A reduced form for perturbed matrix polynomials
Claude-Pierre Jeannerod

To cite this version:

HAL Id: hal-03420227
https://inria.hal.science/hal-03420227
Submitted on 9 Nov 2021
A reduced form for perturbed matrix polynomials

Claude-Pierre Jeannerod
INRIA - Laboratoire LIP - Ecole Normale Supérieure de Lyon
46 Allée d’Italie, 69364 Lyon cedex 07, France
Claude-Pierre.Jeannerod@ens-lyon.fr

ABSTRACT
We show that every perturbation $A(\lambda, \epsilon)$ of an $n \times n$ matrix polynomial $A(\lambda)$ such that det $A(\lambda) = \lambda^m$ with $m \leq n$ can be reduced by equivalence transforms to a perturbed matrix polynomial whose leading matrix has maximal Smith form. This yields a reduced form for square perturbed matrix polynomials from which one can easily recover all the eigenvalue leading terms of the form $\mu \epsilon^d$ with $\beta^{-1} \in \mathbb{N}^*$.

Keywords
Matrix polynomials, perturbed eigenvalues, Smith normal form, Newton diagram.

1. INTRODUCTION
Let $A(\lambda) = \sum_{i=0}^{n} A_i \lambda^i$ be an $n \times n$ matrix polynomial over $\mathbb{Q}$ with det $A(\lambda) \neq 0$. Following [20, 11, 13, 12, 22, 23] we consider the problem of the sensitivity of the eigenvalues of $A(\lambda)$ (i.e. the roots of det $A(\lambda)$) to small perturbations of the matrix coefficients $A_i$. Such perturbations can be represented by perturbed matrix polynomials $A(\lambda, \epsilon) = \sum_{i=0}^{d} A_i(\epsilon) \lambda^i$ where $A_i(\epsilon) = \sum_{j=0}^{d} A_{ij} \epsilon^j$ is an $n \times n$ formal power series matrix over $\mathbb{Q}$ such that $A_i(0) = A_i$ for $0 \leq i \leq d$. Regarding $A(\lambda, \epsilon)$ as a perturbation of its leading matrix $A(\lambda) = A(\lambda, 0)$, we may also write

$$A(\lambda, \epsilon) = A(\lambda) + \epsilon(A_{01} + \sum_{i=1}^{d} A_{i1} \lambda^i) + O(\epsilon^2). \quad (1)$$

The sensitivity of the eigenvalues of $A(\lambda)$ is described by the leading terms $\mu \epsilon^d$ of the formal Puiseux expansions $\lambda(\epsilon) = \mu \epsilon^d + \cdots$ at $\epsilon = 0$ of the solutions of det $A(\lambda, \epsilon) = 0$. (See e.g. [2].) Such a $\lambda(\epsilon)$ will be referred to as an $\epsilon^d$-eigenvalue of $A(\lambda, \epsilon)$.

Assuming that det $A(\lambda) = \lambda^m$, we focus on how the zero eigenvalue of $A(\lambda)$ “splits” into $m$ perturbed eigenvalues of $A(\lambda, \epsilon)$: when $A(\lambda)$ is in Smith form, some sufficient conditions on some entries of $A_{01}$ were given in [11] for $A(\lambda, \epsilon)$ to have $r_1 \epsilon^{1/r_1}$-eigenvalues when $A(\lambda)$ has $r$ invariant polynomials of order $r$; it is also known how to recover the leading coefficients $\mu$ from these particular entries. However, these conditions are not necessary. Consider for example the perturbed matrix polynomials

$$A(\lambda, \epsilon) = \begin{bmatrix} 1 & \epsilon \\ \epsilon & \lambda^2 \end{bmatrix} \quad \text{and} \quad \hat{A}(\lambda, \epsilon) = \begin{bmatrix} \lambda & \epsilon \\ \epsilon & \lambda \end{bmatrix}.$$  

Although the eigenvalues are in both cases $\epsilon$ and $-\epsilon$, the structure $A(\lambda) = \text{diag}[\lambda, \lambda^2]$ does not indicate a priori a splitting into two $\epsilon$-eigenvalues and, in that sense, is misleading. On the other hand, one may easily verify that $\hat{A}(\lambda, \epsilon) = U(\lambda, \epsilon) A(\lambda, \epsilon) V(\lambda, \epsilon)$ with

$$U(\lambda, \epsilon) = \begin{bmatrix} 1 & 0 \\ \frac{\lambda \epsilon}{\lambda^2 - 1} & 1 \end{bmatrix} \quad \text{and} \quad V(\lambda, \epsilon) = \begin{bmatrix} 1 & 0 \\ \frac{\lambda^2}{\lambda - 1} & 1 \end{bmatrix}.$$  

We thus have found a matrix $\hat{A}(\lambda, \epsilon)$ equivalent to $A(\lambda, \epsilon)$ whose leading Smith form $\hat{A}(\lambda) = \text{diag}[\lambda, \lambda]$ is non misleading: we can tell from $\hat{A}(\lambda)$ that $\hat{A}(\lambda, \epsilon)$ admits two $\epsilon$-eigenvalues.

