

PERTURBATION THEORY OF TRANSFER FUNCTION MATRICES

VANNI NOFERINI*, LAURI NYMAN†, JAVIER PÉREZ‡, AND MARÍA C. QUINTANA§

Abstract. Zeros of rational transfer function matrices $R(\lambda)$ are the eigenvalues of associated polynomial system matrices $P(\lambda)$, under minimality conditions. In this paper we define a structured condition number for a simple eigenvalue λ_0 of a (locally) minimal polynomial system matrix $P(\lambda)$, which in turn is a simple zero λ_0 of its transfer function matrix $R(\lambda)$. Since any rational matrix can be written as the transfer function of a polynomial system matrix, our analysis yields a structured perturbation theory for simple zeros of rational matrices $R(\lambda)$. To capture all the zeros of $R(\lambda)$, regardless of whether they are poles or not, we consider the notion of root vectors. As corollaries of the main results, we pay particular attention to the special case of λ_0 being not a pole of $R(\lambda)$ since in this case the results get simpler and can be useful in practice. We also compare our structured condition number with Tisseur’s unstructured condition number for eigenvalues of matrix polynomials, and show that the latter can be unboundedly larger. Finally, we corroborate our analysis by numerical experiments.

Key words. rational matrix, transfer function matrix, polynomial system matrix, rational eigenvalue problem, zeros, poles, root vectors, condition number

AMS subject classifications. 65F15, 15A18, 15A54, 93B20, 93B60

1. Introduction. Given a rational matrix $R(\lambda)$, the Rational Eigenvalue Problem (REP) is often defined as the problem of finding scalars λ_0 such that $R(\lambda_0)$ has finite entries (that is, λ_0 is not a pole of $R(\lambda)$) and that there exist nonzero constant vectors x and y (called eigenvectors) satisfying $R(\lambda_0)x = 0$ and $y^T R(\lambda_0) = 0$, under the assumptions that $R(\lambda)$ is regular, i.e., $R(\lambda)$ is square and its determinant is not identically equal to zero. However, zeros of rational matrices can also be poles: a situation not uncommon in certain applications such as control theory [13, 18, 21]. Thus, in this paper we will define the REP regardless of whether a zero is also a pole or not. For that, we extend the definition of eigenvectors by using the more general notion of root vectors [7, 10, 18]. That is, we define the REP as the problem of finding scalars λ_0 and rational vectors $x(\lambda) \in \mathbb{C}(\lambda)^m$ and $y(\lambda) \in \mathbb{C}(\lambda)^m$ with $x(\lambda_0) \neq 0$ and $y(\lambda_0) \neq 0$ such that $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ and $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0$. The above definition captures all the zeros of regular rational matrices $R(\lambda)$.

It is known [21] that any rational matrix $R(\lambda)$ can be written (or is directly given) as the transfer function matrix of a polynomial system matrix. That is, of the form

$$(1.1) \quad R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda),$$

where $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ are arbitrary polynomial matrices [9], with $A(\lambda)$ regular. Clearly, the expression in (1.1) is not uniquely determined by $R(\lambda)$. Then, under minimality conditions [5, 21], the zeros of $R(\lambda)$ are the eigenvalues of the

*Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland. (vanni.noferini@aalto.fi) Supported by an Academy of Finland grant (Suomen Akatemian päätös 331230).

†Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland. (lauri.s.nyman@aalto.fi).

‡University of Montana, Department of Mathematical Science, 32 Campus Dr, Missoula, MT 59812, United States. (javier.perez-alvaro@mso.umt.edu)

§Corresponding author. Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland. (maria.quintanaponce@aalto.fi) Supported by an Academy of Finland grant (Suomen Akatemian päätös 331230) and by the Agencia Estatal de Investigación of Spain through grant PID2019-106362GB-I00 MCIN/ AEI/10.13039/501100011033/.

associated polynomial matrix

$$(1.2) \quad P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix},$$

which is said to be a polynomial system matrix of $R(\lambda)$. In this paper, we will study the structured conditioning of simple zeros of (1.2), allowing perturbation of each block; if (1.2) is (locally) minimal, the analysis thus provides a structured perturbation theory of simple zeros of rational matrices written as in (1.1).

Rational matrices appear directly from applications, as in linear systems and control theory, or as approximations to nonlinear eigenvalue problems (NLEPs) [11, 15]. An example of rational matrix arising from application [17] is the following:

$$(1.3) \quad R(\lambda) = -K + \lambda M + \lambda^2 \sum_{i=1}^k \frac{1}{\omega_i - \lambda} C_i,$$

which can be easily written as the transfer function of a polynomial system matrix, for example, as follows:

$$(1.4) \quad R(\lambda) = -K + \lambda M + [\lambda I \cdots \lambda I] \begin{bmatrix} (w_1 - \lambda)I & & \\ & \ddots & \\ & & (w_k - \lambda)I \end{bmatrix}^{-1} \begin{bmatrix} \lambda C_1 \\ \vdots \\ \lambda C_k \end{bmatrix}.$$

Another example is the problem loaded `string` from [2]:

$$(1.5) \quad R(\lambda) = A - \lambda B + \frac{k\lambda}{\lambda - k/m} e_n e_n^T,$$

where A, B are certain $n \times n$ matrices, k, m are parameters, and e_n denotes the n th column of the $n \times n$ identity matrix. The rational matrix (1.5) is also easy to write as a transfer function, for instance:

$$(1.6) \quad R(\lambda) = A - \lambda B + \lambda e_n [\lambda - k/m]^{-1} k e_n^T.$$

1.1. Condition numbers. The condition number of a function measures how much the output of the function can change for a small perturbation in the input. To establish a general framework, consider two normed vector spaces $(\mathcal{U}, \|\cdot\|_{\mathcal{U}})$ and $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$ and a function $f : \mathcal{U} \rightarrow \mathcal{V}$, $u \mapsto f(u)$. The (worst-case) *absolute condition number* of f at u is defined as [20]

$$(1.7) \quad \kappa_f(u) = \lim_{\epsilon \rightarrow 0} \sup_{\|e\|_{\mathcal{U}} \leq 1} \frac{\|f(u + \epsilon e) - f(u)\|_{\mathcal{V}}}{\epsilon}.$$

If f is Fréchet differentiable at u , then this definition implies that κ is the operator norm of the Fréchet derivative of f at u . However, (1.7) is valid more generally. The (worst-case) *relative condition number* of f at u is defined analogously as

$$(1.8) \quad \kappa_{f,\text{rel}}(u) = \lim_{\epsilon \rightarrow 0} \sup_{\substack{\|e\|_{\mathcal{U}} \leq 1 \\ \|u\|_{\mathcal{U}} \leq 1}} \frac{\|f(u + \epsilon e) - f(u)\|_{\mathcal{V}}}{\epsilon \|f(u)\|_{\mathcal{V}}} = \kappa_f(u) \cdot \frac{\|u\|_{\mathcal{U}}}{\|f(u)\|_{\mathcal{V}}}.$$

Condition numbers are geometric invariants, meaning that (once an arbitrary, but fixed, choice of norms is made) they are an intrinsic property of the function f and

the point u ; in other words, they do not depend on the algorithm used in practice to compute $f(u)$. In numerical analysis, their use is popular to measure how difficult it is to accurately compute f in finite precision arithmetic. This has been effectively summarized by N. Higham [12, Sec. 1.6] with the “rule of thumb”

$$\text{forward error} \lesssim \kappa \cdot \text{backward error},$$

which comes from a first order expansion and therefore is technically true only for finite condition numbers and small enough values of the backward error. In practice, it is often still valid more generally as an approximate inequality. For a backward stable algorithm, the backward error is guaranteed to be small, but whether this implies that the output of the computation is accurate (that is, whether the forward error is small) also depends on the condition number κ . Finally, *structured* condition numbers can be defined analogously, but with the further restriction that the set of allowed perturbations e must satisfy certain properties. In other words, instead of taking the supremum over the unit ball $\|e\|_{\mathcal{U}} \leq 1$ as in (1.7), for defining a structured condition number one takes the supremum over a certain *subset* of the unit ball. For example, if u is a Hermitian matrix and $\mathcal{U} = \mathbb{C}^{n \times n}$, then one may impose that e is such that $u + \epsilon e$ is also a Hermitian matrix. Note that, by the properties of the supremum, this immediately implies that a structured condition number is less than or equal to the corresponding unstructured condition number; equality is however generally possible.

In this paper, we specialize the general theory to condition numbers of simple zeros (also called eigenvalues) of either rational or polynomial matrices, always assumed to be regular. Suppose that λ_0 is a simple zero for a rational matrix $R(\lambda)$, associated with left and right root vectors [10, 18] $y(\lambda)$ and $x(\lambda)$ respectively. Then, as we prove in Lemma 3.3, for any sufficiently small $\epsilon > 0$ and any perturbation $\Delta R(\lambda, \epsilon) = \epsilon \Delta R(\lambda) + o(\epsilon)$ satisfying certain technical assumptions (corresponding to perturbing the blocks of a polynomial system matrix), the perturbed rational matrix $R(\lambda) + \Delta R(\lambda, \epsilon)$ has a finite zero $\hat{\lambda}_0(\epsilon)$ such that

$$|\hat{\lambda}_0(\epsilon) - \lambda_0| = \epsilon \lim_{\lambda \rightarrow \lambda_0} \left| \frac{y(\lambda)^T \Delta R(\lambda) x(\lambda)}{y(\lambda)^T R'(\lambda) x(\lambda)} \right| + o(\epsilon).$$

From this result, after giving some preliminaries on rational matrices and polynomial system matrices in Section 2, we define in Section 3 a structured condition number κ_S for a locally minimal [5] polynomial system matrix $P(\lambda)$ of a rational matrix $R(\lambda)$, where the structured perturbation $\epsilon \Delta P(\lambda)$ preserves the degrees of the blocks of $P(\lambda)$ in (1.2). For that, we fix the function f to be the input-output function that maps $P(\lambda)$ to λ_0 and $P(\lambda) + \epsilon \Delta P(\lambda)$ to $\hat{\lambda}_0(\epsilon)$. In Section 4, we derive an expression for κ_S . Then, in Section 5, we show that this structured condition number for polynomial system matrices is never larger, and can be much smaller, than Tisseur’s unstructured condition number for matrix polynomials [22]. If we consider the case in which the rational matrix $R(\lambda)$ is expressed in the form $R(\lambda) = D(\lambda) + C(\lambda I_n - A)^{-1} B$, where $D(\lambda)$ is a polynomial matrix and $C(\lambda I_n - A)^{-1} B$ is a minimal state-space realization, it was shown in [8] that there are algorithms to compute the zeros of $R(\lambda)$, via linearization, that guarantee a small (global) backward error, preserving precisely the structure of $R(\lambda)$. Note that this is a particular case in the representation (1.1) where the matrix $A(\lambda)$ is linear and the matrices $B(\lambda)$ and $C(\lambda)$ are constant, but expressing $R(\lambda)$ in that form is valid for any rational matrix and is used in applications [21]. Therefore, in problems where the structured condition number is much smaller than the unstructured one, structured algorithms can thus eventually lead to

significant increase in accuracy. We also include some numerical experiments on this comparison in Section 6. Finally, in Section 7 we give some conclusions and related open problems.

