

On condition numbers of polynomial eigenvalue problems

Nikolaos Papathanasiou* and Panayiotis Psarrakos*

February 4, 2010

Dedicated to the memory of James H. Wilkinson (1919–1986)

Abstract

In this paper, we investigate condition numbers of eigenvalue problems of matrix polynomials with nonsingular leading coefficients, generalizing classical results of matrix perturbation theory. We provide a relation between the condition numbers of eigenvalues and the pseudospectral growth rate. We obtain that if a simple eigenvalue of a matrix polynomial is ill-conditioned in some respects, then it is close to be multiple, and we construct an upper bound for this distance (measured in the euclidean norm). We also derive a new expression for the condition number of a simple eigenvalue, which does not involve eigenvectors. Moreover, an Elsner-like perturbation bound for matrix polynomials is presented.

Keywords: matrix polynomial, eigenvalue, perturbation, condition number, pseudospectrum.

AMS Subject Classifications: 65F15, 15A18.

1 Introduction

The notions of condition numbers of eigenproblems and eigenvalues quantify the sensitivity of eigenvalue problems [4, 6, 10, 11, 16, 19, 20, 23, 25, 26, 27]. They are widely appreciated tools for investigating the behavior under perturbations of matrix-based dynamical systems and of algorithms in numerical linear algebra. An eigenvalue problem is called *ill-conditioned* (resp., *well-conditioned*) if its condition number is sufficiently large (resp., sufficiently small).

In 1965, Wilkinson [25] introduced the condition number of a simple eigenvalue λ_0 of a matrix A while discussing the sensitivity of λ_0 in terms of the associated right and left eigenspaces. Two years later, Smith [19] obtained explicit expressions for certain condition numbers related to the reduction of matrix A to its Jordan canonical form. In early 1970's, Stewart [20] and Wilkinson [26] used the condition number of the simple eigenvalue λ_0 to construct a lower bound and an upper bound for the distance from A to the set of matrices that have λ_0 as a multiple eigenvalue, respectively. Recently, the notion of the condition number of simple eigenvalues of matrices has been extended to multiple eigenvalues of matrices [4, 10, 11] and to eigenvalues of matrix polynomials [11, 23].

In this article, we are concerned with conditioning for the eigenvalue problem of a matrix polynomial $P(\lambda)$ with a nonsingular leading coefficient, generalizing known

*Department of Mathematics, National Technical University of Athens, Zografou Campus, 15780 Athens, Greece (nipapath@mail.ntua.gr, N. Papathanasiou; ppsarr@math.ntua.gr, P. Psarrakos).

results of matrix perturbation theory [4, 7, 10, 19, 26]. In the next section, we give the definitions and the necessary background on matrix polynomials. In Section 3, we investigate the strong connection between the condition numbers of the eigenvalues of $P(\lambda)$ and the growth rate of its pseudospectra. This connection allows us to portrait the abstraction of the condition numbers of eigenvalues. In Section 4, we examine the relation between the condition number of a simple eigenvalue λ_0 of $P(\lambda)$ and the distance from $P(\lambda)$ to the set of matrix polynomials that have λ_0 as a multiple eigenvalue. In particular, we see that if the condition number of λ_0 is sufficiently large, then this eigenvalue is close to be multiple. In Section 5, we provide a new expression for the condition number of a simple eigenvalue λ_0 of $P(\lambda)$, which involves the distances from λ_0 to the rest of the eigenvalues of $P(\lambda)$. Finally, in Section 6, we present an extension of the Elsner Theorem [7, 21, 22] to matrix polynomials. Simple numerical examples are also given to illustrate our results.

2 Preliminaries on matrix polynomials

Consider an $n \times n$ matrix polynomial

$$P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0, \quad (1)$$

where λ is a complex variable and $A_j \in \mathbb{C}^{n \times n}$ ($j = 0, 1, \dots, m$) with $\det A_m \neq 0$. The study of matrix polynomials has a long history, especially with regard to their spectral analysis, which leads to the solutions of higher order linear systems of differential equations. The suggested references on matrix polynomials are [9, 13, 15].

A scalar $\lambda_0 \in \mathbb{C}$ is called an *eigenvalue* of $P(\lambda)$ if the system $P(\lambda_0)x = 0$ has a nonzero solution $x_0 \in \mathbb{C}^n$, known as a *right eigenvector* of $P(\lambda)$ corresponding to λ_0 . A nonzero vector $y_0 \in \mathbb{C}^n$ that satisfies $y_0^* P(\lambda_0) = 0$ is called a *left eigenvector* of $P(\lambda)$ corresponding to λ_0 . The set of all eigenvalues of $P(\lambda)$ is the *spectrum* of $P(\lambda)$, $\sigma(P) = \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\}$, and since $\det A_m \neq 0$, it contains no more than nm distinct (finite) elements. The *algebraic multiplicity* of an eigenvalue $\lambda_0 \in \sigma(P)$ is the multiplicity of λ_0 as a zero of the (scalar) polynomial $\det P(\lambda)$, and it is always greater than or equal to the *geometric multiplicity* of λ_0 , that is, the dimension of the null space of matrix $P(\lambda_0)$.

2.1 Jordan structure and condition number of the eigenproblem

Let $\lambda_1, \lambda_2, \dots, \lambda_r \in \sigma(P)$ be the eigenvalues of $P(\lambda)$, where each λ_i appears exactly k_i times if and only if its geometric multiplicity is k_i ($i = 1, 2, \dots, r$). Suppose also that for an eigenvalue $\lambda_i \in \sigma(P)$, there exist $x_{i,1}, x_{i,2}, \dots, x_{i,s_i} \in \mathbb{C}^n$ with $x_{i,1} \neq 0$, such that

$$\sum_{j=1}^{\xi} \frac{1}{(j-1)!} P^{(j-1)}(\lambda_i) x_{i,\xi-j+1} = 0; \quad \xi = 1, 2, \dots, s_i,$$

where the indices denote the derivatives of $P(\lambda)$ and s_i cannot exceed the algebraic multiplicity of λ_i . Then the vector $x_{i,1}$ is clearly an eigenvector of λ_i , and the vectors $x_{i,2}, x_{i,3}, \dots, x_{i,s_i}$ are known as *generalized eigenvectors*. The set $\{x_{i,1}, x_{i,2}, \dots, x_{i,s_i}\}$ is called a *Jordan chain of length s_i* of $P(\lambda)$ corresponding to the eigenvalue λ_i .

Any eigenvalue of $P(\lambda)$ of geometric multiplicity k has k maximal Jordan chains associated to k linearly independent eigenvectors, with total number of eigenvectors and generalized eigenvectors equal to the algebraic multiplicity of this eigenvalue.

We consider now the $n \times nm$ matrix $X = [x_{1,1} \cdots x_{1,s_1} \ x_{2,1} \cdots x_{r,1} \cdots x_{r,s_r}]$ formed by maximal Jordan chains of $P(\lambda)$, and the associated $nm \times nm$ Jordan matrix $J = J_1 \oplus J_2 \oplus \cdots \oplus J_r$, where each J_i is the Jordan block that corresponds to the

Jordan chain $\{x_{i,1}, x_{i,2}, \dots, x_{i,s_i}\}$ of λ_i . Then the $nm \times nm$ matrix $Q = \begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{m-1} \end{bmatrix}$

is invertible [9], and we can define $Y = Q^{-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_m^{-1} \end{bmatrix}$. The set (X, J, Y) is called a

Jordan triple of $P(\lambda)$, and satisfies $P(\lambda)^{-1} = X(I\lambda - J)^{-1}Y$ for every scalar $\lambda \notin \sigma(P)$ [9]. Motivated by the latter equality and [5], we define the *condition number of the eigenproblem*¹ of $P(\lambda)$ as $k(P) = \|X\| \|Y\|$, where $\|\cdot\|$ denotes the *spectral matrix norm*, i.e., that norm subordinate to the euclidean vector norm.

