Ilchmann, Achim :

Time-varying linear systems and invariants of system equivalence

Zuerst erschienen in:
International Journal of Control 42 (1985), Nr. 4, S. 759-790

INT. J. CONTROL, 1985, VOL. 42, NO. 4, 759–790

Time-varying linear systems and invariants of system equivalence

ACHIM ILCHMANN†

In the paper we consider time-varying systems with coefficients depending meromorphically on time. In differential operator representations these systems are described by matrices over a skew polynomial ring with coefficients in the field of real meromorphic functions. Different kinds of indices (controllability, minimal, geometric and dynamical) are introduced and it is proved that they essentially coincide. The input module and the formal transfer matrix are defined and used for an algebraic description of time-varying systems. A characterization of system equivalence is given in these terms and also a complete list of invariants of similarity for time-varying state-space systems.

1. Introduction

In the present paper we consider time-varying analytic state-space systems of the form

$$\left.\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + E(D)(u) \end{aligned}\right\} \tag{1.1}$$

where the matrices A, B and C are time-varying with entries in

$$\mathscr{A} := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ is real analytic}\}$$

$$E(D) = \sum_{i=0}^{k} E_i D^i \quad E_i \text{ defined over } \mathscr{A}$$

and

$$D : \mathscr{A} \to \mathscr{A}, \; f \mapsto D(f) = \dot{f} \tag{1.2}$$

denotes the usual differential operator.

The associated system matrix of (1.1) is of the form

$$\mathbb{P} = \begin{bmatrix} DI_n - A & -B \\ C & E(D) \end{bmatrix} \in \mathscr{A}[D]^{(n+p)\times(n+m)} \tag{1.3}$$

with

$$\mathscr{A}[D] := \left\{ \sum_{i=0}^{k} f_i D^i \mid f_i \in \mathscr{A}, \quad 0 \leqslant i \leqslant k, \quad k \in \mathbb{N} \right\}$$

More generally we will study system matrices defined over

$$\mathscr{M}[D] := \left\{ \sum_{i=0}^{k} f_i D^i \mid f_i \in \mathscr{M}, \quad 0 \leqslant i \leqslant k, \quad k \in \mathbb{N} \right\}$$

Received 10 January 1984. 5

† Forschungsschwerpunkt Dynamische Systeme, Universität Bremen, Bibliothekstraße, Postfach 330 440, D-2800 Bremen 33, West Germany.

where

$$\mathcal{M} := \{f : \mathbb{R} \to \bar{\mathbb{R}} \mid f \text{ real meromorphic}\}$$

and $D : \mathcal{M} \to \mathcal{M}$, $f \mapsto D(f) = \dot{f}$ is the extension of (1.2) to all of $\mathcal{M}$. $\mathcal{M}[D]$ is an $\mathbb{R}$-vector space. If we identify any $f \in \mathcal{M}$ with the element $g \mapsto f(g) := fg$ of $\mathrm{end}_{\mathbb{R}}(\mathcal{M})$ and define $(Df)(g) = D(fg)$ then

$$(Df)(g) = f\dot{g} + \dot{f}g = (fD + \dot{f})(g) \quad \text{for } f, g \in \mathcal{M} \tag{1.4}$$

Therefore $\mathcal{M}[D]$ can be considered as an $\mathbb{R}$-subalgebra of $\mathrm{end}_{\mathbb{R}}(\mathcal{M})$. From an algebraic point of view, $\mathcal{M}[D]$ is a skew polynomial ring in D with coefficients in $\mathcal{M}$ and the multiplication rule

$$Df = fD + \dot{f} \quad \text{for } f \in \mathcal{M} \tag{1.5}$$

As opposed to time-invariant systems, where the system matrix is defined over the commutative rings $\mathbb{R}[D]$, resp. $\mathbb{C}[D]$, we consider system matrices over the non-commutative ring $\mathcal{M}[D]$ in this paper. Basic results of the theory of skew polynomial rings are given, for example, by Cohn (1971).

It has already been shown in Ilchmann *et al.* (1984) that the skew polynomial ring $\mathcal{M}[D]$ yields an appropriate framework for an algebraic study of time-varying systems. Different frameworks have been suggested by Kamen (1976) and Ylinen (1980). Kamen (1976) considers input–output equations of the form

$$A(z) = B(u) \tag{1.6}$$

where A and B are matrices over a skew polynomial ring $J[p]$, p is a derivative operator and J a left noetherian ring. The Noether condition appears to be somewhat restrictive. The set of real analytic functions is not noetherian. Ylinen (1980) considers equations of the form (1.6) where A and B are defined over a skew polynomial ring with coefficients in any subring of $\mathscr{C}^\infty$ (i.e. the space of infinitely differentiable complex-valued functions on an open real interval) which does not contain zero-divisors.

In the present paper it is important to distinguish between two considerations of the elements

$$P(D) = \sum_{i=0}^{k} P_i D^i \in \mathcal{M}[D]^{r \times n}$$

Let $z \in (\mathscr{C}^\infty)^n$. Then

$$P(D)z = \sum_{i=0}^{k} P_i D^i z \in \mathcal{M}[D]^r$$

is obtained by formal multiplication in $\mathcal{M}[D]$, while

$$P(D)(z) = \sum_{i=0}^{k} P_i z^{(i)} \in \mathcal{M}^r$$

denotes the action of the differential operator $P(D)$ on z.

In § 2 we give some basic results for time-varying systems described by higher-order differential equations of the form

$$\left.\begin{aligned} P(D)(z) &= Q(D)(u) \\ y &= V(D)(z) + W(D)(u) \end{aligned}\right\} \tag{1.7}$$

where the matrices P, Q, V and W are defined over $\mathcal{M}[D]$ (cf. Ilchmann *et al.* (1984)). These equations extend the differential operator representation as introduced by Rosenbrock (1970) to linear time-varying systems.

In § 3 we generalize Rosenbrock's (1970), resp. Kalman's (1971), definition of controllability indices to state-space systems with real analytic coefficients. It is shown that a system of the form (1.1) is controllable if and only if its sum of controllability indices coincides with the dimension of the system.

In § 4 $\mathscr{M}[D]$-right submodules of $\mathscr{M}[D]^r$ are analysed and minimal bases of these modules are characterized. This is an extension of Forney's main theorem (1975, p. 495). Minimal indices of a module are defined. Analogously to the time-invariant case (see Münzner and Prätzel-Wolters (1979)), the set of transformation matrices which transform a minimal basis of a module to another minimal basis is characterized.

In § 5 the mathematical theory of modules over the non-commutative ring $\mathscr{M}[D]$ as developed in § 4 is used for systems-theoretical questions. It is shown that for a system of the form (1.7) the right $\mathscr{M}[D]$-input module

$$\mu(P(D), Q(D)) = \{u \in \mathscr{M}[D]^m \mid \exists\, z \in \mathscr{M}[D]^n : P(D)z = Q(D)u\}$$

is invariant under system equivalence. For analytic state-space systems, we provide a proof that the set of controllability indices and the set of minimal indices of $\mu(DI_n - A, B)$ coincide. From knowledge of the controllability matrix of a controllable system of the form (1.1) a minimal basis of $\mu(DI_n - A, B)$ is constructed.

Brunovský (1970) derives a complete set of invariants for the action of the full feedback group on time-varying state-space systems. These 'geometric indices' are, in general, time-varying. For analytic state-space systems they are constant on $\mathbb{R}\backslash N$, where N is a discrete set. In § 6 it is proved that, on $\mathbb{R}\backslash N$, the set of geometric indices coincide with the set of controllability indices and with the set of the minimal indices of the input module of a given analytic state-space system.

In § 7 we introduce a left skew polynomial field $\mathscr{M}(D)$ of $\mathscr{M}[D]$. This enables us to define a formal transfer matrix $VP^{-1}Q + W$ over $\mathscr{M}(D)$ for systems of the form (1.7). It is invariant under system equivalence. In contrast to time-invariant systems, no interpretation is possible via the Laplace transform. The formal transfer matrices form an $\mathbb{R}$-algebra. In Ilchmann *et al.* (1984, Definition 7.2) an input–output map for systems of the form (1.7) is defined. We prove that for two systems, the formal transfer matrices coincide if and only if the input–output maps coincide.

In § 8 it is shown that the module

$$\{u \in \mathscr{M}[D]^m \mid (VP^{-1}Q + W)u \in \mathscr{M}[D]^p\} \tag{1.8}$$

of an observable system of the form (1.7) coincides with the input module of the system. This is an extension of Forney's (1975) results. The dynamical indices are defined as the minimal indices of the module (1.8).

In § 9 we provide a proof that two systems of the form (1.7) are system-equivalent if and only if their input modules and their formal transfer matrices coincide. Furthermore, as an extension of Popov (1972), we specify a complete set of similarity invariants for controllable analytic state-space systems.

2. Preliminaries

In the present paper we consider time-varying finite-dimensional linear systems in the differential operator representations

$$\left.\begin{aligned} P(D)(z) &= Q(D)(u) \\ y &= V(D)(z) + W(D)(u) \end{aligned}\right\} \tag{2.1}$$

with $P(D)\in\mathscr{M}[D]^{r\times r}$, $Q(D)\in\mathscr{M}[D]^{r\times m}$, $V(D)\in\mathscr{M}[D]^{p\times m}$ and $W(D)\in\mathscr{M}[D]^{p\times m}$

$$u\in\mathscr{U}^m := \{u\in(\mathscr{C}^\infty)^m \mid \operatorname{supp} u \text{ bounded to the left}\}$$

$z\in(\mathscr{C}^\infty)^r$ and $y\in(\mathscr{C}^\infty)^p$.

Following Ilchmann *et al.* (1984) we suppose:

(A1) im $Q(D)\subseteq$ im $P(D)$
i.e. for every input u there exists a solution z with $P(D)(z)=Q(D)(u)$

(A2) $P(D)$ is full w.r.t. $\mathscr{A}$
i.e. $P(D)$ is non-singular and if z is real analytic on a non-void open interval I of $\mathbb{R}$ and $P(D)(z)|_I=0$, then z can be analytically continued to all of $\mathbb{R}$ and $P(D)(z)=0$.

The matrix

$$\mathbb{P}=\begin{bmatrix} P & -Q \\ V & W\end{bmatrix}\in\mathscr{M}[D]^{(r+p)\times(r+m)}$$

is called the system matrix corresponding to equation (2.1) and assumptions (A 1) and (A 2).

For simplicity's sake, we often write P instead of $P(D)$.

The class of systems of the form (2.1) includes (cf. Ilchmann *et al.* 1984):

(i) time-invariant systems in differential operator representation as introduced by Rosenbrock (1970);

(ii) system matrices with $P\in\mathscr{A}[D]^{r\times r}$ non-singular and in normed upper triangular form as dealt with in Ylinen (1980); and

(iii) analytic state-space systems, i.e. systems of the form (1.1).

The system $\dot{x}=Ax+Bu$ is identified with the pair $(A,B)\in\mathscr{A}^{n\times(n+m)}$. For systems

$$\mathbb{P}_i=\begin{bmatrix} P_i & -Q_i \\ V_i & W_i\end{bmatrix}\in\mathscr{M}[D]^{(r_i+p)\times(r_i+m)},\quad i=1,2$$

the concept of system equivalence is introduced (see Ilchmann *et al.* (1984), Proposition 5.3) as follows:

$\mathbb{P}_1$ is *system-equivalent* to $\mathbb{P}_2$, written $\mathbb{P}_1\overset{se}{\sim}\mathbb{P}_2$, iff there exist matrices T, T_1, X, Y over $\mathscr{M}[D]$ of compatible dimension with

$$\begin{bmatrix} T & 0 \\ X & I_p\end{bmatrix}\mathbb{P}_1=\mathbb{P}_2\begin{bmatrix} T_1 & Y \\ 0 & I_m\end{bmatrix} \tag{2.2}$$

and T, P_2, resp. P_1, T_1, are left, resp. right, coprime.

For a given system

$$\mathbb{P}=\begin{bmatrix} P & -Q \\ V & W\end{bmatrix}\in\mathscr{M}[D]^{(n+p)\times(n+m)}$$

the solution space

$$M(P,Q) := \{(z,u)^{\mathrm{T}}\in(\mathscr{C}^\infty)^n\times\mathscr{U}^m \mid P(z)=Q(u)\} \tag{2.3}$$

(where, in order to simplify the notation we use $(z, u)^{\mathrm{T}}$ instead of $(z^{\mathrm{T}}, u^{\mathrm{T}})^{\mathrm{T}}$) can be decomposed into the direct sum of the $\mathbb{R}$-vector space of forced motions starting from zero

$$M_+(P, Q) := \{(z_u, u)^{\mathrm{T}} \in M(P, Q) \cap (\mathscr{U}^n \times \mathscr{U}^m) \tag{2.4}$$

and the $\mathbb{R}$-vector space of free motions

$$\ker P \times \{0\} := \{(z', 0)^{\mathrm{T}} \in M(P, Q)\} \tag{2.5}$$

The assumption that P is full is essential for the decomposition. In contrast to the time-invariant case where $M(P, Q)$ is an $\mathbb{R}[D]$-module, for time-varying systems of the form (2.1) $M(P, Q)$ is in general only an $\mathbb{R}$-vector space and not an $\mathbb{R}[D]$- or $\mathscr{M}[D]$-module.

