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 condition numbers for simple eigenvalues of 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 condition 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

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

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

$$\mathrm{rev}(P(\lambda)) := \lambda^k P(1/\lambda)\,. \tag{1.2}$$

*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).

In order to define finite and infinite eigenvalues in a unified way, it is convenient to express a matrix polynomial in homogeneous form, that is,

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

Then, 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

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

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 $[\alpha_0,\beta_0]^T$ a specific (nonzero) 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 grade 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 nonzero $[\alpha_0,\beta_0]^T$ such that $\lambda_0 = \alpha_0/\beta_0$, including $\lambda_0 = \infty$ for $\beta_0 = 0$.

The numerical solution of the (homogeneous and non-homogeneous) PEP has received considerable attention from many research groups in the last two decades and, as a consequence, several condition numbers for (simple and multiple) eigenvalues of a matrix polynomial P have been defined in the literature to determine the sensitivity of these eigenvalues to perturbations in the coefficients of P [3, 4, 8, 9, 12]. Although several different eigenvalue condition numbers (in the homogeneous and the non-homogeneous formulation) are available in the literature for simple eigenvalues, as far as we know only one definition of condition number for multiple eigenvalues (in the non-homogeneous formulation) has been considered so far. Our goal in this paper is to compare distinct eigenvalue condition numbers. This is the reason why we focus on the condition numbers of simple eigenvalues, aside of the fact that simple eigenvalues are essentially the only ones appearing in the numerical practice due to the effect of the finite arithmetic of the computer.

One of the condition numbers for simple eigenvalues, defined for matrix polynomials expressed in the non-homogeneous form, 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 in a unified framework, other condition numbers are considered in the literature. These condition numbers assume that the matrix polynomial is expressed in homogeneous form.

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 [11, 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 the norm of a differential operator that is constructed making use of the Implicit Function Theorem [3, 4]. 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-homogenous 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 be 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 of simple eigenvalues used in the literature and discuss some of the advantages and disadvantages of each of them. We note that these are normwise condition numbers. When the coefficients of a matrix polynomial have very specific entrywise structures (such as the pencils considered in [13]), it might be more convenient to consider entrywise condition numbers, as those defined in [6] for pencils, so that entrywise structured perturbations are used. Since entrywise condition numbers have not received attention so far in recent research on matrix polynomials, we do not study entrywise condition numbers in this work.

Before recalling the definition of the eigenvalue condition numbers that we study in this paper, we present some notation that will be used throughout the paper.

Let a and b be two integers. We define

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

For any matrix A, $\|A\|_2$ denotes its spectral or 2-norm, i.e., its largest singular value [11]. 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 [12].

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*

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

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*

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

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) $\omega_i = \|B_i\|_2$ for $i = 0, 1, \ldots, k$ (relative coefficient-wise perturbations); 2) $\omega_i = \max_{j=0:k}\{\|B_j\|_2\}$ for $i = 0, 1, \ldots, k$ (relative perturbations with respect to the norm of $P(\lambda)$); 3) $\omega_i = 1$ for $i = 0, 1, \ldots, k$ (absolute perturbations).

An explicit formula for each of the non-homogeneous eigenvalue condition numbers presented above was obtained by Tisseur in [12].

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,*

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

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

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

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

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. As explained in Section 1, one of them is due to Stewart and Sun [11] and the other one, inspired by Shub and Smale's work [10], is due to Dedieu and Tisseur [3, 4]. Neither of these two condition numbers has 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, [5, 7]. 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 [4]. The next theorem provides an explicit formula for this condition number.

THEOREM 2.6. [4, 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*

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

where $D_z \equiv \frac{\partial}{\partial z}$, that is, the partial derivative with respect to $z \in \{\alpha, \beta\}$, 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*

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

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,

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

since

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

We also have

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

The definition of chordal distance is given next.

Definition 2.9. [11, 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. [11, 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 [11, page 294], although an explicit definition is not given in [11]. 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*

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

where $\Delta P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} \Delta B_i$ and ω_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,*

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

where $D_z\equiv\frac{\partial}{\partial z}$, that is, the partial derivative with respect to $z\in\{\alpha,\beta\}$.

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 $\{[\alpha_0,\beta_0]^T,[-\overline{\beta_0},\overline{\alpha_0}]^T\}$ 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

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

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

$$[\Delta\alpha_0,\Delta\beta_0]^T=\mathfrak{h}[-\overline{\beta_0},\overline{\alpha_0}]^T. \tag{2.3}$$

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

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

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 [4, 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,

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

On the other hand,

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

Since, by (2.5),

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

and $\Delta\alpha_0$ and $\Delta\beta_0$ approach zero as $\epsilon \to 0$, 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

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

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

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. [3, 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 $v := [|\beta_0|^k\omega_0, |\alpha_0||\beta_0|^{k-1}\omega_1, \ldots, |\alpha_0|^k\omega_k]^T \in \mathbb{R}^{k+1}$. 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 $\|.\|_1$ denotes the vector 1-norm [11]. The result follows taking into account the fact that $1/\sqrt{k+1} \leq \|v\|_2/\|v\|_1 \leq 1$. ◻

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, $\kappa_a((1/\lambda_0), \mathrm{rev}P)$. For $\lambda_0 = \infty$, this turns into $\kappa_a(0, \mathrm{rev}P)$ 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 $\beta_0 \neq 0$,

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

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

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

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

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

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

$$\Delta\lambda_0 = \frac{\tilde{\alpha_0}\beta_0 - \alpha_0\tilde{\beta_0}}{\beta_0\tilde{\beta_0}} = \frac{\beta_0\Delta\alpha_0 - \alpha_0\Delta\beta_0}{\beta_0(\beta_0+\Delta\beta_0)}. \tag{3.4}$$

Then, we have, by Lemma 2.10,

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

where the second equality follows from (3.4). Then, as $\epsilon \to 0$ 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 $\Delta\lambda_0 \to 0$. 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 rev$P$ and Theorem 2.13, which yield:

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

□

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). We also get

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

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.