2.2 Companion matrix

The (*block*) *companion matrix* of $P(\lambda)$ is the $nm \times nm$ matrix

$$C_P = \begin{bmatrix} 0 & I & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & I \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & \cdots & -A_m^{-1}A_{m-1} \end{bmatrix}.$$

It is straightforward to verify that

$$E(\lambda)(\lambda I - C_P)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{m(n-1)} \end{bmatrix}, \quad (2)$$

where $F(\lambda) = \begin{bmatrix} I & 0 & \cdots & 0 \\ \lambda I & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda^m I & \lambda^{m-1} I & \cdots & I \end{bmatrix}$ and $E(\lambda) = \begin{bmatrix} E_1(\lambda) & E_2(\lambda) & \cdots & E_m(\lambda) \\ -I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & & -I & 0 \end{bmatrix}$

with $E_m(\lambda) = A_m$ and $E_r(\lambda) = A_r + \lambda E_{r+1}(\lambda)$ for $r = m-1, m-2, \dots, 1$. It is also easy to see that $\det F(\lambda) = 1$ and $\det E(\lambda) = \pm \det A_m (\neq 0)$. As a consequence, $\sigma(P)$ coincides with the spectrum of matrix C_P , counting algebraic multiplicities.

¹Note that the definition of the condition number $k(P)$ depends on the choice of the triple (X, J, Y) , but for simplicity, the Jordan triple will not appear explicitly in the notation.

2.3 Weighted perturbations and pseudospectrum

We are interested in perturbations of $P(\lambda)$ of the form

$$Q(\lambda) = P(\lambda) + \Delta(\lambda) = \sum_{j=0}^m (A_j + \Delta_j) \lambda^j, \quad (3)$$

where the matrices $\Delta_0, \Delta_1, \dots, \Delta_m \in \mathbb{C}^{n \times n}$ are arbitrary. For a given parameter $\varepsilon > 0$ and a given set of nonnegative weights $w = \{w_0, w_1, \dots, w_m\}$ with $w_0 > 0$, we define the class of admissible perturbed matrix polynomials

$$\mathcal{B}(P, \varepsilon, w) = \{Q(\lambda) \text{ as in (3)} : \|\Delta_j\| \leq \varepsilon w_j, j = 0, 1, \dots, m\}$$

(recall that $\|\cdot\|$ denotes the spectral matrix norm). The weights w_0, w_1, \dots, w_m allow freedom in how perturbations are measured, and the set $\mathcal{B}(P, \varepsilon, w)$ is convex and compact with respect to the max norm $\|P(\lambda)\|_\infty = \max_{0 \leq j \leq m} \|A_j\|$ [3].

A recently popularized tool for gaining insight into the sensitivity of eigenvalues to perturbations is pseudospectrum; see [3, 8, 12, 24] and the references therein. The ε -pseudospectrum of $P(\lambda)$ (introduced in [17, 24]) is defined by

$$\sigma_\varepsilon(P) = \{\mu \in \sigma(Q) : Q(\lambda) \in \mathcal{B}(P, \varepsilon, w)\} = \{\mu \in \mathbb{C} : s_{\min}(P(\mu)) \leq \varepsilon w(|\mu|)\},$$

where $s_{\min}(\cdot)$ denotes the minimum singular value of a matrix and $w(\lambda) = w_m \lambda^m + w_{m-1} \lambda^{m-1} + \dots + w_1 \lambda + w_0$. The pseudospectrum $\sigma_\varepsilon(P)$ is bounded if and only if $\varepsilon w_m < s_{\min}(A_m)$ [12], and it has no more connected components than the number of distinct eigenvalues of $P(\lambda)$ [3].

2.4 Condition number of a simple eigenvalue

Let $\lambda_0 \in \sigma(P)$ be a simple eigenvalue of $P(\lambda)$ with corresponding right eigenvector $x_0 \in \mathbb{C}^n$ and left eigenvector $y_0 \in \mathbb{C}^n$ (where both x_0 and y_0 are unique up to scalar multiplications). A normwise *condition number* of the eigenvalue λ_0 , originally introduced and studied in [23] (in a slightly different form), is defined by

$$k(P, \lambda_0) = \limsup_{\varepsilon \rightarrow 0} \left\{ \frac{|\delta \lambda_0|}{\varepsilon} : \det Q(\lambda_0 + \delta \lambda_0) = 0, Q(\lambda) \in \mathcal{B}(P, \varepsilon, w) \right\} \quad (4)$$

$$= \frac{w(|\lambda_0|) \|x_0\| \|y_0\|}{|y_0^* P'(\lambda_0) x_0|}. \quad (5)$$

Since λ_0 is also a simple eigenvalue of the companion matrix C_P , we define the *condition number of λ_0 with respect to C_P* as

$$k(C_P, \lambda_0) = \frac{\|\chi_0\| \|\psi_0\|}{|\psi_0^* \chi_0|} \quad (6)$$

(see [19, 26, 27]), where

$$\chi_0 = \begin{bmatrix} x_0 \\ \lambda_0 x_0 \\ \vdots \\ \lambda_0^{m-1} x_0 \end{bmatrix} \quad \text{and} \quad \psi_0 = \begin{bmatrix} E_1(\lambda_0)^* y_0 \\ E_2(\lambda_0)^* y_0 \\ \vdots \\ E_m(\lambda_0)^* y_0 \end{bmatrix} \quad (7)$$

are associated right and left eigenvectors of C_P for the eigenvalue λ_0 , respectively. By straightforward computations, we can see that $\psi_0^* \chi_0 = y_0^* P'(\lambda_0) x_0$. This relation and the definitions (5) and (6) yield the following [14],

$$k(P, \lambda_0) = \frac{w(|\lambda_0|)}{\|\chi_0\| \|\psi_0\|} k(C_P, \lambda_0). \quad (8)$$

2.5 Condition number of a multiple eigenvalue

Suppose that $\lambda_0 \in \sigma(P)$ is a multiple eigenvalue of $P(\lambda)$, and that p_0 is the maximum length of Jordan chains corresponding to λ_0 . Then we can construct a Jordan triple of $P(\lambda)$,

$$(X, J, Y) = \left(\begin{array}{c} [x_{1,1} \cdots x_{1,p_0} \ x_{2,1} \cdots] \\ J_1 \oplus J_2 \oplus \cdots \oplus J_{\kappa_0} \oplus \tilde{J} \end{array} , \begin{array}{c} \begin{bmatrix} y_{1,p_0}^* \\ \vdots \\ y_{1,1}^* \\ y_{2,p_0}^* \\ \vdots \end{bmatrix} \end{array} \right), \quad (9)$$

where $J_1, J_2, \dots, J_{\kappa_0}$ are the $p_0 \times p_0$ Jordan blocks of λ_0 , and \tilde{J} contains all the Jordan blocks of λ_0 of order less than p_0 and all the Jordan blocks that correspond to the rest of the eigenvalues of $P(\lambda)$. Moreover, $x_{1,1}, x_{2,1}, \dots, x_{\kappa_0,1}$ are right eigenvectors of $P(\lambda)$ that correspond to $J_1, J_2, \dots, J_{\kappa_0}$, and $y_{1,1}, y_{2,1}, \dots, y_{\kappa_0,1}$ are the associated left eigenvectors. Following the approach of [4, 10, 11, 16] on multiple eigenvalues, we

consider the matrices $\hat{X} = [x_{1,1} \ x_{2,1} \ \cdots \ x_{\kappa_0,1}] \in \mathbb{C}^{n \times \kappa_0}$ and $\hat{Y} = \begin{bmatrix} y_{1,1}^* \\ y_{2,1}^* \\ \vdots \\ y_{\kappa_0,1}^* \end{bmatrix} \in \mathbb{C}^{\kappa_0 \times n}$,

and define the *condition number of the multiple eigenvalue* λ_0 by

$$\hat{k}(P, \lambda_0) = w(|\lambda_0|) \|\hat{X} \hat{Y}\|. \quad (10)$$

Notice that since the matrices \hat{X} and \hat{Y} are of rank $\kappa_0 \leq n$, the product $\hat{X} \hat{Y}$ is nonzero and $\hat{k}(P, \lambda_0) > 0$ (keeping in mind that $w_0 > 0$). Moreover, if the eigenvalue λ_0 is simple, i.e., $p_0 = \kappa_0 = 1$, then the definitions (5) and (10) coincide [11].

