Stability Theory for High Order Equations

J. C. Willems
Mathematics Institute
University of Groningen
9700 AV Groningen, The Netherlands

and

P. A. Fuhrmann*†
Department of Mathematics
Ben-Gurion University of the Negev
Beer Sheva, Israel

Submitted by Daniel Hershkowitz

ABSTRACT

A Liapunov type stability theory for high order systems of differential equations is developed. This is done by reduction to the classical case, using the theory of polynomial models.

---

1. INTRODUCTION

The problem of the stability of a linear (control) system was one of the first problems of the area of control theory. The interest is usually traced to J. C. Maxwell's theory of governors (1868). However, the problem of root location of polynomials has a longer history. Since, with the work of Galois and Abel, exact determination of zeros of polynomials was proved to be impossible, interest shifted to the problem of localizing the zeros in some region of the complex plane. The unit disc and the major half planes were the regions of

*Earl Katz Family Chair in Algebraic System Theory.

†Partially supported by the Israeli Academy of Sciences.

655 Avenue of the Americas, New York, NY 10010 0024-3795/92/$05.00

greatest interest. The problem of root location was already solved by Hermite (1856). But in this connection the work of Routh (1877) turned out to be important because of the efficiency of the computational algorithm. In the same way the work of Hurwitz (1895) was significant for its connection with topological problems.

In a major contribution to the subject, Liapunov (1893) offered a completely different approach based on energy considerations. In the linear case the Liapunov theory reduces the study of the stability of a system of first order homogeneous constant coefficient differential equations to the positive definiteness of the solution of the celebrated Liapunov equation.

For finite dimensional constant coefficient linear systems

$$\frac{dx}{dt} = Ax \tag{1}$$

with $A \in \mathbf{C}^{n \times n}$, this leads to the theory of the Liapunov equation

$$A^*Q + QA = -C^*C, \tag{2}$$

which relates the system matrix A to the self-adjoint matrix P and $C \in \mathbf{C}^{p \times n}$. The central result in this area is the Wimmer (1974) version of the Liapunov theorem that states the equivalence of the following three conditions:

1. $dx/dt = Ax$ is asymptotically stable.
2. There exists a positive definite Q and C solving (2) with (A, C) observable.
3. For every C such that (A, C) is observable there exists a positive definite Q satisfying (2).

It is clear that Liapunov theory is concerned primarily with stability of systems described by first order differential equations.

As it is trivial to reduce a scalar nth order homogeneous equation to a first order system, it became possible to derive the classical stability criteria for polynomials from Liapunov theory. This was done surprisingly late, and P. C. Parks (1962) is usually considered to have given the first of such derivations. The various reductions seemed to work also for the case of a high order system of equations of the form

$$y^{(n)} + P_{n-1}y^{(n-1)} + \cdots + P_0 y = 0,$$

i.e. for which the polynomial matrix $P(z) = Iz^n + P_{n-1}z^{n-1} + \cdots + P_0$ is monic. Strangely, a gap remained in the theory, and that is finding an algebraic test for the asymptotic stability of solutions of a system of the form

$$P_n y^{(n)} + P_{n-1} y^{(n-1)} + \cdots + P_0 y = 0$$

where $P(z) = P_n z^n + P_{n-1} z^{n-1} + \cdots + P_0$ is an arbitrary nonsingular polynomial matrix. It is our aim in this paper to close this gap.

Our approach to the problem is by reduction to Wimmer's form of Liapunov's theorem. This reduction is achieved via the use of polynomial model theory. We identify the polynomial matrix analogue of the Liapunov equation. With a solution of this equation we construct a two variable polynomial matrix. We study the operator between two polynomial models induced by such a matrix. In a special case of symmetry this map induces a quadratic form on a polynomial model. This leads to the required reduction.

While working on this paper we came upon one of the lesser-known papers of Kalman (1970), which utilizes a similar idea of switching from a polynomial equation in one variable to a polynomial in two variables and its associated quadratic form. The reading of that paper was particularly helpful to us, and we would like to acknowledge this. Similar ideas were used also in Kalman (1969). Both papers deal solely with the scalar case.

The paper is structured as follows. In Section 2 we collect the necessary results from the theory of polynomial models. Section 3 is devoted to the study of two variable polynomial matrices and the operators and Hermitian forms they induce in polynomial models. In Section 4 we prove the main result of this paper, namely a stability test for high order equations. In the last section we use this to derive the classical Hermite stability test in terms of Bezoutians or equivalently in terms of the Hermite-Fujiwara form.

The derivation of classical stability criteria is a well-trodden research area in which it is difficult to do justice to all contributors. We mention such previous work as the classic paper of Krein and Naimark (1936), and the work of Kalman (1969, 1970), which is particularly close in spirit to our approach. In connection with work on high order systems we would like to mention Lerer and Tismenetsky (1982) and some recent work by Lerer, Rodman, and Tismenetsky (1991), though the approach and results in these papers differ from ours.

