

**A Framework for Analyzing Nonlinear  
Eigenproblems and Parametrized Linear Systems**

Laurence Grammont, Nicholas J. Higham and  
Françoise Tisseur

July 2009

MIMS EPrint: **2009.51**

Manchester Institute for Mathematical Sciences  
School of Mathematics

The University of Manchester

Reports available from: <http://www.manchester.ac.uk/mims/eprints>

And by contacting: The MIMS Secretary  
School of Mathematics  
The University of Manchester  
Manchester, M13 9PL, UK

ISSN 1749-9097

# A Framework for Analyzing Nonlinear Eigenproblems and Parametrized Linear Systems<sup>☆</sup>

Laurence Grammont<sup>a</sup>, Nicholas J. Higham<sup>\*,b,1,2</sup>, Françoise Tisseur<sup>b,2</sup>

<sup>a</sup>*LaMUSE (Laboratoire de Mathématiques de L'Université de Saint-Etienne), Université de Saint-Etienne, 23 rue du Dr Paul Michelon, 42023 Saint-Etienne Cedex 2, France*

<sup>b</sup>*School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK*

---

## Abstract

Associated with an  $n \times n$  matrix polynomial of degree  $\ell$ ,  $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ , are the eigenvalue problem  $P(\lambda)x = 0$  and the linear system problem  $P(\omega)x = b$ , where in the latter case  $x$  is to be computed for many values of the parameter  $\omega$ . Both problems can be solved by conversion to an equivalent problem  $L(\lambda)z = 0$  or  $L(\omega)z = c$  that is linear in the parameter  $\lambda$  or  $\omega$ . This linearization process has received much attention in recent years for the eigenvalue problem, but it is less well understood for the linear system problem. We develop a framework in which more general versions of both problems can be analyzed, based on one-sided factorizations connecting a general nonlinear matrix function  $N(\lambda)$  to a simpler function  $M(\lambda)$ , typically a polynomial of degree 1 or 2. Our analysis relates the solutions of the original and linearized problems and in the linear system case indicates how to choose  $c$  and recover  $x$  from  $z$ . For the eigenvalue problem this framework includes many special cases studied in the literature, including the vector spaces of pencils  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$  recently introduced by Mackey, Mackey, Mehl, and Mehrmann and a class of rational problems. We use the framework to investigate the conditioning and stability of the parametrized linear system  $P(\omega)x = b$  and thereby study the effect of scaling, both of the original polynomial and of the pencil  $L$ . Our results identify situations in which scaling can potentially greatly improve the conditioning and stability and our numerical results show that dramatic improvements can be achieved in practice.

*Key words:* nonlinear eigenvalue problem, polynomial eigenvalue problem, rational eigenvalue problem, linearization, quadratization, parametrized linear system, backward error, scaling, companion form

*2000 MSC:* 15A18, 65F15

---

<sup>☆</sup>Version of July 13, 2009.

\*Corresponding author.

*Email addresses:* Laurence.Grammont@univ-st-etienne.fr (Laurence Grammont), higham@ma.man.ac.uk (Nicholas J. Higham), ftisseur@ma.man.ac.uk (Françoise Tisseur)

<sup>1</sup>The work of this author was supported by a Royal Society-Wolfson Research Merit Award.

<sup>2</sup>The work of these authors was supported by Engineering and Physical Sciences Research Council grant EP/D079403.

## 1. Introduction

Consider the matrix polynomial of degree  $\ell \geq 2$ ,

$$P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j, \quad A_j \in \mathbb{C}^{n \times n}, \quad A_\ell \neq 0. \quad (1.1)$$

Associated with  $P$  are two important problems with many practical applications.

- *Polynomial eigenvalue problem* (PEP): find scalars  $\lambda$  and nonzero vectors  $x$  and  $y$  satisfying

$$P(\lambda)x = 0, \quad y^* P(\lambda) = 0; \quad (1.2)$$

$x$  and  $y$  are right and left eigenvectors corresponding to the eigenvalue  $\lambda$ .

- *Parameter-dependent linear system*: find the solution  $x$  to

$$P(\omega)x = b, \quad b \in \mathbb{C}^n, \quad (1.3)$$

for many values of the scalar  $\omega$ , where  $\omega$  is usually either real or pure imaginary with  $|\omega| \in [\omega_l, \omega_h]$ ,  $\omega_l \ll \omega_h$ .

It is common in practice to reformulate these two problems into the two equivalent problems

$$\text{generalized eigenvalue problem:} \quad L(\lambda)z = 0, \quad w^* L(\lambda) = 0, \quad (1.4)$$

$$\text{augmented system:} \quad L(\omega)z = c, \quad (1.5)$$

where  $L(t) = tX + Y$  is now linear in the parameter  $t$ . In the case of (1.4) this allows standard numerical methods (e.g., the QZ algorithm or Krylov subspace methods) to be applied, whereas (1.5) opens up the possibility of employing various techniques that allow substantial savings when solving for many different  $\omega$  [21], [27], [29].

While the eigenvalue problem (1.2) and its linear equivalent are the subject of a large literature [8], [22], [32], the linear system (1.3) has received much less attention from mathematicians. The purpose of this paper is to show that the linear problems (1.4) and (1.5) can be studied in a common framework based on one-sided factorizations that relate  $P$  and  $L$ . In fact, our analysis is phrased in more general terms that make it applicable in a wide variety of situations: we replace  $P$  and  $L$  by arbitrary nonlinear matrix functions  $N$  and  $M$ , respectively, with just the restriction that  $M$  is of dimension at least as large as  $N$ . The generality of these conditions on  $N$  and  $M$  and the one-sided factorizations themselves means that the results we prove apply to many special cases, including linearization of matrix polynomials, the newer concept of quadratization of matrix polynomials [1], [20], and solution of rational eigenproblems via an appropriate form of linearization [30].

As an example, consider the quadratic  $Q(\omega) = \omega^2 A + \omega B + C$  and the associated first companion pencil

$$C_1(\omega) = \omega \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} B & C \\ -I & 0 \end{bmatrix}. \quad (1.6)$$

The solution to  $Q(\omega)x = b$  can be obtained from the solution to the augmented system  $C_1(\omega)z = c$ , where  $c = [b^T \ 0]^T$ . Indeed,  $z = [\omega x^T \ x^T]^T$ , so  $x$  can be recovered from the first  $n$  components of  $z$ , if  $\omega \neq 0$ , or the last  $n$  components. We will show how one-sided factorizations enable a systematic generalization of this example to different degrees and different pencils  $L$  and that the factorizations permit comparisons to be made between sensitivities and backward errors for the original and augmented systems.

The plan of the paper is as follows. In Section 2 we introduce the left- and right-sided factorizations on which our analysis is based and give a variety of examples of such factorizations. In Section 3 we explain the implications of the factorizations for the eigenvalue problem, thereby generalizing recent results in the literature. We turn to the linear system problem in Section 4, where we use the factorizations to derive relations between the solutions of the original system and the augmented system and also to obtain a formula for  $P(\omega)^{-1}$  in terms of  $L(\omega)^{-1}$ . In the rest of the paper we apply our results to linear systems  $P(\omega)x = b$ . Section 5 treats perturbation theory and compares the conditioning of the original and augmented systems, while Section 6 gives an analogous treatment of the backward error. In both cases, a block scaling of the companion forms and a scaling of the original  $P$  are found to be potentially very beneficial. Numerical experiments in Section 7 confirm the value of the analysis and Section 8 contains some concluding remarks.

Finally, we define the notation used throughout this paper. By “matrix function” we mean a rectangular matrix whose elements are a (generally nonlinear) function of a scalar indeterminate,  $\lambda$ . Matrix functions are designated as follows.

- $M(\lambda)$  and  $N(\lambda)$  are matrix functions of size  $r \times r$  and  $n \times n$ , respectively, with  $r \geq n$ .
- $P(\lambda)$  is an  $n \times n$  matrix polynomial of degree  $\ell$ , as in (1.1).
- $L(\lambda)$  is an  $r \times r$  linear matrix polynomial (matrix pencil).
- $Q(\lambda)$  is an  $n \times n$  quadratic matrix polynomial.
- $R(\lambda)$  is an  $n \times n$  rational matrix function of the form

$$R(\lambda) = P(\lambda) + \sum_{j=1}^k \frac{s_j(\lambda)}{q_j(\lambda)} R_j, \quad (1.7)$$

where  $s_j(\lambda)$  and  $q_j(\lambda)$  are scalar polynomials and  $R_j \in \mathbb{C}^{n \times n}$  for all  $j$ .

## 2. One-sided factorizations

Suppose that the  $r \times r$  and  $n \times n$  matrix functions  $M(\lambda)$  and  $N(\lambda)$  satisfy one or both of the one-sided factorizations of the form

$$\text{right-sided factorization} \quad M(\lambda)F(\lambda) = G(\lambda)N(\lambda), \quad (2.1)$$

$$\text{left-sided factorization} \quad E(\lambda)M(\lambda) = N(\lambda)H(\lambda), \quad (2.2)$$

where  $G(\lambda)$ ,  $H(\lambda)^T$ ,  $F(\lambda)$  and  $E(\lambda)^T$  are  $r \times n$  matrix functions. In the following two subsections we show that the conditions (2.1) and (2.2) cover a wide variety of situations

and provide a convenient framework for proving relations between the nonlinear eigenproblem or parametrized problem for  $N(\lambda)$  and the corresponding problem for  $M(\lambda)$ . We are particularly interested in the situation where  $N(\lambda)$  is a matrix polynomial or matrix rational function and  $M(\lambda)$  is a linear or quadratic matrix polynomial.

