

A COMPARISON OF EIGENVALUE CONDITION NUMBERS FOR MATRIX POLYNOMIALS*

LUIS MIGUEL ANGUAS[†], MARÍA ISABEL BUENO[‡], AND FROILÁN M. DOPICO[§]

Abstract. In this paper, we consider the different eigenvalue condition numbers for matrix polynomials used in the literature and we compare them. One of these condition numbers is a generalization of the Wilkinson condition number for the standard eigenvalue problem. This number has the disadvantage of only being defined for finite eigenvalues. In order to give a unified approach to all the eigenvalues of a matrix polynomial, both finite and infinite, two (homogeneous) condition numbers have been defined in the literature. In their definition, very different approaches are used. One of the main goals of this note is to show that, when the matrix polynomial has a moderate degree, both homogeneous numbers are essentially the same and one of them provides a geometric interpretation of the other. We also show how the homogeneous condition numbers compare with the “Wilkinson-like” eigenvalue condition number and how they extend this condition number to zero and infinite eigenvalues.

Key words. Eigenvalue condition number, matrix polynomial, chordal distance, eigenvalue.

AMS subject classifications. 15A18, 15A22, 65F15, 65F35

1. Introduction. Let \mathbb{C} denote the field of complex numbers. A square matrix polynomial of grade k can be expressed (in its non-homogeneous form) as

$$P(\lambda) = \sum_{i=0}^k \lambda^i B_i, \quad B_i \in \mathbb{C}^{n \times n}, \quad (1.1)$$

where the matrix coefficients, including B_k , are allowed to be the zero matrix. In particular, when $B_k \neq 0$, we say that $P(\lambda)$ has degree k . Throughout the paper, the grade of every matrix polynomial $P(\lambda)$ will be assumed to be its degree unless it is specified otherwise.

The (non-homogeneous) polynomial eigenvalue problem (PEP) associated with a regular matrix polynomial $P(\lambda)$, (that is, $\det(P(\lambda)) \neq 0$) consists of finding scalars $\lambda_0 \in \mathbb{C}$ and nonzero vectors $x, y \in \mathbb{C}^n$ satisfying

$$P(\lambda_0)x = 0 \text{ and } y^*P(\lambda_0) = 0.$$

The vectors x and y are called, respectively, a right and a left eigenvector of $P(\lambda)$ corresponding to the eigenvalue λ_0 . In addition, $P(\lambda)$ may have infinite eigenvalues. We say that $P(\lambda)$ has an infinite eigenvalue if 0 is an eigenvalue of the reversal of $P(\lambda)$, where the reversal of a matrix polynomial $P(\lambda)$ of grade k is defined as

$$rev(P(\lambda)) := \lambda^k P(1/\lambda). \quad (1.2)$$

The numerical solution of the PEP has received considerable attention from many research groups in the last two decades and, as a consequence, several condition

*This work was partially supported by the Ministerio de Economía, Industria y Competitividad (MINECO) of Spain through grants MTM2015-65798-P and MTM2017-90682-REDT. The research of L. M. Anguas is funded by the “contrato predoctoral” BES-2013-065688 of MINECO.

[†]Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (languas@math.uc3m.es).

[‡]Department of Mathematics and College of Creative Studies, University of California, Santa Barbara, CA 93106, USA (mbueno@math.ucsb.edu),

[§]Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain (dopico@math.uc3m.es).

numbers for simple eigenvalues of a matrix polynomial $P(\lambda)$ have been defined in the literature to determine the sensitivity of these eigenvalues to perturbations in the coefficients of $P(\lambda)$ [2, 3, 8]. One of these condition numbers is a natural generalization of the Wilkinson condition number for the standard eigenproblem (see Definition 2.1). A disadvantage of this eigenvalue condition number is that it is not defined for infinite eigenvalues. Then, in order to study the conditioning of all the eigenvalues of a matrix polynomial, other condition numbers are considered in the literature. These condition numbers assume that the matrix polynomial is expressed in homogeneous form, that is,

$$P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i. \quad (1.3)$$

If $P(\alpha, \beta)$ is regular, we can consider the corresponding homogeneous PEP that consists in finding pairs of scalars $(\alpha_0, \beta_0) \neq (0, 0)$ and nonzero vectors $x, y \in \mathbb{C}^n$ such that

$$P(\alpha_0, \beta_0)x = 0 \quad \text{and} \quad y^* P(\alpha_0, \beta_0) = 0. \quad (1.4)$$

We note that the pairs (α_0, β_0) satisfying (1.4) are those for which $\det(P(\alpha_0, \beta_0)) = 0$ holds. Notice that (α_0, β_0) satisfies $\det(P(\alpha_0, \beta_0)) = 0$ if and only if $\det(P(c\alpha_0, c\beta_0)) = 0$ for any nonzero complex number c . Therefore, it is natural to define an eigenvalue of $P(\alpha, \beta)$ as any line in \mathbb{C}^2 passing through the origin consisting of solutions of $\det(P(\alpha, \beta)) = 0$. For simplicity, we denote such a line, i.e., an eigenvalue of $P(\alpha, \beta)$, as (α_0, β_0) and by $[\alpha_0, \beta_0] \neq (0, 0)$ a specific representative of this eigenvalue. The vectors x, y in (1.4) are called, respectively, a right and a left eigenvector of $P(\alpha, \beta)$ corresponding to the eigenvalue (α_0, β_0) . For $\beta \neq 0$, we can define $\lambda = \alpha/\beta$ and find a relationship between the homogeneous and the non-homogeneous expressions of a matrix polynomial P of degree k as follows:

$$P(\alpha, \beta) = \beta^k P(\lambda).$$

We note also that, if (y, λ_0, x) is a solution of the non-homogeneous PEP, i.e., an eigentriple of the non-homogeneous PEP, then $(y, (\alpha_0, \beta_0), x)$ is a solution of the corresponding homogeneous PEP, for any $[\alpha_0, \beta_0] \neq (0, 0)$ such that $\lambda_0 = \alpha_0/\beta_0$, including $\lambda_0 = \infty$ for $\beta_0 = 0$.

Two homogeneous eigenvalue condition numbers (well defined for all the eigenvalues of $P(\alpha, \beta)$, finite and infinite) have been presented in the literature. One of them is a natural generalization of the condition number defined by Stewart and Sun in [7, Chapter VI, Section 2.1] for the eigenvalues of a pencil. This condition number is defined in terms of the chordal distance between two lines in \mathbb{C}^2 (see Definition 2.12).

The other homogeneous eigenvalue condition number is defined as the norm of a differential map [2, 3]. The definition of this condition number is very involved and less intuitive than the definition of the other condition numbers that we consider in this paper. For an explicit formula for this condition number, see Theorem 2.6.

In this paper we address the following natural questions:

- how are the two homogeneous eigenvalue condition numbers related? Are they equivalent?
- if λ_0 is a finite nonzero eigenvalue of a matrix polynomial $P(\lambda)$ and (α_0, β_0) is the associated eigenvalue of $P(\alpha, \beta)$ (that is, $\lambda_0 = \alpha_0/\beta_0$), how are the

(non-homogeneous) absolute and relative eigenvalue condition numbers of λ_0 and the (homogeneous) condition numbers of (α_0, β_0) related? Are these two types of condition numbers equivalent in the sense that λ_0 is ill-conditioned if and only if (α_0, β_0) is ill-conditioned?

Partial answers to these questions are scattered in the literature written in an implicit way so that they seem to be unnoticed by most researchers in Linear Algebra. Our goal is to present a complete and explicit answer to these questions. More precisely, we provide an exact relationship between the two homogeneous eigenvalue condition numbers and we use this relationship to prove that they are equivalent. Also, we obtain exact relationships between each of the non-homogeneous (relative and absolute) and the homogeneous eigenvalue condition numbers. From these relationships we prove that the non-homogeneous condition numbers are always larger than the homogeneous condition numbers. This means that non-homogeneous eigenvalues λ_0 are always more sensitive to perturbations than the corresponding homogeneous ones (α_0, β_0) , which is natural since $\lambda_0 = \alpha_0/\beta_0$. Moreover, we will see that non-homogeneous eigenvalues with large or small moduli have much larger non-homogeneous than homogeneous condition numbers. Thus, in these cases, (α_0, β_0) can be very well-conditioned and λ_0 very ill-conditioned. In the context of this discussion, it is important to bear in mind that in most applications of PEPs the quantities of interest are the non-homogeneous eigenvalues, and not the homogeneous ones.

The paper is organized as follows: Section 2 includes the definitions and expressions of the different condition numbers that are used in this work. In section 3, we establish relationships between the condition numbers introduced in section 2, and in section 4 we present a geometric interpretation of these relationships. Section 4 also includes a study of the computability of small and large eigenvalues. Finally, some conclusions are discussed in section 5.

