots & 0 & (z-\lambda_i)^{-1} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \ddots \\ v_{n_i} \end{pmatrix} \right\| \\
&\le \max_{\|v\|=1} \left(\sum_{j=1}^{n_i} \left| v_j \sum_{k=j}^{n_i} (z-\lambda_i)^{-k} \right|^2 \right)^{\frac{1}{2}} \\
&= \max_{\|v\|=1} \left(\sum_{j=1}^{n_i} |v_j|^2 \left| \sum_{k=j}^{n_i} (z-\lambda_i)^{-k} \right|^2 \right)^{\frac{1}{2}}
\end{aligned}$$

Applying the triangle inequality,

$$\begin{aligned}
\frac{1}{\epsilon} &\le \max_{\|v\|=1} \left(\sum_{j=1}^{n_i} |v_j|^2 \sum_{k=j}^{n_i} \left| (z-\lambda_i)^{-k} \right|^2 \right)^{\frac{1}{2}} \\
&\le \max_{\|v\|=1} \left(\sum_{j=1}^{n_i} |v_j|^2 \left(\sum_{k=1}^{n_i} r^{-k} \right)^2 \right)^{\frac{1}{2}} \\
&= \max_{\|v\|=1} \left(\left(\sum_{k=1}^{n_i} r^{-k} \right)^2 \sum_{j=1}^{n_i} |v_j|^2 \right)^{\frac{1}{2}} \\
&= \max_{\|v\|=1} \sum_{k=1}^{n_i} r^{-k} \left(\sum_{j=1}^{n_i} |v_j|^2 \right)^{\frac{1}{2}} \\
&= \sum_{k=1}^{n_i} r^{-k}
\end{aligned}$$

This is the partial sum of a geometric series:

$$\begin{aligned}
\frac{1}{\epsilon} &\le \frac{r^{-n-1} - r^{-1}}{r^{-1} - 1} \\
&= \frac{1 - r^n}{r^n - r^{n+1}} \\
&= f_{n_i}(r)
\end{aligned}$$

Lemma 4.22 states that $\frac{1}{\epsilon} \in Q_{n_i}$ and Lemma 4.23 and $\frac{1}{\epsilon} < f_{n_i}(r)$ together imply that that $f_{n_i}(r) \in Q_{n_i}$ and

$$f_{n_i}^{-1}(\epsilon) < f_{n_i}^{-1}(f_{n_i}(r))$$

Substituting from Equation 5,

$$f_{n_i}^{-1}(\epsilon) < r = |z - \lambda_i|.$$

So

$$z \in D_{f_{n_i}^{-1}(\frac{1}{\epsilon})}(\lambda_i),$$

proving that

$$\sigma_\epsilon(J) \subseteq \bigcup_{i=1}^{m} D_{f_{n_i}^{-1}(\frac{1}{\epsilon})}(\lambda_i).$$

□

This improvement of the bound on the pseudospectrum of a Jordan form matrix has obvious consequences for the general case.

Theorem 4.25. *Let $A \in Mat_{N\times N}(\mathbb{C})$ be a matrix. Let $S, J \in Mat_{N\times N}(\mathbb{C})$ be a corresponding invertible matrix and Jordan form matrix such that $A = SJS^{-1}$. As before, write $J = \bigoplus_{i=1}^{m} J_{n_i}^{\lambda_i}$ and $\eta = \min_i \eta_{n_i}$. Then for $0 < \epsilon < \eta$, we have*

$$\bigcup_{i=1}^{m} D_{(\frac{\epsilon}{\kappa(S)})^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(A) \subseteq \bigcup_{i=1}^{m} D_{f_{n_i}^{-1}(\frac{1}{\kappa(S)\epsilon})}(\lambda_i)$$

Proof. Theorem 4.25 follows straightforward from Lemma 4.16 and Theorem 4.24 as Theorem 4.17 followed from Lemma 4.16 and 4.13. □

Next, we prove that the righthand bound in Theorem 4.24 is optimal.

