# On the Lidskii-Vishik-Lyusternik Perturbation Theory for Eigenvalues of Matrices with Arbitrary Jordan Structure

J. Moro*    J. V. Burke†    M. L. Overton‡

Submitted to *SIAM J. Matrix Anal. Appl.*

*Dedicated to V. B. Lidskii and M. I. Vishik on the respective occasions of their 70th and 75th birthdays.*

**Abstract**

Let $A$ be a complex matrix with arbitrary Jordan structure, and $\lambda$ an eigenvalue of $A$ whose largest Jordan block has size $n$. We review previous results due to Lidskii [10], showing that the splitting of $\lambda$ under a small perturbation of $A$ of order $\varepsilon$ is, generically, of order $\varepsilon^{1/n}$. Explicit formulas for the leading coefficients are obtained, involving the perturbation matrix and the eigenvectors of $A$. We also present an alternative proof of Lidskii's main theorem, based on the use of the Newton diagram. This approach clarifies certain difficulties which arise in the nongeneric case, and leads, in some situations, to the extension of Lidskii's results. These results suggest a new notion of Hölder condition number for multiple eigenvalues, depending only on the conditioning of the associated eigenvectors, not the conditioning of the Jordan vectors.

*Departamento de Matemática Aplicada, Universidad Complutense, Madrid & Courant Institute of Mathematical Sciences, New York University, New York, NY. Supported in part by an Ayuda Complutense Postdoctoral en el Extranjero and by CICYT grant TAP94-115.

†Department of Mathematics, University of Washington, Seattle, WA. Supported in part by NSF grant DMS–9303772.

‡Computer Science Department, Courant Institute of Mathematical Sciences, New York University, New York, NY. Supported in part by NSF grant CCR–9401136

# 1 Introduction

Given a square complex matrix $A$, it is an important question from both the theoretical and the practical point of view to know how the eigenvalues and eigenvectors change when the elements of $A$ are subjected to small perturbations. The usual formulation of the problem introduces a perturbation parameter $\varepsilon$, belonging to some neighborhood of zero, and writes the perturbed matrix as $A + \varepsilon B$ for an arbitrary matrix $B$. In this situation, it is well known ( [1] §9.3.1, [7] §II.1.2) that each eigenvalue or eigenvector of $A+\varepsilon B$ admits an expansion in fractional powers of $\varepsilon$, whose zero-th order term is an eigenvalue or eigenvector of the unperturbed matrix $A$.

In this paper we address the question of determining the first-order term of this expansion, or more precisely, the first term following the terms of order zero. No restriction is imposed on the Jordan structure of $A$ although we assume that this Jordan structure is known from the outset. In Section 2 we present two results, due to Lidskii [10], which provide, under certain nondegeneracy conditions, the leading exponents and leading coefficients of both eigenvalue and eigenvector perturbations. The central idea of the proof is simply to transform the characteristic equation $\det(\omega I - A - \varepsilon B) = 0$ into an equivalent one $Q(\mu, z) = 0$ through a change of variables

$$\begin{aligned} z &= \varepsilon^{1/n} \\ \mu &= \frac{\omega - \lambda}{z} \end{aligned} \tag{1}$$

for a suitable $n$, where $\lambda$ is an eigenvalue of $A$. An appropriate factorization of $Q(\mu, 0)$ leads to the final result.

A specific example may be helpful to give a better idea of these results: consider a $9 \times 9$ Jordan matrix $J$ with a unique zero eigenvalue and four Jordan blocks with respective dimensions $3, 3, 2$ and $1$. Lidskii's results show that, generically, the leading terms in the expansions of the eigenvalues of any perturbation $J+\varepsilon B$ depend *exclusively* on those elements of $B$ marked with a box in the matrix below:

$$B = \left[\begin{array}{ccc|ccc|cc|c}
* & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * \\
\Box & * & * & \Box & * & * & \Box & * & \Box \\ \hline
* & * & * & * & * & * & * & * & * \\
* & * & * & * & * & * & * & * & * \\
\Box & * & * & \Box & * & * & \Box & * & \Box \\ \hline
* & * & * & * & * & * & * & * & * \\
\Box & * & * & \Box & * & * & \Box & * & \Box \\ \hline
\Box & * & * & \Box & * & * & \Box & * & \Box
\end{array}\right].$$

More specifically, let $\Phi_1$ denote the $2 \times 2$ matrix given by the four boxes at the top left, i.e.

$$\Phi_1 = \begin{bmatrix} B_{31} & B_{34} \\ B_{61} & B_{64} \end{bmatrix}$$

and let $\xi_1, \xi_2$ be the eigenvalues of $\Phi_1$. Then, the perturbed matrix $J+\varepsilon B$ has six eigenvalues with leading terms

$$\xi_1^{1/3}\varepsilon^{1/3}, \quad \xi_2^{1/3}\varepsilon^{1/3},$$

using all three cube roots of $\xi_1$ and $\xi_2$. Furthermore, the other three eigenvalues of $J+\varepsilon B$ have, generically, leading terms of order $\varepsilon^{1/2}$ and $\varepsilon$, and their coefficients can be obtained by solving certain polynomial equations which only involve the elements of $B$ marked with boxes above.

In the most general case when $A$ is not in Jordan form, these elements of $B$ must be replaced by products $yBx$, where $x$ (resp. $y$) is a right (resp. left) eigenvector of $A$.

The first results in this direction were obtained by Vishik and Lyusternik [13], motivated by applications to differential operators. Lidskii [10] generalized their results in the finite dimensional case, obtaining simple explicit formulas for the perturbation coefficients and providing, at the same time, a much more elementary proof (which is, essentially, the one we present in Section 2). The results in both [13] and [10] were later refined by Baumgärtel ( [1] §7.4), in the sense of dealing not only with perturbation series for eigenvalues and eigenvectors, but also with the corresponding eigenprojections as functions of $\varepsilon$. Vainberg and Trenogin ( [12] §32), on the other hand, offer a fairly thorough account of similar results, obtained for Fredholm operators by applying the techniques of branching theory. Langer and Najman [9] recently generalized Lidskii's results to matrix pencils $M(\lambda) + N(\lambda, \varepsilon)$, using the local Smith normal form of parameter-dependent matrices (Lidskii's results follow from choosing $M \equiv A - \lambda I$, $N \equiv \varepsilon B$). The fundamental results of Lidskii remain, however, almost completely unknown in the Western literature. The only references to [10] appearing in the Science Citation Index are [3] and [9], and both of these continue earlier work [2] and [8] in which the authors were unaware of [10]. The main purpose of this paper is, therefore, to bring Lidskii's results to the attention of the broad linear algebra community. See [11] for an application to stability theory for Hamiltonian systems.

Section 2 is devoted to reviewing both the results and the proofs given in [10]. We should stress here that, although Lidskii stated his results as being valid for analytic perturbations, we will see that they hold in fact for a more general class of perturbations, namely those of class $C^1$ (see Remark 4 in Section 2). Lidskii's results depend on certain nondegeneracy assumptions, and no information about the leading exponents or coefficients is available in the degenerate case from the approach taken in Section 2. We present in Section 3 a different approach which, apart from providing an alternative proof of Lidskii's main theorem, identifies the difficulties which arise in degenerate cases, and, in some situations, leads to extensions of the results in Section 2. This alternative approach is much in the spirit of [12], since our main tool is the Newton diagram. We end by proposing in Section 4 a new notion of Hölder condition number for multiple eigenvalues, suggested by Lidskii's results. Although it is closely related to previous Hölder condition numbers in the literature ( [4] p.156), its main difference is that it only depends on the conditioning of eigenvectors, not on the conditioning of Jordan vectors.

# 2 Lidskii's perturbation theory

Let $A$ be a complex matrix with distinct eigenvalues $\lambda_1, \ldots, \lambda_\nu$, and a corresponding Jordan form

$$J = P^{-1}\, A\, P = \mathrm{Diag}\,[J_1, \ldots, J_\nu]. \tag{2}$$

We partition

$$P = \left[\begin{array}{c|c|c} P_1 & \ldots & P_\nu \end{array}\right]$$

and

$$P^{-1} = \left[\begin{array}{c} Q_1 \\ \hline \vdots \\ \hline Q_\nu \end{array}\right]$$

conformally with (2). Given $i \in \{1, \ldots, \nu\}$, we write $\lambda \equiv \lambda_i$ and

$$J_i = \mathrm{Diag}(\Gamma_1^1, \ldots, \Gamma_1^{r_1}, \ldots, \Gamma_q^1, \ldots, \Gamma_q^{r_q}), \tag{3}$$

suppressing the index $i$, where, for $j = 1, \ldots, q$,

$$\Gamma_j^1 = \ldots = \Gamma_j^{r_j} = \begin{bmatrix} \lambda & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & \lambda \end{bmatrix}$$

is a Jordan block of dimension $n_j$, repeated $r_j$ times, and ordered so that

$$n_1 > n_2 > \ldots > n_q.$$

The $n_j$ are called the *partial multiplicities* for $\lambda$. The eigenvalue $\lambda$ is semisimple (nondefective) if $q = n_1 = 1$ and nonderogatory if $q = r_1 = 1$. The algebraic and geometric multiplicities of $\lambda$ are, respectively,

$$m = \sum_{j=1}^{q} r_j n_j \qquad \text{and} \qquad g = \sum_{j=1}^{q} r_j.$$

We further partition

$$P_i = \left[ \begin{array}{c|c|c|c|c|c|c} P_1^1 & \ldots & P_1^{r_1} & \ldots & P_q^1 & \ldots & P_q^{r_q} \end{array} \right]$$

conformally with (3), again suppressing the index $i$. The columns of each $P_j^k$ form a right Jordan chain of $A$ with length $n_j$ corresponding to $\lambda$. If we denote by $x_j^k$ the first column of $P_j^k$, each $x_j^k$ is a right eigenvector of $A$ associated with $\lambda$. Analogously, we split

$$Q_i = \left[ \begin{array}{c} Q_1^1 \\ \hline \vdots \\ \hline Q_1^{r_1} \\ \hline \vdots \\ \hline Q_q^1 \\ \hline \vdots \\ \hline Q_q^{r_q} \end{array} \right]$$

also conformally with (3). The rows of each $Q_j^k$ form a left Jordan chain of $A$ of length $n_j$ corresponding to $\lambda$. Hence, if we denote by $y_j^k$ the last (i.e. $n_j$–th) row of $Q_j^k$, each $y_j^k$ is a left eigenvector corresponding to $\lambda$. With these eigenvectors we build up matrices

$$Y_j = \begin{bmatrix} y_j^1 \\ \vdots \\ y_j^{r_j} \end{bmatrix}, \qquad X_j = [x_j^1, \ldots, x_j^{r_j}],$$

and

$$W_j = \begin{bmatrix} Y_1 \\ \vdots \\ Y_j \end{bmatrix}, \qquad Z_j = [X_1, \ldots, X_j]$$

and define square matrices $\Phi_j$ and $E_j$ of dimension $f_j = \sum_{s=1}^{j} r_s$ by

$$\Phi_j = W_j B Z_j, \qquad j = 1, \dots, q,$$

$$E_1 = I, \qquad E_j = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \qquad \text{for } j = 2, \dots, q,$$

where the identity block in $E_j$ has dimension $r_j$. Note that, due to the cumulative definitions of $W_j$ and $Z_j$, every $\Phi_{j-1}$, $j = 2, \dots, q$, is the upper left block of $\Phi_j$.