2. Eigenvalue condition numbers of matrix polynomials. In this section we recall three eigenvalue condition numbers used in the literature and discuss some of the advantages and disadvantages of each of them. Before recalling their definition, we present some notation that will be used throughout the paper.

Let a and b be two integers. We define

$$a : b = \begin{cases} a, a + 1, a + 2, \dots, b, & \text{if } a \leq b, \\ \emptyset, & \text{if } a > b. \end{cases}$$

For any matrix A , $\|A\|_2$ denotes its spectral or 2-norm, i.e., its largest singular value [7]. For any vector x , $\|x\|_2$ denotes its standard Euclidean norm, i.e., $\|x\|_2 = (x^*x)^{1/2}$, where the operator $()^*$ stands for the conjugate-transpose of x .

2.1. Non-homogeneous eigenvalue condition numbers. Next we recall the definition of two versions (absolute and relative) of a normwise eigenvalue condition number introduced in [8].

DEFINITION 2.1. *Let λ_0 be a simple, finite eigenvalue of a regular matrix polynomial $P(\lambda) = \sum_{i=0}^k \lambda^i B_i$ of grade k and let x be a right eigenvector of $P(\lambda)$ associated with λ_0 . We define the normwise absolute condition number $\kappa_a(\lambda_0, P)$ of λ_0 by*

$$\kappa_a(\lambda_0, P) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda_0|}{\epsilon} : [P(\lambda_0 + \Delta\lambda_0) + \Delta P(\lambda_0 + \Delta\lambda_0)](x + \Delta x) = 0, \right. \\ \left. \|\Delta B_i\|_2 \leq \epsilon \omega_i, i = 0 : k \right\},$$

where $\Delta P(\lambda) = \sum_{i=0}^k \lambda^i \Delta B_i$ and ω_i , $i = 0 : k$, are nonnegative weights that allow flexibility in how the perturbations of $P(\lambda)$ are measured.

For $\lambda_0 \neq 0$, we define the normwise relative condition number $\kappa_r(\lambda_0, P)$ of λ_0 by

$$\kappa_r(\lambda_0, P) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta \lambda_0|}{\epsilon |\lambda_0|} : [P(\lambda_0 + \Delta \lambda_0) + \Delta P(\lambda_0 + \Delta \lambda_0)](x + \Delta x) = 0, \right. \\ \left. \|\Delta B_i\|_2 \leq \epsilon \omega_i, i = 0 : k \right\}.$$

We will refer to $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$, respectively, as the *absolute and relative non-homogeneous eigenvalue condition numbers*. Note that the absolute non-homogeneous condition number is not defined for infinite eigenvalues while the relative condition number is not defined for zero or infinite eigenvalues.

REMARK 2.2. In the definitions of $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$, the weights ω_i can be chosen in different ways. The most common ways are: 1) $\omega_i = \|B_i\|_2$ (relative coefficient-wise perturbations); 2) $\omega_i = \max_{i=0:k} \{\|B_i\|_2\}$ (relative perturbations with respect to the norm of $P(\lambda)$); 3) $\omega_i = 1$ (absolute perturbations).

An explicit formula for each of the non-homogeneous eigenvalue condition numbers presented above was obtained by Tisseur in [8].

THEOREM 2.3. *Let $P(\lambda)$ be a regular matrix polynomial of grade k . Let λ_0 be a simple, finite eigenvalue of $P(\lambda)$ and let x and y be, respectively, a right and a left eigenvector of $P(\lambda)$ associated with λ_0 . Then,*

$$\kappa_a(\lambda_0, P) = \frac{(\sum_{i=0}^k |\lambda_0|^i \omega_i) \|y\|_2 \|x\|_2}{|y^* P'(\lambda_0) x|},$$

where $P'(\lambda)$ denotes the derivative of $P(\lambda)$ with respect to λ . For $\lambda_0 \neq 0$,

$$\kappa_r(\lambda_0, P) = \frac{(\sum_{i=0}^k |\lambda_0|^i \omega_i) \|y\|_2 \|x\|_2}{|\lambda_0| |y^* P'(\lambda_0) x|}.$$

The following technical result will be useful for the comparison of the non-homogeneous condition numbers introduced above and the homogeneous condition numbers that we introduce in the next subsection. We will use the concept of reversal of a matrix polynomial defined in (1.2).

LEMMA 2.4. *Let $P(\lambda) = \sum_{i=0}^k \lambda^i B_i$ be a regular matrix polynomial. Let λ_0 be a simple, nonzero, finite eigenvalue of $P(\lambda)$. Then,*

$$\kappa_a\left(\frac{1}{\lambda_0}, revP\right) = \frac{\kappa_r(\lambda_0, P)}{|\lambda_0|} \quad \text{and} \quad \kappa_r\left(\frac{1}{\lambda_0}, revP\right) = \kappa_r(\lambda_0, P).$$

Proof. We only prove the first claim. The second claim follows immediately from the first.

Let x and y be, respectively, a right and a left eigenvector of $P(\lambda)$ associated with λ_0 . It is easy to see that these vectors are also a right and a left eigenvector of $revP$

associated with $\frac{1}{\lambda_0}$. Notice that

$$\begin{aligned} \kappa_a \left(\frac{1}{\lambda_0}, \text{rev}P \right) &= \frac{(\sum_{i=0}^k \left| \frac{1}{\lambda_0} \right|^{k-i} \omega_i) \|x\|_2 \|y\|_2}{|y^*(\text{rev}P)'(\frac{1}{\lambda_0})x|} \\ &= \frac{(\sum_{i=0}^k |\lambda_0|^i \omega_i) \|x\|_2 \|y\|_2}{|\lambda_0|^k |y^*(\text{rev}P)'(\frac{1}{\lambda_0})x|} = \frac{(\sum_{i=0}^k |\lambda_0|^i \omega_i) \|x\|_2 \|y\|_2}{|\lambda_0|^2 |y^* \lambda_0^{k-2} (\text{rev}P)'(\frac{1}{\lambda_0})x|} \\ &= \frac{(\sum_{i=0}^k |\lambda_0|^i \omega_i) \|x\|_2 \|y\|_2}{|\lambda_0|^2 |y^* P'(\lambda_0)x|} = \frac{\kappa_r(\lambda_0, P)}{|\lambda_0|}, \end{aligned}$$

where the fourth equality follows from the facts that $\text{rev}P(\lambda) = \lambda^k P(\frac{1}{\lambda})$ and $P(\lambda_0)x = 0$, and the first and fifth equalities follow from Theorem 2.3. Thus, the claim follows. \square

2.2. Homogeneous eigenvalue condition numbers. As pointed out in the last subsection, neither of the non-homogeneous condition numbers is defined for infinite eigenvalues. Thus, these type of eigenvalues require a special treatment in the non-homogeneous setting. In this section we introduce two condition numbers that allow a unified approach to all eigenvalues, finite and infinite. These condition numbers require the matrix polynomial to be expressed in homogeneous form (see (1.3)). This is the reason why we refer to them as homogeneous eigenvalue condition numbers.

REMARK 2.5. We recall that $(\alpha_0, \beta_0) \neq (0, 0)$ is an eigenvalue of $P(\alpha, \beta)$ if and only if $\lambda_0 := \alpha_0/\beta_0$ is an eigenvalue of $P(\lambda)$, where $\lambda_0 = \infty$ if $\beta_0 = 0$.

Each of the condition numbers presented in this subsection has been defined in the literature with a different approach. One of them is due to Stewart and Sun [7], who define the eigenvalue condition number in terms of the chordal distance between the exact and the perturbed eigenvalues. The other approach is due to Dedieu and Tisseur [2, 3] and makes use of the Implicit Function Theorem to construct a differential operator whose norm is defined to be an eigenvalue condition number. This idea was inspired by Shub and Smale's work [6]. These two condition numbers do not have a specific name in the literature. We will refer to them as the Stewart-Sun condition number and the Dedieu-Tisseur condition number, respectively.

2.2.1. Dedieu-Tisseur condition number. The homogeneous eigenvalue condition number that we present in this section has been often used in recent literature on matrix polynomials as an alternative to the non-homogeneous Wilkinson-like condition number. See, for instance, [4, 5]. We do not include its explicit definition because it is much more involved than Definition 2.1. For the interested reader, the definition can be found in [3]. The next theorem provides an explicit formula for this condition number.