The two main contributions of this paper are as follows. First, we establish for $m \leq n$ that “non misleading” actually means that the leading Smith form is maximal for lexicographic ordering. Second, we prove that the observation made on the above example holds more generally: every perturbation of an $n \times n$ matrix polynomial $A(\lambda)$ such that det $A(\lambda) = \lambda^m$ with $m \leq n$ can be reduced by equivalence transforms to a perturbed matrix polynomial whose leading matrix has maximal Smith form. This yields a reduced form for square perturbed matrix polynomials from which one can easily recover all the eigenvalue leading terms of the form $\mu \epsilon^d$ with $\beta^{-1} \in \mathbb{N}^*$.

These results represent a first step towards the effective formal reduction of matrix polynomials depending on a parameter. There are two main motivations. First, many problems in robust control [3, 1] can be described in terms of parameterized matrix polynomials. On the other hand, using matrix polynomials allows to extend some results obtained for the perturbation $A(\lambda, \epsilon) = \lambda I - J - B(\epsilon)$ of a nilpotent Jordan structure $J$. (See e.g. [14, 16, 15, 10].)

Previous studies about parameterized matrix polynomials were concerned by either determinant degree and Smith

*This work was done while the author was a postdoctoral fellow at the Symbolic Computation Group, University of Waterloo.
In this definition, matrix, of determinant $\Delta$ the sums $\Delta$ of order $L$.

In order to investigate the link between the structure (2) and define introduce the partial sums $b_j \times A$ with $0 < m < n$.

Throughout the paper, $m > n$.

We conclude in Section 5 with two extensions.

Notation. Throughout the paper, $p$ denotes the degree in $\lambda$ of $p(\lambda, \epsilon) = \det A(\lambda, \epsilon)$ and $\Pi$ denotes the matrix set $\Pi = \{ A(\lambda, \epsilon) \in \mathbb{Q}[x][\lambda]^{m \times n} : \det A(\lambda) = \lambda^m \}$ where $A(\lambda) = A(\lambda, 0)$.

2. Perturbation of a Smith Form and Perturbed Eigenvalues

Let $A(\lambda, \epsilon) \in \Pi$ and assume that its leading matrix $A(\lambda)$ is equal to the Smith form

$$S(\lambda) = \text{diag}[1, \ldots, 1, (\lambda^{m_1})^s \cdots (\lambda^{m_q})^s]$$

with $0 < m_1 \leq \cdots \leq m_q$. Here we wrote $(\lambda^n)^b$ for the $b \times b$ diagonal matrix whose all diagonal entries are equal to the monomial $\lambda^n$.

One can think of $m_j$ as the $j$th partial multiplicity of the zero eigenvalue of $A(\lambda)$ repeated $r_j$ times, and note that $r_1 m_1 + \cdots + r_q m_q = m$ and $r_1 + \cdots + r_q \leq n$.

2.1 A perturbation theorem for eigenvalues

In order to investigate the link between the structure (2) of $A(\lambda)$ and the leading terms of the eigenvalues of $A(\lambda, \epsilon)$, introduce the partial sums $s_j = r_j + \cdots + r_q$ for $1 \leq j \leq q$.

and define $L_j$ as the last principal submatrix of $A_{01}$ in (1) of order $s_j$. For $1 \leq j < q$, one has $s_{j+1} = s_j - 1$ and $L_{j+1}$ lies in the lower right corner of $L_j$. The values of the determinants of such submatrices of $A_{01}$ — together with the additional “intermediate” quantities recalled below — were already considered by Langer and Najman [11, 12].

Definition 1. For $1 \leq j \leq q$, let $\Delta_j$ be the determinant of $L_j$. For $1 \leq j < q$ and $0 \leq k \leq r_j$, we further denote by $\Delta_j^{(k)}$ the sum of the principal minors of $L_j$ with order $s_j - k$ that contain $L_{j+1}$.

In this definition, $L_{q+1}$ has been assumed to be the empty matrix, of determinant $\Delta_{q+1} = 1$. Note that $\Delta_j = \Delta_j^{(0)}$ and that $\Delta_{j+1} = \Delta_j^{(r_j)}$ for $1 \leq j \leq q$.

The theorem below provides some sufficient conditions on the sums $\Delta_j^{(0)}$ for the multiple zero eigenvalue of $A(\lambda)$ to split under perturbation according to the number and the order of its partial multiplicities. A proof based on the Newton diagram will be given in Section 2.2.

Theorem 1. Let $j \in \{1, \ldots, q\}$. Assuming that at least one of the sums $\Delta_j^{(k)}$ for $0 \leq k \leq r_j$ 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) $A(\lambda, \epsilon)$ admits $(k_2 - k_1) m_j \epsilon^{1/m_j}$-eigenvalues;

ii) writing $w_j = e^{2i\pi/m_j}$, the leading terms of these eigenvalues are $(\mu_{jk})^{1/m_j} w_i \epsilon^{1/m_j}$, $k = 1, \ldots, k_2 - k_1$, $l = 1, \ldots, m_j$, where the complex numbers $\mu_{jk}$ denote the roots of the polynomial $\sum_{k=k_1}^{k_2} \Delta_j^{(k)} \lambda^{k-k_1}$ and where $(\mu_{jk})^{1/m_j}$ is one of the $m_j$ distinct $m_j$-th roots of $\mu_{jk}$.