3. Controllability indices

In this section we introduce controllability indices and characterize controllability for time-varying state-space systems, whose coefficients depend analytically on time.

For $(A, B) \in \mathscr{A}^{n \times (n+m)}$ and $l \in \mathbb{N}$ we define

$$K^l(A, B) := [(DI_n - A)^0(B), \dots, (DI_n - A)^l(B)] \tag{3.1}$$

where

$$(DI_n - A)^0 := I_n \quad \text{and} \quad (DI_n - A)^i := \underbrace{(DI_n - A) \dots (DI_n - A)}_{i\text{-times}}$$

If the transition matrix of $\dot{x} = Ax$ is denoted by $\phi(t, t_0)$, or briefy by ϕ, the matrix

$$K(A, B) := K^{n-1}(A, B) = \phi[\phi^{-1}B, (\dot{\phi^{-1}B}), \dots, (\phi^{-1}B)^{(n-1)}] \tag{3.2}$$

is said to be the *controllability matrix* of (A, B) (see Silverman and Meadows (1967)).

Definition 3.1

Let R be a ring and $\mathrm{GL}_n(R) := \{A \in R^{n \times n} | \exists A^{-1} \in R^{n \times n} : AA^{-1} = I_n\}$.

Two systems

$$\mathbb{P} = \begin{bmatrix} DI_n - A & -B \\ C & E(D) \end{bmatrix}, \quad \mathbb{P}' = \begin{bmatrix} DI_n - A' & -B' \\ C' & E'(D) \end{bmatrix} \in \mathscr{A}[D]^{(n+p) \times (n+m)}$$

are called (analytically) *similar* (via T), written $\mathbb{P} \overset{s}{\sim} \mathbb{P}'$, if there exists a $T \in \mathrm{GL}_n(\mathscr{A})$ such that

$$AT - TA' = \dot{T}, \quad B' = TB, \quad C' = CT^{-1}, \quad E(D) = E'(D)$$

These equations are equivalent to

$$\begin{bmatrix} T & 0 \\ 0 & I_p \end{bmatrix} \begin{bmatrix} DI_n - A & -B \\ C & E(D) \end{bmatrix} = \begin{bmatrix} DI_n - A' & -B' \\ C' & E'(D) \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & I_m \end{bmatrix}$$

Lemma 3.1

If $\mathbb{P}$ and $\mathbb{P}'$ as given in Definition 3.1 are similar via T it follows that

$$(DI_n - A')^i(B') = T(DI_n - A)^i(B) \quad \text{for every } i \in \mathbb{N}$$

Proof

For $i=0$ the equality holds by Definition 3.1. If the assertion is valid for $i>0$ then

$$\begin{aligned}(DI_n - A')^{i+1}(B') &= (DI_n - A')[T(DI_n - A)^i(B)] \\ &= [T(DI_n - A)][(DI_n - A)^i(B)]\end{aligned} \qquad \square$$

As a consequence, for the controllability matrices of similar systems, (A, B), $(A', B') \in \mathscr{A}^{n\times(n+m)}$ we have

$$T \cdot K(A, B) = K(A', B') \tag{3.3}$$

Applying Rosenbrock's deleting procedure (see Rosenbrock (1970), p. 90) to $K(A, B)$ for a given $(A, B) \in \mathscr{A}^{n\times(n+m)}$ we get

$$H := [b_1, \phi(\phi^{-1}b_1), \ldots, (\phi^{-1}b_1)^{(k_1-1)}, b_2, \ldots, (\phi^{-1}b_m)^{(k_m-1)}] \in \mathscr{A}^{n\times n'} \tag{3.4}$$

with $n' \leqslant n$ where $b_1, \ldots, b_m$ denote the columns of B. If $k_i = 0$ the corresponding column in H is omitted.

Note, if $\phi(\phi^{-1}b_i)^{(j)}$ is linearly dependent on its predecessors then $\phi(\phi^{-1}b_i)^{(j+1)}$ is too. This is not valid, in general, if one constructs a matrix $H(t)$ for fixed $t \in \mathbb{R}$ and considers linear dependency over $\mathbb{R}$. Consider in Example 3.1, $K(A, B)$ at $t=0$ and $t=1$. Thus there is no chance to define time-varying k_i pointwise by the same deleting process.

The numbers $k_1, \ldots, k_m$ are called the *controllability indices* of (A, B) and because of (3.3) they are invariant under similarity.

The following example will also be used later to illustrate new definitions.

Example 3.1

Let

$$(A, B) := \left(0_{3\times3}, \begin{bmatrix} \exp t & -\exp t & 0 \\ t-1 & 1 & t \\ 0 & t & t \end{bmatrix}\right)$$

It is easily computed that

$$K(A, B) = \left[\begin{array}{ccc|ccc|ccc} \exp t & -\exp t & 0 & \exp t & -\exp t & 0 & \exp t & -\exp t & 0 \\ t-1 & 1 & t & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & t & t & 0 & 1 & 1 & 0 & 0 & 0 \end{array}\right]$$

and

$$H = [b_1, \dot{b}_1, b_2] = \begin{bmatrix} \exp t & \exp t & -\exp t \\ t-1 & 1 & 1 \\ 0 & 0 & t \end{bmatrix}$$

Therefore $(k_1, k_2, k_3) = (2, 1, 0)$.

Definition 3.2

A system

$$\begin{bmatrix} DI_n - A & -B \\ C & E(D) \end{bmatrix} \in \mathscr{A}[D]^{(n+p)\times(n+m)}$$

is called *controllable* if for any $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ there exists a $t_1 \geqslant t_0$ and a $u \in \mathscr{U}^m$ with $\operatorname{supp} u \subseteq [t_0, t_1]$ such that

$$\phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \phi(t, s)B(s)u(s)\, ds = 0$$

where $\phi(\cdot, t_0)$ denotes the transition matrix of (A, B).

Other authors (for example, Kalman (1962)) only require that the control functions are piecewise continuous. But for analytic state-space systems it is not restrictive to require that $u \in \mathscr{U}^m$. Moreover, controllability and total controllability (i.e. controllability in every open non-void subinterval of $\mathbb{R}$) coincide for these systems (see Silverman and Meadows (1967) and Ilchmann *et al.* (1984, Appendix)).

Proposition 3.1

Given

$$\mathbb{P} = \begin{bmatrix} DI_n - A & -B \\ C & E(D) \end{bmatrix} \in \mathscr{A}[D]^{(n+p)\times(n+m)}$$

with controllability indices $k_1, \ldots, k_m$,

$$l := \operatorname{rk} B = \sum_{i:k_i>0} 1$$

then the following are equivalent:

(i) $\mathbb{P}$ is controllable

(ii) $\operatorname{rk} K(A, B) = n$

(iii) $\operatorname{rk} K^{n-l}(A, B) = n$

(iv) $\sum_{i=1}^{m} k_i = n$

Proof

As mentioned above, it is not restrictive if $u \in \mathscr{U}^m$, and therefore '(i)$\Leftrightarrow$(ii)' can be proved using Silverman and Meadows (1967, p. 69). '(ii)$\Leftrightarrow$(iv)' and '(iii)$\Leftrightarrow$(ii)' are immediate. It remains to prove '(ii)$\Rightarrow$(iii)': without restriction of generality assume $k_1 \geqslant 1, \ldots, k_l \geqslant 1$, $k_{l+1} = \ldots = k_m = 0$. The assumption that there exists $i \geqslant 1$ such that $k_i > n - l + 1$ leads to the contradiction

$$n = \sum_{i=1}^{l} k_i > l - 1 + n - l + 1 = n$$

Therefore $k_i \leqslant n - l + 1$ for $i = 1, \ldots, m$ and (iii) is proved. □

4. Minimal bases of $\mathscr{M}[D]$-right modules

In this section we analyse submodules of the free $\mathscr{M}[D]$-right module $\mathscr{M}[D]^r$ and characterize their minimal bases.

If m is a right- (left-) $\mathcal{M}[D]$-module, its rank is the cardinality of any maximal right- (left-) linearly independent (over $\mathcal{M}[D]$) subset of element of m (see Cohn (1971, p. 28)). Since $\mathcal{M}[D]$ is a right and left euclidean domain (see Ore 1933) it follows for the free $\mathcal{M}[D]$-right module $\mathcal{M}[D]^r$ that each of its submodules is also free and of rank at most r (see Cohn (1971, p. 46)).

For a matrix $P \in \mathcal{M}[D]^{r \times k}$ the column (row) rank is defined as the rank of the right (left) $\mathcal{M}[D]$-submodule of $\mathcal{M}[D]^r$ ($\mathcal{M}[D]^{1 \times k}$) spanned by the columns (rows) of P. Both ranks coincide (see Cohn (1971, p. 195)).

For $v := (v^1, \ldots, v^r)^T \in \mathcal{M}[D]^r$ let

$$\deg v := \max \{\deg v^i, \quad i = 1, \ldots, r\}$$

where $\deg v^i$ denotes the usual degree of $v^i \in \mathcal{M}[D]$.

For $V = [v_1, \ldots, v_k] \in \mathcal{M}[D]^{r \times k}$ let $\lambda_i := \deg v_i$ be the *ith index* of V $(1 \leqslant i \leqslant k)$ and $\operatorname{ord} V := \sum_{i=1}^{k} \lambda_i$ the *order* of V.

Let m be a right-$\mathcal{M}[D]$ submodule of $\mathcal{M}[D]^r$, written $m \hookrightarrow \mathcal{M}[D]^r$. If m is of rank k then $V \in \mathcal{M}[D]^{r \times k}$ is called a *minimal basis* of m if $m = V \cdot \mathcal{M}[D]^k$, i.e. V is a *basis* of m, and V has the least order among all bases for m.

Let $v_i = \sum_{j=0}^{\lambda_i} D^j v_{ij}$ for $i = 1, \ldots, k$. Then the leading (column) coefficient matrix of V is defined as

$$[V]_l := [v_{1,\lambda_1}, \ldots, v_{k,\lambda_k}]$$

Note that this matrix does not depend on the side on which the coefficients of the column polynomials v_i are written.

The following proposition characterizes a minimal basis of a $\mathcal{M}[D]$-submodule of $\mathcal{M}[D]^r$. This is a generalization of Forney's main theorem (1975, p. 495), see also Münzner and Prätzel-Wolters (1979, p. 293).

Proposition 4.1

Let $m = V \cdot \mathcal{M}[D]^k$ with $V = [v_1, \ldots, v_k] \in \mathcal{M}[D]^{r \times k}$ and $\lambda_1, \ldots, \lambda_k$ denote the indices of V. Then the following are equivalent:

(i) V is a minimal basis of m

(ii) $\operatorname{rk} [V]_l = k$

(iii) For any $x = (x_1, \ldots, x_k)^T \in \mathcal{M}[D]^k \backslash \{0\}$

$$\deg Vx = \max \{\deg x_i + \lambda_i | x_i \neq 0\}$$

(iv) For $d \in \mathbb{N}$, $d \geqslant 0$ the $\mathcal{M}$-vector space

$$m_d := \{v \in m | \deg v \leqslant d\}$$

has dimension $\dim_{\mathcal{M}} m_d = \sum_{i:\lambda_i \leqslant d} (d + 1 - \lambda_i)$

Proof

(i) $\Rightarrow$ (ii): Let $(m_1, \ldots, m_k)^T \in \mathcal{M}^k \backslash \{0\}$ such that $\sum_{i=1}^{k} [v_i]_l m_i = 0$ and λ_p be the maximal index of V with $m_p \neq 0$. Then

$$
\begin{aligned}
v' := \sum_{i=\cancel{0}1}^{k} v_i D^{(\lambda_p - \lambda_i)} m_i &= \sum_{i=\cancel{0}1}^{k} \left(\sum_{j=0}^{\lambda_i - 1} D^j v_{ij} + D^{\lambda_i} v_{i,\lambda_i} \right) D^{(\lambda_p - \lambda_i)} m_i \\
&= \sum_{i=\cancel{0}1}^{k} \sum_{j=0}^{\lambda_i - 1} D^j v_{ij} D^{(\lambda_p - \lambda_i)} m_i + \sum_{i=\cancel{0}1}^{k} D^{\lambda_i} (D^{(\lambda_p - \lambda_i)} v_{i,\lambda_i} + w_i) m_i \\
&\qquad \text{with } w_i \text{ such that } \deg w_i < \lambda_p - \lambda_i \\
&= w + D^{\lambda_p} \sum_{i=\cancel{0}1}^{k} v_{i,\lambda_i} m_i \\
&\qquad \text{with } w \text{ such that } \deg w < \lambda_p \\
&= w
\end{aligned}
$$

Since

$$
v_p = \left(v' - \sum_{i=\cancel{0}1, i \neq p}^{k} v_i D^{(\lambda_p - \lambda_i)} m_i \right) m_p^{-1}
$$

the matrix $[v_1, \ldots, v_{p-1}, v', v_{p+1}, \ldots, v_m]$ is a basis with lower order than V. This contradicts (i).