Remark 3.6. As explained in Section 1, the main motivation behind the definition of the homogeneous eigenvalue condition numbers was to provide a unified way to measure the sensitivity of all the eigenvalues (finite and infinite) of a matrix polynomial to perturbations in its matrix coefficients, since the relative non-homogeneous condition number, being, probably, the most important for practical purposes, is not defined for zero and infinite eigenvalues. However, this non-homogeneous condition number is defined for nonzero eigenvalues whose modulus is as close to 0 and infinity as wished, and note that, according to (3.3), in these cases $\kappa_\theta((\alpha_0, \beta_0), P) \ll \kappa_r(\lambda_0, P)$, because $|\lambda_0|/(1+|\lambda_0|^2) \ll 1$ if $|\lambda_0|$ is very large or very small. Thus, a natural question is to provide an interpretation of $\kappa_\theta((\alpha_0, \beta_0), P)$ in terms of Wilkinson-like eigenvalue condition numbers when $|\lambda_0|$ is very large or very small.

For this purpose, note that from Theorem 3.5, we obtain

$$\frac{\kappa_a(\lambda_0, P)}{2} \leq \kappa_\theta((\alpha_0, \beta_0), P) \leq \kappa_a(\lambda_0, P), \qquad \text{if } |\lambda_0| \leq 1, \tag{3.6}$$

and

$$\frac{\kappa_a\left(\frac{1}{\lambda_0}, \mathrm{rev}P\right)}{2} \leq \kappa_\theta\left((\alpha_0, \beta_0), P\right) \leq \kappa_a\left(\frac{1}{\lambda_0}, \mathrm{rev}P\right), \qquad \text{if } |\lambda_0| \geq 1. \tag{3.7}$$

From (3.6) and (3.7), we state that for "small non-homogeneous eigenvalues", $\kappa_\theta((\alpha_0, \beta_0), P)$ is essentially $\kappa_a(\lambda_0, P)$, while for "large non-homogeneous eigenvalues", $\kappa_\theta((\alpha_0, \beta_0), P)$ is essentially $\kappa_a(1/\lambda_0, \mathrm{rev}P)$. Therefore, $\kappa_\theta((\alpha_0, \beta_0), P)$ measures *absolute* variations of very small non-homogeneous eigenvalues and *absolute* variations of the reciprocal of very large non-homogeneous eigenvalues, while in most applications the relative variations are the ones of interest. These comments reveal that, although the homogeneous eigenvalue condition numbers solve the lack of definition of the relative non-homogeneous condition number for 0 and infinite eigenvalues, and provide a unified way to measure the sensitivity of all the eigenvalues of a matrix polynomial, they achieve these objectives at the cost of becoming an absolute condition number at eigenvalues close to 0 or infinity.

Remark 3.7. Taking into account (3.6) and (3.7), the fact that $\kappa_a(\lambda_0, P) = \kappa_r(\lambda_0, P)|\lambda_0|$, Theorem 3.5, and Lemma 2.4, the following relations are obtained:

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

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

(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 $\kappa_a(1/\lambda_0, \mathrm{rev}P) = \kappa_r(1/\lambda_0, \mathrm{rev}P) \cdot \left|\frac{1}{\lambda_0}\right|$, we get, as $|\lambda_0|$ approaches ∞,

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

Next, we give an example that illustrates results (ii) and (iii) in Remark 3.7 for a quadratic matrix polynomial with very small and very large non-homogeneous eigenvalues. This example will also be useful to illustrate some results in Subsection 4.3, where we study the non-computability of eigenvalues of matrix polynomials with large and small moduli.

EXAMPLE 3.8. Let us consider the quadratic matrix polynomial

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} (\lambda-10^5)(\lambda-10^{10}) & 0 \\ 0 & (\lambda-10^{-5})(\lambda-10^{15}) \end{bmatrix} \\ &= \lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \lambda \begin{bmatrix} -10^5-10^{10} & 0 \\ 0 & -10^{-5}-10^{15} \end{bmatrix} + \begin{bmatrix} 10^{15} & 0 \\ 0 & 10^{10} \end{bmatrix} \\ &=: \lambda^2 B_2 + \lambda B_1 + B_0. \end{aligned}$$

The eigenvalues of $P(\lambda)$ are $\lambda_1 := 10^{-5}$, $\lambda_2 := 10^5$, $\lambda_3 := 10^{10}$, and $\lambda_4 := 10^{15}$. Right and left eigenvectors associated with these eigenvalues are, respectively,

$$x_{\lambda_1} = y_{\lambda_1} = [0,1]^T, \quad x_{\lambda_2} = y_{\lambda_2} = [1,0]^T, \quad x_{\lambda_3} = y_{\lambda_3} = [1,0]^T, \quad x_{\lambda_4} = y_{\lambda_4} = [0,1]^T.$$

In the computation of the condition numbers in this example, we consider the weights $\omega_i = \|B_i\|_2$, $i = 0:2$. We note that $\|B_2\|_2 = 1$, $\|B_1\|_2 = 10^{-5} + 10^{15}$, and $\|B_0\|_2 = 10^{15}$. In the next table we give the order of the value of the three condition numbers $\kappa_\theta((\lambda_0,1),P)$, $\kappa_a(\lambda_0,P)$, and $\kappa_r(\lambda_0,P)$ for each eigenvalue.

