SIAM J. Control Optim.
Vol. 35, No. 3, pp. 1053–1091, May 1997

014

STATE MAPS FOR LINEAR SYSTEMS*

PAOLO RAPISARDA† AND J. C. WILLEMS‡

Abstract. Modeling of physical systems consists of writing the equations describing a phenomenon and yields as a result a set of differential-algebraic equations. As such, state-space models are not a natural starting point for modeling, while they have utmost importance in the simulation and control phase. The paper addresses the problem of computing state variables for systems of linear differential-algebraic equations of various forms. The point of view from which the problem is considered is the behavioral one, as put forward in [J. C. Willems, *Automatica J. IFAC*, 22 (1986), pp. 561–580; *Dynamics Reported*, 2 (1989), pp. 171–269; *IEEE Trans. Automat. Control*, 36 (1991), pp. 259–294].

Key words. behavioral system theory, state map, differential-algebraic equations

AMS subject classifications. 93A30, 93B25

PII. S0363012994268412

1. Introduction.

The usual procedure in modeling consists of *tearing* and *zooming*: a system is viewed as an interconnection of subsystems, and modeling consists of describing the subsystems and the interconnection laws. This procedure is often executed hierarchically, with the subsystems in turn viewed as an interconnection. The net result of such a modeling procedure will be a model which involves *manifest* variables (often called *external* variables), which are the variables whose behavior we try to model, and *latent* variables (often called *internal* variables), which are the variables describing the subsystems. The formalization of this modeling procedure is the philosophy underlying the *behavioral* approach to systems theory. These ideas have been explained in detail in [8, 9, 10].

As should be apparent, the resulting model will typically involve many algebraic relations (for example, interconnection constraints, resistors laws, spring and damper characteristics, kinematic constraints), combined with differential equations. These may be first-order (for example, inductors, capacitors, the dynamics of dampers), second-order (for example, the dynamics of masses), or higher-order (for example, subsystems whose dynamic laws have been obtained from an identification procedure).

A state-space model is hence not the natural end result of the modeling phase, while its importance for simulation or for control design is undisputed. This is one of the reasons why the notion of state is one of the most investigated ones in system theory and why its characterization and construction have been the subject of many papers since the beginning of this discipline. The problem that we deal with in this paper is that of computing state variables, from which a state-space model is easily recovered, starting from an arbitrary set of linear differential-algebraic equations.

The paper is organized as follows. In section 2 a set of definitions and results pertaining to the behavioral framework is introduced. In section 3 the consequences of the property of Markovianity, the key to the notion of state, are worked out. In

*Received by the editors May 23, 1994; accepted for publication (in revised form) April 9, 1996.
http://www.siam.org/journals/sicon/35-3/26841.html

†Department of Electronic, Electrical and Computer Engineering, University of Trieste, I-34127 Trieste, Italy (rapisard@univ.trieste.it).

‡Department of Mathematics, University of Groningen, 9700 AV Groningen, The Netherlands (J.C.Willems@math.rug.nl). The research of this author was supported in part by grant SC1*-CT92-0779 on System Identification of the Science Program of the Commission of the European Community.

section 4 the problem at hand is formally stated. In section 5 operators on polynomials are introduced which will be used in sections 6–9, in which state functions for systems of differential-algebraic equations are computed. As we shall see, the systems may be in kernel, in image, or in hybrid form.

The proofs and some of the notation are collected in the appendices.

2. The behavioral framework. In this section we give a brief introduction to behavioral system theory, with emphasis on the definitions and results pertaining to the problem at hand, referring the reader to [8, 9, 10] for a thorough exposition.

In the behavioral framework a *system* is defined as a triple $\Sigma = (T, W, \mathcal{B})$, with T the time set, W the signal space, and $\mathcal{B}$ the behavior of the system, $\mathcal{B} \subseteq W^T$.

Effectively, a system consists of a family of trajectories which take their value in the signal space. In this paper we consider continuous-time ($T = \mathbb{R}$) systems whose variables take values in a finite-dimensional real vector space, $W = \mathbb{R}^q$. A dynamical system will be called *linear* if $\mathcal{B}$ is a linear vector subspace of $(\mathbb{R}^q)^{\mathbb{R}}$, the latter equipped with the usual vector space structure induced by that of $\mathbb{R}^q$, and *time invariant* if the following holds $\forall\ t \in \mathbb{R}$:

$$(w(\cdot) \ \in \ \mathcal{B}) \Longrightarrow (w(\cdot + t) \ \in \ \mathcal{B}). \tag{2.1}$$

In many instances systems are described by differential equations, say,

$$f_1\left(w, \frac{d}{dt}w, \ldots, \frac{d^L}{dt^L}w\right) = f_2\left(w, \frac{d}{dt}w, \ldots, \frac{d^L}{dt^L}w\right). \tag{2.2}$$

A concrete representation of the behavior of a linear, time-invariant, continuous-time differential system $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ is then given as the solution set of a system of linear, constant coefficient differential equations:

$$R_0 w + R_1 \frac{d}{dt} w + R_2 \frac{d^2}{dt^2} w + \cdots + R_L \frac{d^L}{dt^L} w = 0 \tag{2.3}$$

with constant matrices $R_i \in \mathbb{R}^{\bullet \times q}$. Equation (2.3) is what we call a *kernel representation* of such a system. A shorthand notation for (2.3) is

$$R\left(\frac{d}{dt}\right) w = 0, \tag{2.4}$$

where $R(\xi) := R_0 + R_1\xi + \cdots + R_L\xi^L \in \mathbb{R}^{\bullet \times q}[\xi]$. Note that (2.4) may involve algebraic equations in addition to ordinary differential equations.

The behavioral framework takes into account the nonuniqueness of the representation of behaviors. This is natural, given the connections between this approach and the actual procedure of modeling physical systems, in which different, although equivalent, sets of equations describing the same phenomenon may be produced.

The formalization of this equivalence concept is given as follows. Two kernel representations $R_1(\frac{d}{dt})w = 0$ and $R_2(\frac{d}{dt})w = 0$ with R_1, $R_2 \in \mathbb{R}^{\bullet \times q}[\xi]$ are *equivalent*—that is, the behaviors associated with them are the same—if and only if there exist polynomial matrices F_1, F_2 with a suitable number of columns such that $R_1 = F_1 R_2$ and $R_2 = F_2 R_1$; in particular, if R_1 and R_2 are of full row rank, this means that there exists a unimodular polynomial matrix F such that $R_1 = F R_2$ (see [10, p. 263]).

As already explained in the introduction, (2.4) is not the most natural result of a modeling process, since normally a number of auxiliary latent variables will have been introduced. The natural counterpart of (2.4) to cope with this is

$$R\left(\frac{d}{dt}\right)w = M\left(\frac{d}{dt}\right)\ell, \tag{2.5}$$

where $M \in \mathbb{R}^{\bullet\times d}[\xi]$ and where $\ell \in (\mathbb{R}^d)^{\mathbb{R}}$ are the latent variables. The set of equations (2.5) is called a *latent variable* or a *hybrid representation* of the *latent variable system* $(\mathbb{R}, \mathbb{R}^q, \mathbb{R}^d, \mathcal{B}_f)$, where the *full behavior* $\mathcal{B}_f$ is composed of trajectories (w, ℓ) satisfying (2.5) and inducing the *external* or *manifest behavior* $\mathcal{B}_{ext} := \pi_w \mathcal{B}_f$ by projection on the external variables. Actually the external behavior induced by a latent variable system may be described (modulo some closedness problems pointed out in [6] and discussed in detail in the following) in terms of the external variables only by *eliminating the latent variable*, a procedure discussed in the following.

Of course the problem arises what sort of *solution* we want to use for (2.4) and (2.5). Restricting ourselves to $\mathcal{C}^\infty$ (infinitely differentiable signals) would leave out interesting functions such as steps, etc. The space of distributions is a bit too large, leaving us with the problem of defining the value of a solution at a point. The space $\mathcal{L}_1^{loc}$ of locally integrable functions is large enough to accommodate steps, ramps, and so on and still concrete enough to avoid the problems we would have with distributions. Therefore, in (2.4) and (2.5) *w and ℓ are to be intended in $\mathcal{L}_1^{loc}$ and equality in the sense of distributions.*

Let us focus now on the *elimination of the latent variable* from (2.5).

Hybrid representations involve two kind of variables, namely, the manifest and the latent variables; and associated with a hybrid representation are two behaviors, the full behavior $\mathcal{B}_f$, consisting of trajectories with both the latent and the external variables, and the external behavior $\mathcal{B}_{ext}$, composed of trajectories in the manifest variables only.

At the level of trajectories, the relationship between $\mathcal{B}_f$ and $\mathcal{B}_{ext}$ is the following: any trajectory in $\mathcal{B}_{ext}$ is induced by a trajectory in $\mathcal{B}_f$ via the projection operator $\pi_w((w, \ell)) = w$. At the level of representations and of the equations representing the behaviors, things are more complicated. Take a hybrid representation

$$R\left(\frac{d}{dt}\right)w = M\left(\frac{d}{dt}\right)\ell. \tag{2.6}$$

By premultiplication by a unimodular matrix U we can bring $(\,R \quad | \quad -M\,)$ to the form

$$U\,(\,R \quad | \quad -M\,) = \begin{pmatrix} R_1' & 0 \\ R_2' & -M_2' \end{pmatrix} \tag{2.7}$$

with M_2' of full row rank. Unimodularity of U implies that the full behavior represented by (2.6) is not altered by the change of representation and coincides with the behavior represented by

$$R_1'\left(\frac{d}{dt}\right)w = 0, \tag{2.8}$$

$$R_2'\left(\frac{d}{dt}\right)w = M_2'\left(\frac{d}{dt}\right)\ell. \tag{2.9}$$

A natural candidate for representing the external behavior corresponding to (2.6) would be $R_1'(\frac{d}{dt})w = 0$, since if $w \in \mathcal{B}_{ext}$ of (2.6), then $R_1'(\frac{d}{dt})w = 0$. In fact, for discrete-time systems it has been shown in [10, p. 234] that the analogue of (2.8) in discrete time is indeed a kernel representation of the manifest behavior; this result is referred to in behavioral system theory as the *latent variable elimination theorem*. However, in the continuous-time case there are difficulties. Take, for example, the hybrid representation

$$\begin{aligned} w_1 &= w_2, \\ \frac{d}{dt}w_2 &= \ell. \end{aligned} \tag{2.10}$$

Note that the second equation imposes a smoothness requirement on w_2 not present in the first one: the external behavior does not coincide with the one described by $w_1 = w_2$. When situations like the one exemplified above do not occur, the latent variable ℓ is said to be *properly eliminable* (cf. [6]). Necessary and sufficient conditions for proper eliminability are given in [6].

If the latent variable is not properly eliminable, $\mathcal{B}_{ext}$ is described by $R_1'(\frac{d}{dt})w = 0$ of (2.8) along with some smoothness constraints. These constraints on w cannot be represented by equations involving w alone and the need to circumvent this difficulty arises. The most natural way to do this is to drop them, that is, to consider the closure of $\mathcal{B}_{ext}$ in the topology of $\mathcal{L}_1^{loc}$. This choice has much to recommend it besides its simplicity: it allows to keep $\mathcal{L}_1^{loc}$ as the natural function space in which to operate, and in this way the latent variable can always be eliminated. We summarize this in the following theorem.

THEOREM 2.1. *Let* (2.6) *be a hybrid representation. There exists a unimodular matrix U such that*

$$U\,(\,R \quad | \quad -M\,) = \begin{pmatrix} R_1' & 0 \\ R_2' & -M_2' \end{pmatrix} \tag{2.11}$$

with M_2' of full row rank. Then

$$\begin{aligned} \left\{ w \in \mathcal{L}_1^{loc}(\mathbb{R};\mathbb{R}^q) \mid R_1'\left(\frac{d}{dt}\right)w = 0 \right\} &= \overline{\pi_w(\mathcal{B}_f)}^{closure} \\ &= \overline{\{w \mid \exists\, \ell \text{ s.t. (2.6) holds}\}}^{closure} \end{aligned} \tag{2.12}$$

with the closure taken in the topology of $\mathcal{L}_1^{loc}(\mathbb{R};\mathbb{R}^q)$.

Proof. See the appendix. □

In the following, unless otherwise stated, we will take (2.12) to be the manifest behavior induced by (2.6).

The important notions of controllability and observability emerge in the behavioral context as follows. The time-invariant system $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ is said to be *controllable* if for all w_1, w_2 in $\mathcal{B}$, there exists a $T \geq 0$ and a $w \in \mathcal{B}$ such that $w(t) = w_1(t)$ for $t < 0$ and $w(t + T) = w_2(t)$ for $t \geq 0$. The notion of observability deals with latent variable systems and refers to the possibility of deducing the latent variables from the manifest ones. Thus (2.5) defines an *observable* system if there exists a map $F : (\mathbb{R}^q)^{\mathbb{R}} \mapsto (\mathbb{R}^d)^{\mathbb{R}}$ such that $((w, \ell) \in \mathcal{B}_f) \Longrightarrow (\ell = F(w))$. For linear latent variable systems this is equivalent to $((0, \ell) \in \mathcal{B}_f) \Longrightarrow (\ell = 0)$.

The question when a system (2.4) is controllable can be answered effectively in terms of R. Indeed, (2.4) is controllable if and only if $\text{rank}(R(\lambda)) = \text{rank}(R)$ for all $\lambda \in \mathbb{C}$, as shown in [9, p. 238]. (Here one should view $R(\lambda)$ as a matrix over the field of complex numbers and R as a matrix over the field of real rational functions.) Analogously, (2.5) will be observable if and only if $M(\lambda)$ is right prime (equivalently, if and only if $M(\lambda)$ is of full column rank for all $\lambda \in \mathbb{C}$), as shown in [9, p. 239]. Actually, controllability can also be characterized in terms of (2.5). Take $R = I$ in (2.5), yielding

$$w = M\left(\frac{d}{dt}\right)\ell. \tag{2.13}$$

Let $\mathcal{B}$ be the manifest behavior of (2.13). (More precisely, in view of questions related to closedness, $\mathcal{B} = \overline{M(\frac{d}{dt})\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^d)}^{closure}$, with the closure taken with respect to the topology of $\mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^q)$.) This yields the dynamical system $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$, and for obvious reasons we will call (2.13) an *image representation* of $\mathcal{B}$. By the already mentioned latent variable elimination theorem, $\mathcal{B}$ admits a kernel representation as (2.4). However, not every system (2.4) has an image representation. This is the case if and only if the system is controllable! (See [9, p. 238].)

Finally, let us introduce the notion of *input* and *output*. Consider a system (2.4) with $R(\xi)$ of full row rank. Possibly permuting the components of w, assume R partitioned as $R := (\,P \mid -Q\,)$, with P square, nonsingular, and $P^{-1}Q$ proper. Such a partition of R always exists and can be found as follows. By unimodular premultiplication by a suitable U, bring R in row reduced form (see [3, p. 382]). Let R'_{hc} be the highest row coefficient matrix of $R' := UR$. R'_{hc} has full row rank, and therefore there exists at least one $g \times g$ nonzero minor, corresponding to a choice of the k_1th, k_2th, $\dots, k_g$th column of R_{hc}. The minor of R corresponding to this column selection is of maximal degree among the $g \times g$ minors of R. This implies that if we take P to be the matrix formed by the k_1th, k_2th, $\dots, k_g$th column of R, P is nonsingular and $P^{-1}Q$ is proper, $-Q$ being the complementary matrix of P in R.

