KYBERNETIKA — VOLUME 37 (2001), NUMBER 4, PAGES 397-426

THE ALGEBRAIC STRUCTURE OF DELAY-DIFFERENTIAL SYSTEMS: A BEHAVIORAL PERSPECTIVE

HEIDE GLUESING-LUERSSEN, PAOLO VETTORI AND SANDRO ZAMPIERI

This paper presents a survey on the recent contributions to linear time-invariant delay-differential systems in the behavioral approach. In this survey both systems with commensurate and with noncommensurate delays will be considered. The emphasis lies on the investigation of the relationship between various systems descriptions. While this can be understood in a completely algebraic setting for systems with commensurate delays, this is not the case for systems with noncommensurate delays. In the study of this class of systems functional analytic methods need to be introduced and general convolutional equations have to be incorporated. Whenever it is possible, the results will be linked to the relevant control theoretic notions.

1. INTRODUCTION

Delay systems is a classical topic in the control literature due to the well-known fact that the presence of delays makes the controller synthesis more difficult. In recent years, the theory of delay systems has attracted new attention. This is mainly caused by the fact that the low cost of data transmission makes centralized control strategies more convenient. Indeed, in many practical situations it is now possible to control many remotely positioned coupled plants by means of a unique controller. The use of communication lines usually causes the presence of not negligible delays in the system.

Since delay systems are infinite-dimensional systems, they are usually treated with functional analytic methods, in particular by use of the theory of semigroups. These methods are well-suited for the qualitative analysis of a system, in particular for stability considerations. This analytical theory, however, will not be the topic of this paper and we refer the reader to the vast literature.

In this survey we want to present the state-of-the-art of the *behavioral approach to linear time-invariant delay-differential systems*. This approach is well-suited for the investigation of general structural control theoretic properties like controllability, input/output-structures, causality etc. For the class of linear time-invariant continuous-time systems described by ordinary differential equations (in the sequel simply called purely differential systems) the behavioral approach has been worked out in great detail and proved very successful, see [31]. For linear time-invariant multidimensional systems the behavioral theory is developing thanks to the fundamental paper [28]. According to the behavioral approach, a system is defined as a triple $\Sigma = (T, W, \mathcal{B})$, where T is the time set, W is the alphabet where the signals take on their values, and $\mathcal{B}$ is a subset of the space of all signals $\mathcal{W} := W^T$, which specifies what signals can occur in the given system. This subset is called the behavior of the system; it is the core of the description, since it defines the dynamics by fixing what signals are allowed and what signals are forbidden. In a continuous-time linear time-invariant system we have $T = \mathbb{R}$, $W = \mathbb{R}^q$, and $\mathcal{B}$ is a linear subspace of the signal space $\mathcal{W}$, which is invariant with respect to the forward shift operator. From a conceptual point of view, the difference between the classical approach and the behavioral approach to systems theory is that in the latter one a system is purely defined as the set of its possible trajectories, and not as an operator. However, in order to launch a mathematical control theory, one assumes that the behavior is actually given as the solution space of a system of equations. Hence we have an operator $R : \mathcal{W} \to \mathcal{V}$, with some space $\mathcal{V}$, such that

$$\mathcal{B} = \ker R = \{w \in \mathcal{W} | R(w) = 0\}.$$

The describing operator R is called a kernel representation of $\mathcal{B}$. In the behavioral theory the control theoretic properties of a system are defined purely in terms of its trajectories, that is via the set $\mathcal{B}$. Naturally, the mathematical theory aims at characterizing these properties in terms of kernel representations. At this point it becomes obvious that it is mandatory for the behavioral approach to understand the relationship between operators R and behaviors $\mathcal{B}$. Precisely, which operators give rise to the same behavior? In addition to kernel representations, one might have (or want) a description of the system as, say, the image of an operator $R : \mathcal{V} \to \mathcal{W}$

$$\mathcal{B} = \operatorname{im} R = \{w \in \mathcal{W} \mid \exists\, v \in \mathcal{V} : w = R(v)\}.$$

This is called an image representation and its existence is connected with the controllability of the system. Another representation which is a generalization of both the previous ones is the following

$$\mathcal{B} = R_1^{-1}(\operatorname{im} R_2) = \{w \in \mathcal{W} \mid \exists\, v \in \mathcal{V} :\ R_1(w) = R_2(v)\}$$

where $R_1 : \mathcal{W} \to \mathcal{U}$ and $R_2 : \mathcal{V} \to \mathcal{U}$ are two given operators. Such a description is called a latent variable representation; it does not only involve the so-called manifest variable w of the system (whose trajectories make up the behavior), but also a so-called latent variable v, which is just an auxiliary part of the systems description and whose evolution itself is not relevant for the behavior. Such auxiliary variables might arise directly in the modeling of the system; more importantly, they appear when two systems are interconnected, say in a feedback–loop or in a series interconnection. An important issue in the behavioral approach is the so-called latent variable elimination, which concerns the possibility of obtaining a kernel representation from a latent variable representation. The importance of latent variable representations and of latent variable elimination is widely discussed in [31].

In the behavioral approach to delay-differential systems all the operators arising above are delay-differential operators acting on suitable function spaces. In this survey we shall mainly address the "preliminary" questions raised above: uniqueness of kernel representations, existence of image representations, and latent variable elimination. At the moment the relationship between the various systems descriptions is not completely understood for delay-differential systems. Only partial results are available. Whenever possible, we will also link the systems descriptions to the corresponding control theoretic concepts. As it will turn out, a particular class of delay systems deserves special attention, that is the case where the delays are commensurate. In this case the set of associated operators turns out to carry a nice algebraic structure showing completely how to pass from one systems description to another. How this can be used for the investigation of control theoretic questions will be illustrated by addressing the issue of interconnecting systems.

2. MATHEMATICAL PRELIMINARIES

In this section we shall provide some mathematical notations and preliminary results which will be used in the sequel. For the purpose of this survey, it suffices to restrict to behaviors with signals which are smooth functions, i. e. contained in $\mathcal{E} := \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$. Notice that, as mentioned in the introduction, we consider behaviors defined over the full time axis $\mathbb{R}$. An algebraic theory for systems defined on $\mathbb{R}_+$ is yet unknown and seems unlike harder. With some additional work it is also possible to extend the theory to be presented here to certain larger function spaces (see [35, Sec. 7]).

Several operator algebras on $\mathcal{E}$ will play an important role in the behavioral theory of delay systems. We shall introduce these various algebras by starting from the smallest one, that is the one which is most closely related to the type of delay systems under investigation. It turns out that this algebra is not rich enough for an algebraic theory; motivated by some simple considerations, we shall show how a certain larger algebra naturally arises in this context. This is contained in the algebra of compact support distributions; the latter one will also be helpful for the behavioral theory of systems with delays.

Let the space $\mathcal{E}$ be equipped with the topology of uniform convergence in all derivatives on all compact sets; this turns $\mathcal{E}$ into a Fréchet space. We shall see later on that topological arguments will play a role only for the case of systems with noncommensurate delays. For the commensurate case algebraic arguments will suffice, so that in that case $\mathcal{E}$ can simply be regarded as a module over the ring of delay-differential operators to be introduced next. Let us begin by introducing the shift-operators σ_{t_0}, $t_0 \in \mathbb{R}$, defined as

$$(\sigma_{t_0} f)(t) = f(t - t_0)$$

for any function $f \in \mathcal{E}$. The real number t_0 is also called the delay. Notice that the operator σ_{t_0} can be defined also over the vector valued functions in $\mathcal{E}^q$ and that it induces a linear bijective map. Then the delay-differential operators under consideration are of the form

$$\mathcal{E} \longrightarrow \mathcal{E}, \quad f \longmapsto \sum_{i=0}^{N}\sum_{j=1}^{M} p_{ij} f^{(i)}(\cdot - t_j) = \sum_{i=0}^{N}\sum_{j=1}^{M} p_{ij}\sigma_{t_j} f^{(i)}, \tag{2.1}$$

where N, $M \in \mathbb{N}$, $p_{ij} \in \mathbb{R}$ are constant coefficients, and $t_j \geq 0$ are the delays.

It is standard to rewrite the operator given in (2.1) as follows. Consider the free $\mathbb{Z}$–module $t_1\mathbb{Z} + \ldots + t_M\mathbb{Z} \subseteq \mathbb{R}$ and let $\{\tau_1, \ldots, \tau_k\} \subseteq \mathbb{R}_+$ be a basis of this module. For each $\nu = (\nu_1, \ldots, \nu_k) \in \mathbb{Z}^k$ the composition $\sigma^\nu := \sigma_{\tau_1}^{\nu_1} \circ \ldots \circ \sigma_{\tau_k}^{\nu_k}$ describes the delay operator

$$(\sigma^\nu f)(t) = \left(\sigma_{\tau_1}^{\nu_1} \circ \ldots \circ \sigma_{\tau_k}^{\nu_k} f\right)(t) = f\left(t - \sum_{j=1}^{k} \nu_j \tau_j\right), \quad t \in \mathbb{R}.$$

Together with the ordinary differential operator $D = \frac{\mathrm{d}}{\mathrm{d}t}$, these delays form the operator algebra

$$\mathcal{R} := \mathbb{R}[D, \sigma, \sigma^{-1}] := \mathbb{R}[D, \sigma_{\tau_1}, \ldots, \sigma_{\tau_k}, \sigma_{\tau_1}^{-1}, \ldots, \sigma_{\tau_k}^{-1}].$$

Notice that $\mathcal{R}$ is a commutative subring of $\mathrm{End}_{\mathbb{C}}(\mathcal{E})$, the ring of all endomorphisms on $\mathcal{E}$, and therefore $\mathcal{E}$ is an $\mathcal{R}$–module. Furthermore, the linear independence of $\tau_1, \ldots, \tau_k$ over $\mathbb{Z}$ implies that $\mathbb{R}[D, \sigma] := \mathbb{R}[D, \sigma_{\tau_1}, \ldots, \sigma_{\tau_k}]$ is a polynomial ring in $k+1$ algebraically independent operators. By construction, each operator of the type (2.1) corresponds to an element $a(D, \sigma) \in \mathcal{R}$ and can be written as

$$a(D, \sigma) f := \sum_{i=0}^{N} \sum_{\substack{\nu \in \mathbb{Z}^k \\ \text{finite}}} a_{i\nu} D^i \sigma^\nu f, \tag{2.2}$$

where $a_{i\nu} \in \mathbb{R}$ are constant coefficients. In the case $k = 1$, we call these operators delay–differential operators with commensurate delays; otherwise we say that the delay–differential operator contains noncommensurate delays. The following result of Ehrenpreis will be crucial for the algebraic setting.

Proposition 2.1. [9, p. 697] *Each nonzero operator $a(D, \sigma) \in \mathcal{R}$ is surjective on $\mathcal{E}$.*

As we shall see in the sequel, even though the operator algebra $\mathcal{R}$ is general enough to define delay systems in the behavioral approach, it does not suffice to develop an algebraic theory. Therefore we need to introduce a larger algebra. In order to do so, we associate with each delay–differential operator $a(D, \sigma) \in \mathcal{R}$ its characteristic function, that is, we consider the mapping

$$a(D, \sigma) \longmapsto a(s, e^{-\tau_1 s}, \ldots, e^{-\tau_k s}) =: a^*(s), \tag{2.3}$$

where s is a complex variable. This yields an isomorphism of rings

$$\mathcal{R} \cong \mathcal{R}^* := \mathbb{R}[s, e^{-\tau_1 s}, \ldots, e^{-\tau_k s}, e^{\tau_1 s}, \ldots, e^{\tau_k s}].$$

Hence $\mathcal{R}^*$ is a ring having transcendence degree $k+1$ over $\mathbb{R}$ and being contained in $H(\mathbb{C})$, the ring of entire functions. The importance of the characteristic function $a^* \in H(\mathbb{C})$ rests on its capability to detect the exponential monomials in the solution space

$$\ker a(D,\sigma) := \{f \in \mathcal{E} \mid a(D,\sigma)f = 0\}.$$

Indeed, for each $\lambda \in \mathbb{C}$ and $k \in \mathbb{N}_0$ we have

$$t^k e^{\lambda t} \in \ker a(D,\sigma) \Longleftrightarrow \frac{a^*(s)}{(s-\lambda)^{k+1}} \in H(\mathbb{C}), \tag{2.4}$$

where the right hand side simply says that λ is a zero of a^* of multiplicity at least $k+1$. The equivalence (2.4) immediately implies

$$\ker b(D,\sigma) \subseteq \ker a(D,\sigma) \Longrightarrow \frac{a^*}{b^*} \in H(\mathbb{C}) \tag{2.5}$$

for each a, $b \in \mathcal{R}$. In fact, the converse is true as well and can be deduced from spectral synthesis (see [33, Thm. 5]), but will also follow from our considerations (see Theorem 2.2 (b) below).

Consider now the set

$$\mathcal{H} := \left\{ \frac{a}{b} \,\middle|\, a,\, b \in \mathcal{R},\, b \neq 0,\, \frac{a^*}{b^*} \in H(\mathbb{C}) \right\}, \tag{2.6}$$

which is a subring of the abstract quotient field $\mathbb{R}(D,\sigma)$ of the polynomial ring $\mathbb{R}[D,\sigma]$. Using a result from harmonic analysis, one deduces that $\mathcal{H}$ is an operator algebra contained in $\mathrm{End}_{\mathbb{C}}(\mathcal{E})$ and containing $\mathcal{R}$. We summarize these facts in

Theorem 2.2. The ring $\mathcal{H}$ can be described as follows.
(a) $\mathcal{H} = \left\{ \frac{a}{\phi} \,\middle|\, a \in \mathcal{R},\, \phi \in \mathbb{R}[D]\backslash\{0\},\, \frac{a^*}{\phi^*} \in H(\mathbb{C}) \right\}$.
(b) $\mathcal{H} = \left\{ \frac{a}{b} \in \mathbb{R}(D,\sigma) \,\middle|\, a,\, b \in \mathcal{R},\, \ker b(D,\sigma) \subseteq \ker a(D,\sigma) \right\}$.
(c) Let a, $b \in \mathcal{R}$, $b \neq 0$, and $ab^{-1} \in \mathcal{H}$. Then the mapping

$$\mathcal{E} \longrightarrow \mathcal{E}, \quad f \longmapsto a(D,\sigma)g, \text{ where } b(D,\sigma)g = f$$

is well-defined, $\mathbb{C}$-linear, and depends only on the quotient ab^{-1} (and not on its specific fractional representation via a and b). Denoting this map simply by $\frac{a}{b} : \mathcal{E} \to \mathcal{E}$, the ring $\mathcal{H}$ becomes a commutative subring of $\mathrm{End}_{\mathbb{C}}(\mathcal{E})$. This turns $\mathcal{E}$ into an $\mathcal{H}$-module. We call the operators $ab^{-1} \in \mathcal{H}$ delay-differential operators, too.