THEOREM 2.6. [3, Theorem 4.2] *Let (α_0, β_0) be a simple eigenvalue of the regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$, and let y and x be, respectively, a left and a right eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0) . Then, the Dedieu-Tisseur condition number of (α_0, β_0) is given by*

$$\kappa_h((\alpha_0, \beta_0), P) = \left(\sum_{i=0}^k |\alpha_0|^{2i} |\beta_0|^{2(k-i)} \omega_i^2 \right)^{1/2} \frac{\|y\|_2 \|x\|_2}{|y^*(\beta_0 D_\alpha P(\alpha_0, \beta_0) - \alpha_0 D_\beta P(\alpha_0, \beta_0))x|}, \quad (2.1)$$

where $D_a \equiv \frac{\partial}{\partial a}$, that is, the partial derivative with respect to a and ω_i , $i = 0 : k$ are nonnegative weights that define how the perturbations of the coefficients B_i are measured.

It is important to note that the expression for this eigenvalue condition number does not depend on the choice of representative for the eigenvalue (α_0, β_0) .

2.2.2. Stewart-Sun condition number. Here we introduce another homogeneous eigenvalue condition number. Its definition is easy to convey and to interpret from a geometrical point of view.

We recall that every eigenvalue of a homogeneous matrix polynomial can be seen as a line in \mathbb{C}^2 passing through the origin. The condition number that we present here uses the ‘‘chordal distance’’ between lines in \mathbb{C}^2 to measure the distance between an eigenvalue and a perturbed eigenvalue. This distance is defined on the projective space $\mathbb{P}_1(\mathbb{C})$.

Before introducing the chordal distance, we recall the definition of angle between two lines.

DEFINITION 2.7. *Let x and y be two nonzero vectors in \mathbb{C}^2 and let $\langle x \rangle$ and $\langle y \rangle$ denote the lines passing through zero in the direction of x and y , respectively. We define the angle between the two lines $\langle x \rangle$ and $\langle y \rangle$ by*

$$\theta(\langle x \rangle, \langle y \rangle) := \arccos \frac{|\langle x, y \rangle|}{\|x\|_2 \|y\|_2}, \quad 0 \leq \theta(\langle x \rangle, \langle y \rangle) \leq \pi/2,$$

where $\langle x, y \rangle$ denotes the standard Hermitian inner product, i.e., $\langle x, y \rangle = y^* x$.

REMARK 2.8. We note that $\cos \theta(\langle x \rangle, \langle y \rangle)$ can be seen as the ratio between the length of the orthogonal projection (with respect to the standard inner product in \mathbb{C}^2) of the vector x onto y to the length of the vector x itself, that is,

$$\cos \theta(\langle x \rangle, \langle y \rangle) = \frac{\|proj_y x\|_2}{\|x\|_2},$$

since

$$proj_y x = \frac{\langle x, y \rangle y}{\|y\|_2^2} \quad \text{and} \quad \frac{\|proj_y x\|_2}{\|x\|_2} = \frac{|\langle x, y \rangle|}{\|x\|_2 \|y\|_2}.$$

We also have

$$\sin \theta(\langle x \rangle, \langle y \rangle) = \frac{\|x - proj_y x\|_2}{\|x\|_2}.$$

The definition of chordal distance is given next.

DEFINITION 2.9. [7, Chapter VI, Definition 1.20] *Let x and y be two nonzero vectors in \mathbb{C}^2 and let $\langle x \rangle$ and $\langle y \rangle$ denote the lines passing through zero in the direction of x and y , respectively. The chordal distance between $\langle x \rangle$ and $\langle y \rangle$ is given by*

$$\chi(\langle x \rangle, \langle y \rangle) := \sin(\theta(\langle x \rangle, \langle y \rangle)).$$

Notice that $\chi(\langle x \rangle, \langle y \rangle) \leq \theta(\langle x \rangle, \langle y \rangle)$. Moreover, the chordal distance and the angle are identical asymptotically, that is, when $\theta(\langle x \rangle, \langle y \rangle)$ approaches 0.

The chordal distance between two lines $\langle x \rangle$ and $\langle y \rangle$ in \mathbb{C}^2 can also be expressed in terms of the coordinates of the vectors x and y in the canonical basis for \mathbb{C}^2 . Note that this expression does not depend on the representatives x and y of the lines.

LEMMA 2.10. [7, page 283] *If $\langle \alpha, \beta \rangle$ and $\langle \gamma, \delta \rangle$ are two lines in \mathbb{C}^2 , then*

$$\chi(\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle) = \frac{|\alpha\delta - \beta\gamma|}{\sqrt{|\alpha|^2 + |\beta|^2} \sqrt{|\gamma|^2 + |\delta|^2}}.$$

REMARK 2.11. Notice that $0 \leq \chi(\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle) \leq 1$ for all lines $\langle \alpha, \beta \rangle, \langle \gamma, \delta \rangle$ in \mathbb{C}^2 . Since the line $\langle 1, 0 \rangle$ is identified with the eigenvalue ∞ in PEPs, we see that the chordal distance allows us to measure the distance from ∞ to any other eigenvalue very easily. Moreover, such distance is never larger than one.

Next we introduce the homogeneous eigenvalue condition number in which the change in the eigenvalue is measured using the chordal distance and that we baptize as the Stewart-Sun eigenvalue condition number. This condition number was implicitly introduced for matrix pencils in [7, page 294], although an explicit definition is not given in [7]. See also [1, page 40] for an explicit definition of this condition number for matrix polynomials.

Note that in Definition 2.12 below, $(\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0)$ is the unique simple eigenvalue of $(P + \Delta P)(\alpha, \beta)$ that approaches (α_0, β_0) when ΔP approaches zero.

DEFINITION 2.12. *Let (α_0, β_0) be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let x be a right eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0) . We define*

$$\begin{aligned} \kappa_\theta((\alpha_0, \beta_0), P) := \limsup_{\epsilon \rightarrow 0} & \left\{ \frac{\chi((\alpha_0, \beta_0), (\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0))}{\epsilon} : \right. \\ & [P(\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0) + \Delta P(\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0)](x + \Delta x) = 0, \\ & \left. \|\Delta B_i\|_2 \leq \epsilon \omega_i, i = 0 : k \right\}, \end{aligned}$$

where $\Delta P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} \Delta B_i$ and $\omega_i, i = 0 : k$, are nonnegative weights that allow flexibility in how the perturbations of $P(\alpha, \beta)$ are measured.

As far as we know, no explicit formula for the Stewart-Sun condition number is available in the literature. We provide such an expression next.

THEOREM 2.13. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$, and let x and y be, respectively, a right and a left eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0) . Then,*

$$\kappa_\theta((\alpha_0, \beta_0), P) = \left(\sum_{i=0}^k |\alpha_0|^i |\beta_0|^{k-i} \omega_i \right) \frac{\|y\|_2 \|x\|_2}{|y^* (\overline{\beta_0} D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0} D_\beta P(\alpha_0, \beta_0)) x|}, \quad (2.2)$$

where $D_a \equiv \frac{\partial}{\partial a}$, that is, the partial derivative with respect to a .

Proof. Let $(P + \Delta P)(\alpha, \beta)$ be a small enough perturbation of $P(\alpha, \beta)$. Then, $(P + \Delta P)(\alpha, \beta)$ has a unique simple eigenvalue $(\tilde{\alpha}_0, \tilde{\beta}_0)$, with associated eigenvector $x + \Delta x$, that approaches (α_0, β_0) when $\Delta P(\alpha, \beta)$ approaches zero. Since $\{[\alpha_0, \beta_0], [-\overline{\beta_0}, \overline{\alpha_0}]\}$ is an orthogonal basis for \mathbb{C}^2 (with respect to the standard inner product), we can choose a representative for $(\tilde{\alpha}_0, \tilde{\beta}_0)$ of the form

$$[\alpha_0, \beta_0] + [\Delta\alpha_0, \Delta\beta_0],$$

where $\langle [\alpha_0, \beta_0], [\Delta\alpha_0, \Delta\beta_0] \rangle = 0$ for any given representative $[\alpha_0, \beta_0]$ of (α_0, β_0) . This implies that there exists a scalar \mathfrak{h} such that

$$[\Delta\alpha_0, \Delta\beta_0] = \mathfrak{h}[-\overline{\beta_0}, \overline{\alpha_0}]. \quad (2.3)$$

Expanding for these representatives the left hand side of the constraint

$$[P(\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0) + \Delta P(\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0)](x + \Delta x) = 0$$

in the definition of $\kappa_\theta((\alpha_0, \beta_0), P)$ and keeping only the first order terms, we get

$$[D_\alpha P(\alpha_0, \beta_0)\Delta\alpha_0 + D_\beta P(\alpha_0, \beta_0)\Delta\beta_0]x + P(\alpha_0, \beta_0)\Delta x + \Delta P(\alpha_0, \beta_0)x = O(\epsilon^2).$$

If we multiply the previous equation by y^* on the left, taking into account that y is a left eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0) , we get

$$y^*[D_\alpha P(\alpha_0, \beta_0)\Delta\alpha_0 + D_\beta P(\alpha_0, \beta_0)\Delta\beta_0]x + y^*\Delta P(\alpha_0, \beta_0)x = O(\epsilon^2). \quad (2.4)$$