2. POLYNOMIAL MODELS

If V is a finite dimensional vector space over a field F, then $V[z]$, the space of vector polynomials, is a free finitely generated module over the polynomial ring $F[z]$. Throughout we will assume a basis has been chosen, and thus V will be identified with F^n and similarly $V[z]$ with $F^n[z]$. Also we will identify $F^n[z]$ and $(F[z])^n$ and speak of its elements as polynomial vectors. Similarly, elements of $F^{m\times n}[z]$ will be referred to as polynomial matrices. Because of the nature of the factorization results we are interested in, and for consistency of notation, we will identify the field F with the complex field $\mathbf{C}$, noting that some of the results hold in greater generality.

Our starting point is this basic result about free modules.

THEOREM 2.1. *Let R be a principal ideal domain and M a free left R-module with n basis elements. Then every R-submodule N of M is free and has at most n basis elements.*

The leads to the basic representation theorem for submodules.

THEOREM 2.2. *A subset M of $\mathbf{C}^n[z]$ is a submodule of $\mathbf{C}^n[z]$ if and only if $M = P\mathbf{C}^n[z]$ for some P in $\mathbf{C}^{n\times n}[z]$.*

The following is the basic theorem that relates submodule inclusion to factorization.

THEOREM 2.3. *Let $M = P\mathbf{C}^n[z]$ and $N = Q\mathbf{C}^n[z]$. Then $M \subset N$ if and only if $P = QR$ for some R in $\mathbf{C}^{n\times n}[z]$.*

Let π_+ and π_- denote the projections of $\mathbf{C}^m((z^{-1}))$, the space of truncated Laurent series, on $\mathbf{C}^m[z]$ and $z^{-1}\mathbf{C}[[z^{-1}]]$, the space of formal power series vanishing at infinity, respectively. Since

$$\mathbf{C}^m\big((z^{-1})\big) = \mathbf{C}^m[z] \oplus z^{-1}\mathbf{C}^m\big[[z^{-1}]\big], \tag{3}$$

π_+ and π_- are complementary projections. Given a nonsingular polynomial matrix P in $\mathbf{C}^{m\times m}[z]$, we define two projections, π_P in $\mathbf{C}^m[z]$ and π^P in $z^{-1}\mathbf{C}^m[[z^{-1}]]$, by

$$\pi_P f = P\pi_- P^{-1} f \qquad \text{for} \quad f \in \mathbf{C}^m[z], \tag{4}$$

$$\pi^P f = \pi_- P^{-1}\pi_+ Ph \qquad \text{for} \quad h \in z^{-1}\mathbf{C}^m\big[[z^{-1}]\big], \tag{5}$$

and define two linear subspaces of $\mathbf{C}^m[z]$ and $z^{-1}\mathbf{C}^m[[z^{-1}]]$ by

$$X_P = \operatorname{Im} \pi_P \tag{6}$$

and

$$X^P = \operatorname{Im} \pi^P. \tag{7}$$

An element f of $\mathbf{C}^m[z]$ belongs to X_P if and only if $\pi_+ P^{-1} f = 0$, i.e. if and only if $P^{-1}f$ is a strictly proper rational vector function.

We turn X_P into an $\mathbf{C}[z]$-module by defining

$$p \cdot f = \pi_P p f \qquad \text{for} \quad p \in \mathbf{C}[z], \quad f \in X_P. \tag{8}$$

Since $\operatorname{Ker} \pi_P = P\mathbf{C}^m[z]$, it follows that X_P is isomorphic to the quotient module $\mathbf{C}^m[z]/P\mathbf{C}^m[z]$.

THEOREM 2.4. *With the previously defined module structure X_P is isomorphic to $\mathbf{C}^n[z]/P\mathbf{C}^n[z]$.*

In X_P we will focus on a special map S_P, a generalization of the classical companion matrix, which corresponds to the action of the identity polynomial z, i.e.,

$$S_P f = \pi_P z f \qquad \text{for} \quad f \in P.$$

Thus the module structure in X_P is identical to the module structure induced by S_P through $p \cdot f = p(S_P)f$. With this definition the study of S_P is identical to the study of the module structure of X_P.

The following theorem, which characterizes the spectrum of S_P, is important for the analysis of stability inasmuch as P is stable if and only if S_P is.

THEOREM 2.5. *A complex number α is an eigenvalue of S_P if and only if α is a zero of* $\det P(z)$.

Let X be a finite dimensional complex vector space, and let X^* be its dual space. We will use the notation $\langle x, x^* \rangle$ for the pairing of the two spaces, i.e. for the action of the functional x^* on x. For consistency with complex inner product spaces we will assume $\langle x, x^* \rangle$ is linear in x and antilinear in x^*. Since finite dimensional vector spaces are reflexive, we will identify X^{**} with X. Thus we have

$$\langle x^*, x \rangle = \overline{\langle x, x^* \rangle}\,. \tag{9}$$

Consider now a linear map $T: X \to Y$ where X, Y are complex vector spaces with duals Y^*, X^*, respectively. The adjoint map $T^*: Y^* \to X^*$ is determined through the equality

$$\langle Tx, y^* \rangle = \langle x, T^* y^* \rangle. \tag{10}$$

This makes the following definition natural:

$$\langle x, T^* \rangle = Tx. \tag{11}$$

This in turn leads to the following useful computational rule:

$$\langle x, T^* S^* \rangle = STx = S \langle x, T^* \rangle. \tag{12}$$

We extend now the notion of self-adjointness to this setting. A map $Z: X \to Y$ will be called *self-adjoint* or *Hermitian* if, for all $f, g \in X$,

$$\langle Zf, g \rangle = \langle f, Zg \rangle. \tag{13}$$

We will say that Z is positive if $\langle Zf, f \rangle > 0$ for all nonzero f in X.

