ON THE CONVERGENCE OF THE NEWTON ITERATION FOR SPECTRAL FACTORIZATION

Jan C. Willems*

Abstract

The problem of factoring a polynomial matrix $B = B^*$ into $B(\xi) = H(\xi)^* H(\xi)$, with H Hurwitz, is called the Hurwitz spectral factorization problem. We show that the Newton iteration, applied to the computation of H, converges.

Keywords: Spectral factorization, Newton iteration, quadratic differential forms, two-variable polynomial matrices.

1 Introduction

It is a great pleasure for me to contribute this paper for the Festschrift for Paul Fuhrmann at the occasion of his 60–th birthday. Paul was perhaps the first pure–mathematics trained researcher that I had occasion to collaborate with. At first, this collaboration was mainly in the form of exchanges of ideas, but later on, we co-authored a few papers as well [1], [2], [8]. In many ways, seeing how a mathematics trained researcher as Paul approached problems in systems and control was – well over 25 years ago – a new experience to me, and I remember admiring in a somewhat starry-eyed way, the creativity that the internal logic of mathematics could lead to in a field as systems and control. Paul's virtuosity, his impressive oeuvre and boundless creativity, and his warmth, energy and friendship brought me closer to it than I have ever been.

The purpose of this paper is to explain some ideas on the interaction of one- and two-variable polynomial matrices that have emerged in our present research program and that are more than reminiscent of Paul Fuhrmann's work.

We start with a few items of notation. Throughout, we denote one-variable polynomials with real coefficients by $\mathbb{R}[\xi]$, and two-variable polynomials by $\mathbb{R}[\zeta, \eta]$ (thus ξ is usually the indeterminate in the one-variable case, and ζ, η are the indeterminates in the two-variable case). Polynomial matrices are denoted by

*Mathematics Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands; fax: +31503633976 `e-mail: J.C.Willems@math.rug.nl`.

$\mathbb{R}^{\bullet\times\bullet}[\xi], \mathbb{R}^{\bullet\times n}[\xi]$ (when the number of rows is not specified), etc., with similar notation in the two-variable case. We will use some special operators acting on polynomial matrices: $*, \star, \bullet$ and ∂ :

- $*$ maps $\mathbb{R}^{\bullet\times\bullet}[\xi]$ into itself: if $P \in \mathbb{R}^{n_1\times n_2}[\xi]$, then $P^* \in \mathbb{R}^{n_2\times n_1}[\xi]$ is defined as $P^*(\xi) := P^T(-\xi)$ (^{T} denotes transposition);
- $\star$ maps $\mathbb{R}^{\bullet\times\bullet}[\zeta,\eta]$ into itself: if $P \in \mathbb{R}^{n_1\times n_2}[\zeta,\eta]$, then $P \in \mathbb{R}^{n_2\times n_1}[\zeta,\eta]$ is defined as $P^\star(\zeta,\eta) := P^T(\eta,\zeta)$;
- $\bullet$ maps $\mathbb{R}^{\bullet\times\bullet}[\zeta,\eta]$ into itself: if $P \in \mathbb{R}^{n_1\times n_2}[\zeta,\eta]$, then $\overset{\bullet}{P} \in \mathbb{R}^{n_1\times n_2}[\zeta,\eta]$ is defined as $\overset{\bullet}{P}(\zeta,\eta) := (\zeta+\eta)P(\zeta,\eta)$;
- ∂ maps $\mathbb{R}^{\bullet\times\bullet}[\zeta,\eta]$ into $\mathbb{R}^{\bullet\times\bullet}[\xi]$: if $P \in \mathbb{R}^{n_1\times n_2}[\zeta,\eta]$, then $\partial P \in \mathbb{R}^{n_1\times n_2}[\xi]$ is defined as $\partial P(\xi) := P(-\xi,\xi)$.