The partition of R induces a corresponding partition of w in (y, u) so that (2.4) may be rewritten as

$$P\left(\frac{d}{dt}\right)y = Q\left(\frac{d}{dt}\right)u. \tag{2.14}$$

This is an *input-output* representation of the behavior of (2.4), with y the *output* variables and u the *input* variables.

It is important to note that the selection of P and Q of (2.14) is not unique, in general. This implies that for a system whose behavior is described by (2.4), different selections of inputs and outputs can be given. This corresponds to different selections of R'_{hc} to form a nonzero minor in the procedure sketched above. Anyway, it is possible to prove (see [10, p. 243]) that the number of outputs, and consequently the number of inputs, in any representation (2.4) of the behavior of the system is unique and coincides with rank(R).

3. State, Markovianity, and first-order representations. Let $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ be a time-invariant dynamical system. We will call it *Markovian* if

$$(w_1, w_2 \in \mathcal{B}) \wedge (w_1, w_2 \text{ continuous at } 0) \wedge (w_1(0) = w_2(0)) \tag{3.1}$$

implies $((w_1 \wedge w_2)) \in \mathcal{B})$; $w_1 \wedge w_2$ stands for concatenation:

$$(f_1 \wedge f_2)(t) := \begin{cases} f_1(t), & t < 0, \\ f_2(t), & t \geq 0. \end{cases} \tag{3.2}$$

Thus in a Markovian system trajectories passing in a continuous way through the same point at $t = 0$ can be concatenated. The very much related notion of state refers to systems with latent variables. Thus let $(\mathbb{R}, \mathbb{R}^q, \mathbb{R}^d, \mathcal{B}_f)$ be a time-invariant latent variable system. Then it is a *state system* if

$$\begin{aligned} &((w_1, \ell_1), (w_2, \ell_2) \in \mathcal{B}_f) \wedge (\ell_1(0) = \ell_2(0)) \wedge (\ell_1, \ell_2 \text{ continuous at } t = 0) \\ &\qquad\qquad\qquad \Longrightarrow ((w_1, \ell_1) \wedge (w_2, \ell_2)) \in \mathcal{B}_f). \end{aligned} \tag{3.3}$$

We call (3.3) the *axiom of state*. If (3.3) holds, then the latent variable is called the *state*. Thus in a state model trajectories passing in a continuous way through the same state at $t = 0$ can be concatenated. The continuity requirement is inspired by the fact that we are dealing with solutions of (2.4) and (2.5) in $\mathcal{L}_1^{loc}$, in which case the simple requirement $\ell_1(0) = \ell_2(0)$ is of little consequence.

Usually a Markovian or a state variable is denoted by x. We will do so in the following discussion.

It is easy to prove that if the behavior is described by a set of first-order differential equations, as

$$f\left(x, \frac{d}{dt}x\right) = 0, \tag{3.4}$$

then it is Markovian; similarly, if it can be described by a set of differential equations which is first order in the latent variables and zeroth order in the manifest variables, as

$$f\left(w, x, \frac{d}{dt}x\right) = 0, \tag{3.5}$$

then it is a state model (see [9, p. 191]). For linear differential systems this is, in fact, necessary and sufficient, as shown by the following proposition.

PROPOSITION 3.1. *Let Σ_S be a system as in* (2.5). *Then it is a state-space system if and only if there exist matrices E, F, and G such that $\mathcal{B}_f$ has the kernel representation*

$$Gw + Fx + E\frac{d}{dt}x = 0. \tag{3.6}$$

Analogously, $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ *as in* (2.4) *is Markovian if and only if there exist matrices E and F such that $\mathcal{B}$ has the kernel representation*

$$Fx + E\frac{d}{dt}x = 0. \tag{3.7}$$

Proof. See the appendix. □

Remark 3.1. The above proposition constitutes an example of application of the following fact. Equation (2.4) determines a representation of the behavior $\mathcal{B}$, but it is not the unique possible representation of $\mathcal{B}$. In fact, if U is a unimodular matrix, then UR determines the same behavior. This allows us to obtain representations which put certain properties in evidence, just like the state property above.

Remark 3.2. A state-space system induces, by projection of $\mathcal{B}_f$ on the external variable w, an external behavior $\mathcal{B} := \pi_w \mathcal{B}_f$. Therefore it will be called a *state-space representation* or a *state-space model* of $\mathcal{B}$.

Besides state-space models of $\mathcal{B}$ whose full behavior is described by equations of the form (3.6), *state-space models with driving variables* and *input/state/output models* can be defined.

A *state-space model with driving variables* is described by

$$\frac{d}{dt}x = Ax + Bv,$$
$$w = Cx + Dv, \tag{3.8}$$

where x is a state variable for $\mathcal{B} = \{w \in \mathcal{L}_1^{loc} | \ \exists\ x \in \mathcal{L}_1^{loc}, v \in \mathcal{L}_1^{loc},\ \text{s.t. (3.8) holds}\}$ and v is composed of free but latent variables which generate, together with the initial conditions, the state trajectory and the external signal. We call v the *driving variable.*

By integrating the state property and the input-output structure in the same representation, an *input/state/output representation* is obtained. It can be computed from a state representation (3.6) by partitioning the w variables in inputs u and outputs y and rearranging the equations (3.6) so that a representation

$$\frac{d}{dt}x = Ax + Bu,$$
$$y = Cx + Du \tag{3.9}$$

is obtained.

Let $\Sigma_S = (\mathbb{R}, \mathbb{R}^q, \mathbb{R}^n, \mathcal{B}_f)$ be a state-space system and $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$ be its external (i.e., manifest) behavior. We will call Σ_s *minimal* if whenever $\Sigma_{S'} = (\mathbb{R}, \mathbb{R}^q, \mathbb{R}^{n'}, \mathcal{B}'_f)$ is another state-space model with the same external behavior $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$, then $n \leq n'$. It is possible to prove (see [10, p. 270]) that Σ_S is minimal if and only if it is *observable* (with the state viewed as the latent variable) and *state trim* (meaning that for all $x_0 \in \mathbb{R}^n$ there exists a $(w, x) \in \mathcal{B}_f \bigcap \mathcal{C}^\infty$ such that $x(0) = x_0$). Observability, in particular, implies that there then exists a $F \in \mathbb{R}^{n \times q}[\xi]$ such that $((w, x) \in \mathcal{B}_f) \Longrightarrow (x = F(\frac{d}{dt})w)$. Actually it can further be shown (see [10, p. 271]) that if Σ_S and $\Sigma_{S'} = (\mathbb{R}, \mathbb{R}^q, \mathbb{R}^{n'}, \mathcal{B}'_f)$ are two minimal state space systems with the same external behavior, then there exists a nonsingular matrix $S \in \mathbb{R}^{n \times n}$ such that

$$((w, x) \in \mathcal{B}_S \text{ and } (w, x') \in \mathcal{B}_{S'}) \Longrightarrow (x' = Sx). \tag{3.10}$$

4. Problem statement. The question arises of how to compute a set of state variables when a system is given either in kernel or in hybrid form.

This question may be stated as: *Given a set of equations in either kernel or hybrid form, how do we determine a* state map $X(\frac{d}{dt})$*?* More precisely, given R, determine the integer n and $X \in \mathbb{R}^{n \times q}[\xi]$ such that

$$R\left(\frac{d}{dt}\right)w = 0, \tag{4.1}$$

$$X\left(\frac{d}{dt}\right)w = x \tag{4.2}$$

defines a (minimal) state-space system with external behavior given by (4.1). The problem is to derive X from R. Similarly we want to derive a state map X for the external behavior of a system represented in hybrid form or image form. In this case, in view of the closedness problems discussed in the previous section, we will interpret the external behavior associated with the hybrid or image representation under study, in the sense of Theorem 2.1.

Remark 4.1. It is of utmost importance at this point to note that the external behavior of the state-space system described by (4.1), (4.2) is assumed to be described by (4.1); that is, the equations (4.2) do not impose any smoothness constraint on the trajectories defined by (4.1). Therefore, when dealing with state-space representations of a given external behavior $\mathcal{B}$, we will consider the (latent) state variable x induced by the state map to be *properly eliminable.*

The next section introduces the tools that we will use to deal with the problem of computing state maps for the various sorts of representations introduced so far.

5. Operators on polynomials. The behavioral framework for linear differential systems is intimately connected to polynomial matrix algebra. These connections are also reflected in the results which will be presented in the following sections, related to the characterization of state maps.

This section is devoted to the introduction of some notational conventions related to polynomials and rational functions.

Any rational function can be written in a unique way as the sum of a polynomial and of a strictly proper rational function. That is, given $q \in \mathbb{R}(\xi)$, there exist unique $p \in \mathbb{R}[\xi]$ and $s \in \mathbb{R}_+(\xi)$, the set of strictly proper rational functions, such that $q = p + s$. Now define

$$(\ )_+ : \mathbb{R}(\xi) \mapsto \mathbb{R}[\xi] \tag{5.1}$$

as

$$(q(\xi))_+ := p(\xi). \tag{5.2}$$

On the set of rational functions in the indeterminate ξ, multiplication by ξ^{-1} defines a map $\xi^{-1} : \mathbb{R}(\xi) \mapsto \mathbb{R}(\xi)$ in the obvious way.

DEFINITION 5.1. *The* shift-and-cut *operator σ_+ is defined as*

$$\begin{aligned} &\sigma_+ : \mathbb{R}(\xi) \mapsto \mathbb{R}[\xi], \\ &\sigma_+ := (\ )_+ \circ \xi^{-1}. \end{aligned} \tag{5.3}$$

The definition of σ_+ is extended to vectors and matrices of rational functions in a componentwise manner.

Iterated application of σ_+ will be considered in the following and denoted as

$$\sigma_+^k := \overbrace{\sigma_+ \circ \sigma_+ \circ \cdots \circ \sigma_+}^{k-times}. \tag{5.4}$$

In the following, special importance will be given to the action of σ_+ on vector polynomials. Therefore, let us examine in detail what the result is of the application of σ_+ to a vector polynomial $p \in \mathbb{R}^{1\times q}[\xi]$. Write

$$p(\xi) := p_\delta \xi^\delta + p_{\delta-1}\xi^{\delta-1} + \cdots + p_1\xi + p_0. \tag{5.5}$$

Then

$$\sigma_+(p(\xi)) = p_\delta \xi^{\delta-1} + p_{\delta-1}\xi^{\delta-2} + \cdots + p_1; \tag{5.6}$$

that is,

$$\sigma_+(p(\xi)) = \xi^{-1}(p(\xi) - p_0). \tag{5.7}$$

Now let $R \in \mathbb{R}^{g\times q}[\xi]$ be given, and assume $R := R_0 + R_1\xi + \cdots + R_L\xi^L$. Define R^k, $k = 0, \ldots, L$, as $R^0 := R$ and $R^k := \sigma_+^k R = \sigma_+ R^{k-1}$, $k = 1, \ldots, L$. Define the Ξ-*matrix* R_Ξ as

$$R_\Xi := \operatorname{col}(R^k)_{k=1,\ldots,L} = \begin{pmatrix} R^1 \\ R^2 \\ \vdots \\ R^L \end{pmatrix}. \tag{5.8}$$

Connected to R_Ξ is the important notion of Ξ-space. Let $r_1, r_2, \ldots, r_g$ denote the rows of R. Then define the $\mathbb{R}$-vector space Ξ_R as

$$\Xi_R := \langle \sigma_+^k(r_i) \rangle, \qquad k \in \mathbb{N},\ i = 1, \ldots, g, \tag{5.9}$$

where $\langle\ \rangle$ denotes the span over $\mathbb{R}$. The Ξ-space of R is most easily constructed as the $\mathbb{R}$-vector space generated by the rows of R_Ξ. Note that R_Ξ need not define a basis of the Ξ-space of R.

Introduce now on Ξ_R the equivalence relation $\overset{R}{\sim}$ defined as follows: $p, q \in \Xi_R$ are *equivalent modulo* R, written $p \overset{R}{\sim} q$, if and only if there exists $r(\xi) \in \mathbb{R}^{1\times g}[\xi]$ such that $p(\xi) - q(\xi) = r(\xi)R(\xi)$. It is easily verified that $\overset{R}{\sim}$ is indeed an equivalence relation.

Note that the vector space structure on Ξ_R induces a vector space structure on the set of equivalence classes induced by $\overset{R}{\sim}$ on Ξ_R. We will denote this set of equivalence classes as $\Xi_R \pmod R$. That is,

$$\Xi_R \pmod R = \{[p] \in 2^{\Xi_R} \mid q \in [p] \text{ iff } \exists\, r \in \mathbb{R}^{1\times g}[\xi] \text{ s.t. } p = q + rR\}. \tag{5.10}$$

The following example illustrates the above notions.

Example 5.1. Let

$$R := \begin{pmatrix} \xi^2 + 2\xi - 1 & \xi + 1 \\ \xi - 1 & \xi^2 - 3 \end{pmatrix}, \tag{5.11}$$

and consider its first row, $(\,\xi^2 + 2\xi - 1 \quad \xi + 1\,)$. The shift-and-cut operator acts on this row as

$$\sigma_+ (\,\xi^2 + 2\xi - 1 \quad \xi + 1\,) = (\,\xi + 2 \quad 1\,). \tag{5.12}$$

The Ξ_R space is the vector space spanned by

$$(\,\xi + 2 \quad 1\,),\ (\,1 \quad 0\,),\ (\,1 \quad \xi\,),\ (\,0 \quad 1\,), \tag{5.13}$$

which actually form a basis for this space. It is easily verified that the vectors (5.13), interpreted as representing elements of $\Xi_R \pmod R$, are linearly independent as well, and therefore form a basis of $\Xi_R \pmod R$.

Note that selecting from the rows of R_Ξ a maximal set of linearly independent rows and considering these as representatives of elements of $\Xi_R \pmod R$ do not necessarily yield a basis for $\Xi_R \pmod R$, as made explicit by the following example.

Example 5.2. Let

$$R = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & \xi^3 \\ 0 & 0 & \xi \end{pmatrix}. \tag{5.14}$$

A maximal set of linearly independent rows of R_Ξ is

$$\begin{pmatrix} 0 & 0 & \xi^2 \end{pmatrix}, \begin{pmatrix} 0 & 0 & \xi \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \tag{5.15}$$

but the first and the second element of this set are equivalent to zero modulo R, since $\begin{pmatrix} 0 & 0 & \xi^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & \xi \end{pmatrix} R$ and $\begin{pmatrix} 0 & 0 & \xi \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} R$.

Equipped with these notions, we are now ready to consider the problem of the determination of state-inducing maps for systems in kernel form.

6. State maps for systems in kernel form. The systems we will consider in this section are those described by equations (2.4). For some systems of this kind the problem of computing a state map is trivial, namely, those corresponding to a behavior coinciding with the zero trajectory only. These systems can be characterized as those corresponding to a right prime polynomial matrix R; this can be readily shown by resorting to the Smith form of R. In the following we assume that R is not right prime.