Part (a) is a result about exponential polynomials in harmonic analysis and has been proven in [2]. A slightly simpler proof can be found in [14, Thm. 5.8]. Part (b) can be deduced from (a) by some algebraic arguments along with the obvious fact that $a^*(\phi^*)^{-1} \in H(\mathbb{C})$ iff $\ker \phi(D) \subseteq \ker a(D,\sigma)$ whenever $\phi \in \mathbb{R}[D]$ and $a \in \mathcal{R}$; see also [14]. Part (c) is a simple consequence of (b) and follows by standard calculations;

see [12, Rem. 2.8] or [14, Sec. 3]. Notice that the function g does always exist thanks to Proposition 2.1.

The operator algebra $\mathcal{H}$ has been introduced first for the commensurate case in [12], where it also has been thoroughly studied with respect to its algebraic properties. In completely different shapes and for different purposes it has also appeared in earlier work in [29] and [19]. For the noncommensurate case the algebra $\mathcal{H}$ has been considered first in [14, 26] as well as in [38].

The topological arguments needed for the noncommensurate case will make it necessary to take also the full algebra of convolution operators defined on $\mathcal{E}$ into consideration. Let $\mathcal{D}'$ be the vector-space of complex-valued distributions on the space $\mathcal{D} := \{f \in \mathcal{E} \mid \operatorname{supp} f \text{ is compact}\}$, endowed with the usual inductive limit topology. Here supp denotes the support of a function or distribution. Furthermore, let $\mathcal{D}'_{\mathrm{c}} := \{T \in \mathcal{D}' \mid \operatorname{supp} T \text{ compact}\}$. We shall identify the distributions in $\mathcal{D}'_{\mathrm{c}}$ with $\mathcal{E}'$, that is, with their extension to distributions on $\mathcal{E}$; see [34, Thm. 24.2]. Recall that each distribution $T \in \mathcal{E}'$ induces a convolution operator $f \mapsto T * f$ and thus a continuous map from $\mathcal{E}$ to $\mathcal{E}$. In particular, $\mathcal{E}'$ is (up to isomorphism) contained in $\operatorname{End}_{\mathbb{C}}(\mathcal{E})$. Finally, denote by $\delta_a^{(k)}$ the kth derivative of the Dirac-distribution at $a \in \mathbb{R}$ [41, pp. 124–129]. In this setting, differentiation (resp. forward-shift by τ_j time units) corresponds to convolution with $\delta_0^{(1)}$ (resp. δ_{τ_j}). Precisely, for $a(D,\sigma) \in \mathcal{R}$ and $f \in \mathcal{E}$ we have $a(D,\sigma)f = a(\delta_0^{(1)}, \delta_{\tau_1}, \dots, \delta_{\tau_k}) * f \in \mathcal{E}$. Hence $\mathcal{R}$ is (up to isomorphism) a subring of the domain $\mathcal{E}'$. This observation has already been made in [18], where it was utilized for a transfer function approach to delay-differential systems.

According to a Paley–Wiener Theorem (see [3, pp. 27–28]), one can embed $\mathcal{E}'$ into $H(\mathbb{C})$. Indeed, the Laplace transform $\mathcal{L} : \mathcal{E}' \to H(\mathbb{C})$ which maps $T \in \mathcal{E}'$ onto

$$\mathcal{L}T : \mathbb{C} \longrightarrow \mathbb{C}, \quad s \longmapsto \langle T, e^{-s\cdot} \rangle$$

induces an isomorphism from $\mathcal{E}'$ onto the Paley–Wiener algebra

$$\mathcal{A} := \{f \in H(\mathbb{C}) \mid \exists A,\, B,\, C > 0\ \forall s \in \mathbb{C} : |f(s)| \le A(1+|s|)^B e^{C|\mathrm{Re}(s)|}\}. \tag{2.7}$$

It is not hard to verify [37, Thm. 4.35] that $\frac{a^*}{b^*} \in \mathcal{A}$ for all $\frac{a}{b} \in \mathcal{H}$. Using furthermore the identity $a^* = \mathcal{L}\big(a(\delta_0^{(1)}, \delta_{\tau_1}, \dots, \delta_{\tau_k})\big)$ for $a \in \mathcal{R}$, which is standard in distribution theory, one obtains that $\mathcal{H}$ is (up to the isomorphism $\frac{a}{b} \mapsto \mathcal{L}^{-1}\big(\frac{a^*}{b^*}\big) \in \mathcal{E}'$) a subalgebra of $\mathcal{E}'$ and the map $\frac{a}{b} \in \operatorname{End}_{\mathbb{C}}(\mathcal{E})$ defined in Theorem 2.2 (c) coincides with the associated convolution operator; cf. [11, Thm. 2.8] for the commensurate case. For any $h = \frac{a}{b} \in \mathcal{H}$ we will define $h^* := \frac{a^*}{b^*}$, which coincides with the Laplace transform of the compact support distribution associated with h. We will use the same notation also for matrices with entries in $\mathcal{H}$. Notice that the topology on $\mathcal{E}$ induces a topology on its dual $\mathcal{E}'$ which in turn leads to a topology on $\mathcal{A}$.

In the sequel we shall identify the various objects in the way described above. Hence we arrive at the following embeddings of commutative domains

$$\mathcal{R} \subseteq \mathcal{H} \subseteq \mathcal{A} \subseteq H(\mathbb{C}) \quad \text{and} \quad \mathcal{A} \subseteq \operatorname{End}_{\mathbb{C}}(\mathcal{E}).$$

Since $\mathcal{E}$ and $\mathcal{A}$ are both $\mathcal{A}$-modules, every matrix $R \in \mathcal{A}^{p\times q}$ induces two canonical maps, namely

$$R: \mathcal{E}^q \longrightarrow \mathcal{E}^p \quad \text{and} \quad R: \mathcal{A}^q \longrightarrow \mathcal{A}^p;$$

both of them will be important in the sequel. For ease of notation we shall denote both maps simply by R. The specific meaning will always be clear from the context. However, we shall use the notation

$$\ker_{\mathcal{E}} R \subseteq \mathcal{E}^q,\ \mathrm{im}_{\mathcal{E}} R \subseteq \mathcal{E}^p \ \text{ and } \ \ker_{\mathcal{A}} R \subseteq \mathcal{A}^q,\ \mathrm{im}_{\mathcal{A}} R \subseteq \mathcal{A}^p \tag{2.8}$$

for their respective kernels and images. For matrices R with entries in $\mathcal{H}$ we similarly define $\ker_{\mathcal{H}} R$ and $\mathrm{im}_{\mathcal{H}} R$. Notice that unimodular matrices, i. e. matrices from the group $Gl_p(\mathcal{A}) = \{V \in \mathcal{A}^{p\times p} \mid \det V \text{ is a unit in } \mathcal{A}\}$ or from $Gl_p(\mathcal{H})$, act bijectively on $\mathcal{E}^p$. Consequently, $\ker_{\mathcal{E}} UR = \ker_{\mathcal{E}} R$ whenever U is unimodular. From the inclusion $\mathcal{H} \subseteq \mathbb{R}(D)[\sigma,\sigma^{-1}]$ it is clear that the units of $\mathcal{H}$ are given by the group $\mathcal{H}^\times := \{\alpha\sigma^l \mid \alpha \in \mathbb{R}\backslash\{0\},\ l \in \mathbb{Z}^k\}$. It needs slightly more effort to show that $\mathcal{A}^\times := \{\alpha e^{\lambda s} \mid \alpha \in \mathbb{C}\backslash\{0\},\ \lambda \in \mathbb{C}\}$ are the units of $\mathcal{A}$; see [38, Lem. 2.5].

Finally we introduce the notion of orthogonal subspaces. Given two spaces $\mathcal{B} \subseteq \mathcal{E}^q$ and $\mathcal{N} \subseteq \mathcal{A}^q$, we let

$$\mathcal{B}^\perp := \{a \in \mathcal{A}^q \mid \forall\, w \in \mathcal{B}: \ a^\top w = 0\} \ \text{ and } \ \mathcal{N}^\perp := \{w \in \mathcal{E}^q \mid \forall\, a \in \mathcal{N}: \ a^\top w = 0\}.$$

This induces a lattice antihomomorphism with respect to inclusion. For any shift invariant subspace $\mathcal{B} \subseteq \mathcal{E}^q$ one has $\mathcal{B}^{\perp\perp} = \overline{\mathcal{B}}$. Moreover, for every $R \in \mathcal{A}^{p\times q}$ we have [34, p. 388]

$$\begin{aligned} (\ker_{\mathcal{E}} R)^\perp &= \overline{\mathrm{im}_{\mathcal{A}} R^\top}, \qquad (\mathrm{im}_{\mathcal{E}} R)^\perp = \ker_{\mathcal{A}} R^\top, \\ (\ker_{\mathcal{A}} R)^\perp &= \overline{\mathrm{im}_{\mathcal{E}} R^\top}, \qquad (\mathrm{im}_{\mathcal{E}} R)^\perp = \ker_{\mathcal{A}} R^\top. \end{aligned} \tag{2.9}$$

3. THE ALGEBRAIC SETTING FOR DELAY–DIFFERENTIAL SYSTEMS

Now we are ready to define the dynamical systems or behaviors to be investigated in this paper. The definition of a behavior below is, of course, adapted to our investigation of linear, time-invariant systems with smooth trajectories only.

Definition 3.1. A behavior with q external variables is a linear, shift invariant subspace of $\mathcal{E}^q$. We call a behavior $\mathcal{B} \subseteq \mathcal{E}^q$ a delay-differential behavior (resp. a convolutional behavior) if it is of the form $\mathcal{B} = \ker_{\mathcal{E}} R$, where R is a matrix in $\mathcal{H}^{l\times q}$ (resp. $\mathcal{A}^{l\times q}$) for some $l \in \mathbb{N}$. We will also use the name system in place of behavior.

Several remarks are in order.

Remark 3.2.

(1) By definition a delay–differential or a convolutional behavior does always admit a kernel representation. Actually, the matrix R is said to be a kernel representation of $\mathcal{B}$.

(2) Convolutional behaviors by themselves are not quite the issue of this survey. However, for a detailed understanding of systems with noncommensurate delays it will be necessary to investigate general convolutional behaviors, too.

(3) The reason for defining delay–differential behaviors via matrices with entries in $\mathcal{H}$ instead of in $\mathcal{R}$ can be found in equation (2.5). This implication (which is actually an equivalence, see Theorem 2.2 (b)) shows that systems with kernel representations over $\mathcal{H}$ will naturally enter each behavioral approach for delay–differential systems. Indeed, the most basic problem for a behavioral theory which needs to be solved is the uniqueness of kernel representations. More precisely, given two $l_i \times q$-matrices R_1 and R_2 with entries in $\mathcal{R}$, we need to understand algebraically whether $\ker_{\mathcal{E}} R_1 = \ker_{\mathcal{E}} R_2$. Turning this question around it will be important to understand which transformations on R_1 do not change its solution space $\ker_{\mathcal{E}} R_1$. Only slightly more general is the task of finding algebraic characterizations for the inclusion $\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2$. For purely differential systems, that is, $R_i \in \mathbb{R}[D]^{l_i \times q}$, $i = 1, 2$, this is nicely given as $\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2$ iff $R_2 = XR_1$ for some $X \in \mathbb{R}[D]^{l_2 \times l_1}$ [31, Sec. 3.6]. In other words, in order to compare two purely differential behaviors, it is enough to compare the submodules of $\mathbb{R}[D]^q$ generated by the rows of R_1 and R_2, which is a purely algebraic condition. Now, equation (2.5) shows that the result is not true when $\mathbb{R}[D]$ is replaced by $\mathcal{R}$. As a trivial example, note that for instance $\ker_{\mathcal{E}} D \subseteq \ker_{\mathcal{E}} (\sigma_1 - 1)$, but $\frac{\sigma_1 - 1}{D} \notin \mathcal{R}$. As a consequence, if one aims at an operator algebra which is closed under kernel inclusion in the above sense, then one is forced to take also the quotients occurring in (2.5) into consideration, that is, the operators in $\mathcal{H}$. Notice that part (b) of Theorem 2.2 can be written as $\ker b(D, \sigma) \subseteq \ker a(D, \sigma)$ iff $a = xb$ for some $x \in \mathcal{H}$. This fact can be generalized to operators a, $b \in \mathcal{H}$ and even to matrices as follows.

Proposition 3.3. Let $R_i \in \mathcal{H}^{p_i \times q}$, $i = 1, 2$, be two matrices and let $\operatorname{rk} R_1 = p_1$. Then

$$\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2 \quad \Longleftrightarrow \quad R_2 = XR_1 \text{ for some } X \in \mathcal{H}^{p_2 \times p_1}.$$

Sufficiency of $R_2 = XR_1$ is of course obvious and valid even without the rank condition on R_1. The necessity can be found in [14, Sec. 4]. It is deduced by some fairly standard algebraic arguments from Theorem 2.2 together with the following generalization of the surjectivity result 2.1:

$$\text{for each } A \in \mathcal{H}^{n \times m} \text{ one has: } \operatorname{rk} A = n \Longleftrightarrow \operatorname{im}_{\mathcal{E}} A = \mathcal{E}^n. \tag{3.1}$$

For a general theory of systems described by delay–differential equations it is important to know as to how restrictive the rank condition on R_1 is in the proposition above. Notice that it is not at all clear whether a behavior does always admit a representation $\ker_{\mathcal{E}} R$ with a full row rank matrix R. At this point the theories for the commensurate and the noncommensurate case diverge. In the commensurate case the ring $\mathcal{H}$ enjoys some strong algebraic properties (it turns out to be a Bézout domain) with the consequence that the proposition above holds true even without any rank condition on R_1. This will be shown in Section 5. The algebraic results presented therein have far-reaching consequences for systems described by delay–differential equations with commensurate delays. In essence, a behavioral theory quite parallel to the case of purely differential systems can be developed. On the other hand, for systems with noncommensurate delays Proposition 3.3 fails in general without the rank condition on R_1. In algebraic terms, the difference to the commensurate case is due to the lack of the Bézout property. However, in order to get some information about kernel inclusions in the general case, one has to resort to analytic arguments. While operators with commensurate delays do always have a closed range, this is not true for the general case. The closedness is exactly the requirement to be imposed on R_1 in order that the equivalence above remains valid. However, in this case the matrix X connecting R_1 and R_2 will have entries in $\mathcal{A}$ even though R_1 and R_2 have entries in $\mathcal{H}$ or even in $\mathcal{R}$. For this reason it is natural to formulate the result right away for convolutional behaviors.

