NORTH-HOLLAND

Nongeneric Eigenvalue Perturbations of Jordan Blocks

Yanyuan Ma*
Department of Mathematics
Room 2-333
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307

and

Alan Edelman†
Department of Mathematics
Room 2-380
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139-4307

Submitted by Richard A. Brualdi

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 generalize the result in a number of directions involving multiple Jordan blocks of the same size. © 1998 Elsevier Science Inc.

* E-mail: yanyuan@math.mit.edu: http://www-math.mit.edu/~yanyuan. Supported by NSF grants 9501278-DMS.

† E-mail: edelman@math.mit.edu: www: http://www-math.mit.edu/~edelman. Supported by NSF grants 9501278-DMS and 9404326-CCR.

LINEAR ALGEBRA AND ITS APPLICATIONS 273:45–63 (1998)

0024-3795/98/$19.00
PII S0024-3795(97)00342-X

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 nontypical behavior. We stifle the matrix's ability to form large eigenvalue rings by only allowing perturbations that are upper k-Hessenberg, meaning a matrix containing exactly k subdiagonals below and including the diagonal. The obvious question to ask is what is the typical behavior under this assumption. The result we will show is that the eigenvalue perturbations will then 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 nongeneric behavior occurred some small percentage of the time. Chatelin and Frayssé themselves point out in one example [4, p. 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-pseudospectra of a matrix can depend very much on the sparsity structure of the allowed perturbations. Following an example from Trefethen [9], 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, p. 109; 7, p. 65] that if a matrix A is perturbed by any matrix ϵB, then any multiple eigenvalue splits into rings, and their expansion in ϵ is a Puiseux series, since it is a branch of the solution of a polynomial with analytic coefficients. Unfortunately, the classical references give little information as to how the eigenvalues split as a function of the sparsity structure of the perturbation matrix. Without loss of generality, we will focus on one multiple eigenvalue. Associated with any perturbation B, we may define a partition $\pi(B)$ which contains the sizes (numbers of eigenvalues) of the rings.[1]

[1] The size of a ring is called its "period" in [7, 2]. The eigenvalue functions of ϵ in the same ring constitute a "cycle" in the terminology of these references.

We can quickly summarize most of what is known about the Puiseux series. If A is a single Jordan block of size n, then $\pi(B)$ is almost always $\{n\}$. This happens if and only if the lower left element is nonzero. For more complicated Jordan structure (say A is a nilpotent matrix), $\pi(B)$ is almost always the Segré characteristics of A, i.e., the sizes of the Jordan block structure of A. Lidskii explicitly determined the coefficients of the first order term, and Newton diagram approaches may also be used (see [8] for a discussion).

We used the words "almost always" in the above paragraph. There is an algebraic variety on which different behavior occurs. An ideal mathematical treatment would conveniently categorize all possible behaviors as a function of the perturbation B. This is a very difficult open problem. The only result of which we are aware is given by Burke and Overton [3] and Moro, Burke, and Overton [8]. The former studied when perturbations only yield periods of size 1 and 2 as part of a study of when the perturbations fall to one side, and the latter studied the first order perturbations under generic conditions and addressed some nongeneric situations.

Our approach is to try to identify classes of nongeneric situations where we can explain the typical behavior. We set up hypotheses on the structure of the perturbation, thereby creating nongeneric perturbations. We then ask what is the generic behavior of the eigenvalues given these hypotheses. (To be more precise, unless the perturbation satisfies certain algebraic conditions, the behavior occurs.)

In Section 2 of this paper we explore the case when A is a single Jordan block and B is upper k-Hesseberg. For example, suppose that we perturb an $n \times n$ Jordan block J with a matrix ϵB, where B has the form shown in Figure 1. Here k denotes the number of subdiagonals (including the main diagonal itself) that are not set to zero. If B were dense ($B_{71} \neq 0$), the eigenvalues of $J + \epsilon B$ would split uniformly onto a ring of size $n = 7$ and radius $O(\epsilon^{1/7})$. However, if $k = 4$, we obtain one ring of size 4 with radius $O(\epsilon^{1/4})$ and one ring of size 3 with radius $O(\epsilon^{1/3})$ as illustrated in Figure 2. Figure 3 contains a table of possible ring sizes when $n = 7$ for $k = 1, \ldots, 7$.

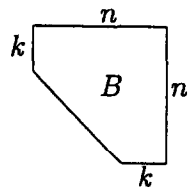

FIG. 1. B is an upper k-Hessenberg matrix.