Theorem 4.26. *Suppose that for some $w > 0$, for all $J \in Mat_{N\times N}(\mathbb{C})$ in Jordan canonical form, that is,*

$$J = J_{n_1}^{\lambda_1} \oplus J_{n_2}^{\lambda_2} \oplus \cdots \oplus J_{n_m}^{\lambda_m},$$

where $n_i \in \mathbb{N}$ and $\lambda_i \in \mathbb{C}$ for each i, $1 \le i \le m \le N$, and $\eta = \min_i \eta_{n_i}$, it is true for all $0 < \epsilon < \eta$ that

$$\sigma_\epsilon(J) \subseteq \bigcup_{i=1}^{m} D_w(\lambda_i).$$

Then

$$w \le f_{n_i}^{-1}(\frac{1}{\epsilon}).$$

Proof. For each $1 \le i \le m$, let $z_i = \lambda_i + w$. From the above proof of Theorem 4.24, we know that

$$\begin{aligned}\|(z_i I - J_{n_i}^{\lambda_i})^{-1}\| &\le f_n(z_i - \lambda_i) \\ &= f_n(w)\end{aligned}$$

Moreover,

$$\begin{aligned}\|(z_i I - J_{n_i}^{\lambda_i})^{-1}\| &= \max_{\|v\|=1} \|(z_i I - J_{n_i}^{\lambda_i})^{-1} v\| \\ &\ge \left\| \begin{pmatrix} (z_i-\lambda_i)^{-1} & (z_i-\lambda_i)^{-2} & \cdots & (z_i-\lambda_i)^{-n_i} \\ 0 & (z_i-\lambda_i)^{-1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & (z_i-\lambda_i)^{-2} \\ 0 & \cdots & 0 & (z_i-\lambda_i)^{-1} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\| \\ &= \sum_{k=1}^{n} (z_i - \lambda_i)^{-k} \\ &= \sum_{k=1}^{n} (w)^{-k} \\ &= f_n(w)\end{aligned}$$

Therefore,

$$\|(z_i I - J_{n_i}^{\lambda_i})^{-1}\| = f_n(w)$$

Since $z \in D_w(\lambda_i)$, $z \in \sigma_\epsilon(J)$. By definition of pseudospectrum,

$$\begin{aligned}\frac{1}{\epsilon} &\le \|(zI - J)^{-1}\| \\ &= \max_i \|(zI - J_{n_i}^{\lambda_i})^{-1}\|\end{aligned}$$

So for at least one i,

$$\begin{aligned}\frac{1}{\epsilon} &\le \|(zI - J_{n_i}^{\lambda_i})^{-1}\| \\ &= f_n(w),\end{aligned}$$

so by Lemma 4.23,

$$w \le f_{n_i}^{-1}(\frac{1}{\epsilon}).$$

□

Remark 4.27. The lower bound,

$$\bigcup_{i=1}^{m} D_{\left(\frac{\epsilon}{\kappa(S)}\right)^{\frac{1}{n_i}}}(\lambda_i) \subseteq \sigma_\epsilon(A)$$

seems to be optimal also, but we know of no proof.

4.5. **Discussion of condition numbers.** By the Jordan Decomposition Theorem, any matrix can be reduced to one of its Jordan Canonical Forms through conjugation with a similarity matrix. That is, for any $A \in \text{Mat}_{n\times n}(\mathbb{C})$, there exists $S \in \text{Mat}_{n\times n}(\mathbb{C})$ such that S is invertible and, for J the Jordan Canonical Form of A, $J = S^{-1}AS$. Our theorem concerning the disk bounds on the ϵ pseudospectra of A involves the condition number $\kappa(S) = ||S||\,||S^{-1}||$. We are interested in minimizing $\kappa(S)$ for a given matrix A.

We note that the matrix S is not unique; we may scale S by a constant. We even have more freedom in choosing S. Let $S_1, S_2 \in \text{Mat}_{n\times n}(\mathbb{C})$ such that

$$\begin{aligned} J &= S_1^{-1}AS_1 \\ &= S_2^{-1}AS_2 \end{aligned}$$

then

$$\begin{aligned} J &= S_1^{-1}S_2JS_2^{-1}S_1 \\ &= (S_2^{-1}S_1)^{-1}J(S_2^{-1}S_1). \end{aligned}$$

It suffices to calculate the degrees of freedom (the number of free parameters) in the choice of $S = S_2^{-1}S_1$, since if given a fixed S_0 such that

$$J = S_0^{-1}AS_0,$$

then conjugating A by S_0S would also result in the matrix J, and as S_0 is invertible, it is one to one, so the degrees of freedom are preserved.