We call an element P of $\mathbb{R}^{\bullet\times\bullet}[\xi]$ *symmetric* if $P = P^*$, and of $R^{\bullet\times\bullet}[\zeta,\eta]$ if $P = P^\star$, and *skew-symmetric* if $P = -P^*$, or $P = -P^\star$. Note that the operator ∂ thus maps (skew-)symmetric elements to (skew-)symmetric elements. Of course, it is possible to view, in an obvious way, the one-variable polynomial matrix $P \in \mathbb{R}^{\bullet\times\bullet}[\zeta]$ as an element of $\mathbb{R}^{\bullet\times\bullet}[\zeta,\eta]$ in which it so happens that no powers of η appear. Thus when the "one-variable" polynomial matrix P is viewed with the indeterminate ζ, $P^\star$ is the "one-variable" polynomial matrix $P^T(\eta)$.

An important relation among the $\bullet$ and the ∂ operators is given in the following proposition (see [9] for a proof).

Proposition 1.1 *The image of the $\bullet$ operator equals the kernel of the ∂ operator. In other words, for $\Phi \in \mathbb{R}^{\bullet\times\bullet}[\zeta,\eta]$,*

$$\frac{\Phi(\zeta,\eta)}{\zeta+\eta}$$

is a polynomial iff $\Phi(\xi,-\xi) = 0$.

The field of rational functions over $\mathbb{R}$ is denoted by $\mathbb{R}(\xi)$; $R^{\bullet\times\bullet}(\xi)$ denotes the set of matrices of rational functions. An element $P \in \mathbb{R}(\xi)$ is said to be *proper* if $P = p_1/p_2$, with the degree of $p_1 \in \mathbb{R}[\xi]$ less than or equal to that of $p_2 \in \mathbb{R}[\xi]$, *bi-proper* if these degrees are equal, and *strictly proper* if "less than" holds. Obviously any $P \in \mathbb{R}(\xi)$ can be written as $P = P_1 + P_2$, with $P_1 \in \mathbb{R}[\xi]$ the polynomial part and $P_2 \in \mathbb{R}(\xi)$ strictly proper. The polynomial part of P is denoted by P_∞. All this can be generalized in an obvious way to $\mathbb{R}^{\bullet\times\bullet}[\xi]$. In particular, $P \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ is bi-proper iff it is square and $P_\infty \in \mathbb{R}^{\bullet\times\bullet}$ is invertible (equivalently, iff P^{-1} exists and is also proper).

A square matrix $P \in \mathbb{R}^{\bullet}[\xi]$ is said to be *Hurwitz* if $det(P)$ is a Hurwitz polynomial, i.e. a non-zero polynomial with all its roots in the open left half of the complex plane.

We are interested in the question of factoring a polynomial matrix $B = B^*$ into the product $B = F^*F$. Such factorization questions, which go under the name of *spectral factorization*, have many applications in control and signal processing, and go back to the work of Wiener in the first half of this century. There have been literally countless articles on this problem since. Formally, let $B = B^* \in \mathbb{R}^{q\times q}[\xi]$. We call the factorization $B = H^*H$ a *Hurwitz spectral factorization* of B if $H \in \mathbb{R}^{q\times q}[\xi]$ is Hurwitz. It is well known when such a factorization exists. We state this for easy reference.

Theorem 1.1 *Let $B = B^* \in \mathbb{R}^{q\times q}[\xi]$. Then there exists $H \in \mathbb{R}^{q\times q}[\xi]$ with H Hurwitz such that*

$$\boxed{B = H^*H} \tag{1}$$

iff

$$\boxed{B(i\omega) > 0 \qquad \forall \omega \in \mathbb{R}} \tag{2}$$

where (2) means that the Hermitian matrix $B(i\omega) \in \mathbb{C}^{q\times q}$ is positive definite for all $\omega \in \mathbb{R}$. This H is unique up to pre-multiplication by an orthogonal matrix.

