



SPECTRAL THEORY FOR SELF-ADJOINT QUADRATIC EIGENVALUE PROBLEMS – A REVIEW*

PETER LANCASTER[†] AND ION ZABALLA[‡]

Abstract. Many physical problems require the spectral analysis of quadratic matrix polynomials $M\lambda^2 + D\lambda + K$, $\lambda \in \mathbb{C}$, with $n \times n$ Hermitian matrix coefficients, M , D , K . In this largely expository paper, we present and discuss canonical forms for these polynomials under the action of both congruence and similarity transformations of a linearization and also λ -dependent unitary similarity transformations of the polynomial itself. Canonical structures for these processes are clarified, with no restrictions on eigenvalue multiplicities. Thus, we bring together two lines of attack: (a) **analytic** via direct reduction of the $n \times n$ system itself by λ -dependent unitary similarity and (b) **algebraic** via reduction of $2n \times 2n$ symmetric linearizations of the system by either *congruence* (Section 4) or *similarity* (Sections 5 and 6) transformations which are independent of the parameter λ . Some new results are brought to light in the process. Complete descriptions of associated canonical structures (over \mathbb{R} and over \mathbb{C}) are provided – including the two cases of real symmetric coefficients and complex Hermitian coefficients. These canonical structures include the so-called *sign characteristic*. This notion appears in the literature with different meanings depending on the choice of canonical form. These sign characteristics are studied here and connections between them are clarified. In particular, we consider which of the linearizations reproduce the (intrinsic) signs associated with the analytic (Rellich) theory (Sections 7 and 9).

Key words. Matrix polynomials, Canonical forms, Linearizations, Eigenvalue functions, Sign characteristic, Elementary divisors.

AMS subject classifications. 7A56, 15A22, 39A99.

1. Introduction. Let $\mathbb{F} = \mathbb{R}$ or \mathbb{C} , the fields of real or complex numbers, respectively. We consider quadratic matrix polynomials P over \mathbb{F} with *real symmetric* or *complex Hermitian matrix coefficients*, $M, D, K \in \mathbb{F}^{n \times n}$ (the linear space of $n \times n$ matrices over \mathbb{F}). Thus,

$$(1) \quad P(\lambda) := M\lambda^2 + D\lambda + K, \quad \lambda \in \mathbb{C},$$

and, on occasion, we will refer to $P(\lambda)$ as a quadratic, Hermitian *system*.

In applications, it is frequently the case that M is positive definite, but this property is often “at risk” (if linearly dependent coordinates are chosen in a finite element scheme, for example) and so, wherever possible, we prefer to work in the more general context of *possibly singular leading coefficient*. It will be assumed throughout that $P(\lambda)$ is a regular matrix polynomial; that is, $\det P(\lambda)$ is not identically zero.

The *spectrum* of P (the set of *eigenvalues* of P) is defined by

$$(2) \quad \sigma(P) := \{\lambda \in \mathbb{C} : \det P(\lambda) = 0\},$$

*Received by the editors on April 16, 2020. Accepted for publication on February 22, 2021. Handling Editor: Froilán Dopico. Corresponding Author: Ion Zaballa. P. Lancaster was supported in part by the Natural Sciences and Engineering Research Council of Canada. I. Zaballa was supported by ‘Ministerio de Ciencia, Innovación y Universidades’ of Spain and ‘Fondo Europeo de Desarrollo Regional (FEDER)’ of EU through grants MTM2017-83624-P and MTM2017-90682-REDT, and by UPV/EHU through grant GIU16/42.

[†]Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta, AB T2N 1N4, Canada (lancaste@ucalgary.ca).

[‡]Departamento de Matemática Aplicada y EIO, Euskal Herriko Unibertsitatea (UPV/EHU), Apdo. Correos 644, 48080 Bilbao, Spain (ion.zaballa@ehu.es).

and recall that the location of the spectrum relative to the real and imaginary axes of the complex plane frequently determines important physical properties of the system.

Linearizations play a central role when developing **the spectral theory** of matrix polynomials (see, for example, [5]). Following, for instance, [6] (Section 7.2) and [15], a *linearization* of $P(\lambda)$ is defined to be a linear pencil $L(\lambda) := \lambda X + Y$, $X, Y \in \mathbb{F}^{2n \times 2n}$ for which

$$(3) \quad E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_n \end{bmatrix},$$

for some $2n \times 2n$ unimodular matrix polynomials $E(\lambda), F(\lambda)$ (i.e. with constant, nonzero determinant). Then the matrix functions $L(\lambda)$ and $\begin{bmatrix} P(\lambda) & 0 \\ 0 & I_n \end{bmatrix}$ are said to be *equivalent* and, what is important for us, $\sigma(L) = \sigma(P)$.

In the spectral theory developed in [6], M is assumed to be the identity matrix, but the whole theory also applies if M is nonsingular ([5] or Chapter 12 of [7]). In all these references, the theory is mostly confined to complex matrix polynomials (see also [14] for the real case). In contrast, for real symmetric or complex Hermitian linear pencils, singular leading coefficients are admitted if we apply the analysis of [12] and [13].

When M is nonsingular, the *companion matrix* of $P(\lambda)$ is defined to be

$$(4) \quad C_P := \begin{bmatrix} 0 & I_n \\ -M^{-1}K & -M^{-1}D \end{bmatrix},$$

the *Jordan canonical form* for $P(\lambda)$ is defined to be that of C_P , and $I_n\lambda - C_P$ is a linearization of $P(\lambda)$. However, in computational practice, it is frequently the case that M is singular (as in [1]) – or is ‘dangerously close’ to a singular matrix – and the more general notion of strong linearization is required.

DEFINITION 1. A linearization $L(\lambda)$, as in (3), is said to be a strong linearization of $P(\lambda)$ if the reverse linear pencil, $\text{rev } L(\lambda) := \lambda Y + X$, is a linearization of the reverse polynomial:

$$\text{rev } P(\lambda) := M + D\lambda + K\lambda^2.$$

Also, $P(\lambda)$ is said to have infinity as an eigenvalue (or that ∞ is an eigenvalue of $P(\lambda)$) if 0 is an eigenvalue of $\text{rev } P(\lambda)$.

Notice that when M is nonsingular all linearizations are *strong*. The reason is that in this case the eigenvalues of $P(\lambda)$ and $\text{rev } P(\lambda)$ (and their linearizations) are the reciprocals of each other.

Whether M is singular or nonsingular (and $P(\lambda)$ is regular), the linear pencil

$$(5) \quad \lambda \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} + \begin{bmatrix} D & K \\ K & 0 \end{bmatrix},$$

preserves symmetry and provides a strong linearization if and only if $\det K \neq 0$. Indeed, this pencil is a member of a family of strong Hermitian linearizations of $P(\lambda)$. It is shown in [8, Sec. 3] (see also Table 1 of [15]) that, for real parameters $a_1, a_2 \in \mathbb{R}$, not both zero, with $-\frac{a_2}{a_1} \notin \sigma(P)$, every pencil of the form

$$(6) \quad L(\lambda) := \lambda \begin{bmatrix} a_1M & a_2M \\ a_2M & a_2D - a_1K \end{bmatrix} + \begin{bmatrix} a_1D - a_2M & a_1K \\ a_1K & a_2K \end{bmatrix},$$

is either a *real symmetric*, or a *complex Hermitian*, strong linearization of $P(\lambda)$.

In particular, the pencil (5) is obtained with $a_1 = 1$, $a_2 = 0$, in which case the condition $-\frac{a_2}{a_1} = 0 \notin \sigma(P)$ is equivalent to $\det K \neq 0$. Observe also that the choice $a_1 = 0$, $a_2 = 1$ determines a real symmetric or complex Hermitian linearization when infinity is not an eigenvalue of $P(\lambda)$; that is, when $\det M \neq 0$.

The family of linearizations (6) will play a key role in this paper and so a specific notation will be used for them according as $P(\lambda)$ is real symmetric or (more generally) complex Hermitian:

(i) If M, D, K are symmetric matrices in $\mathbb{R}^{n \times n}$ and $a_1, a_2 \in \mathbb{R}$ then

$$(7) \quad A_R := \begin{bmatrix} a_1 M & a_2 M \\ a_2 M & a_2 D - a_1 K \end{bmatrix}, \quad B_R := \begin{bmatrix} a_1 D - a_2 M & a_1 K \\ a_1 K & a_2 K \end{bmatrix}.$$

(ii) If M, D, K are Hermitian matrices in $\mathbb{C}^{n \times n}$ and $a_1, a_2 \in \mathbb{R}$ then

$$(8) \quad A_C := \begin{bmatrix} a_1 M & a_2 M \\ a_2 M & a_2 D - a_1 K \end{bmatrix}, \quad B_C := \begin{bmatrix} a_1 D - a_2 M & a_1 K \\ a_1 K & a_2 K \end{bmatrix}.$$

Assuming that a_1 and a_2 are not both zero and $-\frac{a_2}{a_1} \notin \sigma(P)$, $\lambda A_R + B_R$ or $\lambda A_C + B_C$ are strong linearizations of $P(\lambda)$ according as $\mathbb{F} = \mathbb{R}$ or \mathbb{C} . Then, taking into account (3), $P(\lambda)$ is regular if and only if $\lambda A_R + B_R$ (or $\lambda A_C + B_C$, resp.) is regular.

The important role played by these linearizations is apparent from (3) and the easily verified relation that will be important in Sections 8 and 9: For any $\lambda \in \mathbb{C}$

$$(9) \quad (\lambda A_C + B_C) \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix} = \begin{bmatrix} a_1 I_n \\ a_2 I_n \end{bmatrix} P(\lambda),$$

(and similarly for $\lambda A_R + B_R$).

More generally, a computation shows that if $a_1 \neq 0$ or $a_2 \neq 0$ and $-\frac{a_2}{a_1} \notin \sigma(P)$ then, for any pair of matrices $Y_1, Y_2 \in \mathbb{C}^{n \times n}$ for which $\lambda Y_2 - Y_1$ is unimodular, there are linear pencils $X_1(\lambda)$ and $X_2(\lambda)$ such that

$$(10) \quad F(\lambda) := \begin{bmatrix} \lambda I_n & Y_1 \\ I_n & Y_2 \end{bmatrix}, \quad E(\lambda) := \begin{bmatrix} a_1 I_n & X_1(\lambda) \\ a_2 I_n & X_2(\lambda) \end{bmatrix},$$

are unimodular, and

$$(11) \quad (\lambda A_C + B_C) F(\lambda) = E(\lambda) \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_n \end{bmatrix}.$$

This confirms that, if $-\frac{a_2}{a_1} \notin \sigma(P)$, then $\lambda A_C + B_C$ is a linearization of $P(\lambda)$, and similarly for $\lambda A_R + B_R$.

To be precise, in the definition of $E(\lambda)$ in (10), we have linear pencils:

$$\begin{aligned} X_1(\lambda) &= \lambda M(a_1 Y_1 + a_2 Y_2) + a_1 D Y_1 + (a_1 K Y_2 - a_2 M Y_1), \\ X_2(\lambda) &= \lambda \{a_2 D Y_2 - (a_1 K Y_2 - a_2 M Y_1)\} + K(a_1 Y_1 + a_2 Y_2). \end{aligned}$$

And $E(\lambda)$ is unimodular if and only if $\lambda Y_2 - Y_1$ is unimodular and $-\frac{a_2}{a_1} \notin \sigma(P)$ (provided that $a_1 \neq 0$ or $a_2 \neq 0$).

1.1. Canonical forms. The spectral theory of real symmetric or complex Hermitian quadratic matrix polynomials is developed in the literature mostly through canonical forms of their linearizations. This is

what we call the **algebraic approach**. In this approach, there are three interconnected ways of using the linearizations of a given quadratic system with the common goal of reducing them to canonical form by, either, *congruence* or *similarity* transformations which are independent of λ .

The first method consists of reducing the pencil $\lambda A_C + B_C$, or $\lambda A_R + B_R$, by congruence transformations. This method is based on the theory of canonical forms for linear pencils which can be found, for example, in [12, 13]. It will be reviewed in Section 4.

The second method uses the notion of *self-adjoint standard triples* as defined in [14]. These triples can be reduced by similarity to self-adjoint Jordan triples which provide complete spectral information, including *Jordan chains*, concerning the given real symmetric or complex Hermitian quadratic systems. In this method, the condition $\det M \neq 0$ will be required and is discussed in Section 5.

The third method to study the spectral properties of $P(\lambda)$ of (1) is to apply unitary similarity to pairs of matrices (T, H) , where T is a linearization of $P(\lambda)$ and H is an invertible Hermitian (symmetric in the real case) matrix for which T is H -self-adjoint. This theory also requires M nonsingular, and it is reviewed in Section 6.

Another way to study the spectral properties of real symmetric or complex Hermitian quadratic systems is through what we call the **analytic approach**. This is based on the important Rellich theorem (Theorem 2 below), which uses direct reduction of $P(\lambda)$ itself by λ -dependent unitary similarity. This is developed in Sections 2 and 3.

All of the lines of attack (both algebraic and analytic) reveal important properties of the spectrum of $P(\lambda)$. However, reduction by congruence may have the advantages that symmetries of the system are preserved, and a singular leading coefficient may be admissible. In reduction by similarity transformations some symmetries are lost and invertibility of M will be required.

Connections among all these theories are scattered in the literature (including [6, 7] and [14]). One of the main goals of this paper is to bring them together and clarify their connections.

In all cases the notion of a *sign characteristic* associated with the real eigenvalues appears as an invariant for the considered transformation. A thorough analysis of the sign characteristic of Hermitian matrix polynomials was carried out in [16] where applications in Control and Perturbation theory are discussed in Section 1 (see also [2]). It is important to note that, as pointed out in [16], the “sign characteristics” attached to real eigenvalues in different canonical forms may not coincide.

Connections between the sign characteristics appearing in the different canonical forms will be analyzed. This topic is strongly related to the problem of characterizing the strong linearizations of $P(\lambda)$ which preserve the sign characteristic. This question was tackled in [3] (see also [2]). In most cases, including these references, the adopted definition of sign characteristic is *basis-dependent* and this is the reason why canonical forms with different sign characteristics appear in the literature. One of the goals of this paper is to clarify this matter by referring the reader to the only sign characteristic which is *basis-free*: that derived from Rellich’s theorem.

The connection among all analyzed sign characteristics will be clarified in Section 7. The relationship between the eigenvalue functions of quadratic systems and their linearizations of (6) will be investigated in Section 8 where some interlacing inequalities will be disclosed. Section 9 provides, for any given real symmetric or complex Hermitian quadratic matrix polynomial (with possibly singular leading coefficient) strong

linearizations preserving both the symmetry and the sign characteristic, including that of the eigenvalue at infinity. In fact, a characterization of the strong linearizations of (6) which preserve the sign characteristic of such quadratic systems will be given. To our knowledge, these are new result that could serve as a basis for future investigation. When possible, our main results are illustrated with examples and graphics.

A word on notation: we aim to develop the theory for both real symmetric and complex Hermitian systems simultaneously. In order to simplify the notation, unless we must specify that the underlying field is \mathbb{R} or \mathbb{C} , we will call a matrix Hermitian when it is either complex Hermitian ($\mathbb{F} = \mathbb{C}$) or real symmetric ($\mathbb{F} = \mathbb{R}$). So, $A \in \mathbb{F}^{n \times n}$ will be an Hermitian matrix if $A = A^*$ meaning that $a_{ij} = \bar{a}_{ji}$ when $\mathbb{F} = \mathbb{C}$, and $a_{ij} = a_{ji}$ when $\mathbb{F} = \mathbb{R}$. Also, we will say that $P(\lambda)$ is a *quadratic Hermitian system* if its coefficient matrices are Hermitian.

Most results in this paper hold for arbitrary matrix polynomials of any degree with real symmetric or complex Hermitian coefficients. However, quadratic systems arise most frequently and they are worthy of special attention.

2. Parameter dependent unitary similarity and equivalence. An informative geometric perspective on many problems in this context is obtained if we recall that, when M, D , and K are Hermitian, $P(\lambda)$ of (1) is an Hermitian matrix-valued function of $\lambda \in \mathbb{R}$. This is the approach taken in this section where the developments of [16] and [5, 6] or [7] are adapted to quadratic matrix polynomials. The starting point is a theorem of Rellich [17], which will be presented in a general form following [4], and will then be applied to Hermitian quadratic matrix polynomials.

We adopt some conventions of [4]:

- (a) A set $\Omega \subset \mathbb{C}$ is said to be a *domain* if it is nonempty, open, and connected.
- (b) For a domain $\Omega \subset \mathbb{C}$, $H(\Omega)$ denotes the ring of functions f defined on Ω such that, for each $\zeta \in \Omega$, there is an open neighborhood \mathcal{V}_ζ of ζ on which f is analytic. Also, $H(\Omega)^{m \times n}$ will denote the set of $m \times n$ matrices with elements in Ω .
- (c) If $\Omega = \{\zeta\}$ then H_ζ stands for $H(\{\zeta\})$.
- (d) A domain $\Omega \subset \mathbb{C}$ is said to be \mathbb{R} -*symmetric* if $\bar{\Omega} = \Omega$ where $\bar{\Omega} = \{z \in \mathbb{C} : \bar{z} \in \Omega\}$.

Assume that $\Omega \subset \mathbb{C}$ is \mathbb{R} -symmetric. Analytic matrix functions $A(z), U(z) \in H(\Omega)^{n \times n}$ are said to be *Hermitian analytic* on Ω and *unitary analytic* on Ω , respectively, if

$$A(z) = A(\bar{z})^*, \quad U(z)^{-1} = U(\bar{z})^*, \quad \text{for all } z \in \Omega.$$

Note that if Ω is \mathbb{R} -symmetric, $f \in H(\Omega)$ and $g(z) = \overline{f(\bar{z})}$ for all $z \in \Omega$ then $g \in H(\Omega)$.

THEOREM 2 (Rellich's Theorem: [17], [4, Theorem 4.17.2]). *Let $\Omega \subset \mathbb{C}$ be an \mathbb{R} -symmetric domain, let $A(z) \in H(\Omega)^{n \times n}$ be an Hermitian analytic matrix function on Ω , and let J be a real open interval in Ω . Then there exists an \mathbb{R} -symmetric domain Ω_1 such that $J \subset \Omega_1 \subset \Omega$ and the following properties hold:*

1. *The characteristic polynomial of $A(z)$ "splits" in $H(\Omega_1)$. That is, for any $\mu \in \mathbb{C}$, $z \in \Omega_1$*

$$\det(\mu I_n - A(z)) = \prod_{i=1}^n (\mu - \mu_i(z)),$$

where $\mu_1(z), \dots, \mu_n(z) \in H(\Omega_1)$.

2. *There exists a unitary analytic matrix function $U(z) \in H(\Omega_1)^{n \times n}$ such that, for any $z \in \Omega_1$,*

$$(12) \quad A(z) = U(z) \text{Diag}[\mu_1(z), \dots, \mu_n(z)]U(\bar{z})^*.$$

Theorem 2 says that if $A(z)$ is an Hermitian analytic matrix defined on an \mathbb{R} -symmetric domain Ω , then it admits n eigenvalue functions $\mu_1(z), \dots, \mu_n(z)$ which are analytic in an \mathbb{R} -symmetric subdomain $\Omega_1 \subset \Omega$ with $J \subset \Omega_1$. These eigenvalue functions are real when $z \in J$ because $A(z) \in \mathbb{C}^{n \times n}$ is Hermitian for each $z \in J$ and $\mu_1(z), \dots, \mu_n(z)$ are its eigenvalues.