In the particular case where $(k_1, k_2) = (0, r_j)$, this result was obtained by Langer and Najman in [11], describing the typical eigenvalue splitting when $A_{01}$ is dense. Note also that in this case the $\mu_{jk}$’s are up to the sign the eigenvalues of the Schur complement of $L_{j+1}$ in $L_j$. This follows from $\sum_{k=k_1}^{k_2} \Delta_j^{(k)} \lambda^k = \det(L_j + A_{1j})$ with $E_j = \text{diag}[I_{r_j}, O_{n_{j+1}}]$, and from the determinantal equality (see e.g. [5, p. 190])

$$\det(L_j + \lambda E_j) := \det \begin{bmatrix} X + \lambda I & Y \\ Z & L_{j+1} \end{bmatrix} = \det L_{j+1} \det(X - Y L_{j+1}^1 Z + \lambda I).$$

However, though simpler to check, the conditions “$\Delta_j \neq 0$” will prove to be not enough for our purposes.

As an illustration of the above Theorem 1, consider the $6 \times 6$ perturbed matrix polynomial of the form

$$A(\lambda, \epsilon) = \text{diag}[1, 0, 1, (\lambda^3)^6 + \epsilon(A_{01} + A_{11} \lambda)]$$

with $A_{01}$ and $A_{11}$ given by

$$A_{01} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

and

$$A_{11} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Here $q = 2$, $(r_1, m_1) = (3, 1)$, $(r_2, m_2) = (1, 3)$ and

$$L_1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 0 \end{bmatrix}.$$

Hence $\Delta_1^{(0)} = 1, \Delta_1^{(1)} = 1, \Delta_1^{(2)} = 0, \Delta_1^{(3)} = 0$ and it follows from Theorem 1 (applied with $j = 1$) that $A(\lambda, \epsilon)$ has one eigenvalue of the form $\lambda(\epsilon) = -\epsilon + o(\epsilon)$. 
2.2 Newton diagram-based characterization

Following Moro, Burke and Overton [16], let us interpret Theorem 1 in terms of the Newton diagram of $A(\lambda, \epsilon)$. Let $p(\lambda, \epsilon) = \det A(\lambda, \epsilon)$ of degree $l$ in $\lambda$ be written as

$$p(\lambda, \epsilon) = \sum_{i=0}^{l} a_i(\epsilon) \lambda^i.$$ 

For $0 \leq i \leq l$, there exists $(a_i, \alpha_i) \in \mathbb{Q}^* \times \mathbb{N}$ 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 $\{(a, \alpha) : 0 \leq i \leq l\}$ in the cartesian plane is the Newton diagram $\mathcal{N}$ associated with $A(\lambda, \epsilon)$. (See e.g. [16],[4],[2],[24].) 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 rational, possibly infinite $(-\infty)$ slopes and nonzero lengths. In particular, the length of the segment with slope $-\infty$ can be considered as equal to the valuation of $p(\lambda, \epsilon)$ in $\lambda$. We will denote by $\mathcal{N}^+$ the negative part of $\mathcal{N}$, i.e. the subdiagram of $\mathcal{N}$ which consists of negative (finite and infinite) slopes only. Not that $\mathcal{N}$ and $\mathcal{N}^+$ have lengths $l$, $m$ respectively.

Let $(\rho, \nu, \delta) \in \mathbb{N} \times \mathbb{Z} \times \mathbb{N}^*$ be such that $\gcd(\nu, \delta) = 1$. The first main property of the Newton diagram is that $A(\lambda, \epsilon)$ has $\rho \delta^2 \epsilon^{\nu/\delta}$-eigenvalues if and only if its Newton diagram $\mathcal{N}$ has a segment $S$ of slope $-\nu/\delta$ and length $\rho \delta$. Now, letting $x_S$ be the minimum index $i$ so that $(i, \alpha_i)$ belongs to this particular segment $S$, we see that $\delta$ divides $i - x_S$ for all $(i, \alpha_i) \in S$. This allows to associate with each segment $S$ of $\mathcal{N}$ the polynomial

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

of degree $\rho$. (See e.g. [6, p. 137].) The second fundamental property of the Newton diagram is as follows: the leading terms of the $\rho \delta^2 \epsilon^{\nu/\delta}$-eigenvalues of $A(\lambda, \epsilon)$ are

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

where $w = e^{2\pi \i \nu/\delta}$ and where the $\mu_k$’s are the roots of $p_S(\lambda)$.