The main result of this section is a characterization of state-inducing polynomial matrices for systems in kernel form. As a preliminary result, we first consider the conditions under which a trajectory is concatenable with the zero one. These conditions correspond to a system of linear equations involving w and its derivatives and define a polynomial differential operator which, in fact, corresponds to a state map. As we will see, the rows of this polynomial matrix have a nice interpretation in terms of the shift-and-cut operator defined in the previous section.

Before stating Proposition 6.1, we observe the following smoothness result. Consider the polynomial matrices R^1, R^2, $\ldots$, and observe that if $w \in \mathcal{L}_1^{loc}$ is a solution of (2.4) in the sense of distributions, then

$$R^k\left(\frac{d}{dt}\right)w \tag{6.1}$$

is continuous for $k = 1, 2, \ldots$. In order to see this, let

$$R(\xi) = R_0 + R_1\xi + \cdots + R_L\xi^L; \tag{6.2}$$

then (2.4) implies

$$\frac{d}{dt}\left(R_1 + \cdots + R_L\frac{d^{L-1}}{dt^{L-1}}\right)w = -R_0 w. \tag{6.3}$$

Since the right-hand side is in $\mathcal{L}_1^{loc}$, this implies that

$$\left(R_1 + \cdots + R_L\frac{d^{L-1}}{dt^{L-1}}\right)w = R^1\left(\frac{d}{dt}\right)w \tag{6.4}$$

is absolutely continuous. Proceeding recursively yields the absolute continuity of (6.1). Note that this implies that $R_\Xi(\frac{d}{dt})w$ is also absolutely continuous.

PROPOSITION 6.1. *Let a kernel representation as in* (2.4) *be given, and let $\mathcal{B}$ be its behavior. A trajectory $w \in \mathcal{B}$ is concatenable with the zero trajectory; that is, $0 \wedge w \ \in \mathcal{B}$ if and only if*

$$\left(R_\Xi\left(\frac{d}{dt}\right)w\right)(0) = 0. \tag{6.5}$$

Proof. See the appendix. ◻

This yields the main result of this section.

THEOREM 6.2. *The polynomial matrix $X \in \mathbb{R}^{\bullet \times q}[\xi]$ defines a state-inducing map for* (2.4)*; i.e.,*

$$\begin{aligned} R\left(\frac{d}{dt}\right)w &= 0, \\ X\left(\frac{d}{dt}\right)w &= x \end{aligned} \tag{6.6}$$

defines a state-space system with external behavior Ker $R(\frac{d}{dt})$ *if and only if there exists a matrix $A \in \mathbb{R}^{\bullet\times\bullet}$ and a polynomial matrix $B(\xi) \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ such that*

$$R_\Xi(\xi) = AX(\xi) + B(\xi)R(\xi) \tag{6.7}$$

and the latent variable x is properly eliminable from the system with latent variable (6.6).

Proof. See the appendix. ◻

Remark 6.1. Proper eliminability of $x = X(\frac{d}{dt})w$ in the system with latent variable (6.6) can be checked as follows (cf. [6, Theorem 2.5]). Without loss of generality, assume $R(\xi)$ to be of full row rank g, and let $X(\xi)$ have n rows. x is properly eliminable if and only if there exists an $(n+g)\times(n+g)$ submatrix of maximal determinantal degree of

$$\begin{pmatrix} R & 0_{g\times n} \\ X & -I_n \end{pmatrix}, \tag{6.8}$$

which includes the last n columns of (6.8).

Remark 6.2. In the discrete-time case, an analogue of Proposition 6.1 has been given in [7, p. 1075]. A necessary condition for a state map, analogous to (6.7), has been given in the continuous-time case (with a solution space other than $\mathcal{L}_1^{loc}$) in [5, p. 77]. In the context of discrete-time *output nulling representations*

$$\begin{aligned} x(k+1) &= Ax(k) + Bw(k), \\ 0 &= Cx(k) + Dw(k), \end{aligned} \tag{6.9}$$

a procedure similar to using the shift-and-cut operator to obtain R_Ξ from R has been used in [1, p. 3643].

Remark 6.3. If the rows of X of (6.7) are interpreted as representative of elements of Ξ_R (mod R), Theorem 6.2 can be restated as follows: X defines a state-inducing map for (2.4) if and only if the span over $\mathbb{R}$ of its rows contains Ξ_R (mod R) and the latent variable x is properly eliminable from (6.6). This, together with the smoothness result given at the beginning of this section, yields the following corollary of Theorem 6.2.

COROLLARY 6.3. *The polynomial matrix $X \in \mathbb{R}^{\bullet \times q}[\xi]$ defines a minimal state-inducing map for* (2.4) *if and only if its rows, considered as representative of elements of* $\Xi_R \pmod R$*, form a basis for* $\Xi_R \pmod R$.

Remark 6.4. Note that in the scalar case ($R \in \mathbb{R}[\xi]$, $R \neq 0$) the above theorem corresponds to the usual method of stacking the lower-order derivatives of each component to reduce a system of equations of high order to a system of equations of order one, as made explicit by the following example.

Example 6.1. Let $q = 1$ and a system be described by

$$p\left(\frac{d}{dt}\right) w = 0 \tag{6.10}$$

with

$$p(\xi) := p_0 + p_1\xi + \cdots + p_L\xi^L, \tag{6.11}$$

with $p_L \neq 0$.

The Ξ_p space is generated by

$$p_1 + p_2\xi + p_3\xi^2 + \cdots + p_L\xi^{L-1}, p_2 + p_3\xi + \cdots + p_L\xi^{L-2}, \ldots, p_{L-1} + p_L\xi, p_L, \tag{6.12}$$

which in fact constitutes a basis for the space; another basis for Ξ_p could be chosen as

$$\xi^{L-1}, \xi^{L-2}, \ldots, \xi, 1. \tag{6.13}$$

In fact, it can be verified that both (6.12) and (6.13) are bases of $\Xi_p \pmod p$. Therefore a minimal state is induced by the first $L-1$ derivatives of w, as made apparent by (6.13), or by the differential operators associated with the polynomials (6.12). That is, both

$$x := \begin{pmatrix} \frac{d^{L-1}}{dt^{L-1}} w \\ \frac{d^{L-2}}{dt^{L-2}} w \\ \vdots \\ \frac{d}{dt} w \\ w \end{pmatrix} \tag{6.14}$$

and

$$x := \begin{pmatrix} (p_1 + p_2\frac{d}{dt} + p_3\frac{d^2}{dt^2} + \cdots + p_L\frac{d^{L-1}}{dt^{L-1}})w \\ (p_2 + p_3\frac{d}{dt} + \cdots + p_L\frac{d^{L-2}}{dt^{L-2}})w \\ \vdots \\ (p_{L-1} + p_L\frac{d}{dt})w \\ p_L w \end{pmatrix} \tag{6.15}$$

define minimal state variables.

Remark 6.5. Note that computation of a first-order kernel representation of a state system is easy once the state map is given and amounts to solving a linear system of equations. In fact, once the polynomial matrix $X \in \mathbb{R}^{n \times q}[\xi]$ has been computed, the equations may be recovered in the following way. Find matrices E, F in $\mathbb{R}^{(n+g)\times n}$, $G \in \mathbb{R}^{(n+g)\times q}$, and $T \in \mathbb{R}^{(n+g)\times g}[\xi]$ which solve the equation

$$(\xi E + F)X(\xi) + G = T(\xi)R(\xi). \tag{6.16}$$

An input/state/output representation is easily computed from the kernel representation of the state system obtained in this way. Note that for simple cases these computations can be done by inspection.

In the context of the nonuniqueness of representation of behaviors pointed out in section 2, a question arises with respect to Theorem 6.2. That is, given two equivalent kernel representations of the same system, which we assume to correspond to full row rank polynomial matrices, what relationships hold between the corresponding Ξ-spaces?

The following result holds.

PROPOSITION 6.4. *Let a kernel representation* (2.4) *be given with R of full row rank, and let an equivalent representation be obtained as UR, U unimodular. Then there exist a constant full column rank matrix A and a polynomial matrix B such that $(UR)_\Xi = AR_\Xi + BR$.*

Proof. See the appendix. □

COROLLARY 6.5. *Let a kernel representation* (2.4) *be given with R of full row rank, and let an equivalent representation be obtained as UR, U unimodular. For each polynomial matrix $\bar{\Xi}_{UR}$ whose rows form a basis of Ξ_{UR} and every polynomial matrix $\bar{\Xi}_R$ whose rows form a basis of Ξ_R there exists a polynomial matrix C and a constant nonsingular matrix T such that $\bar{\Xi}_{UR} = T\bar{\Xi}_R + CR$.*

Minimal (in the sense of the minimal possible dimension of the state space) states are induced by the choice of a polynomial matrix X whose rows form a basis of Ξ_R (mod R). A natural question arises as to when minimality of the state space is already guaranteed by directly applying σ_+ to the equations describing the system. The following result holds.

PROPOSITION 6.6. *Let a kernel representation such as* (2.4) *be given. Then the nonzero rows of R_Ξ define a basis for Ξ_R* (mod R) *if and only if R is in row reduced form. Whence if R is in row reduced form and X is composed of the nonzero rows of R_Ξ,*

$$R\left(\frac{d}{dt}\right)w = 0,$$

$$X\left(\frac{d}{dt}\right)w = x \tag{6.17}$$

defines a minimal state representation.

Proof. See the appendix. □

COROLLARY 6.7. *Let* (2.4) *be given with R of full row rank. The minimal dimension of the state space of the system associated to R equals the McMillan degree of R, i.e., the maximal degree of the* $\mathrm{rank}(R) \times \mathrm{rank}(R)$ *minors of R. In the row reduced case, this equals the sum of the row degrees of R.*

Let us now give two examples illustrating the procedure of state construction.

Example 6.2. Consider the system with behavior described by

$$\frac{d^n}{dt^n}w_1 + \cdots + p_1\frac{d}{dt}w_1 + p_0 w_1 = q_n\frac{d^n}{dt^n}w_2 + \cdots + q_1\frac{d}{dt}w_2 + q_0 w_2, \tag{6.18}$$

where w_i, $i = 1, 2$, are scalar functions. Defining

$$p^i := \sigma_+^i(p) = \xi^{n-i} + p_{n-1}\xi^{n-i-1} + \cdots + p_i \tag{6.19}$$

and analogously for q^i, it is easy to see that

$$f_i := \begin{pmatrix} p^i & -q^i \end{pmatrix} \tag{6.20}$$

$i = 1, \dots, n$, form a basis for $\Xi_{\begin{pmatrix} p & -q \end{pmatrix}}$ (mod $\begin{pmatrix} p & -q \end{pmatrix}$). Stacking the f^i vectors yields a polynomial minimal state-inducing matrix.

Example 6.3. Let a system be described by

$$\frac{d^2}{dt^2}w_1 - \frac{d}{dt}w_2 = 0, \tag{6.21}$$

which corresponds to $R(\xi) = \begin{pmatrix} \xi^2 & -\xi \end{pmatrix}$. The space Ξ_R is generated by $\begin{pmatrix} \xi & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \end{pmatrix}$, as is easily seen applying the shift-and-cut operator to $R(\xi)$. Then a state is defined as

$$x := \begin{pmatrix} \frac{d}{dt}w_1 - w_2 \\ w_1 \end{pmatrix}, \tag{6.22}$$

which corresponds to the input/state/output equations

$$\begin{aligned} \frac{d}{dt}x &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} w_2. \\ w_1 &= x_2. \end{aligned}$$

We summarize the above results in the following algorithms.

ALGORITHM 1 (Construction of a state map for a kernel representation).

Data: $R \in \mathbb{R}^{g\times q}[\xi]$, of degree L.

Output: $X \in \mathbb{R}^{\bullet\times q}[\xi]$ inducing, through $x = X(\frac{d}{dt})w$, a state for the system described by $R(\frac{d}{dt})w = 0$.

Step 1. Set $R^0 := R$ and compute $R^{k+1} := \sigma_+(R^k)$, for $k = 0, 1, \dots, L-1$.

Step 2. Find $x_1, \dots, x_n \in \mathbb{R}^{1\times q}[\xi]$ such that $\langle x_1, \dots, x_n\rangle$ equals the space spanned by the rows of R_Ξ.

Step 3. $X = \mathrm{col}(x_1, \dots, x_n)$.

Step 4. *Stop.*

ALGORITHM 2 (Construction of a minimal state map for a kernel representation).

Data: $R \in \mathbb{R}^{g\times q}[\xi]$, of degree L.

Output: $X \in \mathbb{R}^{\bullet\times q}[\xi]$ inducing, through $x = X(\frac{d}{dt})w$, a minimal state for the system described by $R(\frac{d}{dt})w = 0$.

Step 1. As in Algorithm 1.

Step 2. Find $x_1, \dots, x_n$ forming a basis of Ξ_R (mod R).

Comment: The computation of the vectors x_i can be accomplished by computing R_Ξ and reducing each of its rows modulo R with standard polynomial operations.

Step 3. $X := \mathrm{col}(x_k)_{k=1,\dots,n}$.

Step 4. *Stop.*

ALGORITHM 3 (Verification of a state map).

Data: $R \in \mathbb{R}^{g\times q}[\xi]$, of degree L and $X \in \mathbb{R}^{n\times q}[\xi]$.

Output: *True* if X is a state map for the system described by R, *False* otherwise.

Step 1. Compute R_Ξ as in Algorithm 1.

Step 2. Find a constant matrix A and a polynomial matrix B such that $R_\Xi(\xi) = AX(\xi) + B(\xi)R(\xi)$.

Comment: The computation of A and B can be accomplished with standard polynomial operations.

Step 3. If A and B exists, then

Step 4. If x is properly eliminable from

$$R\left(\frac{d}{dt}\right)w = 0,$$

$$X\left(\frac{d}{dt}\right)w = x,$$

then Output:=*True* else Output:=*False*.

Comment: Proper eliminability can be checked as described in Remark 6.1.

Step 5. *Stop.*

Remark 6.6. The algorithms described above are of immediate interest for the simulation problem, where by "simulation" we mean a procedure for selecting an arbitrary element of the behavior $\mathcal{B}$, followed by an algorithm for computing it. Before getting to the simulation issue, let us describe how to construct a driving variables representation for a kernel description (2.4). Compute R_Ξ and consider the following set of equations in the unknowns A, B, C, D, $P \in \mathbb{R}^{\bullet\times g}[\xi]$, $P' \in \mathbb{R}^{q\times g}[\xi]$, $U \in \mathbb{R}^{\bullet\times q}[\xi]$:

$$\begin{aligned} \xi R_\Xi &= AR_\Xi + BU + PR, \\ I_q &= CR_\Xi + DU + P'R. \end{aligned} \tag{6.23}$$

Then the equations

$$\begin{aligned} \frac{d}{dt}x &= Ax + Bv, \\ w &= Cx + Dv \end{aligned} \tag{6.24}$$

represent the external behavior of (2.4) with a state-space model with driving variable v, as can be seen applying the latent-variables-elimination theorem. Note that A, B, C, D, U, P, and P' in (6.23) are easily obtained by inspection from R and R_Ξ.

Note that (6.24) leads to the following simulation procedure. Given R, compute (6.24) and choose a vector $x_0 \in \mathbb{R}^n$ and a $v \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^m)$. Generate a trajectory x satisfying the first block of equations of (6.24) with the initial conditions $x(0) = x_0$. Then $w \in \mathcal{B}$ can be computed according to the second block of equations (6.24).