The aim of this paper is to study the convergence of the Newton iteration as an algorithm for computing a Hurwitz spectral factor. The Newton iteration for (1), studied before by Kučera and others ([5, 3, 6]), and the convergence results are stated in the following theorem.

Theorem 1.2 *Assume that $B = B^* \in \mathbb{R}^{q\times q}[\xi]$ satisfies (2). Let $X_0 \in \mathbb{R}^{q\times q}[\xi]$ be Hurwitz and satisfy $((X_0^*)^{-1}BX_0^{-1})_\infty = I$. Then the Newton iteration*

$$\boxed{X_{k+1}^*X_k + X_k^*X_{k+1} = B + X_k^*X_k} \tag{3}$$

with the normalization condition

$$\boxed{(X_{k+1}X_k^{-1})_\infty = I} \tag{4}$$

defines a unique sequence $X_1, X_2, \ldots, X_k, \ldots \in \mathbb{R}^{q\times q}[\xi]$. Moreover, each of the X_k's is Hurwitz, and $X_k \to H$ as $k \to \infty$, with H Hurwitz and satisfying (1). This convergence is quadratic.

2 Quadratic differential forms

Let $\Phi \in \mathbb{R}^{n_1\times n_2}[\zeta,\eta]$, i.e.

$$\Phi(\zeta,\eta) = \sum_{k,\ell} \Phi_{k,\ell}\zeta^k\eta^\ell,$$

with $\Phi_{k\ell} \in \mathbb{R}^{n_1 \times n_2}$. This two-variable polynomial matrix induces the mapping

$$L_\Phi : C^\infty(\mathbb{R}, \mathbb{R}^{n_1}) \times C^\infty(\mathbb{R}, \mathbb{R}^{n_2}) \longrightarrow C^\infty(\mathbb{R}, \mathbb{R}),$$

defined by

$$L_\Phi(v, w) := \sum_{k,\ell} (\frac{d^k}{dt^k} v)^T \Phi_{k\ell} (\frac{d^\ell}{dt^\ell} w).$$

We call L_Φ a *bilinear differential form (BLDF).* For $\Phi \in \mathbb{R}^{n \times n}[\zeta, \eta]$, this leads to the *quadratic differential form (QDF)*

$$Q_\Phi : C^\infty(\mathbb{R}, \mathbb{R}^n) \to C^\infty(\mathbb{R}, \mathbb{R}),$$

defined by $Q_\Phi(w) := L_\Phi(w, w)$. Note that $Q_\Phi = Q_{\Phi^\star} = Q_{\frac{1}{2}(\Phi + \Phi^\star)}$. Accordingly, we only consider QDF's induced by symmetric Φ's. Note that $L_{\overset{\bullet}{\Phi}}(v, w) = \frac{d}{dt} L_\Phi(v, w)$ and $Q_{\overset{\bullet}{\Phi}}(w) = \frac{d}{dt} Q_\Phi(w)$. Following this correspondence of $\Phi = \Phi^\star \in \mathbb{R}^{n \times n}[\zeta, \eta]$ with Q_Φ, we will identify Φ and Q_Φ whenever there is no danger of confusion. We denote $C^\infty(\mathbb{R}, \mathbb{R}^n)$ by C^∞ when the co-domain is obvious from the context (the domain is always $\mathbb{R}$).

We say that Q_Φ (or Φ) is *nonnegative* (denoted $Q \geq 0$) if $Q_\Phi(w) \geq 0$ for all $w \in C^\infty$, and *positive* if in addition $Q_\Phi(w) = 0$ implies $w = 0$. Every $\Phi \geq 0$ can be factored as $\Phi(\zeta, \eta) = M^T(\zeta) M(\eta)$, with $M \in \mathbb{R}^{\bullet \times \bullet}[\xi]$.