3 Condition numbers of eigenvalues and pseudospectral growth

Consider a matrix polynomial $P(\lambda)$ as in (1). Since the leading coefficient of $P(\lambda)$ is nonsingular, for sufficiently small ε , the pseudospectrum $\sigma_\varepsilon(P)$ consists of no more than nm bounded connected components, each one containing a single (possibly multiple) eigenvalue of $P(\lambda)$. By the definition (4) and the proof of Theorem 5 of [23], it follows that any small connected component of $\sigma_\varepsilon(P)$ that contains exactly one simple eigenvalue $\lambda_0 \in \sigma(P)$ is approximately a disc centered at λ_0 . Recall that the Hausdorff distance between two sets $\mathcal{S}, \mathcal{T} \subset \mathbb{C}$ is

$$\mathcal{H}(\mathcal{S}, \mathcal{T}) = \max \left\{ \sup_{s \in \mathcal{S}} \inf_{t \in \mathcal{T}} |s - t|, \sup_{t \in \mathcal{T}} \inf_{s \in \mathcal{S}} |s - t| \right\}.$$

Proposition 1. *If $\lambda_0 \in \sigma(P)$ is a simple eigenvalue of $P(\lambda)$, then as $\varepsilon \rightarrow 0$, the Hausdorff distance between the connected component of $\sigma_\varepsilon(P)$ that contains λ_0 and the disc $\{\mu \in \mathbb{C} : |\mu - \lambda_0| \leq k(P, \lambda_0)\varepsilon\}$ is $o(\varepsilon)$.*

Next we extend this proposition to multiple eigenvalues of the matrix polynomial $P(\lambda)$, generalizing a technique of [10] for matrices (see also [4]).

Theorem 2. *Suppose that λ_0 is a multiple eigenvalue of $P(\lambda)$ and p_0 is the dimension of the maximum Jordan blocks of λ_0 . Then as $\varepsilon \rightarrow 0$, the Hausdorff distance between the connected component of pseudospectrum $\sigma_\varepsilon(P)$ that contains λ_0 and the disc $\{\mu \in \mathbb{C} : |\mu - \lambda_0| \leq (\hat{k}(P, \lambda_0)\varepsilon)^{1/p_0}\}$ is $o(\varepsilon^{1/p_0})$.*

Proof. Consider the Jordan triple (X, J, Y) of $P(\lambda)$ in (9) and the condition number $\hat{k}(P, \lambda_0)$ in (10). For sufficiently small $\varepsilon > 0$, the pseudospectrum $\sigma_\varepsilon(P)$ has a compact connected component \mathcal{G}_ε such that $\mathcal{G}_\varepsilon \cap \sigma(P) = \{\lambda_0\}$. In particular, the eigenvalue λ_0 lies in the (nonempty) interior of \mathcal{G}_ε ; see Corollary 3 and Lemma 8 of [3]. Let also μ be a boundary point of \mathcal{G}_ε . Then it holds that

$$s_{\min}(P(\mu)) = \varepsilon w(|\mu|) \quad \text{and} \quad P(\mu)^{-1} = X(I\mu - J)^{-1}Y.$$

Denote now by N the $p_0 \times p_0$ nilpotent matrix having ones on the super diagonal and zeros elsewhere, and observe that

$$N^{p_0} = 0 \quad \text{and} \quad (I\lambda - N)^{-1} = \begin{bmatrix} \lambda^{-1} & \lambda^{-2} & \dots & \lambda^{-p_0} \\ 0 & \lambda^{-1} & \dots & \lambda^{-p_0+1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^{-1} \end{bmatrix} = \lambda^{-1} \sum_{j=0}^{p_0-1} (\lambda^{-1}N)^j \quad (\lambda \neq 0).$$

As in [4, 10], we verify that

$$\begin{aligned} \frac{|\mu - \lambda_0|^{p_0}}{s_{\min}(P(\mu))} &= |\mu - \lambda_0|^{p_0} \|P^{-1}(\mu)\| = |\mu - \lambda_0|^{p_0} \|X(I\mu - J)^{-1}Y\| \\ &= |\mu - \lambda_0|^{p_0} \left\| X \operatorname{diag} \left\{ (I\mu - J_1)^{-1}, \dots, (I\mu - J_{\kappa_0})^{-1}, (I\mu - \hat{J})^{-1} \right\} Y \right\| \\ &= \left\| (\mu - \lambda_0)^{p_0-1} X \operatorname{diag} \left\{ \sum_{j=0}^{p_0-1} ((\mu - \lambda_0)^{-1}N)^j, \dots \right. \right. \\ &\quad \left. \left. \dots, \sum_{j=0}^{p_0-1} ((\mu - \lambda_0)^{-1}N)^j, (\mu - \lambda_0)(I\mu - \hat{J})^{-1} \right\} Y \right\|. \end{aligned}$$

For each one of the first κ_0 diagonal blocks, we have

$$(\mu - \lambda_0)^{p_0-1} \sum_{j=0}^{p_0-1} ((\mu - \lambda_0)^{-1}N)^j = N^{p_0-1} + O(\mu - \lambda_0).$$

Thus, it follows

$$\begin{aligned}
\frac{|\mu - \lambda_0|^{p_0}}{s_{\min}(P(\mu))} &= \|X \operatorname{diag} \{N^{p_0-1} + O(\mu - \lambda_0), \dots, N^{p_0-1} + O(\mu - \lambda_0), O((\mu - \lambda_0)^{p_0})\} Y\| \\
&= \|X \operatorname{diag} \{N^{p_0-1}, \dots, N^{p_0-1}, 0\} Y\| + O(|\mu - \lambda_0|) \\
&= \left\| \begin{bmatrix} 0 & \dots & x_{1,1} & 0 & \dots & x_{2,1} & 0 & \dots \end{bmatrix} \begin{bmatrix} y_{1,p_0}^* \\ \vdots \\ y_{1,1}^* \\ y_{2,p_0}^* \\ \vdots \end{bmatrix} \right\| + O(|\mu - \lambda_0|),
\end{aligned}$$

where the right eigenvectors $x_{1,1}, x_{2,1}, \dots, x_{\kappa_0,1}$ and the rows $y_{1,1}^*, y_{2,1}^*, \dots, y_{\kappa_0,1}^*$ lie at positions $p_0, 2p_0, \dots, \kappa_0 p_0$, respectively. As a consequence,

$$\frac{|\mu - \lambda_0|^{p_0}}{s_{\min}(P(\mu))} = \|\hat{X} \hat{Y}\| + O(|\mu - \lambda_0|),$$

or

$$\frac{|\mu - \lambda_0|^{p_0}}{\varepsilon w(|\mu|) \|\hat{X} \hat{Y}\|} = 1 + O(|\mu - \lambda_0|),$$

or

$$\frac{|\mu - \lambda_0|}{(\hat{k}(P, \lambda_0) \varepsilon)^{1/p_0}} = 1 + r_\varepsilon,$$

where $r_\varepsilon \in \mathbb{R}$ goes to 0 as $\varepsilon \rightarrow 0$. This means that

$$|\mu - \lambda_0| = (\hat{k}(P, \lambda_0) \varepsilon)^{1/p_0} + o(\varepsilon^{1/p_0}).$$

Since μ lies on the boundary $\partial \mathcal{G}_\varepsilon$, it is easy to see that the Hausdorff distance between \mathcal{G}_ε and the disc $\{\mu \in \mathbb{C} : |\mu - \lambda_0| \leq (\hat{k}(P, \lambda_0) \varepsilon)^{1/p_0}\}$ is $o(\varepsilon^{1/p_0})$. \square

The above two results indicate how the condition number of an eigenvalue of $P(\lambda)$ quantifies the sensitivity of this eigenvalue. Consider, for example, the matrix polynomial

$$P(\lambda) = \begin{bmatrix} (\lambda - 1)^2 & \lambda - 1 & \lambda - 1 \\ 0 & (\lambda - 1)^2 & 0 \\ 0 & \lambda^2 - 1 & \lambda^2 - 1 \end{bmatrix}$$

with $\det(P(\lambda)) = (\lambda - 1)^5(\lambda + 1)$ and $\sigma(P) = \{1, -1\}$. The eigenvalue $\lambda = 1$ has algebraic multiplicity 5 and geometric multiplicity 3, and the eigenvalue $\lambda = -1$ is simple. A Jordan triple of $P(\lambda)$ is given by

$$X = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 2 \end{bmatrix}, \quad J = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & 0 & 0.25 \\ 1 & 0 & -0.5 \\ 0 & 1 & -0.5 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \\ 0 & 0 & -0.25 \end{bmatrix}.$$

The matrices of the eigenvectors that correspond to the maximum Jordan blocks of eigenvalue $\lambda = 1$ are $\hat{X} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix}$ and $\hat{Y} = \begin{bmatrix} 1 & 0 & -0.5 \\ 0 & 1 & 0 \end{bmatrix}$. Thus, for the weights $w_0 = w_1 = w_2 = 1$, we have $\hat{k}(P, 1) = w(1) \|\hat{X} \hat{Y}\| = 4.2426$.