For example, $A(\lambda, \epsilon)$ in (3) has determinant

$$p(\lambda, \epsilon) = (-\epsilon^3 + 2\epsilon^4 + \epsilon^5 \lambda^6 + (3\epsilon^2 + 5\epsilon + 1) - 3\epsilon^3 + 4\epsilon^6 + (\epsilon^2 + 2\epsilon^4 + 3\epsilon^5) \lambda^5 + (7\epsilon^5 + 3\epsilon^4 + 2\epsilon^3 + \epsilon^2 + 4\epsilon^6 + 3\epsilon^2) \lambda^4 + (-4^2 + 2\epsilon^4 + 9\epsilon^6 + \epsilon^2 \lambda^3 + (2\epsilon^4 - 4\epsilon^5) \lambda^2 + (\epsilon^3 + \epsilon^4 + \epsilon^5) \lambda + \epsilon^6.$$ 

Hence the Newton diagram $\mathcal{N}$ shown in Fig. 2. It follows that $A(\lambda, \epsilon)$ has one eigenvalue, three $\epsilon^{1/2}$-eigenvalues, two $\epsilon^{1/2}$-eigenvalues and one $\epsilon^{-1}$-eigenvalue. The corresponding polynomials of type (4a) — built from the boxed terms of $p(\lambda, \epsilon)$ only — are $1 + 1, 1 + \lambda^l, 1 + \lambda^2, 1 + \lambda$ respectively.

Theorem 2 below expresses in terms of the characteristics $(r_1, r_2), \ldots, (r_{m_j}, m_j)$ of the Smith form $A(\lambda)$ what the rightmost possible vertices for the Newton diagram of $A(\lambda, \epsilon)$ are.

This will yield a graphical characterization of the sufficient conditions of Theorem 1 and also allow for a simple proof of this perturbation result. The structure of $A(\lambda)$ and an ordinate $y$ being given, consider all the perturbations of $A(\lambda)$ of the form (1) and denote by $x(y)$ the smallest possible abscissa for a vertex $(x(y), y)$ of the associated Newton diagrams. Additionally, recall that $s_j = r_j + \cdots + r_q$ and define

$$t_j = r_j m_j + \cdots + r_q m_q \quad \text{for } 1 \leq j \leq q.$$ 

**Theorem 2.** Let $j \in \{1, \ldots, q\}$ and $k \in \{0, \ldots, r_j\}$. If $y = s_j - k$ then $x(y) = m - t_j + km_j$ and the coefficient of $\epsilon^{\lambda(\nu)}$ in $p(\lambda, \epsilon)$ is $\Delta_j^{(k)}$.

**Proof.** By definition, $p(\lambda, \epsilon)$ is the sum of all the products (up to the sign) of $\rho$ entries of the matrix $A(\lambda, \epsilon) = A(\lambda) + \epsilon(A_{01} + O(\lambda)) + O(\epsilon^2)$ such that no two entries belong to the same column or row. The entries involved in a product can be $\epsilon$-terms from $A(\lambda, \epsilon) - A(\lambda)$ or $\lambda$-terms from $A(\lambda)$. Let us examine how to form a product of order $\epsilon^{\rho}$ in $\epsilon$ and of minimal order in $\lambda$. For $\epsilon$-terms, we pick $y \leq s_j \leq n$ entries in $\epsilon A_{01}$. It then remains to choose $n - y \lambda$-terms. In order to exhaust the largest powers of $\lambda$ from $A(\lambda)$, these $\epsilon$-terms further lie on the last $s_j + 1$ columns and on any group of $r_j - k$ columns with index between $n - s_j + 1$ and $n - s_j + 1$. The row indices for the $\epsilon$-terms must exactly vary in the same ranges, for otherwise the diagonal structure of $A(\lambda)$ leads to a product equal to zero. Finally, we complete the product by taking the first $n - s_j$ diagonal entries of $A(\lambda)$ together with a group of $k$ invariant polynomials of the form $\lambda^n$. This product thus has order $\lambda^{m - r_j + km_j}$ in $\lambda$, from which $(x(y), y)$ follows. Deleting from $A(\lambda, \epsilon)$ the rows and columns corresponding to each of the above choices of $\lambda$-terms yields the sum $\Delta_j^{(k)}$. \[\square\]

Now, for $1 \leq j \leq q$ and $0 \leq k \leq r_j$, let $p_j^{(k)}$ be the point given by $(x(y), y)$ as in Theorem 2 and write $p_j = p_j^{(0)}$ and $p_{j+1} = p_j^{(e_j)}$ for $1 \leq j \leq q$. In particular, one has $p_{q+1} = (m, 0)$. Using the terminology of Moro, Burke and Overton [16, p. 808], we call Newton envelope associated with the Smith form of $A(\lambda)$ the diagram $\mathcal{E}(A(\lambda))$ obtained by successively connecting the points $p_1, p_2, \ldots, p_{q+1}$. (See Fig. 2 for an example.) It then follows from Theorem 2 that for any perturbation $A(\lambda, \epsilon)$ of $A(\lambda)$ with given structure (2), the negative part $\mathcal{N}^+$ has no point lying below the envelope. Furthermore, the sufficient regularity conditions of Theorem 1 can be characterized graphically as shown below.

**Corollary 1.** Let $j \in \{1, \ldots, q\}$ and $k \in \{0, \ldots, r_j\}$. The sum $\Delta_j^{(k)}$ is nonzero if and only if the point $t_j^{(k)}$ of the envelope is also a point of the Newton diagram.

We end this section by providing a Newton diagram-based proof of Theorem 1.