Using (2.3), (2.4) becomes

$$y^*[D_\alpha P(\alpha_0, \beta_0)\mathfrak{h}\overline{\beta_0} - D_\beta P(\alpha_0, \beta_0)\mathfrak{h}\overline{\alpha_0}]x - y^*\Delta P(\alpha_0, \beta_0)x = O(\epsilon^2).$$

Since $(\alpha_0, \beta_0) \neq 0$ is a simple eigenvalue, by [3, Theorem 3.3],

$$y^*[D_\alpha P(\alpha_0, \beta_0)\overline{\beta_0} - D_\beta P(\alpha_0, \beta_0)\overline{\alpha_0}]x \neq 0.$$

Therefore,

$$\mathfrak{h} = \frac{y^*\Delta P(\alpha_0, \beta_0)x}{y^*[\overline{\beta_0}D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0}D_\beta P(\alpha_0, \beta_0)]x} + O(\epsilon^2). \quad (2.5)$$

On the other hand,

$$\begin{aligned} \frac{\chi((\alpha_0, \beta_0), (\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0))}{\epsilon} &= \frac{|\alpha_0\Delta\beta_0 - \beta_0\Delta\alpha_0|}{\epsilon\sqrt{|\alpha_0|^2 + |\beta_0|^2}\sqrt{|\alpha_0 + \Delta\alpha_0|^2 + |\beta_0 + \Delta\beta_0|^2}} \\ &= \frac{|\mathfrak{h}|}{\epsilon} \frac{\sqrt{|\alpha_0|^2 + |\beta_0|^2}}{\sqrt{|\alpha_0 + \Delta\alpha_0|^2 + |\beta_0 + \Delta\beta_0|^2}}. \end{aligned} \quad (2.6)$$

Since, by (2.5),

$$\frac{|\mathfrak{h}|}{\epsilon} \leq \left(\sum_{i=0}^k |\alpha_0|^i |\beta_0|^{k-i} \omega_i \right) \frac{\|y\|_2 \|x\|_2}{|y^*(\overline{\beta_0}D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0}D_\beta P(\alpha_0, \beta_0))x|} + O(\epsilon) \quad (2.7)$$

and $\Delta\alpha_0$ and $\Delta\beta_0$ approach zero as $\epsilon \rightarrow 0$, from (2.6) and (2.7) we get

$$\kappa_\theta((\alpha_0, \beta_0), P) \leq \left(\sum_{i=0}^k |\alpha_0|^i |\beta_0|^{k-i} \omega_i \right) \frac{\|y\|_2 \|x\|_2}{|y^*(\overline{\beta_0}D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0}D_\beta P(\alpha_0, \beta_0))x|}.$$

Now we need to show that this upper bound on the Stewart-Sun condition number can be attained. Let

$$\Delta B_i = \operatorname{sgn}(\overline{\alpha_0}^{-i}) \operatorname{sgn}(\overline{\beta_0}^{k-i}) \epsilon \omega_i \frac{yx^*}{\|x\|_2 \|y\|_2}, \quad i = 0 : k,$$

where $\operatorname{sgn}(z) = \frac{z}{|z|}$ if $z \neq 0$ and $\operatorname{sgn}(0) = 0$. Note that, with this definition of ΔB_i , we have

$$\|\Delta B_i\|_2 = \epsilon \omega_i, \quad i = 0 : k, \quad \text{and} \quad |y^*\Delta P(\alpha_0, \beta_0)x| = \epsilon \left(\sum_{i=0}^k |\alpha_0|^i |\beta_0|^{k-i} \omega_i \right) \|x\|_2 \|y\|_2.$$

Thus, the inequality in (2.7) becomes an equality and the result follows. \square

3. Comparisons of eigenvalue condition numbers of matrix polynomials. In this section we provide first a comparison between the Dedieu-Tisseur and Stewart-Sun homogeneous condition numbers and, as a consequence, we prove that these condition numbers are equivalent up to a moderate constant depending only on the degree of the polynomial. Then we compare the Stewart-Sun condition number with the non-homogeneous condition number in both its absolute and relative version; as a result, we see that these condition numbers can be very different in certain situations. A simple geometric interpretation of these differences is given in Section 4. In the literature some comparisons can be found, as we will point out, but they provide inequalities among the condition numbers while our expressions are equalities.

3.1. Comparison of the Dedieu-Tisseur and Stewart-Sun condition numbers. As mentioned earlier, the Dedieu-Tisseur and the Stewart-Sun homogeneous condition numbers are defined following a very different approach. So it is a natural question to determine how they are related. We start with a result known in the literature.

THEOREM 3.1. [2, Section 7][1, Corollary 2.6] *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k . Assuming that the same weights ω_i are considered for both condition numbers, we have*

$$\kappa_\theta((\alpha_0, \beta_0), P) \leq C \kappa_h((\alpha_0, \beta_0), P),$$

for some constant C .

To provide an exact relationship between the Dedieu-Tisseur and the Stewart-Sun homogeneous condition numbers, we simply use the explicit formulas given for them in Theorems 2.6 and 2.13, respectively.

THEOREM 3.2. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$. Then,*

$$\kappa_h((\alpha_0, \beta_0), P) = \frac{\left(\sum_{i=0}^k |\alpha_0|^{2i} |\beta_0|^{2(k-i)} \omega_i^2 \right)^{1/2}}{\sum_{i=0}^k |\alpha_0|^i |\beta_0|^{k-i} \omega_i} \kappa_\theta((\alpha_0, \beta_0), P).$$

The following result is an immediate consequence of Theorem 3.2 and shows that, for moderate k , both homogeneous condition numbers are essentially the same.

COROLLARY 3.3. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$. Then,*

$$\frac{1}{\sqrt{k+1}} \leq \frac{\kappa_h((\alpha_0, \beta_0), P)}{\kappa_\theta((\alpha_0, \beta_0), P)} \leq 1.$$

Proof. Let $v := [|\beta_0|^k \omega_0, |\alpha_0| |\beta_0|^{k-1} \omega_1, \dots, |\alpha_0|^k \omega_k]^T \in \mathbb{R}^{k+1}$. From Theorem 3.2, we have

$$\frac{\kappa_h((\alpha_0, \beta_0), P)}{\kappa_\theta((\alpha_0, \beta_0), P)} = \frac{\|v\|_2}{\|v\|_1},$$

where $\|\cdot\|_1$ denotes the vector 1-norm [7]. The result follows taking into account the fact that $1/\sqrt{k+1} \leq \|v\|_2/\|v\|_1 \leq 1$. \square

Since the Dedieu-Tisseur and the Stewart-Sun condition numbers are equivalent and the definition of the Stewart-Sun condition number is much simpler and intuitive,

we do not see any advantage in using the Dedieu-Tisseur condition number. Therefore, in the next subsection, we focus on comparing the Stewart-Sun condition number with the non-homogeneous condition numbers. The corresponding comparisons with the Dedieu-Tisseur condition number follow immediately from Theorem 3.2 and Corollary 3.3.

3.2. Comparison of the homogeneous and non-homogeneous eigenvalue condition numbers. As we mentioned in Section 2, the main drawback of the non-homogeneous condition numbers is that they do not allow a unified treatment of all the eigenvalues of a matrix polynomial since these condition numbers are not defined for the infinite eigenvalues. Thus, some researchers prefer to use a homogeneous eigenvalue condition number instead, although in most applications the non-homogeneous eigenvalues λ are the relevant quantities. In this section, we give an algebraic relationship between the Stewart-Sun condition number and the non-homogeneous (absolute and relative) eigenvalue condition number. We emphasize again that, by Theorem 3.2 and Corollary 3.3, this relation also provides us with a relation between the Dedieu-Tisseur condition number and the non-homogeneous condition numbers.

We start with the only result we have found in the literature on this topic.

THEOREM 3.4. [1, Corollary 2.7] *Let (α_0, β_0) with $\alpha_0 \neq 0$ and $\beta_0 \neq 0$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$. Assuming that the same weights ω_i are considered for both condition numbers, there exists a constant C such that*

$$\kappa_r(\lambda_0, P) \leq C \frac{1 + |\lambda_0|^2}{|\lambda_0|} \kappa_h((\alpha_0, \beta_0), P).$$

The next theorem is the main result in this section and provides exact relationships between the Stewart-Sun condition number and the non-homogeneous condition numbers. Note that in Theorem 3.5 an eigenvalue condition number of the reversal of a matrix polynomial is used, more precisely, $\kappa_a((1/\lambda_0), revP)$. For $\lambda_0 = \infty$, this turns into $\kappa_a(0, revP)$ which can be interpreted as a non-homogeneous absolute condition number of $\lambda_0 = \infty$.

