

A BEHAVIORAL APPROACH TO TIME-VARYING LINEAR SYSTEMS. PART 2: DESCRIPTOR SYSTEMS*

ACHIM ILCHMANN[†] AND VOLKER MEHRMANN[‡]

Abstract. In the sequel to [A. Ilchmann and V. Mehrmann, *SIAM J. Control Optim.*, 44 (2005), pp. 1725–1747], we discuss a behavioral approach for linear, time-varying, differential algebraic (descriptor) systems with real analytic coefficients. The analysis is “almost global” in the sense that the analysis is not restricted to an interval $I \subset \mathbb{R}$ but is allowed for the “time axis” $\mathbb{R} \setminus \mathbb{T}$, where \mathbb{T} is a discrete set of critical points, at which the solution may exhibit a finite escape time. Controllable, observable, and autonomous behavior for linear time-varying descriptor systems is characterized.

Key words. time-varying linear systems, descriptor systems, behavioral approach, controllability, observability, autonomous system

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

DOI. 10.1137/040609021

1. Introduction. In [12], a behavioral approach was developed for linear time-varying systems with real analytic coefficients. In this paper, this approach will be studied for the specific case of linear time-varying descriptor systems described by differential-algebraic equations of the form

$$(1.1) \quad \begin{aligned} E(t) \dot{x}(t) &= A(t)x(t) + B(t)u(t), \\ y(t) &= C(t)x(t) + F(t)u(t), \end{aligned}$$

with real analytic matrices $A \in \mathcal{A}^{l \times n}$, $B \in \mathcal{A}^{l \times m}$, $C \in \mathcal{A}^{p \times n}$, $F \in \mathcal{A}^{p \times m}$, where $E \in \mathcal{A}^{l \times n}$ is allowed to be singular in the sense that $\text{rk } E(t) < \min\{l, n\}$ for some $t \in \mathbb{R}$. Throughout this paper, the nomenclature as introduced and listed in [12] will be used.

As in [12], we make use of the skew-polynomial rings $\mathcal{A}[D]$ and $\mathcal{M}[D]$ (see [6, 13]) of differential polynomials with coefficients in \mathcal{A} , \mathcal{M} , respectively, and indeterminate D representing the differential operator $\frac{d}{dt}$, and the multiplication rule $Df = fD + \dot{f}$. The algebraic object

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

acts on \mathcal{C}^∞ -functions w via

$$R\left(\frac{d}{dt}\right)w(t) = \sum_{i=0}^n R_i(t)w^{(i)}(t).$$

In this notation, time-varying descriptor systems (1.1) may be rewritten as

$$(1.2) \quad R\left(\frac{d}{dt}\right)w = 0,$$

*Received by the editors May 26, 2004; accepted for publication (in revised form) March 8, 2005; published electronically November 23, 2005.

<http://www.siam.org/journals/sicon/44-5/60902.html>

[†]Institut für Mathematik, Technische Universität Ilmenau, Weimarer Straße 25, D-98693 Ilmenau, Germany (ilchmann@mathematik.tu-ilmenau.de).

[‡]Institut für Mathematik, MA 4-5, Technische Universität Berlin, Straße des 17. Juni 136, D-10623 Berlin, Germany (mehrman@math.tu-berlin.de). This research was supported by DFG Research Center MATHEON Mathematics for key technologies in Berlin.

where

$$R(D) = \begin{bmatrix} ED - A & -B & 0 \\ -C & -F & I_p \end{bmatrix}, \quad \text{and} \quad w = [x^T, u^T, y^T]^T.$$

Systems of differential algebraic equations (often called descriptor systems) play an important role in modelling and control of multibody systems, electric circuits, or coupled systems of partial differential equations; see [1, 9].

