On matrix perturbations with minimal leading Jordan structure ⋆

Claude-Pierre Jeannerod

INRIA - Laboratoire LIP - Ecole Normale Supérieure de Lyon
46 Allée d'Italie, 69364 Lyon cedex 07, France

Abstract

We show that any matrix perturbation of an $n \times n$ nilpotent complex matrix is similar to a matrix perturbation whose leading coefficient has minimal Jordan structure. Additionally, we derive the property that, for matrix perturbations with minimal leading Jordan structure, the sufficient conditions of Lidskii's perturbation theorem for eigenvalues are necessary too. It is further shown how minimality can be obtained by computing a similarity transform whose entries are polynomials of degree at most n. This relies on an extension of both Lidskii's theorem and its Newton diagram-based interpretation.

Key words: matrix perturbations, nilpotent Jordan structure, eigenvalues, Newton diagram, matrix similarity
PACS:

1 Introduction

Let $A(\epsilon) = A_0 + A_1\epsilon + O(\epsilon^2)$ be an $n \times n$ matrix perturbation of a nilpotent matrix A_0 over $\mathbb{C}$. Assuming that the Jordan canonical form of the leading matrix A_0 is known from the outset, Lidskii [18] provided sufficient conditions on some entries of A_1 for $A(\epsilon)$ to admit rm eigenvalues with order $O(\epsilon^{1/m})$ (or $\epsilon^{1/m}$-eigenvalues) when A_0 has r nilpotent Jordan blocks with size m; he further showed how to recover the associated leading coefficients from these particular entries exclusively. However, though generically satisfied, Lidskii's conditions are not necessary and the same eigenvalue splitting may occur

⋆ This work was done while the author was a postdoctoral fellow at the Symbolic Computation Group, University of Waterloo.
Email address: `cpjeanne@ens-lyon.fr` (Claude-Pierre Jeannerod).
URL: `http://www.ens-lyon.fr/~cpjeanne` (Claude-Pierre Jeannerod).

for some other leading Jordan structures. In a sense, such leading Jordan structures are misleading. Consider for example the 5×5 matrix perturbations $A(\epsilon)$ and $\tilde{A}(\epsilon) = P(\epsilon)^{-1}A(\epsilon)P(\epsilon)$ given by

$$
A(\epsilon) = \left[\begin{array}{cccc|c} 0 & 1 & & & \\ & 0 & 1 & & \\ \epsilon & & 0 & 1 & \\ & & \epsilon & 0 & \epsilon \\ \hline -\epsilon & & & & 0 \end{array}\right], \tilde{A}(\epsilon) = \left[\begin{array}{ccc|cc} 0 & 1 & & & \\ & 0 & 1 & & \\ \epsilon & & 0 & \epsilon & \\ \hline & & & 0 & 1 \\ & & & \epsilon & 0 \end{array}\right], P(\epsilon) = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \epsilon & \\ & & -1 & & 1 \end{bmatrix}.
$$

Both $A(\epsilon)$ and $\tilde{A}(\epsilon)$ have the same characteristic polynomial $p(\lambda, \epsilon) = (\lambda^3 - \epsilon)(\lambda^2 - \epsilon)$ and, in particular, $A(\epsilon)$ has three $\epsilon^{1/3}$-eigenvalues and two $\epsilon^{1/2}$-eigenvalues. However, using Arnold's compact notation [1] to denote by 0^m an $m \times m$ nilpotent Jordan block and by $(0^m)^r$ the nilpotent Jordan structure that consists of r such blocks, we see that $A_0 = 0^4 0$ does not give any insight on the eigenvalue splitting for $A(\epsilon)$. In that sense, the leading Jordan structure of $A(\epsilon)$ is misleading. On the other hand, this example shows that there is a matrix $\tilde{A}(\epsilon)$ similar to $A(\epsilon)$ whose leading matrix $\tilde{A}_0 = 0^3 0^2$ matches the eigenvalue splitting $(3, 2)$.

The goal of this paper is twofold. First, we want to identify the matrix perturbations whose leading Jordan structure is non misleading, i.e. the matrix perturbations for which Lidskii's sufficient conditions are also necessary. It is shown that this is the class of matrix perturbations whose leading Jordan form is minimal for lexicographic ordering. For an $A(\epsilon)$ in this class, all its eigenvalue first-order terms whose exponent is the inverse of a positive integer can therefore be recovered from A_0 and A_1 by applying (an extension of) Lidskii's perturbation theory. Second, we show that the observation made on the above example is true in the general n-dimensional case: any matrix perturbation $A(\epsilon)$ of a nilpotent matrix A_0 can be reduced by means of a similarity transform $P(\epsilon)$ to a matrix perturbation whose leading Jordan structure is minimal. This defines a reduced form for square matrix perturbations: the invariants of $A(\epsilon)$ displayed by such a reduced form are all the eigenvalue leading terms of $A(\epsilon)$ of the form $\mu\epsilon^\beta$ with $\beta^{-1} \in \mathbb{N}^*$, the set of positive integers. We further establish that, even in the case where $A(\epsilon)$ is a formal power series matrix, it is possible to find a similarity transform $P(\epsilon)$ with polynomial entries in ϵ of degree at most n.

The main tools for deriving our results are 1) the Newton diagram associated with (the characteristic polynomial of) a matrix perturbation and 2) the Newton envelope associated with its leading matrix. The concept of Newton diagram is classic when solving bivariate algebraic equations [28,5,26] and has also proved extremely useful in matrix perturbation theory [3,14,6,20,19,13,12]. On the other hand, the notion of Newton envelope was introduced recently by Moro, Burke and Overton in [20] in order to revisit Lidskii's theory and to handle some nongeneric situations this theory fails to account for. It is the natural complement to the Newton diagram of a matrix perturbation. We extend this notion further by introducing other, more specific, approximations of the Newton diagram which are conceptually close to the discrete orientation polytopes of [15]. However, as far as we know, it is the first time that such geometric objects are used in the context of the eigenvalue perturbation problem. Note also that in his graph theoretical approach [22], Murota uses the Newton diagram but is not concerned with matrix similarity.

In order for our results to hold in the general case of an arbitrary nilpotent leading matrix A_0, we extend both Lidskii's perturbation theory and the Newton diagram-based interpretation by Moro, Burke and Overton. The results presented here therefore improve upon [18,20,12] in the sense that the Jordan canonical form of A_0 needs not be known from the outset. All we suppose is that A_0 is nilpotent, a case to which we can always reduce according to [13].

Our approach is also in contrast to Arnold's normal form [1,16] where $A(\epsilon)$ is normalized by smooth similarities that keep A_0 unchanged. On the contrary, advantage is taken of the various leading Jordan forms that are possible when applying a formal similarity transform to $A(\epsilon)$. Furthermore, we ensure that such a similarity transform does not introduce any fractional power of the parameter ϵ, contrary to the reduction process of [4] in the nilpotent case.

Since we relate eigenvalue leading exponents to the inverse of some Jordan block sizes, we are not concerned with eigenvalues of the form ϵ^{β} where β^{-1} is not an integer; this corresponds to two cases: either $\beta \in \mathbb{N}\backslash\{0, 1\}$ or $\beta = \nu/\delta$ with $\gcd(\nu, \delta) = 1$, $\nu \geq 2$, $\delta \geq 2$. The former case is handled in [13], using techniques developed for computing formal solutions of linear differential systems [21,9,11,17,23,2,24]. The latter case was already mentioned by Wilkinson in [29, p. 80] ($\beta = 2/5$) and is partially covered by the Newton diagram analysis of [20]. We will see however that in both cases the minimal leading Jordan structure has at least one Jordan block of size $\lfloor \beta^{-1} \rfloor$ and one of size $\lceil \beta^{-1} \rceil$, therefore providing the best "approximation" possible for such leading exponents.

We begin our paper considering the structure of the unperturbed matrix A_0 in Section 2.1. This allows to extend Lidskii's theorem and its Newton diagram-based interpretation by Moro, Burke and Overton in sections 2.2 and 2.3 respectively. We then introduce the notion of matrix perturbations with minimal leading Jordan structure in Section 3. Section 3.1 shows that such matrices give an insight into all the eigenvalue leading terms whose exponent is the inverse of a positive integer. A formula for the minimal leading Jordan structure is proposed in Section 3.2. We establish in Section 3.3 that every matrix perturbation is similar to a matrix perturbation whose leading Jordan structure is minimal. Section 4 covers the possibility of polynomial similarity. A graphical minimization scheme is given in Section 4.1. In Section 4.2, we define a set of elementary operations for matrix perturbations. Such operations are used in Section 4.3 to minimize the leading Jordan structure by means of a similarity transform which is a polynomial matrix.

2 Matrix perturbations and leading terms of perturbed eigenvalues

Let Π_n denote the set of $n \times n$ formal power series matrices in ϵ over $\mathbb{C}$. A matrix $A(\epsilon) \in \Pi_n$ thus reads $A(\epsilon) = \sum_{i=0}^{+\infty} A_i \epsilon^i$ where the A_i's are $n \times n$ complex matrices. Consider now the roots $\lambda(\epsilon)$ of the characteristic equation $\det(\lambda I - A(\epsilon)) = 0$. It is known (see e.g. [14] or [3]) that these roots can be expanded at $\epsilon = 0$ into Puiseux series. When concerned with such local expansions, it is therefore natural to view $A(\epsilon)$ as a **matrix perturbation** of its leading term A_0 and we may write $A(\epsilon) = A_0 + O(\epsilon)$. The **leading term** of a perturbed eigenvalue $\lambda(\epsilon)$ is defined by the pair $(\mu, \beta) \in \mathbb{C}\backslash\{0\} \times \mathbb{Q}^+$ such that $\lambda(\epsilon) = \mu\epsilon^\beta + o(\epsilon^\beta)$ and we say that $\lambda(\epsilon)$ is an **ϵ^β-eigenvalue with leading coefficient** μ.

The goal of this section is to provide a extended version of both Lidskii's theorem for the leading terms of the eigenvalues of $A(\epsilon) = A_0 + O(\epsilon)$ [18,20] and its Newton diagram-based interpretation by Moro, Burke and Overton [20]. Such an extension heavily relies on the structure of the unperturbed matrix A_0, which we present first.

