Non-generic Eigenvalue Perturbations of Jordan Blocks

Yanyuan Ma* Alan Edelman†

December 30, 1996

Abstract

We show that if an $n \times n$ Jordan block is perturbed by an $O(\epsilon)$ upper k-Hessenberg matrix (k subdiagonals including the main diagonal), then generically the eigenvalues split into p rings of size k and one of size r (if $r \neq 0$), where $n = pk + r$. This generalizes the familiar result ($k = n, p = 1, r = 0$) that generically the eigenvalues split into a ring of size n. We compute the radii of the rings to first order and the result is generalized in a number of directions involving multiple Jordan blocks of the same size.

Keywords: Eigenvalue, Perturbation, Jordan form

AMS Classfication: 65F15

1 Introduction

Perturb an $n \times n$ Jordan block by order ϵ mathematically or through rounding errors on a computer, and typically the eigenvalues split up into a ring of radius $O(\epsilon^{1/n})$. This paper studies the non-typical behavior. We stifle the matrix's ability to form large eigenvalue rings by

*Department of Mathematics Room 2-333, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, yanyuan@math.mit.edu, http://www-math.mit.edu/~yanyuan.

†Department of Mathematics Room 2-380, Massachusetts Institute of Technology, Cambridge, MA 02139-4307, edelman@math.mit.edu, http://www-math.mit.edu/~edelman, supported by NSF grants 9501278-DMS and 9404326-CCR.

only allowing perturbations that are upper k-Hessenberg, meaning a matrix containing exactly k subdiagonals below and including the diagonal. The result we will show is that generically, the eigenvalue perturbations follow the greediest possible pattern consistent with forming no rings bigger than k. We then generalize and examine some multiple Jordan block cases.

Our interest in this problem came from a perturbation study of Ruhe's matrix [5] using the qualitative approach proposed by Chatelin and Frayssé [4]. We found that non-generic behaviors occurred some small percentage of the time. Chatelin and Frayssé themselves point out in one example [4, page 192] that only 97% of their examples follow the expected behavior. We also became interested in this problem because we wanted to understand how eigenvalues perturb if we move in some, but not all normal directions to the orbit of a matrix with a particular Jordan form such as in Arnold's versal deformation [1, 6]. Such information may be of value in identifying the nearest matrix with a given Jordan structure. Finally, we point out, that the ϵ-pseudo-spectra of a matrix can depend very much on the sparsity structure of the allowed perturbations. Following an example from Trefethen [10], if we take a Jordan block J and then compute in the presence of roundoff error, $A = Q^T J Q$, where Q is a banded orthogonal matrix, then the behavior of $\|A^k\|$, is quite different from what would happen if Q were dense.

It is generally known [2, page 109],[7] that if a matrix A is perturbed by any matrix ϵB, then the eigenvalues split into clusters of rings, and their expansion in ϵ is a Puisseaux series. Unfortunately, the classical references give little information as to how the eigenvalues split into clusters as a function of the sparsity structure of the perturbation matrix. It is obvious generically, that for each Jordan block of size k, the multiple eigenvalue breaks into rings of size k. Explicit determination of the coefficients of the first order term may be found in [8]. A resurrection of both this result and a Newton diagram approach originally found in [11] may be found in [9].

In this paper, we explore the case when B is upper k-Hesseberg employing some techniques by Burke and Overton [9]. For example, suppose that we perturb a 7×7 Jordan block J with a matrix ϵB, where B has the form:

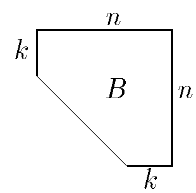

Figure 1: B = upper k-Hessenberg matrix

We assume that k denotes the number of subdiagonals (including the main diagonal itself) that is not set to zero. If B were dense, the eigenvalues would split uniformly onto a ring of radius $O(\epsilon^{\frac{1}{7}})$. However, if $k=4$, we obtain one ring of size 4 with radius $O(\epsilon^{\frac{1}{4}})$ and one ring of size 3 with radius $O(\epsilon^{\frac{1}{3}})$ as illustrated in Figure 2. Figure 3 contains a table of possible ring

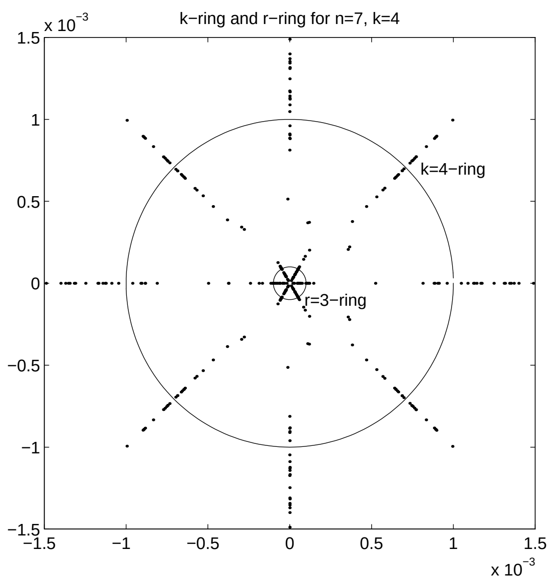

Figure 2: Example rings for $n=7$ and $k=4$. We collected eigenvalues of 50 different random $J_n(0)+\epsilon B$, $\epsilon=10^{-12}$. The figure represents 50 different copies of one 4-ring and 50 copies of one 3-ring. The two circles have radii 10^{-3} and 10^{-4}. If B were a random dense matrix, there would be only one 7-ring with radius $O(10^{-\frac{12}{7}})$.

sizes when $n=7$ for $k=1,\ldots,7$.

	ring size						
k	1	2	3	4	5	6	7
1	7						
2	1	3					
3	1		2				
4			1	1			
5		1			1		
6	1					1	
7							1

Figure 3: Table for one Jordan block of size 7. The entries in each row are the number of rings of a given size when the perturbation is upper k-Hessenberg.