7. State maps for systems in hybrid form. As pointed out in the introduction, hybrid representations are the most natural result of a modeling process. Therefore the characterization of state maps for such representations which we give in this section is especially interesting for applications.

Following Theorem 2.1, given a hybrid representation (2.6), we will consider the problem of computing a state map for the closure of the external behavior described by (2.6). To this purpose, let us make some preliminary comments.

First, note that when considering the simulation or control of a system, the state variables will in general be chosen as function of both the external and the latent variables. In fact, although the former are the quantities we are interested in, in a hybrid representation the two kinds of variables enter the description of the system on an equal footing; this is exemplified by the fact that (2.6) can be considered as a kernel description of the full behavior.

Second, a characterization of w-induced state maps can be given as follows. As discussed in Theorem 2.1, the closure of the external behavior of (2.6) is described by (2.8). This implies that the computation of a w-induced state map for the closure of the external behavior of a hybrid representation (2.6) can be performed as follows. First, the ℓ variable is eliminated by computing a suitable unimodular matrix U such that premultiplying the equations by U yields (2.11). Then a set of generators of the Ξ-space of R_1' of (2.8) is computed, as discussed in section 6.

Note that the computation of w-induced state maps for the closure of the external behavior is obtained by elimination of the latent variables, therefore modifying the original equations. This modification is not a desirable feature of a state construction algorithm: the state variables should reflect as much as possible the physical structure of the system as put in evidence by the original equations.

These considerations motivate us to restrict our attention in the remainder of this section to the characterization of (w, ℓ)-induced state maps.

As a third consideration, note that there are hybrid representations for which the determination of a state variable is trivial, namely, the hybrid representations for which the closure of the external behavior corresponds to $\mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^q)$. These representations are characterized as follows. Note that $\forall\, w \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^q)$ there exists an $\ell \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^d)$ such that (2.6) holds if and only if $M(\xi)$ has full row rank. In the following we will implicitly assume that the systems we are dealing with are not of this kind, and that a nontrivial external behavior corresponds to (2.6).

As a fourth consideration, let us characterize the situations where a state variable for the closure of the external behavior is a state variable for the full behavior as well. This happens if and only if the dimension of the minimal state spaces for the closure of the external and the full behavior are the same. An efficient way of checking this is given in the following proposition.

PROPOSITION 7.1. *Let a system be described in hybrid form as in* (2.6)*, and let ℓ be observable from w. The dimensions of the minimal state spaces for the closure of the external and for the full behavior respectively are the same if and only if there exists an input-output selection in* (w, ℓ) *such that the variables ℓ are all outputs for the full behavior.*

Proof. See the appendix. ∎

Remark 7.1. The above proposition implies that if there exists an input-output selection on (w, ℓ) such that the latent variables can all be chosen as outputs, a polynomial differential operator X is a state map for the closure of the external behavior if and only if it is a state map for the full behavior.

Remark 7.2. Existence of an input-output selection on (w, ℓ) such that ℓ is entirely composed of outputs can be checked as follows. Assume that $(\, R \quad | \quad -M\,)$ has full row rank g. Then ℓ can be chosen as entirely composed of outputs for the full behavior if and only if one of the $\mathrm{rank}((\, R \quad | \quad -M\,)) \times \mathrm{rank}((\, R \quad | \quad -M\,))$ minors of maximal degree among all such minors, contains $-M$ as a submatrix.

This procedure is summarized for future use in the following algorithm.

ALGORITHM ℓ-outputs (Verification if ℓ may be chosen as consisting entirely of outputs for the full system).

Data: $(\, R \;\;|\;\; -M \,) \in \mathbb{R}^{g\times(q+d)}[\xi]$ of full row rank.

Output: *True* if there exists an input-output partition such that ℓ is entirely composed of outputs, *False* otherwise.

Step 1. Compute n, the maximal degree of the nonzero $\mathrm{rank}((\, R \;\;|\;\; -M \,)) \times \mathrm{rank}((\, R \;\;|\;\; -M \,))$ minors of $(\, R \;\;|\;\; -M \,)$.

Step 2. For every subset P_i of columns of R such that $(\, P_i \;\;|\;\; -M \,)$ has $\mathrm{rank}((\, R \;\;|\;\; -M \,))$ columns, compute the maximal degree of its nonzero $\mathrm{rank}((\, R \;\;|\;\; -M \,)) \times \mathrm{rank}((\, R \;\;|\;\; -M \,))$ minors, let it be n_i.

Step 3. If there exists i such that $n_i = n$, then *True* else *False*.

Step 4. *Stop.*

Now assume that ℓ is observable but that in any selection of inputs and outputs for the full system some components of ℓ have to be chosen as inputs. Analogously to what has been done in section 6 for the case of kernel representations, we will first characterize the concatenability of external trajectories. Note that concatenability conditions that involve both w and ℓ are, in general, more restrictive than concatenability conditions involving the external variable only: even if a full trajectory cannot be concatenated with zero at $t = 0$, it could still be possible to concatenate the corresponding external trajectory with the zero one. Therefore, the idea that we pursue in the following is to derive the concatenability conditions for the external trajectories starting from the concatenability conditions for the full trajectories. According to Proposition 6.1, the concatenability conditions of full trajectories can be characterized using the matrix $(\, R \;\; -M \,)_\Xi$. As we will see, to derive concatenability conditions for the external trajectory we project $(\, R \;\; -M \,)_\Xi$ down with a suitably defined linear map. We call this process the *reduction* of $(\, R \;\; -M \,)_\Xi$.

The reduction process involves introducing some new concepts.

Assume that the full row rank matrix $(\, R \;\;|\;\; -M \,)(\xi) = \sum_{j=0}^{L} (\, R_j \;\; -M_j \,)\, \xi^j$ has g rows. Consider

$$\begin{pmatrix} (\, R \;\;|\;\; -M \,)_\Xi \\ 0_{g\times(d+q)} \end{pmatrix} = \mathrm{col}(\sigma_+^k((\, R \;\;|\;\; -M \,)))_{k=1,\dots,L+1} \tag{7.1}$$

and the matrix $T := \mathrm{col}(M_i)_{i=0,\dots,L}$.

Define $E := \{ r \in \mathbb{R}^{1\times(L+1)g} \mid rT = 0\}$. E is the set of constant left annihilators of T; in fact, E is a vector space.

The following example clarifies the notions just introduced.

Example 7.1. Consider

$$(\, R \;\;|\;\; -M \,)(\xi) = \begin{pmatrix} \xi & 1 & | & -1 \\ \xi & \xi^2+1 & | & -\xi+1 \end{pmatrix}. \tag{7.2}$$

In this case

$$(\, R \;\;|\;\; -M \,)_\Xi = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \xi & -1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{7.3}$$

and

$$T = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}. \tag{7.4}$$

The space E is obtained as

$$\begin{aligned} E = \langle (\,0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0\,)\,, (\,0 \quad 0 \quad 0 \quad 0 \quad 1 \quad 0\,)\,, \\ (\,0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1\,)\,, (\,1 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0\,)\,, (\,1 \quad 0 \quad 0 \quad -1 \quad 0 \quad 0\,)\rangle. \end{aligned} \tag{7.5}$$

Let us now examine the conditions under which a trajectory $(w, \ell) \in \mathcal{B}_f$ is externally concatenable with zero; that is, $0 \wedge w \in \mathcal{B}_{ext}$. These conditions correspond to a system of linear equations involving w, ℓ, and their derivatives and define a polynomial differential operator which is in fact a state map. The rows of the corresponding polynomial matrix turn out, as stated in the following proposition, to have an interpretation in terms of a set of generators of E, and of the matrix

$$\begin{pmatrix} (\,R \quad | \quad -M\,)_\Xi \\ 0_{g\times(d+q)} \end{pmatrix}.$$

PROPOSITION 7.2. *Let an hybrid representation* (2.6) *be given with ℓ observable from w. Assume that for every input-output partition of (w, ℓ) there exists at least one component of ℓ chosen as input.*

A trajectory $(w, \ell) \in \mathcal{B}_f$ is externally concatenable with zero, that is, $0 \wedge w \in \mathcal{B}_{ext}$, if and only if given any set $\{v_1, \ldots, v_s\}$ of generators of E, there holds

$$\left(v_i \begin{pmatrix} (\,R \quad | \quad -M\,)_\Xi \\ 0_{g\times(d+q)} \end{pmatrix} \left(\frac{d}{dt}\right) \begin{pmatrix} w \\ \ell \end{pmatrix} \right)(0) = 0, \tag{7.6}$$

$i = 1, \ldots, s$.

Proof. See the appendix. ◻

We can now state the main result regarding systems in hybrid form with ℓ observable from w.

THEOREM 7.3. *Let a system be described in hybrid form as in* (2.6)*, and let ℓ be observable from w. Assume that for every input-output partition of (w, ℓ) there exists at least one component of ℓ chosen as input. The matrix $X \in \mathbb{R}^{\bullet\times(q+d)}[\xi]$ defines a (w, ℓ)-induced state map for the closure of the external system corresponding to* (2.6)*; that is,*

$$\begin{aligned} (\,R \quad -M\,) \left(\frac{d}{dt}\right) \begin{pmatrix} w \\ \ell \end{pmatrix} &= 0, \\ X \left(\frac{d}{dt}\right) \begin{pmatrix} w \\ \ell \end{pmatrix} &= x \end{aligned} \tag{7.7}$$

define a state model for the closure of the external behavior corresponding to (2.6) *if and only if for each constant matrix V whose rows generate E there exist a constant matrix A and a polynomial matrix B such that*

$$V \begin{pmatrix} (\,R \quad | \quad -M\,)_\Xi \\ 0 \end{pmatrix} = AX + B\,(\,R \quad | \quad -M\,)$$

and the variable x does not impose smoothness constraints on the trajectories of the closure of the external behavior of (2.6).

Remark 7.3. To check whether x does not impose smoothness constraints on the trajectories of the closure of the external behavior, we can proceed as follows. Assume without loss of generality that $\begin{pmatrix} R & -M \end{pmatrix}$ is of full row rank g, and note that by unimodular transformations (7.7) can be brought to the form

$$\begin{aligned} R_1'\left(\frac{d}{dt}\right)w &= 0, \\ R_2'\left(\frac{d}{dt}\right)w &= \ell, \\ X_1\left(\frac{d}{dt}\right)w &= -X_2\left(\frac{d}{dt}\right)\ell + x, \end{aligned} \tag{7.8}$$

where $R_1' \in \mathbb{R}^{g' \times q}$, $R_2' \in \mathbb{R}^{d \times q}$, $X_i \in \mathbb{R}^{n \times \bullet}$, $i = 1, 2$, $X = \begin{pmatrix} X_1 & X_2 \end{pmatrix}$, and $g = g' + d$. Again using unimodular transformations, we can modify this description to

$$R_1'\left(\frac{d}{dt}\right)w = 0, \tag{7.9}$$

$$R_2'\left(\frac{d}{dt}\right)w = \ell, \tag{7.10}$$

$$(X_1 + X_2 R_2')\left(\frac{d}{dt}\right)w = x. \tag{7.11}$$

Note that, given (7.9), (7.11), proper eliminability of x could be checked following the procedure illustrated in Remark 6.1. However, this property can be checked on the basis of the original equations, since each $(g' + n) \times (g' + n)$ minor of

$$\begin{pmatrix} R_1' & 0 \\ X_1 + X_2 R_2' & -I_n \end{pmatrix} \tag{7.12}$$

corresponds uniquely to a $(g' + n + d) \times (g' + n + d)$ minor of

$$\begin{pmatrix} R_1' & 0 & 0 \\ R_2' & -I_d & 0 \\ X_1 & X_2 & -I_n \end{pmatrix} \tag{7.13}$$

and therefore to a $(g' + n + d) \times (g' + n + d)$ minor of

$$\begin{pmatrix} R & -M & 0 \\ X_1 & X_2 & -I_n \end{pmatrix}, \tag{7.14}$$

obtained from a submatrix including the d columns corresponding to ℓ (that is, the columns of (7.14) from the $(q+1)$th up to the $(q+d)$th one). Therefore x is properly eliminable from the equations (7.7) if and only if among all $(g+n)\times(g+n)$ submatrices of (7.14) which include the d columns corresponding to ℓ (that is, the columns of (7.14) from the $(q+1)$th up to the $(q+d)$th one), there exists one of maximal determinantal degree which includes all columns corresponding to x (that is, the columns of (7.14) from the $(q + d + 1)$th up to the $(q + d + n)$th one).

Remark 7.4. Theorem 7.3 may be restated as follows: if no input-output partition of (w, ℓ) exists such that ℓ consists entirely of outputs for the full behavior, X defines a state-inducing map if and only if x does not impose any smoothness constraint on the trajectories of the closure of the external behavior, and the span over $\mathbb{R}$ of the rows of X contains the vector space

$$\left\{ r \begin{pmatrix} \mathrm{col}(\sigma_+^i \, (\, R \quad | \quad -M \,))_{i=1,\ldots,L} \\ 0_{g\times(q+d)} \end{pmatrix} \;\; | \;\; r \in E \right\} \; (\mathrm{mod} \; (\, R \quad | \quad -M \,)), \tag{7.15}$$

defined as the set of equivalence classes determined by the equivalence $\overset{(R|-M)}{\sim}$ on the vector space

$$\left\{ r \begin{pmatrix} \mathrm{col}(\sigma_+^i \, (\, R \quad | \quad -M \,))_{i=1,\ldots,L} \\ 0_{g\times(q+d)} \end{pmatrix} \;\; | \;\; r \in E \right\}. \tag{7.16}$$

This equivalent formulation, together with Proposition 7.1, yields the following characterization of minimality:

COROLLARY 7.4. *X defines a (w, ℓ)-induced minimal state map for the external system corresponding to* (2.6) *if and only if either* (1) *there exists an input-output selection in (w, ℓ) in which ℓ is entirely composed of outputs for the full behavior and the rows of X form a basis for* $(\, R \quad | \quad -M \,)_\Xi$ $(\mathrm{mod}\ (\, R \quad -M \,))$ *or* (2) *the rows of X form a basis for the vector space*

$$\left\{ r \begin{pmatrix} \mathrm{col}(\sigma_+^i \, (\, R \quad | \quad -M \,))_{i=1,\ldots,L} \\ 0_{g\times(q+d)} \end{pmatrix} \;\; | \;\; r \in E \right\} \; (\mathrm{mod} \; (\, R \quad | \quad -M \,)). \tag{7.17}$$

Remark 7.5. State-space equations are straightforwardly computed once the state map is given, analogously to the kernel representations case.

The results exposed up to this point suggest the following algorithm for the computation of a state map for a system in hybrid form with ℓ observable from w.

ALGORITHM 4 (Construction of a state map for the external behavior of a system in hybrid form with ℓ observable from w).

Data: $(\, R \quad | \quad -M \,) \in \mathbb{R}^{g\times(q+d)}[\xi]$, of degree L and full row rank, M right prime.

Output:] $X \in \mathbb{R}^{\bullet\times(d+q)}[\xi]$ inducing through $x = X(\frac{d}{dt})\binom{w}{\ell}$ a state for the external behavior of the system described in hybrid form by $(\, R \quad | \quad -M \,)$.

Step 1. Set $(\, R \quad | \quad -M \,)^0 := (\, R \quad | \quad -M \,)$ and compute $(\, R \quad | \quad -M \,)^{k+1} := \sigma_+^k \, (\, R \quad | \quad -M \,)$, $k = 1, \ldots, L+1$.

