Anguas, L. M., Bueno, M. I. & Dopico, F. M. (2019). Conditioning and backward errors of eigenvalues of homogeneous matrix polynomials under Möbius transformations. *Mathematics of Computation*, 89(322), pp. 767–805.

DOI: https://doi.org/10.1090/mcom/3472

CONDITIONING AND BACKWARD ERRORS OF EIGENVALUES OF HOMOGENEOUS MATRIX POLYNOMIALS UNDER MÖBIUS TRANSFORMATIONS*

LUIS MIGUEL ANGUAS†, MARIA ISABEL BUENO‡, AND FROILÁN M. DOPICO§

Abstract. We present the first general study on the effect of Möbius transformations on the eigenvalue condition numbers and backward errors of approximate eigenpairs of polynomial eigenvalue problems (PEPs). By using the homogeneous formulation of PEPs, we are able to obtain two clear and simple results. First, we show that, if the matrix inducing the Möbius transformation is well conditioned, then such transformation approximately preserves the eigenvalue condition numbers and backward errors when they are defined with respect to perturbations of the matrix polynomial which are small relative to the norm of the whole polynomial. However, if the perturbations in each coefficient of the matrix polynomial are small relative to the norm of that coefficient, then the corresponding eigenvalue condition numbers and backward errors are preserved approximately by the Möbius transformations induced by well-conditioned matrices only if a penalty factor, depending on the norms of those matrix coefficients, is moderate. It is important to note that these simple results are no longer true if a non-homogeneous formulation of the PEP is used.

Key words. backward error, eigenvalue condition number, matrix polynomial, Möbius transformation, polynomial eigenvalue problem

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

1. Introduction.

Möbius transformations are a standard tool in the theory of matrix polynomials and in their applications. The use of Möbius transformations of matrix polynomials can be traced back to at least [27, 28], where they are defined for general rational matrices which are not necessarily polynomials. Since Möbius transformations change the eigenvalues of a matrix polynomial in a simple way and preserve most of the properties of the polynomial [26], they have often been used to transform a matrix polynomial with infinite eigenvalues into another polynomial with only finite eigenvalues and for which a certain problem can be solved more easily. Recent examples of this theoretical use can be found, for instance, in [13, 36].

A fundamental property of some Möbius transformations, called Cayley transformations, is to convert matrix polynomials with certain structures arising in control applications into matrix polynomials with other structures that also arise in applications. This allows to translate many properties from one structured class of matrix polynomials into another. The origins of these results on structured problems are found in classical group theory, where Cayley transformations are used, for instance, to transform Hamiltonian into symplectic matrices and vice versa [39]. Such results were extended to Hamiltonian and symplectic matrix pencils, i.e., matrix polynomials of degree one, in [29, 30] (with the goal of relating discrete and continuous control problems) and generalized to several classes of structured matrix polynomials

* This work was partially supported by the Ministerio de Economía, Industria y Competitividad (MINECO) of Spain through grants MTM2012-32542, 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, South Hall 6607, 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)

of degree larger than one in [25]. A thorough treatment of the properties of Möbius transformations of matrix polynomials is presented in a unified way in [26].

The Cayley transformations mentioned in the previous paragraph are not just of theoretical interest, since they, and some variants, have been used explicitly in a number of important numerical algorithms for eigenvalue problems. Some examples are: [3, Algorithm 4.1], where they are used for computing the eigenvalues of a sympletic pencil by transforming such pencil into a Hamiltonian pencil, and then using a structured eigenvalue algorithm for Hamiltonian pencils; [31, Sec. 3.2], where they are used to transform a matrix pencil into another one so that the eigenvalues in the left-half plane of the original pencil are moved into the unit disk, an operation that is a preprocessing before applying an inverse-free disk function method for computing certain stable/un-stable deflating subspaces of the matrix pencil; and [32, Sec. 6], where they are used for transforming palindromic/anti-palindromic pencils into even/odd pencils with the goal of deflating the ± 1 eigenvalues of the palindromic/anti-palindromic pencils via algorithms for deflating the infinite eigenvalues of the even/odd pencils. Other examples can be found in the literature, although, sometimes, the use of the Cayley transformations is not mentioned explicitly. For instance, the algorithm in [8] for computing the structured staircase form of skew-symmetric/symmetric pencils can be used via a Cayley transformation and its inverse for computing a structured staircase form of palindromic pencils, although this is not mentioned in [8].

The numerical use of Möbius transformations in structured algorithms for pencils discussed in the previous paragraph can be extended to matrix polynomials of degree larger than one. Assume that a structured matrix polynomial P is given and we want to solve the corresponding polynomial eigenvalue problem (PEP). Then, the standard procedure is to consider one of the (strong) linearizations L of P of the same structure available in the literature (see, for instance, [7, 11, 25]), assuming it exists. Assume also that a backward stable structured algorithm is available for a certain type of structured pencils and that L can be transformed into a pencil with such structure through a Möbius transformation, M_A. By [26, Corollary 8.6], $M_A(L)$ is a (strong) linearization of $M_A(P)$. However, even if the structured algorithm guarantees that the PEP associated with $M_A(P)$ is solved in a backward stable way [15], it is not guaranteed that it solves the PEP associated with P in a backward stable way as well. Thus, a direct way of checking if this is the case is to analyze how the Möbius transformation affects the backward errors of the computed eigenpairs of the polynomial P.

As illustrated in the previous paragraph, when the numerical solution of a problem is obtained by transforming the problem into another one, a fundamental question is whether or not such transformation deteriorates the conditioning of the problem and/or the backward errors of the approximate solutions, because a significant deterioration of such quantities may lead to unreliable solutions. We have not found in the literature any analysis of this kind concerning the use of Möbius transformations for solving PEPs, apart from a few vague comments in some papers. The results in this paper are a first step in this direction. More specifically, we present the first general study on the effect of Möbius transformations on the eigenvalue condition numbers and backward errors of approximate eigenpairs of PEPs. We are aware that this analysis does not cover all the numerical applications of Möbius transformations that can be found in the literature, since, for instance, the effect on the conditioning of the deflating subspaces of pencils is not covered in our study. For brevity, this and other related problems will be considered in future works.

In this paper, the PEP is formulated in homogeneous form [9, 10, 34] and the corresponding homogeneous eigenvalue condition numbers [10, 1] and backward errors [21] are used. This homogeneous formulation has clear mathematical advantages over the standard non-homogeneous one [10, 1] and has been used recently in the analysis of algorithms for solving PEPs via linearizations [18, 22]. In addition, when a PEP is solved by applying the QZ algorithm to a linearization of the corresponding matrix polynomial, the computed eigenvalues are, in fact, the homogeneous eigenvalues. Note that the non-homogeneous eigenvalues are obtained from the homogeneous ones by the division of its two components, and this operation is performed only after the algorithm QZ has converged. The analysis of the effect of Möbius transformations on the eigenvalue condition numbers of non-homogeneous matrix polynomials is postponed to a future paper for brevity, but also because it is cumbersome since it requires to distinguish several cases. Such complications are related to the fact that, for any Möbius transformation, it is possible to find matrix polynomials for which the modulus of some of its non-homogenous eigenvalues changes wildly under the transformation. This fact has led to the popular belief that any Möbius transformation affects dramatically the conditioning of certain critical eigenvalues, something that is not true in the homogeneous formulation.

By using the homogeneous formulation of PEPs, we are able to obtain, among many others, two clear and simple results that are highlighted in the next lines. First, we show in Theorems 5.6 and 6.2 that, if the matrix inducing the Möbius transformation is well conditioned, then, for any matrix polynomial and simple eigenvalue, such transformation approximately preserves the eigenvalue condition numbers and backward errors when, in the definition of these magnitudes, small perturbations of the matrix polynomial relative to the norm of the whole polynomial are considered. However, if we consider condition numbers and backward errors for which the perturbations in each coefficient of the matrix polynomial are small relative to the norm of that coefficient, then these magnitudes are approximately preserved by the Möbius transformations induced by well-conditioned matrices only if a penalty factor, depending on the norms of the coefficients of the polynomial, is moderate. This is proven in Theorems 5.9 and 6.2.

The paper is organized as follows. In Section 2, we introduce some notation and basic definitions about matrix polynomials. Section 3 contains results about Möbius transformations of homogeneous matrix polynomials. Most of such results are well-known, but the ones in Subsection 3.2 are new, as far as we know. In Section 4, we recall the definitions and expressions of eigenvalue condition numbers and backward errors of PEPs. Sections 5 and 6 include the most important results proven in this paper about the effect of Möbius transformations on eigenvalue condition numbers and backward errors of approximate eigenpairs of PEPs. Numerical experiments that illustrate the theoretical results in previous sections are described in Section 7. Finally, Section 8 discusses the conclusions and some lines of future research.

2. Notation and basic definitions. To begin with, let us introduce some general notation that will be used throughout this paper. Given positive integers a and b, we use the abbreviation

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

For any real number α, $\lfloor \alpha \rfloor$ denotes the largest integer smaller than or equal to α. The field of complex numbers is denoted by $\mathbb{C}$. For any complex vector $x = [x_1, \ldots, x_n]^T \in$

$\mathbb{C}^n$, $\|x\|_p$ denotes its p-norm, i.e., $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$, for $1 \leq p < \infty$. We also denote $\|x\|_\infty := \max_{i=1:n}\{|x_i|\}$. For any complex matrix $A \in \mathbb{C}^{m\times n}$, $\|A\|_2$ denotes its spectral or 2-norm, that is, its largest singular value; $\|A\|_\infty$ denotes its ∞-norm, that is, the maximum row sum of the moduli of its entries; and $\|A\|_1$ denotes its 1-norm, that is, the maximum column sum of the moduli of its entries. Additionally, $\|A\|_M := \max\{|A_{ij}|, i = 1 : m, j = 1 : n\}$ denotes the max norm of A.

Let us present now some basic concepts that will be used frequently in this paper.

A matrix polynomial $P(\alpha, \beta)$ is said to be a *homogeneous matrix polynomial of degree k* if it is of the form

$$P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i, \quad B_i \in \mathbb{C}^{m\times n}, \tag{2.1}$$

where all the matrix coefficients B_i but one are allowed to be zero. If all matrix coefficients B_i are zero, then we say that P has degree $-\infty$ or it is undefined. If $m = n$ and the determinant of $P(\alpha, \beta)$ is not identically equal to zero, P is said to be *regular*. Otherwise, it is said to be *singular*.

Given a regular homogeneous matrix polynomial $P(\alpha, \beta)$, the *(homogeneous) polynomial eigenvalue problem* (PEP) associated to $P(\alpha, \beta)$ consists of finding scalars α_0 and β_0, at least one of them nonzero, and nonzero vectors $x, y \in \mathbb{C}^n$ such that

$$y^* P(\alpha_0, \beta_0) = 0 \quad \text{and} \quad P(\alpha_0, \beta_0)x = 0. \tag{2.2}$$

Note that the equalities in (2.2) hold if and only if the equalities $y^* P(a\alpha_0, a\beta_0) = 0$ and $P(a\alpha_0, a\beta_0)x = 0$ hold for any complex number $a \neq 0$. This equivalence motivates defining the corresponding *eigenvalue* of $P(\alpha, \beta)$ as the set $(\alpha_0, \beta_0) := \{[a\alpha_0, a\beta_0]^T : a \in \mathbb{C}\} \subset \mathbb{C}^2$. The vectors x and y in (2.2) are called, respectively, a *right and a left eigenvector* of $P(\alpha, \beta)$ associated with the eigenvalue (α_0, β_0), and the pairs $(x, (\alpha_0, \beta_0))$ and $(y^*, (\alpha_0, \beta_0))$ are called, respectively, a *right and a left eigenpair* of $P(\alpha, \beta)$. We notice that an eigenvalue can be seen as a line in $\mathbb{C}^2$ passing through the origin whose points are solutions to the equation $\det(P(\alpha, \beta)) = 0$. Throughout the paper, we denote eigenvalues, i.e., lines, as (α_0, β_0) and a specific (nonzero) representative of this eigenvalue, i.e., a specific (nonzero) point on the line (α_0, β_0) in $\mathbb{C}^2$, by $[\alpha_0, \beta_0]^T$. We will also use the notation $\langle x \rangle$, where $x \in \mathbb{C}^2$, to denote the line generated by the vector x in $\mathbb{C}^2$ through scalar multiplication. In particular, $\langle [\alpha_0, \beta_0]^T \rangle = (\alpha_0, \beta_0)$. Notice that all representatives of an eigenvalue of $P(\alpha, \beta)$ are nonzero scalar multiples of each other.

In future sections, we will need to calculate the norm of a homogeneous matrix polynomial $P(\alpha, \beta)$ as in (2.1). In this paper we will use the norm $\|P\|_\infty := \max_{i=0:k}\{\|B_i\|_2\}$.

Some of the results for homogeneous matrix polynomials that will be introduced in Section 3 were proven in the literature for *non-homogeneous matrix polynomials*, that is, matrix polynomials written in the form

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i B_i, \quad B_i \in \mathbb{C}^{m\times n}. \tag{2.3}$$

To extend those results to homogeneous matrix polynomials it will be enough to notice that a homogeneous matrix polynomial $P(\alpha, \beta)$ can be rewritten in non-homogeneous form as follows

$$P(\alpha,\beta) = \begin{cases} \beta^k P(\alpha/\beta), & \text{if } \beta \neq 0, \\ \alpha^k B_k, & \text{if } \beta = 0. \end{cases} \tag{2.4}$$

When $n = m$, we say that a non-homogeneous matrix polynomial $P(\lambda)$ is regular if its determinant is not identically zero. In this case, we can consider the non-homogeneous PEP associated to $P(\lambda)$. As in the homogeneous case, it consists of finding scalars λ_0 and nonzero vectors x and y such that $y^*P(\lambda_0) = 0$ and $P(\lambda_0)x = 0$. The vectors x and y are said to be right and left eigenvectors of $P(\lambda)$ associated with the eigenvalue λ_0, and the pairs (x, λ_0) and (y^*, λ_0) are called, respectively, a right and a left eigenpair of $P(\lambda)$.

The next lemma provides a relationship between the eigenvalues and eigenvectors of a matrix polynomial when expressed in homogeneous and non-homogeneous forms. We omit its proof because it is straightforward.

Lemma 2.1. *A pair $(x, (\alpha_0, \beta_0))$ (resp. $(y^*, (\alpha_0, \beta_0))$) is a right (resp. left) eigenpair for a regular homogeneous matrix polynomial $P(\alpha, \beta)$ if and only if (x, λ_0) (resp. (y^*, λ_0)) is a right (resp. left) eigenpair for the same polynomial when expressed in non-homogeneous form, where $\lambda_0 = \alpha_0/\beta_0$ if $\beta_0 \neq 0$ and $\lambda_0 = \infty$ if $\beta_0 = 0$.*

3. Möbius transformations of homogeneous matrix polynomials. Before introducing the definition of Möbius transformation of matrix polynomials, we present some notation that will be used in this section. We denote by $GL(2, \mathbb{C})$ the set of nonsingular 2×2 matrices with complex entries and by $\mathbb{C}[\alpha, \beta]_k^{m\times n}$ the vector space of $m \times n$ homogeneous matrix polynomials of degree k whose matrix coefficients have complex entries together with the zero polynomial, that is, polynomials of the form (2.1) whose matrix coefficients are allowed to be all zero.

Next we introduce the concept of Möbius transformation.

Definition 3.1. *[26] Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$. Then the Möbius transformation on $\mathbb{C}[\alpha, \beta]_k^{m\times n}$ induced by A is the map $M_A : \mathbb{C}[\alpha, \beta]_k^{m\times n} \to \mathbb{C}[\alpha, \beta]_k^{m\times n}$ given by*

$$M_A\left(\sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i\right)(\gamma, \delta) = \sum_{i=0}^{k} (a\gamma + b\delta)^i (c\gamma + d\delta)^{k-i} B_i. \tag{3.1}$$

The matrix polynomial $M_A(P)(\gamma, \delta)$, that is, the image of $P(\alpha, \beta)$ under M_A, is said to be the Möbius transform of $P(\alpha, \beta)$ under M_A.

It is important to highlight that the Möbius transform of a homogeneous matrix polynomial P of degree k is another homogeneous matrix polynomial of the same degree.

The next example shows that the well-known reversal of a matrix polynomial $P(\alpha, \beta)$ [26] can be seen as a Möbius transform of P.

Example 3.2. Let us consider the Möbius transformation induced by the matrix $R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Given $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$, we have

$$M_R(P)(\gamma, \delta) = \sum_{i=0}^{k} \gamma^{k-i} \delta^i B_i = \text{rev } P(\gamma, \delta).$$

In the next definition, we introduce some Möbius transformations that are useful in converting some types of structured matrix polynomials into others [24, 25, 26, 30].

Definition 3.3. *The Möbius transformations induced by the matrices*

$$A_{+1} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad A_{-1} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \tag{3.2}$$

are called Cayley transformations.

3.1. Properties of Möbius transformations. In this section, we present some properties of the Möbius transformations that were proven in [26] for non-homogeneous matrix polynomials, that is, matrix polynomials of the form (2.3). The proof of the equivalent statement of those properties for homogeneous polynomials follows immediately from the results in [26] and the relationship (2.4) between the homogeneous and non-homogeneous expressions of the same matrix polynomial.

