Definite Matrix Polynomials and their Linearization by Definite Pencils

Higham, Nicholas J. and Mackey, D. Steven and Tisseur, Françoise

2007

MIMS EPrint: **2007.97**

ISSN 1749-9097

DEFINITE MATRIX POLYNOMIALS AND THEIR LINEARIZATION BY DEFINITE PENCILS[*]

NICHOLAS J. HIGHAM[†], D. STEVEN MACKEY[‡], AND FRANÇOISE TISSEUR[†]

Abstract. Hyperbolic matrix polynomials are an important class of Hermitian matrix polynomials that contain overdamped quadratics as a special case. They share with definite pencils the spectral property that their eigenvalues are real and semisimple. We extend the definition of hyperbolic matrix polynomial in a way that relaxes the requirement of definiteness of the leading coefficient matrix, yielding what we call definite polynomials. We show that this class of polynomials has an elegant characterization in terms of definiteness intervals on the extended real line, and that it includes definite pencils as a special case. A fundamental question is whether a definite matrix polynomial P can be linearized in a structure-preserving way. We show that the answer to this question is affirmative: P is definite if and only if it has a definite linearization in $\mathbb{H}(P)$, a certain vector space of Hermitian pencils; and for definite P we give a complete characterization of all the linearizations in $\mathbb{H}(P)$ that are definite. For the important special case of quadratics, we show how a definite quadratic polynomial can be transformed into a definite linearization with a positive definite leading coefficient matrix—a form that is particularly attractive numerically.

Key words. matrix polynomial, hyperbolic matrix polynomial, matrix pencil, definite pencil, structure-preserving linearization, quadratic eigenvalue problem, polynomial eigenvalue problem

AMS subject classifications. 65F15, 15A18

1. Introduction.

Consider the matrix polynomial of degree ℓ,

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

We will assume throughout that P is regular, that is, $\det P(\lambda) \not\equiv 0$. However, we do not insist that A_ℓ is nonzero, so ℓ is part of the problem specification. The polynomial eigenvalue problem is to find scalars λ and nonzero vectors x and y satisfying $P(\lambda)x = 0$ and $y^*P(\lambda) = 0$; x and y are right and left eigenvectors corresponding to the eigenvalue λ.

A standard way of treating the polynomial eigenvalue problem $P(\lambda)x = 0$, both theoretically and numerically, is to convert it into an equivalent linear matrix pencil $L(\lambda) = \lambda X + Y \in \mathbb{C}^{\ell n\times \ell n}$ by the process known as linearization. Formally, L is a linearization of P if it satisfies

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(\ell-1)n} \end{bmatrix}$$

for some unimodular $E(\lambda)$ and $F(\lambda)$. This implies that $c \cdot \det(L(\lambda)) = \det(P(\lambda))$ for some nonzero constant c, so that L and P have the same eigenvalues. The most

[*]Version of September 30, 2008.

[†]School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK (higham@ma.man.ac.uk, http://www.ma.man.ac.uk/~higham/, ftisseur@ma.man.ac.uk, http://www.ma.man.ac.uk/~ftisseur/). The work of both authors was supported by Engineering and Physical Sciences Research Council grant EP/D079403. The work of the first author was also supported by a Royal Society-Wolfson Research Merit Award.

[‡]Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA (steve.mackey@wmich.edu). This work was supported by National Science Foundation grant DMS-0713799.

widely used linearizations in practice are the companion forms [10, Sec. 14.1]

$$C_i(\lambda) = \lambda X_i + Y_i, \quad i = 1, 2$$

defined by

$$X_1 = X_2 = \mathrm{diag}(A_\ell, I_n, \dots, I_n), \tag{1.2a}$$

$$Y_1 = \begin{bmatrix} A_{\ell-1} & A_{\ell-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -I_n & 0 \end{bmatrix}, \quad Y_2 = \begin{bmatrix} A_{\ell-1} & -I_n & \dots & 0 \\ A_{\ell-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \dots & 0 \end{bmatrix}. \tag{1.2b}$$

In recent work [8], [12] two vector spaces of pencils and their intersection have been studied that generalize the companion forms and provide a systematic way of generating a wide class of linearizations. These spaces make it possible to identify linearizations having specific properties such as optimal conditioning [9], optimal backward error bounds [7], and preservation of structure such as symmetry [8] or palindromic or odd-even structure [11]. In this paper we concentrate on matrix polynomials with symmetric or Hermitian matrix coefficients.

Before discussing our aims, we recall some definitions and terminology. A pencil $L(\lambda) = \lambda X + Y$ is called Hermitian if $X, Y \in \mathbb{C}^{n\times n}$ are Hermitian. We write $A > 0$ to denote that the Hermitian matrix A is positive definite. A Hermitian matrix A is definite if either $A > 0$ or $-A > 0$. Two definite matrices have opposite parity if one is positive definite and the other is negative definite. A sequence $A_0, A_1, A_2, \dots$ of definite matrices has alternating parity if A_j and A_{j+1} have opposite parity for all j.

Definition 1.1 (definite pencil). *A Hermitian pencil $L(\lambda) = \lambda X + Y$ is* definite *(or equivalently, the matrices X, Y form a* definite pair*) if*

$$\gamma(X, Y) := \min_{\substack{z\in\mathbb{C}^n \\ \|z\|_2=1}} \sqrt{(z^*Xz)^2 + (z^*Yz)^2} > 0. \tag{1.3}$$

The quantity $\gamma(X, Y)$ is known as the Crawford number of the pencil.

Definition 1.2 (hyperbolic polynomial). *A matrix polynomial $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ is* hyperbolic *if $A_\ell > 0$ and for every nonzero $x \in \mathbb{C}^n$ the scalar equation $x^*P(\lambda)x = 0$ has ℓ distinct real zeros.*

This latter definition, which can be found in Gohberg, Lancaster, and Rodman [4, Sec. 13.4], for example, does not explicitly require the A_i for $i < \ell$ to be Hermitian. However, the fact that the coefficients $a_j(x) = x^*A_jx$, $j = 0{:}\,\ell$, of the scalar polynomial $x^*P(\lambda)x = \sum_{j=0}^{\ell} \lambda^j a_j(x)$ are real-valued functions of $x \in \mathbb{C}^n$ (since their roots are real and the leading coefficient is real) has this implication.

Definite pencils and hyperbolic polynomials share an important spectral property: all their eigenvalues are real and semisimple. In light of this commonality as well as the possibility of giving them analogous definitions, it is natural to wonder whether every hyperbolic P can be linearized by some definite pencil. In particular, can this always be done with one of the pencils in $\mathbb{H}(P)$, a vector space of Hermitian pencils associated with P studied in [8]? In the case of the quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ with A, B, and C Hermitian and $A > 0$, Barkwell and Lancaster [2] and Veselić [16, Thm. A5] show that Q is hyperbolic if and only if the Hermitian pencil

$$L_2(\lambda) = \lambda \begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & C \end{bmatrix} \tag{1.4}$$

is definite. We will show that L_2 is just one of many definite pencils in $\mathbb{H}(Q)$.

Our contributions in this paper are, first, to propose an extension of the definition of hyperbolicity for a matrix polynomial that retains all the spectral properties of hyperbolic polynomials and in the linear case ($\ell = 1$) is equivalent to definiteness of the pencil. Second, we prove that a Hermitian matrix polynomial $P(\lambda)$ has a definite linearization in $\mathbb{H}(P)$ if and only if P is hyperbolic in this extended sense. For these reasons we will refer to this class of extended hyperbolic polynomials as *definite polynomials*. Furthermore, for a definite P we give a complete characterization of all the linearizations in $\mathbb{H}(P)$ that are definite. Finally we specialize to definite quadratics $Q(\lambda)$. In particular, we explain how Q can be linearized into a definite pencil $L(\lambda) = \lambda X + Y$ with $X > 0$, a form that is particularly attractive numerically.

An important theme of this work is the preservation of the definiteness of a polynomial in the process of linearization. As such, this work continues the study of structure-preserving linearizations begun in [8] and [11].

While our extension of the notion of hyperbolicity is mathematically natural, and fruitful in terms of the various properties it yields, it is also of practical relevance. To explain why, we note that in acoustic fluid-structure interaction problems a non-Hermitian generalized eigenvalue problem arises with pencil of the form [1, Chap. 8]

$$\omega \begin{bmatrix} M_s & 0 \\ M_{fs} & M_f \end{bmatrix} + \begin{bmatrix} K_s & -M_{fs}^* \\ 0 & K_f \end{bmatrix},$$

where $M_s, K_s \in \mathbb{C}^{n\times n}$ and $M_f, K_f \in \mathbb{C}^{m\times m}$ are Hermitian positive definite. Multiplying the first block row by $-\omega$ yields the Hermitian quadratic polynomial

$$Q(\omega) = \omega^2 \begin{bmatrix} -M_s & 0 \\ 0 & 0 \end{bmatrix} + \omega \begin{bmatrix} -K_s & M_{fs}^* \\ M_{fs} & M_f \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & K_f \end{bmatrix}.$$

This polynomial is not hyperbolic because its leading coefficient matrix is not positive definite; nor is the reversed polynomial $\omega^2 Q(1/\omega)$ hyperbolic. However, Q *is* a definite polynomial, so that all the theory developed in this paper applies to it. To see why this is so, observe that for sufficiently small, positive ϵ, the matrix

$$Q(-\epsilon) = \begin{bmatrix} -\epsilon^2 M_s + \epsilon K_s & -\epsilon M_{fs}^* \\ -\epsilon M_{fs} & -\epsilon M_f + K_f \end{bmatrix}$$

is congruent to the positive definite matrix

$$\begin{bmatrix} -\epsilon^2 M_s + \epsilon K_s & 0 \\ 0 & -\epsilon M_f + K_f - \epsilon M_{fs}(K_s - \epsilon M_s)^{-1} M_{fs}^* \end{bmatrix},$$

that is, $Q(-\epsilon) > 0$. Similarly, with P the permutation matrix $\left[\begin{smallmatrix} 0 & I_n \\ I_m & 0 \end{smallmatrix}\right]$ we have

$$\epsilon^2 P^T Q(1/\epsilon) P = \begin{bmatrix} \epsilon M_f + \epsilon^2 K_f & \epsilon M_{fs} \\ \epsilon M_{fs}^* & -M_s - \epsilon K_s \end{bmatrix},$$

which is congruent to

$$\begin{bmatrix} \epsilon M_f + \epsilon^2 K_f & 0 \\ 0 & -M_s - \epsilon K_s - \epsilon M_{fs}^* (M_f + \epsilon K_f)^{-1} M_{fs} \end{bmatrix}.$$

For ϵ negative and sufficiently close to 0 this matrix is negative definite, that is, $Q(\epsilon^{-1}) < 0$. Thus there exist distinct $\epsilon_1 < 0$ and $\epsilon_2 < 0$ such that $Q(\epsilon_1) < 0$ and $Q(\epsilon_2) > 0$. It then follows from Theorem 2.6 below that the quadratic Q is definite.

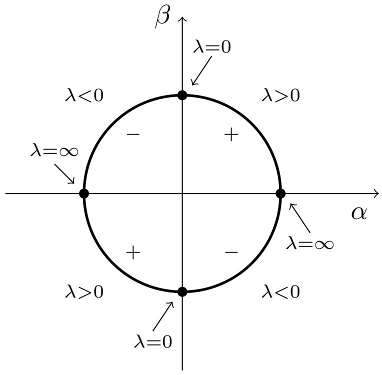


FIG. 2.1. *Correspondence between λ and (α,β). Note the two copies of $\mathbb{R}\cup\{\infty\}$ represented by the upper semicircle and lower semicircle.*

2. Definiteness and hyperbolicity. We will use the homogenous forms of the degree ℓ matrix polynomial $P(\lambda)=\sum_{j=0}^{\ell}\lambda^j A_j$ and pencil $L(\lambda)=\lambda X+Y$, which are given by

$$P(\alpha,\beta)=\sum_{j=0}^{\ell}\alpha^j\beta^{\ell-j}A_j,\qquad L(\alpha,\beta)=\alpha X+\beta Y.$$

Then λ is identified with any pair $(\alpha,\beta)\neq(0,0)$ for which $\lambda=\alpha/\beta$. Without loss of generality we can take $\alpha^2+\beta^2=1$, giving the pictorial representation in Figure 2.1 of the "unit circle" $\mathbb{R}\cup\{\infty\}$ and the correspondence between λ and (α,β). Note that the unit circle contains two copies of $\mathbb{R}\cup\{\infty\}$, since (α,β) and $(-\alpha,-\beta)$ correspond to the same $\lambda\in\mathbb{R}\cup\{\infty\}$.