Theorem 3.4. Let $R_i \in \mathcal{A}^{l_i \times q}$, $i = 1, 2$, be two matrices and assume that $\mathrm{im}_{\mathcal{E}} R_1$ is a closed subset of $\mathcal{E}^{l_1}$. Then

$$\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2 \quad \Longleftrightarrow \quad R_2 = X R_1 \text{ for some } X \in \mathcal{A}^{l_2 \times l_1}.$$

Proof. "$\Leftarrow$" is obvious and clearly holds without the closedness condition. Conversely, the inclusion of the kernels may be restated as $R_2(\ker_{\mathcal{E}} R_1) = \{0\}$, i.e., every row of R_2 is in $(\ker_{\mathcal{E}} R_1)^{\perp}$. However, $\mathrm{im}_{\mathcal{E}} R_1$ is closed if and only if $\mathrm{im}_{\mathcal{A}} {R_1}^{\top}$ is closed [21, Prop. 21.9] and therefore $(\ker_{\mathcal{E}} R_1)^{\perp} = \mathrm{im}_{\mathcal{A}} {R_1}^{\top}$ by use of (2.9). Thus every row of R_2 belongs to $\mathrm{im}_{\mathcal{A}} {R_1}^{\top}$, which establishes the matrix $X \in \mathcal{A}^{l_2 \times l_1}$. □

In [35, Thm. 4.1] it has been shown that the closedness of $\mathrm{im}_{\mathcal{E}} R_1$ is necessary for the equivalence to be valid for *arbitrary* operators R_2. The following example illustrates that the previous theorem is the best which can be obtained for any given pair R_1, R_2.

Example 3.5. Let

$$R_1 = \begin{bmatrix} 1 - \sigma_1 \\ 1 - \sigma_\tau \end{bmatrix} \in \mathcal{R}^2, \; R_2 = [D] \in \mathcal{R}.$$

Notice that if τ is irrational, then the two entries of $R_1^* = \begin{bmatrix} 1 - e^{-s} \\ 1 - e^{-\tau s} \end{bmatrix}$ have $s = 0$ as the unique common zero. Therefore, by the spectral analysis theorem of [33], the solution space $\ker_{\mathcal{E}} R_1$ is given by the constant functions and therefore it coincides with $\ker_{\mathcal{E}} R_2$. However, there is no matrix $X \in \mathcal{H}^{2 \times 1}$ such that $R_2 = X R_1$, since it can be shown that there are no $x_1, x_2 \in \mathcal{H}$ satisfying the Bézout equation $x_1^*(1 - e^{-s}) + x_2^*(1 - e^{-\tau s}) = s$. To see this it is enough to consider the representations of the elements in $\mathcal{H}$ suggested by Theorem 2.2 part (a) and to observe that in the ring $\mathbb{R}(s)[e^{-s}, e^{-\tau s}]$ the ideal generated by $1 - e^{-s}$ and $1 - e^{-\tau s}$ can not be the whole ring. If τ is irrational, then it can be shown that the previous Bézout equation is solvable over $\mathcal{A}$ —and this case corresponds exactly to the case when $\operatorname{im}_{\mathcal{E}} R_1$ is a closed subset of $\mathcal{E}^2$— if and only if τ is a Liouville number [25]. Notice that, since $H(\mathbb{C})$ is a Bézout domain, the equation is always solvable with $x_1, x_2 \in H(\mathbb{C})$.

Our main result of this section shows that the last statement of the previous example remains valid in a very general context. It provides a characterization of kernel inclusions for arbitrary convolutional operators. One should bear in mind, however, that general entire functions do not correspond to operators on $\mathcal{E}$. Hence the matrix X appearing in the theorem below has no operational meaning so that even the direction "$\Leftarrow$" is not trivial.

Theorem 3.6. Let $R_i \in \mathcal{A}^{l_i \times q}$, $i = 1, 2$, be two matrices. Then

$$\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2 \iff R_2 = XR_1 \text{ for some matrix } X \in H(\mathbb{C})^{l_2 \times l_1}.$$

Proof. The proof requires several steps.
1) We firstly provide the according result for kernels consisting of 'polynomial-exponentials' only. Let

$$\mathcal{PE} := \left\{ \sum_{i=1}^{n} p_i(t) e^{\lambda_i t} \,\middle|\, n \in \mathbb{N},\ \lambda_i \in \mathbb{C}, \text{ and } p_i(t) \in \mathbb{C}[t] \right\}.$$

Restricting the kernels of R_i to $\mathcal{PE}^q$ the statement above has been established in [24, p. 278], thus

$$\begin{aligned} \ker_{\mathcal{E}} R_1 \cap \mathcal{PE}^q \subseteq \ker_{\mathcal{E}} R_2 \cap \mathcal{PE}^q \iff & \text{ there exists } X \in H(\mathbb{C})^{l_2 \times l_1} \\ & \text{such that } XR_1 = R_2. \end{aligned} \tag{3.2}$$

2) In order to obtain the equivalence for kernels in $\mathcal{E}^q$, we need the following description of closed shift invariant subspaces of $\mathcal{E}^q$:

$$\mathcal{B} = \overline{\mathcal{B} \cap \mathcal{PE}^q} \text{ for each closed shift invariant subspace } \mathcal{B} \subseteq \mathcal{E}^q. \tag{3.3}$$

This extends the famous result of Schwartz [33] about shift invariant subspaces in $\mathcal{E}$ to the vector case. For the proof of "$\subseteq$" of (3.3) we will employ an analogous result which has been established in the case of continuous functions in [30]. To this end, we introduce the following notation: let $\overline{\cdot}^{\circ}$ and $\to^{\circ}$ denote the closure and the convergence in the space $\mathcal{C} := \mathcal{C}(\mathbb{R}, \mathbb{C})$ of continuous functions, equipped with its usual Fréchet topology (uniform convergence on all compact sets). Recall that without any index we refer to the topology and convergence on $\mathcal{E}$.
We shall first prove the following inclusions:

$$\mathcal{B} \subseteq \overline{\mathcal{B}}^{\circ} = \overline{\overline{\mathcal{B}}^{\circ} \cap \mathcal{PE}^q}^{\circ} \subseteq \overline{\overline{\mathcal{B} \cap \mathcal{PE}^q}}^{\circ} = \overline{\mathcal{B} \cap \mathcal{PE}^q}^{\circ}. \tag{3.4}$$

Since $\overline{\mathcal{B}}^{\circ}$ is a shift invariant and closed subspace in the topology of $\mathcal{C}$, the first identity is just the result proven in [30]. The last equality follows from the simple fact that

$\mathcal{X} \subseteq \overline{\mathcal{X}} \subseteq \overline{\mathcal{X}}^{\circ}$ for each set $\mathcal{X} \subseteq \mathcal{E}$. The proof of the remaining inclusion relies on elementary properties of the compact supported smooth functions $\rho \in \mathcal{D}$, which, as explained in Section 2, can be identified with their Laplace transforms $\rho(s) \in \mathcal{A}$. Let $w \in \overline{\mathcal{B}}^{\circ} \cap \mathcal{PE}^q$. Then there exists a sequence $w_n \in \mathcal{B}$ such that $w_n \to^{\circ} w$. It is easy to check that for any $\rho \in \mathcal{D}$, $\rho(s)w_n$ converges to $\rho(s)w$ with respect to the topology of $\mathcal{E}$; this property of $\mathcal{D}$ is usually called regularization.
Now, given any $a(s) \in \mathcal{B}^{\perp}$, one has $a(s)^{\mathsf{T}}(\rho(s)w_n) = \rho(s)a(s)^{\mathsf{T}}w_n = 0$, i.e. $\rho(s)w_n \in \mathcal{B}^{\perp\perp} = \overline{\mathcal{B}} = \mathcal{B}$. This means that $\rho(s)w \in \overline{\mathcal{B}} = \mathcal{B}$, too. Since $w \in \mathcal{PE}^q$, a simple calculation shows that also $\rho(s)w \in \mathcal{PE}^q$ and therefore $\rho(s)w \in \mathcal{B} \cap \mathcal{PE}^q$.
Taking an approximate identity $\{\rho_k\}$ (see [5, IV.21]), i.e. such that $\rho_k(s) \to 1$ in the topology of $\mathcal{A}$, we get $\rho_k(s)w \to w$, thus $w \in \overline{\mathcal{B} \cap \mathcal{PE}^q}$. Hence $\overline{\mathcal{B}}^{\circ} \cap \mathcal{PE}^q \subseteq \overline{\mathcal{B} \cap \mathcal{PE}^q}$ and (3.4) follows.

It remains to prove (3.3). To this end let $w \in \mathcal{B}$. By (3.4) there is a sequence $w_n \in \mathcal{B} \cap \mathcal{PE}^q$ such that $w_n \to^{\circ} w$. For every $\rho \in \mathcal{D}$ we have $\rho(s)w_n \to \rho(s)w$ by the regularization property and, as before, $\rho(s)w_n \in \mathcal{B} \cap \mathcal{PE}^q$. Thus $\rho(s)w \in \overline{\mathcal{B} \cap \mathcal{PE}^q}$ and, taking an approximate identity $\{\rho_k\}$, we get $\rho_k(s)w \to w$, hence $w \in \overline{\mathcal{B} \cap \mathcal{PE}^q}$.

3) Now we are ready to prove Theorem 3.6. Since $\ker_{\mathcal{E}} R_i$, $i = 1, 2$, are closed shift invariant subspaces of $\mathcal{E}^q$, Equation (3.3) yields $\ker_{\mathcal{E}} R_i = \overline{\ker_{\mathcal{E}} R_i \cap \mathcal{PE}^q}$. This shows that $\ker_{\mathcal{E}} R_1 \subseteq \ker_{\mathcal{E}} R_2$ if and only if $\ker_{\mathcal{E}} R_1 \cap \mathcal{PE}^q \subseteq \ker_{\mathcal{E}} R_2 \cap \mathcal{PE}^q$ which, by (3.2), proves the theorem. □

3.1. Latent variable elimination

In the context of convolutional behaviors (or of delay–differential behaviors) a latent variable representation of a behavior $\mathcal{B} \subseteq \mathcal{E}^{q_1}$ is defined to be a description of the form

$$\mathcal{B} = \{w \in \mathcal{E}^{q_1} \mid R_1 w = R_2 v \text{ for some } v \in \mathcal{E}^{q_2}\}, \tag{3.5}$$

where $R_i \in \mathcal{A}^{l \times q_i}$ (or $R_i \in \mathcal{H}^{l \times q_i}$), $i = 1, 2$. Systems descriptions of this type often arise after interconnecting two systems. In the setting of this survey the following question arises naturally: is $\mathcal{B}$ in (3.5) a convolutional (or delay–differential) behavior in the sense of Definition 3.1, or, in other words, does $\mathcal{B}$ admit a kernel representation $\mathcal{B} = \ker_{\mathcal{E}} R$?

Since an image representation of $\mathcal{B}$, i.e. a representation of the type $\mathcal{B} = \mathrm{im}_{\mathcal{E}} M$, for some matrix M, is a particular form of a latent variable representation (3.5), in this case latent variable elimination can be seen as a way of passing from an image to a kernel representation.

While it is known that latent variable elimination can always be achieved for purely differential systems over $\mathcal{E}$ [31, Thm. 6.2.6], this remains an open problem for general convolutional systems. Only the following partial result is available.

Theorem 3.7. [39] Consider the behavior $\mathcal{B}$ in (3.5) and suppose that $\operatorname{rk} R_2^*(s) = p_2$ for all $s \in \mathbb{C}$, where p_2 is the rank of R_2. Then there exists some matrix $Y \in \mathcal{A}^{t \times q_1}$ (resp. $\mathcal{H}^{t \times q_1}$) such that $\overline{\mathcal{B}} = \ker_{\mathcal{E}} Y$, hence $\overline{\mathcal{B}}$ is a convolutional behavior.

As is shown in [39], the rank condition on R_2 is equivalent to observability of the latent variable representation, that is to say that $R_1w = R_2v_1$ and $R_1w = R_2v_2$ implies $v_1 = v_2$. Thus, the theorem implies that for a behavior $\mathcal{B}$ which admits an observable latent variable representation, the closure $\overline{\mathcal{B}}$ admits a kernel representation.

Finally, we want to address the fact that the latent variable elimination problem for delay-differential systems defined by (3.5) is closely related to the following conjecture.

Shapiro's Conjecture. (see [36]) For every $p, q \in \mathcal{R}$ there exists $r \in \mathcal{R}$ such that $\{\lambda \in \mathbb{C} \mid p^*(\lambda) = q^*(\lambda) = 0\} = \{\lambda \in \mathbb{C} \mid r^*(\lambda) = 0\}$.

While the conjecture is open in this generality, there exist additional sufficient conditions for the statement to be true (see [8]). Using essentially the same proof as for Theorem 3.7 one can show that, if Conjecture 3.8 is true, then the closure of every behavior $\mathcal{B}$ defined by (3.5) admits a kernel representation.

3.2. Controllability of delay–differential systems

In this section we shall investigate controllability for delay–differential behaviors. We shall start by introducing the most fundamental notions of control theory, that are the concepts of inputs and outputs. In behavioral theory, an *input* of a system $\mathcal{B} \subseteq \mathcal{E}^q$ is a maximal subset $w_{i_1}, \dots, w_{i_m}$ of the external variables $\{w_1, \dots, w_q\}$ which can be set freely. Precisely, the map $\mathcal{B} \to \mathcal{E}^m$ which projects a trajectory $w = (w_1, \dots, w_q)^\top \in \mathcal{B}$ onto the components $w_{i_1}, \dots, w_{i_m}$, has to be surjective and no bigger subset with this property must exist; see [31, Def. 3.3.2]. In case an input exists, the collection of the remaining $p := q - m$ external variables is called the *output* of $\mathcal{B}$ and $\mathcal{B}$ is said to be an input/output (i/o-) behavior. It is convenient to assume that, in case an i/o-partition exists, the external variables $w = (w_1, \dots, w_q)^\top$ are reordered in such a way that $w = (u^\top, y^\top)^\top$ where $u \in \mathcal{E}^m$ forms the input and $y \in \mathcal{E}^p$ forms the output. The following theorem provides a simple characterization for the existence of an i/o-partition. For systems with commensurate delays a proof can be found in [11, Thm. 4.2.3], while for the noncommensurate case see [38, Theorem 3.8].