2.1 Structure of the leading matrix

Let us partition the $n \times n$ complex matrix A_0 as

$$A_0 = \begin{bmatrix} X & O \\ Y & Z \end{bmatrix} \tag{1a}$$

where X and Y are a priori dense matrices with respective dimensions $r_0 \times r_0$ and $(n-r_0) \times r_0$, and where Z is the nilpotent Jordan structure of order $n - r_0$

$$Z = (0^{n_1})^{r_1} \cdots (0^{n_q})^{r_q} \qquad \text{with } n_1 > \cdots > n_q. \tag{1b}$$

(It is implicit in the above notation that $r_1, \ldots, r_q$ are nonzero, whereas $r_0 \geq 0$ only.) Two special cases should be noted. In the case where $r_0 = 0$, the matrix

A_0 is the nilpotent Jordan structure Z. This is the context of the Newton diagram-based analysis of Lidskii's theorem provided by Moro, Burke and Overton [20, p. 807]. In the case where $r_0 = n$ (or, equivalently, $q = 0$), the matrix A_0 is a dense matrix whose Jordan form is unknown. The latter situation is important: we may often only know that A_0 is nilpotent, without any a priori knowledge of its Jordan structure.

For the remaining part of the paper, we will refer to (1) as the **structure** of A_0 defined by characteristics $r_0, (r_1, n_1), \ldots, (r_q, n_q)$. Two examples with $n = 8$ are

$$A_0 = \left[\begin{array}{c|cc|cc|c|c|c}
\times & & & & & & & \\ \hline
\times & 0 & 1 & & & & & \\
\times & & 0 & & & & & \\ \hline
\times & & & 0 & 1 & & & \\
\times & & & & 0 & & & \\ \hline
\times & & & & & 0 & & \\ \hline
\times & & & & & & 0 & \\ \hline
\times & & & & & & & 0
\end{array}\right], \qquad \tilde{A}_0 = \left[\begin{array}{ccc|cc|c|c|c}
0 & 1 & & & & & & \\
& 0 & 1 & & & & & \\
& & 0 & & & & & \\ \hline
& & & 0 & 1 & & & \\
& & & & 0 & & & \\ \hline
& & & & & 0 & & \\ \hline
& & & & & & 0 & \\ \hline
& & & & & & & 0
\end{array}\right].$$

Here, the structure of A_0 is given by $q = 2$, $r_0 = 1$, $(r_1, n_1) = (2, 2)$, $(r_2, n_2) = (3, 1)$. For $\tilde{A}_0$, we have $\tilde{q} = 3$, $\tilde{r}_0 = 0$, $(\tilde{r}_1, \tilde{n}_1) = (1, 3)$, $(\tilde{r}_2, \tilde{n}_2) = (1, 2)$, $(\tilde{r}_3, \tilde{n}_3) = (3, 1)$.

2.2 A perturbation theorem for eigenvalues

Let us now investigate the link between the structure of A_0 and the leading terms of the eigenvalues of $A(\epsilon)$. Consider ϵ^0-eigenvalues first. It is known (see e.g. [14]) that X in (1) has x nonzero eigenvalues $\mu_1, \ldots, \mu_x$ if and only if $A(\epsilon)$ has x eigenvalues of the form $\lambda_1(\epsilon) = \mu_1 + O(\epsilon), \ldots, \lambda_x(\epsilon) = \mu_x + O(\epsilon)$. It then follows from Z being nilpotent that the other $n - x$ eigenvalues of $A(\epsilon)$ vanish at $\epsilon = 0$. We may further notice that the number x of the nonzero eigenvalues of X is related to some minors of X as follows. By definition of the determinant of a matrix, the characteristic polynomial of X satisfies

$$\det(\lambda I - X) = \sum_{k=0}^{r_0} (-1)^k \Delta^{(k)} \lambda^{r_0 - k} \tag{2}$$

where $\Delta^{(0)} = 1$ and where, for $1 \leq k \leq r_0$, $\Delta^{(k)}$ is the sum of the principal minors of X with order k. Hence, x is the largest index k such that the sum $\Delta^{(k)}$ is nonzero, and $\mu_1, \ldots, \mu_x$ are the roots of the polynomial $\sum_{k=0}^{x} (-1)^k \Delta^{(k)} \lambda^{x-k}$ of degree x.

In order to get an insight into ϵ^β-eigenvalues with $\beta > 0$, let us do the following. Let C be the $n \times n$ complex matrix whose first r_0 columns are the first r_0 columns of A_0 and whose last $n - r_0$ columns are the last $n - r_0$ columns of A_1. On the other hand, define the partial sums

$$s_j = r_0 + \cdots + r_j \qquad \text{for } 0 \leq j \leq q.$$

Partitioning C conformally with the structure of A_0 into s_q^2 blocks C^{kl} and denoting by c_{kl} the lower left corner of each block C^{kl}, we can form the $s_j \times s_j$ matrices

$$L_j = [c_{kl}]_{1 \leq k,l \leq s_j} \qquad \text{for } 0 \leq j \leq q.$$

In particular, L_0 is the block X of A_0. Note further that L_{j-1} lies in the upper left corner of L_j. In the case where $r_0 = 0$, such submatrices of A_1 already appear in the work of Vishik and Lyusternik on differential operators [27] and, above all, in Lidskii's eigenvalue perturbation theorem [18]. Hence the following definition.

Definition 1 *For $0 \leq j \leq q$, we call j***th Lidskii submatrix** *of $A(\epsilon)$ the matrix L_j. Let also Δ_j be the determinant of L_j and, for $0 \leq k \leq r_j$, denote by $\Delta_j^{(k)}$ the sum of the principal minors of L_j with order $s_{j-1} + k$ that contain L_{j-1}.*

Note that $\Delta_{j-1} = \Delta_j^{(0)}$ and $\Delta_j = \Delta_j^{(r_j)}$. Additionally, it follows from identity (2) that the $\Delta_0^{(k)}$'s and the $\Delta^{(k)}$'s define the same quantities; in particular, $\Delta_0^{(0)} = 1$. Going back to the previous examples, rewrite the first column of A_0 as $[\heartsuit, \cdot, \spadesuit, \cdot, \spadesuit, \diamondsuit, \diamondsuit, \diamondsuit]^T$ - here, the same symbol can represent different numerical values - and partition the two 8×8 matrices A_1 and $\tilde{A}_1$ as

$$A_1 = \left[\begin{array}{c|cc|cc|c|c|c}
\cdot & \spadesuit & \cdot & \spadesuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \spadesuit & \cdot & \spadesuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \spadesuit & \cdot & \spadesuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \diamondsuit & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \diamondsuit & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \diamondsuit & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit
\end{array}\right], \qquad
\tilde{A}_1 = \left[\begin{array}{ccc|cc|c|c|c}
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\heartsuit & \cdot & \cdot & \spadesuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\spadesuit & \cdot & \cdot & \spadesuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\diamondsuit & \cdot & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\diamondsuit & \cdot & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit \\ \hline
\diamondsuit & \cdot & \cdot & \diamondsuit & \cdot & \diamondsuit & \diamondsuit & \diamondsuit
\end{array}\right].$$

It follows that the Lidskii submatrices of $A_0 + \epsilon A_1$ and $\tilde{A}_0 + \epsilon \tilde{A}_1$ are respectively

$$L_0 = \begin{bmatrix} \heartsuit \end{bmatrix}, \qquad L_1 = \begin{bmatrix} \heartsuit & \spadesuit & \spadesuit \\ \spadesuit & \spadesuit & \spadesuit \\ \spadesuit & \spadesuit & \spadesuit \end{bmatrix}, \qquad L_2 = \begin{bmatrix} \heartsuit & \spadesuit & \spadesuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \spadesuit & \spadesuit & \spadesuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \spadesuit & \spadesuit & \spadesuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \end{bmatrix},$$

and

$$\tilde{L}_1 = \begin{bmatrix} \heartsuit \end{bmatrix}, \qquad \tilde{L}_2 = \begin{bmatrix} \heartsuit & \spadesuit \\ \spadesuit & \spadesuit \end{bmatrix}, \qquad \tilde{L}_3 = \begin{bmatrix} \heartsuit & \spadesuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \spadesuit & \spadesuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \\ \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit & \diamondsuit \end{bmatrix}.$$

In this example, the entries of A_0, A_1 and $\tilde{A}_1$ that are relevant when forming the Lidskii submatrices are denoted by $\heartsuit$, $\spadesuit$, $\diamondsuit$, whereas a dot $(\cdot)$ denotes those entries which can be disregarded.

The theorem below provides some sufficient conditions on the sums $\Delta_j^{(k)}$ for the multiple zero eigenvalue of the block Z of A_0 in (1) to split under perturbation according to the number and the size of some Jordan blocks of Z. A Newton diagram-based proof can be found at the end of Section 2.3.

Theorem 2 *Let $j \in \{1, \dots, q\}$. Assuming that at least one of the sums $\Delta_j^{(k)}$ for $0 \le k \le r_j$ is nonzero, let k_1, k_2 be respectively minimal and maximal so that $\Delta_j^{(k_1)} \neq 0$ and $\Delta_j^{(k_2)} \neq 0$. Then*

i) $A(\epsilon)$ admits $(k_2 - k_1)n_j$ ϵ^{1/n_j}-eigenvalues.
ii) Writing $w_j = e^{2i\pi/n_j}$, the leading terms of these eigenvalues are

$$(\mu_{jk})^{1/n_j} w_j^l \epsilon^{1/n_j}\,, \qquad k = k_1 + 1, \dots, k_2\,, \qquad l = 1, \dots, n_j\,,$$

where the μ_{jk}'s denote the roots of the polynomial $\sum_{k=k_1}^{k_2} (-1)^k \Delta_j^{(k)} \lambda^{k_2 - k}$ and where $(\mu_{jk})^{1/n_j}$ is one of the n_j distinct n_jth roots of μ_{jk}.

Lidskii's theorem for eigenvalues [18] (see also [20, Theorem 2.1 p. 798]) follows from Theorem 2 when $r_0 = 0$ and $(k_1, k_2) = (0, r_j)$. These particular values for k_1 and k_2 actually correspond to the typical behavior, for Δ_{j-1} and Δ_j are generically nonzero. This means that in most cases the zero eigenvalue associated with r_j Jordan blocks of order n_j splits under perturbation into $r_j n_j$ ϵ^{1/n_j}-eigenvalues. Furthermore, when $(k_1, k_2) = (0, r_j)$, the polynomial $\sum_{k=0}^{r_j} (-1)^k \Delta_j^{(k)} \lambda^{r_j - k}$ can be rewritten up to the sign as $\det(L_j - \lambda E_j)$ with $E_j = \mathrm{diag}[O_{s_{j-1}}, I_{r_j}]$. The μ_{jk}'s, which define the constants of the eigenvalue leading terms, are therefore the eigenvalues of the Schur complement of L_{j-1} in L_j.