We say that the matrix polynomial $\widetilde{P}(\widetilde{\alpha},\widetilde{\beta})$ is obtained from $P(\alpha,\beta)$ by homogenous rotation if

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix}=\begin{bmatrix}c & -s\\ s & c\end{bmatrix}\begin{bmatrix}\widetilde{\alpha}\\ \widetilde{\beta}\end{bmatrix},\qquad c,s\in\mathbb{R},\qquad c^2+s^2=1 \tag{2.1}$$

and

$$P(\alpha,\beta)=\sum_{j=0}^{\ell}\alpha^j\beta^{\ell-j}A_j=\sum_{j=0}^{\ell}(c\widetilde{\alpha}-s\widetilde{\beta})^j(s\widetilde{\alpha}+c\widetilde{\beta})^{\ell-j}A_j=:\sum_{j=0}^{\ell}\widetilde{\alpha}^j\widetilde{\beta}^{\ell-j}\widetilde{A}_j:=\widetilde{P}(\widetilde{\alpha},\widetilde{\beta}).$$

It is easily checked that $\widetilde{A}_\ell=P(c,s)$ and $\widetilde{A}_0=P(-s,c)$. Further relationships between P and $\widetilde{P}$ are given in the following lemma.

LEMMA 2.1. *Suppose $\widetilde{P}(\widetilde{\alpha},\widetilde{\beta})$ is obtained from $P(\alpha,\beta)$ by homogenous rotation with (α,β) and $(\widetilde{\alpha},\widetilde{\beta})$ related by* (2.1)*. Then*

(a) *P is positive definite at (α,β) if and only if $\widetilde{P}$ is positive definite at $(\widetilde{\alpha},\widetilde{\beta})$. More generally, the signatures of the matrices $P(\alpha,\beta)$ and $\widetilde{P}(\widetilde{\alpha},\widetilde{\beta})$ are the same.*

(b) *$x^*P(\alpha,\beta)x=x^*\widetilde{P}(\widetilde{\alpha},\widetilde{\beta})x$ for all nonzero $x\in\mathbb{C}^n$.*

(c) *The eigenvectors of P and $\widetilde{P}$ are the same, but the corresponding eigenvalues are rotated.*

Proof. Since for any (α, β) and $(\widetilde{\alpha}, \widetilde{\beta})$ related by (2.1), $P(\alpha, \beta)$ and $\widetilde{P}(\widetilde{\alpha}, \widetilde{\beta})$ are exactly the same matrix, the proof is straightforward. $\square$

Since a Hermitian pencil $L(\lambda) = \lambda X + Y$ is definite if its Crawford number $\gamma(X, Y)$ in (1.3) is strictly positive, it is clear that a sufficient condition for definiteness is that one of X and Y is definite. However, it is the definiteness of a suitable linear combination of X and Y that characterizes definiteness of the pair, as shown by the following lemma, which is essentially contained in [14], [15, Thm. 6.1.18] (see [15, p. 290] for references to earlier work on this topic).

THEOREM 2.2. *A Hermitian pencil $L(\lambda) = \lambda X + Y$ is definite if and only if $L(\mu)$ is a definite matrix for some $\mu \in \mathbb{R} \cup \{\infty\}$, or equivalently if $L(\alpha, \beta) > 0$ for some (α, β) on the unit circle.*

Definite pairs have the desirable properties that they are simultaneously diagonalizable under congruence and, in the associated eigenproblem $L(\lambda)x = 0$, the eigenvalues are real and semisimple [15, Cor. 6.1.19].

Recall that, by definition, for a hyperbolic matrix polynomial $P(\lambda) = \sum_{i=0}^{\ell} \lambda^i A_i$ the equation $x^* P(\lambda) x = 0$ has ℓ distinct real zeros for $x \neq 0$, and hence P has real eigenvalues. For such a P let

$$\lambda_1(x) > \lambda_2(x) > \cdots > \lambda_\ell(x)$$

be the roots of $x^* P(\lambda) x$ for some nonzero $x \in \mathbb{C}^n$. Markus [13, §31] (see also [4, Sec. 13.4]) shows that the eigenvalues of P are distributed in ℓ disjoint closed intervals

$$\mathcal{I}_j = \{\lambda_j(x) : \ x \in \mathbb{C}^n, \ \|x\|_2 = 1\}, \qquad j = 1{:}\,\ell. \tag{2.2}$$

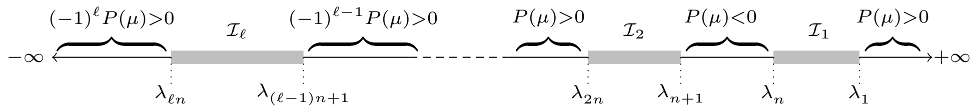
Markus [13, Lem. 31.15] gives, moreover, the following characterization of hyperbolicity. In this result and below, we write $P(\mu)$ when we are considering the definiteness of the matrix $P(\mu) \in \mathbb{C}^{n \times n}$, and reserve the notation $P(\lambda)$ for the matrix polynomial.

THEOREM 2.3 (Markus's characterization of hyperbolicity). *Let $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ be a Hermitian matrix polynomial of degree $\ell > 1$ with $A_\ell > 0$. Then P is hyperbolic if and only if there exist $\mu_j \in \mathbb{R}$ such that*

$$(-1)^j P(\mu_j) > 0, \qquad j = 1{:}\,\ell - 1, \qquad \mu_1 > \mu_2 > \cdots > \mu_{\ell-1}. \tag{2.3}$$

These properties combine to give a useful "definiteness diagram" that summarizes many of the key properties of hyperbolic polynomials.

THEOREM 2.4. *A hyperbolic polynomial $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$, with eigenvalues $\lambda_{\ell n} \leq \cdots \leq \lambda_1$, has the properties displayed in the following diagram, where the closed shaded intervals are the $\mathcal{I}_j$ defined in* (2.2) *and the matrix $P(\mu)$ is indefinite* (*i.e., has both positive and negative eigenvalues*) *on the interiors of these intervals:*



Proof. The proof is essentially a matter of counting sign changes in eigenvalues. Since $A_\ell > 0$, $P(\mu) > 0$ for all sufficiently large μ. Since λ_1 is the largest eigenvalue of the polynomial $P(\lambda)$, $P(\lambda_1)$ is singular and $P(\mu)$ is nonsingular for $\mu > \lambda_1$. Hence we must have $P(\mu) > 0$ for $\mu > \lambda_1$. Likewise, $(-1)^\ell P(\mu) > 0$ for $\mu < \lambda_{\ell n}$.

As μ decreases from λ_1 the inertia of the matrix $P(\mu)$ changes only at an eigenvalue of $P(\lambda)$, and at a k-fold eigenvalue of $P(\lambda)$ the number of negative eigenvalues of $P(\mu)$ can increase by at most k. Hence any number μ_1 for which $P(\mu_1) < 0$ satisfies $\mu_1 < \lambda_n$. Similarly, any number $\mu_{\ell-1}$ for which $(-1)^{\ell-1}P(\mu_{\ell-1}) > 0$ satisfies $\lambda_{(\ell-1)n+1} < \mu_{\ell-1}$. By continuing this argument we find that the points $\mu_{\ell-1}, \dots, \mu_1$ satisfying (2.3) can be accommodated in the diagram only by placing them in the bracketed intervals and that P must be indefinite inside each of the shaded intervals. Finally, it is easily seen that the shaded intervals are precisely the $\mathcal{I}_j$ in (2.2). ☐

Note that hyperbolic pencils $L(\lambda) = \lambda X + Y$ are definite since their coefficient matrices are Hermitian with $X > 0$. However a definite pair is not necessarily hyperbolic in the standard sense since X and Y can both be indefinite. We can now make a definition of definite polynomial that extends the notion of hyperbolicity and is consistent with the definition of definite pencil.

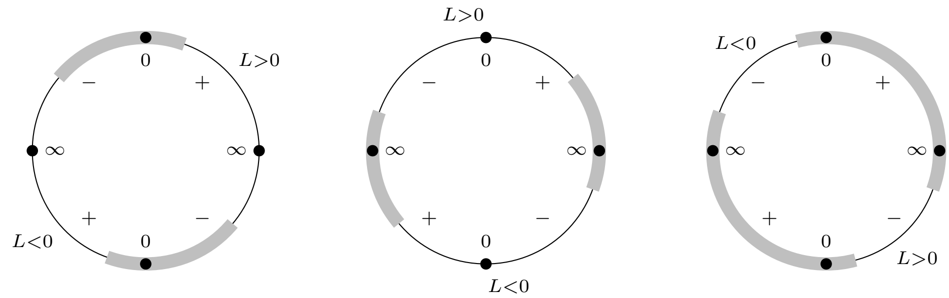
Definition 2.5 (definite polynomial). *A Hermitian matrix polynomial* $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ *is* definite *if there exists* $\mu \in \mathbb{R} \cup \{\infty\}$ *such that the matrix* $P(\mu)$ *is definite and for every nonzero* $x \in \mathbb{C}^n$ *the scalar equation* $x^* P(\lambda) x = 0$ *has* ℓ *distinct zeros in* $\mathbb{R} \cup \{\infty\}$.

To see the consistency of the definition take a definite P and homogeneously rotate P into $\widetilde{P}$ so that μ corresponds to ∞, that is, $\mu = \alpha/\beta$ corresponds to $\widetilde{\mu} = \widetilde{\alpha}/\widetilde{\beta} = 1/0 = \infty$ (this can be done by setting $c = \alpha$ and $s = \beta$ in (2.1)); then by Lemma 2.1, $\widetilde{P}(\widetilde{\mu}) = \widetilde{A}_\ell > 0$. If $x^* P(\lambda) x = 0$ has distinct roots in $\mathbb{R} \cup \{\infty\}$ then $x^* \widetilde{P}(\widetilde{\lambda}) x = 0$ has real distinct zeros (and no infinite root since $\widetilde{A}_\ell > 0$). Here we adopt the convention that $x^* P(\lambda) x = \sum_{j=0}^{\ell} a_j(x) \lambda^j$ has a root at ∞ whenever $a_\ell(x) = 0$. Thus $\widetilde{P}(\widetilde{\lambda})$ is hyperbolic. Hence any definite matrix polynomial is actually a "homogeneously rotated" hyperbolic matrix polynomial. By Lemma 2.1 all the spectral and definiteness properties of hyperbolic polynomials are inherited by definite polynomials as long as we interpret "intervals" *homogeneously* on the unit circle; see Figure 2.2.

Note that for a matrix polynomial $P(\alpha, \beta)$ of degree ℓ in homogeneous variables α, β we have $P(-\alpha, -\beta) = (-1)^\ell P(\alpha, \beta)$. Thus for odd degree Hermitian polynomials $P(-\alpha, -\beta)$ and $P(\alpha, \beta)$ have opposite signature for any (α, β) on the unit circle, and hence an antisymmetric (through the origin) definiteness diagram (see Figure 2.2 (a)). An even degree Hermitian polynomial has $P(-\alpha, -\beta)$ and $P(\alpha, \beta)$ with the same signature for any (α, β) on the unit circle, and hence a symmetric (through the origin) definiteness diagram (see Figure 2.2 (b)). By virtue of this (anti)symmetry the left-most definiteness semi-interval in Theorem 2.4, that is, $\{\mu < \lambda_{\ell n} : \ (-1)^\ell P(\mu) > 0\}$ together with the right-most definiteness semi-interval $\{\mu > \lambda_1 : \ P(\mu) > 0\}$ can be considered as one definiteness interval and any hyperbolic matrix polynomial of degree ℓ can be regarded as having ℓ distinct definiteness intervals (and not $\ell + 1$ as the diagram of Theorem 2.4 might suggest).

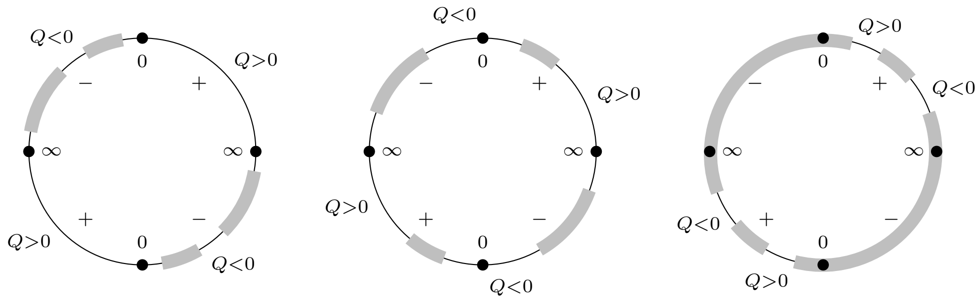
In view of the connection between hyperbolic polynomials and definite polynomials via homogeneous rotation we have the following extension of Markus's characterization of hyperbolic matrix polynomials in Theorem 2.3. Unlike the latter result, ours is valid for $\ell = 1$.