### 2.1. Matrix polynomials

In most practical applications,  $N$  is a matrix polynomial  $P$  of degree  $\ell$  as in (1.1) and  $M = L$  is a pencil. A *two-sided* factorization relating  $L$  and  $P$  arises in the definition of a linearization of a matrix polynomial. An  $\ell n \times \ell n$  pencil  $L(\lambda) = \lambda X + Y$  is a linearization of  $P$  if

$$E_L(\lambda)L(\lambda)F_L(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(\ell-1)n} \end{bmatrix} \quad (2.3)$$

for some unimodular  $\ell n \times \ell n$  matrix polynomials  $E_L(\lambda)$  and  $F_L(\lambda)$  [8, Sec. 7.2], where a unimodular polynomial is one with constant, nonzero determinant. This factorization implies that  $\alpha \det(L(\lambda)) = \det(P(\lambda))$  for some nonzero constant  $\alpha$ , so that  $L(\lambda)$  and  $P(\lambda)$  are (non)singular for precisely the same values of  $\lambda$ . It is easy to show that the two-sided factorization (2.3) implies the existence of the one-sided factorizations (2.1) and (2.2). Indeed if (2.3) holds then

$$L(\lambda)F(\lambda) \equiv L(\lambda) \cdot F_L(\lambda) \begin{bmatrix} I_n \\ 0 \end{bmatrix} = E_L(\lambda)^{-1} \begin{bmatrix} I_n \\ 0 \end{bmatrix} \cdot P(\lambda) \equiv G(\lambda)P(\lambda),$$

where  $G(\lambda)$  is a matrix polynomial since  $\det(E_L(\lambda))$  is a constant. Similarly,

$$E(\lambda)L(\lambda) \equiv [I_n \ 0] E_L(\lambda) \cdot L(\lambda) = P(\lambda) \cdot [I_n \ 0] F_L(\lambda)^{-1} \equiv P(\lambda)H(\lambda),$$

where  $H(\lambda)$  is a matrix polynomial. However, the one-sided factorizations (2.1) or (2.2) may hold without  $L$  being a linearization, as Examples 2.5 and 2.6 below show.

The idea of using one-sided factorizations such as (2.1) and (2.2) originates with Higham, Li, and Tisseur [13], who use the conditions (2.4) in their analysis of how backward errors for  $L$  relate to those for  $P$  in the eigenvalue problem.

While the notion of linearization is of great importance, the two-sided factorization (2.3) itself is of limited use because the matrix polynomials  $E_L(\lambda)$  and  $F_L(\lambda)$  are rarely known explicitly. An advantage of the one-sided factorizations (2.1) and (2.2) is that they are often explicitly known and of a simple form. Moreover, most of the results in this paper require (2.1) and (2.2) to hold at a single point only and do not require that  $L$  be a linearization.

The one-sided factorizations typically hold in the more specialized forms

$$L(\lambda)F(\lambda) = g \otimes P(\lambda), \quad g \in \mathbb{C}^m, \quad (2.4a)$$

$$E(\lambda)L(\lambda) = h^T \otimes P(\lambda), \quad h \in \mathbb{C}^m, \quad (2.4b)$$

where  $\otimes$  denotes the Kronecker product [24, Sec. 12.1]. These forms are special cases of (2.1) and (2.2), as can be seen from

$$g \otimes P(\lambda) = \begin{bmatrix} g_1 P(\lambda) \\ \vdots \\ g_m P(\lambda) \end{bmatrix} = \begin{bmatrix} g_1 I_n \\ \vdots \\ g_m I_n \end{bmatrix} P(\lambda) = (g \otimes I_n) P(\lambda) \equiv G(\lambda)P(\lambda), \quad (2.5)$$

$$h^T \otimes P(\lambda) = [h_1 P \ \dots \ h_m P] = P [h_1 I_n \ \dots \ h_m I_n] = P(h^T \otimes I_n) \equiv PH(\lambda). \quad (2.6)$$

In the rest of this section we show that the factorizations (2.1) and (2.2) hold as identities in  $\lambda$  for many pencils  $L(\lambda)$  that appear in the literature when solving  $P(\lambda)x = 0$  or  $P(\omega)x = b$ .

**Example 2.1 (companion forms)** Associated with  $P$  are two  $\ell n \times \ell n$  companion form pencils,  $C_1(\lambda) = \lambda X_1 + Y_1$  and  $C_2(\lambda) = \lambda X_2 + Y_2$ , called the first and second companion forms [24, Sec. 14.1], respectively, where

$$X_1 = X_2 = \text{diag}(A_\ell, I_n, \dots, I_n), \quad (2.7a)$$

$$Y_1 = \begin{bmatrix} A_{\ell-1} & A_{\ell-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -I_n & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} A_{\ell-1} & -I_n & \dots & 0 \\ A_{\ell-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \dots & 0 \end{bmatrix}. \quad (2.7b)$$

These two companion forms are always linearizations of  $P$ . Note that  $C_2(\lambda) = C_1(\lambda)^{\mathcal{B}}$ , where  $A^{\mathcal{B}}$  denotes the block transpose of  $A$ : if  $A = (A_{ij})$  is a block  $k \times \ell$  matrix with  $m \times n$  blocks  $A_{ij}$ , the block transpose of  $A$  is the block  $\ell \times k$  matrix  $A^{\mathcal{B}}$  with  $m \times n$  blocks defined by  $(A^{\mathcal{B}})_{ij} = A_{ji}$ . With the notation

$$A(\lambda) = A = [\lambda^{\ell-1}, \lambda^{\ell-2}, \dots, 1]^T, \quad (2.8)$$

it is easily checked that  $C_1$  satisfies a right-sided factorization of the form (2.4a) [26],

$$C_1(\lambda)(A \otimes I_n) = e_1 \otimes P(\lambda). \quad (2.9)$$

Block transposing this equation yields a left-sided factorization of the form (2.4b) for  $C_2$ ,

$$(A^T \otimes I_n)C_2(\lambda) = e_1^T \otimes P(\lambda). \quad (2.10)$$

Moreover,  $C_1$  also satisfies a left-sided factorization (2.4b). For  $\ell = 3$ , for example, we have

$$\begin{bmatrix} \lambda^2 I_n & -(A_0 + \lambda A_1) & -\lambda A_0 \\ \lambda I_n & \lambda A_2 + \lambda^2 A_3 & -A_0 \\ I_n & A_2 + \lambda A_3 & A_1 + \lambda A_2 + \lambda^2 A_3 \end{bmatrix} C_1(\lambda) = I_3 \otimes P(\lambda), \quad (2.11)$$

whose block rows yields three different such factorizations. The corresponding relation for general  $\ell$ , and for  $C_2$ , is given in Lemma 5.4 below.

**Example 2.2 (vector spaces  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$ )** Two important vector spaces of pencils that generalize the first and second companion forms have been studied by Mackey, Mackey, Mehl, and Mehrmann [26]. These vector spaces are defined by

$$\mathbb{L}_1(P) = \{ L(\lambda) : L(\lambda)(A \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^\ell \}, \quad (2.12)$$

$$\mathbb{L}_2(P) = \{ L(\lambda) : (A^T \otimes I_n)L(\lambda) = \tilde{v}^T \otimes P(\lambda), \tilde{v} \in \mathbb{C}^\ell \}, \quad (2.13)$$

with  $A$  as in (2.8). From (2.9) and (2.10) we have that  $C_1 \in \mathbb{L}_1(P)$  and  $C_2 \in \mathbb{L}_2(P)$ . Almost all pencils in  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$  are linearizations of  $P$  [26, Prop. 3.2, Prop. 3.12, Thm. 4.7] and if  $L(\lambda) \in \mathbb{L}_1(P)$  with vector  $v$  then  $L(\lambda)^{\mathcal{B}} \in \mathbb{L}_2(P)$  with vector  $\tilde{v} = v$  [14, Thm. 2.2]. From the definition of these spaces we have that (2.4a) and (2.4b) hold for all pencils in  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$ , respectively.

**Example 2.3 (linearizations of Antoniou and Vologiannidis)** With the notation

$$M_0 = \text{diag}(-I_{n(\ell-1)}, A_0), \quad M_\ell = \text{diag}(A_\ell, I_{(\ell-1)n}),$$

$$M_j = \text{diag}\left(I_{n(j-1)}, \begin{bmatrix} -A_{\ell-j} & I_n \\ I_n & 0 \end{bmatrix}, I_{(\ell-j-1)n}\right), \quad j = 1: \ell - 1$$

and any four ordered sets of indices  $\mathcal{I}_k = \{i_{k,1}, i_{k,2}, \dots, i_{k,n_k}\}$ ,  $k = 1:4$ , such that  $\mathcal{I}_i \cap \mathcal{I}_j = \emptyset$  for  $i \neq j$  and  $\cup_{k=1}^4 \mathcal{I}_k = \{1, 2, \dots, \ell - 1\}$ , Antoniou and Vologiannidis [3] show that the matrix pencil

$$L(\lambda) = \lambda M_{\mathcal{I}_1}^{-1} M_\ell M_{\mathcal{I}_2}^{-1} + M_{\mathcal{I}_3} M_0 M_{\mathcal{I}_4} \quad (2.14)$$

is a linearization of  $P$ , where  $M_{\mathcal{I}_k} = M_{i_{k,1}} M_{i_{k,2}} \dots M_{i_{k,n_k}}$  for  $\mathcal{I}_k \neq \emptyset$ . The first and second companion forms of  $P$  are included as special cases:

$$\mathcal{I}_i = \emptyset, \quad i = 1, 2, 4, \quad \mathcal{I}_3 = \{\ell - 1, \dots, 1\}, \quad L(\lambda) = \lambda M_\ell - M_{\ell-1} \dots M_1 M_0 \equiv C_1(\lambda),$$

$$\mathcal{I}_i = \emptyset, \quad i = 1, 2, 3, \quad \mathcal{I}_4 = \{1, \dots, \ell - 1\}, \quad L(\lambda) = \lambda M_\ell - M_0 M_1 \dots M_{\ell-1} \equiv C_2(\lambda),$$

For quadratics  $Q(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ , (2.14) yields four different linearizations that belong to either  $\mathbb{L}_1(Q)$  or  $\mathbb{L}_2(Q)$ , and hence have factorizations (2.4a) or (2.4b), respectively:

$$L_1(\lambda) = \lambda M_1^{-1} M_2 + M_0 = \lambda \begin{bmatrix} 0 & I \\ A_2 & A_1 \end{bmatrix} + \begin{bmatrix} -I & 0 \\ 0 & A_0 \end{bmatrix} \in \mathbb{L}_1(Q) \text{ with vector } v = e_2,$$

$$L_2(\lambda) = \lambda M_2 M_1^{-1} + M_0 = L_1(\lambda)^{\mathcal{B}} \in \mathbb{L}_2(Q) \text{ with vector } v = e_2,$$

$$L_3(\lambda) = \lambda M_2 + M_1 M_0 \equiv C_1(\lambda), \quad L_4(\lambda) = \lambda M_2 + M_0 M_1 \equiv C_2(\lambda).$$

For  $\ell > 2$ , the linearizations in (2.14) do not all belong to  $\mathbb{L}_1(Q)$  or  $\mathbb{L}_2(Q)$ .