Step 2. Invoke algorithm ℓ-outputs (cf. Remark 7.2).

Step 3. If *True* then

Step 4. $X := \mathrm{col}((\, R \quad | \quad -M \,)^k)_{k=1,\ldots,L}$.

Step 5. *Stop.*

Step 6. Else

Step 7. Compute $v_1, \ldots, v_s$ in $\mathbb{R}^{1\times g(L+1)}$ such that $\langle v_1, \ldots, v_s \rangle$ equals the space $\{r \in \mathbb{R}^{1\times(L+1)g} \mid r\mathrm{col}(M_i)_{i=0\ldots L} = 0\}$.

Step 8. $S := \mathrm{col}((\, R \quad | \quad -M \,)^k)_{k=1,\ldots,L+1}$.

Step 9. $X := VS$.

Comment. Due to the smoothness result given at the beginning of section 6, X defines a properly eliminable latent variable.

Step 10. *Stop.*

The following example illustrates the construction of a state-inducing map for a hybrid representation with observable latent variables.

Example 7.2. Let

$$\begin{pmatrix} 0 & 1 & 1 \\ \frac{d}{dt}-1 & \frac{d}{dt}+1 & 1 \\ \frac{d}{dt} & \frac{d}{dt}-1 & \frac{d}{dt} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} \frac{d}{dt}-1 & 0 \\ \frac{d}{dt} & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \ell_1 \\ \ell_2 \end{pmatrix} \tag{7.18}$$

be a hybrid representation of $\mathcal{B}_{ext}$. It is easy to see that ℓ_1 can be chosen as an output for the full system, but in any input-output partition of (w, ℓ), ℓ_2 has to be chosen as an input.

The matrix

$$S = \begin{pmatrix} (\, R \quad | \quad -M \,)_{\Xi} \\ 0_{3\times 5} \end{pmatrix}$$

is

$$S = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 1 & 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{7.19}$$

and the matrix T is

$$T = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{7.20}$$

E can be computed as

$$\begin{aligned} E = \langle (\, 0 \quad 1 \quad -1 \quad 0 \quad 0 \quad 0 \,),\ (\, 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \,),\ (\, 0 \quad 0 \quad 0 \quad 1 \quad -1 \quad 0 \,), \\ (\, 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1 \,) \rangle, \end{aligned} \tag{7.21}$$

and this yields

$$X = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{7.22}$$

as a state-inducing map for the external behavior of the system described by (7.18). By choosing X as $X = \left(\begin{smallmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 & 0 \end{smallmatrix}\right)$ a minimal state is obtained.

Remark 7.6. The results exposed up to this point provide us with a technique for computing state maps for systems in hybrid form for which ℓ is not observable from w. For ease of exposition, we will limit the investigation to the case in which M is of full column rank d; the case in which M has column rank $d' < d$ can be dealt with similarly. Note that any non–right prime matrix M can be factored as $M = \bar{M}F$, with $\bar{M}$ a right prime matrix and F a full row rank matrix;. The following result holds.

PROPOSITION 7.5. *Let a hybrid representation of a latent variable system* (2.6) *be given. Let M be factored as $M = \bar{M}F$, with F a full row rank right divisor of M and $\bar{M}$ right prime. Then the closure of the external behavior of* (2.6) *and the closure of that of*

$$R\left(\frac{d}{dt}\right)w = \bar{M}\left(\frac{d}{dt}\right)\ell \tag{7.23}$$

are the same.

Proof. See the appendix. □

Let us examine the consequences of Proposition 7.5: given a system with ℓ nonobservable from w, factoring out of M an appropriate right divisor F yields a hybrid representation of a system which has the same external behavior of the original one (modulo the usual closedness issues) and the latent variable observable from the manifest ones. This provides us with a technique to tackle the problem of construction of state maps for nonobservable systems. The underlying idea is the following: given $(R \quad | \quad -M)$ with M non–right prime, extract a full row rank right factor G from M, getting a representation $(R \quad -\bar{M})$ with the same external behavior and $\bar{M}$ right prime. Computation of a polynomial matrix X_{obs} that induces a state for this system can be accomplished according to Algorithm 4. Now partition X_{obs} as $X_{obs} := (X_w \quad X_\ell)$.

Now note that $\forall\ \bar{\ell} \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^d)$ such that $(w, \bar{\ell})$ belongs to the full behavior associated with $(R \quad -\bar{M})$ there exists $\ell \in \mathcal{L}_1^{loc}(\mathbb{R}, \mathbb{R}^d)$ such that $G(\frac{d}{dt})\ell = \bar{\ell}$. Moreover, $(w, \ell)^T \in \mathcal{B}_f((R \quad | \quad -M))$.

This suggests that

$$x := X_{obs}\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \bar{\ell} \end{pmatrix} = X_{obs}\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ G(\frac{d}{dt})\ell \end{pmatrix} = (X_w \quad X_\ell G)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}$$

is a good candidate for a state for the closure of the external behavior of the system associated with $(R \quad | \quad -M)$.

PROPOSITION 7.6. *Let $(R \quad | \quad -M)$ be given, with M a non–right prime matrix of full column rank. Let $M = \bar{M}G$, with G a full row rank matrix and $\bar{M}$ right prime.*

Assume that $X_{obs} := (X_w \quad X_l)$ is a state-inducing map for the external behavior of the system in hybrid form associated with $(R \quad -\bar{M})$. Then

$$X := (X_w \quad X_l G) \tag{7.24}$$

defines a state-inducing map for the closure of the external behavior of the system described in hybrid form by $(R \quad | \quad -M)$.

Proof. See the appendix. □

Remark 7.7. Minimal state maps are obtained by choosing X_{obs} to be a minimal state-inducing map for the system associated with $(R \quad -\bar{M})$.

The above proposition is illustrated in the following example.

Example 7.3. Let the following system in hybrid form be given:

$$\begin{pmatrix} \frac{d}{dt} & 1 \\ \frac{d}{dt} & \frac{d}{dt}^3 + 1 \end{pmatrix} w = \begin{pmatrix} \frac{d}{dt} - 1 \\ \frac{d}{dt}^2 - \frac{d}{dt} \end{pmatrix} \ell. \tag{7.25}$$

The system has ℓ nonobservable from w, since $\xi - 1$ is a nontrivial greatest right factor for $M(\xi)$. The associated system with ℓ observable from w is

$$\begin{pmatrix} \frac{d}{dt} & 1 \\ \frac{d}{dt} & \frac{d}{dt}^3 + 1 \end{pmatrix} w = \begin{pmatrix} 1 \\ \frac{d}{dt} \end{pmatrix} \ell. \tag{7.26}$$

Computing a state-inducing map for the full system described by (7.26) yields

$$(R \quad | \quad -M)_\Xi = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \xi^2 & -1 \\ 0 & 0 & 0 \\ 0 & \xi & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{7.27}$$

and the matrix T is

$$T = (1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0 \quad 0)^T. \tag{7.28}$$

T has left nullspace described by

$$\begin{pmatrix} 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{7.29}$$

and therefore a (minimal) state-inducing map for the external observable behavior is obtained multiplying

$$\begin{pmatrix} (R \quad | \quad -M)_\Xi \\ 0_{2\times 5} \end{pmatrix}$$

on the left by the first three rows of (7.29), yielding

$$\begin{pmatrix} 1 & -\frac{d}{dt} & 0 \\ 1 & \frac{d^2}{dt^2} & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{7.30}$$

The last column of this matrix corresponds to $X_{obs,\ell}$ as in Proposition 7.6. The state for the external behavior of the nonobservable system is therefore induced by

$$\begin{pmatrix} 1 & -\frac{d}{dt} & 0 \\ 1 & \frac{d^2}{dt^2} & -(\frac{d}{dt} - 1) \\ 0 & 1 & 0 \end{pmatrix}. \tag{7.31}$$

The following equations can be written for the state induced by the map in (7.31):

$$\begin{pmatrix} \frac{d}{dt} & 1 & 0 \\ 0 & \frac{d}{dt} & 0 \\ 1 & 0 & \frac{d}{dt} \\ 0 & 0 & 1 \end{pmatrix} x + \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} w \\ \ell \end{pmatrix} = 0. \tag{7.32}$$

8. State maps for systems in image form. Image representations

$$w = M\left(\frac{d}{dt}\right)\ell, \tag{8.1}$$

where $w \in (\mathbb{R}^q)^\mathbb{R}$, $\ell \in (\mathbb{R}^d)^\mathbb{R}$, $M \in \mathbb{R}^{q\times d}[\xi]$, of the behavior of a linear time-invariant differential system have been introduced in section 2 in connection with the notion of controllability: the behavior of a system has an image representation if and only if the system is controllable.

In this section we consider the determination of state maps for the external behavior of systems whose full behavior is described by (8.1). These may be considered to be a special case of systems representable in hybrid form, with $R = I_q$. Therefore, before stating the results pertaining to image representations, let us make some important considerations in the light of the results given in the previous section.

Let us restrict attention to the case in which $M(\xi)$ of (8.1) is right prime; i.e., the latent variable ℓ is observable from w. Note that in a system whose behavior is described by (8.1), the latent variables can be chosen as playing the role of outputs in the full behavior. This is most easily seen by considering that for full column rank M a suitable subset R_1 of the columns of the $q \times q$ identity matrix exists such that

$$\begin{pmatrix} R_1 & M \end{pmatrix} \tag{8.2}$$

is nonsingular and, arranging the columns of a complementary canonical basis of R_1 in $\mathbb{R}^q$ in a matrix R_2, we have

$$\begin{pmatrix} R_1 & M \end{pmatrix}^{-1} R_2 \tag{8.3}$$

proper. Then the external variables corresponding to R_2 can be chosen as inputs, while those corresponding to R_1 and the latent variable ℓ can be chosen as outputs for the full behavior.

This result, together with Proposition 7.1, allows us to conclude that for observable image representations the dimensions of the minimal state space for the closure of the external and for the full behavior are equal.

Consider now the problem of determining a state-inducing map for an observable image representation. The following theorem is an immediate consequence of the considerations made so far.

THEOREM 8.1. *Let a system be represented in image form with ℓ observable from w. A polynomial $\bullet \times (q+d)$ matrix X defines a state map for the system* (8.1) *(i.e., $w = M(\frac{d}{dt})\ell$, $x = X(\frac{d}{dt})\binom{w}{\ell}$ defines a state system) if and only if there exist a constant matrix $A \in \mathbb{R}^{\bullet\times\bullet}$ and a polynomial matrix $B \in \mathbb{R}^{\bullet\times q}$ such that $\begin{pmatrix} I_q & -M \end{pmatrix}_{\Xi} = AX + B \begin{pmatrix} I_q & -M \end{pmatrix}$ and the variable x is properly eliminable from*

$$\begin{pmatrix} I_q & -M \end{pmatrix}\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix} = 0, \qquad X\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix} = x. \tag{8.4}$$

Note that $\sigma_+^k \begin{pmatrix} I_q & | & -M \end{pmatrix} = \begin{pmatrix} 0_{q\times q} & -\sigma_+^k M \end{pmatrix}$, $k \in \mathbb{N}$, and therefore any state map for the system (8.1), after suitable rearrangement of the equations, may be considered to be ℓ induced. This suggests the following algorithm for the computation of a state map for the external behavior of the system described by (8.1). Note that it is effectively a restatement of Algorithm 4.

ALGORITHM 5 (Construction of a state map for the external behavior of a system in image form with ℓ observable from w).

Data: $M \in \mathbb{R}^{q\times d}[\xi]$, of degree L, M right prime.

Output: $X \in \mathbb{R}^{\bullet\times d}[\xi]$ inducing through $x = X(\frac{d}{dt})\ell$ a state for the external behavior of the system described in image form by (8.1).

Step 1. Set $M^0 := M$ and compute $M^{k+1} := \sigma_+^k M$, $k = 1, \ldots, L$.

Step 2. $X := \mathrm{col}(M^k)_{k=1,\ldots,L}$.

Step 3. *Stop.*

Remark 8.1. The case in which M of (8.1) is not right prime, i.e., the case in which ℓ is not observable from w, can be dealt with in a manner completely analogous to that described in Remark 7.6.

As noted above, $\sigma_+^k\,(\,I_q \quad | \quad -M\,) = (\,0_{q\times q} \quad -\sigma_+^k M\,)$, $k \in \mathbb{N}$, and therefore the structure of the space Ξ_M is particularly important for the determination of state maps. In view of the results exposed in the next section, let us pursue further investigation of the structure of the space Ξ_M. Without loss of generality (possibly, permuting the rows) consider $M(\xi)$ partitioned as

$$M = \begin{pmatrix} N \\ D \end{pmatrix} \tag{8.5}$$

with D nonsingular and ND^{-1} proper. Equivalently, choose D as a nonsingular $d\times d$ submatrix of M of maximal determinantal degree.

Let us state the following two propositions, which are of independent interest and yield the main result regarding the structure of the space Ξ_M.

PROPOSITION 8.2. *Let $N \in \mathbb{R}^{p\times d}[\xi]$, $D \in \mathbb{R}^{d\times d}[\xi]$, $\det(D) \neq 0$, be two polynomial matrices such that ND^{-1} is proper. Then $\Xi_N \subseteq \Xi_D$.*

Proof. See the appendix. ◻

PROPOSITION 8.3. *Let $D \in \mathbb{R}^{d\times d}[\xi]$ be a nonunimodular polynomial matrix with* $\det(D) \neq 0$. *Then $\Xi_D = \{r \in \mathbb{R}^{1\times d}[\xi] \mid rD^{-1}$ strictly proper$\}$.*

Proof. See the appendix. ◻

The next proposition states the main result regarding the structure of the space Ξ_M.

PROPOSITION 8.4. $\Xi_M = \Xi_D = \{r \in \mathbb{R}^{1\times d}[\xi] \mid rD^{-1}$ *is strictly proper*$\}$.

Proof. See the appendix. ◻

The interest in considering state maps for systems represented in image form arises not only from the controllability issue but also from the connections among image representations and the notion of transfer function as given in the behavioral framework. This is the subject of next section.

9. Transfer functions and state maps. The purpose of this section is to make contact with the algebraic approach to the realization problem, put forward in [4] and extensively studied by Fuhrmann [2].

Consider the input-output system

$$P\left(\frac{d}{dt}\right) y = Q\left(\frac{d}{dt}\right) u \tag{9.1}$$

with $P \in \mathbb{R}^{p\times p}[\xi]$, $Q \in \mathbb{R}^{p\times m}[\xi]$, $\det(P) \neq 0$, and $P^{-1}Q$ proper. The function $G := P^{-1}Q \in \mathbb{R}^{p\times m}(\xi)$ is called the *transfer function* of (9.1).

The latent variable system

$$\begin{aligned} u &= D\left(\frac{d}{dt}\right)\ell, \\ y &= N\left(\frac{d}{dt}\right)\ell \end{aligned} \tag{9.2}$$

with $D \in \mathbb{R}^{m\times m}[\xi]$, $N \in \mathbb{R}^{p\times m}[\xi]$, $\det(D) \neq 0$, and ND^{-1} proper defines an input-output system with transfer function $G = ND^{-1}$.