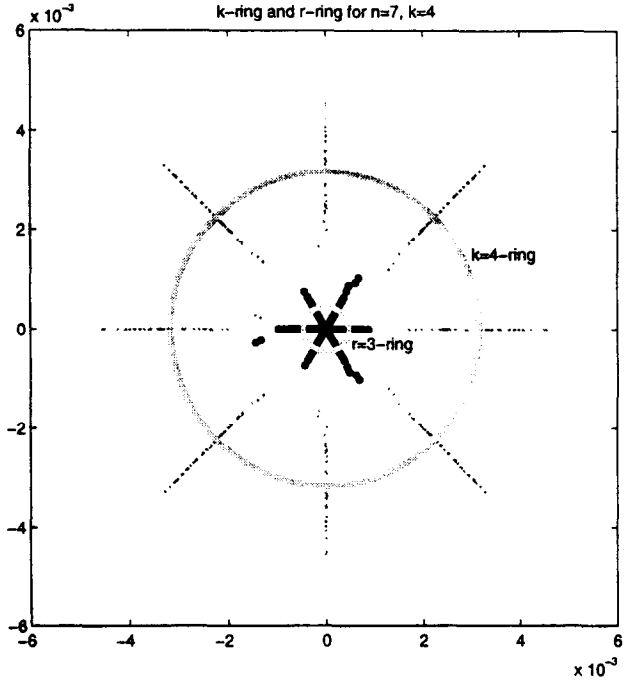

FIG. 2. Example rings for $n = 7$ and $k = 4$. We collected eigenvalues of 50 different random $J + \epsilon B$, $\epsilon = 10^{-12}$. The figure represents 50 different copies of one 4-ring (small dots) and 50 copies of one 3-ring (large dots). The two circles have radii 10^{-3} and 10^{-4}. If B were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

Our main result is that if a Jordan block of size n is perturbed by an upper k-Hessenberg matrix, then the eigenvalues typically split into $\lceil n/k \rceil$ rings, where $p \equiv \lfloor n/k \rfloor$ of them are k-rings with radius $O(\epsilon^{1/k})$, and if k does not divide n, there is typically one remaining r-ring with radius $O(\epsilon^{1/r})$, where $r \equiv n \bmod k$. 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 equal sized Jordan blocks. We only concentrate on the case where all the t blocks have the same eigenvalue λ, since it is well known (see [8]) that the behavior of the perturbation on different eigenvalues splits.

k \ ring size	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

FIG. 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.

Let $J = \operatorname{diag}[J_1, J_2, \ldots, J_t]$, where the J_i's are $n \times n$ Jordan blocks, and we conformally partition

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

Suppose every B_{ij} is an upper k-Hessenberg matrix. We will show in Theorem 2 that generically, the eigenvalues break into $t\lceil n/k \rceil$ rings, tp of them are k-rings, and the remaining t are r-rings if k does not divide n. Here, p and r have 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

$$J = J_7(\lambda) \oplus J_7(\lambda),$$

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 $J + \epsilon B$ will split into four 3-rings centered at λ with radii $O(\epsilon^{1/3})$ and two 1-rings centered at λ 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

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

2. ONE BLOCK CASE

Suppose that the Jordan form of J is simply one Jordan block. We assume that $J = J_n(0)$, which we will perturb with ϵB, where B has the sparsity structure given in Figure 1.

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 on *upper k-Hessenberg matrix*.

DEFINITION 2. Suppose, for ϵ sufficiently small,

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

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

LEMMA 1 [7, p. 65]. *Let λ be a multiple eigenvalue of J with multiplicity s. Then there will be s eigenvalues of $J + \epsilon B$ 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*

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

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

Our Theorem 1 shows how the eigenvalues split into rings, and in Corollary 1 and 2 we analyze the ring constant c. The main idea of the proofs is that only certain terms in the characteristic polynomial of $J + \epsilon B$ influence the ring constants. For the k-rings, we are interested in the terms from $\det(\lambda I - J - \epsilon B)$ of the form $\sum \alpha_i \epsilon^i \lambda^{n-ki}$ and no higher order terms in ϵ. For the r-ring, we are interested in the $O(\epsilon^{p+1})$ term in $\det(A + \epsilon B)$ and the $O(\epsilon^p)$ term multiplying λ^r in the characteristic polynomial of $A + \lambda B$. All of these may be viewed as determinants with entries removed or as bipartite matchings.

The *bipartite graph* associated with an $n \times n$ sparse matrix A is a graph on n *left* vertices and n *right* vertices such that nonzero elements a_{ij} are associated with an edge between left node i and right node j. We find it