Our main result is that if a Jordan block of size n is perturbed by an upper k-Hessenberg matrix, then the eigenvalues split into $\lceil \frac{n}{k} \rceil$ rings, where $p \equiv \lfloor \frac{n}{k} \rfloor$ of them are k-rings with radius $O(\epsilon^{\frac{1}{k}})$, and if k does not divide n, there is one remaining r-ring with radius $O(\epsilon^{\frac{1}{r}})$, where $r \equiv n \mod k$. Moreover, the first order perturbation of the pk eigenvalues in the k-rings only depends on the kth diagonal of B.

In Section 3, we extend these results to the case of t equally sized Jordan blocks. Let $A = \text{Diag}[J_1, J_2, ...J_t]$, where the J_i's are $n \times n$ Jordan blocks, and we conformally partition

$$B = \begin{bmatrix} B_{11} & B_{12} & \dots & B_{1t} \\ B_{21} & B_{22} & \dots & B_{2t} \\ \dots & \dots & \dots & \dots \\ B_{t1} & B_{t2} & \dots & B_{tt} \end{bmatrix}.$$

Suppose every B_{ij} is an upper k-Hessenberg matrix. We will show in Theorem 3 that generically, the eigenvalues break into $t\lceil \frac{n}{k} \rceil$ rings, tp of them are k-rings and the remaining t are r-rings if k does not divide n. Here, p and r has the same meaning as before. Again, the first order perturbation of the first tpk eigenvalues only depends on the kth diagonal of every B_{ij}.

For example, if

$$A = J_7(\lambda_1) \oplus J_7(\lambda_2),$$

so that $n = 7$ and $t = 2$, our block upper k-Hessenberg matrices have the form

$$B = \left[\begin{array}{ccccccc|ccccccc}
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
0 & * & * & * & * & * & * & 0 & * & * & * & * & * & * \\
0 & 0 & * & * & * & * & * & 0 & 0 & * & * & * & * & * \\
0 & 0 & 0 & * & * & * & * & 0 & 0 & 0 & * & * & * & * \\
0 & 0 & 0 & 0 & * & * & * & 0 & 0 & 0 & 0 & * & * & * \\
\hline
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * & * & * & * & * & * \\
0 & * & * & * & * & * & * & 0 & * & * & * & * & * & * \\
0 & 0 & * & * & * & * & * & 0 & 0 & * & * & * & * & * \\
0 & 0 & 0 & * & * & * & * & 0 & 0 & 0 & * & * & * & * \\
0 & 0 & 0 & 0 & * & * & * & 0 & 0 & 0 & 0 & * & * & *
\end{array}\right],$$

i.e. $k = 3$, hence $r = 1$.

In this case, the eigenvalues of $A + \epsilon B$ will split into two 3-rings centered at λ_1 with radii $O(\epsilon^{\frac{1}{3}})$, two 3-rings centered at λ_2 with radii $O(\epsilon^{\frac{1}{3}})$, one 1-ring centered at λ_1 with radius $O(\epsilon)$, and one 1-ring centered at λ_2 with radius $O(\epsilon)$. See Figure 4 for a list of possible rings when $k = 1, \ldots, 7$ and $n = 7$.

	ring size						
k	1	2	3	4	5	6	7
1	14						
2	2	6					
3	2		4				
4			2	2			
5		2			2		
6	2					2	
7							2

Figure 4: Table for two blocks, column index represents size of rings and row index value of k. Entries are number of rings.

2 One Block Case

Suppose that the Jordan form of A is simply one Jordan block. We assume that $A = J_n(0)$, which we will perturb with ϵB, where B has the sparsity structure

$$B = \begin{array}{r} k \left\{ \vphantom{\begin{matrix}1\\1\\1\\1\\1\end{matrix}} \right. \\ n-k \left\{ \vphantom{\begin{matrix}1\\1\\1\\1\\1\end{matrix}} \right. \end{array} \begin{bmatrix} \ast & \ast & . & . & . & . & . & . & . & \ast \\ . & & & & & & & & & . \\ . & & & & & & & & & . \\ . & & & & & & & & & . \\ \ast & & & & & & & & & . \\ 0 & \ast & & & & & & & & . \\ 0 & 0 & . & & & & & & & . \\ . & & . & . & & & & & & . \\ . & & & . & . & & & & & . \\ 0 & . & . & . & 0 & \ast & . & . & . & \ast \end{bmatrix}.$$

Definition 1 *Suppose a matrix has k subdiagonals that are closest to the main diagonal (including the main diagonal), not zero, then we call the matrix an **upper k-Hessenberg matrix**.*

Definition 2 *Suppose for ϵ sufficiently small,*

$$\lambda_j = \lambda + c\epsilon^{\frac{1}{k}}\omega^j + o(\epsilon^{\frac{1}{k}}),$$

*for $j = 0, 1, \ldots, k-1$. We then refer to the set $\{\lambda_1(\epsilon), \ldots, \lambda_k(\epsilon)\}$ as a k-**ring**. Here $\omega = e^{\frac{2\pi i}{k}}$ and we refer to $c \neq 0$ as the ring constant.*

Lemma 1 *[7, page 65] Let λ be a multiple eigenvalue of A with multiplicity s, then there will be s eigenvalues of $A + \epsilon B$ grouped in the manner $\{\lambda_{11}(\epsilon), \ldots, \lambda_{1s_1}(\epsilon)\}$, $\{\lambda_{21}(\epsilon), \ldots, \lambda_{2s_2}(\epsilon)\}$, $\ldots$, and in each group i, the eigenvalues admit the Puiseux series*

$$\lambda_{ih}(\epsilon) = \lambda + \alpha_{i1}\omega_i^h\epsilon^{\frac{1}{s_i}} + \alpha_{i2}\omega_i^{2h}\epsilon^{\frac{2}{s_i}} + \ldots$$

for $h = 1, \ldots, s_i$. Here $\omega_i = e^{\frac{2\pi i}{s_i}}$.