2. Preliminaries.

2.1. Rational matrices and related subspaces. A rational matrix $R(\lambda) \in \mathbb{C}(\lambda)^{p \times m}$ is a matrix whose entries are scalar rational functions in the variable λ . Given $\lambda_0 \in \mathbb{C}$, $R(\lambda)$ is said to be *defined at* λ_0 if $R(\lambda_0) \in \mathbb{C}^{p \times m}$ and is said to be *invertible at* λ_0 if, in addition, $\det R(\lambda_0) \neq 0$. Rational matrices can have zeros and poles, which can be defined through the notion of the local Smith–McMillan form. That is, for any rational matrix $R(\lambda) \in \mathbb{C}(\lambda)^{p \times m}$, with normal rank r , and $\lambda_0 \in \mathbb{C}$ there exist rational matrices $G_1(\lambda)$ and $G_2(\lambda)$, that are invertible at λ_0 , such that

$$(2.1) \quad G_1(\lambda)R(\lambda)G_2(\lambda) = \begin{bmatrix} \text{diag}((\lambda - \lambda_0)^{\nu_1}, \dots, (\lambda - \lambda_0)^{\nu_r}) & 0 \\ 0 & 0_{(p-r) \times (m-r)} \end{bmatrix},$$

where $\nu_1 \leq \dots \leq \nu_r$ are integers. The integers ν_1, \dots, ν_r are uniquely determined by $R(\lambda)$ and λ_0 and are called the *invariant orders at* λ_0 of $R(\lambda)$. The matrix in (2.1) is called the *local Smith–McMillan form of* $R(\lambda)$ at λ_0 .

If $\nu_i > 0$, for some $i = 1, \dots, r$, then λ_0 is a *zero* (or *eigenvalue*) of $R(\lambda)$ with *partial multiplicity* ν_i , and $(\lambda - \lambda_0)^{\nu_i}$ is said to be a *zero elementary divisor of* $R(\lambda)$ at λ_0 . If $\nu_i < 0$, for some $i = 1, \dots, r$, then λ_0 is a *pole* of $R(\lambda)$ with *partial multiplicity* $-\nu_i$, and $(\lambda - \lambda_0)^{-\nu_i}$ is said to be a *pole elementary divisor of* $R(\lambda)$ at λ_0 [13, 23]. A zero of $R(\lambda)$ is said to be *simple* if it only has one positive invariant order and it is equal to 1 (but a simple zero can also be a pole with any pole partial multiplicities). The (global) definition of the Smith–McMillan form of a rational matrix can be found in [13, 21, 23] and was first given by B. McMillan in [16].

If $R(\lambda)$ is a polynomial matrix then the nonzero integers $\nu_i \neq 0$ are all positive and are called *partial multiplicities of* $R(\lambda)$ at λ_0 [9]. In addition, in this case, the diagonal matrix in (2.1) is simply called the *local Smith form of* $R(\lambda)$ at λ_0 and the polynomials $(\lambda - \lambda_0)^{\nu_i}$ with $\nu_i \neq 0$ are called the *elementary divisors of* $R(\lambda)$ at λ_0 . The finite zeros of polynomial matrices are also called eigenvalues.

For rational matrices having full column normal rank, zeros can be determined by the following equivalent definition [3, Chapter 27].

DEFINITION 2.1 (Zero location). *A rational matrix $R(\lambda) \in \mathbb{C}(\lambda)^{p \times m}$ with full column normal rank has a zero at λ_0 if there exists a rational vector $x(\lambda) \in \mathbb{C}(\lambda)^m$ defined at λ_0 such that $x(\lambda_0) \neq 0$ and*

$$(2.2) \quad \lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0.$$

It follows that, in the regular case, i.e., $p = m$ and $\det R(\lambda) \not\equiv 0$, $R(\lambda)$ has a zero at λ_0 if there exist rational vectors $x(\lambda) \in \mathbb{C}(\lambda)^m$ and $y(\lambda) \in \mathbb{C}(\lambda)^m$ defined at λ_0 such that

$$(2.3) \quad \lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0 \quad \text{and} \quad \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0,$$

with $x(\lambda_0) \neq 0$ and $y(\lambda_0) \neq 0$. Then $x(\lambda)$ and $y(\lambda)$ are said to be right and left *root vectors* of $R(\lambda)$ associated with λ_0 , respectively [18]. Their evaluation at λ_0 , that is, $x(\lambda_0)$ and $y(\lambda_0)$, are called right and left *eigenvectors* of $R(\lambda)$ associated with λ_0 .

REMARK 2.2. The REP is often defined in the literature as the problem of finding scalars $\lambda_0 \in \mathbb{C}$ and nonzero vectors $x \in \mathbb{C}^m$ and $y \in \mathbb{C}^m$ such that $R(\lambda_0)x = 0$ and $y^T R(\lambda_0) = 0$. But this definition assumes that λ_0 is a zero of $R(\lambda)$ but not a pole, as $R(\lambda)$ would be defined at λ_0 . In this paper we consider the more general definition in (2.3) to define the REP in order to capture all the zeros of $R(\lambda)$.

REMARK 2.3. The rational vectors $x(\lambda)$ and $y(\lambda)$ appearing in (2.3) are root vectors of $R(\lambda)$ for the special case of a regular rational matrix $R(\lambda)$, see [18, Definition 3.4] for a more general definition of root vectors for rational matrices, not necessarily regular or square, and over arbitrary fields. At least in the regular case, root vectors had been studied in the more general context of nonlinear eigenvalue problems over the complex field: see [10] and the references therein; for the (more involved) nonregular case, a thorough analysis has recently been made for polynomial [7] and rational [18] matrices over an arbitrary field.

In particular, the theory of root vectors for rational functions can be refined [18] so that they not only identify zeros, but also their partial multiplicities (via the concept of a maximal set). As in this paper we will restrict to simple zeros, these details are not important. We refer the reader who would like to know more about maximal sets of root vectors to [7, 10, 18] and the references therein.

EXAMPLE 2.4. Consider the rational matrix $R(\lambda) = \begin{bmatrix} \frac{1}{\lambda-1} & 0 \\ 0 & 1 \end{bmatrix}$. It is clear that $R(\lambda)$ has a pole at $\lambda_0 = 1$, as $R(\lambda)$ is not defined at 1, but it is less clear that $R(\lambda)$ has also a zero at 1. If we consider $x(\lambda) = \begin{bmatrix} \lambda-1 \\ -1 \end{bmatrix}$ we have that $\lim_{\lambda \rightarrow 1} R(\lambda)x(\lambda) = 0$ and $x(1) \neq 0$, i.e., $x(\lambda)$ is a right root vector of $R(\lambda)$ associated with $\lambda_0 = 1$. Thus $\lambda_0 = 1$ is also a zero of $R(\lambda)$. Indeed, the Smith–McMillan form of $R(\lambda)$ is $\begin{bmatrix} \frac{1}{\lambda-1} & 0 \\ 0 & \lambda-1 \end{bmatrix}$, which shows that $R(\lambda)$ has a zero and a pole at $\lambda_0 = 1$.

2.2. Null spaces. For any rational matrix $R(\lambda)$ (regular or singular), we define the following rational vector spaces over $\mathbb{C}(\lambda)$:

$$\ker R(\lambda) = \{x(\lambda) \in \mathbb{C}(\lambda)^{m \times 1} : R(\lambda)x(\lambda) = 0\}, \text{ and} \\ \text{coker } R(\lambda) = \{y(\lambda)^T \in \mathbb{C}(\lambda)^{1 \times p} : y(\lambda)^T R(\lambda) = 0\},$$

which are called the right and left null spaces over $\mathbb{C}(\lambda)$ of $R(\lambda)$, respectively.

For a finite $\lambda_0 \in \mathbb{C}$ that is not a pole of $R(\lambda)$, $\ker R(\lambda_0)$ and $\text{coker } R(\lambda_0)$ denote the right and left null spaces over \mathbb{C} of the constant matrix $R(\lambda_0)$, respectively. Namely,

$$\ker R(\lambda_0) = \{x \in \mathbb{C}^{m \times 1} : R(\lambda_0)x = 0\}, \text{ and} \\ \text{coker } R(\lambda_0) = \{y^T \in \mathbb{C}^{1 \times p} : y^T R(\lambda_0) = 0\}.$$

The vector spaces $\ker R(\lambda_0)$ and $\text{coker } R(\lambda_0)$ are called the right and left null spaces of $R(\lambda)$ at λ_0 , respectively.

The definition of $\ker R(\lambda_0)$ and $\text{coker } R(\lambda_0)$ can be generalized to the case where λ_0 is possibly a pole of $R(\lambda)$ as follows:

$$\ker R(\lambda_0) := \{x \in \mathbb{C}^{m \times 1} : \text{there exists } x(\lambda) \in \mathbb{C}(\lambda)^{m \times 1} \text{ with} \\ \lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0 \text{ and } x(\lambda_0) = x\}, \text{ and}$$

$$\text{coker } R(\lambda_0) := \{y^T \in \mathbb{C}^{1 \times p} : \text{there exists } y(\lambda)^T \in \mathbb{C}(\lambda)^{1 \times p} \text{ with} \\ \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0 \text{ and } y(\lambda_0) = y\}.$$

For regular rational matrices $R(\lambda)$, if λ_0 is a zero of $R(\lambda)$ then $\ker R(\lambda_0)$ and $\text{coker } R(\lambda_0)$ are non trivial and contain the right and left eigenvectors of $R(\lambda)$ associated with λ_0 , respectively. In [18, Definition 3.8 and Remark 3.9] one can find a more general definition for $\ker R(\lambda_0)$ and $\text{coker } R(\lambda_0)$ valid over an arbitrary field \mathbb{F} and that specializes to the one above when $\mathbb{F} = \mathbb{C}$.