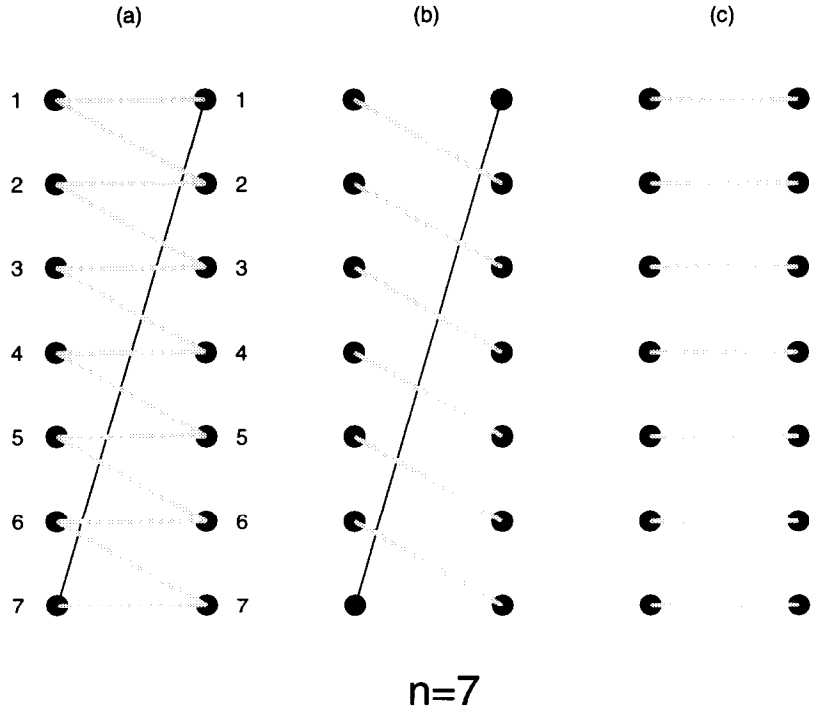

FIG. 5. (a) Bipartite graph of $\lambda I - (J_n + \epsilon B)$ [where $J = J_n(0)$]. (b) Perfect matching defining $\epsilon B_{n,1}$ term. (c) Perfect matching defining λ^n term.

convenient to associate terms in the determinental expansion of $\det(\lambda I - A)$ with subgraphs of the bipartite graph that are perfect matchings.

A simple example where $J = J_n(0)$ and B is nonzero only in the $(n, 1)$ entry is plotted in Figure 5. We call the set in the second column a *laced section*.

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

REMARK 1. Here, generic means that α_p from Equation (3) and γ from Equation (4) are both nonzero. If only $\alpha_p \neq 0$, we have p k-rings, but the

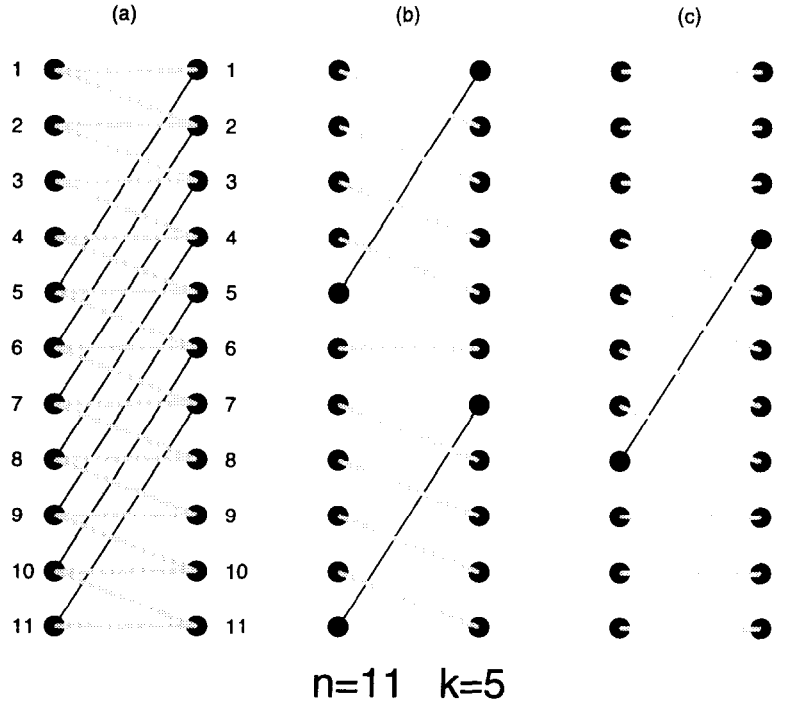

FIG. 6. (a) Bipartite graph of $\lambda I - J - \epsilon B$. (b) Perfect matching with two laced sections. (c) Perfect matching with one laced section.

r-ring is not guaranteed. Some pathological examples that violate the two generic conditions are given in Section 3.