If $\mathscr{B}$ is a basis in X, and $\mathscr{B}^*$ is its dual basis, then the bilinear form $\langle Zf, g \rangle$ can be evaluated as $([Z]_{\mathscr{B}}^{\mathscr{B}^*}[f]^{\mathscr{B}}, [g]^{\mathscr{B}})$. Here (ξ, η) is the standard inner product on $\mathbf{C}^m$. Thus Z is positive if and only if $[Z]_{\mathscr{B}}^{\mathscr{B}^*}$ is a positive definite matrix.

We denote by T^* the adjoint of the matrix T, i.e. $(T\xi, \eta) = (\xi, T^*\eta)$. We define, for an element $A(z) = \sum_{j=-\infty}^{\infty} A_j z^j$ of $\mathbf{C}^{m \times m}((z^{-1}))$, $\tilde{A}$ by

$$\tilde{A}(z) = \sum_{j=-\infty}^{\infty} A_j^* z^j. \tag{14}$$

In $\mathbf{C}^m((z^{-1})) \times \mathbf{C}^m((z^{-1}))$ we define a symmetric bilinear form $[f, g]$ by

$$[f, g] = \sum_{j=-\infty}^{\infty} \left(f_j, g_{-j-1} \right), \tag{15}$$

where $f(z) = \sum_{j=-\infty}^{\infty} f_j z^j$ and $g(z) = \sum_{j=-\infty}^{\infty} g_j z^j$. It is clear that, as both f and g are truncated Laurent series, the sum in (8) is well defined, containing only a finite number of nonzero terms. We denote by T^* the adjoint of a map T relative to the bilinear form of (8), i.e., T^* is the unique map that satisfies

$$[Tf, g] = [f, T^* g] \tag{16}$$

for all $f, g \in \mathbf{C}^m(z^{-1})$). We use this global bilinear form to obtain a concrete representation of X_P^*, the dual space of X_P.

THEOREM 2.6. *X_P^*, the dual space of X_P can be identified with $X_{\tilde{P}}$ under the pairing*

$$\langle f, g\rangle = \left[P^{-1}f, g\right] \tag{17}$$

for $f \in X_P$ and $g \in X_{\tilde{P}}$. Moreover the module structures of X_P and $X_{\tilde{P}}$ are related through

$$S_P^* = S_{\tilde{P}}. \tag{18}$$

3. TWO VARIABLE POLYNOMIALS

We will discuss briefly two variable polynomial matrices and study the naturally induced linear maps and Hermitian forms in complex polynomial models.

Let $\mathbf{C}^{n_1 \times n_2}[z, w]$ denote the $n_1 \times n_2$ complex polynomial matrices in the complex variables z and w. For a matrix A let A^* denote the Hermitian adjoint. For $V \in \mathbf{C}^{n_1 \times n_2}[z, w]$ we define $\tilde{V} \in \mathbf{C}^{n_2 \times n_1}[z, w]$ by

$$\tilde{V}(z, w) = V(\bar{w}, \bar{z})^*.$$

Assume now $M \in \mathbf{C}^{p \times m}[z, w]$. Let $P \in \mathbf{C}^{m \times m}[z]$ and $R \in \mathbf{C}^{p \times p}[z]$ be two nonsingular polynomial matrices. Let X_P and X_R be the associated polynomial models. We extend the definition of the projection π_R to act on polynomial matrices. If needed, in case of two variable polynomial matrices, we will add the variable with respect to which the action takes place. Thus

$$\pi_R^z M(z, w) = R(z)\pi_-^z\left[R(z)^{-1}M(z, w)\right].$$

Analogously we define a right-acting projection map ${}_P^w\pi$ by

$$M(z, w)\,{}_P^w\pi = \left\{\left[M(z, w)P(w)^{-1}\right]\pi_-^w\right\}P(w).$$

With these definitions out of the way, we can define the map $M: X_P \to X_R$ induced by $M(z, w)$ as $(f \in X_P)$

$$Mf(z) = \pi_R^z\left\langle f, \pi_{\tilde{P}}^w \tilde{M}(w, \bar{z})\right\rangle_w. \tag{19}$$

Clearly M is a linear map.

PROPOSITION 3.1. *Let $M \in \mathbf{C}^{p \times m}[z, w]$. Let the map $M : X_P \to X_R$ be defined by* (19). *Then the following statements hold:*

1. *The induced map M satisfies $M = 0$ if and only if there exist $M_i(z, w)$ such that $M(z, w) = R(z, w)M_1(z, w) + M_2(z, w)P(w)$.*
2. *The adjoint map $M^* : X_{\tilde{R}} \to X_{\tilde{P}}$ is given by*

$$M^* g(w) = \pi_{\tilde{P}}^{w}\left\langle g, \pi_{\tilde{R}}^{z} M(z, \overline{w})\right\rangle_z .$$

Proof. 1: Clear.