Proposition 3.4. [26, Proposition 3.5] *For any $A \in GL(2, \mathbb{C})$, M_A is a $\mathbb{C}$-linear operator on the vector space $\mathbb{C}[\alpha, \beta]_k^{m\times n}$, that is, for any $\mu \in \mathbb{C}$ and $P, Q \in \mathbb{C}[\alpha, \beta]_k^{m\times n}$,*

$$M_A(P + Q) = M_A(P) + M_A(Q) \quad \text{and} \quad M_A(\mu P) = \mu M_A(P).$$

Proposition 3.5. [26, Theorem 3.18 and Proposition 3.27] *Let $A, B \in GL(2, \mathbb{C})$ and let I_2 denote the 2×2 identity matrix. Then, when the Möbius transformations are seen as operators on $\mathbb{C}[\alpha, \beta]_k^{m\times n}$, the following properties hold:*

1. *M_{I_2} is the identity operator;*
2. *$M_A \circ M_B = M_{BA}$;*
3. *$(M_A)^{-1} = M_{A^{-1}}$;*
4. *$M_{\mu A} = \mu^k M_A$, for any nonzero $\mu \in \mathbb{C}$;*
5. *If $m = n$, then $\det(M_A(P)) = M_A(\det(P))$, where the Möbius transformation on the right-hand side operates on $\mathbb{C}[\alpha, \beta]_{nk}^{1\times 1}$.*

Remark 3.6. An immediate consequence of Proposition 3.5(5.) is that P is a regular matrix polynomial if and only if $M_A(P)$ is.

The following result provides a connection between the eigenpairs of a regular homogeneous matrix polynomial $P(\alpha, \beta)$ and the eigenpairs of a Möbius transform $M_A(P)(\gamma, \delta)$ of $P(\alpha, \beta)$. As the previous properties, this result was proven for non-homogeneous matrix polynomials in [26]. It is easy to see that an analogous result follows when P is expressed in homogeneous form using (2.4) and Lemma 2.1.

Lemma 3.7. [26, Remark 6.12 and Theorem 5.3] *Let $P(\alpha, \beta)$ be a regular homogeneous matrix polynomial and let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$. If $(x, (\alpha_0, \beta_0))$ (resp. $(y^*, (\alpha_0, \beta_0))$) is a right (resp. left) eigenpair of $P(\alpha, \beta)$, then $(x, \langle A^{-1}[\alpha_0, \beta_0]^T \rangle)$ (resp. $(y^*, \langle A^{-1}[\alpha_0, \beta_0]^T \rangle)$) is a right (resp. left) eigenpair of $M_A(P)(\gamma, \delta)$. Moreover, (α_0, β_0), as an eigenvalue of $P(\alpha, \beta)$, has the same algebraic multiplicity as $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$, when considered an eigenvalue of $M_A(P)(\gamma, \delta)$. In particular, (α_0, β_0) is a simple eigenvalue of $P(\alpha, \beta)$ if and only if $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$ is a simple eigenvalue of $M_A(P)(\gamma, \delta)$.*

Motivated by the previous result, we introduce the following definition.

Definition 3.8. *Let $P(\alpha, \beta)$ be a regular homogeneous matrix polynomial and let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$. Let (α_0, β_0) be an eigenvalue of $P(\alpha, \beta)$ and let $[\alpha_0, \beta_0]^T$ be a representative of (α_0, β_0). Then, we call $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$ the* eigenvalue of $M_A(P)$ associated with the eigenvalue (α_0, β_0) of $P(\alpha, \beta)$ *and we call $A^{-1}[\alpha_0, \beta_0]^T$ the* representative of the eigenvalue of $M_A(P)$ associated with $[\alpha_0, \beta_0]^T$.

In the following remark we explain how to compute an explicit expression for the components of the vector $A^{-1}[\alpha_0, \beta_0]^T$.

Remark 3.9. We recall that, for $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$,

$$A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)}, \tag{3.3}$$

where $\operatorname{adj}(A)$ denotes the adjugate of the matrix A, given by

$$\operatorname{adj}(A) := \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Thus, given a simple eigenvalue (α_0, β_0) of a homogeneous matrix polynomial P and a representative $[\alpha_0, \beta_0]^T$ of (α_0, β_0), the components of the representative $[\gamma_0, \delta_0]^T := A^{-1}[\alpha_0, \beta_0]^T$ of the eigenvalue of $M_A(P)$ associated with $[\alpha_0, \beta_0]^T$ are given by

$$\gamma_0 := \frac{d\alpha_0 - b\beta_0}{\det(A)}, \quad \delta_0 := \frac{a\beta_0 - c\alpha_0}{\det(A)}. \tag{3.4}$$

The following fact, which follows taking into account that $\|\operatorname{adj}(A)\|_\infty = \|A\|_1$ and $\|\operatorname{adj}(A)\|_1 = \|A\|_\infty$, will be used to simplify the bounds on the quotients of condition numbers presented in Section 5:

$$\frac{1}{|\det(A)|} = \frac{\|A^{-1}\|_\infty}{\|A\|_1} = \frac{\|A^{-1}\|_1}{\|A\|_\infty}. \tag{3.5}$$

3.2. The matrix coefficients of the Möbius transform of a matrix polynomial. When comparing the condition number of an eigenvalue of a regular homogeneous matrix polynomial $P(\alpha, \beta)$ with the condition number of the associated eigenvalue of the Möbius transform $M_A(P)$ of P, it will be useful to have an explicit expression for the coefficients of the matrix polynomial $M_A(P)$ in terms of the matrix coefficients of P, as well as an upper bound on the 2-norm of each coefficient of $M_A(P)$ in terms of the norms of the coefficients of P. We provide such expression and upper bound in the following proposition.

Proposition 3.10. *Let* $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i \in \mathbb{C}[\alpha, \beta]_k^{m\times n}$, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$*, and* M_A *be the Möbius transformation induced by* A *on* $\mathbb{C}[\alpha, \beta]_k^{m\times n}$*. Then,* $M_A(P)(\gamma, \delta) = \sum_{\ell=0}^{k} \gamma^\ell \delta^{k-\ell} \widetilde{B}_\ell$*, where*

$$\widetilde{B}_\ell = \sum_{i=0}^{k} \sum_{j=0}^{k-\ell} \binom{i}{j} \binom{k-i}{k-j-\ell} a^{i-j} b^j c^{j+\ell-i} d^{k-j-\ell} B_i, \quad \ell = 0 : k, \tag{3.6}$$

and $\binom{s}{t} := 0$ *for* $s < t$*. Moreover,*

$$\|\widetilde{B}_\ell\|_2 \le \|A\|_\infty^k \binom{k}{\lfloor k/2 \rfloor} \sum_{i=0}^{k} \|B_i\|_2, \quad \ell = 0 : k. \tag{3.7}$$

Proof. By the Binomial Theorem,

$$(a\gamma + b\delta)^i = \sum_{j=0}^{i} \binom{i}{j} (a\gamma)^{i-j} (b\delta)^j, \quad (c\gamma + d\delta)^{k-i} = \sum_{r=0}^{k-i} \binom{k-i}{r} (c\gamma)^{k-i-r} (d\delta)^r.$$

Thus, from (3.1) we get

$$\begin{aligned}
M_A(P)(\gamma,\delta) &= \sum_{i=0}^{k}\sum_{j=0}^{i}\sum_{r=0}^{k-i}\binom{i}{j}\binom{k-i}{r}\gamma^{k-j-r}\delta^{j+r}a^{i-j}b^{j}c^{k-i-r}d^{r}B_i\\
&= \sum_{i=0}^{k}\sum_{j=0}^{i}\sum_{\ell=i-j}^{k-j}\binom{i}{j}\binom{k-i}{k-j-\ell}\gamma^{\ell}\delta^{k-\ell}a^{i-j}b^{j}c^{j+\ell-i}d^{k-j-\ell}B_i\\
&= \sum_{\ell=0}^{k}\sum_{i=0}^{k}\sum_{j=\max\{0,i-\ell\}}^{\min\{i,k-\ell\}}\binom{i}{j}\binom{k-i}{k-j-\ell}\gamma^{\ell}\delta^{k-\ell}a^{i-j}b^{j}c^{j+\ell-i}d^{k-j-\ell}B_i\\
&= \sum_{\ell=0}^{k}\sum_{i=0}^{k}\sum_{j=0}^{k-\ell}\binom{i}{j}\binom{k-i}{k-j-\ell}\gamma^{\ell}\delta^{k-\ell}a^{i-j}b^{j}c^{j+\ell-i}d^{k-j-\ell}B_i,
\end{aligned}$$

where the second equality follows by applying the change of variable $\ell = k-j-r$, and the fourth equality follows because if $i < k-\ell$ and $j > i$, then $\binom{i}{j} = 0$, and if $i-\ell > 0$ and $j < i-\ell$, then $\binom{k-i}{k-\ell-j} = 0$. Hence, (3.6) follows.

From (3.6), we have

$$\begin{aligned}
\|\widetilde{B}_\ell\|_2 &\leq \sum_{i=0}^{k}\sum_{j=0}^{k-\ell}\binom{i}{j}\binom{k-i}{k-j-\ell}|a|^{i-j}|b|^{j}|c|^{j+\ell-i}|d|^{k-j-\ell}\|B_i\|_2\\
&\leq \|A\|_M^k\sum_{i=0}^{k}\binom{k}{k-l}\|B_i\|_2 \leq \|A\|_\infty^k\sum_{i=0}^{k}\binom{k}{k-l}\|B_i\|_2,
\end{aligned}$$

where the second inequality follows from the Chu-Vandermonde identity [2]

$$\sum_{j=0}^{k-\ell}\binom{i}{j}\binom{k-i}{k-\ell-j} = \binom{k}{k-\ell}. \tag{3.8}$$

The inequality in (3.7) follows taking into account

$$\binom{k}{k-\ell} \leq \binom{k}{\lfloor k/2\rfloor}, \quad 0 \leq \ell \leq k,$$

see [5]. □

4. Eigenvalue condition numbers and backward errors of matrix polynomials. To measure the change of the condition number of a simple eigenvalue (α_0, β_0) of a regular homogeneous matrix polynomial $P(\alpha,\beta)$ of degree k when a Möbius transformation is applied to $P(\alpha,\beta)$, two eigenvalue condition numbers may be considered. They were called the *Dedieu-Tisseur* condition number and the *Stewart-Sun* condition number in [1], and were originally introduced in [10] and [34], respectively. In [1, Corollary 3.3], it was proven that the Stewart-Sun and the Dedieu-Tisseur eigenvalue condition numbers differ at most by a factor $\sqrt{k+1}$ and, so, that they are equivalent. Therefore, it is enough to focus on studying the influence of Möbius transformations on just one of these two condition numbers, since the corresponding results for the other one can be immediately obtained from [1, Corollary 3.3]. We focus on the Stewart-Sun condition number in this paper for two reasons: 1) the Stewart-Sun condition number is easier to define than the Dedieu-Tisseur condition number and its definition provides a geometric intuition of the change in the eigenvalue that it measures; 2) the use of the Stewart-Sun condition number will allow us to study easily in the future the effect of Möbius transformations on the Wilkinson-like condition number of a simple eigenvalue of a non-homogeneous matrix polynomial [35]. This is a consequence of Theorem 3.5 in [1], which provides a simple connection between this non-homogeneous condition number and the Stewart-Sun condition number. Such connection is more involved when the Dedieu-Tisseur condition number is considered.

We start by recalling the definition of the Stewart-Sun eigenvalue condition number, which is expressed in terms of the chordal distance whose definition we present next.

DEFINITION 4.1. [34, 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*

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

where

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

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

DEFINITION 4.2. *(Stewart-Sun condition number) 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 degree k and let x be a right eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0). We define*

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

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.

The next theorem presents an explicit formula for this condition number.

THEOREM 4.3. [1, Theorem 2.13] *Let (α_0, β_0) be a simple eigenvalue of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$, 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 Stewart-Sun eigenvalue condition number of (α_0, β_0) is given by*

$$\kappa_\theta((\alpha_0, \beta_0), P) = \left(\sum_{i=0}^{k} |\alpha_0|^i |\beta_0|^{k-i} \omega_i\right) \frac{\|y\|_2 \|x\|_2}{|y^*(\overline{\beta_0} D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0} D_\beta P(\alpha_0, \beta_0))x|}, \tag{4.1}$$

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

It is important to note that the explicit expression for $\kappa_\theta((\alpha_0, \beta_0), P)$ does not depend on the choice of representative of the eigenvalue (α_0, β_0).

In the explicit expression for $\kappa_\theta((\alpha_0, \beta_0), P)$, the weights ω_i can be chosen in different ways leading to different variants of this condition number. In the following definition, we introduce the three types of weights (and the corresponding condition numbers) considered in this paper.

DEFINITION 4.4. *With the same notation and assumptions as in Theorem 4.3:*

1. *The* absolute eigenvalue condition number *of* (α_0, β_0) *is defined by taking* $\omega_i = 1$ *for* $i = 0 : k$ *in* $\kappa_\theta((\alpha_0, \beta_0), P)$ *and is denoted by* $\kappa_\theta^a((\alpha_0, \beta_0), P)$.
2. *The* relative eigenvalue condition number with respect to the norm of P *of* (α_0, β_0) *is defined by taking* $\omega_i = \|P\|_\infty = \max_{j=0:k}\{\|B_j\|_2\}$ *for* $i = 0 : k$ *in* $\kappa_\theta((\alpha_0, \beta_0), P)$ *and is denoted by* $\kappa_\theta^p((\alpha_0, \beta_0), P)$.
3. *The* relative eigenvalue condition number *of* (α_0, β_0) *is defined by taking* $\omega_i = \|B_i\|_2$ *for* $i = 0 : k$ *in* $\kappa_\theta((\alpha_0, \beta_0), P)$ *and is denoted by* $\kappa_\theta^r((\alpha_0, \beta_0), P)$.

The absolute eigenvalue condition number in Definition 4.4 does not correspond to perturbations in the coefficients of P appearing in applications, but it is studied because its analysis is the simplest one. Quoting Nick Higham [20, p. 56], "it is the relative condition number that is of interest, but it is more convenient to state results for the absolute condition number". The relative eigenvalue condition number with respect to the norm of P corresponds to perturbations in the coefficients of P coming from the backward errors of solving PEPs by applying a backward stable generalized eigenvalue algorithm to any reasonable linearization of P [14, 38]. Observe that $\kappa_\theta^p((\alpha_0, \beta_0), P) = \|P\|_\infty \, \kappa_\theta^a((\alpha_0, \beta_0), P)$ and, therefore, one of these condition numbers can be easily computed from the other. Finally, the relative eigenvalue condition number corresponds to perturbations in the coefficients of P coming from an "ideal" coefficientwise backward stable algorithm for the PEP. Unfortunately, nowadays, such "ideal" algorithm exists only for degrees $k = 1$ (the QZ algorithm for generalized eigenvalue problems) and $k = 2$, in this case via linearizations and delicate scalings of P [16, 18, 40]. The recent work [37] shows that there is still some hope of finding an "ideal" algorithm for PEPs with degree $k > 2$.

In this paper, we will also compare the backward errors of approximate right and left eigenpairs of the Möbius transform $M_A(P)$ of a homogeneous matrix polynomial P with the backward errors of approximate right and left eigenpairs of P constructed from those of $M_A(P)$. Next we introduce the definition of backward errors of approximate eigenpairs of a homogeneous matrix polynomial. This definition was presented in [21] based on the original definition given in [35] for non-homogeneous matrix polynomials.

DEFINITION 4.5. *Let* $(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))$ *be an approximate right eigenpair of the regular matrix polynomial* $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i$. *We define the backward error of* $(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))$ *as*

$$\eta_P(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0})) := \min\{\epsilon : (P(\widehat{\alpha_0}, \widehat{\beta_0}) + \Delta P(\widehat{\alpha_0}, \widehat{\beta_0}))\widehat{x} = 0, \|\Delta B_i\|_2 \le \epsilon\, \omega_i, i = 0 : k\},$$

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. Similarly, for an approximate left eigenpair* $(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0}))$, *we define*

$$\eta_P(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0})) := \min\{\epsilon : \widehat{y}^*(P(\widehat{\alpha_0}, \widehat{\beta_0}) + \Delta P(\widehat{\alpha_0}, \widehat{\beta_0})) = 0, \|\Delta B_i\|_2 \le \epsilon\, \omega_i, i = 0 : k\}.$$

In Theorem 4.6, we present the explicit formulas proved in [21, 35] to compute the backward errors introduced in Definition 4.5. Before that, we point out that this definition of backward error does not require that the polynomial $P + \Delta P$ at which the minimum ϵ is attained is regular. This poses a fundamental difficulty that apparently has not been considered before in the literature[1], because for a singular square matrix polynomial $Q(\alpha,\beta)$ it is well-known that neither the equation $\widehat{y}^*Q(\widehat{\alpha_0},\widehat{\beta_0})=0$ nor the equation $Q(\widehat{\alpha_0},\widehat{\beta_0})\widehat{x}=0$ guarantee that $(\widehat{\alpha_0},\widehat{\beta_0})$ is an eigenvalue of $Q(\alpha,\beta)$ [12]. Thus, if we want to say that the approximate right eigenpair $(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ (resp. left eigenpair $(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0})))$ of P is the exact right (resp. left) eigenpair of a polynomial $P+\Delta P$ with the backward errors η_P given in Theorem 4.6, this perturbed polynomial $P+\Delta P$ where the minimum in Definition 4.5 is attained should be regular as well. We have not been able to prove that, but, in Appendix A, we prove that, for any arbitrarily small positive number ϕ, we can find a *regular* matrix polynomial $P+\delta P=\sum_{i=0}^{k}\alpha^i\beta^{k-i}(B_i+\delta B_i)$ for which the approximate right (resp. left) eigenpair of P is an exact eigenpair and such that $\|\delta B_i\|_2 \leq (\eta_P+\phi)\,\omega_i$ for $i=0,\ldots,k$. Thus, the formulas for the backward errors presented in the next theorem are still meaningful as a measure of the backward errors under the additional restriction that $P+\Delta P$ is regular, because ϕ can be chosen smaller than the smallest positive floating point number and its presence does not affect the computed value of the backward errors.