Using again the structures A_0 and $\tilde{A}_0$ already introduced as examples, let us illustrate Theorem 2 by choosing the first column of A_0 and some matrices A_1 and $\tilde{A}_1$ as follows. (Underlined figures denote the entries of A_1 and $\tilde{A}_1$ that correspond to Lidskii submatrices.)

$$A(\epsilon) = \left[\begin{array}{c|cc|cc|c|c|c}
\underline{0} & & & & & & & \\ \hline
3 & 0 & 1 & & & & & \\
\underline{2} & & 0 & & & & & \\ \hline
1 & & & 0 & 1 & & & \\
\underline{2} & & & & 0 & & & \\ \hline
\underline{3} & & & & & 0 & & \\ \hline
\underline{2} & & & & & & 0 & \\ \hline
\underline{0} & & & & & & & 0
\end{array}\right] + \epsilon \left[\begin{array}{c|cc|cc|c|c|c}
0 & \underline{1} & 0 & \underline{0} & 0 & \underline{0} & \underline{0} & \underline{0} \\ \hline
0 & 2 & 1 & 1 & 1 & 2 & 0 & 0 \\
0 & \underline{2} & 3 & \underline{1} & 1 & \underline{0} & \underline{3} & \underline{1} \\ \hline
0 & 1 & 1 & 1 & 3 & 1 & 3 & 1 \\
0 & \underline{3} & 0 & \underline{1} & 0 & \underline{3} & \underline{0} & \underline{0} \\ \hline
0 & \underline{2} & 2 & \underline{1} & 0 & \underline{1} & \underline{1} & \underline{1} \\ \hline
0 & \underline{2} & 0 & \underline{3} & 3 & \underline{1} & \underline{0} & \underline{0} \\ \hline
0 & \underline{0} & 1 & \underline{0} & 0 & \underline{0} & \underline{0} & \underline{0}
\end{array}\right] \qquad (3a)$$

and

$$\tilde{A}(\epsilon) = \left[\begin{array}{ccc|cc|c|c|c} 0 & 1 & & & & & & \\ & 0 & 1 & & & & & \\ & & 0 & & & & & \\ \hline & & & 0 & 1 & & & \\ & & & & 0 & & & \\ \hline & & & & & 0 & & \\ \hline & & & & & & 0 & \\ \hline & & & & & & & 0 \end{array}\right] + \epsilon \left[\begin{array}{ccc|cc|c|c|c} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 2 & 1 & 1 & 1 & 2 & 0 & 0 \\ \underline{2} & 2 & 3 & \underline{1} & 1 & \underline{0} & \underline{3} & \underline{1} \\ \hline 1 & 1 & 1 & 1 & 3 & 1 & 3 & 1 \\ \underline{2} & 3 & 0 & \underline{1} & 0 & \underline{3} & \underline{0} & \underline{0} \\ \hline \underline{3} & 2 & 2 & \underline{1} & 0 & \underline{1} & \underline{1} & \underline{1} \\ \hline \underline{2} & 2 & 0 & \underline{3} & 3 & \underline{1} & \underline{0} & \underline{0} \\ \hline \underline{0} & 0 & 1 & \underline{0} & 0 & \underline{0} & \underline{0} & \underline{0} \end{array}\right]. \tag{3b}$$

It could be easily verified that these two matrix perturbations actually have the same characteristic polynomial and hence the same eigenvalues. However, the structure of $\tilde{A}_0$ is "better" than the structure of A_0, for Theorem 2 allows to recover more eigenvalue leading terms with the second matrix perturbation than with the first one. For $A(\epsilon)$, we have $\Delta_2^{(0)} = 0$, $\Delta_2^{(1)} = -15$, $\Delta_2^{(2)} = 36$ and $\Delta_2^{(3)} = 0$. It then follows from Theorem 2 with $j = 2$ that $A(\epsilon)$ has one ϵ-eigenvalue and that the corresponding leading term is $-12\epsilon/5$. Applying the same theorem to $\tilde{A}(\epsilon)$ yields not only the ϵ-eigenvalue (taking $j = 3$) but also all the $\epsilon^{1/3}$-eigenvalues: one can conclude from $\tilde{\Delta}_1^{(0)} = 1$ and $\tilde{\Delta}_1^{(1)} = 2$ that there are three $\epsilon^{1/3}$-eigenvalues and that their leading terms are given by $2^{1/3}e^{2i\pi l/3}\epsilon^{1/3}$ for $l = 1, 2, 3$. Notice that since $\tilde{r}_0 = 0$, Lidskii's theorem [18] can be applied to $\tilde{A}(\epsilon)$. However, only the $\epsilon^{1/3}$-eigenvalues will be recovered, the ϵ-eigenvalue being missed because $\tilde{\Delta}_3^{(0)} = 0$.

2.3 Newton diagram-based characterization

Following Moro, Burke and Overton [20], we provide an interpretation of Theorem 2 in terms of the Newton diagram of $A(\epsilon)$. Let

$$p(\lambda, \epsilon) = \det(\lambda I - A(\epsilon)) = \sum_{i=0}^{n} a_i(\epsilon)\lambda^{n-i}$$

be the characteristic polynomial of $A(\epsilon)$. One has $a_0(\epsilon) = 1$ and more generally, for $0 \le i \le n$, there exists $(\hat{a}_i, \alpha_i) \in \mathbb{C}\backslash\{0\} \times \mathbb{N}$ such that $a_i(\epsilon) = \hat{a}_i\epsilon^{\alpha_i} + O(\epsilon^{\alpha_i+1})$ providing that $a_i(\epsilon)$ is not identically zero; we further set $\alpha_i = +\infty$ when $a_i(\epsilon)$ is zero. The lower boundary of the convex hull of $\{(i, \alpha_i) : 0 \le i \le n\}$ in the cartesian plane is known as the **Newton dia-**

gram $\mathcal{N}$ associated with $A(\epsilon)$. (See e.g. [20],[6],[3],[26].) Defining the **length** of a line segment as the length of the projection of this segment onto the horizontal axis, the Newton diagram thus consists of a finite number of segments with nonnegative rational (possibly infinite) slopes and nonzero lengths. This is illustrated in Fig. 1 below. Note that the length of the segment with slope $+\infty$ is always equal to $n - \text{rank}A(\epsilon)$ and that $(n, +\infty)$ is the rightmost vertex of $\mathcal{N}$ when $\text{rank}A(\epsilon) < n$.

Let $(\rho, \nu, \delta) \in \mathbb{N} \times \mathbb{N} \times \mathbb{N}^*$ be such that $\gcd(\nu, \delta) = 1$. The first main property of the Newton diagram is that $A(\epsilon)$ has $\rho\delta$ $\epsilon^{\nu/\delta}$-eigenvalues if and only if its Newton diagram $\mathcal{N}$ has a segment $\mathcal{S}$ with slope ν/δ and length $\rho\delta$. Now, $x_{\mathcal{S}}$ being the largest index i so that (i, α_i) belongs to this particular segment $\mathcal{S}$, we see that δ divides $x_{\mathcal{S}} - i$ for all $(i, \alpha_i) \in \mathcal{S}$. This allows to associate with each segment $\mathcal{S}$ of $\mathcal{N}$ the complex polynomial

$$p_{\mathcal{S}}(\lambda) = \sum_{(i,\alpha_i)\in\mathcal{S}} \hat{a}_i \lambda^{(x_{\mathcal{S}}-i)/\delta}$$

of degree ρ. (See e.g. [7, p. 137].) The second fundamental property of the Newton diagram is as follows: the leading terms of the $\rho\delta$ $\epsilon^{\nu/\delta}$-eigenvalues of $A(\epsilon)$ are

$$(\mu_k)^{1/\delta} w^l \epsilon^{\nu/\delta}, \qquad k = 1, \ldots, \rho, \qquad l = 1, \ldots, \delta,$$

where $w = e^{2i\pi/\delta}$ and where the μ_k's are the roots of $p_{\mathcal{S}}(\lambda)$.

For example, if $\mathcal{N}$ has a segment, say $\mathcal{S}$, with slope 0, then $x_{\mathcal{S}}$ is equal to the number x of the nonzero eigenvalues $\mu_1, \ldots, \mu_x$ of the upper left block X of A_0; one may further verify that $p_{\mathcal{S}}(\lambda)$ is the polynomial $\sum_{k=0}^{x} (-1)^k \Delta^{(k)} \lambda^{x-k}$ introduced at the beginning of Section 2.2.

It follows from this simple case that a point of $\mathcal{N}$ with ordinate zero has abscissa at most r_0, whatever the perturbative part $A(\epsilon) - A_0$ of $A(\epsilon)$ may be. Theorem 3 below generalizes this observation by expressing in terms of the characteristics $r_0, (r_1, n_1), \ldots, (r_q, n_q)$ of the structure of A_0 what the rightmost possible vertices for the Newton diagram of $A(\epsilon)$ are. This approach was originally proposed by Moro, Burke and Overton in the case of a nilpotent Jordan structure [20, Theorem 3.1 p. 807]: an ordinate y being given, what is the maximum possible abscissa $x(y)$ for a vertex of $\mathcal{N}$?

In order to give the answer in our more general setting, we recall that $s_j = r_0 + \cdots + r_j$ and we further define

$$t_j = r_0 + r_1 n_1 + \cdots + r_j n_j \quad \text{for } 0 \leq j \leq q.$$

Theorem 3 *Let $j \in \{1, \ldots, q\}$ and $k \in \{0, \ldots, r_j\}$. If $y = s_{j-1} - r_0 + k$ then $x(y) = t_{j-1} + kn_j$ and the coefficient of the monomial of $p(\lambda, \epsilon)$ in $\epsilon^y \lambda^{n-x(y)}$ is $(-1)^{s_{j-1}+k} \Delta_j^{(k)}$.*

PROOF. We use the same techniques as in [20, proof of Theorem 3.1, p. 807]. The power y of ϵ being given, we want to find a sum of minors of $\det(\lambda I - A(\epsilon))$ with the smallest power of λ. There are four categories for the entries of $\lambda I - A(\epsilon)$: the λ's, the $-\times$'s coming from the first r_0 columns of $-A_0$, the -1's coming from the Jordan blocks of the last $n - r_0$ columns of $-A_0$ and the ϵ terms. First, we choose an $A(\epsilon)$ so that ϵ^y can be formed by using y entries of A_1. By definition of the determinant, it thus remains to take $n - y$ terms among the $-\times$, -1 and λ terms in such a way that the number of λ's we use is minimum. One has $n - y = z_{-\times} + z_{-1} + z_\lambda$ where $z_{-\times}$ denotes the number of $-\times$ terms etc. The first time we take a -1 term, this removes two λ terms and then any choice of another -1 term in the same Jordan block removes only one λ term. Furthermore, every choice of a $-\times$ term removes a λ term. If z_b is the number of distinct blocks, one has $z_b + z_{-1} + z_{-\times} + z_\lambda \leq n$ or, equivalently, $z_b \leq y$. Hence, we choose the first y largest Jordan blocks. Since $z_{-\times} \leq r_0$ we then choose r_0 $-\times$ terms. In the case where $k = r_j$, one has $y = r_1 + \cdots + r_j$ and we eventually remove the first $t_j = r_0 + r_1 n_1 + \cdots + r_j n_j$ lambdas of $-\lambda I$. Consequently, $x(y) = t_j$ and the coefficient of $\epsilon^y \lambda^{t_j}$ in $\det(\lambda I - A(\epsilon))$ is $(-1)^{s_j} \Delta_j$. In the case where $k < r_j$, we choose the first $r_1 + \cdots + r_{j-1}$ largest Jordan blocks and then one can complete by taking k Jordan blocks of order n_j. This amounts to removing $t_{j-1} + kn_j$ lambdas and thus $x(y) = t_{j-1} + kn_j$. The several ways of chosing k blocks among the r_j blocks of order n_j lead to the coefficient $(-1)^{s_{j-1}+k} \Delta_j^{(k)}$.