Our Theorem 1 shows how the eigenvalues split into rings, and in Theorem 2 we analyze the ring constant c.

Theorem 1 *Let A, B, $n-k$ and n be given as above. Let r be the remainder of n divided by k, i.e. $n = pk + r$, $0 \leq r < k$. The eigenvalues of $A + \epsilon B$ will then generically split into a) p k-rings and b) one r-ring if $r \neq 0$.*

Proof:

In part a) of our proof, we show that the eigenvalues split into p k-rings. In part b), we prove the statement about the possible existence of one r-ring.

Part a:

First, we study the p k-rings. In this case, we proceed to show by a change of variables that in fact, only the lowest subdiagonal plays a role in the first order perturbation theory.

Let $\lambda = \mu\epsilon^{\frac{1}{k}}$ and $z = \epsilon^{\frac{1}{k}}$. Let

$$L_1 = \text{diag}[z^{-1}, z^{-2}, \ldots, z^{-n}] \tag{1}$$

and

$$R_1 = \mathrm{diag}[1, z^1, \ldots, z^{n-1}] \tag{2}$$

be scaling matrices. Consider $M(z) \equiv L_1(\lambda I - J - \epsilon B)R_1 = (L_1(\mu z I - J - z^k B)R_1)$. At $z = 0$, it has the form

$$M(0) = \begin{bmatrix} \mu & -1 & & & & & \\ & \mu & -1 & & & & \\ & & \cdot & \cdot & & & \\ * & & & \cdot & \cdot & & \\ & * & & & \cdot & \cdot & \\ & & * & & & \mu & -1 \\ & & & * & & & \mu \end{bmatrix}. \tag{3}$$

$M(0)$ has only three diagonals, and the kth subdiagonal has the original entries of B on it.

We claim that $f(\mu) \equiv \det(M(0))$ has the form $\mu^r q(\mu^k)$, where $q(\cdot)$ is a polynomial of order p and its constant term does not vanish generically. Let

$$\omega = e^{2i\pi/k}. \tag{4}$$

Let L_1' be L_1 with z replaced by ω, i.e.

$$L_1' = \mathrm{diag}[\omega^{-1}, \omega^{-2}, \ldots, \omega^{-n}],$$

and let R_1' be R_1 with z replaced by ω, i.e.

$$R_1' = \mathrm{diag}[1, \omega^1, \ldots, \omega^{n-1}],$$

then

$$\omega^{-n} f(\omega\mu) = \omega^{-1-2-\cdots-n} f(\omega\mu)\omega^{0+1+\cdots+(n-1)}$$

$$= \det \left(L_1' \begin{bmatrix} \omega\mu & -1 & & & & \\ & \omega\mu & -1 & & & \\ & & \ddots & \ddots & & \\ * & & & \ddots & \ddots & \\ & * & & & \ddots & \ddots \\ & & * & & \omega\mu & -1 \\ & & & * & & \omega\mu \end{bmatrix} R_1' \right) = \det \begin{bmatrix} \mu & -1 & & & & \\ & \mu & -1 & & & \\ & & \ddots & \ddots & & \\ * & & & \ddots & \ddots & \\ & * & & & \ddots & \ddots \\ & & * & & \mu & -1 \\ & & & * & & \mu \end{bmatrix} = f(\mu).$$

Therefore $f(\mu) = \omega^{-n} f(\omega\mu) = \omega^{-r} f(\omega\mu)$, from which we can see $f(\mu)$ must be of the form $\mu^r q(\mu^k)$.

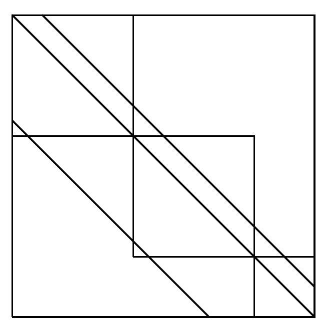

Figure 5: M after it is divided

We now check that the extreme terms of $f(\cdot)$, of degree n and r, do not vanish. The product of the diagonal entries gives the highest order term μ^n in $f(\mu)$. Now consider the μ^r term. Divide the matrix M into p $k \times k$ diagonal blocks and one $r \times r$ block as in Figure 5. We show that the μ^r term generically does not vanish by considering the coefficient of μ^r term of $f(\mu)$ as a polynomial in the $*$'s entries. In the first p blocks, we take all the -1's and the one element at the left bottom corner of every block, in the last block, we take all the μ's, generically, we will get one nonzero coefficient term, hence generically, the coefficient polynomial will not vanish. This proves the claim.

Since the polynomial $q(\cdot)$ has p nonzero roots, which we denote $c_1, c_2, \ldots, c_p$, then the polynomial $f(\mu) = \mu^r q(\mu^k)$ has pk non-zero roots distributed evenly on p circles. From the implicit function theorem, there are pk roots of the determinant of the original matrix near $z = 0$ that have the form $\sqrt[k]{c_i} + o(z)$, for $i = 1, 2 \ldots, p$. Note that $\sqrt[k]{c_i}$ yields k different values $\omega^0, \ldots, \omega^{k-1}$ for every i. This shows the pk eigenvalues form p k-rings, this completes the first half of the proof.

Part b:

We now investigate the remaining r eigenvalues when $r > 0$. Consider the determinant of $J + \epsilon B$, which equals the product of all the eigenvalues. By partitioning $J + \epsilon B$ as in Figure 5, and picking the entries similarly in the first p blocks too, while in the last block, still picking all the -1's and the left bottom corner element, we get an $O(\epsilon^{p+1})$ term generically. Therefore generically, $\det(J + \epsilon B) \geq O(\epsilon^{p+1})$. Since the product of the first pk eigenvalues $= O(\epsilon^p)$, the product of the remaining r eigenvalues $= O(\epsilon)$. From Lemma 1, they form an r-ring. This completes the proof of Theorem 1. $\square$

The analysis of the last r eigenvalues, while straightforward in the one block case, is more difficult when we proceed to the next section. We therefore provide an alternative proof which generalizes.