THEOREM 4.6. [21, 35] *Let $(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$ be, respectively, an approximate right and an approximate left eigenpair of the regular matrix polynomial $P(\alpha,\beta)=\sum_{i=0}^{k}\alpha^i\beta^{k-i}B_i$. Then,*

1. $\eta_P(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))=\dfrac{\|P(\widehat{\alpha_0},\widehat{\beta_0})\widehat{x}\|_2}{(\sum_{i=0}^{k}|\widehat{\alpha_0}|^i|\widehat{\beta_0}|^{k-i}\omega_i)\|\widehat{x}\|_2}$, *and*
2. $\eta_P(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))=\dfrac{\|\widehat{y}^*P(\widehat{\alpha_0},\widehat{\beta_0})\|_2}{(\sum_{i=0}^{k}|\widehat{\alpha_0}|^i|\widehat{\beta_0}|^{k-i}\omega_i)\|\widehat{y}\|_2}$.

As in the case of condition numbers, the weights in Definition 4.5 can be chosen in different ways. We will consider the same three choices as in Definition 4.4, which leads to the following definition.

DEFINITION 4.7. *With the same notation and assumptions as in Definition 4.5:*

1. *The* absolute backward errors *of $(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$ are defined by taking $\omega_i=1$ for $i=0:k$ in $\eta_P(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $\eta_P(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$, and are denoted by $\eta_P^a(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $\eta_P^a(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$.*
2. *The* relative backward errors with respect to the norm of P *of $(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$ are defined by taking $\omega_i=\|P\|_\infty$ for $i=0:k$, and are denoted by $\eta_P^p(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $\eta_P^p(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$.*
3. *The* relative backward errors *of $(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$ are defined by taking $\omega_i=\|B_i\|_2$ for $i=0:k$, and are denoted by $\eta_P^r(\widehat{x},(\widehat{\alpha_0},\widehat{\beta_0}))$ and $\eta_P^r(\widehat{y}^*,(\widehat{\alpha_0},\widehat{\beta_0}))$.*

5. Effect of Möbius transformations on eigenvalue condition numbers. This section contains the most important results of this paper (presented in Subsection 5.2), which are obtained from the key and technical Theorem 5.1 (included in Subsection 5.1). In Subsection 5.3 we present some additional results.

Throughout this section, $P(\alpha,\beta)\in\mathbb{C}[\alpha,\beta]_k^{n\times n}$ is a regular homogeneous matrix polynomial and (α_0,β_0) is a simple eigenvalue of $P(\alpha,\beta)$. Moreover, M_A is a Möbius transformation on $\mathbb{C}[\alpha,\beta]_k^{n\times n}$ and $\langle A^{-1}[\alpha_0,\beta_0]^T\rangle$ is the eigenvalue of $M_A(P)$ associated with (α_0,β_0) introduced in Definition 3.8. We are interested in studying the

[1] This difficulty was pointed out to the authors by an anonymous referee. The authors sincerely thank this referee for the suggestion of studying this question.

influence of the Möbius transformation M_A on the Stewart-Sun eigenvalue condition number, that is, we would like to compare the Stewart-Sun condition numbers of (α_0, β_0) and $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$. More precisely, our goal is to determine sufficient conditions on A, P and $M_A(P)$ so that the condition number of $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$ is similar to that of (α_0, β_0), *independently of the particular eigenvalue* (α_0, β_0) *that is considered.* With this goal in mind, we first obtain upper and lower bounds on the quotient

$$Q_\theta := \frac{\kappa_\theta(\langle A^{-1}[\alpha_0, \beta_0]^T \rangle, M_A(P))}{\kappa_\theta((\alpha_0, \beta_0), P)} \tag{5.1}$$

which are independent of (α_0, β_0) and, then, we find conditions that make these upper and lower bounds approximately equal to one or, more precisely, moderate numbers.

In view of Definition 4.4, three variants of the quotient (5.1), denoted by Q_θ^a, Q_θ^p, and Q_θ^r, are considered, which correspond, respectively, to quotients of absolute, relative with respect to the norm of the polynomial, and relative eigenvalue condition numbers. *The lower and upper bounds for* Q_θ^a *and* Q_θ^p *are presented in Theorems 5.4 and 5.6 and depend only on* A *and the degree* k *of* P. So, these bounds lead to very simple sufficient conditions, valid for *all* polynomials and simple eigenvalues, that allow us to identify some Möbius transformations which do not significantly change the condition numbers. *The lower and upper bounds for* Q_θ^r *are presented in Theorem 5.9 and depend only on* A*, the degree* k *of* P*, and some ratios of the norms of the matrix coefficients of* P *and* $M_A(P)$. These bounds also lead to simple sufficient conditions, valid for *all* simple eigenvalues but only for certain matrix polynomials, that allow us to identify some Möbius transformations which do not significantly change the condition numbers.

The first obstacle we have found in obtaining the results described in the previous paragraph is that a direct application of Theorem 4.3 leads to a very complicated expression for the quotient Q_θ in (5.1). Therefore, in Theorem 5.1 we deduce an expression for Q_θ that depends only on (α_0, β_0), the matrix A inducing the Möbius transformation, and the weights $\widetilde{\omega}_i$ and ω_i used in $\kappa_\theta(\langle A^{-1}[\alpha_0, \beta_0]^T \rangle, M_A(P))$ and $\kappa_\theta((\alpha_0, \beta_0), P)$, respectively. Thus, this expression gets rid of the partial derivatives of P and $M_A(P)$.

5.1. A derivative-free expression for the quotient of condition numbers. The derivative-free expression for Q_θ obtained in this section is (5.2). Before diving into the details of its proof, we emphasize that, even though the formula for the Stewart-Sun condition number is independent of the representative of the eigenvalue, the expression (5.2) is independent of the particular representative $[\alpha_0, \beta_0]^T$ chosen for the eigenvalue (α_0, β_0) of P but *not* of the representative of the associated eigenvalue of $M_A(P)$, which must be $A^{-1}[\alpha_0, \beta_0]^T$. A second remarkable feature of (5.2) is that it depends on the determinant of the matrix A inducing the Möbius transformation. Note also that $\det(A)$ cannot be removed by choosing a different representative of $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$.

THEOREM 5.1. *Let* $P(\alpha, \beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i \in \mathbb{C}[\alpha, \beta]_k^{n \times n}$ *be a regular homogeneous matrix polynomial and let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$*. Let* $M_A(P)(\gamma, \delta) = \sum_{i=0}^k \gamma^i \delta^{k-i} \widetilde{B}_i \in \mathbb{C}[\alpha, \beta]_k^{n \times n}$ *be the Möbius transform of* $P(\alpha, \beta)$ *under* M_A*. Let* (α_0, β_0) *be a simple eigenvalue of* $P(\alpha, \beta)$ *and let* $[\alpha_0, \beta_0]^T$ *be a representative of* (α_0, β_0)*. Let* $[\gamma_0, \delta_0]^T := A^{-1}[\alpha_0, \beta_0]^T$ *be the representative of the eigenvalue of* $M_A(P)$ *associated with* $[\alpha_0, \beta_0]^T$*. Let* Q_θ *be as in (5.1) and let* ω_i *and* $\widetilde{\omega}_i$ *be the* *weights in the definition of the Stewart-Sun eigenvalue condition number associated with the eigenvalues* (α_0, β_0) *and* $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$*, respectively. Then,*

$$Q_\theta = \frac{\sum_{i=0}^{k} |\gamma_0|^i \, |\delta_0|^{(k-i)} \, \widetilde{\omega}_i}{|\det(A)| \sum_{i=0}^{k} |\alpha_0|^i |\beta_0|^{(k-i)} \omega_i} \frac{|\alpha_0|^2 + |\beta_0|^2}{|\gamma_0|^2 + |\delta_0|^2}. \tag{5.2}$$

Moreover, (5.2) is independent of the choice of representative for (α_0, β_0)*.*

Proof. In order to prove the formula (5.2), we compute $\kappa_\theta(\langle A^{-1}[\alpha_0, \beta_0]^T \rangle, M_A(P))$ and $\kappa_\theta((\alpha_0, \beta_0), P)$ separately, and then calculate their quotient. Since the definition of the Stewart-Sun eigenvalue condition number is independent of the choice of representative of the eigenvalue, when computing the condition numbers of (α_0, β_0) and $\langle A^{-1}[\alpha_0, \beta_0]^T \rangle$, we have freedom to choose any representative. In this proof, we choose an arbitrary representative $[\alpha_0, \beta_0]^T$ of (α_0, β_0) and, once $[\alpha_0, \beta_0]^T$ is fixed, we choose $[\gamma_0, \delta_0]^T := A^{-1}[\alpha_0, \beta_0]^T$ as the representative of the eigenvalue of $M_A(P)$ associated with (α_0, β_0).

We first compute $\kappa_\theta((\alpha_0, \beta_0), P)$. Let x and y be, respectively, a right and a left eigenvector of $P(\alpha, \beta)$ associated with (α_0, β_0). We start by simplifying the denominator of (4.1). Note that

$$D_\alpha P(\alpha, \beta) = \sum_{i=1}^{k} i\alpha^{i-1}\beta^{k-i} B_i, \qquad \text{and} \tag{5.3}$$

$$D_\beta P(\alpha, \beta) = \sum_{i=0}^{k-1} (k-i)\alpha^i \beta^{k-i-1} B_i = \sum_{i=0}^{k} (k-i)\alpha^i \beta^{k-i-1} B_i. \tag{5.4}$$

We consider two cases.

Case I: Assume that $\beta_0 \neq 0$. Evaluating (5.3) and (5.4) at $[\alpha_0, \beta_0]^T$, we get

$$\begin{aligned}
&\overline{\beta_0} D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0} D_\beta P(\alpha_0, \beta_0) \\
&= |\beta_0|^2 \sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1} B_i - \overline{\alpha_0} \sum_{i=0}^{k} (k-i)\alpha_0^i \beta_0^{k-i-1} B_i \\
&= (|\beta_0|^2 + |\alpha_0|^2) \sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1} B_i - \overline{\alpha_0} k \sum_{i=0}^{k} \alpha_0^i \beta_0^{k-i-1} B_i.
\end{aligned}$$

Moreover,

$$\begin{aligned}
&|y^*(\overline{\beta_0} D_\alpha P(\alpha_0, \beta_0) - \overline{\alpha_0} D_\beta P(\alpha_0, \beta_0))x| \\
&= \left| y^* \left((|\beta_0|^2 + |\alpha_0|^2) \sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1} B_i - \frac{\overline{\alpha_0} k}{\beta_0} \sum_{i=0}^{k} \alpha_0^i \beta_0^{k-i} B_i \right) x \right| \\
&= (|\beta_0|^2 + |\alpha_0|^2) \left| y^* \left(\sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1} B_i \right) x \right|,
\end{aligned} \tag{5.5}$$

where the last equality follows from $P(\alpha_0, \beta_0)x = 0$. Thus, if $\beta_0 \neq 0$,

$$\kappa_\theta((\alpha_0, \beta_0), P) = \frac{\left(\sum_{i=0}^{k} |\alpha_0|^i |\beta_0|^{(k-i)} \omega_i\right) \|y\|_2 \|x\|_2}{(|\beta_0|^2 + |\alpha_0|^2) \left| y^* \left(\sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1} B_i\right) x \right|}. \tag{5.6}$$

Case II: If $\beta_0 = 0$, evaluating (5.4) at $[\alpha_0, \beta_0]^T$, we get that the denominator of (4.1) is $|\alpha_0|^k |y^* B_{k-1} x|$. Thus,

$$\kappa_\theta((\alpha_0, \beta_0), P) = \left(\sum_{i=0}^{k} |\alpha_0|^i |\beta_0|^{(k-i)} \omega_i \right) \frac{\|y\|_2 \|x\|_2}{|\alpha_0|^k |y^* B_{k-1} x|}. \tag{5.7}$$

Next, we compute $\kappa_\theta(\langle[\gamma_0, \delta_0]^T\rangle, M_A(P))$ and express it in terms of the coefficients of P. As above, we start by simplifying the denominator of (4.1) when $P(\alpha, \beta)$ is replaced by $M_A(P)(\gamma, \delta)$ and $[\alpha_0, \beta_0]^T$ is replaced by $[\gamma_0, \delta_0]^T$. Recall that, by Lemma 3.7, x and y are, respectively, a right and a left eigenvector of $M_A(P)$ associated with $\langle[\gamma_0, \delta_0]^T\rangle$. Note that, since $M_A(P)(\gamma, \delta) = \sum_{i=0}^{k} (a\gamma + b\delta)^i (c\gamma + d\delta)^{k-i} B_i$, we have

$$\begin{aligned} D_\gamma M_A(P)(\gamma, \delta) &= \sum_{i=1}^{k} ai(a\gamma + b\delta)^{i-1}(c\gamma + d\delta)^{k-i} B_i \\ &\quad + \sum_{i=0}^{k} c(k-i)(a\gamma + b\delta)^i (c\gamma + d\delta)^{k-i-1} B_i, \end{aligned} \tag{5.8}$$

$$\begin{aligned} D_\delta M_A(P)(\gamma, \delta) &= \sum_{i=1}^{k} bi(a\gamma + b\delta)^{i-1}(c\gamma + d\delta)^{k-i} B_i \\ &\quad + \sum_{i=0}^{k} d(k-i)(a\gamma + b\delta)^i (c\gamma + d\delta)^{k-i-1} B_i. \end{aligned} \tag{5.9}$$

Again, we consider two cases.

Case I: Assume that $\beta_0 \neq 0$. We evaluate (5.8) and (5.9) at $[\gamma_0, \delta_0]^T = \left[\frac{d\alpha_0 - b\beta_0}{\det(A)}, \frac{a\beta_0 - c\alpha_0}{\det(A)} \right]$, and get

$$\begin{aligned} D_\gamma M_A(P)(\gamma_0, \delta_0) &= \left[a \sum_{i=1}^{k} i\alpha_0^{i-1} \beta_0^{k-i} B_i + c \sum_{i=0}^{k} (k-i) \alpha_0^i \beta_0^{k-i-1} B_i \right] \\ &= \left[(a\beta_0 - c\alpha_0) \sum_{i=1}^{k} i\alpha_0^{i-1} \beta_0^{k-i-1} B_i + ck \sum_{i=0}^{k} \alpha_0^i \beta_0^{k-i-1} B_i \right] \\ &= \left[\det(A)\delta_0 \sum_{i=1}^{k} i\alpha_0^{i-1} \beta_0^{k-i-1} B_i + \frac{ck}{\beta_0} P(\alpha_0, \beta_0) \right]. \end{aligned}$$

An analogous computation shows that

$$D_\delta M_A(P)(\gamma_0, \delta_0) = \left[-\det(A)\gamma_0 \sum_{i=1}^{k} i\alpha_0^{i-1} \beta_0^{k-i-1} B_i + \frac{dk}{\beta_0} P(\alpha_0, \beta_0) \right].$$

This implies that,

$$\begin{aligned} &|y^*(\overline{\delta_0} D_\gamma M_A(P)(\gamma_0, \delta_0) - \overline{\gamma_0} D_\delta M_A(P)(\gamma_0, \delta_0))x| \\ &\qquad = |\det(A)|(|\delta_0|^2 + |\gamma_0|^2) \left| y^* \left(\sum_{i=1}^{k} i\alpha_0^{i-1} \beta_0^{k-i-1} B_i \right) x \right|. \end{aligned} \tag{5.10}$$

Thus, if $\beta_0 \neq 0$,

$$\kappa_\theta(\langle[\gamma_0,\delta_0]^T\rangle, M_A(P)) = \frac{\left(\sum_{i=0}^{k}|\gamma_0|^i|\delta_0|^{(k-i)}\widetilde{\omega}_i\right)\|y\|_2\|x\|_2}{|\det(A)|(|\gamma_0|^2+|\delta_0|^2)\left|y^*\left(\sum_{i=1}^{k} i\alpha_0^{i-1}\beta_0^{k-i-1}B_i\right)x\right|}. \tag{5.11}$$

Case II: If $\beta_0 = 0$, since $A[\gamma_0,\delta_0]^T = [\alpha_0,\beta_0]^T$, we deduce that $c\gamma_0 + d\delta_0 = 0$. Moreover, by (3.4), $\gamma_0 = d\alpha_0/\det(A)$ and $\delta_0 = -c\alpha_0/\det(A)$. Since x is a right eigenvector of $P(\alpha,\beta)$ with eigenvalue $(\alpha_0,0)$, we have that $0 = P(\alpha_0,0)x = \alpha_0^k B_k x$ which implies $B_k x = 0$ since $\alpha_0 \neq 0$. Using all this information and some algebraic manipulations, we get that, if $\beta_0 = 0$,

$$\kappa_\theta(\langle[\gamma_0,\delta_0]^T\rangle, M_A(P)) = \frac{\left(\sum_{i=0}^{k}|\gamma_0|^i|\delta_0|^{(k-i)}\widetilde{\omega}_i\right)\|y\|_2\|x\|_2}{|\det(A)|(|\gamma_0|^2+|\delta_0|^2)|\alpha_0|^{k-2}|y^*B_{k-1}x|}. \tag{5.12}$$

Finally we compute Q_θ. Note that, from (5.6), (5.7), (5.11) and (5.12), we get (5.2), regardless of the value of β_0. Moreover, note that (5.2) does not change if $[\alpha_0,\beta_0]^T$ is replaced by $[t\,\alpha_0, t\,\beta_0]^T$ for any complex number $t \neq 0$. □

In the spirit of Definition 4.4, when comparing the condition number of an eigenvalue (α_0,β_0) of P and the associated eigenvalue of $M_A(P)$, we will consider the three quotients introduced in the next definition.