Now, let $\mathcal{P}_0^{(k)} = (k, 0)$ for $0 \leq k \leq r_0$ and let

$$\mathcal{P}_j^{(k)} = (t_{j-1} + kn_j, s_{j-1} - r_0 + k) \qquad \text{for } 0 \leq k \leq r_j \text{ and } 1 \leq j \leq q.$$

Writing $\mathcal{P}_j = \mathcal{P}_j^{(r_j)}$, the **Newton envelope** $\mathcal{E}$ of A_0 is then defined as the diagram formed by successively connecting the points $(0,0), \mathcal{P}_0, \ldots, \mathcal{P}_q$. (The term of Newton envelope was coined by Moro, Burke and Overton in [20, p. 808].) Thus, Theorem 3 means that, the structure (1) of A_0 being given, the Newton diagram of any perturbation $A(\epsilon)$ of A_0 has no point lying below the Newton envelope $\mathcal{E}(A_0)$ associated with this structure. Furthermore, the regularity of the Lidskii submatrices used in Theorem 2 can be characterized graphically as follows.

Corollary 4 *Let $j \in \{0, \ldots, q\}$ and $k \in \{0, \ldots, r_j\}$. The sum $\Delta_j^{(k)}$ is nonzero if and only if the point $\mathcal{P}_j^{(k)}$ of the envelope also lies on the Newton diagram.*

For example, one can see by computing the characteristic polynomials of $A(\epsilon)$ and $\tilde{A}(\epsilon)$ in (3) that they yield for $0 \leq i \leq n$ the same set of (i, α_i)'s: these pairs are $(0,0)$, $(1,1)$, $(2,1)$, $(3,1)$, $(4,2)$, $(5,3)$, $(6,3)$, $(7,4)$, $(8,6)$. Hence the common Newton diagram shown in Fig. 1 together with Newton envelopes $\mathcal{E} = \mathcal{E}(A_0)$ and $\tilde{\mathcal{E}} = \mathcal{E}(\tilde{A}_0)$. We see on this example that a "better" leading structure corresponds to a higher envelope.

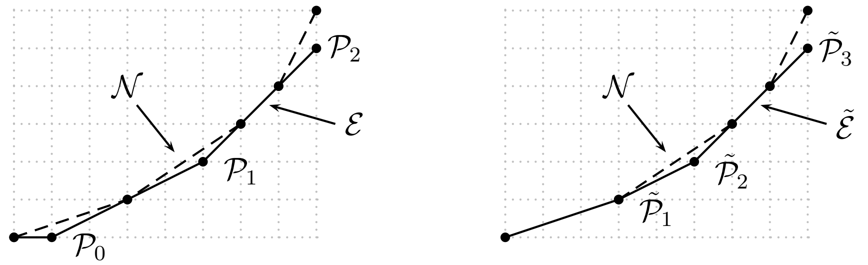


Fig. 1. Newton diagram and envelopes for matrices $A(\epsilon)$ and $\tilde{A}(\epsilon)$ of (3).

Theorem 3 and Corollary 4 further yield the following proof of Theorem 2.

PROOF. [Proof of Theorem 2] The i) part follows from Corollary 4. For the ii) part, we explicit in Theorem 3 the coordinates of the rightmost possible vertices for the Newton diagram and show that they correspond to some particular monomials of the characteristic polynomial of $A(\epsilon)$. It then follows from the definition of Lidskii submatrices that the sum of these monomials for a particular $j \in \{1, \ldots, q\}$ and from k_1 to k_2 is equal, up to the sign, to the polynomial $\sum_{k=k_1}^{k_2} (-1)^k \Delta_j^{(k)} \lambda^{k_2-k}$.

With this approach, A_0 and thus the envelope are given, whereas all the possible perturbations $A(\epsilon)$ with an imposed structure $r_0, (r_1, n_1), \ldots, (r_q, n_q)$ for A_0 and thus all the possible Newton diagrams are considered. This defines $\Pi_n(\mathcal{E})$, the set of $n \times n$ matrix perturbations having the same envelope $\mathcal{E}$. In the next section, we adopt the opposite view: instead of looking for the "lowest possible diagram compatible with the given [Jordan] structure" of A_0 [20, p. 807], we consider the highest possible envelope compatible with the Newton diagram of a given $A(\epsilon)$.

3 Matrix perturbations with minimal leading Jordan structure

Let $\Pi_n(\mathcal{N})$ be the set of $n \times n$ matrix perturbations whose Newton diagram is $\mathcal{N}$. For the remaining part of this paper we will assume that $\mathcal{N}$ has positive slopes only, i.e. the leading matrix of every element of $\Pi_n(\mathcal{N})$ is nilpotent.

Following [8] we associate with an $n \times n$ nilpotent Jordan structure J its Segré characteristics, i.e. the integer partition of n that lists in nonincreasing order the sizes of the blocks of J. We may therefore regard two nilpotent Jordan structures as identical if they differ only in the way their diagonal blocks are ordered. This further allows to compare any two $n \times n$ nilpotent Jordan structures with lexicographic ordering for integer tuples. Consider the set of the leading Jordan structures of the elements of $\Pi_n(\mathcal{N})$. It follows from well-known facts about integer partitions (see e.g. [25, p. 28]) that this set is finite and lower bounded by the $n \times n$ zero matrix. Let us denote by J_{min} its unique minimum. We say that the matrix perturbation set $\Pi_n(\mathcal{N})$ has **minimal leading Jordan structure** J_{min}. We also say that a given $A(\epsilon) \in \Pi_n$ with Newton diagram $\mathcal{N}$ is a **matrix perturbation with minimal leading Jordan structure** when $A_0 = J_{min}$.

3.1 Main property

Matrix perturbations with minimal leading Jordan structure enjoy the property that Lidskii's sufficient conditions of Theorem 2 are also necessary. This result is stated in Corollary 7 below; in order to derive it, let us characterize J_{min} graphically by refining the notion of Newton envelope of Section 2.3. We define the **highest Newton envelope** compatible with $\mathcal{N}$ as follows: one can partition $\mathcal{N}$ into groups of segments with slopes comprised in $[N_1^{-1}, (N_1-1)^{-1}]$, $[N_2^{-1}, (N_2-1)^{-1}], \cdots$ for some positive integers $N_1 > N_2 > \cdots$. The ith group is best lower bounded by two segments, say, $\mathcal{S}_{N_i}$ and $\mathcal{S}_{N_i-1}$, whose slopes are N_i^{-1} and $(N_i-1)^{-1}$ respectively. This is illustrated in Fig. 2 where we denote by ν_{ij}/δ_{ij} and $\rho_{ij}\delta_{ij}$ the slope and length of the jth segment of the ith group. Note that the lengths $N_i x_i$ and $(N_i-1)y_i$ are given by the unique solution $(x_i, y_i) \in \mathbb{N}^* \times \mathbb{N}$ of

$$\begin{bmatrix} 1 & 1 \\ N_i & N_i - 1 \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \sum_j \rho_{ij} \begin{bmatrix} \nu_{ij} \\ \delta_{ij} \end{bmatrix}. \tag{4}$$

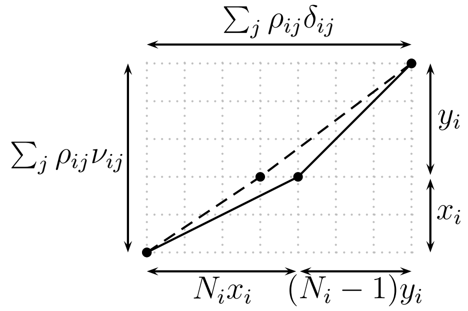


Fig. 2. Best approximation of segments of $\mathcal{N}$ with slopes $\{2/3, 3/4\} \subset [1/2, 1]$ by segments with slope $1/2$ and 1.

Connecting the $\mathcal{S}_{N_i}$'s and $\mathcal{S}_{N_i-1}$'s yields a convex polygon line. Hence the definition below.

Definition 5 *The tightest lower convex approximation of $\mathcal{N}$ by means of segments with slope in $\{\frac{1}{i} : i \in \mathbb{N}^*\}$ is called the* **highest Newton envelope** *compatible with $\mathcal{N}$. We denote it by $\mathcal{E}_{max}$.*

Of course, the maximal value for i in the above set of admissible slopes is $\lceil s^{-1} \rceil \leq n$ where s is the smallest slope of $\mathcal{N}$. Let us now characterize J_{min} graphically.

Theorem 6 *The minimal leading Jordan structure associated with $\Pi_n(\mathcal{N})$ defines the highest Newton envelope compatible with $\mathcal{N}$.*

PROOF. First, there exists $A(\epsilon) \in \Pi_n(\mathcal{N})$ so that $\mathcal{E}(A_0) = \mathcal{E}_{max}$. To see that, it is sufficient to choose for $A(\epsilon)$ the block diagonal matrix whose ith block $A^{(i)}(\epsilon)$ corresponds to the ith group of segments of $\mathcal{N}$ as follows. With the same notation as in (4), let $(x_{ij}, y_{ij}) \in \mathbb{N}^* \times \mathbb{N}$ be the unique solution of

$$\begin{bmatrix} 1 & 1 \\ N_i & N_i - 1 \end{bmatrix} \cdot \begin{bmatrix} x_{ij} \\ y_{ij} \end{bmatrix} = \rho_{ij} \begin{bmatrix} \nu_{ij} \\ \delta_{ij} \end{bmatrix}.$$

Define $A^{(ij)}(\epsilon) = (0^{N_i})^{x_{ij}} (0^{N_i-1})^{y_{ij}} + \epsilon B^{(ij)}$ where $B^{(ij)} = [B^{(ijkl)}]_{1 \leq k,l \leq x_{ij}+y_{ij}}$ is a block matrix partitioned conformally with $A^{(ij)}(0)$ that satisfies: all the blocks $B^{(ijkl)}$ are zero, except when $l = k+1$ or when $(k,l) = (x_{ij} + y_{ij}, 1)$; in these cases, the only nonzero entry of the block is its lower left corner, which is set to 1. Hence $A^{(ij)}(\epsilon)$ has $\rho_{ij}\delta_{ij}$ $\epsilon^{\nu_{ij}/\delta_{ij}}$-eigenvalues and we take $A(\epsilon) = \mathrm{Diag}[A^{(ij)}(\epsilon)]_{i,j}$. Then it is not hard to verify that $\mathcal{E}(A_0) = \mathcal{E}_{max}$ implies $A_0 = J_{min}$.

