International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713393989

Simulation of linear systems and factorization of matrix polynomials†

Paul A. Fuhrmann [a]
[a] Department of Mathematics, Ben Gurion University of the Negev. Beer Sheva. Israel.

To cite this Article: Fuhrmann, Paul A. , 'Simulation of linear systems and factorization of matrix polynomials†', International Journal of Control, 28:5, 689 - 705
To link to this article: DOI: 10.1080/00207177808922490
URL: http://dx.doi.org/10.1080/00207177808922490

INT. J. CONTROL, 1978, VOL. 28, NO. 5, 689–705

Simulation of linear systems and factorization of matrix polynomials†

PAUL A. FUHRMANN‡

Some submodules of the module of formal power series are studied and a lifting theorem derived. The results are applied to the analysis of simulation of linear systems and this in turn to study a problem of factorization of matrix polynomials.

1. Introduction

This paper continues the investigations (Fuhrmann 1976 a, b) into various aspects of linear system theory. As in the other papers, the fundamental idea is the notion of a model of a linear transformation which is similar to the original but in some ways easier to handle.

It is quite well known that given a linear transformation acting as a linear space X over the field F, then an $F[\lambda]$-module structure can be induced on X by letting $p \,.\, x = p(A)x$ for all $p \in F(\lambda)$ and $x \in X$. Of course, the action of A is identical to that of the polynomial $\chi(\lambda) = \lambda$ and the module is a finitely generated torsion module. The model approach reverses this approach. We start with an $F[\lambda]$-module X and define a linear transformation A in X by $Ax = \lambda \,.\, x$. An interesting theory might arise if our choice of module X is well made. As our interest here is strictly in finite dimensional phenomena, then we should restrict ourselves to finitely generated torsion modules over $F[\lambda]$.

From a system theoretic point of view there are two natural choices for our modules. It has been recognized by Kalman (1969) and by Kalman *et al.* (1969) that given a restricted input/output map $f: U[\lambda] \to \lambda^{-1}Y[[\lambda^{-1}]]$, then natural choices for a state-space realization would be $U[\lambda]/\ker f$ and range f. The development of Fuhrmann (1976 a, b) used the first type of representation, whereas here we investigate the dual representation using submodules of $\lambda^{-1}Y[[\lambda^{-1}]]$ as well as the relations between the two types of models. While $F[\lambda]$-module homomorphisms in $U[\lambda]$ are easy to describe, the module $\lambda^{-1}Y[[\lambda^{-1}]]$ is too big for a simple description of all $F[\lambda]$ homomorphism. However, we obtain a certain lifting theorem, Theorem 2.4, which is sufficient for the application to system theory.

In § 3 we introduce a partial order into the sets of rational transfer functions and restricted input/output maps. This is the problem of when one canonical system can be simulated by another canonical system. For further background and results on this problem one should refer to Kalman (1969). Finally, in the last section, we apply the results of simulation to the problem of factoring

---

Received 30 August 1977.

† This research was done while the author was visiting the Electronic Systems Laboratory at the Massachusetts Institute of Technology in Cambridge, Massachusetts 02139. It was supported by the Department of the Army under contract grant DAAG29–76–0008.

‡ Department of Mathematics, Ben Gurion University of the Negev, Beer Sheva, Israel.

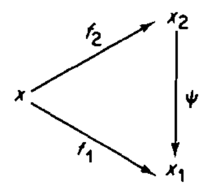
a monic polynomial matrix into monic factors. This gives a system theoretic approach to some results of Gohberg *et al.* (1978), which has the advantage that it holds over every field F. For further results and references on factorizations of matrix polynomials one should consult Langer (1976).

2. On submodules of $\lambda^{-1}Y[[\lambda^{-1}]]$

We begin by stating two elementary lemmas on module homomorphisms which will be used repeatedly through the rest of this work.

Lemma 2.1

Let X, X_1 and X_2 be modules over the ring R and let $f_1: X \to X_1$ and $f_2: X \to X_2$ be R homomorphisms, of which f_2 is assumed to be subjective. Then there exists a uniquely determined R homomorphism $\psi: X_2 \to X_1$ which makes the diagram



(2.1)

commutative if and only if

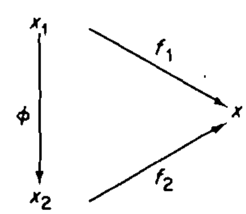
$$\ker f_2 \subset \ker f_1 \tag{2.2}$$

Moreover, ψ is injective if and only if

$$\ker f_2 = \ker f_1 \tag{2.3}$$

Lemma 2.2

Let X, X_1 and X_2 be modules over the ring R and let $f_1: X_1 \to X$ and $f_2: X_2 \to X$ be R homomorphisms of which f_2 is assumed injective. Then there exists a uniquely determined R homomorphism $\phi: X_1 \to X_2$ which makes the diagram



(2.4)

commutative if and only if

$$\text{range } f_1 \subset \text{range } f_2 \tag{2.5}$$

Moreover, ϕ is surjective if and only if

$$\text{range } f_1 = \text{range } f_2 \tag{2.6}$$

We recall now a few notions introduced by Fuhrmann (1976 a). Let F be a field and $F[\lambda]$ the ring of polynomials over F. Let U and Y be finite dimensional vector spaces over F. We denote by $(U, Y)_F$ the space of all F-linear transformations from U to Y. By $U[\lambda]$ we denote the $F[\lambda]$ module of all polynomials with coefficients in U. $U((\lambda^{-1}))$ denotes the set of all truncated Laurent series with coefficients in U, i.e. the space of all series of the form $\sum_{k \leq n} u_{-n}\lambda^{-n}$, where k is any integer. Certainly $U[\lambda]$ is an $F[\lambda]$ submodule of $U((\lambda^{-1}))$ and so we can form the quotient module $U((\lambda^{-1}))/U[\lambda]$. This later module can be identified with $\lambda^{-1}U[[\lambda^{-1}]]$ the set of all formal power series in λ^{-1} with vanishing constant term. Thus the sequence of $F[\lambda]$ homomorphisms $0 \to U[\lambda] \xrightarrow{j} U((\lambda^{-1})) \xrightarrow{\pi_-} \lambda^{-1}U[[\lambda^{-1}]] \to 0$ is a short exact sequence. Here j is the natural injection of $U[\lambda]$ into $U((\lambda^{-1}))$, whereas π_- is the canonical projection of $U((\lambda^{-1}))$ onto $\lambda^{-1}U[[\lambda^{-1}]]$ defined by

$$\pi_- \left(\sum_{k \leq n} u_{-n}\lambda^{-n} \right) = \sum_{1 \leq n} u_{-n}\lambda^{-n}$$

