

## A BEHAVIORAL APPROACH TO TIME-VARYING LINEAR SYSTEMS. PART 1: GENERAL THEORY\*

ACHIM ILCHMANN<sup>†</sup> AND VOLKER MEHRMANN<sup>‡</sup>

**Abstract.** We develop a behavioral approach to linear, time-varying, differential-algebraic systems. The analysis is “almost everywhere” in the sense that the statements hold on  $\mathbb{R} \setminus \mathbb{T}$ , where  $\mathbb{T}$  is a discrete set. Controllability, observability, and autonomy are introduced and related to the behavior of the system. Classical results on the behavior of time-invariant systems are studied in the context of time-varying systems.

**Key words.** time-varying linear systems, behavioral approach, controllability, observability, autonomous system, adjoint system, latent variables

**AMS subject classifications.** 93B11, 93B40, 93B36

**DOI.** 10.1137/S0363012904442239

### Notation.

$I_d$	$:= \text{diag}[1, \dots, 1] \in \mathbb{R}^{d \times d}$
$0_d$	$:= (0, \dots, 0)^T \in \mathbb{R}^d$
$\mathcal{A}$	the ring of real analytic functions $f : \mathbb{R} \rightarrow \mathbb{R}$
$\mathcal{M}$	the field of real meromorphic functions
$\mathcal{A}[D], \mathcal{M}[D]$	the skew polynomial ring of differential polynomials with coefficients in $\mathcal{A}, \mathcal{M}$ , respectively, indeterminate $D$ , and multiplication rule $Df = fD + \dot{f}$
$\mathcal{C}^N(M, \mathbb{R}^q)$	the real vector space of $N$ -times differentiable functions $f : M \rightarrow \mathbb{R}^q$ , $M \subset \mathbb{R}$ an open set, $N \in \mathbb{N} \cup \{\infty\}$
$\mathcal{C}^\omega(\mathbb{I}, \mathbb{R}^q)$	the real vector space of real analytic functions $f : \mathbb{I} \rightarrow \mathbb{R}^q$ , $\mathbb{I} \subset \mathbb{R}$ an open interval
$\mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^q)$	$:= \{w \in \mathcal{C}^\infty(\mathbb{R} \setminus \mathbb{T}, \mathbb{R}^q) \mid \mathbb{T} \subset \mathbb{R} \text{ discrete}\}$
$\mathcal{C}_t^\infty(\mathbb{R}^q)$	$:= \{w \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \mid \mathbb{I} \subset \mathbb{R} \text{ an open interval with } t \in \mathbb{I}, t \in \mathbb{R}\}$
$\text{im}_t M$	$:= \{w \in \mathcal{C}_t^\infty(\mathbb{R}^q) \mid \exists l \in \mathcal{C}_t^\infty(\mathbb{R}^m) \text{ for all } \tau \in \text{dom } w \cap \text{dom } l : w(\tau) = M(\frac{d}{d\tau})l(\tau), t \in \mathbb{R}, M(D) \in \mathcal{M}[D]^{q \times m}\}$
$\text{im } M$	$:= \{w \in \mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^q) \mid \exists l \in \mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^m) \text{ for a.a. } \tau \in \text{dom } w \cap \text{dom } l : w(\tau) = M(\frac{d}{d\tau})l(\tau), M(D) \in \mathcal{M}[D]^{q \times m}\}$
$\ker_t R$	$:= \{w \in \mathcal{C}_t^\infty(\mathbb{R}^q) \mid R(\frac{d}{d\tau})w(\tau) = 0 \text{ for all } \tau \in \text{dom } w, t \in \mathbb{R}, R(D) \in \mathcal{M}[D]^{g \times q}\}$
$\ker R$	$:= \{w \in \mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^q) \mid R(\frac{d}{d\tau})w(\tau) = 0 \text{ for almost all } \tau \in \mathbb{R}, R(D) \in \mathcal{M}[D]^{g \times q}\}$
$\text{dom } w$	the domain of a function $w$

\*Received by the editors March 22, 2004; accepted for publication (in revised form) March 8, 2005; published electronically November 23, 2005.

<http://www.siam.org/journals/sicon/44-5/44223.html>

<sup>†</sup>Institut für Mathematik, Technische Universität Ilmenau, Weimarer Straße 25, 98693 Ilmenau, Germany (achim.ilchmann@tu-ilmenau.de).

<sup>‡</sup>Institut für Mathematik, MA 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany (mehrman@math.tu-berlin.de). This author’s research was supported by DFG Research Center MATHEON *Mathematics for key technologies* in Berlin.

**1. Introduction.**

**1.1. An algebraic approach and solution spaces.** The aim of the present paper is to develop a behavioral approach to linear time-varying systems described by differential-algebraic equations of the form

$$(1.1) \quad R\left(\frac{d}{dt}\right)w = 0,$$

where  $R(D)$  is a  $g \times q$  polynomial matrix in the indeterminate  $D$  with real meromorphic coefficient matrices belonging to  $\mathcal{M}^{g \times q}$ ; we use the notation  $R(D) \in \mathcal{M}^{g \times q}[D]$ .

Instead of considering real meromorphic coefficients of  $R(D)$  on the whole time axis  $\mathbb{R}$ , we also could develop the theory on some open interval  $\mathbb{I} \subset \mathbb{R}$ ; this is omitted.

The ring  $\mathcal{M}[D]$  is endowed with the multiplication rule

$$(1.2) \quad Df = fD + \dot{f}.$$

This is a consequence of assuming the associative rule  $(Df)g = D(fg)$  for all differentiable functions  $f, g$  which yields  $(Df)(g) = \frac{d}{dt}f \cdot g + f \cdot \frac{d}{dt}g = \left(\frac{d}{dt}f + fD\right)(g)$ . The noncommutativity of  $\mathcal{M}[D]$ , in contrast to the commutative ring  $\mathbb{R}[D]$  in the time-invariant case, is crucial in the following.

Note that we distinguish between the algebraic indeterminate  $D$  and the differential operator  $\frac{d}{dt}$ ; for

$$R(D) = \sum_{i=0}^n R_i D^i \in \mathcal{M}[D]^{g \times q} \cong \mathcal{M}^{g \times q}[D],$$

equality in (1.1) means

$$\sum_{i=0}^n R_i(t)w^{(i)}(t) = 0 \quad \text{for almost all } t \in \mathbb{R}.$$

Skew polynomial rings are, for example, treated in the monograph [6]; the ring  $\mathcal{M}[D]$  was introduced in [14] to study linear time-varying systems. We are interested in the behavior introduced by all solutions of (1.1). Since the coefficients of  $R(D)$  are meromorphic functions, we can only expect solutions which are defined “almost globally” (see subsection 1.3). To be more precise, we allow for the solution space

$$\mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^q) = \{w \in \mathcal{C}^\infty(\mathbb{R} \setminus \mathbb{T}, \mathbb{R}^q) \mid \mathbb{T} \subset \mathbb{R} \text{ discrete}\}$$

of piecewise  $\mathcal{C}^\infty$ -functions (see the notation) defined almost everywhere on  $\mathbb{R}$ , and the set

$$\mathcal{C}_t^\infty(\mathbb{R}^q) = \{w \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \mid \mathbb{I} \subset \mathbb{R} \text{ an open interval with } t \in \mathbb{I}\}, \quad t \in \mathbb{R},$$

of  $\mathcal{C}^\infty$ -solution pieces on some open interval including  $t$ .

For  $R(D) \in \mathcal{M}[D]^{g \times q}$ , we study the *almost global behavior* given by the kernel representation

$$\ker R = \{w \in \mathcal{C}_{\text{pw}}^\infty(\mathbb{R}^q) \mid R\left(\frac{d}{d\tau}\right)w(\tau) = 0 \text{ for almost all } \tau \in \mathbb{R}\}$$

and the *local behavior*

$$\ker_t R = \{w \in \mathcal{C}_t^\infty(\mathbb{R}^q) \mid R\left(\frac{d}{d\tau}\right)w(\tau) = 0 \text{ for all } \tau \in \text{dom } w\}, \quad t \in \mathbb{R}.$$

**1.2. Examples of system classes.** Our approach generalizes results on the following subclasses of systems:

- (a) Time-varying state space systems of the form

$$(1.3) \quad \begin{aligned} \frac{d}{dt}x(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t) + F(t)u(t), \end{aligned}$$

with real analytic matrices  $A \in \mathcal{A}^{n \times n}$ ,  $B \in \mathcal{A}^{n \times m}$ ,  $C \in \mathcal{A}^{p \times n}$ , and  $F \in \mathcal{A}^{p \times m}$ , are well studied; see, for example, the standard monograph [30].

- (b) Time-varying descriptor systems of the form

$$(1.4) \quad \begin{aligned} E(t) \frac{d}{dt}x(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t) + F(t)u(t), \end{aligned}$$

with  $A \in \mathcal{A}^{\ell \times n}$ ,  $B \in \mathcal{A}^{\ell \times m}$ ,  $C \in \mathcal{A}^{p \times n}$ ,  $F \in \mathcal{A}^{p \times m}$ , where  $E \in \mathcal{A}^{\ell \times n}$  is allowed to be singular in the sense that  $\text{rk } E(t) < \min\{\ell, n\}$  for some  $t \in \mathbb{R}$ , have been studied by different authors. In [5] controllability and observability were studied in terms of derivative arrays. In [2] a first behavior-like approach to systems (1.4) with analytic coefficients was discussed. A more general approach that allows for larger classes of coefficients and that can be implemented also numerically was introduced in [20] and generalized partially to the nonlinear case in [18]. A completely different approach results from the study of differential-algebraic equations introduced in [1, 9, 19]. A general solvability theory for nonsquare linear time-varying systems was first given in [16] and analyzed for control problems in a behavioral context in [2, 20, 26]; see also [18] for the general nonlinear case.

- (c) In [14] time-varying polynomial systems of the form

$$(1.5) \quad \begin{aligned} P\left(\frac{d}{dt}\right)z(t) &= Q\left(\frac{d}{dt}\right)u(t), \\ y(t) &= V\left(\frac{d}{dt}\right)z(t) + W\left(\frac{d}{dt}\right)u(t), \end{aligned}$$

where  $P(D)$ ,  $Q(D)$ ,  $V(D)$ , and  $W(D)$  are matrices of size  $r \times r$ ,  $r \times m$ ,  $p \times r$ , and  $p \times m$ , respectively, over  $\mathcal{M}[D]$  are studied under the following assumptions:

- $P(D)$  represents a so-called *full* operator, i.e., if  $z$  is a real analytic solution of  $P\left(\frac{d}{dt}\right)z = 0$  on some interval  $\mathbb{I} \subset \mathbb{R}$ , then this solution can be analytically extended to the whole of  $\mathbb{R}$ .
- For every  $u \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^m)$  with bounded support to the left, there exist some  $z \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^r)$  and  $y \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^p)$  so that (1.5) is satisfied.

Time-invariant polynomial (so-called Rosenbrock) systems of the form (1.5)—i.e.,  $P(D)$ ,  $Q(D)$ ,  $V(D)$ , and  $W(D)$  are matrices over  $\mathbb{R}[D]$  and  $\det P(\cdot) \neq 0$ —were introduced in [27] and are well studied; see, for example, [11, 39].

- (d) Time-invariant polynomial systems in the so-called kernel representation

$$(1.6) \quad R\left(\frac{d}{dt}\right)w(t) = 0, \quad R(D) \in \mathbb{R}[D]^{g \times q}$$

were introduced by Willems in [35]; see also [36, 37, 38] and the monograph [24].