Eigenvalues	κ_θ	κ_a	κ_r
10^{-5}	1	1	10^5
10^5	1	10^{10}	10^5
10^{10}	10^{-5}	10^{15}	10^5
10^{15}	10^{-15}	10^{15}	1

The first row of this table illustrates Remark 3.7(ii), while the second, third, and fourth rows illustrate Remark 3.7(iii). The fourth row is particularly interesting because it corresponds to a very large non-homogeneous eigenvalue that is very well-conditioned in a relative sense, since $\kappa_r(10^{15},P) \approx 1$ as a consequence of the fact that $\kappa_\theta((10^{15},1),P)$ is very small. Further discussions on this type of situations are presented in Subsection 4.3.

We end this section with a comment on the impact of the results in Theorem 3.5 on the comparison of the condition numbers of some $\lambda_0 \in \mathbb{C}$ when seen as an eigenvalue of a matrix polynomial P and as an eigenvalue of a linearization of P.

We recall that the usual way to solve a polynomial eigenvalue problem $P(\lambda)x = 0$ is to construct a linearization of $P(\lambda)$ and to solve the associated generalized eigenvalue problem. However, it is important to ensure that the condition numbers of the eigenvalues of P are not significantly magnified by this procedure. More precisely, given a linearization $L(\lambda)$ of $P(\lambda)$ and a simple, finite, nonzero eigenvalue λ_0 of $P(\lambda)$, we would like that $\kappa(\lambda_0, L)/\kappa(\lambda_0, P) = O(1)$ [2, 7]. We note that, depending on which definition of condition number is used, this quotient might, in principle, produce very different results. However, Theorem 3.5 shows that, for the three types of condition numbers considered in this subsection, the quotients will be the same. Namely,

$$\frac{\kappa_\theta((\lambda_0,1),L)}{\kappa_\theta((\lambda_0,1),P)} = \frac{\kappa_a(\lambda_0,L)}{\kappa_a(\lambda_0,P)} = \frac{\kappa_r(\lambda_0,L)}{\kappa_r(\lambda_0,P)}.$$

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 $\frac{1}{1+\left|\frac{1}{\lambda_0}\right|^2}$, 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 3.7 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 $\frac{1}{1+|\lambda_0|^2}$, $\frac{1}{1+\left|\frac{1}{\lambda_0}\right|^2}$ 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

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

(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 Remarks 3.6 and 3.7 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,

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

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 number[1] of this function is

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

which is huge when θ approaches zero. Moreover, the relative-absolute condition number[2] of the cotangent function is given by

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

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

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

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,*

[1]When 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.

[2]When 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.

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

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

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

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

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

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

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 $|\cot(\theta)|$, which depends only on the properties of $\cot(\theta)$ 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 $P(\lambda)$, we have already mentioned that 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 non-homogeneous eigenvalues.

4.3. Computability of small and large non-homogeneous eigenvalues of matrix polynomials. By Remark 3.7, if $|\lambda_0|$ is very large, we have

$$\kappa_\theta((\lambda_0, 1), P) \ll \kappa_r(\lambda_0, P) \ll \kappa_a(\lambda_0, P), \tag{4.6}$$

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

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

These relations suggest that the non-homogeneous eigenvalue condition numbers $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ may often be very large when either $|\lambda_0|$ is very large or very close to 0. However, we know from the eigenvalue $\lambda_0 = 10^{15}$ in Example 3.8 that this is not always the case, since it is possible that a very large eigenvalue λ_0 has $\kappa_r(\lambda_0, P) \approx 1$. In this section we provide sufficient conditions under which $\kappa_r(\lambda_0, P) \gg 1$ or $\kappa_a(\lambda_0, P) \gg 1$ for very large or small eigenvalues, which show that this happens in many situations. In these cases, the non-homogeneous eigenvalues are very ill-conditioned and could be computed with such huge errors by the available algorithms that it could be simply said that they are not computable.

In the rest of this section, we provide lower bounds on both $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ that will allow us to determine sufficient conditions for λ_0 to be ill-conditioned under any of the two non-homogeneous condition numbers. We focus first on the behavior of the condition numbers of the eigenvalues of pencils (that is, matrix polynomials of grade 1) with very small or very large modulus. This case is very important since, as explained in Section 3, the most common approach to computing the eigenvalues of a matrix polynomial is to use a linearization. To keep in mind that we are not working with general matrix polynomials in the first part of this section, we will use the notation $L(\lambda)$ instead of $P(\lambda)$ to denote a pencil. Moreover, we will focus on eigenvalue condition numbers with weights ω_i corresponding to the relative with-respect-to-the-norm-of-P and relative coefficient-wise perturbations since these correspond to the backward errors of current algorithms for generalized eigenvalue problems (recall Remark 2.2).

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,

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

The following result, which is an immediate consequence of the previous inequality, provides lower bounds on the non-homogeneous condition numbers that will be used in Remark 4.5 to identify sufficient conditions for the non-homogeneous eigenvalues of a pencil $L(\lambda)$ with large or small modulus to not be computable, i.e., to be very ill-conditioned.

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 $L(\lambda)$. Let ω_1, ω_0 be the weights used in the definition of the non-homogeneous condition numbers of λ_0. Let*

$$\rho_0 := \frac{\|B_0\|_2}{\max\{\|B_0\|_2, \|B_1\|_2\}}, \quad \rho_1 := \frac{\|B_1\|_2}{\max\{\|B_0\|_2, \|B_1\|_2\}}.$$

1. *If $\omega_0 = \omega_1 = \max\{\|B_1\|_2, \|B_0\|_2\}$, then*

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

where the lower bound for $\kappa_r(\lambda_0, L)$ holds only if $\lambda_0 \neq 0$.