**Example 2.4 (linearizations of Amiraslani, Corless, and Lancaster)** Consider a sequence of polynomials  $\{\phi_j(\lambda)\}_{j=0}^\infty$  with  $\phi_0(\lambda) \equiv 1$  and  $\phi_j(\lambda)$  of degree  $j$  satisfying a three-term recurrence relation and rewrite the  $n \times n$  matrix polynomial  $P(\lambda)$  of degree  $\ell$  as

$$P(\lambda) = \phi_\ell(\lambda) B_\ell + \dots + \phi_1(\lambda) B_1 + \phi_0(\lambda) B_0. \quad (2.15)$$

Amiraslani, Corless, and Lancaster [2] construct pencils  $L(\lambda) = \lambda X + Y$  that are defined in terms of the  $B_i$  and the coefficients of the recurrence and are linearizations of  $P$ . They satisfy (2.4b) with

$$E(\lambda) = \Phi^T(\lambda) \otimes I, \quad h^T = \alpha_{\ell-1} e_\ell^T, \quad (2.16)$$

where  $\Phi(\lambda)^T = [\phi_0(\lambda), \phi_1(\lambda), \dots, \phi_{\ell-1}(\lambda)]$  and  $\alpha_{\ell-1} \neq 0$  is the leading coefficient of  $\phi_{\ell-1}(x)$ .

**Example 2.5 (factorization of  $P$ )** One approach to our two problems (1.2) and (1.3) is to try to factorize the matrix polynomial  $P$ . Consider the quadratic case ( $\ell = 2$ ). If a solvent can be found, that is, a matrix  $S \in \mathbb{C}^{n \times n}$  such that  $P(S) = A_2 S^2 + A_1 S + A_0 = 0$ , then  $P(\lambda) = (\lambda A_2 + A_1 + A_2 S)(\lambda I - S)$  and clearly both (2.1) and (2.2) hold with  $r = n$  and  $G(\lambda) = H(\lambda) \equiv I_n$ . This factorization approach is the basis of

some numerical methods for solving the PEP (1.2): they compute a solvent and thereby reduce the problem to solving one standard eigenvalue problem and one generalized eigenvalue problem [10], [12]. Note that this example is rather different from the others: neither factor is a linearization, as it does not have the correct dimensions to satisfy (2.3). Moreover, unlike in the examples above,  $E(\lambda)$  and  $F(\lambda)$  here are rank deficient for certain  $\lambda$ , namely half of the eigenvalues of  $P$ .

**Example 2.6 (quadratization of  $P$ )** A quadratization of a matrix polynomial  $P$  of even degree  $\ell = 2d > 2$  is a quadratic matrix polynomial  $Q$  that is unimodularly equivalent to  $\begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix}$  for an appropriately sized identity matrix  $I$ . Quadratizations are of particular interest for structured polynomials  $P$  when a correspondingly structured  $Q$  can be found and efficient numerical methods are available for  $Q$  (see, for example, [9], [19]), or when there is no structured linearization in the class of interest (for examples of which see [25]). For palindromic matrix polynomials of even degree, Huang, Lin, and Su [20] show how to build palindromic quadratizations that satisfy one-sided factorizations of the form (2.4a) and (2.4b). As an example, the  $*$ -palindromic quartic polynomial  $P(\lambda) = \lambda^4 A_2 + \lambda^3 A_1 + \lambda^2 A_0 + \lambda A_1^* + A_2^*$  with  $A_0^* = A_0$  and  $A_2$  nonsingular can be quadratized into

$$Q(\lambda) = \lambda^2 \begin{bmatrix} A_1 & A_2 \\ I & 0 \end{bmatrix} + \lambda \begin{bmatrix} A_0 - I - A_2 A_2^* & 0 \\ 0 & -I \end{bmatrix} + \begin{bmatrix} A_1^* & I \\ A_2^* & 0 \end{bmatrix},$$

which satisfies

$$Q(\lambda) \begin{bmatrix} \lambda I \\ \lambda^2 I + A_2^* \end{bmatrix} = e_1 \otimes P(\lambda), \quad [\lambda I \quad I + \lambda^2 A_2] Q(\lambda) = e_1^T \otimes P(\lambda).$$

## 2.2. Matrix rational functions

Rational eigenproblems  $R(\lambda)x = 0$ , where the  $n \times n$  rational function  $R(\lambda)$  has the form (1.7) occur in a variety of physical applications [6]. The matrices  $R_j$  are usually of low rank. Using the process of minimal realization [4, pp. 91–98] together with rank-revealing factorizations of the  $R_j$ , Su and Bai [30] show how to rewrite  $R(\lambda)$  as

$$R(\lambda) = P(\lambda) + U(C - \lambda D)^{-1} V^*,$$

where  $U$  and  $V$  are  $n \times m$  and  $C, D$  are  $m \times m$  and the value of  $m$  depends on the degree of the polynomials  $q_j(\lambda)$  and the rank of the matrices  $R_j$ . Now take any linearization  $L_1(\lambda) = \lambda X_1 + Y_1 \in \mathbb{L}_1(P)$  with  $v = e_1$  and premultiply  $R(\lambda)x = 0$  by  $e_1$  in the Kronecker sense to obtain

$$(L_1(\lambda) + (e_1 \otimes U)(C - \lambda D)^{-1}(e_\ell^T \otimes V^*))(A \otimes x) = 0,$$

which becomes a linear eigenvalue problem  $L(\lambda)z = 0$  with

$$L(\lambda) = \lambda \begin{bmatrix} X_1 & 0 \\ 0 & D \end{bmatrix} + \begin{bmatrix} Y_1 & e_1 \otimes U \\ e_\ell^T \otimes V^* & -C \end{bmatrix}, \quad z = \begin{bmatrix} A \otimes x \\ (C - \lambda D)^{-1} V^* x \end{bmatrix}.$$

It is then easy to check that (2.1) holds with

$$M(\lambda) = L(\lambda), \quad F(\lambda) = \begin{bmatrix} A \otimes I_n \\ (C - \lambda D)^{-1} V^* \end{bmatrix}, \quad G(\lambda) = \begin{bmatrix} e_1 \otimes I_n \\ 0 \end{bmatrix}, \quad N(\lambda) = R(\lambda).$$

### 3. Eigenvalue and eigenvector relations

We first determine what the factorizations (2.1) and (2.2) imply about the relation between the eigensystem of  $M$  and that of  $N$ . While the spectra are not necessarily identical, there are nevertheless close connections between eigenpairs of  $M$  and eigenpairs of  $N$ , as the next result shows.

**Theorem 3.1** *Let  $M(\lambda)$  and  $N(\lambda)$  be matrix functions of dimensions  $r \times r$  and  $n \times n$ , respectively, with  $r \geq n$ .*

(a) *Let  $\lambda \in \mathbb{C}$  and assume that (2.1) holds at  $\lambda$  with  $F(\lambda)$  and  $G(\lambda)$  of full rank. Then*

(i)  *$F(\lambda)x$  is a right eigenvector of  $M$  with eigenvalue  $\lambda$  if and only if  $x$  is a right eigenvector of  $N$  with eigenvalue  $\lambda$ .*

(ii) *If  $w \in \mathbb{C}^r$  is a left eigenvector of  $M$  with eigenvalue  $\lambda$  then  $G(\lambda)^*w$  is a left eigenvector of  $N$  with eigenvalue  $\lambda$  provided that it is nonzero.*

(b) *Let  $\lambda \in \mathbb{C}$  and assume that (2.2) holds at  $\lambda$  with  $E(\lambda)$  and  $H(\lambda)$  of full rank. Then*

(i) *If  $z \in \mathbb{C}^r$  is a right eigenvector of  $M$  with eigenvalue  $\lambda$  then  $H(\lambda)z$  is a right eigenvector of  $N$  with eigenvalue  $\lambda$  provided that it is nonzero.*

(ii)  *$E(\lambda)^*y$  is a left eigenvector of  $M$  with eigenvalue  $\lambda$  if and only if  $y$  is a left eigenvector of  $N$  with eigenvalue  $\lambda$ .*

**Proof.** The four parts follow, respectively, from the relations

$$M(\lambda)F(\lambda)x = G(\lambda)N(\lambda)x, \quad w^*M(\lambda)F(\lambda) = w^*G(\lambda)N(\lambda), \quad (3.1a)$$

$$E(\lambda)M(\lambda)z = N(\lambda)H(\lambda)z, \quad y^*E(\lambda)M(\lambda) = y^*N(\lambda)H(\lambda). \quad \square \quad (3.1b)$$

It is instructive to apply Theorem 3.1 to some of the examples given in the previous section, in particular when  $M(\lambda) = P(\lambda)$  is a matrix polynomial and  $M(\lambda) = L(\lambda)$  is a pencil. For  $L \in \mathbb{L}_1(P)$  and  $L \in \mathbb{L}_2(P)$  in Example 2.2 we have  $F(\lambda) = \Lambda \otimes I_n$ ,  $G(\lambda) = v \otimes I$ , and  $E(\lambda) = \Lambda^T \otimes I_n$ ,  $H(\lambda) = \tilde{v}^T \otimes I$ , respectively, which are of full rank for all  $\lambda$  assuming  $v$  and  $\tilde{v}$  are nonzero, and by exploiting this special structure the conclusions of the theorem can be strengthened in two respects. First, the phrase “provided that it is nonzero” can be removed in parts (a)(ii) and (b)(i) under the assumption that  $P$  is regular (that is,  $\det(P(\lambda)) \not\equiv 0$ ). Second, under the assumption that  $L$  is a linearization of  $P$ , every right eigenvector of  $L$  and left eigenvector of  $P$  can be shown to be of the forms given in parts (a)(i) and (a)(ii), and similarly for (b)(i) and (b)(ii). For proofs, see [26, Thms. 3.8, 3.14, 4.4] [13, Thms. 3.2, 3.3]. It is worth noting that in the case of  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$ , the eigenvector relations in parts (a)(ii) and (b)(i), respectively, of Theorem 3.1 are not found in [26], but were first identified in [13]. The systematic use of one-sided factorizations makes it easier to identify such relations in their full generality.

For the pencils in Example 2.4 arising from the basis of polynomials satisfying a three-term recurrence,  $E(\lambda) = \Phi^T(\lambda) \otimes I$  has full rank for all  $\lambda$ . Moreover, since  $h^T = \alpha_{\ell-1}e_\ell^T$ , the vector  $H(\lambda)z = (h^T \otimes I_n)z$  in part (b)(i) of Theorem 3.1 is always nonzero. Moreover,

it is shown in [2] that these pencils are strong linearizations and hence from the arguments in [13, Thm. 3.3] and [26, Thm. 3.14, 4.4] it follows that every right eigenvector of  $P$  and left eigenvector of  $L$  is of the forms given in parts (b)(i) and (b)(ii). The same results apply to the pencils for the Bernstein basis on  $[a, b]$  when  $\lambda \neq b$  [2, Sec. 4].