2.3. Polynomial system matrices and transfer function matrices. It is known that any rational matrix $R(\lambda) \in \mathbb{C}(\lambda)^{p \times m}$ can be written with an expression of the form

$$(2.4) \quad R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda),$$

where $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ are matrix polynomials, with $A(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ regular. In addition, the representation (2.4) is not unique. The associated polynomial matrix

$$(2.5) \quad P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix}$$

is said to be a polynomial system matrix of the rational matrix $R(\lambda)$ in (2.4), and the rational matrix $R(\lambda)$ is called the transfer function matrix of $P(\lambda)$ (see [21]). Note that $R(\lambda)$ is the Schur complement of $A(\lambda)$ in $P(\lambda)$. The polynomial system matrix $P(\lambda)$ in (2.5), with $n > 0$, is said to be minimal if the polynomial matrices

$$(2.6) \quad \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A(\lambda) & B(\lambda) \end{bmatrix}$$

have no eigenvalues in \mathbb{C} [21]. An important consequence of minimality is that, if a polynomial system matrix $P(\lambda)$ is minimal, the poles of its transfer function matrix $R(\lambda)$ are the zeros of $A(\lambda)$, and the zeros of $R(\lambda)$ are the zeros of $P(\lambda)$, together with their partial multiplicities [21, Chapter 3, Theorem 4.1]. The notion of minimality was extended in a local sense in [5]. For that, if the matrices in (2.6) have no eigenvalues in a nonempty subset $\Omega \subseteq \mathbb{C}$, $P(\lambda)$ is said to be minimal in Ω [5, Definition 3.1]. Then, if a polynomial system matrix is minimal in a set Ω , the poles of $R(\lambda)$ in Ω are the zeros of $A(\lambda)$ in Ω , and the zeros of $R(\lambda)$ in Ω are the zeros of $P(\lambda)$ in Ω , together with their partial multiplicities [5, Theorem 3.5]. If $P(\lambda)$ is minimal in $\Omega := \{\lambda_0\}$, with $\lambda_0 \in \mathbb{C}$, then $P(\lambda)$ is said to be minimal at λ_0 .

Given a polynomial system matrix $P(\lambda)$ minimal at λ_0 , the following Proposition 2.5 establishes the relation between the right and left null spaces of $P(\lambda)$ and its transfer function matrix $R(\lambda)$ at λ_0 , and it is valid for arbitrary rational matrices $R(\lambda)$ (regular or singular). This result was stated in [6, Lemma 2.1] in the (easier) special case when λ_0 is not a pole of $R(\lambda)$, which follows from [4, Proposition 5.1] and [4, Proposition 5.2]. Here, we give a proof of the general case.

PROPOSITION 2.5. *Let*

$$P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} \in \mathbb{C}[\lambda]^{(n+p) \times (n+m)}$$

be a polynomial system matrix, with $A(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ regular, and transfer function matrix $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda) \in \mathbb{C}(\lambda)^{p \times m}$. Let $\lambda_0 \in \mathbb{C}$. If $P(\lambda)$ is minimal at λ_0 then the following statements hold:

(a) *The linear map*

$$L_r : \ker R(\lambda_0) \longrightarrow \ker P(\lambda_0), \quad x_0 \longmapsto \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1} B(\lambda) x(\lambda) \\ x_0 \end{bmatrix}$$

is a bijection, where $x(\lambda) \in \mathbb{C}(\lambda)^m$ is such that $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ and $x(\lambda_0) = x_0$.

(b) *The linear map*

$$L_\ell : \operatorname{coker} R(\lambda_0) \longrightarrow \operatorname{coker} P(\lambda_0), \quad y_0^T \longmapsto \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T C(\lambda) A(\lambda)^{-1} & y_0^T \end{bmatrix}$$

is a bijection, where $y(\lambda) \in \mathbb{C}(\lambda)^p$ is such that $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0$ and $y(\lambda_0) = y_0$.

Proof. We only prove part (a) since part (b) is analogous.

First, we prove that L_r is well-defined. Let $x_0 \in \ker R(\lambda_0)$. Then there exists a rational vector $x(\lambda) \in \mathbb{C}(\lambda)^m$ such that $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ and $x(\lambda_0) = x_0$, by definition. Since $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ we have that

$$(2.7) \quad \lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} \begin{bmatrix} A(\lambda)^{-1} B(\lambda) x(\lambda) \\ x(\lambda) \end{bmatrix} = 0.$$

Then, note that (2.7) can be rewritten as

$$(2.8) \quad \lim_{\lambda \rightarrow \lambda_0} \left(\begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} A(\lambda)^{-1} B(\lambda) x(\lambda) + \begin{bmatrix} B(\lambda) \\ D(\lambda) \end{bmatrix} x(\lambda) \right) = 0.$$

Since $P(\lambda)$ is minimal at λ_0 , we know that $\begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix}$ has a rational left inverse that is defined at λ_0 . In other words, there exists a rational matrix $H(\lambda)$ such that λ_0 is not a pole of $H(\lambda)$ and $H(\lambda) \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} = I$ (the claim follows immediately from the observation that the local Smith form of $\begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix}$ at λ_0 is $\begin{bmatrix} I \\ 0 \end{bmatrix}$). Taking into account (2.8), we obtain that the limit

$$\begin{aligned} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1} B(\lambda) x(\lambda) &= \lim_{\lambda \rightarrow \lambda_0} H(\lambda) \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} A(\lambda)^{-1} B(\lambda) x(\lambda) \\ &= - \lim_{\lambda \rightarrow \lambda_0} H(\lambda) \begin{bmatrix} B(\lambda) \\ D(\lambda) \end{bmatrix} x(\lambda) = H(\lambda_0) \begin{bmatrix} B(\lambda_0) \\ D(\lambda_0) \end{bmatrix} x \end{aligned}$$

is defined. Thus, $\lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} A(\lambda)^{-1} B(\lambda) x(\lambda) \\ x(\lambda) \end{bmatrix} \in \ker P(\lambda_0)$. This proves that L_r is well-defined. Next, note that if $\begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1} B(\lambda) x(\lambda) \\ x_0 \end{bmatrix} = 0$ then $x_0 = 0$, and this establishes the injectivity of L_r . Finally, to see that L_r is a bijection, we prove that $\dim \ker R(\lambda_0) = \dim \ker P(\lambda_0)$. Denote by s the number of positive invariant orders of $R(\lambda)$ at λ_0 ; observe that s is also equal to the number of partial multiplicities of $P(\lambda)$ at λ_0 , by [5, Theorem 3.5], since $P(\lambda)$ is minimal at λ_0 . By [18, Theorem 3.10],

$$\dim \ker R(\lambda_0) = \dim \ker R(\lambda) + s, \quad \dim \ker P(\lambda_0) = \dim \ker P(\lambda) + s.$$

On the other hand, by [19, Lemma 2.1], $\dim \ker P(\lambda) = \dim \ker R(\lambda)$, and hence $\dim \ker P(\lambda_0) = \dim \ker R(\lambda_0)$. \square

REMARK 2.6 (Notation). If λ_0 is a pole of $R(\lambda)$, then $A(\lambda)^{-1}$ is not defined at λ_0 . However, as we proved in Proposition 2.5, if $P(\lambda)$ is minimal at λ_0 the limit $\lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda)$ is defined at λ_0 , even if λ_0 is a pole of $R(\lambda)$. Thus, in what follows, for a given $x(\lambda) \in \mathbb{C}(\lambda)^m$ such that $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$, we will slightly abuse notation and write $A(\lambda_0)^{-1}B(\lambda_0)x(\lambda_0)$ as a shorthand for $\lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda)$; this should be interpreted as the rational vector $A(\lambda)^{-1}B(\lambda)x(\lambda)$ evaluated at λ_0 (this evaluation is guaranteed to yield a complex vector by Proposition 2.5). The same notation will be used for the limit $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T C(\lambda)A(\lambda)^{-1}$ for any $y(\lambda) \in \mathbb{C}(\lambda)^p$ such that $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0$, i.e., $y(\lambda_0)^T C(\lambda_0)A(\lambda_0)^{-1}$.

We know that if λ_0 is a zero of a rational matrix $R(\lambda)$ then λ_0 is also a zero (an eigenvalue) of any polynomial system matrix $P(\lambda)$ of $R(\lambda)$, with same partial multiplicities, whenever $P(\lambda)$ is minimal at λ_0 [5, Theorem 3.5]. Then Proposition 2.5 shows a corresponding relation between the eigenvectors, even if λ_0 is not only a zero but also a pole of $R(\lambda)$. Moreover, since the maps in Proposition 2.5 are bijections, they preserve linear independence. Thus, one can recover a basis of $\ker P(\lambda_0)$ (resp. $\text{coker } P(\lambda_0)$) from a basis of $\ker R(\lambda_0)$ (resp. $\text{coker } R(\lambda_0)$), and conversely.

EXAMPLE 2.7. Consider the rational matrix $R(\lambda)$ in Example 2.4, and the right root vector $x(\lambda)$ of $R(\lambda)$ associated with $\lambda_0 = 1$. Define the polynomial system matrix

$$P(\lambda) := \left[\begin{array}{c|cc} \lambda - 1 & 1 & 0 \\ \hline 0 & 1 & 0 \\ -1 & 0 & 1 \end{array} \right] =: \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix},$$

whose transfer function matrix is $R(\lambda)$ and is minimal at $\lambda_0 = 1$. Then $\lambda_0 = 1$ is also a zero of $P(\lambda)$ and $v = \begin{bmatrix} A(\lambda_0)^{-1}B(\lambda_0)x(\lambda_0) \\ x(\lambda_0) \end{bmatrix} = [-1 \quad 0 \quad -1]^T$ is a right eigenvector of $P(\lambda)$ associated with $\lambda_0 = 1$.

3. Perturbation expansion for simple zeros. Let $R(\lambda)$ be the transfer function matrix of a polynomial system matrix $P(\lambda)$. In this section, we assume that $R(\lambda)$ is regular (so that $P(\lambda)$ is also regular) and that λ_0 is a simple finite zero of $R(\lambda)$. If $P(\lambda)$ is minimal at λ_0 , λ_0 is also a simple eigenvalue of $P(\lambda)$ [5, Theorem 3.5]. We first study how analytic perturbations of $P(\lambda)$ affect to simple eigenvalues and local minimality.

THEOREM 3.1. Let $P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix}$ be a regular polynomial system matrix, with $A(\lambda)$ regular. Let $\lambda_0 \in \mathbb{C}$ be a finite simple eigenvalue of $P(\lambda)$, with right (resp. left) eigenvector v (resp. w), and assume that $P(\lambda)$ is minimal at λ_0 . Let

$$\widehat{P}(\lambda, \epsilon) = \begin{bmatrix} -\widehat{A}(\lambda, \epsilon) & \widehat{B}(\lambda, \epsilon) \\ \widehat{C}(\lambda, \epsilon) & \widehat{D}(\lambda, \epsilon) \end{bmatrix}$$

be a matrix-valued function that can be written as $\widehat{P}(\lambda, \epsilon) = \sum_{i=0}^d P_i(\epsilon)\lambda^i$ where each $P_i(\epsilon)$ is an analytic matrix function of ϵ in some open ball centered at 0. Assume that $\widehat{P}(\lambda, 0) = P(\lambda)$. Then there exists a constant $r > 0$ such that the following statements hold:

- (a) There exist functions $\widehat{\lambda}_0(\epsilon)$, $\widehat{v}(\epsilon)$ and $\widehat{w}(\epsilon)$ such that (i) $\widehat{\lambda}_0(\epsilon)$, $\widehat{v}(\epsilon)$ and $\widehat{w}(\epsilon)$ are analytic in ϵ for all $|\epsilon| < r$; and (ii) $\widehat{\lambda}_0(0) = \lambda_0$, $\widehat{w}(0) = w$ and $\widehat{v}(0) = v$.