2. *If $\omega_i = \|B_i\|_2$ for $i = 0, 1$, then*

$$\kappa_a(\lambda_0, L) \geq \max\left\{\frac{1}{2}, \frac{|\lambda_0|}{2}\right\}(\rho_0 + |\lambda_0|\rho_1), \text{ and} \tag{4.9}$$

$$\kappa_r(\lambda_0, L) \geq \max\left\{\frac{1}{2|\lambda_0|}, \frac{1}{2}\right\}(\rho_0 + |\lambda_0|\rho_1),$$

where the lower bound for $\kappa_r(\lambda_0, L)$ holds only if $\lambda_0 \neq 0$.

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 $i = 0, 1$. Then,

$$\kappa_\theta((\lambda_0, 1), L) \geq \frac{\rho_0 + \rho_1|\lambda_0|}{1 + |\lambda_0|}.$$

By Theorem 3.5, we have

$$\kappa_a(\lambda_0, L) \geq \left(\frac{1 + |\lambda_0|^2}{1 + |\lambda_0|}\right)(\rho_0 + |\lambda_0|\rho_1),$$

and the result follows for $\kappa_a(\lambda_0, L)$ by noticing that

$$\frac{1 + |\lambda_0|^2}{1 + |\lambda_0|} \geq \begin{cases} 1/2, & \text{if } |\lambda_0| \leq 1 \\ |\lambda_0|/2, & \text{if } |\lambda_0| > 1. \end{cases}$$

The result for $\kappa_r(\lambda_0, L)$ follows from (4.9) taking into account that $|\lambda_0|\kappa_r(\lambda_0, L) = \kappa_a(\lambda_0, L)$. □

Remark 4.5. From Proposition 4.4, we obtain the following sufficient conditions for the non-computability of simple eigenvalues with small and large moduli.

When relative perturbations with respect to the norm of $L(\lambda)$ are considered,

1. $\kappa_a(\lambda_0, L) \gg 1$ if $|\lambda_0| \gg 1$,
2. $\kappa_r(\lambda_0, L) \gg 1$ if $|\lambda_0| \gg 1$ or $|\lambda_0| \ll 1$.

Now we consider relative coefficient-wise perturbations. Note that $\rho_0, \rho_1 \leq 1$ and $\rho_0 + \rho_1 \geq 1$. Thus,

- if $|\lambda_0| \leq 1$,

$$\kappa_a(\lambda_0, L) \geq \frac{\rho_0 + \rho_1|\lambda_0|}{2} \quad \text{and} \quad \kappa_r(\lambda_0, L) \geq \frac{\rho_0}{2|\lambda_0|} + \frac{\rho_1}{2}.$$

- if $|\lambda_0| > 1$,

$$\kappa_a(\lambda_0, L) \geq \frac{|\lambda_0|}{2}(\rho_0 + \rho_1|\lambda_0|) \geq \frac{|\lambda_0|}{2} \quad \text{and} \quad \kappa_r(\lambda_0, L) \geq \frac{\rho_0 + \rho_1|\lambda_0|}{2}.$$

Therefore,

1. $\kappa_a(\lambda_0, L) \gg 1$ if $|\lambda_0| \gg 1$.
2. $\kappa_r(\lambda_0, L) \gg 1$ if any of the following conditions hold:
   (a) $\|B_0\|_2 \approx \|B_1\|_2$, and $|\lambda_0| \gg 1$ or $|\lambda_0| \ll 1$;
   (b) $\|B_0\|_2 \gg \|B_1\|_2$ and $|\lambda_0| \ll 1$;
   (c) $\|B_0\|_2 \gg \|B_1\|_2$, $|\lambda_0| \gg 1$ and $\|B_1\|_2|\lambda_0| \gg \|B_0\|_2$;
   (d) $\|B_1\|_2 \gg \|B_0\|_2$, $|\lambda_0| \ll 1$, and $\|B_0\|_2 \gg \|B_1\|_2|\lambda_0|$;
   (e) $\|B_1\|_2 \gg \|B_0\|_2$ and $|\lambda_0| \gg 1$.

   Notice that conditions (b) and (c) (resp., (d) an (e)) are still sufficient conditions for $\kappa_r(\lambda_0, L) \gg 1$ with the less restrictive inequality $\|B_0\|_2 \geq \|B_1\|_2$ (resp., $\|B_1\|_2 \geq \|B_0\|_2$). However, we have chosen the more strict inequalities so that cases (a)-(e) are "disjoint" in the sense that each of them considers different scenarios. For example, if $\|B_0\|_2 \gtrsim \|B_1\|_2$, the condition $\|B_0\|_2 \geq \|B_1\|_2$ and $|\lambda_0| \ll 1$ (less restrictive version of (b)) is already included in (a).

We admit a certain degree of ambiguity in the meaning of the expression "$\|B_1\|_2 \approx \|B_0\|_2$" used above. 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 lead to lower bounds for $\kappa_r(\lambda_0, L)$ depending on C. We have preferred the simpler but somewhat ambiguous statement given above.

We also note that, for a generic matrix pencil, we can expect $\|B_0\|_2 \approx \|B_1\|_2$, which implies that, for these pencils, both the eigenvalues with very large and very small moduli, are not computable when relative coefficient-wise perturbations are considered, since $\kappa_r(\lambda_0, L) \gg 1$ according to the condition (a) above.

In Remark 4.5 we have given sufficient conditions for eigenvalues of pencils with very small or very large moduli to be ill-conditioned. However, these conditions are not necessary. For brevity, this is illustrated in Examples 4.6 and 4.7 by providing counterexamples only for two of these conditions. The interested reader is invited to find counterexamples for the remaining ones.