(ii)$\Rightarrow$(iii): Let $x = (x_1, \ldots, x_k)^{\mathrm{T}} \in \mathcal{M}[D]^k \backslash \{0\}$. Then

$$
\deg Vx = \deg \sum_{i=1}^{k} v_i x_i \leq \max \{\deg x_i + \lambda_i | x_i \neq 0\} =: a
$$

Let $l_i := \deg x_i$ for $i = 1, \ldots, k$ and $N := \{i \in \{1, \ldots, k\} | l_i + \lambda_i = a\}$. Then

$$
\begin{aligned}
Vx &= \sum_{i=0}^{k} \sum_{j=0}^{\lambda_i} D^j v_{ij} \sum_{\mu=0}^{l_i} D^{\mu} x_{i\mu} \\
&= \sum_{i=0}^{k} \sum_{j=0}^{\lambda_i} D^j \sum_{\mu=0}^{l_i} (D^{\mu} v_{ij} + y_{\mu ij}) x_{i\mu} \\
&\qquad\qquad \text{with } y_{\mu ij} \text{ such that } \deg y_{\mu ij} < \mu \\
&= D^a \sum_{i \in \mathbf{N}} v_{i\lambda_i} x_{il_i} + y \\
&\qquad\qquad \text{with } y \text{ such that } \deg y < a
\end{aligned}
$$

(iii)$\Rightarrow$(iv): Use similar arguments as in Münzner and Prätzel-Wolters (1979, p. 294).

(iv)$\Rightarrow$(i): For $d \in \mathbb{N}$ let $h(d) := \sum_{i:\lambda_i = d} 1$, i.e. the number of indices of V equal to d. (iv) yields

$$
\begin{aligned}
h(d) &= \sum_{i:\lambda_i \leqslant d} (d + 1 - \lambda_i) + (d - 1 - \lambda_i) - 2(d - \lambda_i) - \sum_{i:\lambda_i = d} (d - 1 - \lambda_i) \\
&= \sum_{i:\lambda_i \leqslant d} (d + 1 - \lambda_i) + \sum_{i:\lambda_i \leqslant d} (d - 2 + 1 - \lambda_i) - \sum_{i:\lambda_i = d} (d - 1 - \lambda_i) \\
&\quad - 2 \sum_{i:\lambda_i \leqslant d} (d - 1 + 1 - \lambda_i) \\
&= \dim_{\mathcal{M}} \mathscr{m}_d + \dim_{\mathcal{M}} \mathscr{m}_{d-2} - 2 \dim_{\mathcal{M}} \mathscr{m}_{d-1}
\end{aligned}
$$

If (iv) is valid, $h(d)$ is only determined by the module, not by the specific basis. All bases which satisfy (iv) have the same order $\lambda := \sum_{i=1}^{k} \lambda_i = \sum_{d=1}^{\infty} d \cdot h(d)$. Since for a minimal basis of $\mathscr{m}$ (iv) is fulfilled, it follows that V is a minimal basis. $\square$

Remark 4.1

The last part of the above proof shows that two minimal bases of a given submodule $\mathscr{m} \hookrightarrow \mathscr{M}[D]^r$ have the same set of indices. We call the indices of a minimal basis of $\mathscr{m}$ the minimal *indices* of $\mathscr{m}$.

Remark 4.2

Given $\mathscr{m} \hookrightarrow \mathscr{M}[D]^r$ of rank k one can select k vectors which form a basis of $\mathscr{m}$ (see Rosenbrock (1970, p. 96)). Part '(i)$\Rightarrow$(ii)' of the above proof leads to an algorithm which starts with an arbitrary basis of $\mathscr{m}$ and constructs a minimal basis in a finite number of steps.

Definition 4.1

Let $\mathscr{m}, \mathscr{m}' \hookrightarrow \mathscr{M}[D]^r$. A right-$\mathscr{M}[D]$ homomorphism $\psi : \mathscr{m} \to \mathscr{m}'$ is called *degree-preserving* if $\deg v = \deg \psi(v)$ for every $v \in \mathscr{m}$. Let

$$\Gamma_{\mathscr{m}} := \{\psi : \mathscr{m} \to \mathscr{m} \mid \psi \text{ is a right-}\mathscr{M}[D] \text{ isomorphism}\}$$

Fixing a minimal basis V of a given submodule $\mathscr{m} \hookrightarrow \mathscr{M}[D]^r$ one obtains a bijective map:

$$h : \{V' \mid V' \text{ is a minimal basis of } \mathscr{m}\} \to \Gamma_{\mathscr{m}}$$

$$V' \mapsto h(V') : Vx \mapsto V'x$$

(cf. Münzner and Prätzel-Wolters (1979)).

To every $h(V') \in \Gamma_{\mathscr{m}}$ one can assign a unique basis transformation matrix $T \in \mathrm{GL}_n(\mathscr{M}[D])$ with $h(V') = VT$. If $\bar{V}$ and V are two minimal bases of $\mathscr{m}$ with ordered indices $\lambda_1 \geqslant \ldots \geqslant \lambda_k$ and $\bar{V} = VT$, then T is an element of

$$\mathscr{T}_{\mathscr{m}} := \left\{ T \in \mathrm{GL}_k(\mathscr{M}[D]) \,\middle|\, \begin{array}{ll} \deg t_{ij} \leqslant \lambda_j - \lambda_i & \text{for } \lambda_i \leqslant \lambda_j \\ t_{ij} = 0 & \text{for } \lambda_i > \lambda_j \end{array} \right\}$$

This implies that T is of the following form:

$$T = \begin{bmatrix} \boxed{*} & & 0 \\ & \ddots & \\ * & & \boxed{*} \end{bmatrix}$$

with square diagonal block corresponding to the columns of V with the same degree.

Proposition 4.2

Let V be a minimal basis of a submodule $\mathscr{m} \hookrightarrow \mathscr{M}[D]^r$ of rank k with ordered indices $\lambda_1 \geqslant \ldots \geqslant \lambda_k$. Then $\bar{V} = VT$ is a minimal basis of $\mathscr{m}$ if and only if $T \in \mathscr{T}_{\mathscr{m}}$.

Proof

Let $V=[v_1,\dots,v_k]$, $\bar{V}=[\bar{v}_1,\dots,\bar{v}_k]$ and $T=(t_{ij})_{1\leqslant i,j\leqslant k}$.

'$\Rightarrow$': Use the same arguments as in Münzner and Prätzel-Wolters (1979, p. 295).

'$\Leftarrow$': First we show that the indices of $\bar{V}$ coincide with those of V. Consider a subset $J:=\{p,\dots,l\}$ of the minimal indices with $\lambda_{p-1}<\lambda_p=\dots=\lambda_l<\lambda_{l+1}$. Then

$$\deg \bar{v}_j=\deg \sum_{i=1}^{k} v_i t_{ij}\leqslant \max\{\deg v_i t_{ij} \mid 1\leqslant i\leqslant k\}\leqslant \max\{(\lambda_i+(\lambda_j-\lambda_i)) \mid 1\leqslant i\leqslant k\}=\lambda_j$$

Therefore for $j\in J$ there exists $i_0\in J$ such that $t_{i_0,j}\neq 0$ (since $T\in\mathscr{T}_m$, the diagonal blocks are invertible over $\mathscr{M}$). This implies

$$\deg \bar{v}_j=\deg Vt_j=\max\{(\deg t_{ij}+\lambda_i) \mid t_{ij}\neq 0\}\geqslant \deg t_{i_0,j}+\lambda_{i_0}=\lambda_j$$

which proves that

$$\deg \bar{v}_i=\deg v_i \quad \text{for } 1\leqslant i\leqslant k$$

If $j\in J$ we have the representation

$$\bar{v}_j=v_p t_{pj}+\dots+v_l t_{lj}+v_{l+1}t_{l+1,j}+\dots+v_k t_{kj}$$

Since $t_{pj},\dots,t_{lj}\in\mathscr{M}$ and $\deg v_i t_{ij}\leqslant\lambda_j$ for $i\in\{l+1,\dots,k\}$ we conclude that $[\bar{v}_j]_l=\sum_{i=1}^{k}[v_i]_l t_{ij}^*$ where t_{ij}^* is the coefficient of $D^{\lambda_j-\lambda_i}$ in t_{ij}. Therefore $[\bar{V}]_l=[V]_l T^*$ and T^* is invertible since the diagonal blocks of T and T^* coincide and $T\in \mathrm{GL}_k(\mathscr{M}[D])$. So $[\bar{V}]_l$ has full rank and $\bar{V}$ is a minimal basis. $\square$

5. The input module and its minimal indices

For a system

$$\mathbb{P}=\begin{bmatrix} P & -Q \\ V & W \end{bmatrix}\in\mathscr{M}[D]^{(n+p)\times(n+m)}$$

Ilchmann *et al.* (1984) analyse the differential equation $\mathbb{P}((z,u)^{\mathrm{T}})=(0,y)^{\mathrm{T}}$ for $(z,u,y)\in(\mathscr{C}^\infty)^n\times\mathscr{U}^m\times(\mathscr{C}^\infty)^p$. In this section the algebraic equation

$$Pz=Qu \quad \text{for } (z,u)\in\mathscr{M}[D]^n\times\mathscr{M}[D]^m$$

is considered.

Definition 5.1

Given $\mathbb{P}$ as above and the projection

$$\pi_2:\mathscr{M}[D]^n\times\mathscr{M}[D]^m\to\mathscr{M}[D]^m,\quad (z,u)\mapsto u$$

the *input module* of $\mathbb{P}$ is defined as

$$\mu(P,Q):=\pi_2(\ker[P,-Q])\hookrightarrow\mathscr{M}[D]^m$$

where

$$\ker[P,-Q]:=\{x\in\mathscr{M}[D]^{n+m} \mid [P,-Q]x=0\}$$

We use this notation in order to show the close connection with time-invariant systems as analysed in Münzner and Prätzel-Wolters (1979). Note that $\mu(P,Q)$

coincides with the kernel of the *controllability map*

$$\varphi: \mathcal{M}[D]^m \to \mathcal{M}[D]^n / P \cdot \mathcal{M}[D]^n, \quad u \mapsto Qu + P \cdot \mathcal{M}[D]^n$$

The following lemma shows that the minimal indices of the input module are invariant under system equivalence.

Lemma 5.1

For two system matrices

$$\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathcal{M}[D]^{(n_i+p)\times(n_i+m)}, \quad i = 1, 2$$

with $\mathbb{P}_1 \overset{\text{se}}{\approx} \mathbb{P}_2$ it follows that $\mu(P_1, Q_1) = \mu(P_2, Q_2)$.

Proof

Use Ilchmann *et al.* (1984, Proposition 5.3) to describe the system isomorphism. Then the proof is similar to Münzner and Prätzel-Wolters (1979, Proposition 4). □

A consequence of the following proposition will be that the input module of an analytic state-space system can be characterized in terms of the matrices $(DI_n - A)^i(B)$, see (3.1).

Proposition 5.1

Let $(A, B) \in \mathscr{A}^{n\times(n+m)}$ and $u = \sum_{i=0}^{r} D^i(-1)^i u_i = \sum_{i=0}^{r} (-1)^i \tilde{u}_i D^i \in \mathcal{M}[D]^m$. Then the following are equivalent:

(i) $u \in \mu(DI_n - A, B)$

(ii) $K^r(A, B)\cdot(u_0, \ldots, u_r)^{\mathrm{T}} = \sum_{i=0}^{r} (DI_n - A)^i(B)u_i = 0$

(iii) $\sum_{i=0}^{r} (DI_n - A)^i(B\tilde{u}_i) = 0$

Proof

We make us of two multiplication rules which can easily be proved by induction. Let $N \in \mathcal{M}[D]^{n\times m}$, $k \in \mathbb{N}$ and $x_0, \ldots, x_k \in \mathcal{M}^n$. Then

$$ND^k = \sum_{i=0}^{k} (-1)^i \binom{k}{i} D^{k-i} N^{(i)} \tag{5.1}$$

$$N \sum_{i=0}^{k} D^i x_i = \sum_{i=0}^{k} D^i \sum_{\lambda=i}^{k} (-1)^{\lambda-i} \binom{\lambda}{\lambda - i} N^{(\lambda-i)} x_\lambda \tag{5.2}$$

(i)⇔(ii): If $\phi(\cdot, t_0)$ denotes the transition matrix of $\dot{x} = Ax$ then $[DI_n - A]x = Bu$ for $x \in \mathcal{M}[D]^n$ is equivalent to

$$DI_n \phi^{-1} x = \phi^{-1} Bu \tag{5.3}$$

Let $\bar{x} := \sum_{i=0}^{k-1} D^i \bar{x}_i := \phi^{-1}x$ and $\bar{B} := \phi^{-1}B$. Use of (5.2) yields that (5.3) is equivalent to

$$\sum_{i=1}^{k} D^i \bar{x}_{i-1} = \sum_{i=0}^{k} D^i \sum_{\lambda=i}^{k} (-1)^{\lambda-i} \binom{\lambda}{\lambda-i} \bar{B}^{(\lambda-i)}(-1)^{\lambda} u_{\lambda}$$

Comparing the coefficients we get for $i = 1, \ldots, k$

$$0 = \sum_{\lambda=0}^{k} \bar{B}^{(\lambda)} u_{\lambda} \quad \text{and} \quad \bar{x}_{i-1} = \sum_{\lambda=i}^{k} (-1)^{i} \binom{\lambda}{\lambda-i} \bar{B}^{(\lambda-i)} u_{\lambda} \tag{5.4}$$

By Lemma 3.1, the first equation in (5.4) is equivalent to (ii). On the other hand, let u satisfy ~~(i)~~(ii). Then, using similarity and the second equation in (5.4), $x \in \mathcal{M}[D]^n$ is defined such that $[DI_n - A]x = Bu$.