Definition 5.2. *With the same notation and assumptions as in Theorem 5.1, we define the following three quotients of condition numbers:*

1. $Q_\theta^a := \dfrac{\kappa_\theta^a(\langle A^{-1}[\alpha_0,\beta_0]^T\rangle, M_A(P))}{\kappa_\theta^a((\alpha_0,\beta_0),P)}$, *which is called the absolute quotient.*
2. $Q_\theta^p := \dfrac{\kappa_\theta^p(\langle A^{-1}[\alpha_0,\beta_0]^T\rangle, M_A(P))}{\kappa_\theta^p((\alpha_0,\beta_0),P)}$, *which is called the relative quotient with respect to the norms of $M_A(P)$ and P.*
3. $Q_\theta^r := \dfrac{\kappa_\theta^r(\langle A^{-1}[\alpha_0,\beta_0]^T\rangle, M_A(P))}{\kappa_\theta^r((\alpha_0,\beta_0),P)}$, *which is called the relative quotient.*

Combining Definition 4.4 and the expression (5.2), we obtain immediately expressions for Q_θ^a, Q_θ^p, and Q_θ^r as explained in the following corollary.

Corollary 5.3. *With the same notation and assumptions as in Theorem 5.1:*

1. *Q_θ^a is obtained from* (5.2) *by taking $\omega_i = \widetilde{\omega}_i = 1$ for $i = 0:k$.*
2. *Q_θ^p is obtained from* (5.2) *by taking $\omega_i = \max_{j=0:k}\{\|B_j\|_2\}$ and $\widetilde{\omega}_i = \max_{j=0:k}\{\|\widetilde{B}_j\|_2\}$ for $i = 0:k$.*
3. *Q_θ^r is obtained from* (5.2) *by taking $\omega_i = \|B_i\|_2$ and $\widetilde{\omega}_i = \|\widetilde{B}_i\|_2$ for $i = 0:k$.*

5.2. Eigenvalue-free bounds on the quotients of condition numbers. The first goal of this section is to find lower and upper bounds on the quotients Q_θ^a, Q_θ^p, and Q_θ^r, introduced in Definition 5.2, that are independent of the considered eigenvalues. The second goal is to provide simple sufficient conditions guaranteeing that the obtained bounds are moderate numbers, i.e., not far from one. The bounds on Q_θ^a are obtained from the expression (5.2) in Theorem 5.1. The proofs of the bounds on Q_θ^p and Q_θ^r also require (5.2), but, in addition, Proposition 3.10 is used.

The bounds on Q_θ^p and Q_θ^r can be expressed in terms of the condition number of the matrix $A \in GL(2,\mathbb{C})$ that induces the Möbius transformation. We will use the infinite condition number of A, that is,

$$\mathrm{cond}_\infty(A) := \|A\|_\infty \|A^{-1}\|_\infty.$$

In contrast, the bounds on Q_θ^a are expressed in terms of $\|A^{-1}\|_\infty$ and $\|A\|_\infty$, which can be considered as the "absolute" condition numbers of the matrices A and A^{-1}, respectively. In order to see this recall [19, Theorem 6.5] that

$$\frac{1}{\mathrm{cond}_\infty(A)} = \min\left\{\frac{\|\Delta A\|_\infty}{\|A\|_\infty} :\ A+\Delta A \text{ singular}\right\}, \tag{5.13}$$

that is, $\mathrm{cond}_\infty(A)$ is the reciprocal of the relative distance of A to the set of singular matrices. From (5.13), we also get

$$\frac{1}{\|A^{-1}\|_\infty} = \min\left\{\|\Delta A\|_\infty :\ A+\Delta A \text{ singular}\right\}.$$

Thus, $\|A^{-1}\|_\infty$ is the reciprocal of the absolute distance of A to the set of singular matrices and we can see it as an absolute condition number of A. In summary, the conditioning of A with respect to the singularity, either absolute or relative, has a key influence on Q_θ^a, Q_θ^p and Q_θ^r.

It is interesting to highlight that the bounds on the quotients Q_θ^a, Q_θ^p, and Q_θ^r in Theorems 5.4, 5.6, and 5.9 will require different types of proofs for polynomials of degree $k=1$ and for polynomials of degree $k\geq 2$. In fact, this is related to actual differences in the behaviours of these quotients for polynomials of degree 1 and larger than 1 when the matrix A inducing the Möbius transformation is ill-conditioned. These questions are studied in Subsection 5.3.

The next theorem presents the announced upper and lower bounds on Q_θ^a.

THEOREM 5.4. *Let $P(\alpha,\beta)\in\mathbb{C}[\alpha,\beta]_k^{n\times n}$ be a regular homogeneous matrix polynomial and let $A\in GL(2,\mathbb{C})$. Let (α_0,β_0) be a simple eigenvalue of $P(\alpha,\beta)$ and let $\langle A^{-1}[\alpha_0,\beta_0]^T\rangle$ be the eigenvalue of $M_A(P)(\gamma,\delta)$ associated with (α_0,β_0). Let Q_θ^a be the absolute quotient in Definition 5.2(1.) and let $S_k := 4(k+1)$.*

1. *If $k=1$, then*

$$\frac{1}{2\|A\|_\infty} \leq Q_\theta^a \leq 2\|A^{-1}\|_\infty.$$

2. *If $k\geq 2$, then*

$$\frac{\|A^{-1}\|_\infty}{S_k\,\|A\|_\infty^{k-1}} \leq Q_\theta^a \leq S_k\,\frac{\|A^{-1}\|_\infty^{k-1}}{\|A\|_\infty}.$$

Proof. Let $A=\begin{bmatrix} a & b\\ c & d\end{bmatrix}$. As in the proof of Theorem 5.1, we choose an arbitrary representative $[\alpha_0,\beta_0]^T$ of (α_0,β_0), and the associated representative $[\gamma_0,\delta_0]^T := A^{-1}[\alpha_0,\beta_0]^T = \left[\frac{d\alpha_0-b\beta_0}{\det(A)},\frac{a\beta_0-c\alpha_0}{\det(A)}\right]$ of the eigenvalue of $M_A(P)$. We obtain first the upper bounds.

If $k=1$, then, from Corollary 5.3(1.) and (3.5), and recalling that $\frac{1}{\sqrt{2}}\|x\|_1\leq\|x\|_2\leq\|x\|_1$ for every 2×1 vector x, we get

$$Q_\theta^a = \frac{1}{|\det(A)|}\frac{(|\gamma_0|+|\delta_0|)(|\alpha_0|^2+|\beta_0|^2)}{(|\alpha_0|+|\beta_0|)(|\gamma_0|^2+|\delta_0|^2)} = \frac{1}{|\det(A)|}\frac{\|[\gamma_0,\delta_0]^T\|_1\|[\alpha_0,\beta_0]^T\|_2^2}{\|[\alpha_0,\beta_0]^T\|_1\|[\gamma_0,\delta_0]^T\|_2^2} \tag{5.14}$$

$$\leq \frac{2}{|\det(A)|}\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1} \leq \frac{2\|A\|_1}{|\det(A)|} = 2\|A^{-1}\|_\infty. \tag{5.15}$$

If $k \geq 2$, then by using again Corollary 5.3(1.) and (3.5), and recalling that $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{2}\|x\|_\infty$ for every 2×1 vector x, we have

$$Q_\theta^a \leq \frac{(k+1)}{|\det(A)|} \frac{\max\{|\gamma_0|^k, |\delta_0|^k\}}{(|\alpha_0|^k + |\beta_0|^k)} \frac{\|[\alpha_0, \beta_0]^T\|_2^2}{\|[\gamma_0, \delta_0]^T\|_2^2} \tag{5.16}$$

$$\leq \frac{2(k+1)}{|\det(A)|} \frac{\|[\gamma_0, \delta_0]^T\|_\infty^{k-2}}{\|[\alpha_0, \beta_0]^T\|_\infty^{k-2}} \tag{5.17}$$

$$\leq 2(k+1) \frac{\|A^{-1}\|_\infty^{k-2}}{|\det(A)|} = 2(k+1) \frac{\|A^{-1}\|_\infty^{k-1}}{\|A\|_1}, \tag{5.18}$$

and the upper bound for Q_θ^a follows taking into account that $\|A\|_1 \geq \frac{\|A\|_\infty}{2}$. To obtain the lower bounds, note that Proposition 3.5(2.) implies

$$\frac{1}{Q_\theta^a} = \frac{\kappa_\theta^a((\alpha_0, \beta_0), P)}{\kappa_\theta^a(\langle A^{-1}[\alpha_0, \beta_0]^T\rangle, M_A(P))} = \frac{\kappa_\theta^a((\alpha_0, \beta_0), M_{A^{-1}}(M_A(P)))}{\kappa_\theta^a(\langle A^{-1}[\alpha_0, \beta_0]^T\rangle, M_A(P))}. \tag{5.19}$$

The previously obtained upper bounds can be applied to the right-most quotient in (5.19) with A and A^{-1} interchanged. This leads to the lower bounds for Q_θ^a. □

Remark 5.5. (Discussion on the bounds in Theorem 5.4)

$$\|A^{-1}\|_\infty \approx 1 \quad \text{and} \quad \|A\|_\infty \approx 1 \tag{5.20}$$

are sufficient to imply that all the bounds in Theorem 5.4 are moderate numbers since the factor in the bounds depending on k is small for moderate k. Therefore, the conditions (5.20), which involve only A, guarantee that the Möbius transformation M_A *does not significantly change the absolute eigenvalue condition number of any simple eigenvalue of any matrix polynomial*. Observe that (5.20) implies, in particular, that $\text{cond}_\infty(A) \approx 1$, although the reverse implication does not hold.

For $k = 1$ and $k > 2$, the conditions (5.20) are also necessary for the bounds in Theorem 5.4 to be moderate numbers. This is obvious for $k = 1$. For $k > 2$, note that $\|A^{-1}\|_\infty / \|A\|_\infty^{k-1} \approx 1$ and $\|A^{-1}\|_\infty^{k-1} / \|A\|_\infty \approx 1$ imply $\|A\|_\infty^{k^2-2k} \approx 1$ and $\|A^{-1}\|_\infty^{k^2-2k} \approx 1$, and, thus, $\|A\|_\infty \approx 1$ and $\|A^{-1}\|_\infty \approx 1$.

However, the quadratic case $k = 2$ is different because the bounds in Theorem 5.4 can be moderate in cases in which the conditions (5.20) are not satisfied. For $k = 2$, the lower and upper bounds are $\|A^{-1}\|_\infty / (12\,\|A\|_\infty)$ and $12\,\|A^{-1}\|_\infty / \|A\|_\infty$, which are moderate under the unique necessary and sufficient condition $\|A^{-1}\|_\infty \approx \|A\|_\infty$.

Notice that the very important Cayley transformations introduced in Definition 3.3 satisfy $\|A\|_\infty = 2$ and $\|A^{-1}\|_\infty = 1$ and, so, they satisfy (5.20). The same happens for the reversal Möbius transformation in Example 3.2 since $\|R\|_\infty = \|R^{-1}\|_\infty = 1$.

The next theorem presents the bounds on Q_θ^p. As explained in the proof, these bounds can be readily obtained from combining Theorem 5.4 and Proposition 3.10.

Theorem 5.6. *Let $P(\alpha, \beta) \in \mathbb{C}[\alpha, \beta]_k^{n \times n}$ be a regular homogeneous matrix polynomial and let $A \in GL(2, \mathbb{C})$. Let (α_0, β_0) be a simple eigenvalue of $P(\alpha, \beta)$ and let $\langle A^{-1}[\alpha_0, \beta_0]^T\rangle$ be the eigenvalue of $M_A(P)(\gamma, \delta)$ associated with (α_0, β_0). Let Q_θ^p be the relative quotient with respect to the norms of $M_A(P)$ and P in Definition 5.2(2.) and let $Z_k := 4(k+1)^2\binom{k}{\lfloor k/2 \rfloor}$.*

1. *If $k = 1$, then*

$$\frac{1}{4\,\text{cond}_\infty(A)} \leq Q_\theta^p \leq 4\,\text{cond}_\infty(A).$$

2. If $k \geq 2$, then

$$\frac{1}{Z_k \operatorname{cond}_\infty(A)^{k-1}} \leq Q_\theta^p \leq Z_k \operatorname{cond}_\infty(A)^{k-1}.$$

Proof. We only prove the upper bounds, since the lower bounds can be obtained from the upper bounds using an argument similar to the one used in (5.19). Notice that parts (1.) and (2.) in Corollary 5.3 and Proposition 3.10 imply

$$Q_\theta^p = Q_\theta^a \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}} \leq Q_\theta^a\,(k+1)\binom{k}{\lfloor k/2 \rfloor}\|A\|_\infty^k. \tag{5.21}$$

Now, the upper bounds follow from the upper bounds on Q_θ^a in Theorem 5.4. □

Remark 5.7. (Discussion on the bounds in Theorem 5.6) The first observation on the bounds presented in Theorem 5.6 is that the factor Z_k, depending only on the degree k of P, becomes very large even for moderate values of k (consider, for instance, $k = 15$). This fact makes the lower and upper bounds very different from each other, even for matrices A whose condition number is close to 1, and, so, Theorem 5.6 is useless for large k from a strictly rigurous point of view. However, even for such large k, the bounds in Theorem 5.6 reveal that the main source of potential instability, with respect to the conditioning of eigenvalues, of applying a Möbius transformation to any matrix polynomial is the possible ill-conditioning of A. In this context, it is worth emphasizing that the bounds in Theorem 5.6 are extreme *a priori* bounds, which do not include *any* information on the polynomial $P(\alpha, \beta)$ or on the considered eigenvalues and, so, we cannot expect that they are precise. In Theorem 5.11, we will provide much sharper (and much more complicated) *a posteriori* bounds at the cost of including the eigenvalues and the coefficients of both P and $M_A(P)$ in the expression of the bounds. The presence of the large factor Z_k in the bounds of Theorem 5.6 may be seen at the light of the "classic philosophy" of Jim Wilkinson about error analysis [17, p. 64]: "There is still a tendency to attach too much importante to the precise error bounds obtained by an *a priori* error analysis. In my opinion, the bound itself is usually the least important part of it. The main object of such an analysis is to expose the potential instabilities, if any, of an algorithm ... Usually the bound itself is weaker than it might have been because of the necessity of restricting the mass of detail to a reasonable level..." As we will comment below, this point of view is fully supported in our case by the numerical experiments in Section 7, for which the factor Z_k is very pessimistic. These numerical experiments and the discussion in this paragraph motivate us to consider the factor Z_k as the less important part of the bounds in Theorem 5.6 and to call the part depending on $\operatorname{cond}_\infty(A)$ the *essential part* of these bounds.

In addition to the discussion in the previous paragraph, we stress that in many important applications of matrix polynomials, k is very small and so is Z_k [4]. For instance, the linear case $k = 1$ (generalized eigenvalue problem) and the quadratic case $k = 2$ (quadratic eigenvalue problem) are particularly important. Therefore, in these important cases, we can state that Theorem 5.6 proves rigorously that the ill-conditioning of A is the only potential source of a significant change of the eigenvalue condition numbers under Möbius transformations.

In the rest of the discussion on the bounds of Theorem 5.6, we focus on the essential part of these bounds (i.e., the part depending on $\operatorname{cond}_\infty(A)$) and, so, we emphasize once again that such discussion is strictly rigorous *only for small values of*

k. In our opinion, Theorem 5.6 is the most illuminating result in this paper because it refers to the comparison of condition numbers that are very interesting in numerical applications (recall the comments in the paragraph just after Definition 4.4) and also because it delivers a very clear sufficient condition that guarantees that the Möbius transformation M_A *does not significantly change the relative eigenvalue condition number with respect to the norm of the polynomial of any simple eigenvalue of any matrix polynomial of small degree* k. This sufficient condition is simply that the matrix A is well-conditioned, since $\mathrm{cond}_\infty(A) \approx 1$ if and only if the lower and upper bounds in Theorem 5.6 are moderate numbers. Notice that the very important Cayley transformations in Definition 3.3 satisfy $\mathrm{cond}_\infty(A) = 2$ and that the reversal Möbius transformation in Example 3.2 satisfies $\mathrm{cond}_\infty(R) = 1$.

Remark 5.8. (Conjectures) We will see in all of the many numerical experiments in Section 7 that the factor Z_k in the bounds of Theorem 5.6 is very pessimistic, i.e., although there is an observable dependence on the true values of the quotient Q_θ^p (and also of Q_θ^r) on k, such dependence is much smaller than the one predicted by Z_k. Thus, it is tempting to conjecture that Z_k can be replaced by a much smaller constant C_k for moderate or large values of k. However, in our opinion, such a conjecture is audacious at the current stage of knowledge, since many other numerical tests with families of matrix polynomials generated in highly nonstandard manners are still possible and, perhaps, Z_k can be almost attained for a very particular family of matrix polynomials. Therefore, at present, we simply conjecture in a vague probabilistic sense that, for *most* matrix polynomials with moderate or large degree and for most simple eigenvalues, the constant Z_k can be replaced by a much smaller quantity while the bounds still hold.

In the last part of this subsection, we present and discuss the bounds on Q_θ^r. As previously announced, these bounds depend on A, P, and $M_A(P)$ and, so, are qualitatively different from the bounds on Q_θ^a and Q_θ^p presented in Theorems 5.4 and 5.6, which only depend on A and the degree k of P. In order to simplify the bounds, we will assume that the matrix coefficients with indices 0 and k of P and $M_A(P)$ (i.e. B_0, B_k, $\widetilde{B}_0$ and $\widetilde{B}_k$) are different from zero, which covers the most interesting cases in applications.