The theorem also implies that, for $z \in \Omega_1$, there are corresponding (mutually orthogonal) analytic “eigenvector functions” $x_j(z)$ with values in \mathbb{C}^n . In fact, if $x_j(z)$ is the j th column of $U(z)$, then

$$(13) \quad A(z)x_j(z) = \mu_j(z) x_j(z), \quad j = 1, 2, \dots, n, \quad z \in \Omega_1,$$

and the orthogonality conditions hold:

$$(14) \quad x_j(\bar{z})^* x_k(z) = \delta_{jk}, \quad j, k = 1, 2, \dots, n, \quad z \in \Omega_1,$$

where δ_{jk} is the Kronecker delta.

More generally, and in contrast with Theorem 2, basic spectral data for matrix functions in $H(\Omega)$ (with or without symmetries) can be revealed by applying *equivalence* transformations to $A(z)$. In fact, for any nonempty connected set $\Omega \subset \mathbb{C}$ (not necessarily a domain), $H(\Omega)$ is an elementary divisor domain ([4, Theorem 1.5.3]). Commutative rings of this kind were introduced by Kaplansky in [10] and have the important property that any *possibly rectangular* matrix is equivalent to a matrix whose only nonzero elements are in the first diagonal positions. This “diagonal” matrix is called the *Smith normal form* of $A(z)$ ([4, Theorem 1.14.1] and [20]).

Specifically, let $A(z), B(z) \in H(\Omega)^{m \times n}$ then $A(z)$ and $B(z)$ are said to be *equivalent* if there are invertible matrices $F_1(z) \in H(\Omega)^{m \times m}$ and $F_2(z) \in H(\Omega)^{n \times n}$ (analytic matrix functions whose determinants are nonzero on Ω) such that $B(z) = F_1(z)A(z)F_2(z)$. Then, we have

THEOREM 3. *Let $\Omega \subset \mathbb{C}$ be a nonempty connected set and $A(z) \in H(\Omega)^{m \times n}$. Then $A(z)$ is equivalent to a matrix with the following diagonal form:*

$$(15) \quad D(z) := \begin{bmatrix} \text{Diag}[d_1(z), \dots, d_\rho(z)] & 0 \\ 0 & 0 \end{bmatrix},$$

where $\rho = \text{rank } A(z)$ is the (normal) rank of $A(z)$ (i.e., the order of the biggest nonzero minor of $A(z)$), $d_i(z) \in H(\Omega)$ for $i = 1, \dots, \rho$ and $d_i(z) \mid d_{i+1}(z)$, for $i = 1, \dots, \rho - 1$.

The analytic functions $d_1(z), \dots, d_\rho(z)$ are the *invariant factors* of $A(z)$ and they are uniquely determined by $A(z)$ up to products by units of $H(\Omega)$ (i.e., analytic functions which have no zeros in Ω). The diagonal matrix function, $D(z)$ of (15), is called the *Smith normal form* of $A(z)$.

If $z_0 \in \Omega$ then for $i = 1, \dots, \rho$ there is a nonnegative integer $m_i \geq 0$ such that $d_i(z) = (z - z_0)^{m_i} c_i(z)$ with $c_i(z) \in H(\Omega)$ and $c_i(z_0) \neq 0$. If $m_i = 0$, then $d_i(z_0) \neq 0$ and if $m_i > 0$ then $\text{rank } A(z_0) < \rho$ and so z_0 is an *eigenvalue* of $A(z)$ ([16, Section 2])¹. In particular, if $\rho = m = n$, then $\det A(z) = \prod_{i=1}^n d_i(z)$ up to a product by a unit of $H(\Omega)$, and so $\det A(z_0) = 0$ if and only if $d_i(z_0) = 0$ for some $i = 1, \dots, n$. The nonnegative integers $0 \leq m_1 \leq \dots \leq m_n$ different from 0 are called the *partial multiplicities* of $A(z)$ for its eigenvalue z_0 . For notational simplicity, we will admit zeros as partial multiplicities when needed.

¹If Ω is a domain then the set of eigenvalues of $A(z)$ is a countable set in Ω whose accumulation points are on the boundary of Ω .

Now, it follows from $A(z) \in H(\Omega)^{m \times n}$ that $A(z) \in H_{z_0}^{m \times n}$ for all $z_0 \in \Omega$; i.e. $A(z)$ is an analytic matrix function at z_0 . The Smith normal form of $A(z) \in H_{z_0}^{m \times n}$ is particularly simple (see [7, Theorem A.6.4] for example). In fact, if, as above, we write $d_j(z) = (z - z_0)^{m_j} c_j(z)$ with $c_j(z_0) \neq 0$, $j = 1, \dots, \rho$, then

$$G(z) := \text{Diag}[c_1(z), \dots, c_\rho(z), 1, \dots, 1],$$

is invertible in $H_{z_0}^{m \times m}$. Since $A(z)$ and $D(z)$ are equivalent in $H_{z_0}^{m \times m}$, we conclude that

$$G_2(z)^{-1}D(z) = \begin{bmatrix} \text{Diag}[(z - z_0)^{m_1}, \dots, (z - z_0)^{m_\rho}] & 0 \\ 0 & 0 \end{bmatrix},$$

is equivalent to $A(z)$ in $H_{z_0}^{m \times n}$. Now, the divisibility condition $d_1(z) \mid d_2(z) \mid \dots \mid d_\rho(z)$ implies $0 \leq m_1 \leq \dots \leq m_\rho$. Hence $G_2(z)^{-1}D(z) \in H_{z_0}^{m \times n}$ is a matrix in Smith normal form; it is called the *local Smith form* of $A(z)$ at z_0 . It is completely determined by the rank of $A(z)$ and its partial multiplicities (including zeros) at z_0 .

When $A(z)$ is an Hermitian analytic matrix function, the partial multiplicities of its eigenvalues can be obtained from the eigenvalue functions. Specifically, let $\Omega \subset \mathbb{C}$ be an \mathbb{R} -symmetric domain and let $A(z) \in H(\Omega)^{n \times n}$ be an Hermitian analytic matrix. According to Rellich's theorem, there is an \mathbb{R} -symmetric domain Ω_1 and a unitary analytic matrix $U(z) \in H(\Omega_1)^{n \times n}$ such that (12) holds. If $\text{rank } A(z) = \rho < n$ and $\mu_1(z), \dots, \mu_n(z)$ are its eigenvalue functions, we can assume without loss of generality that, for $i = 1, 2, \dots, \rho$, $\mu_i(z)$ is not the zero function in Ω_1 and $\mu_i(z) = 0$ for $i = \rho + 1, \dots, n$ and all $z \in \Omega_1$.

PROPOSITION 4. *With the above notation, let $z_0 \in \Omega_1$ be an eigenvalue of $A(z) \in H(\Omega)^{n \times n}$, let $\rho = \text{rank } A(z)$ and let (m_1, \dots, m_ρ) be the partial multiplicities (including zeros) for the eigenvalue z_0 . Then there is a permutation (k_1, \dots, k_ρ) of $(1, \dots, \rho)$ such that*

$$\mu_{k_j}(z) = (z - z_0)^{m_j} \nu_{k_j}(z), \quad j = 1, \dots, \rho, \quad \nu_{k_j}(z_0) \neq 0,$$

and $\nu_{k_j}(z)$ is analytic in a neighbourhood of z_0 .

Proof. It follows from item 2 of Theorem 2 that, for $z \in \Omega_1$, $A(z)$ is equivalent to $M(z) := \text{Diag}[\mu_1(z), \dots, \mu_n(z)]$ in $H(\Omega_1)^{n \times n}$. Write

$$\mu_j(z) = (z - z_0)^{s_j} \nu_j(z), \quad \nu_j(z_0) \neq 0, \quad j = 1, \dots, \rho.$$

As seen above, $G(z) = \text{Diag}[\nu_1(z), \dots, \nu_\rho(z), 1, \dots, 1]$ is invertible in $H_{z_0}^{n \times n}$ and

$$G(z)^{-1}M(z) = \text{Diag}[(z - z_0)^{s_1}, \dots, (z - z_0)^{s_\rho}, 0, \dots, 0].$$

Now, there is an $n \times n$ permutation matrix P such that

$$S(z) = PG(z)^{-1}M(z)P^T = \text{Diag}[(z - z_0)^{s_{k_1}}, \dots, (z - z_0)^{s_{k_\rho}}, 0, \dots, 0],$$

where $s_{k_1} \leq \dots \leq s_{k_\rho}$ is the sequence (s_1, \dots, s_ρ) arranged in nondecreasing order. It follows from $(z - z_0)^{s_{k_1}} \mid \dots \mid (z - z_0)^{s_{k_\rho}}$ that $S(z)$ is the local Smith form of $A(z)$ at z_0 . Hence, $s_{k_j} = m_j$ for $j = 1, \dots, \rho$ and the Proposition follows. \square

2.1. The quadratic case. Now we apply the Rellich technique to the Hermitian quadratic matrix polynomial of (1). Here, attention is confined to the special case $\lambda \in \mathbb{R}$.

THEOREM 5. If M , D , and K of (1) are Hermitian matrices then the following properties hold:

(i) There are n real-valued analytic functions $v_1(\lambda), \dots, v_n(\lambda)$, $\lambda \in \mathbb{R}$, such that

$$\det(\mu I_n - P(\lambda)) = \prod_{i=1}^n (\mu - v_i(\lambda)).$$

(ii) There is an $n \times n$ matrix-valued analytic function $U(\lambda)$ for which $U(\lambda)^* U(\lambda) = I_n$, and

$$(16) \quad P(\lambda) = U(\lambda) \text{Diag}[v_1(\lambda), v_2(\lambda), \dots, v_n(\lambda)] U(\lambda)^*, \quad \lambda \in \mathbb{R}.$$

(iii) If $\lambda_1 < \lambda_2 < \dots < \lambda_r$ are the real eigenvalues of $P(\lambda)$ (if any) and for $1 \leq j \leq r$, $m_{j1} \leq \dots \leq m_{jg_j}$ are the partial multiplicities of the eigenvalue λ_j , then there are g_j eigenvalue functions $v_{i_1}(\lambda), \dots, v_{i_{g_j}}(\lambda)$ such that for $k = 1, \dots, g_j$ and $j = 1, \dots, r$

$$(17) \quad v_{i_k}(\lambda) = (\lambda - \lambda_j)^{m_{jk}} \nu_{ji_k}(\lambda), \quad \nu_{ji_k}(\lambda_j) \neq 0, \quad \lambda \in \mathbb{R}.$$

Proof. Since M , D and K are $n \times n$ Hermitian matrices, $P(\lambda) = M\lambda^2 + D\lambda + K$ is an Hermitian analytic matrix for $\lambda \in \mathbb{R}$. More generally, we will prove properties (i) and (ii) for any matrix $A(\lambda) \in H(J)^{n \times n}$ which is Hermitian analytic in an open interval $J \subseteq \mathbb{R}$. In fact, for such a matrix we have $A(\lambda)^* = A(\lambda)$ for $\lambda \in J$ and for each $\lambda_0 \in J$ there is an open ball $\mathcal{B}_{\lambda_0} \subset \mathbb{C}$ with $A(z)$ analytic for all $z \in \mathcal{B}_{\lambda_0}$. Let $\Omega = \cup_{\lambda_0 \in J} \mathcal{B}_{\lambda_0}$. Then Ω is open, connected, \mathbb{R} -symmetric, $J \subset \Omega$ and $A(z), A(\bar{z})^* \in H(\Omega)^{n \times n}$.

Now let us show that $A(z) = A(\bar{z})^*$; i.e. $A(z)$ is Hermitian analytic in Ω . For $1 \leq i \leq j \leq n$ let $b_{ij}(z) = a_{ij}(z) - \bar{a}_{ji}(\bar{z})$, $z \in \Omega$. Then, $b_{ij}(z)$ is an analytic function in Ω and so ([18, Theorem 10.18]) its set of zeros is either Ω or a countable set with no limit point in Ω . Since $b_{ij}(z) = 0$ for all $z \in J$, we must have $b_{ij}(z) = 0$ for all $z \in \Omega$ as claimed.

Properties (i) and (ii) for $A(\lambda)$ follow at once from Theorem 2. Also, if λ_j is a real eigenvalue of $A(\lambda)$ and $m_1 \leq \dots \leq m_g$ are its partial multiplicities, then by Proposition 4, there are g eigenvalue functions $v_{i_1}(\lambda), \dots, v_{i_g}(\lambda)$ such that for $k = 1, \dots, g$

$$(18) \quad v_{i_k}(\lambda) = (\lambda - \lambda_j)^{m_{jk}} \nu_{ji_k}(\lambda), \quad \nu_{ji_k}(\lambda_j) \neq 0.$$

Since $P(\lambda)$ is a matrix polynomial, it has a finite number of real eigenvalues. Property (iii) follows by applying (18) to each of them. \square

REMARK 6. The proof of our Theorem 2 in [4] follows the ideas of [11, Chapter 2, Theorem 6.1] and [20]. In both cases, Theorem 2 is first proved for real $\lambda \in J \subset \Omega$. (See [6, Theorem S6.3] for a different, but closely related, proof.) Analytic continuation is then used in [4] and [11] to extend the result to an open connected set $\Omega_1 \subset \Omega$ containing J . Using that approach, the proof that $P(\lambda)$ satisfies properties (i) and (ii) of Theorem 5 is straightforward. Proposition 4 is still needed to prove property (iii). \square

DEFINITION 7. (a) If $\lambda_1, \lambda_2, \dots, \lambda_r$ are the distinct eigenvalues of $P(\lambda)$ and for $j = 1, \dots, r$, $m_{j1} \leq \dots \leq m_{jg_j}$ are the partial multiplicities of the eigenvalue λ_j , then $\sum_{k=1}^{g_j} m_{jk}$ is the algebraic multiplicity of λ_j and g_j is its geometric multiplicity. (b) For $j = 1, \dots, r$ and $i = 1, \dots, g_j$, $(\lambda - \lambda_j)^{m_{ij}}$ is said to be an elementary divisor of $P(\lambda)$ for the eigenvalue λ_j with partial multiplicity m_{ij} .

Now we introduce the concept of a *sign characteristic* for the real eigenvalues of $P(\lambda)$.

DEFINITION 8. If $\lambda_1 < \lambda_2 < \dots < \lambda_r$ are the real eigenvalues of $P(\lambda)$, the family of +1's and -1's determined in equation (17) by the sequence $(\text{sign}(\nu_{ji_k}(\lambda_j)), 1 \leq i_1 < i_2 < \dots < i_{g_j} \leq n)$ for each $j = 1, \dots, r$ is called the sign characteristic of the real eigenvalue λ_j . It is uniquely determined for each λ_j up to a permutation of signs associated with equal elementary divisors.

In other words, the sign characteristic of λ_j is the *family* of signs of the m_{jk} th derivatives of $v_{i_k}(\lambda)$ evaluated at λ_j : $\text{sign } v_{i_k}^{(m_{jk})}(\lambda_j)$, $k = 1, \dots, g_j$. Note that, for a *semisimple* real eigenvalue λ_j , the sign characteristic has a geometric interpretation in terms of the collection of positive, or negative, slopes of eigenvalue functions $v_{i_k}(\lambda)$, which vanish at λ_j (see (18)).

DEFINITION 9. We will use the term sign characteristic of $P(\lambda)$ to refer to the sequence of sign characteristics of all real eigenvalues of $P(\lambda)$ in some order (for instance, arranging the distinct real eigenvalues in increasing order).

With some abuse of language, we will often refer to **the** sign characteristic of $P(\lambda)$, although it is uniquely determined only up to reordering of its real eigenvalues and the signs associated with equal elementary divisors at each real eigenvalue.

More generally, the sign characteristic for the eigenvalues of analytic Hermitian matrix functions $A(\lambda)$, $\lambda \in \mathbb{R}$, can be defined in the same way using (18) (see [16, Def. 2.3]. See also [6, Sec. 12.3] or [7, Sec. 12.5] for a different but equivalent approach).

3. Examples and discussion.

EXAMPLE 10. (a) Figure 1 depicts the graphics of the eigenvalue functions of the real symmetric matrix polynomial

$$(19) \quad P(\lambda) = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \lambda^2 + \begin{bmatrix} 1 & 4 & 1 \\ 4 & -3 & 4 \\ 1 & 4 & 1 \end{bmatrix} \lambda + \begin{bmatrix} 7 & 8 & 7 \\ 8 & -6 & 8 \\ 7 & 8 & 7 \end{bmatrix},$$

whose real eigenvalues are the points where the eigenvalue functions cut or touch the real axis. In fact (to four decimal places):

$$\sigma(P) = \{-1.2496, 0, 3.7220, -2.0695 \pm 2.8947i\}.$$

There are two real eigenvalues of partial multiplicity 1 (simple eigenvalues), namely, -1.2496 and 3.7220 . The eigenvalue 0 has algebraic multiplicity 2 and geometric multiplicity 1. So the elementary divisor associated with the eigenvalue 0 is λ^2 . The sign characteristic of the eigenvalue -1.2496 is -1 and that of 3.7220 is $+1$.

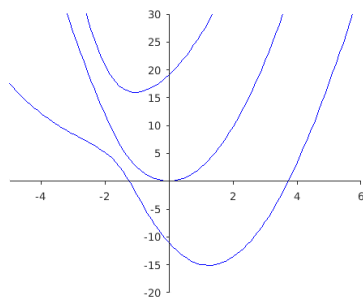


FIGURE 1. Eigenvalue functions of the quadratic matrix polynomial of Example 10 (a).

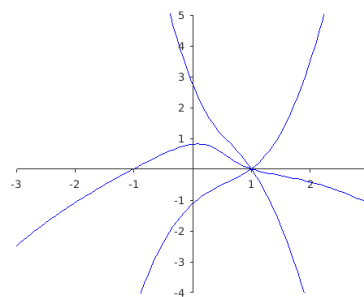


FIGURE 2. Eigenvalue functions of the quadratic matrix polynomial of Example 10 (b).

Also, the eigenvalue function corresponding to the eigenvalue 0 can be written as $v(\lambda) = \lambda^2\nu(\lambda)$ with $\nu(0) \neq 0$. The graph of this curve in Figure 1 shows that $\nu(0) > 0$ and so (see Definition 8 and (17)) the sign characteristic of eigenvalue 0 and partial multiplicity 2 is also +1. Therefore, the sign characteristic of $P(\lambda)$ is $((-1), (+1), (+1))$.

(b) Consider the following real symmetric quadratic matrix polynomial

$$(20) \quad P(\lambda) = \begin{bmatrix} -0.5 & 2 & 0.5 \\ 2 & 1.4 & -0.5 \\ 0.5 & -0.5 & -0.5 \end{bmatrix} \lambda^2 + \begin{bmatrix} -1.0 & -3.7 & 0 \\ -3.7 & -1.8 & 0 \\ 0 & 0 & 0 \end{bmatrix} \lambda + \begin{bmatrix} 1.5 & 1.7 & -0.5 \\ 1.7 & 0.4 & 0.5 \\ -0.5 & 0.5 & 0.5 \end{bmatrix}.$$

Its elementary divisors are $(\lambda - 1), (\lambda - 1), (\lambda - 1), (\lambda + 1)$ and the eigenvalue functions are shown in Figure 2. The sign characteristic for the simple eigenvalue -1 is +1 and that of the semisimple eigenvalue $+1$ is $(-1, -1, +1)$ (but also $(-1, +1, -1)$ or $(+1, -1, -1)$). There is one conjugate pair of non-real eigenvalues. A sign characteristic for $P(\lambda)$ is, for instance, $((+1), (-1, +1, -1))$. \square

We see that the real zeros, λ_j , (if any) of the real eigenvalue functions $v_j(\lambda)$ determine the real eigenvalues of $P(\lambda)$, and the nature of such a zero λ_j is dependent on the *partial multiplicities* and the associated *sign characteristics*, $\eta_j = \pm 1$; one for each partial multiplicity (or, equivalently, each elementary divisor). With these concepts we can add some geometric intuition to the discussion of canonical forms.