Example 4.6. We have shown in Remark 4.5 that $\kappa_a(\lambda_0, L) \gg 1$ if $|\lambda_0| \gg 1$ for the relative coefficient-wise weights $\omega_i = \|B_i\|_2$ for $i = 0, 1$. Here we present an example of a pencil with a small eigenvalue whose "absolute" non-homogeneous condition number is also large.

Let $L(\lambda) = \lambda \begin{bmatrix} 10^{14} & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 10^{-7} \end{bmatrix}$. Note that $\lambda_0 = 10^{-7}$ is a simple eigenvalue of $L(\lambda)$ and $[0, 1]^T$ is both a left and a right eigenvector of $L(\lambda)$ associated with λ_0. Moreover, from the first equality in (4.8) with $\omega_0 = \|B_0\|_2 = 1$ and $\omega_1 = \|B_1\|_2 = 10^{14}$, we get

$$\kappa_\theta((\lambda_0, 1), L) = \frac{1 + 10^7}{1 + 10^{-14}} \approx 10^7,$$

which implies that $\kappa_a(\lambda_0, L) = (1 + |\lambda_0|^2)\kappa_\theta((\lambda_0, 1), L) \approx 10^7$.

Example 4.7. We have shown in the condition (d) of Remark 4.5 that for the weights $\omega_i = \|B_i\|_2$ for $i = 0, 1$, if a pencil $L(\lambda) = \lambda B_1 + B_0$ has a simple eigenvalue λ_0 whose modulus is very close to 0 and the coefficients of $L(\lambda)$ satisfy the conditions $\|B_1\|_2 \geq \|B_0\|_2$ and $\|B_0\|_2 \gg \|B_1\|_2|\lambda_0|$, then $\kappa_r(\lambda_0, L) \gg 1$. Note that in Example 4.6, the pencil $L(\lambda)$ has a very small eigenvalue (namely, 10^{-7}), $\|B_1\|_2 \geq \|B_0\|_2$, $\|B_0\|_2 < \|B_1\|_2|\lambda_0|$, and $\kappa_r(\lambda_0, L) \approx 10^{14}$.

Our next goal is to extend the results about non-computability of eigenvalues from pencils to general matrix polynomials of arbitrary grade. These results and its consequences are presented, respectively, in Theorem 4.8 and Remark 4.9. It is worth mentioning that the bounds in Theorem 4.8 obtained for pencils (polynomials of grade 1) are not the ones in Proposition 4.4 since the manipulations used to prove Theorem 4.8 are different than those used in Proposition 4.4. In fact, both types of bounds complement each other at some extent.

The proof of Theorem 4.8 has two steps, as the one of Proposition 4.4. First, we start by providing a lower bound for $\kappa_\theta((\lambda_0, 1), P)$ for a general matrix polynomial $P(\lambda)$ and, then, we use Theorem 3.5 to obtain lower bounds for $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$. The first step requires simplifying the expression in the denominator of the formula (2.2) for $\kappa_\theta((\lambda_0, 1), P)$. Let $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$. Notice that

$$D_\alpha P(\alpha, \beta) = \sum_{i=1}^{k} i\alpha^{i-1}\beta^{k-i} B_i, \qquad \text{and} \tag{4.10}$$

$$D_\beta P(\alpha, \beta) = \sum_{i=0}^{k-1} (k-i)\alpha^i \beta^{k-i-1} B_i = \sum_{i=0}^{k} (k-i)\alpha^i \beta^{k-i-1} B_i. \tag{4.11}$$

Let $(\lambda_0, 1)$ be a simple eigenvalue of $P(\alpha, \beta)$. Evaluating (4.10) and (4.11) at the representative $[\lambda_0, 1]^T$, we get

$$\begin{aligned} D_\alpha P(\lambda_0, 1) - \overline{\lambda_0} D_\beta P(\lambda_0, 1) &= \sum_{i=1}^{k} i\lambda_0^{i-1} B_i - \overline{\lambda_0} \sum_{i=0}^{k} (k-i)\lambda_0^i B_i \\ &= (1 + |\lambda_0|^2) \sum_{i=1}^{k} i\lambda_0^{i-1} B_i - \overline{\lambda_0} k \sum_{i=0}^{k} \lambda_0^i B_i \\ &= (1 + |\lambda_0|^2) \sum_{i=1}^{k} i\lambda_0^{i-1} B_i - \overline{\lambda_0} k P(\lambda_0, 1). \end{aligned}$$

Let x and y be, respectively, a right and a left eigenvector of $P(\alpha, \beta)$ associated with $(\lambda_0, 1)$. Then, taking into account that $P(\lambda_0, 1)x = 0$, the denominator of $\kappa_\theta((\lambda_0, 1), P)$ in (2.2) is given by

$$|y^*(D_\alpha P(\lambda_0, 1) - \overline{\lambda_0} D_\beta P(\lambda_0, 1))x| = (1 + |\lambda_0|^2) \left| y^* \left(\sum_{i=1}^{k} i\lambda_0^{i-1} B_i \right) x \right|. \tag{4.12}$$

Hence,

$$\kappa_\theta((\lambda_0, 1), P) = \frac{(\sum_{i=0}^{k} |\lambda_0|^i \omega_i)\|x\|_2\|y\|_2}{(1 + |\lambda_0|^2)|y^*(\sum_{i=1}^{k} i\lambda_0^{i-1} B_i)x|} \geq \frac{(\sum_{i=0}^{k} |\lambda_0|^i \omega_i)}{(1 + |\lambda_0|^2)(\sum_{i=1}^{k} i|\lambda_0|^{i-1}\|B_i\|_2)}. \tag{4.13}$$