The analysis of the behavior of (1.1) has to cope with three essential difficulties. First, the solutions of time-varying systems may exhibit critical points, i.e., a finite escape time. Second, descriptor systems behave quite differently from classical state space systems (i.e., $E = I_n$ in (1.1)). For state space systems, the function $u(\cdot)$ can be considered as an input function free to choose, and initial conditions can be arbitrary. This is in general not true for descriptor systems (1.1), since descriptor systems may contain algebraic constraints, which restrict the solutions, the set of possible inputs, and also the initial values to some manifold. Third, some of the constraints that arise (the hidden constraints) are not explicit and thus it is not clear how to choose the underlying spaces for the descriptor variables x, u, y . Finally, the analytic property of the solution or behavior is local, which is in contrast to the global algebraic properties of $R(D)$. These difficulties are illustrated by the following example.

Example 1.

- (i) The scalar differential equations $t\dot{x} = -x$, $t^2\dot{x} = -x$, $t\dot{x} = x$, have local solutions $t \mapsto t^{-1}$, $e^{1/t}$, t , respectively. Hence at $t = 0$ the solution might be rational with a pole, not even analytic, or does not have any pole, respectively.
- (ii) The variables $x_1, \dots, x_4, u_1, u_2$ of the descriptor system (1.1) with

$$E = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$C = [0 \ 0 \ 0 \ 1], \quad F = 0_{1 \times 2}$$

satisfy the equivalent description

$$u_2 = 0, \quad \dot{x}_2 = x_1, \quad y = x_4, \quad \dot{x}_3 = x_2 + u_1.$$

Thus, u_2 is constrained to be 0 and cannot be freely chosen, as it could in the case of state space systems. The variables x_1 and x_4 can be viewed as input or state variables; the system description does not determine this.

Note also that if we choose the input u_1 as a step function, then either x_3 is chosen as input as well to compensate the delta distribution in u_1 (since $x_1 = \dot{x}_2 = \dot{x}_3 - \dot{u}_1$), or we may have to enlarge the solution space to allow that x_1 is a delta distribution. But even if we do so, then we have the problem that x_1 is not observable from the output y , which means that internally the system has impulsive parts of the solution, which are not observed. For many types of practical systems, such as, for example, mechanical systems, this would be a disaster: impulses in the solution cannot be tolerated. \square

Example 1 indicates that the behavioral viewpoint, where state-, output-, and input-variables are not distinguished, seems the appropriate concept for the analysis of descriptor systems. The behavioral approach has been introduced by Willems [25, 26, 27, 28]; see also the textbook [21] and [12] for a general presentation.

Motivated by Example 1 and as introduced in [12], we study, for $R(D) \in \mathcal{M}[D]^{q \times q}$, local solutions of $R(\frac{d}{dt})w = 0$ belonging to

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

as the *almost global behavior* given by the kernel representation

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

The *local behavior*

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

becomes a *real vector space* if endowed, for $w_1, w_2 \in \ker_t R$, with addition

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

and obvious scalar multiplication.

We also have to consider those points of the real axis, where the local solution is no longer extendable.

DEFINITION 1.1. *Consider the descriptor system (1.2). The set of critical points, where the solution is not defined, is given by*

$$(1.3) \quad \mathbb{T}_R^{\text{crit}} := \left\{ t' \in \mathbb{R} \left| \begin{array}{l} \text{there exists, for some } \varepsilon > 0, \text{ a } \mathcal{C}^\infty \text{ function} \\ w : (t' - \varepsilon, t') \rightarrow \mathbb{R}^q \text{ or } w : (t', t' + \varepsilon) \rightarrow \mathbb{R}^q \\ \text{which solves (1.2) and cannot be extended to} \\ (t' - \varepsilon, t'] \text{ or } [t', t' + \varepsilon), \text{ respectively.} \end{array} \right. \right\}$$

Note that for the three differential equations in Example 1(i), the sets of critical points are $\{0\}$, $\{0\}$, \emptyset , respectively.

Since E in (1.1) is real analytic, it follows that for almost all $\hat{t} \in \mathbb{R}$, the rank of the matrix $E(\hat{t}) \in \mathbb{R}^{l \times m}$ is equal to $\text{rk}_A E$, and the set of critical points is a discrete set. It is an open problem to characterize the set of critical points. However, we will determine discrete sets which include all critical points.