Theorem 5.9. *Let* $P(\alpha,\beta) = \sum_{i=0}^k \alpha^i \beta^{k-i} B_i \in \mathbb{C}[\alpha,\beta]_k^{n\times n}$ *be a regular homogeneous matrix polynomial and let* $A \in GL(2,\mathbb{C})$. *Let* (α_0,β_0) *be a simple eigenvalue of* $P(\alpha,\beta)$ *and let* $\langle A^{-1}[\alpha_0,\beta_0]^T\rangle$ *be the eigenvalue of* $M_A(P)(\gamma,\delta) = \sum_{i=0}^k \gamma^i \delta^{k-i} \widetilde{B}_i$ *associated with* (α_0,β_0). *Let* Q_θ^r *be the relative quotient in Definition 5.2(3.) and let* $Z_k := 4(k+1)^2\binom{k}{\lfloor k/2\rfloor}$. *Assume that* $B_0 \neq 0, B_k \neq 0, \widetilde{B}_0 \neq 0$, *and* $\widetilde{B}_k \neq 0$ *and define*

$$\rho := \frac{\max\limits_{i=0:k}\{\|B_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}}, \quad \widetilde{\rho} := \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|\widetilde{B}_0\|_2, \|\widetilde{B}_k\|_2\}}. \tag{5.22}$$

1. *If* $k = 1$, *then*

$$\frac{1}{4\,\mathrm{cond}_\infty(A)\,\widetilde{\rho}} \leq Q_\theta^r \leq 4\,\mathrm{cond}_\infty(A)\,\rho.$$

2. *If* $k \geq 2$, *then*

$$\frac{1}{Z_k\,\mathrm{cond}_\infty(A)^{k-1}\,\widetilde{\rho}} \leq Q_\theta^r \leq Z_k\,\mathrm{cond}_\infty(A)^{k-1}\,\rho.$$

Proof. We only prove the upper bounds, since the lower bounds can be obtained from the upper bounds using a similar argument to that used in (5.19).

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Select an arbitrary representative $[\alpha_0, \beta_0]^T$ of (α_0, β_0), and consider the representative $[\gamma_0, \delta_0]^T := A^{-1}[\alpha_0, \beta_0]^T = \left[\frac{d\alpha_0 - b\beta_0}{\det(A)}, \frac{a\beta_0 - c\alpha_0}{\det(A)}\right]$ of the eigenvalue of $M_A(P)$ associated with (γ_0, δ_0).

If $k = 1$, then, from Corollary 5.3(3.), (5.14) and (5.15), we obtain

$$Q_\theta^r \leq \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}} Q_\theta^a \leq \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}} 2\,\|A^{-1}\|_\infty.$$

Proposition 3.10 implies $\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\} \leq 2\,\|A\|_\infty \max\limits_{i=0:k}\{\|B_i\|_2\}$, which combined with the previous inequality yields the upper bound for $k = 1$.

If $k \geq 2$, then, from Corollary 5.3(3.) and the inequalities (5.16) and (5.18), we get

$$\begin{aligned} Q_\theta^r &\leq \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}} \frac{(k+1)}{|\det(A)|} \frac{\max\{|\gamma_0|^k, |\delta_0|^k\}}{(|\alpha_0|^k + |\beta_0|^k)} \frac{2\max\{|\beta_0|^2, |\alpha_0|^2\}}{\max\{|\delta_0|^2, |\gamma_0|^2\}} \\ &\leq \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}} 2(k+1) \frac{\|A^{-1}\|_\infty^{k-1}}{\|A\|_1}, \end{aligned}$$

which combined with Proposition 3.10 and $\|A\|_1 \geq \|A\|_\infty/2$ yields the upper bound for $k \geq 2$. □

Remark 5.10. (Discussion on the bounds in Theorem 5.9) The only difference between the bounds in Theorem 5.9 and those in Theorem 5.6 is that the former can be obtained from the latter by multiplying the upper bounds by ρ and dividing the lower bounds by $\widetilde{\rho}$. Moreover, since $\rho \geq 1$ and $\widetilde{\rho} \geq 1$, the bounds in Theorem 5.9 are moderate numbers if and only if the ones in Theorem 5.6 are and $\rho \approx 1 \approx \widetilde{\rho}$. Thus, as long as the degree k of the polynomial is small, the three conditions $\mathrm{cond}_\infty(A) \approx 1$, $\rho \approx 1$, and $\widetilde{\rho} \approx 1$ are sufficient to imply that all the bounds in Theorem 5.9 are moderate numbers and guarantee that the Möbius transformation M_A *does not significantly change the relative eigenvalue condition numbers of any eigenvalue of a matrix polynomial* P *satisfying* $\rho \approx 1$ *and* $\widetilde{\rho} \approx 1$. Note that the presence of ρ and $\widetilde{\rho}$ is natural, since ρ has appeared previously in a number of results that compare the relative eigenvalue condition numbers of a matrix polynomial and of some of its linearizations [22, 6]. As in the case of the bounds in Theorem 5.6, all the numerical experiments presented in Section 7 (as well as many others) indicate that for large values of k the factor Z_k is also very pessimistic for the bounds in Theorem 5.9, therefore, a probabilistic conjecture similar to that in Remark 5.8 can be stated for the bounds in Theorem 5.9.

5.3. Bounds involving eigenvalues for Möbius transformations induced by ill-conditioned matrices. The bounds in Theorem 5.4 on Q_θ^a are very satisfactory for low degree matrix polynomials under the sufficient conditions $\|A\|_\infty \approx \|A^{-1}\|_\infty \approx 1$ since, then, the lower and upper bounds are moderate numbers not far from one for small values of k. The same happens with the bounds on Q_θ^p in Theorem 5.6 under the sufficient condition $\mathrm{cond}_\infty(A) \approx 1$, and for the bounds in Theorem 5.9 on Q_θ^r, with the two additional conditions $\rho \approx \widetilde{\rho} \approx 1$. Obviously, these bounds are no longer satisfactory for any value of the degree k if $\mathrm{cond}_\infty(A) \gg 1$, i.e., if the Möbius transformation is induced by an ill-conditioned matrix, since the lower and upper bounds are very different from each other and do not give any information about the true values of Q_θ^a, Q_θ^p, and Q_θ^r. Note, in particular, that, for any ill-conditioned A, the upper bounds in Theorems 5.6 and 5.9 are much larger than 1, while the lower bounds are much smaller than 1.

Although we do not know any Möbius transformation M_A with $\mathrm{cond}_\infty(A) \gg 1$ that is useful in applications and we do not see currently any reason for using such transformations, we consider them in this section for completeness and in connection with the attainability of the bounds in such situation.

For brevity, we limit our discussion to the bounds on the quotients Q_θ^a and Q_θ^p, since the presence of ρ and $\widetilde{\rho}$ in Theorem 5.9 complicates the discussion on Q_θ^r.

We start by obtaining in Theorem 5.11 sharper upper and lower bounds on Q_θ^a and Q_θ^p at the cost of involving the eigenvalues and the coefficients of P and $M_A(P)$ in the expressions of the new bounds. The reader will notice that in Theorem 5.11, we are using the 1-norm for degree $k = 1$ and the ∞-norm for degree $k \geq 2$. The reason for these different choices of norms is that they lead to sharper bounds in each case. Obviously, in the case $k = 1$, we can also use the ∞-norm at the cost of worsening somewhat the bounds on Q_θ^a and Q_θ^p.

THEOREM 5.11. *Let $P(\alpha,\beta) = \sum_{i=0}^k \alpha^i\beta^{k-i}B_i \in \mathbb{C}[\alpha,\beta]_k^{n\times n}$ be a regular homogeneous matrix polynomial and let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2,\mathbb{C})$. Let $M_A(P)(\gamma,\delta) = \sum_{i=0}^k \gamma^i\delta^{k-i}\widetilde{B}_i \in \mathbb{C}[\alpha,\beta]_k^{n\times n}$ be the Möbius transform of $P(\alpha,\beta)$ under M_A. Let (α_0,β_0) be a simple eigenvalue of $P(\alpha,\beta)$ and let $\langle A^{-1}[\alpha_0,\beta_0]^T\rangle$ be the eigenvalue of $M_A(P)$ associated with (α_0,β_0). Let $[\alpha_0,\beta_0]^T$ be an arbitrary representative of (α_0,β_0) and let $[\gamma_0,\delta_0]^T := A^{-1}[\alpha_0,\beta_0]^T$ be the associated representative of $\langle A^{-1}[\alpha_0,\beta_0]^T\rangle$. Let Q_θ^a and Q_θ^p be the quotients in Definition 5.2(1.) and (2.), respectively.*

1. *If $k = 1$, then*

$$\frac{1}{2|\det(A)|}\,\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1} \leq Q_\theta^a \leq \frac{2}{|\det(A)|}\,\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1},$$

$$\frac{1}{2|\det(A)|}\,\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1}\,\frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}} \leq Q_\theta^p \leq \frac{2}{|\det(A)|}\,\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1}\,\frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}}.$$

2. *If $k \geq 2$, then*

$$\frac{1}{2(k+1)\,|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2} \leq Q_\theta^a \leq \frac{2(k+1)}{|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2},$$

$$\frac{1}{2(k+1)\,|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2}\frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}} \leq Q_\theta^p,$$

$$Q_\theta^p \leq \frac{2(k+1)}{|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2}\frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}}.$$

Moreover, the bounds in this theorem are sharper than those in Theorems 5.4 and 5.6. That is, each upper (resp. lower) bound in the previous inequalities is smaller (resp. larger) than or equal to the corresponding upper (resp. lower) bound in Theorems 5.4 and 5.6.

Proof. We only need to prove the bounds for Q_θ^a. The bounds for Q_θ^p follow immediately from the bounds for Q_θ^a and the equality in (5.21). For $k=1$, the upper bound on Q_θ^a can be obtained from (5.15); the lower bound follows easily from (5.14) through an argument similar to the one leading to the upper bound. For $k \geq 2$, the upper bound on Q_θ^a is just (5.17), and the lower bound follows easily from Corollary 5.3(1.) and (5.2) through an argument similar to the one leading to the upper bound.

Next, we prove that the bounds in this theorem are sharper than those in Theorems 5.4 and 5.6. For the upper bounds in Theorem 5.4, this follows from (5.15) for $k=1$ and the inequality (5.18) for $k \geq 2$. The corresponding results for the lower bounds in Theorem 5.4 follow from a similar argument. Note that the bounds in Theorem 5.6 can be obtained from the ones in this theorem in two steps: first bounds on $\|[\alpha_0,\beta_0]^T\|_1/\|[\gamma_0,\delta_0]^T\|_1$, for $k=1$, and on $\|[\gamma_0,\delta_0]^T\|_\infty/\|[\alpha_0,\beta_0]^T\|_\infty$, for $k \geq 2$, are obtained and, then, upper and lower bounds on $\max_{i=0:k}\{\|\widetilde{B}_i\|_2\}/\max_{i=0:k}\{\|B_i\|_2\}$ are obtained from Proposition 3.10 (the lower bounds are obtained by interchanging the roles of B_i and $\widetilde{B}_i$ and by replacing A by A^{-1}, since $P = M_{A^{-1}}(M_A(P))$). This proves that the bounds in this theorem are sharper than those in Theorems 5.4 and 5.6. □

Observe that, in contrast with Theorems 5.4 and 5.6, the lower and upper bounds in Theorem 5.11 only differ by the *linear in the degree* constant $2(k+1)$, which is very moderate from a numerical point of view (apart from being pessimistic according to our numerical tests). Thus, from Theorem 5.11, we obtain the following approximate (up to the mentioned linear constant) equalities:

$$Q_\theta^a \approx \frac{1}{|\det(A)|}\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1},\ Q_\theta^p \approx \frac{1}{|\det(A)|}\frac{\|[\alpha_0,\beta_0]^T\|_1}{\|[\gamma_0,\delta_0]^T\|_1}\frac{\max_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max_{i=0:k}\{\|B_i\|_2\}}, \quad \text{for } k=1, \tag{5.23}$$

and

$$Q_\theta^a \approx \frac{1}{|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2}, \qquad \text{for } k \geq 2, \tag{5.24}$$

$$Q_\theta^p \approx \frac{1}{|\det(A)|}\left(\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty}\right)^{k-2}\frac{\max_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max_{i=0:k}\{\|B_i\|_2\}}, \qquad \text{for } k \geq 2. \tag{5.25}$$

The approximate equalities (5.23), (5.24), and (5.25) are much simpler than the exact expressions for the quotients Q_θ^a and Q_θ^p given by (5.2), using the appropriate weights (see Corollary 5.3), and reveal clearly when the *essential parts* (i.e., those depending on $\|A\|_\infty$, $\|A^{-1}\|_\infty$, or $\text{cond}_\infty(A)$, but not containing the factors S_k or Z_k) of the lower and upper bounds in Theorems 5.4 and 5.6 are attained. An analysis of these approximate expressions leads to some interesting conclusions that are informally discussed below. Throughout this discussion, we often use expressions similar to "this bound is essentially attained" with the meaning that the essential part of that bound is attained. Note that, this discussion will illustrate, among other things, that the factor $\text{cond}_\infty(A)^{k-1}$ in Theorem 5.6 is necessary in the bounds for any value of $k \geq 2$. However, we emphasize that this discussion proves rigorously that the actual lower and upper bounds in Theorems 5.4 and 5.6 are (almost) attained *only for small values of the degree* k, i.e., when the factors S_k and Z_k are moderate. Also, for brevity, when comparing the bounds for $k=1$ and $k \geq 2$ in our analysis, we use the fact

$$\frac{\|[\alpha_0, \beta_0]^T\|_1}{\|[\gamma_0, \delta_0]^T\|_1} \approx \frac{\|[\alpha_0, \beta_0]^T\|_\infty}{\|[\gamma_0, \delta_0]^T\|_\infty}$$

without saying it explicitly.

1. *The bounds in Theorem 5.4 on Q_θ^a are essentially optimal* in the following sense: for a fixed matrix A (which is otherwise arbitrary, and so, it may be very ill-conditioned), it is always possible to find regular matrix polynomials with simple eigenvalues for which the upper bounds are essentially attained; the same happens with the lower bounds. Next we show these facts.

   For $k = 1$, (5.23) implies that the upper (resp. lower) bound in Theorem 5.4 is essentially attained for any regular pencil with a simple eigenvalue (α_0, β_0) satisfying

$$\frac{\|[\alpha_0, \beta_0]^T\|_1}{\|[\gamma_0, \delta_0]^T\|_1} = \frac{\|AA^{-1}[\alpha_0, \beta_0]^T\|_1}{\|A^{-1}[\alpha_0, \beta_0]^T\|_1} = \|A\|_1. \tag{5.26}$$

   In contrast, the lower bound is essentially attained by any regular pencil with a simple eigenvalue (α_0, β_0) such that

$$\frac{\|A^{-1}[\alpha_0, \beta_0]^T\|_1}{\|[\alpha_0, \beta_0]^T\|_1} = \|A^{-1}\|_1. \tag{5.27}$$

   Note that, for any positive integer n, a regular pencil of size $n \times n$ can be easily constructed satisfying (5.26) (resp. (5.27)): just take a diagonal pencil with a main-diagonal entry having the desired eigenvalue as a root.

   For $k = 2$, (5.24) implies that $Q_\theta^a \approx 1/|\det(A)| = \|A^{-1}\|_\infty/\|A\|_1$. So, the quotient Q_θ^a is independent of the eigenvalue and the polynomial's matrix coefficients, and is essentially always equal to both the lower and upper bounds.

   Finally, for $k > 2$, (5.24) implies that the upper (resp. lower) bound in Theorem 5.4 is essentially attained if the right (resp. left) inequality in

$$\frac{1}{\|A\|_\infty} \leq \frac{\|[\gamma_0, \delta_0]^T\|_\infty}{\|[\alpha_0, \beta_0]^T\|_\infty} \leq \|A^{-1}\|_\infty$$

   is an equality. Again, for any size $n \times n$, regular matrix polynomials with simple eigenvalues satisfying either of the two conditions can be easily constructed as diagonal matrix polynomials of degree k having a main diagonal entry with the desired eigenvalue as a root.

2. From (5.23) and (5.24), and the discussion above, we see that, for a fixed ill-conditioned matrix A (which implies that the upper and lower bounds on Q_θ^a in Theorem 5.4 are very far apart), *the behaviours of Q_θ^a for $k = 1$, $k = 2$, and $k > 2$ are very different from each other* in the following sense: If the lower (resp. upper) bound on Q_θ^a given in Theorem 5.4 is essentially attained for an eigenvalue (α_0, β_0) when $k = 1$, then the upper (resp. lower) bound on Q_θ^a is attained for the same eigenvalue when $k > 2$ (recall that the expression for Q_θ^a only depends on A, the eigenvalue (α_0, β_0) and the degree k of the matrix polynomial; also recall that for any k we can construct a matrix polynomial of degree k with a simple eigenvalue equal to (α_0, β_0)). When $k = 2$, the true value of Q_θ^a does not depend (essentially) on (α_0, β_0), according to (5.24). In this sense, the behaviours for $k = 1$ and $k > 2$ are opposite from each other, while the one for $k = 2$ can be seen as "neutral".