It is important to note that this technique does not require the invertibility of M . In particular, the whole argument holds true when $M = 0$; that is to say, for Hermitian pencils. Thus, if $A, B \in \mathbb{F}^{n \times n}$ are Hermitian and $L(\lambda) = \lambda A + B$, we can apply Rellich's theorem (Theorem 2) to $L(\lambda)$, and so we have for $L(\lambda)$ the notions of eigenvalue functions and sign characteristics for each eigenvalue of Definition 8 and those of partial, algebraic and geometric multiplicities of Definition 7. (See also [7, Theorem 5.11.1].)

Figure 3 shows the eigenvalue functions of the Hermitian linearization of (7) for the values $a_1 = 0$ and $a_2 = 1$ corresponding to the symmetric quadratic matrix polynomial of (19). Notice that this is a linearization of $P(\lambda)$ of the form (6) because, as $\det M \neq 0$, infinity is not an eigenvalue of $P(\lambda)$. Observe also that the predicted eigenvalues are consistent with those of Figure 1. Moreover, the sign characteristics of the real eigenvalues are preserved.

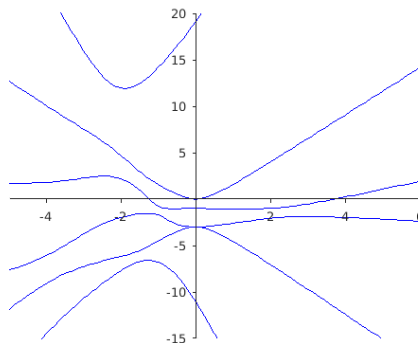


FIGURE 3. Eigenvalue functions of $L(\lambda) = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$ where M, D , and K are the coefficients of the quadratic system of Example 10(a).

TABLE 1

Sign characteristics of the eigenvalue $\mu_0 = \frac{1}{\lambda_0}$ of $\text{rev } P(\lambda)$ when λ_0 is the corresponding eigenvalue of $P(\lambda) = M\lambda^2 + D\lambda + K$ with sign characteristic η_0 for the partial multiplicity m_0 .

| | $\mu_0 = \frac{1}{\lambda_0}$ | $\mu_0 = 0$ | $\mu_0 = \infty$ |
|------------|-------------------------------|-------------|------------------|
| m_0 even | η_0 | $-\eta_0$ | $-\eta_0$ |
| m_0 odd | $-\eta_0$ | $-\eta_0$ | $-\eta_0$ |

TABLE 2

Sign characteristics of the eigenvalue $\mu_0 = \frac{1}{\lambda_0}$ of $\text{rev } L(\lambda)$ when λ_0 is the corresponding eigenvalue of $L(\lambda) = A\lambda + B$ with sign characteristic η_0 for the partial multiplicity m_0 .

| | $\mu_0 = \frac{1}{\lambda_0}$ | $\mu_0 = 0$ | $\mu_0 = \infty$ |
|------------|--|-------------|------------------|
| m_0 even | $\text{sign}\left(\frac{1}{\lambda_0}\right)\eta_0$ | $-\eta_0$ | $-\eta_0$ |
| m_0 odd | $-\text{sign}\left(\frac{1}{\lambda_0}\right)\eta_0$ | $-\eta_0$ | $-\eta_0$ |

On the other hand, if M is not the zero matrix but singular ($\det M = 0$), then $P(\lambda) = M\lambda^2 + D\lambda + K$ has an eigenvalue at infinity (or infinity as an eigenvalue). The same is true for any linear matrix pencil $L(\lambda) = \lambda A + B$ if $A \neq 0$ and $\det A = 0$. More generally, the partial multiplicities and the algebraic and geometric multiplicities of the eigenvalue at infinity are defined to be those of 0 as an eigenvalue of the reverse polynomial $\text{rev } P(\lambda) = K\lambda^2 + D\lambda + M$. Also, the elementary divisors of $P(\lambda)$ at infinity are those of $\text{rev } P(\lambda)$ at 0.

The sign characteristic of the eigenvalue at infinity has been studied in [16]. It is defined to be the sign characteristic of the eigenvalue 0 of $S(\lambda) = -\text{rev } P(\lambda)$ (notice the minus sign preceding $\text{rev } P(\lambda)$). With this definition, the relationship between the sign characteristics of the real eigenvalues of $P(\lambda) = M\lambda^2 + D\lambda + K$ and of $\text{rev } P(\lambda) = K\lambda^2 + D\lambda + M$ is illustrated in Example 3.6 of [16]. The following relevant results, which will be used later on, have been extracted from that Example:

- (a) Assume that λ_0 is a finite or infinite eigenvalue of $P(\lambda)$. Let m_0 be one of its partial multiplicities and let η_0 be the sign characteristic of λ_0 associated with m_0 . Then the sign characteristics for the partial multiplicity m_0 of the corresponding eigenvalue μ_0 of $\text{rev } P(\lambda)$ are shown in Table 1 (recall that $\lambda_0 \neq 0, \infty \Leftrightarrow \mu_0 = \frac{1}{\lambda_0}$, $\lambda_0 = 0 \Leftrightarrow \mu_0 = \infty$, and $\lambda_0 = \infty \Leftrightarrow \mu_0 = 0$).
- (b) Similarly, the relationship between the sign characteristics of the real eigenvalues of $L(\lambda) = A\lambda + B$ and $\text{rev } L(\lambda) = B\lambda + A$ is shown in Table 2.

EXAMPLE 11. The curves of Figure 4 are the eigenvalue functions of the reversal of $P(\lambda)$ in (19) – Example 10; that is,

$$(21) \quad \text{rev } P(\lambda) = \begin{bmatrix} 7 & 8 & 7 \\ 8 & -6 & 8 \\ 7 & 8 & 7 \end{bmatrix} \lambda^2 + \begin{bmatrix} 1 & 4 & 1 \\ 4 & -3 & 4 \\ 1 & 4 & 1 \end{bmatrix} \lambda + \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

and should be compared with those of Figure 1.

Recall that $\sigma(P) = \{0, -2.0695 \pm 2.8947i, -1.2496, 3.7220\}$ so that, approximately, $\sigma(\text{rev } P) = \{\infty, -0.1634 \pm 0.2286i, -0.8003, 0.2687\}$. Recall also that the sign characteristic of -1.2496 is -1 , that of 3.7220 is $+1$ and

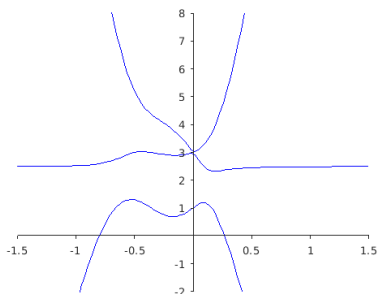


FIGURE 4. Eigenvalue functions of the reversal of the quadratic system of (19).

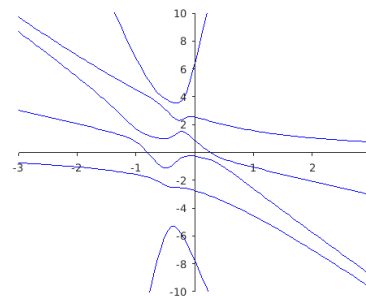


FIGURE 5. Eigenvalue functions of the reversal of the linear system of Figure 3.

that of the eigenvalue 0 and partial multiplicity 2 is also +1. By definition, the eigenvalue infinity has partial multiplicity 2. The partial multiplicities of the nonzero finite eigenvalues are all equal to 1. It follows from the graphics of the eigenvalue functions of Figure 4 that the sign characteristic of -0.8003 is +1 and that of 0.2687 is -1 . This is consistent with the sign characteristics in Table 1. The sign characteristic of the eigenvalue of $\text{rev}P$ at infinity and partial multiplicity 2 is not apparent in Figure 4. However, it follows from the sign characteristics in Table 1 that it must be -1 . \square

EXAMPLE 12. Concerning Example 10, a similar analysis can be made regarding the matrix pencil $L(\lambda)$ of Figure 3 and its real symmetric reversal $\text{rev} L(\lambda) = \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} \lambda + \begin{bmatrix} 0 & M \\ M & D \end{bmatrix}$ where M , D , and K are the coefficient matrices of $P(\lambda)$ of (19). The curves of the eigenvalue functions of $\text{rev} L(\lambda)$ are exhibited in Figure 5. We can see that, in agreement with the signs in Table 2, the sign characteristics of the two eigenvalues, -0.8003 and 0.2687 , are both -1 . According to Table 2, the sign characteristic of the eigenvalue at infinity and partial multiplicity 2 is -1 . \square

Although $L(\lambda) = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$ is a linearization of the form (6) (with $a_1 = 0, a_2 = 1$) for $P(\lambda) = M\lambda^2 + D\lambda + K$ when $0 \notin \sigma(P)$, $\text{rev} L(\lambda)$ is *not* a pencil of the form (6) (with respect to $\text{rev} P(\lambda)$). However, we have the simple relation

$$\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \left(\begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} \lambda + \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \right) \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} = \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix} \lambda + \begin{bmatrix} D & M \\ M & 0 \end{bmatrix}.$$

The pencil on the right-hand side of this identity is a strong linearization of $\text{rev} P(\lambda)$ of the form (6) (for $a_1 = 1$ and $a_2 = 0$) because $0 = -\frac{a_2}{a_1} \notin \sigma(\text{rev} P)$. Since this pencil and $\text{rev} L(\lambda)$ are equivalent matrix polynomials through the constant matrix $\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$ (i.e. they are *strictly equivalent*), $\text{rev} L(\lambda)$ is also a strong linearization of $\text{rev} P(\lambda)$.

On the other hand, we can see in Example 11 that, while the sign characteristics of the real eigenvalues of $P(\lambda)$ and $L(\lambda)$ coincide, this is no longer true for the sign characteristics of the real eigenvalues of $\text{rev} P(\lambda)$ and $\text{rev} L(\lambda)$. In other words, $\text{rev} L(\lambda)$ is a strong linearization of $\text{rev} P(\lambda)$ which *does not preserve* the sign characteristic. As mentioned in Section 1.1, the problem of characterizing the Hermitian linearizations that preserve the sign characteristic of a given Hermitian matrix polynomial was studied in [3] for matrices of arbitrary degree and nonsingular leading coefficient. However, the definition of sign characteristic in [3] is not that of Definition 8 derived from Rellich's theorem. We will analyze the relation between both definitions in Section 7. We will see in Section 9 how to obtain *strong linearizations which*, when $\det M = 0$, *preserve the sign characteristics for the real eigenvalues as well as for the eigenvalue at infinity*. We first need to

analyze the relationship between the sign characteristic of Hermitian pencils and the signs appearing in their canonical form under congruence ([12, Theorems 6.1 and 9.2]).

4. Reduction of a linear pencil by congruence. Linearizations of the form (6) preserve spectral properties of $P(\lambda)$. Another way to reveal these properties is to reduce the two coefficients of (6) to canonical form by congruence transformations: $\lambda A_1 + B_1$ and $\lambda A_2 + B_2$ are congruent linear pencils if there exists a nonsingular matrix X such that $X^*(\lambda A_1 + B_1)X = \lambda A_2 + B_2$. General results of this kind (together with comments on the long history of this topic) can be found in [12] (Theorems 9.2 and 6.1, respectively), and we present them concisely here (without proof) in the regular case. Interesting connections are made with earlier, closely related work of Thompson [19], and more recent work of Mehrmann *et al* [16].

4.1. Regular linear pencils. Canonical forms of Hermitian linear pencils are direct sums of blocks associated with the elementary divisors. When the associated eigenvalue is real, each of these blocks has a sign +1 or -1 attached to it (see (23) or (24) for example). Whether these signs coincide with those of the linear pencil as defined in Definition 8 will be investigated in this section. We will see that the canonical forms of reference [12] can be slightly modified in such a way that the two collections of signs coincide.

The canonical forms under congruence are formulated in terms of some fixed elementary matrices. Define real symmetric $m \times m$ matrices:

$$F_m = \begin{bmatrix} 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & & 1 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 1 & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} = F_m^{-1},$$

$$(22) \quad G_m = \begin{bmatrix} 0 & \cdots & \cdots & 1 & 0 \\ \vdots & & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & 0 & 0 \\ 1 & 0 & & \vdots & \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} = \begin{bmatrix} F_{m-1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \tilde{G}_m = \begin{bmatrix} 0 & 0 \\ 0 & F_{m-1} \end{bmatrix},$$

and the $2m \times 2m$ symmetric matrix (with 2×2 nonzero blocks on the counter diagonal):

$$H_{2m} = \begin{bmatrix} 0 & 0 & \cdots & & 1 & 0 \\ 0 & 0 & \cdots & & 0 & -1 \\ & & & 1 & 0 & 0 \\ & & & 0 & -1 & 0 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ 1 & 0 & \cdots & & 0 & 0 \\ 0 & -1 & \cdots & & 0 & 0 \end{bmatrix}.$$

In the following theorems (and throughout the paper)m, finite **real eigenvalues** of matrix polynomials are denoted by $\alpha_1, \dots, \alpha_q$ and **nonreal eigenvalue pairs** are denoted by $\mu_1 \pm i\nu_1, \dots, \mu_s \pm i\nu_s$, with possible repetitions in both cases. In the canonical forms of (23) and (24) (below), the sums to r , q , and s refer to, respectively: (1) the “singular structure” (eigenvalues at infinity), (2) the real eigenvalues, and (3) the non-real eigenvalues.

THEOREM 13 ([12, Theorem 9.2] (“Real” case)). *Let A, B be $n \times n$ real symmetric matrices such that $\det(\lambda A + B)$ is not identically zero. Then, $\lambda A + B$ is congruent (over \mathbb{R}) to a pencil of the form*

$$(23) \quad \widehat{L}_R(\lambda) = \bigoplus_{j=1}^r \widehat{\delta}_j (F_{k_j} + \lambda G_{k_j}) \oplus \bigoplus_{j=1}^q \widehat{\eta}_j ((\lambda - \alpha_j) F_{l_j} + G_{l_j}) \\ \oplus \bigoplus_{j=1}^s \left((\lambda - \mu_j) F_{2m_j} - \nu_j H_{2m_j} + \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix} \right),$$

where $k_1 \leq \dots \leq k_r$, $l_1 \leq \dots \leq l_q$, and $m_1 \leq \dots \leq m_s$ are positive integers, and $\widehat{\delta}_1, \dots, \widehat{\delta}_r, \widehat{\eta}_1, \dots, \widehat{\eta}_q$ are equal to either +1 or -1.

We have a similar result for complex Hermitian pencils:

THEOREM 14 ([12, Theorem 6.1] (“Complex” case)). *Let A, B be $n \times n$ complex Hermitian matrices such that $\det(\lambda A + B)$ does not vanish identically. Then, $\lambda A + B$ is congruent (over \mathbb{C}) to a pencil of the form*

$$(24) \quad \widehat{L}_C(\lambda) = \bigoplus_{j=1}^r \widehat{\delta}_j (F_{k_j} + \lambda G_{k_j}) \oplus \bigoplus_{j=1}^q \widehat{\eta}_j ((\lambda - \alpha_j) F_{l_j} + G_{l_j}) \\ \oplus \bigoplus_{j=1}^s \left(\begin{bmatrix} 0 & (\lambda - \mu_j - i\nu_j) F_{m_j} \\ (\lambda - \mu_j + i\nu_j) F_{m_j} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{m_j} \\ G_{m_j} & 0 \end{bmatrix} \right),$$

where $k_1 \leq \dots \leq k_r$, $l_1 \leq \dots \leq l_q$, and $m_1 \leq \dots \leq m_s$ are positive integers, and $\widehat{\delta}_1, \dots, \widehat{\delta}_r, \widehat{\eta}_1, \dots, \widehat{\eta}_q$ are equal to either +1 or -1.

In both theorems, the sizes of the blocks must satisfy

$$\sum_{i=1}^r k_i + \sum_{j=1}^q l_j + 2 \sum_{k=1}^s m_k = n.$$

To illustrate some of the structure in (23) and (24), we note that:

$$F_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \quad G_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad H_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}.$$

Thus, for a real eigenvalue α with a partial multiplicity four there will be a term in the second summation in (23) or (24):

$$(\lambda - \alpha) F_4 + G_4 = \begin{bmatrix} 0 & 0 & 1 & \lambda - \alpha \\ 0 & 1 & \lambda - \alpha & 0 \\ 1 & \lambda - \alpha & 0 & 0 \\ \lambda - \alpha & 0 & 0 & 0 \end{bmatrix}.$$

For a nonreal conjugate pair of eigenvalues $\mu \pm i\nu$ with a partial multiplicity two, we have in the third summation of (23):

$$(\lambda - \mu) F_4 - \nu H_4 + \begin{bmatrix} F_2 & 0 \\ 0 & 0_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -\nu & \lambda - \mu \\ 1 & 0 & \lambda - \mu & \nu \\ -\nu & \lambda + \mu & 0 & 0 \\ \lambda - \mu & \nu & 0 & 0 \end{bmatrix}.$$

And in (24) :

$$\begin{bmatrix} 0 & (\lambda - \mu - i\nu)F_2 \\ (\lambda - \mu + i\nu)F_2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_2 \\ G_2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & \lambda - \mu - i\nu \\ 0 & 0 & \lambda - \mu - i\nu & 0 \\ 1 & \lambda - \mu + i\nu & 0 & 0 \\ \lambda - \mu + i\nu & 0 & 0 & 0 \end{bmatrix}.$$

In a paper of 1991, [19], R. C. Thompson has also investigated the canonical forms of real symmetric and complex Hermitian matrix pencils under congruence. His canonical forms differ from those of Theorems 13 and 14 in the constituent blocks. They are, of course, intimately connected (the “ones” are in the counter-subdiagonal). For the purpose of comparison, we state Thompson’s results (recall the definition of \tilde{G}_m in (22)).

THEOREM 15 ([19, Section 6] (“Real” case)). *Under the hypotheses of Theorem 13, $\lambda A + B$ is congruent (over \mathbb{R}) to a pencil of the form*

$$\begin{aligned} \tilde{L}_R(\lambda) &= \bigoplus_{j=1}^r \tilde{\delta}_j (F_{k_j} - \lambda \tilde{G}_{k_j}) \oplus \bigoplus_{j=1}^q \tilde{\eta}_j ((\alpha_j - \lambda)F_{l_j} + \tilde{G}_{l_j}) \\ &\oplus \bigoplus_{j=1}^s \left((\mu_j - \lambda)F_{2m_j} + \nu_j H_{2m_j} + \begin{bmatrix} 0_2 & 0 \\ 0 & F_{2m_j-2} \end{bmatrix} \right), \end{aligned} \tag{25}$$

where $k_1 \leq \dots \leq k_r$, $l_1 \leq \dots \leq l_q$, and $m_1 \leq \dots \leq m_s$ are positive integers, and $\tilde{\delta}_1, \dots, \tilde{\delta}_r, \tilde{\eta}_1, \dots, \tilde{\eta}_q$ are equal to either +1 or -1.