- (b) $\widehat{P}(\lambda, \epsilon)$ has a simple zero at $\widehat{\lambda}_0(\epsilon)$, for all $|\epsilon| < r$, with associated right (resp. left) eigenvector $\widehat{v}(\epsilon)$ (resp. $\widehat{w}(\epsilon)$).
- (c) $\widehat{P}(\lambda, \epsilon)$, viewed as a polynomial matrix in λ , is a polynomial system matrix minimal at both λ_0 and $\widehat{\lambda}_0(\epsilon)$ for all $|\epsilon| < r$.
- (d) The transfer function matrix $\widehat{R}(\lambda, \epsilon)$ of $\widehat{P}(\lambda, \epsilon)$ has a simple zero at $\widehat{\lambda}_0(\epsilon)$ for all $|\epsilon| < r$.

Proof. Items (a) and (b) are classical results, see for example [1, 14].

For (c), we only prove that for sufficiently small ϵ , $\widehat{P}(\lambda, \epsilon)$ must be minimal at $\widehat{\lambda}_0(\epsilon)$. Indeed, minimality is equivalent to $\begin{bmatrix} -\widehat{A}(\widehat{\lambda}_0(\epsilon), \epsilon) & \widehat{B}(\widehat{\lambda}_0(\epsilon), \epsilon) \end{bmatrix}$ having full row rank and $\begin{bmatrix} -\widehat{A}(\widehat{\lambda}_0(\epsilon), \epsilon) \\ \widehat{C}(\widehat{\lambda}_0(\epsilon), \epsilon) \end{bmatrix}$ having full column rank. On the other hand, $\widehat{P}(\widehat{\lambda}_0(\epsilon), \epsilon)$ is analytic in ϵ ; hence, so is any of its submatrices and the statement follows by the continuity of singular values [14].

Finally, (d) follows from (b) and (c). That is, since $\widehat{P}(\lambda, \epsilon)$ is minimal at $\widehat{\lambda}_0(\epsilon)$ and has a simple zero at $\widehat{\lambda}_0(\epsilon)$, by [5, Theorem 3.5], it implies that $\widehat{\lambda}_0(\epsilon)$ is also a simple zero of $\widehat{R}(\lambda, \epsilon)$. \square

REMARK 3.2. For a given rational matrix $R(\lambda)$, not every small perturbation of $R(\lambda)$ has a polynomial system matrix $\widehat{P}(\lambda, \epsilon)$ satisfying the assumptions in Theorem 3.1. For instance, consider the rational matrix

$$R(\lambda) = \begin{bmatrix} \frac{1}{\lambda} & \frac{1}{\lambda^3} \\ 0 & \frac{1}{\lambda} \end{bmatrix}.$$

The invariant orders of $R(\lambda)$ at 0 are $-3, 1$, and there are not any other finite zeros or poles. For any $\epsilon > 0$, let

$$R(\lambda) + \epsilon \Delta R(\lambda) = \frac{1}{\lambda^3} \begin{bmatrix} \lambda^2 + \epsilon & 1 \\ 0 & \lambda^2 - \epsilon \end{bmatrix}, \quad \text{where} \quad \Delta R(\lambda) = \begin{bmatrix} \frac{1}{\lambda^3} & 0 \\ 0 & -\frac{1}{\lambda^3} \end{bmatrix}.$$

The invariant orders of $R(\lambda) + \epsilon \Delta R(\lambda)$ at 0 are $-3, -3$. Then, 0 is a pole. Moreover, the zeros of $R(\lambda) + \epsilon \Delta R(\lambda)$ are *not* analytic in ϵ : indeed, they can be shown to be equal to $\pm \epsilon^{1/2}$ and $\pm i \epsilon^{1/2}$. Thus, the statement of Theorem 3.1(d) is not satisfied for this particular perturbation. However, this is not a counterexample to the theorem because the assumptions are also false: suppose indeed that $P(\lambda) = \widehat{P}(\lambda, 0)$ such that $P(\lambda)$ and $\widehat{P}(\lambda, \epsilon)$ are polynomial system matrices of $R(\lambda)$ and $R(\lambda) + \epsilon \Delta R(\lambda)$, respectively, both minimal at 0. Then, by [5, Theorem 3.5], the matrices $A(\lambda)$ and $\widehat{A}(\lambda, \epsilon)$ must have determinant, respectively, $p(\lambda)\lambda^3$ and $f(\lambda, \epsilon)\lambda^6$ where $p(\lambda)$ is a polynomial with $p(0) \neq 0$ and $f(\lambda, \epsilon)$ is a function polynomial in λ , analytic in ϵ with $f(0, \epsilon) \neq 0$. But then, if $P(\lambda) = \widehat{P}(\lambda, 0)$, by the continuity of the determinant we would have that $f(\lambda, 0)\lambda^6 = p(\lambda)\lambda^3$, which is impossible. Although every rational matrix admits minimal polynomial system matrices, the crux here is the requirement $\widehat{P}(\lambda, 0) = P(\lambda)$. For example,

$$P(\lambda) = \left[\begin{array}{c|cc} \lambda^3 & \lambda^2 & 1 \\ \hline 1 & 0 & 0 \\ \lambda^2 & -\lambda & 0 \end{array} \right] \quad \text{and} \quad Q(\lambda, \epsilon) = \left[\begin{array}{c|cc} \lambda^3 & 0 & \epsilon & 0 \\ \hline 0 & \lambda^3 & \lambda^2 & 1 \\ 0 & 1 & 0 & 0 \\ \epsilon & \lambda^2 & -\lambda & 0 \end{array} \right]$$

are minimal polynomial system matrices of $R(\lambda)$ and $R(\lambda) + \epsilon \Delta R(\lambda)$, respectively, but clearly $Q(\lambda, 0) \neq P(\lambda)$ (as the sizes are different).

Nevertheless, in this paper we consider perturbed rational matrices having polynomial system matrices that always satisfy the assumptions in Theorem 3.1. More precisely, we consider rational matrices $R(\lambda)$ expressed as in (2.4), i.e., $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$, for arbitrary polynomial matrices $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$, and its associated polynomial system matrix $P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix}$. Then, we consider a perturbation of $P(\lambda)$ of the form:

$$(3.1) \quad P(\lambda) + \epsilon \Delta P(\lambda) := \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} + \epsilon \begin{bmatrix} -\Delta A(\lambda) & \Delta B(\lambda) \\ \Delta C(\lambda) & \Delta D(\lambda) \end{bmatrix},$$

where $\Delta A(\lambda)$, $\Delta B(\lambda)$, $\Delta C(\lambda)$ and $\Delta D(\lambda)$ are polynomial matrices. It is clear that $\hat{P}(\lambda, \epsilon) := P(\lambda) + \epsilon \Delta P(\lambda)$ satisfies the assumptions in Theorem 3.1 as it is a linear perturbation in ϵ of $P(\lambda)$. In order to study perturbations of simple eigenvalues, we will use Lemma 3.3, which is known for general polynomial matrices [22, Proof of Theorem 5] but that we state here for the particular case of polynomial system matrices.

LEMMA 3.3. *Let $\lambda_0 \in \mathbb{C}$ be a simple eigenvalue of a regular polynomial system matrix $P(\lambda)$; and let v and w be, respectively, right and left eigenvectors of $P(\lambda)$ associated with λ_0 . Consider the perturbed polynomial system matrix $P(\lambda) + \epsilon \Delta P(\lambda)$. Then, for $\epsilon > 0$ small enough, there exist $\Delta \lambda_0$ and Δv satisfying*

$$[P(\lambda_0 + \Delta \lambda_0) + \epsilon \Delta P(\lambda_0 + \Delta \lambda_0)] [v + \epsilon \Delta v] = 0.$$

That is, $\lambda_0 + \Delta \lambda_0$ is an eigenvalue for the perturbed polynomial system matrix $P(\lambda) + \epsilon \Delta P(\lambda)$. In addition,

$$(3.2) \quad |\Delta \lambda_0| = \epsilon \left| \frac{w^T \Delta P(\lambda_0) v}{w^T P'(\lambda_0) v} \right| + o(\epsilon).$$

Now, we consider the perturbed transfer function matrix of the perturbed polynomial system matrix in (3.1). Namely,

$$(3.3) \quad R(\lambda) + \Delta R(\lambda, \epsilon) := D(\lambda) + \epsilon \Delta D(\lambda) + \\ (C(\lambda) + \epsilon \Delta C(\lambda))(A(\lambda) + \epsilon \Delta A(\lambda))^{-1}(B(\lambda) + \epsilon \Delta B(\lambda)).$$

Then, we set

$$(3.4) \quad \Delta R(\lambda) := \left[\frac{\partial \Delta R(\lambda, \epsilon)}{\partial \epsilon} \right]_{\epsilon=0} = \Delta D(\lambda) + \Delta C(\lambda)A(\lambda)^{-1}B(\lambda) + C(\lambda)A(\lambda)^{-1}\Delta B(\lambda) \\ - C(\lambda)A(\lambda)^{-1}\Delta A(\lambda)A(\lambda)^{-1}B(\lambda).$$

Note that (3.4) implies the formal¹ expansion $\Delta R(\lambda, \epsilon) = \epsilon \Delta R(\lambda) + o(\epsilon)$.

In Corollary 3.6, we write (3.2) in terms of $R(\lambda)$ and $\Delta R(\lambda)$ by using root vectors [18]. For that, we need Lemma 3.4.

¹While one can always write such a formal expansion, when $\lambda \rightarrow \lambda_0$ and λ_0 is a pole of $R(\lambda)$ then the radius of convergence (in ϵ) of such an expansion may tend to 0. For example, consider

$$R(\lambda) + \Delta R(\lambda, \epsilon) = \frac{1}{\lambda - \epsilon} = \frac{1}{\lambda} \sum_{k=0}^{\infty} \frac{\epsilon^k}{\lambda^k}, \text{ where } R(\lambda) = \frac{1}{\lambda} \text{ and } \Delta R(\lambda, \epsilon) = \frac{\epsilon}{\lambda^2} + o(\epsilon).$$

For any fixed $\lambda \neq 0$, the radius of convergence of the power series in ϵ is $|\lambda|$. Hence, there is no uniform radius of convergence for λ in a neighbourhood of the pole $\lambda_0 = 0$.