The complementary projection to π_-, i.e. $I - \pi_-$, is denoted by π_+. Clearly

$$\pi_+ \sum_{k \leq n} u_{-n}\lambda^{-n} = \sum_{k \leq n \leq 0} u_{-n}\lambda^{-n}$$

Any $F[\lambda]$ submodule M of $U[\lambda]$ has the form $M = DU[\lambda]$ for some $D \in (U, U)_F[\lambda]$. The quotient module $U[\lambda]/DU[\lambda]$ is a finitely generated torsion module if and only if D is non-singular over $F[\lambda]$ or equivalently if and only if det D is not the zero polynomial. For a non-singular $D \in (U, U)_F[\lambda]$ we define the map $\pi_D : U[\lambda] \to U[\lambda]$ by

$$\pi_D p = D\pi_- D^{-1} p \tag{2.7}$$

for all $p \in U[\lambda]$. Denote K_D the range of the projection π_D and induce an $F[\lambda]$-module structure in K_D by letting a polynomial $p \in F[\lambda]$ act on $u \in K_D$ by $u \to \pi_D(pu)$. In particular, if we let χ denote the polynomial $\chi(\lambda) = \lambda$, then the operator $S(D) : K_D \to K_D$ defined by

$$S(D)u = \pi_D(\chi u), \quad u \in K_D \tag{2.8}$$

figured prominently in the model approach to linear systems (Fuhrmann 1976 a).

It is our purpose now to derive a set of analogous results where the setting is $\lambda^{-1}Y[[\lambda^{-1}]]$ rather than $U[\lambda]$.

Thus, let $D \in (Y, Y)_F[\lambda]$ be non-singular. We define a map

$$\pi^D : \lambda^{-1}[[\lambda^{-1}]] \to \lambda^{-1}Y[[\lambda^{-1}]]$$

by

$$\pi_- D^{-1} \pi_+ D y \tag{2.9}$$

Obviously π^D is a projection operator in $\lambda^{-1}Y[[\lambda^{-1}]]$ but it is not an $F[\lambda]$ homomorphism. However, $L_D = \text{range } \pi^D$ is a submodule of $\lambda^{-1}Y[[\lambda^{-1}]]$, as can easily be checked. We have the following counterpart of Theorem 3.1 in Fuhrmann (1976 a).

Theorem 2.3

A subset of M of $\lambda^{-1}Y[[\lambda^{-1}]]$ is a finitely generated torsion submodule if and only if

$$M = L_D = \text{range } \pi^D \tag{2.10}$$

for some non-singular $D \in (Y, Y)_F[\lambda]$.

Proof

Let $M = L_D$ for some non-singular D. By Cramer's rule $DE = (\det D) \, . \, I$, where E is the co-factor matrix of D. This implies that d annihilates all of M, i.e. M is a torsion submodule of $\lambda^{-1}Y[[\lambda^{-1}]]$. As M is a finite dimensional over F it is clearly finitely generated over $F[\lambda]$.

Conversely, assume M is a finitely generated torsion submodule of $\lambda^{-1}Y[[\lambda^{-1}]]$. There exists a polynomial $p \in F[\lambda]$ which annihilates all of M. Consider next the set J defined by

$$J = \{A \in (Y, Y)_F[\lambda] \mid \pi(Ay) = 0 \quad \text{for all } y \in M\}$$

then clearly J is a left ideal in $(Y, Y)_F[\lambda]$ and so, by Theorem 3.1 of Fuhrmann (1976 a), has the form $J = (Y, Y)_F[\lambda]D$ for some $D \in (Y, Y)_F[\lambda]$. Since $p \, . \, I \in J$, it follows that D is necessarily non-singular.

Define now a map ρ_D : range $\pi^D \to$ range π_D by

$$\rho_D y = Dy \tag{2.11}$$

Since for every polynomial

$$\begin{aligned} p \in F[\lambda] \rho_D(p \, . \, y) &= \rho_D \pi_-(py) = D\pi_-(py) = D\pi_- D^{-1} p Dy \\ &= \pi_D p(Dy) = \pi_D p(\rho_D y) = p \, . \, p_D y \end{aligned}$$

it follows that ρ_D is an $F[\lambda]$ homomorphism that maps M into a submodule of K_D. But submodules of K_D correspond in a bijective way to left factors of D, hence necessarily $M = \text{range } \pi^D$.

Given now two finitely generated torsion submodules M and M_1 of $\lambda^{-1}Y[[\lambda^{-1}]]$ and $\lambda^{-1}Y_1[[\lambda^{-1}]]$ respectively, we want to characterize the set of all $F[\lambda]$ homomorphisms from M into M_1. The situation is analogous to that of Theorem 4.5 of Fuhrmann (1976 a) and so is the result.

Theorem 2.4

Let D and D_1 be non-singular elements of $(Y, Y)_G[\lambda]$ and $(Y_1, Y_1)_F[\lambda]$ respectively. A map $\psi_0 : L_D \to L_{D_1}$ is an $F[\lambda]$ homomorphism if and only if there exist Ψ and Ψ_1 in $(Y, Y_1)_F[\lambda]$ satisfying

$$\Psi D = D_1 \Psi_1 \tag{2.12}$$

and for which

$$\psi_0(y) = \pi_-(\Psi_1 y) \tag{2.13}$$

Proof

Assume there exist Ψ and Ψ_1 satisfying (2.12) and let ψ_0 be defined by (2.13). Let $y \in L_D$, so Dy is in $Y[\lambda]$. Now $\psi_0(y) = \pi(\Psi_1 y)$ and

$$D_1 \psi_0(y) = D_1 \pi_-(\Psi_1 y) = D_1 \pi_- D_1^{-1} \, D_1 \Psi_1 y = \pi_{D_1}(D_1 \Psi_1 y) = \pi_{D_1}(\Psi D y)$$

and clearly $\pi_{D_1}(\Psi Dy) \in Y_1[\lambda]$. Therefore $\psi_0(y) \in L_{D_1}$. To show that ψ_0 is an $F[\lambda]$ homomorphism, let p be any polynomial in $F[\lambda]$. Then

$$\psi_0(\pi_-(py)) = \pi_-(\Psi_1 \pi_-(py)) = \pi_-(\Psi_1 py) \sim \pi_-(p\Psi_1 y) = \pi_-(p\pi_-(\Psi_1 y)) = \pi_-(p\psi_0 y))$$