(i)$\Leftrightarrow$(iii): Let $x = \sum_{i=0}^{k-1} x_i D^i \in \mathcal{M}[D]^r$ such that

$$(DI_n - A)x = \sum_{i=0}^{k-1} (x_i DI_n + (DI_n - A)(x_i))D^i = B \sum_{i=0}^{k} (-1)^i \tilde{u}_i D^i$$

By comparing the coefficients we get

$$\left.\begin{aligned}
(DI_n - A)(x_0) &= B\tilde{u}_0 \\
x_0 + (DI_n - A)(x_1) &= B\tilde{u}_1(-1) \\
&\vdots \\
x_{k-2} + (DI_n - A)(x_{k-1}) &= B\tilde{u}_{k-1}(-1)^{k-1} \\
x_{k-1} + (DI_n - A)(x_k) &= B\tilde{u}_k(-1)^k
\end{aligned}\right\} \tag{5.5}$$

Substitution yields

$$0 = B\tilde{u}_0 - (DI_n - A)(B\tilde{u}_1(-1) - (DI_n - A)(\ldots)) = \sum_{i=0}^{k} (DI_n - A)^i (B\tilde{u}_i)$$

On the other hand, if (iii) is valid, by (5.5) $x \in \mathcal{M}[D]^n$ is defined such that $[DI_n - A]x = Bu$. This proves (i). □

For $(A, B) \in \mathcal{A}^{n \times (n+m)}$ the map

$$\hat{K}_{A,B} : \mathcal{M}^m[D] \to \mathcal{M}^n$$

$$\sum_{i=0}^{r} D^i(-1)^i u_i \mapsto K^r(A, B)(u_0, \ldots, u_r)^{\mathrm{T}}$$

is a right $\mathcal{M}$-homomorphism. Using this notation we obtain the following.

Corollary 5.1

(i) The map $\psi : \ker [DI_n - A, -B] \to \ker \hat{K}_{A,B}$
$$(z, u)^{\mathrm{T}} \mapsto u$$
is a degree-preserving $\mathcal{M}[D]$-right isomorphism.

(ii) $\ker \hat{K}_{A,B} = \mu(DI_n - A, B)$

Proof

$(z, u)^{\mathrm{T}} \in \ker [DI_n - A, -B]$ implies that $\deg z < \deg u$ and therefore ψ is degree-preserving. That ψ is surjective and injective is a direct consequence of the proof of Proposition 5.1 '(i)$\Leftrightarrow$(ii)'. □

The following lemma will be used to prove that the set of controllability indices of a state-space system coincides with the set of minimal indices of its input module.

Lemma 5.2

Let $(A, B) \in \mathscr{A}^{n\times(n+m)}$ be controllable with controllability indices $k_1, \dots, k_m$. There exist $T \in \mathrm{GL}_n(\mathscr{M})$ and $U \in \mathrm{GL}_{n+m}(\mathscr{M})$ such that

(i) $$T\cdot[DI_n - A, -B]\cdot\begin{bmatrix} T^{-1} & 0 \\ 0 & I_m \end{bmatrix} = [DI_n - A', -B']$$

where

$$B' = \begin{bmatrix} B_1' \\ \vdots \\ B_m' \end{bmatrix}, \quad B_i' = (-1)^{k_i-1}\begin{array}{c} \begin{bmatrix} & & & 0_{k_i-1,m} & & & \\ 0 & \dots & 0 & 1 & * & \dots & * \end{bmatrix} \\ \uparrow \\ i\text{th column} \end{array}$$

and B_i' is omitted if $k_i = 0$

$$A' = \begin{array}{c} \left[\begin{array}{cccc|ccc} 0 & 1 & & & & & \\ & \ddots & \ddots & & & 0 & \\ & & 0 & 1 & & & \\ * & \cdots & & * & * & \cdots & * \\ \hline & & & \ddots & & & \\ \hline & & & & 0 & 1 & \\ & 0 & & & & \ddots & \ddots \\ & & & & & 0 & 1 \\ * & \cdots & & * & * & \cdots & * \end{array}\right] \end{array} \begin{array}{l} \\ \\ \\ \leftarrow s_1\text{th row} \\ \\ \\ \\ \\ \leftarrow s_m\text{th row} \end{array}$$

$s_i := k_1 + \cdots + k_i$ for $i = 1, \dots, m$ and $k_i > 0$, and the '$*$'s are elements of $\mathscr{M}$.

(ii) $$T[DI_n - A, -B]\begin{bmatrix} T^{-1} & 0 \\ 0 & I_m \end{bmatrix}U = [\operatorname{diag}(L_1, \dots, L_l), 0_{n\times(m-l)}]$$

where

$$l := \sum_{i:k_i>0} 1 = \mathrm{rk}_{\mathscr{M}}\, B$$

and

$$L_i := \begin{bmatrix} D & -1 & & \\ & \ddots & \ddots & \\ & & D & -1 \end{bmatrix} \in \mathscr{A}[D]^{k_i\times(k_i+1)} \quad \text{for } k_i > 0$$

Proof

First we show an elementary property which will be used in the proof. Let

$g^{\mathrm{T}}, b \in \mathcal{M}^n$; $r \in \mathbb{N}$ be such that

$$gb^{(\nu)} = \begin{cases} 0 & \text{for } \nu \leqslant r \\ 1 & \text{for } \nu = r+1 \end{cases}$$

Then

$$g^{(\mu)} b^{(\nu)} = \begin{cases} 0 & \text{for } \mu + \nu \leqslant r \\ (-1)^{\mu} & \text{for } \mu + \nu = r+1 \end{cases} \tag{5.6}$$

This is proved by induction on μ. Assume that the assertion is valid for $\mu < r+1$.

Let $\mu + \nu = r$. By assumption, $g^{(\mu)} b^{(\nu+1)} = (-1)^{\mu}$, and since $g^{(\mu)} b^{(\nu)} = 0$ we have $0 = (g^{(\mu)} b^{(\nu)})^{\cdot} = g^{(\mu+1)} b^{(\nu)} + g^{(\mu)} b^{(\nu+1)} = g^{(\mu+1)} b^{(\nu)} + (-1)^{\mu}$. Therefore $g^{(\mu+1)} b^{(\nu)} = (-1)^{\mu+1}$.

Let $\mu + \nu < r$. Then by assumption $g^{(\mu)} b^{(\nu)} = g^{(\mu)} b^{(\nu+1)} = 0$. Thus $0 = (g^{(\mu)} b^{(\nu)})^{\cdot} = g^{(\mu+1)} b^{(\nu)} + 0$. This completes the proof.

We proceed in several steps.

(α) Without restriction of generality we may assume that $A = 0$. Let $B = [b_1, \dots, b_m]$. For H as given in (3.4) use the representation $I_n = H^{-1} H$ in the following form

$$[b_1, \dots, b_1^{(k_1-1)}, \dots, b_i, \dots, b_i^{(k_i-1)}, \dots, b_m, \dots, b_m^{(k_m-1)}]$$

$$\left[\begin{array}{c} g_1 \\ \vdots \\ g_{s_1} \\ \hline \vdots \\ \hline g_{s_{i-1}+1} \\ \vdots \\ g_{s_i} \\ \hline \vdots \\ \hline g_{s_{m-1}+1} \\ \vdots \\ g_{s_m} \end{array}\right] \begin{array}{ccccc} \begin{array}{|ccc|} 1 & & \\ & \ddots & \\ & & 1 \\ \hline \end{array} & & & & \\ & \ddots & & & \\ & & \begin{array}{|ccc|} \hline 1 & & \\ & \ddots & \\ & & 1 \\ \hline \end{array} & & \\ & & & \ddots & \\ & & & & \begin{array}{|ccc|} \hline 1 & & \\ & \ddots & \\ & & 1 \\ \hline \end{array} \end{array}$$

If $k_i = 0$ the matrix $[g_{s_{i-1}+1}, \dots, g_{s_i}]^{\mathrm{T}}$ is omitted.

Furthermore define

$$H_1 := \left[\begin{array}{c} g_{s_1} \\ \dot{g}_{s_1} \\ \vdots \\ g_{s_1}^{(k_1-1)} \\ \hline \vdots \\ \hline g_{s_m} \\ \vdots \\ g_{s_m}^{(k_m-1)} \end{array}\right]$$

(β) We prove: $B' = H_1 B$. It suffices to show for $i \in \{1, \ldots, m\}$ and $k_i > 0$ that

$$\begin{bmatrix} g_{s_i} \\ \vdots \\ g_{s_i}^{(k_i-1)} \end{bmatrix} [b_1, \ldots, b_i, \ldots, b_m] = \left[\begin{array}{ccc|cccc} & 0 & & & 0 & & \\ \hline 0 & \ldots & 0 & 1 & * & \ldots & * \end{array}\right] (-1)^{k_i-1}$$

$\uparrow$
ith column

(β_1) It is proved that

$$\begin{bmatrix} g_{s_i} \\ \vdots \\ g_{s_i}^{(k_i-1)} \end{bmatrix} [b_1, \ldots, b_{i-1}] = 0_{k_i \times (i-1)}$$

Let $j \in \{1, \ldots, i-1\}$. Then by (α) it is known that $g_{s_i} b_j^{(\nu)} = 0$ for $\nu = 0, \ldots, k_j - 1$ and $k_j > 0$. If $k_j - 1 < k_i - 1$ let

$$b_j^{(k_j)} = K_0 u_0 + \ldots + K_{k_j-1} u_{\underset{k_j-1}{\cancel{j-1}}} + K_{k_j}(u_{1,k_j}, \ldots, u_{j-1,k_j}, 0, \ldots, 0)^{\mathrm{T}}$$

with $u_{i,j} = 0$ if $b_i^{(j)}$ is not a column of H and $K_i := (DI_n - A)^i(B)$. Using (α) we conclude $g_{s_i} b_j^{(k_j)} = 0$. Proceeding in this way we obtain $g_{s_i} b_j^{(\nu)} = 0$ for $\nu = 0, \ldots, k_i - 1$ and $k_j > 0$. If $k_j = 0$ let $b_j = \sum_{\lambda=1}^{j-1} b_\lambda u_\lambda$ with $u_\lambda = 0$ if $k_\lambda = 0$. Then $g_{s_i} b_j^{(\nu)} = g_{s_i} \sum_{\lambda=1}^{j-1} \sum_{\mu=0}^{\nu} \binom{\nu}{\mu} b_\lambda^{(\nu-\mu)} u_\lambda^{(\mu)}$ and the foregoing implies that $g_{s_i} b_j^{(\nu)} = 0$ for $\nu = 0, \ldots, k_i - 1$.

For $j \in \{1, \ldots, i-1\}$ use of (5.6) yields $g_{s_i}^{(\mu)} b_j^{(\nu)} = 0$ for $\mu + \nu \leq k_i - 1$. This proves the assertion.

(β_2) We show that

$$\begin{bmatrix} g_{s_i} \\ \vdots \\ g_{s_i}^{(k_i - \overset{2}{\cancel{1}})} \end{bmatrix} [b_i, \ldots, b_m] = 0_{(k_i-1) \times (m-i+1)}$$

Let $j \in \{i, \ldots, m\}$. (α) implies $g_{s_i} b_j^{(\nu)} = 0$ for $\nu = 0, \ldots, k_j - 2$ and $k_j > 0$. Using the same arguments as in (β_1) we obtain $g_{s_i}^{(\mu)} \cdot b_j^{(\nu)} = 0$ for $\mu + \nu \leqslant k_i - 2$.

(β_3) It remains to show that $g_{s_i}^{(k_i-1)}[b_i, \ldots, b_m] = (1, *, \ldots, *)(-1)^{k_i-1}$. With ($\alpha$) and ($\beta_2$) one concludes that

$$g_{s_i} b_i^{(\nu)} = \begin{cases} 0 & \text{for } \nu \leqslant k_i - 2 \\ 1 & \text{for } \nu = k_i - 1 \end{cases}$$

(5.6) implies

$$g_{s_i}^{(\mu)} b_i^{(\nu)} = \begin{cases} 0 & \text{for } \mu + \nu \leqslant k_i - 2 \\ (-1)^\mu & \text{for } \mu + \nu = k_i - 1 \end{cases}$$

(γ) It is proved that $H_1 \in \mathrm{GL}_n(\mathscr{M})$. Let $\mathrm{v}^{\mathrm{T}} \in \mathscr{M}^n$ and $\mathrm{v}H_1 = 0$. Using (β) yields: $0 = \mathrm{v}H_1 B = \mathrm{v}B'$. By the special form of B' it follows that $\mathrm{v}_{s_1} = 0, \ldots, \mathrm{v}_{s_m} = 0$. Using the results of ($\beta$) we conclude

$$H_1 \dot{B} = \begin{bmatrix} \bar{B}_1 \\ \vdots \\ \bar{B}_m \end{bmatrix} \quad \text{with } \bar{B}_i = \begin{bmatrix} & & & 0_{k_i-2,m} & & & \\ 0 & \ldots & 0 & 1 & * & \ldots & * \\ * & & & \ldots & & & * \end{bmatrix} \cdot (-1)^{k_i - \overset{2}{\cancel{1}}}$$

$\uparrow$
ith column

and $\bar{B}_i$ is omitted if $k_i = 0$. $\mathrm{v} \cdot H_1 \dot{B} = 0$ implies $\mathrm{v}_{s_1 - 1} = \ldots = \mathrm{v}_{s_m - 1} = 0$. Proceeding in this way we have $\mathrm{v} = 0$, which proves the claim.