LEMMA 3.4. Let $P(\lambda) \in \mathbb{C}[\lambda]^{(n+p) \times (n+p)}$ be a polynomial system matrix as in (2.5), with $A(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ regular, and transfer function matrix $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda) \in \mathbb{C}(\lambda)^{p \times p}$. Assume that $P(\lambda)$ is minimal at λ_0 . Let $v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \ker P(\lambda_0)$ and $w^T = \begin{bmatrix} w_1^T & w_2^T \end{bmatrix} \in \text{coker } P(\lambda_0)$, partitioned conformably to the blocks of $P(\lambda)$, with $v \neq 0$ and $w \neq 0$. Then, there exist right and left root vectors $x(\lambda) \in \mathbb{C}(\lambda)^p$ and $y(\lambda) \in \mathbb{C}(\lambda)^p$ of $R(\lambda)$, both associated with λ_0 , such that

$$(3.5) \quad w^T P'(\lambda_0)v = \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R'(\lambda)x(\lambda),$$

with $x(\lambda_0) = v_2$ and $y(\lambda_0) = w_2$, and that

$$(3.6) \quad w^T \Delta P(\lambda_0)v = \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T \Delta R(\lambda)x(\lambda),$$

where $\Delta R(\lambda)$ is defined as in (3.4).

Proof. By Proposition 2.5, we know that if $v \in \ker P(\lambda_0)$ and $w \in \text{coker } P(\lambda_0)$ then $v = \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda) \\ x(\lambda_0) \end{bmatrix}$ and $w^T = \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T C(\lambda)A(\lambda)^{-1} & y(\lambda_0)^T \end{bmatrix}$ for some $x(\lambda) \in \mathbb{C}(\lambda)^p$ and $y(\lambda) \in \mathbb{C}(\lambda)^p$ with $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ and $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R(\lambda) = 0$, respectively. Then

$$\begin{aligned} & w^T P'(\lambda_0)v \\ &= \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T C(\lambda)A(\lambda)^{-1} & y(\lambda_0)^T \end{bmatrix} \begin{bmatrix} -A'(\lambda_0) & B'(\lambda_0) \\ C'(\lambda_0) & D'(\lambda_0) \end{bmatrix} \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda) \\ x(\lambda_0) \end{bmatrix} \\ &= \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T (D'(\lambda) + C'(\lambda)A(\lambda)^{-1}B(\lambda) + C(\lambda)A(\lambda)^{-1}B'(\lambda) \\ & \quad - C(\lambda)A(\lambda)^{-1}A'(\lambda)A(\lambda)^{-1}B(\lambda))x(\lambda) \\ &= \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R'(\lambda)x(\lambda) \end{aligned}$$

and

$$\begin{aligned} & w^T \Delta P(\lambda_0)v \\ &= \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T C(\lambda)A(\lambda)^{-1} & y(\lambda_0)^T \end{bmatrix} \begin{bmatrix} -\Delta A(\lambda_0) & \Delta B(\lambda_0) \\ \Delta C(\lambda_0) & \Delta D(\lambda_0) \end{bmatrix} \begin{bmatrix} \lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda) \\ x(\lambda_0) \end{bmatrix} \\ &= \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T (\Delta D(\lambda) + \Delta C(\lambda)A(\lambda)^{-1}B(\lambda) + C(\lambda)A(\lambda)^{-1}\Delta B(\lambda) \\ & \quad - C(\lambda)A(\lambda)^{-1}\Delta A(\lambda)A(\lambda)^{-1}B(\lambda))x(\lambda) \\ &= \lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T \Delta R(\lambda)x(\lambda), \end{aligned}$$

with $x(\lambda_0) = v_2$ and $y(\lambda_0) = w_2$. We still have to prove that $x(\lambda_0) \neq 0$ and $y(\lambda_0) \neq 0$. We only prove that $x(\lambda_0) \neq 0$, since $y(\lambda_0) \neq 0$ can be proved analogously. Since $\lim_{\lambda \rightarrow \lambda_0} R(\lambda)x(\lambda) = 0$ we have that

$$\lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} A(\lambda)^{-1}B(\lambda)x(\lambda) + \begin{bmatrix} B(\lambda) \\ D(\lambda) \end{bmatrix} x(\lambda) = 0.$$

By contradiction, if we assume that $x(\lambda_0) = 0$, we get that $\lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} B(\lambda) \\ D(\lambda) \end{bmatrix} x(\lambda) = 0$ and, therefore, $\lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} A(\lambda)^{-1}B(\lambda)x(\lambda) = 0$. Taking into account that $P(\lambda)$

is minimal at λ_0 , this is, that $\lim_{\lambda \rightarrow \lambda_0} \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} = \begin{bmatrix} -A(\lambda_0) \\ C(\lambda_0) \end{bmatrix}$ has full column rank, we obtain that $\lim_{\lambda \rightarrow \lambda_0} A(\lambda)^{-1}B(\lambda)x(\lambda) = 0$, which is a contradiction since this implies that $v = 0$. \square

REMARK 3.5. We remark that $w^T P'(\lambda_0)v \neq 0$ by [1, Theorem 3.2], for right and left eigenvectors v and w of $P(\lambda)$, respectively, associated with a simple eigenvalue λ_0 . Taking into account Lemma 3.4, this also implies that $\lim_{\lambda \rightarrow \lambda_0} y(\lambda)^T R'(\lambda)x(\lambda) \neq 0$, for right and left root vectors $x(\lambda)$ and $y(\lambda)$ of $R(\lambda)$, respectively, associated with a simple zero λ_0 of $R(\lambda)$.

COROLLARY 3.6. *Assume that the polynomial system matrix $P(\lambda)$ in Lemma 3.3 is minimal at λ_0 , and let $R(\lambda)$ be the transfer function matrix of $P(\lambda)$. Let $R(\lambda) + \Delta R(\lambda, \epsilon)$ be the transfer function matrix of the perturbed polynomial system matrix $P(\lambda) + \epsilon \Delta P(\lambda)$ in (3.1). Define $\Delta R(\lambda)$ as in (3.4). Then, for $\epsilon > 0$ small enough, $\lambda_0 + \Delta \lambda_0$ is also a zero for $R(\lambda) + \Delta R(\lambda, \epsilon)$. In addition, there exist right and left root vectors $x(\lambda) \in \mathbb{C}(\lambda)^p$ and $y(\lambda) \in \mathbb{C}(\lambda)^p$ of $R(\lambda)$ associated with λ_0 , respectively, such that*

$$(3.7) \quad |\Delta \lambda_0| = \epsilon \lim_{\lambda \rightarrow \lambda_0} \left| \frac{y(\lambda)^T \Delta R(\lambda)x(\lambda)}{y(\lambda)^T R'(\lambda)x(\lambda)} \right| + o(\epsilon).$$

4. Condition number formulae. In this section, we define absolute condition numbers for simple zeros of minimal polynomial system matrices as in (1.2). For that, we express the block polynomial matrices $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ in terms of the monomial basis. Namely,

$$D(\lambda) = \sum_{i=0}^{k_D} D_i \lambda^i \in \mathbb{C}[\lambda]^{p \times p}, \quad C(\lambda) = \sum_{i=0}^{k_C} C_i \lambda^i \in \mathbb{C}[\lambda]^{p \times n},$$

$$A(\lambda) = \sum_{i=0}^{k_A} A_i \lambda^i \in \mathbb{C}[\lambda]^{n \times n}, \quad \text{and} \quad B(\lambda) = \sum_{i=0}^{k_B} B_i \lambda^i \in \mathbb{C}[\lambda]^{n \times p}.$$

Then, we perturb the matrix coefficients of each matrix polynomial, and the perturbed rational matrix we get is (3.3) with

$$\Delta D(\lambda) = \sum_{i=0}^{k_D} \Delta D_i \lambda^i \in \mathbb{C}[\lambda]^{p \times p}, \quad \Delta C(\lambda) = \sum_{i=0}^{k_C} \Delta C_i \lambda^i \in \mathbb{C}[\lambda]^{p \times n},$$

$$\Delta A(\lambda) = \sum_{i=0}^{k_A} \Delta A_i \lambda^i \in \mathbb{C}[\lambda]^{n \times n}, \quad \text{and} \quad \Delta B(\lambda) = \sum_{i=0}^{k_B} \Delta B_i \lambda^i \in \mathbb{C}[\lambda]^{n \times p}.$$

With the above representation, the goal is to study how (simple) zeros of $P(\lambda)$ change to first order in ϵ . Recall that the perturbed rational matrix (3.3) is the transfer function matrix of the perturbed polynomial system matrix (3.1) and, by the discussion in Section 3, it is clear that it is equivalent to define such a structured condition number for the PEP referring to simple eigenvalues of (structured perturbations of) $P(\lambda)$ or simple zeros of the corresponding perturbations of its transfer function $R(\lambda)$.

DEFINITION 4.1 (Absolute structured condition number of a simple zero). *Let $R(\lambda)$ be a regular rational matrix as in (2.4), and let $P(\lambda)$ be a regular polynomial system matrix for $R(\lambda)$ as in (2.5). Assume that $\lambda_0 \in \mathbb{C}$ is a simple finite eigenvalue*

of $R(\lambda)$ and that $P(\lambda)$ is minimal at λ_0 . Then, we define the structured absolute condition number of the eigenvalue λ_0 as follows:

$$\kappa_S(\lambda_0) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda_0|}{\epsilon} : [P(\lambda_0 + \Delta\lambda_0) + \epsilon\Delta P(\lambda_0 + \Delta\lambda_0)][v + \epsilon\Delta v] = 0 \quad \text{with} \right. \\ \left. \|\Delta D_i\|_2 \leq d_i, \|\Delta A_i\|_2 \leq a_i, \|\Delta B_i\|_2 \leq b_i, \|\Delta C_i\|_2 \leq c_i \right\},$$

where the a_i, b_i, c_i, d_i are nonnegative parameters, and v is a right eigenvector of $P(\lambda)$ associated with λ_0 . Equivalently, $\kappa_S(\lambda_0)$ can be defined by using the transfer function matrix $R(\lambda)$ of $P(\lambda)$ as follows:

$$\kappa_S(\lambda_0) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda_0|}{\epsilon} : \lambda_0 + \Delta\lambda_0 \text{ is a zero of } R(\lambda) + \Delta R(\lambda, \epsilon) \quad \text{with} \right. \\ \left. \|\Delta D_i\|_2 \leq d_i, \|\Delta A_i\|_2 \leq a_i, \|\Delta B_i\|_2 \leq b_i, \|\Delta C_i\|_2 \leq c_i \right\}.$$

The nonnegative parameters a_i, b_i, c_i, d_i will be called weights and provide some freedom in how perturbations to the rational matrix $R(\lambda)$ are measured. For example, two natural choices are either $a_i = b_i = c_i = d_i = 1$ (the coefficients are allowed to be perturbed by perturbations of uniform absolute magnitude) or $a_i = \|A_i\|, b_i = \|B_i\|, c_i = \|C_i\|, d_i = \|D_i\|$ (the coefficients are allowed to be perturbed by perturbations of uniform relative magnitude). We prefer to keep the weights as free parameters in our analysis as, in principle, a particular problem may suggest more peculiar choices. In Theorem 4.4 we derive a formula for $\kappa_S(\lambda_0)$.