Let $\mathcal{S}(A, J)$ be the space

$$\mathcal{S}(A, J) = \{S,\ SA = JS\}$$

and let

$$\mathcal{S}^*(A, J) = \mathcal{S}(A, J) \cap GL_n(\mathbb{C})$$

where $GL_n(\mathbb{C})$ is the general linear group of complex matrices of degree n. Notice that fixing an S_0 as above, if $T \in \mathcal{S}^*(A, J)$, then we see that

$$T^{-1}AT = J \text{ and } S_0^{-1}AS_0 = J$$

and hence

$$((S_0^{-1}T)^{-1})^{-1}J(S_0^{-1}T)^{-1}) = J$$

so $\mathcal{S}^*(A,J) \subseteq S_0\mathcal{S}^*(J,J)$. Now, as noted earlier, for $S \in \mathcal{S}^*(J,J)$, conjugation of A by S_0S also results in J, so we have $\mathcal{S}^*(A,J) = S_0\mathcal{S}^*(J,J)$. Thus, it suffices to calculate the degrees of freedom for $\mathcal{S}^*(J,J)$. We make a further observation that the degrees of freedom belonging to $\mathcal{S}(A,J)$ and $\mathcal{S}^*(A,J)$ are the same; the Lebesgue measure belonging to the set $\mathcal{GL}_n(\mathbb{C})$ is full measure. Thus, an intersection of this set with $\mathcal{S}(A,J)$ will either annihilate the entire set or preserve the same degrees of freedom; since $I \in \mathcal{S}(A,J)$, the degrees of freedom are preserved. We define the degrees of freedom of the choice in S to mean the number of free parameters in which we may choose a value for each parameter from an infinite set. We are interested in finding the degrees of freedom corresponding to the set $\mathcal{S}^*(A,J)$, or equivalently, to the set $\mathcal{S}(A,J)$, as a means to find the infimum of $\kappa(S)$ for all $S \in \mathcal{S}^*(A,J)$.

We proceed with a few lemmas.

Lemma 4.28. *Let $J_n(\lambda)$ be a Jordan block of dimension n and let $T \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. Then $TJ = JT$ if and only if T is an Upper Toeplitz matrix.*

Proof. ($\Rightarrow$) Suppose T is an Upper Toeplitz matrix. Then

$$JT = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} & a_n \\ 0 & a_0 & \ddots & \cdots & a_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_0 & a_1 \\ 0 & \cdots & 0 & 0 & a_0 \end{pmatrix}$$

$$= \begin{pmatrix} \lambda a_0 & \lambda a_1 + a_0 & \lambda a_2 + a_1 & \cdots & \lambda a_n + a_{n-1} \\ 0 & \lambda a_0 & \lambda a_1 + a_0 & \cdots & \lambda a_{n-1} + a_{n-2} \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & \lambda a_0 & \lambda a_1 + a_0 \\ 0 & \cdots & 0 & 0 & \lambda a_0 \end{pmatrix}$$

and similarly

$$TJ = \begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} & a_n \\ 0 & a_0 & \ddots & \cdots & a_{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_0 & a_1 \\ 0 & \cdots & 0 & 0 & a_0 \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

$$= \begin{pmatrix} \lambda a_0 & a_0 + \lambda a_1 & a_1 + \lambda a_2 & \cdots & a_{n-1} + \lambda a_n \\ 0 & \lambda a_0 & a_0 + \lambda a_1 & \cdots & a_{n-2} + \lambda a_{n-1} \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 0 & \lambda a_0 & a_0 + \lambda a_1 \\ 0 & \cdots & 0 & 0 & \lambda a_0 \end{pmatrix}$$

So $JT = TJ$.