When (2.1) and (2.2) correspond to factorization of  $P$  (Example 2.5), the vectors  $G(\lambda)^*w$  and  $H(\lambda)z$  in parts (a)(ii) and (b)(i) are just  $w$  and  $z$ , and so are automatically nonzero. Since in a factorization  $P(\lambda) = L(\lambda)F(\lambda)$  the factors  $L$  and  $F$  may have eigenvalues in common, a stronger result is obtained (in part (a)(i), for example) by replacing the assumption that  $F(\lambda)$  is of full rank by the assumption that  $x$  is not in the null space of  $F(\lambda)$ .

Higham, Mackey, and Tisseur [15] investigate the conditioning of linearizations from the vector space  $\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P)$ . The analysis in that paper can be generalized by using the conditions (2.1) and (2.2) in place of the conditions defining  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$ . To indicate the key idea, let  $x$  and  $y$  denote right and left eigenvectors of  $P$  and let  $z$  and  $w$  denote right and left eigenvectors of  $L$ , all corresponding to a simple, nonzero, finite eigenvalue  $\lambda$ . Eigenvalue condition numbers are given, in the 2-norm, by the following expressions:

$$\kappa_P(\lambda) = \frac{(\sum_{j=0}^{\ell} |\lambda|^j \|A_j\|_2) \|y\|_2 \|x\|_2}{|\lambda| |y^* P'(\lambda) x|}, \quad \kappa_L(\lambda) = \frac{(|\lambda| \|X\|_2 + \|Y\|_2) \|w\|_2 \|z\|_2}{|\lambda| |w^* L'(\lambda) z|}. \quad (3.2)$$

These condition numbers measure the sensitivity of the eigenvalue  $\lambda$  of  $P$  and  $L$ , respectively to small perturbations of  $P$  and  $L$  measured in a normwise relative fashion [31, Thm. 5]. Ideally when solving (1.2) via (1.4) we would like  $\kappa_L(\lambda) \approx \kappa_P(\lambda)$ . The following lemma shows that our factorizations (2.1) and (2.2) imply a close relation between these condition numbers.

**Lemma 3.2** *Let  $\lambda$  and  $x$  be a finite eigenvalue and corresponding right eigenvector of  $P$  and let  $w$  be a left eigenvector of  $L$  corresponding to  $\lambda$ . Assume that for  $M(\lambda) \equiv L(\lambda)$  and  $N(\lambda) \equiv P(\lambda)$ , (2.1) holds with  $F(\lambda)$  of full rank and  $y = G(\lambda)^*w \neq 0$ . Then  $z = F(\lambda)x$  is a right eigenvector of  $L$ ,  $y$  is a left eigenvector of  $P$ , and*

$$w^* L'(\lambda) z = y^* P'(\lambda) x.$$

**Proof.** By Theorem 3.1 (a)(i),  $z = F(\lambda)x$  is a right eigenvector of  $L$  and by Theorem 3.1 (a)(ii)  $y$  is a left eigenvector of  $P$ . Differentiating  $L(\lambda)F(\lambda) = G(\lambda)P(\lambda)$  with respect to  $\lambda$  gives

$$L'(\lambda)F(\lambda) + L(\lambda)F'(\lambda) = G'(\lambda)P(\lambda) + G(\lambda)P'(\lambda).$$

Evaluating at  $\lambda$ , premultiplying by  $w^*$ , and postmultiplying by  $x$  gives

$$w^* L'(\lambda) z = w^* L'(\lambda) F(\lambda) x = w^* G(\lambda) P'(\lambda) x = y^* P'(\lambda) x. \quad \square$$

An entirely analogous result holds for (2.2).

From Lemma 3.2 it follows that for a simple, finite eigenvalue  $\lambda$ ,

$$\frac{\kappa_L(\lambda)}{\kappa_P(\lambda)} = \frac{|\lambda| \|X\|_2 + \|Y\|_2}{\sum_{j=0}^{\ell} |\lambda|^j \|A_j\|_2} \cdot \frac{\|w\|_2 \|z\|_2}{\|y\|_2 \|x\|_2}.$$

This expression can now be used to investigate the size of  $\kappa_L(\lambda)/\kappa_P(\lambda)$  as  $L$  varies, for fixed  $P$ , where the  $L$ -dependent terms are  $X$ ,  $Y$ ,  $w$ , and  $z$ . This is done in [15] for pencils  $L \in \mathbb{DL}(P)$ , where minimization of the ratio over  $L$  is considered. The same can be done for the other special cases described in Section 2, but we will not pursue this here.

In the rest of the paper we concentrate on linear systems, which are much less well studied than the polynomial eigenvalue problem.

#### 4. Parametrized linear system relations

Now we turn to linear systems and show that by using the factorization (2.1) or (2.2) we can identify an augmented system  $M(\omega)z = c$  whose solution is related in a well-defined way to that of the original system  $N(\omega)x = b$ .

**Theorem 4.1** *Let  $\omega \in \mathbb{C}$ ,  $r \geq n$ , and let the  $r \times r$  matrix  $M(\omega)$  and  $n \times n$  matrix  $N(\omega)$  be nonsingular.*

(a) *Assume that (2.1) holds at  $\omega$  with  $G(\omega)$  of full rank. Then  $x$  is the unique solution to  $N(\omega)x = b$  if and only if  $z = F(\omega)x$  is the unique solution to  $M(\omega)z = G(\omega)b$ .*

(b) *Assume that (2.2) holds at  $\omega$  with  $E(\omega)$  of full rank. If  $z$  is the unique solution to  $M(\omega)z = c$  for some right hand side  $c$  satisfying  $E(\omega)c = \gamma b$ , where  $0 \neq \gamma \in \mathbb{C}$ , then  $x = \gamma^{-1}H(\omega)z$  is the unique solution to  $N(\omega)x = b$ .*

**Proof.** (a) Let  $x$  be the unique solution to  $N(\omega)x = b$  and  $z$  the unique solution to  $M(\omega)z = G(\omega)b$ . Then multiplying (2.1) on the right by  $x$  gives

$$M(\omega)F(\omega)x = G(\omega)N(\omega)x = G(\omega)b. \quad (4.1)$$

It follows that  $z = F(\omega)x$ .

(b) Let  $z$  be the unique solution to  $M(\omega)z = c$ . Using (2.2) we have

$$N(\omega)H(\omega)z = E(\omega)M(\omega)z = E(\omega)c = \gamma b.$$

It follows that  $x = \gamma^{-1}H(\omega)z$ .  $\square$

Theorem 4.1 shows that if the right-sided factorization (2.1) holds then the right-hand side  $c$  of a suitable augmented system  $M(\omega)z = c$  is easy to construct. However, the solution  $x$  may not be easy to recover from  $z = F(\omega)x$  unless  $F(\omega)$  has a simple form. For the left-sided factorization (2.2) it is the right-hand side that is harder to construct, but recovery of  $x$  is trivial.

Note that if  $L$  is a linearization of a matrix polynomial  $P$  then  $P(\omega)$  nonsingular implies  $L(\omega)$  nonsingular, by (2.3).

Now we examine how Theorem 4.1 specializes for the vector spaces of  $\ell n \times \ell n$  pencils  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$ . Let  $\Lambda = \Lambda(\omega) = [\omega^{\ell-1} \ \omega^{\ell-2} \ \dots \ 1]^T$ . For  $z \in \mathbb{C}^{\ell n}$  we write  $z_j \equiv z((j-1)n+1:jn)$ ,  $j = 1:\ell$ .

**Corollary 4.2** *Let  $\omega \in \mathbb{C}$  and let  $P(\omega)$  and the  $\ell n \times \ell n$  matrix  $L(\omega)$  be nonsingular.*

(a) If  $L \in \mathbb{L}_1(P)$  with vector  $v \in \mathbb{C}^\ell$  then the unique solutions of  $P(\omega)x = b$  and  $L(\omega)z = v \otimes b$  are related by  $z = \Lambda \otimes x$ .

(b) If  $L \in \mathbb{L}_2(P)$  with vector  $\tilde{v} \in \mathbb{C}^\ell$ , and if  $u \in \mathbb{C}^\ell$  is such that  $\Lambda^T u \neq 0$ , then the unique solutions of  $P(\omega)x = b$  and  $L(\omega)z = u \otimes b$  are related by  $x = (\Lambda^T u)^{-1} \sum_{j=1}^\ell \tilde{v}_j z_j$ .

**Proof.** The proof is straightforward, on noting in (b) that we can take  $c = u \otimes b$  in Theorem 4.1 (b).  $\square$

Corollary 4.2 shows that for  $L \in \mathbb{L}_1(P)$ ,  $x$  can be recovered from the solution  $z = \Lambda \otimes x$  of the augmented system in many ways. Although the vector  $z$  for  $L \in \mathbb{L}_2(P)$  with right-hand side of the form  $u \otimes b$  does not have special structure in general, we do have freedom in the choice of  $u$ .

Note that  $L \in \mathbb{L}_1(P)$  with vector  $v$  implies that

$$L(\omega)(\Lambda \otimes I_n) = v \otimes P(\omega) \iff (\Lambda \otimes I_n)P(\omega)^{-1} = L(\omega)^{-1}(v \otimes I_n). \quad (4.2)$$

Multiplying (4.2) on the left by  $f^* \otimes I$  with  $f \in \mathbb{C}^\ell$  such that  $\Lambda^* f \neq 0$  leads to an interesting formula for  $P(\omega)^{-1}$  that will be used in Section 5. An analogous formula is obtained in a similar way for  $\mathbb{L}_2$ .

**Lemma 4.3** (a) Let  $L(\omega) \in \mathbb{L}_1(P)$  with vector  $v$ . For any  $f \in \mathbb{C}^\ell$  such that  $f^* \Lambda \neq 0$ ,

$$P(\omega)^{-1} = \frac{1}{f^* \Lambda} (f^* \otimes I_n) L(\omega)^{-1} (v \otimes I_n). \quad (4.3)$$

(b) Let  $L(\omega) \in \mathbb{L}_2(P)$  with vector  $\tilde{v}$ . For any  $f \in \mathbb{C}^\ell$  such that  $f^* \Lambda \neq 0$ ,

$$P(\omega)^{-1} = \frac{1}{f^* \Lambda} (\tilde{v}^T \otimes I_n) L(\omega)^{-1} (\bar{f} \otimes I_n).$$

In the following two sections we use these results to compare the sensitivity of the augmented system with that of the original system and to understand how backward errors for the augmented system propagate into backward errors for the original system. The focus of our analysis will be on the companion form pencil  $C_1$  and understanding the effects of scaling  $C_1$  or  $P$ , but our analysis could equally well be used to guide the choice of pencil  $L$  when, as is the case for the spaces  $\mathbb{L}_1(P)$  and  $\mathbb{L}_2(P)$  for example, there is a parametrized family of possible choices.