2: We will base our proof on the fact that for $M \in \mathbf{C}^{p \times m}[z, w]$ we have a (nonunique) representation

$$\pi_{\tilde{R}}^{z} M(z, w)_{P}^{w}\pi = \sum_i R_i(z) P_i(w)$$

with $R_i \in X_R$ and $P_i^T \in X_P$. Thus it suffices to show the result for the case $M(z,w) = R_i(z)P_i(w)$. This we easily compute

$$\begin{aligned}
\langle Mf, g\rangle &= \left\langle \left\langle f, \tilde{M}(\cdot, \bar{z})\right\rangle_w, g\right\rangle_z = \left\langle \left\langle f, \tilde{P}_i(\cdot)\tilde{R}_i(\bar{z})\right\rangle_w, g\right\rangle_z \\
&= \left\langle R_i(z)\left\langle f, \tilde{P}_i\right\rangle_w, g\right\rangle_z = \langle f, \tilde{P}_i\rangle_w \langle R_i, g\rangle_z \\
&= \left\langle f, \overline{\langle R_i, g\rangle_z}\tilde{P}_i\right\rangle = \left\langle f, \left\langle g, R_i\right\rangle_z \tilde{P}_i\right\rangle_w \\
&= \left\langle f, \left\langle g, R_i\right\rangle_w \tilde{P}_i\right\rangle_z \\
&= \left\langle f, \left\langle g, R_i(\cdot)P_i(\bar{z})\right\rangle_w\right\rangle_z \\
&= \left\langle f, \tilde{P}_i\left\langle g, R_i\right\rangle_w\right\rangle_z .
\end{aligned}$$

So

$$(M^* g)(z) = \left\langle g, M(\cdot, \bar{z})\right\rangle_w . \qquad \blacksquare$$

An element $V \in \mathbf{C}^{n \times n}[z, w]$ will be called *Hermitian* if

$$\tilde{V}(z, w) = V(z, w).$$

It is easy to see that this is equivalent to the conditions $V_{ij} = V_{ji}^*$.

PROPOSITION 3.2. *Let $M(z,w)$ be Hermitian. Then the induced map $M: X_P \to X_{\tilde{P}}$ is given by*

$$Mf = \pi_{\tilde{P}}^{z}\left\langle f, \pi_{\tilde{P}}^{w}\tilde{M}(\cdot, \bar{z})\right\rangle_{w} \tag{20}$$

and is a Hermitian map.

An element $M \in \mathbf{C}^{n\times n}[z,w]$ with $M(z,w) = \sum M_{ij} z^{i-1} w^{j-1}$ will be called *nonnegative*, and denoted $M \geqslant 0$, if and only if for all $\xi_i \in \mathbf{C}^n$ we have

$$\sum_{i,j}\left(M_{ij}\xi_j, \xi_i\right) \geqslant 0.$$

It is easily seen that $M \geqslant 0$ if and only there exists some $C \in \mathbf{C}^{k\times n}[z]$ such that

$$M(z,w) = \tilde{C}(z)C(w).$$

In the case of a scalar polynomial $M(z,w)$ we can interpret the matrix (m_{ij}) as a matrix representation of the map M.

PROPOSITION 3.3. *We have*

$$M = [M]_{\mathrm{co}}^{\mathrm{st}}$$

Proof. Let $\{e_1, \dots, e_n\}$ be the control basis. Putting $f = e_k$ we have $f_i = \delta_{ki}$. So

$$\begin{aligned} Me_k &= \left\langle e_k, \tilde{M}(\cdot, \bar{z})\right\rangle_w \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}(e_k)_j z^{i-1} = \sum_{i=1}^{n}\sum_{j=1}^{n} m_{ij}\,\delta_{kj} z^{i-1} \\ &= \sum_{i=1}^{n} m_{ik} z^{i-1}, \end{aligned}$$

which proves the claim. ∎

Consider now the special case of the Hermitian polynomial matrix $M(z, w) = \tilde{R}(z)R(w)$ with $R \in X_P$, i.e. with RP^{-1} strictly proper. We define $M : X_P \to X_{\tilde{P}}$ by

$$\begin{aligned} Mf &= \left\langle f, \tilde{M}(\cdot, \bar{z})\right\rangle \\ &= \left\langle f, \tilde{R}(\cdot)R(\bar{z})\right\rangle = R(\bar{z})^*\langle f, \tilde{R}\rangle \\ &= \tilde{R}\langle f, \tilde{R}\rangle. \end{aligned}$$

This implies

$$\langle Mf, f\rangle = \left\langle \tilde{R}\left\langle f, \tilde{R}\right\rangle, f\right\rangle = \langle f, \tilde{R}\rangle\langle \tilde{R}, f\rangle = \left|\langle f, \tilde{R}\rangle\right|^2.$$