If $\omega_j = \max_{i=0:k}\{\|B_i\|_2\}$ for $j = 0, 1, 2, \ldots, k$, from (4.13), we get

$$\kappa_\theta((\lambda_0, 1), P) \geq \frac{\sum_{i=0}^{k} |\lambda_0|^i}{(1 + |\lambda_0|^2)(\sum_{i=1}^{k} i|\lambda_0|^{i-1})} \frac{\max_{i=0:k}\{\|B_i\|_2\}}{\max_{i=1:k}\{\|B_i\|_2\}}. \tag{4.14}$$

If $\omega_j = \|B_j\|_2$ for $j = 0, 1, 2, \ldots, k$, then, from (4.13), we get

$$\kappa_\theta((\lambda_0, 1), P) \geq \frac{\sum_{i=0}^{k} |\lambda_0|^i \|B_i\|_2}{(1 + |\lambda_0|^2)(\sum_{i=1}^{k} i|\lambda_0|^{i-1}\|B_i\|_2)}. \tag{4.15}$$

In the results that we present next we will use the following notation:

$$\mathfrak{h} := \frac{\|B_0\|_2}{\max_{i=1:k}\{\|B_i\|_2\}}, \quad \mu := \frac{\max_{i=0:k}\{\|B_i\|_2\}}{\max_{i=1:k}\{\|B_i\|_2\}} = \max\{1, \mathfrak{h}\}.$$

With the previous inequalities and definitions at hand, we are now in the position of providing in Theorem 4.8 the announced lower bounds for $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ when P is a matrix polynomial with arbitrary grade.

THEOREM 4.8. *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ be a regular matrix polynomial and let λ_0 be a finite, simple eigenvalue of $P(\lambda)$.*

1. *If* $\omega_j = \max_{i=0:k}\{\|B_i\|_2\}$ *for* $j = 0, 1, 2, \ldots, k$, *then*

$$\kappa_a(\lambda_0, P) \geq \max\left\{\frac{1}{k^2}, \frac{|\lambda_0|}{k}\right\}\mu, \quad \kappa_r(\lambda_0, P) \geq \max\left\{\frac{1}{k^2|\lambda_0|}, \frac{1}{k}\right\}\mu,$$

where the lower bound for $\kappa_r(\lambda_0, P)$ *holds only if* $\lambda_0 \neq 0$.

2. *If* $\omega_j = \|B_j\|_2$ *for* $j = 0, 1, 2, \ldots, k$, *then*

$$\kappa_a(\lambda_0, P) \geq \frac{\mathfrak{h}}{k^2 \max\{1, |\lambda_0|^{k-1}\}} + \frac{|\lambda_0|}{k}, \quad \kappa_r(\lambda_0, P) \geq \frac{\mathfrak{h}}{k^2 \max\{|\lambda_0|, |\lambda_0|^k\}} + \frac{1}{k},$$

where the lower bound for $\kappa_r(\lambda_0, P)$ *holds only if* $\lambda_0 \neq 0$.

Proof. Let us start with the relative with-respect-to-the-norm-of-the-polynomial case, that is, let $\omega_j = \max_{i=0:k}\{\|B_i\|_2\}$ for $j = 0, 1, \ldots, k$. Then, by Theorem 3.5 and (4.14),

$$\kappa_a(\lambda_0, P) \geq \frac{\sum_{i=0}^k |\lambda_0|^i}{\sum_{i=1}^k i|\lambda_0|^{i-1}}\mu \geq \frac{1 + |\lambda_0| \sum_{i=1}^k |\lambda_0|^{i-1}}{k \sum_{i=1}^k |\lambda_0|^{i-1}}\mu = \left(\frac{1}{k \sum_{i=1}^k |\lambda_0|^{i-1}} + \frac{|\lambda_0|}{k}\right)\mu.$$

Taking into account that, if $|\lambda_0| \leq 1$, then $\sum_{i=1}^k |\lambda_0|^{i-1} \leq k$, and if $|\lambda_0| > 1$, then $\sum_{i=1}^k |\lambda_0|^{i-1} \leq k|\lambda_0|^{k-1}$, we get

$$\kappa_a(\lambda_0, P) \geq \left(\frac{1}{k^2 \max\{1, |\lambda_0|^{k-1}\}} + \frac{|\lambda_0|}{k}\right)\mu.$$

The results for the relative with-respect-to-the-norm-of-the-polynomial case follow easily from this expression and $\kappa_r(\lambda_0, P) = \kappa_a(\lambda_0, P)/|\lambda_0|$.

Let us consider now the relative coefficient-wise case, that is, let $\omega_j = \|B_j\|_2$, for $j = 0, 1, \ldots, k$. Then, by Theorem 3.5 and (4.15),

$$\begin{aligned}\kappa_a(\lambda_0, P) &\geq \frac{\sum_{i=0}^k |\lambda_0|^i \|B_i\|_2}{(\sum_{i=1}^k i|\lambda_0|^{i-1}\|B_i\|_2)} \geq \frac{\|B_0\|_2 + |\lambda_0| \sum_{i=1}^k |\lambda_0|^{i-1}\|B_i\|_2}{k \sum_{i=1}^k |\lambda_0|^{i-1}\|B_i\|_2} \\ &\geq \frac{\|B_0\|_2}{k \max_{i=1:k}\{\|B_i\|_2\} \sum_{i=1}^k |\lambda_0|^{i-1}} + \frac{|\lambda_0|}{k}.\end{aligned}$$

The results follow from $\sum_{i=1}^k |\lambda_0|^{i-1} \leq k \max\{1, |\lambda_0|^{k-1}\}$ and $\kappa_r(\lambda_0, P) = \kappa_a(\lambda_0, P)/|\lambda_0|$.
□

Remark 4.9. The lower bounds on $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ presented in Theorem 4.8 allow to determine, in a straightforward way, the following sufficient conditions for the non-computability of the eigenvalue λ_0, when either $|\lambda_0| \gg 1$ or $|\lambda_0| \ll 1$. In order to obtain these conditions, note that, if $\mu > 1$, then $\mu = \mathfrak{h}$.