## 5. Sensitivity of parametrized linear systems

We denote by  $\|\cdot\|$  any vector norm and the corresponding subordinate matrix norm. The norm  $\|\cdot\|_D$  dual to a given vector norm  $\|\cdot\|$  on  $\mathbb{C}^n$  is defined by

$$\|x\|_D = \max_{z \neq 0} \frac{|z^* x|}{\|z\|}$$

and we say that  $z$  is a vector dual to  $y$  if  $z^* y = \|z\|_D \|y\| = 1$ . For complex  $\alpha$  we define

$$\text{sign}(\alpha) = \begin{cases} \bar{\alpha}/|\alpha|, & \alpha \neq 0, \\ 0, & \alpha = 0. \end{cases}$$

We assume throughout this section that  $P(\omega) = \sum_{i=0}^{\ell} \omega^i A_i$  is nonsingular, that is, the parameter  $\omega$  is not an eigenvalue of  $P$ . For notational convenience we define  $\Delta P(\omega) = \sum_{i=0}^{\ell} \omega^i \Delta A_i$ .

A normwise condition number of the solution  $x$  to  $P(\omega)x = b$  can be defined by

$$\kappa_{P,b}(\omega, x) := \limsup_{\epsilon \rightarrow 0} \left\{ \frac{\|\Delta x\|}{\epsilon \|x\|} : \left( (P(\omega) + \Delta P(\omega))(x + \Delta x) = b + \Delta b, \right. \right. \\ \left. \left. \|\Delta A_i\| \leq \epsilon \alpha_i, \quad i = 0: \ell, \quad \|\Delta b\| \leq \epsilon \beta \right) \right\}. \quad (5.1)$$

The  $\alpha_i$ ,  $i = 0: \ell$  and  $\beta$  are nonnegative weights, included to allow flexibility in how the perturbations are measured; in particular,  $\Delta A_i$  can be forced to zero by setting  $\alpha_i = 0$ . The normwise relative measure of the perturbations corresponds to

$$\alpha_i = \|A_i\|, \quad i = 0: \ell, \quad \beta = \|b\|. \quad (5.2)$$

The following result provides a perturbation bound.

**Theorem 5.1** *Let  $P(\omega)x = b$  and  $(P(\omega) + \Delta P(\omega))y = b + \Delta b$ , where  $\|\Delta A_i\| \leq \epsilon \alpha_i$ ,  $i = 0: \ell$  and  $\|\Delta b\| \leq \epsilon \beta$ . Assume that  $\epsilon \left( \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right) \|P(\omega)^{-1}\| < 1$ . Then*

$$\frac{\|x - y\|}{\|x\|} \leq \frac{\epsilon \|P(\omega)^{-1}\| \left( \beta / \|x\| + \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right)}{1 - \epsilon \|P(\omega)^{-1}\| \sum_{i=0}^{\ell} |\omega|^i \alpha_i}, \quad (5.3)$$

and this bound is attainable to first order in  $\epsilon$ .

**Proof.** It is straightforward to obtain  $y - x = P(\omega)^{-1}(\Delta b - \Delta P(\omega)x) + P(\omega)^{-1}\Delta P(\omega)(x - y)$ . Taking norms yields  $\|y - x\| \leq \epsilon \|P(\omega)^{-1}\| \left( \beta + \sum_{i=0}^{\ell} |\omega|^i \alpha_i (\|x\| + \|y - x\|) \right)$ , which yields the bound on rearranging. It is straightforward to show that the bound is attained to first order for the perturbations

$$\Delta A_i = -\epsilon \text{sign}(\omega^i) \alpha_i \|x\| z v^*, \quad i = 0: \ell, \quad \Delta b = \epsilon \beta z, \quad (5.4)$$

where  $\|z\| = 1$ ,  $\|P(\omega)^{-1}z\| = \|P(\omega)^{-1}\|$ , and  $v$  is a vector dual to  $x$ .  $\square$

An explicit formula for the condition number can now be identified.

**Corollary 5.2** *The condition number  $\kappa_{P,b}(\omega, x)$  is given by*

$$\kappa_{P,b}(\omega, x) = \|P(\omega)^{-1}\| \left( \frac{\beta}{\|x\|} + \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right). \quad (5.5)$$

For the rest of this section we specialize to normwise relative perturbations (see (5.2)). The dominant term in (5.5) is then

$$\kappa_P(\omega) := \|P(\omega)^{-1}\| \sum_{i=0}^{\ell} |\omega|^i \|A_i\| \in \left[ \frac{1}{2} \kappa_{P,b}(\omega, x), \kappa_{P,b}(\omega, x) \right], \quad (5.6)$$

since  $\|b\| = \|P(\omega)x\| \leq \sum_{i=0}^{\ell} |\omega|^i \|A_i\| \|x\|$ . Hence which right-hand side we choose for the linearized system has little effect on the conditioning of the system when  $P$  is subject to perturbation. Our aim in the rest of this section is to compare  $\kappa_{L,c}(\omega, z)$  to  $\kappa_{P,b}(\omega, x)$  and to derive sufficient conditions on the coefficient matrices and parameters defining  $L$  for  $\kappa_{L,c}(\omega, z) \approx \kappa_{P,b}(\omega, x)$  to hold.

For the 2-norm, (5.6) implies that

$$\frac{\kappa_{L,c}(\omega, z)}{\kappa_{P,b}(\omega, x)} \approx \frac{\kappa_L(\omega)}{\kappa_P(\omega)} = \frac{\|L(\omega)^{-1}\|_2 \|\omega\| \|X\|_2 + \|Y\|_2}{\|P(\omega)^{-1}\|_2 \sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2}. \quad (5.7)$$

We will need a result from [13, Lem. 3.5] that is useful when taking norms of block matrices.

**Lemma 5.3** *For any block  $\ell \times m$  matrix  $B = (B_{ij})$  we have  $\max_{i,j} \|B_{ij}\|_2 \leq \|B\|_2 \leq \sqrt{\ell m} \max_{i,j} \|B_{ij}\|_2$ .*

We will concentrate on the companion form pencils,  $C_1(\omega) = \omega X_1 + Y_1$  and  $C_2(\omega) = \omega X_2 + Y_2$  given by (2.7). For  $k = 1, 2$ ,  $\|X_k\|_2 = \max(\|A_\ell\|_2, 1)$  and from Lemma 5.3,  $\max(1, \max_{i=0:\ell-1} \|A_i\|_2) \leq \|Y_k\|_2 \leq \ell \max(1, \max_{i=0:\ell-1} \|A_i\|_2)$ . Hence

$$\frac{|\omega| \|X_k\|_2 + \|Y_k\|_2}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \geq \frac{|\omega| \max(1, \|A_\ell\|_2) + \max(1, \max_{i=0:\ell-1} \|A_i\|_2)}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \geq \frac{1}{\|A\|_1}. \quad (5.8)$$

As an upper bound we obtain

$$\frac{|\omega| \|X_k\|_2 + \|Y_k\|_2}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \leq (|\omega| + \ell) \frac{\max(1, \max_i \|A_i\|_2)}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \leq \frac{\max(1, \max_i \|A_i\|_2) (|\omega| + \ell)}{\min(\|A_0\|_2, \|A_\ell\|_2) (1 + |\omega|^\ell)}. \quad (5.9)$$

We now need the following result from [13, Lem. 3.4] in order to bound the ratio  $\|C_k(\omega)^{-1}\|_2 / \|P(\omega)^{-1}\|_2$ . Recall that  $A^{\mathcal{B}}$  denotes the block transpose of  $A$ .

**Lemma 5.4** *For the first and second companion forms  $C_1$  and  $C_2$  there exists a block  $\ell \times \ell$  matrix polynomial  $R(\lambda) \in \mathbb{C}^{\ell n \times \ell n}$  such that*

$$R(\lambda)C_1(\lambda) = I_\ell \otimes P(\lambda) = C_2(\lambda)R(\lambda)^{\mathcal{B}}, \quad (5.10)$$

where the  $n \times n$  blocks of  $R(\lambda)$  are given by

$$[R(\lambda)]_{ij} = \begin{cases} \lambda^{\ell-i} I, & 1 \leq i \leq \ell, j = 1, \\ -\lambda^{j-i} \sum_{k=0}^{\ell-j} \lambda^{k-1} A_k, & 1 \leq i < j, 1 < j \leq \ell, \\ \lambda^{j-i} \sum_{k=\ell-j+1}^{\ell} \lambda^{k-1} A_k, & 1 < j \leq i \leq \ell. \end{cases} \quad (5.11)$$

(Note that (2.11) is the special case with  $\ell = 3$ .)

Thus

$$\|A\|_2 \leq \frac{\|C_k(\omega)^{-1}\|_2}{\|P(\omega)^{-1}\|_2} \leq \|R(\omega)\|_2 \leq \ell \|A\|_1 \max\left(1, \max_i \|A_i\|_2\right), \quad k = 1, 2, \quad (5.12)$$

where the lower bound is from Lemma 4.3 (a) with  $v = e_1$  and  $f = A$  and the upper bounds are from Lemma 5.3 and Lemma 5.4. Hence, combined with (5.7)–(5.9) and (A.1), this yields

$$\frac{1}{\ell^{1/2}} \leq \frac{\kappa_{C_k}(\omega)}{\kappa_P(\omega)} \leq \ell^3 \frac{\max(1, \max_i \|A_i\|_2)^2}{\min(\|A_0\|_2, \|A_\ell\|_2)}, \quad k = 1, 2. \quad (5.13)$$

When the coefficient matrices of  $P$  have norms that differ widely, the companion matrices  $C_k(\lambda)$ ,  $k = 1, 2$  are badly scaled and the bounds above signal that  $\kappa_{C_k} \gg \kappa_P$  is possible. For this reason we next analyze the effect on the conditioning of scaling the identity blocks of  $C_1$  and  $C_2$ , using a scaling proposed by Higham, Li, and Tisseur [13, Sec. 3.3].