If we drop the assumption that $R \in X_P$ and still take $M(z, w) = \tilde{R}(z)R(w)$, the map $M : X_P \to X_{\tilde{P}}$ can be defined by

$$\begin{aligned} Mf &= \pi_{\tilde{P}}^{z}\left\langle f, \pi_{\tilde{P}}^{\bar{z}}\tilde{M}(\cdot, \bar{z})\right\rangle \\ &= \left\langle f, \pi_{\tilde{P}}^{\bar{z}}\tilde{M}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi\right\rangle \end{aligned}$$

Here the projection ${}_P\pi$ is defined by

$$N(w, z)\, {}_{P}^{z}\pi = \left[N(w, z)P(z)^{-1}\pi_-\right]P(z).$$

In our case

$$\begin{aligned} Mf &= \pi_{\tilde{P}}^{z}\left\langle f, \pi_{\tilde{P}}^{w}\tilde{R}(w)R(\bar{z})\right\rangle \\ &= \pi_{\tilde{P}}^{z}R(\bar{z})^*\left\langle f, \pi_{\tilde{P}}^{w}\tilde{R}(w)\right\rangle \\ &= \pi_{\tilde{P}}^{z}\tilde{R}(z)\left\langle f, \pi_{\tilde{P}}^{w}\tilde{R}\right\rangle \end{aligned}$$

We can view this from a slightly different perspective. Let $Z : \mathbf{C} \to X_{\tilde{P}}$ be given by

$$Z\alpha = \pi_{\tilde{P}}\tilde{R}\alpha.$$

Then $Z^* : X_P \to \mathbf{C}$ is given by

$$Z^*f = \langle f, \pi_{\tilde{P}}\tilde{R}\rangle.$$

A nonnegative polynomial in two variables does not necessarily induce a nonnegative form in X_P. We will say that a Hermitian polynomial $V(z, w)$ is *P-positive*, denoted by $V >_P 0$, if the induced Hermitian map V is positive, that is, $\langle Vf, f\rangle > 0$ for all nonzero $f \in X_P$. P-positivity has also a time domain interpretation. We will not go into the proof of the next result. We do point out however that a time domain interpretation of polynomial models can be found in Hinrichsen and Prätzel-Wolters (1980).

PROPOSITION 3.4. *Let $P \in \mathbf{C}^{m\times m}[z]$ be nonsingular, and let $V \in \mathbf{C}^{m\times m}[z, w]$ be Hermitian. Then the following conditions are equivalent:*

1. *V is P-positive.*
2. *There exist a basis $\{C_i\}$ in X_P for which*

$$V(z,w) = \sum_i \tilde{C}_i(z)C_i(w).$$

3. *For any solution w of $P(d/dt)w = 0$ that satisfies*

$$\sum_{i,j}\left(V_{ij}\frac{d^{(i-1)}w}{dt^{i-1}}, \frac{d^{(j-1)}w}{dt^{j-1}}\right) = 0$$

we have $w = 0$.

4. STABILITY FOR HIGH ORDER EQUATIONS

In this section we extend Liapunov's theorem to higher order systems of equations.

We proceed now to the establishment of stability criteria for complex polynomials. This is done by reduction to the standard Liapunov theorem.

We recall that a complex matrix A is called stable if all its eigenvalues lie in the open left half plane. Clearly a matrix A is stable if and only if its characteristic polynomial is stable.

Next we state a slightly modified version of Wimmer's form of the theorem of Liapunov (1893).

THEOREM 4.1. *Let X and U be two complex vector spaces. Let $A: X \to X$ and $C: X \to U$ be linear transformations such that the pair (A, C) is observable. Then A is stable if and only if the* Liapunov equation

$$A^*Q + QA = -C^*C \tag{21}$$

has a unique solution $Q: X \to X^*$ for which the form $\langle Qx, x\rangle$ is positive definite.

We proceed to give our main theorem.

THEOREM 4.2. *Let P be a real, nonsingular polynomial matrix. Then P is stable if and only if, for any polynomial matrix R for which P and R are right coprime and RP^{-1} is strictly proper, there exists a solution Q of the equation*

$$\tilde{P}(-s)Q(s) + \tilde{Q}(-s)P(s) = \tilde{R}(-s)R(s) \tag{22}$$

for which the quadratic form induced in X_P by

$$V(z,w) = \frac{\tilde{P}(z)Q(w) + \tilde{Q}(z)P(w) - \tilde{R}(w)R(z)}{z+w} \tag{23}$$

is positive definite.

Proof. Assume P is stable and R is left coprime with P. The coprimeness condition implies (see Fuhrmann, 1976) that the pair (A, B) defined, in the state space X_P, by

$$\begin{aligned} A &= S_P, \\ Cf &= \langle f, \tilde{R} \rangle \end{aligned} \tag{24}$$

is observable.