($\Leftarrow$) Now suppose J commutes with T. Then

$$JT = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1}, & a_{n,n} \end{pmatrix}$$

$$= \begin{pmatrix} \lambda a_{1,1} + a_{2,1} & \lambda a_{1,2} + a_{2,2} & \cdots & \cdots & \lambda a_{1,n} + a_{2,n} \\ \lambda a_{2,1} + a_{3,1} & \lambda a_{2,2} + a_{3,2} & \cdots & \cdots & \lambda a_{2,n} + a_{3,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda a_{n-1,1} + a_{n,1} & \lambda a_{n-1,2} + a_{n,2} & \cdots & \cdots & \lambda a_{n-1,n} + a_{n,n} \\ \lambda a_{n,1} & \lambda a_{n,2} & \cdots & \cdots & \lambda a_{n,n} \end{pmatrix}$$

and

$$TJ = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n-1} & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1}, & a_{n,n} \end{pmatrix} \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

$$= \begin{pmatrix} \lambda a_{1,1} & a_{1,1} + \lambda a_{1,2} & \cdots & \cdots & a_{1,n-1} + \lambda a_{1,n} \\ \lambda a_{2,1} & a_{2,1} + \lambda a_{2,2} & \cdots & \cdots & a_{2,n-1} + \lambda a_{2,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda a_{n-1,1} & a_{n-1,1} + \lambda a_{n-1,2} & \cdots & a_{n-1,n-2} + \lambda a_{n-1,n-1} & a_{n-1,n-1} + \lambda a_{n-1,n} \\ \lambda a_{n,1} & a_{n,1} + \lambda a_{n,2} & \cdots & a_{n,n-2} + \lambda a_{n,n-1} & a_{n,n-1} + \lambda a_{n,n} \end{pmatrix}$$

Equating the coefficients completes the proof. □

We generalize Lemma 4.28 from a single Jordan block to any matrix in Jordan Canonical Form. Let $\{J_{n_i}^{\lambda_i}\}_{i=1}^m$ be Jordan blocks where $J_{n_i}^{\lambda_i}$ has dimension n_i and eigenvalue λ_i. Let $J = \bigoplus_{i=1}^m J_{n_i}^{\lambda_i}$. Define the matrix $T_{i,j}$ to be a square matrix of dimensions equal to $\min(n_i, n_j)$ such that $T_{i,j}$ is an upper Toeplitz matrix if $\lambda_i = \lambda_j$ and a zero matrix otherwise. Now define $\widetilde{T}_{i,j}$ to be an $n_i \times n_j$ matrix obtained by setting the entries of the top right $\min(n_i, n_j) \times \min(n_i, n_j)$ square equal to the corresponding entries of $T_{i,j}$ and setting all other entries equal to 0. Finally, consider

$$T = \begin{pmatrix} \widetilde{T}_{1,1} & \widetilde{T}_{1,2} & \cdots & \widetilde{T}_{1,m} \\ \widetilde{T}_{2,1} & \widetilde{T}_{2,2} & \cdots & \widetilde{T}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{T}_{m,1} & \widetilde{T}_{m,2} & \cdots & \widetilde{T}_{m,m} \end{pmatrix}$$

We can now state our result:

Proposition 4.29. *A square matrix commutes with J if and only if it is in the same form as T.*

Proof. Let J be an $n \times n$ matrix.

($\Leftarrow$) Suppose T is the matrix defined above. One can verify computationally that,

$$TJ = \begin{pmatrix} \widetilde{T}_{1,1} & \widetilde{T}_{1,2} & \cdots & \widetilde{T}_{1,m} \\ \widetilde{T}_{2,1} & \widetilde{T}_{2,2} & \cdots & \widetilde{T}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{T}_{m,1} & \widetilde{T}_{m,2} & \cdots & \widetilde{T}_{m,m} \end{pmatrix} \begin{pmatrix} J_{n_1}^{\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{n_2}^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_m}^{\lambda_m} \end{pmatrix}$$

$$= \begin{pmatrix} \widetilde{T}_{1,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{1,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{1,m}J_{n_m}^{\lambda_m} \\ \widetilde{T}_{2,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{2,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{2,m}J_{n_m}^{\lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{T}_{m,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{m,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{m,m}J_{n_m}^{\lambda_m} \end{pmatrix}$$