It is easy to see that time-varying descriptor systems (1.4) or, if  $E = I_n$  and  $n = \ell$ , state space systems (1.3) are special cases of time-varying Rosenbrock systems

(1.5). Furthermore, time-varying Rosenbrock systems of the form (1.5) are a special case of systems in kernel representation (1.1): set  $w = [z^T, u^T, y^T]^T$  and

$$(1.7) \quad R(D) = [R_1(D), R_2(D)], \quad R_1(D) = \begin{bmatrix} P(D) \\ V(D) \end{bmatrix}, \quad R_2(D) = \begin{bmatrix} -Q(D), & 0 \\ W(D), & -I_p \end{bmatrix}.$$

**1.3. Examples of time-varying scalar differential equations.** In the following, we present some prototypical scalar differential equations which illustrate how time-varying coefficients may affect the solutions in very different ways. Set, for  $r(D) \in \mathcal{M}[D]$  and  $\mathcal{W}$  a suitable solution space to be specified,

$$\ker_{\mathcal{W}} r\left(\frac{d}{dt}\right) := \{w \in \mathcal{W} \mid r\left(\frac{d}{dt}\right)w = 0\}.$$

- (i) Let  $r(D) = tD + 1$ . Then the function  $t \mapsto w(t) = t^{-1}$  is a meromorphic solution of  $r\left(\frac{d}{dt}\right)w = t\frac{d}{dt}w + w = 0$ .

The point 0 is the only zero of the leading coefficient  $t \mapsto t$  of  $r(D)$ , and 0 is also a pole of  $t \mapsto w(t)$ . Therefore,

$$\ker_{\mathcal{A}} r\left(\frac{d}{dt}\right) = \ker_{C^\infty(\mathbb{R}, \mathbb{R})} r\left(\frac{d}{dt}\right) = \{0\},$$

but, for every interval  $\mathbb{I} \subset \mathbb{R}$  with  $0 \notin \mathbb{I}$ ,

$$\dim \ker_{\mathcal{M}} r\left(\frac{d}{dt}\right) = \dim \ker_{\mathcal{A}_{\mathbb{I}}} r\left(\frac{d}{dt}\right) = 1 = \deg r(D).$$

In this example, in the meromorphic case the dimension of the solution space equals the degree of  $r(D)$ . This is not true in general, as illustrated by the following example.

- (ii) Let  $r(D) = t^2D + 1$ . Then the function  $t \mapsto w(t) = e^{1/t}$  solves  $r\left(\frac{d}{dt}\right)w = 0$ . The point 0 is again the only zero of the leading coefficient  $t \mapsto t^2$  of  $r(D)$ , and 0 is also a pole of  $t \mapsto w(t)$ . But  $w$  is not meromorphic and the singularity at  $t = 0$  differs from (i) as follows: no matter whether the solution  $w$  in (i) approaches 0 from the left or right, the limit at  $t = 0$  does not exist; whereas, for the solution  $w$  in the present example, we have  $\lim_{t \rightarrow 0^-} w(t) = 0$  and  $\lim_{t \rightarrow 0^+} w(t) = \infty$ . Hence,

$$\ker_{\mathcal{M}} r\left(\frac{d}{dt}\right) = \{0\}.$$

For every interval  $\mathbb{I} \subset \mathbb{R}$  with  $0 \notin \mathbb{I}$  we have

$$\dim \ker_{\mathcal{M}_{\mathbb{I}}} r\left(\frac{d}{dt}\right) = 1 = \deg r(D).$$

- (iii) Let  $r(D) = tD - 1$ . Then the function  $t \mapsto w(t) = t$  solves  $r\left(\frac{d}{dt}\right)w = 0$  and

$$\dim \ker_{\mathcal{A}} r\left(\frac{d}{dt}\right) = 1 = \deg r(D).$$

Note that again the point  $t = 0$  is the only zero of the leading coefficient  $t \mapsto t$  of  $r(D)$ , but this time the zero does not produce a pole of the solution, the solution  $w$  is even a real analytic function on  $\mathbb{R}$ . However, the solution is not as arbitrary as for time-invariant systems, since  $w(0) = 0$  is the only value at  $t = 0$ .

- (iv) Let  $r(D) = 2tD - 1$ . Then the functions  $t \mapsto w_+(t) = \sqrt{t}$  and  $t \mapsto w_-(t) = \sqrt{-t}$  solve  $r(\frac{d}{dt})w = 0$  on  $(0, \infty)$  and  $(-\infty, 0)$ , respectively. For every interval  $\mathbb{I} \subset \mathbb{R}$  with  $0 \notin \mathbb{I}$ , we have

$$\dim \ker_{\mathcal{A}_{\mathbb{I}}} r(\frac{d}{dt}) = 1 = \deg r(D).$$

However,

$$\ker_{\mathcal{M}} r(\frac{d}{dt}) = \{0\}.$$

The real analytic solution  $w_+$  on  $(0, \infty)$  cannot be continued to  $(-\varepsilon, \infty)$  for any  $\varepsilon > 0$ .

This also proves that the attempt to connect real analytic solutions between critical points by cutting the neighborhood and going into the complex sphere, as suggested by Ilchmann et al. [13], does not work.<sup>1</sup>

- (v) Let  $r(D) = (1 - t^2)^2 D + 2t$ . Then the function

$$t \mapsto w(t) = \begin{cases} \exp\{-(1 - t^2)^{-1}\}, & t \in (-1, 1), \\ 0, & t \in \mathbb{R} \setminus (-1, 1), \end{cases}$$

satisfies  $w \in \ker_{\mathcal{C}^\infty} r(\frac{d}{dt})$ , is not real analytic, and has compact support. This is impossible for time-invariant scalar differential equations.

**1.4. An example of a mobile manipulator.** Systems of differential-algebraic equations play an important role in modeling multibody systems, electric circuits, or coupled systems of partial differential equations; see [1, 10]. We present an application which first shows that modeling does not necessarily lead to a state space system; second, it illustrates a simple system where the notion of input, output, and state is not a priori clear; and third, the example serves to illustrate the concepts introduced in the following sections. Consider a simplified, linearized model of a two-dimensional, three-link constrained mobile manipulator [12] as depicted in Figure 1.

The Lagrangian equations of motion take the form

$$(1.8) \quad \begin{aligned} M(\theta) \ddot{\theta} + D(\theta, \dot{\theta}) \dot{\theta} + K(\theta) &= u + F^T(\theta)\mu, \\ \psi(\theta) &= 0, \end{aligned}$$

where  $\theta = [\theta_1, \theta_2, \theta_3]^T$  is the vector of joint displacements,  $u \in \mathbb{R}^3$  is the vector of control torques applied at the joints, and the maps  $M : \mathbb{R}^3 \rightarrow \mathbb{R}^{3 \times 3}$ ,  $D : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^{3 \times 3}$ , and  $K : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  model the mass, centrifugal and Coriolis forces, gravity, respectively.  $l_1, l_2, l_3, l > 0$  are the lengths of the robot arms. The nonlinear constraint function is  $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ ,  $F = \frac{\partial \psi}{\partial \theta}$ , and  $\mu \in \mathbb{R}^2$  represents the Lagrange multipliers and  $F^T(\theta)\mu$  is the generalized constraint force. We are interested in the behavior, i.e., local solutions  $t \mapsto [\theta(t)^T, u(t)^T]$  of (1.8). It can be shown that  $u(\cdot)$  is a latent variable; for its definition, see [24, sect. 6.2]. Under suitable smoothness assumptions of the involved functions, it can be shown (see, for example, [25, p. 62]) that there exists a local (possibly global) solution  $\theta(\cdot)$  of (1.8) on some open interval  $\mathbb{I}$ . Linearizing along this trajectory [4] and rewriting the system in Cartesian coordinates yields a model of the form

$$\begin{aligned} M_0(t) \ddot{z}(t) + D_0(t) \dot{z}(t) + K_0(t) z(t) &= S_0 u(t) + F_0^T(t) \mu, \\ F_0(t) z(t) &= 0, \end{aligned}$$

<sup>1</sup>We are indebted to the anonymous referee of an earlier version of the present paper for pointing out this example to us.

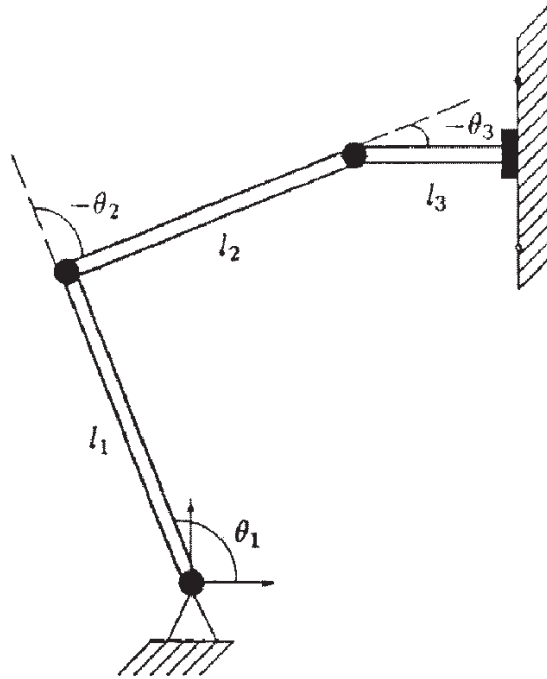


FIG. 1. Three-link constrained mobile manipulator.

where  $M_0, D_0, K_0 \in \mathcal{C}^\omega(\mathbb{I}, \mathbb{R}^{3 \times 3})$  and  $S_0 \in \mathbb{R}^{3 \times 3}$ ,  $F_0^T \in \mathbb{R}^{3 \times 2}$  with  $S_0$  having full row rank. Introducing the eight-dimensional variable  $x(t) = [z(t)^T, \dot{z}(t)^T, \mu(t)^T]^T$  results in the equivalent descriptor system description of the form

$$(1.9) \quad \begin{aligned} E(t) \frac{d}{dt} x(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t), \end{aligned}$$

where

$$E(t) := \begin{bmatrix} I_3 & 0 & 0 \\ 0 & M_0(t) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A(t) := \begin{bmatrix} 0 & I_3 & 0 \\ -K_0(t) & -D_0(t) & F_0^T(t) \\ F_0(t) & 0 & 0 \end{bmatrix}, \quad B(t) := \begin{bmatrix} 0 \\ S_0 \\ 0 \end{bmatrix},$$

and  $C(\cdot)$  denotes a matrix with appropriate format; see [12] for explicit data. Actually, in this example  $F_0$  does not depend on  $t$ .

**1.5. Literature survey.** The crucial difference between time-varying and time-invariant ordinary, linear differential equations is that the solutions behave qualitatively considerably different. Whereas any local solution of a time-invariant system is always extendable to a global analytic solution, solutions of time-varying systems may have finite escape times. Simple examples have been presented in subsection 1.3. All algebraic contributions to time-varying systems struggle with this difficulty.

Early algebraic contributions to time-varying systems in polynomial descriptions are given in [15, 40, 41]; however, the assumptions on the system classes are rather restrictive.

In [7], matrices over the ring of linear differential operators  $k[D]$  are considered, where  $k$  is a differential field. Linear dynamics are finitely generated left  $k[D]$ -modules. This contribution is rather on the algebraic side; the solution space is not specified. In [29] contributions to duality of systems in the setup of [7] for systems in generalized state space representation are given; however, the solution space is not specified either.