THEOREM 16 ([19, Section 8] (“Complex” case)). *Under the hypotheses of Theorem 14, $\lambda A + B$ is congruent (over \mathbb{C}) to a pencil of the form*

$$\begin{aligned} \tilde{L}_C(\lambda) &= \bigoplus_{j=1}^r \tilde{\delta}_j (F_{k_j} - \lambda \tilde{G}_{k_j}) \oplus \bigoplus_{j=1}^q \tilde{\eta}_j ((\alpha_j - \lambda)F_{l_j} + \tilde{G}_{l_j}) \\ &\oplus \bigoplus_{j=1}^s \left(\begin{bmatrix} 0 & (\mu_j + i\nu_j - \lambda)F_{m_j} \\ (\mu_j - i\nu_j - \lambda)F_{m_j} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \tilde{G}_{m_j} \\ \tilde{G}_{m_j} & 0 \end{bmatrix} \right), \end{aligned} \tag{26}$$

where $k_1 \leq \dots \leq k_r$, $l_1 \leq \dots \leq l_q$, and $m_1 \leq \dots \leq m_s$ are positive integers, and $\tilde{\delta}_1, \dots, \tilde{\delta}_r, \tilde{\eta}_1, \dots, \tilde{\eta}_q$ are equal to either +1 or -1.

In all of the canonical forms (23), (24), (25), (26), the positive integers $k_1 \leq \dots \leq k_r$ are the exponents of the elementary divisors at infinity while $l_1 \leq \dots \leq l_q$, and $m_1 \leq \dots \leq m_s$ are, respectively, the exponents of the elementary divisors for the real eigenvalues and for pairs of complex conjugate eigenvalues of $\lambda A + B$.

The collection of signs $(\widehat{\delta}_1, \dots, \widehat{\delta}_r; \widehat{\eta}_1, \dots, \widehat{\eta}_q)$ in Theorems 13 and 14 is called in [12] the *sign characteristic* of $\lambda A + B$. And the collection of signs $(\widetilde{\delta}_1, \dots, \widetilde{\delta}_r; \widetilde{\eta}_1, \dots, \widetilde{\eta}_q)$ in Theorems 15 and 16 is called in [19] the *inertial signature* of $\lambda A + B$. We will see that these collections of signs may not coincide with the sign characteristic of Definition 8. So we call:

- $(\widehat{\delta}_1, \dots, \widehat{\delta}_r; \widehat{\eta}_1, \dots, \widehat{\eta}_q)$ of Theorems 13 and 14 the *Lancaster-Rodman sign characteristic (LR-sign characteristic, for short)* of $\lambda A + B$.
- $(\widetilde{\delta}_1, \dots, \widetilde{\delta}_r; \widetilde{\eta}_1, \dots, \widetilde{\eta}_q)$ of Theorems 15 and 16 the *inertial signature* of $\lambda A + B$.

We reserve the name *sign characteristic* of $\lambda A + B$ for the sign characteristic of Definition 8 (obtained by applying unitary similarity to $\lambda A + B$ as an Hermitian analytic matrix function of real variable λ ; the “Rellich strategy”).

4.2. No real eigenvalues. It is a direct consequence of Theorems 13 and 14 that, in the absence of real and infinite eigenvalues, two real symmetric or complex Hermitian linear pencils are congruent if and only if they have the same elementary divisors for the nonreal eigenvalues; that is, if they are strictly equivalent. Specifically:

COROLLARY 17. *Let $\lambda A_1 + B_1$ and $\lambda A_2 + B_2$ be $n \times n$ Hermitian linear pencils. If $\det A_i \neq 0$ and $\sigma(\lambda A_i + B_i) \subset \mathbb{C} \setminus \mathbb{R}$, $i = 1, 2$, then $\lambda A_1 + B_1$ and $\lambda A_2 + B_2$ are congruent if and only if they have the same elementary divisors.*

Since blocks associated with the nonreal eigenvalues of $\widehat{L}_R(\lambda)$ and $\widetilde{L}_R(\lambda)$, namely,

$$(\lambda - \mu_j)F_{2m_j} - \nu_j H_{2m_j} + \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix},$$

and

$$(\mu_j - \lambda)F_{2m_j} + \nu_j H_{2m_j} + \begin{bmatrix} 0_2 & 0 \\ 0 & F_{2m_j-2} \end{bmatrix},$$

have one and the same elementary divisor, $(\lambda^2 - 2\mu_j\lambda + \mu_j^2 + \nu_j^2)^{m_j}$, it follows from Corollary 17 that they are congruent.

Also, the blocks

$$\begin{bmatrix} 0 & (\lambda - \mu_j - i\nu_j)F_{m_j} \\ (\lambda - \mu_j + i\nu_j)F_{m_j} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{m_j} \\ G_{m_j} & 0 \end{bmatrix},$$

and

$$\begin{bmatrix} 0 & (\mu_j + i\nu_j - \lambda)F_{m_j} \\ (\mu_j - i\nu_j - \lambda)F_{m_j} & 0 \end{bmatrix} + \begin{bmatrix} 0 & \widetilde{G}_{m_j} \\ \widetilde{G}_{m_j} & 0 \end{bmatrix},$$

of $\widehat{L}_C(\lambda)$ and $\widetilde{L}_C(\lambda)$, respectively, are congruent (see (24) and (26) and definition of G_m and \widetilde{G}_m in (22)).

However, we will see that $F_{k_j} + \lambda G_{k_j}$ and $F_{k_j} - \lambda \widetilde{G}_{k_j}$, on one hand, and $(\lambda - \alpha_j)F_{l_j} + G_{l_j}$ and $(\alpha_j - \lambda)F_{l_j} + \widetilde{G}_{l_j}$, on the other hand, may *not* be congruent. Actually, we aim to provide a new canonical form for $\lambda A + B$ under congruence where the signs accompanying the blocks associated with the real and infinite eigenvalues coincide with the sign characteristic of $\lambda A + B$. Then we will apply our findings to the linearizations given in (6) for $P(\lambda)$ of (1) to see whether they have the same sign characteristic as $P(\lambda)$.

4.3. General (real and complex) spectrum. In what follows we will repeatedly use the fact that the sign characteristic of an Hermitian matrix polynomial is invariant under constant congruence. That is to say, for any $n \times n$ matrix polynomial $P(\lambda)$ and any invertible matrix $R \in \mathbb{F}^{m \times n}$, the sign characteristics of the eigenvalues of $P(\lambda)$ and $R^*P(\lambda)R$ coincide. This is a direct consequence of the stronger result [16, Theorem 2.7] concerning the finite eigenvalues (see also [6, Theorem 12.4] although the notion of sign characteristic in this reference has a different but equivalent meaning to that of Definition 8). Also, by Theorem 3, the rank and finite elementary divisors of $P(\lambda)$ and $R^*P(\lambda)R$ are identical. As far as the eigenvalue at infinity is concerned the sign characteristic and elementary divisors of $P(\lambda)$ at infinity are the sign characteristic of $-\text{rev } P(\lambda)$ and elementary divisors of $\text{rev } P(\lambda)$ for the eigenvalue 0, respectively. But $\text{rev}(R^*P(\lambda)R) = R^*\text{rev } P(\lambda)R$. Therefore, the sign characteristic and elementary divisors of $P(\lambda)$ and $R^*P(\lambda)R$ coincide for the eigenvalue at infinity as well. As a consequence, the sign characteristic of $\lambda A + B$ and its canonical forms (those of Theorems 13 and 15, or of Theorems 14 and 16) are the same.

That the LR-sign characteristic of $\lambda A + B$ (of Theorems 13 or 14) may not be its sign characteristic was proved in [16]. The proof is based on the following result:

THEOREM 18 ([16, Theorem 4.11]). (a) For $\alpha \in \mathbb{R}$, the real symmetric pencil $(\lambda - \alpha)F_k + G_k$ has the unique real eigenvalue α with partial multiplicity k and the sign characteristic $(-1)^{k+1}$.

(b) The real symmetric pencil $F_k + \lambda G_k$ has infinity as its unique eigenvalue. It has partial multiplicity k and the sign characteristic $(-1)^k$.

Bearing in mind that, as mentioned, the sign characteristic is invariant under congruence, we obtain:

COROLLARY 19. Let $\lambda A + B$ be an $n \times n$ Hermitian matrix pencil and let $\widehat{L}_R(\lambda)$, $\widehat{L}_C(\lambda)$ be the canonical forms (under congruence) of Theorem 13 or 14, respectively, according as A and B are real or complex matrices. Then,

- (i) the sign characteristic of a real eigenvalue α_j with partial multiplicity l_j is $(-1)^{l_j+1}\widehat{\eta}_j$, $j = 1, \dots, q$, and
- (ii) the sign characteristic of the eigenvalue at infinity with partial multiplicity k_j is $(-1)^{k_j}\widehat{\delta}_j$, $j = 1, \dots, r$.

Proof. We will focus on the real case; the proof in the complex case is similar. First, $\lambda A + B$ and $\widehat{L}_R(\lambda)$ have the same sign characteristic because they are congruent. Second, by applying Rellich's theorem (Theorem 2) to each block of $\widehat{L}_R(\lambda)$, it is easily seen that the sign characteristic of the eigenvalue α_i and partial multiplicity l_i for $\widehat{L}_R(\lambda)$ is the sign characteristic of the eigenvalue α_i and partial multiplicity l_i for $\widehat{\eta}_i((\lambda - \alpha_i)F_{l_i} + G_{l_i})$. According to Theorem 18 this is $(-1)^{l_i+1}\widehat{\eta}_i$.

The sign characteristic associated with an eigenvalue at infinity with partial multiplicity k_j is, by definition, the sign characteristic of the eigenvalue 0 with partial multiplicity k_j of $-\text{rev } \widehat{L}_R(\lambda)$. As above, this is the sign characteristic of the eigenvalue 0 and partial multiplicity k_j of $-\widehat{\delta}_j(\lambda F_{k_j} + G_{k_j})$. By Theorem 18, this is $(-1)^{k_j+2}\widehat{\delta}_j$, as claimed. \square

The canonical forms under congruence of Theorems 13 or 14 can be slightly modified so that the signs associated with each block agree with the sign characteristics of the corresponding elementary divisors of $\lambda A + B$. We start with the real case.

THEOREM 20. Let A, B be $n \times n$ real symmetric matrices such that $\det(\lambda A + B)$ is not identically equal to zero. Then, $\lambda A + B$ is congruent (over \mathbb{R}) to a pencil of the form

$$(27) \quad L_R(\lambda) = \bigoplus_{j=1}^r \delta_j (\lambda G_{k_j} - F_{k_j}) \oplus \bigoplus_{j=1}^q \eta_j ((\lambda - \alpha_j)F_{l_j} - G_{l_j}) \\ \oplus \bigoplus_{j=1}^s \left((\lambda - \mu_j)F_{2m_j} - \nu_j H_{2m_j} - \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix} \right),$$

where $(\delta_1, \dots, \delta_r; \eta_1, \dots, \eta_q)$ is the sign characteristic of $\lambda A + B$.

Proof. We will prove the following properties:

- (i) $\widehat{\eta}_j((\lambda - \alpha_j)F_{l_j} + G_{l_j})$ is congruent to $\widehat{\eta}_j((\lambda - \alpha_j)F_{l_j} - G_{l_j})$ if l_j is odd.
- (ii) $\widehat{\eta}_j((\lambda - \alpha_j)F_{l_j} + G_{l_j})$ is congruent to $-\widehat{\eta}_j((\lambda - \alpha_j)F_{l_j} - G_{l_j})$ if l_j is even.
- (iii) $\widehat{\delta}_j(F_{k_j} + \lambda G_{k_j})$ is congruent to $\widehat{\delta}_j(\lambda G_{k_j} - F_{k_j})$ if k_j is even.
- (iv) $\widehat{\delta}_j(F_{k_j} + \lambda G_{k_j})$ is congruent to $-\widehat{\delta}_j(\lambda G_{k_j} - F_{k_j})$ if k_j is odd.

Let k be a nonnegative integer and define

$$S_k = \text{Diag}[-1, +1, -1, +1, \dots, (-1)^k] \in \mathbb{R}^{k \times k}.$$

A simple computation shows that if l_j is odd, then $S_{l_j}((\lambda - \alpha_j)F_{l_j} + G_{l_j})S_{l_j} = (\lambda - \alpha_j)F_{l_j} - G_{l_j}$. And if l_j is even, then $S_{l_j}((\lambda - \alpha_j)F_{l_j} + G_{l_j})S_{l_j} = -(\lambda - \alpha_j)F_{l_j} + G_{l_j} = -((\lambda - \alpha_j)F_{l_j} - G_{l_j})$. Similarly, if k_j is odd, then $S_{k_j}(F_{k_j} + \lambda G_{k_j})S_{k_j} = F_{k_j} - \lambda G_{k_j} = -(\lambda G_{k_j} - F_{k_j})$. And if k_j is even, then $S_{k_j}(F_{k_j} + \lambda G_{k_j})S_{k_j} = -F_{k_j} + \lambda G_{k_j} = \lambda G_{k_j} - F_{k_j}$. Therefore, properties (i)–(iv) hold true.

Define $\eta_j = \widehat{\eta}_j$ if l_j is odd and $\eta_j = -\widehat{\eta}_j$ otherwise. Also, let $\delta_j = \widehat{\delta}_j$ if k_j is even and $\delta_j = -\widehat{\delta}_j$ if k_j is odd. Then, $\widehat{L}_R(\lambda)$ of (23) and $L_R(\lambda)$ of (27) are congruent matrix pencils. In addition, $\delta_j = (-1)^{k_j} \widehat{\delta}_j$ $j = 1, \dots, r$ and $\eta_j = (-1)^{l_j+1} \widehat{\eta}_j$ for $j = 1, \dots, q$. According to Corollary 19, η_j is the sign characteristic of the eigenvalue α_j and partial multiplicity l_j , and δ_j is the sign characteristic of the eigenvalue at infinity and partial multiplicity k_j . \square

For complex Hermitian matrices we have a similar result. The proof follows the same lines and is omitted.

THEOREM 21. *Let A, B be $n \times n$ complex Hermitian matrices such that $\det(\lambda A + B) \neq 0$. Then, $\lambda A + B$ is congruent (over \mathbb{C}) to a pencil of the form*

$$(28) \quad L_C(\lambda) = \bigoplus_{j=1}^r \delta_j (\lambda G_{k_j} - F_{k_j}) \oplus \bigoplus_{j=1}^q \eta_j ((\lambda - \alpha_j)F_{l_j} - G_{l_j}) \\ \oplus \bigoplus_{j=1}^s \left(\begin{bmatrix} 0 & (\lambda - \mu_j - i\nu_j)F_{m_j} \\ (\lambda - \mu_j + i\nu_j)F_{m_j} & 0 \end{bmatrix} - \begin{bmatrix} 0 & G_{m_j} \\ G_{m_j} & 0 \end{bmatrix} \right),$$

where $(\delta_1, \dots, \delta_r; \eta_1, \dots, \eta_q)$ is the sign characteristic of $\lambda A + B$.

Now we can clarify how the canonical forms given by Thompson ((25)–(26)) and those of Theorems 20 and 21 are related. Note first that for any nonnegative integer k

$$(29) \quad F_k F_k F_k = F_k, \quad F_k G_k F_k = \widetilde{G}_k,$$

and $(\mu_j - \lambda)F_{2m_j} + \nu_j H_{2m_j} + \begin{bmatrix} 0 & 0 \\ 0 & F_{2m_j-2} \end{bmatrix}$ is, by Corollary 17, congruent to

$$(\mu_j - \lambda)F_{2m_j} + \nu_j H_{2m_j} + \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix}.$$

Therefore, $-\widetilde{L}_R(\lambda)$ and $L_R(\lambda)$ are congruent matrix pencils and so $\widetilde{\delta}_j = -\delta_j$, $j = 1, \dots, r$, and $\widetilde{\eta}_j = -\eta_j$, $j = 1, \dots, q$. In other words, for any real symmetric linear pencil $\lambda A + B$, Thompson's *inertial signature* is the “opposite” of the *sign characteristic* of Definition 8. The same properties hold for the canonical forms (26) and (28) of complex Hermitian pencils.

5. Reduction of a linearization by similarity. The notion of a *sign characteristic* plays an important role in the context of self-adjoint standard triples and also in the context of H -self-adjoint matrices. In this section, we review the main concepts of these theories in the context of quadratic eigenvalue problems. Basic references are [7] and [14] where proofs and other relevant related results can be found.

5.1. Canonical forms for similar self-adjoint standard triples.

DEFINITION 22. *Let $T, \widehat{T} \in \mathbb{F}^{2n \times 2n}$, $X, \widehat{X} \in \mathbb{F}^{n \times 2n}$ and $Y, \widehat{Y} \in \mathbb{F}^{2n \times n}$. Then*

- (a) (X, T) is said to be a **standard pair** over \mathbb{F} if $\begin{bmatrix} X \\ X T \end{bmatrix}$ is invertible.
- (b) (X, T, Y) is a **standard triple** over \mathbb{F} if (X, T) is a standard pair and there is an invertible matrix $Q \in \mathbb{F}^{n \times n}$ such that $Y = \begin{bmatrix} X \\ X T \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ Q \end{bmatrix}$.

(c) A standard triple (X, T, Y) is said to be **self-adjoint** if there is a nonsingular Hermitian matrix $H \in \mathbb{F}^{2n \times 2n}$ for which

$$(30) \quad Y^* = XH^{-1}, \quad T^* = HTH^{-1}, \quad X^* = HY.$$

(When $\mathbb{F} = \mathbb{R}$ we will sometimes emphasize that (X, T, Y) is a **real self-adjoint standard triple**.)

(d) Two triples (X, T, Y) and $(\widehat{X}, \widehat{T}, \widehat{Y})$ are said to be **similar** over \mathbb{F} if there is an invertible matrix $S \in \mathbb{F}^{2n \times 2n}$ such that $\widehat{T} = STS^{-1}$, $\widehat{X} = XS^{-1}$, and $\widehat{Y} = SY$.

It is easily seen that if (X, T, Y) is a standard triple and it is similar to $(\widehat{X}, \widehat{T}, \widehat{Y})$, then the latter is also a standard triple. Thus, according to the definition of similarity of triples (item (d)), condition (30) means that (X, T, Y) and (Y^*, T^*, X^*) are similar standard triples with an Hermitian matrix H as similarity transformation. It turns out that a standard triple (X, T, Y) is self-adjoint if and only if it is similar to (Y^*, T^*, X^*) . This is Theorem 3.4 of [14]. As a consequence, if (X, T, Y) and $(\widehat{X}, \widehat{T}, \widehat{Y})$ are similar triples and one of them is self-adjoint then the other triple is also self-adjoint. This property will be used below.