(δ) To prove part (i) of Lemma 5.2 define

$$[DI_n - A', -B'] := H_1[DI_n, -B]\begin{bmatrix} H_1^{-1} & 0 \\ 0 & I_m \end{bmatrix}$$

Since $\dot{H}_1^{\mathrm{T}} = H_1^{\mathrm{T}}(H_1^{\mathrm{T}-1})^{\cdot}(-H_1^{\mathrm{T}})$ it follows that for $H_1^{\mathrm{T}} A'^{\mathrm{T}} = \dot{H}_1^{\mathrm{T}}$, A' has the claimed form.

(ε) Part (ii) of Lemma 5.2 is easily proved: One can choose an elementary matrix U' over $\mathscr{M}$ such that in $[DI_n - A', -B']U'$ the '*' of A' are annulled. Multiplying the columns of B' by units and exchanging the columns of the matrix gives the claimed form. □

Note, that in Lemma 5.2 we do not call (A', B') a normal form of (A, B) since T is not necessarily an element of $\mathrm{GL}_n(\mathscr{A})$.

Proposition 5.2

Let $(A, B) \in \mathscr{A}^{n \times (n+m)}$ be controllable. Then

(i) The set of controllability indices of (A, B) coincides with the set of minimal indices of $\mu(DI_n - A, B)$.

(ii) $\dim_{\mathscr{M}[D]} \mu(DI_n - A, B) = m$.

Proof

By analogy to Rosenbrock (1970, pp. 96 and 97) it can be proved that the minimal indices of $\ker [DI_n - A, B]$ are invariant under transformations as considered in Lemma 5.2. Then the proposition is an immediate consequence of Lemma 5.2. □

Now we discuss the relations between the input modules of time-varying and time-invariant systems.

For $[P, -Q] \in \mathscr{M}[D]^{n \times (n+m)}$ let

$$\mu_{\mathbb{R}[D]}(P, Q) := \{u \in \mathbb{R}[D]^m \mid \exists\, z \in \mathbb{R}[D]^n : Pz = Qu\}$$

Consider a time-invariant controllable state-space system $(A, B) \in \mathbb{R}^{n \times (n+m)}$. Then Proposition 5.2 and Theorem 1.1 in Rosenbrock (1970, p. 96) yields:

$$\dim_{\mathscr{M}[D]} \mu(DI_n - A, B) = m = \dim_{\mathbb{R}[D]} \mu(DI_n - A, B)$$

Proposition 4.2 implies that for every minimal basis $\tilde{U}$ of $\mu(DI_n - A, B)$ there exists a transformation matrix $T \in \mathscr{T}_m$ such that $\tilde{U} = UT$ and $U \in \mathbb{R}[D]^{m \times m}$ with constant coefficients. For an arbitrary controllable system $(A, B) \in \mathscr{A}^{n \times (n+m)}$ the module $\mu(DI_n - A, B)$ does not necessarily possess a minimal basis over $\mathbb{R}[D]$. Consider for example

$$U := \begin{bmatrix} 1 + D(t-1) + D^2(t-2) & 1 + Dt & 1 \\ 0 & 1 + Dt & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

which is a minimal basis of $\mu(DI_3, B)$, where B is defined as in Example 3.1. It can easily be concluded that U cannot be transformed to a minimal basis over $\mathbb{R}[D]$.

Example 5.1

We illustrate how (A, B) as given in Example 3.1 is transformed to (A', B') as given in Lemma 5.2. We only calculate the matrix B'; A' is left to the reader.

$$H^{-1} = (t-2)^{-1} \begin{bmatrix} -\exp(-t) & 1 & -2t^{-1} \\ \exp(-t)(t-1) & -1 & 1 \\ 0 & 0 & (t-2)t^{-1} \end{bmatrix}$$

Since $s_1 = 2$ and $s_2 = 3$, it is calculated that H_1 as defined in (α) of the proof of Lemma 5.2 is of the form

$$H_1 = \begin{bmatrix} \dfrac{(t-1)}{t-2}\exp(-t) & -(t-2)^{-1} & (t-2)^{-1} \\ \dfrac{-t^2+3t-3}{(t-2)^2}\exp(-t) & (t-2)^{-2} & -(t-2)^{-2} \\ 0 & 0 & t^{-1} \end{bmatrix}$$

and

$$B' = H_1 B = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

In the remainder of this section we construct a minimal basis of the input module of a controllable state-space system. (Cf. Kalman (1971) for the time-invariant case.) Let $(A, B) \in \mathscr{A}^{n\times(n+m)}$ be controllable and use the notation as given in (3.1), (3.2) and (3.4). For $i \in \{1, \ldots, m\}$ there exist unique elements $\lambda_{i,j}$ of $\mathscr{M}$ such that

$$\phi(\phi^{-1}b_i)^{(k_i)} = B(\lambda_{1,0}, \ldots, \lambda_{m,0})^{\mathrm{T}} + \cdots + [(DI_n - A)^{k_i-1}(B)](\lambda_{1,k_i-1}, \ldots, \lambda_{m,k_i-1})^{\mathrm{T}} + [\phi(\phi^{-1}b_1)^{(k_i)}, \ldots, \phi(\phi^{-1}b_{i-1})^{(k_i)}](\lambda_{1,k_i}, \ldots, \lambda_{i-1,k_i})^{\mathrm{T}}$$

with

$$\lambda_{i,j} = 0 \quad \text{if } \phi(\phi^{-1}b_i)^{(j)} \notin H$$

For $i = 1, \ldots, m$ define

$$u_i = \sum_{j=0}^{k_i-1} D^j(-1)^j(\lambda_{1,j}, \ldots, \lambda_{m,j})^{\mathrm{T}} + D^{k_i}(-1)^{k_i}(\lambda_{1,k_i}, \ldots, \lambda_{i-1,k_i}, -1, 0, \ldots, 0)^{\mathrm{T}} \tag{5.7}$$

Using this notation we obtain the following.

Proposition 5.3

For a controllable $(A, B) \in \mathscr{A}^{n\times(n+m)}$ the matrix $U := [u_1, \ldots, u_m] \in \mathscr{M}[D]^{m\times m}$ with u_i as defined in (5.7) is a minimal basis of $\mu(DI_n - A, B)$.

Proof

Let $V = [v_1, \ldots, v_m]$ be a minimal basis of $\mu(DI_n - A, B)$. Assume without restriction of generality that $k_1 \geqslant \ldots \geqslant k_m$. Since the minimal and the controllability indices coincide, let $\deg v_i = k_i$. By construction, we have $u_i \in \ker \hat{K}_{A,B} = \mu(DI_n - A, B)$.

Thus there exists a $T\in\mathcal{M}[D]^{m\times m}$ such that $U=VT$. By Proposition 4.2 it remains to show that $T\in\mathcal{T}_{\mu(DI_n-A,B)}$. By Proposition 4.1 (iii) and construction of u_i it follows that

$$k_i=\deg u_i=\max\{\deg t_{ji}+\deg v_j|j:t_{ji}\neq 0\}\quad\text{for } i=1,\ldots,m$$

This implies $\deg t_{ji}\leqslant k_i-k_j$ for $k_j\leqslant k_i$ and $t_{ji}=0$ for $k_j>k_i$. Since

$$[U]_l=\begin{bmatrix}(-1)^{k_1+1} & & \\ & \ddots & * \\ 0 & & (-1)^{k_m+1}\end{bmatrix}$$

the columns of U are linearly independent over $\mathcal{M}[D]$, see proof '(i)$\Rightarrow$(ii)' of Proposition 4.1. Therefore T is non-singular. Due to the special structure of T, its diagonal blocks are non-singular over $\mathcal{M}$. Thus $T\in\mathrm{GL}_m(\mathcal{M}[D])$. □

Example 5.2

We compute a minimal basis for the input module of the system considered in Example 3.1. Using the procedure given in Proposition 5.3 one obtains

$$b_1^{(2)}=B(1,0,0)^{\mathrm{T}}+(\dot{B}-AB)(1-t,0,0)^{\mathrm{T}}\cdot(2-t)^{-1}$$

and

$$u_1=(1+D(t-1)+D^2(t-2),0,0)^{\mathrm{T}}(2-t)^{-1}\in\mu(DI_3,B)$$

Furthermore

$$b_2^{(1)}=B(1,1,0)^{\mathrm{T}}+(\dot{B}-AB)(-t,0,0)^{\mathrm{T}}\cdot t^{-1}$$

and

$$u_2=(1+Dt,1+Dt,0)^{\mathrm{T}}t^{-1}$$

The matrix

$$V=\begin{bmatrix}0 & 0 & D\exp(t)(2-t)+\exp(t)(t-3)\\ 0 & -t^2 & D(1-t)(t-2)+2(t-2)-(1-t)^2\\ 0 & -t^2 & 0\\ 1 & 1+Dt & 1+D(t-1)+D^2(t-2)\\ 1 & 1+Dt & 0\\ -1 & 0 & 0\end{bmatrix}$$

is a minimal basis of $\ker[DI_3,-B]$. This is true by construction and in addition because of Proposition 4.1 since

$$\mathrm{rk}\,[V]_l=\mathrm{rk}\begin{bmatrix} & 0_{3\times 3} & \\ 1 & t & t-2\\ 1 & t & 0\\ -1 & 0 & 0\end{bmatrix}=3$$

At least

$$\mu(DI_3, B) = \begin{bmatrix} 1 & 1+Dt & 1+D(t-1)+D^2(t-2) \\ 1 & 1+Dt & 0 \\ -1 & 0 & 0 \end{bmatrix} \cdot \mathcal{M}[D]^3$$

with minimal indices (0, 1, 2).

6. Geometric indices and controllability

Following Brunovksý (1970, p. 179) we introduce a third class of indices.

Let $(A, B) \in \mathcal{A}^{n\times(n+m)}$ and consider $K^l(A, B)$ as in (3.1) for fixed $t \in \mathbb{R}$, i.e. $K^l(A(t), B(t)) \in \mathbb{R}^{n\times(l+1)m}$. Define

$$r_i(t) := \mathrm{rk}_{\mathbb{R}}\, K^i(A(t), B(t)) - \mathrm{rk}_{\mathbb{R}}\, K^{i-1}(A(t), B(t)) \tag{6.1}$$

for $i = 0, \ldots, n-1$ and $rk_{\mathbb{R}}\, K^{-1}(A(t), B(t)) := 0$.

Let $\alpha_i(t)$ be the number of $r_j(t)$s which are greater or equal to i, i.e.

$$\alpha_i(t) := \sum_{j: r_j(t) \geq i} 1 \quad \text{for } i = 1, \ldots, m \tag{6.2}$$

Since (A, B) is analytic, it follows that $r_i(t) = \text{const}$ for $t \in \mathbb{R}\backslash N$, where N is a discrete set, and $\alpha_i(t) = \text{const}$ for $t \in \mathbb{R}\backslash M$, $M \subseteq N$. Then on $\mathbb{R}\backslash N$ we have

$$0 \leq r_{n-1}(t) \leq \ldots \leq r_0(t) \leq \mathrm{rk}_{\mathbb{R}}\, B(t) \leq m$$

and

$$0 \leq \alpha_m(t) \leq \ldots \leq \alpha_1(t) \leq n$$

The functions $\alpha_1, \ldots, \alpha_m$ are called the *geometric indices* of (A, B).

The following example demonstrates that the information on (A, B) contained in the $r_i(t)$ may be lost if we consider $\alpha_i(t)$.

Example 6.1

Consider the system given in Example 3.1 (resp. Example 5.2). Then

$$r_0(t) = \mathrm{rk}_{\mathbb{R}}\, B(t) = \begin{cases} 1 & \text{for } t = 0 \\ 2 & \text{for } t \neq 0 \end{cases}$$

$$r_1(t) = \mathrm{rk}_{\mathbb{R}}\, [B(t), \dot{B}(t)] - \mathrm{rk}_{\mathbb{R}}\, B(t) = 3 - r_0(t) = \begin{cases} 2 & \text{for } t = 0 \\ 1 & \text{for } t \neq 0 \end{cases}$$

$$r_2(t) = \mathrm{rk}_{\mathbb{R}}\, K(A(t), B(t)) - \mathrm{rk}_{\mathbb{R}}\, [B(t), \dot{B}(t)] = 0$$

and $(\alpha_1(t), \alpha_2(t), \alpha_3(t)) = (2, 1, 0)$.

Proposition 6.1

For $(A, B) \in \mathcal{A}^{n\times(n+m)}$ the set of geometric indices of (A, B) coincide with the set of minimal indices of $\mu(DI_n - A, B)$ on $\mathbb{R}\backslash N$, where N is a discrete set.

Proof

The proof follows the proof of Münzner and Prätzel-Wolters (1979, p. 298) for the time-invariant case. Let $V_d := \{v \in \mathscr{M}^m[D] \mid \deg v \leqslant d\}$, then

$$\hat{K}_{A,B}(V_d) = \{K^d(A, B)[u_0, \ldots, u_d]^{\mathrm{T}} \mid u_0, \ldots, u_d \in \mathscr{M}^m\}$$

The map

$$f: \hat{K}_{A,B}(V_d) \to V_d / V_d \cap \ker \hat{K}_{A,B}$$

$$K^d[u_0, \ldots, u_d]^{\mathrm{T}} \mapsto u_0 + \ldots + D^d u_d + (V_d \cap \ker \hat{K}_{A,B})$$

is a $\mathscr{M}$-right homomorphism.