An important contribution by Fröhler and Oberst [8] has the following background: Consider the simple examples given in subsection 1.3. It can be shown that the local solution  $(t \mapsto 1/t) \in \ker(t \frac{d}{dt} + 1)$  can be extended to a distribution belonging to  $\mathcal{D}'(\mathbb{R}, \mathbb{R})$ ; however,  $(t \mapsto \exp(1/2t^2)) \in \ker(t^3 \frac{d}{dt} + 1)$  cannot be extended to a distribution belonging to  $\mathcal{D}'(\mathbb{R}, \mathbb{R})$ . Hence enlarging the solution space to allow for distributions on  $\mathbb{R}$  does not necessarily resolve the problem, even in the simple case when the coefficients of the time-varying systems are polynomials. However, if the solution space is enlarged even further to allow for Sato's hyperfunctions, i.e., generalized distributions introduced in [31, 32], then [8] considers systems of the form (1.1), respectively, behavior in the kernel representation  $\ker R$ , where the coefficient matrices of the polynomial  $R(D)$  are defined over rational analytic functions

$$\frac{f(\cdot)}{g(\cdot)} \quad \text{for } f, g \in \mathbb{C}[t] \quad \text{with } g(t) \neq 0 \text{ for all } t \in \mathbb{I}.$$

Note that by multiplication with a least common multiple of all denominators of the coefficients, the coefficients of  $R(D)$  are polynomials. Based on the seminal paper [22], where an algebraic analytic approach is developed to show a categorical duality between the solution spaces of linear partial differential equations with constant coefficients and certain polynomial modules associated to them, a generalization to time-varying but ordinary differential equations is achieved in [8].

The skew polynomial ring  $\mathcal{M}[D]$  was first exploited by [14] to describe time-varying linear systems. This ring is nice in the sense that it is simple (i.e., the only two-sided ideals are the trivial ones) and admits right- and left-Euclidean division. Hence matrices over the ring can be transformed into the Teichmüller–Nakayama normal form; see section 2. The latter is the essential tool in [14] to study time-varying Rosenbrock systems of the form (1.5). The solution space is the set of  $C^\infty$ -functions on the whole time axis; this is ensured by the assumption that  $\text{im } Q(\frac{d}{dt}) \subset \text{im } P(\frac{d}{dt})$  and, most importantly, that  $P(D)$  is a “full” operator, i.e., every local analytic solution of  $P(\frac{d}{dt})z = 0$  is extendable to a global analytic solution on the whole of  $\mathbb{R}$ . The latter is a rather restrictive assumption. To overcome this assumption, in [13] a first approach in the spirit of the present paper was presented for scalar systems. A behavioral approach to a certain class of time-varying systems was presented in [3].

A completely different approach results from the study of differential-algebraic equations introduced in [1, 9]. A general solvability theory for nonsquare linear time-varying systems was first given in [16] and analyzed for control problems in a behavioral context in [2, 20, 26]; see also [18] for the general nonlinear case, and a latest monograph [19].

This paper is organized as follows. In section 2, the algebraic tools, such as the Teichmüller–Nakayama normal form, and some facts on the behavior are collected. In section 3, we introduce and characterize algebraically the concept of controllable behavior for the kernel and image representation. The relationship between behavior, controllable, and autonomous behavior is investigated in section 4. In section 5, observability is defined, it is related via the adjoint of the kernel representation to the controllable behavior, and it is characterized algebraically. Finally, in section 6 we

investigate the elimination of latent variables.

**2. Behavior.** In this section we present the Teichmüller–Nakayama normal form for matrices over  $\mathcal{M}[D]$ . This will be the main tool for analyzing  $\ker_t R$ . To this end we recall some results on matrices over the skew polynomial ring  $\mathcal{M}[D]$ ; a standard reference for this is [6].  $\mathcal{M}[D]$  is *simple*, i.e., the only ideals which are right and left ideals at the same time are the trivial ones; the rank of a matrix over  $\mathcal{M}[D]$  is unambiguous, since column rank and row rank coincide; the Teichmüller–Nakayama normal form is the analogue of the Smith normal form for matrices over the commutative ring  $\mathbb{R}[D]$ ; it is simpler for matrices over  $\mathcal{M}[D]$ , since the class of transformations is larger.  $W(D) \in \mathcal{M}[D]^{n \times n}$  is called *unimodular* if and only if there exists some  $W(D)^{-1} \in \mathcal{M}[D]^{n \times n}$  such that  $W(D)W(D)^{-1} = I_n$ ; two elements  $q_1, q_2 \in \mathcal{M}[D]$  are *similar* if and only if  $q_1 a = b q_2$  for some  $a, b \in \mathcal{M}[D]$  for which  $q_1$  and  $b$  ( $q_2$  and  $a$ ) are left (right) coprime. For example,  $a(D) = D$  and  $b(D) = D - 1/t$  are similar:  $[D + (t^2 - 1)/t]a(D) = b(D)[D + t]$  and  $D + (t^2 - 1)/t, b(D)$  are right coprime,  $a(D), D + t$  are left coprime. Moreover, this example shows that a unique factorization of the ring elements cannot be expected. However, Ore [23] has shown that the degree of similar polynomials coincide. The latter property is crucial for determining dimensions of solution spaces.

A proof and an interesting historical description of the following normal form can be found in [6, Chap. 8]. The proof is constructive, using elementary matrices and Euclidean division. So if the coefficients consist of real polynomials  $\mathbb{R}[t]$ , then it is possible to calculate a normal form by means of computer algebra.

**THEOREM 2.1** (Teichmüller–Nakayama normal form). *Any  $R(D) \in \mathcal{M}[D]^{g \times q}$  with  $\text{rk}_{\mathcal{M}[D]} R(D) = \ell$  can be factorized into*

$$(2.1) \quad R(D) = U(D)^{-1} \begin{bmatrix} I_{\ell-1} & 0 & 0 \\ 0 & r(D) & 0 \\ 0 & 0 & 0_{(g-\ell) \times (q-\ell)} \end{bmatrix} V(D)^{-1},$$

where  $U(D)$  and  $V(D)$  are  $\mathcal{M}[D]$ -unimodular matrices of sizes  $g$  and  $q$ , respectively, and  $r(D) \in \mathcal{M}[D]$  is nonzero, unique up to similarity, and of unique degree.

*Remark 1.* Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  and consider the factorization (2.1).

(i) Then we have, for almost all  $t \in \mathbb{R}$ ,

$$\text{for all } w \in C_t^\infty(\mathbb{R}^q) : \left[ w \in \ker_t R \iff w \in \ker_t \left( \left[ \begin{array}{c|c} I_{\ell-1} & \\ \hline & r \end{array} \right] 0_{\ell \times q} \right] V^{-1} \right].$$

Hence we may assume, without restriction of generality, that  $R(D)$  has full row rank.

(ii) The set  $\ker_t R$  becomes a *real vector space* if endowed, for  $w_1, w_2 \in \ker_t R$ , with addition

$$(w_1 + w_2)(\tau) := w_1(\tau) + w_2(\tau) \quad \text{for all } \tau \in \text{dom } w_1 \cap \text{dom } w_2$$

and obvious scalar multiplication. The dimension of this vector space is defined as



$$\dim \ker_t R := \sup \left\{ k \in \mathbb{N} \mid \exists w_1, \dots, w_k \in \ker_t R \text{ linearly independent on } \bigcap_{i=1}^k \text{dom } w_i \right\}.$$

Furthermore,

$$\dim \ker_t R = \begin{cases} \deg r(D) & \text{for almost all } t \in \mathbb{R} & \text{if } \text{rk } R(D) = q, \\ \infty & \text{for all } t \in \mathbb{R} & \text{if } \text{rk } R(D) < q. \end{cases}$$

The latter is a simple consequence of (2.1) and the fact that the set of  $t$  where  $r(\frac{d}{dt})\varphi(t) = 0$  does not have a solution is a subset of  $\{t \in \mathbb{R} \mid r_N(t) = 0\}$ , where  $r(D) = \sum_{i=0}^N r_i(t)D^i$ ,  $r_N \neq 0$ . To see this, use the canonical transformation to a vector-valued differential equation of first order; see, for example, [34, Chap. IV].

- (iii) Let  $\mathbb{T} = \mathbb{T}(R, U, V, r)$  denote the union of all zeros and poles of the meromorphic coefficients in all nonzero entries of  $U(D)$ ,  $U(D)^{-1}$ ,  $V(D)$ ,  $V(D)^{-1}$ , and  $r(D)$ . Certainly,  $\mathbb{T}$  is a discrete set which depends on the factorization and hence is not unique.  $\mathbb{T}$  encompasses all possible critical points where a finite escape may occur (see the examples in subsection 1.3); however,  $\mathbb{T}$  might be much larger. We gain system theoretic information from the normal form but may also hide information: consider, for example, a state space system of the form (1.3). Then this system does not have any critical points; however, taking it into a normal form may introduce a possibly nonempty set  $\mathbb{T}$ . It is an open problem to determine an algorithm for the transformation into the Teichmüller–Nakayama normal form which produces a “minimal” set  $\mathbb{T}$ .

However, there are situations where it is possible to determine a set including all critical points without invoking algebraic transformations, as in the Teichmüller–Nakayama normal form: For general linear and nonlinear descriptor systems, it has been shown in [16, 17, 18, 20] that for sufficiently often differentiable coefficient functions there exist invariants (corresponding to ranks of submatrices) which are independent of the choice of transformation matrices, and the set of points where these quantities jump includes all critical points.

- (iv) If  $R(D)$  is not left invertible, then the set of points where the local behavior is nontrivial, i.e.,  $\{t \in \mathbb{R} \mid \ker_t R \neq \{0\}\}$ , is discrete.

*Remark 2.* Suppose that  $R(D)$  has constant coefficients, i.e.,  $R(D) \in \mathbb{R}[D]^{g \times q}$ .

- (i) If the class of unimodular transformations for the computation of the normal form (2.1) is restricted to  $\mathbb{R}[D]$ -unimodular matrices, then we arrive at the Smith normal form

$$(2.2) \quad R(D) = U(D)^{-1} \begin{bmatrix} \text{diag}\{r_1(D), \dots, r_\ell(D)\} & 0_{\ell \times (q-\ell)} \\ 0_{(g-\ell) \times \ell} & 0_{(g-\ell) \times (q-\ell)} \end{bmatrix} V(D)^{-1},$$

where  $U(D)$  and  $V(D)$  are  $\mathbb{R}[D]$ -unimodular matrices of sizes  $g$  and  $q$ , respectively, and  $r_i(D) \in \mathbb{R}[D]$  are nonzero monic polynomials with  $r_i \mid r_{i+1}$ ,  $i = 1, \dots, \ell - 1$ , where  $\ell = \text{rk}_{\mathbb{R}[D]} R(D)$  and  $r_i(D) = \psi_i(D)/\psi_{i-1}(D)$ ,  $\psi_0(\cdot) \equiv 1$ , and  $\psi_i(D)$  is the greatest common divisor of minors of order  $i$  of  $R(D)$ ; see, for example, [28, pp. 91–93].

Note that due to the smaller class of transformations, the Smith normal form is less simple than the Teichmüller–Nakayama normal form.