Theorem 2.6 (characterization of definite matrix polynomial). *A Hermitian matrix polynomial* $P(\lambda) = \sum_{j=0}^{\ell} \lambda^j A_j$ *is definite if and only if there exist* $\gamma_j \in \mathbb{R} \cup \{\infty\}$ *with* $\gamma_0 > \gamma_1 > \gamma_2 > \cdots > \gamma_{\ell-1}$ *(*$\gamma_0 = \infty$ *being possible) such that* $P(\gamma_0)$, $P(\gamma_1)$, …, $P(\gamma_{\ell-1})$ *are definite matrices with alternating parity.*



$X > 0$, Y indefinite $\qquad$ X indefinite, $Y > 0$ $\qquad$ X, Y both indefinite

(a) Definite pencils $L(\alpha, \beta) = \alpha X + \beta Y$.



$A > 0$, $B > 0$, $C > 0$ (overdamped) $\qquad$ $A > 0$, $C < 0$ $\qquad$ A, C both indefinite

(b) Definite quadratics $Q(\alpha, \beta) = \alpha^2 A + \alpha\beta B + \beta^2 C$.

FIG. 2.2. *Examples of definiteness diagrams for* (*a*) *definite pencils and* (*b*) *definite quadratics. The shaded arcs are the arcs of indefiniteness.*

3. Hermitian linearizations. We recall some definitions and results from [8], [12]. With the notation

$$\Lambda = [\lambda^{\ell-1}, \lambda^{\ell-2}, \dots, 1]^T \in \mathbb{F}^\ell, \tag{3.1}$$

where $\ell = \deg(P)$ and $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$, define two vector spaces of $\ell n \times \ell n$ pencils $L(\lambda) = \lambda X + Y$:

$$\mathbb{L}_1(P) = \{\, L(\lambda) : L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda),\ v \in \mathbb{F}^\ell \,\}, \tag{3.2}$$

$$\mathbb{L}_2(P) = \{\, L(\lambda) : (\Lambda^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda),\ w \in \mathbb{F}^\ell \,\}, \tag{3.3}$$

where $\otimes$ is the Kronecker product [10, Chap. 12]. The vectors v and w are referred to as "right ansatz" and "left ansatz" vectors, respectively. It is easily checked that for the companion forms in (1.2), $C_1(\lambda) \in \mathbb{L}_1(P)$ with $v = e_1$ and $C_2(\lambda) \in \mathbb{L}_2(P)$ with $w = e_1$, where e_i denotes the ith column of I_ℓ. For any regular P almost all pencils in $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are linearizations of P [12, Thm. 4.7]. The intersection

$$\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P) \tag{3.4}$$

is of particular interest, because there is a simultaneous correspondence via Kronecker products between left and right eigenvectors of P and those of pencils in $\mathbb{DL}(P)$. Two key facts are that $L \in \mathbb{DL}(P)$ if and only if L satisfies the conditions in (3.2) and (3.3) with $w = v$, and that every $v \in \mathbb{F}^\ell$ uniquely determines X and Y such that $L(\lambda) = \lambda X + Y$ is in $\mathbb{DL}(P)$ [8, Thm. 3.4], [12, Thm. 5.3]. Thus $\mathbb{DL}(P)$ is an ℓ-dimensional space of pencils associated with P. Just as for $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$, almost all pencils in $\mathbb{DL}(P)$ are linearizations when P is regular [12, Thm. 6.8]. Define the block Hankel, block $j \times j$ matrices

$$\mathcal{L}_j = \begin{bmatrix} & & & A_\ell \\ & & \cdot^{\cdot^{\cdot}} & A_{\ell-1} \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ A_\ell & A_{\ell-1} & \dots & A_{\ell-j+1} \end{bmatrix}, \qquad \mathcal{U}_j = \begin{bmatrix} A_{j-1} & \dots & A_1 & A_0 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ A_1 & \cdot^{\cdot^{\cdot}} & & \\ A_0 & & & \end{bmatrix}.$$

It is shown in [8, Thm. 3.5] that the jth standard basis pencil in $\mathbb{DL}(P)$ with ansatz vector e_j ($j = 1\colon \ell$) can be expressed as

$$L_j(\lambda) = \lambda X_j - X_{j-1}, \qquad X_j = \begin{bmatrix} \mathcal{L}_j & 0 \\ 0 & -\mathcal{U}_{\ell-j} \end{bmatrix}. \tag{3.5}$$

($\mathcal{L}_j$ and $\mathcal{U}_j$ are taken to be void when $j = 0$.)

Now for a Hermitian matrix polynomial $P(\lambda)$ of degree ℓ, let

$$\mathbb{H}(P) := \left\{ \lambda X + Y \in \mathbb{L}_1(P) : X^* = X,\ Y^* = Y \right\} \tag{3.6}$$

denote the set of all Hermitian pencils in $\mathbb{L}_1(P)$. It is shown in [8, Thm. 6.1] that $\mathbb{H}(P)$ is the subset of all pencils in $\mathbb{DL}(P)$ with a real ansatz vector. In other words, for each vector $v \in \mathbb{R}^\ell$ there is a unique Hermitian pencil in $\mathbb{H}(P)$ defined by $\sum_{j=1}^{\ell} v_j(\lambda X_j + X_{j-1})$ with X_j as in (3.5), and every pencil in $\mathbb{H}(P)$ can be written in this way.

4. Definite linearizations. In this section, $L(\lambda) = \lambda X + Y$ is an element of $\mathbb{H}(P)$ with ansatz vector $v \in \mathbb{R}^\ell$, where ℓ is the degree of P. We begin with the statement of our two main results.

THEOREM 4.1 (existence of definite linearizations). *A Hermitian matrix polynomial $P(\lambda)$ has a definite linearization in $\mathbb{H}(P)$ if and only if P is definite.*

We denote by $\mathcal{D}(P)$ the subset of all definite pencils in $\mathbb{H}(P)$, i.e.,

$$\mathcal{D}(P) = \{\, L \in \mathbb{H}(P) :\ L \text{ is a definite pencil} \,\} \subseteq \mathbb{H}(P).$$

Note that since every definite pencil is regular, by [12, Thm. 4.3] any pencil in $\mathcal{D}(P)$ is automatically a definite linearization of P. With a vector $v \in \mathbb{R}^\ell$ we associate a scalar polynomial

$$\mathsf{p}(\lambda; v) := v^T \Lambda = \sum_{i=1}^{\ell} v_i \lambda^{\ell-i},$$

referred to as the v-polynomial. We adopt the convention that $\mathsf{p}(\lambda; v)$ has a root at ∞ whenever $v_1 = 0$.

THEOREM 4.2 (characterization of $\mathcal{D}(P)$). *Suppose the Hermitian matrix polynomial P of degree ℓ is definite and $L(\lambda) = \lambda X + Y \in \mathbb{H}(P)$ with ansatz vector $v \in \mathbb{R}^\ell$. Then $L \in \mathcal{D}(P)$ if and only if the roots of $\mathsf{p}(x; v)$, including ∞ if $v_1 = 0$, are real, simple* (*i.e., of multiplicity* 1), *and lie in distinct definiteness intervals for P. Moreover, $L(\alpha, \beta) = \alpha X + \beta Y$ is a definite matrix if and only if $L \in \mathcal{D}(P)$ and (α, β) lies in the one definiteness interval for P that is not occupied by a root of $\mathsf{p}(x; v)$.*

The rest of this section is devoted to the proofs of these two theorems.

4.1. Ansatz vector conditions. The first task is to eliminate from further consideration any ansatz vector $v \in \mathbb{R}^\ell$ such that $\mathsf{p}(x; v)$ has either a complex (nonreal) root or a real root (including ∞) with multiplicity 2 or greater. Synthetic division carried out by Horner's method is one of the key ingredients for this task.

LEMMA 4.3. *Let $p(x) = a_0 + a_1 x + \cdots + a_m x^m$ and denote by $q_i(s)$, $i = 0{:}\,m$, the scalars generated by Horner's method for evaluating the polynomial p at a point s, that is,*

$$q_0(s) = a_m, \quad q_i(s) = s q_{i-1}(s) + a_{m-i}, \quad i = 1{:}\,m.$$

Define the degree $m-1$ polynomial

$$\widehat{q}_s(x) := q_{m-1}(s) + q_{m-2}(s)x + \cdots + q_k(s)x^{m-k-1} + \cdots + q_1(s)x^{m-2} + q_0(s)x^{m-1}.$$

Then

(a) *$p(x) = (x-s)\widehat{q}_s(x) + p(s)$.*

(b) *If s is a root of $p(x)$ then $p(x) = (x-s)\widehat{q}_s(x)$ and $\widehat{q}_s(x) = a_m \prod_{i=1}^{m-1}(x - r_i)$, where $r_1, r_2, \ldots, r_{m-1}$ are the other roots of p.*

(c) *If r and s are distinct roots of $p(x)$, then $\widehat{q}_s(r) = 0$.*

(d) *s is a root of p with multiplicity at least two if and only if $\widehat{q}_s(s) = 0$.*

Proof. (a) is a standard identity for Horner's method; see, e.g., [6, Sec. 5.2]. (a) $\Rightarrow$ (b) is immediate, and (b) implies (c) and (d). ☐

In what follows we often use Λ in (3.1) with an argument:

$$\Lambda(r) = [r^{\ell-1}, r^{\ell-2}, \ldots, 1]^T.$$

We will need to refer to the following result from [12, Lem. 6.5].

LEMMA 4.4. *Suppose that $L(\lambda) \in \mathbb{DL}(P)$ with ansatz vector v and $\mathsf{p}(x; v)$ is the* v*-polynomial of v. Let Y_j denote the jth block column of Y in $L(\lambda) = \lambda X + Y$, where $1 \le j \le \ell - 1$. Then*

$$\big(\Lambda^T(x) \otimes I\big) Y_j = q_{j-1}(x; v) P(x) - x\,\mathsf{p}(x; v) P_{j-1}(x),$$

where $q_{j-1}(x; v)$ and $P_{j-1}(x)$ are the scalar and matrix generated by Horner's method for evaluating $\mathsf{p}(x; v)$ and $P(x)$, respectively, as in Lemma 4.3.

We recall an operation on block matrices introduced in [12] that is useful for constructing pencils in $\mathbb{L}_1(P)$. For block $\ell \times \ell$ matrices X and Y with $n \times n$ blocks X_{ij} and Y_{ij}, the column-shifted sum $X \boxplus\!\!\rightarrow Y$ of X and Y is defined by

$$X \boxplus\!\!\rightarrow Y := \begin{bmatrix} X_{11} & \ldots & X_{1\ell} & 0 \\ \vdots & & \vdots & \vdots \\ X_{\ell 1} & \ldots & X_{\ell\ell} & 0 \end{bmatrix} + \begin{bmatrix} 0 & Y_{11} & \ldots & Y_{1\ell} \\ \vdots & \vdots & & \vdots \\ 0 & Y_{\ell 1} & \ldots & Y_{\ell\ell} \end{bmatrix} \in \mathbb{F}^{\ell n \times \ell(n+1)},$$

where the zero blocks are $n \times n$. It is shown in [12, Lem. 3.4] that

$$L(\lambda) \in \mathbb{L}_1(P) \text{ with ansatz vector } v \in \mathbb{F}^\ell \;\Leftrightarrow\; X \boxplus\!\!\rightarrow Y = v \otimes [A_\ell \; A_{\ell-1} \; \ldots \; A_0]. \tag{4.1}$$

We can now prove the following result.