Alternative proof of Theorem 1 Part b:

Let

$$L_2 = \mathrm{diag}[D_L, z^{-r} I_{k-r}, z^{-r} D_L, z^{-2r} I_{k-r}, \ldots, z^{-pr} D_L], \tag{5}$$

where

$$D_L = \mathrm{diag}[z^{-1}, \ldots, z^{-r}].$$

Let

$$R_2 = \mathrm{diag}[D_R, z^r I_{k-r}, z^r D_R, z^{2r} I_{k-r}, \ldots, z^{pr} D_R], \tag{6}$$

where

$$D_R = \mathrm{diag}[z^0, \ldots, z^{r-1}].$$

Also as with our original proof, we make a change of variables by setting $\lambda = \mu z$ and $z = \epsilon^{\frac{1}{r}}$. Let

$$N(z) = L_2(\lambda I - J - \epsilon B)R_2 = L_2(\mu z I - J - z^r B)R_2.$$

Figure 6 illustrates $N(0)$.

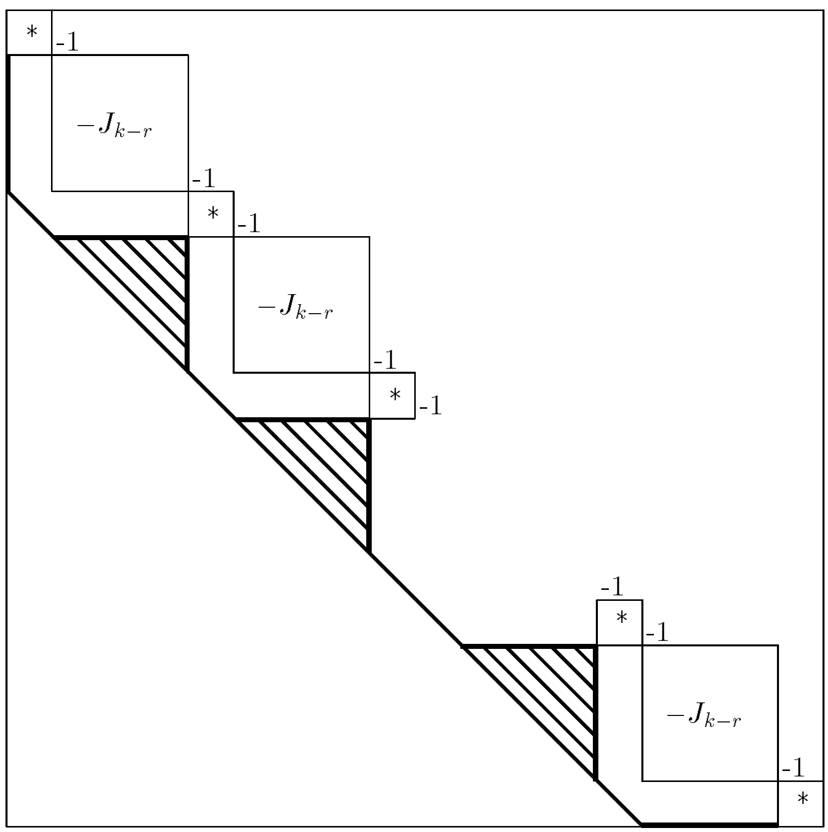

Figure 6: Shaded triangles and thick lines contain the original entries of B. Squares marked with $*$'s are $\lambda I - J_r$ blocks.

Again we claim that $g(\mu) \equiv \det(N(0))$ is a polynomial of μ^r. Let

$$\omega = e^{\frac{2\pi i}{r}} \tag{7}$$

and let $L_2'{=}L_2$ with z replaced by ω and $R_2'{=}R_2$ with z replaced by ω. Replace μ in $g(\mu)$ by $\omega\mu$, then we get

$$g(\omega\mu) = \omega^{-r(p+1)}g(\omega\mu) = \det(L_2' N_{\omega\mu}(0) R_2') = g(\mu).$$

Here $N_{\omega\mu}(0)$ represents the matrix $N(0)$ with $\omega\mu$ instead of μ on the main diagonal. Therefore, g is a polynomial of μ^r, say $g(\mu) = h(\mu^r)$. By taking all the -1's and the left bottom element of every block, we can see the constant term is generically not zero. By taking the same entries of the first p blocks and all the μ's of the last block, we obtain a nonzero μ^r term generically. Hence there are at least r roots of $h(\mu^r)$, and they are the rth root of some constant c. By the implicit function theorem there are at least r eigenvalues having the expression $\sqrt[r]{c}\omega^j\epsilon^{\frac{1}{r}} + o(\epsilon^{\frac{1}{r}})$, with $j = 0, \ldots, r-1$. They form an r-ring. This is as many as we can get since we already have pk of the eigenvalues from Part a). $\square$

Theorem 2 *The kth power of the ring constants for the k-rings are the roots of $q(z)$, where*

$$q(z) = z^p + \alpha_1 z^{p-1} + \cdots + \alpha_i z^{p-i} + \cdots + \alpha_p, \tag{8}$$

and

$$\alpha_i = \sum_{l_{j+1}-l_j \geq k} \beta_{l_1}\beta_{l_2}\ldots\beta_{l_i} \tag{9}$$

for $i = 1, \ldots, p$, where $\beta_1, \ldots, \beta_{n-k+1}$ denote the elements on the kth diagonal of B. So long as $\alpha_p \neq 0$, we obtain the generic behavior described in Theorem 1.