**Theorem 1 (Lidskii [10])** *Let $j \in \{1, \dots, q\}$ be given, and assume that, if $j > 1$, $\Phi_{j-1}$ is nonsingular. Then there are $r_j n_j$ eigenvalues of the perturbed matrix $A + \varepsilon B$ admitting a first-order expansion*

$$\lambda_j^{kl}(\varepsilon) = \lambda + (\xi_j^k)^{1/n_j} \varepsilon^{1/n_j} + o(\varepsilon^{1/n_j}) \tag{4}$$

*for $k = 1, \dots, r_j$, $l = 1, \dots, n_j$, where*

*i) the $\xi_j^k$, $k = 1, \dots, r_j$ are the roots of equation*

$$\det(\Phi_j - \xi E_j) = 0 \tag{5}$$

*or, equivalently, the eigenvalues of the Schur complement of $\Phi_{j-1}$ in $\Phi_j$ (if $j = 1$, the $\xi_1^k$ are just the $r_1$ eigenvalues of $\Phi_1$);*

*ii) the different values $\lambda_j^{kl}$ for $l = 1, \dots, n_j$ are defined by taking the $n_j$ distinct $n_j$-th roots of $\xi_j^k$.*

*If, in addition, the $r_j$ solutions $\xi_j^k$ of (5) are all distinct, then the eigenvalues (4) can be expanded locally in power series of the form*

$$\lambda_j^{kl}(\varepsilon) = \lambda + (\xi_j^k)^{1/n_j} \varepsilon^{1/n_j} + \sum_{s=2}^{\infty} a_{js}^{kl} \varepsilon^{s/n_j}, \tag{6}$$

*$k = 1, \dots, r_j$, $l = 1, \dots, n_j$.*

**Remark:** Two special cases of Theorem 1 are well known. In the case that $\lambda$ is semisimple, i.e. $q = n_1 = 1$ with multiplicity $r_1$, equation (4) reduces to

$$\lambda_1^{k1}(\varepsilon) = \lambda + \xi_1^k \varepsilon + o(\varepsilon),$$

where the $\xi_1^k$ are the eigenvalues of the $r_1$ by $r_1$ matrix $Y_1 B X_1$ (cf. [7] §II.2.3). In the case that $\lambda$ is nonderogatory, i.e. $q = r_1 = 1$ with multiplicity $n_1$, equation (4) reduces to

$$\lambda_j^{1l}(\varepsilon) = \lambda + (\xi_1^1)^{1/n_1} \varepsilon^{1/n_1} + o(\varepsilon^{1/n_1}),$$

where $\xi_1^1 = y_1^1 B x_1^1$. These two cases coincide when $\lambda$ is simple.

**Theorem 2 (Lidskii [10])** *Let $\Phi_s$ be nonsingular for $s = 1, \ldots, q$, and let $j \in \{1, \ldots, q\}$ be such that the $r_j$ roots of (5) are different. Then the corresponding eigenvalues (6) of $A + \varepsilon B$ are simple for $\varepsilon$ small enough and the associated right eigenvectors admit a power series expansion*

$$v_j^{kl}(\varepsilon) = u_j^k + \sum_{s=1}^{\infty} w_{js}^{kl}\, \varepsilon^{s/n_j} \tag{7}$$

$k = 1, \ldots, r_j; \quad l = 1, \ldots, n_j,$ *where*

$$u_j^k = \sum_{p=1}^{f_j} c_j^{kp}\, x_j^p; \qquad k = 1, \ldots, r_j$$

*and the column vector*

$$c_j^k = \begin{bmatrix} c_j^{k1} \\ \vdots \\ c_j^{kf_j} \end{bmatrix}$$

*satisfies*

$$(\Phi_j - \xi_j^k\, E_j)\, c_j^k = 0. \tag{8}$$

*PROOF of Theorem 1:*

We may suppose, for the sake of simplicity, that $A$ has only one eigenvalue $\lambda$. The general case may be reduced to this one using appropriate Riesz projections (we refer to Lidskii's original paper [10] pp. 83-84, or [1] §3.9.1 for the details). We are interested in the roots $\omega$ of the characteristic equation

$$\det C(\omega, \varepsilon) \equiv \det(\omega I - J - \varepsilon \widetilde{B}) = 0, \qquad \widetilde{B} = P^{-1} B P. \tag{9}$$

As announced in the introduction, we perform on $C(\omega, \varepsilon)$ the change of variables (1), choosing $n = n_j$, the partial multiplicity corresponding to $j$. This leads to a polynomial equation

$$\det C(\mu, z) = \det[(\lambda + \mu z) I - J - z^{n_j} \widetilde{B}] = 0$$

in the new variables. Since we are mainly concerned with solutions which are close to $z = 0$, it will prove convenient to multiply $C(\mu, z)$ by the following diagonal matrices $L(z)$ and $R(z)$, partitioned conformally with $J$,

$$L(z) = \mathrm{Diag}\,[L_1^1, \ldots, L_1^{r_1}, \ldots, L_q^1, \ldots, L_q^{r_q}],$$
$$R(z) = \mathrm{Diag}\,[R_1^1, \ldots, R_1^{r_1}, \ldots, R_q^1, \ldots, R_q^{r_q}],$$

where

$$L_i^1(z) = \ldots = L_i^{r_i}(z) = \mathrm{diag}\,[z^{-1}, z^{-2}, \ldots, z^{-n_i}] \qquad \text{if } i \geq j,$$

$$L_i^1(z) = \ldots = L_i^{r_i}(z) = \mathrm{diag}\,[\underbrace{1, \ldots, 1}_{n_i - n_j}, z^{-1}, z^{-2}, \ldots, z^{-n_j}] \qquad \text{if } i < j$$

and

$$R_i^1(z) = \ldots = R_i^{r_i}(z) = \mathrm{diag}\,[1, z, z^2, \ldots, z^{n_i - 1}] \qquad \text{if } i \geq j$$

$$R_i^1(z) = \ldots = R_i^{r_i}(z) = \mathrm{diag}\,[\underbrace{1, \ldots, 1}_{n_i - n_j}, 1, z, z^2, \ldots, z^{n_j - 1}] \qquad \text{if } i < j$$

for $i = 1, \ldots, q$ (note that $n_i \geq n_j$ if and only if $i \leq j$). We now introduce the matrix $F(\mu, z) = L(z)\, C(\mu, z)\, R(z)$ and define

$$Q(\mu, z) = \det\, F(\mu, z).$$

The nonsingularity of both $L(z)$ and $R(z)$ implies that, for any given $z \neq 0$,

$$\det\, C(\mu, z) = 0 \;\Leftrightarrow\; Q(\mu, z) = 0,$$

although, of course, the condition numbers of $L(z)$, $R(z)$ diverge to $\infty$ as $z \to 0$.

Let us show that $Q$ is a polynomial in $\mu$ and $z$. For this, we split $F(\mu, z) = G(\mu, z) + H(z)$, where

$$G(\mu, z) = L(z)\, [(\lambda + \mu z)\, I - J]\, R(z)$$

is block-diagonal and

$$H(z) = -z^{n_j}\, L(z)\, \widetilde{B}\, R(z).$$

We write $\Gamma_s^k = \lambda\, I + N_s, \quad k = 1, \ldots, r_s$ where

$$N_s = \begin{bmatrix} 0 & 1 & & & \\ & \cdot & \cdot & & \\ & & \cdot & \cdot & \\ & & & \cdot & 1 \\ & & & & 0 \end{bmatrix},$$

for $s = 1, \ldots, q$, and use two straightforward properties of the matrices $L_i^k$ and $R_i^k$, namely, that

$$L_i^k(z)\, N_i\, R_i^k(z) = N_i, \qquad i = 1, \ldots, q, \quad k = 1, \ldots, r_i;$$

$$L_i^k(z)\, R_i^k(z) = z^{-1}\, I \qquad \text{whenever} \quad n_i \leq n_j$$

to check that the diagonal blocks of $G(\mu, z)$ are

$$G_i^k(\mu, z) = L_i^k\,(\mu z\, I - N_i)\, R_i^k = \begin{cases} \mu\, I - N_i & \text{if } i \geq j, \\ \operatorname{diag}\,[\underbrace{\mu z, \ldots, \mu z}_{n_i - n_j}, \mu, \ldots, \mu] - N_i & \text{if } i < j. \end{cases} \tag{10}$$

Hence, all powers of $z$ in $G$ are non-negative, and the same applies to $H$, since no negative powers appear either in $z^{n_j}\, L_i^k$ or in $R_i^k$. This proves our claim that $Q(\mu, z)$ is a polynomial.

Let us now examine $F(\mu, 0) = G(\mu, 0) + H(0)$: the block-diagonal matrix $G(\mu, 0)$ is given by equations (10) with $z = 0$. To give a careful description of $H(0)$ we need to partition $\widetilde{B}$, $F$ and $H$ conformally with $J$. We denote by

$$\widetilde{B}\,{}^{k_1 k_2}_{j_1 j_2}, \qquad j_i = 1, \ldots, q; \quad k_i = 1, \ldots, r_{j_i}; \quad i = 1, 2$$

the $n_{j_1} \times n_{j_2}$ block of $\widetilde{B}$ lying on the same rows as $\Gamma_{j_1}^{k_1}$ and on the same columns as $\Gamma_{j_2}^{k_2}$ (the corresponding blocks of $H$ and $F$ are defined likewise). Following this notation, we have

$$H^{k_1 k_2}_{j_1 j_2}(z) = -z^{n_j}\, L^{k_1}_{j_1}(z)\, \widetilde{B}^{k_1 k_2}_{j_1 j_2}\, R^{k_2}_{j_2}(z),$$

which implies, in the first place, that

$$H^{k_1 k_2}_{j_1 j_2}(0) = 0 \qquad \text{if } n_{j_1} < n_j, \tag{11}$$

due to the vanishing of $z^{n_j}\, L^{k_1}_{j_1}$ at $z = 0$. On the other hand, if $j_1 \leq j$ and $j_2 \leq j$, an elementary calculation shows that

$$H^{k_1 k_2}_{j_1 j_2}(0) = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ -\beta^{k_1 k_2}_{j_1 j_2} & * & \cdots & * & 0 & \cdots & 0 \end{bmatrix}, \tag{12}$$

where the $n_{j_2} - n_j$ elements marked with an asterisk are irrelevant to our argument, and $\beta^{k_1 k_2}_{j_1 j_2}$ is the element in the lower left corner of the block $\widetilde{B}^{k_1 k_2}_{j_1 j_2}$. The same structure (12) applies to the case $j_1 \leq j$, $j_2 > j$, but with zeros instead of asterisks. The main point of (12) is that every $\beta^{k_1 k_2}_{j_1 j_2} = y^{k_1}_{j_1}\, B\, x^{k_2}_{j_2}$ is an element of $\Phi_s = Y_s B X_s$ for $s = \max\{j_1, j_2\}$. In other words, we may find all the elements of the matrix $\Phi_j$ by looking at the lower left corners of the blocks (12) for $j_1$, $j_2 \leq j$ (or, equivalently, for $n_{j_1}, n_{j_2} \geq n_j$).