Proof. If B only has elements on the kth subdiagonal, it is easy to study the p k-rings. The situation is illustrated in Figure 6. Every term in the characteristic polynomial must correspond to a union of laced sections of size k and horizontal lines. We therefore have that

$$\det[\lambda I - (J + \epsilon B)] = \lambda^n + \alpha_1 \epsilon \lambda^{n-k} + \alpha_2 \epsilon^2 \lambda^{n-2k} + \cdots + \alpha_p \epsilon^p \lambda^{n-pk},$$

where

$$\alpha_i = (-1)^i \sum_{l_{j+1} - l_j \geq k} B_{l_1+k-1, l_1} B_{l_2+k-1, l_2} \cdots B_{l_i+k-1, l_i} \tag{1}$$

for $i = 1, \ldots, p$, and $B_{k,1}, \ldots, B_{n,n-k+1}$ denote the elements on the kth diagonal of B. Therefore $J + \epsilon B$ has r eigenvalues equal to 0 up to $O(\epsilon^{1/k})$ and p k-rings with radii the kth powers of the zeros of

$$q(z) = z^p + \alpha_1 z^{p-1} + \cdots + \alpha_p.$$

Now if B is upper k-Hessenberg and we wish to study the $O(\epsilon^{1/k})$ eigenvalues, only the lowest subdiagonal elements matter to first order. To see this, note that $\det[\lambda I - (J + \epsilon B)]$ has no laced sections of size $> k$, and that any of size $< k$ has too many λ's for dominant balance. Therefore, only the laced sections of size k remain.

Alternatively this result may be obtained following the Lidskii approach of letting $\lambda = \mu\epsilon^{1/k}$, $z = \epsilon^{1/k}$, $L_1 = \operatorname{diag}[z^{-1}, z^{-2}, \ldots, z^{-n}]$, $R_1 = \operatorname{diag}[1, z, \ldots, z^{n-1}]$ and studying the limit of $L_1(\mu z I - J - z^k B)R_1$ as $z \to 0$.

We now turn to the r-ring. Readers familiar with the Newton diagram [2, 8] can easily see that one r-ring remains, because the Newton diagram consists of one line segment from $(0, 0)$ to (pk, p) and a second from (pk, p) to $(n, p + 1)$ when $r \neq 0$ (see Corollary 3 in [8]). Using the bipartite graph approach, it is easy to see that the typical term consists of $p + 1$ laced sections each of size at most k. This may also be obtained from a Lidskii style argument as the determinant in Figure 7. ■

COROLLARY 1. *The kth powers of the ring constants for the k-rings are the roots of* $q(z)$, *where*

$$q(z) = z^p + \alpha_1 z^{p-1} + \cdots + \alpha_i z^{p-i} + \cdots + \alpha_p \tag{2}$$

and

$$\alpha_i = (-1)^i \sum_{l_{j+1} - l_j \geq k} B_{l_1+k-1, l_1} B_{l_2+k-1, l_2} \cdots B_{l_i+k-1, l_i} \tag{3}$$

for $i = 1, \ldots, p$, *where* $B_{k,1}, \ldots, B_{n,n-k+1}$ *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.

COROLLARY 2. *The rth power of the ring constant for the r-ring is the root of*

$$\alpha_p z + (-1)^{p+1}\gamma = 0.$$

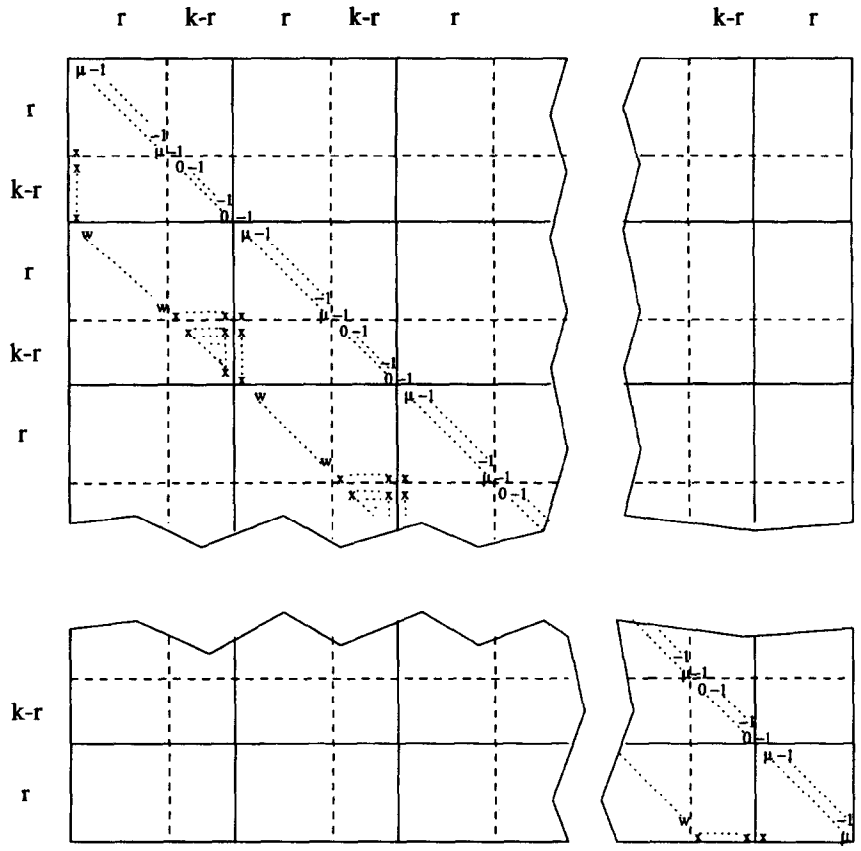

FIG. 7. A picture of the effective matrix in Lidskii approach. The blocks with μ's on the diagonals have sizes $r \times r$. The blocks with 0's on the diagonals have sizes $(k - r) \times (k - r)$. The x's and w's represent the original entries of the matrix B at the same position. Dotted lines represent a repetition of the format.