We will deal with continuous-time real linear time-invariant differential systems, as discussed in [7], and amply elaborated in our recent work. Thus the time axis is $\mathbb{R}$, the signal space is $\mathbb{R}^q$ (the number of variables may of course depend on the case at hand), and the behavior $\mathfrak{B}$ is the solution set of a system of linear constant coefficient differential equations

$$R(\frac{d}{dt}) w = 0, \tag{5}$$

where $R \in \mathbb{R}^{\bullet \times q}[\xi]$. We denote the resulting system by $\Sigma_R = (\mathbb{R}, \mathbb{R}^q, \mathfrak{B}_R)$ and its behavior by

$$\mathfrak{B}_R = \{ w \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^q) \mid R(\frac{d}{dt}) w = 0 \}. \tag{6}$$

This class of systems, or alternatively its behaviors, is denoted by $\mathfrak{L}^q$. For obvious reasons we refer to (5) as a *kernel representation* of Σ_R.

We say that $\mathfrak{B} \in \mathfrak{L}^q$ is *autonomous* if $w_1, w_2 \in \mathfrak{B}$, $w_1(t) = w_2(t)$ for $t < 0$ implies $w_1 = w_2$. Autonomous systems are those that allow a representation (5) with $R \in \mathbb{R}^{q \times q}[\xi]$ nonsingular: $\det(R) \neq 0$. An autonomous system $\mathfrak{B} \in \mathfrak{L}^q$ is said to be *asymptotically stable* if $w \in \mathfrak{B}$ implies $w(t) \to 0$ as $t \to \infty$. Asymptotically stable systems correspond to those for which R can be taken to be Hurwitz.

Let $\Sigma \in \mathfrak{L}^q$ be represented by (5), and let $S \in \mathbb{R}^{\bullet \times \bullet}[\xi]$. Then

$$x = S(\frac{d}{dt})w \tag{7}$$

is called a *state map* for (5) if (5,7) jointly define a state system, i.e. if whenever $w_1, w_2 \in \mathfrak{B}_R$ satisfy $S(\frac{d}{dt})w_1(0) = S(\frac{d}{dt})w_2(0)$, then $w_1 \wedge w_2$ ($\wedge$ denotes concatenation at 0) is a weak solution of (5) and $x_1 \wedge x_2$ is absolutely continuous. A state map is *minimal* if for all $a \in \mathbb{R}^{rowdim(S)}$, there exists $w \in \mathfrak{B}_R$ such that $S(\frac{d}{dt})w(0) = a$, in other words if the map $w \in \mathfrak{B}_R \mapsto S(\frac{d}{dt})w(0)$ is surjective. For autonomous systems, the minimal state map construction can be carried out as follows. Consider the set $\{f \in \mathbb{R}^{1\times q}[\xi] \mid fR^{-1}\text{is strictly proper}\}$. This set is actually a finite-dimensional vector space. Let $S = col(f_1, f_2, \ldots, f_n)$ be a basis for this vector space. This S defines a minimal state map for (5).

We will encounter the question when two systems admit the same state map.

Lemma 2.1 *Assume that R_1, $R_2 \in \mathbb{R}^{q\times q}[\xi]$, with* $\det(R_1) \neq 0$, *and* $\det(R_2) \neq 0$, *are such that $R_2 R_1^{-1}$ is bi-proper. Then $x = S(\frac{d}{dt})w$ is a state map for $R_1(\frac{d}{dt})w = 0$ iff it is a state map for $R_2(\frac{d}{dt})w = 0$.*

Proof: This lemma follows immediately from the fact that, with $f \in \mathbb{R}^{1\times q}[\xi]$, fR^{-1} is strictly proper iff fR_2^{-1} is. ■

Let $\mathfrak{B} \in \mathfrak{L}^q$. We will call the QDF Q_Φ $\mathfrak{B}$-*nonnegative* (denoted $\Phi \overset{\mathfrak{B}}{\geq} 0$) if $Q_\Phi(w) \geq 0$ for all $w \in \mathfrak{B}$ and $\mathfrak{B}$-*positive* (denoted $\Phi \overset{\mathfrak{B}}{>} 0$) if $\Phi \overset{\mathfrak{B}}{\geq} 0$ and if $Q_\Phi(w) = 0$ implies $w = 0$. In [9] $\mathfrak{B}$ positivity is studied in depth. The following proposition from there plays a role in the sequel.