It can be shown that two systems have the same transfer function if and only if they have the same controllable part (see [9, p. 248]). The (unique) controllable system which has a given transfer function $G \in \mathbb{R}_+^{p\times m}(\xi)$ can be obtained by making a left coprime factorization $G = P^{-1}Q$ of G and considering (9.1) or by making a right factorization $G = ND^{-1}$ and considering (9.2); if the latter factorization is right coprime, then (9.2) will be an observable image representation of the controllable system with transfer function G (see [9, pp. 249, 250]).

Note that the algorithms described in sections 6 and 8 can be directly applied in order to obtain a state-space realization of a system with a given transfer function.

Transfer functions play a prominent role in control theory, since they provide a natural framework in many engineering applications. The concept of realization as put forward in [4], associated with the notion of an input-output map, is intimately connected with the notion of transfer function. Not surprisingly, therefore, many formalizations of the notion of state starting from an input-output or transfer function point of view have been given in the past.

The algorithms proposed in our paper are very akin to those of Fuhrmann [2]. The module structure on which his approach is based, has many connections with left and right factorizations of transfer functions. In particular, the state space corresponding to a right factorization ND^{-1} of a transfer function is defined therein to be isomorphic to the vector space K_D defined as

$$K_D := \{f \in \mathbb{R}^{1\times d}[\xi] \mid fD^{-1} \in \mathbb{R}_+^{1\times d}(\xi)\} \tag{9.3}$$

(cf. [2, Lemma 2-15, p. 11, and Theorem 10-2, p. 41]). The connection with the result of Proposition 8.4 is evident.

10. Conclusions. In modeling physical systems the most natural way of proceeding is to write a set of high-order differential equations possibly with algebraic constraints among the variables. When it comes to simulation of the corresponding system, however, state-space equations are the most natural representation to use. Therefore the need arises to compute the latter from the former. In this paper a characterization of state-inducing maps has been given for systems given in kernel or in hybrid representations. This characterization suggests immediately algorithms to actually perform a computation of the state function from which state-space equations are easily recovered.

Appendix A. Notation.

$\mathbb{N}$ natural numbers (0 is not included).
$\mathbb{Z}_+$ nonnegative integers.
$\mathbb{R}$ real numbers.
2^A set whose members are the subsets of A.
$\mathbb{R}[\xi]$ polynomials with real coefficients.
$\mathbb{R}_+(\xi)$ proper rational functions.
e_i the ith vector of a canonical basis vector in $\mathbb{R}^{1\times\bullet}$.
$\mathbb{R}^{g\times q}$ $g \times q$ real matrices.
$\mathbb{R}^{\bullet\times q}$ real matrices with q columns.
$\mathrm{col}(r_1, \ldots, r_n)$ the matrix

$$\begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix}.$$

$\mathrm{diag}(x_k)_{k=1,\ldots,r}$ $r \times r$ diagonal matrix with diagonal elements x_k.

$\mathbb{R}^{g\times q}[\xi]$ $g \times q$ polynomial matrices in the indeterminate ξ.
$\mathbb{R}^{\bullet\times q}$ polynomial matrices in the indeterminate ξ with q columns.
$\mathbb{R}_+^{g\times q}(\xi)$ $g \times q$ matrices of strictly proper rational functions.
$(W)^T$ maps from W to T.
$\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^q)$ infinitely differentiable functions from $\mathbb{R}$ to $\mathbb{R}^q$.
$\mathcal{L}_1^{loc}(\mathbb{R}; \mathbb{R}^q)$ locally integrable functions from $\mathbb{R}$ to $\mathbb{R}^q$.
$\langle r_1, \dots, r_n\rangle$ space spanned by the vectors r_i.
π_w projection on the w variables: $\pi_w(w, \ell) := w$.
$\circ$ composition of maps.
$[p]$ equivalence class with representative p.

Appendix B. Proofs.

B.1. Proof of Theorem 2.1. That $w \in \overline{\pi_w(\mathcal{B}_f)}^{closure}$ implies $R_1'(\frac{d}{dt})w = 0$ is easy to see. To prove the converse, let $\mathcal{B}_1$ be the behavior of $R_1'(\frac{d}{dt})w = 0$, and observe that $\mathcal{B}_1 \bigcap \mathcal{C}^\infty$ is dense in $\mathcal{B}_1$. Let $M_2' \in \mathbb{R}^{n_1\times n_2}[\xi]$. Obviously $M_2'(\frac{d}{dt})$ maps $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{n_2})$ into $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{n_1})$. Since M_2' is of full row rank, this map is surjective. (In order to see this, use the Smith form of M_2'.) Hence for all $w \in \mathcal{B}_1 \bigcap \mathcal{C}^\infty$ there exists a $(w, \ell) \in \mathcal{B}_f \bigcap \mathcal{C}^\infty$. This shows $\mathcal{B}_1 = \overline{\pi_w(\mathcal{B}_f)}^{closure}$. ◻

B.2. Proof of Proposition 3.1. We will prove only the Markovian case, the state-space case being entirely equivalent. The "if" part is trivial. To show the "only if" case, assume that (2.4) satisfies the concatenability condition. Without loss of generality we can assume that R has full row rank. Also, there exists a unimodular $U \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ such that $R' := UR$ is in row reduced form, meaning that the matrix formed by the coefficients of the highest powers in ξ of the rows of $R'(\xi)$ has full row rank. It is easy to see that systems with kernel representations defined by R and R' are the same. We will now show that R' is a first order polynomial matrix. Assume the contrary. Write R' in input-output form:

$$P\left(\frac{d}{dt}\right) w_1 = Q\left(\frac{d}{dt}\right) w_2 \tag{B.1}$$

with $\det(P) \neq 0$ and $P^{-1}Q$ proper. The assumption that R' is not first order implies that P is not. From the assumption that $R(\frac{d}{dt})w = 0$ is Markovian, it follows that also $P(\frac{d}{dt})w_1 = 0$ is. Now let w_1', w_1'' be solutions of $P(\frac{d}{dt})w = 0$ with $w_1'(0) = w_1''(0)$. Since $\det(P) \neq 0$, w_1' and w_1'' are also $\mathcal{C}^\infty$ and by the state property are concatenable. In order to obtain a contradiction it suffices therefore to prove Proposition 3.1 for autonomous systems. This, however, is an immediate consequence of the following lemma.

LEMMA B.1. *Let the autonomous system $R(\frac{d}{dt})w = 0$ with $R \in \mathbb{R}^{q\times q}[\xi]$, $\det R \neq 0$, be Markovian. Then this system admits the kernel representation*

$$Fw + E\frac{d}{dt}w = 0 \tag{B.2}$$

with E, $F \in \mathbb{R}^{q\times q}$ and $\det(E\xi + F) \neq 0$

Proof. Let $\mathcal{B}$ be the behavior of $R(\frac{d}{dt})w = 0$. Let $P(\frac{d}{dt})w = 0$ be the corresponding representation in row reduced form, as in the above proof. Write it as

$$P_0 w + P_1\frac{d}{dt}w + \cdots + P_L\frac{d}{dt}w = 0. \tag{B.3}$$

We need to prove that it is first order. Assume that this is not the case and that $P_L \neq 0$ and $L \geq 2$.

Denote with L_k, $k = 1, \dots, q$, the highest order of differentiation of w_k in (B.3). Note that there is at least one $L_k \geq 2$. Introduce the auxiliary variables z_i^k defined as

$$z_i^k := \frac{d^i w_k}{dt^i}, \tag{B.4}$$

$k = 1, \dots, q$, $i = 0, \dots, L_k - 1$, and define

$$z := \begin{pmatrix} z_0^1 & z_1^1 & \dots & z_{L_1-1}^1 & \dots & z_0^q & \dots & z_{L_q-1}^q \end{pmatrix}. \tag{B.5}$$

Now consider the system with latent variable z, described by the equations

$$\begin{aligned} \frac{d}{dt} z &= F z, \\ w_k &= z_0^k, \quad k = 1, \dots, q, \end{aligned} \tag{B.6}$$

where the entries of the $\sum_{k=1}^q L_k \times \sum_{k=1}^q L_k$ matrix F are determined from (B.3) and the definitions (B.4). Equations (B.6) represent a system in hybrid form with latent variable z; it external behavior coincides with that described by (B.3), as can be checked by applying the latent-variable-elimination theorem.

However, the external behavior of (B.6) does not enjoy the Markovianity property. In fact, (B.6) has exactly one solution (w, z) for each initial condition vector

$$\begin{pmatrix} z_0^1(0) & z_1^1(0) & \dots & z_{L_1-1}^1(0) & \dots & z_0^q(0) & \dots & z_{L_q-1}^q(0) \end{pmatrix}. \tag{B.7}$$

This contradicts Markovianity, since two solutions (w, z), (w', z') of (B.6) with $z_0^k(0) = z_0'^k(0)$, $k = 1, \dots, q$, cannot be concatenated unless also $z_j^k(0) = z_j'^k(0)$, $j = 1, \dots, L_k - 1$, $k = 1, \dots, q$. $\square$

B.3. Proof of Proposition 6.1. The behavior described by $R(\frac{d}{dt})w(t) = 0$ with $w \in \mathcal{L}_1^{loc}$ is the set of all w for which

$$\int_{-\infty}^{+\infty} w^T(t) \left(R\left(-\frac{d}{dt}\right)^T \right) f(t) dt = 0 \tag{B.8}$$

for all testing functions $f(t)$ (that is, f is a $\mathcal{C}^\infty$ vector-valued function with compact support).

(Only if) Assume that $w \in \mathcal{B}$ and $0 \wedge w \in \mathcal{B}$. Define $R^k := \sigma_+^k(R)$. We will show that $(R^k(\frac{d}{dt})w)(0) = 0$ for $k = 1, 2, \dots$. Consider

$$\int_{-\infty}^{+\infty} (0 \wedge w)^T(t) \left(R\left(-\frac{d}{dt}\right)^T \right) f(t) dt. \tag{B.9}$$

This obviously equals

$$\int_{0}^{+\infty} w^T(t) \left(R\left(-\frac{d}{dt}\right)^T \right) f(t) dt. \tag{B.10}$$

Since $R^k(\frac{d}{dt})w$ is locally integrable $\forall\ k = 1, \dots, \deg(R)$, (B.10) may be integrated by parts and equals

$$\int_0^{+\infty} \left(R\left(\frac{d}{dt}\right) w(t) \right)^T f(t) dt + \sum_{k=1}^{L} \sum_{j=k}^{L} (-1)^{k-1} (w^{(j-k)}(0))^T R_j^T f^{(k-1)})(0). \tag{B.11}$$

But since $w \in \mathcal{B}$, the integral in (B.11) is zero. Since $0 \wedge w \in \mathcal{B}$, (B.9) and hence (B.11) are also zero, and therefore so is the double sum in (B.11). Hence, due to the arbitrariness of the testing function f,

$$\left(R^k\left(\frac{d}{dt}\right)w\right)(0) = 0 \tag{B.12}$$

$\forall\ k = 1, \dots, \deg(R)$.

(If) Assume that $w \in \mathcal{B}$ satisfies $(R^k(\frac{d}{dt})w)(0) = 0$ for $k = 1, 2, \dots$. We want to show that $0 \wedge w \in \mathcal{B}$. To prove this, we have to prove that the integral (B.9) is zero for all testing functions f. Proceeding as above, integrating by parts, (B.11) is obtained. Now the claim is obtained by noting that since $w \in \mathcal{B}$, $R(\frac{d}{dt})w = 0$ holds and therefore the integral is zero, while the double sum in (B.11) is zero by assumption. ∎

B.4. Proof of Theorem 6.2. Let us first prove the following

LEMMA B.2. *Let $\mathcal{B}$ be the behavior of* (2.4). *Let X_1, $X_2 \in \mathbb{R}^{\bullet\times q}[\xi]$. Assume that for all $w \in \mathcal{B} \bigcap \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^q)$ there holds*

$$\left\{\left(X_1\left(\frac{d}{dt}\right)w\right)(0) = 0\right\} \Longrightarrow \left\{\left(X_2\left(\frac{d}{dt}\right)w\right)(0) = 0\right\}. \tag{B.13}$$

Then there exist $A \in \mathbb{R}^{\bullet\times\bullet}$ and $B \in \mathbb{R}^{\bullet\times\bullet}[\xi]$ such that

$$X_2(\xi) = AX_1(\xi) + B(\xi)R(\xi). \tag{B.14}$$

Proof. We will prove this lemma only in the case that (2.4) defines a controllable system. The general case is left to the reader. Using the Smith form for R it follows that there exist unimodular matrices U and V such that $URV = (\, I \quad 0\,)$. Let $v := V^{-1}(\frac{d}{dt})w$. Then $w \in \mathcal{B}$ if and only if $(\, I \quad 0\,)\,v = 0$. Define $X_1' := X_1V$ and $X_2' := X_2V$. Partition v, X_1', and X_2' as $\binom{v_1}{v_2}$, $(\, X_{11}' \quad X_{12}'\,)$, $(\, X_{21}' \quad X_{22}'\,)$, with the partition induced by $(\, I \quad 0\,)$. Then for any $v_2 \in \mathcal{C}^\infty$ there holds

$$\left\{\left(X_{12}'\left(\frac{d}{dt}\right)v_2\right)(0) = 0\right\} \Longrightarrow \left\{\left(X_{22}'\left(\frac{d}{dt}\right)v_2\right)(0) = 0\right\}. \tag{B.15}$$

This implies that there exists a matrix $A \in \mathbb{R}^{\bullet\times\bullet}$ such that

$$X_{22}'\left(\frac{d}{dt}\right) = AX_{12}'\left(\frac{d}{dt}\right). \tag{B.16}$$

This yields that X_2' is of the form

$$X_2'(\xi) = AX_1'(\xi) + B(\xi)\,(\, I \quad 0\,)\,. \tag{B.17}$$

Now postmultiply by V^{-1}. ∎

This lemma yields the claim of the theorem as follows.

(Only if) Assume that $X(\frac{d}{dt})$ is a state map. Then $x = X(\frac{d}{dt})w$ is properly eliminable from

$$\begin{aligned} R\left(\frac{d}{dt}\right)w &= 0, \\ X\left(\frac{d}{dt}\right)w &= x. \end{aligned} \tag{B.18}$$

Moreover, $(X(\frac{d}{dt})w)(0) = 0$ implies that w is concatenable with the zero trajectory. Proposition 6.1 states that concatenability with zero is equivalent to $(R_\Xi(\frac{d}{dt})w)(0) = 0$. One has to apply now the above lemma with $X_1 = X$ and $X_2 = R_\Xi$.

(If) Assume that (6.7) holds. Recall (cf. the beginning of section 6) that $R_\Xi(\frac{d}{dt})w$ is absolutely continuous. Now consider $w \in \mathcal{B}$ such that $X(\frac{d}{dt})w$ is continuous at $t = 0$. Then $R_\Xi = AX + BR$ implies that $(BR)(\frac{d}{dt})w$ is continuous at $t = 0$, so that $(R_\Xi(\frac{d}{dt})w)(0) = (AX(\frac{d}{dt})w)(0) + (BR(\frac{d}{dt})w)(0)$ and $(X(\frac{d}{dt})w)(0) = 0$ imply $(R_\Xi(\frac{d}{dt})w)(0) = 0$ since $(BR)(\frac{d}{dt})w = 0$ and $(BR)(\frac{d}{dt})w$ is continuous at $t = 0$. Therefore by Proposition 6.1 one concludes that $0 \wedge w \in \mathcal{B}$. By assumption, $x = X(\frac{d}{dt})w$ is properly eliminable from (B.18), and therefore (B.18) defines a state-space system with external behavior Ker $R(\frac{d}{dt})$. ◻

B.5. Proof of Proposition 6.4. The claim follows directly by applying the second of the lemmas below. To get to that result, let us first consider the following lemma.