• For the weights $\omega_j = \max_{i=0:k}\{\|B_i\|_2\}$, $j = 0, 1, \ldots, k$:
(a) $\kappa_a(\lambda_0, P) \gg 1$ if $|\lambda_0| \gg k$ or if $\|B_0\|_2 \gg k^2 \max_{i=1:k}\{\|B_i\|_2\}$.
(b) $\kappa_r(\lambda_0, P) \gg 1$ if $|\lambda_0| \ll \frac{1}{k^2}$ or if $\|B_0\|_2 \gg k \max_{i=1:k}\{\|B_i\|_2\}$.
• For the weights $\omega_j = \|B_j\|_2$ for $j = 0, 1, \ldots, k$:
(c) $\kappa_a(\lambda_0, P) \gg 1$ if $|\lambda_0| \gg k$ or if $\|B_0\|_2 \gg k^2 \max_{i=1:k}\{\|B_i\|_2\} \max\{1, |\lambda_0|^{k-1}\}$.
(d) $\kappa_r(\lambda_0, P) \gg 1$ if $\|B_0\|_2 \gg k^2 \max_{i=1:k}\{\|B_i\|_2\} \max\{|\lambda_0|, |\lambda_0|^k\}$.

Next, we give an example to illustrate some of the results in Theorem 4.8 and Remark 4.9.

EXAMPLE 4.10. Let us consider the quadratic matrix polynomial given in Example 3.8, whose eigenvalues are $10^{-5}, 10^5, 10^{10}$, and 10^{15}. Note that $\mathfrak{h} = \frac{10^{15}}{10^{-5}+10^{15}} \approx 1$. The sufficient conditions in Remark 4.9 imply, when $\omega_i = \|B_i\|_2$ for $i = 0, 1, 2$, that

$$\kappa_a(10^5, P), \kappa_a(10^{10}, P), \kappa_a(10^{15}, P) \gg 1,$$

and

$$\kappa_r(10^{-5}, P) \gg 1.$$

Notice that the result for $\lambda_0 = 10^{-5}$ follows from the fact that the condition $\|B_0\|_2 \gg k^2 \max_{i=1:k}\{\|B_i\|_2\} \max\{|\lambda_0|, |\lambda_0|^k\}$, which is equivalent to $10^{15} \gg 4\,|10^{-5}+10^{15}|\,10^{-5}$, is true. These conclusions can be confirmed by the values of the condition numbers in the table in Example 3.8. We stress that the conditions in Remark 4.9 are not necessary. For brevity, we illustrate this observation with some of the results obtained in Example 3.8. Note that $\kappa_r(10^5, P)$ and $\kappa_r(10^{10}, P)$ are both large but $\|B_0\|_2 < k^2 \max_{i=1:k}\{\|B_i\|_2\} \max\{|\lambda_0|, |\lambda_0|^k\}$ in both cases.

We highlight that whenever $\kappa_\theta((\lambda_0, 1), P)$ is moderate (of order one) and $|\lambda_0|$ is large, then $\kappa_a(\lambda_0, P)$ and $\kappa_r(\lambda_0, P)$ are both always large, as exemplified by the eigenvalue 10^5 of this example and guaranteed by Theorem 3.5 and its consequence (4.6). However, we also highlight that it is possible to have $\kappa_r(\lambda_0, P) \approx 1$ for very large $|\lambda_0|$, as exemplified by the eigenvalue 10^{15} in this example, since it is possible to have $\kappa_\theta((\lambda_0, 1), P) \ll 1$ in these situations (recall again the table in Example 3.8). In the jargon we are using in this section, this means that there exist very large non-homogenous eigenvalues which are computable and that this happens if and only if the corresponding homogeneous condition number is very close to zero as a consequence of (3.3).

Finally, we present Corollary 4.11, which is an informal result that summarizes in a concise way all the sufficient conditions obtained in this section for the non computability of eigenvalues of pencils with small or large moduli. This corollary is, in fact, obtained from the results in Remarks 4.5 and 4.9, when particularized for $k = 1$. It is worth highlighting that the sufficient conditions that, in Corollary 4.11, guarantee that $\kappa_r(\lambda_0, L) \gg 1$ for the weights $\omega_i = \|B_i\|_2$, $i = 0, 1$, show that the ratio of the magnitudes of the monomials $|\lambda_0|\|B_1\|_2$ and $\|B_0\|_2$ plays a relevant role in determining when the relative non-homogenous condition numbers of very large and small eigenvalues of pencils are very large.

COROLLARY 4.11. *Let $L(\lambda) = \lambda B_1 + B_0$ be a regular pencil and let λ_0 be a simple, finite eigenvalue of $L(\lambda)$.*

1. *If $\omega_i = \max\{\|B_0\|_2, \|B_1\|_2\}$ for $i = 0, 1$, or $\omega_i = \|B_i\|_2$ for $i = 0, 1$, then*

   $\kappa_a(\lambda_0, L) \gg 1$, *if* $|\lambda_0| \gg 1$, *or* $|\lambda_0| \ll 1$ *and* $\|B_0\|_2 \gg \|B_1\|_2$.

2. *If $\omega_i = \max\{\|B_0\|_2, \|B_1\|_2\}$ for $i = 0, 1$ and $\lambda_0 \neq 0$, then*

   $\kappa_r(\lambda_0, L) \gg 1$, *if* $|\lambda_0| \gg 1$ *or* $|\lambda_0| \ll 1$.