Theorem 3.9. A delay-differential behavior $\mathcal{B} \subseteq \mathcal{E}^{m+p}$ with external variables $w = (u^\top, y^\top)^\top$ defines an i/o-behavior with input $u \in \mathcal{E}^m$ and output $y \in \mathcal{E}^p$ if and only if $\mathcal{B} = \ker_{\mathcal{E}} [P, -Q]$ for some $P \in \mathcal{H}^{l\times m}$ and $Q \in \mathcal{H}^{l\times p}$, where $p = \operatorname{rk}[P, -Q] = \operatorname{rk} Q$. In this case, there exists a matrix $H \in \mathbb{R}(D, \sigma)^{p\times m}$ such that $QH = P$ and it is called the formal transfer function of $\mathcal{B}$. As a consequence, each delay–differential behavior admits an i/o-partition (where the extreme case of $m = 0$ inputs is included).

The extreme case of behaviors with no inputs can be characterized in terms of the trajectories of the behavior in an alternative way. In order to do so, we need the notation $w|_{(-\infty,0]}$ for the restriction of the function w, defined on $\mathbb{R}$, to the closed left half line $(-\infty, 0]$. Then one has just like for purely differential behaviors [31, Thm. 3.2.5] the following characterization.

Proposition 3.10. A delay–differential behavior $\mathcal{B} \subseteq \mathcal{E}^q$ has no inputs if and only if for every $w \in \mathcal{B}$ the condition $w|_{(-\infty,0]} = 0$ implies $w = 0$. A behavior with this property is said to be autonomous.

Again, for systems with commensurate delays a proof is given in [11, Prop. 4.2.7]. For the general case this can be seen as a consequence of Titchmarsh-Lions theorem on supports of distributions [24, p. 277].

The above simply says that in an autonomous behavior the future of a trajectory is completely determined by its past. At this point one should recall that behaviors $\mathcal{B} \subseteq \mathcal{E}^q$ are by definition shift invariant, that is $\sigma_{t_0}(\mathcal{B}) = \mathcal{B}$ for all $t_0 \in \mathbb{R}$. Therefore the time instant $t_0 = 0$ occurring in the definition of $w|_{(-\infty,0]}$ is just a matter of choice and has no specific meaning by itself.

We now turn to another central notion of control theory, that is controllability. In behavioral theory, controllability can be defined purely in terms of the trajectories of the behavior and independently of any systems description. It expresses the capability of the system to steer each of its trajectories into every other within finite time. Put another way, controllability describes the possibility to combine any past of the behavior with any desired (far) future of the behavior. This can be made precise in the following way.

Definition 3.11. [31, Def. 5.2.2]

(a) For w, $w' \in \mathcal{E}^q$ and $t_0 \in \mathbb{R}$ define the concatenation of w and w' at time T as the function $w \wedge_T w' : \mathbb{R} \to \mathbb{C}^q$ given by

$$(w \wedge_T w')(t) := w(t) \;\text{ for } t < T \;\text{ and }\; (w \wedge_T w')(t) := w'(t) \;\text{ for } t \geq T.$$

(b) A behavior $\mathcal{B} \subseteq \mathcal{E}^q$ is called controllable if for all w, $w' \in \mathcal{B}$ there exists $T \geq 0$ and a function $c : [0, T) \to \mathbb{C}^q$ such that $w \wedge_0 c \wedge_T \sigma_T w' \in \mathcal{B}$.

Again, due to shift invariance of the behaviors under consideration, the particular time instants $T \geq t_0 := 0$ for the concatenation in (b) can be replaced by any other choices. Notice that the requirement $w \wedge_0 c \wedge_T \sigma_T w' \in \mathcal{B}$ implies in particular that the concatenation is smooth. Since $\sigma_T w'(T) = w'(0)$, the concatenation switches exactly from $w(0)$ to $w'(0)$ but allows for some finite time $T \geq 0$ in order to make the switching smooth and the trajectory be contained in $\mathcal{B}$.

A variety of different characterizations for controllability are known for purely differential systems [31]. In trying to extend these results, we arrive at the following.

Theorem 3.12. Let $\mathcal{B} = \ker_{\mathcal{E}} R$, where $R \in \mathcal{H}^{l \times q}$ has rank p. Consider the following properties.

(a) $\operatorname{rk} R^*(s) = p$ for all $s \in \mathbb{C}$,

(b) $\mathcal{B} = \overline{\mathcal{B} \cap \mathcal{D}^q}$, where, again, $\mathcal{D}$ is the space of all functions in $\mathcal{E}$ with compact support and $\overline{\cdot}$ denotes the closure with respect to the topology of $\mathcal{E}$,

(c) $\mathcal{B} = \overline{\mathrm{im}_{\mathcal{E}} T}$ for some $T \in \mathcal{H}^{q\times(q-p)}$,

(d) $\mathcal{A}^q/\mathcal{B}^{\perp}$ is a torsion free $\mathcal{A}$-module,

(e) Let $R = [P, -Q]$, where $Q \in \mathcal{H}^{l\times p}$ has rank $p = \mathrm{rk}\, Q = \mathrm{rk}\, R$, and let $H \in \mathbb{R}(D,\sigma)^{p\times m}$ be such that $QH = P$. If $\hat{R} = [\hat{P}, -\hat{Q}] \in \mathcal{H}^{\hat{p}\times(m+p)}$ is such that $p = \mathrm{rk}\, \hat{Q}$ and $\hat{Q}H = \hat{P}$, then $\mathcal{B} \subseteq \ker_{\mathcal{E}} [\hat{P}, -\hat{Q}]$. In other words, $\mathcal{B}$ is a subbehavior of each behavior having the same formal transfer function.

(f) $\mathcal{B}$ is controllable,

(g) $\mathcal{B}$ has an image representation, that is, $\mathcal{B} = \mathrm{im}_{\mathcal{E}} T$ for some $T \in \mathcal{H}^{q\times(q-p)}$,

(h) R has a generalized inverse over $\mathcal{H}$, that is, there exists a matrix $G \in \mathcal{H}^{q\times l}$ such that $RGR = R$.

Then the properties (a), (b), (c), (d), (e) are all equivalent and (a)$\Leftarrow$(f)$\Leftarrow$(g)$\Leftarrow$(h). Moreover, (a)$\not\Rightarrow$(f).

Proof. The implications (a)$\Leftrightarrow$(b)$\Leftrightarrow$(c)$\Leftarrow$(f)$\Leftarrow$(g)$\Leftarrow$(h) have already been established in [38, 39]. The fact that (a)$\not\Rightarrow$(f) will be shown by an example in Section 4. It remains to prove (c)$\Leftrightarrow$(d)$\Leftrightarrow$(e).
(c) $\Rightarrow$ (d): If $a \in \mathcal{A}\backslash\{0\}$ and $x \in \mathcal{A}^q$ are such that $ax \in \mathcal{B}^{\perp}$, then we have $ax^{\top}Tv = 0$ for all $v \in \mathcal{E}^{q-p}$. This implies that $ax^{\top}T = 0$, and, since $\mathcal{A}$ is a domain, $x^{\top}T = 0$. Thus $x \in (\mathrm{im}_{\mathcal{E}} T)^{\perp} = \mathcal{B}^{\perp}$.
(d) $\Rightarrow$ (e): Let $\mathcal{F}$ be the field of fractions of $\mathcal{A}$. The full row rank of the matrices $Q^{\top}$ and $\hat{Q}^{\top}$ implies $\mathrm{im}_{\mathcal{F}} R^{\top} = \mathrm{im}_{\mathcal{F}} \hat{R}^{\top}$. In order to show that $\ker_{\mathcal{E}} R \subseteq \ker_{\mathcal{E}} \hat{R}$ it suffices to establish $\mathrm{im}_{\mathcal{A}} \hat{R}^{\top} \subseteq \overline{\mathrm{im}_{\mathcal{A}} R^{\top}} = \mathcal{B}^{\perp}$, see (2.9). Let $x \in \mathrm{im}_{\mathcal{A}} \hat{R}^{\top}$. From $\mathrm{im}_{\mathcal{F}} R^{\top} = \mathrm{im}_{\mathcal{F}} \hat{R}^{\top}$ one deduces that there exists $a \in \mathcal{A}\backslash\{0\}$ and $y \in \mathcal{A}^p$ such that $ax^{\top} = y^{\top}R$. This implies that $ax \in \mathcal{B}^{\perp}$ and from (d) we get $x \in \mathcal{B}^{\perp}$.
(e) $\Rightarrow$ (c): Let $T \in \mathcal{H}^{q\times(q-p)}$ be any full column rank matrix such that $RT = 0$. Then it is clear that $\ker_{\mathcal{E}} R \supseteq \overline{\mathrm{im}_{\mathcal{E}} T}$ and the converse inclusion remains to be proven. Let $x \in \ker_{\mathcal{A}} T^{\top} \subseteq \mathcal{A}^q$. By the rank conditions on R and T we can deduce the existence of some $a \in \mathcal{A}\backslash\{0\}$ such that $ax \in \mathrm{im}_{\mathcal{A}} R^{\top} \subseteq \mathcal{B}^{\perp}$. This implies $\ker_{\mathcal{E}} ax^{\top} \supseteq \ker_{\mathcal{E}} R$ and thus by Theorem 3.6 there exists a vector $h \in H(\mathbb{C})^l$ such that $ax^{\top} = h^{\top}R$. Partition $x = ({x_1}^{\top}, {x_2}^{\top})^{\top}$, where $x_1 \in \mathcal{A}^{q-p}$ and $x_2 \in \mathcal{A}^p$ and define $\hat{R} = [\hat{P}, \hat{Q}] := \begin{bmatrix} P & Q \\ {x_1}^{\top} & {x_2}^{\top} \end{bmatrix}$. Notice that $a{x_1}^{\top} = ah^{\top}P = ah^{\top}QH = a{x_2}^{\top}H$, which yields ${x_1}^{\top} = {x_2}^{\top}H$. This shows that $\hat{R}$ satisfies the hypothesis of condition (e) and hence $\ker_{\mathcal{E}} R \subseteq \ker_{\mathcal{E}} \hat{R} \subseteq \ker_{\mathcal{E}} x^{\top}$, implying $x \in \mathcal{B}^{\perp}$. Hence we derived $\ker_{\mathcal{A}} T^{\top} \subseteq \mathcal{B}^{\perp}$ and taking orthogonals leads to $\ker_{\mathcal{E}} R \subseteq \overline{\mathrm{im}_{\mathcal{E}} T}$ as desired. □

The fact that the conditions of the previous theorem are not all equivalent suggests to call a behavior $\mathcal{B} \subseteq \mathcal{E}^q$ satisfying $\mathcal{B} = \overline{\mathcal{B} \cap \mathcal{D}^q}$ a weakly controllable behavior. Moreover, we define the weakly controllable subbehavior of a given behavior $\mathcal{B} = \ker_{\mathcal{E}} R$, where $R \in \mathcal{H}^{p\times q}$, as the space

$$\mathcal{B}_c := \overline{\mathcal{B} \cap \mathcal{D}^q}.$$

The weakly controllable subbehavior $\mathcal{B}_c$ of $\mathcal{B}$ is a (closed) behavior in the sense of Definition 3.1. However, it is not known whether $\mathcal{B}_c$ is a delay-differential or convolutional behavior, precisely, whether $\mathcal{B}_c$ admits a kernel representation. We can prove, however, that the weakly controllable subbehavior of a delay-differential behavior always admits a dense image representation like the one introduced in the previous theorem.

Proposition 3.13. Let $\mathcal{B} = \ker_{\mathcal{E}} R$, where $R \in \mathcal{H}^{l\times q}$ has rank p and let $\mathcal{B}_c$ be its weakly controllable subbehavior. Then there exists a full column rank matrix $M \in \mathcal{H}^{q\times(q-p)}$ such that

$$\mathcal{B}_c = \overline{\mathrm{im}_{\mathcal{E}} M}.$$

Furthermore, $\mathcal{B}' \subseteq \mathcal{B}_c$ for every weakly controllable behavior $\mathcal{B}'$ contained in $\mathcal{B}$.

Proof. Let $M \in \mathcal{H}^{q\times(q-p)}$ be any matrix of full column rank such that $RM = 0$. Define $\mathcal{B}_{cs} := \mathcal{B} \cap \mathcal{D}^q$. It is clear that $\mathrm{im}_{\mathcal{D}} M \subseteq \mathcal{B}_{cs}$. By continuity of the operator M we can argue that $\mathrm{im}_{\overline{\mathcal{D}}} M \subseteq \overline{\mathrm{im}_{\mathcal{D}} M}$. Using $\overline{\mathcal{D}} = \mathcal{E}$ this implies $\overline{\mathrm{im}_{\mathcal{E}} M} \subseteq \overline{\mathrm{im}_{\mathcal{D}} M} \subseteq \overline{\mathcal{B}_{cs}} = \mathcal{B}_c$ and it remains to prove the other inclusion. To this aim we shall show

$$(\mathrm{im}_{\mathcal{E}} M)^{\perp} \subseteq \mathcal{B}_{cs}^{\perp}, \tag{3.6}$$

from which $\mathcal{B}_c \subseteq \overline{\mathrm{im}_{\mathcal{E}} M}$ follows by taking orthogonals. As for (3.6), let $a \in (\mathrm{im}_{\mathcal{E}} M)^{\perp}$. Then for each $v \in \mathcal{E}^{q-p}$ we have $a^{\top} M v = 0$ showing that $a^{\top} M = 0$. The rank condition on M ensures that there exist $\alpha \in \mathcal{A}\backslash\{0\}$ and $b \in \mathcal{A}^l$ such that $\alpha a^{\top} = b^{\top} R$. Pick now $w \in \mathcal{B}_{cs}$. Then $\alpha a^{\top} w = b^{\top} R w = 0$. Using the identifications of Section 2, we have $\mathcal{B}_{cs} \subseteq \mathcal{D}^q \subseteq (\mathcal{E}')^q = \mathcal{A}^q$, thus $a^{\top} w$ is an element of the domain $\mathcal{A}$ and the identity $\alpha a^{\top} w = 0$ implies $a^{\top} w = 0$. This shows that $a \in \mathcal{B}_{cs}^{\perp}$ and thus (3.6) is proved. The last assertion of the proposition is now straightforward. □

The following theorem shows that any delay-differential behavior can be decomposed as the sum of its weakly controllable subbehavior and an autonomous subbehavior.