LEMMA 4.5. *Suppose P is a Hermitian matrix polynomial and $L \in \mathbb{H}(P)$ with ansatz vector v. If either*

(a) *the* v*-polynomial $\mathsf{p}(x; v)$ has a* (*nonreal*) *complex conjugate pair of roots s and $\overline{s}$, or*

(b) $\mathsf{p}(x;v)$ *has a real root* s *with multiplicity at least* 2, *or*

(c) ∞ *is a root of* $\mathsf{p}(x;v)$ *with multiplicity at least* 2, *i.e.,* $v_1 = v_2 = 0$,

then L *is not a definite pencil.*

Proof. To prove parts (a) and (b) we use a congruence transformation $S(\alpha X + \beta Y)S^*$ so as to preserve and reveal the nondefiniteness of $L(\alpha,\beta) = \alpha X + \beta Y$. Let

$$S = \left[\begin{array}{c} I_{\ell-1} \quad 0 \\ \hline \Lambda^T(s) \end{array}\right] \otimes I_n,$$

where s is the given complex (or multiple real) root of $\mathsf{p}(x;v)$. Using Lemma 4.4 together with $\mathsf{p}(s;v) = 0$, we see that the bottom block row of SY has the form

$$(SY)_{\ell,:} = [\, q_0(s;v)P(s) \quad q_1(s;v)P(s) \quad \dots \quad q_{\ell-2}(s;v)P(s) \quad \boxed{*} \,],$$

where the scalars $q_j(s;v)$, $j = 0\colon \ell-2$ are generated by Horner's method for $\mathsf{p}(s;v)$ as in Lemma 4.3. Observe that $S \cdot L(\lambda)$ is no longer in $\mathbb{H}(P)$ but is still in $\mathbb{L}_1(P)$, since

$$\begin{aligned} S\cdot L(\lambda)(\Lambda \otimes I) = S\cdot (v \otimes P(\lambda)) &= \left(\left[\begin{array}{c} I_{\ell-1} \quad 0 \\ \hline \Lambda^T(s) \end{array}\right] \otimes I_n \right)(v \otimes P(\lambda)) \\ &= \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{\ell-1} \\ \mathsf{p}(s;v) \end{bmatrix} \otimes P(\lambda) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_{\ell-1} \\ 0 \end{bmatrix} \otimes P(\lambda), \end{aligned}$$

so that we can use the column shifted sum to deduce the structure of SX and SY. Since the right ansatz vector of $S\cdot L(\lambda)$ is $[v_1,\dots,v_{\ell-1},0]^T$, we know from (4.1) that the bottom block row of the shifted sum $SX \boxplus\!\!\rightarrow SY$ must be zero. Thus the bottom block rows of SX and SY are

$$\begin{aligned} (SX)_{\ell,:} &= -\,[\,0 \quad q_0(s;v)P(s) \quad q_1(s;v)P(s) \quad \dots \quad q_{\ell-2}(s;v)P(s)\,], \\ (SY)_{\ell,:} &= [\,q_0(s;v)P(s) \quad q_1(s;v)P(s) \quad \dots \quad q_{\ell-2}(s;v)P(s) \quad 0\,]. \end{aligned}$$

From this we can now compute the (ℓ,ℓ)-blocks of SXS^* and SYS^*,

$$\begin{aligned} (SXS^*)_{\ell,\ell} &= (SX)_{\ell,:}\, S^*_{:,\ell} = (SX)_{\ell,:}\,(\Lambda(\overline{s}) \otimes I_n) = -\widehat{q}_s(\overline{s};v)P(s), \\ (SYS^*)_{\ell,\ell} &= (SY)_{\ell,:}\,(\Lambda(\overline{s}) \otimes I_n) = \overline{s}\,\widehat{q}_s(\overline{s};v)P(s), \end{aligned}$$

where $\widehat{q}_s(x;v) = \sum_{j=0}^{\ell-2} q_j(s;v)x^{\ell-j-2}$. But $\widehat{q}_s(\overline{s};v) = 0$ by Lemma 4.3 (c) for part (a) or Lemma 4.3 (d) for part (b). Thus $[\,S(\alpha X + \beta Y)S^*\,]_{\ell,\ell} = 0$ for all (α,β) on the unit circle, showing that $\alpha X + \beta Y$ is not a definite pencil.

For the proof of part (c) we observe from (3.5) that for the standard basis pencils $L_3, L_4, \dots, L_\ell$ for $\mathbb{H}(P)$ with ansatz vectors $e_3, e_4, \dots, e_\ell$, the $(1,1)$ block is identically 0. Thus for any ansatz vector $v \in \mathbb{R}^\ell$ with $v_1 = v_2 = 0$, the corresponding $L(\lambda) = \lambda X + Y \in \mathbb{H}(P)$ has zero $(1,1)$ blocks in X and Y, so that $(\alpha X + \beta Y)_{1,1} \equiv 0$ for all α,β. Hence $\alpha X + \beta Y$ is not a definite matrix for any (α,β) on the unit circle and L is therefore not a definite pencil. ▯

In light of Lemma 4.5, we now assume throughout the remainder of section 4 that the ansatz vector $v \in \mathbb{R}^\ell$ is such that $\mathsf{p}(x;v)$ has $\ell - 1$ distinct real roots (including possibly ∞ when $v_1 = 0$).

The second major step in the proof of Theorems 4.1 and 4.2 is to block-diagonalize the pencil $\alpha X+\beta Y$ by a "definiteness-revealing" congruence. The generic case $v_1 \neq 0$ is treated first in its entirety; then we come back to look at $v_1 = 0$, $v_2 \neq 0$, and see which parts of the argument for $v_1 \neq 0$ must be modified to handle this case.

4.2. Case 1: $\boldsymbol{v_1 \neq 0}$. As before, $L(\lambda) \in \mathbb{H}(P) \subset \mathbb{DL}(P)$. Let $r_1, r_2, \ldots, r_{\ell-1}$ be the finite real and distinct roots of the v-polynomial $\mathsf{p}(x;v)$. We start the reduction of $\alpha X+\beta Y$ to block diagonal form by nonsingular congruence with the matrix

$$S = \begin{bmatrix} e_1 & \Lambda(r_1) & \Lambda(r_2) & \ldots & \Lambda(r_{\ell-1}) \end{bmatrix}^T \otimes I_n. \tag{4.2}$$

It is easy to verify that $S \cdot L(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector $v_1 e_1$. SY has the form

$$SY = \begin{bmatrix} v_1A_{\ell-1} - v_2A_\ell & v_1A_{\ell-2} - v_3A_\ell & \ldots & v_1A_1 - v_\ell A_\ell & v_1A_0 \\ q_0(r_1;v)P(r_1) & q_1(r_1;v)P(r_1) & \ldots & q_{\ell-2}(r_1;v)P(r_1) & \boxed{*} \\ q_0(r_2;v)P(r_2) & q_1(r_2;v)P(r_2) & \ldots & q_{\ell-2}(r_2;v)P(r_2) & \boxed{*} \\ \vdots & \vdots & & \vdots & \vdots \\ q_0(r_{\ell-1};v)P(r_{\ell-1}) & q_1(r_{\ell-1};v)P(r_{\ell-1}) & \ldots & q_{\ell-2}(r_{\ell-1};v)P(r_{\ell-1}) & \boxed{*} \end{bmatrix};$$

for block rows $2\colon\ell$ this is obtained by using Lemma 4.4 repeatedly on block columns $1\colon\ell-1$ of Y, while the form of the first block row follows from that of Y on using the basis elements in (3.5) (since $L(\lambda) \in \mathbb{DL}(P)$). Combining this with the shifted sum property in (4.1), $SX \boxplus\!\!\rightarrow SY = v_1 e_1 \otimes [\, A_\ell \quad \ldots \quad A_0 \,]$, we find that

$$SX = \begin{bmatrix} v_1A_\ell & v_2A_\ell & \ldots & v_\ell A_\ell \\ 0 & -q_0(r_1;v)P(r_1) & \ldots & -q_{\ell-2}(r_1;v)P(r_1) \\ \vdots & \vdots & & \vdots \\ 0 & -q_0(r_{\ell-1};v)P(r_{\ell-1}) & \ldots & -q_{\ell-2}(r_{\ell-1};v)P(r_{\ell-1}) \end{bmatrix},$$

$$SY = \begin{bmatrix} v_1A_{\ell-1} - v_2A_\ell & \ldots & v_1A_1 - v_\ell A_\ell & v_1A_0 \\ q_0(r_1;v)P(r_1) & \ldots & q_{\ell-2}(r_1;v)P(r_1) & 0 \\ \vdots & & \vdots & \vdots \\ q_0(r_{\ell-1};v)P(r_{\ell-1}) & \ldots & q_{\ell-2}(r_{\ell-1};v)P(r_{\ell-1}) & 0 \end{bmatrix}.$$

Completing the congruence by right multiplication with S^* yields the two Hermitian matrices

$$SXS^* = \begin{bmatrix} v_1A_\ell & 0 & 0 & \ldots & 0 \\ 0 & -\widehat{q}_{r_1}(r_1)P(r_1) & -\widehat{q}_{r_1}(r_2)P(r_1) & \ldots & -\widehat{q}_{r_1}(r_{\ell-1})P(r_1) \\ 0 & -\widehat{q}_{r_2}(r_1)P(r_2) & -\widehat{q}_{r_2}(r_2)P(r_2) & \ldots & -\widehat{q}_{r_2}(r_{\ell-1})P(r_2) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & -\widehat{q}_{r_{\ell-1}}(r_1)P(r_{\ell-1}) & -\widehat{q}_{r_{\ell-1}}(r_2)P(r_{\ell-1}) & \ldots & -\widehat{q}_{r_{\ell-1}}(r_{\ell-1})P(r_{\ell-1}) \end{bmatrix},$$

$$SYS^* = \begin{bmatrix} v_1A_{\ell-1} - v_2A_\ell & v_1P(r_1) & v_1P(r_2) & \ldots & v_1P(r_{\ell-1}) \\ v_1P(r_1) & r_1\widehat{q}_{r_1}(r_1)P(r_1) & r_2\widehat{q}_{r_1}(r_2)P(r_1) & \ldots & r_{\ell-1}\widehat{q}_{r_1}(r_{\ell-1})P(r_1) \\ v_1P(r_2) & r_1\widehat{q}_{r_2}(r_1)P(r_2) & r_2\widehat{q}_{r_2}(r_2)P(r_2) & \ldots & r_{\ell-1}\widehat{q}_{r_2}(r_{\ell-1})P(r_2) \\ \vdots & \vdots & \vdots & & \vdots \\ v_1P(r_{\ell-1}) & r_1\widehat{q}_{r_{\ell-1}}(r_1)P(r_{\ell-1}) & r_2\widehat{q}_{r_{\ell-1}}(r_2)P(r_{\ell-1}) & \ldots & r_{\ell-1}\widehat{q}_{r_{\ell-1}}(r_{\ell-1})P(r_{\ell-1}) \end{bmatrix}.$$

But by Lemma 4.3 (c), $\widehat{q}_{r_j}(r_i) = 0$ for $i \neq j$ so that all the off-diagonal blocks in the trailing principal block $(\ell-1) \times (\ell-1)$ submatrices of SXS^* and SYS^* are zero. Combining these results gives the arrowhead form

$$S(\alpha X + \beta Y)S^* = \begin{bmatrix} M_0 & v_1\beta P(r_1) & v_1\beta P(r_2) & \dots & v_1\beta P(r_{\ell-1}) \\ v_1\beta P(r_1) & \mu_1 P(r_1) & 0 & \dots & 0 \\ v_1\beta P(r_2) & 0 & \mu_2 P(r_2) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ v_1\beta P(r_{\ell-1}) & 0 & \dots & 0 & \mu_{\ell-1} P(r_{\ell-1}) \end{bmatrix},$$

where

$$M_0 = \alpha v_1 A_\ell + \beta(v_1 A_{\ell-1} - v_2 A_\ell), \tag{4.3}$$

$$\mu_i = (r_i\beta - \alpha)\widehat{q}_{r_i}(r_i), \qquad i = 1\colon \ell-1. \tag{4.4}$$

From the form of the diagonal blocks $\mu_k P(r_\ell)$ we can already deduce two *necessary conditions* for $\alpha X + \beta Y$ to be definite:

1. (α, β) must be distinct from the roots $r_k = (r_k, 1)$ in the homogeneous sense.
2. Each root r_k must lie in some definiteness interval for P.