REMARK 4.2. For any choice of positive weights, Definition 4.1 yields an absolute structured condition number (1.7) for the eigenvalue-computing function $f : P(\lambda) \mapsto \lambda_0$ defined in Section 1.1. Here, \mathcal{U} is the vector space of polynomial system matrices of the form (1.2) such that $A(\lambda), B(\lambda), C(\lambda)$ and $D(\lambda)$ are matrix polynomials of degree at most, resp., k_A, k_B, k_C and k_D . In particular, Definition 4.1 is the condition number induced by the choice of the norm

$$(4.1) \quad P(\lambda) \mapsto \|P(\lambda)\| = \max \left\{ \max_{i=0}^{k_A} \frac{\|A_i\|}{a_i}, \max_{i=0}^{k_B} \frac{\|B_i\|}{b_i}, \max_{i=0}^{k_C} \frac{\|C_i\|}{c_i}, \max_{i=0}^{k_D} \frac{\|D_i\|}{d_i} \right\},$$

with the convention that if, e.g., $B(\lambda) = 0$ then $k_B = -\infty$ and $\max_{i=0}^{k_B} \frac{\|B_i\|}{b_i} = 0$. We leave as an exercise to the reader the straightforward verification that (4.1) is indeed a norm, for any choice of the parameters.

REMARK 4.3. A subtler problem is to assess whether Definition 4.1 yields a condition number for the eigenvalue-computing function $f : R(\lambda) \mapsto \lambda_0$: indeed, (4.1) is clearly *not* a norm on the vector space of rational functions. On the other hand, it is worth observing that the set of allowed perturbation is highly structured with respect to the original REP, i.e., not every rational small perturbation of $R(\lambda)$ can be written in this form (in fact, in a precise geometric sense, only very few can). Indeed, note that the tangent space at $\epsilon = 0$ to the set of perturbations that we consider is

$$\{\Delta D(\lambda) + \Delta C(\lambda)A(\lambda)^{-1}B(\lambda) + C(\lambda)A(\lambda)^{-1}\Delta B(\lambda) - C(\lambda)A(\lambda)^{-1}\Delta A(\lambda)A(\lambda)^{-1}B(\lambda)\}$$

which is a finite dimensional vector space over \mathbb{C} . Indeed, even neglecting the issue of the potential non-unicity of the representation, its dimension is readily verified

to be $\leq pm(\eta_A + \eta_B + \eta_C + \eta_D + 4)$ where $\eta_D = k_D$ if $D(\lambda) \neq 0$ or $\eta_D = -1$ otherwise, and similarly for k_A, k_B, k_C . In contrast, the vector space of $p \times m$ rational matrices has (uncountably) infinite dimension over \mathbb{C} . Nevertheless, we believe that our definition provide a sensible practical tool and that restricting to this set of perturbation is justified when working with the representation of rational matrices that we are considering in this paper.

Now, we derive a formula for the condition number introduced in Definition 4.1.

THEOREM 4.4 (Condition number formula). *Let $P(\lambda) \in \mathbb{C}[\lambda]^{(n+p) \times (n+p)}$ be a regular polynomial system matrix as in (2.5), with $A(\lambda) \in \mathbb{C}[\lambda]^{n \times n}$ regular, and transfer function matrix $R(\lambda)$. Assume that $\lambda_0 \in \mathbb{C}$ is a simple eigenvalue of $R(\lambda)$ and that $P(\lambda)$ is minimal at λ_0 . Then the structured absolute condition number $\kappa_S(\lambda_0)$ in Definition 4.1 is given by*

$$\kappa_S(\lambda_0) = \frac{1}{K} \left[\|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \quad \|y(\lambda_0)\|_2 \right] S(\lambda_0) \left[\frac{\|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2}{\|x(\lambda_0)\|_2} \right],$$

by using the notation in Remark 2.6, where $x(\lambda)$ and $y(\lambda)$ are, respectively, right and left root vectors of $R(\lambda)$ associated with λ_0 ,

$$K := \lim_{\lambda \rightarrow \lambda_0} |y(\lambda)^T R'(\lambda) x(\lambda)| \quad \text{and} \quad S(\lambda) := \begin{bmatrix} \sum_{i=0}^{k_A} a_i |\lambda|^i & \sum_{i=0}^{k_B} b_i |\lambda|^i \\ \sum_{i=0}^{k_C} c_i |\lambda|^i & \sum_{i=0}^{k_D} d_i |\lambda|^i \end{bmatrix}.$$

Proof. By using Lemma 3.3 and Proposition 2.5, and ignoring $o(\epsilon)$ terms, we get

$$\begin{aligned} |\Delta \lambda_0| &\leq \frac{\epsilon}{K} \left(\|x(\lambda_0)\|_2 \|y(\lambda_0)\|_2 \sum_{i=0}^{k_D} d_i |\lambda_0|^i + \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2 \|y(\lambda_0)\|_2 \sum_{i=0}^{k_C} c_i |\lambda_0|^i + \right. \\ &\quad \left. \|x(\lambda_0)\|_2 \|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \sum_{i=0}^{k_B} b_i |\lambda_0|^i + \right. \\ &\quad \left. \|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2 \sum_{i=0}^{k_A} a_i |\lambda_0|^i \right) = \\ &\frac{\epsilon}{K} \left[\|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \quad \|y(\lambda_0)\|_2 \right] S(\lambda_0) \left[\frac{\|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2}{\|x(\lambda_0)\|_2} \right]. \end{aligned}$$

This provides an upper bound for the condition number $\kappa_S(\lambda_0)$. We need to show that this upper bound is sharp, i.e., there is a perturbation of the form that we consider and that achieves the bound. Let us define

$$\mu = \begin{cases} 0 & \text{if } \lambda_0 = 0; \\ \frac{\lambda_0}{|\lambda_0|} & \text{otherwise,} \end{cases}$$

and let us consider the following rank-1 matrix perturbations:

•

$$\Delta D_i = d_i \mu^i \frac{y(\lambda_0) x(\lambda_0)^T}{\|x(\lambda_0)\|_2 \|y(\lambda_0)\|_2},$$

for $i = 0, 1, \dots, k_D$.

•

$$\Delta C_i = c_i \mu^i \frac{y(\lambda_0) (A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0))^T}{\|y(\lambda_0)\|_2 \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2},$$

if $A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0) \neq 0$ or $\Delta C_i = 0$ otherwise, for $i = 0, 1, \dots, k_C$.

•

$$\Delta B_i = b_i \mu^i \frac{(y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1})^T x(\lambda_0)^T}{\|x(\lambda_0)\|_2 \|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2},$$

if $y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1} \neq 0$ or $\Delta B_i = 0$ otherwise, for $i = 0, 1, \dots, k_B$, and

•

$$\Delta A_i = -a_i \mu^i \frac{(y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1})^T (A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0))^T}{\|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2},$$

if $A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0) \neq 0$ and $y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1} \neq 0$ or $\Delta A_i = 0$ otherwise, for $i = 0, 1, \dots, k_A$.

It is clear that these matrix perturbations satisfy $\|\Delta D_i\| \leq d_i$, $\|\Delta B_i\| \leq b_i$, $\|\Delta C_i\| \leq c_i$ and $\|\Delta A_i\| \leq a_i$. Then we have

$$y(\lambda_0)^T \left(\epsilon \sum_{i=0}^{k_D} \lambda_0^i \Delta D_i \right) x(\lambda_0) = \epsilon \|x(\lambda_0)\|_2 \|y(\lambda_0)\|_2 \left(\sum_{i=0}^{k_D} |\lambda_0|^i d_i \right),$$

$$y(\lambda_0)^T \left(\epsilon \sum_{i=0}^{k_C} \lambda_0^i \Delta C_i \right) (A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)) = \epsilon \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2 \|y(\lambda_0)\|_2 \sum_{i=0}^{k_C} c_i |\lambda_0|^i,$$

$$(y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}) \left(\epsilon \sum_{i=0}^{k_B} \lambda_0^i \Delta B_i \right) x(\lambda_0) = \epsilon \|x(\lambda_0)\|_2 \|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \sum_{i=0}^{k_B} b_i |\lambda_0|^i, \blacksquare$$

and

$$(y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}) \left(\epsilon \sum_{i=0}^{k_A} \lambda_0^i \Delta A_i \right) (A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)) = \\ - \epsilon \|y(\lambda_0)^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \|A(\lambda_0)^{-1} B(\lambda_0) x(\lambda_0)\|_2 \sum_{i=0}^{k_A} a_i |\lambda_0|^i.$$

Hence, for these perturbations all the norm inequalities are satisfied as equalities and, so, the upper bound is tight. \square

4.1. Condition number formula for zeros of $R(\lambda)$ that are not poles.

Theorem 4.4 can be simplified for those zeros λ_0 such that $\det A(\lambda_0) \neq 0$. Note that, if $\det A(\lambda_0) \neq 0$, the polynomial system matrix $P(\lambda)$ is minimal at λ_0 . It is worth emphasizing that if $\det A(\lambda_0) = 0$ not necessarily λ_0 is a pole of $A(\lambda)$, if we do not assume minimality on $P(\lambda)$ at λ_0 , but all the poles of $R(\lambda)$ satisfy that $\det A(\lambda_0) = 0$. So if $\det A(\lambda_0) \neq 0$ then λ_0 is not a pole of $R(\lambda)$, and the root vectors in Theorem 4.4 can be considered to be constant.

COROLLARY 4.5 (Condition number formula for a zero λ_0 with $\det A(\lambda_0) \neq 0$).