Matrix S of item (d) is uniquely determined by either (X, T) and $(\widehat{X}, \widehat{T})$ or (T, Y) and $(\widehat{T}, \widehat{Y})$ (see [6, Theorem 1.25]):

$$(31) \quad S = \begin{bmatrix} \widehat{X} \\ \widehat{X}\widehat{T} \end{bmatrix}^{-1} \begin{bmatrix} X \\ XT \end{bmatrix} = [\widehat{Y} \quad \widehat{T}\widehat{Y}] [Y \quad TY]^{-1}.$$

Therefore, if (X, T, Y) is a self-adjoint standard triple and H is the invertible Hermitian matrix of (30), then it is uniquely determined (see [7, Section 12.4] or [14, Proposition 3.6]):

$$(32) \quad H = \begin{bmatrix} Y^* \\ Y^*T^* \end{bmatrix}^{-1} \begin{bmatrix} X \\ XT \end{bmatrix} = [X^* \quad T^*X^*] [Y \quad TY]^{-1}.$$

Conversely, if (X, T) is a standard pair and H is an invertible Hermitian matrix such that $T^* = HTH^{-1}$, then $(X, T, H^{-1}X^*)$ is a self-adjoint standard triple. In what follows, we may use either of the notations, (X, T, Y) or $(X, T, H^{-1}X^*)$, for a self-adjoint standard triple.

5.2. Application to quadratic problems. As in (1) and (4), let $P(\lambda) = M\lambda^2 + D\lambda + K$ with $M, D, K \in \mathbb{F}^{n \times n}$, and $\det M \neq 0$. If

$$(33) \quad X_0 = [I_n \quad 0], \quad C_P = \begin{bmatrix} 0 & I_n \\ -M^{-1}K & -M^{-1}D \end{bmatrix}, \quad Y_0 = \begin{bmatrix} 0 \\ M^{-1} \end{bmatrix}$$

then (X_0, C_P, Y_0) is a standard triple. If, in addition, $M = M^*$, $D = D^*$, and $K = K^*$, then (X_0, C_P, Y_0) is a self-adjoint standard triple. Indeed, if

$$(34) \quad A = \begin{bmatrix} D & M \\ M & 0 \end{bmatrix}$$

then A is invertible and Hermitian,

$$(35) \quad C_P^* A = AC_P = \begin{bmatrix} -K & 0 \\ 0 & M \end{bmatrix},$$

and $AY_0 = X_0^*$. In effect, $\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} (\lambda A - AC_P) \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix}$ is the linearization of (8) with $a_1 = 0$ and $a_2 = 1$.

Recall that triples similar to self-adjoint standard triples are themselves self-adjoint.

DEFINITION 23. Let $P(\lambda)$ be the Hermitian quadratic matrix polynomial of (1) with M nonsingular. The triple $(X_0, C_P, Y_0) = (X_0, C_P, A^{-1}X_0^*)$ is called the primitive self-adjoint standard triple of $P(\lambda)$. And any triple (X, T, Y) similar to (X_0, C_P, Y_0) is said to be a self-adjoint standard triple of $P(\lambda)$.

In particular, if (X, T, Y) is a self-adjoint standard triple of $P(\lambda)$, then $\lambda I_{2n} - T$ is a linearization of $P(\lambda)$ because T and C_P are similar matrices. Conversely, if $\lambda I_{2n} - T$ is a linearization of $P(\lambda)$ and $T = SC_P S^{-1}$ then $(X_0 S^{-1}, T, SY_0)$ is a self-adjoint standard triple of $P(\lambda)$.

Canonical forms for the similarity of self-adjoint standard triples over \mathbb{C} and \mathbb{R} of a given Hermitian matrix polynomial with nonsingular leading coefficient are provided in Theorems 4.3 and 4.4 of [14], respectively. They are obtained by using the following result that is stated here for quadratic matrix polynomials and partially proved in [14] for matrix polynomials of arbitrary degree. The proof accompanying the following theorem can be easily extended to matrix polynomials of any degree.

THEOREM 24. Let M be nonsingular, $(X, T, H^{-1}X^*)$ be a self-adjoint standard triple (as in Definition 22(c)), and $S = \begin{bmatrix} X_0 \\ XT \end{bmatrix}$. Then, with the definitions of (33) and (34), the triples $(X_0, C_P, A^{-1}X_0^*)$ and $(X, T, H^{-1}X^*)$ are similar if and only if

$$(36) \quad \lambda H - HT = S^*(\lambda A - AC_P)S.$$

Proof. If $(X_0, C_P, A^{-1}X_0^*)$ and $(X, T, H^{-1}X^*)$ are similar then there is an invertible matrix $R \in \mathbb{F}^{2n \times 2n}$ such that

$$T = R^{-1}C_P R, \quad X = X_0 R, \quad H^{-1}X^* = R^{-1}A^{-1}X_0^*.$$

Bearing in mind (31) and $\begin{bmatrix} X_0 \\ X_0 C_P \end{bmatrix} = I_{2n}$, we get

$$R = \begin{bmatrix} X_0 \\ X_0 C_P \end{bmatrix}^{-1} \begin{bmatrix} X \\ XT \end{bmatrix} = \begin{bmatrix} X \\ XT \end{bmatrix} = S.$$

Thus, it follows from $X = X_0 S$ that $H^{-1}X^* = S^{-1}A^{-1}X_0^* = S^{-1}A^{-1}S^{-*}X^*$. Also, using this identity, $T^* = HTH^{-1}$, $T = S^{-1}C_P S$ and $C_P^* = AC_P A^{-1}$, we obtain

$$\begin{aligned} H^{-1}T^*X^* &= TH^{-1}X^* = TS^{-1}A^{-1}S^{-*}X^* = S^{-1}C_P A^{-1}S^{-*}X^* \\ &= S^{-1}A^{-1}C_P^* S^{-*}X^* = S^{-1}A^{-1}S^{-*}T^*X^*. \end{aligned}$$

Hence

$$H^{-1} \begin{bmatrix} X^* & T^*X^* \end{bmatrix} = S^{-1}A^{-1}S^{-*} \begin{bmatrix} X^* & T^*X^* \end{bmatrix}.$$

Since $S^* = \begin{bmatrix} X^* & T^*X^* \end{bmatrix}$ is invertible, we must have $H = S^*AS$. Also, $HT = S^*AST = S^*AC_P S$ because $T = S^{-1}C_P S$. Therefore, $\lambda H - HT = S^*(\lambda A - AC_P)S$ as desired.

Conversely, if $H = S^*AS$ and $HT = S^*AC_P S$, then $HT = S^*ASS^{-1}C_P S = HS^{-1}C_P S$. Since H is nonsingular, $T = S^{-1}C_P S$. In addition,

$$XS^{-1} = X \begin{bmatrix} X \\ XT \end{bmatrix}^{-1} = \begin{bmatrix} I_n & 0 \end{bmatrix} = X_0.$$

Finally, $H^{-1}X^* = S^{-1}A^{-1}S^{-*}S^*X_0^* = S^{-1}A^{-1}X_0^*$. Hence, $(X_0, C_P, A^{-1}X_0^*)$ and $(X, T, H^{-1}X^*)$ are similar self-adjoint standard triples. \square

A simple consequence of Theorem 24 is

COROLLARY 25. Let $T, H, S \in \mathbb{F}^{2n \times 2n}$ be an arbitrary matrix, an invertible Hermitian matrix, and an invertible matrix, respectively. Then (36) holds if and only if $T = S^{-1}C_P S$ and $H = S^*AS$. In addition, if one of these two equivalent conditions holds and $X = [I_n \ 0]S$, then S is the unique matrix for which the following relations hold:

$$(37) \quad T = S^{-1}C_P S, \quad X = X_0 S, \quad H^{-1}X^* = S^{-1}A^{-1}X_0^*.$$

Proof. That $\lambda H - HT = S^*(\lambda A - AC_P)S$ implies $T = S^{-1}C_P S$ and $H = S^*AS$ has been shown in the second part of the proof of Theorem 24. The converse is trivial. Now, if $T = S^{-1}C_P S$, $H = S^*AS$ and $X = [I_n \ 0]S = X_0 S$, then in order to prove (37), we only have to see that $H^{-1}X^* = S^{-1}A^{-1}X_0^*$. But $H^{-1}X^* = (S^*AS)^{-1}(X_0 S)^* = S^{-1}A^{-1}S^{-*}S^*X_0^* = S^{-1}A^{-1}X_0^*$ as desired. Thus, $(X, T, H^{-1}X^*)$ and $(X_0, C_P, A^{-1}X_0^*)$ are similar and so $(X, T, H^{-1}X^*)$ is a self-adjoint standard triple. Hence, there is a unique matrix S satisfying (37). \square

Given $P(\lambda) = M\lambda^2 + D\lambda + K$ with $M^* = M$, $D^* = D$, $K^* = K$, and $\det M \neq 0$ we can obtain a canonical representative for the class of similar self-adjoint standard triples of $P(\lambda)$ as follows (see [14]):

1. Form the Hermitian linear pencil $\lambda A - AC_P$ of (34) and (35). We have already seen that this pencil is a linearization of $P(\lambda)$.
2. Compute the canonical form $L_C(\lambda)$ or $L_R(\lambda)$ of $\lambda A - AC_P$ of Theorems 21 and 20 according as A and AC_P are complex Hermitian or real symmetric matrices. Assume that $(\lambda - \alpha_j)^{l_j}$ are, for $j = 1, \dots, q$, the elementary divisors of $P(\lambda)$ (and of $\lambda A - AC_P$) for the real eigenvalues, and $(\lambda - \mu_j - i\nu_j)^{m_j}$ and $(\lambda - \mu_j + i\nu_j)^{m_j}$ are, for $j = 1, \dots, s$, the elementary divisors for the nonreal eigenvalues of $P(\lambda)$ (and of $\lambda A - AC_P$). Let ε_j be the sign characteristic for the elementary divisor $(\lambda - \alpha_j)^{l_j}$ of $\lambda A - AC_P$, $j = 1, \dots, q$. Define

$$(38) \quad P_{\varepsilon, J} = \bigoplus_{j=1}^q \varepsilon_j F_{l_j} \oplus \bigoplus_{j=1}^s F_{2m_j},$$

$$(39) \quad J_C = \bigoplus_{j=1}^q (\alpha_j I_{l_j} + F_{l_j} G_{l_j}) \oplus \bigoplus_{j=1}^s \begin{bmatrix} (\mu_j - i\nu_j) I_{m_j} + F_{m_j} G_{m_j} & 0 \\ 0 & (\mu_j + i\nu_j) I_{m_j} + F_{m_j} G_{m_j} \end{bmatrix},$$

or

$$(40) \quad J_R = \bigoplus_{j=1}^q (\alpha_j I_{l_j} + F_{l_j} G_{l_j}) \oplus \bigoplus_{j=1}^s \left(\mu_j I_{2m_j} + \nu_j F_{2m_j} H_{2m_j} + F_{2m_j} \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix} \right),$$

according as A and AC_P are complex Hermitian or real symmetric matrices. Then an easy computation shows that

$$L_C(\lambda) = \lambda P_{\varepsilon, J} - P_{\varepsilon, J} J_C \quad \text{and} \quad L_R(\lambda) = \lambda P_{\varepsilon, J} - P_{\varepsilon, J} J_R.$$

That is,

THEOREM 26. With the above notation, $\lambda P_{\varepsilon, J} - P_{\varepsilon, J} J_R$ and $\lambda P_{\varepsilon, J} - P_{\varepsilon, J} J_C$ are the canonical forms under congruence of Theorems 20 and 21, respectively, for the Hermitian pencil $\lambda A - AC_P$.

As a consequence $(\varepsilon_1, \dots, \varepsilon_q)$ is the sign characteristic of $\lambda A - AC_P$.

3. Let S be an invertible matrix for which $\lambda P_{\varepsilon,J} - P_{\varepsilon,J} J_C = S^*(\lambda A - AC_P)S$ or $\lambda P_{\varepsilon,J} - P_{\varepsilon,J} J_R = S^*(\lambda A - AC_P)S$ according as $\lambda A - AC_P$ is a complex Hermitian or real symmetric pencil, and let

$$(41) \quad X_\varepsilon = [I_n \quad 0] S.$$

A straightforward consequence of Corollary 25 is the following result. Note that $P_{\varepsilon,J}^{-1} = P_{\varepsilon,J}$.

THEOREM 27. Let A , C_P , $P_{\varepsilon,J}$, J_C , J_R , and X_ε be the matrices of (34), (35), (38), (39), (40), and (41) respectively. Then $(X_\varepsilon, J_C, P_{\varepsilon,J} X_\varepsilon^*)$ ($(X_\varepsilon, J_R, P_{\varepsilon,J} X_\varepsilon^*)$ in the real case) is similar over \mathbb{C} (over \mathbb{R} , respectively) to $(X_0, C_P, A^{-1} X_0^*)$ with X_0 of (33).

Note also that, while J_C (or J_R) and $P_{\varepsilon,J}$ are uniquely determined, up to a permutation of blocks corresponding to equal elementary divisors, by A and C_P (and then by $P(\lambda)$), X_ε depends on S which, in general, is not unique. This lack of uniqueness is related to the fact that the columns of X_ε are (generalized) eigenvectors of $P(\lambda)$. In particular, if the eigenvalues of $P(\lambda)$ are all semisimple, then J_C is diagonal and the columns of X_ε form a complete system of right eigenvectors of $P(\lambda)$ while the “rows” of $P_{\varepsilon,J} X_\varepsilon^*$ are a complete system of left eigenvectors of $P(\lambda)$ (see Sections 1.8 and 2.1 of [6]).

In general, if J_C is the complex Jordan matrix of (39), then the columns of X_ε can be partitioned into blocks $X_\varepsilon = [X_1 \quad \cdots \quad X_q \quad X_{q+1} \quad X_{q+2} \quad \cdots \quad X_{q+2s}]$ consistent with the Jordan blocks of J_C . Then the columns of each X_j form a right Jordan chain of $P(\lambda)$ for the eigenvalue of the corresponding Jordan block. Also, (see [6, Section 2.1]) the rows of $P_{\varepsilon,J} X_\varepsilon^*$, partitioned consistently with the partition of J_C and taken in reverse order in each block, form left Jordan chains of $P(\lambda)$.

DEFINITION 28. Self-adjoint standard triples of the form $(X_\varepsilon, J_C, P_{\varepsilon,J} X_\varepsilon^*)$ and $(X_\varepsilon, J_R, P_{\varepsilon,J} X_\varepsilon^*)$ are called, respectively, self-adjoint Jordan triples and real self-adjoint Jordan triples of $P(\lambda)$.

Self-adjoint Jordan triples act as canonical representatives of the self-adjoint standard triples of $P(\lambda)$ under similarity.

The collection of signs $(\varepsilon_1, \dots, \varepsilon_q)$ is called in [14] the sign characteristic of $P(\lambda)$. It is, in fact, uniquely determined by $P(\lambda)$ because, according to Theorems 20 and 21, it is the sign characteristic of $\lambda A - AC_P$. However, we want to reserve the name of “sign characteristic of $P(\lambda)$ ” for that of Definition 8 when applied to $P(\lambda)$ and, at this point, we do not know whether both coincide. Therefore,

DEFINITION 29. We call the collection of signs $(\varepsilon_1, \dots, \varepsilon_q)$ of (38) the Standard Triple sign characteristic (ST-sign characteristic, for short) of $P(\lambda)$.

Thus, the ST-sign characteristic of $P(\lambda)$ is the sign characteristic of $\lambda A - AC_P$. We state this result as a proposition for later use.

PROPOSITION 30. Let $P(\lambda) = M\lambda^2 + D\lambda + K$ be an $n \times n$ Hermitian quadratic matrix polynomial with M nonsingular, and let C_P and A be the matrices of (33) and (34), respectively. Then, the ST-sign characteristic of $P(\lambda)$ is the sign characteristic of $\lambda A - AC_P$.

6. Canonical forms for unitarily similar H -self-adjoint matrices.

DEFINITION 31 ([7]). Let $T, \widehat{T} \in \mathbb{F}^{n \times n}$ and let $H, \widehat{H} \in \mathbb{F}^{n \times n}$ be invertible Hermitian matrices.

- (a) T is said to be H -self-adjoint if $T^* H = HT$.
- (b) The pairs (T, H) and $(\widehat{T}, \widehat{H})$ are said to be unitarily similar over the field \mathbb{F} if there is an invertible matrix $S \in \mathbb{F}^{n \times n}$ such that $\widehat{T} = S^{-1} T S$ and $\widehat{H} = S^* H S$.

If $(X, T, H^{-1}X^*)$ is a self-adjoint standard triple then, by definition (see (30)), T is H -self-adjoint. In particular, if $P(\lambda) = M\lambda^2 + D\lambda + K$ is an Hermitian quadratic matrix polynomial and C_P and A are the matrices of (33) and (34), respectively, then C_P is A -self-adjoint (see (35)). Also, if $(X_1, T_1, H_1^{-1}X_1^*)$ and $(X_2, T_2, H_2^{-1}X_2^*)$ are similar self-adjoint standard triples then, by [14, Proposition 3.9], (T_1, H_1) and (T_2, H_2) are unitarily similar.

Canonical forms for unitarily similar pairs over \mathbb{C} and over \mathbb{R} are given in Theorems 5.1.1 and 6.1.5 of [7], respectively. We state these results for the readers' convenience.

THEOREM 32 ([7, Theorem 5.1.1] “complex case”). *Let $T, H \in \mathbb{C}^{n \times n}$ with H an invertible Hermitian matrix and let T be H -self-adjoint over \mathbb{C} . For $j = 1, \dots, q$ let $(\lambda - \alpha_j)^{l_j}$ be the elementary divisors of T for the real eigenvalues, and for $j = 1, \dots, s$ let $(\lambda - \mu_j - i\nu_j)^{m_j}$ and $(\lambda - \mu_j + i\nu_j)^{m_j}$ be the elementary divisors of T for the nonreal eigenvalues. Then there is an invertible matrix $S \in \mathbb{C}^{n \times n}$ such that $\widehat{J}_C = S^{-1}TS$ and $P_{\widehat{\epsilon}, J} = S^*HS$ as in (38). Thus,*

$$(42) \quad P_{\widehat{\epsilon}, J} = \bigoplus_{j=1}^q \widehat{\epsilon}_j F_{l_j} \oplus \bigoplus_{j=1}^s F_{2m_j},$$

$$(43) \quad \widehat{J}_C = \bigoplus_{j=1}^q (\alpha_j I_{l_j} + G_{l_j} F_{l_j}) \oplus \bigoplus_{j=1}^s \begin{bmatrix} (\mu_j + i\nu_j) I_{m_j} + G_{m_j} F_{m_j} & 0 \\ 0 & (\mu_j - i\nu_j) I_{m_j} + G_{m_j} F_{m_j} \end{bmatrix},$$

and $(\widehat{\epsilon}_1, \dots, \widehat{\epsilon}_q)$ is an ordered family of signs $+1$ or -1 uniquely determined by (T, H) up to a permutation of signs corresponding to equal Jordan blocks.

THEOREM 33 ([7, Theorem 6.1.5] “real case”). *Let $T, H \in \mathbb{R}^{n \times n}$ with H an invertible symmetric matrix and let T be H -self-adjoint over \mathbb{R} . For $j = 1, \dots, q$ let $(\lambda - \alpha_j)^{l_j}$ be the elementary divisors of T for the real eigenvalues, and for $j = 1, \dots, s$ let $(\lambda - \mu_j - i\nu_j)^{m_j}$ and $(\lambda - \mu_j + i\nu_j)^{m_j}$ be the elementary divisors of T for the nonreal eigenvalues. Then there is an invertible matrix $S \in \mathbb{R}^{n \times n}$ such that $\widehat{J}_R = S^{-1}TS$ and $\widehat{P}_{\widehat{\epsilon}, J} = S^*HS$ where $P_{\widehat{\epsilon}, J}$ is the matrix of (42),*

$$(44) \quad \widehat{J}_R = \bigoplus_{j=1}^q (\alpha_j I_{l_j} + G_{l_j} F_{l_j}) \oplus \bigoplus_{j=1}^s \left(\mu_j I_{2m_j} + \nu_j H_{2m_j} F_{2m_j} + \begin{bmatrix} F_{2m_j-2} & 0 \\ 0 & 0_2 \end{bmatrix} F_{2m_j} \right),$$

and $(\widehat{\epsilon}_1, \dots, \widehat{\epsilon}_q)$ is an ordered family of signs $+1$ or -1 uniquely determined by (T, H) up to a permutation of signs corresponding to equal Jordan blocks.