Proposition 2.1 *Let $\mathfrak{B} \in \mathfrak{L}^q$. Then $\mathfrak{B}$ is asymptotically stable if there exists $\Psi = \Psi^* \in \diagdown^{q\times q}[\zeta, \eta]$ such that $\Psi \overset{\mathfrak{B}}{\geq} 0$ and $\overset{\bullet}{\Psi} \overset{\mathfrak{B}}{<} 0$.*

3 The polynomial matrix Lyapunov equation

The analysis of (3) leads to the analysis of the (linear) polynomial equation

$$\boxed{X^* A + A^* X = B} \tag{8}$$

with A, $B = B^*$, $X \in \mathbb{R}^{q\times q}[\xi]$. In (8) we view A, B as given and X as the unknown. Because of its similarity to the case in which A, B, X are ordinary matrices, we call (8) the *polynomial matrix Lyapunov equation*. It is of much interest to study this equation in its full generality. However, we only consider the case that corresponds to the situation that is encountered in the Newton iteration (3).

Theorem 3.1 *Assume that $B = B^* \in \mathbb{R}^{q\times q}[\xi]$, and that $A \in \mathbb{R}^{q\times q}[\xi]$ is Hurwitz. Assume further that $(A^*)^{-1}BA^{-1}$ is bi-proper, normalized to $((A^*)^{-1}BA^{-1})_\infty = 2I$. Then (8) admits a solution such that XA^{-1} is bi-proper and a unique such solution with*

$$\boxed{(XA^{-1})_\infty = I.} \tag{9}$$

This unique solution satisfies $((X^*)^{-1}BX^{-1})_\infty = 2I$

Proof: Consider the more general version of (8) $YA + A^*X = B$, with X, $Y \in \mathbb{R}^{q\times q}[\xi]$ unknown. To show the existence of a solution, observe that using the Smith form for polynomial matrices, we may assume with loss of generality that A is diagonal. This equation then reduces to q^2 equations of the form $ax + b^*y = c$, with a, b, $c \in \mathbb{R}[\xi]$, a, b Hurwitz, in the unknowns $x, y \in \mathbb{R}[\xi]$. These Bezout-type equations have a solution since a, b Hurwitz implies that a, b^* are co-prime. Now use $B = B^*$ to show that $X^*A + A^*Y^* = B$ and conclude that $(X + Y^*)/2$ solves (8).

It is easy to see [3, 9], that if X is a solution, then all the solutions are generated by $X \mapsto X + SA$, where $S \in \mathbb{R}^{q\times q}[\xi]$ ranges over the skew-symmetric elements. Note further that the polynomial part of $(A^*)^{-1}X^* + XA^{-1}$ equals $2I$. The solution $X' = X + SA$, with S such that the polynomial part of $XA^{-1} + S$ equals I, yields a solution such that XA^{-1} is bi-proper with polynomial part I. The above representation of all solutions also yields the uniqueness of this solution.

To prove the normalization of the polynomial part of $((X^*)^{-1}BX^{-1})_\infty$, note that $(X^*)^{-1}A^* + AX^{-1} = (X^*)^{-1}BX^{-1}$. Since XA^{-1} bi-proper implies AX^{-1} bi-proper, and since their polynomial parts are each other's inverse, it follows that $(X^*)^{-1}BX^{-1}$ is also bi-proper, with polynomial part $2I$. ■

The following refinement of the above theorem plays also an important role in our analysis of the Newton iteration (3).