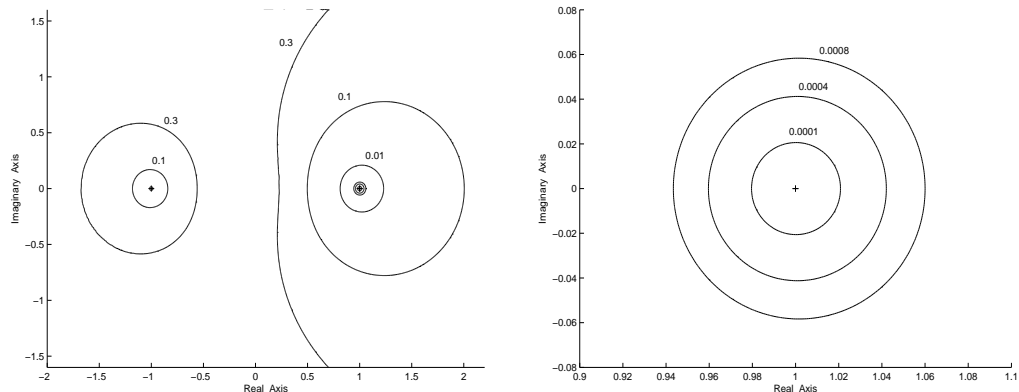


Figure 1: The boundaries $\partial\sigma_\varepsilon(P)$ for $\varepsilon = 10^{-4}, 4 \cdot 10^{-4}, 8 \cdot 10^{-4}, 10^{-2}, 10^{-1}, 3 \cdot 10^{-1}$.

The boundaries of the pseudospectra $\sigma_\varepsilon(P)$, $\varepsilon = 10^{-4}, 4 \cdot 10^{-4}, 8 \cdot 10^{-4}, 10^{-2}, 10^{-1}, 3 \cdot 10^{-1}$, are illustrated in the left part of Figure 1. The eigenvalues of $P(\lambda)$ are marked with +’s, and the components of the simple eigenvalue $\lambda = -1$ are visible only for $\varepsilon = 10^{-2}, 10^{-1}, 3 \cdot 10^{-1}$. The components of the multiple eigenvalue $\lambda = 1$ are magnified in the right part of the figure, and they are very close to circular discs centered at $\lambda = 1$ of radii $(\hat{k}(P, 1) 10^{-4})^{1/2} = 0.0206$, $(\hat{k}(P, 1) 4 \cdot 10^{-4})^{1/2} = 0.0412$ and $(\hat{k}(P, 1) 8 \cdot 10^{-4})^{1/2} = 0.0583$, confirming Theorem 2.

4 Distance from a given simple eigenvalue to multiplicity

Let $P(\lambda)$ be a matrix polynomial as in (1), and let λ_0 be a simple eigenvalue of $P(\lambda)$. In the sequel, we generalize a methodology of Wilkinson [26] in order to obtain a relation between the condition number $k(P, \lambda_0)$ and the distance from $P(\lambda)$ to the matrix polynomials that have λ_0 as a multiple eigenvalue, namely,

$$\text{dist}(P, \lambda_0) = \inf \{ \varepsilon > 0 : \exists Q(\lambda) \in \mathcal{B}(P, \varepsilon, w) \text{ with } \lambda_0 \text{ as a multiple eigenvalue} \}.$$

The next proposition is a known result (see [1, Theorem 3.2] and [3, Proposition 16]). Here, we give a new proof, which motivates the proof of the main result of this section (Theorem 4) and is necessary for the remainder.

Proposition 3. *Let $P(\lambda)$ be a matrix polynomial as in (1), $\lambda_0 \in \sigma(P) \setminus \sigma(P')$ and $y_0, x_0 \in \mathbb{C}^n$ be corresponding left and right unit eigenvectors, respectively. If $y_0^* P'(\lambda_0) x_0 = 0$, then λ_0 is a multiple eigenvalue of $P(\lambda)$.*

Proof. By Schur's triangularization, and without loss of generality, we may assume that the matrix $P(\lambda_0)$ has the following form,

$$P(\lambda_0) = \begin{bmatrix} 0 & b^* \\ 0 & B \end{bmatrix}; \quad b \in \mathbb{C}^{n-1}, \quad B \in \mathbb{C}^{(n-1) \times (n-1)}.$$

Moreover, since $P(\lambda_0)x_0 = 0$, we can set $x_0 = e_1 = [1 \ 0 \ \dots \ 0]^T$. Then we have that $y_0^*P'(\lambda_0)e_1 = 0$, and hence, $y_0^*P'(\lambda_0) = [0 \ w^*]$ for some $0 \neq w \in \mathbb{C}^{n-1}$.

Since $\lambda_0 \notin \sigma(P')$ and $y_0^*P(\lambda_0) = 0$, it follows

$$y_0^*P'(\lambda_0) \{ [P'(\lambda_0)]^{-1}P(\lambda_0) \} = 0,$$

or equivalently,

$$y_0^*P'(\lambda_0) \left\{ [P'(\lambda_0)]^{-1} \begin{bmatrix} 0 & b^* \\ 0 & B \end{bmatrix} \right\} = 0,$$

or equivalently,

$$[0 \ w^*] \begin{bmatrix} 0 & a^* \\ 0 & A \end{bmatrix} = 0,$$

where $a \in \mathbb{C}^{n-1}$ and $A \in \mathbb{C}^{(n-1) \times (n-1)}$. As a consequence, $w^*A = 0$ and the matrix A has 0 as an eigenvalue. Thus, 0 is a multiple eigenvalue of the matrix $[P'(\lambda_0)]^{-1}P(\lambda_0)$. We consider two cases:

(i) If the geometric multiplicity of $0 \in \sigma([P'(\lambda_0)]^{-1}P(\lambda_0))$ is greater than or equal to 2, then $\text{rank}(P(\lambda_0)) \leq n - 2$, and hence, λ_0 is a multiple eigenvalue of $P(\lambda)$.

(ii) Suppose that the geometric multiplicity of the eigenvalue $0 \in \sigma([P'(\lambda_0)]^{-1}P(\lambda_0))$ is equal to 1 and its algebraic multiplicity is greater than or equal to 2. Then, keeping in mind that $[P'(\lambda_0)]^{-1}P(\lambda_0)e_1 = 0$, we verify that there exists a vector $z_1 \in \mathbb{C}^n$ such that $[P'(\lambda_0)]^{-1}P(\lambda_0)z_1 = e_1$, or equivalently, $P(\lambda_0)(-z_1) + P'(\lambda_0)e_1 = 0$. Thus, λ_0 is a multiple eigenvalue of $P(\lambda)$ with a Jordan chain of length at least 2. \square

Recall that the condition number of an invertible matrix A is defined by $c(A) = \|A\| \|A^{-1}\|$ and it is always greater than or equal to 1.

Theorem 4. *Let $P(\lambda)$ be a matrix polynomial as in (1), $\lambda_0 \in \sigma(P) \setminus \sigma(P')$ be a simple eigenvalue of $P(\lambda)$, and $y_0, x_0 \in \mathbb{C}^n$ be corresponding left and right unit eigenvectors, respectively. If the vector $[P'(\lambda_0)]^*y_0$ is not a scalar multiple of x_0 , then*

$$\text{dist}(P, \lambda_0) \leq \frac{c(P'(\lambda_0)) \|P(\lambda_0)\|}{k(P, \lambda_0) \left(\|y_0^*P'(\lambda_0)\|^2 - |y_0^*P'(\lambda_0)x_0|^2 \right)^{1/2}}.$$

Proof. As in the proof of the previous proposition, without loss of generality, we may assume that

$$P(\lambda_0) = \begin{bmatrix} 0 & b^* \\ 0 & B \end{bmatrix}; \quad b \in \mathbb{C}^{n-1}, \quad B \in \mathbb{C}^{(n-1) \times (n-1)}$$

and $x_0 = e_1$. If we denote $\delta = y_0^*P'(\lambda_0)x_0 = y_0^*P'(\lambda_0)e_1 \neq 0$, then it is clear that

$$y_0^*P'(\lambda_0) = [\delta \ w^*],$$

for some $w \in \mathbb{C}^{n-1}$. Furthermore, $w \neq 0$ because $|\delta| < \|y_0^* P'(\lambda_0)\|$.