An example of highest envelope is the Newton envelope $\tilde{\mathcal{E}}$ in Fig. 1. The matrix perturbation given by (3b) has therefore minimal leading Jordan structure.

By definition of $\mathcal{E}_{max}$, every segment of $\mathcal{N}$ whose slope is the inverse of a positive integer belongs to a segment of $\mathcal{E}_{max}$. Combining Corollary 4 and Theorem 6 therefore yields the result below.

Corollary 7 *[Converse of Theorem 2 i)] Let $A(\epsilon) \in \Pi_n$ be a matrix perturbation with minimal leading Jordan structure of the form (1b). Then, if $A(\epsilon)$ has rm $\epsilon^{1/m}$-eigenvalues, there exists $j \in \{1, \ldots, q\}$ such that $m = n_j$ and $r = k_2 - k_1$ with k_1 minimal and k_2 maximal so that $\Delta_j^{(k_1)} \neq 0$ and $\Delta_j^{(k_2)} \neq 0$.*

Let $v(\mathcal{N})$ be the set of vertices of the Newton diagram $\mathcal{N}$. In order to explicit the minimal leading Jordan structure J_{min} in terms of $v(\mathcal{N})$, we introduce the integer sequence $(\sigma_i)_{i\in\mathbb{N}^*}$ given by

$$\sigma_i = 2\nu_i - \nu_{i+1} - \nu_{i-1}\,, \qquad \nu_i = \min\{n - k + i\alpha_k \,:\, (k,\alpha_k) \in v(\mathcal{N})\}. \quad (5)$$

Graphically, the ν_i's can be interpreted as follows. Let $\mathcal{N}_i$ be the diagram deduced from $\mathcal{N}$ by transforming $(k,\alpha_k) \in v(\mathcal{N})$ into $(k, n-k+i\alpha_k) \in v(\mathcal{N}_i)$. Applying the same transform to the vertices of $\mathcal{E} = \mathcal{E}(J_{min})$, we obtain the highest Newton envelope $\mathcal{E}_i$ compatible with $\mathcal{N}_i$. A segment of $\mathcal{N}$ or $\mathcal{E}$ with slope s is thus transformed into a segment of $\mathcal{N}_i$ or $\mathcal{E}_i$ with the same length but slope $is-1$. By definition of the highest envelope, ν_i is therefore the smallest ordinate for the vertices of both $\mathcal{N}_i$ and $\mathcal{E}_i$.

Consider now the σ_i's. They are nonnegative, for the sequence $(\nu_{i+1}-\nu_i)_{i\in\mathbb{N}}$ is nonincreasing. Additionally, it follows from the Lemma 8 below that $(i\sigma_i)_{i\in\mathbb{N}^*}$ defines an integer partition of n and that $(\sigma_i)_{i\in\mathbb{N}^*}$ is zero almost everywhere. Theorems 9 and 10 also rely on this lemma.

Lemma 8 *The sequence $(\sigma_i)_{i\in\mathbb{N}^*}$ given by (5) satisfies*

$$\sum_{i=j+1}^{\infty}(i-j)\sigma_i = n - \nu_j \quad \textit{for all } j \in \mathbb{N}.$$

PROOF. Let $s > 0$ be the smallest slope of $\mathcal{N}$ and let $N = \lceil s^{-1}\rceil$. For $i \geq N$, the diagram $\mathcal{N}_i$ therefore has nonnegative slopes only. Equivalently, $\nu_i = n$ for all $i \geq N$. Hence $\sigma_i = 0$ for $i > N$ and intermediate cancellations in the sum yield $\sum_{i=j+1}^{N}(i-j)\sigma_i = -\nu_j + (N-j+1)\nu_N - (N-j)\nu_{N+1}$. This is $n - \nu_j$.

Theorem 9 *Let $\sigma_{n_1}, \ldots, \sigma_{n_q}$ be the nonzero values of $(\sigma_i)_{i\in\mathbb{N}^*}$ numbered so that $n_1 > \cdots > n_q$. The minimal leading Jordan structure associated with $\Pi_n(\mathcal{N})$ is $J_{min} = (0^{n_1})^{\sigma_{n_1}} \cdots (0^{n_q})^{\sigma_{n_q}}$.*

PROOF. Let $J_{min} = (0^{m_1})^{r_1} \cdots (0^{m_p})^{r_p}$ with $m_1 > \cdots > m_p$. For $j \in \{1,\ldots,p\}$, we prove that $\sigma_{m_j} = r_j$ as follows. Denote by (x_1,y_1), (x_2,y_2) the cartesian coordinates of the vertices defining the segment of $\mathcal{E} = \mathcal{E}(J_{min})$ with slope $1/m_j$. When successively looking at the transformed envelopes $\mathcal{E}_{m_j-1}$, $\mathcal{E}_{m_j}$ and $\mathcal{E}_{m_j+1}$, we see that

$$
\begin{aligned}
\nu_{m_j-1} &= n - x_2 + (m_j - 1)y_2; \\
\nu_{m_j} &= n - x_i + m_j y_i \quad \text{for } i = 1, 2; \\
\nu_{m_j+1} &= n - x_1 + (m_j + 1)y_1.
\end{aligned}
$$

Writing $(x_2, y_2) = (x_1 + r_j m_j, y_1 + r_j)$ then yields $\sigma_{m_j} = y_2 - y_1 = r_j$. On the other hand, it follows from the r_j's being positive and from $m_1 > \cdots > m_p$ that the σ_{m_j}'s define p out of the q nonzero values of $(\sigma_i)_{i\in\mathbb{N}^*}$. Since $\sum_{i\in\{m_1,\ldots,m_p\}} i\sigma_i = n$, we conclude that $p = q$ and $m_j = n_j$ for all $j \in \{1, \ldots, q\}$.

For example, the Newton diagram $\mathcal{N}$ of Fig. 1 has vertices $(0, 0)$, $(3, 1)$, $(6, 3)$, $(7, 4)$, $(8, 6)$. It follows from (5) that $\sigma_1 = 3$, $\sigma_2 = 1$, $\sigma_3 = 1$, $\sigma_4 = \sigma_5 = \cdots = 0$. The minimal leading Jordan structure associated with $\Pi_8(\mathcal{N})$ is thus $J_{min} = 0^3 0^2 (0)^3$, i.e. $\tilde{A}_0$ in (3b).

3.3 Minimization by similarity

Let $A(\epsilon) \in \Pi_n$ be given, with Newton diagram $\mathcal{N}$. The goal of this section is to show that the minimal leading Jordan structure J_{min} associated with $\Pi_n(\mathcal{N})$ is actually the leading coefficient of a matrix perturbation similar to $A(\epsilon)$. Recall that $\Pi_n = \mathbb{C}[[\epsilon]]^{n\times n}$ and consider $A(\epsilon) \in \Pi_n$ as an element of $\mathbb{C}[[\epsilon]][\epsilon^{-1}]^{n\times n}$. The orbit of $A(\epsilon)$ is

$$
\mathcal{O}(A(\epsilon)) = \{P(\epsilon)^{-1}A(\epsilon)P(\epsilon) \,:\, P(\epsilon) \in \Pi_n, \det P(\epsilon) \neq 0\},
$$

for the poles of $P(\epsilon)$ can be removed by multiplying $P(\epsilon)$ with a suitable power of ϵ. The set of the **matrix perturbations similar** to $A(\epsilon)$ is therefore $\mathcal{O}(A(\epsilon)) \cap \Pi_n(\mathcal{N})$. (This set simply consists of the elements of $\mathcal{O}(A(\epsilon))$ which have no pole in ϵ.) In this paper, we usually denote such matrices by $\tilde{A}(\epsilon)$.

Let $p(\lambda, \epsilon) = \sum_{k=0}^{n} a_k(\epsilon)\lambda^{n-k}$ be the characteristic polynomial associated with $\mathcal{O}(A(\epsilon)) \cap \Pi_n(\mathcal{N})$. Denoting by val_ϵ the valuation function with respect to ϵ, let further $\alpha_k = \mathrm{val}_\epsilon\, a_k(\epsilon)$ for $0 \leq k \leq n$. It follows from the definition of the Newton diagram that the ν_i's in (5) are also equal to $\min\{n - k + i\alpha_k \,:\, 0 \leq k \leq n\}$. Equivalently,

$$
\nu_i = \mathrm{val}_\epsilon\, p(\lambda\, \mathrm{e}, \epsilon^i) \quad \text{for all } i \in \mathbb{N}.
$$

This equality, together with Theorem 9 and Lemma 8, leads to the following reduction theorem.