and

$$JT = \begin{pmatrix} J_{n_1}^{\lambda_1} & 0 & \cdots & 0 \\ 0 & J_{n_2}^{\lambda_2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_{n_m}^{\lambda_m} \end{pmatrix} \begin{pmatrix} \widetilde{T}_{1,1} & \widetilde{T}_{1,2} & \cdots & \widetilde{T}_{1,m} \\ \widetilde{T}_{2,1} & \widetilde{T}_{2,2} & \cdots & \widetilde{T}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{T}_{m,1} & \widetilde{T}_{m,2} & \cdots & \widetilde{T}_{m,m} \end{pmatrix}$$

$$= \begin{pmatrix} J_{n_1}^{\lambda_1}\widetilde{T}_{1,1} & J_{n_1}^{\lambda_1}\widetilde{T}_{1,2} & \cdots & J_{n_1}^{\lambda_1}\widetilde{T}_{1,m} \\ J_{n_2}^{\lambda_2}\widetilde{T}_{2,1} & J_{n_2}^{\lambda_2}\widetilde{T}_{2,2} & \cdots & J_{n_2}^{\lambda_2}\widetilde{T}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ J_{n_m}^{\lambda_m}\widetilde{T}_{m,1} & J_{n_m}^{\lambda_m}\widetilde{T}_{m,2} & \cdots & J_{n_m}^{\lambda_m}\widetilde{T}_{m,m} \end{pmatrix}.$$

Note that, for any i and j, $1 \leq i, j \leq m$, we have that $J_{n_i}^{\lambda_i}\widetilde{T}_{i,j}$ and $\widetilde{T}_{i,j}J_{n_j}^{\lambda_j}$ are both $n_i \times n_j$ matrices (note that this multiplication of matrices depends only on the structure of the matrix J, as it has matrix blocks along the diagonal and zeros elsewhere, so we may later partition an arbitrary matrix and multiply that matrix with J in a similar manner). Thus, we need only verify that $J_{n_i}^{\lambda_i}\widetilde{T}_{i,j} = \widetilde{T}_{i,j}J_{n_j}^{\lambda_j}$.

There are two cases: if $\lambda_i \neq \lambda_j$, then $\widetilde{T}_{i,j} = 0$, so

$$J_{n_i}^{\lambda_i}\widetilde{T}_{i,j} = \widetilde{T}_{i,j}J_{n_j}^{\lambda_j} = 0.$$

If $\lambda_i = \lambda_j$, then we note that multiplying $J_{n_i}^{\lambda_i}\widetilde{T}_{i,j}$ or $\widetilde{T}_{i,j}J_{n_j}^{\lambda_j}$, the resulting matrix will have zeros in all entries outside of the top right $\min(n_i, n_j) \times \min(n_i, n_j)$ square, and the $\min(n_i, n_j) \times \min(n_i, n_j)$ square will have entries equal to the corresponding entries of the matrices $J_{n_i}^{\lambda_i}T_{i,j}$ and $T_{i,j}J_{n_j}^{\lambda_j}$. By the Lemma 4.28, these matrices are equal, so

$$J_{n_i}^{\lambda_i}\widetilde{T}_{i,j} = \widetilde{T}_{i,j}J_{n_j}^{\lambda_j}.$$

Since for each partition $J_{n_i}^{\lambda_i}\widetilde{T}_{i,j}$ and $\widetilde{T}_{i,j}J_{n_j}^{\lambda_j}$ are equal, and they are partitions of the same size in the same part of the matrices TJ and JT, we have shown that $TJ = JT$.

($\Rightarrow$) Now let T be any arbitrary $n \times n$ matrix and partition T as above into $T_{1,1}, \ldots, T_{m,m}$. Then as we commented earlier, the structure of J allows us to multiply the Jordan blocks of J with these partitions of T, thus

$$JT = \begin{pmatrix} J_{n_1}^{\lambda_1}\widetilde{T}_{1,1} & J_{n_1}^{\lambda_1}\widetilde{T}_{1,2} & \cdots & J_{n_1}^{\lambda_1}\widetilde{T}_{1,m} \\ J_{n_2}^{\lambda_2}\widetilde{T}_{2,1} & J_{n_2}^{\lambda_2}\widetilde{T}_{2,2} & \cdots & J_{n_2}^{\lambda_2}\widetilde{T}_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ J_{n_m}^{\lambda_m}\widetilde{T}_{m,1} & J_{n_m}^{\lambda_m}\widetilde{T}_{m,2} & \cdots & J_{n_m}^{\lambda_m}\widetilde{T}_{m,m} \end{pmatrix} \text{ and}$$