Theorem 3.14. Consider a behavior $\mathcal{B} = \ker_{\mathcal{E}} R$, where $R \in \mathcal{H}^{l\times q}$ and let $\mathcal{B}_c$ be its weakly controllable subbehavior. Assume that $R = [P, Q]$ where $Q \in \mathcal{H}^{l\times p}$ has rank $p = \mathrm{rk}\, R$. Then

$$\mathcal{B} = \mathcal{B}_c + \mathcal{B}_a,$$

where $\mathcal{B}_a$ is the autonomous subbehavior defined as $\mathcal{B}_a := \left\{(0\ y)^{\top} \,\middle|\, y \in \ker_{\mathcal{E}} Q\right\} = $ $= \ker_{\mathcal{E}} \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix}$.

Proof. Only "$\subseteq$" requires proof. To this end, put $m := q-p$ and let $M \in \mathcal{H}^{q\times m}$ be a full column rank matrix satisfying the assertion of Proposition 3.13. Partition

$$M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix},$$

where $M_1 \in \mathcal{H}^{m\times m}$ and $M_2 \in \mathcal{H}^{p\times m}$. The rank conditions imposed on M and Q imply that M_1 is nonsingular and thus is surjective as operator from $\mathcal{E}^m$ to $\mathcal{E}^m$, see (3.1).
Now, let $w = (u^\top, y^\top)^\top \in \mathcal{B}$. Pick $v \in \mathcal{E}^m$ such that $u = M_1 v$ and put $\tilde{y} := M_2 v$. Then $w = w_1 + (w - w_1)$, where

$$w_1 := \begin{pmatrix} u \\ \tilde{y} \end{pmatrix} \in \operatorname{im}_{\mathcal{E}} M \subseteq \mathcal{B}_c \quad \text{and} \quad w - w_1 = \begin{pmatrix} 0 \\ y - \tilde{y} \end{pmatrix} \in \ker_{\mathcal{E}} \begin{bmatrix} I & 0 \\ 0 & Q \end{bmatrix} = \mathcal{B}_a,$$

since $R(w - w_1) = Rw - Rw_1 = 0$. □

We close this section with a list of open problems.

Open problems

(1) The main open problem concerns the latent variable elimination, whose general solution is connected with Shapiro's conjecture.
(2) Another open problem which can be shown to be related to Shapiro's conjecture concerns the equivalence of kernel representations. We conjecture that every delay–differential behavior admits a kernel representation with a full row rank matrix.
(3) In the context of controllability analysis we know that conditions (a) and (f) of Theorem 3.12 are not equivalent but we do not know yet whether (f), (g), and (h) are equivalent. In other words, we have no algebraic characterization of controllable behaviors.
(4) For a certain class of behaviors we know that all the conditions of Theorem 3.12 are equivalent (see [38, 15]). This class includes systems in state space form. It would be important to continue this investigation and to understand how pathological are the behaviors for which this equivalence does not hold.
(5) The theory of behavioral control by interconnection is completely open for delay–differential systems with noncommensurate delays.

4. A COUNTEREXAMPLE

In this section we present an example of a delay–differential behavior satisfying condition (a) of Theorem 3.12 but which is not controllable and hence violates condition (f) of that theorem.

Let $\tau \in \mathbb{R}_+$ be a Liouville number, that is, a transcendental number satisfying the following condition [27, p. 91]: for every positive integer $K \in \mathbb{N}$ there exist an infinite number of pairs $(n, d) \in \mathbb{N}^2$ such that $|d\tau - n| \le d^{-K}$. Consider the matrices

$$R = [a,\, b] \in \mathcal{H}^{1\times 2},\ \ M = [-b,\, a]^\top \in \mathcal{H}^{2\times 1},\ \text{ where } a = \frac{1 - \sigma_\tau}{D},\ b = 1 - \sigma_1 \in \mathcal{H}. \tag{4.1}$$

The behavior serving as an example is given by $\mathcal{B} := \ker_{\mathcal{E}} R$. Since τ is an irrational number $\mathcal{B}$ is a delay-differential behavior with two noncommensurate delays. Observe that the characteristic functions $a^*(s) = (1 - e^{-s\tau})/s$ and $b^*(s) = 1 - e^{-s} \in \mathcal{A}$

have no common zeros, and therefore condition (a) of Theorem 3.12 is satisfied. Furthermore, the proof of the equivalence of (a) and (c) of that theorem yields

$$\mathcal{B} = \ker_{\mathcal{E}} R = \overline{\operatorname{im}_{\mathcal{E}} M}. \tag{4.2}$$

We shall construct a function $w \in \ker_{\mathcal{E}} R$ such that there is no trajectory steering from zero to w in finite time, and thus the behavior is not controllable. This also proves that $\mathcal{B}$ does not admit an image representation since, by Theorem 3.12 that would imply controllability (for a more direct proof in a similar case see [38]).

The specific property of the delay τ is essential for the proof. We will start with some preparation. By the Liouville property it is possible to find a strictly increasing sequence $d_k \in \mathbb{N} \setminus \{0\}$ such that

$$\forall\, k \in \mathbb{N} \setminus \{0\}\ \exists\, n_k \in \mathbb{N} : |d_k\tau - n_k| \le d_k^{1-2k}. \tag{4.3}$$

The monotonicity of d_k allows us to define another monotonic sequence:

$$c_l = \begin{cases} 1 & \text{if } l = 0; \\ k & \text{if } l = d_k; \\ c_{l-1} & \text{if } l \neq d_k\ \forall k \in \mathbb{N}. \end{cases}$$

Note that also c_l is divergent and moreover

$$c_{d_k} = k. \tag{4.4}$$

We will construct a function $w : \mathbb{R} \to \mathbb{R}^2$ having the following structure:

$$w = \begin{pmatrix} 0 \\ x \end{pmatrix}, \ x \in \mathcal{E}. \tag{4.5}$$

The condition $w \in \ker_{\mathcal{E}} R$ is equivalent to say that $x \in \ker_{\mathcal{E}} b$, which by definition of b means that x is periodic with period one. It is known that every sufficiently regular 1-periodic function can be written as a Fourier series [10, p. 46]

$$x(t) = \sum_{l\in\mathbb{Z}} x_l(t) = \sum_{l\in\mathbb{Z}} u_l e^{j2\pi l t}, \text{ where } u_l = \int_0^1 x(t)e^{-j2\pi l t}\mathrm{d}t. \tag{4.6}$$

Moreover [10, p. 42], any series $x(t) = \sum_{l\in\mathbb{Z}} u_l e^{j2\pi l t}$ with $u_l \in \mathbb{C}$ defines a function $x \in \mathcal{E}$ if and only if $\lim_{l\to\infty} l^n|u_l| = 0$ for all $n \in \mathbb{N}$. We have $x \in \mathcal{C}^\infty(\mathbb{R},\mathbb{R})$ if and only if $u_{-l} = \overline{u_l}$ for each $l \in \mathbb{Z}$. Let us define a real-valued function x by choosing the coefficients $u_l = \overline{u_{-l}}$ in such a way that

$$|u_l| = l^{-c_l}\ \forall\, l \in \mathbb{N}. \tag{4.7}$$

Then $l^n|u_l| = l^n l^{-c_l} = l^{n-c_l} \to 0$ since c_l is increasing and thus the exponent is negative for sufficiently big l. As a consequence, $x \in \mathcal{E}$ and the function w in (4.5) belongs actually to $\mathcal{B} = \ker_{\mathcal{E}} R$.

We wish to show now that the behavior $\mathcal{B}$ in (4.2) is not controllable, that is, $\mathcal{B}$ does not satisfy Theorem 3.12 (f). We proceed by contradiction. Then there exists some $T > 0$ and some function $c : [0, T) \to \mathbb{C}^2$ such that the trajectory $\tilde{w} := 0 \wedge_0 c \wedge_T (\sigma_T w)$ is in $\mathcal{B}$. By possibly shifting the trajectory $\tilde{w}$ in forward direction, we can assume $T \in \mathbb{N}$ and the periodicity of x implies $\tilde{w} = 0 \wedge_0 c \wedge_T w$. Denoting the two components of c by c_1 and c_2, we have $\tilde{w} = (\tilde{w}_1, \tilde{w}_2)^{\mathsf{T}}$, where

$$\tilde{w}_1 := 0 \wedge_0 c_1 \wedge_T 0, \ \tilde{w}_2 = \tilde{w}_{21} + \tilde{w}_{22} \ \text{ with } \ \tilde{w}_{21} = 0 \wedge_0 c_2 \wedge_T 0 \ \text{ and } \ \tilde{w}_{22} = 0 \wedge_T x.$$

Notice that $\tilde{w}_1$ and $\tilde{w}_{21}$ have compact support and thus can be regarded as elements of $\mathcal{E}'$. Using the identifications of Section 2 (that is, identifying distributions with compact support with their Laplace transform), we have $\tilde{w}_1(s), \tilde{w}_{21}(s) \in \mathcal{A}$. Furthermore, it is easy to check that $\sigma_{-T} b \tilde{w}_{22} = 0 \wedge_0 x \wedge_1 0$. Hence this function is in $\mathcal{E}'$, too, and its Laplace transform is given by $X \in \mathcal{A}$ where

$$X(s) = \int_0^1 x(t) e^{-st} \mathrm{d}t.$$

Since $\tilde{w} \in \ker_{\mathcal{E}} R = \ker_{\mathcal{E}} [a, b]$ we have $R\tilde{w} = a\tilde{w}_1 + b\tilde{w}_{21} + b\tilde{w}_{22} = 0$, which after Laplace transformation can be rewritten as

$$a^*(s)\tilde{w}_1(s) + b^*(s)\tilde{w}_{21}(s) = -e^{-sT} X(s).$$

By rearranging the equation, this yields

$$a^*(s) f(s) + b^*(s) g(s) = X(s), \tag{4.8}$$

for some $f, g \in \mathcal{A}$.

We want to evaluate equation (4.8) at $s = j2\pi d_k$. First note that $b^*(j2\pi l) = 0$ and that, by (4.6), $X(j2\pi l) = u_l$ for every $l \in \mathbb{Z}$. Therefore, since from (4.7) and (4.4) it follows that $|u_{d_k}| = d_k^{-c_{d_k}} = d_k^{-k}$, we get

$$|a^*(j2\pi d_k)||f(j2\pi d_k)| = d_k^{-k}. \tag{4.9}$$

The growth condition (4.7) implies for $f \in \mathcal{A}$ the existence of A, $B > 0$ such that $|f(jy)| \le A(1 + |y|)^B$ for all $y \in \mathbb{R}$. Moreover, by definition of a and (4.3), we get

$$\begin{aligned} |a^*(j2\pi d_k)| &= \frac{|1 - e^{-j2\pi d_k \tau}|}{|j2\pi d_k|} = \frac{|e^{-j\pi d_k \tau}||e^{j\pi d_k \tau} - e^{-j\pi d_k \tau}|}{2\pi d_k} = \frac{|\sin \pi d_k \tau|}{\pi d_k} \\ &= \frac{|\sin \pi (d_k \tau - n_k)|}{\pi d_k} \le \frac{|\pi (d_k \tau - n_k)|}{\pi d_k} \le d_k^{-2k}. \end{aligned}$$

Now, upon using (4.9) we obtain the contradiction

$$1 = d_k^k |a^*(j2\pi d_k)||f(j2\pi d_k)| \le d_k^{-k} A(1 + 2\pi d_k)^B \sim A(2\pi)^B d_k^{B-k} \to 0, \text{ for } k \to \infty,$$

since B is a fixed constant depending only on f. This shows that the behavior $\mathcal{B} = \ker_{\mathcal{E}} R$ is not controllable and hence that we have found a delay-differential behavior satisfying (a) but not (f) of Theorem 3.12.

5. DELAY-DIFFERENTIAL SYSTEMS WITH COMMENSURATE DELAYS

In this section we shall concentrate on the case where only commensurate delays occur in the delay-differential operator (2.1) or (2.2). Hence $\sigma := \sigma_{\tau_1}$ is the forward shift of length $\tau_1 > 0$, which after suitable rescaling of the time axis can be assumed to be $\tau_1 = 1$. Thus, throughout this section $\mathbb{R}[D, \sigma]$ is a bivariate polynomial ring and, according to Theorem 2.2 (a), the operator algebra to be considered in this context is given by

$$\mathcal{H} = \left\{ \frac{a}{\sigma^l \phi} \,\middle|\, a \in \mathbb{R}[D,\sigma],\ l \in \mathbb{N}_0,\ \phi \in \mathbb{R}[D],\ \frac{a^*}{\phi^*} \in H(\mathbb{C}) \right\}.$$

In Section 5.1 we shall present some of the algebraic properties of $\mathcal{H}$. It will turn out that $\mathcal{H}$ enjoys properties which let its matrices behave almost like matrices over Euclidean domains like, say, over $\mathbb{R}[D]$. This will lead to a refinement of the results about general delay-differential behaviors presented in Section 3. Indeed, we will derive a correspondence between delay-differential behaviors and their kernel representations which is quite similar to the one for purely differential systems. In Section 5.2. we shall see how this machinery can be used to obtain the equivalence of all the properties connected to controllability in Theorem 3.12. Finally, we will address the issue of interconnection of delay-differential systems with commensurate delays.