Theorem 3.2 *Assume that $B = B^* \in \mathbb{R}^{q\times q}[\xi]$, that $A \in \mathbb{R}^{q\times q}[\xi]$ is Hurwitz, and that $(A^*)^{-1}BA^{-1}$ is bi-proper, normalized to $((A^*)^{-1}BA^{-1})_\infty = 2I$. Assume further that (2) holds. Let X be the unique solution of (8,9), identified in theorem 3.1. Then X is Hurwitz.*

In the proof we use the following lemmas. The first lemma is proven in [9].

Lemma 3.1 *Let $B = B^* \in R^{q\times q}[\xi]$. Then there exists $\Phi = \Phi^* \in R^{q\times q}[\zeta, \eta]$ such that $\partial\Phi = B$ and $\Phi > 0$ iff (2) holds.*

Lemma 3.2 *Let the assumptions of theorem 3.2 be in force, and let $X \in \mathbb{R}^{q\times q}[\xi]$ be the solution to (8,9). Define $\Psi \in \mathbb{R}^{q\times q}[\zeta, \eta]$ by :*

$$\Psi(\zeta, \eta) = \frac{A^T(\zeta)X(\eta) + X^T(\zeta)A(\eta) - \Phi(\zeta, \eta)}{\zeta + \eta} \tag{10}$$

Let $x = S(\frac{d}{dt})w$ be a minimal state map of the system defined by $A(\frac{d}{dt})w = 0$. Then there exists $K = K^T$ such that $\Psi(\zeta, \eta) = S^T(\zeta)KS(\eta)$.

Proof: Let $\Phi(\zeta,\eta) = M^T(\zeta)M(\eta)$. Use (8) to deduce that MA^{-1} is proper. Therefore $(\zeta+\eta)\Psi(\zeta,\eta)A^{-1}(\eta)$, viewed as a matrix of rational functions in η with coefficients in $\mathbb{R}[\zeta]$, is proper. Factor $\Psi(\zeta,\eta)$ as $N^T(\zeta)LN(\eta)$ with $L = L^T \in \mathbb{R}^{\bullet\times\bullet}$. Let $(NA^{-1})_\infty(\eta) = N_k\eta^k + \ldots + N_0$. Then $N^T(\zeta)N_k = 0$. Proceed recursively and obtain $N^T(\zeta)(N_k\eta^k+\ldots+N_0) = 0$. This yields $\Psi(\zeta,\eta) = N^T(\zeta)L\tilde{N}^T(\eta)$ with $\tilde{N}A^{-1}$ strictly proper, and, by a symmetric argument, $\Psi(\zeta,\eta) = \tilde{N}^T(\eta)L\tilde{N}(\eta)$, with $\tilde{N}A^{-1}$ strictly proper. By lemma 2.1 $\tilde{N}(\xi) = FS(\xi)$ for some $F \in \mathbb{R}^{\bullet\times\bullet}$. The result follows. ■

Proof of Theorem 3.2: The proof that X is Hurwitz is based on the Lyapunov theory for high-order differential equations discussed in [9]. Let Ψ be defined by 10, where Φ is as in lemma 3.1. That Ψ is indeed a matrix of polynomials follows from proposition 1.1 and equation (8). Let $\mathfrak{B}_X \in \mathfrak{L}^q$ be the behavior defined by $X(\frac{d}{dt})w = 0$. ¿From the definition of Ψ, it follows that $\overset{\bullet}{\Psi} \overset{\mathfrak{B}_X}{=} -\Phi$. Obviously, therefore, $\overset{\bullet}{\Psi} \overset{\mathfrak{B}_X}{<} 0$. Next, we show that $\Psi \overset{\mathfrak{B}_X}{\geq} 0$. Indeed, for all $w \in C^\infty$ that converge to zero together with all its derivatives, there holds

$$Q_\Psi(w)(0) = \int_0^\infty (Q_\Phi(w) + 2 < A(\frac{d}{dt})w, X(\frac{d}{dt})w >)dt.$$