**Proof.** [Proof of Theorem 1] It follows from Corollary 1 that the Newton diagram of $A(\lambda, \epsilon)$ has a segment $S_j$
We define the \( N \)-gram whose length is \( m \).

Now, let \( k \in \{k_1, \ldots, k_s\} \) be such that \( p_{\alpha}(x) = (x(y), y) \) lies on \( S_j \) — or equivalently, \( \Delta_j(\alpha) \neq 0 \). One has \( \alpha_\nu(y) = y \) and it follows from Theorem 2 that the coefficient of \( \rho_\nu(y) \) is \( \tilde{\alpha}_\nu(y) = \Delta_j(\nu) \). On the other hand, one may easily verify that \( x(y) - x_{\nu} = (k - k_1)m_j \).

Hence \( p_{\nu}(\lambda) = \sum_{k = k_1}^{k_2} \Delta_j(\nu) \lambda^{-k-1} \).

With the approach above, the Smith form \( A(\lambda) \) and thus \( \mathcal{E}(A(\lambda)) \) are given, and we identified among all the perturbations \( A(\lambda, \epsilon) \) compatible with \( A(\lambda) \) the ones that allow to recover as much of \( \mathcal{N} \) as possible. In the next section, the question we consider is the opposite: \( A(\lambda, \epsilon) \) and thus \( \mathcal{N} \) being given, what is the “best” possible Smith form for \( A(\lambda) \) compatible with \( \mathcal{N} \)?
Lemma 1. The sequence \((\sigma_i)_{i \in \mathbb{N}_0}\) given by (6) satisfies
\[
\sum_{i=\sigma_i}^{\infty} (i-j)\sigma_i = m - \nu_j \quad \text{for all } j \in \mathbb{N}.
\]

Proof. Let \(s < 0\) be the largest slope of \(N^-\) and let \(N = [-s^{-1}]\). For \(i \geq N\), the diagram \(N^-\) therefore has slopes in \(Q^-\) only. Since \((m,0)\) is a point of \(N^-\), it follows that \(\nu_i = m_0\) for all \(i \geq N\). Hence \(\sigma_i = 0\) for \(i > N\) and intermediate cancellations in the sum yield \(\sum_{i=\sigma_i}^{\infty} (i-j)\sigma_i = -\nu_j + (N-j+1)\nu_N - (N-j)\nu_{N+1}\). This is \(m - \nu_j\).

Theorem 3. Let \(\sigma_{N+1}, \ldots, \sigma_m\) denote the nonzero values of \((\sigma_i)_{i \in \mathbb{N}_0}\), numbered so that \(m_1 < \cdots < m_q\). If additionally one defines \(Q_0 = (0,n_0 + \sum_{k=1}^{q} m_k)\) and, for \(1 \leq j \leq q+1\), \(Q_j = (i_0 + \sum_{k=1}^{j-1} m_k, \sum_{k=0}^{q} m_k)\), then \(E_{max}\) is the diagram obtained by successively connecting the points \(Q_0, Q_1, \ldots, Q_{q+1}\).

Proof. \(E_{max}\) consists of \(p+1\) segments \(S_0, S_1, \ldots, S_p\), where \(S_0\) is the segment of slope \(-1\) corresponding to the slope \(-\infty\) of \(N^-\) and where, for \(1 \leq j \leq p\), \(S_j\) has slope \(-1/n_j\) and length \(r_jn_j\) with \(n_1 < \cdots < n_p\). For \(j \in \{1, \ldots, p\}\) we prove that \(\sigma_{n_j} = r_j\) as follows. Denote by \((x_1, y_1), (x_2, y_2)\) the cartesian coordinates of the extreme points of \(S_j\). When successively looking at the transformed envelopes \(E_i\) for \(i = 1, n_j, n_j+1\), we see that
\[
\begin{align*}
\nu_{n_j-1} & = x_1 + (n_j - 1)y_1; \\
\nu_{n_j} & = x_1 + n_jy_1; \quad \text{for } i = 1, 2; \\
\nu_{n_j+1} & = x_2 + (n_j + 1)y_2.
\end{align*}
\]
Writing \((x_2, y_2) = (x_1 + r_jn_j, y_1 + r_j)\) then yields \(\sigma_{n_j} = y_1 - y_2 = r_j\). On the other hand, it follows from the \(r_j\)'s being positive that the \(\sigma_i\)'s define \(p\) out of the \(q\) nonzero values of \((\sigma_i)_{i \in \mathbb{N}_0}\). For \(i = 0\), it follows from (6b) that the length of \(S_0\) is equal to \(n_0\). Hence \(m = i_0 + \sum_{i \in \{n_1, \ldots, n_p\}} i\sigma_i\). Since \(n_1 < \cdots < n_p\), Lemma 1 then implies that \(p = q\) and \(n_j = n_j\) for all \(j \in \{1, \ldots, q\}\). Hence the coordinates of the \(Q_j\)'s, noticing that \((m,0)\) is the rightmost point of \(E_{max}\).