Theorem 10 *For all $A(\epsilon) \in \Pi_n$ there exists a nonsingular $P(\epsilon) \in \Pi_n$ so that $P(\epsilon)^{-1}A(\epsilon)P(\epsilon)$ is a matrix perturbation with minimal leading Jordan structure.*

PROOF. Following Moser [21, p. 388], we first reduce to the case where $A(\epsilon)$ has a single invariant factor. Consider $A(\epsilon) \in \Pi_n$ as an $n \times n$ matrix over the quotient field $\mathbb{C}[[\epsilon]][\epsilon^{-1}]$ and let $F(\epsilon) = \mathrm{diag}[C_{f_1}(\epsilon), \ldots, C_{f_l}(\epsilon)]$ be its Frobenius normal form. Here $C_{f_k}(\epsilon)$ denotes the companion matrix defined by the kth invariant factor $f_k(\lambda, \epsilon)$ of $A(\epsilon)$. (See e.g. [10, p. 948-949] or [16]). On the other hand, denote by $(\sigma_{ik})_{i \in \mathbb{N}^*}$ the sequence (5) associated with factor $f_k(\lambda, \epsilon)$. It follows from $p(\lambda, \epsilon) = \prod_{k=1}^{l} f_k(\lambda, \epsilon)$ that $\mathrm{val}_\epsilon\, p(\lambda\epsilon, \epsilon^i) = \sum_{k=1}^{l} \mathrm{val}_\epsilon\, f_k(\lambda\epsilon, \epsilon^i)$. Hence $\sigma_i = \sum_{k=1}^{l} \sigma_{ik}$ and it is possible to treat each block of the Frobenius form independently. Finally, a matrix being similar to its transpose, we are reduced to the case

$$A(\epsilon) = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -a_n(\epsilon) & \cdots & -a_2(\epsilon) & -a_1(\epsilon) \end{bmatrix}.$$

Let us now explicit a similarity $D(\epsilon)$ that transforms $A(\epsilon)$ into $B(\epsilon) = D(\epsilon)^{-1}A(\epsilon)D(\epsilon)$ so that $B(0)$ coincides with J_{min} of Theorem 9. Let

$$D(\epsilon) = \mathrm{diag}[D_q^{(\sigma_{n_q})}(\epsilon), \ldots, D_q^{(1)}(\epsilon) \mid \cdots \mid D_1^{(\sigma_{n_1})}(\epsilon), \ldots, D_1^{(1)}(\epsilon)]\,,$$

where

$$D_j^{(k)}(\epsilon) = \epsilon^{\sigma_{n_j} + \cdots + \sigma_{n_q} - k + 1} \cdot I_{n_j} \qquad \text{for } 1 \leq j \leq q \text{ and } 1 \leq k \leq \sigma_{n_j}.$$

It is not hard to see that the first $n-1$ rows of $B(\epsilon)$ coincide, at $\epsilon = 0$, with the first $n-1$ rows of J_{min} (up to ordering). It remains to prove that the last row of $B(\epsilon)$ is identically zero at $\epsilon = 0$. To do so, partition the last row $\vec{a}(\epsilon) = [a_n(\epsilon), \ldots, a_1(\epsilon)]$ of $-A(\epsilon)$ conformally with the block structure of J_{min}:

$$\vec{a}(\epsilon) = [\vec{a}_q^{(\sigma_{n_q})}(\epsilon), \ldots, \vec{a}_q^{(1)}(\epsilon) \mid \cdots \mid \vec{a}_1^{(\sigma_{n_1})}(\epsilon), \ldots, \vec{a}_1^{(1)}(\epsilon)].$$

When applied to $A(\epsilon)$, the similarity $D(\epsilon)$ transforms each subrow $-\vec{a}_j^{(k)}(\epsilon)$ of its last row $-\vec{a}(\epsilon)$ into $\vec{b}_j^{(k)}(\epsilon) = -\vec{a}_j^{(k)}(\epsilon) \cdot \epsilon^{-(\sigma_{n_1} + \cdots + \sigma_{n_{j-1}} - k - 1)}$. Now, J_{min}

defines an envelope for the Newton diagram of $B(\epsilon)$. Hence $\mathrm{val}_\epsilon\, \vec{a}_j^{(k)}(\epsilon) \geq \sigma_{n_1} + \cdots + \sigma_{n_{j-1}} + k$ and thus $\mathrm{val}_\epsilon\, \vec{b}_j^{(k)}(\epsilon) > 0$. This completes the proof.

Theorem 10 therefore defines a reduced form for square matrix perturbations. The invariants of $A(\epsilon)$ displayed by such a reduced form are all the eigenvalue leading terms of $A(\epsilon)$ of the form $\mu\epsilon^\beta$ with $\beta^{-1} \in \mathbb{N}^*$. (Of course, this reduced form is highly non-unique, for only its leading matrix is required to be normalized as J_{min}.) For example, $A(\epsilon)$ in (3a) can be reduced to $\tilde{A}(\epsilon)$ in (3b) with $P(\epsilon) = \mathrm{diag}[\epsilon, I_7]$. When $A(\epsilon)$ has polynomial entries, its Frobenius normal form is a polynomial matrix. Finding a reduction matrix $P(\epsilon)$ that is also polynomial is therefore not surprising. This is however still possible in the general case where $A(\epsilon)$ has an infinite number of matrix coefficients. We explain such a possibility in the next section.

4 Minimizing the leading Jordan structure by polynomial similarity

In this section, we restrict ourselves to polynomial similarity: the similarity transforms are nonsingular elements of $\mathbb{C}[\epsilon]^{n\times n}$ exclusively.

4.1 Computing the highest Newton envelope iteratively

For $A(\epsilon) \in \Pi_n$, let $\mathcal{E}^{(0)}$ be the Newton envelope associated with the structure (1) of A_0. In the general case where A_0 is dense, $\mathcal{E}^{(0)}$ clearly reduces to a segment with slope 0 and length n. With the definition below, we further introduce quantities that give a tighter approximation of the Newton diagram $\mathcal{N}$ of $A(\epsilon)$.

Definition 11 *For $i \geq 1$, let $\mathcal{E}^{(i)}$ be the i***th Newton envelope** *of $\mathcal{N}$, that is, the tightest lower convex approximation of $\mathcal{N}$ by means of segments with slope in $\{0\} \cup \{\frac{1}{i}, \frac{1}{i-1}, \ldots, \frac{1}{2}, 1\}$.*

Writing $N = \lceil s^{-1} \rceil$ with s the smallest slope of $\mathcal{N}$, one sees that $\mathcal{E}^{(N)}$ is the highest Newton envelope $\mathcal{E}_{max}$. Furthermore, the smallest slope of $\mathcal{E}^{(i)}$ is $1/i$ for all $1 \leq i \leq N$, and $\mathcal{E}^{(i+1)} = \mathcal{E}^{(i)}$ if and only if $\mathcal{E}^{(i+1)}$ has no segment of slope $1/(i+1)$. Equivalently,

$$\mathcal{E}^{(i+1)} = \mathcal{E}^{(i)} \text{ if and only if } i \geq N.$$

Starting with $\mathcal{E}^{(0)}$, one can therefore compute $\mathcal{E}_{max}$ iteratively in a finite number of steps as described below:

$i \leftarrow 0$;
Compute $\mathcal{E}^{(1)}$;
while $\mathcal{E}^{(i+1)} \neq \mathcal{E}^{(i)}$ do
  $i \leftarrow i+1$;
  Compute $\mathcal{E}^{(i+1)}$;
Return $\mathcal{E}^{(i)}$.

A way of computing $\mathcal{E}^{(i+1)}$ for $i \geq 0$ is to deduce $\mathcal{E}^{(i+1)}$ from $\mathcal{E}^{(i)}$ as follows: we "cut off" as much as possible the corner defined by the segments of slopes 0 and $1/i$, therefore making room for a new segment whose slope is $1/(i+1)$ and whose length is as large as possible. Denoting by $r_{0,i}$ the length of the segment of $\mathcal{E}^{(i)}$ with slope 0, one can think of this process as moving the parametrized line $\mathcal{L}_{i,h}$ of equation $x-(i+1)y = r_{0,i} - h \in \mathbb{N}$ towards $\mathcal{N}$ by increasing h until a vertex or a whole segment of $\mathcal{N}$ belongs to $\mathcal{L}_{i,h}$. Denoting by $\mathcal{E}^{(i,h)}$ the Newton envelope obtained after intermediate step (i,h), the step "Compute $\mathcal{E}^{(i+1)}$" in the above algorithm thus decomposes into

$$\mathcal{E}^{(i)} = \mathcal{E}^{(i,0)} \to \mathcal{E}^{(i,1)} \to \cdots \to \mathcal{E}^{(i,h)} \to \mathcal{E}^{(i,h+1)} \to \cdots \to \mathcal{E}^{(i+1)}. \tag{6a}$$

An example is shown in Fig. 3 for $i=1$. More precisely, $\mathcal{E}^{(1)} = \mathcal{E}^{(1,0)} \to \mathcal{E}^{(1,1)} \to \mathcal{E}^{(1,2)} = \mathcal{E}^{(2)}$, where the three stages correspond to the parametrized line of equation $x-2y = 3-h$ with $h \in \{0,1,2\}$.

From this corner cutting process, it follows that

$$\mathcal{E}^{(i,h)} = \mathcal{E}^{(i+1)} \quad \Longleftrightarrow \quad \mathcal{N} \cap (\text{segment of } \mathcal{E}^{(i,h)} \text{ with slope } \tfrac{1}{i+1}) \neq \emptyset, \tag{6b}$$

and, recalling that $\mathcal{N}$ has no zero slope when A_0 is nilpotent,

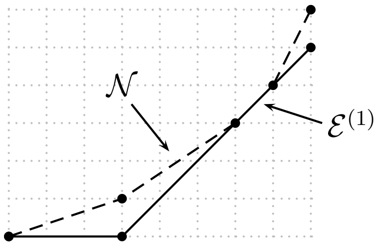
$$\mathcal{E}^{(i)} = \mathcal{E}_{max} \quad \Longleftrightarrow \quad \mathcal{E}^{(i)} \text{ has no segment with slope } 0. \tag{6c}$$

Finally, although the segment of $\mathcal{E}^{(i)}$ with slope $1/i$ may completely disappear when deducing $\mathcal{E}^{(i+1)}$, it should be noted that the segments of slope greater than $1/i$ are the same for $\mathcal{E}^{(i)}$ and $\mathcal{E}^{(i+1)}$.

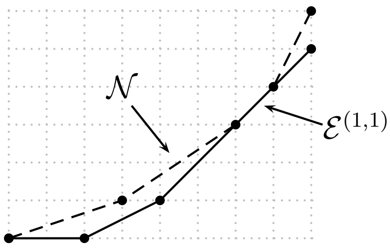
Recalling that the sum $\Delta_j^{(k)}$ was introduced in Definition 1, we end this section by providing the matrix counterparts of (6b) and (6c).