The stability of P on the other hand implies that there exists a solution to the equation (22). That this is indeed so can be seen from the following. As P is stable, so is $p = \det P$, and the polynomials p and $p_*(s) = p(-s)$ are coprime. Therefore the equation

$$p(s)K(s) + L(s)p(-s) = \tilde{R}(-s)R(s)$$

is solvable, and the solution unique if we assume L is reduced modulo p and K modulo p_*. By a symmetry argument, $K = L_*$ and hence

$$p(s)L_*(s) + L(s)p(-s) = R(s)R_*(s).$$

Using Cramer's rule, $p = P \operatorname{adj} P$, we have

$$P(s)\left[\operatorname{adj} P(s)L_*(s)\right] + \left[L(s) \operatorname{adj} P_*(s)\right]P_*(s) = R(s)R_*(s).$$

Also

$$\left[\operatorname{adj} P(s) L_*(s)\right]_* = L(s)\left[\operatorname{adj} P(s)\right]_* = L(s) \operatorname{adj} P_*(s).$$

Without loss of generality we can assume QP^{-1} to be strictly proper. As a consequence of (22), $V(z,w)$ defined by (23) is a polynomial matrix in two variables. Let

$$M(z,w) = (z+w)V(z,w) = \tilde{P}(z)Q(w) + \tilde{Q}(z)P(w) - \tilde{R}(w)R(z). \tag{25}$$

The polynomial matrix M induces a Hermitian form $\mathscr{M}$ on X_P by Equation (20). Clearly both $V(z,w)$ and $M(z,w)$ are Hermitian. Moreover we have

$$\pi_{\tilde{P}}^{z} M(z,w)\, {}_{P}^{w}\pi = -\tilde{R}(z)R(w). \tag{26}$$

Therefore

$$\langle \mathscr{M}f, f\rangle = -\langle f, \tilde{R}\rangle \tilde{R} = -C^*Cf. \tag{27}$$

On the other hand we can evaluate $\langle \mathscr{M}f, f\rangle$ differently. Using the fact that $V(z,w)$ is Hermitian,

$$\begin{aligned}
\langle \mathscr{M}f, f\rangle &= \left\langle \left\langle f, \pi_{\tilde{P}}^{w} \tilde{M}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&= \left\langle \left\langle f, \pi_{\tilde{P}}^{w} w \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z + \left\langle \left\langle f, \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, \bar{z}\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&= \left\langle \left\langle f, \pi_{\tilde{P}}^{w} w \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&\quad + \left\langle \left\langle f, \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi\, \bar{z}\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&= \left\langle \left\langle f, S_{\tilde{P}} \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&\quad + \left\langle S_{\tilde{P}} \left\langle f, \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&= \left\langle \left\langle S_P f, \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&\quad + \left\langle S_{\tilde{P}} \left\langle f, \pi_{\tilde{P}}^{w} \tilde{V}(\cdot, \bar{z})\, {}_{P}^{\bar{z}}\pi \right\rangle_w, f \right\rangle_z \\
&= \langle VS_P f, f\rangle + \langle S_{\tilde{P}} V f, f\rangle \\
&= \left\langle \left(S_{\tilde{P}} V + VS_P\right) f, f \right\rangle = \left\langle \left(A^* V + VA\right) f, f \right\rangle.
\end{aligned} \tag{28}$$

Combining the two equalities, we get

$$A^*V + VA = -C^*C. \tag{29}$$

Thus by Liapunov's theorem the Hermitian form $\langle Vf, f\rangle$ is positive definite, or alternatively $V >_P 0$.

Conversely, assume Q is a solution of (22) and the quadratic form induced by $V(z,w)$ as defined in (23) is positive definite. Then, since $S_{\tilde{P}} = S_P^*$, as before,

$$S_P^*V + VS_P = -C^*C. \tag{30}$$

This implies, by Liapunov's theorem, the stability of S_P and hence, by Theorem 2.5, that of P. ■

We wish to remark that with the same technique we could derive an analogue of the Ostrowski-Schneider (1962) inertia theorem.

5. DERIVATION OF CLASSICAL STABILITY CRITERION THEORY

We can use the results of the previous section to obtain an easy derivation of the classical stability criteria for real and complex scalar polynomials.

THEOREM 5.1. *A necessary and sufficient condition for a complex polynomial p to be a Hurwitz polynomial is that the* Hermite-Fujiwara *form, which is defined as the Hermitian form with generating function*

$$\frac{\tilde{p}(z)p(w) - p(-z)\tilde{p}(-w)}{z+w}, \tag{31}$$

is positive definite.

Proof. Without loss of generality we can assume p is monic. We split the proof into two cases.

Case I: deg p *is odd.* In this case we set $r(z) = p(z) + \tilde{p}(-z)$. Clearly deg $r <$ deg p. We observe that $q(z) = p(z) + \tilde{p}(-z)$ is a solution of the equation

$$\tilde{p}(s)q(-s) + \tilde{q}(s)p(-s) = \tilde{r}(s)r(-s).$$

Indeed,

$$
\begin{aligned}
&\tilde{p}(s)q(-s) + \tilde{q}(s)q(-s) \\
&\quad = \tilde{p}(s)[p(-s) + \tilde{p}(s)] + [p(-s) + \tilde{p}(s)]p(-s) \\
&\quad = [p(-s) + \tilde{p}(s)][p(-s) + \tilde{p}(s)] \\
&\quad = \tilde{r}(s)r(-s).
\end{aligned}
$$