Conversely, let $\psi_0 : L_D \to L_{D_1}$ be an $F[\lambda]$ homomorphism. Since $\rho_D : L_D \to K_D$ and $\rho_{D_1} \to K_{D_1}$ defined by (2.11) are $F[\lambda]$ homomorphisms, then $\psi : K_D \to K_{D_1}$ defined by $\psi = \rho_{D_1}\psi_0\rho_D^{-1}$ is also an $F[\lambda]$ homomorphism. We can now apply Theorem 4.5 of Fuhrmann (1976 a) which characterizes these homomorphisms. Thus there exist Ψ and Ψ_1 in $(Y, Y_1)_F[\lambda]$ satisfying (2.12) and for which ψ is given by

$$\psi(u) = \pi_{D_1}(\Psi u)$$

As $\psi_0 = \rho_{D_1}^{-1}\,\psi\rho_D$ we obtain for $y \in L_D$

$$\begin{aligned}\psi_0(y) &= \rho_D^{-1}\,\psi\rho_D y = D_1^{-1}\,\pi_{D_1}\Psi Dy = D_1^{-1}\,D_1\pi_- D_1^{-1}\,\Psi Dy \\ &= \pi_-(D_1^{-1}\,\Psi D)y = \pi_-(\Psi_1 y)\end{aligned}$$

by virtue of relation (2.12).

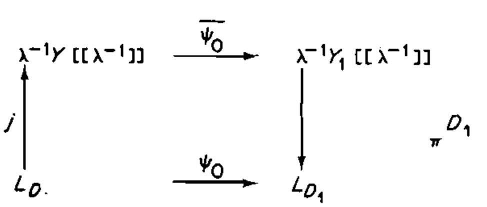
Now the map $y \to \pi(\Psi_1 y)$ is clearly an $F[\lambda]$ homomorphism of $\lambda^{-1}Y[[\lambda^{-1}]]$ into $\lambda^{-1}Y_1[[\lambda^{-1}]]$ and so, simply by (2.13), we obtain a lifting theorem in this setting.

Theorem 2.5

Let $M = L_D$ and $M_1 = L_{D_1}$ be finitely generated torsion submodules of $\lambda^{-1}Y[[\lambda^{-1}]]$ and $\lambda^{-1}Y_1[[\lambda^{-1}]]$ respectively and let $\psi_0 : M \to M_1$ be an $F[\lambda]$ homomorphism. Then there exists an $F[\lambda]$ homomorphism

$$\overline{\psi_0} : \ \lambda^{-1}Y[[\lambda^{-1}]] \to \lambda^{-1}Y_1[[\lambda^{-1}]]$$

such that the diagram



(2.14)

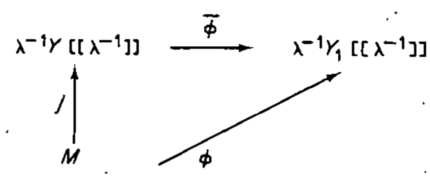
is commutative.

Easily obtained is also the dual version of Lemma 4.6 of Fuhrmann (1976 a).

Corollary 2.6

Let M be a finitely generated torsion submodule of $\lambda^{-1}Y[[\lambda^{-1}]]$ and let $\phi : M \to \lambda^{-1}Y_1[[\lambda^{-1}]]$ be an $F[\lambda]$ homomorphism. Then there exists an

$F[\lambda]$ homomorphism $\bar{\phi}: \lambda^{-1}Y[[\lambda^{-1}]] \to \lambda^{-1}Y_1[[\lambda^{-1}]]$ which makes the following diagram



(2.15)

commutative.

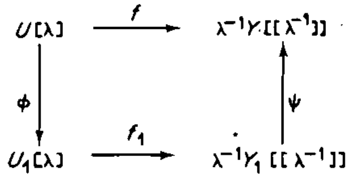
3. Simulation of linear systems

In this section we adopt the approach of Kalman (1969) and Kalman *et al.* (1969) to the description of linear systems. For the relation of this to co-prime factorizations of transfer functions one should consult Fuhrmann (1976 a, b) and Hautus and Heymann (1976).

Consider two F-vector spaces U and Y. A restricted input/output map f is an $F[\lambda]$ homomorphism $f: U[\lambda] \to \lambda^{-1}Y[[\lambda^{-1}]]$. A realization of f is a factorization $f = OR$, where X is an $F[\lambda]$ module and $R: U[\lambda] \to X$ and $O: X \to \lambda^{-1}Y[[\lambda^{-1}]]$ are $F[\lambda]$ homomorphisms. A realization is finite dimensional if X is finite dimensional as a vector space over F which is equivalent to X being a finitely generated torsion module. The realization is reachable if R is surjective, observable if O is injective and canonical if it is both reachable and observable. An input/output map f has a finite dimensional realization if and only if there exists a rational function $T \in \lambda^{-1}(U, Y)_F[[\lambda^{-1}]]$ for which

$$f(u) = \pi_{-}(Tu) \tag{3.1}$$

holds with π the canonical projection of $Y((\lambda^{-1}))$ onto $\lambda^{-1}Y[[\lambda^{-1}]]$. We introduce a partial order in the set of restricted input/output maps and the set of transfer function by the following definitions. A function written $f|f_1$, $f: U[\lambda] \to \lambda^{-1}Y[[\lambda^{-1}]]$ is *simulated* by a function $f_1: U_1[\lambda] \to \lambda^{-1}Y_1[[\lambda^{-1}]]$ if there exist $F[\lambda]$ homomorphisms ϕ and ψ which make the following diagram commutative :



(3.2)

Let T and T_1 be elements of $\lambda^{-1}(U_1 Y)_F[[\lambda^{-1}]]$ and $\lambda^{-1}(U_1, Y_1)_F[[\lambda^{-1}]]$ respectively. Then we say that T *divides* T_1, written $T|T_1$, if there exist Φ, Ψ and Π in $(U, U_1)_F[\lambda]$, $(Y_1, Y)_F[\lambda]$ and $(U, Y)_F[\lambda]$ respectively for which

$$T = \Psi T_1 \Phi + \Pi \tag{3.3}$$

holds. It is clear that both relations are reflexive and transitive.

The first result relates the notion of simulation with that of divisibility.

Theorem 3.1

Let f and f_1 be two restricted input/output maps having finite dimensional realizations and let T and T_1 be their corresponding transfer functions. Then f is simulated by f_1 if and only if T divides T_1.