$$TJ = \begin{pmatrix} \widetilde{T}_{1,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{1,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{1,m}J_{n_m}^{\lambda_m} \\ \widetilde{T}_{2,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{2,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{2,m}J_{n_m}^{\lambda_m} \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{T}_{m,1}J_{n_1}^{\lambda_1} & \widetilde{T}_{m,2}J_{n_2}^{\lambda_2} & \cdots & \widetilde{T}_{m,m}J_{n_m}^{\lambda_m} \end{pmatrix}.$$

Now, by previous comments, $JT = TJ$ only when the corresponding partitions are equal, that is, when $J_{n_i}^{\lambda_i}\widetilde{T}_{i,j} = \widetilde{T}_{i,j}J_{n_j}^{\lambda_j}$. Computing the matrix product and equating the entries, we find that T has the desired form. □

Note that each $\widetilde{T}_{i,j}$ has an embedded upper Toeplitz matrix $T_{i,j}$ with a number of degrees of freedom equal to its dimension. Therefore, T has d degrees of freedom where

$$d = \sum_{T_{i,j} \neq 0} \dim(T_{i,j}).$$

Also note that reordering the Jordan blocks in the matrix J does not change the number of degrees of freedom. Given this fact, we can reorder the Jordan blocks so that all blocks with nondistinct eigenvalues are arranged together in increasing block size; this will allow us to formulate an equation on the degrees of freedom of the space $\mathcal{S}(J, J)$.

Theorem 4.30. *Let $\lambda_1, \ldots, \lambda_\ell$ be the distinct eigenvalues of a matrix A. Let $n_{i,1}, n_{i,2} \ldots, n_{i,k_i}$ be the size of the Jordan blocks associated to the eigenvalue λ_i such that $n_{i,1} \leq n_{i,2} \leq \ldots \leq n_{i,k_i}$. Then the number of degrees of freedom of $\mathcal{S}(A, J)$, d, is given by*

$$d = \sum_{i=1}^{p} \sum_{j=1}^{k_i} n_{i,j}(2k_i - (2j - 1)).$$

Proof. Let J be a Jordan Canonical Form of A such that the order of the Jordan blocks along the diagonal is $J^{\lambda_1}_{n_{1,1}}, J^{\lambda_1}_{n_{1,2}}, \ldots, J^{\lambda_1}_{n_{1,k_1}}, J^{\lambda_2}_{n_{2,1}}, \ldots, J^{\lambda_2}_{n_{2,k_2}}, \ldots, J^{\lambda_\ell}_{n_{\ell,k_\ell}}$, that is, Jordan blocks with nondistinct eigenvalues are arranged consecutively in ascending block size. It suffices to calculate the degrees of freedom of $\mathcal{S}(J, J)$ since $\mathcal{S}(A, J)$ has the same degrees of freedom. By Proposition 4.29, for $S \in \mathcal{S}(A, J)$, we see that the only nonzero partitions of S are in the same of the square partitions corresponding to $J^{\lambda_1}_{n_{1,1}}, \ldots, J^{\lambda_\ell}_{n_{\ell,k_\ell}}$. Thus, we need only to consider the "fitted" Upper Toeplitz matrices in the positions of these partitions; it suffices to consider an arbitrary eigenvalue and its associated Jordan blocks, and then sum up the deg! rees of freedom contributed by all eigenvalues.