For example, \(N^-\) in Fig. 2 has vertices \((0, 4), (1, 3), (4, 1), (6, 0)\) and it follows from (6) that \(\sigma_1 = 2, \sigma_2 = 2, \sigma_3 = \sigma_4 = \cdots = 0\). The corresponding highest envelope \(E_{max}\) is represented on the second graphic.

Figure 2: Newton diagram and Newton envelopes for the matrix \(A(\lambda, \epsilon)\) of (3).

3.2 Recovery of some eigenvalue leading terms

When \(m \leq n\), recovering some eigenvalue leading terms is made possible by the following relation between the highest envelope and the maximal leading Smith form. (We discuss the case where \(m > n\) in Section 4.2.)

Theorem 4. Let \(m \leq n\) and let \(A(\lambda, \epsilon) \in \Pi_m^n\) with Newton diagram \(N\). Then \(E_{max}\) is the Newton envelope associated with \(S(\lambda)_{max}\).

Proof. Let \(S(\lambda) \in \Sigma_m^n(N^-)\). If \(S(\lambda) < S(\lambda)_{max}\) then \(E(S(\lambda)) = \Pi(\lambda)\) has at least one vertex strictly above \(E(S(\lambda))\), and thus \(E(S(\lambda)) \neq E_{max}\).

In order to prove that the converse holds true, let us first construct \(T(\lambda, \epsilon) \in \Pi_m^n(N^-)\) such that \(T(\lambda)\) is in Smith form and \(E(T(\lambda)) = E_{max}\). Let \(T(\lambda, \epsilon) = \text{diag}(1, \ldots, 1, B(\lambda, \epsilon))\) where we choose for \(B(\lambda, \epsilon)\) the block diagonal matrix whose \(k\)th block corresponds to a segment (or a group of segments) of \(N^-\) as follows. To the infinite slope of \(N^-\) of length \(\rho\), we associate the block \(B(1, \lambda, \epsilon) = \lambda I_\rho\). To a segment of \(N^-\) with slope \(-1/N\) and length \(\rho\), we associate the \(\rho \times \rho\) matrix \(B(\lambda, \epsilon) = \text{diag}(\lambda^{\rho-1} \cdot i \in S\leq \epsilon\). With the same notation as before, consider now a group of \(s\) segments with slopes \(-1/N < \beta_1 < \cdots < \beta_s < -1/(N+1)\). For \(1 \leq j \leq s\), let \((x_j, y_j) \in N^* \times N^*\) be the solution of the linear system deduced from (5) by replacing the right hand side with \(\nu_j|v_j^T, \beta_j|^{T}\). Additionally, let \(C_{j}^h(\lambda, \epsilon) = (\lambda^{N^+}y_j^T(N^+)^y + C_{j}^h(\epsilon)\) with \(C_{j}^h(\epsilon)\) the \(\rho_j \nu_j \times \rho_j \nu_j\) matrix of the form
\[
\begin{bmatrix}
0 & \epsilon \\
\vdots & \ddots & \ddots \\
\epsilon & \ddots & \ddots & 0
\end{bmatrix}
\]

Clearly, \(C_{j}^h(\lambda, \epsilon) = \text{diag}(C_{j}^h(\lambda, \epsilon), \ldots, C_{j}^h(\lambda, \epsilon))\) has \(\rho_j \delta_j\) - eigenvalues for \(1 \leq j \leq s\). Since \(x = \sum_{j=1}^{s} x_j\) and \(y = \sum_{j=1}^{s} y_j\), we take for \(B(\lambda, \epsilon)\) the matrix deduced from \(C_{j}^h(\lambda, \epsilon)\) by permuting some rows and columns in such a way that the new leading matrix is \(B(\lambda) = (\lambda N^+)y_j^T(N^+)^y\). Note that the size of \(B(\lambda, \epsilon)\) is at most \(n\) by assumption on the length \(m\) of \(N^-\).

Now, assume that \(E(S(\lambda)) \neq E_{max}\). Considering the first vertex of \(E(S(\lambda))\) that lies strictly below \(E_{max}\) and since \(T(\lambda)\) is an element of \(\Sigma_m^n(N^-)\) whose envelope is \(E_{max}\), we obtain \(S(\lambda) < T(\lambda)\). Hence \(S(\lambda) < S(\lambda)_{max}\).

Theorem 4 has two consequences: first, it can be combined with Theorem 3 in order to yield the formula
\[
S(\lambda)_{max} = \text{diag}(1, \ldots, 1, (\lambda^{m_1})^{e_1} \cdots (\lambda^{m_q})^{e_q})
\]
when \(m \leq n\). For example, the matrix of (3) satisfies \(m = 6\) and we deduce from \(E_{max}\) in Fig. 2 that \(m_0 = 0, q = 2, (m_1, \sigma_{m_1}) = (1, 2), (m_2, \sigma_{m_2}) = (2, 2)\) and, therefore, that \(S(\lambda)_{max} = \text{diag}(1, \ldots, 1, (\lambda^2)^2, (\lambda^3)^3)\). Second, using Corollary 1, it follows that the sufficient conditions of Theorem 1 can be necessary too:

Corollary 2. (Converse of Theorem 1 i)) Let \(m \leq n\) and let \(A(\lambda, \epsilon) \in \Pi_m^n\) be a perturbed matrix polynomial with
maximal leading Smith form (2). Then, if \( A(\lambda, \epsilon) \) has vi \( \epsilon^{1/2} \)-eigenvalues, there exists \( j \in \{1, \ldots, q\} \) such that \( i = m_j \) and \( r = k_2 - k_1 \) with \( k_1 \) minimal and \( k_2 \) maximal so that \( \Delta^{(k_1)} \neq 0 \) and \( \Delta^{(k_2)} \neq 0 \).