LEMMA B.3. *Let $p \in \mathbb{R}^{1\times q}[\xi]$, $R \in \mathbb{R}^{q\times g}[\xi]$. Then*

$$\sigma_+(pR) = (\sigma_+ p)R + p(0)(\sigma_+ R). \tag{B.19}$$

Proof. Let $p = \begin{pmatrix} p_1 & \dots & p_q \end{pmatrix}$, $p_i = \sum_{j=0}^n p_{ji}\xi^j$, and $R = \mathrm{col}(R_i)_{i=1,\dots,q}$, $R_i \in \mathbb{R}^{1\times g}[\xi]$. Note that

$$pR = \sum_{i=1}^q \left(\sum_{j=0}^n p_{ji}\xi^j R_i\right) \tag{B.20}$$

and therefore that

$$\sigma_+(pR) = \sigma_+\left(\sum_{i=1}^q \left(\sum_{j=0}^n p_{ji}\xi^j R_i\right)\right), \tag{B.21}$$

which is equivalent to $\sum_{i=1}^q \sigma_+(\sum_{j=0}^n p_{ji}\xi^j R_i)$. This is equivalent to

$$\sum_{i=1}^q \left(\sum_{j=1}^n p_{ji}\xi^{j-1} R_i + p_{0i}\sigma_+(R_i)\right), \tag{B.22}$$

which yields $\sum_{i=1}^q (\sigma_+ p_i)R_i + \sum_{i=1}^q p_{0i}(\sigma_+ R_i)$ and the claim. ◻

This lemma explains how σ_+ acts on vector multiples of a given matrix. The next one shows how σ_+ acts on unimodular matrix multiples.

LEMMA B.4. *Let R, R' be matrices related as*

$$R = UR' \tag{B.23}$$

for a unimodular U. Then

$$R_\Xi = VR'_\Xi + BR' \tag{B.24}$$

with V a constant full column rank matrix and B a polynomial matrix.

Proof. The proof follows trivially from Lemma B.3 and the fact that a unimodular matrix U has $\det(U(0)) \neq 0$. ◻

B.6. Proof of Proposition 6.6. Let us first prove that the nonzero rows of R_Ξ form a basis of Ξ_R if and only if R is in row reduced form.

Define R'_Ξ to be the submatrix of R_Ξ consisting of the nonzero rows of R_Ξ. Let us prove sufficiency. Assume that R'_Ξ has not full row rank. Then there exists at least one row which is a linear combination of the others. Since $R_\Xi := \mathrm{col}(R^k)_{k=1,\ldots,L}$, the highest coefficient vector of this row is the same as that of the corresponding row of R, and is a linear combination of the highest coefficient vectors of the other rows of R. But this contradicts row reducedness. The proof of necessity goes along the same lines.

Note that the result just proven implies necessity of the claim of the proposition. Let us prove sufficiency. Let module(R) be the module of $R^{1\times q}[\xi]$ generated by the rows of R. We have to prove that the intersection of Ξ_R and module(R) consists of the zero vector only.

We prove this as follows. Let ν_i, $i = 1, \ldots, g$, be the degree of the ith row R_i of R, and assume that the rows of R have been ordered so that $\nu_1 = \nu_2 = \cdots = \nu_{g'} > \nu_{g'+1} \geq \cdots \geq \nu_g$.

Let now $y \in \Xi_R \bigcap \mathrm{module}(R)$. Note first that, since $y \in \Xi_R$, $\deg(y) \leq \nu_1 - 1$. Then note that $y \in \mathrm{module}(R)$ implies $y = xR$ for some $x = (\, x_1, \ldots, x_g \,) \in \mathbb{R}^{1\times g}[\xi]$. From the predictable degree property of R [3, p. 387], we conclude that $\deg(y) = \max_{1\leq i\leq g,\ x_i\neq 0}\{\deg(x_i) + \nu_i\}$ and therefore, since $\deg(y) \leq \nu_1 - 1$, that $x_i = 0$, $1 \leq i \leq g'$.

Assume now that $x_i = 0$ for $g' + 1 \leq i \leq \bar{g} < g$, $x_{\bar{g}+1} \neq 0$. By the predictable degree property of R, $\deg(y) \geq \nu_{\bar{g}+1}$. Since $y \in \Xi_R$, $y = \sum_{j=1}^{g}\sum_{i=1}^{\nu_j} \alpha_{ij}(\sigma_+^i R_j)$, for suitable scalars $\alpha_{ij} \in \mathbb{R}$. Since $\deg(y) \geq \nu_{\bar{g}+1}$, at least one of the α_{ij}'s with $1 \leq j \leq \bar{g}$ must be nonzero, since the only generators of Ξ_R of degree $\geq \nu_{\bar{g}+1}$ are to be found among the vectors $(\sigma_+^i R_j)$, $1 \leq j \leq \bar{g}$. This implies that the highest coefficient of y is a linear combination of the first $\bar{g}$ rows of R_{hc}, the highest row coefficient matrix of R. On the other hand, since $x_i = 0$ for $1 \leq i \leq \bar{g}$, and $x_{\bar{g}+1} \neq 0$, the highest coefficient of $y = xR$ is a linear combination of the last $g - \bar{g}$ rows of R_{hc}. But this implies that the first $\bar{g}$ and the last $g - \bar{g}$ rows of R_{hc} generate the same vector; this, by row reducedness, is possible if and only if this vector is zero. Therefore $x_i = 0$ for all i; that is, $y = 0$ as was to be proven. □

B.7. Proof of Proposition 7.1. The system represented by (2.6) has ℓ observable from w. Therefore it allows a representation of the form

$$N\left(\frac{d}{dt}\right)w = \ell,$$

$$R_1'\left(\frac{d}{dt}\right)w = 0, \tag{B.25}$$

with R_1' of full row rank. Now partition R_1' as $(\, P_1 \quad Q_1 \,)$ with $P_1^{-1}Q_1$ proper. This induces a partition (y, u) on w in inputs u and outputs y. The proposition follows immediately from examining the minors of

$$\begin{pmatrix} P_1 & Q_1 & 0 \\ N_1 & N_2 & -I \end{pmatrix}. \tag{B.26}$$

In fact, sufficiency can be proven as follows. Since (y, ℓ) consists of outputs for the full system, it follows that

$$\begin{pmatrix} P_1 & 0 \\ N_1 & -I \end{pmatrix} \tag{B.27}$$

has maximal degree among the minors of (B.26). By Corollary 6.7 this implies that the minimal dimension of the state space for the full behavior equals $\deg(\det(\left(\begin{smallmatrix} P_1 & 0 \\ N_1 & -I \end{smallmatrix}\right))) = \deg(\det(P_1))$. Now note that $\deg(\det(P_1))$ equals the minimal dimension of the state space for the external behavior since $\det(P_1)$ has maximal degree among the minors of R_1'. This yields the claim.

As for necessity, note that if the minimal dimensions of the state space for the external and the full behavior are the same, $\deg(\det(P_1))$ equals the maximal degree of the minors of (B.26). Since $\det(\left(\begin{smallmatrix} P_1 & 0 \\ N_1 & -I \end{smallmatrix}\right))$ has degree $\deg(\det(P_1))$, it has maximal degree among the minors of (B.26) and therefore the corresponding partition of the (w, ℓ) variables, that is, (y, ℓ), is a set of outputs for the full system. ∎

B.8. Proof of Proposition 7.2. Before proving the proposition, let us point out the following general result. Assume that in a hybrid representation ℓ is observable from w; then $w(t) = 0 \ \forall \ t < 0$ implies $\ell(t) = 0 \ \forall \ t < 0$. This is proven as follows. Observability implies that there exists a polynomial differential operator $F(\frac{d}{dt})$ such that $F(\frac{d}{dt})w = \ell$; then for $t < 0$ $(F(\frac{d}{dt})w)(t) = (F(\frac{d}{dt})0)(t) = 0$.

Let us turn to the proof of the proposition.

(Only if) External concatenability with zero is equivalent to

$$\int_{-\infty}^{+\infty} \begin{pmatrix} 0 \wedge w \\ \ell \end{pmatrix}^T (t) \left(\begin{pmatrix} R & -M \end{pmatrix}^T \left(-\frac{d}{dt} \right) \right) f(t) dt = 0 \tag{B.28}$$

for every testing function f.

Observability of ℓ from w and the remark made above imply that integration can be considered in $[0, +\infty)$ only:

$$\int_{0}^{+\infty} \begin{pmatrix} w \\ \ell \end{pmatrix}^T (t) \left(\begin{pmatrix} R & -M \end{pmatrix}^T \left(-\frac{d}{dt} \right) \right) f(t) dt = 0. \tag{B.29}$$

Note that since ℓ is a locally integrable function, there exists an absolutely continuous function $\mathcal{L}$ such that $\frac{d}{dt}\mathcal{L} = \ell$ almost everywhere.

Then $R(\frac{d}{dt})w = M(\frac{d}{dt})\ell$ may be written as

$$R\left(\frac{d}{dt}\right) w = M\left(\frac{d}{dt}\right) \frac{d}{dt}\mathcal{L}, \tag{B.30}$$

and if

$$\begin{pmatrix} R & -M \end{pmatrix} = \begin{pmatrix} R_0 & -M_0 \end{pmatrix} + \begin{pmatrix} R_1 & -M_1 \end{pmatrix} \xi + \cdots + \begin{pmatrix} R_L & -M_L \end{pmatrix} \xi^L, \tag{B.31}$$

(B.30) corresponds to the polynomial matrix

$$\begin{aligned} \begin{pmatrix} R & -M \end{pmatrix}' := \begin{pmatrix} R_0 & 0 \end{pmatrix} + \begin{pmatrix} R_1 & -M_0 \end{pmatrix} \xi + \cdots \\ + \begin{pmatrix} R_L & -M_{L-1} \end{pmatrix} \xi^L + \begin{pmatrix} 0 & -M_L \end{pmatrix} \xi^{L+1}, \end{aligned} \tag{B.32}$$

which is more conveniently written as $\sum_{j=0}^{L+1} \begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \xi^j$, defining $M_{-1} := 0$, $R_{L+1} := 0$.

Equation (B.29) corresponds then to

$$\int_{0}^{+\infty} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^T (t) \left(\sum_{j=0}^{L+1} (-1)^j \begin{pmatrix} R_j & -M_{j-1} \end{pmatrix}^T \frac{d^j}{dt^j} f(t) \right) dt = 0. \tag{B.33}$$

Analogously to what has been done at the beginning of section 6 it is possible to prove that the Ξ-matrix of (B.32) induces an absolutely continuous function. It is then possible to integrate by parts the left-hand side of (B.33), and this yields

$$\begin{aligned}
&\int_0^{+\infty} \left(\sum_{j=0}^{L+1} \begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j)} (t) \right)^T f(t)dt \\
&- \sum_{j=1}^{L+1} \left(\left(\begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j-1)} \right) (0) \right)^T f(t) \\
&+ \sum_{j=2}^{L+1} \left(\left(\begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j-2)} \right) (0) \right)^T \frac{d}{dt} f(t) + \cdots \\
&+ (-1)^{L+1} \left(\begin{pmatrix} 0 & -M_L \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix} (0) \right)^T f^{(L)}(0),
\end{aligned} \tag{B.34}$$

where $g^{(j)}$ denotes the jth derivative of a function g (and the function itself in case $j = 0$).

Equation (B.34) can be rewritten as

$$\begin{aligned}
&\int_0^{+\infty} \left(\sum_{j=0}^{L+1} \begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j)} (t) \right)^T f(t)dt \\
&+ \sum_{k=1}^{L+1} \left(\sum_{j=k}^{L+1} (-1)^k \left(\begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j-k)} (0) \right)^T \right) f^{(k-1)}(0).
\end{aligned} \tag{B.35}$$

Since $(0 \wedge w, \ell) \in \mathcal{B}_f$, (B.35) equals zero for all testing functions f. Now note that the integral in (B.35) is zero, since $(w, \ell) \in \mathcal{B}_f$. Therefore $(0 \wedge w, \ell) \in \mathcal{B}_f$ implies

$$\sum_{k=1}^{L+1} \left(\sum_{j=k}^{L+1} (-1)^k \left(\begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j-k)} (0) \right)^T \right) f^{(k-1)}(0) = 0. \tag{B.36}$$

Arbitrariness of the testing function f then yields the set of equations

$$\sum_{j=k}^{L+1} (-1)^k \left(\begin{pmatrix} R_j & -M_{j-1} \end{pmatrix} \begin{pmatrix} w \\ \mathcal{L} \end{pmatrix}^{(j-k)} \right) (0) = 0, \tag{B.37}$$

$k = 1, \ldots, L+1$, which is more conveniently written as

$$\begin{aligned}
\sum_{j=k}^{L} \left(\begin{pmatrix} R_j & -M_j \end{pmatrix} \begin{pmatrix} w \\ \ell \end{pmatrix}^{(j-k)} \right) (0) &= M_{k-1}\mathcal{L}(0), \qquad k = 1, \ldots, L, \\
0 &= M_L \mathcal{L}(0).
\end{aligned} \tag{B.38}$$

In matrix form, (B.38) reads as

$$\left(\left(\begin{matrix} \begin{pmatrix} R & -M \end{pmatrix}_\Xi \\ 0 \end{matrix} \right) \left(\frac{d}{dt} \right) \begin{pmatrix} w \\ \ell \end{pmatrix} \right) (0) = \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix} \mathcal{L}(0). \tag{B.39}$$

Now let $V := \mathrm{col}(v_1, \ldots, v_s)$, with $\{v_i\}_{i=1,\ldots,s}$ a set of generators of E. Multiply both sides of (B.39) by V. This yields

$$V\left(\left(\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = \left(\left(V\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = 0, \tag{B.40}$$

which is the claim of the proposition.

(If) Assume that

$$\left(\left(V\left(\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\left(\frac{d}{dt}\right)\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = 0, \tag{B.41}$$

where V is a set of generators of E.