Since $\lambda_0 \notin \sigma(P')$ and $y_0^* P(\lambda_0) = 0$, it follows

$$y_0^* P'(\lambda_0) \{ [P'(\lambda_0)]^{-1} P(\lambda_0) \} = 0,$$

or equivalently,

$$y_0^* P'(\lambda_0) \left\{ [P'(\lambda_0)]^{-1} \begin{bmatrix} 0 & b^* \\ 0 & B \end{bmatrix} \right\} = 0,$$

or equivalently,

$$[\delta \ w^*] \begin{bmatrix} 0 & a^* \\ 0 & A \end{bmatrix} = 0$$

for some $a \in \mathbb{C}^{n-1}$ and $A \in \mathbb{C}^{(n-1) \times (n-1)}$. If $a = 0$, then $w^* A = 0$, and the proof of Proposition 3 implies that λ_0 is a multiple eigenvalue of $P(\lambda)$; this is a contradiction. As a consequence, $a \neq 0$. Moreover,

$$w^* A + \delta a^* = 0,$$

and hence,

$$w^* \left(A + \frac{\delta}{w^* w} w a^* \right) = 0.$$

This means that if we consider the (perturbation) matrix $E = \begin{bmatrix} 0 & 0 \\ 0 & \frac{\delta}{w^* w} w a^* \end{bmatrix}$, then the matrix

$$[P'(\lambda_0)]^{-1} P(\lambda_0) + E = [P'(\lambda_0)]^{-1} [P(\lambda_0) + P'(\lambda_0) E]$$

has 0 as a multiple eigenvalue.

We define the $n \times n$ matrices

$$\hat{\Delta} = P'(\lambda_0) E \quad \text{and} \quad \hat{Q} = P(\lambda_0) + \hat{\Delta},$$

and the matrix polynomial $\Delta(\lambda) = \sum_{j=0}^m \Delta_j \lambda^j$ with coefficients

$$\Delta_j = \left(\frac{\overline{\lambda_0}}{|\lambda_0|} \right)^j \frac{w_j}{w(|\lambda_0|)} \hat{\Delta}; \quad j = 0, 1, \dots, m,$$

where (by convention) we assume that $\overline{\lambda_0}/\lambda_0 = 0$ whenever $\lambda_0 = 0$. Then, denoting $\phi = \frac{w'(|\lambda_0|)}{w(|\lambda_0|)} \frac{\overline{\lambda_0}}{|\lambda_0|}$, one can verify that

$$\Delta(\lambda_0) = \hat{\Delta} \quad \text{and} \quad \Delta'(\lambda_0) = \phi \hat{\Delta}.$$

We define also the matrix polynomial $Q(\lambda) = P(\lambda) + \Delta(\lambda)$, and consider two cases:

(i) Suppose that the geometric multiplicity of $0 \in \sigma([P'(\lambda_0)]^{-1} \hat{Q})$ is greater than or equal to 2. Then $\text{rank}(\hat{Q}) = \text{rank}(Q(\lambda_0)) \leq n - 2$, or equivalently, λ_0 is a multiple eigenvalue of the matrix polynomial $Q(\lambda)$ of geometric multiplicity at least 2.

(ii) Suppose now that the geometric multiplicity of the eigenvalue $0 \in \sigma([P'(\lambda_0)]^{-1} \hat{Q})$ is equal to 1, and its algebraic multiplicity is greater than or equal to 2. Then, keeping in mind that $\hat{Q} e_1 = 0$, there is a vector $z_1 \in \mathbb{C}^n$ such that

$$[P'(\lambda_0)]^{-1} \hat{Q} z_1 = e_1,$$

or equivalently,

$$\hat{Q}(-z_1) + P'(\lambda_0)e_1 = 0. \quad (11)$$

We observe that $\Delta'(\lambda_0)e_1 = \phi\hat{\Delta}e_1 = \phi P'(\lambda_0)Ee_1 = 0$. As a consequence, (11) is written in the form

$$Q(\lambda_0)(-z_1) + Q'(\lambda_0)e_1 = 0.$$

Thus, λ_0 is a multiple eigenvalue of $Q(\lambda)$ with a Jordan chain of length at least 2.

In both cases above, we have proved that λ_0 is a multiple eigenvalue of $Q(\lambda)$. Furthermore, we see that

$$\begin{aligned} \|E\| &= \left\| \begin{bmatrix} 0 & 0 \\ 0 & \frac{\delta}{w^*w}wa^* \end{bmatrix} \right\| = \left\| \frac{\delta}{w^*w}wa^* \right\| = \frac{|\delta|}{\|w\|} \|a\| \\ &\leq \frac{|\delta|}{\|w\|} \left\| \begin{bmatrix} 0 & a \\ 0 & A \end{bmatrix} \right\| = \frac{|\delta|}{\|w\|} \|[P'(\lambda_0)]^{-1}P(\lambda_0)\| \\ &\leq \frac{|\delta|}{\|w\|} \|[P'(\lambda_0)]^{-1}\| \|P(\lambda_0)\|. \end{aligned}$$

Hence, for every $j = 0, 1, \dots, m$,

$$\begin{aligned} \|\Delta_j\| &= \frac{w_j}{w(|\lambda_0|)} \|\hat{\Delta}\| = \frac{w_j}{w(|\lambda_0|)} \|P'(\lambda_0)E\| \\ &\leq \frac{w_j}{w(|\lambda_0|)} \|P'(\lambda_0)\| \|E\| \\ &\leq \frac{w_j}{w(|\lambda_0|)} \frac{|\delta|}{\|w\|} \|[P'(\lambda_0)]^{-1}\| \|P'(\lambda_0)\| \|P(\lambda_0)\| \\ &= w_j \frac{c(P'(\lambda_0)) \|P(\lambda_0)\|}{k(P, \lambda_0) \left(\|y_0^* P'(\lambda_0)\|^2 - \delta^2 \right)^{1/2}}, \end{aligned}$$

and the proof is complete. \square

The spectrum of the matrix polynomial

$$P(\lambda) = I\lambda^2 + \begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 0.25 \\ 0 & 0 & -0.5 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 & 8 \\ 0 & 25 & -i \\ 0 & 0 & 15.25 \end{bmatrix}$$

is $\sigma(P) = \{0, -1, 0.25 \pm i3.8971, \pm i5\}$. For the weights $w_2 = \|A_2\| = 1$, $w_1 = \|A_1\| = 2.2919$ and $w_0 = \|A_0\| = 25.0379$, the above theorem implies $\text{dist}(P, -1) \leq 0.4991$. If we estimate the same distance using the method proposed in [18], then we see that $\text{dist}(P, -1) \leq 0.5991$. On the other hand, for the eigenvalue $0.25 - i3.8971$, Theorem 4 yields $\text{dist}(P, 0.25 - i3.8971) \leq 0.1485$, and the method of [18] implies $\text{dist}(P, 0.25 - i3.8971) \leq 0.1398$. At this point, it is necessary to note that the methodology of [18] is applicable to every complex number and not only to simple eigenvalues of $P(\lambda)$.

Remark 5. The proofs of Proposition 3 and Theorem 4 do not depend on the companion matrix C_P , and do not require the invertibility of the leading coefficient A_m .

Furthermore, the definitions (4) and (5) are also valid for the finite eigenvalues of matrix polynomials with singular leading coefficients [11, 23]. As a consequence, Proposition 3 and Theorem 4 hold also in the case where A_m is singular (for the finite eigenvalues of $P(\lambda)$).

5 An expression of $k(P, \lambda_0)$ without eigenvectors

In this section, we derive a new expression of the condition number $k(P, \lambda_0)$ that involves the distances from $\lambda_0 \in \sigma(P)$ to the rest of the eigenvalues of the matrix polynomial $P(\lambda)$, instead of the left and right eigenvectors of λ_0 . The next three lemmas are necessary for our discussion. The first lemma is part of the proof of Theorem 2 in [19], the second lemma follows readily from the singular value decomposition, and the third lemma is part of Theorem 4 in [19].