THEOREM 3.5. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, where $\lambda_0 = \infty$ if $\beta_0 = 0$. Assume that the same weights ω_i are considered in the definition of all the condition numbers appearing below, i.e., $\|\Delta B_i\|_2 \leq \epsilon \omega_i, i = 0 : k$, in all of them. Then,*

(i) if $\beta_0 \neq 0$,

$$\kappa_h((\alpha_0, \beta_0), P) = \kappa_a(\lambda_0, P) \frac{1}{1 + |\lambda_0|^2}; \quad (3.1)$$

(ii) if $\alpha_0 \neq 0$,

$$\kappa_\theta((\alpha_0, \beta_0), P) = \kappa_a\left(\frac{1}{\lambda_0}, revP\right) \frac{1}{1 + \left|\frac{1}{\lambda_0}\right|^2}; \quad (3.2)$$

(iii) if $\alpha_0 \neq 0$ and $\beta_0 \neq 0$,

$$\kappa_\theta((\alpha_0, \beta_0), P) = \kappa_r(\lambda_0, P) \frac{|\lambda_0|}{1 + |\lambda_0|^2}. \quad (3.3)$$

Proof. Assume first that $\beta_0 \neq 0$, which implies that λ_0 is finite. Let $(\tilde{\alpha}_0, \tilde{\beta}_0) := (\alpha_0 + \Delta\alpha_0, \beta_0 + \Delta\beta_0)$ be a perturbation of (α_0, β_0) small enough so that $\beta_0 + \Delta\beta_0 \neq 0$, and let $\lambda_0 + \Delta\lambda_0 := \frac{\tilde{\alpha}_0}{\tilde{\beta}_0}$. Note that

$$\Delta\lambda_0 = \frac{\tilde{\alpha}_0\beta_0 - \alpha_0\tilde{\beta}_0}{\beta_0\tilde{\beta}_0} = \frac{\beta_0\Delta\alpha_0 - \alpha_0\Delta\beta_0}{\beta_0(\beta_0 + \Delta\beta_0)}. \quad (3.4)$$

Then, we have, by Lemma 2.10,

$$\begin{aligned} \chi((\alpha_0, \beta_0), (\tilde{\alpha}_0, \tilde{\beta}_0)) &= \frac{|\alpha_0\Delta\beta_0 - \beta_0\Delta\alpha_0|}{|\beta_0|\sqrt{1+|\lambda_0|^2}|\beta_0 + \Delta\beta_0|\sqrt{1+|\lambda_0 + \Delta\lambda_0|^2}} \\ &= \frac{|\Delta\lambda_0|}{\sqrt{1+|\lambda_0|^2}\sqrt{1+|\lambda_0 + \Delta\lambda_0|^2}} \end{aligned} \quad (3.5)$$

where the second equality follows from (3.4). Then, as $\epsilon \rightarrow 0$ in the definition of the condition numbers (which implies that $|\Delta\alpha_0|$ and $|\Delta\beta_0|$ approach 0 as well, using a continuity argument), we have $\Delta\lambda_0 \rightarrow 0$. Thus, bearing in mind Definitions 2.1 and 2.12, (3.1) follows from (3.5).

Assume now that $\alpha_0 \neq 0$ and $\beta_0 \neq 0$, that is, $\lambda_0 \neq 0$ and λ_0 is not an infinite eigenvalue. Then, from (3.5), we get

$$\chi((\alpha_0, \beta_0), (\tilde{\alpha}_0, \tilde{\beta}_0)) = \frac{|\Delta\lambda_0|}{|\lambda_0|} \frac{|\lambda_0|}{\sqrt{1+|\lambda_0|^2}\sqrt{1+|\lambda_0 + \Delta\lambda_0|^2}}.$$

This implies (3.3). By Lemma 2.4, (3.2) follows from (3.3) for finite, nonzero eigenvalues. It only remains to prove (3.2) for $\lambda_0 = \infty$, which corresponds to $(\alpha_0, \beta_0) = (1, 0)$. This follows from Theorem 2.3 applied to $revP$ and Theorem 2.13, which yield:

$$\kappa_a(0, revP) = \frac{\omega_k \|x\|_2 \|y\|_2}{|y^* B_{k-1} x|} = \kappa_\theta((1, 0), P).$$

□

From Theorem 3.5, we immediately get

$$\kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_r(\lambda_0, P) \quad \text{for } 0 < |\lambda_0| < \infty,$$

as a consequence of (3.3), and

$$\kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_a(\lambda_0, P) \quad \text{for } 0 \leq |\lambda_0| < \infty,$$

as a consequence of (3.1). In addition, Theorem 3.5 guarantees that there exist values of λ_0 for which $\kappa_\theta((\alpha_0, \beta_0), P)$ and $\kappa_r(\lambda_0, P)$ are very different and also values for which they are very similar. The same happens for $\kappa_\theta((\alpha_0, \beta_0), P)$ and $\kappa_a(\lambda_0, P)$. In the rest of this subsection we explore some of these scenarios.

The next result follows directly from Theorem 3.5 and says that for small $|\lambda_0|$, $\kappa_\theta((\alpha_0, \beta_0), P)$ is essentially $\kappa_a(\lambda_0, P)$, while for $|\lambda_0| \geq 1$, $\kappa_\theta((\alpha_0, \beta_0), P)$ is essentially $\kappa_a(1/\lambda_0, revP)$.

COROLLARY 3.6. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, where $\lambda_0 = \infty$ if $\beta_0 = 0$. Assume that the same weights ω_i are considered in the definition of the condition numbers that appear below. Then*

(i) If $0 \leq |\lambda_0| \leq 1$, we have

$$\frac{\kappa_a(\lambda_0, P)}{2} \leq \kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_a(\lambda_0, P). \quad (3.6)$$

(ii) If $|\lambda_0| \geq 1$, we have

$$\frac{\kappa_a\left(\frac{1}{\lambda_0}, \text{rev}P\right)}{2} \leq \kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_a\left(\frac{1}{\lambda_0}, \text{rev}P\right). \quad (3.7)$$

From Corollary 3.6, we can informally state that for $0 \leq |\lambda_0| \leq 1$, $\kappa_\theta((\alpha_0, \beta_0), P)$ behaves more as an absolute than as a relative eigenvalue condition number of P . The same happens for $|\lambda_0| \geq 1$ but with respect to $\text{rev}P$ and $1/\lambda_0$.

Taking into account Lemma 2.4 and the fact that $\kappa_a(\lambda_0, P) = \kappa_r(\lambda_0, P) \cdot |\lambda_0|$, the observations in (3.6) and (3.7) lead to the following conclusions.

COROLLARY 3.7. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, where $\lambda_0 = \infty$ if $\beta_0 = 0$. Assume that the same weights ω_i are considered in the definition of the condition numbers appearing below. Then,*

(i) If $0 < |\lambda_0| \leq 1$, we have

$$\kappa_r(\lambda_0, P) \frac{|\lambda_0|}{2} \leq \kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_r(\lambda_0, P) |\lambda_0|,$$

(ii) If $1 \leq |\lambda_0| < \infty$, we have

$$\kappa_r(\lambda_0, P) \frac{1}{2|\lambda_0|} \leq \kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_r(\lambda_0, P) \frac{1}{|\lambda_0|}.$$

REMARK 3.8. Using the results in this subsection, we can give a description (in function of $|\lambda_0|$) of the behavior of the Stewart-Sun condition number in terms of the non-homogeneous condition numbers. For this purpose, we use $\kappa_a(\lambda_0, P) = \kappa_r(\lambda_0, P) |\lambda_0|$ to obtain the following relations:

(i) From (3.1), we get that, if $|\lambda_0| = 1$, then

$$\kappa_\theta((\alpha_0, \beta_0), P) = \frac{\kappa_a(\lambda_0, P)}{2} = \frac{\kappa_r(\lambda_0, P)}{2}.$$

(ii) From (3.6), we get, as $|\lambda_0|$ approaches 0,

$$\kappa_\theta((\alpha_0, \beta_0), P) \approx \kappa_a(\lambda_0, P) \ll \kappa_r(\lambda_0, P).$$

(iii) From (3.7), Lemma 2.4, and the fact that $\kappa_a(1/\lambda_0, \text{rev}P) = \kappa_r(1/\lambda_0, \text{rev}P) \cdot \left|\frac{1}{\lambda_0}\right|$, we get, as $|\lambda_0|$ approaches ∞ ,

$$\kappa_\theta((\alpha_0, \beta_0), P) \approx \kappa_a\left(\frac{1}{\lambda_0}, \text{rev}P\right) \ll \kappa_r\left(\frac{1}{\lambda_0}, \text{rev}P\right) = \kappa_r(\lambda_0, P) \ll \kappa_a(\lambda_0, P).$$

4. A geometric interpretation of the relationship between homogeneous and non-homogeneous condition numbers. In Theorem 3.5, we provided an exact relationship between the Stewart-Sun homogeneous condition number and the non-homogeneous condition numbers. This relationship involves the factors

$\frac{1}{1+|\lambda_0|^2}$ and $\frac{1}{1+|\frac{1}{\lambda_0}|^2}$, when the Stewart-Sun condition number is compared with the absolute non-homogeneous condition number, and involves the factor $\frac{|\lambda_0|}{1+|\lambda_0|^2}$, when the Stewart-Sun condition number is compared with the relative non-homogeneous condition number. In this section we give a geometric interpretation of these factors, which leads to a natural understanding of the situations discussed in Remark 3.8 where the homogeneous and non-homogeneous condition numbers can be very different. The reader should bear in mind, once again, that in most applications of PEPs, the quantities of interest are the non-homogeneous eigenvalues, which can be accurately computed with the current algorithms only if the non-homogeneous condition numbers are moderate.