We define the appropriate behavior, i.e., the solution space, of (1.2) on the time-axis $\mathbb{R} \setminus \mathbb{T}$, where \mathbb{T} is discrete and includes the set of critical points of (1.2). Controllability and observability are defined in terms of trajectories (descriptor variables), which is a conceptual generalization of controllability and observability for state space systems. For these systems in [5] controllability and observability have been studied in terms of derivative arrays. In [4] a first behavior-like approach for analytic coefficients has been discussed. A more general approach that allows for larger classes of coefficients and that can also be implemented numerically has been introduced in [16]. In [11] a first approach in the spirit of the present paper was presented for scalar systems. A completely different approach results from the study of differential-algebraic equations; see [1, 8, 17]. A general solvability theory for nonsquare linear time-varying systems was first given in [15] and analyzed for control problems in a behavioral context in [4, 18, 22]; see also [16] for the general nonlinear case. In these papers, however, mainly the concept of regularization has been discussed, i.e., the problem of finding appropriate feedback that makes the system regular and also decreases the index. Here we consider controllability and observability in the behavioral context.

This paper is organized as follows. In section 2, we define critical points and follow the concepts of [15, 22] by deriving condensed forms for time-varying descriptor systems (1.2) to determine sets covering the critical points. In section 3, controllability is defined, algebraically characterized, and related to the well-known concepts of controllability. In section 4, we apply results from [12] and briefly discuss autonomous behavior and observability for descriptor systems.

2. Condensed forms. In this section, condensed forms with respect to state and input transformations are studied for time-varying descriptor systems (1.2). The condensed form allows us to classify the solution sets and to identify the constraint manifolds for the variables. These forms are akin to the forms derived in [4, 17, 18].

The construction of the condensed forms is based on the computation of analytic singular value decompositions that were introduced in [2] for analytic matrices and that are also valid for real analytic matrices. This result states that for a matrix function $A \in \mathcal{A}^{l \times n}$ there exist real orthogonal matrix functions $U \in \mathcal{A}^{l \times l}, V \in \mathcal{A}^{n \times n}$ and a diagonal matrix $\Sigma \in \mathcal{A}^{r \times r}$, where $r = \text{rk}A$, such that

$$U^T A V = \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix}.$$

It should be noted, though, that, in contrast to the usual singular value decomposition for matrices, the diagonal elements of $\Sigma(t)$, in general, cannot be chosen positive or in descending order. In this way, however, the analytic singular value decomposition is not uniquely defined. Essentially, there is freedom to perform orthogonal transformations in the spaces associated with multiple singular values. This freedom can be removed by choosing minimal variation curves or by always choosing the analytic singular value decomposition to be closest to a reference point [3, 20].

THEOREM 2.1. *Consider a time-varying descriptor system of the form (1.2) with*

$$R(D) = \begin{bmatrix} ED - A & -B & 0 \\ -C & -F & I_p \end{bmatrix} \in \mathcal{A}[D]^{(l+p) \times (n+m+p)}.$$

(i) *There exist orthogonal matrices $U_1 \in \mathcal{A}^{l \times l}, V_1 \in \mathcal{A}^{n \times n}$ so that*

$$(2.1) \quad \begin{bmatrix} U_1 & 0 \\ 0 & I_p \end{bmatrix} R(D) \begin{bmatrix} V_1 & 0 \\ 0 & I_{m+p} \end{bmatrix}$$

corresponds to the descriptor system

$$(2.2) \quad \begin{aligned} \Sigma_d \dot{x}_1 &= A_{11} x_1 + A_{12} x_2 + A_{13} x_3 + B_1 u, \\ 0 &= A_{21} x_1 + \Sigma_a x_2 + B_2 u, \\ 0 &= A_{31} x_1 + B_3 u, \\ 0 &= A_{41} x_1, \\ 0 &= 0_{l-\nu}, \\ y &= C_1 x_1 + C_2 x_2 + C_3 x_3 + F u, \end{aligned}$$

where $\Sigma_d \in \mathcal{A}^{d \times d}, \Sigma_a \in \mathcal{A}^{a \times a}$ are diagonal and invertible over \mathcal{M} with $d = \text{rk}E$, and $B_3 \in \mathcal{A}^{\gamma \times m}, A_{41} \in \mathcal{A}^{f \times d}$ with full row rank; i.e., $\gamma = \text{rk}B_3$, $f = \text{rk}A_{41}$, and $\nu = d + a + \gamma + f$. All matrices are real analytic and of conforming formats.