In particular, along solutions of $A(\frac{d}{dt})w = 0$, we have

$$Q_\Psi(w)(0) = \int_0^\infty Q_\Phi(w)dt$$

Hence $\Psi \overset{\mathfrak{B}_A}{\geq} 0$. By lemma 3.2, $\Psi(\zeta,\eta) = S^T(\zeta)KS(\eta)$, with $K = K^T \in R^{\bullet\times\bullet}$, and S a minimal state map for $\mathfrak{B}_A$, and thus, by lemma 2.1, for $\mathfrak{B}_X$. Since the map that takes $w \in \mathfrak{B}_A$ to $X(\frac{d}{dt})w(0)$ is surjective, $\Psi \overset{\mathfrak{B}_A}{\geq} 0$. This implies that $K = K^T \geq 0$. Hence $\Psi \overset{\mathfrak{B}_X}{\geq} 0$ and $\overset{\bullet}{\Psi} \overset{\mathfrak{B}_X}{<} 0$, and therefore by proposition 2.1, X is indeed Hurwitz. ■

4 Convergence analysis

In this section, we will prove theorem 1.2. The result of theorem 3.2 allows to conclude that if $X_0 \in \mathbb{R}^{q\times q}[\xi]$ is Hurwitz, with $((X_0^*)^{-1}BX_0)_\infty = I$, then (3) generates an unique sequence such that $(X_{k+1}X_k^{-1})_\infty = I$. We will now prove that the limit of these X_k's exists and that this limit is a Hurwitz spectral factorization of B.

In order to do this, we consider the sequence of two-variable polynomial matrices $\Psi_1, \Psi_2, \cdots$ with Ψ_k defined by:

$$\Psi_{k+1}(\zeta,\eta) = \frac{X_{k+1}^T(\zeta)X_k(\eta) + X_k^T(\zeta)X_{k+1}(\eta) - \Phi(\zeta,\eta) - X_k^T(\zeta)X_k(\eta)}{\zeta+\eta} \qquad (11)$$

where Φ is obtained from B using lemma 3.1.

Let $x = S\left(\frac{d}{dt}\right) w$ be a minimal state map for $X_0\left(\frac{d}{dt}\right) w = 0$. Use lemma 2.1 and the fact that $X_k X_0^{-1}$ is bi-proper to conclude that $S\left(\frac{d}{dt}\right) w$ is hence a minimal state map for all the systems $X_k\left(\frac{d}{dt}\right) w = 0$. Hence, by lemma 3.2, each of the $\Psi_k(\zeta, \eta)$'s is of the form $X_0^T(\zeta) K_k X_0(\eta)$ for suitable $K_k = K_k^T \in \mathbb{R}^{\bullet\times\bullet}$. We will show that $\Psi_1 \geq \Psi_2 \geq \cdots \geq \Psi_k \geq \Psi_{k+1} \geq \cdots \geq 0$. In order to see this, observe that $\Delta_k = \Psi_k - \Psi_{k+1}$ satisfies (in the obvious notation)

$$\overset{\bullet}{\Delta}_k = X_k^* X_{k-1} + X_{k-1}^* X_k - X_{k-1}^* X_{k-1} + X_{k+1}^* X_k + X_k^* X_{k+1} - X_k^* X_k$$

Denote the behavior of $X_k\left(\frac{d}{dt}\right) w = 0$ by $\mathfrak{B}_k$. It follows from the above equation that

$$Q_{\Delta_k}(w)(0) \overset{\mathfrak{B}_k}{=} \int_0^{+\infty} \| X_{k-1}(\frac{d}{dt}) w \|^2 \, dt$$

Using the surjectivity of the map $w \in \mathfrak{B}_k \mapsto X(\frac{d}{dt}) w(0)$, this yields $\Delta_k \geq 0$. To show that $\Psi_k \geq 0$, use the same reasoning after observing that

$$Q_{\Psi_{k+1}}(w)(0) \overset{\mathfrak{B}_k}{=} \int_0^{\infty} Q_\Phi(w) dt.$$

The monotone convergence and boundedness from below, imply that the K_k's and hence the Ψ_k's converge.