- (ii) Suppose in addition that  $\text{rk}_{\mathbb{R}[D]} R(D) = q$ . Then every local solution  $w \in \mathcal{C}_t^N(\mathbb{R}^q)$  of  $R(\frac{d}{dt})w = 0$ , where  $N$  is sufficiently large depending on  $\text{deg } R(D)$  and the degrees of the transformation matrices, can be continued to a global solution on  $\mathbb{R}$  and is even real analytic. This follows immediately from the Smith normal form (2.2) and the theory of linear time-invariant differential equations. Therefore, we may identify  $\ker_t R = \ker R$  for any  $t \in \mathbb{R}$ , and it follows that  $\dim \ker_t R = \sum_{i=1}^{\ell} \text{deg } r_i(D)$  for all  $t \in \mathbb{R}$ .

*Remark 3.* Suppose that  $R(D) \in \mathcal{M}[D]^{g \times q}$  has full rank and  $g \leq q$ . Let  $R(D)$  be factorized as in (2.1) and differently into

$$(2.3) \quad R(D) = \bar{U}(D)^{-1} \left[ \begin{array}{c|c} I_{g-1} & \\ \hline \bar{r}(D) & \end{array} \right]_{0_{g \times (q-g)}} \bar{V}(D)^{-1}.$$

Then a simple algebraic manipulation shows that

$$(2.4) \quad \bar{V}(D)^{-1}V(D) = \left[ \begin{array}{cc} W_1(D) & 0 \\ W_3(D) & W_4(D) \end{array} \right],$$

where  $W_1(D) \in \mathcal{M}[D]^{g \times g}$  and  $W_4(D) \in \mathcal{M}[D]^{(q-g) \times (q-g)}$  are unimodular, and where  $W_3(D) \in \mathcal{M}[D]^{(q-g) \times g}$ .

**3. Controllability.** In this section we introduce, study, and characterize the concept of controllability of systems (1.1). This is a generalization of the behavioral concept introduced by Willems [35]; see also [24].

**DEFINITION 3.1.** *Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  and  $t \in \mathbb{R}$ . A local subbehavior  $\mathfrak{B}_t$  of  $\ker_t R$ , i.e., a subset  $\mathfrak{B}_t \subset \ker_t R$ , is called locally controllable at  $t \in \mathbb{R}$  if and only if for every  $w^1, w^2 \in \mathfrak{B}_t$  and every  $t_0 \in (-\infty, t) \cap \text{dom } w^1$  there exist  $t_1 \in \text{dom } w^2 \cap (t, \infty)$  and  $w \in \mathfrak{B}_t$  such that*

$$w(t) = \begin{cases} w^1(t), & t \in (-\infty, t_0] \cap \text{dom } w^1, \\ w^2(t), & t \in [t_1, \infty) \cap \text{dom } w^2. \end{cases}$$

A behavior  $\mathfrak{B} = \bigcup_{t \in \mathbb{R}} \mathfrak{B}_t$ ,  $\mathfrak{B}_t \subset \ker_t R$ , is called controllable almost everywhere if and only if  $\mathfrak{B}_t$  is locally controllable for almost all  $t \in \mathbb{R}$ . Since  $\ker_t R$  is a real vector space by Remark 1(ii), the family of its linear subspaces may be partially ordered by inclusion, and thus constitutes a lattice with respect to  $+$  and  $\cap$ . Hence  $\ker_t^{\text{contr}} R \subset \ker_t R$  as largest controllable local behavior of  $\ker_t R$  is well defined. The set  $\ker^{\text{contr}} R = \bigcup_{t \in \mathbb{R}} \ker_t^{\text{contr}} R$  is called the largest controllable behavior of  $\ker R$ .

This concept is illustrated in Figure 2.

*Remark 4.*

- (i) Loosely speaking, controllability means that any two trajectories  $w^1, w^2 \in \ker_t R$  can be connected by another trajectory  $w \in \ker_t R$  so that in finite time  $w^1$  moves via  $w$  into  $w^2$ . A similar notion of controllability via trajectories was introduced in [11] for time-invariant Rosenbrock systems of the form (1.5). For time-invariant state space systems of the form (1.3), the concept of controllability coincides with the one introduced in [24, sect. 5.2].
- (ii) Since  $\ker_t R$  is a linear subspace, the trajectory  $w^2$  in Definition 3.1 may be replaced, without restriction of generality, by  $w^2 = 0$ .

We are now in position to prove the main theorem of this section, which characterizes controllability in algebraic terms. Recall that  $R(D)$  is called *right invertible* if and only if there exists some  $R^\#(D) \in \mathcal{M}[D]^{q \times g}$  such that  $R(D)R^\#(D) = I_g$ . In view of Remark 1(i) we assume that  $R(D)$  has full row rank.

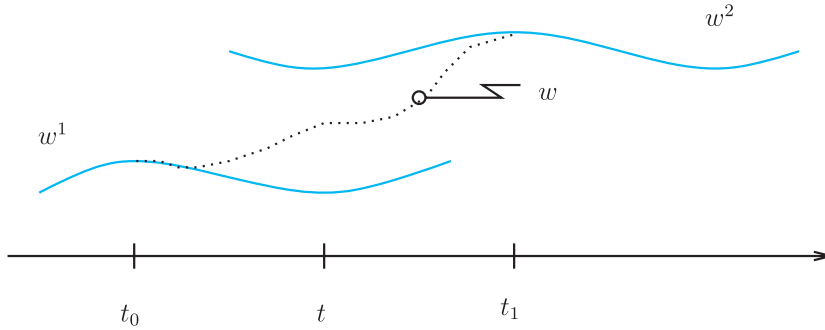


FIG. 2. Local controllability at  $t$ .

**THEOREM 3.2.** *Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  have full row rank. Then the behavior  $\ker R$  is controllable almost everywhere if and only if  $R(D)$  is right invertible.*

*Proof.* Suppose that  $R(D)$  is factorized as in (2.1) and let  $\mathbb{T} = \mathbb{T}(R, U, V, r)$  denote the discrete set given in Remark 1(iii). Then it remains to show that  $\ker_t R$  is locally controllable at  $t \in \mathbb{R} \setminus \mathbb{T}$  if and only if  $r(D)$  is a nonzero meromorphic function.

“ $\Rightarrow$ ”: Suppose that  $\deg r(D) \geq 1$  and  $t \in \mathbb{R} \setminus \mathbb{T}$ . By [34, Chap. IV] there exists an open interval  $\mathbb{I} \subset \mathbb{R} \setminus \mathbb{T}$  with  $t \in \mathbb{I}$  and some nonzero real analytic solution  $\varphi : \mathbb{I} \rightarrow \mathbb{R}$  which solves  $r(\frac{d}{dt})\varphi = 0$ . By the construction of  $\mathbb{T}$  and letting  $e_g$  denote the  $g$ th canonical basis vector in  $\mathbb{R}^q$ , it follows that

$$\hat{w}^1 := V(\frac{d}{dt})\varphi e_g \in \mathcal{C}^\omega(\mathbb{I}, \mathbb{R}^q)$$

and solves  $R(\frac{d}{dt})\hat{w}^1 = 0$ .

Seeking a contradiction, suppose that  $\ker_t R$  were locally controllable at  $t \in \mathbb{R}$ . Let  $t_0 \in (-\infty, t) \cap \mathbb{I}$ . Then there exist  $t_1 \in (t, \infty)$  and  $w \in \ker_t R$  such that

$$(3.1) \quad w(t) = \begin{cases} w^1(t), & t \in (-\infty, t_0] \cap \text{dom } w^1, \\ 0, & t \in [t_1, \infty). \end{cases}$$

Therefore,

$$\text{diag}\{1, \dots, 1, r(\frac{d}{dt}), 0, \dots, 0\}V(\frac{d}{dt})^{-1}w = 0 \quad \text{for all } t \in \text{dom } w,$$

which yields

$$V(\frac{d}{dt})^{-1}w =: [0, \dots, 0, \varphi_g, \dots, \varphi_q]^T \in \mathcal{C}^\omega(\text{dom } w, \mathbb{R}^q)$$

and  $r(\frac{d}{dt})\varphi_g(t) = 0$  for all  $t \in \text{dom } w$ . By (3.1) we have  $\varphi_g(t) = 0$  for all  $t \in [t_1, \infty)$ , and since  $\varphi_g$  is real analytic, the identity property of real analytic functions gives  $\varphi \equiv \varphi_g \equiv 0$ , which is a contradiction.

“ $\Leftarrow$ ”: Let  $t \in \mathbb{R} \setminus \mathbb{T}$ , let  $r(D)$  be meromorphic and nonzero, and let  $w^1 \in \ker_t R$ . Then there exists some open interval  $\mathbb{I} := (\tau_0, \tau_1) \subset (\mathbb{R} \setminus \mathbb{T}) \cap \text{dom } w^1$  with  $t \in \mathbb{I}$  such that

$$w^1 =: V(\frac{d}{dt})[0, \dots, 0, \varphi_{g+1}, \dots, \varphi_q]^T \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q).$$

Choose  $\delta \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$  such that

$$\delta(t) = \begin{cases} 1, & t \leq \tau_0, \\ 0, & t \geq \tau_1. \end{cases}$$

Then

$$w := V\left(\frac{d}{dt}\right) \delta [0, \dots, 0, \varphi_{g+1}, \dots, \varphi_q]^T \in C^\infty(\mathbb{I}, \mathbb{R}^q)$$

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

$$w(t) = \begin{cases} w^1(t), & t \leq \tau_0, \\ 0, & t \geq \tau_1. \end{cases}$$

This completes the proof.  $\square$

For time-invariant systems (1.1), Theorem 3.2 is derived differently in [24, Thm. 5.2.10].

*Remark 5.* For time-varying systems (1.3) or (1.5), it is well known that controllability of the system yields that it can be controlled in arbitrary short time. The proof of Theorem 3.2, in particular the choice of  $(\tau_0, \tau_1)$  and  $\delta$ , shows that this is also valid for the behavior  $\ker_t R$ : If  $\ker_t R$  is controllable, then  $t_0 < t$  and  $t_1 > t$  in Definition 3.1 can be replaced by any  $t'_0 < t < t'_1$  arbitrary close to  $t$ .

In the following remark we recall the classical concept of controllability for time-varying state space systems and clarify the set of admissible input functions.

*Remark 6.* Controllability for state space systems (1.3) means (see, for example, [33, Def. 3.1.6]) that for any  $x^0, x^1 \in \mathbb{R}^n$  and  $t_0 \in \mathbb{R}$ , there exist  $t_1 > t_0$  and a continuous function  $u : [t_0, t_1] \rightarrow \mathbb{R}^m$  such that

$$x(t) = (Lu)(t) := \Phi(t, t_0)x^0 + \int_{t_0}^t \Phi(t, \tau)B(\tau)u(\tau)d\tau, \quad t \in [t_0, t_1],$$

satisfies  $x(t_1) = x^1$ . Here  $\Phi$  denotes the transition matrix of the homogeneous system  $\dot{x} = Ax$ .

Using the fact that the set of  $C^\infty$ -functions with support in  $[t_0, t_1]$  lies dense, with respect to the  $\mathcal{L}^1$ -norm, in the set of piecewise continuous functions with support included in  $[t_0, t_1]$ , it follows from a straightforward modification of the proof of Lemma A2 in [14] that, for all  $t \in (t_0, t_1)$ ,

$$\begin{aligned} & \{(Lu)(t) \mid u \in C^\infty((t_0, t_1), \mathbb{R}^m)\} \\ &= \{(Lu)(t) \mid u : [t_0, t_1] \rightarrow \mathbb{R}^m \text{ piecewise continuous with } \text{supp } u \subset [t_0, t_1]\}. \end{aligned}$$

Therefore, although in the original definition  $u$  is required to be continuous, we may choose, without any restriction of generality,  $u \in C^\infty(\mathbb{R}, \mathbb{R}^m)$  with  $\text{supp } u \subset [t_0, t_1]$ .