Let $R(\lambda)$ be a regular rational matrix as in (2.4), and assume that $\lambda_0 \in \mathbb{C}$ is a simple eigenvalue of $R(\lambda)$ with $\det A(\lambda_0) \neq 0$. Then the structured absolute condition number $\kappa_S(\lambda_0)$ in Definition 4.1 is given by

$$\kappa_S(\lambda_0) = \frac{1}{|y^T R'(\lambda_0) x|} \left[\|y^T C(\lambda_0) A(\lambda_0)^{-1}\|_2 \quad \|y\|_2 \right] S(\lambda_0) \begin{bmatrix} \|A(\lambda_0)^{-1} B(\lambda_0) x\|_2 \\ \|x\|_2 \end{bmatrix},$$

where x and y are, respectively, right and left eigenvectors of $R(\lambda)$ associated with λ_0 and

$$S(\lambda) := \begin{bmatrix} \sum_{i=0}^{k_A} a_i |\lambda|^i & \sum_{i=0}^{k_B} b_i |\lambda|^i \\ \sum_{i=0}^{k_C} c_i |\lambda|^i & \sum_{i=0}^{k_D} d_i |\lambda|^i \end{bmatrix}.$$

5. Comparison with Tisseur's unstructured condition number. As we explain in Remark 4.2, Definition 4.1 is a structured absolute condition number for the eigenvalue computing function $f : P(\lambda) \mapsto \lambda_0$, where $P(\lambda)$ is a polynomial system matrix as in (2.5) for a rational transfer function matrix $R(\lambda)$. That is, due to the fact that we perturb each block $A(\lambda)$, $B(\lambda)$, $C(\lambda)$ and $D(\lambda)$ of $P(\lambda)$ and the allowed perturbations respect, separately, the degrees of each block: the approach of studying this set of structured perturbation is natural in the context of the representation of $R(\lambda)$ that we assume to be given. In contrast, the condition number for matrix polynomials defined by F. Tisseur in [22] considers perturbations that do not respect the degrees of particular blocks of the corresponding polynomial matrix $P(\lambda)$ but the overall degree of $P(\lambda)$. In other words, the more general theory developed in [22] corresponds to studying unstructured condition numbers, as opposed to our structured condition numbers.

In this section, we compare our structured condition number with the unstructured condition number for $P(\lambda)$ obtained from [22]. Let us first recall Tisseur's result. For that, we expand $P(\lambda)$ on the monomial basis, i.e., $P(\lambda) = P_d \lambda^d + \dots + P_1 \lambda + P_0$, where $d = \max\{k_A, k_B, k_C, k_D\}$. Let us now perturb each matrix coefficient P_i of $P(\lambda)$ with a perturbation $\epsilon \Delta P_i$ such that $\|\Delta P_i\|_2 \leq p_i$, then the condition number formula in [22, Theorem 5] for a simple eigenvalue λ_0 of $P(\lambda)$ is

$$(5.1) \quad \kappa_U(\lambda_0) = \frac{\|w\|_2 \|v\|_2 \left(\sum_{i=0}^d |\lambda_0|^i p_i \right)}{|w^T P'(\lambda_0) v|},$$

where w and v are right and left eigenvectors of $P(\lambda)$ associated with λ_0 , respectively.

REMARK 5.1. In [22, Theorem 5] the definition of condition number, and hence its formula, is scaled by dividing by $|\lambda_0|$. Dividing by $|\lambda_0|$ is more suitable if one wants to obtain a relative condition number, as done in [22]. However, we prefer to present our results for absolute condition numbers as this simplifies the exposition. Note that a relative condition number can always be obtained from an absolute one via (1.8). We do not divide by $|\lambda_0|$ in (5.1) so to obtain an absolute condition number, for the purpose of a fair comparison with Definition 4.1.

We now note that if each block polynomial matrix of $P(\lambda)$ is perturbed so that $\|\Delta D_i\|_2 \leq d_i$, $\|\Delta A_i\|_2 \leq a_i$, $\|\Delta B_i\|_2 \leq b_i$ and $\|\Delta C_i\|_2 \leq c_i$ then the sharp upper bound for $\|\Delta P_i\|_2$ is $\|\Delta P_i\|_2^2 \leq \max_{v^2+w^2=1} (a_i v + b_i w)^2 + (c_i v + d_i w)^2$. In other words, when we compare the structured and unstructured condition numbers, we should set

$$(5.2) \quad p_i := \left(\max_{v^2+w^2=1} (a_i v + b_i w)^2 + (c_i v + d_i w)^2 \right)^{1/2}$$

as this yields the smallest set of unstructured perturbations that includes the set of structured perturbations. Moreover, we can see that

$$\max\{a_i, b_i, c_i, d_i\}^2 \leq \max_{v^2+w^2=1} (a_i v + b_i w)^2 + (c_i v + d_i w)^2 \leq 4 \max\{a_i, b_i, c_i, d_i\}^2,$$

where the lower and upper bounds are reached for some values of a_i, b_i, c_i, d_i . A special case of particular interest is when $a_i = b_i = c_i = d_i = 1$; in this case, (5.2) yields $p_i = 2$. If we bound $\|\Delta P_i\|_2$ by p_i as given in (5.2), we then expect that the unstructured absolute condition number $\kappa_U(\lambda_0)$ in (5.1) is larger than the structured

absolute condition number $\kappa_S(\lambda_0)$ in Definition 4.1 or Theorem 4.4. This indeed follows by definition of supremum, because the set of allowed perturbations for κ_S is a subset of the set of allowed perturbations for κ_U .

THEOREM 5.2. *Let $P(\lambda)$ be a polynomial system matrix as in (2.5) with transfer function matrix $R(\lambda)$. Assume that $\lambda_0 \in \mathbb{C}$ is a simple zero of $R(\lambda)$ and that $P(\lambda)$ is minimal at λ_0 . Assume that $\|\Delta D_i\|_2 \leq d_i$, $\|\Delta A_i\|_2 \leq a_i$, $\|\Delta B_i\|_2 \leq b_i$, $\|\Delta C_i\|_2 \leq c_i$ and let p_i be defined as in 5.2. Then $\kappa_S(\lambda_0) \leq \kappa_U(\lambda_0)$.*

In the following example, we see that κ_U can be unboundedly larger than κ_S . Then, in some applications restricting to structured perturbations can make a very significant difference.

EXAMPLE 5.3. *Consider the rational matrix*

$$R(\lambda; \alpha, \beta, k) := \begin{bmatrix} \lambda - \alpha & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda^k & \beta \\ \beta & \lambda^k \end{bmatrix}^{-1}$$

for some parameters α, β and k such that $\alpha \neq 0$ and $k > 0$. We can set $A(\lambda; \beta, k) := \begin{bmatrix} \lambda^k & \beta \\ \beta & \lambda^k \end{bmatrix}$, $C(\lambda; \alpha) := \begin{bmatrix} \lambda - \alpha & 0 \\ 0 & 1 \end{bmatrix}$, $B := I_2$ and $D := 0_2$ such that

$$P(\lambda; \alpha, \beta, k) := \begin{bmatrix} -A(\lambda; \beta, k) & B \\ C(\lambda; \alpha) & D \end{bmatrix}$$

is a polynomial system matrix of $R(\lambda; \alpha, \beta, k)$. Note that $R(\lambda; \alpha, \beta, k)$ has a zero at α , which is also a zero of $P(\lambda; \alpha, \beta, k)$. A right eigenvector for P associated with α is $v = [1 \ 0 \ \alpha^k \ \beta]^T$ and an associated left eigenvector is $w = [0 \ 0 \ 1 \ 0]^T$. Then, we compute the condition numbers $\kappa_U(\alpha)$ and $\kappa_S(\alpha)$ given in (5.1) and Theorem 4.4, respectively. We remark that, although the expression for κ_S in Theorem 4.4 involves the inverse of A and right and left root vectors of R , κ_S can be computed by using right and left eigenvectors of P taking into account Proposition 2.5 and (3.5). For $\kappa_S(\alpha)$, we choose the weights $d_0 = 0$ and $a_i = b_i = c_i = 1$ for all $i \leq k_A, k_B, k_C$, respectively, and we obtain $\kappa_S(\alpha) = |\alpha| + 1$. For $\kappa_U(\alpha)$, we choose the weights p_i according to (5.2) for all $i \leq k$. This yields $p_0 = (1 + \sqrt{5})/2$, $p_1 = \sqrt{2}$ and $p_i = 1$ for $1 < i \leq k$ if $k > 1$. We obtain $\kappa_U(\alpha) = \sum_{i=0}^k p_i |\alpha|^i \sqrt{1 + |\alpha|^{2k} + |\beta|^2}$. Therefore, when $|\alpha| > 1$, $\kappa_U(\alpha) \rightarrow \infty$ as $k \rightarrow \infty$ or $\beta \rightarrow \infty$, while $\kappa_S(\alpha)$ keeps constant. Moreover, $\kappa_U(\alpha)/\kappa_S(\alpha) \rightarrow \infty$ as $\alpha \rightarrow \infty$. Finally, note that α is both a zero and a pole when $\beta = \pm \alpha^k$.

6. Numerical experiments. In this section, we explore the relationship between the unstructured and structured condition numbers κ_U and κ_S , respectively, via numerical experiments. We will show that, also in actual applications, κ_U may be much larger than κ_S . Moreover, the choice of the representation of the rational matrix can influence the ratio between κ_U and κ_S .

Experiment 1. We first consider a REP from the real-life application (1.3). One natural choice for the weights for κ_S is to choose $a_i = b_i = c_i = d_i = 1$ for all $i \leq k_A, k_B, k_C, k_D$, respectively, and otherwise the weights are set to 0. We choose the weights p_i for κ_U according to (5.2) for all $i \leq d$. More realistically, we might want to perturb only the actual data. In the representation given in (1.4), this corresponds to setting $c_0 = c_1 = b_0 = a_1 = 0$ and $d_0 = d_1 = b_1 = a_0 = 1$. If we use the representation in (1.4), the ratio of the condition numbers can be bounded above by 2 when using

the former way of choosing the weights. In our numerical experiments, the ratio is bounded by a relatively low number also in the case when only the data is perturbed.

It is also possible to represent (1.3) in different ways. For example, we can move one factor λ from $B(\lambda)$ to $C(\lambda)$, or, in other words, consider the representation

$$(6.1) \quad R(\lambda) = -K + \lambda M + [\lambda^2 I \cdots \lambda^2 I] \begin{bmatrix} (w_1 - \lambda)I & & \\ & \ddots & \\ & & (w_k - \lambda)I \end{bmatrix}^{-1} \begin{bmatrix} C_1 \\ \vdots \\ C_k \end{bmatrix}$$

rather than (1.4). The structured condition number κ_S is unchanged by this change of representation. However, the corresponding unstructured condition number κ_U is not. In the representation (6.1), perturbing only the real data corresponds to setting $c_0 = c_1 = c_2 = a_1 = 0$ and $d_0 = d_1 = b_0 = a_0 = 1$. This yields $p_0 = (1 + \sqrt{5})/2$, $p_1 = 1$, and $p_2 = 0$. Choosing $a_i = b_i = c_i = d_i = 1$ for all $i \leq k_A, k_B, k_C, k_D$ yields $p_0 = 2$, $p_1 = (1 + \sqrt{5})/2$, and $p_2 = 1$. For this representation, we compute the condition numbers κ_U and κ_S given in (5.1) and in Theorem 4.4, respectively, for the largest finite eigenvalue in magnitude. We set $K, M, C_i \in \mathbb{C}^{200 \times 200}$ and $k = 2$, where we generate the entries of the matrices as well as the poles ω_i randomly from a normal distribution of mean 0 and variance 1. We generate 10^3 realizations in the described way and compute the ratio for the condition numbers. The results are shown in Figure 6.1. We can see that, in both cases, the ratio can grow large although higher ratios are increasingly less probable.