Corollary 12 *Let $A(\epsilon) \in \Pi_n$ with A_0 nilpotent and let $i \geq 0$, $h \geq 1$ be given.*

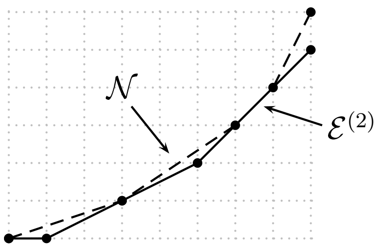
i) Assume that $\mathcal{E}(A_0) = \mathcal{E}^{(i,h)}$. One has $\mathcal{E}(A_0) = \mathcal{E}^{(i+1)}$ if and only if there



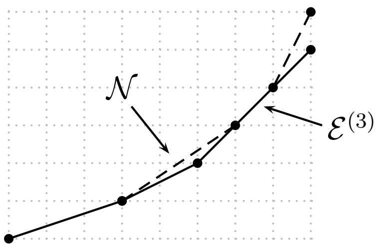
$$A_0^{(1)} = \left[\begin{array}{ccc|ccccc}
\times & \times & \times & & & & & \\
\times & \times & \times & & & & & \\
\times & \times & \times & & & & & \\
\hline
\times & \times & \times & 0 & & & & \\
\times & \times & \times & & 0 & & & \\
\times & \times & \times & & & 0 & & \\
\times & \times & \times & & & & 0 & \\
\times & \times & \times & & & & & 0
\end{array}\right]$$



$$A_0^{(1,1)} = \left[\begin{array}{cc|cc|cccc}
\times & \times & & & & & & \\
\times & \times & & & & & & \\
\hline
\times & \times & 0 & 1 & & & & \\
\times & \times & & 0 & & & & \\
\hline
\times & \times & & & 0 & & & \\
\times & \times & & & & 0 & & \\
\times & \times & & & & & 0 & \\
\times & \times & & & & & & 0
\end{array}\right]$$



$$A_0^{(2)} = \left[\begin{array}{c|cc|cc|ccc}
\times & & & & & & & \\
\hline
\times & 0 & 1 & & & & & \\
\times & & 0 & & & & & \\
\hline
\times & & & 0 & 1 & & & \\
\times & & & & 0 & & & \\
\hline
\times & & & & & 0 & & \\
\times & & & & & & 0 & \\
\times & & & & & & & 0
\end{array}\right]$$



$$A_0^{(3)} = \left[\begin{array}{ccc|cc|ccc}
0 & 1 & & & & & & \\
& 0 & 1 & & & & & \\
& & 0 & & & & & \\
\hline
& & & 0 & 1 & & & \\
& & & & 0 & & & \\
\hline
& & & & & 0 & & \\
& & & & & & 0 & \\
& & & & & & & 0
\end{array}\right]$$

Fig. 3. Successive envelopes and corresponding leading matrix structures for the Newton diagram of (3).

exists $k \in \{1,\dots,r_1\}$ such that $\Delta_1^{(k)} \neq 0$.

ii) Assume that $\mathcal{E}(A_0)=\mathcal{E}^{(i)}$. One has $A_0 = J_{min}$ if and only if $r_0 = 0$.

The matrix counterpart of $\mathcal{E}^{(i,h)} \to \mathcal{E}^{(i,h+1)}$ in (6a) is less immediate and will be given in Theorem 13. This result relies on the elementary operations we propose in the next section.

4.2 Elementary operations for matrix perturbations

We define some constant similarities that can be applied to $A(\epsilon)$ in order to modify its Lidskii submatrices in various ways without changing the structure (1) of its leading matrix. Such similarities P thus have the double property that A_0 and $P^{-1}A_0P$ have the same structure (1) and $P^{-1}A_1P$ is "simpler" than A_1 in the sense of the three paragraphs below.

• **Eliminate on a row of a Lidskii submatrix ("elim").** For $j \in \{1,\dots,q\}$, consider the Lidskii submatrix L_j and two entries $L_j[i,k] \neq 0$, $L_j[i,l]$ such that $k < l$. Our goal is to show that there exists a constant similarity P that reduces $A(\epsilon)$ to $\tilde{A}(\epsilon) = P^{-1}A(\epsilon)P$ so that $\tilde{A}_0$ and A_0 have the same structure and $\tilde{L}_j[i,l]=0$. It is sufficient for our purpose to assume that k corresponds to a Jordan block of size greater than 1.

Let $\alpha = L_j[i,l]/L_j[i,k]$ and apply to $A(\epsilon)$ the similarity operation defined by

$$\mathrm{col}_{a(l)} \leftarrow \mathrm{col}_{a(l)} - \alpha\, \mathrm{col}_{a(k)}, \tag{7a}$$
$$\mathrm{row}_{a(k)} \leftarrow \mathrm{row}_{a(k)} + \alpha\, \mathrm{row}_{a(l)}, \tag{7b}$$

where the ith row or column of L_j has index $a(i)$ in $A(\epsilon)$. Let us now carefully examine the effect of (7) on both L_j and the structure of A_0. First, (7a) zeroes $L_j[i,l]$ without modifying the structure (1) of A_0. In particular, if $k > r_0$, the $a(k)$th column of A_0 is identically zero. On the other hand, (7b) does not perturb the zeroed entry $L_j[i,l]$, for $a(k) \neq a(i)$ because of the assumption on the block size. However, (7b) does modify the structure of A_0 by adding α to its $(a(k), a(l)+1)$ entry (which is initially zero), i.e. to the $(1,2)$ entry of an off-diagonal block of Z. This "perturbation" of A_0 can be removed by constant similarity as follows: applying to $A(\epsilon)$ the similarity transform given by

$$\mathrm{col}_{a(l)+1} \leftarrow \mathrm{col}_{a(l)+1} - \alpha\, \mathrm{col}_{a(k)+1}, \tag{8a}$$
$$\mathrm{row}_{a(k)+1} \leftarrow \mathrm{row}_{a(k)+1} + \alpha\, \mathrm{row}_{a(l)+1}, \tag{8b}$$

shifts α from the $(a(k), a(l)+1)$ entry of A_0 to its $(a(k)+1, a(l)+2)$ entry.

Furthermore, this does not modify the current L_j, for its entries stem from rows and columns that differ from those involved in (8). Since $k < l$ and since the last row of a nilpotent Jordan block is identically zero, it thus suffices to iterate with (8) until the first bottom row is reached. This allows to recover the initial structure of A_0.

• **Zero the last row of a Lidskii submatrix ("zero").** For a singular Lidskii submatrix L_j, let us show that there exists a constant similarity P so that $\tilde{A}(\epsilon) = P^{-1}A(\epsilon)P$ satisfies: $\tilde{A}_0$ and A_0 have the same structure and the last row of $\tilde{L}_j$ is identically zero. Since there exists some α_k's so that $L_j[s_j, *] = \sum_{k=1}^{s_j-1} \alpha_k L_j[k, *]$, we apply to $A(\epsilon)$ the similarity

$$\mathrm{row}_{a(i)} \leftarrow \mathrm{row}_{a(i)} - \alpha_k \, \mathrm{row}_{a(k)}, \tag{9a}$$
$$\mathrm{col}_{a(k)} \leftarrow \mathrm{col}_{a(k)} + \alpha_k \, \mathrm{col}_{a(i)}, \tag{9b}$$

for $1 \le k \le s_j$ and $k \ne i$. With arguments similar to those used for "elim", we obtain the desired result providing operations analogous to (8) are applied to $A(\epsilon)$ after (9). Note that these techniques also allow to zero the last row of any submatrix of L_j which has not full row rank.

• **Swap two entries of a Lidskii submatrix ("swap").** This can be done by swapping two diagonal entries or Jordan blocks of A_0 by means of constant similarities applied to $A(\epsilon)$. This follows from the classic fact that for arbitrary square matrices A, B, C, D, E one has $P^{-1}\mathrm{diag}[A, B, C, D, E]P = \mathrm{diag}[A, D, C, B, E]$ where

$$P = \begin{bmatrix} I_a & & & & \\ & & & I_b & \\ & & I_c & & \\ & I_d & & & \\ & & & & I_e \end{bmatrix} \quad \text{and} \quad P^{-1} = \begin{bmatrix} I_a & & & & \\ & & & I_d & \\ & & I_c & & \\ & I_b & & & \\ & & & & I_e \end{bmatrix},$$

a denoting the dimension of the square matrix A etc.

4.3 Polynomial reduction via elementary operations

The goal here is to establish that any matrix perturbation can be reduced to a matrix perturbation with minimal leading Jordan structure by means of a *polynomial* similarity transform. Such a transform is obtained after a finite number of elementary operations "elim", "zero" and "swap". We start with the matrix counterpart of $\mathcal{E}^{(i,h)} \to \mathcal{E}^{(i,h+1)}$ in (6a).

Theorem 13 *For $A(\epsilon) \in \Pi_n$, assume that $\mathcal{E}(A_0) = \mathcal{E}^{(i,h)}$ for some (i,h). If $\mathcal{E}^{(i,h)} \neq \mathcal{E}^{(i+1)}$ then one can construct a nonsingular polynomial matrix $P(\epsilon)$ of degree at most 1 so that $\tilde{A}(\epsilon) = P(\epsilon)^{-1}A(\epsilon)P(\epsilon)$ satisfies $\mathcal{E}(\tilde{A}_0) = \mathcal{E}^{(i,h+1)}$.*

PROOF. Note first that A_0 is given by (1) with $(r_1, n_1) = (h, i+1)$ and $n_2 = i$. In order for this proof to handle the case where $h = 0$ too, we allow r_1 to be zero in (1). Let us rewrite the assumptions in terms of the regularity/singularity of Lidskii submatrices L_0, L_1, L_2. From Corollary 4, these assumptions imply (10) hereafter:

$$\Delta_1^{(k)} = 0 \quad \text{for } 0 \leq k \leq r_1. \tag{10a}$$

$$\text{There exists } k \in \{1, \ldots, r_2\} \text{ such that } \Delta_2^{(k)} \neq 0. \tag{10b}$$

The proof can be decomposed into three main steps.

1. Let us show that fully taking advantage of (10) allows to simplify A_1 without modifying the structure (1) of A_0. We will see that such a reduction relies exclusively on the elementary operations introduced in the previous section. Assume that the first r_0 rows of L_1 have full rank. (The simpler case where they are linearly dependent can be treated using similar arguments.) Since $\Delta_1^{(r_1)} = \det L_1 = 0$, one can zero the last row of L_1 using the "zero" operation. It follows from (10b) and the definition of the $\Delta_j^{(k)}$'s that at least one of the last r_2 entries of the (r_0+r_1)th row of L_2 is nonzero. Using "swap", we can ensure in particular that the (r_0+r_1, r_0+r_1+1)th entry of L_q (i.e. of the largest Lidskii submatrix, which contains all the Lidskii submatrices) is nonzero. We write $y = L_q[r_0+r_1, r_0+r_1+1]$ to denote this entry. For $0 \leq j \leq r_1$, let $L_1^{(j)}$ be the jth principal submatrix of L_1 that contains L_0 in its upper left corner. It follows from (10b) that condition $\Delta_1^{(r_1-1)} = 0$ now reads $\det L_1^{(r_1-1)} = 0$, i.e. the sum $\Delta_1^{(r_1-1)}$ reduces to a single term. If the first r_0 rows of $L_1^{(r_1-1)}$ are linearly dependent, set $r = 1$; otherwise, zero the last row of $L_1^{(r_1-1)}$ as done before with $L_1 = L_1^{(r_1)}$. It follows that $\Delta_1^{(r_1-2)} = \det L_1^{(r_1-2)} = 0$. We repeat the process until we find $1 \leq r \leq r_1$ such that the first r_0 rows of $L_1^{(r_1-r)}$ are linearly dependent. Such a minimal r will always be found, for L_0 is singular according to (10a). Let us now zero the r_0th row of $L_1^{(r_1-r)}$ by using the "zero" elementary operation. On the other hand, it follows from r being minimal that at least one of the last r entries of the r_0th row of L_1 is nonzero; one can zero $r-1$ among these r entries using "elim" and "swap" with the first nonzero entry. We therefore ensure that the (r_0, r_0+r_1) entry of L_q is nonzero and we write $x = L_q[r_0, r_0+r_1]$. Note that y is not modified by

these operations. To summarize, the elementary operations used so far define a constant similarity transform S that reduces $A(\epsilon)$ to $\tilde{A}^{(1)}(\epsilon) = S^{-1}A(\epsilon)S$ with the following properties: first, $\tilde{A}_0^{(1)}$ has the same structure as A_0; second, the r_0th and $(r_0 + r_1)$th rows of $\tilde{L}_q^{(1)}$ (i.e. of the largest Lidskii submatrix for $\tilde{A}^{(1)}(\epsilon)$) are zero everywhere except for the (r_0, r_0+r_1) and (r_0+r_1, r_0+r_1+1) entries, respectively equal to x and y. (When $r_1 = 0$, (x, y) should be replaced with y. When $q = 1$, we see that "elim" is not used; hence the assumption on the block size when defining "elim" in the previous section.)