Assume now that p is stable. This implies the coprimeness of $p(s)$ and $\tilde{p}(-s)$ and hence that of p and r. By Theorem 4.2 the Hermitian form induced by

$$
\frac{\tilde{p}(z)q(w) + \tilde{q}(z)p(w) - \tilde{r}(z)r(w)}{z + w} \tag{32}
$$

is positive definite. We compute now

$$
\begin{aligned}
&\tilde{p}(z)q(w) + \tilde{q}(z)p(w) - \tilde{r}(z)r(w) \\
&\quad = \tilde{p}(z)[p(w) + \tilde{p}(-w)] + [p(-z) + \tilde{p}(z)]p(w) \\
&\qquad - [p(-z) + \tilde{p}(z)][p(w) + \tilde{p}(-w)] \\
&\quad = \tilde{p}(z)p(z) - p(-z)\tilde{p}(-w).
\end{aligned} \tag{33}
$$

Thus the two Hermitian forms

$$
\frac{\tilde{p}(z)q(w) + \tilde{q}(z)p(w) - \tilde{r}(z)r(w)}{z + w}
$$

and

$$
\frac{\tilde{p}(z)p(w) - p(-z)\tilde{p}(-w)}{z + w}
$$

are equal. In particular the last form is positive definite.

Conversely, assume the form (31) is positive definite. Then so is (32), with q and r given as before. The positive definiteness of (31) implies the coprimeness of $p(s)$ and $\tilde{p}(-s)$.

To see this we note that if (32) is positive definite, then by a change of variable $w = -\zeta$ the bilinear form

$$\frac{\tilde{p}(z)p(-\zeta) - p(-z)\tilde{p}(\zeta)}{z-\zeta}, \tag{34}$$

which is just the Bezoutian of $\tilde{p}(z)$ and $p(-z)$, is nonsingular. Thus $p(z)$ and $p(-s)$ are coprime. We can apply now Theorem 4.2 to infer the stability of p.

Case II: deg p *is even*. In this case we set $r(z) = p(z) - \tilde{p}(-z)$. Clearly deg $r <$ deg p. We observe that $q(z) = p(z) - \tilde{p}(-z)$ is a solution of the equation

$$\tilde{p}(-s)q(s) + \tilde{q}(-s)p(s) = \tilde{r}(-s)r(s),$$

for

$$\begin{aligned}\tilde{p}(-s)q(s) + \tilde{q}(-s)q(s) &= \tilde{p}(-s)[p(s) + \tilde{p}(-s)] + [\tilde{p}(-s) - p(s)]\,p(s)\\ &= [\tilde{p}(-s) - p(s)][p(s) + \tilde{p}(-s)]\\ &= \tilde{r}(-s)r(s).\end{aligned}$$

We compute now

$$\begin{aligned}&\tilde{p}(z)q(w) + \tilde{q}(z)p(w) - \tilde{r}(z)r(w)\\ &\quad= \tilde{p}(z)[p(w) - \tilde{p}(-w)] + [\tilde{p}(z) - p(-z)]\,p(w)\\ &\qquad -[\tilde{p}(z) - p(-z)][p(w) - \tilde{p}(-w)]\\ &\quad= \tilde{p}(z)p(w) - p(-z)\tilde{p}(-w).\end{aligned} \tag{35}$$

The rest of the proof goes as before. ∎

In the case of real polynomials the Hermite-Fujiwara form admits a further reduction. To this end we introduce the end and odd parts of a real polynomial. Given a real polynomial $p(z) = \sum_{j\geqslant 0} p_j z^j$, then the polynomials p_+ and p_- are defined by

$$p_+(z) = \sum_{j\geqslant 0} p_{2_j} z^j$$

and

$$p_-(z) = \sum_{j \geqslant 0} p_{2_{j+1}} z^j.$$

Then

$$\begin{aligned} p(z) &= p_+(z^2) + zp_-(z^2), \\ p_*(z) &= p_+(z^2) - zp_-(z^2). \end{aligned} \tag{36}$$

We observe that the Bezoutian $B(p, p_*)$ and the Hermite-Fujiwara form have useful decompositions.

THEOREM 5.2. *Let p be a real polynomial.*

1. *We have the following isomorphism of quadratic forms. For the Hermite-Fujiwara form $H(p)$,*

$$H(p) = 2B(zp_-, p_+) \oplus 2(-1)B(p_-, p_+) = 2B(zp_-, p_+) \oplus 2B(p_+, p_-),$$

whereas for the Bezoutian $B(p, p_*)$,

$$B(p, p_*) = 2B(zp_-, p_+) \oplus 2B(p_-, p_+).$$

2. *The Hermite-Fujiwara form is positive definite if and only if the two Bezoutians $B(q_+, q_-)$ and $B(zq_-, q_+)$ are positive definite.*

Proof. 1: Since p is real, we have $p = \tilde{p}$. From (36) it follows that $p(-z) = p_+(z^2) - zp_-(z^2)$. Therefore

$$\begin{aligned} &\frac{p(z)p(w) - p(-z)p(-w)}{z+w} \\ &= \frac{\begin{array}{c}[p_+(z^2) + zp_-(z^2)][p_+(w^2) + wp_-(w^2)] \\ -[p_+(z^2) - zp_-(z^2)][p_+(w^2) - wp_-(w^2)]\end{array}}{z+w} \\ &= 2\frac{zp_-(z^2)p_+(w^2) + p_+(z^2)wp_-(w^2)}{z+w} \end{aligned}$$

$$= 2\frac{z^2p_-(z^2)p_+(w^2) - p_+(z^2)w^2p_-(w^2)}{z^2 - w^2}$$

$$- 2zw\frac{p_-(z^2)p_+(w^2) - p_+(z^2)p_-(w^2)}{z^2 - w^2}$$

One form contains only even terms and the other only odd ones, so this proves the first statement.