This means that such perturbed matrix polynomials are precisely those for which all the eigenvalue leading terms of the form \( \mu^d \) with \( \beta^{-1} \in \mathbb{N}^* \) can be recovered via Theorem 1. For example, \( A(\lambda) \) in (3) is not maximal, for otherwise applying Theorem 1 to \( A(\lambda, \epsilon) \) would have revealed not only its \( \epsilon \)-eigenvalue but also its two \( \epsilon^{1/2} \)-eigenvalues. Nevertheless, as shown in the next section, \( A(\lambda, \epsilon) \) can be reduced to an equivalent perturbed matrix polynomial whose leading Smith form is maximal.

4. MAXIMIZATION BY EQUIVALENCE TRANSFORMS

4.1 Reduction theorem

Theorem 5. Let \( m \leq n \). Then for every \( A(\lambda, \epsilon) \in \mathbb{P}_m^n \), one can define a pair of equivalence transforms \( U(\lambda, \epsilon) \) and \( V(\lambda, \epsilon) \) such that \( U(\lambda, \epsilon)A(\lambda, \epsilon)V(\lambda, \epsilon) \) is a perturbed matrix polynomial with maximal leading Smith form.

Proof. Over \( \mathbb{F}[\lambda] \), the matrix polynomial \( A(\lambda, \epsilon) \) is equivalent to its Smith form \( \text{diag}(I_{n-m}, s_1(\lambda, \epsilon), \ldots, s_r(\lambda, \epsilon)) \) with \( s_k \) monic in \( \lambda \) and of positive degree for \( 1 \leq k \leq x \). Multiplying the last \( x \) rows by suitable constants of \( \mathbb{F} \) yields an equivalent matrix

\[
P(\lambda, \epsilon) = \text{diag}[I_{n-x}, p_1(\lambda, \epsilon), \ldots, p_x(\lambda, \epsilon)]
\]

with no negative power of \( \epsilon \) and such that

\[
p(\lambda, \epsilon) = \prod_{k=1}^{x} p_k(\lambda, \epsilon).
\]

Let \( m_k = \text{deg} p_k(\lambda, 0) \). Since \( m \leq n \), one can partition \( n \) as \( n_1 + \cdots + n_x \) so that \( m_k \leq m_k \) for \( 1 \leq k \leq x \). The matrix \( A(\lambda, \epsilon) \) is therefore equivalent to \( \text{diag}[P(\lambda, \epsilon)_{1, \leq x}] \) with \( P(\lambda, \epsilon) = \text{diag}[I_{n-s}, p(\lambda, \epsilon)] \).

Now let \( N_\lambda, E_{k, \text{max}} = E_{k, \text{max}}(N^-) \) and \( S_{k, \text{max}}(\lambda) \) be respectively the Newton diagram, the Newton envelope and the maximal leading Smith form associated with \( P(\lambda, \epsilon) \). It follows from the equations of (5) being linear that the equality \( E_{\text{max}}(N^-) = \cup E_{k, \text{max}}(N^-) \) holds. Here, the union denotes the convex diagram obtained by merging convex diagrams for \( 1 \leq k \leq x \). For example \( \cup N_k = N^- \). Since \( m_k \leq n_k \) for \( 1 \leq k \leq x \), we deduce \( S(\lambda)_{\text{max}} \) from \( \text{diag}[S_{k, \text{max}}(\lambda)]_{1, \leq x} \) by row and column permutations. The \( P_k \)'s can thus be handled independently and we may assume without loss of generality that \( A(\lambda, \epsilon) = \text{diag}[I_{n-x}, p(\lambda, \epsilon)] \).

Let us now reduce \( A(\lambda, \epsilon) \) to an equivalent \( \tilde{A}(\lambda, \epsilon) \) such that \( \tilde{A}(\lambda) \) is equal to the maximal Smith form of (7). Then \( r = n_0 + \sum_{j=1}^{q} \sigma_{m_j} \). It suffices to find a pair of \( \tau \times \tau \) unimodular matrices \( U(\lambda, \epsilon), V(\lambda, \epsilon) \) such that \( UBV = \widetilde{B} \) with \( B \) given by \( A = \text{diag}[I_{n-x}, B] \) and \( \widetilde{B} \) is the lower right block of \( A \) partitioned conformally. From now we focus on the case where \( n_0 = 0 \). (When \( n_0 > 0 \), it suffices to replace \( m_1 \) with \( n_0 \) in the argument below.) When \( r = 1 \), one has \( UV = V = 1 \); otherwise, writing \( p = r + s \) with \( r(\lambda, \epsilon) = \sum_{i \geq m_k} a_i(\epsilon)\lambda^i \) and \( s(\lambda, \epsilon) = \sum_{i \geq m_k} a_i(\epsilon)\lambda^i \), we notice that all the \( a_i(\epsilon)'s \) in the “regular” part \( r \) have valuation in \( \epsilon \) greater than \( r - 1 \) whereas the “singular” part \( s \) has at least one coefficient of valuation \( \tau - 1 \) or less. This allows to define