Before our final examination of $Q(\mu, 0)$, let us briefly turn to the diagonal blocks of $F(\mu, 0)$ of size $n_j$, i.e.

$$F_{jj}^{kk}(\mu,0)=\begin{bmatrix} \mu & -1 & 0 & \cdots & 0 \\ 0 & \mu & -1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ -\beta_{jj}^{kk} & 0 & 0 & \cdots & \mu \end{bmatrix}$$

The determinant of this block does not change if we add to its first column the products of its second column by $\mu$, of its third column by $\mu^2,\ldots,$ and of its last column by $\mu^{n_j-1}$. Neither does the whole determinant $Q(\mu,0)$ if we perform identical operations on the same columns of the whole matrix $F(\mu,0)$, since, according to (11) and (12), the $n_j-1$ columns of $F$ which are being multiplied by powers of $\mu$ have no non-zero elements outside $F_{jj}^{kk}$. This amounts to replacing every block $F_{jj}^{kk}$ with the block

$$\begin{bmatrix} 0 & -1 & 0 & \cdots & 0 \\ 0 & \mu & -1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ -\beta_{jj}^{kk}+\mu^{n_j} & 0 & 0 & \cdots & \mu \end{bmatrix}.$$

Let us now prove that the determinant $Q(\mu,0)$ of the matrix we finally obtain can be written in the form

$$Q(\mu,0)=\pm\,\mu^{\alpha}\ \det\left(\Phi_j-\mu^{n_j}\,E_j\right) \tag{13}$$

for a suitable $\alpha\geq 0$. Although elementary, the proof is quite messy in the general case, so we will instead illustrate the strategy on a specifical example. Take, for instance, the case $q=3,\ j=2,\ n_1=4,\ n_2=3,\ n_3=2,\ r_1=1,\ r_2=2,\ r_3=1$, i.e., $Q(\mu,0)$ is the determinant of the $12\times 12$ matrix

$$\left[\begin{array}{cccc|ccc|ccc|cc}
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\beta_{11}^{11} & * & 0 & \mu & -\beta_{12}^{11} & * & 0 & -\beta_{12}^{12} & 0 & 0 & -\beta_{13}^{11} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \mu & -1 & 0 & 0 & 0 & 0 & 0 \\
-\beta_{21}^{11} & 0 & 0 & \mu & -\beta_{22}^{11}+\mu^3 & 0 & \mu & -\beta_{22}^{12} & 0 & 0 & -\beta_{23}^{11} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -1 & 0 & 0 \\
-\beta_{21}^{21} & 0 & 0 & \mu & -\beta_{22}^{21} & 0 & 0 & -\beta_{22}^{22}+\mu^3 & 0 & \mu & -\beta_{23}^{21} & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mu
\end{array}\right]$$

There are four rows in this matrix (first, fifth, eighth and twelfth) containing one single non-zero element. In calculating the determinant of the matrix we may, therefore, remove the rows and columns corresponding to these four elements. This leaves an $8\times 8$ matrix $M_1(\mu)$ such that $Q(\mu,0)=\mu\,\det\,M_1(\mu)$. Because of the previous deletions, there are again four rows in $M_1$ with one single non-zero element. Eliminating the appropriate rows and columns of $M_1$ we get a $4\times 4$ matrix $M_2$ and, finally, deleting one row and one column of $M_2$, a $3\times 3$ matrix

$$M_3=\begin{bmatrix} -\beta_{11}^{11} & -\beta_{12}^{11} & -\beta_{12}^{12} \\ -\beta_{21}^{11} & -\beta_{22}^{11}+\mu^3 & -\beta_{22}^{12} \\ -\beta_{21}^{21} & -\beta_{22}^{21} & -\beta_{22}^{22}+\mu^3 \end{bmatrix}$$

such that $Q(\mu,0)=\mu^2\,\det\,M_3(\mu)$. But we know from (12) that the $\beta_{j_1j_2}^{k_1k_2}$ are just the elements of $\Phi_2$. Hence, $M_3(\mu)=-\Phi_2+\mu^3\,E_2$ and we obtain (13) with $\alpha=2$. The same procedure goes through to the general case exploiting in much the same way our knowledge of the block structure of the modified matrix $F(\mu,0)$.

Once we have $Q(\mu,0)$ factorized as in (13), we note that its second factor $\det\,(\Phi_j-\mu^{n_j}\,E_j)$ is a polynomial of degree $r_j$ in $\mu^{n_j}$. This is trivial if $j=1$ and a consequence of the nonsingularity of $\Phi_{j-1}$ if $j>1$. We now take $Q(\mu,z)$ as a polynomial in $\mu$ whose coefficients are continuous $z$-dependent functions. Continuous dependence of the roots of $Q$ upon its coefficients guarantees the existence of exactly $r_jn_j$ continuous functions

$$\mu_j^{kl}(z)=(\xi_j^k)^{1/n_j}+o(1),\qquad k=1,\ldots,r_j;\quad l=1,\ldots,n_j,$$

describing all solutions of $Q(\mu(z),z)=0$ for $z$ small enough (recall that some roots $\xi_j^k$ of (5) might be zero if $\Phi_j$ is singular). Expansion (4) is obtained by returning to the original variables $(\lambda,\varepsilon)$.

Finally, if all $r_j$ roots $\xi_j^k$ of equation (5) are known to be distinct, we may apply the implicit function theorem to (13) to show that the $\mu_j^{kl}(z)$ are in fact analytic functions of $z$, thus giving rise to the power series (6). ■

*PROOF of Theorem 2:*

In the conditions of Theorem 2 it is clear that, for $\varepsilon$ small enough, no eigenvalue (6) corresponding to $j$ can possibly coincide with any of the eigenvalues (4) corresponding to Jordan blocks of different dimensions.

Furthermore, given one of these simple eigenvalues $\lambda_j^{kl}(\varepsilon)$, a right eigenvector associated with it may be constructed by taking as its components the $m$ cofactors of the elements of a row of $A+\varepsilon B-\lambda_j^{kl}I$. This implies the analyticity of the eigenvector, since the elements of this latter matrix are analytic functions of $\varepsilon$ and the cofactors are simply sums of products of these elements (we recall that the eigenvector is unique up to constant multiples due to the simplicity of the eigenvalue).

Finally, let $e_j^{kl}(z)$ be a vector in the null space of $F_j^{kl}(z)=F(\mu_j^{kl}(z),z)$, where $\mu_j^{kl}(z)=(\lambda_j^{kl}(z)-\lambda)/z$. Dropping for simplicity both sub and superscripts, we have

$$F(z)e(z) = L(z)C(\mu(z), z)R(z)e(z) = 0,$$

which, due to the nonsingularity of $L(z)$ for $z \neq 0$, shows that $R(z)e(z)$ is a right eigenvector of $J + z^{n_j}\widetilde{B}$ associated with $\lambda_j^{kl}$. Hence, the first term $R(0)e(0)$ of its $\varepsilon$-expansion must be, up to a constant, equal to the zero-th order term $u_j^k$ of expansion (7). Equation (8) is finally obtained by applying to the linear system $F(0)e = 0$ the same ideas we used to simplify $Q(\mu, 0)$ in the proof of Theorem 1. ■

**Remarks:**

1. A proof of the existence of power series expansions (6) and (7) under the conditions of Theorem 2 goes back to Vishik and Lyusternik ( [13], Theorem 6, Appendix I). Their approach, however, is radically different from Lidskii's, since they *impose* both expansions (6) and (7) as formal series at the outset, recursively find all coefficients and finally prove the convergence of the series on some nontrivial interval around $\varepsilon = 0$. In their setting, the assumption that all Schur complements have non-zero distinct eigenvalues arises as a solvability condition on the system of infinitely many equations determining the coefficients of the series. Lidskii's approach in [10], on the other hand, concentrates only on the leading term, regardless of the rest of the expansion. This allows him to get more general results, avoiding at the same time the issue of convergence of the series: in those cases where a power series expansion is obtained, its convergence is a consequence of the function theoretical results invoked in the proof.

2. The computation of Schur complements is equivalent to (and may be replaced by) choosing the eigenvector matrices $X_j$ and $Y_j$ in a special way. Suppose, for instance, that $j = 2$ and $\Phi_1 = Y_1 B X_1$ is nonsingular. A straightforward calculation shows that the columns of

$$\widetilde{X}_2 = X_2 - X_1 \Phi_1^{-1} Y_1 B X_2 \tag{14}$$

and the rows of

$$\widetilde{Y}_2 = Y_2 - Y_2 B X_1 \Phi_1^{-1} Y_1 \tag{15}$$

are, respectively, right and left eigenvectors of $A$, corresponding to Jordan chains of the same length as the eigenvectors given by the rows and columns of $X_2$ and $Y_2$. Furthermore, we have $Y_1 B \widetilde{X}_2 = 0$ and $\widetilde{Y}_2 B X_1 = 0$. Hence, it suffices to define

$$\widetilde{W}_2 = \begin{bmatrix} Y_1 \\ \widetilde{Y}_2 \end{bmatrix}, \qquad \widetilde{Z}_2 = \begin{bmatrix} X_1, \ \widetilde{X}_2 \end{bmatrix} \quad \text{and} \quad \widetilde{\Phi}_2 = \widetilde{W}_2 B \widetilde{Z}_2$$

to obtain a block-diagonal matrix $\widetilde{\Phi}_2$ whose lower right block $\widetilde{Y}_2 B \widetilde{X}_2$ is precisely the Schur complement of $\Phi_1$ in the old matrix $\Phi_2$. Replacement of $X_3$ and $Y_3$ by suitable matrices leads to block-diagonalization of $\Phi_3$, provided that $\widetilde{\Phi}_2$ is nonsingular. It should be noted that, although only one of the two matrices (14) or (15) is required to reproduce the Schur complement for $j = 2$ (by block-triangularizing $\Phi_2$), both of them are needed to carry on to the following step $j = 3$.