This can be proved by a combinatorial argument as follows:

Proof:

Consider the matrix $M(0)$ as in Equation (3) and the μ^{n-jk} term in its determinant $f(\mu)$, where $j = 0, \ldots, p$. We write out the indices of the columns of M in a column on the left and the indices of the rows of M in a column on the right, and connect the indices iff there is a structurally nonzero entry in the position of $M(0)$ as in Figure 7. Then, every perfect

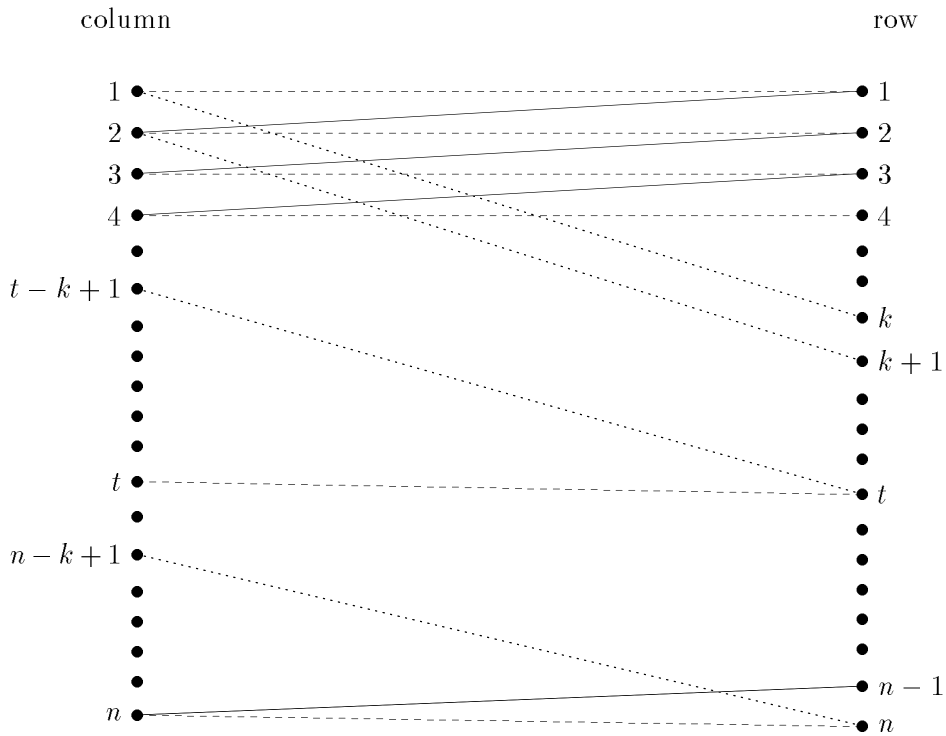

Figure 7: Bipartite graph of M. Only important lines for the argument are shown.

matching obtained from a subset of this graph will give us a term in the determinant of M. The horizontal lines correspond to the diagonal entries μ, so we are interested in matchings with $n - jk$ horizontal lines. The downward sloping lines correspond to the kth subdiagonal, the upperward sloping lines are the superdiagonal consisting of -1's. Assume that the first downward sloping line (there must be one in the perfect matching except when $j = 0$) that we pick connects column $t - k + 1$ to row t, then all the numbers before t in the second column will be paired with the same number or just one step larger in the first column, which means all the first t numbers in the second column will be paired with the first t numbers in the first column. We proceed to show that every downward sloping line is matched with $k - 1$ upward sloping lines and that the set together takes the place of k horizontal lines. To be precise, the first $t - k$ lines must be horizontal lines, and the remaining $k - 1$ are upward sloping. This leaves us with the numbers from $t + 1$ to n in both columns and the situation returns exactly to what we began with, we may proceed by induction. Also, since the downward sloping lines must be at least k apart, we have $l_{j+1} - l_j \geq k$ as in Equation (9). Also, we find that every downward sloping line is associated with $k-1$ upperward sloping lines which means a loss of k horizontal lines from a total of n. Once we succeed in finding j downward sloping lines, we will lose a total of jk horizontal lines and we get the term μ^{n-jk}. This completes the proof. $\square$

When $k > \frac{n}{2}$, Equation (9) simplifies to: the sum of the elements on the kth subdiagonal is not zero. In such cases, J. Burke and M. Overton([3, Theorem 4]) gave a general result on the characteristic polynomial of $A + \epsilon B$: the coefficient of every term $\epsilon\lambda^i$ is the sum of the elements on the $(n-i)$th subdiagonal for $i = 0, 1, \ldots, n-1$. From this theorem, if we assume that the last subdiagonal that does not sum up to zero is the kth subdiagonal, for $k > \frac{n}{2}$, using a Newton diagram (see Figure 8 for an example of a Newton diagram), it can be easily seen that the eigenvalues split into one k-ring and one $(n-k)$-ring.