By a change of variable $w = -\zeta$ the Hermite-Fujiwara form transforms into

$$\frac{\tilde{p}(z)p(-\zeta) - p(-z)\tilde{p}(\zeta)}{z - \zeta}, \tag{37}$$

which is just the Bezoutian of $\tilde{p}(z)$ and $p(-z)$. However, this change of variable affects only the terms in

$$2zw\frac{p_-(z^2)p_+(w^2) - p_+(z^2)p_-(w^2)}{z^2 - w^2},$$

and only by a change of sign.

Part 2 follows from the direct sum representation of the Hermite-Fujiwara form. ■

As a direct corollary of this we get the following classical result.

THEOREM 5.3. *Let $p(z)$ be monic of degree n. Then the following statements are equivalent*:

(i) *$p(z)$ is a stable, or Hurwitz, polynomial.*
(ii) *The Hermite-Fujiwara form is positive definite.*
(iii) *The two Bezoutians $B(p_+, p_-)$ and $B(zp_-, p_+)$ are positive definite.*

REFERENCES

Coppel, W. A. 1974a. Matrix quadratic equations, *Bull. Austral. Math. Soc.* 10:377–401.

Coppel, W. A. 1974b. Matrices of rational functions, *Bull. Austral. Math. Soc.* 11:89–113.

Fuhrmann, P. A. 1976. Algebraic system theory: An analyst's point of view, *J. Franklin Inst.* 301:521–540.

Fuhrmann, P. A. 1977. On strict system equivalence and similarity, *Internat. J. Control* 25:5–10.
Fuhrmann, P. A. 1981. Polynomial models and algebraic stability criteria, in *Proceedings of Joint Workshop on Synthesis of Linear and Nonlinear Systems*, Bielefeld, June 1981, pp. 78–90.
Gantmacher, F. R. 1959. *The Theory of Matrices*, Chelsea, New York.
Gruber, M. 1968. Stability Analysis Using Exact Differentials, MIT Report ESL-TM-369.
Helmke, U. and Fuhrmann, P. A. 1989. Bezoutians, *Linear Algebra Appl.* 122–124:1039–1097.
Hinrichsen, D. and Pratzel-Wolters, D. 1980. Solution models and system equivalence, *Internat. J. Control* 32:777–802.
Hermite, C. 1856. Sur le nombre des racines d'une équation algébrique comprise entre des limites donnés, *J. Reine Angew. Math.* 52:39–51.
Hurwitz, A. 1895. Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reelen Teilen besitzt, *Math. Ann.* 46:273–284.
Kalman, R. E. 1969. Algebraic characterization of polynomials whose zeros lie in algebraic domains, *Proc. Nat. Acad. Sci. U.S.A.* 64:818–823.
Kalman, R. E. 1970. New algebraic methods in stability theory, in *Proceedings V. International Conference on Nonlinear Oscillations*, Kiev.
Krein, M. G. and Naimark, M. A. 1936. The method of symmetric and Hermitian forms in the theory of the separation of the roots of algebraic equations, English transl. *Linear and Multilinear Algebra* 10:265–308 (1981).
Lerer, L., Rodman, L., and Tismenetsky, M. 1991. Inertia theorems for matrix polynomials, *Linear and Multilinear Algebra* 30:157–182.
Lerer, L. and Tismenetsky, M. 1982. The Bezoutian and the eigenvalue-separation problem for matrix polynomials, *Integral Equations Operator Theory* 5:386–445.
Liapunov, A. M. 1893. Problème général de la stabilité de mouvement, *Ann. Fac. Sci. Toulouse* 9:203–474 (1907) (French transl. of the Russian paper published in *Comm. Soc. Math. Kharkow*).
Maxwell, J. C. 1868. On governors, *Proc. Roy. Soc. London Ser. A* 16:270–283.
Ostrowski, A. and Schneider, H. 1962. Some theorems on the inertia of general matrices, *J. Math. Anal. Appl.* 4:72–84.
Parks, P. C. A new proof of the Routh-Hurwitz stability criterion using the second method of Lyapunov, *Proc. Cambridge Philos. Soc.* 58:694–702.
Routh, E. J. 1877. *A Treatise on the Stability of a Given State of Motion*, Macmillan, London.
Wimmer, H. 1974. Inertia theorems for matrices, controllability and linear vibrations, *Linear Algebra Appl.* 8:337–343.

Received 11 April 1991; final manuscript accepted 22 October 1991