Then in particular

$$\left(V\left(\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = V\begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix}\mathcal{L}(0), \tag{B.42}$$

with $\mathcal{L}$ such that $\frac{d}{dt}\mathcal{L} = \ell$ almost everywhere. Equation (B.42) is equivalent to

$$V\left(\left(\left(\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) - \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix}\mathcal{L}(0)\right) = 0 \tag{B.43}$$

and therefore

$$\left(\left(\begin{pmatrix} R & -M \\ & 0 \end{pmatrix}_{\Xi}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) - \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix}\mathcal{L}(0) \tag{B.44}$$

belongs to the vector space generated by the columns of

$$\begin{pmatrix} M_0 \\ \vdots \\ M_L \end{pmatrix};$$

that is, there exist $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, d$, such that (B.44) equals

$$\begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix}\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}. \tag{B.45}$$

Consider now the function $\bar{\mathcal{L}}$ defined as follows: $\bar{\mathcal{L}}(t) := \mathcal{L}(t)\ \forall t \neq 0$ and

$$\bar{\mathcal{L}}(0) := \mathcal{L}(0) + \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_d \end{pmatrix}.$$

Then $\frac{d}{dt}\bar{\mathcal{L}} = \ell$ almost everywhere and

$$\left(\left(\begin{matrix} (\,R & -M\,)_\Xi \\ & 0 \end{matrix}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) - \begin{pmatrix} M_0 \\ M_1 \\ \vdots \\ M_L \end{pmatrix} \bar{\mathcal{L}}(0) = 0. \tag{B.46}$$

Now consider $(0 \wedge w, 0 \wedge \ell)$. To show that it belongs to the full behavior, we prove that

$$R\left(\frac{d}{dt}\right)(0 \wedge w) = M\left(\frac{d}{dt}\right)\frac{d}{dt}(0 \wedge \bar{\mathcal{L}}). \tag{B.47}$$

Using the notation introduced in (B.32), note that (B.47) holds if and only if

$$\int_{-\infty}^{+\infty} \begin{pmatrix} 0 \wedge w \\ 0 \wedge \bar{\mathcal{L}} \end{pmatrix}^T (t) \left(\sum_{j=0}^{L+1} (-1)^j \,(\, R_j \quad -M_{j-1}\,)^T \frac{d^j f}{dt^j}(t)\right) dt = 0, \tag{B.48}$$

that is, if and only if

$$\int_{0}^{+\infty} \begin{pmatrix} w \\ \bar{\mathcal{L}} \end{pmatrix}^T (t) \left(\sum_{j=0}^{L+1} (-1)^j \,(\, R_j \quad -M_{j-1}\,)^T \frac{d^j f}{dt^j}(t)\right) dt = 0. \tag{B.49}$$

The Ξ-matrix of (B.32) induces an absolutely continuous function. Therefore (B.49) can be integrated by parts and, with manipulations completely analogous to those of the necessity part of the proof, this yields

$$\begin{aligned} &\int_{0}^{+\infty} \left(\sum_{j=0}^{L+1} (\, R_j \quad -M_{j-1}\,) \begin{pmatrix} w \\ \bar{\mathcal{L}} \end{pmatrix}^{(j)} (t)\right)^T f(t)dt \\ &+ \sum_{k=1}^{L+1} \left(\sum_{j=k}^{L+1} (-1)^k \left((\, R_j \quad -M_{j-1}\,) \begin{pmatrix} w \\ \bar{\mathcal{L}} \end{pmatrix}^{(j-k)} (0)\right)^T\right) f^{(k-1)}(0). \end{aligned} \tag{B.50}$$

The integral is zero, since $\frac{d}{dt}\bar{\mathcal{L}} = \ell$ and $(w, \ell) \in \mathcal{B}_f$ by assumption. The double sum is zero, since by assumption each addendum of the outermost sum is zero (cf. (B.46)). The claim follows. ☐

B.9. Proof of Theorem 7.3. Note first that Lemma B.2 holds also for the kernel representation $(\, R \quad -M\,)(\frac{d}{dt})\binom{w}{\ell} = 0$ of the full behavior. Let us now prove necessity. If $X(\frac{d}{dt})$ is a (w, ℓ)-induced state map for the external behavior, then x is properly eliminable, and $(X(\frac{d}{dt})\binom{w}{\ell})(0) = 0$ implies external concatenability with zero. Proposition 7.2 states that external concatenability with zero is equivalent to

$$\left(V \left(\begin{matrix} (\,R & -M\,)_\Xi \\ & 0 \end{matrix}\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = 0.$$

Now apply Lemma B.2 with $X_1 = X$ and

$$X_2 = V \left(\begin{matrix} (\,R & -M\,)_\Xi \\ & 0 \end{matrix}\right).$$

Sufficiency is proven as follows.

$$\left(V\left(\begin{matrix}(\,R & -M\,)_\Xi \\ & 0\end{matrix}\right)\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}$$

is an absolutely continuous function (cf. the remark made at the beginning of section 6). Since

$$\left(V\left(\begin{matrix}(\,R & -M\,)_\Xi \\ & 0\end{matrix}\right)\right)\left(\frac{d}{dt}\right) = AX\left(\frac{d}{dt}\right) + B\,(\,R \quad -M\,)\left(\frac{d}{dt}\right), \tag{B.51}$$

for each $(w,\ell) \in \mathcal{B}_f$ such that $(X(\frac{d}{dt})\binom{w}{\ell})(0) = 0$ and such that $X(\frac{d}{dt})\binom{w}{\ell}$ is continuous at $t = 0$,

$$\left(V\left(\begin{matrix}(\,R & -M\,)_\Xi \\ & 0\end{matrix}\right)\right)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}(0) = 0$$

holds, since

$$B\,(\,R \quad -M\,)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix} = V\left(\begin{matrix}(\,R & -M\,)_\Xi \\ & 0\end{matrix}\right)\Bigg)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix} - AX\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}$$

is continuous at $t = 0$. Then Proposition 7.2 can be applied, and external concatenability with zero follows. Moreover, since x is properly eliminable, the external behavior of the state-space representation is the same as that of the original hybrid representation. This concludes the proof. ∎

B.10. Proof of Proposition 7.5. Consider the system described by (2.6). Following Theorem 2.1, computation of a kernel description of (the closure of) its external behavior is done by determining a unimodular matrix U such that $UM = \binom{0}{M_2'}$, with M_2' of full row rank. Partitioning U, R, and M according to the number of rows of M_2' as

$$U := \begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}, \quad M := \begin{pmatrix} M_1 \\ M_2 \end{pmatrix}, \quad R := \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}, \tag{B.52}$$

the description of the external behavior is given as $R_1'(\frac{d}{dt})w = 0$, with $R_1' = U_{11}R_1 + U_{12}R_2$.

Now assume $M = \bar{M}F$, with F a full row rank right factor of M. Note that

$$0 = U_{11}M_1 + U_{12}M_2 = U_{11}\bar{M}_1F + U_{12}\bar{M}_2F = (U_{11}\bar{M}_1 + U_{12}\bar{M}_2)F \tag{B.53}$$

if and only if $U_{11}\bar{M}_1 + U_{12}\bar{M}_2 = 0$ by the fact that F has full row rank. Therefore U eliminates the latent variable in the description $R(\frac{d}{dt})w = \bar{M}(\frac{d}{dt})\ell$, and $R_1'(\frac{d}{dt})w = 0$ describes the closure of the external behavior of this system as well. ∎

B.11. Proof of Proposition 7.6. G is a nonsingular $d\times d$ matrix and therefore $\forall\, \bar{\ell} \in \mathcal{L}_1^{loc}(\mathbb{R};\mathbb{R}^d)\ \exists\, \ell \in \mathcal{L}_1^{loc}(\mathbb{R};\mathbb{R}^d)$ such that $\bar{\ell} = G(\frac{d}{dt})\ell$. Therefore

$$\begin{aligned}\left(X_{obs}\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \bar{\ell} \end{pmatrix}\right)(0) = 0 &\Longleftrightarrow \left(X_{obs}\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ G\left(\frac{d}{dt}\right)\ell \end{pmatrix}\right)(0) = 0 \\ &\Longleftrightarrow \left((\,X_{obs,w} \quad X_{obs,l}G\,)\left(\frac{d}{dt}\right)\begin{pmatrix} w \\ \ell \end{pmatrix}\right)(0) = 0.\end{aligned}$$

Since X_{obs} induces a state map, $(X_{obs}(\frac{d}{dt})\binom{w}{\bar{\ell}})(0) = 0$ implies that $(w, \bar{\ell})$ is externally concatenable with zero; moreover, external concatenability in zero for $(w, \bar{\ell})$ is equivalent to external concatenability in zero for (w, ℓ), as $\bar{\ell} = G(\frac{d}{dt})\ell$. Let us now prove that $x = X(\frac{d}{dt})\binom{w}{\ell}$ does not impose additional smoothness constraints on the trajectories of the external behavior of $\begin{pmatrix} R & -M \end{pmatrix}(\frac{d}{dt})\binom{w}{\ell} = 0$. By unimodular transformations, which preserve the proper eliminability of a latent variable, we can bring the equations

$$\begin{aligned} R\left(\frac{d}{dt}\right) w &= M\left(\frac{d}{dt}\right)\ell, \\ X_w\left(\frac{d}{dt}\right) &= -(X_\ell G)\left(\frac{d}{dt}\right)\ell + x \end{aligned} \tag{B.54}$$

to the form

$$\begin{aligned} R_1'\left(\frac{d}{dt}\right) w &= 0, \\ R_2'\left(\frac{d}{dt}\right) w &= G\left(\frac{d}{dt}\right)\ell, \\ X_w\left(\frac{d}{dt}\right) &= -(X_\ell G)\left(\frac{d}{dt}\right)\ell + x \end{aligned} \tag{B.55}$$

and, again by unimodular operations, to

$$\begin{aligned} R_1'\left(\frac{d}{dt}\right) w &= 0, \\ R_2'\left(\frac{d}{dt}\right) w &= G\left(\frac{d}{dt}\right)\ell, \\ (X_w + X_\ell R_2')\left(\frac{d}{dt}\right) w &= x. \end{aligned} \tag{B.56}$$

Observe that $x = (X_w + X_\ell R_2')(\frac{d}{dt})w$ is a state variable for the behavior Ker $R_1'(\frac{d}{dt})$ and therefore that it is properly eliminable from

$$\begin{aligned} R_1'\left(\frac{d}{dt}\right) w &= 0, \\ (X_w + X_\ell R_2')\left(\frac{d}{dt}\right) w &= x; \end{aligned} \tag{B.57}$$

the claim follows. □

B.12. Proof of Proposition 8.2. Let N_i be the ith row of N. Since $N_i D^{-1}$ is proper, there exists a rational vector $n_i := \sum_{k=0}^{\infty} n_{ik}\xi^{-k}$ such that $N_i D^{-1} = n_i$. Write $D = D_0 + D_1\xi + \cdots + D_L\xi^L$ and $N_i = N_{i0} + N_{i1}\xi + \cdots + N_{iL'}$, $L' \leq L$. $N_i = n_i D$ yields the following equalities:

$$\begin{aligned} N_{i0} &= n_{i0}D_0 + n_{i1}D_1 + \cdots + n_{iL}D_L, \\ N_{i1} &= n_{i0}D_1 + n_{i1}D_2 + \cdots + n_{iL-1}D_L, \\ &\vdots \\ N_{iL'} &= n_{i0}D_{L'} + n_{i1}D_{L'+1} + \cdots + n_{iL-L'}D_L. \end{aligned} \tag{B.58}$$

These equalities imply $N_i = n_{i0}D + n_{i1}\sigma_+(D) + \cdots + n_{iL}\sigma_+^L(D)$ and therefore $\sigma_+(N_i) = n_{i0}\sigma_+(D) + n_{i1}\sigma_+^2(D) + \cdots + n_{iL-1}\sigma_+^L(D)$. Then $\sigma_+(N_i) \in \Xi_D$, and the same holds for $\sigma_+^2(N_i)$, $\sigma_+^3(N_i), \dots$. This yields the claim. ∎

B.13. Proof of Proposition 8.3. The inclusion $\Xi_D \subseteq \{r \mid rD^{-1}$ is strictly proper$\}$ can be proven as follows. Let D_i be the ith row of D, $D_i = \sum_{k=0}^{L} D_{ik}\xi^L$. Then $\sigma_+ D_i = \xi^{-1}D_i - \xi^{-1}D_{i0}$, and therefore $\sigma_+ D_i D^{-1} = \xi^{-1}e_i - \xi^{-1}D_{i0}D^{-1} \in \mathbb{R}_+^{1\times d}(\xi)$. Analogously, $\sigma_+^2 D_i = \xi^{-2}D_i - \xi^{-2}D_{i0} - \xi^{-1}D_{i1}$ and therefore $\sigma_+^2 D_i D^{-1} = \xi^{-2}e_i - \xi^{-2}D_{i0}D^{-1} - \xi^{-1}D_{i1}D^{-1} \in \mathbb{R}_+^{1\times d}(\xi)$ and similarly for all iterations of σ_+ and for all rows of D.

The opposite inclusion may be proven as follows. Take $r \in \{r' \mid r'D^{-1}$ is strictly proper$\}$. Then there exists $n \in \mathbb{R}_+^{1\times d}(\xi)$ such that $r = nD$. Write

$$n = \sum_{k=1}^{\infty} n_k \xi^{-k}, \tag{B.59}$$

$n_k \in \mathbb{R}^{1\times d}$, and denote

$$D := D_0 + D_1\xi + \cdots + D_L\xi^L \tag{B.60}$$

and

$$r := r_0 + r_1\xi + \cdots + r_{L'}\xi^{L'}, \tag{B.61}$$

where without loss of generality we can assume $L' \leq L-1$. $r = nD$ yields, equating powers of ξ,

$$\begin{aligned} r_0 &= n_1D_1 + n_2D_2 + \cdots + n_LD_L, \\ r_1 &= n_1D_2 + n_2D_3 + \cdots + n_{L-1}D_L, \\ &\ \vdots \\ r_j &= n_1D_{j+1} + n_2D_{j+2} + \cdots + n_{L-j}D_L, \\ &\ \vdots \end{aligned} \tag{B.62}$$

and this implies

$$r = n_1\sigma_+(D) + n_2\sigma_+^2(D) + \cdots + n_L\sigma_+^L(D) \tag{B.63}$$

and therefore that $r \in \Xi_D$, as we were to prove. ∎

B.14. Proof of Proposition 8.4. The first equality follows from the fact that $\Xi_M = \Xi_N + \Xi_D$ and from Proposition 8.2. The second equality can be proven by applying Proposition 8.3. ∎

Acknowledgment. The first author thanks the Department of Mathematics of the University of Groningen and the Dutch Systems and Control Theory Network for the hospitality and financial support granted for visits to Groningen during which the present results were worked out.

REFERENCES

[1] R. D'Andrea and F. Paganini, *Interconnection of uncertain behavioral systems for robust control*, in Proc. 32nd IEEE Conf. on Decision and Control, San Antonio, TX, 1993, pp. 3642–3647.

[2] P. A. Fuhrmann, *Linear Systems and Operators in Hilbert Space*, McGraw-Hill, New York, 1981.

[3] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[4] R. Kalman, P. Falb, and M. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, New York, 1969.

[5] W. Neven, *Polynomial methods in system theory*, Master's thesis, Dept. of Mathematics and Computing Science, Eindhoven University of Technology, Eindhoven, the Netherlands, 1988.

[6] J. W. Polderman, *Proper elimination of latent variables*, in Proc. 12th World Congress IFAC, 1992, pp. 73–76.

[7] C. Praagman, *Inputs, outputs and states in the representation of time series*, in Analysis and Optimization of Systems, A. Bensoussan and J. Lions, eds., Springer-Verlag, Berlin, 1988, pp. 1069–1078.

[8] J. C. Willems, *From time series to linear system. Part* 1*—finite dimensional linear time invariant systems*, Automatica J. IFAC, 22 (1986), pp. 561–580.

[9] J. C. Willems, *Models for dynamics*, Dynamics Reported, 2 (1989), pp. 171–269.

[10] J. C. Willems, *Paradigms and puzzles in the theory of dynamical systems*, IEEE Trans. Automat. Control, 36 (1991), pp. 259–294.