Following [7]:

DEFINITION 34. *The collection of signs $(\widehat{\epsilon}_1, \dots, \widehat{\epsilon}_q)$ of (42) is called the sign characteristic of (T, H) under unitary similarity.*

Notice that $\widehat{J}_C = J_C^*$ and $\widehat{J}_R = J_R^*$ where J_C and J_R are the matrices of (39) and (40), respectively. However, an easy computation yields

$$(45) \quad \left. \begin{aligned} & F_k^{-1} = F_k^* = F_k \\ & F_k F_k F_k = F_k \\ & F_{2k} \begin{bmatrix} (\mu_j + i\nu_j)I_k + G_k F_k & 0 \\ 0 & (\mu_j - i\nu_j)I_k + G_k F_k \end{bmatrix} F_{2k} \\ & = \begin{bmatrix} (\mu_j - i\nu_j)I_k + F_k G_k & 0 \\ 0 & (\mu_j + i\nu_j)I_k + F_k G_k \end{bmatrix} \end{aligned} \right\},$$

for any positive integer k . Hence, if $Q = \bigoplus_{j=1}^q F_{l_j} \oplus \bigoplus_{j=1}^s F_{2m_j}$, then $QP_{\widehat{\varepsilon},J}Q = P_{\widehat{\varepsilon},J}$, $Q\widehat{J}_C Q = J_C$ and $Q\widehat{J}_R Q = J_R$. This means that $(J_C, P_{\widehat{\varepsilon},J})$ and $(J_R, P_{\widehat{\varepsilon},J})$ are unitarily similar over \mathbb{C} and \mathbb{R} to $(\widehat{J}_C, P_{\widehat{\varepsilon},J})$ and $(\widehat{J}_R, P_{\widehat{\varepsilon},J})$, respectively. Therefore,

THEOREM 35 (“complex case”). *Let T be H -self-adjoint, as in Theorem 32. Then there is an invertible matrix $S \in \mathbb{C}^{n \times n}$ such that $J_C = S^{-1}TS$ and $P_{\widehat{\varepsilon},J} = S^*HS$ where J_C and $P_{\widehat{\varepsilon},J}$ are the matrices of (39) and (42), respectively, and $(\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_q)$ is an ordered family of signs $+1$ or -1 uniquely determined by (T, H) up to a permutation of signs corresponding to equal Jordan blocks.*

THEOREM 36 (“real case”). *Let $T, H \in \mathbb{R}^{n \times n}$ be H -self-adjoint as in Theorem 33. Then there is an invertible matrix $S \in \mathbb{R}^{n \times n}$ such that $J_R = S^{-1}TS$ and $P_{\widehat{\varepsilon},J} = S^*HS$ where J_R and $P_{\widehat{\varepsilon},J}$ are the matrices of (40) and (42) and $(\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_q)$ is an ordered family of signs $+1$ or -1 uniquely determined by (T, H) up to a permutation of signs corresponding to equal Jordan blocks.*

It follows immediately from Theorems 35 and 36 that $J_C^* P_{\widehat{\varepsilon},J} = P_{\widehat{\varepsilon},J} J_C$ and $J_R^* P_{\widehat{\varepsilon},J} = P_{\widehat{\varepsilon},J} J_R$. This means (Definition 31) that J_C and J_R are $P_{\widehat{\varepsilon},J}$ -self-adjoint over \mathbb{C} and \mathbb{R} , respectively.

6.1. Application to the quadratic problem. Now let $P(\lambda) = M\lambda^2 + D\lambda + K$ be an Hermitian quadratic matrix polynomial (with M nonsingular) and let C_P and A be the matrices of (33) and (34), respectively. Assume, as in the previous section, that $(\lambda - \alpha_j)^{l_j}$ are, for $j = 1, \dots, q$, the elementary divisors of $P(\lambda)$ (and of $\lambda A - AC_P$) for the *real* eigenvalues, and $(\lambda - \mu_j - i\nu_j)^{m_j}$ and $(\lambda - \mu_j + i\nu_j)^{m_j}$ are, for $j = 1, \dots, s$, the elementary divisors for the *nonreal* eigenvalues of $P(\lambda)$ (and of $\lambda A - AC_P$). Then, by Theorem 26, $\lambda A - AC_P$ is congruent to $\lambda P_{\widehat{\varepsilon},J} - P_{\widehat{\varepsilon},J} J_C$ or $\lambda P_{\widehat{\varepsilon},J} - P_{\widehat{\varepsilon},J} J_R$ according as $\lambda A - AC_P$ is complex Hermitian or real symmetric, respectively. Now, $C_P^* A = AC_P$ (see (35)); that is, C_P is A -self-adjoint (over \mathbb{C} or \mathbb{R}). Also, by Corollary 25 there is an invertible matrix $S \in \mathbb{F}^{2n \times 2n}$ such that $P_{\widehat{\varepsilon},J} = S^* A S$, $J_C = S^{-1} C_P S$ or $J_R = S^{-1} C_P S$ according as $\mathbb{F} = \mathbb{C}$ or \mathbb{R} , respectively. This implies that J_C and J_R are $P_{\widehat{\varepsilon},J}$ -self-adjoint over \mathbb{C} and \mathbb{R} , respectively, and, also, that (C_P, A) and $(J_C, P_{\widehat{\varepsilon},J})$ are unitarily similar over \mathbb{C} , and (C_P, A) and $(J_R, P_{\widehat{\varepsilon},J})$ are unitarily similar over \mathbb{R} .

On the other hand, it follows from Theorems 35 or 36 that (C_P, A) is either unitarily similar to $(J_C, P_{\widehat{\varepsilon},J})$ over \mathbb{C} or unitarily similar to $(J_R, P_{\widehat{\varepsilon},J})$ over \mathbb{R} .

We see, as a consequence of the invariance of the sign characteristic of (C_P, A) under unitary similarity, that $\varepsilon = \widehat{\varepsilon}$ up to a permutation of signs corresponding to equal elementary divisors. In conclusion, bearing in mind Proposition 30,

PROPOSITION 37. *Let $P(\lambda) = M\lambda^2 + D\lambda + K$ be an $n \times n$ Hermitian quadratic matrix polynomial with M nonsingular, and let C_P and A be the matrices of (33) and (34), respectively. Then the sign characteristic of (C_P, A) is the ST -sign characteristic of $P(\lambda)$ and this is the sign characteristic of $\lambda A - AC_P$.*

REMARK 38. That the sign characteristic of (C_P, A) is the same as the sign characteristic of $\lambda A - AC_P$ can be obtained directly from Theorem 21 and [7, Theorem 5.10.1] in the complex case and Theorem 20 and [7, Theorem 6.3.1] in the real case. For completeness, we have also studied their relation with the ST-sign characteristic of $P(\lambda)$. \square

7. The quadratic problem: Making the connection.

7.1. Nonsingular leading coefficient. We will show in this section that in Proposition 37 (where $\det M \neq 0$), the sign characteristics of $P(\lambda)$ and $\lambda A - AC_P$ as matrix polynomials coincide. As a consequence, the elementary divisors associated with the real eigenvalues of $P(\lambda)$ and their sign characteristics can be obtained by either a direct application of Rellich's Theorem to $P(\lambda)$ (Theorem 2) or by computing the appropriate canonical form ($L_C(\lambda)$ of (28)), or $L_R(\lambda)$ of (27)) for $\lambda A - AC_P$. Keep in mind that, as we are assuming that $\det M \neq 0$, $P(\lambda)$ has no infinite elementary divisors.

The Gohberg, Lancaster, Rodman approach to the sign characteristic of complex Hermitian and real symmetric matrix polynomials takes the sign characteristic of the pair (C_P, A) as reference (see, for example, [5, Section 1.3], [6, 10.5] or [7, Section 12.4]). This is also the case for [3] and [2]. We will use [7, Section 12.4] because it includes both the complex and real cases and, for convenience, we will restrict the theory to either complex Hermitian or real symmetric quadratic matrix polynomials. Generalization to matrix polynomials of any degree is straightforward.

Given $P(\lambda)$ of (1) with $M^* = M$, $D^* = D$, $K^* = K$, and $\det M \neq 0$, let $(X, T, H^{-1}X^*)$ be a self-adjoint standard tripe of $P(\lambda)$ (see Definition 23) Then (see item (c) of Definition 22) $HT = T^*H$, i.e., T is H -self-adjoint.

In [7, Section 12.4], the “sign characteristic” of $P(\lambda)$ is defined to be the sign characteristic of (T, H) under unitary similarity (see Definition 34). Let us call this collection of signs the *Gohberg-Lancaster-Rodman sign characteristic* (*GLR-sign characteristic, for short*) of $P(\lambda)$. Note that it is well defined because if $(X_1, T_1, H_1^{-1}X_1^*)$ and $(X_2, T_2, H_2^{-1}X_2^*)$ are standard self-adjoint triples of $P(\lambda)$, then, by Definition 23, they are similar. This implies that (T_1, H_1) and (T_2, H_2) are unitarily similar (see Section 6) and so they have the same sign characteristic under unitary similarity (Definition 34 and Theorems 32 and 33).

In summary, we have several “sign characteristics” associated with an Hermitian quadratic matrix polynomial. In fact, let X_0 , C_P , and A be the matrices of (33) and (34). If $(\lambda - \alpha_j)^{l_j}$, $j = 1, \dots, q$, are the elementary divisors of $P(\lambda)$ for the real eigenvalues, we have:

- The *sign characteristic* of $P(\lambda)$ as an analytic matrix function for $\lambda \in \mathbb{R}$, (ρ_1, \dots, ρ_q) , under parameter dependent unitary similarity (Definition 8).
- The *LR-sign characteristic* of the linearization $\lambda A - AC_P$, $(\widehat{\eta}_1, \dots, \widehat{\eta}_q)$, under congruence (Theorems 13 and 14).
- The *inertial signature* of $\lambda A - AC_P$, $(\widetilde{\eta}_1, \dots, \widetilde{\eta}_q)$, under congruence (Theorems 15 and 16).
- The *sign characteristic* of $\lambda A - AC_P$ as an analytic matrix function for $\lambda \in \mathbb{R}$, (η_1, \dots, η_q) , under parameter dependent unitary similarity (Definition 8).
- The *ST-sign characteristic* of $P(\lambda)$, $(\varepsilon_1, \dots, \varepsilon_q)$, under similarity of self-adjoint standard triples (Definition 29).
- The *GLR-sign characteristic* of $P(\lambda)$, $(\widehat{\varepsilon}_1, \dots, \widehat{\varepsilon}_q)$, which is the sign characteristic of (C_P, A) under unitary similarity (Definition 34).

We are now in position to link together all these “sign characteristics.”

First, we have seen in Section 4.3, as a consequence of Corollary 19 and Theorems 20 and 21, that

$$(46) \quad -\tilde{\eta}_j = \eta_j = \begin{cases} \widehat{\eta}_j & \text{if } l_j \text{ odd} \\ -\widehat{\eta}_j & \text{if } l_j \text{ even} \end{cases}, \quad j = 1, \dots, q.$$

On the other hand, by Proposition 30,

$$(47) \quad \varepsilon_j = \eta_j, \quad j = 1, \dots, q.$$

Also, since $(X_0, C_P, A^{-1}X_0)$ is a self-adjoint standard triple of $P(\lambda)$, the *GLR*-sign characteristic of $P(\lambda)$ is the sign characteristic of (C_P, A) which, in turn, by Proposition 37, is the *ST*-sign characteristic of $P(\lambda)$. Hence

$$(48) \quad \widehat{\varepsilon}_j = \varepsilon_j, \quad j = 1, \dots, q.$$

Finally, Theorem 12.5.2 of [7] (see also [5, Theorem 3.7] and [6, Theorem 12.5]) ensures that the *GLR*-sign characteristic of $P(\lambda)$ is the sign characteristic of Definition 8 applied to $P(\lambda)$. We state this result for the readers’ convenience (compare with Theorem 5).

THEOREM 39. *Let $P(\lambda)$ be the Hermitian matrix polynomial of (1) with $\det M \neq 0$ and let $v_1(\lambda), \dots, v_n(\lambda)$ be real analytic functions for real λ such that*

$$\det(v_j(\lambda)I_n - P(\lambda)) = 0, \quad j = 1, \dots, n.$$

Let $\alpha_1 < \alpha_2 < \dots < \alpha_r$ be the distinct real eigenvalues of $P(\lambda)$. For every $i = 1, \dots, r$, there is an α_i such that

$$v_j(\lambda) = (\lambda - \alpha_i)^{m_{ij}} \nu_{ij}(\lambda), \quad m_{ij} \geq 0,$$

*where $\nu_{ij}(\alpha_i) \neq 0$ is real. Then the nonzero numbers among m_{i1}, \dots, m_{in} are the partial multiplicities of $P(\lambda)$ associated with α_i , and $\text{sign } \nu_{ij}(\alpha_i)$ (for $m_{ij} \neq 0$) is the sign attached to the partial multiplicity m_{ij} of $P(\lambda)$ at α_i in its *GLR*-sign characteristic.*

Hence,

$$(49) \quad \widehat{\varepsilon}_j = \rho_j, \quad j = 1, \dots, q.$$

A straightforward consequence of (46)–(49) is that $\eta_j = \rho_j$, $j = 1, \dots, q$. That is to say:

THEOREM 40. *When $\det M \neq 0$, the sign characteristics of $P(\lambda)$ and $\lambda A - AC_P$, as analytic matrix functions for $\lambda \in \mathbb{R}$ (Definition 8), coincide.*

Recalling that in [3] and [2] the sign characteristic of (C_P, A) is taken as definition of the sign characteristic of $P(\lambda)$, we can recover Theorem 5.3 of [3] for the particular case of **quadratic** Hermitian matrix polynomials with nonsingular leading coefficient:

THEOREM 41. *A $2n \times 2n$ Hermitian linear pencil $L(\lambda)$ is an Hermitian strong linearization of $P(\lambda)$ of (1) with $M^* = M$, $D^* = D$, $K^* = K$, and $\det M \neq 0$ that preserves its sign characteristic if and only if it is congruent to $\lambda A - AC_P$.*

Let us recall now that

$$(50) \quad \widehat{L}_P(\lambda) := \lambda A - AC_P = \begin{bmatrix} D & M \\ M & 0 \end{bmatrix} \lambda + \begin{bmatrix} K & 0 \\ 0 & -M \end{bmatrix},$$

is not of the form (6) for any choice of real parameters a_1 and a_2 . However, it is congruent to

$$(51) \quad \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \widehat{L}_P(\lambda) \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix} =: L_P(\lambda),$$

which is, indeed, the linearization of $P(\lambda)$ of the form (6) with $a_1 = 0$ and $a_2 = 1$. It follows then that both $\widehat{L}_P(\lambda)$ and $L_P(\lambda)$ have the same elementary divisors and the same sign characteristic. Actually, it is easily seen using Rellich's theorem (Theorem 2) and (51) that both, $L_P(\lambda)$ and $\widehat{L}_P(\lambda)$ (as Hermitian analytic matrices of $\lambda \in \mathbb{R}$) have the same eigenvalue functions. Thus, when $\det M \neq 0$, $L_P(\lambda)$ of (51) is a linearization of $P(\lambda)$ with the form (6) which preserves its sign characteristic.

Figure 3 depicts the eigenvalue functions of the linearization $L_P(\lambda)$ (and $\widehat{L}_P(\lambda)$) for the real symmetric quadratic matrix polynomial $P(\lambda)$ of (19) whose eigenvalue functions are illustrated in Figure 1.

One may also ask whether all strong linearizations of the form (6) are congruent for any given Hermitian quadratic matrix polynomial. The answer is in the negative. Figures 6 and 7 depict the eigenvalue functions of the linearizations $\lambda A_R + B_R$ of $P(\lambda)$ of (19) with $a_1 = 1/4$, $a_2 = 1/2$ and $a_1 = 1/2$, $a_2 = -1$, respectively. It is apparent that both linearizations do not have the same sign characteristic and so they are not congruent. In particular, if we compare the graphics in Figures 1 and 7, we see that the linearization with $a_1 = 1/2$, $a_2 = -1$ does not reproduce the sign characteristic of $P(\lambda)$.

7.2. Singular leading coefficient. So far (in this section), we have been assuming $\det M \neq 0$ and in this case we have found that the linearization $L_P(\lambda)$ of (51) (or $\widehat{L}_P(\lambda)$ of (50)) preserves the sign characteristic of $P(\lambda)$. In addition, $L_P(\lambda)$ is of the form (6) with $a_1 = 0$ and $a_2 = 1$. Consider now the case $\det M = 0$ but $M \neq 0$ and recall that the pencils of (6) are strong linearizations of $P(\lambda)$ provided that a_1 and a_2 are not both zero and $-\frac{a_2}{a_1} \notin \sigma(P)$. When $\det M = 0$, and $a_1 = 0$, $a_2 = 1$ the latter condition is not satisfied because $P(\lambda)$ has infinity as an eigenvalue. In this case, it follows from [15, Theorem 4.3 (Strong Linearization Theorem)] that $L_P(\lambda)$ is *not* a linearization of $P(\lambda)$.

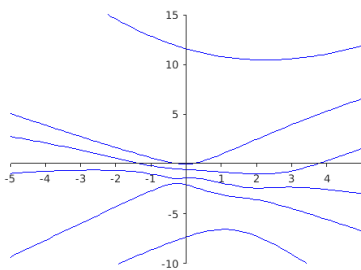


FIGURE 6. Eigenvalue functions of the linearization (6) of $P(\lambda)$ in (19) with $a_1 = \frac{1}{4}$ and $a_2 = \frac{1}{2}$.

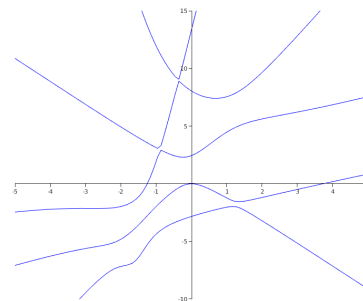


FIGURE 7. Eigenvalue functions of the linearization (6) of $P(\lambda)$ in (19) with $a_1 = \frac{1}{2}$ and $a_2 = -1$.

Nevertheless, Rellich's Theorem 2 can be used to provide strong linearizations of $P(\lambda)$, with the form of (6), which preserve its sign characteristic for *both* the finite elementary divisors associated with real eigenvalues and the infinite elementary divisors. This is the goal of Section 9 where a characterization of such linearizations will be obtained (Theorem 44). This characterization is based on an observation about the relationship between the eigenvalue functions of quadratic Hermitian matrix polynomials and their linearizations of (6) (see (59)) which we analyze in the next section.