4.1. Geometric interpretation in terms of the chordal distance. The factors $\frac{1}{1+|\lambda_0|^2}$, $\frac{1}{1+|\frac{1}{\lambda_0}|^2}$ and $\frac{|\lambda_0|}{1+|\lambda_0|^2}$ appearing in (3.1), (3.2) and (3.3), respectively, can be interpreted in terms of the chordal distance as we show in the next theorem. We do not include a proof of this result since it can be immediately obtained from the definition of chordal distance (Definition 2.9) and Remark 2.8. Note that $\chi((\alpha_0, \beta_0), (1, 0))$ can be seen as the chordal distance from “ $\lambda_0 = \frac{\alpha_0}{\beta_0}$ to ∞ ”, while $\chi((\alpha_0, \beta_0), (0, 1))$ can be seen as the chordal distance from “ $\lambda_0 = \frac{\alpha_0}{\beta_0}$ to 0”.

PROPOSITION 4.1. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, where $\lambda_0 = \infty$ if $\beta_0 = 0$. Let θ denote the angle between (α_0, β_0) and $(1, 0)$. Then,*

(i) *If $\beta_0 \neq 0$, then*

$$\frac{1}{1+|\lambda_0|^2} = \chi((\alpha_0, \beta_0), (1, 0))^2 = \sin^2(\theta).$$

(ii) *If $\alpha_0 \neq 0$, then*

$$\frac{1}{1+|\frac{1}{\lambda_0}|^2} = \chi((\alpha_0, \beta_0), (0, 1))^2 = \cos^2(\theta).$$

(iii) *If $\alpha_0 \neq 0$ and $\beta_0 \neq 0$, then*

$$\frac{|\lambda_0|}{1+|\lambda_0|^2} = \chi((\alpha_0, \beta_0), (1, 0)) \chi((\alpha_0, \beta_0), (0, 1)) = \sin(\theta) \cos(\theta).$$

REMARK 4.2. We notice that the conditions $0 \leq |\lambda_0| \leq 1$ and $|\lambda_0| \geq 1$ used in Corollaries 3.6 and 3.7 and in Remark 3.8 are equivalent, respectively, to $\chi((\alpha_0, \beta_0), (1, 0)) \geq 1/\sqrt{2}$ and $\chi((\alpha_0, \beta_0), (0, 1)) \geq 1/\sqrt{2}$, or, in other words, to $\sin(\theta) \geq 1/\sqrt{2}$ and $\cos(\theta) \geq 1/\sqrt{2}$.

Combining Proposition 4.1 (iii) with Theorem 3.5 (iii), we see that either when the angle between the lines (α_0, β_0) and $(1, 0)$ is very small or when the angle between the lines (α_0, β_0) and $(0, 1)$ is very small, $\kappa_\theta((\alpha_0, \beta_0), P) \ll \kappa_r(\lambda_0, P)$, i.e., even in the case $\kappa_\theta((\alpha_0, \beta_0), P)$ is moderate and the line (α_0, β_0) changes very little under perturbations, the quotient $\lambda_0 = \alpha_0/\beta_0$ can change a lot in a relative sense. This is immediately understood geometrically in \mathbb{R}^2 . The combination of the remaining parts of Theorem 3.5 and Proposition 4.1 lead to analogous discussions. In the next section, we make an analysis of these facts from another perspective.

4.2. Geometric interpretation in terms of the condition number of the cotangent function. Let $\alpha_0, \beta_0 \in \mathbb{C}$ with $\beta_0 \neq 0$, let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, and let $\theta := \theta((\alpha_0, \beta_0), (1, 0))$, that is, let θ denote the angle between the lines (α_0, β_0) and $(1, 0)$. From Proposition 4.1,

$$\cos \theta = \frac{|\lambda_0|}{\sqrt{1 + |\lambda_0|^2}}, \quad \sin \theta = \frac{1}{\sqrt{1 + |\lambda_0|^2}}.$$

Thus,

$$|\lambda_0| = \frac{|\alpha_0|}{|\beta_0|} = \cotan \theta \quad \text{and} \quad \frac{1}{|\lambda_0|} = \tan \theta.$$

Note that this is also the standard definition of the cotangent and tangent functions in the first quadrant of \mathbb{R}^2 . The cotangent function is differentiable in $(0, \pi/2)$. Thus, the absolute condition number¹ of this function is

$$\kappa_{a,ct}(\theta) := |\cotan'(\theta)| = |1 + \cotan^2(\theta)| = 1 + |\lambda_0|^2, \quad (4.1)$$

which is huge when θ approaches zero. Moreover, the relative-absolute condition number² of the cotangent function is given by

$$\kappa_{r,ct}(\theta) := \frac{|\cotan'(\theta)|}{|\cotan(\theta)|} = \frac{|1 + \cotan^2(\theta)|}{|\cotan(\theta)|} = \frac{1 + |\lambda_0|^2}{|\lambda_0|}, \quad (4.2)$$

which is huge when θ approaches either zero or $\pi/2$.

The tangent function is also differentiable in $(0, \pi/2)$ and the absolute condition number of this function is

$$\kappa_{a,t}(\theta) := |\tan'(\theta)| = |1 + \tan^2(\theta)| = 1 + \left| \frac{1}{\lambda_0} \right|^2. \quad (4.3)$$

From (3.1), (3.2), (3.3), (4.1), (4.2), and (4.3), we obtain the following result.

THEOREM 4.3. *Let $(\alpha_0, \beta_0) \neq (0, 0)$ be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$ of grade k and let $\lambda_0 := \frac{\alpha_0}{\beta_0}$, where $\lambda_0 = \infty$ if $\beta_0 = 0$. Let $\theta := \theta((\alpha_0, \beta_0), (1, 0))$. Assume that the same weights are considered in the definitions of all the condition numbers appearing below. Then,*

(i) *If $\beta_0 \neq 0$, then*

$$\kappa_a(\lambda_0, P) = \kappa_\theta((\alpha_0, \beta_0), P) \kappa_{a,ct}(\theta). \quad (4.4)$$

(ii) *If $\alpha_0 \neq 0$, then*

$$\kappa_a\left(\frac{1}{\lambda_0}, \text{rev}P\right) = \kappa_\theta((\alpha_0, \beta_0), P) \kappa_{a,t}(\theta).$$

(iii) *If $\alpha_0 \neq 0$ and $\beta_0 \neq 0$, then*

$$\kappa_r(\lambda_0, P) = \kappa_\theta((\alpha_0, \beta_0), P) \kappa_{r,ct}(\theta). \quad (4.5)$$

¹When we refer to the *absolute* condition number of a function of θ , we mean that we are measuring the changes both in the function and in θ in an absolute sense.

²When we mention the *relative-absolute* condition number of a function of θ , we consider that we are measuring the change in the function in a relative sense and the one in θ in an absolute sense.

Since for lines (α_0, β_0) and $(\tilde{\alpha}_0, \tilde{\beta}_0)$ very close to each other, $\chi((\alpha_0, \beta_0), (\tilde{\alpha}_0, \tilde{\beta}_0)) \approx \theta((\alpha_0, \beta_0), (\tilde{\alpha}_0, \tilde{\beta}_0)) = |\theta((\alpha_0, \beta_0), (1, 0)) - \theta((\tilde{\alpha}_0, \tilde{\beta}_0), (1, 0))|$, equations (4.4) and (4.5) express the non-homogeneous condition numbers of λ_0 as a combination of two effects: the change of the homogeneous eigenvalue measured by $\theta((\alpha_0, \beta_0), (\tilde{\alpha}_0, \tilde{\beta}_0))$ as a consequence of perturbations in the coefficients of $P(\alpha, \beta)$ and the alteration that this change produces in $|\cotan(\theta)|$, which depends only on the properties of $\cotan(\theta)$ and not on $P(\lambda)$. In fact, with this idea in mind, Theorem 4.3 can also be obtained directly from the definitions of the involved condition numbers.