5.1. A Galois-correspondence between systems and operators

The specific algebraic feature of the commensurate case is that $\mathcal{H} \subseteq \mathbb{R}(D)[\sigma, \sigma^{-1}]$ where

$$\mathbb{R}(D)[\sigma,\sigma^{-1}] = \left\{ \frac{a}{\sigma^l \phi} \,\middle|\, a \in \mathbb{R}[D,\sigma],\ l \in \mathbb{N}_0,\ \phi \in \mathbb{R}[D] \right\}$$

is a univariate Laurent polynomial ring over a field and thus a Euclidean domain. Hence one can perform long division within $\mathbb{R}(D)[\sigma,\sigma^{-1}]$. Furthermore, for each $a\sigma^{-l}\phi^{-1} \in \mathbb{R}(D)[\sigma,\sigma^{-1}]$, the meromorphic function $a^*(\phi^*)^{-1}$ has only finitely many poles in the complex plane (this is also true in the noncommensurate case). A careful combination of these two facts allows one to perform certain calculations of $\mathbb{R}(D)[\sigma,\sigma^{-1}]$ even within the subring $\mathcal{H}$. In other words, one calculates in $\mathbb{R}(D)[\sigma,\sigma^{-1}]$ and at the same time controls the possibly arising poles. As a result one obtains the following strong algebraic properties of $\mathcal{H}$.

Theorem 5.1.

(a) $\mathcal{H}$ is a Bézout domain, that is, each two elements $a, b \in \mathcal{H}$ have a greatest common divisor $d \in \mathcal{H}$, which is unique up to units in $\mathcal{H}$ and can be expressed as a linear combination $d = xa + yb$ with suitable coefficients x, $y \in \mathcal{H}$.

(b) Each matrix is left equivalent to an upper triangular matrix. Precisely, for each matrix $R \in \mathcal{H}^{p\times q}$ there exists a matrix $U \in Gl_p(\mathcal{H})$ such that UR is upper triangular.

(c) $\mathcal{H}$ is an elementary divisor domain, that is by definition, for each matrix $R \in \mathcal{H}^{p\times q}$ with rank ρ there exist $V \in Gl_p(\mathcal{H})$ and $W \in Gl_q(\mathcal{H})$ such that

$$VRW = \mathrm{diag}_{p\times q}(r_1, \dots, r_\rho)$$

where the symbol $\mathrm{diag}_{p\times q}(r_1, \dots, r_\rho)$ means a $p \times q$ matrix having $r_1, \dots, r_\rho$ as the first ρ elements on the diagonal and all the other entries equal to zero. The elements $r_1, \dots, r_\rho \in \mathcal{H}$ are the invariant factors of R. Hence they are unique up to units in $\mathcal{H}$ and r_i divides r_{i+1} in $\mathcal{H}$ for $i = 1, \dots, \rho - 1$. In other words, matrices over $\mathcal{H}$ admit a Smith-form.

(d) Let $A \in \mathcal{H}^{n\times q}$ and $B \in \mathcal{H}^{m\times q}$ be two matrices of full row rank. Let $\mathrm{rk}\,[A^\mathsf{T}, B^\mathsf{T}] = r$. Then A and B have a greatest common divisor $D \in \mathcal{H}^{r\times q}$, denoted by $D = \mathrm{gcrd}(A, B)$, which has full row rank, is unique up to left equivalence, and can be expressed as a linear combination $D = XA + YB$ for some matrices X and Y with entries in $\mathcal{H}$.
Moreover, A and B have a least common left multiple $M \in \mathcal{H}^{(n+m-r)\times q}$ of full row rank which is unique up to left equivalence and denoted by $M = \mathrm{lclm}(A, B)$. In case $r = n + m$, the matrix M is the empty matrix.

Part (a) has been proven in [12, Prop. 3.1, Thm. 3.2]. In special cases, basically, if the factors are coprime and one of the factors is monic in s, a Bézout identity has been earlier derived in a fairly different setting, see [29, Sec. 4] and [19, (3.2),(4.14)]. In [4, Prop. 7.8] a Bézout identity $1 = \sum_{j=1}^{n} f_j g_j$ has been obtained for exponential polynomials $f_j \in \mathbb{C}[s, e^{is}]$ with coefficients g_j in the corresponding Paley–Wiener algebra. The parts (b) and (d) are valid for every commutative Bézout domain. Part (c) follows from a certain factorization property in $\mathcal{H}$ called adequateness [12, Lem. 3.4]. It is a classical result of ring theory [16, 20] that each adequate commutative Bézout domain is an elementary divisor domain. It is also worth mentioning that it is still an open conjecture whether even *every* commutative Bézout domain is an elementary divisor domain, see [7, p. 492, ex. 7] and [23].

Remark 5.2.

(1) The properties above imply that matrices over $\mathcal{H}$ behave almost like matrices over a Euclidean domain. In particular, from (c) it follows that $R \in \mathcal{H}^{p\times q}$ has a right inverse $T \in \mathcal{H}^{q\times p}$ if and only if $[I_p, 0]$ is a Smith-form of R and this in turn is equivalent to R being completable to a unimodular matrix $[R^\mathsf{T}, S^\mathsf{T}]^\mathsf{T} \in Gl_q(\mathcal{H})$. All this is equivalent to the property $\mathrm{rk}\, R^*(s) = p$ for all $s \in \mathbb{C}$. Furthermore, using a Smith-form one observes that each full row rank matrix $R \in \mathcal{H}^{p\times q}$ can be factored as $R = BR_c$ where $B \in \mathcal{H}^{p\times p}$ is nonsingular (i. e. $\det B \neq 0$) and $R_c \in \mathcal{H}^{p\times q}$ is right invertible over $\mathcal{H}$.

(2) It should be mentioned that $\mathcal{H}$ is not a principal ideal domain since it contains ideals which are not finitely generated. In other words, $\mathcal{H}$ is not factorial and not Noetherian [12, Prop. 3.1]

(3) Part (a) and hence the other assertions fail in the noncommensurate case [14, Exa. 5.13]; see also Example 3.5.

Before we illustrate how to compute in practice a Bézout identity for given functions a, $b \in \mathcal{H}$, we wish to present the immediate consequences of the theorem for systems with commensurate delays. First of all, using left equivalent triangular forms, one observes that each delay-differential behavior admits a full row rank kernel representation. As a consequence, the rank condition on R_1 in Proposition 3.3 can be dropped and one arrives at the complete delay-differential analogue of the characterization given in [31, Sec. 3.6] for purely differential behaviors. Some more detailed arguments even lead to a Galois-correspondence, that is, an anti-isomorphism between the lattice of delay-differential behaviors in $\mathcal{E}^q$ on the one hand and the lattice of finitely generated submodules of $\mathcal{H}^q$ on the other. We summarize as follows.

Theorem 5.3. Let $R_i \in \mathcal{H}^{l_i \times q}$, $i = 1, 2$, be two matrices and put $\mathcal{B}_i := \ker_{\mathcal{E}} R_i$. Then

(1) $\mathcal{B}_1 \subseteq \mathcal{B}_2 \Longleftrightarrow R_2 = XR_1$ for some matrix $X \in \mathcal{H}^{l_2 \times l_1}$.
In particular, if $\operatorname{rk} R_i = l_i$, then $\mathcal{B}_1 = \mathcal{B}_2$ iff $l_1 = l_2$ and $R_2 = XR_1$ for some $X \in Gl_{l_1}(\mathcal{H})$.

Let $\operatorname{rk} R_i = l_i$ for $i = 1, 2$. Then

(2) $\mathcal{B}_1 \cap \mathcal{B}_2 = \ker_{\mathcal{E}} \operatorname{gcrd}(R_1, R_2)$,

(3) $\mathcal{B}_1 + \mathcal{B}_2 = \ker_{\mathcal{E}} \operatorname{lclm}(R_1, R_2)$.

As a consequence, the maps

$$\mathcal{B} \longmapsto \mathcal{B}^{\perp} \cap \mathcal{H}^q = \{h \in \mathcal{H}^q \mid \forall\, w \in \mathcal{B} : h^{\mathsf{T}} w = 0\},$$
$$\mathcal{M} \longmapsto \mathcal{M}^{\perp} = \{w \in \mathcal{E}^q \mid \forall\, h \in \mathcal{M} : h^{\mathsf{T}} w = 0\}$$

are inverses of each other and form anti-isomorphisms between the lattice of all delay-differential behaviors $\mathcal{B}$ in $\mathcal{E}^q$ and the lattice of all finitely generated submodules $\mathcal{M}$ of $\mathcal{H}^q$.

The parts (2) and (3) are standard consequences of (1) together with the properties of the greatest common right divisor and least common left multiple over a commutative Bézout domain and the surjectivity in (3.1). Notice that the anti-isomorphism maps a finitely generated submodule of $\mathcal{H}^q$ onto its solution space in $\mathcal{E}^q$ and a behavior onto its annihilator in $\mathcal{H}^q$. For details see [12, Prop. 4.4] and [13, Sec. 4.1].

Along the same line of general algebraic arguments one can also show that the latent variable elimination problem (see (3.5)) is always solvable. Precisely we have

Theorem 5.4.

(a) The image of a delay-differential behavior under a delay-differential operator is a delay-differential behavior again. Precisely, if $R_i \in \mathcal{H}^{l_i \times q}$ are two matrices, then $R_1(\ker_{\mathcal{E}} R_2) = \ker_{\mathcal{E}} X$ for some matrix X with entries in $\mathcal{H}$. In particular, $\operatorname{im}_{\mathcal{E}} R_1$ is a delay-differential behavior and thus a closed subspace of $\mathcal{E}^{l_1}$.

(b) For two matrices $R_i \in \mathcal{H}^{l \times q_i}$ the space $\mathcal{B} := \{w \in \mathcal{E}^{q_1} \mid R_1 w \in \operatorname{im}_{\mathcal{E}} R_2\}$ is a delay-differential behavior, that is, $\mathcal{B} = \ker_{\mathcal{E}} Y$ for some matrix $Y \in \mathcal{H}^{t \times q_1}$.

Part (a) follows upon noticing that $w \in R_1(\ker_{\mathcal{E}} R_2)$ if and only if $(w^\top, 0)^\top \in \mathrm{im}_{\mathcal{E}}[R_1{}^\top, R_2{}^\top]^\top$ and resorting to a left equivalent triangular form for $[R_1{}^\top, R_2{}^\top]^\top$ along with the surjectivity in (3.1). Part (b) is a direct consequence of (a) because in this case we have $\mathcal{B} = [I_{q_1}, 0](\ker_{\mathcal{E}} [R_1, -R_2])$.

This result shows that for commensurate delays the class of delay–differential behaviors in the sense of Definition 3.1 is not as restrictive as it appears on first sight. Images of operators and projections of delay–differential behaviors, extracting the desired (manifest) variables, are delay–differential behaviors again.

At the end of this section we want to address the computability of the various objects arising in the previous theorems. In practice one wishes to know, of course, whether (and how) a kernel representation of, say, $\ker_{\mathcal{E}} R_1 \cap \ker_{\mathcal{E}} R_2$ or $R_1(\ker_{\mathcal{E}} R_2)$ can actually be computed from the given data R_1 and R_2. A brief study of the corresponding constructions reveals that this question reduces in essence to the computation of a greatest common divisor along with a representing Bézout identity for given operators in $\mathcal{H}$. The following example should illustrate how to proceed for calculating a Bézout identity from the given data.

Example 5.5.

(a) Let $a = \sigma + 1$ and $b = D + 1 \in \mathbb{Q}[D, \sigma] \subseteq \mathcal{H}$. Then a and b are coprime in $\mathcal{H}$ and for a Bézout identity $1 = xa + yb$ in $\mathcal{H}$ one needs $y^* = (1 - x^* a^*)\,(s+1)^{-1} \in H(\mathbb{C})$ in order to have $y \in \mathcal{H}$. Since $a^*(s) = e^{-s} + 1$ (recall that σ is the forward shift of *unit* length) this leads to the sole condition $x^*(-1) = (e+1)^{-1}$ for $x \in \mathcal{H}$. Thus one obtains the Bézout identity $1 = xa + yb$ where $x = (e+1)^{-1} \in \mathbb{R}$ is a constant and $y = \big(1 - (e+1)^{-1}(\sigma + 1)\big)\,(D+1)^{-1} \in \mathcal{H}$.

(b) Let $a = (\sigma - e)\,(D+1)^{-1}$, $b = D + \sigma \in \mathcal{H}$. Again, a and b are coprime in $\mathcal{H}$. In order to obtain a Bézout identity one first observes that a and b are coprime also in the Euclidean domain $\mathbb{R}(D)[\sigma]$. In this larger ring a denominator free version of a Bézout identity is given by

$$x_1 a + y_1 b = D + e \in \mathbb{R}[D] \tag{5.1}$$

where $x_1 = -(D+1)$ and $y_1 = 1$. Now one has to adjust the coefficients x_1 and y_1 in such a way that they become divisible by $D + e$ within the ring $\mathcal{H}$. Precisely, one wants some $h \in \mathcal{H}$ such that

$$x = \frac{x_1 + hb}{D + e} \text{ and } y = \frac{y_1 - ha}{D + e} \text{ are in } \mathcal{H}, \tag{5.2}$$

for then $1 = xa + yb$ forms a desired Bézout equation. The function $h \in \mathcal{H}$ can be found as follows. Equation (5.1) implies

$$\begin{pmatrix} x_1^*(-e) \\ y_1^*(-e) \end{pmatrix} \in \ker_{\mathbb{R}}[a^*(-e), b^*(-e)] = \mathrm{im}_{\mathbb{R}} \begin{bmatrix} -b^*(-e) \\ a^*(-e) \end{bmatrix}.$$

Indeed, with the given data a, b, x_1, and $y_1 \in \mathcal{H}$ one can check that

$$\begin{pmatrix} x_1^*(-e) \\ y_1^*(-e) \end{pmatrix} = \begin{bmatrix} -b^*(-e) \\ a^*(-e) \end{bmatrix} h, \text{ where } h = \frac{1 - e}{e^e - e} \in \mathbb{R}.$$

With this choice for the function $h \in \mathbb{R} \subseteq \mathcal{H}$ we obtain (5.2) and thus the Bézout identity $1 = xa + yb$ in $\mathcal{H}$.

The example above is typical for the general situation in the way how to proceed for deriving a Bézout identity. The only difference is that in general several steps are needed in order to eliminate the zeros (like $-e$ in (b)) of a Bézout identity in $\mathbb{R}(D)[\sigma]$; see [11, Rem. 2.5] for a general procedure. The example is however not typical in the sense that in both cases above a greatest common divisor of the given elements was simply found by inspection. This is of course not always possible. In any case, use of the fact that $\mathbb{R}(D)[\sigma]$ is Euclidean along with a careful handling of the denominators which arise when calculating in that ring, one can build even a procedure which upon any input a, $b \in \mathcal{H}$ produces a greatest common divisor of a and b along with a representing Bézout identity; for details see [13, Thm. 3.1.5].