2. As a second step, define the nonsingular diagonal matrix

$$D(\epsilon) = \operatorname{diag}[I_{r_0-1}, \epsilon \,|\, I_{r_1 n_1 - 1}, \epsilon \,|\, I_{n-r_0-r_1 n_1}].$$

This matrix transforms $\tilde{A}^{(1)}(\epsilon)$ into $\tilde{A}^{(2)}(\epsilon) = D(\epsilon)^{-1}\tilde{A}^{(1)}(\epsilon)D(\epsilon)$ with leading matrix having the following structure:

$$\tilde{A}_0^{(2)} = \begin{bmatrix} X^{(2)} & O \\ Y^{(2)} & Z^{(2)} \end{bmatrix}, \quad Z^{(2)} = J + B,$$

where $X^{(2)}$ and $Y^{(2)}$ are a priori dense matrices with respective dimensions $(r_0 - 1) \times (r_0 - 1)$ and $(n - r_0 + 1) \times (r_0 - 1)$. Additionally, J is the nilpotent Jordan structure

$$J = \operatorname{diag}[0, (0^{n_1})^{r_1-1}, 0^{n_2}, 0, (0^{n_2})^{r_2} \cdots (0^{n_q})^{r_q}] \quad \text{with } n_2 = n_1 - 1,$$

and B is the $(n-r_0+1)\times(n-r_0+1)$ matrix whose only nonzero rows are rows 1 and $r_1 n_1+1$. In particular, $B[1, (r_1-1)n_1+2] = x$ and $B[r_1 n_1+1, r_1 n_1+2] = y$, and all the other entries that also belong to a Lidskii submatrix are zero. This is because of step 1.

3. The third and last step of the proof is to show that the complex matrix $Z^{(2)}$ is similar to the nilpotent Jordan structure $(0^{n_1})^{r_1+1}(0^{n_2})^{r_2-1}(0^{n_3})^{r_3}\cdots(0^{n_q})^{r_q}$. The nonzero entries of B other than x and y can easily be zeroed by using the 1's of the Jordan blocks of J: applying to $\tilde{A}^{(2)}(\epsilon)$ the similarity transform $(\text{row}_{r_0} \leftarrow \text{row}_{r_0} - \alpha\, \text{row}_j, \text{col}_j \leftarrow \text{col}_j + \alpha\, \text{col}_{r_0})$ sets $B[1, j+1] = \alpha$ to zero, for the r_0th column of $\tilde{A}_0^{(2)}$ is identically zero. The same applies to the other nonzero row of B, noting that the $(r_0 + r_1 n_1)$th column of $\tilde{A}_0^{(2)}$ is identically zero. Thus, the only nonzero entries of the new B are $B[1, (r_1 - 1)n_1 + 2] = x$ and $B[r_1 n_1 + 1, r_1 n_1 + 2] = y$. Since $n_2 = n_1 - 1$, we see that $y \neq 0$ and the 0^{n_2} block immediately below merge into a new 0^{n_1} block. On the other hand, if $M = \operatorname{diag}[0, (0^{n_1})^{r_1-1}, 0^{n_1-1}] + C$ with C zero everywhere except for $C[1, (r_1-1)n_1+2] = x \neq 0$, then one can easily find a constant similarity that

transforms M into $(0^{n_1})^{r_1}$. In summary, we have shown in step 3 that there exists a constant similarity T such that $\tilde{A}^{(3)}(\epsilon) = T^{-1}\tilde{A}^{(2)}(\epsilon)T$ satisfies

$$A_0^{(3)} = \begin{bmatrix} X^{(3)} & O \\ Y^{(3)} & Z^{(3)} \end{bmatrix}, \quad Z^{(3)} = (0^{n_1})^{r_1+1}(0^{n_2})^{r_2-1}(0^{n_3})^{r_3}\cdots(0^{n_q})^{r_q},$$

where $X^{(3)}$ and $Y^{(3)}$ are a priori dense matrices with respective dimensions $(r_0-1)\times(r_0-1)$ and $(n-r_0+1)\times(r_0-1)$. The proof follows from writing $P(\epsilon) = SD(\epsilon)T$ and recalling that $(r_1, n_1) = (h, i+1)$ and $n_2 = i$.

An illustration of successive applications of Theorem 13 is given in Fig. 3. Here, the Newton diagram is the one of our matrix example (3), but it has been assumed that the initial leading matrix A_0 is dense and that its Jordan form is not known from the outset. Nevertheless, we eventually obtain $\tilde{A}_0 = A_0^{(3)} = J_{min}$.

The result below follows from Section 4.1, Corollary 12 and Theorem 13.

Corollary 14 *For all $A(\epsilon) \in \Pi_n$ one can construct a nonsingular polynomial matrix $P(\epsilon)$ over $\mathbb{C}$ of degree at most n, such that $P(\epsilon)^{-1}A(\epsilon)P(\epsilon)$ is a matrix perturbation with minimal leading Jordan structure.*

References

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

[2] M. Barkatou and E. Pflügel, *An algorithm computing the regular singular formal solutions of a linear differential system*, J. Symbolic Computation, 28(4-5) (1999), pp. 568–587.

[3] H. Baumgärtel, *Analytic perturbation theory for matrices and operators*, Birkhäuser Verlag, Basel, 1985.

[4] V. Bogaevski and A. Povzner, *Algebraic methods in nonlinear perturbation theory*, Springer-Verlag, New York, 1991.

[5] E. Brieskorn and H. Knörrer, *Plane algebraic curves*, Birkhäuser Verlag, Basel, 1986.

[6] J. Burke and M. Overton, *Stable perturbations of nonsymmetric matrices*, Linear Algebra Appl., 171 (1992), pp. 249–273.

[7] D. Duval, *Rational Puiseux expansions*, Compositio Mathematica, 70 (1989), pp. 119–154.

[8] A. Edelman, E. Elmroth, and B. Kågström, *A geometric approach to perturbation theory of matrices and matrix pencils: Part II: Stratifications*, SIAM J. Matrix Anal. Appl., 20(3) (1999), pp. 667–699.

[9] R. Gérard and A. Levelt, *Invariants mesurant l'irrégularité en un point singulier des systèmes différentiels linéaires*, Ann. Inst. Fourier, 23 (1973), pp. 157–195.

[10] M. Giesbrecht, *Nearly optimal algorithms for canonical matrix forms*, SIAM J. Computing, 24 (1995), pp. 948–969.

[11] A. Hilali and A. Wazner, *Calcul des invariants de Malgrange et de Gérard et Levelt d'un système différentiel linéaire en un point singulier irrégulier*, Jour. Diff. Equ., 69 (1987).

[12] C.-P. Jeannerod, *An algorithm for the eigenvalue perturbation problem: reduction of a κ-matrix to a Lidskii matrix*, in Proceedings of ISSAC 2000, C. Traverso, ed., St Andrews, Scotland, 2000, ACM Press, pp. 184–191.

[13] C.-P. Jeannerod and E. Pflügel, *A reduction algorithm for matrices depending on a parameter*, in Proceedings of ISSAC 1999, S. Dooley, ed., Vancouver, Canada, 1999, ACM Press, pp. 121–128.

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

[15] J. Klosowski, M. Held, J. Mitchell, H. Sowizral, and K. Zikan, *Efficient detection collision using bounding volume hierarchies of k-DOPs*. Full paper submitted for journal publication - available at `http://www.ams.sunysb.edu/ jsbm/publications.html`, 1996.

[16] S. Krein and A. Tovbis, *Linear singular differential equations in finite-dimensional spaces and Banach spaces*, Leningrad Math. J., 2(5) (1991), pp. 931–985.

[17] A. Levelt, *Stabilizing differential operators: a method for computing invariants at irregular singularities*, in Differential equations and computer algebra, M. Singer, ed., Academic Press, 1991, pp. 181–228.

[18] V. Lidskii, *Perturbation theory of non-conjugate operators*, U.S.S.R. Comput. Math. and Math. Phys., 6(1) (1966), pp. 73–85. (Zh. Vychisl. Mat. i Mat. Fiz., volume 6 (1966), number 1 (January-February), pp. 52–60).

[19] Y. Ma and A. Edelman, *Nongeneric perturbations of Jordan blocks*, Linear Algebra Appl., 273 (1998), pp. 45–63.

[20] J. Moro, J. Burke, and M. Overton, *On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 793–817.

[21] J. Moser, *The order of a singularity in Fuchs' theory*, Math. Zeitschr., 72 (1960), pp. 379–398.

[22] K. Murota, *Computing Puiseux-series solutions to determinantal equations via combinatorial relaxation*, SIAM J. Comput., 19(6) (1990), pp. 1132–1161.

[23] E. Pflügel, *Résolution symbolique des systèmes différentiels linéaires*, PhD thesis, Université Joseph Fourier, Grenoble, France, 1998.

[24] ——, *Effective formal reduction of linear differential systems*, Appl. Alg. in Eng., Com. and Comp., 2 (2000).

[25] R. Stanley, *Enumerative Combinatorics*, Cambridge University Press, 1997.

[26] M. Vainberg and V. A. Trenogin, *Theory of branching of solutions of non-linear equations*, Noordhoff, Leyden, 1974.

[27] M. Vishik and L. 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.

[28] R. Walker, *Algebraic curves*, Dover, 1962.

[29] J. Wilkinson, *The algebraic eigenvalue problem*, Oxford, 1965.