A sequence of $\ell - 1$ more congruences eliminates the rest of the off-diagonal blocks. This begins with a congruence by

$$\left[\begin{array}{c|cccc} \mu_1 & -v_1\beta & 0 & \dots & 0 \\ \hline 0 & & & I_{\ell-1} & \end{array}\right] \otimes I_n =: S_1$$

in order to eliminate the $(1,2)$ and $(2,1)$ blocks:

$$S_1 S(\alpha X + \beta Y)S^* S_1^* = \begin{bmatrix} M_1 & 0 & v_1\beta\mu_1 P(r_2) & \dots & v_1\beta\mu_1 P(r_{\ell-1}) \\ 0 & \mu_1 P(r_1) & & & \\ v_1\beta\mu_1 P(r_2) & & \mu_2 P(r_2) & & \\ \vdots & & & \ddots & \\ v_1\beta\mu_1 P(r_{\ell-1}) & & & & \mu_{\ell-1} P(r_{\ell-1}) \end{bmatrix},$$

where $M_1 = \mu_1^2 M_0 - v_1^2 \beta^2 \mu_1 P(r_1)$. With

$$S_j = \left[\begin{array}{c|ccccccc} \mu_j & 0 & \dots & 0 & \sigma_j & 0 & \dots & 0 \\ \hline 0 & & & & I_{\ell-1} & & & \end{array}\right] \otimes I_n, \qquad j \geq 2,$$

where $\sigma_j = -v_1\beta \prod_{k=1}^{j-1} \mu_k$ is the $(1, j+1)$ entry of S_j, it can be proved by induction that after k such "eliminations-by-congruence" we have

$$S_k \dots S_1 S(\alpha X + \beta Y)S^* S_1^* \dots S_k^* =$$

$$\begin{bmatrix} M_k & 0 & \dots & 0 & v_1\beta\prod_{j=1}^{k}\mu_j \cdot P(r_{k+1}) & \dots & v_1\beta\prod_{j=1}^{k}\mu_j \cdot P(r_{\ell-1}) \\ 0 & \mu_1 P(r_1) & & & & & \\ \vdots & & \ddots & & & & \\ 0 & & & \mu_k P(r_k) & & & \\ v_1\beta\prod_{j=1}^{k}\mu_j \cdot P(r_{k+1}) & & & & \mu_{k+1}P(r_{k+1}) & & \\ \vdots & & & & & \ddots & \\ v_1\beta\prod_{j=1}^{k}\mu_j \cdot P(r_{\ell-1}) & & & & & & \mu_{\ell-1}P(r_{\ell-1}) \end{bmatrix},$$

where

$$M_k = \prod_{j=1}^{k}\mu_j \cdot \left[\prod_{j=1}^{k}\mu_j \ \cdot M_0 - v_1^2\beta^2\sum_{i=1}^{k}\Bigg(\prod_{\substack{j=1\\ j\neq i}}^{k}\mu_j\Bigg)P(r_i)\right]. \tag{4.5}$$

Thus after completing all $\ell-1$ eliminations-by-congruence we have the block diagonal form

$$S_{\ell-1}\dots S_1 S(\alpha X+\beta Y)S^* S_1^* \dots S_{\ell-1}^* = \begin{bmatrix} M & & & \\ & \mu_1 P(r_1) & & \\ & & \ddots & \\ & & & \mu_{\ell-1}P(r_{\ell-1}) \end{bmatrix}, \tag{4.6}$$

where $M = M_{\ell-1}$ is given by (4.5) with k replaced by $\ell-1$. Quite remarkably, M simplifies to just a scalar multiple of $P(\alpha,\beta)$:

$$M = v_1\left[\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\right]^2 \prod_{j=1}^{\ell-1}(\alpha - r_j\beta)\cdot P(\alpha,\beta).$$

The proof of this simplification for M, which involves some tedious calculations, is left to Appendix A. The block diagonal form in (4.6) can be simplified even further; a scaling congruence removes the squared term $\left[\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\right]^2$ from the $(1,1)$ block, and v_1 can be factored out of each $\mu_i = (r_i\beta-\alpha)\widehat{q}_{r_i}(r_i)$ since from Lemma 4.3 (b), $\widehat{q}_{r_i}(r_i) = v_1\prod_{j\neq i}(r_i - r_j)$. Thus we see that $\alpha X+\beta Y$ is congruent to the block diagonal form

$$v_1 \cdot \begin{bmatrix} \prod_{j=1}^{\ell-1}(\alpha - r_j\beta)\cdot P(\alpha,\beta) & & & \\ & (r_1\beta-\alpha)\prod_{j\neq 1}(r_1 - r_j)\cdot P(r_1) & & \\ & & \ddots & \\ & & & (r_{\ell-1}\beta-\alpha)\prod_{j\neq \ell-1}(r_{\ell-1}-r_j)\cdot P(r_{\ell-1}) \end{bmatrix}. \tag{4.7}$$

This block diagonalization allows us to make the connection between definiteness of the polynomial P and definiteness of the matrix $\alpha X+\beta Y$.

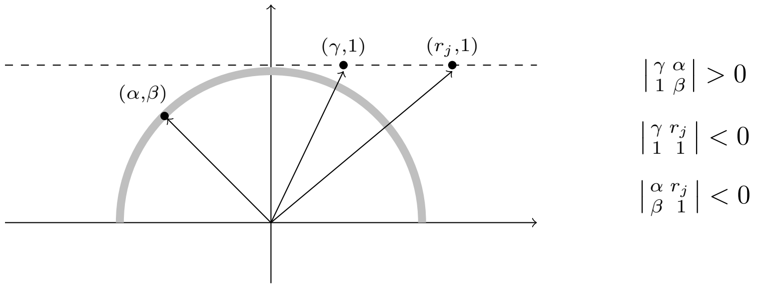


FIG. 4.1. *Pictorial representation of* (4.8).

We first make the convention that (α,β) lies in the upper half-circle, since replacing (α,β) by $(-\alpha,-\beta)$ changes the signs of all the diagonal blocks, and hence does not affect the definiteness or indefiniteness of $\alpha X+\beta Y$. The finite roots $r_1,\ldots,r_{\ell-1}$ can be viewed in homogeneous terms as vectors $(r_i,1)$ in the upper (α,β)-plane. Each diagonal block in (4.7) is a scalar multiple of P evaluated at one of the constants in the set $\mathcal{R}=\{(\alpha,\beta),r_1,r_2,\ldots,r_{\ell-1}\}$. For any $\gamma\in\mathcal{R}$, the scalar in front of $P(\gamma)$ can be interpreted as a product of $\ell-1$ factors in which γ is compared with each of the other $\ell-1$ constants in $\mathcal{R}$ via 2×2 determinants:

$$\begin{aligned} &\begin{vmatrix}\gamma & \alpha\\ 1 & \beta\end{vmatrix} = \gamma\beta-\alpha; \quad \begin{vmatrix}\gamma & r_j\\ 1 & 1\end{vmatrix} = \gamma - r_j;\\ \text{or}\quad &\begin{vmatrix}\alpha & r_j\\ \beta & 1\end{vmatrix} = \alpha - r_j\beta \quad \text{when } \gamma=(\alpha,\beta).\end{aligned} \tag{4.8}$$

The sign of any of these determinants is positive for any constant in $\mathcal{R}$ that lies counterclockwise from γ, and negative for any that lie clockwise from γ; see Figure 4.1.

Consequently the sign of the whole scalar multiple of $P(\gamma)$ reveals the parity of the number of constants in $\mathcal{R}$ that lie clockwise from γ. Hence from (4.7) we see that when P is definite any choice of (α,β) and finite $r_1,r_2,\ldots,r_{\ell-1}$ to be placed, one in each of the ℓ distinct definiteness intervals for P in the upper half-circle, will result in a definite matrix $\alpha X+\beta Y$, and hence a definite pencil $L(\lambda)=\lambda X+Y\in\mathbb{H}(P)$. This proves the "if" part of Theorem 4.1.

Now suppose there exists a definite pencil $\alpha X+\beta Y$ in $\mathbb{H}(P)$. We first assume that $v_1\neq 0$. By a final permutation congruence, we rearrange the diagonal blocks in (4.7) so that the vectors (α,β), $(r_i,1)$, $i=1{:}\,\ell-1$ at which P is evaluated are encountered in counterclockwise order (starting with ∞ if $(\alpha,\beta)=\infty$) as we descend the diagonal. With this reordering of blocks, the scalar coefficient of P in the $(1,1)$ block will be positive, and the rest of the scalar coefficients will have alternating signs as we descend the diagonal. Thus in order for $\alpha X+\beta Y$ to be a definite matrix, the definiteness parity of the matrices $P(r_i)$, $P(\alpha,\beta)$ must also alternate as we descend the diagonal. Thus by Theorem 2.6, we see that the existence of a definite pencil in $\mathbb{H}(P)$ with $v_1\neq 0$ implies that P must be definite. Now if there exists a definite pencil in $\mathbb{H}(P)$ with $v_1=0$, $v_2\neq 0$ then, since pencils in $\mathbb{H}(P)$ vary continuously with the ansatz vector $v\in\mathbb{R}^\ell$, and definite pencils form an open subset of $\mathbb{H}(P)$, a sufficiently small perturbation of v_1 away from zero will result in a definite pencil in $\mathbb{H}(P)$ with $v_1\neq 0$, and thereby imply the definiteness of P. Thus the existence of any definite pencil in $\mathbb{H}(P)$ implies the definiteness of P. This completes the proof of Theorem 4.1.

To complete the proof of Theorem 4.2 characterizing the set of all definite pencils in $\mathbb{H}(P)$ we need to consider the case where one of the roots r_j of $\mathsf{p}(x;v)$ is ∞ (or equivalently, $v_1 = 0$), assuming that one of the definiteness intervals of P contains ∞.

4.3. Case 2: $v_1 = 0$, $v_2 \neq 0$. We can no longer start the block diagonalization of $\alpha X + \beta Y$ with S as in (4.2), since one of the roots r_i is ∞. Instead we use all available finite (real) roots $r_1, r_2, \ldots, r_{\ell-2}$ and let

$$\widetilde{S} = [\, e_1 \quad e_2 \quad \Lambda(r_1) \quad \Lambda(r_2) \quad \ldots \quad \Lambda(r_{\ell-2})\,]^T \otimes I_n.$$

By arguments similar to those used in subsection 4.2 we find that

$$\widetilde{S}X = \begin{bmatrix} 0 & v_2A_\ell & v_3A_\ell & \ldots & v_\ell A_\ell \\ v_2A_\ell & v_2A_{\ell-1}+v_3A_\ell & v_3A_{\ell-1}+v_4A_\ell & \ldots & v_\ell A_{\ell-1} \\ 0 & -q_0(r_1;v)P(r_1) & -q_1(r_1;v)P(r_1) & \ldots & -q_{\ell-2}(r_1;v)P(r_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & -q_0(r_{\ell-2};v)P(r_{\ell-2}) & -q_1(r_{\ell-2};v)P(r_{\ell-2}) & \ldots & -q_{\ell-2}(r_{\ell-2};v)P(r_{\ell-2}) \end{bmatrix},$$

$$\widetilde{S}Y = \begin{bmatrix} -v_2A_\ell & -v_3A_\ell & \ldots & -v_\ell A_\ell & 0 \\ -v_3A_\ell & v_2A_{\ell-2}-v_3A_{\ell-1}-v_4A_\ell & \ldots & v_2A_1-v_\ell A_{\ell-1} & v_2A_0 \\ q_0(r_1;v)P(r_1) & q_1(r_1;v)P(r_1) & \ldots & q_{\ell-2}(r_1;v)P(r_1) & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ q_0(r_{\ell-2};v)P(r_{\ell-2}) & q_1(r_{\ell-2};v)P(r_{\ell-2}) & \ldots & q_{\ell-2}(r_{\ell-2};v)P(r_{\ell-2}) & 0. \end{bmatrix}.$$

Note that $q_0(x;v) = v_1 = 0$ and $q_1(x;v) = v_1x + v_2 = v_2$. Using these and completing the congruence by right multiplication with $\widetilde{S}^*$ yields

$$\widetilde{S}X\widetilde{S}^* = \begin{bmatrix} 0 & v_2A_\ell & 0 & \ldots & 0 \\ v_2A_\ell & v_2A_{\ell-1}+v_3A_\ell & 0 & \ldots & 0 \\ 0 & 0 & -\widehat{q}_{r_1}(r_1)P(r_1) & \ldots & -\widehat{q}_{r_1}(r_{\ell-2})P(r_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & -\widehat{q}_{r_{\ell-2}}(r_1)P(r_{\ell-2}) & \ldots & -\widehat{q}_{r_{\ell-2}}(r_{\ell-2})P(r_{\ell-2}) \end{bmatrix},$$

$$\widetilde{S}Y\widetilde{S}^* = \begin{bmatrix} -v_2A_\ell & -v_3A_\ell & 0 & \ldots & 0 \\ -v_3A_\ell & v_2A_{\ell-2}-v_3A_{\ell-1}-v_4A_\ell & v_2P(r_1) & \ldots & v_2P(r_{\ell-2}) \\ 0 & v_2P(r_1) & r_1\widehat{q}_{r_1}(r_1)P(r_1) & \ldots & r_{\ell-2}\widehat{q}_{r_1}(r_{\ell-2})P(r_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & v_2P(r_{\ell-2}) & r_1\widehat{q}_{r_{\ell-2}}(r_1)P(r_{\ell-2}) & \ldots & r_{\ell-2}\widehat{q}_{r_{\ell-2}}(r_{\ell-2})P(r_{\ell-2}) \end{bmatrix}.$$

But by Lemma 4.3 (c), all the off-diagonal blocks in the bottom right $(\ell-2)\times(\ell-2)$ block submatrices of SXS^* and SYS^* are zero. Hence

$$\widetilde{S}(\alpha X + \beta Y)\widetilde{S} = \begin{bmatrix} -v_2\beta A_\ell & (v_2\alpha - v_3\beta)A_\ell & 0 & \ldots & 0 \\ (v_2\alpha - v_3\beta)A_\ell & N_0 & v_2\beta P(r_1) & \ldots & v_2\beta P(r_{\ell-2}) \\ 0 & v_2\beta P(r_1) & \mu_1 P(r_1) & & \\ \vdots & \vdots & & \ddots & \\ 0 & v_2\beta P(r_{\ell-2}) & & & \mu_{\ell-2}P(r_{\ell-2}) \end{bmatrix},$$

where $N_0 = \alpha(v_2A_{\ell-1} + v_3A_\ell) + \beta(v_2A_{\ell-2} - v_3A_{\ell-1} - v_4A_\ell)$, and the $\mu_i = (r_i\beta - \alpha)\widehat{q}_{r_i}(r_i)$ are the same as in case 1. Note that because of $-v_2\beta A_\ell$ in the $(1,1)$ block, choosing $\beta = 0$ results in $\alpha X + \beta Y$ not being a definite matrix. Thus we cannot choose (α, β) to be ∞, which is entirely consistent with case 1 where we had to choose (α, β) to be distinct from all the r_i. From now on, then, we assume that $\beta \neq 0$. Note that the blocks $\mu_i P(r_i)$ on the diagonal of this condensed form once again show that each r_i must lie in some definiteness interval for P, and that (α, β) must be chosen distinct from all the finite roots $r_1, r_2, \ldots, r_{\ell-2}$; otherwise $\alpha X + \beta Y$ will not be a definite matrix.