Proof

Assume $T|T_1$. Thus there exist Φ and Ψ such that (2.12) holds. Define $F[\lambda]$ homomorphisms $\phi: U[\lambda]\to U_1[\lambda]$ and $\psi: \lambda^{-1}Y_1[[\lambda^{-1}]]\to\lambda^{-1}Y[[\lambda^{-1}]]$ by $\phi(u)=\Phi u$ and $\psi(y)=\pi(\Psi y)$. Then for $u\in U[\lambda]$ we have

$$f(u)=\pi_-(Tu)=\pi_-((\Psi T_1\Phi+\Pi)u)=\pi_-(\Psi\pi_-(T_1\Phi u))=\psi f_1\phi(u) \quad \text{or} \quad f|f_1$$

Conversely, assume $f|f_1$, that is $f=\psi f_1\phi$. Now every $F[\lambda]$ homomorphism $\phi: U[\lambda]\to U_1[\lambda]$ is of the form $\phi(u)=\Phi u$ for some $\Phi\in(U, U_1)_F[\lambda]$. As for ψ, we restrict it to the range of f_1 which is a finitely generated torsion submodule of $\lambda^{-1}Y_1[[\lambda^{-1}]]$. We now apply Theorem 2.5 to obtain an extension $\bar{\psi}$ to all of $\lambda^{-1}Y_1[[\lambda^{-1}]]$ which has the form $\bar{\psi}(y)=\pi(\Psi y)$ for some $\Psi\in(Y_1, Y)_F[\lambda]$. Clearly $f=\psi f_1\phi=\bar{\psi}f_1\phi$ and so for $u\in U[\lambda]$ we have

$$f(u)=\pi(Tu)=\pi(\Psi\pi(T_1\Phi u))=\pi(\Psi T_1\Phi u)$$

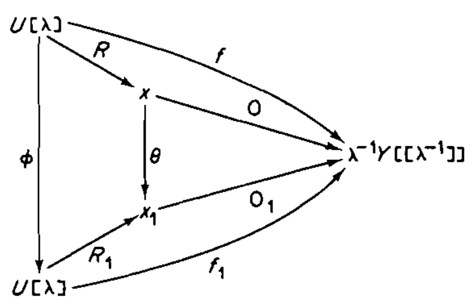
and this implies (3.3).

Remark

Condition (3.3) cannot be replaced by the stronger condition $T=\Psi T_1\Phi$ as claimed in Kalman (1969). The following simple counter example has been furnished by K. Overtoom. Let $T(\lambda)=-\lambda^{-2}$ and $T_1(\lambda)=\lambda^{-1}-\lambda^{-2}$. With $\Psi(\lambda)=\lambda+1$ we have $f=\psi f_1$ but there exist no Φ_1, Ψ_1 such that $T=\Psi_1T_1\Phi_1$.

Theorem 3.2

Let $f: U[\lambda]\to\lambda^{-1}Y[[\lambda^{-1}]]$ and $f_1: U_1[\lambda]\to\lambda^{-1}Y[[\lambda^{-1}]]$ be two restricted input/output maps having finite dimensional canonical factorizations $f: OR$ and $f_1=O_1R_1$ through the $F[\lambda]$ modules X and X_1 respectively. Then there exists an injective $F[\lambda]$ homomorphism $\Theta: X\to X_1$ which makes the diagram



(3.4)

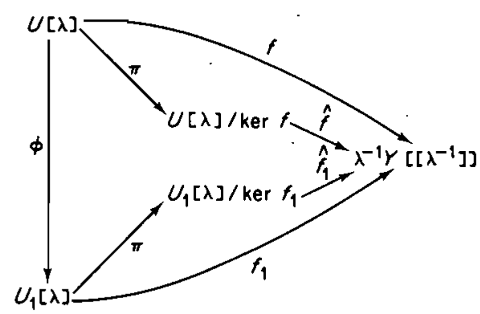
commutative if and only if

$$\text{range } f\subset\text{range } f_1 \tag{3.5}$$

Proof

Assume such a homomorphism Θ exists. Since the factorizations of f and f_1 are canonical we have range $O=$ range f and range $O_1=$ range f_1. Since $O=O_1\Theta$ it follows that range $O\subset$ range O_1 and so (3.5) follows.

Conversely, assume (3.5) holds and consider the homomorphisms $\hat{f}$ and $\hat{f}_1$ induced by f and f_1 in $U[\lambda]/\ker f$ and $U_1[\lambda]/\ker f_1$ respectively. Clearly $\hat{f}$ and $\hat{f}_1$ are injective and range $\hat{f} \subset$ range $\hat{f}_1$. By Lemma 2.2 there exists an injective homomorphism $\hat{\phi}: U[\lambda]/\ker f \to U_1[\lambda]/\ker f_1$ for which $\hat{f} = \hat{f}_1 \hat{\phi}$. Applying Theorem 4.5 in Fuhrmann (1976) we can lift $\hat{\phi}$ to an $F[\lambda]$ homomorphism $\phi: U[\lambda] \to U_1[\lambda]$ which makes the diagram



(3.6)

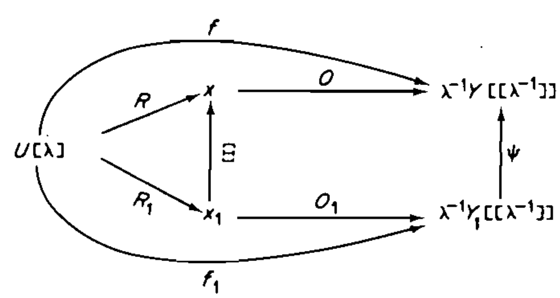
commutative.

By Lemma 2.1 there exists a uniquely determined $F[\lambda]$ homomorphism $\Theta: X \to X_1$ which satisfies $R_1\phi = \Theta R$. This implies $O_1\Theta R = O_1 R_1 \phi = f_1\phi = f = OR$ as R is surjective we obtain $O_1\Theta = 0$. Since O is injective it follows that Θ is injective too.

The dual result to the previous theorem is the following.

Theorem 3.3

Let $f: U[\lambda] \to \lambda^{-1}Y[[\lambda^{-1}]]$ and $f_1: U[\lambda] \to \lambda^{-1}Y_1[[\lambda^{-1}]]$ be two restricted input/output maps having finite dimensional canonical factorizations $f = OR$ and $f_1 = O_1 R_1$ through the $F[\lambda]$ modules X and X_1 respectively. Then there exists a surjective $F[\lambda]$ homomorphism $\Xi: X_1 \to X$ which makes the diagram



(3.7)

commutative if and only if

$$\ker f_1 \subset \ker f \tag{3.8}$$

Proof

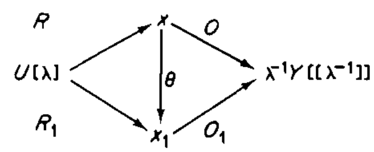
Suppose such a homomorphism Ξ exists. Since $\Xi R_1 = R$ it follows that $\ker R_1 \subset \ker R$. Since the two factorizations of f and f_1 are canonical we have $\ker f = \ker R$ and $\ker f_1 = \ker R_1$. Thus (3.8) follows.