Let us now turn to a different aspect of the example. Assume we are interested in *symbolic computability* of Bézout identities (and consequently of upper triangular forms etc.), that is, we wish *exact* computations, not numerical. For the notion of symbolic computability (also known as effectiveness or decidability) we refer the reader to standard literature of computer algebra, for instance [6, 1]. As an indispensable prerequisite for symbolic computations one needs, of course, a way to represent the objects on a computer. It turns out that this part is the main (and only) obstacle for the symbolic computability of Bézout identities in $\mathcal{H}$. We wish to briefly illustrate the problem arising in this context. Since rational numbers (as opposed to arbitrary real numbers) are symbolically representable on a computer, it is reasonable to investigate the issue for functions with coefficients in $\mathbb{Q}$. Consider now Example 5.5 again. In Part (a) we started with two polynomials in $\mathbb{Q}[D, \sigma]$ and derived a Bézout equation where the constants are in the extension field $\mathbb{Q}(e)$ of $\mathbb{Q}$. In the second example we were given two functions with coefficients in $\mathbb{Q}(e)$; in that case we were forced to pass to the even larger coefficient field $\mathbb{Q}(e, e^e)$ in order to derive a Bézout identity. One should have in mind that such successive Bézout identities (using the output of one equation as input for the next one) are for instance to be computed for the transformation of a matrix into triangular form. For symbolic computability, in fact for the symbolic representation, it is important to have some information about the algebraic structure of the coefficient fields involved. While this is completely understood for the field $\mathbb{Q}(e)$, since e is transcendental, this is not at all clear for the field $\mathbb{Q}(e, e^e)$. Indeed, it seems to be unknown whether the transcendence degree of $\mathbb{Q}(e, e^e)$ is two, which is what one would expect. This is a very specific case of a more general conjecture in transcendental number theory attributed to Schanuel.

Schanuel's Conjecture. (see [22, p. 687]) If $\lambda_1, \dots, \lambda_l$ are complex numbers, linearly independent over $\mathbb{Q}$, then the transcendence degree of $\mathbb{Q}(\lambda_1, \dots \lambda_l, e^{\lambda_1}, \dots, e^{\lambda_l})$ is at least l.

Notice that in the special case where $\lambda_1, \dots, \lambda_l$ are algebraic numbers, the well-known Theorem of Lindemann–Weierstrass [17, p. 277] tells us that the transcendence degree of $\mathbb{Q}(\lambda_1, \dots \lambda_l, e^{\lambda_1}, \dots, e^{\lambda_l})$ is even equal to l. A verification of the conjecture would answer a lot of questions concerning the algebraic independence of given transcendental numbers, like, say, e and π (where it is in fact even unknown whether $e + \pi$ is irrational!), or e and e^e.

As for the calculations in $\mathcal{H}$, it can be shown that an affirmative answer of Schanuel's conjecture would imply the symbolic computability of a greatest common divisor along with a Bézout identity for any finite set of operators in $\mathcal{H}$ with coefficients in a computable field. The key point is that for the calculation of a Bézout identity starting with operators having coefficients in a computable field F (like, say, $\mathbb{Q}$ or $\mathbb{Q}(e)$), one has to successively adjoin elements $\lambda \in \mathbb{C}$ which are algebraic over F along with the element $e^\lambda \in \mathbb{C}$. This leads to field extensions of $\mathbb{Q}$ of the type considered by Schanuel. Thanks to the fact that the elements $\lambda \in \mathbb{C}$ do not contribute to the transcendence degree, the conjecture would yield the *exact* transcendence degree and even a transcendence basis. It can be shown that this suffices for symbolic representation and computability. The lengthy details of this topic are elaborated in [13, Sec. 3.5]. Needless to say that these considerations are still fairly theoretical, since in general the symbolic terms needed even for a single Bézout identity turn easily into rather huge expressions.

5.2. Controllability and interconnections

In this section we utilize the machinery of Section 5.1 to launch a behavioral control theory for systems with commensurate delays. The detailed elaboration of this section, performed completely in the algebraic context of the commensurate case, can be found in [13, Ch. 4].

First observe that, since we can assume full row rank kernel representations, each i/o–behavior $\mathcal{B}$ has a kernel representation $\mathcal{B} = \ker_{\mathcal{E}} [P, Q]$ for some $P \in \mathcal{H}^{p\times m}$ and some nonsingular $Q \in \mathcal{H}^{p\times p}$, where $p = \operatorname{rk}[P, Q] = \operatorname{rk} Q$. Hence, the formal transfer function of $\mathcal{B}$ is the matrix $Q^{-1}P \in \mathbb{R}(D, \sigma)^{p\times m}$.

In the commensurate case it is possible to analyze causality relations (with respect to time) between the external variables. The corresponding notion is called nonanticipation in behavioral control theory.

Theorem 5.7. Let $[P, Q] \in \mathcal{H}^{p\times(m+p)}$ and $\det Q \neq 0$. Hence $\mathcal{B} := \ker_{\mathcal{E}} [P, Q] \subseteq \mathcal{E}^{m+p}$ is an i/o–behavior with input $u \in \mathcal{E}^m$ and output $y \in \mathcal{E}^p$. The following are equivalent.

(a) For all $u \in \mathcal{E}^m$ satisfying $u|_{(-\infty,0]} = 0$ there exists $y \in \mathcal{E}^p$ such that $y|_{(-\infty,0]} = 0$ and $(u^\top, y^\top)^\top \in \mathcal{B}$.

(b) $Q^{-1}P \in \mathbb{R}(D)[\![\sigma]\!]^{p\times m}$, that is, the entries of $Q^{-1}P$ are formal power series in σ with coefficients in the field $\mathbb{R}(D)$.

If one of these conditions is satisfied, the delay–differential behavior $\mathcal{B}$ is said to be nonanticipating.

Recall that the formal transfer function $Q^{-1}P$ does always exist in $\mathbb{R}(D, \sigma)^{p\times m}$. Since $\mathbb{R}(D, \sigma) \subseteq \mathbb{R}(D)((\sigma))$, the space of formal Laurent series in σ with coefficients in $\mathbb{R}(D)$, part (b) above simply requires that $Q^{-1}P$ does not contain any negative

powers of σ, hence no backward shifts.
The most convenient way for proving the theorem is by interpreting $Q^{-1}P \in \mathbb{R}(D,\sigma)^{p\times m}$ as a map from $\mathcal{E}_+^m$ to $\mathcal{E}_+^p$, where $\mathcal{E}_+$ denotes the space of functions in $\mathcal{E}$ with support bounded on the left. This is indeed possible since one can canonically embed $\mathbb{R}(D,\sigma)$ in the space of all distributions having support bounded on the left; for details see [11, Thm. 2.6] and [13, Sec. 4.2].

The result above might surprise on first sight if applied to purely differential behaviors, i. e. to an operator $[P,Q] \in \mathbb{R}[D]^{p\times(m+p)}$. In that case, it simply says that every i/o-behavior is nonanticipating. No properness of the associated formal transfer function $Q^{-1}P$ arises. This is due to the fact that only $\mathcal{C}^\infty$-trajectories are being considered and has been pointed out already in [40, p. 333]. Only if more general function spaces, say L^1_{loc}, are taken into consideration, the properness of $Q^{-1}P$ is of specific importance, see [40] for purely differential behaviors and [13, Rem. 4.2.4] for the case of (commensurate) delay-differential systems.

We now turn to controllability for delay-differential behaviors. The following theorem shows that, in the case of commensurate delays, all the conditions of Theorem 3.12 are equivalent; recall that this is not the case for the noncommensurate case. Thanks to left equivalent upper triangular forms it suffices, again, to restrict to full row rank kernel representations.

Theorem 5.8. Let $\mathcal{B} = \ker_{\mathcal{E}} R$, where $R \in \mathcal{H}^{p\times q}$ has rank p. Then the following are equivalent.
(a) $\operatorname{rk} R^*(s) = p$ for all $s \in \mathbb{C}$,
(b) $\mathcal{H}^q/(\mathcal{H}^q \cap \mathcal{B}^\perp) = \mathcal{H}^q/\operatorname{im}_{\mathcal{H}} R^\top$ is a free $\mathcal{H}$-module,
(c) $\mathcal{B}$ is a subbehavior of each delay-differential behavior having the same formal transfer function,
(d) $\mathcal{B}$ is controllable,
(e) $\mathcal{B}$ has an image representation,
(f) R has a right inverse over $\mathcal{H}$.

Proof. The implications (a)$\Leftrightarrow$(c)$\Leftarrow$(d)$\Leftarrow$(e)$\Leftarrow$(f) result from Theorem 3.12. Part (b) above is the analogue of part (d) in Theorem 3.12 together with the fact that any finitely generated torsion-free module over a Bézout domain is free. As for (f), notice that for full row rank matrices the notions of generalized inverses and right inverses coincide. The part (a)$\Rightarrow$(f) has been discussed in Remark 5.2. □

Remark 5.9. A detailed study of the equivalences above (see, e. g., the direct proof of (a)$\Leftrightarrow$(d) in [12, Sec. 5]) reveals that controllability of $\mathcal{B}$ is equivalent to the capability of steering each trajectory in finite time to zero. Precisely, $\mathcal{B}$ is controllable if and only if for all $w \in \mathcal{B}$ there exists $T \geq 0$ and $c : [0,T) \to \mathbb{C}^q$ such that $w \wedge_0 c \wedge_T 0 \in \mathcal{B}$.

The following theorem shows that in the commensurate case the weakly controllable subbehavior is actually a controllable delay-differential behavior which, moreover, enjoys a simple description.

Theorem 5.10. Let $R \in \mathcal{H}^{p\times q}$ be a matrix with rank p and put $\mathcal{B} = \ker_{\mathcal{E}} R$. Factor R as $R = BR_c$ where $B \in \mathcal{H}^{p\times p}$ and $R_c \in \mathcal{H}^{p\times q}$ is right invertible. Then the weakly controllable subbehavior $\mathcal{B}_c = \overline{\mathcal{B} \cap \mathcal{D}^q}$ of $\mathcal{B}$ is a controllable delay-differential behavior and given by $\mathcal{B}_c = \ker_{\mathcal{E}} R_c$. Moreover, if $M \in \mathcal{H}^{q\times(q-p)}$ is such that $RM = 0$, then $\mathcal{B}_c = \mathrm{im}_{\mathcal{E}} M$. Finally, $\mathcal{B}' \subseteq \mathcal{B}_c$ for every controllable delay-differential behavior $\mathcal{B}'$ contained in $\mathcal{B}$. We call $\mathcal{B}_c$ the controllable subbehavior of $\mathcal{B}$.

Proof. Let $\hat{\mathcal{B}} := \ker_{\mathcal{E}} R_c$, thus $\hat{\mathcal{B}}_c$ is controllable. From Proposition 3.13 we know that $\hat{\mathcal{B}}_c \subseteq \mathcal{B}_c$. For the converse inclusion, pick $w \in \mathcal{B} \cap \mathcal{D}^q$. Then $BR_c w = 0$ and thus $R_c w \in \ker_{\mathcal{E}} B \cap \mathcal{D}^q$. Since $\ker_{\mathcal{E}} B$ is autonomous, we obtain $R_c w = 0$ (see Theorem 3.9 and Proposition 3.10). Thus $\mathcal{B} \cap \mathcal{D}^q \subseteq \hat{\mathcal{B}}_c$ and consequently $\mathcal{B}_c \subseteq \hat{\mathcal{B}}_c$. The image representation $\mathcal{B}_c = \mathrm{im}_{\mathcal{E}} M$ follows now from Proposition 3.13 (and its proof) together with the fact that $\mathcal{B}_c$ is a closed space. □

So far we have only been concerned with the analysis of a single behavior. Now we shall direct our attention to the interconnection of two behaviors, one of which being regarded the given plant, the other one the to-be-designed controller. Indeed, a controller does constitute a behavior itself. It processes (part of) the output of the to-be-controlled system and computes (part of) the inputs for that system with the purpose to achieve certain desired properties of the overall behavior, like for instance stability. Thus, the plant and the controller are interconnected to form a new system. In the behavioral framework the interconnection can be written as the intersection of two suitably defined behaviors. The underlying idea is simply that the trajectories of the overall system have to satisfy both sets of equations, those governing the plant behavior and those imposed by the controller behavior. In order to obtain an efficient controller one has to add some regularity condition on the interconnection.

Definition 5.11. (see [40, p. 332]) The interconnection of two delay-differential behaviors $\mathcal{B}_i = \ker_{\mathcal{E}} R_i \subseteq \mathcal{E}^q$, $i = 1, 2$, where $R_i \in \mathcal{H}^{p_i \times q}$, is defined to be the delay-differential behavior $\mathcal{B} := \mathcal{B}_1 \cap \mathcal{B}_2$. The interconnection is called regular if $\mathrm{rk}\,[R_1{}^{\top}, R_2{}^{\top}]^{\top} = \mathrm{rk}\, R_1 + \mathrm{rk}\, R_2$.

The concept of a regular interconnection is rather natural in the behavioral setting as it can be seen by Theorem 3.9. Indeed, the number q of external variables minus the rank of a kernel representation represents the number of input variables of a behavior. If one thinks of one of the interconnecting components as the controller, it is natural to require that each linearly independent equation of the controller should put a restriction onto one additional input channel, for otherwise the controller would be inefficient. As a consequence, the resulting interconnection $\mathcal{B} = \ker_{\mathcal{E}} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$ of $\mathcal{B}_1$ and $\mathcal{B}_2$ is left with $q - \mathrm{rk}\, R_1 - \mathrm{rk}\, R_2$ input variables, which is exactly the regularity condition.

Obviously, an interconnection is a subbehavior of either of its components. It is fairly simple to characterize *algebraically* those subbehaviors of a given behavior, which can be achieved as regular interconnections from that given behavior. But it is also not hard to give a *dynamical* characterization purely in terms of the trajectories involved.

Theorem 5.12. Let $\hat{\mathcal{B}} \subseteq \mathcal{B} \subseteq \mathcal{E}^q$ be two delay–differential behaviors and assume $\hat{\mathcal{B}} = \ker_{\mathcal{E}} \hat{R}$ where $\hat{R} \in \mathcal{H}^{\hat{p}\times q}$ is a matrix with rank $\hat{p}$. Then the following statements are equivalent.