3. *The bounds in Theorem 5.6 on Q_θ^p are essentially optimal* in the following sense: if the matrix A is fixed, then it is always possible to find regular matrix polynomials with a simple eigenvalue for which the upper bounds on Q_θ^p are essentially attained; the same happens with the lower bounds.
Here we only discuss our claim for the upper bounds on Q_θ^p. Then, to show that the lower bounds on Q_θ^p can be attained, an argument similar to that in (5.19) can be used.
From (5.23) and (5.25), for the upper bound on Q_θ^p to be essentially attained, both the upper bound on Q_θ^a (presented in Theorem 5.4) and the upper bound on $\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}/\max\limits_{i=0:k}\{\|B_i\|_2\}$ (given in Proposition 3.10) must be essentially attained. Thus, we need to construct a regular matrix polynomial with a simple eigenvalue for which both bounds are attained simultaneously. In our discussion in item 1. above we discussed how to find an eigenvalue (α_0,β_0) attaining the upper bound on Q_θ^a for each value of k. In order to construct a regular matrix polynomial P with (α_0,β_0) as a simple eigenvalue and such that $\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}/\max\limits_{i=0:k}\{\|B_i\|_2\} \approx \|A\|_\infty^k$ we proceed as follows:
Let $q(\alpha,\beta)$ be any nonzero scalar polynomial of degree k such that $q(\alpha_0,\beta_0)=0$ and define $P(\alpha,\beta)=\operatorname{diag}(\varepsilon q(\alpha,\beta),Q(\alpha,\beta)) =: \sum_{i=0}^k \alpha^i\beta^{k-i}B_i$, where $\varepsilon>0$ is an arbitrarily small parameter and $Q(\alpha,\beta)=\sum_{i=0}^k \alpha^i\beta^{k-i}C_i \in \mathbb{C}[\alpha,\beta]_k^{(n-1)\times(n-1)}$ is a regular matrix polynomial. Then, $P(\alpha,\beta)$ is regular, and has (α_0,β_0) as a simple eigenvalue if (α_0,β_0) is not an eigenvalue of $Q(\alpha,\beta)$. Moreover, if ε is sufficiently small and $\|C_\ell\|_2 := \max\limits_{i=0:k}\{\|C_i\|_2\}$, then $\max\limits_{i=0:k}\{\|B_i\|_2\} = \|B_\ell\|_2$. Let us assume, for simplicity, that $(\alpha_0,\beta_0)\neq(1,0)$ and $(\alpha_0,\beta_0)\neq(0,1)$, although such assumption is not essential. Next, we explain how to construct $Q(\alpha,\beta)$ depending on which entry of A has the largest modulus.
• If $\|A\|_M=|a|$, let $Q(\alpha,\beta) := \alpha^k C_k$, where C_k is an arbitrary $(n-1)\times(n-1)$ nonsingular matrix such that, for ε small enough, $P(\alpha,\beta)$ satisfies $\|B_k\|_2 \gg \|B_i\|_2$ for $i\neq k$, and so, (3.6) implies $\widetilde{B}_k \approx a^k B_k$. Hence $\max_{i=0:k}\{\|\widetilde{B}_i\|_2\} \geq \|\widetilde{B}_k\|_2 \approx |a|^k\|B_k\|_2$. By (3.7), we have

$$\frac{1}{2^k}\,\|A\|_\infty^k \leq \|A\|_M^k \leq \frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}} \leq \|A\|_\infty^k(k+1)\binom{k}{\lfloor k/2\rfloor}.$$

Thus, we deduce that $\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}/\max\limits_{i=0:k}\{\|B_i\|_2\} \approx \|A\|_\infty^k$, up to a factor depending on k. Note that (α_0,β_0) is not an eigenvalue of $Q(\alpha,\beta)$ by construction.
• If $\|A\|_M=|b|$, then the same conclusion follows by taking again $Q(\alpha,\beta)=\alpha^k C_k$, since (3.6) implies $\widetilde{B}_0 \approx b^k B_k$.
• If $\|A\|_M=|c|$, we get the desired result from taking $Q(\alpha,\beta)=\beta^k C_0$, since (3.6) implies $\widetilde{B}_k \approx c^k B_0$.
• If $\|A\|_M=|d|$, take again $Q(\alpha,\beta)=\beta^k C_0$, since (3.6) implies $\widetilde{B}_0 \approx d^k B_0$.
4. From (5.23) and (5.25), we see that, for a fixed ill-conditioned A, *the behaviours of Q_θ^p for $k=1$ and $k>2$ are very different from each other* in the following sense: the eigenvalues (α_0,β_0) for which Q_θ^p essentially attains the upper (resp. lower) bound given in Theorem 5.6 for $k>2$, do not attain the upper (resp. lower) bound on Q_θ^p for $k=1$. Notice, for example, that if Q_θ^p attains the upper bound for some polynomial of degree $k>2$ having (α_0,β_0) as a simple eigenvalue, then $\frac{\|[\gamma_0,\delta_0]^T\|_\infty}{\|[\alpha_0,\beta_0]^T\|_\infty} \approx \|A^{-1}\|_\infty$ and $\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}/\max\limits_{i=0:k}\{\|B_i\|_2\} \approx \|A\|_\infty^k$, which implies

that $Q_\theta^p \approx \mathrm{cond}_\infty(A)^{k-1}$ by (3.5) while, in this case, the value of Q_θ^p associated with a polynomial of degree 1 is of order 1 (by (5.23)), which is not close to the upper bound $4\,\mathrm{cond}_\infty(A)$ since A is ill-conditioned. Note also that, in contrast to the discussion for Q_θ^a, we cannot state that such behaviours are opposite from each other. In our example, the lower bound for Q_θ^p with $k = 1$ is much smaller than 1 when $\mathrm{cond}_\infty(A)$ is very large while Q_θ^p may be of order 1. These different behaviours have been very clearly observed in the numerical experiments presented in Section 7 as it is explained in the next paragraph.

5. For a fixed ill-conditioned A, we have observed numerically that the eigenvalues (α_0, β_0) of randomly generated matrix polynomials $P(\alpha, \beta)$ of any degree almost always satisfy that

$$\frac{\|[\gamma_0, \delta_0]^T\|_\infty}{\|[\alpha_0, \beta_0]^T\|_\infty} = \theta \|A^{-1}\|_\infty, \tag{5.28}$$

with θ not too close to 0. This is naturally expected because "random" vectors $[\alpha_0, \beta_0]^T$, when expressed in the (orthonormal) basis of right singular vectors of A^{-1}, have non-negligible components on the vector corresponding to the largest singular value. We have also observed that randomly generated polynomials $P(\alpha, \beta)$ of moderate degree almost always satisfy

$$\frac{\max\limits_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max\limits_{i=0:k}\{\|B_i\|_2\}} = \xi \|A\|_\infty^k, \tag{5.29}$$

with ξ not far from 1. Combining (5.28), (5.29), (5.23), and (5.25) we get that, for randomly generated polynomials, the following conditions almost always hold: $Q_\theta^p \approx \xi/\theta \approx 1$ for $k = 1$; $Q_\theta^p \approx \xi\|A\|_\infty^2/|\det(A)| \approx \mathrm{cond}_\infty(A)$ for $k = 2$; and $Q_\theta^p \approx \xi\theta^{k-2}\|A^{-1}\|_\infty^{k-2}\|A\|_\infty^k/|\det(A)| \approx \mathrm{cond}_\infty(A)^{k-1}$ for $k > 2$. This explains why in random numerical tests for $k = 1$ the quotient Q_θ^p is almost always close to 1 and seems to be insensitive to the conditioning of A, as we will check numerically in Section 7. However, remember, that both the upper and lower bounds in Theorem 5.6 can be essentially attained for any fixed A.

We finish this section by remarking that the differences mentioned above between the degrees $k = 1$ and $k \geq 2$ are also observed numerically for the relative quotients Q_θ^r as shown in Section 7, although the differences are somewhat less clear. It is also possible to justify that the lower and upper bounds in Theorem 5.9 can be essentially attained, but the arguments need to take into account the factors ρ and $\widetilde{\rho}$ and are more complicated. We have performed numerical tests that confirm that those bounds are approximately attainable.

6. Effect of Möbius transformations on backward errors of approximate eigenpairs. The scenario in this section is the following: we want to compute eigenpairs of a regular homogeneous matrix polynomial $P(\alpha, \beta) \in \mathbb{C}[\alpha, \beta]_k^{n \times n}$, but, for some reason, it is advantageous to compute eigenpairs of its Möbius transform $M_A(P)(\gamma, \delta)$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$. A motivation for this might be, for instance, that $P(\alpha, \beta)$ has a certain structure that can be used for computing very efficiently and/or accurately its eigenpairs, but there are no specific algorithms available for such structure, although there are for the structured polynomial $M_A(P)(\gamma, \delta)$. Note that if $(\widehat{x}, (\widehat{\gamma}_0, \widehat{\delta}_0))$ and $(\widehat{y}^*, (\widehat{\gamma}_0, \widehat{\delta}_0))$ are computed *approximate* right and left eigenpairs of $M_A(P)$, and $(\widehat{\alpha_0}, \widehat{\beta_0}) := (a\widehat{\gamma_0} + b\widehat{\delta_0}, c\widehat{\gamma_0} + d\widehat{\delta_0})$ then, because of Proposition 3.5 and Lemma 3.7, $(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))$ and $(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0}))$ can be considered approximate right and left eigenpairs of $P(\alpha, \beta)$. Assuming that $(\widehat{x}, (\widehat{\gamma_0}, \widehat{\delta_0}))$ and $(\widehat{y}^*, (\widehat{\gamma_0}, \widehat{\delta_0}))$ have been computed with small backward errors in the sense of Definition 4.5, a natural question in this setting is whether $(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))$ and $(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0}))$ are also approximate eigenpairs of P with small backward errors. This would happen if the quotients

$$Q_{\eta,\text{right}} := \frac{\eta_P(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))}{\eta_{M_A(P)}(\widehat{x}, (\widehat{\gamma_0}, \widehat{\delta_0}))}, \quad Q_{\eta,\text{left}} := \frac{\eta_P(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0}))}{\eta_{M_A(P)}(\widehat{y}^*, (\widehat{\gamma_0}, \widehat{\delta_0}))} \tag{6.1}$$

are moderate numbers not much larger than one. In this section we provide upper bounds on the quotients in (6.1) that allow us to determine simple sufficient conditions that guarantee that such quotients are not large numbers. For completeness, we also provide lower bounds for these quotients, although they are less interesting than the upper ones in the scenario described above.

Note that, from Theorem 4.6, we can easily deduce that the backward error is independent of the choice of representative of the approximate eigenvalue.

The first result in this section is Theorem 6.1, which proves that the quotients in (6.1) are equal and provides an explicit expression for them.

THEOREM 6.1. *Let $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i \in \mathbb{C}[\alpha, \beta]_k^{n\times n}$ be a regular homogeneous matrix polynomial, let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{C})$, and let $M_A(P)(\gamma, \delta) = \sum_{i=0}^{k} \gamma^i \delta^{k-i} \widetilde{B}_i$ be the Möbius transform of $P(\alpha, \beta)$ under M_A. Let $(\widehat{x}, (\widehat{\gamma_0}, \widehat{\delta_0}))$ and $(\widehat{y}^*, (\widehat{\gamma_0}, \widehat{\delta_0}))$ be approximate right and left eigenpairs of $M_A(P)$, and let $[\widehat{\alpha_0}, \widehat{\beta_0}]^T := A[\widehat{\gamma_0}, \widehat{\delta_0}]^T$. Let $Q_{\eta,\text{right}}$ and $Q_{\eta,\text{left}}$ be as in (6.1) and let ω_i and $\widetilde{\omega_i}$ be the weights used in the definition of the backward errors for P and $M_A(P)$, respectively. Then,*

$$Q_{\eta,\text{right}} = Q_{\eta,\text{left}} = \frac{\sum_{i=0}^{k} |\widehat{\gamma_0}|^i |\widehat{\delta_0}|^{k-i} \widetilde{\omega_i}}{\sum_{i=0}^{k} |\widehat{\alpha_0}|^i |\widehat{\beta_0}|^{k-i} \omega_i}. \tag{6.2}$$

Moreover, (6.2) is independent of the choice of representative for $(\widehat{\gamma_0}, \widehat{\delta_0})$.

Proof. Since the backward error does not depend on the choice of representative of approximate eigenvalues, we choose an arbitrary representative $[\widehat{\gamma_0}, \widehat{\delta_0}]^T$ of $(\widehat{\gamma_0}, \widehat{\delta_0})$, and, once $[\widehat{\gamma_0}, \widehat{\delta_0}]^T$ is fixed, we choose $[\widehat{\alpha_0}, \widehat{\beta_0}]^T := A[\widehat{\gamma_0}, \widehat{\delta_0}]^T$ as representative of the approximate eigenvalue of P. For these representatives note that $M_A(P)(\widehat{\gamma_0}, \widehat{\delta_0}) = \sum_{i=0}^{k} (a\widehat{\gamma_0} + b\widehat{\delta_0})^i (c\widehat{\gamma_0} + d\widehat{\delta_0})^{k-i} B_i = P(\widehat{\alpha_0}, \widehat{\beta_0})$. Thus, Theorem 4.6 implies (6.2). □

Analogously to the quotients of condition numbers in Definition 5.2, we can consider absolute, relative with respect to the norm of the polynomial, and relative quotients of backward errors. They are defined, taking into account Definition 4.7, as

$$Q^s_{\eta,\text{right}} := \frac{\eta^s_P(\widehat{x}, \langle A[\widehat{\gamma_0}, \widehat{\delta_0}]^T \rangle)}{\eta^s_{M_A(P)}(\widehat{x}, (\widehat{\gamma_0}, \widehat{\delta_0}))}, \quad Q^s_{\eta,\text{left}} := \frac{\eta^s_P(\widehat{y}^*, \langle A[\widehat{\gamma_0}, \widehat{\delta_0}]^T \rangle)}{\eta^s_{M_A(P)}(\widehat{y}^*, (\widehat{\gamma_0}, \widehat{\delta_0}))}, \text{ for } s = a, p, r. \tag{6.3}$$

Theorem 6.2 provides upper and lower bounds on the quotients in (6.3).

THEOREM 6.2. *With the same notation and hypotheses of Theorem 6.1, let $Y_k := (k+1)^2 \binom{k}{\lfloor k/2 \rfloor}$. Then*

1. $\dfrac{1}{(k+1)\, \|A\|_\infty^k} \leq Q^a_{\eta,\text{right}} = Q^a_{\eta,\text{left}} \leq (k+1)\, \|A^{-1}\|_\infty^k.$

2. $\dfrac{1}{Y_k \operatorname{cond}_\infty(A)^k} \leq Q^p_{\eta,\text{right}} = Q^p_{\eta,\text{left}} \leq Y_k \operatorname{cond}_\infty(A)^k$.

3. If $B_0 \neq 0, B_k \neq 0, \widetilde{B}_0 \neq 0$, *and* $\widetilde{B}_k \neq 0$, *and* ρ *and* $\widetilde{\rho}$ *are defined as in* (5.22), *then*

$$\frac{1}{Y_k \operatorname{cond}_\infty(A)^k \, \widetilde{\rho}} \leq Q^r_{\eta,\text{right}} = Q^r_{\eta,\text{left}} \leq Y_k \operatorname{cond}_\infty(A)^k \, \rho.$$

Proof. We only prove the upper bounds since the lower bounds can be obtained in a similar way. Moreover, we only need to pay attention to the quotients for right eigenpairs, taking into account (6.2). Let us start with the absolute quotients. From (6.2) with $\omega_i = \widetilde{\omega}_i = 1$, we obtain

$$\begin{aligned} Q^a_{\eta,\text{right}} &\leq (k+1) \frac{\|[\widehat{\gamma}_0, \widehat{\delta}_0]^T\|^k_\infty}{\|A[\widehat{\gamma}_0, \widehat{\delta}_0]^T\|^k_\infty} \\ &= (k+1) \frac{\|A^{-1}A[\widehat{\gamma}_0, \widehat{\delta}_0]^T\|^k_\infty}{\|A[\widehat{\gamma}_0, \widehat{\delta}_0]^T\|^k_\infty} \leq (k+1)\|A^{-1}\|^k_\infty. \end{aligned} \tag{6.4}$$

The upper bound on $Q^p_{\eta,\text{right}}$ follows from combining $Q^p_{\eta,\text{right}} = Q^a_{\eta,\text{right}} \frac{\max_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\max_{i=0:k}\{\|B_i\|_2\}}$, which is obtained from (6.2), the upper bound on $Q^a_{\eta,\text{right}}$ obtained above, and (3.7).

The upper bound on $Q^r_{\eta,\text{right}}$ can be obtained noting that (6.2) and (3.7) imply

$$\begin{aligned} Q^r_{\eta,\text{right}} &\leq \frac{\max_{i=0:k}\{\|\widetilde{B}_i\|_2\}}{\min\{\|B_0\|_2, \|B_k\|_2\}} \frac{\sum_{i=0}^k |\widehat{\gamma}_0|^i |\widehat{\delta}_0|^{k-i}}{|a\widehat{\gamma}_0 + b\widehat{\delta}_0|^k + |c\widehat{\gamma}_0 + d\widehat{\delta}_0|^k} \\ &\leq (k+1)^2 \binom{k}{\lfloor k/2 \rfloor} \|A\|^k_\infty \frac{\|[\widehat{\gamma}_0, \widehat{\delta}_0]^T\|^k_\infty}{\|[a\widehat{\gamma}_0 + b\widehat{\delta}_0, c\widehat{\gamma}_0 + d\widehat{\delta}_0]^T\|^k_\infty} \, \rho \\ &\leq (k+1)^2 \binom{k}{\lfloor k/2 \rfloor} \|A\|^k_\infty \|A^{-1}\|^k_\infty \, \rho, \end{aligned}$$

where the last inequality is obtained as in (6.4). □

Remark 6.3. The bounds in Theorem 6.2 on the quotients of backward errors have the same flavor as those in Theorems 5.4, 5.6, and 5.9 on the quotients of condition numbers. However, note that in Theorem 6.2 there is no need to make a distinction between the bounds for $k = 1$ and $k \geq 2$, in contrast with Theorems 5.4, 5.6, and 5.9, since the bounds for the quotients of backward errors are obtained in the same way for all k. This has numerical consequences since the differences discussed in Section 5.3, and shown in practice in some of the tests in Section 7, between the quotients of condition numbers for $k = 1$ and $k \geq 2$ when $\operatorname{cond}_\infty(A) \gg 1$ do not exist for the quotients of backward errors.