Lemma 6. *For any matrices $C, R, W \in \mathbb{C}^{n \times n}$, $R \operatorname{adj}(WCR)W = \det(WR) \operatorname{adj}(C)$.*

Lemma 7. *Let A be an $n \times n$ matrix with 0 as a simple eigenvalue, $s_1 \geq s_2 \geq \dots \geq s_{n-1} > s_n = 0$ be the singular values of A , and $u_n, v_n \in \mathbb{C}^n$ be left and right singular vectors of $s_n = 0$, respectively. Then u_n and v_n are also left and right eigenvectors of A corresponding to 0 , respectively.*

Lemma 8. *Let A be an $n \times n$ matrix with 0 as a simple eigenvalue. If $s_1 \geq s_2 \geq \dots \geq s_{n-1} > s_n = 0$ are the singular values of A , then $\|\operatorname{adj}(A)\| = s_1 s_2 \dots s_{n-1}$.*

The following theorem is a direct generalization of Theorem 2 of [19].

Theorem 9. *Let $P(\lambda)$ be a matrix polynomial as in (1) with spectrum $\sigma(P) = \{\lambda_1, \lambda_2, \dots, \lambda_{nm}\}$, counting algebraic multiplicities. If λ_i is a simple eigenvalue, then*

$$k(P, \lambda_i) = \frac{w(|\lambda_i|) \|\operatorname{adj}(P(\lambda_i))\|}{|\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|}.$$

Proof. For the simple eigenvalue $\lambda_i \in \sigma(P)$, consider a singular value decomposition of matrix $P(\lambda_i)$,

$$P(\lambda_i) = U \Sigma V^* = U \operatorname{diag}\{s_1, \dots, s_{n-1}, 0\} V^*.$$

Then we have

$$\begin{bmatrix} U^* & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \begin{bmatrix} P(\lambda_i) & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \begin{bmatrix} V & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} = \begin{bmatrix} \Sigma & 0 \\ 0 & I_{n(m-1)} \end{bmatrix},$$

and Lemma 6 implies

$$\begin{aligned} & \begin{bmatrix} V & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \operatorname{adj} \left(\begin{bmatrix} \Sigma & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right) \begin{bmatrix} U^* & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \\ &= \det(U^*V) \operatorname{adj} \left(\begin{bmatrix} P(\lambda_0) & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right), \end{aligned} \tag{12}$$

where $|\det(U^*V)| = 1$.

Let $u_n, v_n \in \mathbb{C}^n$ be the last columns of U and V , respectively, i.e., they are left and right singular vectors of the zero singular value of $P(\lambda_i)$. Then by Lemma 7, $y_i = u_n$ and $x_i = v_n$ are left and right unit eigenvectors of $\lambda_i \in \sigma(P)$, respectively. Let also ψ_i and χ_i be the associated left and right eigenvectors of C_P for the eigenvalue λ_i given by (7). Then by (8), [19, Theorem 2], Lemma 6, (2) and (12) (applied in this specific order), it follows

$$\begin{aligned}
k(P, \lambda_i) &= \frac{w(|\lambda_i|)}{\|\chi_i\| \|\psi_i\|} k(C_P, \lambda_i) \\
&= \frac{w(|\lambda_i|)}{\|\chi_i\| \|\psi_i\|} \frac{\|\text{adj}(\lambda_i I - C_P)\|}{\prod_{j \neq i} |\lambda_j - \lambda_i|} \\
&= \frac{w(|\lambda_i|)}{\|\chi_i\| \|\psi_i\|} \frac{\|F(\lambda_i) \text{adj}(E(\lambda_i)(\lambda_i I - C_P)F(\lambda_i)) E(\lambda_i)\|}{|\det(F(\lambda_i)E(\lambda_i))| \prod_{j \neq i} |\lambda_j - \lambda_i|} \\
&= \frac{w(|\lambda_i|)}{\|\chi_i\| \|\psi_i\|} \frac{\left\| F(\lambda_i) \text{adj} \left(\begin{bmatrix} P(\lambda_i) & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right) E(\lambda_i) \right\|}{|\det(F(\lambda_i)E(\lambda_i))| \prod_{j \neq i} |\lambda_j - \lambda_i|} \\
&= \frac{w(|\lambda_i|)}{\|\chi_i\| \|\psi_i\|} \frac{\left\| F(\lambda_i) \begin{bmatrix} V & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \text{adj} \left(\begin{bmatrix} \Sigma & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right) \begin{bmatrix} U^* & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} E(\lambda_i) \right\|}{|\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|}.
\end{aligned}$$

Thus,

$$k(P, \lambda_i) = \frac{w(|\lambda_i|) \|G\|}{\|\chi_i\| \|\psi_i\| |\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|}, \quad (13)$$

where

$$G = F(\lambda_i) \begin{bmatrix} V & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \text{adj} \left(\begin{bmatrix} \Sigma & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right) \begin{bmatrix} U^* & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} E(\lambda_i).$$

Moreover,

$$\text{adj} \left(\begin{bmatrix} \Sigma & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \right) = \begin{bmatrix} S & 0 \\ 0 & 0_{n(m-1)} \end{bmatrix},$$

where $S = s_1 s_2 \cdots s_{n-1} \text{diag}\{0, \dots, 0, 1\}$. As a consequence, the matrix G is written

$$\begin{aligned}
G &= F(\lambda_i) \begin{bmatrix} V & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} \begin{bmatrix} S & 0 \\ 0 & 0_{n(m-1)} \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & I_{n(m-1)} \end{bmatrix} E(\lambda_i) \\
&= F(\lambda_i) \begin{bmatrix} V S U^* & 0 \\ 0 & 0_{n(m-1)} \end{bmatrix} E(\lambda_i) \\
&= \begin{bmatrix} V S U^* & 0 & \cdots & 0 \\ \lambda_i V S U^* & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i^{m-1} V S U^* & 0 & \cdots & 0 \end{bmatrix} E(\lambda_i) \\
&= s_1 s_2 \cdots s_{n-1} \begin{bmatrix} v_n u_n^* & 0 & \cdots & 0 \\ \lambda_i v_n u_n^* & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i^{m-1} v_n u_n^* & 0 & \cdots & 0 \end{bmatrix} E(\lambda_i)
\end{aligned}$$

$$\begin{aligned}
&= s_1 s_2 \cdots s_{n-1} \begin{bmatrix} x_i y_i^* & 0 & \cdots & 0 \\ \lambda_i x_i y_i^* & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_i^{m-1} x_i y_i^* & 0 & \cdots & 0 \end{bmatrix} E(\lambda_i) \\
&= s_1 s_2 \cdots s_{n-1} \begin{bmatrix} x_i y_i^* E_1(\lambda_i) & \cdots & x_i y_i^* E_m(\lambda_i) \\ \lambda_i x_i y_i^* E_1(\lambda_i) & \cdots & \lambda_i x_i y_i^* E_m(\lambda_i) \\ \vdots & \ddots & \vdots \\ \lambda_i^{m-1} x_i y_i^* E_1(\lambda_i) & \cdots & \lambda_i^{m-1} x_i y_i^* E_m(\lambda_i) \end{bmatrix} \\
&= s_1 s_2 \cdots s_{n-1} \begin{bmatrix} x_i \\ \lambda_i x_i \\ \vdots \\ \lambda_i^{m-1} x_i \end{bmatrix} \begin{bmatrix} E_1(\lambda_i)^* y_i \\ E_2(\lambda_i)^* y_i \\ \vdots \\ E_m(\lambda_i)^* y_i \end{bmatrix}^* \\
&= s_1 s_2 \cdots s_{n-1} (\chi_i \psi_i^*).
\end{aligned}$$

Hence, by (13) and Lemma 8, it follows

$$k(P, \lambda_i) = \frac{w(|\lambda_i|) (s_1 s_2 \cdots s_{n-1}) \|\chi_i \psi_i^*\|}{\|\chi_i\| \|\psi_i\| |\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|} = \frac{w(|\lambda_i|) \|\text{adj}(P(\lambda_i))\|}{|\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|} \frac{\|\chi_i \psi_i^*\|}{\|\chi_i\| \|\psi_i\|}.$$

Since $\|\chi_i \psi_i^*\| = \|\chi_i\| \|\psi_i\|$, the proof is complete. \square

The next corollary follows readily.