Let $m_d := V_d \cap \ker \hat{K}_{A,B}$. Then

$$\dim_{\mathscr{M}} \hat{K}_{A,B}(V_d) = m(d+1) - \dim_{\mathscr{M}} m_d \tag{6.3}$$

From now on we consider the system on a non-void open interval where the $r_i(t)$s ($i = 0, \ldots, n-1$) as defined in (6.1) are constant. Let $h(d)$ denote the number of indices of $\mu(DI_n - A, B)$ equal to $d \in \mathbb{N}$. Using the proof '(iv)$\Rightarrow$(i)' in Proposition 4.1 we have:

$$h(d) = \dim_{\mathscr{M}} m_d + \dim_{\mathscr{M}} m_{d-2} - 2 \dim_{\mathscr{M}} m_{d-1} \tag{6.4}$$

It remains to show that $h(d) = k(d)$ for $d \in \mathbb{N}$, where $k(d) := \sum_{l:\alpha_l = d} 1$ and $\alpha_1, \ldots, \alpha_m$ denote the set of geometric indices of (A, B). Since $\sum_{d:d \geqslant l} k(d) = \sum_{i:\alpha_i \geqslant l} 1 = \sum_{i:1 \leqslant i \leqslant r_{l-1}} 1 = r_{l-1}$ it follows that

$$k(d) = r_{d-1} - r_d = 2\,\mathrm{rk}_{\mathscr{M}} K^{d-1}(A, B) - \mathrm{rk}_{\mathscr{M}} K^{d-2}(A, B) - \mathrm{rk}_{\mathscr{M}} K^d(A, B)$$

The equations $\mathrm{rk}_{\mathscr{M}} K^d(A, B) = \dim_{\mathscr{M}} \hat{K}_{A,B}(V_{d-1})$, (6.3) and (6.4) yield

$$\begin{aligned} k(d) &= 2 \dim_{\mathscr{M}} \hat{K}_{A,B}(V_{d-1}) - \dim_{\mathscr{M}} \hat{K}_{A,B}(V_{d-2}) - \dim_{\mathscr{M}} \hat{K}_{A,B}(V_d) \\ &= \dim_{\mathscr{M}} m_d + \dim_{\mathscr{M}} m_{d-2} - 2 \dim_{\mathscr{M}} m_{d-1} = h(d) \end{aligned}$$ □

Now we are able to characterize controllability by all of the indices discussed before.

Proposition 6.2

For $(A, B) \in \mathscr{A}^{n \times (n+m)}$ the following are equivalent:

(i) (A, B) is controllable.

(ii) The sum of the controllability indices of (A, B) is n.

(iii) The sum of the minimal indices of $\mu(DI_n - A, B)$ is n.

(iv) There exists a non-void open interval $I \subseteq \mathbb{R}$ such that the sum of geometric indices of (A, B) on I is n.

Proof

For '(i)$\Leftrightarrow$(ii)$\Rightarrow$(iii)$\Rightarrow$(iv)' see Propositions 3.1, 5.2 and 6.1. We prove '(iv)$\Rightarrow$(i)'. Let $\sum_{i=1}^{m} \alpha_i(t) = n$ on I. Since $\sum_{i=1}^{m} \alpha_i(t) = \sum_{j=0}^{n-1} \sum_{i:i \leqslant r_j(t)} 1 = \sum_{j=0}^{n-1} r_j(t) = \mathrm{rk}_{\mathbb{R}} K(A(t), B(t))$ on I, it follows from Proposition 3.1 that (A, B) is controllable. □

In Ilchmann *et al.* (1984, Definition 6.1) controllability for systems in differential operator representation is defined as follows.

Definition 6.1

$\mathbb{P}=\begin{bmatrix} P & -Q \\ V & W \end{bmatrix}\in \mathcal{M}[D]^{(n+p)\times(n+m)}$ is called *controllable* on $[t_0, t_1]$ if for any $z^0 \in \ker P$ there exists a control $u \in \mathcal{U}^m$ such that

$$(z^0 + z_u)(t) = \begin{cases} z^0(t) & \text{for } t \leqslant t_0 \\ 0 & \text{for } t \geqslant t_1 \end{cases}$$

where z_u denotes the forced motion starting from zero under control u, see (2.4).

Since controllability (cf. Ilchmann *et al.* (1984), Remark 6.2) and the input module is invariant under system equivalence (see Lemma 5.1) one can use a state-space representation for the system matrices (cf. Ilchmann *et al.* (1984), Proposition 5.7) and the foregoing Proposition 6.2 to prove the following.

Proposition 6.3

$\mathbb{P}=\begin{bmatrix} P & -Q \\ V & W \end{bmatrix}\in \mathcal{M}[D]^{(r+p)\times(r+m)}$ is controllable on an interval $I \in \mathbb{R}$ if and only if the sum of minimal indices of $\mu(P, Q)$ coincides with $\dim \ker_{\mathcal{A}} P$.

This proposition also shows that the input module is an appropriate tool to generalize invariant indices for time-varying systems in differential operator representation.

In the following it will be explained how different definitions of controllability in the case of state-space systems are related.

Definition 6.2

A system $(A, B) \in \mathcal{A}^{n\times(n+m)}$ is said to be *uniformly controllable* if $\mathrm{rk}_{\mathbb{R}}\, K(A(t), B(t)) = n$ for every $t \in \mathbb{R}$.

While in the single-input case this condition is equivalent to controllability and constant $r_i(t)$s, for multi-input systems uniform controllability is not sufficient to guarantee the existence of a normal form. Consider, for example, the system

$$\left(0, \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}\right) \in \mathcal{A}^{2\times 4}$$

and assume that there exists a $T \in \mathcal{M}^{2\times 2}$ such that

$$T\begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

with $b \in \mathcal{A}$. Then this equality implies

$$T = \begin{bmatrix} t^{-1} & c \\ 0 & t^{-1} \end{bmatrix} \notin \mathrm{GL}_2(\mathcal{A}), \quad c \in \mathcal{M}$$

Remark 6.1

For $(A, B) \in \mathscr{A}^{n \times (n+m)}$ let $H \in \mathscr{A}^{n \times n'}$ be given as in (3.4). The value of H at $t \in \mathbb{R}$ will be denoted by $H(t)$. Then the following implications hold due to the construction of H:

$$\begin{aligned} &\mathrm{rk}_{\mathbb{R}}\, H(t) = \mathrm{rk}_{\mathscr{M}} H \quad \text{for every } t \in \mathbb{R} \\ &\qquad \Rightarrow r_i(t) = \text{const} \quad \text{for } i = 0, \ldots, n-1 \\ &\qquad \Rightarrow \mathrm{rk}_{\mathbb{R}}\, K(A(t), B(t)) = \sum_{i=1}^{n-1} r_i = \sum_{i=1}^{m} \alpha_i = \mathrm{rk}_{\mathscr{M}} H \quad \text{for every } t \in \mathbb{R} \end{aligned}$$

In general the inverse conclusions are false. To see this, consider $(A, B) = \left(0, \begin{bmatrix} t & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\right) \in \mathscr{A}^{2 \times 5}$. Then $r_0(t) = \mathrm{rk}_{\mathbb{R}}\, B(t) = 2$ and $r_1(t) = 0$, but $H = \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}$

For the inversion of the second implication consider $(A, B) = \left(0, \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}\right) \in \mathscr{A}^{2 \times 4}$. Then

$$\mathrm{rk}_{\mathbb{R}}\, K(A(t), B(t)) = \mathrm{rk}_{\mathbb{R}} \begin{bmatrix} t & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = 2 = \mathrm{rk}_{\mathscr{M}} \begin{bmatrix} t & 0 \\ 0 & 1 \end{bmatrix}$$

and, in addition, (A, B) is uniformly controllable. But

$$r_0(t) = \mathrm{rk}_{\mathbb{R}}\, B(t) = \begin{cases} 2 & \text{for } t \neq 0 \\ 1 & \text{for } t = 0 \end{cases}$$

7. Formal transfer matrix

In Ilchmann *et al.* (1984, Definition 7.2) the *input–output map* G of a system $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathscr{M}[D]^{(n+p) \times (n+m)}$ is introduced:

$$\left.\begin{aligned} G: \mathscr{U}^m &\to \mathscr{U}^p \\ u &\mapsto V(z_u) + W(u) \end{aligned}\right\} \tag{7.1}$$

where z_u is the forced motion starting from zero under control u, see (2.4).

In the following we introduce the formal transfer matrix $\hat{G} = VP^{-1}Q + W$ and analyse its connection with G. For this, let

$$\mathscr{M}(D) := \{p^{-1}q \mid p \in \mathscr{M}[D]^*, q \in \mathscr{M}[D]\}$$

denote the left-skew field of fractions of $\mathscr{M}[D]$. This field is constructed as follows (cf. Cohn (1971), p. 20): For pairs $(p, q) \in \mathscr{M}[D]^* \times \mathscr{M}[D]$ we define an equivalence relation between them by the condition: $(p_1, q_1) = (p_2, q_2)$ iff there exist $u_1, u_2 \in \mathscr{M}[D]^*$ such that

$$u_2 p_1 = u_1 p_2 \quad \text{and} \quad u_2 q_1 = u_1 q_2$$

The equivalence class containing a pair (p, q) is denoted by $p^{-1}q$. The multiplication

$$p_1^{-1}q_1 \cdot p_2^{-1}q_2 = (u_2 p_1)^{-1}(u_1 q_2) \quad \text{with } u_1, u_2 \in \mathscr{M}[D]^* \text{ such that } u_1 p_2 = u_2 q_1$$

depends only on the equivalence classes of the factors and is associative.

For $P \in \mathscr{M}(D)^{n \times m}$ it can be proved (analogous to the commutative case) that there exist $U \in \mathrm{GL}_n(\mathcal{M}(D))$ and $U' \in \mathrm{GL}_m(\mathcal{M}(D))$ such that

$$P = U \operatorname{diag}(1, \ldots, \underset{\substack{\uparrow \\ r\text{th element}}}{1}, 0, \ldots, 0)U'$$

The number r is called the *rank of* P, $\operatorname{rk} P = r$, and it coincides with the maximal number of right linear independent columns or left independent rows (over $\mathcal{M}(D)$) of P. Furthermore for $n \leqslant m$ $(n \geqslant m)$ we have

$$P \text{ is right- (left-) invertible iff } \operatorname{rk} P = n\ (=m)$$

Definition 7.1

Let $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)}$. Then $\hat{G} = VP^{-1}Q + W \in \mathcal{M}(D)^{p\times m}$ is called the *formal transfer matrix* of $\mathbb{P}$.

$\hat{G}$ can be associated with an operator acting on $\mathcal{U}^m$ in the following way (see (2.4))

$$\hat{G} : \mathcal{U}^m \to \mathcal{U}^p$$
$$u \mapsto (VP^{-1}Q + W)(u)$$
$$\text{with } (P^{-1}Q)(u) := z_u \quad \text{and} \quad (z_u, u)^{\mathrm{T}} \in M_+(P, Q)$$

Therefore $G(u) = \hat{G}(u)$ for every $u \in \mathcal{U}^m$.

If $\mathbb{P} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)}$ is an analytic state-space system with constant free motion it follows that

$$[C \quad (DI_n)^{-1}B + E(D)](u(t)) = C(t)\int_{-\infty}^{t} B(s)u(s)\, ds + E(D)(u(t))$$

for every $u \in \mathcal{U}^m$.

Unfortunately the formal transfer matrices of time-varying systems do not form an $\mathbb{R}$-algebra with respect to the usual multiplication of matrices over a skew field: If $P \in \mathcal{M}[D]^{n\times n}$ and $Q \in \mathcal{M}[D]^{r\times m}$, then $P^{-1}Q \in \mathcal{M}(D)^{n\times m}$ can be interpreted as an operator on $\mathcal{U}^m$ only if P is full and $\operatorname{im} Q \subseteq \operatorname{im} P$. But the set $\{P^{-1}Q \in \mathcal{M}(D)^{n\times p} \mid P \in \mathcal{M}[D]^{n\times n}$ full w.r.t. $\mathcal{A}$, $Q \in \mathcal{M}[D]^{n\times p}$ and $\operatorname{im} Q \subseteq \operatorname{im} P\}$ do not form an $\mathbb{R}$-algebra. If $P_1^{-1}Q_1 \cdot P_2^{-1}Q_2 = P^{-1}Q$ and P_1, P_2 are full, in general P is not full.

Lemma 7.1

$$\left\{ VP^{-1}Q + W \in \mathcal{M}(D)^{p\times m} \,\middle\|\, \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)} \text{ is a system matrix, } n \in \mathbb{N} \right\}$$

is an $\mathbb{R}$-algebra with respect to the following multiplication and addition:

$$(V_1P_1^{-1}Q_1 + W_1) \odot (V_2P_2^{-1}Q_2 + W_2) = [0, V_2, W_2P^{-1}] \begin{bmatrix} -Q \\ 0 \\ W_1 \end{bmatrix},$$

$$P := \begin{bmatrix} P_1 & 0 & 0 \\ 0 & P_2 & -Q_2 \\ V_1 & 0 & -I_k \end{bmatrix}$$

$$(V_1P_1^{-1}Q_1+W_1)\oplus(V_2P_2^{-1}Q_2+W_2)=[-V_1, V_2]\begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}^{-1}\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}+W_1+W_2$$

These operations correspond to series and parallel connections of systems, see Rosenbrock (1970, p. 125).