Here, α_p *is defined in Equation* (3), *and*

$$\gamma = \sum B_{i_1, i_0+1} B_{i_2, i_1+1} \cdots B_{i_{p+1}, i_p+1}. \tag{4}$$

B_{i_1, i_0+1}, $B_{i_2, i_1+1}, \ldots, B_{i_{p+1}, i_p+1}$ *are the entries of* B *in the* x *position of Figure* 7, $i_0 = 0$, $i_{p+1} = n$, *and they satisfy*

$$r \leqslant i_{m+1} - i_m \leqslant k \tag{5}$$

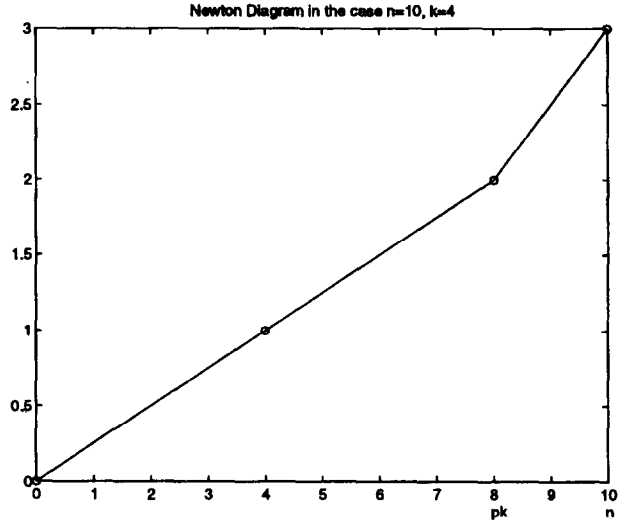

FIG. 8. Newton diagram.

for

$$m = 0, 1, 2, \ldots, p$$

3. PATHOLOGICAL EXCEPTIONS

When $k > n/2$, $p = 1$, the $q(z)$ in Corollary 1 is $q(z) = z + \alpha_1$, where $\alpha_1 = -\Sigma_{l_1} B_{l_1+k-1,l_1}$. Therefore $\alpha_1 \neq 0$ is the generic case. 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 to zero is the kth subdiagonal, for $k > n/2$, using a Newton diagram [8, 2] (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.

We can argue similarly for $k < n/2$. When $\alpha_p = 0$, we generically lose one k-ring and the r-ring. Consider the Newton diagram: the (pk, p) point moves up and the whole diagram generically breaks into three segments, one with slope $1/k$ of length $(p - 1)k$, one with slope $1/(k - 1)$ of length $k - 1$, and one with slope $1/(r + 1)$ of length $r + 1$. This means it has

$(p-1)k$ eigenvalues forming $p-1$ k-rings, $k-1$ 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 $k=r+1$; then the whole diagram breaks into only two segments, the first one remains untouched, and the second one has slope $2/(k+r)$ and length $k+r$. When γ in Equation (4) is zero, the r-ring will be lost.

The following are three examples that violate the two generic conditions.

EXAMPLE 1. $n=9$, $k=4$, $p=2$, $r=1$ (see Figure 9):

$$B=\begin{bmatrix} -11 & 1 & -1 & -14 & 22 & 2 & -6 & -13 & -9 \\ -8 & -2 & 2 & -4 & 3 & 10 & -15 & 7 & -10 \\ 4 & -3 & -1 & -5 & 9 & 12 & -1 & -14 & -1 \\ 1 & -7 & 17 & 18 & 7 & -5 & 6 & -13 & -24 \\ 0 & -1 & 16 & 8 & 6 & 9 & 1 & -6 & -7 \\ 0 & 0 & 6 & 1 & 10 & -2 & 16 & -15 & -14 \\ 0 & 0 & 0 & -3 & 13 & -3 & -3 & 6 & 3 \\ 0 & 0 & 0 & 0 & 0 & 5 & 8 & -3 & 6 \\ 0 & 0 & 0 & 0 & 0 & 9 & -8 & -13 & 1 \end{bmatrix}.$$

EXAMPLE 2. $n=5$, $k=2$, $p=2$, $r=1$ (see Figure 10):

$$B=\begin{bmatrix} 16 & 2 & 4 & -6 & -10 \\ 4 & -7 & 2 & -7 & 8 \\ 0 & 8 & -17 & -2 & -1 \\ 0 & 0 & -3 & -6 & -4 \\ 0 & 0 & 0 & 1 & -10 \end{bmatrix}.$$