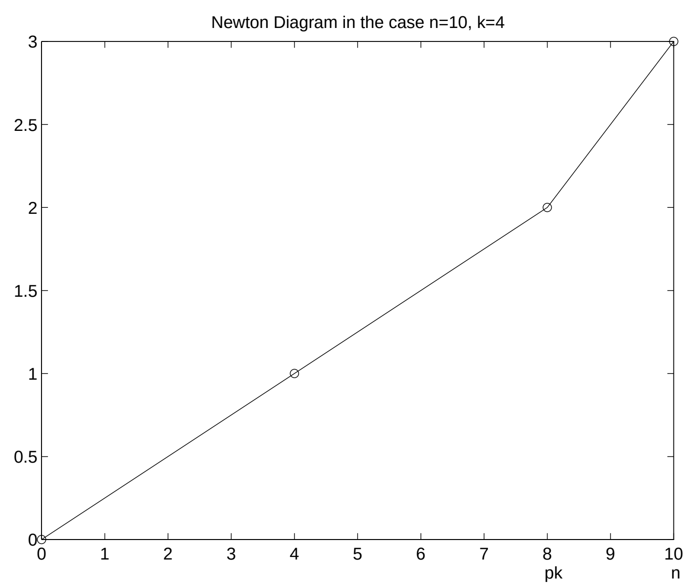

Figure 8: Newton diagram

Actually, we can argue similarly for the $k < \frac{n}{2}$ case. When the sum α_p is zero, the constant term of $q(\cdot)$ in Equation (8) is zero. Generically, this results in the loss of one k-ring. Consider the Newton diagram: the (pk, k) point moves up and the whole diagram generically breaks into three segments, one with slope $\frac{1}{k}$, of length $(p-1)k$ and one with slope $\frac{1}{k-1}$, of length $k-1$ and one with slope $\frac{1}{r+1}$, of length $r+1$. This means it has $(p-1)k$ eigenvalues forming $p-1$ k-rings and $k-1$ eigenvalues forming one $(k-1)$-ring and $r+1$ eigenvalues forming an $(r+1)$-ring. There are two special cases when this does not happen. One is when $k-1 = r+1$, then the last two segments combine into one segment. The other is when $k-1 < r+1$ which can happen when $k = r+1$, and the whole diagram breaks into only two segments, the first one remains untouched, and the second one has slope $\frac{2}{k+r}$, length $k+r$.

3 t Block Case (All B_{ij}'s are upper k-Hessenberg matrices)

We now study the case when the Jordan form of A has t blocks all with the same size n. Here, we only consider the admittedly special case where the perturbation matrix B has the block upper k-Hessenberg form: if we divide B into $n \times n$ blocks, every B_{ij} is an upper k-Hessenberg matrix. In this special case, we have

Theorem 3 *Let A, B, $n-k$ and n be given as above and let r be the remainder of n divided by k, i.e. $n = pk + r$, $0 \le r < k$. The eigenvalues of $A + \epsilon B$ will then split into tp k-rings and t r-rings if $r \neq 0$.*

Proof: The proof follows closely that of the proof of Theorem 1.

Part a:

Let L_1 be a block diagonal matrix with t blocks and every block has the form as in Equation (1). Let R_1 be a block diagonal matrix with t blocks and every block has the form in Equation (2). Let

$$M(z) = L_1(\lambda I - J - \epsilon B)R_1.$$

Then $M(0)$ breaks into t^2 $n \times n$ blocks. All of the diagonal blocks have the same form as in Equation (3) and the form of the off diagonal blocks results from replacing the μ's and -1's with 0's in the diagonal blocks. Call the resulting matrix $M(0)$. We can reach the same claim that

$$f(\mu) \equiv \det(M(0)) = \mu^{r_0} q(\mu^k),$$

where $r_0 \equiv nt \mod k$. By considering the diagonal blocks we can see that generically the terms μ^{nt} and μ^{rt} appear. This can be shown simply by using the same ω as in Equation (4) and constructing L_1' and R_1' by replacing the z's in L_1 and R_1 with ω's, and going through exactly the same procedure. Thus, we will have at least $nt - rt = tpk$ eigenvalues yielding the form $\sqrt[k]{c_i}\epsilon^{\frac{1}{k}} + o(\epsilon^{\frac{1}{k}})$, $i = 1, ...tp$. Note that every $\sqrt[k]{c_i}$ gives k values. They form pt k-rings.

Part b:

Let L_2 be a block diagonal matrix with t blocks and every block has the form as in Equation (5). Let R_2 be a block diagonal matrix with t blocks and every block has the form as in Equation (6). Let

$$N(z) = L_2(\lambda I - J - \epsilon B)R_2.$$

Then $N(0)$ breaks into t^2 $n \times n$ blocks. All of the diagonal blocks have the same form as $N(0)$ in Figure 6 and the form of the off diagonal blocks results from replacing the μ's and -1's with 0's in the diagonal blocks. We can reach the same claim that

$$g(\mu) \equiv \det(N(0)) = h(\mu^r)$$

and by considering the diagonal blocks we can see generically the term μ^0 and μ^{rt} appear. This can be shown simply by using the same ω as in Equation (7) and construct L_2' and R_2' by replacing the z's in L_2 and R_2 with ω's, and going through exactly the same procedure. Thus, we will have at least rt eigenvalues yielding the form $\sqrt[r]{c_l}\epsilon^{\frac{1}{r}} + o(\epsilon^{\frac{1}{r}})$, here $l = 1, ...t$. Note that every $\sqrt[r]{c_l}$ gives r values. They form t r-rings.

Since the matrix $J + \epsilon B$ has only nt eigenvalues, it must have exactly tpk and tr of each. This completes the proof of the theorem. $\square$

4 t block case (Every B_{ij} is an upper K_{ij}-Hessenberg matrix)

When the number of subdiagonals in each B_{ij} differs, the situation becomes much more complicated, the general problem remains open. We have some observations in four special cases. Let K_{ij} = the number of subdiagonals of B_{ij}, for $1 \leq i, j \leq n$, i.e., B_{ij} is an upper K_{ij}-Hessenberg matrix.

Theorem 4

Case 1:

Let $K_{\max} = \max(K_{ij})$, $i = 1, \ldots, n$, $j = 1, \ldots, n$. If $K_{11} = K_{22} = \ldots = K_{tt} = K_{\max}$ then Theorem 3 holds by replacing k with $K_{\max}$.

Case 2:

Let $K_{\max} = \max(K_{ij})$. If we can find t K_{ij}'s equal to $K_{\max}$ s.t. no two of them are in the same row or column, then the result from Theorem 3 holds for $K_{\max}$.

Case 3:

When $K_{ii} \geq K_{ij}$, $K_{ii} \geq K_{ji}$ for all i and j, and $K_{ii} \geq \frac{n}{2}$, then the resulting eigenvalue behavior looks like putting the t diagonal blocks together, i.e, $J + \epsilon B$ has K_{ii} eigenvalues that form one K_{ii}-ring for $i = 1, \ldots, t$. It also has $n - K_{ii}$ eigenvalues that form one $(n - K_{ii})$-ring for $i = 1, \ldots, t$.