We notice that the expressions in (4.4) and (4.5) can be interpreted as follows: Given a matrix polynomial $P(\lambda)$, the usual way to solve the polynomial eigenvalue problem is to use a linearization $L(\lambda) = \lambda L_1 - L_0$ of $P(\lambda)$. A standard algorithm to solve the generalized eigenvalue problem associated with $L(\lambda)$ is the QZ algorithm. This algorithm computes first the generalized Schur decomposition of L_1 and L_0 , that is, these matrix coefficients are factorized in the form $L_1 = QSZ^*$ and $L_0 = QTZ^*$, where Q and Z are unitary matrices and S and T are upper-triangular matrices. The pairs (T_{ii}, S_{ii}) , where S_{ii} and T_{ii} denote the main diagonal entries of S and T in position (i, i) , respectively, are the ‘‘homogeneous’’ eigenvalues of $L(\lambda)$ (and, therefore, of $P(\lambda)$). In order to obtain the non-homogeneous eigenvalues of $P(\lambda)$, one more step is necessary, namely, to divide T_{ii}/S_{ii} . The expressions in (4.4) and (4.5) say that, even if $\kappa_\theta((T_{ii}, S_{ii}), P)$ is moderate and the pair (T_{ii}, S_{ii}) is ‘‘accurately computed’’, the quotient $\lambda_i := T_{ii}/S_{ii}$ may be ‘‘inaccurately computed’’ when S_{ii} is very close to zero (that is, when $|\lambda_i|$ is very large) or when T_{ii} is close to zero (that is, when $|\lambda_i|$ is close to zero) since $|\lambda_i|$ will have a huge non-homogeneous condition number. More precisely, for the large eigenvalues, both $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ will be much larger than $\kappa_\theta((\alpha_0, \beta_0), P)$, and for the small eigenvalues, $\kappa_r(\lambda_0, P)$ will be much larger than $\kappa_\theta((\alpha_0, \beta_0), P)$. This observation brings up the question of the computability of small and large eigenvalues.

4.3. Computability of small and large eigenvalues of matrix polynomials. By Remark 3.8, if $|\lambda_0|$ is very large, we have

$$\kappa_\theta((\lambda_0, 1), P) \ll \kappa_r(\lambda_0, P) \ll \kappa_a(\lambda_0, P), \quad (4.6)$$

and, if $|\lambda_0|$ is very close to 0, then

$$\kappa_\theta((\lambda_0, 1), P) \approx \kappa_a(\lambda_0, P) \ll \kappa_r(\lambda_0, P). \quad (4.7)$$

The question that we study in this section is whether or not the absolute non-homogeneous condition number $\kappa_a(\lambda_0, P)$ is always very large when $|\lambda_0|$ is very large, or the relative non-homogeneous condition number $\kappa_r(\lambda_0, P)$ is always very large when either $|\lambda_0|$ is very close to 0 or is very large. If this was the case, then these types of non-homogeneous eigenvalues would be always very ill-conditioned and would be computed with such huge errors by the available algorithms that it could be simply said that they are not computable.

We focus our answer to this question on the behavior of the eigenvalues of pencils (that is, we focus on matrix polynomials $P(\lambda)$ of grade 1). This is reasonable since, when computing the eigenvalues of a matrix polynomial, the most common approach is to use a linearization. In order to keep in mind that we are not working with general matrix polynomials, we will use the notation $L(\lambda)$ to denote a pencil instead of $P(\lambda)$. Moreover, we will focus on eigenvalue condition numbers with weights ω_i corresponding to the backward errors of current algorithms for generalized eigenvalue problems.

Let $L(\lambda) := \lambda B_1 + B_0$ be a regular pencil and let λ_0 be a finite, simple eigenvalue of $L(\lambda)$. Notice that, by Theorem 2.13,

$$\kappa_\theta((\lambda_0, 1), L) = \frac{(\omega_0 + |\lambda_0| \omega_1) \|y\|_2 \|x\|_2}{|y^*(B_1 - \lambda_0 B_0)x|} \geq \frac{(\omega_0 + |\lambda_0| \omega_1)}{\|B_1\|_2 + |\lambda_0| \|B_0\|_2}. \quad (4.8)$$

The following result is an immediate consequence of the previous inequality and shows some important cases in which the eigenvalues of a pencil $L(\lambda)$ with large or small modulus are not computable.

PROPOSITION 4.4. *Let $L(\lambda) = \lambda B_1 + B_0$ be a regular pencil and let λ_0 be a finite, simple eigenvalue of $P(\lambda)$. Let ω_1, ω_0 be the weights used in the definition of the non-homogeneous condition number of λ_0 . Then,*

$$\kappa_a(\lambda_0, L) \geq 1 + |\lambda_0|^2 \quad \text{and} \quad \kappa_r(\lambda_0, L) \geq \frac{1 + |\lambda_0|^2}{|\lambda_0|},$$

(where the second inequality holds only if $\lambda_0 \neq 0$) if any of the following conditions holds:

1. $\omega_0 = \omega_1 = \max\{\|B_1\|_2, \|B_0\|_2\}$;
2. $\omega_i = \|B_i\|_2$ for $i = 0, 1$, $|\lambda_0| < 1$ and $\|B_1\|_2 \leq \|B_0\|_2$;
3. $\omega_i = \|B_i\|_2$ for $i = 0, 1$, $|\lambda_0| > 1$ and $\|B_0\|_2 \leq \|B_1\|_2$;
4. $\omega_i = \|B_i\|_2$ for $i = 0, 1$ and $\|B_1\|_2$ and $\|B_0\|_2$ are similar.

Proof. First assume that $\omega_1 = \omega_0 = \max\{\|B_0\|_2, \|B_1\|_2\}$. From (4.8), we get

$$\kappa_\theta((\lambda_0, 1), L) \geq 1,$$

which implies the result by Theorem 3.5. Assume now that $\omega_i = \|B_i\|_2$ for $i = 0, 1$. If $|\lambda_0| < 1$ and $\|B_1\|_2 \leq \|B_0\|_2$, then

$$(1 - |\lambda_0|) \|B_1\|_2 \leq (1 - |\lambda_0|) \|B_0\|_2,$$

or equivalently,

$$\frac{\|B_0\|_2 + |\lambda_0| \|B_1\|_2}{\|B_1\|_2 + |\lambda_0| \|B_0\|_2} \geq 1$$

and the result follows from (4.8), which implies $\kappa_\theta((\lambda_0, 1), L) \geq 1$, and Theorem 3.5.

The proof of the result assuming 3. follows from applying 2. to $revL$ and $\frac{1}{|\lambda_0|}$ and taking into account Lemma 2.4. Finally, assume that $\|B_1\|_2 \approx \|B_0\|_2$. Then, the result follows from 2. and 3. \square

REMARK 4.5. We admit a certain degree of ambiguity in the meaning of “ $\|B_1\|_2$ and $\|B_0\|_2$ are similar” in the fourth set of assumptions in Proposition 4.4. It is possible to make a quantitative assumption of the type $\frac{1}{C} \leq \frac{\|B_1\|_2}{\|B_0\|_2} \leq C$ for some constant $C > 1$, which would more laboriously lead to more complicated lower bounds for $\kappa_a(\lambda_0, L)$ and $\kappa_r(\lambda_0, L)$ which decrease as C increases. We have preferred the simpler but somewhat ambiguous statement in Proposition 4.4.

Taking into account Proposition 4.4, the only case left to study of the potential non-computability of eigenvalues with small and large absolute values is when $|\lambda_0| \ll 1$ and $\|B_0\|_2 \ll \|B_1\|_2$. The case in which $|\lambda_0| \gg 1$ and $\|B_0\|_2 \gg \|B_1\|_2$ follows by using the reversal of $L(\lambda)$ and the fact that

$$\kappa_r(\lambda_0, L) = \kappa_r\left(\frac{1}{\lambda_0}, revL\right).$$

Notice that when $\|B_0\|_2 \ll \|B_1\|_2$, in practice, $L(\lambda)$ is essentially λB_1 , and since $L(\lambda)$ is regular, B_1 is generically nonsingular, which implies that, when $\frac{\|B_0\|_2}{\|B_1\|_2}$ approaches 0, all the eigenvalues of $L(\lambda)$ are very close to 0, independently of the specific matrices B_0 and B_1 . Thus, in this trivial case, we should expect that the eigenvalues are all well-conditioned. Next we show an example that illustrates this observation.