3. Both matrices $\Phi_{j-1}$ and $\Phi_j$ need to be nonsingular in order to obtain the leading nonzero terms in all expansions (4). That is probably why Lidskii's original statement of Theorem 1 imposed nonsingularity of *both* $\Phi_{j-1}$ and $\Phi_j$. Nevertheless, as we have seen in the proof of the Theorem, only $\Phi_{j-1}$ need be nonsingular for the eigenvalue expansions (4) to hold: suppose that $\Phi_{j-1}$ is nonsingular and $\Phi_j$ is singular. Then we have

$$\det(\Phi_j - \xi E_j) = \xi^\beta q(\xi), \qquad q(0) \neq 0,$$

for a certain $\beta > 0$, i.e. (5) has $\beta$ zero and $r_j - \beta$ nonzero solutions. Hence, $(r_j - \beta)n_j$ expansions (4) have a nonzero first-order term, while all we can say about the remaining $\beta n_j$ eigenvalues is that they are of the form $\lambda_j^k = \lambda + o(\varepsilon^{1/n_j})$. This strongly suggests the possibility of interaction with eigenvalues associated with Jordan blocks of size less than $n_j$. These interactions will become much clearer in the next section with the use of the Newton diagram.

4. Lidskii pointed out in [10] that his results concerning perturbations $A + \varepsilon B$ could be readily extended to the case of analytic perturbations $A + \sum_{k=1}^{\infty} \varepsilon^k B_k$. We may further improve this statement by noting that *Theorem 1 remains valid even if the perturbation is of class* $C^1$, i.e. if the perturbed matrix is of the form

$$A + \varepsilon B + o(\varepsilon).$$

In fact, the same proof holds with minor changes. First observe that the continuity of the Riesz projections only depends on the continuity of the perturbation [7, Theorem 5.1]. Next, using a bar to denote matrices arising in this new setting, we obtain that

$$\det \overline{C}(\omega, \varepsilon) \equiv \det(\omega I - A - \varepsilon B - o(\varepsilon)) = \det C(\omega, \varepsilon) + o(\varepsilon)$$

where $C(\omega, \varepsilon)$ is given by (9). After changing to variables $\mu, z$, we have

$$\det \overline{C}(\mu, z) = \det C(\mu, z) + o(z^{n_j}),$$

so, multiplying by $L(z)$ and $R(z)$, we obtain

$$\overline{F}(\mu, z) = L(z)\overline{C}(\mu, z)R(z) = F(\mu, z) + \overline{S}(z)$$

for $\overline{S}(z) = L(z)\, M\, R(z)$, where $M$ is $o(z^{n_j})$. Now, recall that no negative power of $z$ in $L(z)$ has absolute value larger than $n_j$. This means that $\overline{S}(z) = o(1)$, so

$$\overline{Q}(\mu, 0) = \det\, \overline{F}(\mu, 0) = Q(\mu, 0)$$

and the factorization (13) still holds. Finally, although in this case $\overline{Q}(\mu, z)$ will no longer be a polynomial in both variables $\mu$ and $z$, it is a polynomial in $\mu$ whose coefficients are continuous functions of $z$. Thus, we may still guarantee that the roots of $\overline{Q}(\mu, z)$ depend continuously on $z$ to conclude the proof.

**Example:**

We consider the simplest case of a matrix having an eigenvalue which is neither semisimple nor nonderogatory: let $A$ be a $3\times 3$ matrix with a triple eigenvalue $\lambda$ of geometric multiplicity two (in our notation, $q = 2$, $n_1 = 2$, $n_2 = 1$, $r_1 = r_2 = 1$). Dropping the superscripts, we denote the two left eigenvectors by $y_1$, $y_2$, and the two right eigenvectors by $x_1$, $x_2$. We find $\Phi_1 = \beta_{11} = y_1 B x_1$ and

$$\Phi_2 = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}, \quad \beta_{ij} = y_i B x_j, \quad i, j = 1, 2.$$

We always get two eigenvalues

$$\lambda_1^l = \lambda \pm \sqrt{\beta_{11}}\, \varepsilon^{1/2} + o(\varepsilon^{1/2}), \qquad l = 1, 2.$$

Furthermore. if both $\beta_{11}$ and $\det\, \Phi_2$ are different from zero, the third eigenvalue is $\lambda + \xi_2 \varepsilon + o(\varepsilon)$, where $\xi_2 = \det\, \Phi_2/\beta_{11}$ is the solution of

$$\det(\Phi_2 - \xi E_2) = \det\, \Phi_2 - \beta_{11}\, \xi = 0.$$

Note that, if $\Phi_1$ is zero, we only know that two of the eigenvalues are $o(\varepsilon^{1/2})$ perturbations of $\lambda$, without any further indication of their asymptotic order.

Note that even in this quite simple example, it is unclear which leading powers of $\varepsilon$ are to be expected when some $\Phi_j$ is singular. Lidskii [10] gives an example where all three eigenvalues above are perturbed to order $\varepsilon^{2/3}$. One may also find in Wilkinson [14] §2.22 a similar example of a 5 by 5 matrix $A$ with Jordan blocks of sizes 3 and 2, whose perturbed eigenvalues are of order $\varepsilon^{2/5}$ for a conveniently chosen perturbation. None of these leading exponents can be explained, in principle, by any of the above results. It seems that the information that $Q(\mu, 0)$ provides in the case of singular $\Phi_j$ is only helpful in predicting which powers of $\varepsilon$ *cannot* appear in the eigenvalue expansions. In the following section we present an alternative approach that will improve our understanding of Theorem 1, giving us a global picture of what happens in the degenerate case when some $\Phi_j$ is singular.

# 3 Application of Newton's diagram

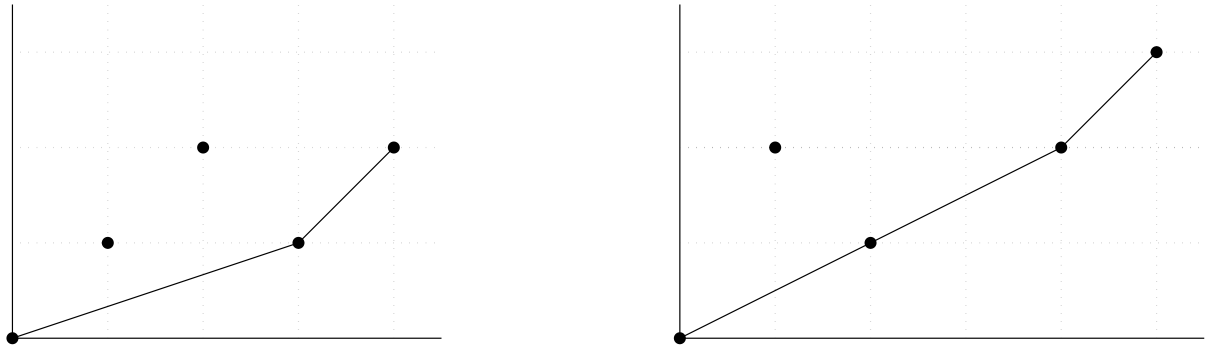
Consider a complex polynomial equation[1]

$$P(\lambda, \varepsilon) = \lambda^m + \alpha_1(\varepsilon)\lambda^{m-1} + \ldots + \alpha_{m-1}(\varepsilon)\lambda + \alpha_m(\varepsilon) = 0 \tag{16}$$

in $\lambda$, with analytic coefficients

$$\alpha_k(\varepsilon) = \hat{\alpha}_k \varepsilon^{a_k} + \ldots, \qquad k = 1, \ldots, m,$$

where $a_k$ is the leading exponent and $\hat{\alpha}_k$ the leading coefficient of $\alpha_k(\varepsilon)$ (i.e. $\hat{\alpha}_k \neq 0$ and no term of order lower than $a_k$ appears in the expansion of $\alpha_k(\varepsilon)$). It is well known ( [1], [7]) that the roots of (16) are given by expansions in fractional powers of $\varepsilon$. The leading exponents of these expansions can be easily found through the following elementary geometrical construction: we plot the values $a_k$ versus $k$ for $k = 1, \ldots, m$ together with the point $(0, 0)$ corresponding to $\lambda^m$ (if $\alpha_j(\varepsilon) \equiv 0$, the corresponding point is disregarded). Then we draw the segments on the lower boundary of the convex hull of the plotted points. These segments constitute the so-called *Newton diagram* associated with $P(\lambda, \varepsilon)$ (see Fig. 1 for two specific examples). One can prove ( [1] Appendix A7, [2], [12]) that the slopes of the different segments



a) b)

Figure 1: *Newton diagrams associated with the polynomials: a)* $\lambda^4 + (2\varepsilon - \varepsilon^2)\lambda^3 + \varepsilon^2\lambda^2 + (\varepsilon - \varepsilon^3)\lambda + \varepsilon^2$; *b)* $\lambda^5 - \varepsilon^2\lambda^4 + (\varepsilon - 3\varepsilon^2)\lambda^3 + \varepsilon^2\lambda - \varepsilon^3$

of the Newton diagram are precisely the leading powers of the $\varepsilon$-expansions of the roots $\lambda = \lambda(\varepsilon)$ of (16). The number of roots corresponding to each slope equals the length of the projection on the horizontal axis of the segment with that particular slope. The underlying idea is to substitute an *Ansatz*

[1] In this section the symbol $\lambda$ is used as a parameter, not a fixed value as earlier.

$$\lambda(\varepsilon) = \mu\varepsilon^{\beta} + \dots \tag{17}$$

into (16), with $\mu$, $\beta$ to be determined. Every point $(k, a_k)$ plotted in the diagram produces an $\varepsilon^{a_k+(m-k)\beta}$ term. If $\lambda$ is a root of (16), all the terms we obtain from this substitution must cancel each other. Hence, at least two terms of the lowest order in $\varepsilon$ must be present, and this lowest order is clearly to be found among the exponents $\{a_1 + (m-1)\beta, a_2 + (m-2)\beta, \dots, a_m\}$. Consider the segment $S$ of the Newton diagram with the smallest slope $s$, and choose $\beta = s$ in (17). All points $(k, a_k)$ lying on $S$ give rise to terms with the same exponent, since $a_k + (m-k)s$ is constant on $S$. The fact that no point $(k, a_k)$ lies below $S$ implies that no other term of the expansion can be of a lower order in $\varepsilon$. Hence, the leading coefficients $\mu$ are determined as the solutions of

$$\sum_{(k,a_k)\in S} \mu^{m-k}\hat{\alpha}_k = 0. \tag{18}$$

We get the leading terms of the remaining roots of (16) by repeating the same argument for the increasingly larger slopes appearing in the Newton diagram.