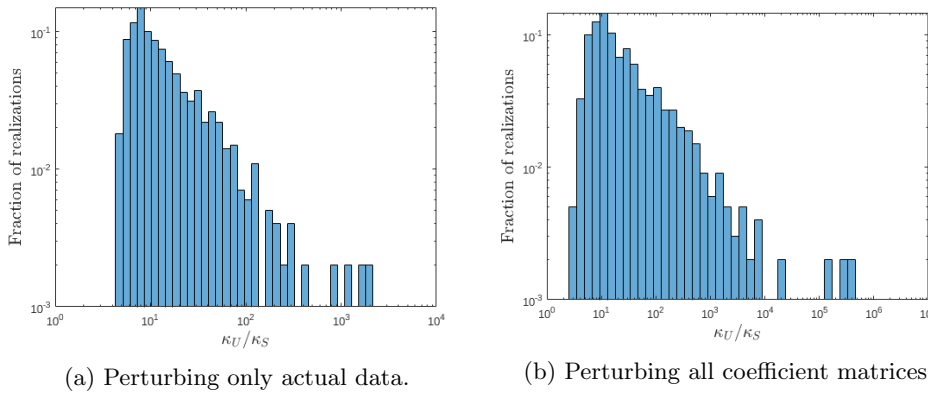


Fig. 6.1: For (6.1), the ratio κ_U/κ_S of the condition numbers for 10^3 realizations.

Experiment 2. Let us next consider the real-life application (1.5). In particular, we consider two different ways to represent this rational matrix. The first representation is (1.6) while the second representation is

$$(6.2) \quad R(\lambda) = A - \lambda B + k\lambda e_n [\lambda - k/m]^{-1} e_n^T.$$

For perturbations, we set $a_i = b_i = c_i = d_i = 1$ for all $i \leq k_A, k_B, k_C, k_D$, respectively, and otherwise we set the weights to 0. Again, we choose the weights p_i for κ_U according to (5.2). Let us fix $n = 10$ and $m = 1$, and let k vary. We compute the

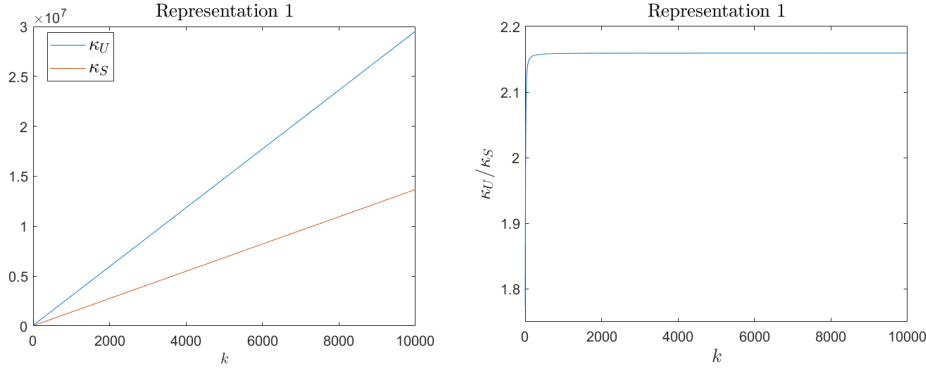


Fig. 6.2: For (1.6), the condition numbers κ_U and κ_S as well as their ratio as a function of the parameter k .

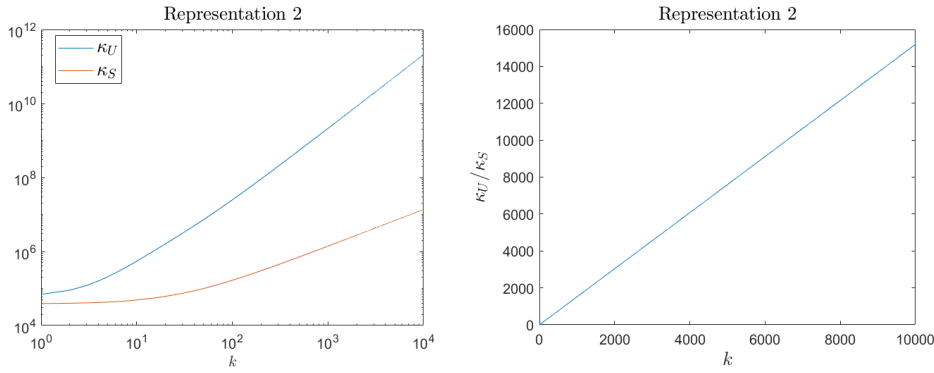


Fig. 6.3: For (6.2), the condition numbers κ_U and κ_S as well as their ratio as a function of the parameter k .

condition numbers κ_U and κ_S for the largest finite eigenvalue in magnitude for both representations. The results are in Figures 6.2 and 6.3.

In Figure 6.2, both κ_U and κ_S are linear in k . However, in Figure 6.3 the slope for κ_U in the loglog-plot is close to 2, which implies that κ_U is quadratic in k ; whereas the slope for κ_S is close to 1, which implies that κ_S is linear in k . As a result, the ratio κ_U/κ_S diverges as k approaches infinity. In addition, we have that the structured condition number κ_S is the same for the two representations (1.6) and (6.2), whereas the corresponding unstructured condition number κ_U does depend on this choice.

7. Conclusions and open problems. We defined a structured condition number κ_S for eigenvalues λ_0 of a (locally) minimal polynomial system matrix, which are the zeros of the corresponding transfer function $R(\lambda)$. We then derived a computable expression for κ_S , that is valid even in the case of λ_0 being also a pole of $R(\lambda)$. For that, we used the notion of root vectors. We finally compared our structured condition number κ_S with the standard condition number for polynomial matrices defined by F. Tisseur in [22], showing that κ_S is never larger and that may actually be much smaller in some applications, taking into account different representations of $R(\lambda)$.

An open problem is to study which representations of rational matrices have favourable properties with respect to conditioning. In addition, nowadays the standard approach to solve the REP is via linearizations. That is, transforming the REP into a generalized eigenvalue problem in such a way that the spectral information is preserved. It is also an open to study which linearizations have favourable properties with respect to conditioning. Another related problem is to derive a computable expression for local backward errors for pairs of approximate eigenvalues and eigenvectors of rational matrices.

Acknowledgements. VN and MQ thank Froilán Dopico for enlightening comments on the topic of this paper and for chupitos, both offered during a research visit at Universidad Carlos III de Madrid.

REFERENCES

- [1] A. L. Andrew, K. W. Eric Chu, P. Lancaster, Derivatives of eigenvalues and eigenvectors of matrix functions, *SIAM J. Matrix Anal. Appl.* 14 (4) (1993), 903–926.
- [2] T. Betcke, N. J. Higham, V. Mehrmann, C. Schroeder, F. Tisseur, NLEVP: A Collection of Nonlinear Eigenvalue Problems, *ACM Trans. Math. Soft.* 39(2) (2013), 7:1–7:28.
- [3] M. Dahleh, M. A. Dahleh, G. Verghese, Lectures on Dynamic Systems and Control, Massachusetts Institute of Technology, 2011. Book in progress. Lecture notes available at <https://ocw.mit.edu/courses/6-241j-dynamic-systems-and-control-spring-2011/>
- [4] F. M. Dopico, S. Marcaida, M. C. Quintana, Strong linearizations of rational matrices with polynomial part expressed in an orthogonal basis, *Linear Algebra Appl.* 570 (2019), 1–45.
- [5] F. M. Dopico, S. Marcaida, M. C. Quintana, P. Van Dooren, Local linearizations of rational matrices with application to rational approximations of nonlinear eigenvalue problems, *Linear Algebra Appl.* 604 (2020), 441–475.
- [6] F. M. Dopico, S. Marcaida, M. C. Quintana, P. Van Dooren, Block full rank linearizations of rational matrices, *Linear and Multilinear Algebra* (2022) DOI:10.1080/03081087.2022.2033151.
- [7] F. M. Dopico, V. Noferini, Root polynomials and their role in the theory of matrix polynomials, *Linear Algebra Appl.* 584 (2020), 37–78.
- [8] F. M. Dopico, M. C. Quintana, P. Van Dooren, Structural backward stability in rational eigenvalue problems solved via block Kronecker linearizations, submitted. Available as arXiv:2103.16395v1.
- [9] I. Gohberg, P. Lancaster, L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, PA, USA, 2009 (unabridged republication of book first published by Academic Press in 1982).
- [10] S. Güttel, L. Tisseur, The Nonlinear Eigenvalue Problem, *Acta Numer.* 26 (2017), 1–94.
- [11] S. Güttel, R. Van Beeumen, K. Meerbergen, W. Michiels, NLEIGS: A class of fully rational Krylov methods for nonlinear eigenvalue problems, *SIAM J. Sci. Comput.* 36(6) (2014), A2842–A2864.
- [12] N. J. Higham, *Accuracy and stability of numerical algorithms*, SIAM, Philadelphia (PA), 1996.
- [13] T. Kailath, *Linear Systems*, Prentice Hall, New Jersey, 1980.
- [14] T. Kato, *Perturbation Theory for Linear Operators*, Springer, 1980.
- [15] P. Lietaert, J. Pérez, B. Vandereycken, K. Meerbergen, Automatic rational approximation and linearization of nonlinear eigenvalue problems, *IMA J. Numer. Anal.* 42(2) (2022), 1087–1115.
- [16] B. McMillan, Introduction to formal realizability theory II, *Bell System Tech. J.* 31 (1952), 541–600.
- [17] V. Mehrmann, H. Voss, Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods, *GAMM Mitteilungen* 27(2) (2005), 121–152.
- [18] V. Noferini, P. Van Dooren, Root vectors of polynomial and rational matrices: theory and computation, submitted. Available as arXiv:2204.10955v1.
- [19] J. Pérez, M. C. Quintana, Linearizations of rational matrices from general representations, *Linear Algebra Appl.* 647 (2022), 89–126.
- [20] J. R. Rice, A theory of condition, *SIAM J. Numer. Anal.* 3 (1966), 287–310.
- [21] H. H. Rosenbrock, *State-space and Multivariable Theory*, Thomas Nelson and Sons, London, 1970.
- [22] F. Tisseur, Backward error and condition of polynomial eigenvalue problems, *Linear Algebra Appl.* 309 (2000), 339–361.
- [23] A. I. G. Vardulakis, *Linear Multivariable Control*, John Wiley and Sons, New York, 1991.