The next step in the reduction is to eliminate the blocks in the second block row and column that are of the form $v_2 \beta P(r_i)$, using $\ell - 2$ congruences analogous to the ones used in case 1. The first of these is by the matrix

$$\widetilde{S}_1 = \left[\begin{array}{c|cccc} 1 & & & & \\ \hline & \mu_1 & -v_2\beta & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{array}\right] \otimes I_n,$$

yielding

$$\widetilde{S}_1 \widetilde{S}(\alpha X + \beta Y)\widetilde{S}^* \widetilde{S}_1^* =$$
$$\begin{bmatrix} -v_2\beta A_\ell & \mu_1(v_2\alpha - v_3\beta)A_\ell & 0 & \ldots & \ldots & 0 \\ \mu_1(v_2\alpha - v_3\beta)A_\ell & N_1 & 0 & v_2\beta\mu_1 P(r_2) & \ldots & v_2\beta\mu_1 P(r_{\ell-1}) \\ 0 & 0 & \mu_1 P(r_1) & 0 & \ldots & 0 \\ \vdots & v_2\beta\mu_1 P(r_2) & 0 & \mu_2 P(r_2) & & \vdots \\ \vdots & \vdots & \vdots & & \ddots & 0 \\ 0 & v_2\beta\mu_1 P(r_{\ell-1}) & 0 & \ldots & 0 & \mu_{\ell-2} P(r_{\ell-2}) \end{bmatrix},$$

where $N_1 = \mu_1^2 N_0 - v_2^2 \beta^2 \mu_1 P(r_1)$. Continuing in analogous fashion we ultimately obtain

$$\widetilde{S}_{\ell-2} \ldots \widetilde{S}_1 \widetilde{S}(\alpha X + \beta Y)\widetilde{S}^* \widetilde{S}_1^* \ldots \widetilde{S}_{\ell-2}^* =$$
$$\begin{bmatrix} -v_2\beta A_\ell & \prod_{i=1}^{\ell-2} \mu_i \cdot (v_2\alpha - v_3\beta)A_\ell & & & \\ \prod_{i=1}^{\ell-2} \mu_i \cdot (v_2\alpha - v_3\beta)A_\ell & N_{\ell-2} & & & \\ & & \mu_1 P(r_1) & & \\ & & & \ddots & \\ & & & & \mu_{\ell-2} P(r_{\ell-2}) \end{bmatrix},$$

where $N_{\ell-2} = \prod_{j=1}^{\ell-2} \mu_j \cdot \left[\prod_{j=1}^{\ell-2} \mu_j \cdot N_0 - v_2^2 \beta^2 \sum_{i=1}^{\ell-2} \left(\prod_{\substack{j=1 \\ j \neq i}}^{\ell-2} \mu_j \right) P(r_i) \right]$. One more congruence completes the block diagonalization, namely congruence by

$$E = \left[\begin{array}{cc|c} 1 & 0 & \\ \prod_{j=1}^{\ell-2} \mu_j \cdot (v_2\alpha - v_3\beta) & v_2\beta & \\ \hline & & I_{\ell-2} \end{array}\right] \otimes I_n.$$

Note that E is nonsingular because $\beta \neq 0$. This gives

$$E\widetilde{S}_{\ell-2} \ldots \widetilde{S}_1 \widetilde{S}(\alpha X + \beta Y)\widetilde{S}^* \widetilde{S}_1^* \ldots \widetilde{S}_{\ell-2}^* E^* =$$

$$\begin{bmatrix} -v_2\beta A_\ell & & & & \\ & N & & & \\ & & \mu_1 P(r_1) & & \\ & & & \ddots & \\ & & & & \mu_{\ell-2}P(r_{\ell-2}) \end{bmatrix}, \tag{4.9}$$

where

$$N = v_2\beta \prod_{j=1}^{\ell-2} \mu_j \cdot \Bigg[\prod_{j=1}^{\ell-2} \mu_j \cdot (v_2\alpha - v_3\beta)^2 A_\ell + v_2\beta \prod_{j=1}^{\ell-2} \mu_j \cdot N_0 - v_2^3\beta^3 \sum_{i=1}^{\ell-2} \Bigg(\prod_{\substack{j=1 \\ j\neq i}}^{\ell-2} \mu_j \Bigg) P(r_i) \Bigg]. \tag{4.10}$$

It is shown in Appendix A.2 that

$$N = v_2 \Bigg[v_2 \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \Bigg]^2 \beta \prod_{j=1}^{\ell-2} (\alpha - r_j\beta) \cdot P(\alpha,\beta).$$

The block diagonal form in (4.9) is simplified further by factoring out $-v_2$ and by a scaling congruence to remove the squared term $\big[v_2 \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j)\big]^2$ from the $(2,2)$ block. Thus $\alpha X + \beta Y$ is congruent to the block diagonal form

$$-v_2 \cdot \begin{bmatrix} \beta A_\ell & & & & \\ & -\beta \prod_{j=1}^{\ell-2} (\alpha - r_j\beta) \cdot P(\alpha,\beta) & & & \\ & & -(r_1\beta - \alpha) \prod_{j\neq 1} (r_1 - r_j) \cdot P(r_1) & & \\ & & & \ddots & \\ & & & & -(r_{\ell-2}\beta - \alpha) \prod_{j\neq \ell-2} (r_{\ell-2} - r_j) \cdot P(r_{\ell-2}) \end{bmatrix}.$$

With the block diagonalization of $\alpha X + \beta Y$ written in this particular form, it now becomes possible to give a common conceptual interpretation for all the diagonal blocks that is similar to the one we gave for case 1. We just point out the differences. Recall that when $v_1 = 0$ and $v_2 \neq 0$, $\mathsf{p}(x; v)$ has $\ell - 2$ finite roots $r_1, r_2, \ldots, r_{\ell-2}$ and one which is infinite, $r_{\ell-1} = \infty$, i.e., $(1,0)$ in homogeneous form. Then we have $P(r_{\ell-1}) \equiv P(\infty) \equiv P(1,0) = A_\ell$. Hence each diagonal block is a scalar multiple of P evaluated at one of the constants in the set $\mathcal{R} = \{(\alpha,\beta), r_1, r_2, \ldots, r_{\ell-1}\}$. For any $\gamma \in \mathcal{R}$, the scalar in front of $P(\gamma)$ can be interpreted as a product of $\ell - 1$ factors in which γ is compared with each of the other $\ell - 1$ constants in $\mathcal{R}$ via 2×2 determinants:

- $\left|\begin{smallmatrix} \gamma & r_j \\ 1 & 1 \end{smallmatrix}\right| = \gamma - r_j$, $\left|\begin{smallmatrix} \gamma & \alpha \\ 1 & \beta \end{smallmatrix}\right| = \gamma\beta - \alpha$, or $\left|\begin{smallmatrix} \gamma & 1 \\ 1 & 0 \end{smallmatrix}\right| = -1$ if γ is finite,
- $\left|\begin{smallmatrix} 1 & r_j \\ 0 & 1 \end{smallmatrix}\right| = +1$, or $\left|\begin{smallmatrix} 1 & \alpha \\ 0 & \beta \end{smallmatrix}\right| = \beta$ if $\gamma = (1,0)$ is infinite,
- $\left|\begin{smallmatrix} \alpha & r_j \\ \beta & 1 \end{smallmatrix}\right| = \alpha - r_j\beta$ and $\left|\begin{smallmatrix} \alpha & 1 \\ \beta & 0 \end{smallmatrix}\right| = -\beta$ for $\gamma = (\alpha,\beta)$ otherwise.

Recall our convention that each $\gamma \in \mathcal{R}$ lies in the strict upper (α,β) half-plane $\cup$ $(1,0)$. Then the sign of any of these determinants is positive for any constant in $\mathcal{R}$ that lies counterclockwise from γ, and negative for any that lie clockwise from γ. Consequently the sign of the whole scalar multiple of $P(\gamma)$ reveals the parity of the number of constants in $\mathcal{R}$ that lie clockwise from γ. If we re-order the blocks (via permutation congruence) so that as we go down the diagonal we encounter the constants from $\mathcal{R}$ in the evaluated P's in strict counterclockwise order (starting with ∞ whenever $\infty \in \mathcal{R}$), then the scalar multiples will have alternating sign, starting with a positive sign in the $(1,1)$ block. We then see that $\alpha X + \beta Y$ is a definite matrix if and only if the matrices $P(\gamma)$ have strictly alternating definiteness parity as we descend the diagonal. This completes the characterization of the set of all definite pencils in $\mathbb{H}(P)$ as given in Theorem 4.2.

5. Application to quadratics. We now concentrate our attention on quadratic polynomials, $Q(\lambda) = \lambda^2 A + \lambda B + C$ with Hermitian A, B, and C. For $x \in \mathbb{C}^n$ let

$$q_x(\lambda) = x^* Q(\lambda) x = \lambda^2 (x^* A x) + \lambda (x^* B x) + x^* C x = \lambda^2 a_x + \lambda b_x + c_x$$

be the scalar section of Q at x. The discriminant of Q at x is the discriminant of $q_x(\lambda)$:

$$\mathcal{D}_x := b_x^2 - 4 a_x c_x = \mathcal{D}_x(Q).$$

The following result is specific to quadratics.

THEOREM 5.1. *A Hermitian quadratic matrix polynomial $Q(\lambda)$ is definite if and only if any two (and hence all) of the following properties hold:*

(a) $Q(\mu) > 0$ *for some* $\mu \in \mathbb{R} \cup \{\infty\}$.

(b) $\mathcal{D}_x = (x^* B x)^2 - 4(x^* A x)(x^* C x) > 0$ *for all nonzero* $x \in \mathbb{C}^n$.

(c) $Q(\gamma) < 0$ *for some* $\gamma \in \mathbb{R} \cup \{\infty\}$.

Proof. (a) and (b) are equivalent to Q being definite, directly from Definition 2.5. (a) and (c) are equivalent to Q being definite by Theorem 2.6. Suppose (b) and (c) hold and let $\widetilde{Q}(\lambda) = -Q(\lambda)$. Then $\widetilde{Q}(\gamma) > 0$ and $\mathcal{D}_x(\widetilde{Q}) = \mathcal{D}_x(Q)$, so $\widetilde{Q}$ satisfies (a) and (b) and so is definite. But then by Theorem 2.6, $\widetilde{Q}(\mu) < 0$ for some μ, which means that $Q(\mu) > 0$, so Q satisfies (a) and (b) and hence is definite. Finally, Q being definite implies (b) and (c) hold by Definition 2.5 and Theorem 2.6. ☐

Here is a simple example where A, B, and C are all indefinite but properties (a) and (c) of Theorem 5.1 hold, so that $Q(\lambda)$ is definite:

$$Q(\lambda) = \lambda^2 \begin{bmatrix} -3 & -1 \\ -1 & 2 \end{bmatrix} + \lambda \begin{bmatrix} 6 & 3 \\ 3 & -10 \end{bmatrix} + \begin{bmatrix} 0 & -2 \\ -2 & 9 \end{bmatrix}, \quad Q(1) > 0, \quad Q(3) < 0.$$

The definiteness diagram for Q has the form of the last diagram in Figure 2.2 (b).

The standard basis pencils of $\mathbb{H}(Q)$ with ansatz vectors e_1 and e_2 are given by

$$L_1(\lambda) = \lambda \begin{bmatrix} A & 0 \\ 0 & -C \end{bmatrix} + \begin{bmatrix} B & C \\ C & 0 \end{bmatrix}, \qquad L_2(\lambda) = \lambda \begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & C \end{bmatrix}.$$

$L_1(\lambda)$ is a linearization if the trailing coefficient matrix C is nonsingular. Since the root of $\mathsf{p}(x; e_1)$ is 0, Theorem 4.2 implies that if $Q(\lambda)$ is definite with $C = Q(0)$ definite then $L_1(\lambda)$ is definite. Similarly, $L_2(\lambda)$ is a linearization if the leading coefficient matrix A is nonsingular, and since the root of $\mathsf{p}(x; e_2)$ is ∞, Theorem 4.2 implies that if $Q(\lambda)$ is definite with $A = Q(\infty)$ definite then $L_2(\lambda)$ is definite.