(a) There exists a delay–differential behavior $\mathcal{B}' \subseteq \mathcal{E}^q$ such that $\hat{\mathcal{B}} = \mathcal{B} \cap \mathcal{B}'$ is a regular interconnection of $\mathcal{B}$ and $\mathcal{B}'$,

(b) the image $\hat{R}(\mathcal{B}) \subseteq \mathcal{E}^{\hat{p}}$ of $\mathcal{B}$ is controllable,

(c) $\mathcal{B} = \mathcal{B}_c + \hat{\mathcal{B}}$, where $\mathcal{B}_c$ denotes the controllable subbehavior of $\mathcal{B}$,

(d) $\mathcal{B}$ is $\hat{\mathcal{B}}$–controllable, that is, for each $w \in \mathcal{B}$ there exist $T \geq 0$, $\hat{w} \in \hat{\mathcal{B}}$, and a function $c : [0,T) \to \mathbb{C}^q$ such that $w \wedge_0 c \wedge_T \hat{w} \in \mathcal{B}$.

If any of these equivalent conditions is satisfied, the subbehavior $\hat{\mathcal{B}}$ is said to be achievable via interconnection from $\mathcal{B}$.

From a behavioral point of view, part (d) is the most important characterization for it provides us with an intrinsic criterion for regular interconnections; it is purely in terms of trajectories and does not resort to any kind of representation of the behaviors. Observe that $\hat{\mathcal{B}}$–controllability can be understood as the capability to steer every trajectory of $\mathcal{B}$ into the subspace $\hat{\mathcal{B}}$ in finite time. In light of Remark 5.9 we see that controllability in the sense of the previous section is the same as $\{0\}$–controllability. The characterization above is close to what has been obtained for multidimensional systems in [32, Thm. 4.2].

Proof of Theorem 5.12. Let $\mathcal{B} = \ker_{\mathcal{E}} R$ for some $R \in \mathcal{H}^{p\times q}$ having full row rank. The inclusion $\hat{\mathcal{B}} \subseteq \mathcal{B}$ implies a relation $X\hat{R} = R$ where $X \in \mathcal{H}^{p\times\hat{p}}$ is a full row rank matrix. One easily verifies that $\hat{R}(\mathcal{B}) = \ker_{\mathcal{E}} X$.

(a) $\Rightarrow$ (b) Let $\mathcal{B}' = \ker_{\mathcal{E}} R'$ where $R' \in \mathcal{H}^{p'\times q}$ has rank p'. Then $\hat{\mathcal{B}} = \ker_{\mathcal{E}} \begin{bmatrix} R \\ R' \end{bmatrix} = \ker_{\mathcal{E}} \hat{R}$ and $\hat{p} = p + p'$ by regularity of the interconnection. Hence Theorem 5.3 (1) yields that the matrices $\hat{R}$ and $\begin{bmatrix} R \\ R' \end{bmatrix}$ are left equivalent. This shows that X is a block row of a unimodular matrix and therefore $\ker_{\mathcal{E}} X = \hat{R}(\mathcal{B})$ is controllable by virtue of Theorem 5.8 (f) and Remark 5.2.

(b) $\Rightarrow$ (a) follows by completing X to a unimodular matrix $[X^\top, Y^\top]^\top$ and defining $R' = Y\hat{R}$.

(b) $\Rightarrow$ (c) Let $R = BR_c$ be factored as in Theorem 5.10, thus $\mathcal{B}_c = \ker_{\mathcal{E}} R_c$ is the controllable subbehavior of $\mathcal{B}$. Then $\mathcal{B} = \mathcal{B}_c + \hat{\mathcal{B}}$ is equivalent to $\mathrm{lclm}(R_c, \hat{R}) = R$ (up to unimodular left factors), see Theorem 5.3 (3). But the latter follows from the right invertibility of X, since every $\mathrm{lclm}(R_c, \hat{R})$ is of the form $L = A\hat{R} \in \mathcal{H}^{p\times q}$ and a right divisor of $R = X\hat{R} = BR_c$.

(c) $\Rightarrow$ (d) Choose $w = w_c + \hat{w} \in \mathcal{B}$ where $w_c \in \mathcal{B}_c$ and $\hat{w} \in \hat{\mathcal{B}}$. Controllability of $\mathcal{B}_c$ implies the existence of a trajectory $v := w_c \wedge_0 c \wedge_T 0 \in \mathcal{B}_c$. As a consequence,

$v + \hat{w} = w \wedge_0 c' \wedge_T \hat{w} \in \mathcal{B}$, which proves (d).

(d) $\Rightarrow$ (b) Let $v = \hat{R}w \in \hat{R}(\mathcal{B})$ for some $w \in \mathcal{B}$. By assumption there exists a trajectory $\hat{w} \in \hat{\mathcal{B}}$ such that $w_1 := w \wedge_0 c \wedge_T \hat{w} \in \mathcal{B}$ for some $T > 0$ and a suitable function c defined on $[0, T)$. Using [12, Lem. 5.3] one obtains $\hat{R}w_1 = \hat{R}w \wedge_0 c' \wedge_{T_1} \hat{R}\hat{w} \in \hat{R}(\mathcal{B})$ for some $T_1 \geq 0$ and a function c' (here one has to assume that $\hat{R}$ does not contain any negative powers of σ for otherwise the first concatenation would occur at a negative time instant; but this can indeed be assumed without loss of generality, since σ is a bijection on $\mathcal{E}$). Since $\hat{R}\hat{w} = 0$, the last part shows that every trajectory in $\hat{R}(\mathcal{B})$ can be steered to zero, which by Remark 5.9 is equivalent to controllability of $\hat{R}(\mathcal{B})$. □

Since the image of a controllable behavior is controllable again [12, Lem. 5.4], the following additional characterization is immediate from the above theorem. Notice that by part (b) below the term controllability can now be understood in a twofold way. Firstly, it describes the ability to steer trajectories (Definition 3.11), and secondly, it expresses the achievability of *all* subbehaviors via regular interconnections. In other words, it guarantees the very existence of controllers.

Corollary 5.13. The following conditions on a delay–differential behavior $\mathcal{B} \subseteq \mathcal{E}^q$ are equivalent.

(a) $\mathcal{B}$ is controllable,

(b) each subbehavior $\hat{\mathcal{B}} \subseteq \mathcal{B}$ can be achieved via a regular interconnection from $\mathcal{B}$,

(c) $\{0\} \subseteq \mathcal{B}$ can be achieved via a regular interconnection from $\mathcal{B}$.

We close this section on delay–differential behaviors with commensurate delays with a brief outlook at some

Open problems

(1) First of all, from a control theoretic point of view it would be interesting to develop a theory for behaviors where the trajectories have their components in more general functions spaces, say in the space L^1_{loc}. In [31] this has been elaborated for purely differential behaviors. While for purely differential systems every sufficiently smooth weak solution is even a strong one [31, Thm. 2.3.11], it is not clear how strong and weak solutions are related for delay–differential systems. This, however, would be a helpful information for extending the results presented in this paper to larger function spaces. The results in [35, Sec. 7] might also be helpful in this regard. Furthermore, it is obvious how to define a behavior in L^1_{loc} via kernel representations over $\mathcal{H}$, but it seems to be fairly difficult to characterize controllability for these behaviors.

(2) For more general function spaces like L^1_{loc} the properness of the associated formal transfer function plays a fundamental role. With the methods presented in this section it is possible to explain this relationship, if one considers inputs with components in L^1_{loc} having support bounded to the left. This in turn leads to a more involved notion of regular interconnection where the properness is taken into consideration as well, the so-called regular feedback–interconnection,

see [40, p. 334] for purely differential behaviors. The question of achievability via regular feedback–interconnections is completely unsolved even for purely differential behaviors.

(3) For any kind of underlying function space the concept of stabilizability remains to be investigated. Only partial results are available in this regard.

(Received November 22, 2000.)

REFERENCES

[1] T. Becker and V. Weispfennig: Gröbner Bases: A Computational Approach to Commutative Algebra. Springer, New York 1993.

[2] C. A. Berenstein and M. A. Dostal: The Ritt theorem in several variables. Ark. Mat. *12* (1974), 267–280.

[3] C. A. Berenstein and D. C. Struppa: Complex analysis and convolution equations. Several complex variables. Encyclopedia Math. Sci. *54* (1993), 1–108.

[4] C. A. Berenstein and A. Yger: Ideals generated by exponential-polynomials. Adv. in Math. *60* (1986), 1–80.

[5] H. Brezis: Analyse fonctionnelle: theorie et applications. Masson, Paris 1983.

[6] A. M. Cohen, H. Cuypers, and H. Sterk (eds.): Some Tapas of Computer Algebra. Springer, Berlin 1999.

[7] P. M. Cohn: Free Rings and Their Relations. Academic Press, London 1985. Second edition.

[8] A. Diab: Sur les zéros communs des polynômes exponentiels. C. R. Acad. Sci. Paris Sér. A *281* (1975), 757–758.

[9] L. Ehrenpreis: Solutions of some problems of division. Part III. Division in the spaces $\mathcal{D}'$, $\mathcal{H}$, $\mathcal{Q}_A$, $\mathcal{O}$. Amer. J. Math. *78* (1956), 685–715.

[10] G. B. Folland: Fourier Analysis and its Applications. Wadsworth & Brooks, Pacific Grove 1992.

[11] H. Gluesing–Luerssen: A convolution algebra of delay-differential operators and a related problem of finite spectrum assignability. Math. Control Signal Systems *13* (2000), 22–40.

[12] H. Gluesing–Luerssen: A behavioral approach to delay differential equations. SIAM J. Control Optim. *35* (1997), 480–499.

[13] H. Gluesing–Luerssen: Linear delay-differential systems with commensurate delays: An algebraic approach. Habilitationsschrift at the University of Oldenburg 2000. Accepted for publication as Lecture Notes in Mathematics, Springer.

[14] L. C. G. J. M. Habets: System equivalence for AR-systems over rings – With an application to delay-differential systems. Math. Control Signal Systems *12* (1999), 219–244.

[15] L. C. G. J. M. Habets and S. J. L. Eijndhoven: Behavioral controllability of time-delay systems with incommensurate delays. In: Proc. IFAC Workshop on Linear Time Delay Systems (A. M. Perdon, ed.), Ancona 2000, pp. 195–201.

[16] O. Helmer: The elementary divisor theorem for certain rings without chain condition. Bull. Amer. Math. Soc. *49* (1943), 225–236.

[17] N. Jacobson: Basic Algebra I. Second edition. W. H. Freeman, New York 1985.

[18] E. W. Kamen: On an algebraic theory of systems defined by convolution operators. Math. Systems Theory *9* (1975), 57–74.

[19] E. W. Kamen, P. P. Khargonekar, and A. Tannenbaum: Proper stable Bezout factorizations and feedback control of linear time-delay systems. Internat. J. Control *43* (1986), 837–857.

[20] I. Kaplansky: Elementary divisors and modules. Trans. Amer. Math. Soc. *66* (1949), 464–491.
[21] J. L. Kelley and I. Namioka: Topological Vector Spaces. Van Nostrand, 1963.
[22] S. Lang: Algebra. Second edition. Addison–Wesley, Reading, N.J. 1984.
[23] O. Lezama and O. Vasquez: On the simultaneous basis property in Prüfer domains. Acta Math. Hungar. *80* (1998), 169–176.
[24] B. Malgrange: Existence et approximations des solutions des équations aux dérivées partielles et des équations de convolution. Ann. Inst. Fourier *6* (1955/1956), 271–355.
[25] G. H. Meisters: Periodic distributions and non-Liouville numbers. J. Funct. Anal. *26* (1977), 68–88.
[26] H. Mounier: Algebraic interpretations of the spectral controllability of a linear delay system. Forum Math. *10* (1998), 39–58.
[27] I. Niven: Irrational Numbers. Wiley, New York 1956.
[28] U. Oberst: Multidimensional constant linear systems. Acta Appl. Math. *20* (1990), 1–175.
[29] A. W. Olbrot and L. Pandolfi: Null controllability of a class of functional differential systems. Internat. J. Control *47* (1988), 193–208.
[30] F. Parreau and Y. Weit: Schwartz's theorem on mean periodic vector-valued functions. Bull. Soc. Math. France *117* (1989), 3, 319–325.
[31] J. W. Polderman and J. C. Willems: Introduction to Mathematical Systems Theory. A behavioral approach. Springer, Boston 1998.
[32] P. Rocha and J. Wood: Trajectory control and interconnection of 1D and nD systems. SIAM J. Control Optim. *40* (2001), 107–134.
[33] L. Schwartz: Théorie génerale des fonctions moyennes-périodiques. Ann. of Math. (2) *48* (1947), 857–929.
[34] F. Treves: Topological Vector Spaces, Distributions and Kernels. Academic Press, New York 1967.
[35] S. J. L. van Eijndhoven and L. C. G. J. M. Habets: Equivalence of Convolution Systems in a Behavioral Framework. Report RANA 99-25. Eindhoven University of Technology 1999.
[36] A. J. van der Poorten and R. Tijdeman: On common zeros of exponential polynomials. Enseign. Math. (2) *21* (1975), 57–67.
[37] P. Vettori: Delay Differential Systems in the Behavioral Approach. Ph. D. Thesis, Università di Padova 1999.
[38] P. Vettori and S. Zampieri: Controllability of systems described by convolutional or delay-differential equations. SIAM J. Control Optim. *39* (2000), 728–756.
[39] P. Vettori and S. Zampieri: Some results on systems described by convolutional equations. IEEE Trans. Automat Control *AC-46* (2001), 793–797.
[40] J. C. Willems: On interconnection, control, and feedback. IEEE Trans. Automat. Control *AC-42* (1997), 326–339.
[41] A. H. Zemanian: Distribution Theory and Transform Analysis. McGraw–Hill, New York 1965.

Dr. Heide Gluesing–Luerssen, Fachbereich Mathematik, Universität Oldenburg, P. O. Box 2503, D-26 111 Oldenburg. Germany.
email: gluesing@mathematik.uni-oldenburg.de

Dr. Paolo Vettori and Dr. Sandro Zampieri, Dipartimento di Elettronica e Informatica, Università di Padova, via Gradenigo 6/a, 35131 Padova. Italy.
e-mail: p.vettori@dei.unipd.it, zampi@dei.unipd.it