In the following proposition, it is shown how controllability encompasses other well-established controllability concepts.

**PROPOSITION 3.3.** *Consider a time-varying Rosenbrock system of the form (1.5) with corresponding  $R(D)$  as defined in (1.7), and suppose that  $R(D)$  has full row rank. Then the following conditions are equivalent:*

- (i)  $\ker R$  is controllable almost everywhere.
- (ii)  $[P(D), -Q(D)]$  is right invertible.
- (iii)  $\ker[P, Q]$  is controllable almost everywhere.
- (iv) (1.5) is controllable in the sense defined in [14].
- (v) If  $R(D)$  represents a time-invariant Rosenbrock system (1.5), then (1.5) is controllable in the sense defined in [11].

(vi) If  $R(D)$  represents a state space system (1.3) with corresponding  $R(D)$  as defined in (1.7), then (1.3) is controllable in the classical sense as, for example, given in [33, Def. 3.1.6].

*Proof.* The equivalences “(i)  $\Leftrightarrow$  (ii)  $\Leftrightarrow$  (iii)” follow from Theorem 3.2 and simple algebraic manipulations; “(ii)  $\Leftrightarrow$  (iv)” follows from [14, Thm. 6.4]. “(ii)  $\Leftrightarrow$  (v)” follows from [11, Cor. 7.3]. It remains to prove that the classical concept of controllability as given in Remark 6 is encompassed in the behavioral setup. It is easy to see that (iii) implies (vi), and we omit the proof. To prove the converse, suppose that (vi) holds. Then for given

$$(x^i, u^i) \in C^\infty(\mathbb{R}, \mathbb{R}^n) \times C^\infty(\mathbb{R}, \mathbb{R}^m) \quad \text{such that} \quad \frac{d}{dt}x^i(t) = A(t)x^i(t) + B(t)u^i(t), \quad i = 1, 2,$$

and given  $t_0 \in \mathbb{R}$ , we need to find

$$(x, u) \in C^\infty(\mathbb{R}, \mathbb{R}^n) \times C^\infty(\mathbb{R}, \mathbb{R}^m), \quad \text{so that} \quad \frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t),$$

and  $t_1 > t_0$  such that

$$(3.2) \quad (x(t), u(t)) = \begin{cases} (x^1(t), u^1(t)) & \text{for all } t \leq t_0, \\ (x^2(t), u^2(t)) & \text{for all } t \geq t_1. \end{cases}$$

Let  $\bar{x}^1 = x^1(t_0)$  and, for arbitrary but fixed  $t_1 > t_0$ , let  $\bar{x}^2 = x^2(t_1)$ . Then by (vi) we may choose  $\hat{u} \in C^\infty(\mathbb{R}, \mathbb{R}^m)$  with  $\text{supp } \hat{u} \subset [t_0, t_1]$  such that

$$x(t) = \Phi(t, t_0)\bar{x}_1 + \int_{t_0}^t \Phi(t, \tau)B(\tau)\hat{u}(\tau)d\tau \quad \text{satisfies} \quad x(t_2) = \bar{x}^2.$$

Define, for all  $t \in \mathbb{R}$ ,

$$u(t) = \begin{cases} u^1(t) & \text{for all } t \leq t_0, \\ \hat{u}(t) & \text{for all } t \in (t_0, t_1), \\ u^2(t) & \text{for all } t \geq t_1 \end{cases} \quad \text{and} \quad x(t) = \Phi(t, t_0)\bar{x}_1 + \int_{t_0}^t \Phi(t, \tau)B(\tau)u(\tau)d\tau.$$

Then  $(x, u)$  satisfies  $\dot{x} = Ax + Bu$  and (3.2). The function  $u$  is in general not infinitely many times differentiable at  $t_0$  or at  $t_1$ , but applying Remark 6, one may replace  $\hat{u}$  so that  $u \in C^\infty(\mathbb{R}, \mathbb{R}^m)$ . This completes the proof.  $\square$

Next we study, for  $R(D) \in \mathcal{M}[D]^{g \times q}$ , the relationship between the local kernel representation  $\ker_t R$  and the local image representation at  $t \in \mathbb{R}$ , i.e., for some  $M(D) \in \mathcal{M}[D]^{q \times m}$ , the real vector space

$$\text{im}_t M := \{w \in C_t^\infty(\mathbb{R}^q) \mid \exists l \in C_t^\infty(\mathbb{R}^m) \text{ for all } \tau \in \text{dom } w \cap \text{dom } l : w(\tau) = M(\frac{d}{dt})l(\tau)\}.$$

**PROPOSITION 3.4.** *Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  have full row rank.  $\ker R$  is controllable almost everywhere if and only if there exist  $m \in \mathbb{N}$  and  $M(D) \in \mathcal{M}[D]^{q \times m}$  such that  $\ker_t R = \text{im}_t M$  for almost all  $t \in \mathbb{R}$ .*

*Proof.* Suppose  $R(D)$  is factorized as in (2.1) and let  $\mathbb{T}$  denote the discrete set given in Remark 1. By Theorem 3.2 it remains to show that  $r(D)$  is a nonzero meromorphic function if and only if  $\ker_t R = \text{im}_t M$  for all  $t \in \mathbb{R} \setminus \mathbb{T}$ .

“ $\Rightarrow$ ”: Set

$$M(D) := V(D) \begin{bmatrix} 0_{g \times (q-g)} \\ I_{q-g} \end{bmatrix}.$$

Then  $\text{im}_t M \subset \ker_t R$  for all  $t \in \mathbb{R} \setminus \mathbb{T}$  is immediate. If  $w \in \ker_t R$  for  $t \in \mathbb{R} \setminus \mathbb{T}$ , then  $r(D)$  being nonzero and meromorphic yields

$$[I_g \mid 0_{g \times (q-g)}] V(\frac{d}{dt})^{-1} w(t) = 0 \quad \text{for all } t \in \text{dom } w \cap (\mathbb{R} \setminus \mathbb{T}),$$

and so there exists  $l \in \mathcal{C}_t^\infty(\mathbb{R}^m)$  such that

$$V(\frac{d}{dt})^{-1} w = \begin{bmatrix} 0_{g \times (q-g)} \\ I_{(q-g)} \end{bmatrix} l.$$

“ $\Leftarrow$ ”: Let  $t \in \mathbb{R} \setminus \mathbb{T}$  and choose an open interval  $\mathbb{I} \subset (\mathbb{R} \setminus \mathbb{T})$  with  $t \in \mathbb{I}$ . Seeking a contradiction, by Theorem 3.2 one may assume that  $\text{deg } r(D) \geq 1$ . Comparing the  $g$ th components of the identical vector spaces

$$\left\{ w \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \mid \begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} \begin{bmatrix} 0_{g \times (q-g)} \\ I_{(q-g)} \end{bmatrix} V(\frac{d}{dt})^{-1} w = 0 \right\}$$

and  $\{w \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \mid \exists l \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^m) : w = M(\frac{d}{dt})l\}$  yields that

$$\begin{aligned} \dim\{(V(\frac{d}{dt})^{-1} w(t))_g \mid w \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \wedge r(\frac{d}{dt})(V(\frac{d}{dt})^{-1} w(t))_g = 0\} \\ = \dim\{(V(\frac{d}{dt})^{-1} M(\frac{d}{dt})l(t))_g \mid l \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^m)\}. \end{aligned}$$

However, the former has finite dimension  $\text{deg } r(D) \geq 1$ , while the latter is zero-dimensional or has infinite dimension. This is a contradiction, and hence the proof of the proposition is complete.  $\square$

Proposition 3.4 is known for time-invariant systems; see [24, Thm. 6.6.1]. However, the different proof presented here might also be of interest in the time-invariant case.

In the following proposition we show how to present the largest controllable behavior in terms of the nonunique factorization (2.1).

PROPOSITION 3.5. *If  $R(D) \in \mathcal{M}[D]^{g \times q}$  is factorized as in (2.1), then we have*

$$\ker_t^{\text{contr}} R = \{w \in \ker_t R \mid [I_g, 0_{g \times (q-g)}] V(\frac{d}{dt})^{-1} w = 0\} \quad \text{for almost all } t \in \mathbb{R}.$$

*Proof.* Since  $[I_g, 0_{g \times (q-g)}] V(D)^{-1}$  is right invertible, it follows from Theorem 3.2 that

$$\ker_t^c R := \{w \in \ker_t R \mid [I_g, 0] V(\frac{d}{dt})^{-1} w = 0\}$$

is a controllable behavior almost everywhere. Therefore, we have to show that  $\ker_t^{\text{contr}} R \subset \ker_t^c R$  almost everywhere. Let  $\mathbb{T}$  denote the union of all zeros and poles of the meromorphic coefficients in all entries of  $U(D)$ ,  $U(D)^{-1}$ ,  $V(D)$ ,  $V(D)^{-1}$ ,  $r(D)$ ,  $\bar{U}(D)$ ,  $\bar{U}(D)^{-1}$ ,  $\bar{V}(D)$ ,  $\bar{V}(D)^{-1}$ , and  $\bar{r}(D)$ . Then  $\mathbb{T}$  is a discrete set. Let  $w \in \ker_t^{\text{contr}} R$  for  $t \in \mathbb{R} \setminus \mathbb{T}$ . Choose an open interval  $\mathbb{I} \subset \mathbb{T}$  with  $t \in \mathbb{I}$ . Then

$$V(\frac{d}{dt})^{-1} w =: [0, \dots, 0, \varphi_g, \dots, \varphi_q]^T \in \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \quad \text{and} \quad r(\frac{d}{dt})\varphi_g = 0.$$

The function  $\varphi_g$ , as a solution of a linear ordinary differential equation with real analytic coefficients on  $\mathbb{I}$ , is real analytic on  $\mathbb{I}$  itself. Therefore, the normal form (2.1) and the identity property of analytic function yields  $\varphi_g \equiv 0$ . This proves  $w(t) = V(\frac{d}{dt}) [0, \dots, 0, \varphi_{g+1}, \dots, \varphi_q]^T \in \ker_t^c R$ .

If  $R(D)$  is factorized as in (2.3), then by Remark 3 one concludes that

$$[I_g, 0] \bar{V} \left(\frac{d}{dt}\right)^{-1} w = [I_g, 0] \begin{bmatrix} W_1\left(\frac{d}{dt}\right) & 0 \\ W_3\left(\frac{d}{dt}\right) & W_4\left(\frac{d}{dt}\right) \end{bmatrix} V \left(\frac{d}{dt}\right)^{-1} w = [W_1\left(\frac{d}{dt}\right), 0] V \left(\frac{d}{dt}\right)^{-1} w,$$

and the result follows, since  $W_1(D)$  is unimodular. This completes the proof.  $\square$

*Example 1.* Revisiting example (1.9), we now can show that this system is locally controllable almost everywhere.