Now if Q is definite with $A > 0$ and $C < 0$ then $\left[\begin{smallmatrix} A & 0 \\ 0 & -C \end{smallmatrix}\right] > 0$. Thus the eigenvalues of Q can be computed by the Cholesky–QR method on either $L_1(\lambda)$ or $\lambda L_2(1/\lambda)$. Note that the Cholesky–QR method [3] has several advantages over the QZ algorithm for the numerical solution of hyperbolic quadratics. First, the Cholesky–QR method takes advantage of the symmetry of the pencil, which results in a reduction in both the storage requirement and the computational cost. Second, it guarantees to produce real eigenvalues and therefore preserves this spectral property of definite quadratics. This is not necessarily the case for the QZ algorithm.

We now show how to transform definite quadratics into special forms. Three different cases are considered.

Case (a) Suppose that Q is hyperbolic, so that $A > 0$, and we wish to transform Q into $\widehat{Q}$ with $\widehat{A} > 0$ and $\widehat{C} < 0$. An ordinary translation suffices to achieve this as long as we know one value $\gamma \in \mathbb{R}$ such that $Q(\gamma) < 0$. Then

$$\begin{aligned}\widehat{Q}(\lambda) := Q(\lambda + \gamma) &= \lambda^2 A + \lambda(B + 2\gamma A) + C + \gamma B + \gamma^2 A \\ &\equiv \lambda^2 \widehat{A} + \lambda \widehat{B} + \widehat{C}\end{aligned}$$

with $\widehat{Q}(0) = Q(\gamma) = \widehat{C} < 0$ and $\widehat{A} = A > 0$.

Case (b) Suppose Q is definite and we wish to transform it into a hyperbolic $\widetilde{Q}$ with $\widetilde{A} > 0$. This can be done by a homogeneous rotation provided we know one value $\mu \in \mathbb{R} \cup \{\infty\}$ for which $Q(\mu) > 0$. We need to make μ for Q correspond to ∞ for $\widetilde{Q}$. For this we express μ in homogeneous coordinates: $\mu = (c, s)$ with $c^2 + s^2 = 1$. Recall that $\infty = (1, 0)$ in homogeneous coordinates. From (2.1) we see that $\left[\begin{smallmatrix}\alpha \\ \beta\end{smallmatrix}\right] = \left[\begin{smallmatrix}c \\ s\end{smallmatrix}\right] = \mu$ will correspond to $\left[\begin{smallmatrix}\widetilde{\alpha} \\ \widetilde{\beta}\end{smallmatrix}\right] = \left[\begin{smallmatrix}1 \\ 0\end{smallmatrix}\right]$. Thus homogenous rotation with these (c, s)-values will give $\widetilde{Q}$ such that $\widetilde{A} = Q(c, s) > 0$.

Case (c) Suppose Q is definite and we wish to transform Q into Q' so that $A' > 0$ and $C' < 0$. To do this we need to know a $\mu \in \mathbb{R} \cup \{\infty\}$ such that $Q(\mu) > 0$ and a $\gamma \in \mathbb{R} \cup \{\infty\}$ such that $Q(\gamma) < 0$. With these two numbers in hand we first do the rotation of case (b) to obtain $A' > 0$ and then the translation of case (a) to obtain $C' < 0$.

Finally, we note that an efficient algorithm for testing whether a Hermitian quadratic is hyperbolic is developed by Guo, Higham, and Tisseur [5]. In the case of an affirmative test this algorithm provides a μ such that $Q(\mu) < 0$.

Appendix A.

This appendix deals with the simplification of one of the diagonal blocks obtained during the block diagonalizations of $\alpha X + \beta Y$ for both case 1 ($v_1 \neq 0$) and case 2 ($v_1 = 0$, $v_2 \neq 0$). The next lemma is the main technical result needed to achieve these simplifications.

LEMMA A.1. *Let $p_m(x; f)$ be the polynomial of degree $m-1$ that interpolates the function f at the m distinct points $r_1, r_2, \ldots, r_m$. Rewrite f as*

$$f(x) = \mathcal{E}(x; f) + p_m(x; f), \tag{A.1}$$

where $\mathcal{E}(x; f)$ is the error in interpolation. Then

(a) $\mathcal{E}(x; x^m) = \displaystyle\prod_{i=1}^{m}(x - r_i)$,

(b) $\mathcal{E}(x; x^{m+1}) = \Big(x + \displaystyle\sum_{i=1}^{m} r_i\Big) \prod_{i=1}^{m}(x - r_i)$,

(c) $\mathcal{E}(x; x^{m+2}) = \Big(x^2 + x\displaystyle\sum_{i=1}^{m} r_i + \Big(\sum_{i=1}^{m} r_i\Big)^2 - \sum_{\substack{i,j=1 \\ i<j}}^{m} r_i r_j\Big) \prod_{i=1}^{m}(x - r_i)$.

Proof. Recall the Lagrange form of the interpolating polynomial to f at the m distinct points $r_1, r_2, \ldots, r_m$,

$$p_m(x; f) = \sum_{j=1}^{m} f(r_j) \prod_{\substack{i=1 \\ i \neq j}}^{m} \Big(\frac{x - r_i}{r_j - r_i} \Big).$$

To avoid clutter in the proof we define

$$s := \sum_{i=1}^{m} r_i, \quad \widetilde{s} := \sum_{\substack{i,j=1 \\ i<j}}^{m} r_i r_j.$$

We will use repeatedly the following fact: if two monic polynomials (p and q) of degree m agree at m distinct points then they are identically equal (since $p - q$ is a degree $m-1$ polynomial with m zeros.)

(a) $f(x) = x^m$ and $q_0(x) = \prod_{i=1}^{m}(x - r_i) + p_m(x; f)$ are monic polynomials that agree at the m points $r_1, r_2, \ldots, r_m$. Hence $q_0(x) = x^m$ and the expression for $\mathcal{E}(x; x^m)$ in (a) follows. This result can also be obtained from the standard formula for the error in polynomial interpolation. Observe that equating coefficients of the degree $m-1$ terms in $x^m = q_0(x)$ gives the identity

$$\sum_{j=1}^{m} \Big(\frac{r_j^m}{\prod_{i \neq j}(r_j - r_i)} \Big) = s. \tag{A.2}$$

(b) Note that x^{m+1} and $q_1(x) = (x + \sum_{i=1}^{m} r_i) \cdot \prod_{i=1}^{m}(x - r_i) + p_m(x; x^{m+1})$ are monic degree $m+1$ polynomials that agree at the m points $r_1, r_2, \ldots, r_m$. Thus an $(m+1)$th point is needed to prove that $q_1(x) = x^{m+1}$. Now,

$$\begin{aligned}
q_1(0) &= s \prod_{i=1}^{m}(-r_i) + \sum_{j=1}^{m} \Big[r_j^{m+1} \prod_{\substack{i=1 \\ i \neq j}}^{m} \Big(\frac{-r_i}{r_j - r_i} \Big) \Big] \\
&= s \prod_{i=1}^{m}(-r_i) + \sum_{j=1}^{m} \Big[\prod_{i=1}^{m}(-r_i) \Big] \frac{-r_j^m}{\prod_{i \neq j}(r_j - r_i)} \\
&= \prod_{i=1}^{m}(-r_i) \cdot \Bigg(s - \sum_{j=1}^{m} \Big[\frac{r_j^m}{\prod_{i \neq j}(r_j - r_i)} \Big] \Bigg) \\
&= 0
\end{aligned}$$

by (A.2). Thus $q_1(x) = x^{m+1}$ whenever $r_1, r_2, \ldots, r_m$ are all nonzero. Now suppose one of the points, r_m say, is zero, so that the above argument is not valid. In this case we view q_1 as a function of x and $r_1, r_2, \ldots, r_m$, and observe that q_1 is continuous in all these variables, as long as the r_i remain distinct. We perturb r_m away from zero, keeping it distinct from all the other r_i. Then for any fixed but arbitrary x we have $q_1(x, r_1, r_2, \ldots, r_m) = x^{m+1}$ when $r_m \neq 0$, and by continuity we have the same equality as $r_m \to 0$. Thus $q_1(x) \equiv x^{m+1}$ holds for any set of distinct $r_1, r_2, \ldots, r_m$, even if one of them is zero.

(c) We begin by computing the three highest order terms of $(x^2 + sx + s^2 - \widetilde{s}) \prod_{i=1}^{m}(x - r_i) + p_m(s; x^{m+2}) =: q_2(x)$, which come solely from the first expression since $p_m(x;x^{m+2})$ is of degree $m-1$. We have

$$
\begin{aligned}
(x^2+sx+s^2-\widetilde{s})\prod_{i=1}^{m}(x-r_i) &= (x^2+sx+s^2-\widetilde{s})(x^m-sx^{m-1}+\widetilde{s}x^{m-2}+\cdots)\\
&= x^{m+2}+0\cdot x^{m+1}+0\cdot x^m+\cdots.
\end{aligned}
$$

Thus $h(x) := q_2(x)-x^{m+2}$ is actually a degree $m-1$ polynomial, and it is easy to see that $h(x)$ has the m distinct zeros $r_1, r_2, \dots, r_m$, so that $h(x)=0$, i.e., $q_2(x)\equiv x^{m+2}$ and the expression for $\mathcal{E}(x;x^{m+2})$ follows. □

Using the Lagrange form of the interpolating polynomial and letting $x=\alpha/\beta$ with $\beta\neq 0$, $(a), (b), (c)$ of Lemma A.1 together with (A.1) yield the following identities in the homogeneous variables (α,β):

$$
\alpha^m = \prod_{i=1}^{m}(\alpha-r_i\beta)+\sum_{i=1}^{m}(r_i\beta)^m\prod_{\substack{j=1\\ j\neq i}}^{m}\left(\frac{\alpha-r_j\beta}{r_i\beta-r_j\beta}\right), \tag{A.3}
$$

$$
\alpha^{m+1} = (\alpha+\sum_{i=1}^{m}\beta r_i)\prod_{i=1}^{m}(\alpha-r_i\beta)+\sum_{i=1}^{m}(r_i\beta)^{m+1}\prod_{\substack{j=1\\ j\neq i}}^{m}\left(\frac{\alpha-r_j\beta}{r_i\beta-r_j\beta}\right), \tag{A.4}
$$

$$
\begin{aligned}
\alpha^{m+2} = \Big(\alpha^2+\alpha\sum_{i=1}^{m}\beta r_i+\Big(\sum_{i=1}^{m}\beta r_i\Big)^2-\beta^2\sum_{\substack{i,j=1\\ i<j}}^{m} r_ir_j\Big)\prod_{i=1}^{m}(\alpha-r_i\beta)&\\
+\sum_{i=1}^{m}(r_i\beta)^{m+2}\prod_{\substack{j=1\\ j\neq i}}^{m}\left(\frac{\alpha-r_j\beta}{r_i\beta-r_j\beta}\right)&.
\end{aligned} \tag{A.5}
$$

With these results in hand we can now return to the simplification of the blocks $M=M_{\ell-1}$ in (4.5) and N in (4.10).

A.1. Simplification of M. Recall from (4.5) that $M=M_{\ell-1}=\prod_{j=1}^{\ell-1}\mu_j\cdot\widetilde{M}$ with

$$
\widetilde{M} = \prod_{j=1}^{\ell-1}\mu_j\cdot M_0 - v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)P(r_i), \tag{A.6}
$$

where, from (4.3) and (4.4), $M_0=\alpha v_1A_\ell+\beta(v_1A_{\ell-1}-v_2A_\ell)$ and $\mu_i=(r_i\beta-\alpha)\widehat{q}_{r_i}(r_i)$, $i=1\colon \ell-1$.