Returning to the eigenvalue problem, our main goal in this section is to establish the relationship between the quantities $\det \Phi_j$, $j = 1, \dots, q$, and the Newton diagram associated with the characteristic polynomial of $A+\varepsilon B$. We recall that $A$ is a matrix with only one eigenvalue $\lambda$ of multiplicity $m$, with partial multiplicities $n_j$, each repeated $r_j$ times, for $j = 1, \dots, q$. With no loss of generality, we may assume that this eigenvalue is zero. In that case, the characteristic polynomial $p(\lambda, \varepsilon) = \det(\lambda I - A - \varepsilon B) = \det(\lambda I - J - \varepsilon \tilde{B})$ can be written in the form (16), with $m = \sum_1^q r_j n_j$. In order to draw the Newton diagram associated with $p(\lambda, \varepsilon)$, we need to know the exponents $a_k$ for $k = 1, \dots, m$. This is quite easy if the eigenvalue is semisimple, since the Jordan form $J$ of $A$ is zero and each $\alpha_k(\varepsilon)$ equals $\varepsilon^{m-k}$ multiplied by a certain sum of minors of $\tilde{B}$ of dimension $m-k$. In this case, the Newton diagram is formed by one single segment of slope $s = 1$. If the eigenvalue is not semisimple, some nontrivial Jordan block appears in $J$, which means that, apart from the $O(\varepsilon^{m-k})$ terms, each $\alpha_k(\varepsilon)$ will typically contain terms of lower order generated by the ones appearing above the diagonal of $J$. This clearly shows that the effect of nontrivial Jordan blocks is to introduce in the Newton diagram line segments with slopes less than 1, with the smallest possible slope corresponding to the case of a nonderogatory eigenvalue (one single segment of slope $1/m$) and the largest possible one to the semisimple case. All possible Newton diagrams for the given multiplicity $m$ lie between these two extremal segments.

We now must carefully determine which points $(k, a_k)$ may appear on the Newton diagram for a particular Jordan structure. To that purpose, it will be useful to find the lowest possible diagram compatible with the given Jordan structure: to do this, we fix every exponent $l$ of $\varepsilon$ and find the largest possible $k = k(l)$ such that there exists a perturbation matrix $B$ for which $a_{k(l)} = l$. This amounts to fixing a height $l$ on the vertical axis of the Newton diagram, and determining the rightmost possible point $(k, a_k)$ in the diagram. The following theorem gives us the values $k(l)$ for the exponents $l = 1, \ldots, f_q$ which are relevant to our argument (we recall that $f_j = r_1 + \ldots + r_j$) and, more importantly, also provides some coefficients of the characteristic polynomial which are crucial to determine the Newton diagram.

**Theorem 3** *For every $l$, $l = 1, \ldots, f_q$, the corresponding $k(l)$ is equal to the sum of the dimensions of the $l$ largest Jordan blocks of $J$. More precisely, write $f_0 = 0$ and suppose $l = f_{j-1} + \rho$, for some $j = 1, \ldots, q$ and $0 < \rho \leq r_j$. Then,*

$$k(l) = r_1 n_1 + \ldots + r_{j-1} n_{j-1} + \rho\, n_j$$

*and the coefficient of $\varepsilon^l$ in $\alpha_{k(l)}$ is equal to $(-1)^l$ multiplied by the sum of all principal minors of $\Phi_j$ corresponding to submatrices of dimension $l$ that contain the upper left block $\Phi_{j-1}$ of $\Phi_j$ (if $j = 1$,* all *principal minors of dimension $l$ are to be considered). If, in particular, $l = f_j$ for some $j \in \{1, \ldots, q\}$, then the coefficient of $\varepsilon^{f_j}$ in $\alpha_{k(f_j)}$ is $(-1)^{f_j} \det \Phi_j$.*

*PROOF of Theorem 3:*

The characteristic polynomial of $A + \varepsilon B$ is a linear combination with coefficients $\pm 1$ of all possible products of $m$ elements of the matrix $\lambda I - J - \varepsilon \widetilde{B}$, with the restriction that no two factors can be on the same row or the same column.

It is clear that the only way to obtain a product of order $\varepsilon^l$ is to choose exactly $m - l$ factors containing $\varepsilon$-independent terms (i.e. "lambdas" or "minus ones"). Furthermore, we should try to include as few lambdas as possible among these factors, since we are looking for the smallest possible power of $\lambda$. However, we are not free to make whatever choices we want: due to the special position of the $\varepsilon$-independent terms, every time we choose a minus one we are, at the same time, excluding from the product those lambdas which lie on the same row or the same column. Let us examine the restrictions: suppose an admissible choice (i.e. a choice producing a term of order $\varepsilon^l$) contains $\beta$ minus ones. These $\beta$ choices remove a certain number of lambdas, which depends on the number of Jordan blocks these $\beta$ minus ones are sampled from. This is due to the fact that the *first* minus one we choose from a particular block excludes two lambdas, while any further minus one from the same block removes only one. Suppose the $\beta$ minus ones were taken from $\gamma$ different blocks. These $\beta$ choices exclude $\beta + \gamma$ lambdas, which, together with the $m - l - \beta$ lambdas which were actually chosen in the product, cannot exceed the total number $m$ of available lambdas. We conclude that $\gamma \leq l$, i.e., we are allowed to sample minus ones from at most $l$ Jordan blocks. Hence, to produce the lowest possible power of $\lambda$ we must exhaust all minus ones from the $l$ largest possible Jordan blocks of $J$, and only then complete with lambdas until we have the $m - l$ factors.

Suppose, in the first place, that $l = f_j$ for some $j = 1, \ldots, q$. Then, there is only one way of choosing these $m-l$ factors: we must choose the $\beta = r_1(n_1-1) + \ldots + r_j(n_j-1)$ minus ones from the $l$ largest Jordan blocks, plus the $m-l-\beta = r_{j+1}n_{j+1} + \ldots + r_q n_q$ diagonal lambdas from the remaining Jordan blocks. Thus, we get $k(l) = l + \beta = r_1 n_1 + \ldots + r_j n_j$. Note that if we delete from $\lambda I - J - \varepsilon \tilde{B}$ the rows and columns corresponding to these $m-l$ elements, the $f_j \times f_j$ remaining matrix is precisely $-\varepsilon \Phi_j$, which proves our claim.

The same argument is valid in the case $\rho < r_{j+1}$, although in this case there is more than one way of building up the products: each one corresponds to a different choice of $\rho$ blocks among the $r_{j+1}$ Jordan blocks of dimension $n_{j+1}$, generating a different principal minor of $\Phi_j$ to be included in the sum. ■

Let us now introduce the following definition:

**Definition 4** *Denote* $\mathcal{P}_j \equiv (k(f_j), f_j)$, *and let* $S_j$ *be the segment of slope* $1/n_j$ *connecting* $\mathcal{P}_{j-1}$ *with* $\mathcal{P}_j$, *for* $j = 1, \ldots, q$. *We define the* Newton envelope corresponding to the Jordan structure of $J$ *as the diagram obtained by successively joining the segments* $S_1$, $S_2, \ldots, S_q$.

As a first consequence of Theorem 3, note that all points $(k(l), l)$ for $l$ between $f_{j-1}$ and $f_j$ lie along the corresponding segment $S_j$. Hence, the Newton envelope is indeed the lowest Newton diagram we were looking for. This is not, however, its most interesting feature: keep in mind that, *given a particular* $B$, only those points $(k(l), l)$ from the envelope such that $a_{k(l)} = l$ will actually be plotted in the Newton diagram. This means, in particular, that a corner point $\mathcal{P}_j$ of the Newton envelope appears on the Newton diagram only if the perturbation $B$ is such that the corresponding coefficient $\pm \det \Phi_j$ is nonzero. In other words, the Newton envelope displays the *generic* behavior of the eigenvalues of $A$ under perturbation, in the sense that it coincides with the Newton diagram in all situations except in those nongeneric cases in which the perturbation $B$ causes one of the $\Phi_j$ to be singular.

Theorem 3 largely explains the importance of the $\det \Phi_j$ in obtaining the exponents $1/n_j$ in the eigenvalue expansions. Furthermore, it opens the way for an independent proof of Lidskii's Theorem 1:

*PROOF of Theorem 1:*

Let us suppose first that $j \in \{2, \ldots, q\}$ is such that $\Phi_{j-1}$ and $\Phi_j$ are nonsingular (the case $j = 1$ is completely analogous). Then, both $\mathcal{P}_{j-1}$ and $\mathcal{P}_j$ are among the points plotted to construct the Newton diagram, and $S_j$ is one of the segments in the diagram (no point $(k, a_k)$ can lie below $S_j$). Thus, we obtain the leading exponent of expansion (4). We also get the leading coefficient by carefully examining, for the segment $S_j$, equation (18). We first note that the only candidates $(k, a_k)$ to lie on

$S_j$ are the intermediate points $\mathcal{Q}_t \equiv (k(f_j - t), f_j - t)$, $t = 1, \ldots, r_j - 1$. The fact that the $\mathcal{Q}_t$ are separated from each other by a distance $n_j$ on the horizontal axis implies that equation (18) depends on $\mu$ only through $\mu^{n_j}$. More precisely, let $T$ be the set of values $t \in \{1, \ldots, r_{j-1}\}$ such that $\mathcal{Q}_t$ appears in the Newton diagram. Then,

$$\sum_{(k,a_k)\in S_j} \mu^{m-k}\hat{\alpha}_k = \mu^{m-k(f_j)} \left[ \mu^{n_j r_j}\hat{\alpha}_{k(f_{j-1})} + \sum_{t\in T} \hat{\alpha}_{k(f_j-t)}\mu^{t\,n_j} + \hat{\alpha}_{k(f_j)} \right] = 0.$$

where the expression in brackets is a polynomial in $\mu^{n_j}$. Now, we recall from Theorem 3 that for each $l = f_j - t$ with $t \in T$, the corresponding $\hat{\alpha}_{k(l)}$ is (up to the sign) the sum of all principal minors of $\Phi_j$ of dimension $l$ containing $\Phi_{j-1}$, which is precisely the way the coefficients of the powers of $\xi$ are obtained in $\det(\Phi_j - \xi E_j)$. This implies that we get the nonzero solutions of (18) by solving $\det(\Phi_j - \mu^{n_j} E_j) = 0$.