Without loss of generality, we may assume that the coordinate system for the Lagrange multipliers is such that  $F_0 = [F_1 \ 0]$  with nonsingular  $F_1 \in \mathbb{R}^{2 \times 2}$ , and we partition

$$-K_0 = \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}, \quad M_0 = \begin{bmatrix} M_{11}(t) & M_{12}(t) \\ M_{21}(t) & M_{22}(t) \end{bmatrix}, \quad -D_0 = \begin{bmatrix} D_{11}(t) & D_{12}(t) \\ D_{21}(t) & D_{22}(t) \end{bmatrix},$$

and

$$S_0 = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix},$$

with  $K_{11}(t), M_{11}(t), D_{11}(t) \in \mathbb{R}^{2 \times 2}$ ,  $S_1 \in \mathbb{R}^{2 \times 3}$ , and all other formats accordingly. Then the system (1.9), for  $t \in \mathbb{I}$ , may be written as

$$\begin{bmatrix} I_2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & M_{11}(t) & M_{12}(t) & 0 \\ 0 & 0 & M_{21}(t) & M_{22}(t) & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & I_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ K_{11}(t) & K_{12}(t) & D_{11}(t) & D_{12}(t) & F_1^T \\ K_{21}(t) & K_{22}(t) & D_{21}(t) & D_{22}(t) & 0 \\ F_1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ S_1 \\ S_2 \\ 0 \end{bmatrix} u.$$

Since  $F_1$  is nonsingular and  $S_1, S_2$  are constant matrices of full row rank, it follows that  $x_1 = 0$  and  $\dot{x}_1 = 0$ , whence  $x_3 = 0$ . Therefore, (1.9) is equivalent to

$$\begin{bmatrix} D & -1 & 0_{2 \times 1} & 0 \\ -K_{12}(t) & M_{12}(t) & -F_1 & S_1 \\ -K_{22}(t) & M_{22}(t) & 0 & S_2 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \\ x_5 \\ u \end{bmatrix} = 0,$$

with corresponding right invertible matrix  $R(D)$ . By Theorem 3.2, the system (1.9) is locally controllable almost everywhere on  $\mathbb{I}$ .

**4. Autonomous behavior.** In this section we show that the local behavior (in the sense almost everywhere) can be decomposed into the direct sum of the controllability subspace and an autonomous subspace.

**DEFINITION 4.1.** Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  and  $t \in \mathbb{R}$ . A local subbehavior  $\mathfrak{B}_t \subset \ker_t R$  is called autonomous if and only if for any  $w^1, w^2 \in \mathfrak{B}_t$  with  $w^1 \equiv w^2$  on some open interval  $\mathbb{I} \subset \text{dom } w^1 \cap \text{dom } w^2$  with  $t \in \mathbb{I}$  it follows that  $w^1 \equiv w^2$  on  $\text{dom } w^1 \cap \text{dom } w^2$ .

A behavior  $\mathfrak{B} = \bigcup_{t \in \mathbb{R}} \mathfrak{B}_t \subset \ker_t R$  is called autonomous if and only if  $\mathfrak{B}_t$  is autonomous for almost all  $t \in \mathbb{R}$ .

The above definition is a generalization of autonomous subbehavior of time-invariant systems as, for example, defined in [24, p. 67].

PROPOSITION 4.2. Consider  $R(D) \in \mathcal{M}[D]^{g \times q}$  with factorization (2.1) and  $\text{rk } R(D) = g$ . Then for any autonomous behavior  $\ker^{\text{aut}} R$ , the following properties hold:

- (i)  $\ker_t^{\text{aut}} R \cap \ker_t^{\text{contr}} R = \{0\}$  for almost all  $t \in \mathbb{R}$ .
- (ii) If  $w \in \ker_t^{\text{aut}} R$ , then

$$\begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} V(\frac{d}{dt})^{-1} w = 0.$$

(iii)

$$\left\{ w \in \mathcal{C}_t^\infty(\mathbb{R}^q) \mid \begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} V(\frac{d}{dt})^{-1} w = 0 \right\}$$

is an autonomous behavior for almost all  $t \in \mathbb{R}$ .

- (iv) The behavior  $\ker R$  is autonomous if and only if  $R(D)$  has full column rank.

*Proof.* (i) If  $w \in \ker^{\text{aut}} R$  and  $w \neq 0$ , then it cannot belong to the controllable behavior; otherwise Definition 4.1 would be violated.

(ii) By (i) and Proposition 3.5, any  $w \in \ker_t^{\text{aut}} R$  satisfies  $[0_{(q-g) \times g}, I_{q-g}] V(\frac{d}{dt})^{-1} w = 0$ . Hence (ii) follows from (2.1).

(iii) Let  $\mathbb{T}$  denote the discrete set given in Remark 1 and let  $t \in \mathbb{R} \setminus \mathbb{T}$ . If  $w \in \ker_t R$  and satisfies

$$\begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} V(\frac{d}{dt})^{-1} w = 0,$$

then (2.1) yields that  $w$  is of the form

$$w = V(\frac{d}{dt}) [0, \dots, 0, \varphi_g, 0, \dots, 0]^T$$

for some  $\varphi_g \in \mathcal{C}_t^\infty(\mathbb{R})$  with  $r(\frac{d}{dt})\varphi_g = 0$ . Since  $r$  has real analytic coefficients, the solution is real analytic, too, and the identity property of real analytic functions ensures local uniqueness of  $w$  as in Definition 4.1. This completes the proof.

(iv) This statement follows immediately from the definition and from Theorem 2.1.  $\square$

Note that the representation of the autonomous behavior in Proposition 4.2(iii) is not uniquely defined; it depends on the factorization (2.1). This holds true already for time-invariant systems; see [24, Rem. 5.2.15]. However, the dimension of this autonomous behavior is unique; this follows from the fact that  $r(D)$  is unique up to similarity, and the latter preserves the degree; see Theorem 2.1. For time-invariant systems (1.1), the results of Proposition 4.2 can be found in [24, sect. 5.2]. More importantly, the sum of an autonomous behavior and the controllable behavior is indeed uniquely defined. In the following we generalize this result to time-varying systems.



THEOREM 4.3. Consider the system  $R(\frac{d}{dt})w = 0$  with  $R(D) \in \mathcal{M}[D]^{g \times q}$  and  $\text{rk } R(D) = g$ , factorizations (2.1), (2.3), and define, for all  $t \in \mathbb{R}$ ,

$$\begin{aligned} \ker_t^{\text{contr}} R &= \{w \in \ker_t R \mid [I_g, \quad 0_{g \times (q-g)}] V(\frac{d}{dt})^{-1}w = 0\}, \\ \ker_t^{\text{aut}} R &= \left\{w \in \ker_t R \mid \begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} V(\frac{d}{dt})^{-1}w = 0\right\}, \\ \overline{\ker}_t^{\text{aut}} R &= \left\{w \in \ker_t R \mid \begin{bmatrix} I_{g-1} & \\ & \bar{r}(\frac{d}{dt}) \end{bmatrix} \bar{V}(\frac{d}{dt})^{-1}w = 0\right\}, \end{aligned}$$

where the latter is defined with respect to (2.3). Then

$$(4.1) \quad \ker_t R = \ker_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R = \overline{\ker}_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R \quad \text{for almost all } t \in \mathbb{R}.$$

*Proof.* Let  $\mathbb{T}$  denote the union of all zeros and poles of the meromorphic coefficients in all entries of  $U(D)$ ,  $U(D)^{-1}$ ,  $V(D)$ ,  $V(D)^{-1}$ ,  $r(D)$  and  $\bar{U}(D)$ ,  $\bar{U}(D)^{-1}$ ,  $\bar{V}(D)$ ,  $\bar{V}(D)^{-1}$ ,  $\bar{r}(D)$ .  $\mathbb{T}$  is a discrete set. In the following we consider  $t \in \mathbb{R} \setminus \mathbb{T}$  and an open interval  $\mathbb{I} \subset \mathbb{T}$  with  $t \in \mathbb{I}$ . We proceed in several steps.

*Step 1.* By Proposition 4.2(i) the sums in (4.1) are direct sums.

*Step 2.* The inclusion

$$\ker_t R \supset \ker_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R$$

follows from the definition of  $\ker_t^{\text{aut}} R$  and  $\ker_t^{\text{contr}} R$ .

*Step 3.* We show

$$(4.2) \quad \ker_t R \subset \ker_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R.$$

Let  $w \in \ker_t R$  and set

$$[\varphi_1, \dots, \varphi_q]^T := V(\frac{d}{dt})^{-1}w \in \mathcal{C}^\infty(\mathbb{I}; \mathbb{R}^q).$$

Then

$$\begin{bmatrix} I_{g-1} & \\ & r(\frac{d}{dt}) \end{bmatrix} V(\frac{d}{dt})^{-1}w = 0,$$

and hence

$$[\varphi_1, \dots, \varphi_q]^T = [0, \dots, 0, \varphi_g, 0, \dots, 0]^T \quad \text{with } r(\frac{d}{dt})\varphi_g = 0.$$

Finally

$$\begin{aligned} w_1 &:= V(\frac{d}{dt})^{-1} [0, \dots, 0, \varphi_g, 0, \dots, 0]^T \in \ker_t^{\text{aut}} R, \\ w_2 &:= V(\frac{d}{dt})^{-1} [0, \dots, 0, \varphi_{g+1}, \dots, \varphi_q]^T \in \ker_t^{\text{contr}} R \end{aligned}$$

yields  $w_1 + w_2 = w$ , whence (4.2).

*Step 4.* We show

$$\ker_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R \subset \overline{\ker}_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R.$$

Let  $w_1 \in \ker_t^{\text{aut}} R$  and  $w_2 \in \ker_t^{\text{contr}} R$ . Then

$$\begin{aligned} [0, \dots, 0, \varphi_g, 0, \dots, 0]^T &:= V(\frac{d}{dt})^{-1}w_1 \in \mathcal{C}^\infty(\mathbb{I}; \mathbb{R}^q) \quad \text{with } r(\frac{d}{dt})\varphi_g = 0, \\ [0, \dots, 0, \varphi_{g+1}, \dots, \varphi_q]^T &:= V(\frac{d}{dt})^{-1}w_2 \in \mathcal{C}^\infty(\mathbb{I}; \mathbb{R}^q). \end{aligned}$$

Since  $w := w_1 + w_2 \in \ker_t R$ , it follows from (2.3) that

$$\bar{V}\left(\frac{d}{dt}\right)^{-1}w = [0, \dots, 0, \bar{\varphi}_g, \dots, \bar{\varphi}_q]^T \in \mathcal{C}^\infty(\mathbb{I}; \mathbb{R}^q) \quad \text{with } r\left(\frac{d}{dt}\right)\bar{\varphi}_g = 0.$$

Finally, setting

$$\begin{aligned} \bar{w}_1 &:= \bar{V}\left(\frac{d}{dt}\right)^{-1} [0, \dots, 0, \bar{\varphi}_g, 0, \dots, 0]^T \in \overline{\ker}^{\text{aut}} R, \\ \bar{w}_2 &:= \bar{V}\left(\frac{d}{dt}\right)^{-1} [0, \dots, 0, \bar{\varphi}_{g+1}, \dots, \bar{\varphi}_q]^T \in \overline{\mathfrak{B}}_R^{\text{contr}} \end{aligned}$$

shows  $w = \bar{w}_1 + \bar{w}_2 \in \overline{\ker}_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R$ .

*Step 5.* The inclusion

$$\ker_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R \supset \overline{\ker}_t^{\text{aut}} R \oplus \ker_t^{\text{contr}} R$$

follows by symmetry as in Step 4. This completes the proof of the theorem.  $\square$

**5. Observability.** In this section, we study how one behavior can be observed from another. Essential for this are the concepts of adjoints of matrices over  $\mathcal{M}[D]$  and the adjoint of a kernel representation  $\ker R$ .