Conversely, assume (3.8) holds. Range f and range f_1 are finitely generated torsion submodules of $\lambda^{-1}Y[[\lambda^{-1}]]$ and $\lambda^{-1}Y_1[[\lambda^{-1}]]$ respectively. By Lemma 2.1 there exists an $F[\lambda]$-module homomorphism $\hat{\psi}$: range $f_1 \to$ range f which satisfies $\psi\hat{f}_1 = f$. By Theorem 2.5, $\hat{\psi}$ can be lifted to an $F[\lambda]$ homomorphism $\psi : \lambda^{-1}Y_1[[\lambda^{-1}]] \to \lambda^{-1}Y[[\lambda^{-1}]]$ which still satisfies $\psi f_1 = f$. From this we obtain range $\psi O_1 \subset$ range O, and as O is injective it follows from Lemma 2.2 that there exists an $F[\lambda]$ homomorphism $\Xi : X_1 \to X$ for which $O\Xi = \psi O_1$. Finally $O\Xi R_1 = \psi O_1 R_1 = \psi f_1 = f = OR$. As O is injective it follows that $\Xi R_1 = R$. Thus diagram (3.7) is commutative and as R is surjective Ξ must be surjective too.

As a corollary to Theorems 3.2 and 3.3 we obtain the state-space isomorphism theorem (Kalman *et al.* 1969).

Theorem 3.4

Let $f = OR$ and $f = O_1R_1$ be two finite dimensional canonical factorizations of a restricted input/output map f through the $F[\lambda]$ modules X and X_1 respectively. Then there exists an $F[\lambda]$ isomorphism $\Theta : X \to X_1$ which makes the diagram



(3.8)

commutative.

Proof

By Theorems 3.2 and 3.3 there exist an injective homomorphism $\Theta : X \to X_1$ and a surjective homomorphism $\Xi : X_1 \to X$ which satisfy $\Theta R = R_1$, $O_1\Theta = O$, $\Xi R_1 = R$ and $O\Xi = O_1$. It follows that $\Xi\Theta R = \Xi R_1 = R$ and by the surjectivity of R that $\Xi\Theta = I_X$.

Similarly, $\Theta\Xi R_1 = \Theta R = R_1$ and so $\Theta\Xi = I_{X_1}$. These two relations show that Θ and Ξ are isomorphisms.

4. Factorizations of matrix polynomials

This section is devoted to an application of the results on simulation of linear systems to the problem of factoring matrix polynomials. They give system theoretic proofs to results first obtained by Gohberg *et al.* (1978).

Specifically, let X be an n-dimensional vector space over the field F. A matrix polynomial is, by a slight abuse of language, an element of $(X_1X)_F[\lambda]$. If L is a matrix polynomial, then $L(\lambda) = \sum_{j=0}^{l} A_j\lambda^j$. It is of degree l if $A_l \neq 0$ and it is monic of degree l if $A_l = I$. A polynomial L is a right factor, or right divisor, of L_1 if there exists a polynomial Φ for which $L_1 = \Phi L$ and a left factor of L_1 if there exists a polynomial Ψ for which $L_1 = L\Psi$.

Assume L is a monic matrix polynomial of degree l, that is

$$L(\lambda) = I\lambda^l + L_{l-1}\lambda^{l-1} + \dots + L_0 \tag{4.1}$$

Then L^{-1} is a proper rational function having the series representation $L(\lambda)=\sum_{j=l}^{\infty} A_j\lambda^{-j}$ with $A_l=I$.

If we consider the projection π_L acting in $X[\lambda]$, then clearly

$$\text{range}\ \pi_L=\{\xi_0+\ \dots\ +\xi_{l-1}\lambda^{l-1}\mid \xi_j\in X\}$$

If we identify the polynomial $\xi_0+\ \dots\ +\xi_{l-1}\lambda^{l-1}$ with the vector $\begin{bmatrix}\xi_0\\ \vdots\\ \xi_{l-1}\end{bmatrix}$ in X^l, the the action of $\chi(\lambda)=\lambda$ has the matrix representation

$$\begin{bmatrix}0 & & -L_0\\ I & & \vdots\\ & \ddots & \vdots\\ & I & -L_{l-1}\end{bmatrix} \tag{4.2}$$

Thus the state space for a canonical realization of L^{-1} is ln-dimensional and may be identified with X^l. Now every canonical factorization of the input/output map that is induced by L^{-1} can be described by a triple (A, B, C) with $A: X^l\to X^l$, $B: X\to X^l$ and $C: X^l\to XF$ linear map. A similar realization holds for a monic L_1 of degree k in the state space X^k.

Before proceeding with the problem of polynomial factorization we prove a useful representation for matrix polynomials. This result is due to Gohberg *et al.* (1978) who give a slightly different exposition.

Theorem 4.1

(*a*) Let $L(\lambda)=I\lambda^k+D_{k-1}\lambda^{k-1}+\ \dots\ +D_0\in(X_1X)_F[\lambda]$ be monic of degree k and let (A, B, C) be a canonical realization of L^{-1}. Then

$$\text{rank}\ (B\ \ AB\ \dots\ A^{k+1}B)=nk \tag{4.3}$$

and

$$\text{rank}\begin{bmatrix}C\\ CA\\ \vdots\\ CA^{k-1}\end{bmatrix}=nk \tag{4.4}$$

where $n=\dim X$, and L has the representation

$$L(\lambda)=I\lambda^k-CA^k(U_0+U_1\lambda+\ \dots\ +U_{k-1}\lambda^{k-1}) \tag{4.5}$$

where $U_j\in(X_1X^k)_F$ satisfies

$$\begin{bmatrix}C\\ CA\\ \vdots\\ CA^{k-1}\end{bmatrix}^{-1}=(U_0\ \dots\ U_{k-1}) \tag{4.6}$$

(*b*) Let (A, C) with $A\in(X^k, X^k)_F$ and $C\in(X^k, X)_F$ satisfy (4.4). Then there exists a uniquely determined map $B\in(X, X^k)_F$ and a monic polynomial $L\in(X, X)_F[\lambda]$ such that (A, B, C) is a canonical realization of L^{-1}. L itself is given by (4.5).