Ignoring the factors depending only on the degree k, Theorem 6.2 guarantees that the quotients of backward errors are moderate numbers under the same sufficient conditions under which Theorems 5.4, 5.6, and 5.9 guarantee that the quotients of condition numbers are moderate numbers. That is: $\|A\|_\infty \approx \|A^{-1}\|_\infty \approx 1$ implies that $Q^a_{\eta,\text{right}} = Q^a_{\eta,\text{left}}$ is a moderate number, $\operatorname{cond}_\infty(A) \approx 1$ implies that $Q^p_{\eta,\text{right}} = Q^p_{\eta,\text{left}}$ is a moderate number, and $\operatorname{cond}_\infty(A) \approx 1$ and $\rho \approx \widetilde{\rho} \approx 1$ imply that $Q^r_{\eta,\text{right}} = Q^r_{\eta,\text{left}}$ is a moderate number.

7. Numerical experiments. In this section, we present a few numerical experiments that compare the exact values of the quotients Q_θ^p, Q_θ^r, $Q_{\eta,\text{right}}^p$, and $Q_{\eta,\text{right}}^r$ with the bounds on these quotients obtained in Sections 5 and 6. Observe that, implicitly, these experiments also compare the exact values of $Q_{\eta,\text{left}}^p$ and $Q_{\eta,\text{left}}^r$ with the bounds on these quotients as a consequence of Theorem 6.1. We do not present experiments on Q_θ^a and $Q_{\eta,\text{right}}^a$ for brevity and also because the weights corresponding to these quotients are not interesting in applications, as it was explained after Definition 4.4. We remark that many other numerical tests have been performed, in addition to the ones presented in this section, and that all of them confirm the theory developed in this paper.

The results in Sections 5 and 6 prove that eigenvalue condition numbers and backward errors of approximate eigenpairs can change significantly under Möbius transformations induced by ill-conditioned matrices. Therefore, the use of such Möbius transformations is not recommended in numerical practice. As a consequence most of our numerical experiments consider Möbius transformations induced by matrices A such that $\text{cond}_2(A) = 1$, which implies $1 \leq \text{cond}_\infty(A) \leq 2$. The only exception is Experiment 3.

Next we explain the goals of each of the numerical experiments in this section. Experiment 1 illustrates that the factor Z_k appearing in the bounds on Q_θ^p and Q_θ^r in Theorems 5.6 and 5.9 is very pessimistic in practice. This is a very important fact since Z_k is very large for moderate values of k and, if its effect was observed in practice, then even Möbius transformations induced by well-conditioned matrices would not be recommendable for matrix polynomials with moderate degree. Experiment 2 illustrates that Q_θ^r indeed depends on the factor ρ defined in (5.22) and, so, that the bounds in Theorem 5.9 reflect often the behaviour of Q_θ^r when ρ is large. Experiment 3 is mainly of academic interest, since it considers Möbius transformations induced by ill-conditioned matrices. The goal of this experiment is to illustrate the results presented in Subsection 5.3, in particular, the different typical behaviors of the quotients Q_θ^p for $k = 1$ and $k \geq 2$ when the polynomials are randomly generated. Experiments 4 and 5 are the counterparts of Experiments 1 and 2, respectively, for the quotients of backward errors.

All the experiments have been run on MATLAB-R2018a. Since in these experiments we have sometimes encountered badly scaled matrix polynomials (that is, polynomials with matrix coefficients whose norms vary widely), ill-conditioned eigenvalues have appeared. These eigenvalues could potentially be computed inaccurately and spoil the comparison between the results in the experiments and the theory. To avoid this problem, all the computations in Experiments 1, 2, and 3 have been done using variable precision arithmetic with 40 decimal digits of precision. To obtain the eigenvalues of each matrix polynomial P in these experiments, the function `eig` in MATLAB has been applied to the first Frobenius companion form of P. In Experiments 4 and 5, we have also used variable precision arithmetic with 40 decimal digits of precision for computing the Möbius transforms of the generated polynomials, but, since we are dealing with backward errors, the eigenvalues have been computed in the standard double precision of MATLAB with the command `polyeig`.

Experiment 1. In this experiment, we generate random matrix polynomials $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$ by using the MATLAB's command `randn` to generate the matrix coefficents B_i. Then, for each polynomial $P(\alpha, \beta)$, a random 2×2 matrix A is constructed as the unitary Q matrix produced by the command `qr(randn(2))`, which guarantees that $\text{cond}_2(A) = \|A\|_2 \|A^{-1}\|_2 = 1$. Finally, the Möbius transform

$M_A(P)$ is computed. We have worked with degrees $k = 1 : 15$ and, for each degree k, we have generated n_k matrix polynomials of size 5×5, where the values of n_k can be found in the following table:

n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	n_9	n_{10}	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}
75	37	25	18	15	12	10	9	8	7	7	6	5	5	5

(7.1)

For each pair (P, A) and each (simple) eigenvalue (α_0, β_0) of $P(\alpha, \beta)$, we compute two quantities: the exact value of Q_θ^p (through the formula (5.2) with the weights in Definition 4.4(2)) and the upper bound on this quotient given in Theorem 5.6, which depends only on $\mathrm{cond}_\infty(A)$ and Z_k. These quantities are shown in the left plot of Figure 7.1 as a function of k: the exact values of Q_θ^p are represented with the marker $*$ while the upper bounds use the marker $\circ$. Note that in this plot the scale of the vertical axis is logarithmic. This experiment confirms the (in Remark 5.7 anticipated) fact that the factor Z_k is very pessimistic, since we observe in the plot that, although the quotients Q_θ^p typically increase slowly with the degree k, they are much smaller than the corresponding upper bounds. A closer look at the exact values of Q_θ^p shows that most of them are larger than one, some considerably larger, and that the very few which are smaller than one are very close to one. We have observed this typical behavior of Q_θ^p (and also of Q_θ^r) in all our *random* numerical experiments, but we stress that it is easy to produce tests with the opposite behavior by interchanging the roles of P and $M_A(P)$ and of A and A^{-1}, respectively. Note that, in this case, the set of random matrix polynomials $M_A(P)$ is very different that the one produced by generating the matrix coefficients with the command `randn`.

We have performed an experiment similar to the one described in the previous paragraphs for confirming that Z_k is also pessimistic in the bounds in Theorem 5.9 on Q_θ^r. In this case, we have scaled the coefficients of the randomly generated matrix polynomials in such a way that the factor ρ in (5.22) is always equal to 10^3. The plot for the obtained exact values of Q_θ^r and their upper bounds is essentially the one on the left of Figure 7.1 with the vertical coordinates of all the markers multiplied by 10^3. For brevity, this plot is omitted.

Experiment 2. In this experiment, we have generated 30 random matrix polynomials of size 5×5 and degree 2 for which the factor ρ defined in (5.22) equals 10^t, where t has been randomly chosen for each polynomial by using the MATLAB's command `randi([0 10])`. More precisely, the matrix coefficients B_0, B_1, B_2 of these matrix polynomials with $\rho = 10^t$ have been generated with the next procedure. First, we generated matrix polynomials of size 5×5 and degree 2 by generating the matrix coefficients B_0', B_1', and B_2' with MATLAB's command `randn`. For each of these polynomials, we determined $\rho_T := \max_{i=0:2}\{\|B_i'\|_2\}/\min\{\|B_0'\|_2, \|B_2'\|_2\}$ and the coefficient B_s' such that $\|B_s'\|_2 = \max_{i=0:2}\{\|B_i'\|_2\}$. Then, the matrix coefficients B_0', B_1' and B_2' were scaled (obtaining new coefficients B_0, B_1, B_2) to get a new polynomial with the desired ρ, using the following criteria: If $\min\{\|B_0'\|_2, \|B_2'\|_2\} = \|B_0'\|_2$ and

(a) $\|B_0'\|_2 = \|B_1'\|_2 = \|B_2'\|_2$, then `q := randi([0 2])`, $B_q := \rho B_q'$ and $B_i := B_i'$ for $i \neq q$.

(b) $\|B_0'\|_2 = \|B_1'\|_2 = \|B_2'\|_2$ does not hold and $s = 1$, then:
  - (b1) If $\rho_T \leq \rho$, then $B_0 := \rho_T B_0'$, $B_1 := \rho B_1'$, and $B_2 := \rho_T B_2'$.
  - (b2) If $\rho_T > \rho$, then $B_0 := \rho_T B_0'$, $B_1 := \rho B_1'$, and $B_2 := \rho(\|B_1'\|_2/\|B_2'\|_2)\, B_2'$.

(c) $\|B_0'\|_2 = \|B_1'\|_2 = \|B_2'\|_2$ does not hold and $s \neq 1$ (which means $s = 2$), then:
  - (c1) If $\rho_T \leq \rho$, then $B_0 := B_0'$, $B_1 := B_1'$, and $B_2 := (\rho/\rho_T)\, B_2'$.
  - (c2) If $\rho_T > \rho$, then $B_0 := \rho_T B_0'$, $B_1 := \rho(\|B_2'\|_2/\|B_1'\|_2)\, B_1'$, and $B_2 := \rho B_2'$.

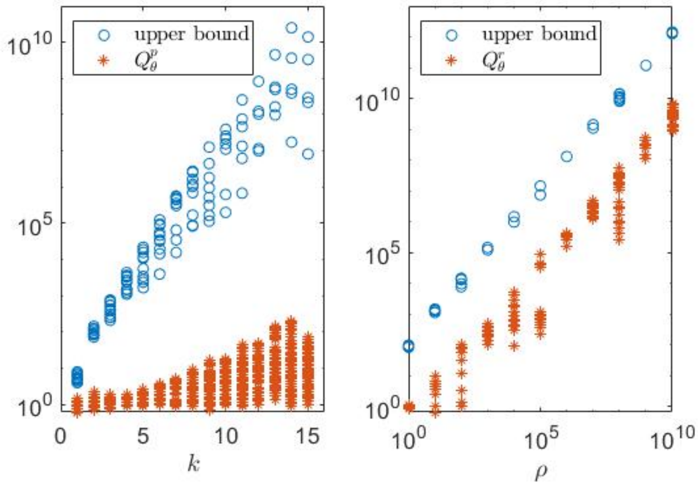

FIG. 7.1. *On the left results of Experiment 1, i.e., plot of Q_θ^p versus the degree k for Möbius transformations induced by matrices A with* $\mathrm{cond}_2(A) = 1$. *On the right results of Experiment 2, i.e., plot of Q_θ^r versus ρ for Möbius transformations of matrix polynomials with degree* 2 *induced by matrices A with* $\mathrm{cond}_2(A) = 1$.

If $\min\{\|B_0'\|_2, \|B_2'\|_2\} = \|B_2'\|_2$, then one proceeds in the same way but interchanging the roles of B_0' and B_2'.

For each matrix polynomial P generated as above, a random 2×2 matrix A with $\mathrm{cond}_2(A) = 1$ was constructed as in Experiment 1 and, then, $M_A(P)$ was computed. Finally, for each pair (P, A) and each (simple) eigenvalue (α_0, β_0) of P, we computed two quantities: the exact value of Q_θ^r, from the formula (5.2) with the weights in Definition 4.4(3), and the upper bound for this quotient in Theorem 5.9, which depends only on $\mathrm{cond}_\infty(A)$, ρ, and Z_2. These quantities are shown in the right plot of Figure 7.1 as a function of ρ: the markers of the exact values of Q_θ^r are $*$ and the markers of the upper bounds are $\circ$. Note that, in this plot, the scale of both the horizontal and vertical axes are logarithmic. It can be observed that many of the exact values of Q_θ^r essentially attain the upper bounds (recall that here $Z_2 = 72$), and, so, that Q_θ^r typically increases proportionally to ρ for the random matrix polynomials that we have generated. We report that, if in this set of random polynomials the roles of P and $M_A(P)$ and the roles of A and A^{-1} are interchanged, and the results are graphed against the factor $\widetilde{\rho}$ in (5.22), then the exact values of Q_θ^r essentially attain the *lower* bounds in Theorem 5.9. This plot is omitted for brevity.

Experiment 3. In this experiment, we generated random matrix polynomials P by generating their coefficients with MATLAB's command `randn`. In particular, we generated 30 matrix polynomials of degree $k = 1$ and sizes 5×5, 10×10, and 15×15; 20 matrix polynomials of degree $k = 2$ and sizes 5×5 and 10×10; and 20 matrix polynomials of degree $k = 3$ and sizes 5×5 and 8×8 (more precisely, 10 matrix polynomials of each pair degree-size). For each polynomial P, a random 2×2 matrix $A := U\mathrm{diag}(r, r/10^s)W$ was constructed, where U and W are random orthogonal matrices generated as the unitary Q matrices produced by the application of the MATLAB command `qr(randn(2))` twice; r = `randn`, and s = `randi([0 10])`, which implies $\mathrm{cond}_2(A) = 10^s$. Then the Möbius transform $M_A(P)$ of each polynomial P was computed.

For each pair (P, A) and each (simple) eigenvalue (α_0, β_0) of P, we computed two quantities: the exact value of Q_θ^p (from the formula (5.2) with the weights in Definition 4.4(2)) and $\mathrm{cond}_\infty(A)$. The quotients Q_θ^p are graphed (using the marker $*$) in the plots in Figure 7.2 as a function of $\mathrm{cond}_\infty(A)$: the figure on the left corresponds to the polynomials of degree 1, the figure in the middle corresponds to the polynomials of degree 2, and the figure on the right corresponds to the polynomials of degree 3. Observe that in these plots the scales of both axes are logarithmic and that solid lines corresponding to the upper bounds in Theorem 5.6 are also drawn. As announced and explained in Section 5.3, (recall, in particular, the fourth and fifth points) the differences between the behaviours of Q_θ^p for degrees $k = 1$ and $k \geq 2$ and the considered random polynomials are striking: typically, when $k = 2$ or $k = 3$, the exact values of Q_θ^p grow proportionally to $\mathrm{cond}_\infty(A)^{k-1}$ and are close to the upper bounds in Theorem 5.6, but, for $k = 1$, Q_θ^p remains close to 1 even when the matrix A is extremely ill-conditioned. However, the reader should bear in mind that for any given matrix A, it is always possible (and easy) to construct regular matrix polynomials of degree 1 (pencils) with eigenvalues for which the upper bound on Q_θ^p in Theorem 5.6 is essentially attained, as it was explained in the third point in Section 5.3. We have generated pencils of this type but the results are not shown for brevity. Again, we report that, for degrees 2 and 3, if in these sets of random polynomials the roles of P and $M_A(P)$ and the roles of A and A^{-1} are interchanged, then the exact values of Q_θ^p essentially attain the *lower* bounds in Theorem 5.6. These plots are also omitted for brevity.

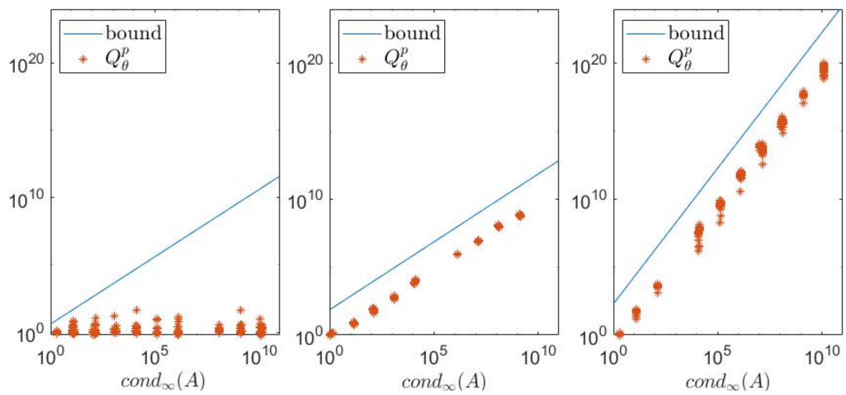

FIG. 7.2. *Results of Experiment 3: plots of* Q_θ^p *versus* $\mathrm{cond}_\infty(A)$ *for degrees* $k = 1$ *(on the left),* $k = 2$ *(on the middle), and* $k = 3$ *(on the right).*

We performed an experiment analogous to Experiment 3 but where all matrix polynomials were generated so that the value of ρ (as in (5.22)) equaled 10^3. The exact values of the quotients Q_θ^r and the upper bounds in Theorem 5.9 were then computed. The obtained plots are essentially the ones in Figure 7.2 with the vertical coordinates of the quotients and the upper bounds multiplied by 10^3.

Experiment 4. This experiment is the counterpart for backward errors of Experiment 1 and, as a consequence, is described very briefly. We generated a set of random matrix polynomials P and their Möbius transforms $M_A(P)$ exactly as in Experiment 1. Therefore, $\mathrm{cond}_2(A) = 1$ for all the matrices A in this test. Then, for each pair (P, A), we computed the (approximate) right eigenpairs of $M_A(P)(\gamma, \delta)$ in floating point arithmetic with the command `polyeig`. For each of these computed eigenpairs, we computed two quantities: $Q^p_{\eta,\text{right}}$ (from the expression (6.2) with the weights in Definition 4.7(2)) and the upper bound on this quotient obtained in Theorem 6.2, which depends only on $\text{cond}_\infty(A)$ and Y_k. These quantities are shown in the left plot of Figure 7.3 as functions of the degree k of P. We observe the same behaviour as in the left plot of Figure 7.1 and similar comments are valid. Therefore, it can be deduced that the factor Y_k in the bounds on the quotients of the backward errors is very pessimistic.

Experiment 5. This experiment is the counterpart of Experiment 2 for backward errors. We generated a set of random matrix polynomials P of degree 2 and their Möbius transforms $M_A(P)$ exactly as in Experiment 2. For each pair (P, A) and each right eigenpair of $M_A(P)(\gamma, \delta)$, computed in floating point arithmetic with `polyeig`, two quantities are computed: $Q^r_{\eta,\text{right}}$ (from the expression (6.2) with the weights in Definition 4.7(3)) and the upper bound for this quotient in Theorem 6.2, which depends only on $\text{cond}_\infty(A)$, ρ, and Y_2. These two quantities are shown in the right plot of Figure 7.3 as functions of ρ. The same behaviour as in the right plot of Figure 7.1 is observed and similar comments remain valid. Therefore, it can be deduced that the quotients $Q^r_{\eta,\text{right}}$ of the backward errors typically grow proportionally to ρ.