Proof

We prove that P is full w.r.t. $\mathscr{A}$. The state-space representation for system matrices (see Ilchmann *et al.* (1984), p. 353) yields

$$\begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \overset{\text{se}}{\sim} \begin{bmatrix} DI_{n_i} & -B_i \\ C_i & E_i(D) \end{bmatrix} \in \mathscr{A}[D]^{(n_i+p)\times(n_i+m)}, \quad n_i=\dim\ker_{\mathscr{A}} P_i, \quad i=1,2$$

Using this representation and the fact that the full matrix

$$\begin{bmatrix} DI_{n_i} & 0 & 0 \\ -B_2C_1 & DI_{n_2} & 0 \\ C_1 & 0 & -I_k \end{bmatrix} \text{ is equivalent to } \begin{bmatrix} DI_{n_1} & 0 & 0 \\ 0 & DI_{n_2} & -B_2 \\ C_1 & 0 & -I_k \end{bmatrix} := P'$$

we conclude

$$\left[\begin{array}{c|c} & -Q \\ P & 0 \\ & W_1 \\ \hline 0, V_2, W_2 & 0 \end{array}\right] \overset{\text{se}}{\sim} \left[\begin{array}{c|c} & -B_1 \\ P' & 0 \\ & E_1(D) \\ \hline 0, C_2, E_2(D) & 0 \end{array}\right]$$

The property full is preserved under system equivalence. Thus P is full. □

Lemma 7.2

Both the formal transfer matrix and the input–output map of a system matrix are invariant under system equivalence.

Proof

For the latter, see Ilchmann *et al.* (1984, Proposition 7.3 (*a*)). We give a proof of the first statement. Let

$$\mathbb{P}_i=\begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix}\in[D]^{(n_i+p)\times(n_i+m)}, \quad i=1,2 \quad \text{and} \quad \mathbb{P}_1 \overset{\text{se}}{\sim} \mathbb{P}_2$$

With the notation of (2.2) we obtain

$$XP_1+V_1=V_2T_1, \quad -XQ_1+W_1=V_2Y+W_2, \quad TP_1=P_2T_1, \quad -TQ_1=P_2Y-Q_2$$

Using the equations in this sequence we conclude

$$\begin{aligned} V_1P_1^{-1}Q_1+W_1 &= (V_2T_1P_1^{-1}-X)Q_1+W_1 = V_2Y+V_2T_1P_1^{-1}Q_1+W_2 \\ &= V_2P_2^{-1}(P_2Y+TQ_1)+W_2 = V_2P_2^{-1}Q_2+W_2 \end{aligned}$$ □

The following lemma is used to give a proof of the main result of this section.

Lemma 7.3

Let $\mathbb{P} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathscr{A}[D]^{(n+p)\times(n+m)}$ with input–output map G given. Then there exist $k(D) \in \mathscr{M}[D]^*$ and $N(D) \in \mathscr{M}[D]^{p\times m}$ such that $(k(D)I_p \circ G)(u) = N(D)(u)$ for every $u \in \mathscr{U}^m$.

Proof

(i) For $c, b \in \mathscr{A}$ we prove: $(c \circ D^{-1}b)(u) = \left(\left(D - \frac{\dot{c}}{c}\right)^{-1} cb\right)(u)$ for every $u \in \mathscr{U}$. Multiplication in $\mathscr{M}(D)$ implies $cD^{-1} = \left(D - \frac{\dot{c}}{c}\right)^{-1} c$. Let $D^{-1}b(u) = \bar{z}_u$ with $(\bar{z}_u, u) \in M_+(D, b)$ and $\left(\left(D - \frac{\dot{c}}{c}\right)^{-1} cb\right)(u) = z_u$ with $(z_u, u) \in M_+\left(D - \frac{\dot{c}}{c}, cb\right)$. It remains to show that $c\bar{z}_u = z_u$. Since the kernel of $\left(D - \frac{\dot{c}}{c}\right)$ consists only of analytic functions (see Ilchmann *et al.* (1984), Lemma 2.5) and z_u and $\bar{z}_u$ have bounded support, the proof is achieved if $\left(D - \frac{\dot{c}}{c}\right)(c\bar{z}_u - z_u) = 0$. The latter follows from $\left(D - \frac{\dot{c}}{c}\right)(c\bar{z}_u) = (Dc - \dot{c})(\bar{z}_u) = cD(\bar{z}_u) = cbu = \left(D - \frac{\dot{c}}{c}\right)(z_u)$.

(ii) Let $((c_{ij})) = C$ and $k = k_{ij}\left(D - \frac{\dot{c}_{ij}}{c_{ij}}\right) \in \mathscr{A}[D]$ be a least common left multiple of $\left\{D - \frac{\dot{c}_{ij}}{c_{ij}} \middle| c_{ij} \neq 0, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant m\right\}$.

Applying (i) we conclude

$$\begin{aligned} (kI_p \circ G)(u) &= (kI_p \circ ((c_{ij})) \circ D^{-1}B + kI_p \circ E(D))(u) \\ &= \left(kI_p \circ \left(\left(D - \frac{\dot{c}_{ij}}{c_{ij}}\right)^{-1} c_{ij}\right)B + kI_p E(D)\right)(u) \\ &:= N(D)(u) \end{aligned}$$

□

Proposition 7.1

Let $\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathscr{M}[D]^{(n_i+p)\times(n_i+m)}$ with input–output map G_i and formal transfer matrix $\hat{G}_i$ be given, $i = 1, 2$. Then $\hat{G}_1 = \hat{G}_2$ iff $G_1(u) = G_2(u)$ for every $u \in \mathscr{U}^m$.

Proof

Because of Lemma 7.2 and the state-space representation for system matrices (see Ilchmann *et al.* (1984), Proposition 5.7) we may assume without restriction of generality that

$$\mathbb{P}_\lambda = \begin{bmatrix} DI_{n\lambda} & -B_\lambda \\ C_\lambda & E_\lambda(D) \end{bmatrix} \in \mathscr{A}[D]^{(n_\lambda+p)\times(n_\lambda+m)}, \quad \lambda = 1, 2$$

Let $((c_{ij}^{\lambda})) = C_\lambda$ for $\lambda = 1, 2$ and $k(D)$ be a least common left multiple of

$$\left\{ \left(D - \frac{\dot{c}_{ij}^{\lambda}}{c_{ij}^{\lambda}} \right) \middle| c_{ij}^{\lambda} \neq 0, 1 \leqslant i \leqslant p, 1 \leqslant j \leqslant n, \lambda = 1, 2 \right\}$$

$N_\lambda(D) := k(D) I_p (C_\lambda D^{-1} I_{n_\lambda} B_\lambda + E_\lambda(D)) \in \mathscr{M}[D]^{p \times m}$ for $\lambda = 1, 2$. Use of Lemma 7.3 implies $(k(D) \circ (G_1 - G_2))(u) = (N_1(D) - N_2(D))(u)$. From Ilchmann *et al.* (1984, Appendix) the following implication is known for arbitrary $N(D) \in \mathscr{M}[D]^{p \times m}$: If $N(D)(u) = 0$ for every $u \in \mathscr{U}^m$ then $N(D) = 0$. This completes the proof. □

8. Dynamical indices

The introduction of the formal transfer matrix in § 7 enables us to define a fourth class of indices.

Definition 8.1

Let $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathscr{M}[D]^{(n+p) \times (n+m)}$ and $\hat{G} = VP^{-1}Q + W$. The minimal indices of the $\mathscr{M}[D]$-right module

$$m_{\hat{G}} := \{ u \in \mathscr{M}[D]^m \mid \hat{G}u \in \mathscr{M}[D]^p \}$$

are called the *dynamical indices* of $\hat{G}$.

Forney (1975, Chap. 7) considers proper rational input–output maps $G(s): \mathbb{R}(s)^m \to \mathbb{R}(s)^p$ and the minimal indices of the rational vector space $\{(u(s), G(s)u(s))^{\mathrm{T}} \mid u(s) \in \mathbb{R}(s)^m\}$ which he calls the 'dynamical indices' of $G(s)$. Münzner and Prätzel-Wolters (1979) show that these indices coincide with those of the module $\{u(s) \in \mathbb{R}[s]^m \mid G(s)u(s) \in \mathbb{R}[s]^p\}$. Forney (1975) proves that the set of dynamical indices of $G(s)$ and the set of controllability indices of a certain realization of $G(s)$ coincide. This result generalizes as follows.

Definition 8.2

$\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in [D]^{(n+p) \times (n+m)}$ is called *observable* if V acts as a monomorphism on $\ker_{\mathscr{A}} P$.

It can be shown that $\mathbb{P}$ is observable if and only if P and V are right coprime, i.e. there exist $R \in \mathscr{M}[D]^{n \times n}$ and $S \in \mathscr{M}[D]^{n \times p}$ such that $RP + SV = I_n$ (see Ilchmann *et al.* (1984) § 6).

Proposition 8.1

Let $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathscr{M}[D]^{(n+p) \times (n+m)}$ be observable. Then $m_{\hat{G}} = \mu(P, Q)$.

Proof

We use the notation as above.

'$\subseteq$': Let $u \in m_{\hat{G}}$. Then

$$z := P^{-1}Qu = (RP + SV)P^{-1}Qu = RQu + S(\hat{G} - W)u \in \mathscr{M}[D]^n.$$

'$\supseteq$': For $u \in \mathcal{M}[D]^m$ there exists $z \in \mathcal{M}[D]^n$ such that $Pz = Qu$. Then

$$\hat{G}u = Vz + Wu \in \mathcal{M}[D]^p. \qquad \square$$

Corollary 8.1

Let $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)}$ be observable and $\hat{G} = VP^{-1}Q + W$. Then the set of dynamical indices of $\hat{G}$ and the set of minimal indices of $\mu(P, Q)$ coincide.

9 Characterization of system equivalence and a complete set of invariants

The analysis of the input module enables us to give a characterization of system equivalent controllable time-varying systems. These results are known for time-invariant state-space systems, cf. Popov (1972).

Proposition 9.1

Let $\mathbb{P} = \begin{bmatrix} P & -Q \\ V & W \end{bmatrix} \in \mathcal{M}[D]^{(r+p)\times(r+m)}$ and $\mathbb{P}' = \begin{bmatrix} P' & -Q' \\ V' & W' \end{bmatrix} \in \mathcal{M}[D]^{(r'+p)\times(r'+m)}$

be both controllable. Then

$$\mathbb{P} \overset{se}{\approx} \mathbb{P}' \quad \text{iff} \quad \mu(P, Q) = \mu(P', Q') \quad \text{and} \quad VP^{-1}Q \overset{W\,=}{+} V'P'^{-1}Q' + W'$$

Proof

For every system matrix there exists a system-equivalent state-space representation (see Ilchmann *et al.* (1984), Proposition 5.7). Furthermore controllability, input module and formal transfer matrix are invariant under system equivalence. Therefore we assume without restriction of generality that

$$\mathbb{P} = \begin{bmatrix} DI_n & -B \\ C & E(D) \end{bmatrix} \in \mathcal{M}[D]^{(n+p)\times(n+m)} \quad \text{and} \quad \mathbb{P}' = \begin{bmatrix} DI_{n'} & -B' \\ C' & E'(D) \end{bmatrix} \in \mathcal{M}[D]^{(n'+p)\times(n'+m)}$$

'$\Rightarrow$': Immediately by Lemma 5.1.

'$\Leftarrow$':

(i) We prove that $\ker \hat{K}_{0,B} = \ker \hat{K}_{0,B'}$ implies $(0, B) \overset{s}{\sim} (0, B')$. Let $B = [b_1, \dots, b_m]$ and $B' = [b'_1, \dots, b'_m]$ for $i = 1, \dots, m$. k_i (resp. k'_i) denote the controllability indices of $(0, B)$ (resp. $(0, B')$). Then

$$H = [b_1, \dot{b}_1, \dots, b_1^{(k_1-1)}, b_2, \dots, b_m^{(k_m-1)}] \in \mathrm{GL}_n(\mathcal{M})$$

and

$$H' = [b'_1, \dot{b}'_1, \dots, b_1'^{(k'_1-1)}, b'_2, \dots, b_m'^{(k'_m-1)}] \in \mathrm{GL}_{n'}(\mathcal{M})$$

Since $\ker \hat{K}_{0,B} = \ker \hat{K}_{0,B'}$ we have $\{k_1, \dots, k_m\} = \{k'_1, \dots, k'_m\}$. Let

$$u = \sum_{\lambda=0}^{k_i-1} D^\lambda (-1)^\lambda u_\lambda \quad \text{with} \quad u_\lambda = (0, \dots, 0, \underset{\substack{\uparrow \\ i\text{th row}}}{a_\lambda}, 0, \dots, 0)^{\mathrm{T}} \in \mathcal{M}^m$$

for $i \in \{1, \dots, m\}$ and $a_\lambda \neq 0$ for some $\lambda \in \{0, \dots, k_i - 1\}$. Then $u \notin \ker \hat{K}_{0,B}$ and so $u \notin \ker \hat{K}_{0,B'}$, which implies $k_i \leqslant k'_i$. On the other hand $k'_i \leqslant k_i$. Thus $k_i = k'_i$ for $1 \leqslant i \leqslant m$.