Case 4:

If we can find t numbers $K_{i_1 j_1}, K_{i_2 j_2}, \ldots, K_{i_t j_t}$, all $\geq \frac{n}{2}$, such that $K_{i_s j_s} \geq K_{i_s l}$ and $K_{i_s j_s} \geq K_{m j_s}$ for any l and m, $s = 1, \ldots, t$, and $i_s \neq i_{s'}$, $j_s \neq j_{s'}$, when $s \neq s'$, then the results in Case 3 hold.

Proof of Case 1:

This can be checked simply by replacing all of the K_{ij}'s with $K_{\max}$ and noticing that the proof of Theorem 3 is still valid, in that the genericity condition is the same even if some of the off diagonal entries are zero. $\square$

Proof of Case 2:

We also replace all the K_{ij}'s with $K_{\max}$. The proof of Theorem 3 remains valid with a minor modification. While some terms in $p(\mu)$ may be 0 in one block, one can always obtain non-zero terms in each block row and column in the block with $K_{ij} = K_{\max}$. This will guarantee the same nonzero terms generically. $\square$

The following is an example where $t = 2$:

$$
\left[\begin{array}{cccccccc|cccccccc}
\mu & -1 & & & & & & & & & & & & & & \\
 & \mu & -1 & & & & & & & & & & & & & \\
\clubsuit & & \mu & -1 & & & & & & \heartsuit & & & & & & \\
 & \clubsuit & & \mu & -1 & & & & & & \heartsuit & & & & & \\
 & & \clubsuit & & \mu & -1 & & & & & & \heartsuit & & & & \\
 & & & \clubsuit & & \mu & -1 & & & & & & \heartsuit & & & \\
 & & & & \clubsuit & & \mu & -1 & & & & & & \heartsuit & & \\
 & & & & & \clubsuit & & \mu & & & & & & & \heartsuit & \\
\hline
 & & & & & & & & \mu & -1 & & & & & & \\
 & & & & & & & & & \mu & -1 & & & & & \\
 & \heartsuit & & & & & & & \clubsuit & & \mu & -1 & & & & \\
 & & \heartsuit & & & & & & & \clubsuit & & \mu & -1 & & & \\
 & & & \heartsuit & & & & & & & \clubsuit & & \mu & -1 & & \\
 & & & & \heartsuit & & & & & & & \clubsuit & & \mu & -1 & \\
 & & & & & \heartsuit & & & & & & & \clubsuit & & \mu & -1 \\
 & & & & & & \heartsuit & & & & & & & \clubsuit & & \mu
\end{array}\right]
$$

This is an example in which instead of taking $\clubsuit$'s which may be all zeros, we take $\heartsuit$'s which are nonzero generically.

Proof of Case 3:

For any K_{ii}, let L_{1_i} be a diagonal matrix formed by t blocks of size $n \times n$. For block j, if $K_{jj} \leq K_{ii}$, then the block will be

$$\mathrm{diag}[z^{-1}, z^{-2}, ... z^{-n}],$$

if $K_{jj} \geq K_{ii}$, then the block will be

$$\mathrm{diag}[D_{l_i}, z^{-n+K_{ii}} I_{K_{ii}-n+K_{jj}}, z^{-n+K_{ii}} D_{l_j}], \tag{10}$$

where

$$D_{l_i} = \mathrm{diag}[z^{-1}, \ldots, z^{-n+K_{ii}}] \tag{11}$$

and

$$D_{l_j} = \mathrm{diag}[z^{-1}, \ldots, z^{-n+K_{jj}}].$$

Let R_{1_i} be a diagonal matrix formed by t blocks of size $n \times n$. For block j, if $K_{jj} \leq K_{ii}$, then the block will be

$$\mathrm{diag}[z^0, z^1, ... z^{n-1}],$$

if $K_{jj} \geq K_{ii}$, then the block will be

$$\mathrm{diag}[D_{r_i}, z^{K_{ii}} I_{K_{ii}+K_{jj}-n}, z^{K_{ii}} D_{r_{ij}}], \tag{12}$$

where

$$D_{r_i} = \mathrm{diag}[z^0, \ldots, z^{K_{ii}-1}], \tag{13}$$

and

$$D_{r_{ij}} = \mathrm{diag}[z^0, \ldots, z^{2n-K_{ii}-K_{jj}-1}].$$

Let $\lambda = \mu z$, $z = \epsilon^{\frac{1}{K_{ii}}}$ and $M_i(z) = L_{1_i}(\lambda I - J - \epsilon B)R_{1_i}$. Then $M_i(0)$ is a $t \times t$ block matrix where the jth diagonal block looks like either the $M(0)$ in Equation (3) for $k = K_{ii}$ or the block has the form

$$\begin{bmatrix} \mu & -1 & & & & & & & & & & \\ & \cdot & \cdot & & & & & & & & & \\ & & \mu & -1 & & & & & & & & \\ * & & & 0 & -1 & & & & & & & \\ * & & & & 0 & -1 & & & & & & \\ \cdot & & & & & \cdot & \cdot & & & & & \\ \cdot & & & & & & \cdot & \cdot & & & & \\ * & & & & & & & 0 & 1 & & & \\ & * & & & & & & & \mu & -1 & & \\ & & \cdot & & & & & & & \cdot & \cdot & \\ & & & \cdot & & & & & & & \cdot & \cdot \\ & & & & * & & & & & & & \mu \end{bmatrix}$$

Here, the $*$'s on the first column appear from the K_{ii}th row to K_{jj}th row. The off diagonal block $M(0)_{i_{lm}}$ looks the same as in Equation (3) with $k = \max(l, m)$. Replacing z in L_{1_i} and R_{1_i} with ω, which is $e^{\frac{1}{K_{ii}}}$, to get L'_{1_i} and R'_{1_i}, we get the same conclusion that $\det(M_i(0))$ is of the form $\mu^{n-K_{ii}}p(\mu^{K_{ii}})$ and by extracting the constant terms from the diagonal blocks with the new form above and the $\mu^{n-K_{ii}}$ terms and the μ^n terms from all the other diagonal blocks, we get the result that there are at least $t_i K_{ii}$ eigenvalues forming t_i K_{ii}-rings. Here, t_i is the number of times K_{ii} appears on the diagonal.