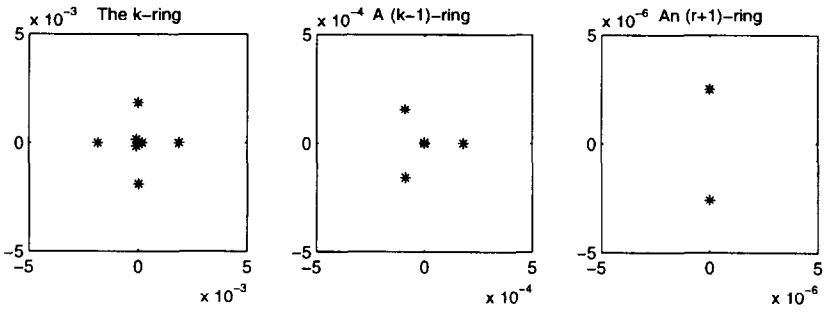

FIG. 9. Example 1: $\alpha_p=0$; we lose one k-ring and the r-ring. The last k-ring becomes an $(k-1)$-ring, and the remaining $r+1$ eigenvalues form an $(r+1)$-ring.

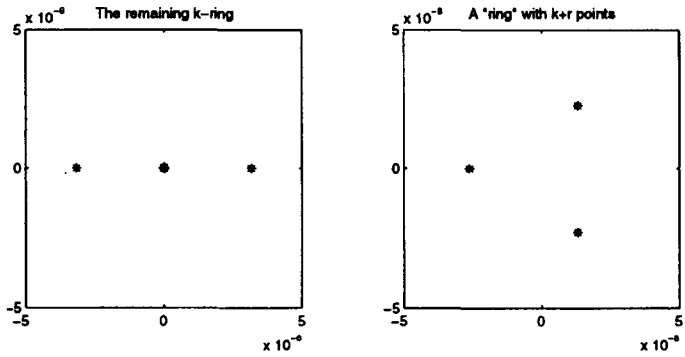

FIG. 10. Example 2: $\alpha_p = 0$; we lose one k-ring and the r-ring. The remaining $k + r$ eigenvalues still spread evenly on a circle, but they do not form a ring.

EXAMPLE 3. $n = 5$, $k = 3$, $p = 1$, $r = 2$ (see Figure 11):

$$B = \begin{bmatrix} 1 & 10 & 3 & -5 & 5 \\ -72 & 19 & -6 & -1 & 9 \\ -12 & 17 & -9 & -4 & 26 \\ 0 & 7 & 5 & 6 & -10 \\ 0 & 0 & 1 & -6 & -10 \end{bmatrix}.$$

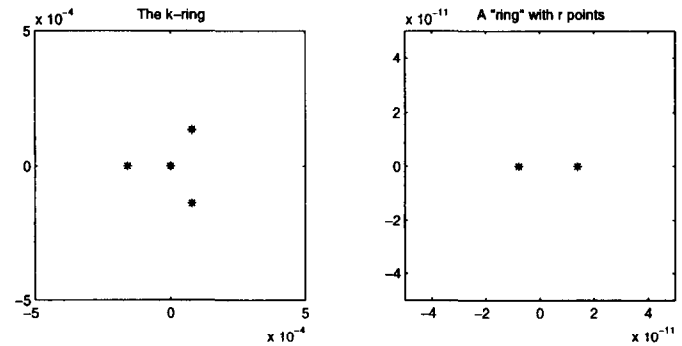

FIG. 11. Example 3: $\gamma = 0$; the r eigenvalues still spread evenly on a circle, but they do not form a ring.

4. t BLOCK CASE (ALL B_{ij}'S ARE UPPER k-HESSENBERG MATRICES)

We now study the case when the Jordan form of J has t blocks all with the same size n. We found the case when J has a Jordan structure of different size blocks too complicated for general analysis, though individual cases are easily examined. In this section, we only consider the admittedly special case where the perturbation matrix B has the block upper k-Hessenberg form obtained by dividing B into $n \times n$ blocks and every B_{ij} is an upper k-Hessenberg matrix. In this special case, we have

THEOREM 2. *Let J, B, n, and k be given as above, and let r be the remainder of n divided by k, i.e., $n = pk + r$, $0 \leqslant r < k$. The eigenvalues of $J + \epsilon B$ will then split into tp k-rings and t r-rings if $r \neq 0$.*

Proof. The proof follows closely that of Theorem 1, but we now imagine that B has only elements on the kth subdiagonal of each block. Every term in the characteristic polynomial must correspond to a union of possibly deformed laced sections of size k (see Figure 12) and horizontal lines.