EXAMPLE 4.6. Let $\epsilon < 1$. Consider the regular pencil

$$L(\lambda) = \lambda B_1 + B_0 := \lambda \begin{bmatrix} 1/\epsilon & 1/\epsilon \\ 1/\epsilon & 1 \end{bmatrix} + \begin{bmatrix} \epsilon & 1 \\ 1 & 1 \end{bmatrix}.$$

We note that the standard condition numbers for inversion, i.e., $\|B_i\|_2 \|B_i^{-1}\|_2$, of the matrix coefficients B_0 and B_1 are bounded by

$$\frac{1}{1-\epsilon} \leq \text{cond}(B_0) = \text{cond}(B_1) \leq \frac{4}{1-\epsilon}$$

and

$$\frac{2}{\sqrt{2}} \leq \|B_0\|_2 \leq 2, \quad \frac{2}{\sqrt{2}\epsilon} \leq \|B_1\|_2 \leq \frac{2}{\epsilon}.$$

Thus, as ϵ approaches 0, $\text{cond}(B_0) = \text{cond}(B_1) \approx 1$ and $\|B_0\|_2 \ll \|B_1\|_2$. Moreover, the eigenvalues of $L(\lambda)$, given by

$$\lambda_0 = \frac{\epsilon - \epsilon^3 + \sqrt{\epsilon^6 - 6\epsilon^4 + 8\epsilon^3 - 3\epsilon^2}}{2\epsilon - 2}, \quad \lambda_1 = \frac{\epsilon - \epsilon^3 - \sqrt{\epsilon^6 - 6\epsilon^4 + 8\epsilon^3 - 3\epsilon^2}}{2\epsilon - 2}$$

are both of order ϵ as $\epsilon \rightarrow 0$. Additionally, the right (resp. left) eigenvectors of $L(\lambda)$ associated with λ_0 and λ_1 are denoted, respectively, by x_{λ_0} and x_{λ_1} (resp. y_{λ_0} and y_{λ_1}) are given by

$$x_{\lambda_0} = \overline{y_{\lambda_0}} = \left(1, \frac{-\lambda_0 - \epsilon^2}{\lambda_0 + \epsilon} \right)^T, \quad x_{\lambda_1} = \overline{y_{\lambda_1}} = \left(1, \frac{-\lambda_1 - \epsilon^2}{\lambda_1 + \epsilon} \right)^T.$$

Observe that, as $\epsilon \rightarrow 0$, we have $\|x_{\lambda_0}\|_2 = \|y_{\lambda_0}\|_2 \approx \|y_{\lambda_1}\|_2 = \|x_{\lambda_1}\|_2 \approx \sqrt{2}$. Moreover, we have

$$\kappa_\theta((\lambda_0, 1), L) = \frac{(|\lambda_0| \|B_1\|_2 + \|B_0\|_2) \|y_{\lambda_0}\|_2 \|x_{\lambda_0}\|_2}{|y_{\lambda_0}^*(B_1 - \overline{\lambda_0} B_0)x_{\lambda_0}|}.$$

Some computations show that

$$\begin{aligned} y_{\lambda_0}^*(B_1 - \overline{\lambda_0} B_0)x_{\lambda_0} &= \frac{1}{\epsilon} - \overline{\lambda_0} \epsilon + \left(\frac{1}{\epsilon} + \overline{\lambda_0} \epsilon \right) \left(\frac{-\lambda_0 - \epsilon^2}{\lambda_0 + \epsilon} \right) \\ &\quad + \left[\frac{1}{\epsilon} - \overline{\lambda_0} + \frac{(-\lambda_0 - \epsilon^2)(1 + \overline{\lambda_0})}{\lambda_0 + \epsilon} \right] \left(\frac{-\epsilon - \epsilon^2}{2\epsilon} \right). \end{aligned}$$

Combining all this information, we deduce that $\kappa_\theta((\lambda_0, 1), L) = O(\epsilon)$ as $\epsilon \rightarrow 0$, noticing that the numerator tends to a constant while the denominator is $O(1/\epsilon)$. Then, from Theorem 3.5 and from $\frac{1+|\lambda_0|^2}{|\lambda_0|} = O(\frac{1}{\epsilon})$, we have that $\kappa_r(\lambda_0, L) = O(1)$. It is clear that a similar conclusion can be obtained for $\kappa_r(\lambda_1, L)$.

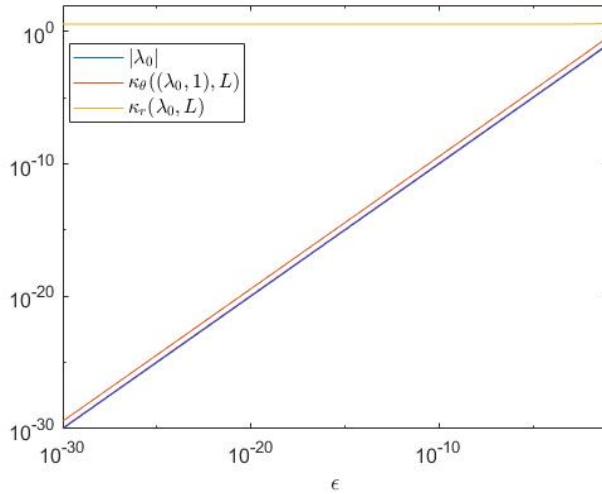


FIG. 4.1. Plot of $|\lambda_0|$, $\kappa_\theta((\lambda_0, 1), L)$ and $\kappa_r(\lambda_0, L)$ against ϵ in Example 4.6.

In the Figure 4.1, we plot the values of $|\lambda_0|$, $\kappa_\theta((\lambda_0, 1), L)$ and $\kappa_r(\lambda_0, L)$ against ϵ . We can observe that, as $\epsilon \rightarrow 0$, both $|\lambda_0|$ and $\kappa_\theta((\lambda_0, 1), L)$ tend to 0 at the same rate but that $\kappa_r(\lambda_0, L)$ remains essentially constant and close to 1. A similar graph can be obtained for λ_1 .

In summary, we have proven that, as $\epsilon \rightarrow 0$, the pencil $L(\lambda) = \lambda B_1 + B_0$ in Example 4.6 has well-conditioned matrix coefficients with very different norms, it has two eigenvalues close to 0 and they are both computable, since their non-homogeneous relative condition numbers are approximately 1. It is easy to check that $revL$ is an example of a pencil that has computable very large eigenvalues. As commented before, these pencils illustrate the only situation in which non-homogeneous eigenvalues with very small or large modulus can be accurately computed, which corresponds to pencils with highly unbalanced coefficients.

5. Conclusions. We have gathered together the definitions of (non-homogeneous and homogeneous) eigenvalue condition numbers of matrix polynomials that were scattered in the literature. We have also derived for the first time an exact formula to compute one of these condition numbers (the homogeneous condition number that is based on the chordal distance, also called Stewart-Sun condition number). On the one hand, we have determined that the two homogeneous condition numbers studied in this paper differ at most by a factor $\sqrt{k+1}$, where k is the grade of the polynomial, and so are essentially equal in practice. Since the definition of the homogeneous condition number based on the chordal distance is considerably simpler, we believe that its use should be preferred among the homogeneous condition numbers. On the other hand, we have proved exact relationships between each of the non-homogeneous condition numbers and the homogeneous condition number based on the chordal distance. This result would allow us to extend results that have been proved for the non-homogeneous condition numbers to the homogeneous condition numbers (and vice versa). Besides, we have provided geometric interpretations of the factor that appears in these exact relationships, which explain transparently when and why the non-homogeneous condition numbers are much larger than the homogeneous ones. Fi-

nally, we have used these relationships to analyze for which cases the large and small non-homogeneous eigenvalues of a matrix polynomial are computable with some accuracy and we have seen that this is only possible in some very particular situations.

REFERENCES

- [1] M. BERHANU, *The polynomial eigenvalue problem*. Thesis (Ph.D.) - The University of Manchester (United Kingdom). 2005. 221 pp. ISBN: 978-0355-47610-1.
- [2] J. P. DEDIEU, *Condition operators, condition numbers, and condition number theorem for the generalized eigenvalue problem*, Linear Algebra Appl., 263 (1997), 1-24.
- [3] J.P. DEDIEU AND F. TISSEUR, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), 71-94.
- [4] S. HAMMARLING, C.J. MUNRO, AND F. TISSEUR, *An algorithm for the complete solution of quadratic eigenvalue problems*, ACM Trans. Math. Softw., 39,3 (2013), 18:1-18:19.
- [5] N.J. HIGHAM, D. S. MACKAY, AND F. TISSEUR, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28, 4 (2006), 1005-1028.
- [6] M. SHUB AND S. SMALE, *Complexity of Bezout's theorem IV: Probability of success; extensions*, SIAM J. Numer. Anal., 33,1 (1996), 128-148.
- [7] G. W. STEWART AND J.-G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
- [8] F. TISSEUR, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), 339-361.