8. The eigenvalue functions of $P(\lambda)$ and those of a linearization. Let $P(\lambda)$ be the Hermitian quadratic matrix polynomial of (1). Looking at $P(\lambda)$ and any of its linearizations of (6) as Hermitian analytic matrices of real variable $\lambda \in \mathbb{R}$, we can use Theorem 2 to obtain two Rellich reductions: one for $P(\lambda)$, and one for its a_1, a_2 -dependent, $2n \times 2n$ linearization $\lambda A_C + B_C$ (of equation (8)) or $\lambda A_R + B_R$ (of equation (7)). Recall that we are under the assumptions $a_1 \neq 0$ or $a_2 \neq 0$ and $-\frac{a_2}{a_1} \notin \sigma(P)$. Thus, we have (replace $\lambda A_C + B_C$ by $\lambda A_R + B_R$ in what follows if the coefficient matrices of $P(\lambda)$ are real):

$$(52) \quad P(\lambda) = U(\lambda)V(\lambda)U(\lambda)^* \quad \text{and} \quad \lambda A_C + B_C = H(\lambda)D_C(\lambda)H(\lambda)^*,$$

where $V(\lambda), D_C(\lambda)$ are analytic diagonal matrix functions, and

$$U(\lambda)^*U(\lambda) = I_n, \quad H(\lambda)^*H(\lambda) = I_{2n} \quad \text{for all } \lambda \in \mathbb{R}.$$

Substituting in (9) we obtain

$$(53) \quad H(\lambda)D_C(\lambda)H(\lambda)^* \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix} = \begin{bmatrix} a_1 I_n \\ a_2 I_n \end{bmatrix} U(\lambda)V(\lambda)U(\lambda)^*,$$

Multiply this relation on the left and right by

$$\begin{bmatrix} U(\lambda)^* & 0 \\ 0 & U(\lambda)^* \end{bmatrix} \quad \text{and} \quad U(\lambda),$$

respectively, to obtain

$$(54) \quad \begin{bmatrix} U(\lambda)^* & 0 \\ 0 & U(\lambda)^* \end{bmatrix} H(\lambda)D_C(\lambda)H(\lambda)^* \begin{bmatrix} U(\lambda) & 0 \\ 0 & U(\lambda) \end{bmatrix} \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix} = \begin{bmatrix} a_1 I_n \\ a_2 I_n \end{bmatrix} V(\lambda).$$

Now define the $2n \times 2n$ matrix function

$$(55) \quad G(\lambda) := H(\lambda)^* \begin{bmatrix} U(\lambda) & 0 \\ 0 & U(\lambda) \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

We see that $G(\lambda)^*G(\lambda) = I_{2n}$, ($G(\lambda)$) is analytic unitary) and (53) implies

$$(56) \quad G(\lambda)^*D_C(\lambda)G(\lambda) \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix} = \begin{bmatrix} a_1 I_n \\ a_2 I_n \end{bmatrix} V(\lambda), \quad \lambda \in \mathbb{R}.$$

This statement already provides a link between the n Rellich eigenvalue functions of $P(\lambda)$ (diagonal entries of $V(\lambda)$ on the right) and the $2n$ eigenvalue functions of its linearization $\lambda A_C + B_C$ (diagonal entries of $D_C(\lambda)$ on the left), see (52), but we can do better:

THEOREM 42. Given $P(\lambda)$ of (1) with $M^* = M$, $D = D^*$, $K = K^*$, and the two Rellich reductions of (52), define the $2n \times n$ matrix function

$$(57) \quad W(\lambda) := \frac{1}{\sqrt{1+\lambda^2}} H(\lambda)^* \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix} U(\lambda), \quad \lambda \in \mathbb{R}.$$

Then $W(\lambda)^* W(\lambda) = I_n$ for all $\lambda \in \mathbb{R}$ and the diagonal matrix functions $V(\lambda)$, $D_C(\lambda)$ of (52) satisfy

$$(58) \quad W(\lambda)^* D_C(\lambda) W(\lambda) = \frac{a_1 \lambda + a_2}{1 + \lambda^2} V(\lambda), \quad \lambda \in \mathbb{R}.$$

Proof. Using the fact that, when $\lambda \in \mathbb{R}$, $U(\lambda)^* U(\lambda) = I_n$ and $H(\lambda) H(\lambda)^* = I_{2n}$ the unitary-like property $W(\lambda)^* W(\lambda) = I_n$ is easily verified. Then, using (52) and (57), with $\lambda \in \mathbb{R}$,

$$W(\lambda)^* D_C(\lambda) W(\lambda) = \frac{1}{1 + \lambda^2} \begin{bmatrix} \lambda U(\lambda)^* & U(\lambda)^* \end{bmatrix} H(\lambda) D_C(\lambda) H(\lambda)^* \begin{bmatrix} \lambda U(\lambda) \\ U(\lambda) \end{bmatrix}.$$

Recalling the definition (55), we obtain

$$W(\lambda)^* D_C(\lambda) W(\lambda) = \frac{1}{1 + \lambda^2} \begin{bmatrix} \lambda I_n & I_n \end{bmatrix} G(\lambda)^* D_C(\lambda) G(\lambda) \begin{bmatrix} \lambda I_n \\ I_n \end{bmatrix},$$

with $G(\lambda)^* G(\lambda) = I_{2n}$. Thus, using (56), when $\lambda \in \mathbb{R}$,

$$(59) \quad W(\lambda)^* D_C(\lambda) W(\lambda) = \frac{1}{1 + \lambda^2} \begin{bmatrix} \lambda I_n & I_n \end{bmatrix} \begin{bmatrix} a_1 I_n \\ a_2 I_n \end{bmatrix} V(\lambda) = \frac{a_1 \lambda + a_2}{1 + \lambda^2} V(\lambda),$$

as required. □

A consequence of (58) is that for each $\lambda \in \mathbb{R}$ the eigenvalues of $\lambda A_C + B_C$ and $\frac{a_1 \lambda + a_2}{1 + \lambda^2} P(\lambda)$ interlace (see Corollary 43 below) in a way that we will explain now. This result will not be explicitly used in the developments to come, but it will form the basis for a relationship between the sign characteristics of $P(\lambda)$ and its linearizations of (6) in Section 9.

In equations (52) (cf. equations (16) and (17)), we write the diagonal matrix of eigenvalue functions of $P(\lambda)$ and $\lambda A_C + B_C$ in the form

$$(60) \quad V(\lambda) = \text{Diag}[v_1(\lambda), v_2(\lambda), \dots, v_n(\lambda)],$$

and

$$(61) \quad D_C(\lambda) = \text{Diag}[d_1(\lambda), d_2(\lambda), \dots, d_{2n}(\lambda)],$$

respectively. For each $\lambda \in \mathbb{R}$ let $\delta_1(\lambda) \geq \delta_2(\lambda) \geq \dots \geq \delta_{2n}(\lambda)$ be the ordered eigenvalues of $\lambda A_C + B_C$. That is to say, $\delta_1(\lambda) \geq \delta_2(\lambda) \geq \dots \geq \delta_{2n}(\lambda)$ is for each $\lambda \in \mathbb{R}$ the sequence $(d_1(\lambda), d_2(\lambda), \dots, d_{2n}(\lambda))$ arranged in non-increasing order. Similarly, let $\mu_1(\lambda) \geq \mu_2(\lambda) \geq \dots \geq \mu_n(\lambda)$ be the sequence $\frac{a_1 \lambda + a_2}{1 + \lambda^2} (v_1(\lambda), v_2(\lambda), \dots, v_n(\lambda))$ arranged, for each $\lambda \in \mathbb{R}$, in nonincreasing order. Again, $\mu_1(\lambda) \geq \mu_2(\lambda) \geq \dots \geq \mu_n(\lambda)$ are for each $\lambda \in \mathbb{R}$ the ordered eigenvalues of the Hermitian matrix $\frac{a_1 \lambda + a_2}{1 + \lambda^2} P(\lambda)$. Notice that $\delta_i(\lambda)$ and $\mu_j(\lambda)$ may not be analytic functions for $\lambda \in \mathbb{R}$. For instance, in Figure 8, which shows the eigenvalues of the linearization $\lambda A_R + B_R$ of $P(\lambda)$ of Example 10(b), $\delta_1(\lambda)$, $\delta_2(\lambda)$, $\delta_4(\lambda)$, and $\delta_6(\lambda)$ are analytic functions in \mathbb{R} but $\delta_3(\lambda)$ and $\delta_5(\lambda)$ are not analytic at 1.

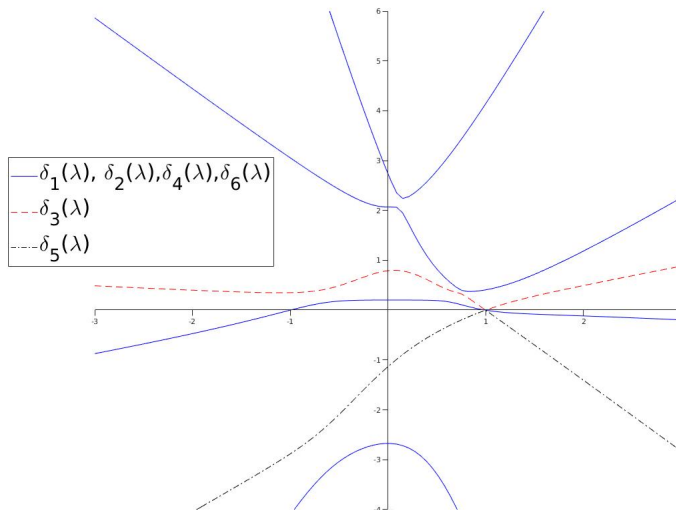


FIGURE 8. Eigenvalues of $\lambda A_R + B_R = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$ where M , D , and K are the coefficients of the quadratic system of Example 10(b).

COROLLARY 43. With the above notation, for each $\lambda \in \mathbb{R}$, the following interlacing inequalities between the eigenvalues of $\lambda A_C + B_C$ and $\frac{a_1 \lambda + a_2}{1 + \lambda^2} P(\lambda)$ are satisfied:

$$(62) \quad \delta_i(\lambda) \geq \mu_i(\lambda) \geq \delta_{i+n}(\lambda), \quad 1 \leq i \leq n.$$

Proof. Condition (58) means that $\frac{a_1 \lambda + a_2}{1 + \lambda^2} V(\lambda)$ is, for each $\lambda \in \mathbb{R}$, the $n \times n$ leading principal submatrix of a $2n \times 2n$ matrix, $A(\lambda)$ say, which is unitarily similar to $D_C(\lambda)$. Hence, $A(\lambda)$ is Hermitian and unitarily similar to $\lambda A_C + B_C$ for each $\lambda \in \mathbb{R}$. Therefore, $\delta_1(\lambda) \geq \dots \geq \delta_{2n}(\lambda)$ are its eigenvalues. On the other hand, $\mu_1(\lambda) \geq \dots \geq \mu_n(\lambda)$ are the eigenvalues of $\frac{a_1 \lambda + a_2}{1 + \lambda^2} V(\lambda)$ for each $\lambda \in \mathbb{R}$. Using the *Poincaré separation theorem* ([9, Corollary 4.3.16]), we get the desired interlacing inequalities. \square

We illustrate with the Examples 10(a) and 10(b) and their linearizations defined by $a_1 = 1/2$, $a_2 = -1$ and $a_1 = 0$, $a_2 = 1$, respectively. The eigenvalue functions of the corresponding $\frac{a_1 \lambda + a_2}{1 + \lambda^2} P(\lambda)$ and $\lambda A_R + B_R$ are depicted in Figures 9 and 10, respectively. It can be seen that the interlacing inequalities (62) are satisfied.

As remarked at the end of Section 7.1, $P(\lambda)$ of Example 10 (a) and its linearization $\lambda A_R + B_R$ of (7) with $a_1 = 1/2$, $a_2 = -1$ do not have the same sign characteristic for the eigenvalues -1.2496 and 0 (see Figures 1 and 7). However, Figure 9 shows that $\frac{a_1 \lambda + a_2}{1 + \lambda^2} P(\lambda)$ and $\lambda A_R + B_R$ do have the same sign characteristic attached to all real eigenvalues. This is not a particular property of these matrices, but a general property of any quadratic $P(\lambda)$ and its linearizations $L(\lambda)$ of (6), as we see in the next section where we explore the relationship between the sign characteristics of $(a_1 \lambda + a_2)P(\lambda)$ and that of the linearizations $L(\lambda)$ of $P(\lambda)$ of the form (6). This will allow us to obtain strong linearizations of $P(\lambda)$ of the form (6) which preserves the sign characteristic of the finite real eigenvalues as well as those at infinity.

9. Strong linearizations of quadratic Hermitian systems preserving the sign characteristic.

Let $P(\lambda)$ be the matrix polynomial of (1) with $M = M^*$, $D = D^*$, and $K = K^*$. We return to the notation

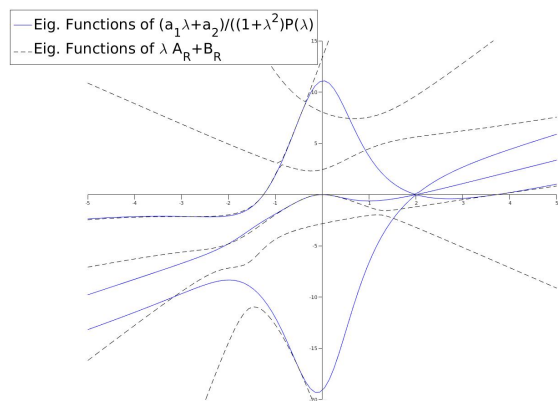


FIGURE 9. Eigenvalue functions of $\frac{a_1\lambda+a_2}{1+\lambda^2}P(\lambda)$ and $\lambda A_R + B_R$ where $P(\lambda)$ is the quadratic matrix polynomial of Example 10 (a) and $\lambda A_R + B_R$ is its linearization of (7) with $a_1 = 1/2$, $a_2 = -1$.

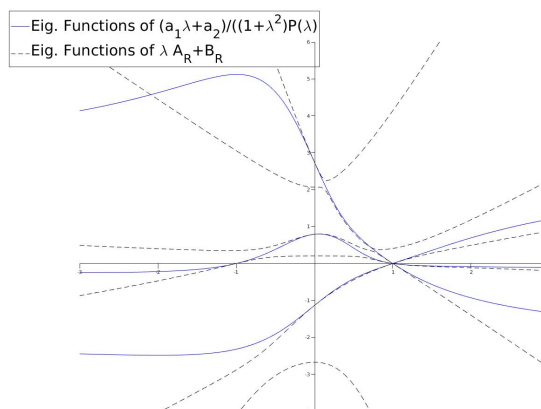


FIGURE 10. Eigenvalue functions of $\frac{a_1\lambda+a_2}{1+\lambda^2}P(\lambda)$ and $\lambda A_R + B_R$ where $P(\lambda)$ is the quadratic matrix polynomial of Example 10 (b) and $\lambda A_R + B_R$ is its linearization of (7) with $a_1 = 0$ and $a_2 = 1$.

introduced in Section 4 and used in the subsequent sections: Thus, $\alpha_1, \dots, \alpha_q \in \mathbb{R}$ denote the real eigenvalues of $P(\lambda)$, including possible repetitions; $(\lambda - \alpha_1)^{l_1}, \dots, (\lambda - \alpha_q)^{l_q}$ are the elementary divisors for the real eigenvalues; and $k_1 \leq k_2 \leq \dots \leq k_r$ are the exponents of the elementary divisors at infinity (or, equivalently, the partial multiplicities of the eigenvalue at infinity).

We recall that if $L(\lambda)$ is a linear pencil of the form (6) then it is a strong linearization of $P(\lambda)$ provided that a_1 and a_2 are not simultaneously zero and $-\frac{a_2}{a_1} \notin \sigma(P)$. In this case, $L(\lambda)$ and $P(\lambda)$ have the same finite and infinite elementary divisors. Let $\epsilon_i(P)$ and $\epsilon_i(L)$ be the sign characteristic of an elementary divisor $(\lambda - \alpha_i)^{l_i}$ for $P(\lambda)$ and $L(\lambda)$, respectively, and let $\epsilon_{i\infty}(P)$ and $\epsilon_{i\infty}(L)$ be the sign characteristics of the elementary divisor at infinity with exponent k_i for $P(\lambda)$ and $L(\lambda)$, respectively.

THEOREM 44. *Let $P(\lambda)$ be the Hermitian quadratic matrix polynomial of (1) and let $L(\lambda)$ be a strong linearization of (6). Then for each real eigenvalue $\alpha_1, \dots, \alpha_q \in \sigma(P) \cap \mathbb{R} = \sigma(L) \cap \mathbb{R}$ the following property holds:*

$$(63) \quad \epsilon_i(L) = \text{sign}(a_1\alpha_i + a_2)\epsilon_i(P), \quad i = 1, \dots, q,$$

and for the eigenvalue at infinity,

$$(64) \quad \epsilon_{i\infty}(L) = \text{sign}(a_1)\epsilon_{i\infty}(P), \quad i = 1, \dots, r.$$

Proof. Rellich's Theorem (Theorem 2) will be applied to both $P(\lambda)$ and $L(\lambda)$. Hence, without further notice, we view these polynomial matrices as Hermitian analytic matrices of the real parameter $\lambda \in \mathbb{R}$.

We assume first that $a_1 \neq 0$ and $a_2 \neq 0$ and recall that $-\frac{a_2}{a_1} \notin \sigma(P)$. Put $X = -\frac{a_1}{a_2}I_n$ and $Y = -\left(\frac{a_1}{a_2}\right)^2 P\left(-\frac{a_2}{a_1}\right)$. Then Y is an $n \times n$ constant invertible matrix and a computation shows that

$$\begin{bmatrix} \lambda I_n & I_n \\ I_n & X \end{bmatrix} L(\lambda) \begin{bmatrix} \lambda I_n & I_n \\ I_n & X \end{bmatrix} = (\lambda a_1 + a_2) \begin{bmatrix} P(\lambda) & 0 \\ 0 & Y \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

Applying Rellich's Theorem to $P(\lambda)$ (see also (52)) there exists a unitary analytic matrix $U(\lambda)$ and an analytic matrix $V(\lambda) = \text{Diag}(v_1(\lambda), \dots, v_n(\lambda))$ such that $P(\lambda) = U(\lambda)V(\lambda)U(\lambda)^*$. Let

$$R(\lambda) = \begin{bmatrix} \lambda I_n & I_n \\ I_n & X \end{bmatrix} \begin{bmatrix} U(\lambda) & 0 \\ 0 & U(\lambda) \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

and observe that $R(\lambda)$ is an Hermitian analytic matrix, $\det R(\lambda) \neq 0$ for $\lambda \neq -\frac{a_2}{a_1}$, and

$$R(\lambda)^* L(\lambda) R(\lambda) = (a_1 \lambda + a_2) \begin{bmatrix} V(\lambda) & 0 \\ 0 & U(\lambda)^* Y U(\lambda) \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

It follows from [6, Theorem 12.4] or [16, Theorem 2.7] that, since $-\frac{a_2}{a_1} \notin \sigma(L)$, $L(\lambda)$ and $R(\lambda)^* L(\lambda) R(\lambda)$ have the same sign characteristic for the finite real eigenvalues of $L(\lambda)$. On the other hand, applying Rellich's Theorem to $U(\lambda)^* Y U(\lambda)$, we conclude that the eigenvalue functions of this analytic matrix function are those of Y . Since this matrix is Hermitian all its eigenvalue functions are constant; actually, they are the eigenvalues of Y . This implies that the sign characteristic of $L(\lambda)$ and $(a_1 \lambda + a_2)V(\lambda)$ coincide and so (63) follows.