Now, suppose that $\Phi_j$ is singular. Then, the corresponding point $\mathcal{P}_j$ no longer belongs to the diagram, implying the loss of some of the expansions (4) or, equivalently, the loss of part of the segment $S_j$. The part of $S_j$ that actually remains depends upon the nullity of $\Phi_j$ : if rank $\Phi_j = f_j - \beta$, there are $\beta$ rows or columns of $\Phi_j$ that depend linearly on the remaining ones. This means, on one hand, that no point $\mathcal{Q}_t \equiv (k(f_j - t), f_j - t)$ appears in the diagram for $1 \le t < \beta$, and, on the other hand, that $\mathcal{Q}_\beta$ does appear (each principal minor of $\Phi_j$ of dimension $f_j - \beta$ either vanishes or takes a common nonzero value, since only $f_j - \beta$ columns of $\Phi_j$ are linearly independent). We conclude that the part of $S_j$ remaining in the Newton diagram is the segment connecting $\mathcal{P}_{j-1}$ with $\mathcal{Q}_\beta = (k(f_j - \beta), f_j - \beta)$ (see Fig. 2). This accounts for $(r_j - \beta)n_j$ expansions (4), whose leading coefficients are, reasoning as above, the $n_j$-th roots of the $r_j - \beta$ nonzero solutions of equation (5). As to the $\beta n_j$ remaining eigenvalues, they correspond to segments whose slope is strictly larger than $1/n_j$. Hence, expansion (4) is still valid, since they correspond to the $\beta$ null solutions of equation (5). ∎

A further consequence of this Newton diagram approach is that the hypotheses of Theorem 2 can be weakened in the sense that, if $j$ is such that all roots of (5) are distinct, we only need $\Phi_1, \ldots, \Phi_j$ to be nonsingular to guarantee the simplicity of the $r_j n_j$ eigenvalues $\lambda_j^{kl}$: the slope of any segment of the diagram lying to the right of $\mathcal{P}_j$ is strictly larger than $1/n_j$ regardless of the singularity of $\Phi_s$, $s = j+1, \ldots, q$. Hence, no eigenvalue corresponding to a Jordan block of size less than $n_j$ can coincide with any $\lambda_j^{kl}$ if $\varepsilon$ is small enough.

Apart from recovering the results of Section 2, the approach through the Newton diagram is quite helpful in getting a better understanding of the nongeneric case when some $\Phi_j$ is singular: the fact that, in that case, $\mathcal{P}_j$ does not belong to the Newton diagram implies that neither the complete segment $S_j$ nor the complete

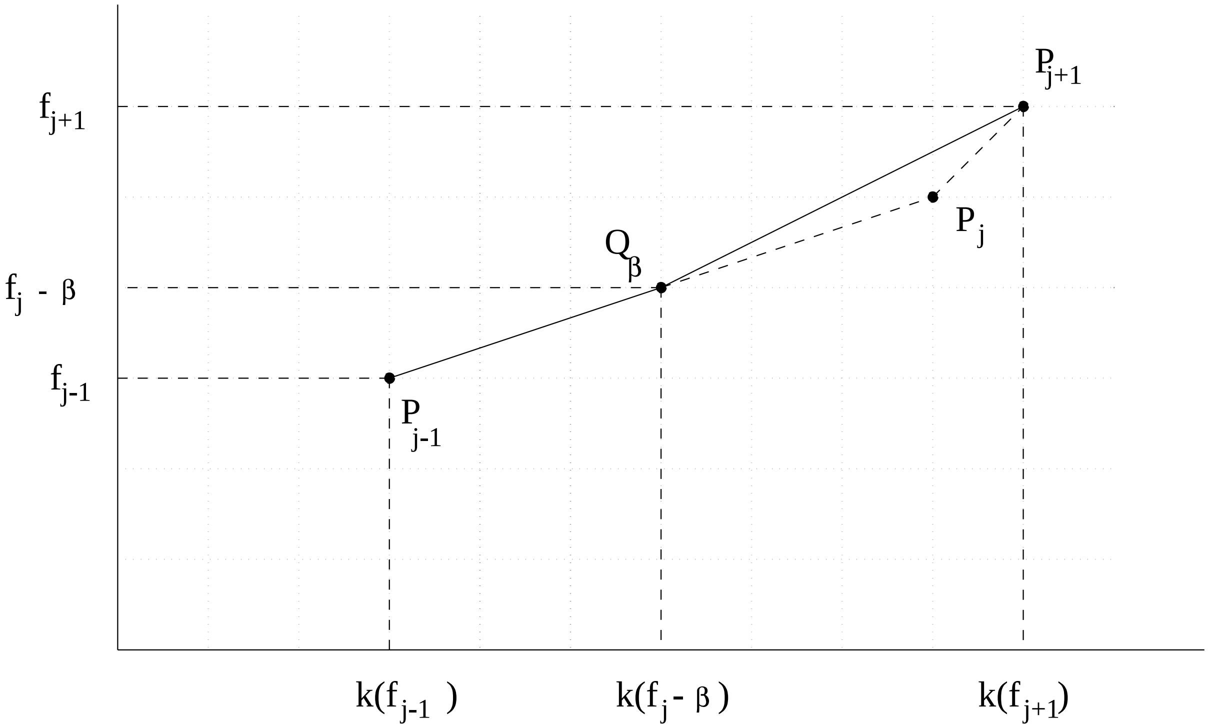



Figure 2: *The Newton diagram is shown as a solid line, and the envelope as a dashed line.*

$S_{j+1}$ appear on the Newton diagram. This indicates some kind of interaction between the eigenvalues associated with blocks of size $n_j$ and those associated with blocks of size $n_{j+1}$. We may, in fact, use the Newton diagram as a tool to quantify this interaction, actually finding both leading exponents and coefficients of the missing eigenvalue expansions in simple situations. The range of possibilities is easily visualized with the aid of the Newton envelope. For example, if $(h_1, k_1)$ and $(h_2, k_2)$ are two points that are known to lie on both the Newton diagram and the Newton envelope, then the segment of the Newton diagram between $h_1$ and $h_2$ must necessarily lie between the chord connecting $(h_1, k_1)$ to $(h_2, k_2)$ and the Newton envelope. Thus, to determine the Newton diagram one need only focus on the integer lattice points trapped between this chord and the Newton envelope. As an illustration of the power of this observation in the nongeneric case, we provide the following corollary. In this corollary, we identify a case in which no integer lattice points can lie between the chord and the Newton envelope.

**Corollary 5** *Let $0 \leq \beta \leq r_j$ and $0 \leq \alpha \leq r_{j+1}$. Suppose that $\mathcal{Q}^j_\beta = (k(f_j - \beta), f_j - \beta)$ and $\hat{\mathcal{Q}}^j_\alpha = (k(f_j + \alpha), f_j + \alpha)$ are two points lying on the Newton diagram, while the points $\mathcal{Q}^j_s$ for $s = \beta - 1, \ldots, 1$, $\mathcal{P}_j$, and $\hat{\mathcal{Q}}^j_t$ for $t = 1, 2, \ldots, \alpha - 1$ do not lie* *on the Newton diagram. Set* $p = \beta n_j + \alpha n_{j+1}$ *and* $\sigma = \frac{\beta+\alpha}{p}$. *If*

$$(\sigma n_j - 1)\beta \leq \min(\sigma, 1-\sigma), \tag{19}$$

*then there are* $p$ *eigenvalues of* $A + \epsilon B$ *of the form*

$$\lambda^l(\epsilon) = \lambda + \eta^{1/p}\epsilon^\sigma + o(\epsilon^\sigma), \qquad l = 1, 2, \ldots, p, \tag{20}$$

*where* $\eta \neq 0$. *Moreover, if either*

*i) the inequality in (19) is strict*

*or*

*ii)* $(\sigma n_j - 1)\beta < \sigma$ *and* $\alpha = n_{j+1} = 1$,

*then*

$$\eta = -\frac{\hat{\alpha}_{k(f_j+\alpha)}}{\hat{\alpha}_{k(f_j-\beta)}} . \tag{21}$$

*PROOF of Corollary 5:*

As noted above the Newton diagram must lie between the chord connecting $\mathcal{Q}_\beta^j$ to $\hat{\mathcal{Q}}_\alpha^j$ and the Newton envelope on the interval $[k(f_j - \beta), k(f_j + \alpha)]$. Furthermore, since none of the points $\mathcal{Q}_s^j$ for $s = \beta - 1, \ldots, 1$, $P_j$, and $\hat{\mathcal{Q}}_t^j$ for $t = 1, 2, \ldots, \alpha - 1$ lie on the Newton diagram, the Newton envelope and diagram only coincide at the end points of this interval. Thus, the expansions (20) will be valid if we can show that there are no integer lattice points strictly between the chord and the Newton diagram on the interval $[k(f_j - \beta), k(f_j + \alpha)]$. In order to do this we need only show that the lattice points $(k(f_j) - 1, f_j)$ and $(k(f_j) + 1, f_j + 1)$ lie on or above the chord. The condition that $(k(f_j) - 1, f_j)$ lies on or above the chord yields the inequality $(\sigma n_j - 1)\beta \leq \sigma$, while the condition that $(k(f_j) + 1, f_j + 1)$ lies on or above the chord yields the inequality $(\sigma n_j - 1)\beta \leq 1 - \sigma$. Note that this second condition is no longer needed if $\alpha = n_{j+1} = 1$, since in that case $(k(f_j) + 1, f_j + 1)$ coincides with $\hat{\mathcal{Q}}_\alpha^j$. Thus, under either condition *i)* or *ii)*, $\mathcal{Q}_\beta^j$ and $\hat{\mathcal{Q}}_\alpha^j$ are the only integer lattice points on the chord and so (21) follows from(18). ∎

It is interesting to consider a few special cases of the above result. Note that if $\beta = 0$, then $\alpha$ can take any of the values $0, 1, \ldots, r_{j+1}$, and if $\alpha = 0$, then $\beta$ can take any of the values $0, 1, \ldots, r_j$. The case $\beta = 0$ reaffirms the third remark at the end of Section 2, while the case $\alpha = 0$ illustrates that expansions with power $1/n_j$ are possible even if $\Phi_{j-1}$ is singular. Finally, if one is given a fixed value for either $\alpha$ or $\beta$, then simple bounds on the other value are easily obtained from (19). For example, if $\alpha = 1$, then the restriction (19) yields the inequality $\beta(n_j - n_{j+1} - 1) \leq 1$. That is, if $n_j = n_{j+1} + 1$, then $\beta$ can take on any of the values $0, 1, \ldots, r_j$; if $n_j = n_{j+1} + 2$, then $\beta$ can only take the values 0 and 1, and if $n_j > n_{j+1} + 2$, then $\beta$ must be zero.