Proof

(*a*) Let (A, B, C) be a canonical realization of L^{-1}. Since L is monic of degree k, L^{-1} has the expansion

$$L(\lambda)^{-1} = \sum_{j=k}^{\infty} A_j \lambda^{-j} \quad \text{with } A_k = I \tag{4.7}$$

As $A_j = CA^{j-1}B$ for all $j \geqslant 1$ it follows that

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} (B \;\; AB \; \ldots \; A^{k-1}B) = \begin{bmatrix} 0 & & I \\ & \cdot^{\cdot^{\cdot}} & A_{k+1} \\ \text{SC} & \cdot^{\cdot^{\cdot}} & \\ I \;\; A_{k+1} & & A_{2k-1} \end{bmatrix} \tag{4.8}$$

which implies (4.3) and (4.4). To get the representation (4.5) we use $L(\lambda)L(\lambda)^{-1} = I$. By equating coefficients we easily obtain

$$(D_0 \ldots D_{k-1}) \begin{bmatrix} CB & CAB & CA^{k-1}B \\ CAB & \cdot^{\cdot^{\cdot}} & \\ \vdots & & \\ CA^{k-1}B & & CA^{2k-2}B \end{bmatrix} = -(CA^kB \ldots CA^{2k-1}B) \tag{4.9}$$

The coefficient matrix in this system of equation is the Hankel matrix which has the factorization (4.8). Thus, since

$$-(CA^kB \ldots CA^{2k-1}B) = -CA^k(B \;\; AB \; \ldots \; A^{k-1}B)$$

we obtain

$$(D_0 \ldots D_{k-1}) = -CA^k \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}^{-1} \tag{4.10}$$

which implies the result.

(*b*) Assume the pair (A, C) satisfies (4.4). Define the map $U_j \in (X, X^k)_F$ by (4.6) and the monic polynomial L by (4.5). Also let $B \in (X, X^k)_F$ be defined by $B = U_{k-1}$. Then (4.6) implies

$$\begin{bmatrix} CU_0 & \ldots & CU_{k-1} \\ CAU_0 & & CAU_{k-1} \\ \vdots & & \vdots \\ CA^{k-1}U_0 & & CA^{k-1}U_{k-1} \end{bmatrix} = \begin{bmatrix} I & & 0 \\ & \ddots & \\ 0 & & I \end{bmatrix}$$

and hence

$$CA^jB = \begin{cases} 0 & 0 \leqslant j < k-1 \\ I & j = k-1 \end{cases} \tag{4.11}$$

If we define Aj by

$$A_j = CA^{j-1}B \quad \text{for all } j \geqslant 1 \tag{4.12}$$

Then (4.8) follows from (4.11). In particular the rank condition (4.3) is satisfied. Thus (A, B, C) is a canonical triple. Now define the map $D_j \in (X, X)_F$ by

$$(D_0 \ldots D_{k-1}) = -CA^{k-1}(U_0 \ldots U_{k-1}) \tag{4.13}$$

and the monic polynomial L by (4.5). It remains to show that

$$L(\lambda) \sum_{j=k}^{\infty} A_j \lambda^{-j} = I \tag{4.14}$$

From (4.8) we obtain

$$(U_0 \ldots U_{k-1}) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}^{-1} = (B \ AB \ldots A^{k-1}B) \begin{bmatrix} 0 & & I \\ & \cdot^{\cdot^{\cdot}} & A_{k+1} \\ & & \vdots \\ I & A_{k+1} \ldots & A_{2k-1} \end{bmatrix}^{-1}$$

and so

$$-CA^k(U_0 \ldots u_{k-1}) = -(CA^kB \ CA^{k+1}B \ldots CA^{2k-1}B) \begin{bmatrix} 0 & & I \\ & \cdot^{\cdot^{\cdot}} & A_{k+1} \\ & & \vdots \\ I & A_{k+1} \ldots & A_{2k-1} \end{bmatrix}^{-1}$$

Next we use the definition (4.13) of the D_i to get

$$(D_0 \ldots D_{k-1}) \begin{bmatrix} 0 & & I \\ & \cdot^{\cdot^{\cdot}} & A_{k+1} \\ & & \vdots \\ I & A_{k+1} \ldots & A_{2k-1} \end{bmatrix} = -(A_k \ldots A_{2k})$$

This last equality can be rewritten as

$$(D_0 \ldots D_{k-1}I) \begin{bmatrix} 0 & I \\ \cdot^{\cdot^{\cdot}} & A_{k+1} \\ I & \cdot^{\cdot^{\cdot}} \\ A_{k+1} & A_{2k} \end{bmatrix} = (0 \ldots 0)$$

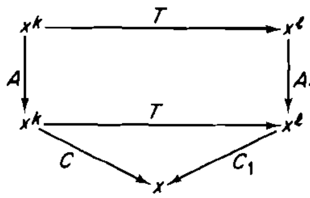
and as the Hankel matrix has maximal rank we actually have

$$(D_0 \ldots K_{k-1}I) \begin{bmatrix} & & I & A_{k+1} & A_{k+2} \cdots \\ & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \\ I & A_{k+1} & \cdot^{\cdot^{\cdot}} & & \\ A_{k+1} & A_{k+2} & & & \end{bmatrix} = (0 \ldots 0 \ 0)$$

and this condition is equivalent to (4.14).

Theorem 4.2

Let L_1 and L be monic operator polynomials of degrees R and k respectively. Let (A_1, B_1, C_1) be a canonical realization of L_1^{-1} in X^l. Then L is a monic right divisor of L_1 if and only if there exists an injective map $T : X^R \to X^k$ whose range is an A_1-invariant subspace such that the maps A and C defined by the commutative diagram

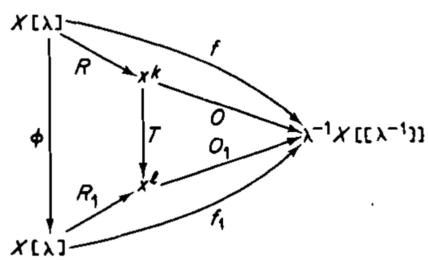


(4.15)

determined a canonical realization of L^{-1}.