DEFINITION 5.1. *The adjoint for matrices over  $\mathcal{M}[D]$  is defined as*

$$\cdot^{\text{ad}} : \mathcal{M}^{n \times m}[D] \rightarrow \mathcal{M}^{m \times n}[D], \quad \sum_{i=0}^k P_i D^i \mapsto \left( \sum_{i=0}^k P_i D^i \right)^{\text{ad}} := \sum_{i=0}^k (-1)^i D^i P_i^T.$$

PROPOSITION 5.2. *The adjoint is an anti-isomorphism; i.e., it is surjective, injective, and satisfies, for arbitrary matrices  $P(D), Q(D)$  over  $\mathcal{M}[D]$  with appropriate formats,*

$$(5.1) \quad [P(D) + Q(D)]^{\text{ad}} = P(D)^{\text{ad}} + Q(D)^{\text{ad}},$$

$$(5.2) \quad [P(D) \cdot Q(D)]^{\text{ad}} = Q(D)^{\text{ad}} \cdot P(D)^{\text{ad}}.$$

*Proof.* Surjectivity, injectivity, and addition are straightforward. It remains to prove the antimultiplication rule (5.2). This is well known in the scalar case; see, for example, [21, p. 25]. To prove the matrix case, denote the entries of  $P(D) \in \mathcal{M}^{n \times m}[D], Q(D) \in \mathcal{M}^{m \times l}[D]$  by  $p_{ij}(D), q_{ij}(D)$ , respectively. Then

$$P(D)^{\text{ad}} = (p_{ji}(D)^{\text{ad}})_{1 \leq i \leq n, 1 \leq j \leq m}, \quad Q(D)^{\text{ad}} = (q_{ji}(D)^{\text{ad}})_{1 \leq i \leq m, 1 \leq j \leq l}$$

and applying this to

$$(P(D) \cdot Q(D))_{ij} = \sum_{\lambda=1}^k p_{i\lambda}(D) q_{\lambda j}(D)$$

and using the antimultiplication rule (5.2) for scalar polynomials yield the result. This completes the proof.  $\square$

DEFINITION 5.3. *Let  $R(D) \in \mathcal{M}[D]^{g \times q}$  and let  $t \in \mathbb{R}$ . The local adjoint of the kernel representation  $\ker_t R$  is the image representation  $\text{im}_t R^{\text{ad}}$ , i.e.,  $(\ker_t R)^{\text{ad}} = \text{im}_t R^{\text{ad}}$ .*

Certainly, the projection onto the first  $q$  components of the kernel representation

$$\left\{ (\tilde{w}, l) \in \mathcal{C}_t^\infty(\mathbb{R}^q) \times \mathcal{C}_t^\infty(\mathbb{R}^g) \mid \text{for all } \tau \in \text{dom } \tilde{w} \cap \text{dom } l : \left[ I_q, R\left(\frac{d}{d\tau}\right)^{\text{ad}} \right] \begin{bmatrix} \tilde{w}(\tau) \\ l(\tau) \end{bmatrix} = 0 \right\}$$

yields the image representation  $\text{im}_t R^{\text{ad}}$ .

The following definition is a straightforward generalization of observability for time-invariant systems in the behavioral setup; see [24, Def. 5.3.2].

DEFINITION 5.4. *Let  $[R_1(D), R_2(D)] \in \mathcal{M}[D]^{g \times (q_1 + q_2)}$  and let  $t \in \mathbb{R}$ . Then  $w_2 \in \mathcal{C}_t^\infty(\mathbb{R}^{q_2})$  is called locally observable at  $t \in \mathbb{R}$  from  $w_1 \in \mathcal{C}_t^\infty(\mathbb{R}^{q_1})$  for  $t \in \mathbb{R}$  if and only if*

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \begin{bmatrix} w_1 \\ \tilde{w}_2 \end{bmatrix} \in \ker_t [R_1, R_2]$$

implies that

$$\text{for all } \tau \in \text{dom } w_2 \cap \text{dom } \tilde{w}_2 : w_2(\tau) = \tilde{w}_2(\tau).$$

An algebraic characterization of observability is given in the following theorem.

THEOREM 5.5. *Let  $[R_1(D), R_2(D)] \in \mathcal{M}[D]^{g \times (q_1 + q_2)}$ . Then  $w_2$  is locally observable almost everywhere from  $w_1$  if and only if  $R_2(D)$  is left invertible.*

*Proof.* First note that in view of the linearity of the system, it remains to show that for almost all  $t \in \mathbb{R}$  we have

$$[w_2 \in \ker_t R_2 \implies w_2 = 0] \iff R_2(D) \text{ is left invertible.}$$

“ $\Leftarrow$ ” is immediate.

“ $\Rightarrow$ ”: Let  $\mathbb{T}$  denote the discrete set of the union of all zeros and poles of the meromorphic coefficients in all entries of  $U_2(D), U_2(D)^{-1}, V_2(D), V_2(D)^{-1}, r_2(D)$  which take  $R_2(D)$  into a normal form (2.1).

Seeking a contradiction, suppose  $R_2(D)$  is not left invertible and let  $t \in \mathbb{T}$ . Now either  $\text{rk}_{\mathcal{M}[D]} R_2(D) < q_2$  (in which case the normal form (2.1) applied to  $R_2(D)$  yields the existence of some  $w_2 \in \ker_t R_2$  with  $w_2 \neq 0$ ) or, again by Theorem 2.1, there exist  $r_2(D) \in \mathcal{M}[D]$  with  $\deg r_2(D) \geq 1$  and unimodular  $U_2(D) \in \mathcal{M}[D]^{g \times g}$ ,  $V_2(D) \in \mathcal{M}[D]^{q_2 \times q_2}$  such that

$$(5.3) \quad U_2(D)^{-1} R_2(D) V_2(D)^{-1} = \begin{bmatrix} I_{q_2-1} & 0_{(q_2-1) \times 1} \\ 0_{1 \times (q_2-1)} & r_2(D) \\ 0_{(g-q_2) \times (q_2-1)} & 0 \end{bmatrix}.$$

By  $\deg r_2(D) \geq 1$  there exists  $\varphi \in \mathcal{C}_t^\infty(\mathbb{R}) \setminus \{0\}$  such that  $r_2(\frac{d}{dt})\varphi = 0$ . Therefore  $w_2 := (0, \dots, 0, \varphi)^T \in \ker_t R_2$ , which is a contradiction. This completes the proof.  $\square$

The following theorem relates the concepts of controllability and observability.

THEOREM 5.6. *For  $R(D) \in \mathcal{M}[D]^{g \times q}$  the following two statements are equivalent:*

- (i) *The behavior  $\ker R$  is locally controllable almost everywhere.*
- (ii) *The variable  $l$  is locally observable almost everywhere from  $w$  with respect to the behavior induced by*

$$[I_q, R^{\text{ad}}] \begin{pmatrix} w \\ l \end{pmatrix} = 0.$$

*Proof.* By Theorem 3.2, statement (i) is equivalent to  $R(D)$  being right invertible, which, by Proposition 5.2, is equivalent to  $R(D)^{\text{ad}}$  being left invertible. The latter is, by invoking Proposition 5.5, equivalent to assertion (ii). This completes the proof of the theorem.  $\square$

In order to relate the classical concepts of observability known in the literature to observability as introduced above, we have to permute the columns in the presentation (1.5), (1.7) in the following proposition.

PROPOSITION 5.7. *For a time-varying Rosenbrock system of the form (1.5) represented in the form*

$$R(D) = [R_1(D), R_2(D)], \quad R_1(D) = \begin{bmatrix} -Q(D), & 0 \\ W(D), & -I_p \end{bmatrix}, \quad R_2(D) = \begin{bmatrix} P(D) \\ V(D) \end{bmatrix},$$

the following conditions are equivalent:

- (i)  $w_2$  is locally observable from  $w_1$  almost everywhere with respect to the behavior induced by

$$[R_1(\frac{d}{dt}), R_2(\frac{d}{dt})] \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = 0.$$

- (ii)  $R_2(D)$  is left invertible.
- (iii)  $[R_1(D), R_2(D)]$  is observable in the sense defined in [14].
- (iv) If  $R(D)$  represents a time-invariant Rosenbrock system, then it is observable in the sense defined in [11].
- (v) If  $R(D)$  represents a state space system (1.3) in the form

$$R_1(D) = \begin{bmatrix} -B & 0 \\ -F & I_p \end{bmatrix}, \quad R_2(D) = \begin{bmatrix} DI_n - A \\ -C \end{bmatrix},$$

then it is observable in the classical sense; see, for example, [30].

*Proof.* The equivalence “(i)  $\Leftrightarrow$  (ii)” follows from Theorem 5.5. The equivalences “(ii)  $\Leftrightarrow$  (iii)” and “(ii)  $\Leftrightarrow$  (iv)” follow from [14, Thm. 6.5] and [11, Cor. 7.6], respectively. They all can be shown directly, but only for state space systems we prove “(i)  $\Leftrightarrow$  (v)” directly; it shows how observability in the classical sense and in the behavioral setup are related. Note that in the case of time-varying state space systems and time-invariant Rosenbrock systems the set of critical points  $\mathbb{T}$  is empty, and the system is defined on the whole time axis.

Complete observability for time-varying state space systems of the form (1.3) means (see [30, Def. 9.7]) that for any open interval  $\mathbb{I} \subset \mathbb{R}$  we have

$$(5.4) \quad \begin{bmatrix} \frac{d}{dt}I_n - A(t) \\ -C(t) \end{bmatrix} z(t) = 0 \quad \text{for all } t \in \mathbb{I} \quad \implies \quad z(t) = 0 \quad \text{for all } t \in \mathbb{I}.$$

(5.4) is equivalent to  $R_2(D)$  being left invertible, and hence “(i)  $\Leftrightarrow$  (v)” follows from Theorem 5.5. This completes the proof of the theorem.  $\square$

*Example 2.* Revisiting Example 1, see also (1.9), with

$$(5.5) \quad C = \begin{bmatrix} 0 & 0 & I_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

corresponding to measuring the positions, we see that the resulting matrix

$$\begin{bmatrix} E(t)D - A(t) \\ C \end{bmatrix}$$

is left invertible if and only if the matrix

$$\begin{bmatrix} D & -1 & 0 \\ -K_{12}(t) & M_{12}(t)D & -F_1 \\ -K_{22}(t) & M_{22}(t)D & 0 \end{bmatrix}$$

is left invertible over the ring  $\mathcal{M}[D]$ , which holds if and only if  $K_{22}(t)$  is nonzero. The latter is typically the case in practice, since the stiffness matrix  $K_0(t)$  is symmetric and positive definite. An application of Theorem 5.5 yields the following:  $x$  is locally observable from  $(u, y)$  at  $t$  with respect to the system (1.9), (5.5) if and only if  $K_{22}(t)$  is nonzero.

**6. Latent variables and elimination.** In [24, sect. 6.2], full and manifest behavior is considered for time-invariant systems. We do not repeat these definitions for time-varying systems but show a time-varying version of the crucial Theorem 6.2.6 in [24].