The situation becomes more complicated with the introduction of more integer lattice points between the chord and the Newton envelope since more possibilities for the Newton diagram arise. But, in many cases, these possibilities can be delineated by considering certain key lattice points as was done in the proof of the above Corollary. Indeed, this approach can provide a fairly complete picture in many particular cases. Let us conclude by applying the ideas of this section to some specific examples:

**Example 1.** We first turn to our example in the preceding section of a $3 \times 3$ matrix with $q = 2,\ n_1 = 2,\ n_2 = 1,\ r_1 = r_2 = 1$. The expansions we obtained in the previous section correspond to a Newton diagram with two segments: $S_1$ connecting $(0, 0)$ with $\mathcal{P}_1 = (2, 1)$, and $S_2$ joining $\mathcal{P}_1$ and $\mathcal{P}_2 = (3, 2)$. Note that this is precisely the Newton envelope corresponding to the given Jordan structure (see Fig. 3 a)).

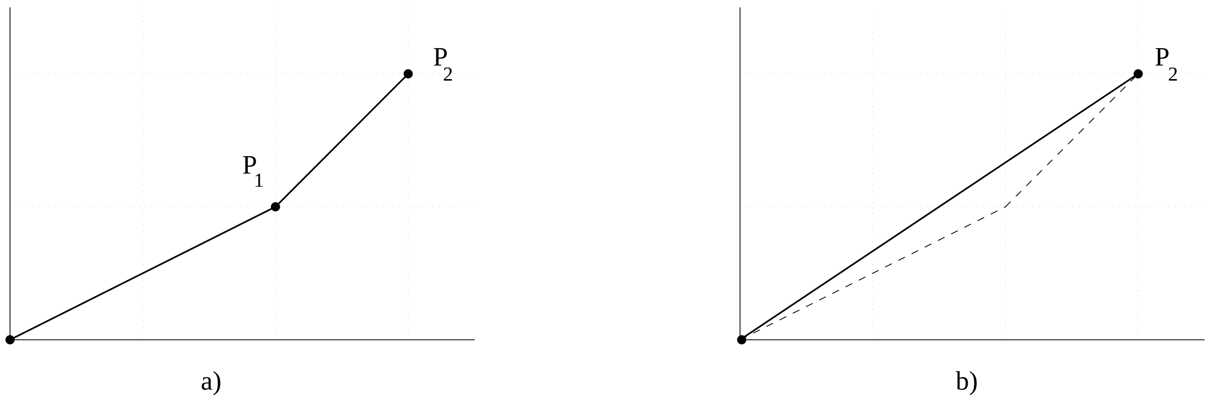



Figure 3: *Newton diagrams corresponding to Example 1. In Fig. a), the Newton diagram and envelope coincide. In Fig. b), the diagram is shown as a solid line and the envelope as a dashed line.*

Now, suppose that $\Phi_1 = 0$ with $\det \Phi_2 \neq 0$. This means that $\mathcal{P}_1$ no longer is plotted, so the diagram consists of a single segment of slope $2/3$ joining $(0, 0)$ with $\mathcal{P}_2$ (see Fig. 3 b)). If both $\Phi_1$ and $\Phi_2$ are singular, we obtain one single segment of slope $1$ as long as $B$ is nonsingular.

**Example 2.** We consider a $5 \times 5$ matrix with one zero eigenvalue and $q = 2,\ n_1 = 2,\ n_2 = 1,\ r_1 = 2,\ r_2 = 1$. We assume, for the sake of simplicity, that $A$ is already in Jordan form, i.e.

$$A = \left[\begin{array}{cc|cc|c} 0 & 1 & & & \\ & 0 & & & \\ \hline & & 0 & 1 & \\ & & & 0 & \\ \hline & & & & 0 \end{array}\right]$$

If $B = (b_{ij})_{i,j=1}^{5}$, then

$$\Phi_1 = \begin{bmatrix} b_{21} & b_{23} \\ b_{41} & b_{43} \end{bmatrix}, \qquad \Phi_2 = \left[\begin{array}{cc|c} b_{21} & b_{23} & b_{25} \\ b_{41} & b_{43} & b_{45} \\ \hline b_{51} & b_{53} & b_{55} \end{array}\right].$$

We consider the different possibilities:

- Suppose $\det \Phi_1 \neq 0$, so that $P_1 = (4,2)$ appears in the diagram. Then, the perturbed matrix has four eigenvalues

  $$\lambda_1^{kl} = (\xi_1^k)^{1/2} \varepsilon^{1/2} + o(\varepsilon^{1/2}), \qquad k,l = 1,2,$$

  where $\xi_1^1$, $\xi_1^2$ are the eigenvalues of $\Phi_1$;

  - if, additionally, $\det \Phi_2 \neq 0$, so that $P_2 = (5,3)$ also appears in the diagram, then the fifth eigenvalue of $A + \varepsilon B$ is $\lambda_2^{11} = \xi_2 \varepsilon + o(\varepsilon)$ for

    $$\xi_2 = b_{55} - \begin{bmatrix} b_{51} & b_{53} \end{bmatrix} \Phi_1^{-1} \begin{bmatrix} b_{25} \\ b_{45} \end{bmatrix}.$$

    In this case, the Newton diagram and envelope coincide (see Fig. 4 a);
  - if, on the other hand, $\Phi_2$ is singular, $P_2$ does not appear in the diagram. In that case, the order of the fifth eigenvalue is at least $O(\varepsilon^2)$, corresponding to a segment joining $\mathcal{P}_1 = (4,2)$ with $(5,4)$. However, higher slopes might appear in some cases.
- Suppose now that $\Phi_1$ is singular. Then the point $\mathcal{P}_1$ no longer appears in the Newton diagram, and to determine the order of the perturbations we need to know whether or not the coefficient

  $$\delta = -\operatorname{tr} \Phi_1 = -b_{21} - b_{43}$$

  of $\varepsilon$ in $\alpha_{k(1)}$ is different from zero, i.e., whether or not $\mathcal{Q}_1 = (2,1)$ is among the points plotted in the Newton diagram:

  - if $\delta \neq 0$, then $a_{k(1)} = 1$ and there is a segment of slope $1/2$ in the Newton diagram connecting $(0,0)$ and $\mathcal{Q}_1$. If, additionally, $\det \Phi_2 \neq 0$, there is a second segment of slope $2/3$ between $\mathcal{Q}_1$ and $\mathcal{P}_2 = (5,3)$ (see Fig. 4 b)). Hence, two of the eigenvalues are

    $$\lambda_1^l = \delta^{1/2} \varepsilon^{1/2} + o(\varepsilon^{1/2}),$$

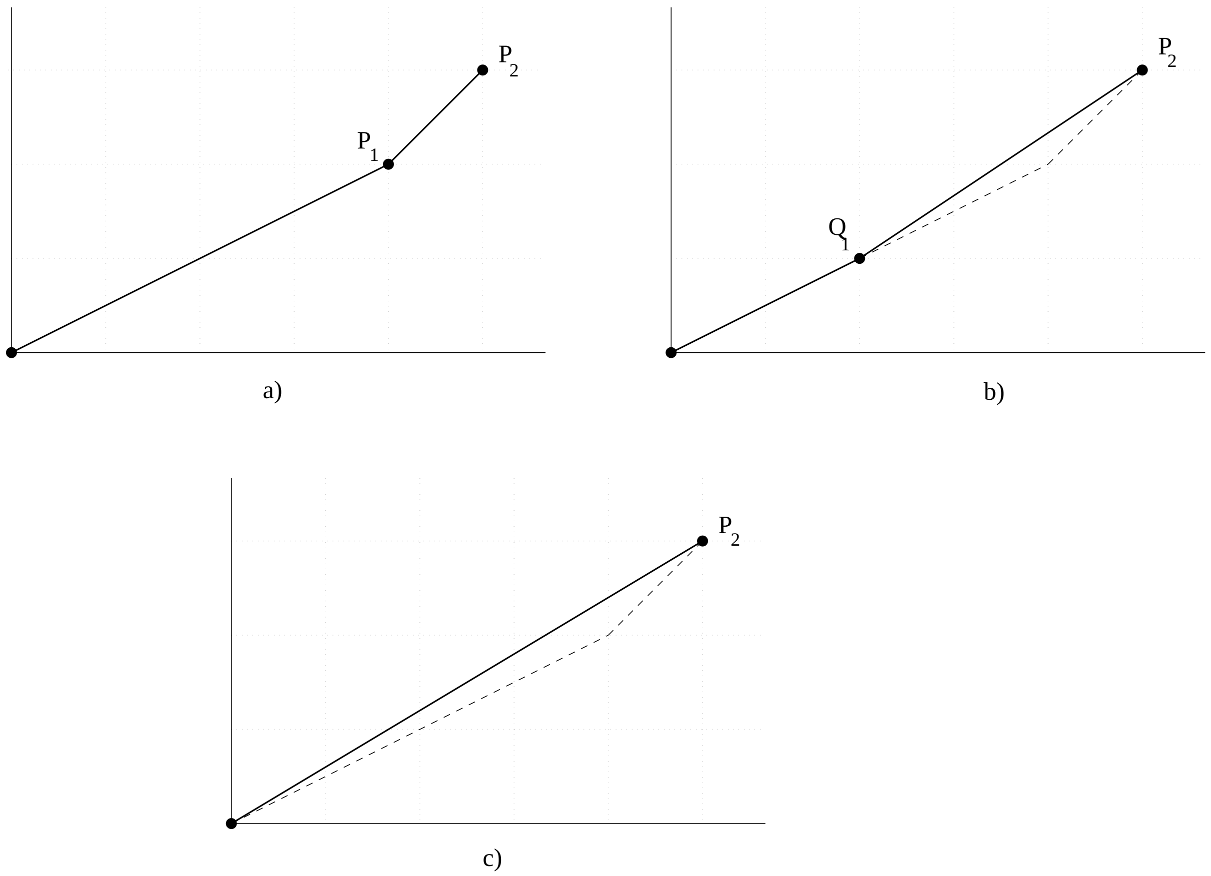



Figure 4: *Newton diagrams corresponding to Example 2. In Fig. a), the Newton diagram and envelope coincide. In Figs. b) and c), the diagram is shown as a solid line and the envelope as a dashed line.*

since $\delta = \hat{\alpha}_{k(1)}$ is the unique nonzero eigenvalue of $\Phi_1$. The other three eigenvalues of $A + \varepsilon B$ may be expanded as

$$\lambda_2^l = \eta^{1/3}\varepsilon^{2/3} + o(\varepsilon^{2/3}), \qquad \text{for } \eta = \frac{\det \Phi_2}{\delta},$$

applying Corollary 5 *ii)* (in this case $\alpha = n_2 = 1$ and $(\sigma n_1 - 1)\beta = 1/3 < 2/3 = \sigma$.)