Note that the X_k's were defined by (3,4), without involving the Ψ_k's. Further, (11) shows how to compute the Ψ_k's using the X_k's. However, it is also possible to deduce the X_k's from the Ψ_k's. Since we know already that the Ψ_k's converge, this will allow us to conclude that the X_k's also converge. In order to obtain this desired relation, write (11) as

$$\begin{aligned}&(\zeta + \eta) S^T(\zeta) K_k S(\eta) = \\ &X_k^T(\zeta) X_{k-1}(\eta) + X_{k-1}^T(\zeta) X_k(\eta) - M^T(\zeta) M(\eta) - X_{k-1}^T(\zeta) X_{k-1}(\eta)\end{aligned}$$

where $M^T(\zeta) M(\eta) = \Phi(\zeta, \eta)$. Now pre-multiply by $(X_0^T(\zeta))^{-1}$ and keep the polynomial part of the left and the right hand side, viewed as rational functions with coefficients in $\mathbb{R}^{\bullet\times\bullet}[\eta]$. Note that, as a consequence of (3), $M X_k^{-1}$ is proper. We obtain $Y_\infty^T K_k S(\eta) = X_k(\eta) - M_\infty^T M(\eta)$, where $Y_\infty = (\xi S(\xi) X_0^{-1}(\xi))_\infty$, and $M_\infty = (M X_0^{-1})_\infty$. This relation and the convergence of the Ψ_k's imply that the X_k's also converge. Let $\tilde{X}$ be this limit. Obviously, $(\tilde{X} X_0^{-1})_\infty = I$. Use this and the fact that each of the X_k's is Hurwitz to conclude that the limit $\tilde{X}$ is also Hurwitz.

That the convergence is quadratic is a general property of convergent Newton iterations.

This ends the proof of theorem 1.2. ■

It is worthwhile to note the more than casual similarity between the above proof and the proof of the convergence of the Newton iteration for computing the solution of the Algebraic Riccati Equation [4].

5 Conclusions

In this paper we have given a proof for the convergence of the Newton iteration for spectral factorization studied before in [5, 3, 6]. The proof involves an interesting interplay between one- and two-variable polynomial matrices.The crucial step of the algorithm is the solution of the matrix polynomial Lyapunov equation at each iteration. In [3] a very nice recursive implementation for giving this solution is given, using the Routh array. We are presently investigating whether this equation can be solved using fast algorithms for polynomial equations, similar to FFT algorithms.

References

[1] P.A. Fuhrmann and J.C. Willems, Factorization indices at infinity for rational matrix functions, *Integral Equations and Operator Theory*, volume 2/3, pages 287-301, 1979.

[2] P.A. Fuhrmann and J.C. Willems, A study of (A,B)-invariant subspaces via polynomial models, *International Journal on Control*, volume 31, pages 467-494, 1980.

[3] J. Ježek and V. Kučera, Efficient algorithm for matrix spectral factorization, *Automatica*, volume 21, pages 663-669, 1985.

[4] D.L. Kleinman, On an iterative technique for Riccati equation computations, *IEEE Transactions on Automatic Control*, volume 13, pages 114-115, 1968.

[5] V. Kučera, *Discrete Linear Control: The Polynomial Equation Approach,* Wiley, 1979.

[6] V. Kučera, *Analysis and Design of Discrete Linear Control Systems,* Prentice Hall, 1991.

[7] J.C. Willems, Paradigms and puzzles in the theory of dynamical systems, *IEEE Transactions on Automatic Control*, volume 36, pages 259-294, 1991.

[8] J.C. Willems and P.A. Fuhrmann, Stability theory of high order equations, *Linear Algebra and its Applications*, volume 167, pages 131-150, 1992.

[9] J.C. Willems and H.L. Trentelman, On quadratic differential forms, *SIAM Journal on Control and Optimization*, submitted.