We therefore have that

$$\det[\lambda I - (J + \epsilon B)] = \lambda^{nt} + \alpha_1 \epsilon \lambda^{nt-k} + \alpha_2 \epsilon^2 \lambda^{nt-2k} + \cdots + \alpha_{pt} \epsilon^{pt} \lambda^{nt-pkt},$$

where instead of Equation (3) in Section 2, we have

$$\alpha_i = (-1)^i \sum \det B_{l_1, l_2, \ldots, l_i}.$$

Here, $B_{l_1, l_2, \ldots, l_i}$ represents the matrix formed by extracting the entries at rows $l_1 + k - 1$, $l_2 + k - 1, \ldots, l_i + k - 1$ and columns $l_1, l_2, \ldots, l_i$, with $l_{j+1} - l_j \geqslant k$. Therefore $J + \epsilon B$ has rt eigenvalues equal to 0 up to $O(\epsilon^{1/k})$, and pt k-rings with radii the kth powers of the zeros of

$$q(z) = z^{pt} + \alpha_1 z^{pt-1} + \cdots + \alpha_{pt}.$$

Now if B has the block upper k-Hessenberg form and we wish to study the $O(\epsilon^{1/k})$ eigenvalues, only the lowest subdiagonal elements matter to first order, for the same reason as in the proof of Theorem 1.

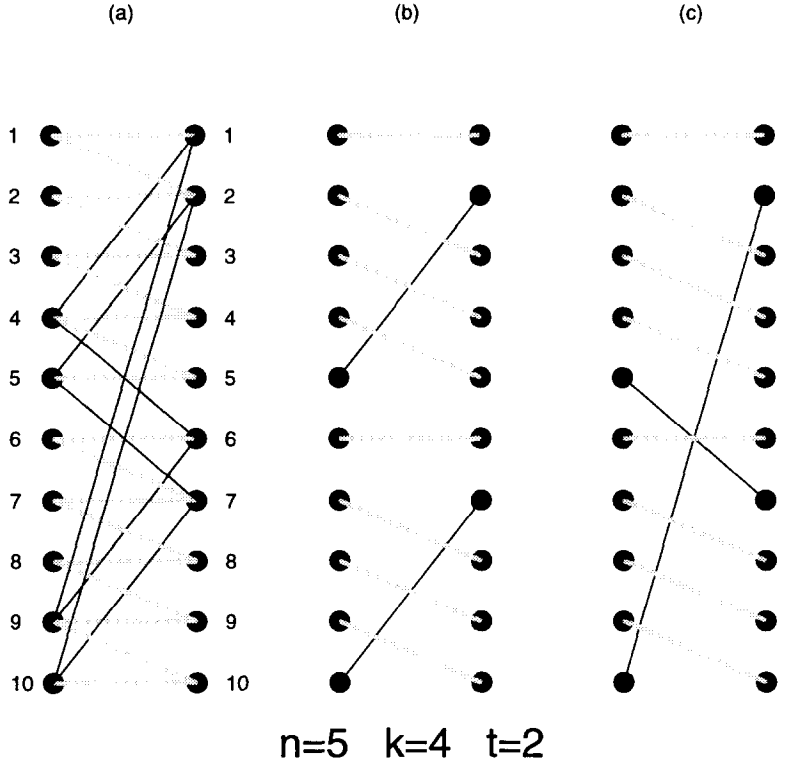

FIG. 12. (a) Bipartite graph of $\lambda I - (J + \epsilon B)$. (b) Perfect matching with two laced sections of size 4 contributing to $\epsilon^3\lambda^2$ term. (c) Perfect matching with two deformed laced sections of size 4 contributing to $\epsilon^2\lambda^2$ term.

Alternatively, this result may be obtained following the Lidskii approach of letting $\lambda = \mu\epsilon^{1/k}$, $z = \epsilon^{1/k}$,

$$L_1 = \operatorname{diag}\left[\operatorname{diag}[z^{-1}, z^{-2}, \ldots, z^{-n}], \ldots, \operatorname{diag}[z^{-1}, z^{-2}, \ldots, z^{-n}]\right],$$

$$R_1 = \operatorname{diag}\left[\operatorname{diag}[1, z, \ldots, z^{n-1}], \ldots, \operatorname{diag}[1, z, \ldots, z^{n-1}]\right]$$

and studying the limit of $L_1(\mu zI - J - z^k B)R_1$ as $z \to 0$.

We now turn to the r-rings. Although the Newton diagram approach can not be applied in an obvious way here, with the bipartite graph approach or the Lidskii approach it can be seen that the typical terms consist of $tp + i$ possibly deformed laced sections each of size between r and k, with $i = 0, 1, \ldots, t$. ■

COROLLARY 3. *The kth powers of the ring constants for the k-rings are the roots of*

$$q(z) = z^{pt} + \alpha_1 z^{pt-1} + \alpha_2 z^{pt-2} + \cdots + \alpha_i z^{pt-i} + \cdots + \alpha_{pt},$$

with

$$\alpha_i = (-1)^i \det B_{l_1, l_2, \ldots, l_i},$$

where $B_{l_1, l_2, \ldots, l_i}$ *represents the matrix formed by extracting the entries at rows* $l_1 + k - 1$, $l_2 + k - 1, \ldots, l_i + k - 1$ *and columns* $l_1, l_2, \ldots, l_i$, *with* $l_{j+1} - l_j \geqslant k$. *So long as* $\alpha_{pt} \neq 0$, *we obtain the generic behavior described in Theorem* 2.