3. *If $\omega_i = \|B_i\|_2$ for $i = 0, 1$ and $\lambda_0 \neq 0$, then*

   $\kappa_r(\lambda_0, L) \gg 1$, *if any of the following two conditions hold*

   (i) $|\lambda_0| \gg 1$ *and either* $|\lambda_0|\|B_1\|_2 \gg \|B_0\|_2$ *or* $|\lambda_0|\|B_1\|_2 \ll \|B_0\|_2$;
   (ii) $|\lambda_0| \ll 1$ *and* $|\lambda_0|\|B_1\|_2 \ll \|B_0\|_2$.

Proof. We omit the proofs of 1. and 2. since they are elementary and focus on the proof of 3. We notice that, in order to prove 3., it is enough to show that conditions (i) and (ii) in 3. are equivalent to conditions (a)-(e) in Remark 4.5 and condition (d) in Remark 4.9. Note that, since in this corollary we only consider eigenvalues with very large or very small moduli, we can express (d) in Remark 4.9 as

(d1) $|\lambda_0| \gg 1$ and $\|B_0\|_2 \gg \|B_1\|_2|\lambda_0|$;

(d2) $|\lambda_0| \ll 1$ and $\|B_0\|_2 \gg \|B_1\|_2|\lambda_0|$.

We also introduce a separate notation for each condition in (i) in 3.:

(i)-1 $|\lambda_0| \gg 1$ and $|\lambda_0|\|B_1\|_2 \gg \|B_0\|_2$;

(i)-2 $|\lambda_0| \gg 1$ and $|\lambda_0|\|B_1\|_2 \ll \|B_0\|_2$.

It is clear that (ii) in 3. is equivalent to (d2) and (i)-2 is equivalent to (d1). Next we show that conditions (i) and (ii) are also equivalent to conditions (a)-(e) in Remark 4.5. We consider three scenarios:

Case I: Assume that $\|B_0\|_2 \approx \|B_1\|_2$. Then, condition (i)-1 is equivalent to $|\lambda_0| \gg 1$ and condition (ii) in 3. is equivalent to $|\lambda_0| \ll 1$. These two conditions together are equivalent to condition (a) in Remark 4.5.

Case II: Assume that $\|B_0\|_2 \gg \|B_1\|_2$. Then, (i)-1 is equivalent to $|\lambda_0| \gg 1$ and $|\lambda_0|\|B_1\|_2 \gg \|B_0\|_2$, which is equivalent to case (c) in Remark 4.5. Condition (ii) in 3. is equivalent to $|\lambda_0| \ll 1$, which is equivalent to case (b) in Remark 4.5.

Case III: Assume that $\|B_1\|_2 \gg \|B_0\|_2$. Then, condition (i)-1 is equivalent to $|\lambda_0| \gg 1$, which is equivalent to case (e) in Remark 4.5. Condition (ii) in 3. is equivalent to $|\lambda_0| \ll 1$ and $|\lambda_0|\|B_1\|_2 \ll \|B_0\|_2$, which is equivalent to (d) in Remark 4.5. □

5. Conclusions and future work. 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 proven exact relationships between each of the non-homogeneous condition numbers and the homogeneous condition number based on the chordal distance. This result will allow to extend results that have been proven 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. Finally, we have used these relationships to analyze cases for which very large and very small non-homogeneous eigenvalues of matrix polynomials are computable with some accuracy, i.e., are not very ill-conditioned, and we have seen that this is only possible in some rather particular situations.

Some possible future research work related to the results in this manuscript are discussed in this paragraph. Since we have only considered normwise eigenvalue condition numbers, a natural and interesting problem is to extend the results in this paper to entrywise eigenvalue condition numbers. Note, also, that we have only studied condition numbers of simple eigenvalues. Therefore, a comparison of condition numbers of multiple eigenvalues of matrix polynomials similar to the one that we have presented for simple eigenvalues is another natural question that can be analyzed in the future. Observe that this extension would require first to introduce definitions of condition numbers for multiple eigenvalues of homogeneous matrix polynomials, because the definitions currently available in the literature are only valid for non-homogeneous polynomial eigenvalue problems [8, 9].

Acknowledgements. The authors sincerely thank two anonymous referees for their positive comments and for providing helpful suggestions which contributed to improve the paper significantly.

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] M. I. Bueno, F. M. Dopico, S. Furtado, and L. Medina, *A block-symmetric linearization of odd-degree matrix polynomials with optimal eigenvalue condition number and backward error*, Calcolo, 55 (2018), 32:1-32:43.

[3] J. P. Dedieu, *Condition operators, condition numbers, and condition number theorem for the generalized eigenvalue problem*, Linear Algebra Appl., 263 (1997), 1-24.

[4] J. P. Dedieu and F. Tisseur, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), 71-94.

[5] 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.

[6] D. J. Higham and N. J. Higham, *Structured backward error and condition of generalized eigenvalue problems*, SIAM J. Matrix Anal. Appl., 20 (1998), 493-512.

[7] N. J. Higham, D. S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28, 4 (2006), 1005-1028.

[8] D. Kressner, M. J. Peláez, and J. Moro, *Structured Hölder condition numbers for multiple eigenvalues*, SIAM J. Matrix Anal. Appl., 31 (2009), 175-201.

[9] N. Papathanasiou and P. Psarrakos, *On condition numbers of polynomial eigenvalue problems*, Appl. Math. Comput., 216 (2010), 1194-1205.

[10] M. Shub and S. Smale, *Complexity of Bezout's theorem IV: Probability of success; extensions*, SIAM J. Numer. Anal., 33,1 (1996), 128-148.

[11] G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.

[12] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), 339-361.

[13] M. Van Barel and F. Tisseur, *Polynomial eigenvalue solver based on tropically scaled Lagrange linearization*, Linear Algebra Appl., 542 (2018), 186-208.