The first step is to break apart every instance of P into three pieces:

$$
P(\lambda)=\lambda^\ell A_\ell+\lambda^{\ell-1}A_{\ell-1}+\widetilde{P}(\lambda). \tag{A.7}
$$

Then we rewrite M_0 so as to eliminate v_2 and group A_ℓ and $A_{\ell-1}$ together. For this, note that, since the r_i are the roots of the v-polynomial, $v_2/v_1=-(r_1+r_2+\cdots+r_{\ell-1}) =: -s$ so that $v_2=-v_1\,s$ and

$$
M_0=v_1(\alpha+s\beta)A_\ell+v_1\beta A_{\ell-1}. \tag{A.8}
$$

Substituting (A.7) and (A.8) into (A.6) and grouping all the A_ℓ and $A_{\ell-1}$ together yields

$$
\widetilde{M}=A_\ell\Big[v_1(\alpha+s\beta)\prod_{j=1}^{\ell-1}\mu_j-v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)r_i^\ell\Big] \tag{A.9}
$$

$$+ A_{\ell-1}\Big[v_1\beta\prod_{j=1}^{\ell-1}\mu_j - v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)r_i^{\ell-1}\Big] - v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)\widetilde{P}(r_i).$$

We now simplify each of these three pieces in turn. Since $\mu_j = (r_j\beta - \alpha)\widehat{q}_{r_j}(r_j)$, by (4.4),

$$\prod_{j=1}^{\ell-1}\mu_j = (-1)^{\ell-1}\prod_{j=1}^{\ell-1}(\alpha - r_j\beta)\widehat{q}_{r_j}(r_j). \tag{A.10}$$

Also, from Lemma 4.3 (b),

$$\widehat{q}_{r_j}(r_j) = v_1\prod_{i\neq j}(r_j - r_i). \tag{A.11}$$

Substituting (A.10) first and then (A.11) in the coefficient of A_ℓ gives

$$\begin{aligned}
&v_1(\alpha+s\beta)\prod_{j=1}^{\ell-1}\mu_j - v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)r_i^{\ell}\\
&= v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\Big[(\alpha+s\beta)\prod_{j=1}^{\ell-1}(\alpha - r_j\beta) + \beta^2\sum_{i=1}^{\ell-1}r_i^{\ell}\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\Big(\frac{\alpha - r_j\beta}{r_i - r_j}\Big)\Big]\\
&= v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\alpha^{\ell},
\end{aligned}$$

where we used (A.4) for the last equality. The simplification of the coefficient of $A_{\ell-1}$ is very similar to that of A_ℓ. On using (A.10) and (A.11) we obtain

$$\begin{aligned}
&v_1\beta\prod_{j=1}^{\ell-1}\mu_j - v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)r_i^{\ell-1}\\
&= v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\beta\Big[\prod_{j=1}^{\ell-1}(\alpha - r_j\beta) + \beta\sum_{i=1}^{\ell-1}r_i^{\ell-1}\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\Big(\frac{\alpha - r_j\beta}{r_i - r_j}\Big)\Big]\\
&= v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\alpha^{\ell-1}\beta,
\end{aligned}$$

where we used (A.3) in the last equality. The rest of $\widetilde{M}$ is simplified as follows:

$$-v_1^2\beta^2\sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)\widetilde{P}(r_i) = v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\beta^2\Big[\sum_{i=1}^{\ell-1}\widetilde{P}(r_i)\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\Big(\frac{\alpha - r_j\beta}{r_i - r_j}\Big)\Big].$$

But for $\beta \neq 0$ and $x = \alpha/\beta$,

$$\sum_{i=1}^{\ell-1}\widetilde{P}(r_i)\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\Big(\frac{\alpha - r_j\beta}{r_i - r_j}\Big) = \beta^{\ell-2}\Big[\sum_{i=1}^{\ell-1}\widetilde{P}(r_i)\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\Big(\frac{x - r_j}{r_i - r_j}\Big)\Big],$$

and the expression inside the square bracket is the Lagrange form of $\widetilde{P}(x)$ since $\widetilde{P}$ is of degree $\ell-2$ and the $\ell-1$ points r_i are distinct. Hence, for $\beta \neq 0$,

$$-v_1^2\beta^2 \sum_{i=1}^{\ell-1}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)\widetilde{P}(r_i) = v_1(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\cdot\beta^2\widetilde{P}(\alpha,\beta).$$

Finally note that the last equality also holds for $\beta = 0$ by continuity. Putting these three simplifications back into (A.9) yields

$$\widetilde{M} = v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\,P(\alpha,\beta),$$

and on using (A.10), $M = M_{\ell-1}$ in (4.5) becomes

$$\begin{aligned} M &= \prod_{j=1}^{\ell-1}\mu_j\cdot v_1\Big[(-1)^{\ell-1}\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]\cdot P(\alpha,\beta)\\ &= v_1\Big[\prod_{j=1}^{\ell-1}\widehat{q}_{r_j}(r_j)\Big]^2\Big[\prod_{j=1}^{\ell-1}(\alpha-r_j\beta)\Big]\,P(\alpha,\beta). \end{aligned}$$

A.2. Simplification of N. To simplify

$$\begin{aligned} N = v_2\beta\prod_{j=1}^{\ell-2}\mu_j\Big[&\prod_{j=1}^{\ell-2}\mu_j\cdot(v_2a-v_3\beta)^2A_\ell\\ &+v_2\beta\prod_{j=1}^{\ell-2}\mu_j\cdot N_0 - v_2^2\beta^2\sum_{i=1}^{\ell-2}\Big(\prod_{\substack{j=1\\ j\neq i}}^{\ell-1}\mu_j\Big)P(r_i)\Big], \end{aligned} \tag{A.12}$$

where $N_0 = \alpha(v_2A_{\ell-1}+v_3A_\ell)+\beta(v_2A_{\ell-2}-v_3A_{\ell-1}-v_4A_\ell)$ and $\mu_i = (r_i\beta-\alpha)\widehat{q}_{r_i}(r_i)$, we this time break apart every instance of P into four pieces

$$P(\lambda) = \lambda^\ell A_\ell + \lambda^{\ell-1}A_{\ell-1} + +\lambda^{\ell-2}A_{\ell-2} + \widehat{P}(\lambda). \tag{A.13}$$

Leaving aside the product $v_2\beta\Big(\prod_{j=1}^{\ell-2}\mu_j\Big)$ at the beginning of the expression for M and focusing on the quantity inside the square brackets, we now simplify the coefficients of the A_ℓ, $A_{\ell-1}$, $A_{\ell-2}$ and $\widehat{P}(\lambda)$-terms.

The A_ℓ-term.

$$v_2^2\Big[\prod_{i=1}^{\ell-2}\mu_i\,\Big(\alpha-\frac{v_3}{v_2}\beta\Big)^2 + \beta\prod_{i=1}^{\ell-2}\mu_i\,\Big(\frac{v_3}{v_2}\alpha-\frac{v_4}{v_2}\beta\Big) - v_2\beta^3\sum_{j=1}^{\ell-2}r_j^\ell\Big(\prod_{i\neq j}\mu_i\Big)\Big].$$

Defining $\widehat{s} = r_1+r_2+\cdots+r_{\ell-2}$ and $\theta = \sum_{1\le i<j\le \ell-2} r_ir_j$ we note that $v_3/v_2 = -\widehat{s}$ and $v_4/v_2 = \theta$. From these notations and the definition of the μ_i we have

$$v_2^2\Big[(-1)^{\ell-2}(\alpha+\widehat{s}\beta)^2\prod_{j=1}^{\ell-2}(\alpha-r_j\beta)\widehat{q}_{r_j}(r_j) + (-1)^{\ell-2}\beta(-\widehat{s}\alpha-\theta\beta)\prod_{j=1}^{\ell-2}(\alpha-r_j\beta)\widehat{q}_{r_j}(r_j)$$

$$
\begin{aligned}
&\qquad +(-1)^{\ell-2} v_2 \beta^3 \sum_{i=1}^{\ell-2} \Big(r_i^\ell \prod_{j\neq i} (\alpha - r_j\beta) \widehat{q}_{r_i}(r_i) \Big) \Bigg] \\
&= v_2^2 (-1)^{\ell-2} \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \cdot \Bigg[\big(\alpha^2 + \widehat{s}\alpha\beta + (\widehat{s}^2 - \theta)\beta \big) \prod_{j=1}^{\ell-2} (\alpha - r_j\beta) \\
&\qquad\qquad + \beta^3 \sum_{i=1}^{\ell-2} r_i^\ell \cdot \prod_{j\neq i} \left(\frac{\alpha - r_j\beta}{r_i - r_j} \right) \Bigg] \\
&= v_2^2 (-1)^{\ell-2} \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \cdot \alpha^\ell,
\end{aligned}
$$

where we used the identity (A.6) to obtain the last equality.

The $A_{\ell-1}$-term.

$$
v_2^2 \beta \Big[\prod_{i=1}^{\ell-2} \mu_i \left(\alpha - \frac{v_3}{v_2} \beta \right) - v_2 \beta^2 \sum_{j=1}^{\ell-2} r_j^{\ell-1} \prod_{\substack{i=1\\ i\neq j}}^{\ell-2} \mu_i \Big].
$$

We use again the fact that $sh = -v_3/v_2$ to rewrite this expression as

$$
v_2^2 \beta \Big[(\alpha + \widehat{s}\beta) \prod_{i=1}^{\ell-2} \mu_i \ - v_2 \beta^2 \sum_{j=1}^{\ell-2} r_j^{\ell-1} \prod_{\substack{i=1\\ i\neq j}}^{\ell-2} \mu_i \Big]. \tag{A.14}
$$

A similar analysis as in the simplification of the A_ℓ-coefficient in case 1, using (A.4), simplifies (A.14) to

$$
v_2^2 \beta (-1)^{\ell-2} \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \cdot \alpha^{\ell-1} \beta.
$$

The $A_{\ell-2}$-term.

$$
v_2^2 \beta^2 \Bigg[\prod_{i=1}^{\ell-2} \mu_i \ - v_2 \beta \sum_{j=1}^{\ell-2} r_j^{\ell-2} \prod_{\substack{i=1\\ i\neq j}}^{\ell-2} \mu_i \Bigg].
$$

The simplification of this term is completely analogous to the treatment of the coefficient of $A_{\ell-1}$ in case 1. We obtain

$$
v_2^2 (-1)^{\ell-2} \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \cdot \alpha^{\ell-2} \beta^2.
$$

The $\widehat{P}(\lambda)$-term. Substituting for the μ_i and factoring out all the $\widehat{q}_{r_i}(r_i)$ terms leads to

$$
-v_2^3 \beta^3 \sum_{j=1}^{\ell-2} \Bigg(\widehat{P}(r_j) \prod_{\substack{i=1\\ i\neq j}}^{\ell-2} \mu_i \Bigg) = v_2^2 (-1)^{\ell-2} \prod_{j=1}^{\ell-2} \widehat{q}_{r_j}(r_j) \cdot \beta^3 \sum_{i=1}^{\ell-2} \widehat{P}(r_i) \prod_{\substack{j=1\\ j\neq i}}^{\ell-2} \left(\frac{\alpha - r_j\beta}{r_i - r_j} \right),
$$

which by the Lagrange interpolation formula simplifies to

$$v_2^2(-1)^{\ell-2}\prod_{j=1}^{\ell-2}\widehat{q}_{r_j}(r_j)\cdot\beta^3\widehat{P}(\alpha,\beta).$$

Bringing these four simplifications all together we have

$$N = v_2\left[v_2\prod_{j=1}^{\ell-2}\widehat{q}_{r_j}(r_j)\right]^2\beta\prod_{j=1}^{\ell-2}(\alpha-r_j\beta)\cdot P(\alpha,\beta).$$

REFERENCES

[1] ANSYS, Inc. *Theory Manual.* SAS IP, Inc., twelfth edition.

[2] Lawrence Barkwell and Peter Lancaster. Overdamped and gyroscopic vibrating systems. *Trans. AME: J. Applied Mechanics*, 59:176–181, 1992.

[3] Philip I. Davies, Nicholas J. Higham, and Françoise Tisseur. Analysis of the Cholesky method with iterative refinement for solving the symmetric definite generalized eigenproblem. *SIAM J. Matrix Anal. Appl.*, 23(2):472–493, 2001.

[4] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Indefinite Linear Algebra and Applications.* Birkhäuser, Basel, Switzerland, 2005.

[5] Chun-Hua Guo, Nicholas J. Higham, and Françoise Tisseur. Detecting and solving hyperbolic quadratic eigenvalue problems. MIMS EPrint 2007.117, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, September 2007. Revised August 2008. To appear in SIAM J. Matrix Anal. Appl.

[6] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, second edition, 2002.

[7] Nicholas J. Higham, Ren-Cang Li, and Françoise Tisseur. Backward error of polynomial eigenproblems solved by linearization. *SIAM J. Matrix Anal. Appl.*, 29(4):1218–1241, 2007.

[8] Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159, 2006.

[9] Nicholas J. Higham, D. Steven Mackey, and Françoise Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.

[10] Peter Lancaster and Miron Tismenetsky. *The Theory of Matrices.* Academic Press, London, second edition, 1985.

[11] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051, 2006.

[12] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.

[13] A. S. Markus. *Introduction to the Spectral Theory of Polynomial Operator Pencils.* American Mathematical Society, Providence, RI, USA, 1988.

[14] G. W. Stewart. Perturbation bounds for the definite generalized eigenvalue problem. *Linear Algebra Appl.*, 23:69–85, 1979.

[15] G. W. Stewart and Ji-guang Sun. *Matrix Perturbation Theory.* Academic Press, London, 1990.

[16] Krešimir Veselić. A Jacobi eigenreduction algorithm for definite matrix pairs. *Numer. Math.*, 64:241–269, 1993.