Consider λ_i, $1 \leq i \leq \ell$. Since the size of the Jordan blocks are arranged in ascending order, we see that the first Jordan block, with size $n_{i,1}$, by Proposition 4.29, results in the commuting matrix having $n_{i,1}(2k_i - 1)$ degrees of freedom. In a similar manner, the j^{th} Jordan block has size $n_{i,j}$ and results in $n_{i,j}(2k_i - (2j - 1))$ more degrees of freedom. Thus, λ_i contributes to the commuting matrix

$$\sum_{j=1}^{k_i} n_{i,j}(2k_i - (2j - 1))$$

degrees of freedom. Summing over all eigenvalues, we obtain that

$$d = \sum_{i=1}^{p} \sum_{j=1}^{k_i} n_{i,j}(2k_i - (2j - 1))$$

completing the proof. ∎

Ultimately, we want the $L = \inf_{S\in\mathcal{S}^*(A,J)} \kappa(S)$ to depend on n, or some properties of A, such that we can improve our bounds on the ϵ - pseudospectrum. Our results were an attempt to understand the space $\mathcal{S}^*(A, J)$ in terms of degrees of freedom, though we have not yet completed our analysis. We do note, however, that our result has significantly simplified the calculation of L for the Jordan form of J used in Proposition 4.30: $S \in \mathcal{S}^*(A, J)$ is also a matrix with blocks along the diagonal (corresponding to the Jordan blocks of J), and so by Lemma 4.12, we need only use Proposition 4.29 on each block of the matrix S corresponding to each distinct eigenvalue Jordan blocks, compute their norm, and find the maximum over each distinct eigenvalue partition. Since we have found the degrees of freedom of each distinct eigenvalue partition, this will increase the efficiency of computationally finding L.

References

[1] Chatelin, F., Braconnier, T., *About the qualitative computation of Jordan forms* Z. Angew. Math. Mech., 74 no. 2, 105–113 (1994).

[2] Contedini, M., Embree, M.,Trefethen, L. N., *Spectra, pseudospectra, and localization for random bidiagonal matrices*, Comm. Pure Appl. Math. 54, 595–623 (2001).

[3] Davies, E.B., *Pseudo-spectra, the harmonic oscillator and complex resonances*, The Royal Society, Proc. Mathematical, Physical and Engineering Sciences, Vol. 455, No. 1982, 585-599 (1999).

[4] Dummit and Foote

[5] Embree, M.,Trefethen, L.N., *Generalizing Eigenvalue Theorems to Pseudospectra Theorems.* SIAM J. Sci. Comput. Vol. 23, No. 2, 583-590 (2001).

[6] Goldsheid, I.Y., Khoruzhenko, B.: *Eigenvalue curve of asymmetric tridiagonal random matrices.* Electr.Journ.Prob. 5(16), 1–28 (2000).

[7] Goldsheid, I.Y., Khoruzhenko, B.: *Regular spacings of complex eigenvalues in the one-dimensional non-Hermitian Anderson model.* Commun. Math. Phys. 238, 505–524 (2003).

[8] Hatano, N., Nelson, D.R.: *Localization transitions in non-Hermitian quantum mechanics.* Phys. Rev. Lett. 77, 570-573 (1996).

[9] Hatano, N., Nelson, D.R.: *Vortex Pinning and Non-Hermitian quantum mechanics.* Phys. Rev. B56, 8651–8673 (1997).

[10] Horn, R., Johnson, C., *Matrix Analysis*, Cambridge University Press, New York, (1985).

[11] Kato, T., *Perturbation theory for linear operators. Reprint of the 1980 edition*, Classics in Mathematics, Springer-Verlag, Berlin, xxii+619 pp, (1995).

[12] Shnreb, N.M., Nelson, D.R.:*Non-Hermitian localization and population biology.* Phys. Rev. B58, 1383-1403 (1998).

[13] Teschl, Gerald, *Functional Analysis*, available online at http://www.mat.univie.ac.at/ gerald/ftp/book-fa/.

[14] Trefethen, L. N., Embree, M., *Spectra and pseudospectra. The behavior of nonnormal matrices and operators*, Princeton University Press, Princeton, NJ, xviii+606 pp. (2005).

Matthew Coudron
University of Minnesota
Minneapolis, MN 55199
coudr003@umn.edu

Amalia Culiuc
Mount Holyoke College
South Hadley, MA 01075
culiu22a@mtholyoke.edu

Philip Vu
Williams College
Williamstown, MA 01267
pvv1@williams.edu

Stephen Webster
Williams College
Williamstown, MA 01267
stephendwebster@gmail.com