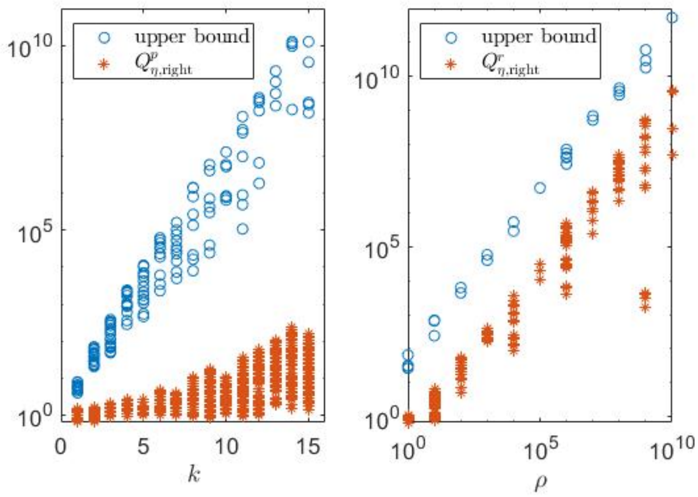

FIG. 7.3. *On the left results of Experiment 4, i.e., plot of $Q^p_{\eta,\text{right}}$ versus the degree k for Möbius transformations induced by matrices A such that $\text{cond}_2(A) = 1$. On the right results of Experiment 5, i.e., plot of $Q^r_{\eta,\text{right}}$ versus ρ for Möbius transformations of matrix polynomials with degree 2 induced by matrices A such that $\text{cond}_2(A) = 1$.*

Finally, we report that, for the quotients of backward errors $Q^p_{\eta,\text{right}}$, we have also performed an experiment analogous to the Experiment 3. The corresponding plots are not presented in this paper for brevity. However, we stress that the plot corresponding to the degree $k = 1$ is remarkably different from the left plot in Figure 7.2, since it shows that $Q^p_{\eta,\text{right}}$ typically increases proportionally to $\text{cond}_\infty(A)$ and, therefore, no difference of behavior is observed in this respect between the quotients of backward errors for degrees $k = 1$ and $k \geq 2$. This fact was pointed out and explained in Remark 6.3.

8. Conclusions and future work. In this paper, we have studied the influence of Möbius transformations on the (Stewart-Sun) eigenvalue condition number and backward errors of approximate eigenpairs of regular homogeneous matrix polynomials. More precisely, we have given sufficient conditions, independent of the eigenvalue, for the condition number of a simple eigenvalue of a polynomial P and the condition number of the associated eigenvalue of a Möbius transform of P to be close. Similarly, we have given sufficient conditions for the backward error of an approximate eigenpair of a Möbius transform of P and the associated approximate eigenpair of P to be close. In doing this analysis, we considered three variants of the Stewart-Sun condition number and of backward errors, depending on the selection of weights involved in their definitions, that we called absolute, relative with respect to the norm of the polynomial, and relative.

The most important conclusion of our study is that in the relative-to-the-norm-of-the-polynomial case, if the matrix A that defines the Möbius transformation is well-conditioned and the degree of P is moderate, then the Möbius transformation preserves approximately the conditioning of the simple eigenvalues of P, and the backward errors of the computed eigenpairs of P are similar to the backward errors of the computed eigenpairs of $M_A(P)$. In the relative case, these conclusions hold as well if, additionally, we assume that the matrix coefficients of P (resp., the matrix coefficients of $M_A(P)$) have similar norms. Furthermore, we have provided some insight on the behavior of the quotients of eigenvalue condition numbers when the matrix A defining the Möbius transformation is ill-conditioned. Our study shows that, in this case, a significantly different typical behavior of the quotients of eigenvalue condition numbers can be expected when the matrix polynomial has degree 1, 2 or larger than 2.

We must point out that the simple sufficient conditions for the approximate preservation of the eigenvalue condition numbers after the application of a Möbius transformation to a homogeneous matrix polynomial cannot be immediately extrapolated to the non-homogeneous case, which will be studied in a separate paper. In this case, special attention must be paid to eigenvalues with very large modulus or modulus close to 0.

In this paper, we have only considered the effect of Möbius transformations on the condition numbers of simple eigenvalues of a matrix polynomial. An interesting future line of research may be to extend our results to multiple eigenvalues by taking as starting point the condition numbers defined in [23, 33].

As explained in the introduction, in some relevant applications, the Möbius transformations are used to compute invariant or deflating subspaces associated with eigenvalues with certain properties. Thus, studying how a Möbius transformation affects the condition numbers of eigenvectors and invariant/deflating subspaces is an interesting problem that we will also address separately.

Appendix A. Backward errors attained at regular matrix polynomials. The next theorem proves that the formulas for the backward errors of approximate right and left eigenpairs of regular homogeneous matrix polynomials presented in Theorem 4.6 are a meaningful measure of the backward errors under the additional constraint that the perturbed polynomials considered in Definition 4.5 are regular. In order to keep the proof of Theorem A.1 simple, we restrict ourselves to consider any of the sets of weights from Definition 4.7 and that the coefficients B_0 and B_k of the unperturbed polynomial $P(\alpha, \beta)$ are different from zero. This last restriction is very mild but guarantees that the weights ω_0 and ω_k are both nonzero, which simplifies the proof of Theorem A.1. Otherwise, a proof is still possible but is more complicated.

THEOREM A.1. *Let $(\widehat{x}, (\widehat{\alpha_0}, \widehat{\beta_0}))$ and $(\widehat{y}^*, (\widehat{\alpha_0}, \widehat{\beta_0}))$ be an approximate right and an approximate left eigenpair, respectively, of a regular matrix polynomial $P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$ with $B_0 \neq 0$ and $B_k \neq 0$. Then,*

(1) *For any positive number ϕ (which, therefore, can be chosen arbitrarily small), there exists a regular matrix polynomial $P(\alpha,\beta) + \delta P(\alpha,\beta)$ of degree k, where $\delta P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} \delta B_i$, such that*
  (a) $(P(\widehat{\alpha_0}, \widehat{\beta_0}) + \delta P(\widehat{\alpha_0}, \widehat{\beta_0}))\widehat{x} = 0$, *and*
  (b)

$$\|\delta B_i\|_2 \leq \left(\frac{\|P(\widehat{\alpha_0}, \widehat{\beta_0})\widehat{x}\|_2}{(\sum_{i=0}^{k} |\widehat{\alpha_0}|^i |\widehat{\beta_0}|^{k-i} \omega_i)\|\widehat{x}\|_2} + \phi \right) \omega_i, \quad i = 0, 1, \ldots, k,$$

where the weights ω_i, $i = 0 : k$, are any of the ones in Definition 4.7.

(2) *For any positive number ϕ (which, therefore, can be chosen arbitrarily small), there exists a regular matrix polynomial $P(\alpha,\beta) + \delta P(\alpha,\beta)$ of degree k, where $\delta P(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} \delta B_i$, such that*
  (a) $\widehat{y}^*(P(\widehat{\alpha_0}, \widehat{\beta_0}) + \delta P(\widehat{\alpha_0}, \widehat{\beta_0})) = 0$, *and*
  (b)

$$\|\delta B_i\|_2 \leq \left(\frac{\|\widehat{y}^* P(\widehat{\alpha_0}, \widehat{\beta_0})\|_2}{(\sum_{i=0}^{k} |\widehat{\alpha_0}|^i |\widehat{\beta_0}|^{k-i} \omega_i)\|\widehat{y}\|_2} + \phi \right) \omega_i, \quad i = 0, 1, \ldots, k,$$

where the weights ω_i, $i = 0 : k$, are any of the ones in Definition 4.7.

Proof. We only prove Part (1) since Part (2) can be proved similarly.

By Theorem 4.6, there exists

$$Q(\alpha, \beta) := P(\alpha, \beta) + \Delta P(\alpha, \beta)$$

of degree k that satisfies (a) and (b) with $\delta P = \Delta P$ and $\phi = 0$. If $Q(\alpha, \beta)$ is regular, then the result is proven. Therefore, we assume in the rest of the proof that $Q(\alpha, \beta)$ is singular.

If $Q(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} Q_i$ is singular, then the matrix coefficient Q_0 must be singular since, otherwise, $\det(Q(0,1)) = \det(Q_0) \neq 0$ and $Q(\alpha, \beta)$ would be regular. Analogously, the matrix coefficient Q_k must be singular as well.

Consider the regular non-homogeneous pencil $L_0(\lambda) = \lambda I_n + Q_0$. Since Q_0 is singular, 0 is an eigenvalue of $L_0(\lambda)$. Moreover, either 0 is the only eigenvalue of Q_0 or there exists a nonzero eigenvalue λ_0 of $L_0(\lambda)$ with smallest modulus among the nonzero eigenvalues of $L_0(\lambda)$. If 0 is the only eigenvalue of Q_0, then the matrix $sI_n + Q_0$ is nonsingular for all nonzero s, i.e., for all s such that $0 < |s|$. If Q_0 has nonzero eigenvalues, then $sI_n + Q_0$ is nonsingular for all s such that $0 < |s| < |\lambda_0|$.

Analogously, for the pencil $L_k(\lambda) = \lambda I_n + Q_k$ it can be proven that $sI_n + Q_k$ is nonsingular for all s such that $0 < |s|$ if 0 is the only eigenvalue of Q_k or it is nonsingular for all s such that $0 < |s| < |\mu_0|$, where μ_0 is the nonzero eigenvalue of $L_k(\lambda)$ with smallest modulus. Thus there exists a positive number t such that $sI_n + Q_0$ and $sI_n + Q_k$ are both nonsingular for all s such that $0 < |s| < t$.

Let $\mu := \frac{t}{\max\{\omega_0, \ldots, \omega_k\}}$ and take an arbitrary positive number ϕ satisfying $0 < \phi < \mu$. Next, we construct the scalar polynomial

$$q(\alpha, \beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} q_i := \frac{\phi \min\{\omega_0, \omega_k\}}{\max\{|\widehat{\alpha_0}|^k, |\widehat{\beta_0}|^k\}} \left(\beta^k \, \widehat{\alpha_0}^k - \alpha^k \, \widehat{\beta_0}^k \right),$$

which satisfies

(a) $q(\widehat{\alpha_0}, \widehat{\beta_0}) = 0$;

(b) $q_k \neq 0$ or $q_0 \neq 0$ (or both) and $q_i = 0$ for $i \neq 0, k$;

(c) $|q_i| \leq \phi\, \omega_i < t$ for $i = 0, 1, \ldots, k$.

Finally, consider the matrix polynomial

$$Q(\alpha, \beta) + q(\alpha, \beta)I_n = P(\alpha, \beta) + \Delta P(\alpha, \beta) + q(\alpha, \beta)I_n.$$

Let $\delta P(\alpha, \beta) := \Delta P(\alpha, \beta) + q(\alpha, \beta)I_n$. Then, by property (a) of $q(\alpha, \beta)$ and the definition of $Q(\alpha, \beta)$, we have

$$(P(\widehat{\alpha_0}, \widehat{\beta_0}) + \delta P(\widehat{\alpha_0}, \widehat{\beta_0}))\widehat{x} = 0;$$

and, for $i = 0, 1, \ldots, k$,

$$\|\delta B_i\|_2 = \|\Delta B_i + q_i I_n\|_2 \leq \|\Delta B_i\|_2 + \phi\, \omega_i \leq \left(\frac{\|P(\widehat{\alpha_0}, \widehat{\beta_0})\widehat{x}\|_2}{(\sum_{i=0}^{k} |\widehat{\alpha_0}|^i |\widehat{\beta_0}|^{k-i}\omega_i)\|\widehat{x}\|_2} + \phi \right) \omega_i.$$

Moreover, $P(\alpha, \beta) + \delta P(\alpha, \beta)$ is regular because either $B_0 + \delta B_0 = B_0 + \Delta B_0 + q_0 I_n = Q_0 + q_0 I_n$ or $B_k + \delta B_k = B_k + \Delta B_k + q_k I_n = Q_k + q_k I_n$ or both are invertible by properties (b) and (c) of $q(\alpha, \beta)$. Since ϕ is any number satisfying $0 < \phi < \mu$, it can be chosen arbitrarily small. Moreover, once the inequality in (b) in Part (1) of the statement holds for some ϕ, it holds for any larger value of ϕ. Thus the result is proved for any positive number ϕ. ∎

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

REFERENCES

[1] L.M. Anguas, M.I. Bueno, and F.M. Dopico, *A comparison of eigenvalue condition numbers for matrix polynomials*, Linear Algebra Appl., 564 (2018), pp. 170–200.

[2] W.N. Bailey, *Generalised Hypergeometric Series*, Cambridge University Press, Cambridge, 1935.

[3] P. Benner, V. Mehrmann, and H. Xu, *A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils*, Numer. Math., 78 (1998), 329–358.

[4] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur, *NLEVP: A collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), 7:1-7:28.

[5] R.A. Brualdi, *Introductory Combinatorics*, Prentice Hall, New Jersey, 2010.

[6] 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.

[7] M.I. Bueno, K. Curlett, and S. Furtado, *Structured linearizations from Fiedler pencils with repetition I.*, Linear Algebra Appl., 460 (2014), 51-80.

[8] R. Byers, V. Mehrmann, and H. Xu, *A structured staircase algorithm for skew-symmetric/symmetric pencils*, Electron. Trans. Numer. Anal., 26 (2007), 1-33.

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

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

[11] F. De Terán, F.M. Dopico, and D.S. Mackey, *Palindromic companion forms for matrix polynomials of odd degree*, J. Comput. Appl. Math., 236 (2011), 1464-1480.

[12] F. De Terán, F.M. Dopico, and D.S. Mackey, *Spectral equivalence of matrix polynomials and the index sum theorem*, Linear Algebra Appl., 459 (2014), 264-333.

[13] F. De Terán, F.M. Dopico, and P. Van Dooren, *Matrix polynomials with completely prescribed eigenstructure*, SIAM J. Matrix Anal. Appl., 36 (2015), 302-328.
[14] F.M. Dopico, P.W. Lawrence, J. Pérez, and P. Van Dooren, *Block Kronecker linearizations of matrix polynomials and their backward errors*, Numer. Math., 140 (2018), 373-426.
[15] F.M. Dopico, J. Pérez, and P. Van Dooren, *Structured backward error analysis of linearized structured polynomial eigenvalue problems*, Math. Comp., 317 (2019), 1189-1228.
[16] H.-Y. Fan, W.-W. Lin, and P. Van Dooren, *Normwise scaling of second order polynomial matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), 252-256.
[17] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, 1996.
[18] S. Hammarling, C.J. Munro, and F. Tisseur, *An algorithm for the complete solution of quadratic eigenvalue problems*, ACM Trans. Math. Software, 39 (2013), 18:1-18:19.
[19] N.J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd ed., SIAM, Philadelphia, 2002.
[20] N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM, Philadelphia, 2008.
[21] N.J. Higham, R.-C. Li, and F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29 (2007), 1218-1241.
[22] N.J. Higham, D.S. Mackey, and F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), 1005-1028.
[23] D. Kressner, M.J. Peláez, and J. Moro, *Structured Hölder condition numbers for multiple eigenvalues*, SIAM J. Matrix Anal. Appl., 31 (2009), 175-201.
[24] P. Lancaster and L. Rodman, *The Algebraic Riccati Equation*, Oxford University Press, Oxford, 1995.
[25] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Structured polynomial eigenvalue problems: good vibrations from good linearizations*, SIAM J. Matrix Anal. Appl., 28 (2006), 1029-1051.
[26] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Möbius transformations of matrix polynomials*, Linear Algebra Appl., 470 (2015), 120-184.
[27] B. McMillan, *Introduction to formal realizability theory. I*, Bell System Tech. J., 31 (1952), 217-279.
[28] B. McMillan, *Introduction to formal realizability theory. II*, Bell System Tech. J., 31 (1952), 541-600.
[29] V. Mehrmann, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, Lecture Notes in Control and Information Sciences, vol. 163, Springer-Verlag, Heidelberg, 1991.
[30] V. Mehrmann, *A step toward a unified treatment of continuous and discrete time control problems*, Linear Algebra Appl., 241-243 (1996), 749-779.
[31] V. Mehrmann and F. Poloni, *A generalized structured doubling algorithm for the numerical solution of linear quadratic optimal control problems*, Numer. Linear Algebra Appl., 20 (2013), 112-137.
[32] V. Mehrmann and H. Xu, *Structure preserving deflation of infinite eigenvalues in structured pencils*, Electron. Trans. Numer. Anal., 44 (2015), 1–24.
[33] N. Papathanasiou and P. Psarrakos, *On condition numbers of polynomial eigenvalue problems*, Appl. Math. Comp., 216 (2010), 1194-1205.
[34] G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[35] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), 339-361.
[36] F. Tisseur and I. Zaballa, *Triangularizing quadratic matrix polynomials*, SIAM J. Matrix Anal. Appl., 34 (2013), 312-337.
[37] M. Van Barel and F. Tisseur, *Polynomial eigenvalue solver based on tropically scaled Lagrange linearization*, Linear Algebra Appl., 542 (2018), 186-208.
[38] P. Van Dooren and P. Dewilde, *The eigenstructure of an arbitrary polynomial matrix: computational aspects*, Linear Algebra Appl., 50 (1983), 545-579.
[39] H. Weyl, *The Classical Groups*, Princeton University Press, Princeton, NJ, 1973.
[40] L. Zeng and Y. Su, *A backward stable algorithm for quadratic eigenvalue problems*, SIAM J. Matrix Anal. Appl., 35 (2014), 499-516.