COROLLARY 4. *The rth powers of the ring constants for the r-rings are the roots of*

$$g(z) = \gamma_0 z^t + \gamma_1 z^{t-1} + \gamma_2 z^{t-2} + \cdots + \gamma_i z^{t-i} + \cdots + \gamma_t,$$

where

$$\gamma_i = (-1)^{pt+i} \sum \det B_{l_0, l_1, l_2, \ldots, l_{pt+i}}$$

with

$$r \leqslant i_{m+1} - i_m \leqslant k$$

for $i = 0, 1, \ldots, t$. *Here,* $B_{l_0, l_1, l_2, \ldots, l_{pt+i}}$ *represents a matrix obtained by extracting the entries on rows* $l_1, l_2, \ldots, l_{pt+i}$ *and columns* $l_0 + 1$, $l_1 + 1, \ldots, l_{pt+i-1} + 1$ *from the matrix formed by repeating Figure* 7 *t times on the diagonal and with* μ*'s and* -1*'s replaced by* 0*'s on the off diagonals, and we have* $l_0 = 0$, $l_{pt+t} = nt$.

5. 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, and the general problem remains open. We have some observations in two special cases. Let K_{ij} be the number of subdiagonals of B_{ij} for $1 \leqslant i, j \leqslant n$, i.e., B_{ij} is an upper K_{ij}-Hessenberg matrix.

THEOREM 3

Case 1: *Let* $K_{\max} = \max(K_{ij})$. *If we can find* t K_{ij}*'s equal to* $K_{\max}$ *such that no two of them are in the same row or column, then the result from Theorem* 2 *holds upon taking* $k = K_{\max}$.

Case 2: *If we can find* t *numbers* $K_{i_1j_1}, K_{i_2j_2}, \ldots, K_{i_tj_t}$, *all* $\geq n/2$, *such that* $K_{i_sj_s} \geq K_{i_sl}$ *and* $K_{i_sj_s} \geq K_{mj_s}$ *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* $J + \epsilon B$ *has* $K_{i_sj_s}$ *eigenvalues that form one* $K_{i_sj_s}$*-ring for* $s = 1, \ldots, t$. *The remaining* $n - K_{i_sj_s}$ *eigenvalues form an* $(n - K_{i_sj_s})$*-ring for* $s = 1, \ldots, t$.

Proof of Case 1. Simply replace all the K_{ij}'s with $K_{\max}$ and notice that the proof of Theorem 2 is still valid with k replaced by $K_{\max}$, with the same genericity condition. ∎

Proof of Case 2. Suppose all $K_{i_1j_1}, K_{i_2j_2}, \ldots, K_{i_tj_t}$ are obtained from the matrices on the main block diagonal and the matrices on the off diagonals are all zero matrices; then the results hold obviously. Now assume the off diagonal matrices are the ones satisfying the conditions in case 2. The way to change the Newton diagram is through getting a new nonzero term from the off diagonal elements, which is not possible in this case. The result still holds without the assumption that $K_{i_1j_1}, K_{i_2j_2}, \ldots, K_{i_tj_t}$ are obtained from the main block diagonal matrices: notice that we still get the same nonzero terms generically from those off diagonal matrices where the $K_{i_1j_1}, K_{i_2j_2}, \ldots, K_{i_tj_t}$ instead come from diagonal matrices. ∎

REFERENCES

1 V. I. Arnold, On matrices depending on parameters, *Russian Math. Surveys*, 26:29–43 (1971).

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

3 J. V. Burke and M. L. Overton. Stable perturbations of nonsymmetric matrices, *Linear Algebra Appl.* 171:249–273 (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, *SIAM J. Math. Anal Appl.*, to appear.

6 A. Edelman, E. Elmroth, and B. Kågström. A geometric approach to perturbation theory of matrices and matrix pencils: Part 1: Versal deformations, 18:653–692 (1997).
7 T. Kato, *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1980.
8 J. Moro, J. V. Burke, and M. L. Overton, On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure, *SIAM J. Matrix Anal. Appl.*, 18:793–817 (1997).
9 L. Reichel and L. N. Trefethen, Eigenvalues and pseudoeigenvalues of Toeplitz matrices. *Linear Algebra Appl.* 162–164:153–185 (1992).

Received 6 January 1997; final manuscript accepted 28 July 1997