For any K_{ii}, let L_{2_i} be a diagonal matrix formed by t blocks of size $n \times n$. For block j, if $K_{jj} \leq K_{ii}$, then the block will be

$$\operatorname{diag}[D_{l_i}, z^{-n+K_{ii}} I_{2K_{ii}-n}, z^{-n+K_{ii}} D_{l_i}],$$

where D_{l_i} is given by Equation (11), if $K_{jj} \geq K_{ii}$, then the block will be the same as in Equation (10). Let R_{2_i} be a diagonal matrix formed by t blocks of size $n \times n$. Let

$$D_{ni} = \mathrm{diag}[z^0, \ldots, z^{n-K_{ii}-1}]. \tag{14}$$

For block j, if $K_{jj} \leq K_{ii}$, then the block will be

$$\mathrm{diag}[D_{ni}, z^{n-K_{ii}} I_{2K_{ii}-n}, z^{n-K_{ii}} D_{ni}],$$

if $K_{jj} \geq K_{ii}$, then the block will be

$$\mathrm{diag}[D_{ni}, z^{n-K_{ii}} I_{K_{jj}+K_{ii}-n}, z^{n-K_{ii}} D_{nj}],$$

where D_{nj} and D_{ni} follows the definition in Equation (14). Let $\lambda = \mu z$ and $z = \epsilon^{\frac{1}{n-K_{ii}}}$. It can be checked that $L_{2_i}(\lambda I - J - \epsilon B) R_{2_i}$ at $z = 0$ is a $t \times t$ block matrix $N(0)_i$ while the jth diagonal block looks like

$$\left[\begin{array}{ccccccccccc}
\mu & -1 & & & & & & & & & \\
 & \cdot & \cdot & & & & & & & & \\
 & & \cdot & \cdot & & & & & & & \\
* & & \mu & -1 & & & & & & & \\
* & & & 0 & -1 & & & & & & \\
\cdot & & & & \cdot & \cdot & & & & & \\
\cdot & & & & & \cdot & \cdot & & & & \\
* & & & & & & 0 & -1 & & & \\
 & * & & & & & & \mu & -1 & & \\
 & & \cdot & & & & & & \cdot & \cdot & \\
 & & & \cdot & & & & & & \cdot & \cdot \\
 & & & & * & & & & & & \mu
\end{array}\right].$$

For $K_{jj} \leq K_{ii}$, the $*$ in the first column goes from the $(n - K_{ii})$'th row to the K_{ii}th row, while for $K_{jj} \geq K_{ii}$, the $*$ in first column goes from the $(n - K_{ii})$th row to the K_{jj}th row. For the off diagonal blocks, if $l < m$, then $N(0)_{lm}$ has exactly the same form as $N(0)_{ll}$ with μ

and -1 replaced by 0. If $l > m$, then it has the form $M(0)_{mm}$ with only the $*$'s on the first column remaining. Taking L'_{2_i} and R'_{2_i} as L_{2_i} and R_{2_i} with z replaced by $\omega = e^{\frac{2\pi i}{n-K_{ii}}}$, we find that $\det(N_i(0))$ is $f(\mu^{n-K_{ii}})$ and the constant term and $\mu^{(n-K_{ii})t_i}$ term appear generically by inspecting the diagonal blocks only. So $J + \epsilon B$ has at least $(n - K_{ii})t_i$ eigenvalues forming t_i $(n - K_{ii})$-rings. Comparing the total number of eigenvalues of $J + \epsilon B$, we reach the conclusion.
$\square$

Proof of Case 4:

This can be proved by treating the $K_{i_1j_1}, K_{i_2j_2}, ...K_{i_tj_t}$ as $K_{11}, K_{22}, ...K_{tt}$'s as in Case 3 and going through the same proof, applying the same permutation as in Case 2. $\square$

References

[1] V. I. Arnold. On Matrices Depending on Parameters. *Russian Math. Surveys*, pages 29–43, 1971.

[2] H. Baumgärtel. *Analytic Perturbation Theory for Matrices and Operators.* Birkhäuser, Basel, 1985.

[3] James V. Burke and Michael L. Overton. Stable Perturbations of Nonsymmetric Matrices. *Linear Algebra Appl.*, pages 249–273, 1992.

[4] F. Chaitin-Chatelin and V. Frayssé. *Lectures on Finite Precision Computations.* SIAM, Philadelphia, 1996.

[5] A. Edelman, E. Elmroth, and B. Kågström. A Geometric Approach to Perturbation Theory of Matrices and Pencils: Part 2: Stratifications. *Submitted to SIAM J. Mat. Anal Appl.*

[6] A. Edelman, E. Elmroth, and B. Kågström. A Versal Deformation Approach To Perturbation Theory of Matrices and Pencils: Part 1: Versal Deformations. *SIAM J. Mat. Anal Appl., to appear.*

[7] T. Kato. *Perturbation Theory for Linear Operators.* Springer, Berlin, 1980.

[8] V. B. Lidskii. Pertuabation Theory of Non-conjugate Operators. *U.S.S.R. Comput. Maths. Math. Phys.*, pages 73–85, 1965.

[9] Julio Moro, James V. Burke, and L. Overton. On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of Matrices with Arbitrary Jordan Structure. *SIAM J. Matrix Anal. Appl., to appear.*

[10] Lothar Reichel and Lloyd N. Trefethen. Eigenvalues and Pseudo-Eigenvalues of Toeplitz Matrices. *Linear Algebra Appl.*, pages 153–185, 1992.

[11] M. M. Vainberg and L. A. Trenogin. *Theory of Branching of Solutions of Non-linear Equations.* Noordhoff, Leyden, 1974.