Corollary 10. *Let $P(\lambda)$ be a matrix polynomial as in (1) with spectrum $\sigma(P) = \{\lambda_1, \lambda_2, \dots, \lambda_{nm}\}$, counting algebraic multiplicities. If λ_i is a simple eigenvalue of $P(\lambda)$ with $y_i, x_i \in \mathbb{C}^n$ associated left and right unit eigenvectors, respectively, then*

$$\min_{j \neq i} |\lambda_j - \lambda_i| \leq \left(\frac{w(|\lambda_i|) \|\text{adj}(P(\lambda_i))\|}{k(P, \lambda_i) |\det A_m|} \right)^{\frac{1}{nm-1}} = \left(\frac{|y_i^* P'(\lambda_i) x_i| \|\text{adj}(P(\lambda_i))\|}{|\det A_m|} \right)^{\frac{1}{nm-1}}.$$

Moreover, if the vector $[P'(\lambda_0)]^* y_0$ is not a scalar multiple of x_0 , then

$$\text{dist}(P, \lambda_i) \leq \frac{c(P'(\lambda_i)) \|P(\lambda_i)\| |\det A_m|}{w(|\lambda_i|) \|\text{adj}(P(\lambda_i))\| \left(\|y_i^* P'(\lambda_i)\|^2 - |y_i^* P'(\lambda_i) x_i|^2 \right)^{1/2} \prod_{j \neq i} |\lambda_j - \lambda_i|}.$$

It is remarkable that for the simple eigenvalue $\lambda_i \in \sigma(P)$, Theorem 9 and the definition (5) yield

$$|y_i^* P'(\lambda_i) x_i| = \frac{|\det A_m| \prod_{j \neq i} |\lambda_j - \lambda_i|}{\|\text{adj}(P(\lambda_i))\|} \neq 0 \quad (\|x_i\| = \|y_i\| = 1).$$

Thus, Proposition 3 follows as a corollary of Theorem 9, and the size of the angle between the vectors $[P'(\lambda_i)]^* y_i$ and x_i is partially expressed in algebraic terms such as determinants and eigenvalues. Note also that λ_i is relatively close to some other eigenvalues of $P(\lambda)$ if and only if $k(P, \lambda_i)$ is sufficiently greater than the quantity $w(|\lambda_i|) \|\text{adj}(P(\lambda_i))\| |\det A_m|^{-1}$. Furthermore, the condition number $k(P, \lambda_i)$ is

relatively large (and λ_i is an ill-conditioned eigenvalue) if and only if the product $\prod_{j \neq i} |\lambda_j - \lambda_i|$ is sufficiently less than $w(|\lambda_i|) \|\text{adj}(P(\lambda_i))\| |\det A_m|^{-1}$.

To illustrate numerically the latter remark, consider the matrix polynomial

$$P(\lambda) = \begin{bmatrix} 0.001 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} -0.003 & 0 \\ 0 & -7 \end{bmatrix} \lambda + \begin{bmatrix} 0.002 & 0.001 \\ 0 & 12 \end{bmatrix}$$

with (well separated) simple eigenvalues 1, 2, 3 and 4, and set $w_0 = w_1 = w_2 = 1$. Then it is straightforward to see that for the eigenvalues $\lambda = 1$ and $\lambda = 2$,

$$k(P, 1) \cong 3000, \quad |2 - 1| |3 - 1| |4 - 1| = 6 \quad \text{and} \quad \frac{w(1) \|\text{adj}(P(1))\|}{|\det A_m|} \cong 18000,$$

and

$$k(P, 2) \cong 7000, \quad |1 - 2| |3 - 2| |4 - 2| = 2 \quad \text{and} \quad \frac{w(2) \|\text{adj}(P(2))\|}{|\det A_m|} \cong 14000.$$

On the other hand, for the eigenvalue $\lambda = 4$, we have

$$k(P, 4) \cong 21.2897, \quad |1 - 4| |2 - 4| |3 - 4| = 6 \quad \text{and} \quad \frac{w(4) \|\text{adj}(P(4))\|}{|\det A_m|} \cong 127.738.$$

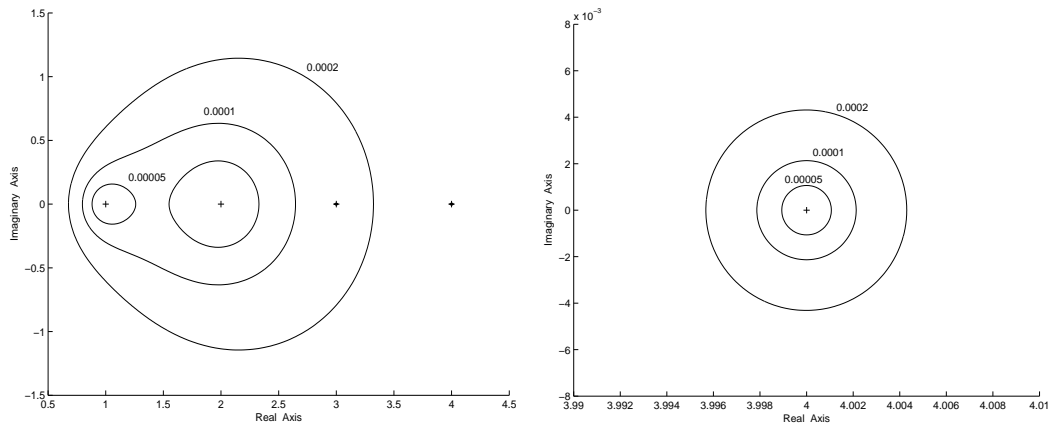


Figure 2: The boundaries $\partial\sigma_\varepsilon(P)$ for $\varepsilon = 5 \cdot 10^{-5}, 10^{-4}, 2 \cdot 10^{-4}$.

The left part of Figure 2 indicates the boundaries of the pseudospectra $\sigma_\varepsilon(P)$ for $\varepsilon = 5 \cdot 10^{-5}, 10^{-4}, 2 \cdot 10^{-4}$. The eigenvalues of $P(\lambda)$ are marked with '+'s. The small components of $\sigma_{5 \cdot 10^{-5}}(P)$, $\sigma_{10^{-4}}(P)$ and $\sigma_{2 \cdot 10^{-4}}(P)$ that correspond to the eigenvalue $\lambda = 4$ are not visible in the left part of the figure, and they are magnified in the right part. Note that these components almost coincide with circular discs centered at $\lambda = 4$ of radii $k(P, 4)\varepsilon$, $\varepsilon = 5 \cdot 10^{-5}, 10^{-4}, 2 \cdot 10^{-4}$, as expected from Proposition 1. It is also apparently confirmed that the eigenvalue $\lambda = 2$ is more sensitive than the eigenvalue $\lambda = 1$ (more particularly, one may say that the eigenvalue $\lambda = 2$ is more than twice as sensitive as $\lambda = 1$), and that both of them are much more sensitive than the eigenvalue $\lambda = 4$.

6 An Elsner-like bound

In this section, we apply the Elsner technique [7] (see also [21]) to obtain a perturbation result for matrix polynomials. This technique allows large perturbations, yielding error bounds, and it does not distinguish between ill-conditioned and well-conditioned eigenvalues.

Theorem 11. *Consider a matrix polynomial $P(\lambda)$ as in (1) and a perturbation $Q(\lambda) \in \mathcal{B}(P, \varepsilon, w)$ as in (3). For any $\mu \in \sigma(Q) \setminus \sigma(P)$, it holds that*

$$\min_{\lambda \in \sigma(P)} |\mu - \lambda| \leq \left(\frac{\varepsilon w(|\mu|)}{|\det A_m|} \right)^{\frac{1}{mn}} \|P(\mu)\|^{1 - \frac{1}{mn}}.$$

Proof. Let $\sigma(P) = \{\lambda_1, \lambda_2, \dots, \lambda_{nm}\}$, counting algebraic multiplicities, and suppose $\mu \in \sigma(Q) \setminus \sigma(P)$. Then

$$\min_{\lambda \in \sigma(P)} |\mu - \lambda|^{nm} \leq \prod_{i=1}^{nm} |\mu - \lambda_i| = \frac{|\det P(\mu)|}{|\det A_m|}.$$

Let now $U = [u_1 \ u_2 \ \dots \ u_n]$ be an $n \times n$ unitary matrix such that $Q(\mu)u_1 = 0$. By Hadamard's inequality [21] (see also [22, Theorem 2.4]), it follows

$$\begin{aligned} \min_{\lambda \in \sigma(P)} |\mu - \lambda|^{nm} &\leq \frac{|\det P(\mu)|}{|\det A_m|} = \frac{|\det P(\mu)| |\det U|}{|\det A_m|} \\ &\leq \frac{1}{|\det A_m|} \prod_{i=1}^{nm} \|P(\mu)u_i\| \\ &= \frac{1}{|\det A_m|} \|P(\mu)u_1\| \prod_{i=2}^{nm} \|P(\mu)u_i\| \\ &= \frac{1}{|\det A_m|} \|P(\mu)u_1 - Q(\mu)u_1\| \prod_{i=2}^{nm} \|P(\mu)u_i\| \\ &\leq \frac{1}{|\det A_m|} \|\Delta(\mu)u_1\| \|P(\mu)\|^{nm-1} \\ &\leq \frac{\varepsilon w(|\mu|)}{|\det A_m|} \|P(\mu)\|^{nm-1}, \end{aligned}$$

and the proof is complete. \square

Recently, the classical Bauer-Fike Theorem [2, 22] has been generalized to the case of matrix polynomials [5]. Applying the arguments of the proof of Theorem 4.1 in [5], it is easy to verify the “weighted version” of the result.