Let  $D = \text{diag}(d) \otimes I_n$ , where  $d \in \mathbb{R}^\ell$  with  $d_1 = 1$  and  $d_i = \max_j \|A_j\|_2$ ,  $i = 2: \ell$ . It is easy to check that  $DC_1(\omega) = \omega \tilde{X}_1 + \tilde{Y}_1 = \tilde{C}_1 \in \mathbb{L}_1(P)$  with  $v = e_1$  and  $C_2(\omega)D = \omega \tilde{X}_2 + \tilde{Y}_2 = \tilde{C}_2 \in \mathbb{L}_2(P)$  with  $\tilde{v} = e_1$ , and that we have, using (A.1) again,

$$\frac{|\omega| + 1}{\sum_{i=0}^{\ell} |\omega|^i} \leq \frac{|\omega| \|\tilde{X}_k\|_2 + \|\tilde{Y}_k\|_2}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \leq \frac{(|\omega| + \ell) \max_i \|A_i\|_2}{(|\omega|^\ell + 1) \min(\|A_0\|_2, \|A_\ell\|_2)} \leq \frac{\ell^2}{\|A\|_1} \rho \quad (5.14)$$

for  $k = 1, 2$ , where

$$\rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_\ell\|_2)}. \quad (5.15)$$

Also, since (5.10) becomes  $R(\omega)D^{-1} \cdot DC_1(\omega) = I_\ell \otimes P(\omega) = C_2(\omega)D \cdot D^{-1}R(\omega)^B$ , the bounds in (5.12) translate to

$$\|A\|_2 \leq \frac{\|\tilde{C}_k(\omega)^{-1}\|_2}{\|P(\omega)^{-1}\|_2} \leq \ell \|A\|_1, \quad k = 1, 2,$$

so that, overall, using (A.2) for the lower bound,

$$\frac{2\ell^{1/2}}{\ell + 1} \leq \frac{\kappa_{\tilde{C}_k}(\omega)}{\kappa_P(\omega)} \leq \ell^3 \rho, \quad k = 1, 2. \quad (5.16)$$

The upper bound in (5.16) is much smaller than that in (5.13) if  $\max_i \|A_i\|_2 \gg 1$  or  $\max_i \|A_i\|_2 \ll 1$ , showing that block scaling has potentially a very beneficial effect on the conditioning of the companion pencils.

One way to reduce  $\rho$  is by scaling the parameter  $\omega$ . Write  $\omega = \theta\mu$  and

$$P(\omega) = \tilde{P}(\mu) = \sum_{i=0}^{\ell} \mu^i \tilde{A}_i, \quad \tilde{A}_i = \theta^i A_i.$$

For the scaled polynomial  $\tilde{P}$ , we have

$$\rho_{\tilde{P}} = \rho_{\tilde{P}}(\theta) = \frac{\max_i \theta^i \|A_i\|_2}{\min(\|A_0\|_2, \theta^\ell \|A_\ell\|_2)}.$$

Betcke [5] shows that the unique minimizer of  $\rho_{\tilde{P}}(\theta)$  over all  $\theta > 0$  is  $\theta_{\text{opt}} = (\|A_0\|_2 / \|A_\ell\|_2)^{1/2}$ . For  $\ell = 2$ , Fan, Lin, and Van Dooren [7] suggest the same scaling and also multiply all three coefficient matrices by another scalar, chosen to bring the norms as

close as possible to 1. This double scaling is used in [16], for example. It yields  $2/3 \leq \max(\|\tilde{A}_2\|_2, \|\tilde{A}_1\|_2, \|\tilde{A}_0\|_2) \leq 2$  and  $\rho_{\tilde{P}}(\theta_{\text{opt}}) = \max(1, \|\tilde{A}_1\|_2 / (\|\tilde{A}_0\|_2 \|\tilde{A}_2\|_2)^{1/2})$ , and so  $\rho_{\tilde{P}}(\theta_{\text{opt}}) \approx 1$  if the quadratic is elliptic [17], [23] or, in the terminology of quadratics arising from mechanical systems with damping, the system is not too heavily damped. It is easy to see that for normwise relative perturbations,  $\kappa_P$  is invariant under the Fan, Lin, and Van Dooren scaling.

## 6. Backward error for linear systems

Suppose that  $\hat{x}$  is an approximate solution to  $P(\omega)x = b$ . We can interpret  $\hat{x}$  as the exact solution of a perturbed system

$$(P(\omega) + \Delta P(\omega))\hat{x} = \left( \sum_{i=0}^{\ell} \omega^i (A_i + \Delta A_i) \right) \hat{x} = b + \Delta b,$$

where there are many possible choices of  $\Delta A_i$ ,  $i = 0:\ell$  and  $\Delta b$ . We define the backward error to be the smallest of all such perturbations, in the following sense:

$$\eta_{P,b}(\omega, \hat{x}) = \min\{\epsilon : (P(\omega) + \Delta P(\omega))\hat{x} = b + \Delta b, \|\Delta A_i\| \leq \epsilon \alpha_i, i = 0:\ell, \|\Delta b\| \leq \epsilon \beta\}.$$

We denote by  $\eta_P(\omega, \hat{x})$  the backward error with unperturbed right hand side ( $\beta = 0$ ).

From a straightforward modification of a result of Rigal and Gaches on the normwise backward error for a linear system [11, Thm. 7.1], [28] we derive an explicit expression for  $\eta_{P,b}(\omega, \hat{x})$ .

**Theorem 6.1** *The normwise backward error  $\eta_{P,b}(\omega, \hat{x})$  is given by*

$$\eta_{P,b}(\omega, \hat{x}) = \frac{\|b - P(\omega)\hat{x}\|}{\left( \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right) \|\hat{x}\| + \beta}. \quad (6.1)$$

**Proof.** It is straightforward to show that the right hand side of (6.1) is a lower bound for  $\eta_{P,b}(\omega, \hat{x})$ . This lower bound is attained for the perturbations

$$\Delta A_i = -\frac{\text{sign}(\omega^i) \alpha_i \|\hat{x}\| r z^*}{\left( \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right) \|\hat{x}\| + \beta}, \quad i = 0:\ell, \quad \Delta b = \frac{\beta r}{\left( \sum_{i=0}^{\ell} |\omega|^i \alpha_i \right) \|\hat{x}\| + \beta},$$

where  $r = b - P(\omega)\hat{x}$  is the residual vector and  $z$  is a vector dual to  $\hat{x}$ .  $\square$

A straightforward modification of [11, Prob. 7.7] yields the following result for normwise relative perturbations.

**Lemma 6.2** *Let  $\alpha_i \equiv \|A_i\|$  and  $\beta = \|b\|$ . Then*

$$\eta_{P,b}(\omega, \hat{x}) \leq \eta_P(\omega, \hat{x}) \leq \frac{2\eta_{P,b}(\omega, \hat{x})}{1 - \eta_{P,b}(\omega, \hat{x})}.$$

Hence if  $\eta_{P,b}(\omega, \hat{x}) \ll 1$  then the normwise relative backward error  $\eta_P(\omega, \hat{x})$  with unperturbed right hand side is within a small factor of  $\eta_{P,b}(\omega, \hat{x})$ . For this reason we consider only  $\eta_P(\omega, \hat{x})$  in the rest of this section. We concentrate on the 2-norm from this point on and set  $\alpha_i \equiv \|A_i\|_2$ .

To relate backward errors for  $L$  and  $P$  we need to assume that the pencil  $L$  satisfies a *left*-sided factorization (2.2), with  $E(\omega)$  of full rank. Recalling Theorem 4.1(b), let  $\hat{z}$  be the computed solution to  $L(\omega)z = c$  with  $c \in \mathbb{C}^r$  such that  $E(\omega)c = \gamma b$  for some nonzero scalar  $\gamma$ . Then,

$$E(\omega)(L(\omega)\hat{z} - c) = P(\omega)H(\omega)\hat{z} - \gamma b.$$

So if we recover  $\hat{x}$  from  $\hat{z}$  as  $\hat{x} = \gamma^{-1}H(\omega)\hat{z}$  we have a well-defined relation between the residual for the linearized system and the residual for the original problem. Let us assume that  $\hat{z}$  is computed exactly from this expression; certainly in the common case where  $H(\omega) = h^T \otimes I$  and  $h$  is a unit vector  $e_k$  (see (2.6)),  $H(\omega)\hat{z} = (e_k^T \otimes I)\hat{z} = \hat{z}_k$  is obtained exactly. In particular,

$$\|P(\omega)\hat{x} - b\|_2 \leq \gamma^{-1}\|E(\omega)\|_2\|L(\omega)\hat{z} - c\|_2.$$

From (6.1) we have

$$\begin{aligned} \frac{\eta_P(\omega, \hat{x})}{\eta_L(\omega, \hat{z})} &= \frac{\|P(\omega)\hat{x} - b\|_2}{\|L(\omega)\hat{z} - c\|_2} \cdot \frac{|\omega|\|X\|_2 + \|Y\|_2}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \frac{\|\hat{z}\|_2}{\|\hat{x}\|_2} \\ &\leq \|E(\omega)\|_2 \cdot \frac{|\omega|\|X\|_2 + \|Y\|_2}{\sum_{i=0}^{\ell} |\omega|^i \|A_i\|_2} \frac{\|\hat{z}\|_2}{\|H(\omega)\hat{z}\|_2}. \end{aligned}$$

From Lemma 5.4 and (2.6) it follows that for  $C_1$  the left-sided factorization (2.2) holds for

$$E_k(\omega) = (e_k^T \otimes I_n)R(\omega), \quad H(\omega) = e_k^T \otimes I, \quad k = 1:\ell, \quad (6.2)$$

and for  $c = e_1 \otimes b$  we have that  $E_k(\omega)c = \omega^{\ell-k}b$ . So  $\gamma = \omega^{\ell-k}$  and  $\hat{x} = \hat{z}_k/\omega^{\ell-k}$ . Using  $\|E_k(\omega)\|_2 \leq \ell^{1/2}\|A\|_1 \max(1, \max_i \|A_i\|_2)$ , (5.9), and (A.1), we obtain

$$\frac{\eta_P(\omega, \hat{x})}{\eta_{C_1}(\omega, \hat{z})} \leq \ell^{5/2} \frac{\max(1, \max_i \|A_i\|_2)^2}{\min(\|A_0\|_2, \|A_\ell\|_2)} \frac{\|\hat{z}\|_2}{\|\hat{z}_k\|_2}. \quad (6.3)$$

Note that the second,  $A_i$ -dependent term in the bound is the same as that in the bound (5.13) for the ratio of condition numbers.

If  $\|\hat{z}_k\|_2 \ll \|\hat{z}\|_2$  this bound is large, reflecting the fact that the computed  $\hat{x}$  is likely to suffer from damaging cancellation. For  $C_1$ , and more generally any  $L \in \mathbb{L}_1(P)$ , Corollary 4.2(a) shows that the exact  $z$  has the form  $\Lambda \otimes x$ . To minimize the bound in (6.3) we should choose  $k = \ell$  if  $|\omega| \ll 1$  and  $k = 1$  if  $|\omega| \gg 1$ , whence  $\|z\|_2/\|z_k\|_2 \leq \ell^{1/2}$ .