- If $\delta = 0$, then $a_{k(1)} > 1$ and $\mathcal{Q}_1$ does not appear in the diagram. In the case when $\det \Phi_2 \neq 0$, the Newton diagram consists of a single segment of slope $3/5$ connecting the origin with $\mathcal{P}_2$ (see Fig. 4 c)). The five eigenvalues are of the form

$$(\det \Phi_2)^{1/5}\varepsilon^{3/5} + o(\varepsilon^{3/5}).$$

Finally, if both $\delta$ and $\det \Phi_2$ are zero, the actual slopes of the Newton diagram depend on the vanishing of the four-dimensional minors of $B$.

# 4 Spectral condition numbers

The results of Section 2 lead immediately to the proposition of a new notion of condition number for multiple eigenvalues:

**Definition 6** *Define the* Hölder condition number *of the eigenvalue* $\lambda$ *by*

$$\operatorname{cond}(\lambda) = (n_1, \alpha),$$

*where, as before, $n_1$ is the dimension of the largest Jordan block associated with $\lambda$, and*

$$\alpha = \max_{||B|| \leq 1} \operatorname{spr}(Y_1 B X_1),$$

*where* spr *denotes spectral radius and, as earlier, the $r_1$ columns of $X_1$ (rows of $Y_1$) are independent right (left) eigenvectors of $\lambda$, each corresponding to a Jordan chain of greatest length $n_1$.*

The motivation for this definition is that $1/n_1$ is the smallest possible power of $\varepsilon$ in the expansion of the eigenvalues of any perturbation $A + \varepsilon B$, while $\alpha^{1/n_1}$ is the largest possible magnitude of the coefficient of $\varepsilon^{1/n_1}$ in such expansions. Clearly, it follows from Theorem 1 that for all $c > 1$, the eigenvalues $\lambda'$ of $A + \varepsilon B$ converging to $\lambda$ as $\varepsilon \downarrow 0$ satisfy

$$|\lambda' - \lambda| \leq c \alpha^{1/n_1} \varepsilon^{1/n_1}$$

for all sufficiently small positive $\varepsilon$.

Note that the definition depends on the choice of matrix norm $|| \cdot ||$. We shall restrict our attention to unitarily invariant norms ( [6], p. 308).

**Theorem 7** *If the condition number* $\operatorname{cond}(\lambda) = (n_1, \alpha)$ *is defined by any unitarily invariant matrix norm $|| \cdot ||$, then*

$$\alpha = ||X_1 Y_1||_2.$$

*PROOF*:

One has

$$\begin{aligned} \max_{||B|| \leq 1} \operatorname{spr}(Y_1 B X_1) &= \max_{||B|| \leq 1} \operatorname{spr}(B X_1 Y_1) \\ &\leq \max_{||B|| \leq 1} ||B X_1 Y_1||_2 \\ &\leq ||X_1 Y_1||_2, \end{aligned}$$

where the final inequality follows because $||B||_2 \leq ||B||$ for any unitarily invariant norm. To see that equality holds, note the following. Let the scalar $\sigma_1$, the row vector $u_1$ and the column vector $v_1$ be respectively the largest singular value and the corresponding left and right singular vectors of $X_1Y_1$, so that $||u_1|| = ||v_1|| = 1$ and $u_1X_1Y_1v_1 = \sigma_1 = ||X_1Y_1||_2$. Setting $B = v_1u_1$ gives $\mathrm{spr}(BX_1Y_1) = \mathrm{spr}(u_1X_1Y_1v_1) = \sigma_1$. ∎

Note the following:

- In the case that $\lambda$ is simple, we have $n_1 = r_1 = 1$, so $\mathrm{cond}(\lambda) = (1, \alpha)$, where $\alpha = ||xy|| = ||x||\ ||y||$, with the column vector $x$ and the row vector $y$ respectively right and left eigenvectors for $\lambda$, normalized so that $yx = 1$. Without loss of generality, one can take $||x|| = 1$, so $\alpha = ||y||$. Thus the condition number reduces to the standard definition ( [4], p. 152); see also [5], p. 202, where the normalization used is $||x|| = ||y|| = 1$, giving the definition $1/(yx)$ for the condition number.
- In the case that $\lambda$ is nonderogatory, we have $r_1 = 1$, so $\mathrm{cond}(\lambda) = (n_1, \alpha)$, where $\alpha = ||xy|| = ||x||\ ||y||$, with $x$ and $y$ respectively right and left eigenvectors. However, when $n_1 > 1$, we have $yx = 0$ (directly from the Jordan form). The normalization in this case is $Q_1^1P_1^1 = I$, where the columns of $P_1^1$ (rows of $Q_1^1$) are right (left) Jordan chains for $\lambda$, $x$ being the first column of $P_1^1$ and $y$ the last row of $Q_1^1$. If $A$ is in Jordan form, then $P_1^1 = Q_1^1 = I$, so $\mathrm{cond}(\lambda) = (n_1, 1)$.

For example, take

$$A = \begin{bmatrix} \delta & 1 \\ 0 & -\delta \end{bmatrix}$$

and consider the eigenvalue $\lambda = \delta$. For $\delta > 0$, one has $x = [1 \ \ 0]^T$, $y = [1 \ \ 1/(2\delta)]$, so $\mathrm{cond}(\delta) = (1, \alpha)$ with $\alpha = ||x||\ ||y|| = \sqrt{1 + 1/(4\delta^2)}$. Since the eigenvalue is simple, this condition number coincides with those given by [4] and [5]. As $\delta \downarrow 0$, the coefficient $\alpha$ in $\mathrm{cond}(\delta)$ diverges to $+\infty$. For $\delta = 0$, the eigenvalue $\lambda = \delta$ has multiplicity two, so the definitions given in [4] (p. 152) and [5] do not apply. In this case, $A$ is in Jordan form, so one has $x = [1 \ \ 0]^T$, $y = [0 \ \ 1]$, and $\mathrm{cond}(0) = (2, 1)$. Thus, although the condition number is not a continuous function of $\delta$ in the conventional sense, the divergence of $\alpha$ as $\delta \downarrow 0$ is reflected by the drop in the power $1/n_1$ at the limit point.

Chatelin ( [4], p. 156) has also introduced a closely related Hölder condition number in the more general context of clusters of eigenvalues. Let us restrict our attention to the case where the cluster consists of one multiple eigenvalue $\lambda$. Chatelin defines a Hölder condition number $\mathrm{csp}(\lambda) = (n_1, \beta)$, with a coefficient $\beta$ which depends on the conditioning of the transformation reducing the matrix to Jordan form. An important advantage of $\mathrm{cond}(\lambda) = (n_1, \alpha)$ over $\mathrm{csp}(\lambda) = (n_1, \beta)$ is that the coefficient $\alpha$ depends only on the conditioning of the *eigenvectors*, *not* the conditioning of the Jordan vectors. For example, consider

$$A = \begin{bmatrix} 0 & \delta \\ 0 & 0 \end{bmatrix}$$

with a double eigenvalue $\lambda = 0$. When $\delta > 0$, the Jordan form of $A$ is given by

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = P^{-1}AP, \qquad P = \begin{bmatrix} 1 & 0 \\ 0 & 1/\delta \end{bmatrix}$$

with right eigenvector equal to the first column of $P$, i.e. $x = [1 \quad 0]^T$, and left eigenvector equal to the second row of $P^{-1}$, $y = [0 \quad \delta]$. Thus, for $\delta > 0$, $\mathrm{cond}(0) = (2, \alpha)$ with $\alpha = ||x||\,||y|| = \delta$. The Chatelin condition number is $\mathrm{csp}(0) = (2, \beta)$ with $\beta = 1/\delta$ because the condition number of $P$ is $1/\delta$. As $\delta \downarrow 0$, the coefficient $\alpha$ in $\mathrm{cond}(\lambda) = (2, \alpha)$ converges to zero, as it should since, at the limit point, the power $1/n_1$ *increases* to 1, giving the perfect condition number $\mathrm{cond}(0) = (1, 1)$. However, the coefficient $\beta$ in $\mathrm{csp}(0) = (2, \beta)$ diverges to $+\infty$, even though the condition number in the limiting case $\delta = 0$ is also $\mathrm{csp}(0) = (1, 1)$.

The condition number $\mathrm{cond}(\lambda)$ is trivially extended to clusters of eigenvalues by defining it to be the lexicographic maximum of the ordered pairs defining the condition number for each element of the cluster.

**Acknowledgments:** The third author thanks Prof. A. Seyranian for bringing [13] to his attention several years ago and for suggesting the dedication of this paper.

# References

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

[2] J. V. BURKE AND M. L. OVERTON, *Stable perturbations of nonsymmetric matrices*, Linear Algebra Appl., 171 (1992), pp. 249-273.

[3] J. V. BURKE AND M. L. OVERTON, *Differential properties of the spectral abscissa and the spectral radius for analytic matrix-valued mappings*, Nonlinear Analysis, 23 (1994), pp. 467-488.

[4] F. CHATELIN, *Eigenvalues of Matrices*, Wiley, 1993.

[5] G. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, First ed., Johns Hopkins University Press, Baltimore, 1983.

[6] R. HORN AND C. R. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.

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

[8] H. LANGER AND B. NAJMAN, *Remarks on the perturbation of analytic matrix functions III*, Integr. Equat. Oper. Th., 15 (1992), pp. 796-806.

[9] H. LANGER AND B. NAJMAN, *Leading coefficients of the eigenvalues of perturbed analytic matrix functions*, Integr. Equat. Oper. Th., 16 (1993), pp. 600-604.

[10] V. B. LIDSKII, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Maths. Math. Phys., 1 (1965), pp. 73-85 (*Zh. vychisl. Mat. mat. Fiz., 6 (1965) pp. 52-60*).

[11] J. H. MADDOCKS AND M. L. OVERTON, *Stability theory for dissipatively perturbed Hamiltonian systems.* To appear in Comm. Pure Appl. Math.

[12] M. M. VAINBERG AND V. A. TRENOGIN, *Theory of Branching of Solutions of Non-linear Equations*, Noordhoff, Leyden, 1974.

[13] M. I. VISHIK AND L. A. LYUSTERNIK, *The solution of some perturbation problems for matrices and selfadjoint or non-selfadjoint differential equations I*, Russian Math. Surveys, 15 (1960), pp. 1-74 (*Uspekhi Mat. Nauk, 15 (1960), pp. 3-80*).

[14] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.