Since $\mathbb{P}$ and $\mathbb{P}'$ are controllable it follows that $n = \sum_{i=1}^{m} k_i = \sum_{i=1}^{m} k_i' = n'$. Let $T := H'H^{-1} \in \mathrm{GL}_n(\mathscr{M})$. Then $H' = TH$, i.e. $b_i'^{(j)} = Tb_i^{(j)}$ for $1 \leqslant i \leqslant m$, $j = 0, \ldots, k_i - 1$. Since there exists a unique $u_i \in \mathscr{M}^n$ such that $Hu_i = b_i^{(k_i)}$, the assumption implies $H'u_i = b_i'^{(k_i)}$. Therefore $Tb_i^{(k_i)} = THu_i = H'u_i = b_i'^{(k_i)}$ for $1 \leqslant i \leqslant m$. This yields $(Tb_i^{(j)})^{\cdot} = T(b_i^{(j)})^{\cdot}$ for $1 \leqslant i \leqslant m$, $j = 0, \ldots, k_i - 1$ and $\dot{T}H = (TH)^{\cdot} - T\dot{H} = 0$. In particular $TB = B'$. Since $\dot{T} = 0$ we have proved $(0, B) \overset{s}{\sim} (0, B')$.

(ii) It remains to prove: $\ker \hat{K}_{0,B} = \ker \hat{K}_{0,B'}$ and $C(DI_n)^{-1}B + E(D) = C'(DI_n)^{-1}B' + E'(D)$ imply $\mathbb{P} \overset{se}{\sim} \mathbb{P}'$. By (i) we have $C(DI_n)^{-1}B + E(D) = C'(DI_n)^{-1}TB + E'(D)$. Since $\mathbb{P}$ is controllable there exist $S \in \mathscr{M}[D]^{n \times n}$ and $R \in \mathscr{M}[D]^{m \times n}$ such that $DI_nS + BR = I_n$ (see Ilchmann *et al.* (1984), Theorem 6.4). Therefore $(CD^{-1}B + E(D))R = (C'D^{-1}TB + E'(D))R$ or equivalently $(CD^{-1}BR - C'D^{-1}TBR) = (E'(D) - E(D))R$. Since T is constant we have $D^{-1}T = TD^{-1}$. Thus the following equivalent equations are valid:

$$(C - C'T)D^{-1}(I_n - DS) = (E'(D) - E(D))R$$

$$(C - C'T)(I_n - SD) = (E'(D) - E(D))RD$$

$$(C - C'T) = ((E'(D) - E(D))R + (C - C'T)S)D$$

Comparing the coefficients we conclude $C = C'T$ and $E(D) = E'(D)$. □

If no outputs are considered the following corollary clarifies the relation between the input module and the solution vector space of a system, see (2.3).

Corollary 9.1

Let $\mathbb{P}_i = [P_i, Q_i] \in \mathscr{M}[D]^{n \times (n+m)}$ for $i = 1, 2$ be both controllable. Then the following are equivalent:

(i) $\mathbb{P}_1 \overset{se}{\sim} \mathbb{P}_2$

(ii) $\mu(P_1, Q_1) = \mu(P_2, Q_2)$

(iii) there exists a map

$$f : M(P_1, Q_1) \to M(P_2, Q_2)$$

$$\begin{pmatrix} z \\ u \end{pmatrix} \mapsto \begin{bmatrix} T_1 & Y \\ 0 & I \end{bmatrix} \begin{pmatrix} z \\ u \end{pmatrix}$$

with $T_1 \in \mathrm{GL}_n(\mathscr{M}[D])$, $Y \in \mathscr{M}[D]^{n \times m}$

Proof

Use Ilchmann *et al.* (1984, Proposition 4.2(c)). □

In Ilchmann *et al.* (1984, Proposition 7.3) system equivalence is characterized in terms of the input–output map. Using Proposition 7.1 this can now be carried out in the following form.

Remark 9.1

Let $\mathbb{P}_i = \begin{bmatrix} P_i & -Q_i \\ V_i & W_i \end{bmatrix} \in \mathscr{M}[D]^{(n+p) \times (n+m)}$, $(i = 1, 2)$ be both controllable and observable. Then

$$\mathbb{P}_1 \overset{\text{se}}{\sim} \mathbb{P}_2 \quad \text{iff } V_1 P_1^{-1} Q_1 + W_1 = V_2 P_2^{-1} Q_2 + W_2$$

For controllable state-space systems a complete set of invariants of similarity is given in the following.

Remark 9.2

Let $(A, B) \in \mathscr{A}^{n \times (n+m)}$ be a controllable system with controllability indices $k_1, \ldots, k_m$. We use the notation as in (3.4). Then there exists a unique matrix $\tilde{U} = [\tilde{u}_1, \ldots, \tilde{u}_m] \in \mathscr{M}^{n \times m}$ such that

$$[\phi(\phi^{-1} b_1)^{(k_1)}, \ldots, \phi(\phi^{-1} b_m)^{(k_m)}] = H\tilde{U}$$

Let $\tilde{u}_i = ((u_i^1)_0, \ldots, (u_i^1)_{k_1-1}, \ldots, (u_i^m)_0, \ldots, (u_i^m)_{k_m-1})^{\mathrm{T}} \in \mathscr{M}^n$ where $(u_i^j)_\lambda$ is omitted if $k_j = 0$, $\lambda \in \mathbb{N}$. Then by construction of H we know that necessarily

$$(u_i^j)_\lambda = 0 \quad \text{for } \begin{cases} k_i + 1 \leqslant \lambda \leqslant k_j - 1 \\ \lambda = k_j \text{ and } i < j \end{cases} \tag{9.1}$$

Using this notation we obtain:

Proposition 9.2

Let $(A, B), (A', B') \in \mathscr{A}^{n \times (n+m)}$ both be controllable. Then $(A, B) \overset{s}{\sim} (A', B')$ iff $\tilde{u}_i = \tilde{u}_i'$ and $k_i = k_i'$ for $i = 1, \ldots, m$.

Proof

'$\Rightarrow$': Let (A, B) and (A', B') be similar via $T \in \mathrm{GL}_n(\mathscr{A})$. Lemma 3.1 yields $TH\hat{u}_i = T\phi(\phi^{-1} b_i)^{(k_i)} = \phi'(\phi'^{-1} b_i')^{(k_i)} = H'\tilde{u}_i'$ for $1 \leqslant i \leqslant m$.

'$\Leftarrow$': Without restriction of generality assume $A = A' = 0$. Let $T := H'H^{-1}$. Then $Tb_i^{(k_i)} = TH\tilde{u}_i = H'\tilde{u}_i' = b_i'^{(k_i)}$ for $1 \leqslant i \leqslant m$. Arguing as in (i) of the proof of Proposition 9.1 completes the proof. □

The following example illustrates how to construct a system with a presented list of invariants.

Example 9.1

Let $(k_1, k_2, k_3) = (2, 0, 1)$ and $U = [u_1, u_2, u_3] \in \mathscr{A}^{3 \times 3}$. Because of (9.1) let $\tilde{U}$ be of the following structure

$$\tilde{U} = \begin{bmatrix} a & d & e \\ b & 0 & 0 \\ c & 0 & f \end{bmatrix}.$$

Define

$$H := [e_{10}, e_{11}, e_{30}] := I_3$$

and

$$[e_{1k_1}, e_{2k_2}, e_{3k_3}] := \tilde{U}, K_0 := B := [e_{10}, e_{20}, e_{30}] = \begin{bmatrix} 1 & d & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$K_1 := [e_{11}, e_{21}, e_{31}] = \begin{bmatrix} 0 & & e \\ 1 & e_{21} & 0 \\ 0 & & f \end{bmatrix} = -AK_0 + \dot{K}_0$$

which implies

$$a_1 = -\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, a_3 = -\begin{bmatrix} e \\ 0 \\ f \end{bmatrix}$$

where

$$A := [a_1, a_2, a_3]$$

and

$$e_{21} = -d \cdot a_1 + \begin{bmatrix} d \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} d \\ d \\ 0 \end{bmatrix}$$

$$K_2 := [e_{12}, e_{22}, e_{32}] = \begin{bmatrix} a & & \\ b, e_{22}, e_{32} & & \\ c & & \end{bmatrix} = -AK_1 + \dot{K}_1 = -A\begin{bmatrix} 0 & d & e \\ 1 & d & 0 \\ 0 & 0 & f \end{bmatrix} + \begin{bmatrix} 0 & \ddot{d} & \dot{e} \\ 0 & \dot{d} & 0 \\ 0 & 0 & \dot{f} \end{bmatrix}$$

which implies

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = -a_2$$

So (A, B) with invariants $(2, 0, 1)$ is determined.

In general one obtains:

Proposition 9.3

Let $k_1, \ldots, k_m \in \mathbb{N}$ with $\sum_{i=1}^{m} k_i = n$ be given and $\tilde{U} \in \mathscr{A}^{n\times m}$ which satisfies (9.1). Then there exists a controllable system $(A, B) \in \mathscr{A}^{n\times(n+m)}$ with invariants $k_1, \ldots, k_m$.

Proof

Let $H := I_n := [e_{10}, \ldots, e_{1,k_1-1}, e_{20}, \ldots, e_{m,k_m-1}]$

$$[e_{1,k_1}, \ldots, e_{m,k_m}] := U, K_0 := B := [e_{10}, \ldots, e_{m0}]$$

Now successively the columns of AH are defined:

$$K_1 := [e_{11}, \ldots, e_{m1}] = -AK_0 + \dot{K}_0$$
$$\vdots$$
$$K_d := [e_{1d}, \ldots, e_{md}] = -AK_{d-1} + \dot{K}_{d-1}$$

where $d = \max\{k_i \mid 1 \leqslant i \leqslant m\}$. The so constructed system (A, B) has the invariants $k_1, \dots, k_m$. □

Remark 9.3

It is important to choose $\hat{U}$ in Proposition 9.3 with entries in $\mathscr{A}$, not in $\mathscr{M}$. Otherwise in general for arbitrary given indices $k_1, \dots, k_m \in \mathbb{N}$ there does not exist an analytic system $(A, B) \in \mathscr{A}^{n \times (n+m)}$. Let for example $n = m = k_1 = 1$ and $u_1 = t^{-1}$. Then $H = 1 = K_0 = B$ and $K_1 = t^{-1} = -(-t^{-1}) \cdot 1 + 0$.

Therefore $(A, B) = (-t^{-1}, 1)$. But the system $(D + t^{-1})(z) = u$ is not of interest since $\ker_{\mathscr{M}}(D + t^{-1}) = \{rt^{-1} \mid r \in \mathbb{R}\}$. That means $D + t^{-1}$ is not full w.r.t. $\mathscr{A}$.

ACKNOWLEDGMENT

Very special thanks for helpful suggestions are due to Diederich Hinrichsen (Bremen) and Wiland Schmale (Oldenburg) and to Sue Irwin (Oldenburg) for correcting the English.

REFERENCES

BRUNOVSKÝ, P., 1970, *Kybernet. Čislo*, **3,** 173.
COHN, P. M., 1971, *Free Rings and their Relations* (London and New York: Academic Press).
FORNEY, G. D., JR., 1975, *SIAM J. Control*, **13,** 493.
HINRICHSEN, D., and PRÄTZEL-WOLTERS, D., 1980, *Int. J. Control*, **32,** 777.
ILCHMANN, A., NÜRNBERGER, I., and SCHMALE, W., 1984, *Int. J. Control*, **40,** 329.
KALMAN, R. E., 1962, *Proc. Nat. Acad. Sci.* (U.S.A.), **48,** 596; 1971, Kronecker invariants and feedback. Ordinary Differential Equations, edited by L. Weiss. *Proc. Conf. Ordinary Differential Equations*, Washington D.C.
KAMEN, E. W., 1976, *J. Franklin Inst.*, **301,** 559.
MÜNZNER, H. F., and PRÄTZEL-WOLTERS, D., 1979, *Int. J. Control*, **30,** 291.
ORE, O., 1933, *Ann. Math.*, **34,** 480.
POPOV, V. M., 1972, *SIAM J. Control*, **10,** 252.
PRÄTZEL-WOLTERS, D., 1981 a, *I.E.E.E. Trans. autom. Control*, **26,** 429; 1981 b, Feedback morphisms between unified linear systems. Thesis, Report 39, Forschungsschwerpunkt Dynamische Systeme, Universität Bremen.
ROSENBROCK, H. H., 1970, *State-Space and Multivariable Theory* (London: Nelson and Sons Ltd.).
SILVERMAN, L. M., 1966, *I.E.E.E. Trans. autom. Control*, **11,** 300.
SILVERMAN, L. M., and MEADOWS, H. E., 1967, *SIAM J. Control*, **5,** 64.
WOLOVICH, W. A., 1974, *Linear Multivariable Systems* (New York: Springer).
YLINEN, R., 1980, *Act. Polyt. Scand., Math. Comp. Science Series*, Vol. 32.