For the block scaled companion pencil  $\tilde{C}_1$ , for which  $E(\omega)$  is replaced by  $E(\omega)D^{-1}$ , we have, using (5.14), the bound

$$\frac{\eta_P(\omega, \hat{x})}{\eta_{\tilde{C}_1}(\omega, \hat{z})} \leq \ell^{5/2} \rho \frac{\|\hat{z}\|_2}{\|\hat{z}_k\|_2}, \quad (6.4)$$

which is much smaller than (6.3) when  $\max_i \|A_i\|_2 \gg 1$  or  $\max_i \|A_i\|_2 \ll 1$ .

For the second companion linearization, the factorization (2.2) holds for

$$E(\omega) = \Lambda^T \otimes I, \quad H(\omega) = e_1^T \otimes I, \quad (6.5)$$

and for  $c = u \otimes b$  for any  $u \in \mathbb{C}^n$  such that  $\Lambda^T u \neq 0$ , Corollary 4.2(b) gives  $\hat{x} = \hat{z}_1 / (\Lambda^T u)$ . Using (5.9) and (A.1) we obtain

$$\frac{\eta_P(\omega, \hat{x})}{\eta_{C_2}(\omega, \hat{z})} \leq \ell^2 \frac{\max(1, \max_i \|A_i\|_2) \|\hat{z}\|_2}{\min(\|A_0\|_2, \|A_\ell\|_2) \|\hat{z}_1\|_2}. \quad (6.6)$$

For the block scaled second companion pencil this bound improves to

$$\frac{\eta_P(\omega, \hat{x})}{\eta_{\tilde{C}_2}(\omega, \hat{z})} \leq \ell^2 \rho \frac{\|\hat{z}\|_2}{\|\hat{z}_1\|_2}.$$

An important conclusion that can be drawn from (5.16) and (6.4) is that for the block scaled first companion pencil, the desirable relations  $\kappa_{\tilde{C}_1}(\omega) \approx \kappa_P(\omega)$  and  $\eta_{\tilde{C}_1}(\omega, \hat{z}) \approx \eta_P(\omega, \hat{x})$  hold for a suitable choice of  $k$ , provided that  $\rho \approx 1$ .

## 7. Numerical experiments

To illustrate the theory we report experiments with linear systems  $P(\omega)x = b$  corresponding to three quadratic matrix polynomials from the NLEVP collection [6]. In each case we use the first companion linearization  $C_1$  and the augmented system with  $c = e_1 \otimes b$  and solve the problem in three forms: with  $C_1$  and  $P$  both unscaled, with  $C_1$  having the block scaling and  $P$  unscaled, and with  $C_1$  unscaled but  $P$  having the Fan, Lin, and Van Dooren scaling. We evaluate the condition numbers and backward errors for normwise relative perturbations (thus with the parameters (5.2)). We report condition numbers and backward errors over frequencies  $\omega = 2\pi i 10^t$  with  $t$  taking ten equally spaced values on  $[-3, 3]$ . The right-hand side  $b$  has equally spaced entries on the interval  $[-2, 1]$ .

Recall that key quantities are

$$\theta = \frac{\max(1, \max_i \|A_i\|_2)^2}{\min(\|A_0\|_2, \|A_\ell\|_2)}, \quad \rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_\ell\|_2)}, \quad (7.1)$$

since  $\theta$  is a factor in the condition number and backward error bounds (5.13) and (6.3) for the unscaled  $C_1$ , while  $\rho$  is a factor in the condition number and backward error bounds (5.16) and (6.4) for the block scaled  $C_1$ ,

The first test is `nlevp('railtrack')`, which is a badly scaled quadratic of dimension 1005 arising from a model of the vibration of rail tracks under the excitation of high speed trains [18], [25]. The condition numbers  $\kappa_P$  and  $\kappa_{C_1}$  are plotted in Figure 7.1. We have  $\|A_2\|_1 = 4.2 \times 10^{10}$ ,  $\|A_1\|_1 = 1.9 \times 10^{11}$ ,  $\|A_0\|_1 = 3.1 \times 10^{10}$ ,  $\theta = 1.2 \times 10^{12}$  and  $\rho = 6.2$  for the original  $P$ , and  $\theta = 8.9$  and  $\rho = 5.3$  for  $P$  after the Fan-Lin-Van Dooren scaling. Hence our theory suggests that  $\kappa_{C_1}(\omega)$  may be up to a factor  $10^{12}$  larger than  $\kappa_P(\omega)$ , but that with block scaling of  $C_1$  or the Fan-Lin-Van Dooren scaling of  $P$ , the condition numbers must be of the same order of magnitude. This behavior is confirmed by Figure 7.1.

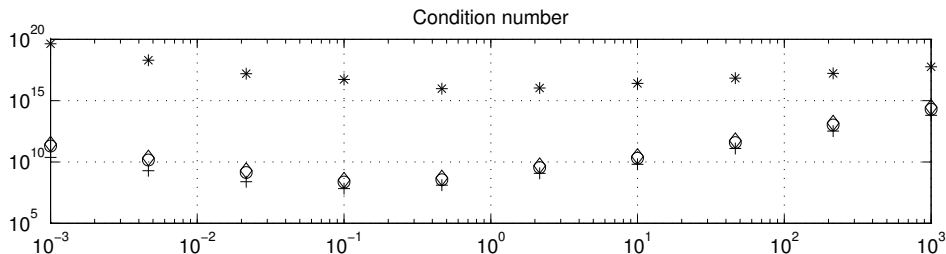


Figure 7.1: Railtrack problem: frequency  $\omega$  against condition number  $\kappa_{C_1}$  or  $\kappa_P$ . Key: unscaled  $C_1$  (“\*”), block scaled  $C_1$  (“◇”),  $C_1$  with Fan-Lin-Van Dooren scaling of  $P$  (“o”), and  $P$  (“+”).

The second problem is `nlevp('cd_player')`, which is a quadratic of dimension 60 arising in the study of a CD player control task. Figure 7.2 plots the condition numbers  $\kappa_P$  and  $\kappa_{C_1}$  along with the ratios  $\eta_{P,b}(\omega, \hat{x})/\eta_{C_1}(\omega, \hat{z})$  of backward errors, where  $\hat{x}$  is recovered from  $\hat{z}$  as described just after (6.2), with  $k$  chosen to maximize  $\|z_k\|_2$ , and  $\hat{z}$  is computed via the MATLAB backslash operator or via the MATLAB `gmres` function with no restarts, a convergence tolerance  $10^{-4}$ , and a random starting vector within relative distance  $10^{-2}$  of the true solution. We have  $\|A_2\|_1 = 1.0$ ,  $\|A_1\|_1 = 1.1 \times 10^7$ ,  $\|A_0\|_1 = 2.5 \times 10^5$ ,  $\theta = 1.2 \times 10^{14}$  and  $\rho = 1.1 \times 10^7$  for the original  $P$ , and  $\theta = 4.3 \times 10^4$  and  $\rho = 4.4 \times 10^4$  for  $P$  after the Fan-Lin-Van Dooren scaling. Again, we see scaling bringing improvements consistent with the bounds. However, for  $|\omega| > 10$  block scaling produces a slight worsening in the conditioning of  $C_1$ , and the backward error for backslash is worsened by the Fan-Lin-Van Dooren scaling for most  $\omega$ ; this behavior is within the freedom of a factor  $10^4$  afforded by the bounds.

Our final example is `nlevp('damped_beam')`, which is a quadratic from a finite element model of a beam clamped at both ends with a damper in the middle, and which is analyzed in detail by Higham, Mackey, Tisseur, and Garvey [16] with respect to the eigenvalue problem. Here,  $\|A_2\|_1 = 6.7 \times 10^{-3}$ ,  $\|A_1\|_1 = 5.0$ ,  $\|A_0\|_1 = 1.8 \times 10^9$ ,  $\theta = 4.6 \times 10^{20}$  and  $\rho = 2.6 \times 10^{11}$  for the original  $P$ , and  $\theta = 2.0$  and  $\rho = 1.0$  for  $P$  after the Fan-Lin-Van Dooren scaling. For computations analogous to those in the second example, the results are shown in Figure 7.3. Surprisingly, the condition numbers of  $C_1$  are in several cases very close to those of  $P$  when no scaling is used, despite the large values of  $\theta$  and  $\rho$ . Since the Fan-Lin-Van Dooren scaling produces  $\theta$  and  $\rho$  of order 1 it guarantees the ideal behavior that  $\eta_{P,b} \approx \eta_{C_1}$ , which is verified by the second and third plots since the corresponding ratios are of order 1. Note that there is some growth of the ratios with  $|\omega|$  for block scaling, but that this is well within the freedom afforded by the bounds given that  $\rho$  is of order  $10^{11}$ .

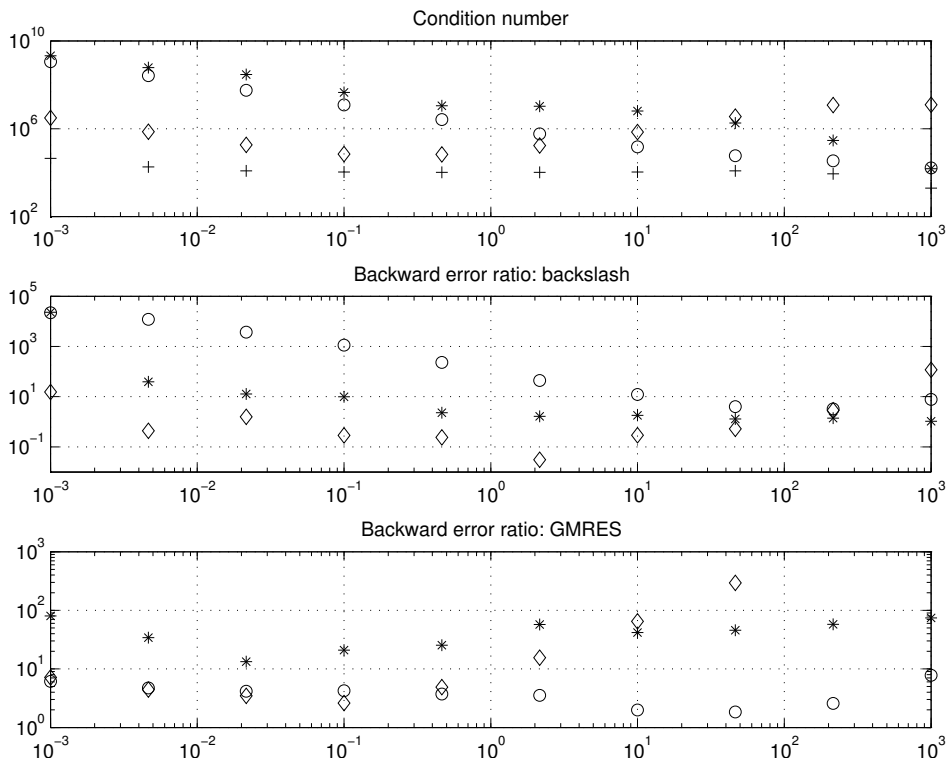


Figure 7.2: CD player problem.: frequency  $\omega$  against condition number  $\kappa_{C_1}$  or  $\kappa_P$  or backward error ratio  $\eta_{P,b}/\eta_{C_1,c}$ . Key: unscaled  $C_1$  (“\*”), block scaled  $C_1$  (“◇”),  $C_1$  with Fan-Lin-Van Dooren scaling of  $P$  (“o”), and  $P$  (“+”).