Proof

Assume $L_1 = \Phi L$ is a factorization of L_1, then it follows that $L_1^{-1}\,\Phi = L^{-1}$. Let $f : X[\lambda] \to \lambda^{-1}X[[\lambda^{-1}]]$ and $f_1 : X[\lambda] \to \lambda^{-1}X[[\lambda^{-1}]]$ be the restricted input/output maps that correspond to the transfer functions L^{-1} and L_1^{-1} respectively. As we observed, L^{-1} and L_1^{-1} have canonical realizations in X^k and X^l respectively given by (A, B, C) and (A_1, B_1, C_1) respectively. Of course X^k and X^l are $F[\lambda]$ modules where the module structure is given by $p \,.\, \xi = p(A)\xi$ in X^k and $p \,.\, \xi = p(A_1)\xi$ in X^l. An $F[\lambda]$ homomorphism T from X^k into X^l is an F linear map which satisfies $A_1 T = TA$. This last relation also implies that range T is an A_1-invariant subspace. Since $L_1^{-1}\Phi = L$ it follows that $f_1\phi = f$, where $\phi : X[\lambda] \to X[\lambda]$ is given by $\phi(x) = \Phi x$. Thus the system f is simulated by f_1. Apply now Theorem 3.2 to infer the existence of an injective $F[\lambda]$ homomorphism $T : X^k \to X^l$ which makes the diagram



(4.16)

commutative. Obviously T is linear and $TA = A_1 T$. Since $O_1 T = 0$ it follows that for each $x \in X^k$

$$\sum_{j=0}^{\infty} \frac{C_1 A_1^{\,j}\, Tx}{\lambda^{j+1}} = \sum_{j=0}^{\infty} \frac{CA^j x}{\lambda^{j+1}}$$

or $C_1 A_1^{\,j}\, T = CA^j$ for all $j \geqslant 0$. In particular this shows the commutativity of the diagram (4.15).

Conversely, let $T : X^k \to X^l$ be injective and assume range T is an A_1-invariant subspace of X^l. To show that the diagram (4.15) actually defines a pair (A, C) we apply Lemma 2.2 with an F-vector space structure. By assumption range $A_1 T \subset$ range T and so there exists a linear map $A : X^k \to X^k$ for which $A_1 T = TA$. $C : X^k \to X$ is defined by $C = C_1 T$. From Theorem 4.1 it follows, as (A, C) determines a realization of L^{-1} in X^k, that actually the rank condition

$$\operatorname{rank}\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} = nk \tag{4.17}$$

holds.

Now let f and f_1 be the restricted input/output maps corresponding to the transfer functions L^{-1} and L_1^{-1} respectively. Let $f = OR$ and $f_1 = O_1 R_1$ be the factorizations of f and f_1 through X^k and X^l respectively. Since $O = O, T$, range f = range O and range f_1 = range O_1 it follows that range $f \subset$ range f_1. By Theorem 3.2 there exists an $F[\lambda]$ homomorphism $\phi : X[\lambda] \to X[\lambda]$ for which $f = f_1 \phi$. Since ϕ is given by multiplication by some $\Phi' \in (X, X)_F[\lambda]$, Φ' is of course not uniquely defined; however, from $f = f_1 \phi$ we obtain $\pi_-(L_1^{-1} \Phi' - L^{-1})x = 0$ for all $x \in X[\lambda]$. This shows that $L_1^{-1} \Phi' - L^{-1} = M$ for some $M \in (X, X)_F[\lambda]$. Thus $\Phi' L = L_1 + L_1 ML$ and hence $\Phi L = L_1$, where Φ is defined by $\Phi = \Phi' - L_1 M$. Thus $L_1 = \Phi L$ follows and this proves the theorem.

Corollary 4.3

Let L_1 be a monic polynomial of degree l and let (A_1, B_1, C_1) be a canonical realization of L_1^{-1} in X^l. Let $\mathscr{L}$ be a kn-dimensional A_1-invariant subspace of X^l such that the map

$$\left.\begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{k-1} \end{bmatrix}\right| \mathscr{L} \tag{4.18}$$

is an invertible map from $\mathscr{L}$ into X^k. Then L_1 is factorable as $L_1 = \Phi L$ for some monic polynomial L of degree k. A realization of L^{-1} is determined by the pair (A, C), where $A = A_1 | \mathscr{L}$ and $C = C_1 | \mathscr{L}$.

Conversely, if $L_1 = \Phi L$ for some monic polynomial L of degree k, then there exists a kn-dimensional A_1-invariant subspace $\mathscr{L}$ of X^l s.t. the operator defined by (4.18) is invertible.

Proof

Assume $\mathscr{L}$ is a kn-dimensional A_1-invariant subspace of X^l for which the operator (4.18) is invertible. This means that the pair (A, C) satisfies the rank condition (4.17). By Theorem 4.1 it uniquely determines a map B such that (A, B, C) is a canonical realization of L^{-1} for some uniquely determined monic polynomial L. Let J be the injection of $\mathscr{L}$ into X^l, then by Theorem 3.2 there exists a homomorphism $\phi : X[\lambda] \to X[\lambda]$ for which $f = f_1 \phi$ and this leads to the factorization of L_1 as in the previous theorem.

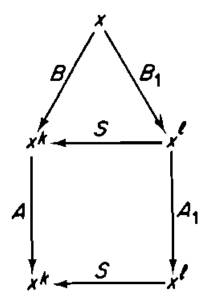
To prove the converse let $L_1 = \Phi L$ and let (A, B, C) be a canonical realization of L^{-1} in X^k. If f and f_1 are the input/output maps corresponding to the transfer functions L^{-1} and L_1^{-1}, then the diagram (4.16) is commutative for some uniquely determined injective homomorphism $T: X^k \to X^l$. Let $\mathscr{L} = \text{range } T$, then $\mathscr{L}$ being a submodule is clearly an A_1-invariant subspace. Moreover since $O = O$, T is injective we must have that $O_1|(\mathscr{L} = \text{range } T)$ is injective. This is clearly equivalent to the invertibility of

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} \left. \begin{bmatrix} C_1 \\ C_1A_1 \\ \vdots \\ C_1A_1^{k-1} \end{bmatrix} \right| \mathscr{L} \tag{4.19}$$

We proceed now to derive the dual result concerning the existence of monic left factors.

Theorem 4.4

Let L_1 and L be monic matrix polynomials of degrees l and k respectively. Let (A_1, B_1, C_1) be a canonical realization of L_1^{-1} in X^l. Then L is a left divisor of L_1 if and only if there exists a surjective map $S: X^l \to X^k$ whose kernel is A_1 invariant such that the maps A and B defined by the commutative diagram



(4.19)

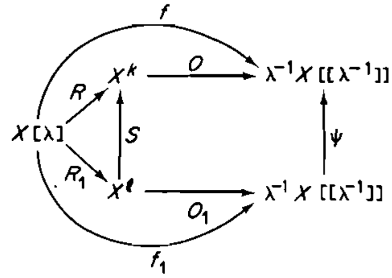
determine a canonical realization of L^{-1}.