\[
U_1^{(0)} = \begin{bmatrix}
1 & \lambda^{m_1} & \epsilon^{r-1} & I_{r-2} & 1
\end{bmatrix}
\]

and

\[
V_1^{(0)} = \begin{bmatrix}
\lambda^{m_1} & \epsilon^{r-1} & I_{r-2} & -\epsilon^{r-1} & 0
\end{bmatrix}
\]

and to transform

\[
B = \begin{bmatrix}
1 & I_{r-2} & p(\lambda, \epsilon)
\end{bmatrix}
\]

into

\[
U_1^{(0)}BV_1^{(0)} = \begin{bmatrix}
\lambda^{m_1} & \epsilon^{r-1} & I_{r-2} & \lambda^{m_1} & s(\lambda, \epsilon)
\end{bmatrix}
\]

We have obtained the first diagonal entry of the target leading matrix \( B(\lambda) \). More generally, let \( \tau_j = \sigma_{m_j} + \cdots + \sigma_{m_q} \) for \( 1 \leq j \leq q \). One can then repeat the process and successively multiply on the left and on the right, respectively, \( \tau \times \tau \) unimodular matrices \( \text{diag}[I_{n-s}, U_j^{(k)}] \) and \( \text{diag}[I_{n-s}, V_j^{(k)}] \) where \( U_j^{(k)}, V_j^{(k)} \) are defined for \( 0 \leq k \leq \sigma_{m_j} \) and \( 1 \leq j \leq q \) on the same model as \( U_1^{(0)} \), \( V_1^{(0)} \). (In particular, \( U_j^{(0)} = U_j^{(\sigma_{m_j})} \) and \( V_j^{(0)} = V_j^{(\sigma_{m_j})} \).) It follows that defining \( U_{q+1} = U_{q}^{(\sigma_{m_q})} \cdots U_1^{(0)} \) and \( V = V_1^{(0)} \cdots V_q^{(\sigma_{m_q})} \) allows to transform \( B \) into \( \tilde{B} = UBV \). Note that the last diagonal entry of \( B(\lambda) \) stems from the term \( \lambda^{m_q} \) of \( p(\lambda, 0) \).

Theorem 5 thus defines a rational reduced form for equivalent square perturbed matrix polynomials over \( \mathbb{F}[\lambda] \). The invariants of \( A(\lambda, \epsilon) \) it gives access to are all the eigenvalue leading terms of \( A(\lambda, \epsilon) \) of the form \( \mu^d \) with \( \beta^{-1} \in \mathbb{N}^* \). Of course, this reduced form is not unique, for only its leading matrix is required to be normalized as \( S(\lambda)_{\text{max}} \).
4.2 Case where \( m > n \)
In this case the maximal leading Smith form can still be misleading. Consider for example \( A(\lambda, \epsilon) = \lambda^2 - \epsilon^2 \). Here \((m, n) = (2, 1)\) and \( S(\lambda)_{\text{max}} = A(\lambda) \) but Theorem 1 fails to recover the leading terms of the two \( \epsilon \)-eigenvalues. There is thus no point in trying to systematically reduce \( A(\lambda, \epsilon) \) itself. It is however always possible to reduce \( B(\lambda, \epsilon) = \text{diag}(I_{\lambda}, \ldots, A(\lambda, \epsilon)) \). Returning to our example, it is not hard to see that \( B(\lambda, \epsilon) = \text{diag}(1, \lambda^2 - \epsilon^2) \) would be transformed into the matrix \( A(\lambda, \epsilon) \) shown in introduction. The possible need for embedding can be explained as follows: the perturbation theory of Section 2 is linear in the sense that only \( A(\lambda, \epsilon) \) mod \( \epsilon^2 \) is used. Embedding may thus be needed to “linearize” \( A(\lambda, \epsilon) \) into an augmented reduced matrix to which Theorem 1 applies successfully.

5. CONCLUDING REMARKS
This study suggests two natural extensions. First, our results can easily be translated to hold in the general case where \( A(\lambda) \) has some nonzero eigenvalues. In that case we can consider the local Smith form of \( A(\lambda) \) at \( \lambda = 0 \) ([9, p. 330]) and use local equivalence of matrix functions rather than matrix polynomials. Second, the constructive proof of Theorem 5 yields a naive reduction algorithm that requires the computation of the Smith form \( S(\lambda, \epsilon) \) of \( A(\lambda, \epsilon) \) with \( A(\lambda, \epsilon) \) viewed as a polynomial matrix over \( \mathbb{F} \). However, it would be interesting to know how to compute an equivalent perturbed matrix polynomial with maximal leading Smith form without resorting to the full knowledge of \( S(\lambda, \epsilon) \). This is left for further research.

6. REFERENCES