As for the eigenvalue at infinity, we must check the sign characteristics, $\epsilon_{i0}(\widehat{P})$ and $\epsilon_{i0}(\widehat{L})$, of the elementary divisors λ^{k_i} of $\widehat{P}(\lambda) = \text{rev } P(\lambda)$ and $\widehat{L}(\lambda) = \text{rev } L(\lambda)$, respectively, $i = 1, \dots, r$. First,

$$\begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} \text{rev } L(\lambda) \begin{bmatrix} 0 & I_n \\ I_n & 0 \end{bmatrix} = \begin{bmatrix} a_2 K & a_1 K \\ a_1 K & a_1 D - a_2 M \end{bmatrix} \lambda + \begin{bmatrix} a_2 D - a_1 K & a_2 M \\ a_2 M & a_1 M \end{bmatrix}$$

which is the linearization of (6) for $\text{rev } P(\lambda) = K\lambda^2 + D\lambda + M$ with the roles of a_1 and a_2 exchanged. Hence, we can apply (63) to \widehat{P} and \widehat{L} for the eigenvalue 0 to obtain

$$\epsilon_{i0}(\widehat{L}) = \text{sign}(a_1) \epsilon_{i0}(\widehat{P}), \quad i = 1, \dots, r.$$

Bearing in mind that the sign characteristic of the eigenvalue at infinity and order k of any matrix polynomial $Q(\lambda)$ is, by definition (see Section 3), the sign characteristic of the elementary divisor λ^k of $-\text{rev } Q(\lambda)$, we get $\epsilon_{i\infty}(L) = -\epsilon_{i0}(\widehat{L}) = \text{sign}(a_1)(-\epsilon_{i0}(\widehat{P})) = \text{sign}(a_1) \epsilon_{i\infty}(P)$. Thus, condition (64) follows.

Assume now that $a_1 = 0$ and $a_2 \neq 0$ (recall that $a_1 = a_2 = 0$ is not possible). Then ∞ is not an eigenvalue of $P(\lambda)$ because $-\frac{a_2}{a_1} \notin \sigma(P)$. Therefore, $\det M \neq 0$. Observe that, in (6), $L(\lambda) = a_1 L_1(\lambda) + a_2 L_2(\lambda)$ where

$$(65) \quad L_1(\lambda) = \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix} \lambda + \begin{bmatrix} D & K \\ K & 0 \end{bmatrix}, \quad L_2(\lambda) = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix} \lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}.$$

Thus, in this case, (6) reduces to $L(\lambda) = a_2 L_2(\lambda)$. It follows from Theorem 40 and (51) that $L_2(\lambda)$ and $P(\lambda)$ have the same sign characteristics for each (finite) elementary divisor. Then, $\epsilon_i(L) = \text{sign}(a_2) \epsilon_i(L_2) = \text{sign}(a_2) \epsilon_i(P)$, $i = 1, \dots, q$. That is to say, (63) is satisfied and (64) does not apply because there are no elementary divisors at infinity.

Finally, consider the case $a_2 = 0$ and $a_1 \neq 0$. Then $0 \notin \sigma(P)$ and so, K is an invertible matrix. Now $L(\lambda) = a_1 L_1(\lambda)$ and then $\widehat{L}(\lambda) = a_1 \widehat{L}_1(\lambda)$ where

$$\widehat{L}_1(\lambda) = \text{rev } L_1(\lambda) = \begin{bmatrix} D & K \\ K & 0 \end{bmatrix} \lambda + \begin{bmatrix} M & 0 \\ 0 & -K \end{bmatrix}.$$

By Theorem 40, $\widehat{L}_1(\lambda)$ has the same sign characteristics for the finite eigenvalues $\frac{1}{\alpha_1}, \dots, \frac{1}{\alpha_q}$ and, if M is singular, 0 as $\widehat{P}(\lambda) = \text{rev } P(\lambda) = K\lambda^2 + D\lambda + M$. According to Tables 1 and 2:

- (i) for $i = 1, \dots, q$, $\epsilon_i(P) = \epsilon_i(\widehat{P})$ if l_i is even and $\epsilon_i(P) = -\epsilon_i(\widehat{P})$ if l_i is odd. Also, $\epsilon_i(L_1) = \text{sign}(\alpha_i)\epsilon_i(\widehat{L}_1)$ if l_i is even and $\epsilon_i(L_1) = -\text{sign}(\alpha_i)\epsilon_i(\widehat{L}_1)$ if l_i is odd, and
- (ii) if M is singular, $\epsilon_{i\infty}(P) = -\epsilon_{i0}(\widehat{P})$ and $\epsilon_{i\infty}(L_1) = -\epsilon_{i0}(\widehat{L}_1)$, $i = 1, \dots, r$.

Therefore, bearing in mind that the sign characteristics of $\widehat{P}(\lambda)$ and $\widehat{L}_1(\lambda)$ coincide,

$$\epsilon_{i\infty}(L) = -\epsilon_{i0}(\widehat{L}) = -\text{sign}(a_1)\epsilon_{i0}(\widehat{L}_1) = -\text{sign}(a_1)\epsilon_{i0}(\widehat{P}) = \text{sign}(a_1)\epsilon_{i\infty}(P).$$

Also, using item (i) above, for $i = 1, \dots, q$,

$$\epsilon_i(P) = \text{sign}(\alpha_i)\epsilon_i(L_1) = \text{sign}(\alpha_i)\text{sign}(a_1)\epsilon_i(L) = \text{sign}(a_1\alpha_i)\epsilon_i(L).$$

The theorem follows. □

As a simple consequence of Theorem 44, we can obtain a characterization of the strong linearizations of $P(\lambda)$ of the form (6) which preserve the sign characteristics of its real eigenvalues and the eigenvalue at infinity. The minimum real eigenvalue of $P(\lambda)$ plays an important role in this characterization. Since the focus is being placed on the sign characteristic, we are implicitly assuming that the eigenvalues of $P(\lambda)$ have been already computed (through a convenient strong linearization, for instance).

COROLLARY 45. *Let $P(\lambda)$ be the Hermitian quadratic matrix polynomial of (1) and let $L(\lambda)$ be its strong linearization of (6) such that $-\frac{a_2}{a_1} \notin \sigma(P)$. Let $m = \min\{\lambda \in \sigma(P) \cap \mathbb{R} : \lambda < \infty\}$. Then $L(\lambda)$ and $P(\lambda)$ have the same sign characteristics for the real eigenvalues and for the eigenvalue at infinity if and only if*

$$(66) \quad a_1 > 0 \text{ and } a_1m + a_2 > 0.$$

Proof. If $L(\lambda)$ and $P(\lambda)$ have the same sign characteristic for the infinite elementary divisor of order k_i , $i = 1, \dots, r$, then $\epsilon_{i\infty}(L) = \epsilon_{i\infty}(P)$. It follows from (64) that $a_1 > 0$. Similarly, for each finite elementary divisor $(\lambda - \alpha_i)^{l_i}$, $\epsilon_i(L) = \epsilon_i(P)$, $i = 1, \dots, q$. By (63) $a_1\alpha_i + a_2 > 0$, $i = 1, \dots, q$. In particular, $a_1m + a_2 > 0$.

Conversely, if $a_1 > 0$, then it follows from (64) that $\epsilon_{i\infty}(L) = \epsilon_{i\infty}(P)$. On the other hand, if $a_1m + a_2 > 0$ then $a_1\alpha_i + a_2 \geq a_1m + a_2 > 0$ and then (63) implies $\epsilon_i(L) = \epsilon_i(P)$. □

It was seen in (51) that when $\det M \neq 0$ the pencil $L_P(\lambda) = \begin{bmatrix} 0 & M \\ M & D \end{bmatrix}\lambda + \begin{bmatrix} -M & 0 \\ 0 & K \end{bmatrix}$ is a linearization of $P(\lambda)$ of (1) which preserves its sign characteristic for the finite elementary divisors. However, by [15, Theorem 4.3 (Strong Linearization Theorem)], $L_P(\lambda)$ is not a linearization of $P(\lambda)$ when $\det M = 0$. It is an interesting and easy consequence of Corollary 45 that this can be seen as a “limit situation.” Specifically, taking $a_2 = 1$ in Corollary 45, and recalling (6), we get:

COROLLARY 46. *With the notation of Corollary 45 let $a \in \mathbb{R}$ be any positive real number satisfying $ma > -1$. Then*

$$(67) \quad L_a(\lambda) = \begin{bmatrix} aM & M \\ M & D - aK \end{bmatrix}\lambda + \begin{bmatrix} aD - M & aK \\ aK & K \end{bmatrix},$$

is an Hermitian strong linearization of $P(\lambda)$ of (1) which preserves the sign characteristics for the elementary divisors associated with real eigenvalues and with the eigenvalue at infinity. In addition,

$$(68) \quad \lim_{a \rightarrow 0^+} L_a(\lambda) = L_P(\lambda),$$

where $L_P(\lambda)$ is the linear pencil of (51).

It is noteworthy that $L_a(\lambda)$ is a strong linearization of $P(\lambda)$ of the form (6) with $a_2 = 1$ and $a_1 = a > 0$, while $L_P(\lambda)$ is a pencil of the form (6) with $a_2 = 1$ and $a_1 = 0$ which *is not* a linearization of $P(\lambda)$ when $\det M = 0$. The limit in (68) tells us that for values of (a_1, a_2) close enough to $(0, 1)$ we can always find strong linearizations of $P(\lambda)$ of the form (6) which preserve its sign characteristics for the finite and infinite elementary divisors. The neighborhood where such values of (a_1, a_2) can be found depends on the smallest real eigenvalue of $P(\lambda)$.

We illustrate the results in this section with an example.

EXAMPLE 47. Consider the reversal of the quadratic system of Example 10(a):

$$Q(\lambda) := \begin{bmatrix} 7 & 8 & 7 \\ 8 & -6 & 8 \\ 7 & 8 & 7 \end{bmatrix} \lambda^2 + \begin{bmatrix} 1 & 4 & 1 \\ 4 & -3 & 4 \\ 1 & 4 & 1 \end{bmatrix} \lambda + \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

which has singular leading coefficient. This is the quadratic system of (21) whose eigenvalues, $\sigma(Q) = \{\infty, -0.1634 \pm 0.2286i, -0.8003, 0.2687\}$, were computed in Example 11. Its minimal real eigenvalue is, approximately, -0.8003 and so, according to Corollary 46, for any $0 < a < \frac{1}{0.8003} \approx 1.2496$, $L_a(\lambda)$ of (67) is a strong linearization of $Q(\lambda)$ which preserves the sign characteristics of the elementary divisors associated with its finite real eigenvalues and with the eigenvalue at infinity.

Let us take $a = 1$ in (67):

$$(69) \quad L_1(\lambda) = \begin{bmatrix} M & M \\ M & D - K \end{bmatrix} \lambda + \begin{bmatrix} D - M & K \\ K & K \end{bmatrix},$$

with

$$M = \begin{bmatrix} 7 & 8 & 7 \\ 8 & -6 & 8 \\ 7 & 8 & 7 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 4 & 1 \\ 4 & -3 & 4 \\ 1 & 4 & 1 \end{bmatrix}, \quad K = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}.$$

The eigenvalue functions of $Q(\lambda)$ for the finite elementary divisors associated with real eigenvalues are depicted in Figure 4 (and in Figure 12 for reader's convenience). And those of $L_1(\lambda)$ in Figure 11. It is clear that the sign characteristic of the eigenvalues -0.8003 and 0.2687 are, for both $Q(\lambda)$ and $L_1(\lambda)$, $+1$ and -1 , respectively.

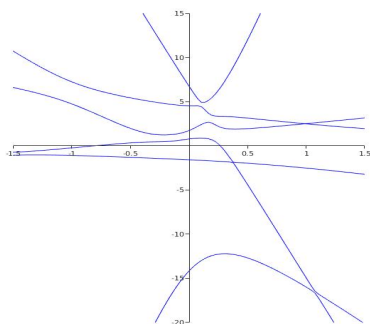


FIGURE 11. Eigenvalue functions of the linearization $L_1(\lambda)$ (cf. (69)) of the symmetric quadratic matrix polynomial $Q(\lambda)$ of Example 47.

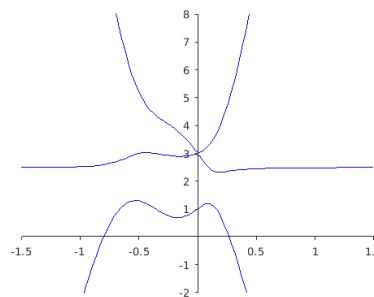


FIGURE 12. Eigenvalue functions of the symmetric quadratic matrix polynomial $Q(\lambda)$ of Example 47.

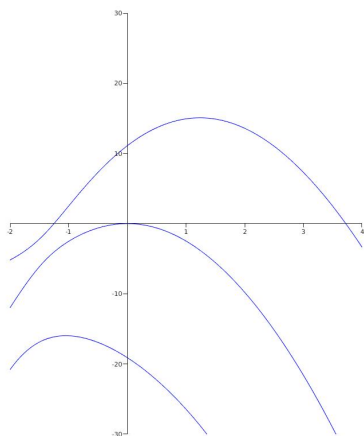


FIGURE 13. *Eigenvalue functions of $-\text{rev } Q(\lambda)$ where $Q(\lambda)$ is the symmetric quadratic matrix polynomial of Example 47.*

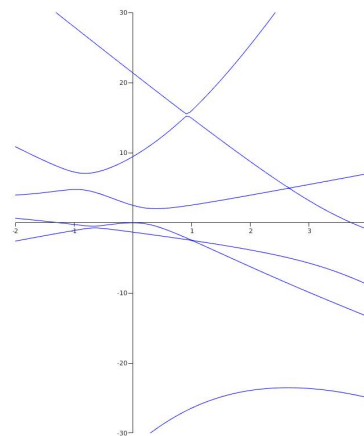


FIGURE 14. *Eigenvalue functions of $-\text{rev } L_1(\lambda)$ where $L_1(\lambda)$ is the linearization in (69) of $Q(\lambda)$ of Example 47.*

For the elementary divisors at infinity, we must compute the elementary divisors at 0 of $-\text{rev } Q(\lambda)$. But $-\text{rev } Q(\lambda) = -P(\lambda)$ where $P(\lambda)$ is the matrix of (19). We saw in Example 10 that 0 is an eigenvalue of $P(\lambda)$ with λ^2 as elementary divisor. Hence, λ^2 is also an elementary divisor of $-\text{rev } Q(\lambda)$ and so $Q(\lambda)$ has an elementary divisor at infinity of order 2. Figures 13 and 14 depict the graphics of the eigenvalue functions of $-\text{rev } Q(\lambda)$ and $-\text{rev } L_1(\lambda)$, respectively. It is apparent that the sign characteristics for the elementary divisor λ^2 of $-\text{rev } Q(\lambda)$ and $-\text{rev } L_1(\lambda)$ coincide and are equal to -1 . Hence, this is the sign characteristic for the elementary divisor at infinity of $Q(\lambda)$ and $L_1(\lambda)$. \square

10. Conclusions. A major technique for both spectral and numerical analysis of quadratic eigenvalue problems requires the formulation of a *linear* eigenvalue problem which, as far as possible, retains spectral properties of the quadratic system. However, there are also methods of spectral and numerical analysis which work *directly* with the quadratic system. The methods of “perturbation theory” are also strongly connected. In this paper, we have made connections between several such lines of attack on these problems.

Eigenvalue problems for symmetric quadratic matrix-valued functions of the form (1) can be approached in two particular ways: either directly, by searching for the values of λ for which $P(\lambda)$ is singular or, indirectly, by first formulating an isospectral symmetric “linearization” of $P(\lambda)$, say $L(\lambda)$ as in (6), and then searching for the values of λ for which $L(\lambda)$ is singular.

In both cases, one may apply symmetry preserving reduction methods to the matrix function (either $P(\lambda)$ or $L(\lambda)$). Namely, either:

- (1) unitary similarity applied to the Hermitian quadratic function $P(\lambda)$ itself (Theorem 2), or
- (2) congruence transformations applied to the Hermitian linear function $L(\lambda)$ (Theorems 20 and 21).

When $\det M \neq 0$, Theorem 40 ensures that the “sign characteristics” (as well as the eigenvalues) of the chosen linearization agree with those of $P(\lambda)$ itself. When M is singular, Corollary 45 provides a characterization of the linearizations which preserves the sign characteristics for both finite and infinite real eigenvalues.

Corollary 43 discloses interesting interlacing inequalities between the eigenvalues of the Hermitian matrices $\frac{a_1\lambda+a_2}{1+\lambda^2}P(\lambda)$ and the linearization $\lambda A_C + B_C$ (or $\lambda A_R + B_R$ if $\mathbb{F} = \mathbb{R}$) of $P(\lambda)$ for each real λ .

Acknowledgments. The authors thank anonymous referees for some suggestions that helped to improve the exposition.

REFERENCES

- [1] S.P. Bhat and D.S. Bernstein. Second-order systems with singular mass matrix and an extension of Guyan reduction. *SIAM J. Matrix Anal. Appl.*, 17:649–657, 1996.
- [2] M.I. Bueno, J. Breen, S. Ford, and S. Furtado. On the sign characteristic of Hermitian linearizations in $DL(P)$. *Linear Algebra Appl.*, 519: 73–101, 2017.
- [3] M.I. Bueno, F.M. Dopico, and S. Furtado. Linearizations of Hermitian matrix polynomials preserving the sign characteristic. *SIAM J. Matrix Anal. Appl.*, 38:249–272, 2017.
- [4] S. Friedland. *Matrices. Algebra, Analysis and Applications*. World Scientific, New Jersey, 2016.
- [5] I. Gohberg, P. Lancaster, and L. Rodman. Spectral analysis of adjoint matrix polynomials. *Ann. Math.*, 112:33–71, 1980.
- [6] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, New York, 1982 and SIAM, Philadelphia, 2009.
- [7] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser, Basel, 2005.
- [8] N. Higham, D.S. Mackey, N. Mackey, and F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159, 2006.
- [9] R.A. Horn and Ch.R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, 1985.
- [10] I. Kaplansky. Elementary divisors and modules. *Trans. Amer. Math. Soc.*, 66:464–491, 1949.
- [11] T. Kato. *Perturbation Theory for Linear Operators*. Springer-Verlag, New York, 1984.
- [12] P. Lancaster and L. Rodman. Canonical forms for Hermitian matrix pairs under strict equivalence and congruence. *SIAM Rev.*, 47(3):407–443, 2005.
- [13] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra Appl.*, 406:1–76, 2005.
- [14] P. Lancaster and I. Zaballa. A review of canonical forms for self-adjoint matrix polynomials. *Operator Theory Adv. Appl.*, 218:425–443, 2012.
- [15] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28:971–1004, 2006.
- [16] V. Mehrmann, V. Noferini, F. Tisseur, and H. Xu. On the sign characteristics of Hermitian matrix polynomials. *Linear Algebra Appl.*, 511:328–364, 2016.
- [17] F. Rellich. *Perturbation Theory of Eigenvalue Problems*. Gordon and Breach, New York, 1969.
- [18] W. Rudin. *Real and Complex Analysis. Third Edition*. McGraw Hill, New York, 1987.
- [19] R.C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.
- [20] H. K. Wimmer. Rellich’s perturbation theorem on Hermitian matrices of holomorphic functions. *J. Math. Anal. Appl.*, 114:52–54, 1986.