Proof

Assume $L_1 = L\Psi$ is a factorization of L_1 with L monic, then $L^{-1} = \Psi L_1^{-1}$. Let $f, f_1 : X[\lambda] \to \lambda^{-1}X[[\lambda^{-1}]]$ be the restricted input/output maps that correspond to the transfer functions L^{-1} and L_1^{-1} respectively and let

$$\psi: \lambda^{-1}X[[\lambda^{-1}]] \to \lambda^{-1}X[[\lambda^{-1}]]$$

be the $F[\lambda]$ homomorphism defined by $\psi(x) = \pi(\Psi x)$. The system f is simulated by the system f_1, thus by Theorem 3.3 there exists a surjective $F[\lambda]$ homomorphism S which makes the diagram



(4.20)

commutative. Clearly S is linear and $SA_1 = AS$. This in turn implies that $\ker S$ is A_1 invariant. Also from $SR_1 = R$ it follows, restricting ourselves to zero degree polynomials, that $SB = B_1$ and this shows the commutativity of diagram (4.19).

Conversely let $S: X^l \to X^k$ be surjective and $\ker SA_1$ invariant. Since S is surjective and $\ker S \subset \ker SA_1$ it follows by Lemma 2.1 that there exists a linear map $A: X^k \to X^k$ for which $SA_1 = AS$. Define B by $B = SB_1$, then the pair (A, B), which is assumed to realize L^{-1} in X^k, must satisfy the rank condition

$$\operatorname{rank}(B \ AB \dots A^{k-1}B) = nk \tag{4.21}$$

If f and f_1 are the restricted input/output maps corresponding to L^{-1} and L_1^{-1} respectively, then if R and R_1 are the reachability maps of (A, B) and (A_1, B_1) respectively we clearly have $R = SR_1$ and so $\ker R_1 \subset \ker R$, which implies $\ker f_1 \subset \ker f$. By Theorem 3.3 there exists an $F[\lambda]$ homomorphism $\psi: \lambda^{-1}X[[\lambda^{-1}]] \to \lambda^{-1}X[[\lambda^{-1}]]$ which is of the form $\psi(x) = \pi_-(\Psi' x)$ for some $\Psi' \in (X, X)_F[\lambda]$, that satisfies $f = \psi f_1$. This last equality implies that

$$\pi_-(L^{-1}x - \Psi' \pi_- L_1^{-1} x) = \pi_-(L^{-1} - \Psi' L_1^{-1})x = 0$$

for all $x \in X[\lambda]$. So $L_1^{-1} - \Psi' L_1^{-1} = N$ for some $N \in (X, X)_F[\lambda]$. This yields $L_1 - L\Psi' = LNL_1$ or $L_1 = L\Psi$ with $\Psi = \Psi' + NL_1$, which proves the theorem.

Corollary 4.5

Let L_1 be a monic matrix polynomial of degree l and let (A_1, B_1, C_1) be a canonical realization of L_1^{-1} in X^l. Let $\mathscr{L}$ be an $(l-k)n$-dimensional A_1-invariant subspace of X^l for which the map

$$(B_1 \ A_1B_1 \dots A_1^{k-1} B_1) \tag{4.22}$$

is an invertible map from X^k onto a complementary subspace M of $\mathscr{L}$ in X^l. Then L_1 is factorable as $L_1 = L\Psi$ for some monic polynomial $\mathscr{L}$ of degree k.

Conversely, if $L_1 = L\Psi$ for some monic polynomial L of degree k, then there exists an $(l-k)n$-dimensional A_1-invariant subspace $\mathscr{L}$ of X^l such that the map (4.22) is an invertible map from X^k onto a subspace M complementary to $\mathscr{L}$ in X^l.

Proof

Let $\mathscr{L}$ be an $(l-k)n$-dimensional A_1-invariant subspace of X^l and let $(B_1 \ A_1B_1 \dots A_1^{k-1} B_1): X^k \to X^l$ be a surjective map onto a complementary subspace M of $\mathscr{L}$, that is $X^l = M \oplus \mathscr{L}$. This means that the pair (A, B) defined by diagram (4.19) satisfies the rank condition (4.21). By Theorem 4.1 it determines uniquely a map $C: X^k \to X$ such that (A, B, C) is a canonical realization of L^{-1}, for some uniquely determined monic polynomial L. By Theorem 3.3 there exists a homomorphism $\psi: \lambda^{-1}X[[\lambda^{-1}]] \to \lambda^{-1}X[[\lambda^{-1}]]$ satisfying $f = \psi f_1$ and this leads to the factorization $L_1 = L\Psi$ as in the proof of the preceding theorem.

Conversely, let us assume L_1 has a factorization $L_1 = L\psi$ with L monic of degree k. Let (A, B, C) be a canonical realization of L^{-1} in X^k. With f and f_1 the input/output maps associated with L^{-1} and L_1^{-1} there exists a surjective homomorphism $S : X^l \to X^k$ which makes diagram (4.19) commutative. Let $\mathscr{L} = \ker S$, then clearly being a submodule it is A_1 invariant and moreover as S is surjective we must have dim $\mathscr{L} = (l-k)n$. Also as

$$(B \ \ AB \ \ A^{k-1}B) : \ X^k \to X^k$$

is surjective it follows from $R = SR_1$ that $M = \text{range}\,(B_1 \ \ A_1B_1 \ldots A_1^{k-1} B_1)$ is a complementary subspace to $\mathscr{L}$ in X^l.

Acknowledgment

The help of S. K. Mitter, D. S. Flamm and K. Overtoom in pointing out some mistakes in an earlier version of this paper is gratefully acknowledged.

References

Fuhrmann, P. A., 1976 a, *J. Franklin Inst.*, **301,** 521 ; 1976 b, Math. Report No. 143, Ben Gurion University of the Negev, Beer Sheva, Israel.

Gohberg, I., Lancaster, P., and Rodman, L., 1978, *Linear Alg. Appl.*, **20,** 1.

Hautus, M. L. J., and Heymann, M., 1976, *J. SIAM Control*, **16,** 83.

Kalman, R. E., 1969, *Lectures on Controllability and Observability*, CIME Summer Course, 1968, Cremonese, Roma.

Kalman, R. E., Falb, P., and Arbib, M. A., 1969, *Topics in Mathematical System Theory* (New York : McGraw-Hill).

Langer, H., 1976, *Acta sci. Math.*, **38,** 83.