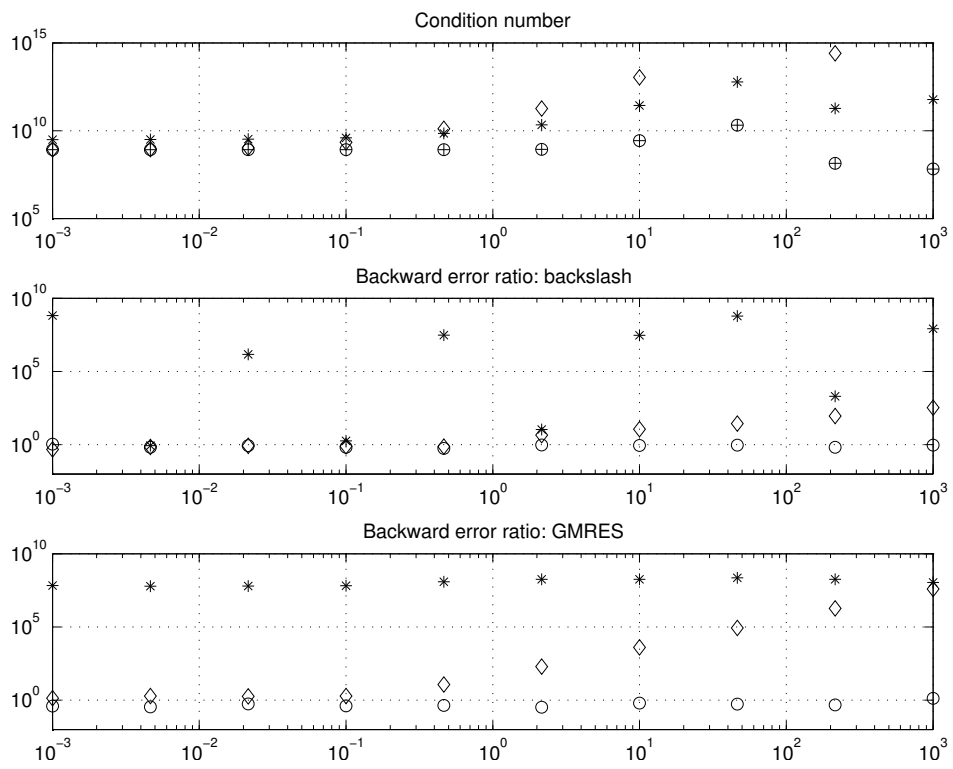


Figure 7.3: Damped beam problem: frequency  $\omega$  against condition number  $\kappa_{C_1}$  or  $\kappa_P$  or backward error ratio  $\eta_{P,b}/\eta_{C_1,c}$ . Key: unscaled  $C_1$  (“\*”), block scaled  $C_1$  (“◇”),  $C_1$  with Fan-Lin-Van Dooren scaling of  $P$  (“+”), and  $P$  (“+”).

## 8. Concluding remarks

A general technique for solving nonlinear eigenvalue problems and parametrized linear systems is to reduce the nonlinear problem to a larger but simpler (usually linear) problem. For polynomial eigenvalue problems various classes of linearizations have been derived and analyzed. In particular, analysis in [13], [15] compares the sensitivity of the original and certain linearized problems and connects the backward errors of their approximate solutions. In this work we have introduced a way to treat general nonlinear matrix functions  $N(\lambda)$  in terms of one-sided factorizations relating  $N(\lambda)$  to the simpler function  $M(\lambda)$ . We have shown that such factorizations hold in many important special cases and that they imply close relations between the eigensystems of  $N$  and  $M$  (Theorem 3.1) and between the solutions of the parametrized linear systems  $N(\omega)x = b$  and the augmented systems  $M(\omega)z = c$  (Theorem 4.1). We have developed the theory in some detail for parametrized linear systems, which have received little attention in the literature to date, but our techniques are equally applicable to rational and general nonlinear eigenproblems. The one-sided factorization framework provides a balance between simplicity (so that the factorizations can be found) and utility (so that informative results can be proved), and we intend to explore its use further in future work.

## A. Appendix

We need the following two pairs of bounds, which are proved in [15, Lem. A.1]:

$$1 \leq \frac{(1+x^2)(1+x^2+x^4+\dots+x^{2(\ell-1)})}{1+x^{2\ell}} \leq \ell, \quad (\text{A.1})$$

$$\frac{2\ell^{1/2}}{\ell+1} \leq \frac{(1+x)(1+x^2+x^4+\dots+x^{2(\ell-1)})^{1/2}}{1+x+x^2+\dots+x^\ell} \leq 1. \quad (\text{A.2})$$

## References

- [1] M. Al-Ammari and F. Tisseur. On quadratization of matrix polynomials. MIMS EPrint 2009.xx, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2009. In preparation.
- [2] A. Amiraslani, R. M. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29:141–157, 2009.
- [3] E. N. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electron. J. Linear Algebra*, 11:78–87, 2004.
- [4] A. C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2005. ISBN 0-89871-529-6. xxv+479 pp.
- [5] T. Betcke. Optimal scaling of generalized and polynomial eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 30(4):1320–1338, 2008.
- [6] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur. NLEVP: A collection of nonlinear eigenvalue problems. <http://www.mims.manchester.ac.uk/research/numerical-analysis/nlevp.html>.
- [7] H.-Y. Fan, W.-W. Lin, and P. Van Dooren. Normwise scaling of second order polynomial matrices. *SIAM J. Matrix Anal. Appl.*, 26(1):252–256, 2004.
- [8] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982. ISBN 0-12-287160-X. xiv+409 pp.
- [9] C.-H. Guo, N. J. Higham, and F. Tisseur. Detecting and solving hyperbolic quadratic eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 30(4):1593–1613, 2009.

- [10] C.-H. Guo and P. Lancaster. Algorithms for hyperbolic quadratic eigenvalue problems. *Math. Comp.*, 74(252):1777–1791, 2005.
- [11] N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2002. ISBN 0-89871-521-0. xxx+680 pp.
- [12] N. J. Higham and H.-M. Kim. Numerical analysis of a quadratic matrix equation. *IMA J. Numer. Anal.*, 20(4):499–519, 2000.
- [13] N. J. Higham, R.-C. Li, and F. Tisseur. Backward error of polynomial eigenproblems solved by linearization. *SIAM J. Matrix Anal. Appl.*, 29(4):1218–1241, 2007.
- [14] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159, 2006.
- [15] N. J. Higham, D. S. Mackey, and F. Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.
- [16] N. J. Higham, D. S. Mackey, F. Tisseur, and S. D. Garvey. Scaling, sensitivity and stability in the numerical solution of quadratic eigenvalue problems. *Internat. J. Numer. Methods Eng.*, 73(3): 344–360, 2008.
- [17] N. J. Higham, F. Tisseur, and P. M. Van Dooren. Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems. *Linear Algebra Appl.*, 351–352:455–474, 2002.
- [18] A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In P. Neittaanmäki, T. Rossi, S. Korotov, E. Oñate, J. Périaux, and D. Knörzer, editors, *Proceedings of the European Congress on Computational Methods in Applied Sciences and Engineering (EC-COMAS 2004)*, Jyväskylä, Finland, 2004. <http://www.mit.jyu.fi/eccomas2004/proceedings/proceed.html>.
- [19] T.-M. Huang, W.-W. Lin, and J. Qian. Structure-preserving algorithms for palindromic quadratic eigenvalue problems arising from vibration of fast trains. *SIAM J. Matrix Anal. Appl.*, 30(4): 1566–1592, 2009.
- [20] T.-M. Huang, W.-W. Lin, and W.-S. Su. Palindromic quadratization and structure-preserving algorithm for palindromic matrix polynomials of even degree. NCTS preprints in mathematics, National Center for Theoretical Sciences, National Tsing Hua University, Taiwan, June 2009. 20 pp.
- [21] M. Kuzuoglu and R. Mittra. Finite element solution of electromagnetic problems over a wide frequency range via the Padé approximation. *Comput. Methods Appl. Mech. Engrg.*, 169(3-4): 263–277, 1999.
- [22] P. Lancaster. *Lambda-Matrices and Vibrating Systems*. Pergamon Press, Oxford, 1966. ISBN 0-486-42546-0. xiii+196 pp. Reprinted by Dover, New York, 2002.
- [23] P. Lancaster. Quadratic eigenvalue problems. *Linear Algebra Appl.*, 150:499–506, 1991.
- [24] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, London, second edition, 1985. ISBN 0-12-435560-9. xv+570 pp.
- [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(4):1029–1051, 2006.
- [26] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.
- [27] K. Meerbergen. Fast frequency response computation for Rayleigh damping. *Internat. J. Numer. Methods Eng.*, 73(1):96–106, 2008.
- [28] J. L. Rigal and J. Gaches. On the compatibility of a given solution with the data of a linear system. *J. Assoc. Comput. Mach.*, 14(3):543–548, 1967.
- [29] V. Simoncini and F. Perotti. On the numerical solution of  $(\lambda^2 A + \lambda B + C)x = b$  and application to structural dynamics. *SIAM J. Sci. Comput.*, 23(6):1875–1897, 2002.
- [30] Y. Su and Z. Bai. Solving rational eigenvalue problems via linearization. Technical Report CSE-2008-13, Department of Computer Science, UC Davis, Davis, CA, USA, Sept. 2008. 19 pp. Available from <http://www.cs.ucdavis.edu/research/tech-reports/index.html>.
- [31] F. Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra Appl.*, 309:339–361, 2000.
- [32] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.