Theorem 12. *Consider a matrix polynomial $P(\lambda)$ as in (1) and a perturbation $Q(\lambda) \in \mathcal{B}(P, \varepsilon, w)$ as in (3), and let (X, J, Y) be a Jordan triple of $P(\lambda)$. For any $\mu \in \sigma(Q) \setminus \sigma(P)$, it holds that $\min_{\lambda \in \sigma(P)} |\mu - \lambda| \leq \max\{\vartheta, \vartheta^{1/p}\}$, where $\vartheta = p k(P) \varepsilon w(|\mu|)$ and p is the maximum dimension of the Jordan blocks of J .*

To compare these two bounds, we consider the matrix polynomial

$$P(\lambda) = I\lambda^3 + \begin{bmatrix} 0 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \lambda$$

(see [5, Example 1] and [9, Example 1.5]) with $\det P(\lambda) = \lambda^2(\lambda + 1)^2(\lambda - 1)^2$. A Jordan triple (X, J, Y) of $P(\lambda)$ is given by

$$X = \begin{bmatrix} 1 & 0 & -\sqrt{2} + 1 & \sqrt{2} - 2 & \sqrt{2} + 1 & \sqrt{2} + 2 \\ 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$$\text{and } Y^T = \frac{1}{4} \begin{bmatrix} 0 & -4 & \sqrt{2} + 2 & -\sqrt{2} - 1 & -\sqrt{2} + 2 & -\sqrt{2} + 1 \\ 4 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

The associated condition number of the eigenproblem of $P(\lambda)$ is $k(P) = 6.4183$. For $\varepsilon = 0.3$ and $w = \{w_0, w_1, w_2, w_3\} = \{0.1, 1, 1, 0\}$, the matrix polynomial

$$Q(\lambda) = I\lambda^3 + \begin{bmatrix} i0.3 & \sqrt{2} \\ \sqrt{2} & -i0.3 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & -0.7 \\ 0.7 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 0.01 & 0 \\ 0 & 0.03 \end{bmatrix}$$

lies on the boundary of $\mathcal{B}(P, 0.3, w)$ and has $\mu = 0.5691 + i0.0043$ as an eigenvalue. Then $\min_{\lambda \in \sigma(P)} |\mu - \lambda| = |0.5691 + i0.0043 - 1| = 0.4309$, the upper bound of Theorem 11 is 0.8554, and the upper bound of Theorem 12 is 3.8240.

It is clear that the Elsner-like upper bound is tighter than the upper bound of Theorem 12 when $\|P(\mu)\|$ is sufficiently small; this is the case in the above example, where $\|P(0.5691 + i0.0043)\| = 1.0562$. In particular, if we define the quantity

$$\Omega(P, \varepsilon, \mu) = \begin{cases} |\det A_m| (p k(P))^{mn} (\varepsilon w(|\mu|))^{mn-1}, & \text{when } p k(P) \varepsilon w(|\mu|) \geq 1 \\ |\det A_m| (p k(P))^{\frac{mn}{p}} (\varepsilon w(|\mu|))^{\frac{mn}{p}-1}, & \text{when } p k(P) \varepsilon w(|\mu|) < 1 \end{cases},$$

then it is straightforward to see that the bound of Theorem 11 is better than the bound of Theorem 12 if and only if $\|P(\mu)\| < \Omega(P, \varepsilon, \mu)^{\frac{1}{mn-1}}$.

References

- [1] A.L. Andrew, K.-W.E. Chu and P. Lancaster, Derivatives of eigenvalues and eigenvectors of matrix functions, *SIAM J. Matrix Anal. Appl.*, 14 (1993) 903–926.
- [2] F.L. Bauer and C.T. Fike, Norms and exclusion theorems, *Numer. Math.*, 2 (1960) 137–144.
- [3] L. Boulton, P. Lancaster and P. Psarrakos, On pseudospectra of matrix polynomials and their boundaries, *Math. Comp.*, 77 (2008) 313–334.
- [4] J.V. Burke, A.S. Lewis and M.L. Overton, Spectral conditioning and pseudospectral growth, *Numer. Math.*, 107 (2007) 27–37.

- [5] K.-W.E. Chu, Perturbation of eigenvalues for matrix polynomials via the Bauer-Fike theorems, *SIAM J. Matrix Anal. Appl.*, 25 (2003) 551–573.
- [6] J.W. Demmel, On condition numbers and the distance to the nearest ill-posed problem, *Numer. Math.*, 51 (1987) 251–289.
- [7] L. Elsner, An optimal bound for the spectral variation of two matrices, *Linear Algebra Appl.*, 71 (1985) 77–80.
- [8] M. Embree and L.N. Trefethen, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.
- [9] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.
- [10] M. Karow, *Geometry of Spectral Valued Sets*, PhD Thesis, University of Bremen, 2003.
- [11] D. Kressner, M.J. Peláez and J. Moro, Structured Hölder condition numbers for multiple eigenvalues, *SIAM J. Matrix Anal. Appl.*, 31 (2009) 175–201.
- [12] P. Lancaster and P. Psarrakos, On the pseudospectra of matrix polynomials, *SIAM J. Matrix Anal. Appl.*, 27 (2005) 115–129.
- [13] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, 2nd edition, Academic Press, Orlando, 1985.
- [14] D.S. Mackey, *Structured Linearizations for Matrix Polynomials*, PhD Thesis, University of Manchester, 2006.
- [15] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, Amer. Math. Society, Providence, Translations of Math. Monographs, Vol. 71, 1988.
- [16] J. Moro, J.V. Burke and M.L. Overton, On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure, *SIAM J. Matrix Anal. Appl.*, 18 (1997) 793–817.
- [17] R.G. Mosier, Root neighbourhoods of a polynomial, *Math. Comp.*, 47 (1986) 265–273.
- [18] N. Papathanasiou and P. Psarrakos, The distance from a matrix polynomial to matrix polynomials with a prescribed multiple eigenvalue, *Linear Algebra Appl.*, 429 (2008) 1453–1477.
- [19] R.A. Smith, The condition numbers of the matrix eigenvalue problem, *Numer. Math.*, 10 (1967) 232–240.
- [20] G.W. Stewart, Error bounds for approximate invariant subspaces of closed linear operators, *SIAM J. Numer. Anal.*, 8 (1971) 796–808.
- [21] G.W. Stewart, An Elsner-like perturbation theorem for generalized eigenvalues, *Linear Algebra Appl.*, 390 (2004) 1–5.
- [22] G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
- [23] F. Tisseur, Backward error and condition of polynomial eigenvalue problems, *Linear Algebra Appl.*, 309 (2000) 339–361.

- [24] F. Tisseur and N.J. Higham, Structured pseudospectra for polynomial eigenvalue problems with applications, *SIAM J. Matrix Anal. Appl.*, 23 (2001) 187–208.
- [25] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.
- [26] J.H. Wilkinson, Note on matrices with a very ill-conditioned eigenproblem, *Numer. Math.*, 19 (1972) 175–178.
- [27] J.H. Wilkinson, On neighbouring matrices with quadratic elementary divisors, *Numer. Math.*, 44 (1984) 1–21.