**THEOREM 6.1.** *Let  $[R(D), S(D)] \in \mathcal{M}[D]^{g \times (q+s)}$ . Then there exists  $R'(D) \in \mathcal{M}[D]^{g' \times q}$  such that, for almost all  $t \in \mathbb{R}$ ,*

$$(6.1) \quad \ker_t R' = \{w \in C_t^\infty(\mathbb{R}^q) \mid \exists l \in C_t^\infty(\mathbb{R}^q) \text{ for all } \tau \in \text{dom } w \cap \text{dom } l : R(\frac{d}{d\tau})w(\tau) = S(\frac{d}{d\tau})l(\tau)\}.$$

*Proof.* By Theorem 2.1, there exists some unimodular  $U(D) \in \mathcal{M}[D]^{g \times g}$  such that

$$U(D)R(D) = \begin{bmatrix} R'(D) \\ R''(D) \end{bmatrix}, \quad U(D)S(D) = \begin{bmatrix} 0 \\ S''(D) \end{bmatrix},$$

where  $R'(D) \in \mathcal{M}[D]^{g' \times q}$ ,  $R''(D) \in \mathcal{M}[D]^{g'' \times q}$ ,  $S''(D) \in \mathcal{M}[D]^{g'' \times s}$ , and  $\text{rk}_{\mathcal{M}[D]} S''(D) = g''$ .

Applying Theorem 2.1 again, there exist  $\mathcal{M}[D]$ -unimodular matrices  $U(D)$  and  $V(D)$  of sizes  $g''$  and  $s$ , and  $r(D) \in \mathcal{M}[D]$  such that

$$S''(D) = U(D)^{-1} \left[ \begin{array}{c|c} I_{g''-1} & 0 \\ \hline 0 & r(D) \end{array} \right] 0_{g'' \times (q-g'')} V(D)^{-1}.$$

Choose  $\mathbb{T}$  as the discrete set of the union of all zeros and poles of the meromorphic coefficients in all entries of  $U(D), V(D), U(D)^{-1}, V(D)^{-1}, r(D)$ . Let  $\mathbb{I}$  be an open interval with  $\mathbb{I} \subset \mathbb{R} \setminus \mathbb{T}$  and  $t \in \mathbb{I}$ .

Then, for all  $\tau \in \mathbb{I}$ ,

$$R(\frac{d}{d\tau})w(\tau) = S(\frac{d}{d\tau})l(\tau) \iff \begin{bmatrix} R'(\frac{d}{d\tau}) & 0 \\ R''(\frac{d}{d\tau}) & S''(\frac{d}{d\tau}) \end{bmatrix} \begin{bmatrix} w(\tau) \\ l(\tau) \end{bmatrix}.$$

Hence the inclusion “ $\supset$ ” in (6.1) is obvious. To show “ $\subset$ ” in (6.1), let  $w \in \ker_t R'$  for  $t \in \mathbb{I}$ . Let  $\tilde{l}_{g''} \in C^\infty(\mathbb{I}, \mathbb{R})$  denote the solution of

$$r(\frac{d}{d\tau})\tilde{l}_{g''}(\tau) = (U(\frac{d}{d\tau})S''(\frac{d}{d\tau})w(\tau))_{g''} \quad \text{on } \mathbb{I}.$$

This solution exists; see, for example, [34, Chap. IV]. Setting

$$l := V[0, \dots, 0, \tilde{l}_{g''}]^T$$

yields

$$U(\frac{d}{d\tau})R''(\frac{d}{d\tau})w(\frac{d}{d\tau}) = \left[ \begin{array}{c|c} I_{g''-1} & 0 \\ \hline 0 & r(D) \end{array} \right] 0_{g'' \times (q-g'')} V(\frac{d}{d\tau})^{-1}l(\tau),$$

which is equivalent to  $R(\frac{d}{d\tau})w(\tau) = S(\frac{d}{d\tau})l(\tau)$ . This completes the proof of the theorem.  $\square$

As an “inverse” to Proposition 3.4, we show that any image representation of a behavior may be written as a kernel representation.

**COROLLARY 6.2.** *Let  $M(D) \in \mathcal{M}[D]^{q \times m}$ . Then there exist  $g \in \mathbb{N}$  and  $R'(D) \in \mathcal{M}[D]^{g \times q}$  such that*

$$\text{im}_t M = \ker_t R' \quad \text{for almost all } t \in \mathbb{R}.$$

*Proof.* Apply Theorem 6.1 to  $[R(D), S(D)] = [I_q, M(D)]$ .  $\square$

**Acknowledgment.** We are indebted to Jan C. Willems and anonymous referees for their constructive criticisms of an earlier version of this paper.

While the results of the present paper were reviewed, Eve Zerz [42] wrote, based on our findings, a much more elegant algebraic approach where she partially achieves the present results with shorter proofs and also characterizes behaviors included in each other.

#### REFERENCES

- [1] K. E. BRENNAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations*, Elsevier Science, North-Holland, New York, 1996.
- [2] R. BYERS, P. KUNKEL, AND V. MEHRMANN, *Regularization of linear descriptor systems with variable coefficients*, SIAM J. Control Optim., 35 (1997), pp. 117–133.
- [3] K. ÇAMLIBEL, M. N. BELUR, A. J. SASANE, AND J. C. WILLEMS, *On a class of time varying behaviors*, in Proceedings of the 15th Symposium on the Mathematical Theory of Networks and Systems, Notre Dame, IN, 2002, Session FA5.
- [4] S. L. CAMPBELL, *Linearization of DAE's along trajectories*, Z. Angew. Math. Phys., 46 (1995), pp. 70–84.
- [5] S. L. CAMPBELL, N. K. NICHOLS, AND W. J. TERRELL, *Duality, observability, and controllability for linear time-varying descriptor systems*, Circuits Systems Signal Process., 10 (1991), pp. 455–470.
- [6] P. M. COHN, *Free Rings and Their Relations*, Academic Press, London, New York, 1971.
- [7] M. FLIESS, *Some basic structural properties of generalized linear systems*, Systems Control Lett., 15 (1990), pp. 391–396.
- [8] S. FRÖHLER AND U. OBERST, *Continuous time-varying linear systems*, Systems Control Lett., 35 (1998), pp. 97–110.
- [9] E. GRIEPENTROG AND R. MÄRZ, *Differential-Algebraic Equations and Their Numerical Treatment*, Teubner Verlag, Leipzig, 1986.
- [10] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Springer-Verlag, Berlin, 1996.
- [11] D. HINRICHSSEN AND D. PRÄTZEL-WOLTERS, *Solution modules and system equivalence*, Internat. J. Control, 32 (1980), pp. 777–802.
- [12] M. HOU AND P. C. MÜLLER, *LQ and Tracking Control of Descriptor Systems with Application to Constrained Manipulator*, Technical report, Sicherheitstechnische Regelungs- und Meßtechnik, Universität Wuppertal, Germany, 1994.
- [13] A. ILCHMANN, Y. KUANG, M. KUIJPER, AND C. ZHANG, *Continuous time-varying scalar systems—a behavioural approach*, in Proceedings of the 3rd Third Asian Control Conference, Shanghai, 2000, pp. 429–433.
- [14] A. ILCHMANN, I. NÜRNBERGER, AND W. SCHMALE, *Time-varying polynomial matrix systems*, Internat. J. Control, 40 (1984), pp. 329–362.
- [15] E. W. KAMEN, *Representation and realization of operational differential equations with time-varying coefficients*, J. Franklin Inst., 301 (1976), pp. 559–570.
- [16] P. KUNKEL AND V. MEHRMANN, *A new look at pencils of matrix valued functions*, Linear Algebra Appl., 212/213 (1993), pp. 215–248.
- [17] P. KUNKEL AND V. MEHRMANN, *Local and global invariants of linear differential-algebraic equations and their relation*, Electron. Trans. Numer. Anal., 4 (1996), pp. 138–157.

- [18] P. KUNKEL AND V. MEHRMANN, *Analysis of over- and underdetermined nonlinear differential-algebraic systems with application to nonlinear control problems*, Math. Control Signals Systems, 14 (2001), pp. 233–256.
- [19] P. KUNKEL AND V. MEHRMANN, *Differential-Algebraic Equations—Analysis and Numerical Solution*, EMS Publishing House, Zürich, Switzerland, 2006, to appear.
- [20] P. KUNKEL, V. MEHRMANN, AND W. RATH, *Analysis and numerical solution of control problems in descriptor form*, Math. Control Signals Systems, 14 (2001), pp. 29–61.
- [21] K. S. MILLER, *Linear Differential Equations*, Routledge & Kegan Paul, London, 1964.
- [22] U. OBERST, *Multidimensional constant linear systems*, Acta Appl. Math., 20 (1990), pp. 1–175.
- [23] O. ORE, *Theory of non-commutative polynomials*, Ann. of Math. (2), 34 (1933), pp. 480–508.
- [24] J. W. POLDERMAN AND J. C. WILLEMS, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Springer-Verlag, New York, 1998.
- [25] P. J. RABIER AND W. C. RHEINOLDT, *Nonholonomic Motion of Rigid Mechanical Systems from a DAE Viewpoint*, SIAM, Philadelphia, 2000.
- [26] W. RATH, *Feedback Design and Regularization for Linear Descriptor Systems with Variable Coefficients*, Ph.D. thesis, TU Chemnitz-Zwickau, 1996; Shaker Verlag, Aachen, 1997.
- [27] H. H. ROSENBROCK, *State-Space and Multivariable Theory*, John Wiley, New York, 1970.
- [28] H. H. ROSENBROCK AND C. STOREY, *Mathematics of Dynamical Systems*, Nelson, London, 1970.
- [29] J. RUDOLPH, *Duality in time-varying linear systems: A module theoretic approach*, Linear Algebra Appl., 245 (1996), pp. 83–106.
- [30] W. J. RUGH, *Linear Systems Theory*, Prentice-Hall, Upper Saddle River, NJ, 1996.
- [31] M. SATO, *Theory of hyperfunctions I*, J. Fac. Sci. Univ. Tokyo Sect. I, 8 (1959), pp. 139–193.
- [32] M. SATO, *Theory of hyperfunctions II*, J. Fac. Sci. Univ. Tokyo Sect. I, 8 (1960), pp. 387–436.
- [33] E. D. SONTAG, *Mathematical Control Theory*, 2nd ed., Springer-Verlag, New York, 1998.
- [34] W. WALTER, *Ordinary Differential Equations*, Springer-Verlag, New York, 1998.
- [35] J. C. WILLEMS, *System theoretic models for the analysis of physical systems*, Ricerche Automat., 10 (1981), pp. 71–106.
- [36] J. C. WILLEMS, *From time series to linear system, I: Finite dimensional linear time invariant systems*, Automatica J. IFAC, 22 (1986), pp. 561–580.
- [37] J. C. WILLEMS, *From time series to linear system, II: Exact modelling*, Automatica J. IFAC, 22 (1986), pp. 675–694.
- [38] J. C. WILLEMS, *From time series to linear system, III: Approximate modelling*, Automatica J. IFAC, 23 (1987), pp. 87–115.
- [39] W. A. WOLOVICH, *Linear Multivariable Systems*, Springer-Verlag, New York, 1974.
- [40] R. YLINEN, *On the Algebraic Theory of Linear Differential and Difference Systems with Time-Varying or Operator Coefficients*, Systems Theory Laboratory Report, B23, Helsinki University, Helsinki, 1975.
- [41] R. YLINEN, *An algebraic theory for analysis and synthesis of time-varying linear systems*, Acta Polytech. Scand. Math. Comput. Sci. Ser., 32 (1980), pp. 1–61.
- [42] E. ZERZ, *An algebraic analysis approach to linear time-varying systems*, IMA J. Math. Control Inform., to appear.