- (ii) *There exist orthogonal matrices $U_2 \in \mathcal{A}^{l \times l}, V_2 \in \mathcal{A}^{n \times n}, W \in \mathcal{A}^{p \times p}, Z \in \mathcal{A}^{m \times m}$ so that*

$$(2.3) \quad \begin{bmatrix} U_2 & 0 \\ 0 & W \end{bmatrix} R(D) \begin{bmatrix} V_2 & 0 & 0 \\ 0 & Z & 0 \\ 0 & 0 & I_p \end{bmatrix}$$

corresponds to the following descriptor system in condensed form:

$$(2.4) \quad \begin{aligned} \Sigma_d \dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + A_{13}x_3 + A_{14}x_4 + A_{15}x_5 + B_{11}u_1 + B_{12}u_2, \\ 0 &= A_{21}x_1 + \Sigma_a x_2 + B_{21}u_1 + B_{22}u_2, \\ 0 &= A_{31}x_1 + \Sigma_\gamma u_1, \\ 0 &= \Sigma_f x_5, \\ 0 &= 0_{l-\nu}, \\ y_1 &= C_{11}x_1 + C_{12}x_2 + \Sigma_\omega x_3 + C_{15}x_5 + F_{11}u_1 + F_{12}u_2, \\ y_2 &= C_{21}x_1 + C_{22}x_2 + C_{25}x_5 + F_{21}u_1 + F_{22}u_2, \end{aligned}$$

where $\Sigma_d, \Sigma_a, \Sigma_\gamma, \Sigma_f, \Sigma_\omega$ are diagonal matrices that are invertible over \mathcal{M} and have sizes d, a, γ, f, ω , respectively. Furthermore, $\nu = d + a + \gamma + f$ and all matrices are real analytic and of conforming formats.

- (iii) *There exist matrices $U \in \mathcal{A}^{(l-p) \times (l-p)}, V \in \mathcal{M}^{n \times n}$ invertible over $\mathcal{M}, X \in \mathcal{M}^{p \times (l-p)}, W \in \mathcal{A}^{p \times p}$ orthogonal, $Z \in \mathcal{A}^{m \times m}$ orthogonal, a scalar function $\sigma \in \mathcal{A}$, and a permutation matrix $P \in \mathcal{A}^{(n+m) \times (n+m)}$ so that*

$$(2.5) \quad \tilde{R}(D) := \begin{bmatrix} U & 0 \\ X & W \end{bmatrix} R(D) \begin{bmatrix} P & 0 \\ 0 & I_p \end{bmatrix}$$

$$= \left[\begin{array}{ccc|ccc|cc} \sigma DI_d - \tilde{A}_{11} & -\tilde{A}_{13} & -\tilde{A}_{14} & -\tilde{B}_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Sigma_a^{-1} \tilde{B}_{22} & I_a & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I_f & 0 & 0 & 0 \\ \Sigma_\gamma^{-1} \tilde{A}_{31} & 0 & 0 & 0 & 0 & 0 & I_\gamma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & -\Sigma_\omega & 0 & -\sigma^{-1} \tilde{F}_{12} & 0 & 0 & 0 & I_\omega & 0 \\ -\sigma^{-1} \tilde{C}_{21} & 0 & 0 & -\sigma^{-1} \tilde{F}_{22} & 0 & 0 & 0 & 0 & I_p \end{array} \right]$$

corresponds to the meromorphic descriptor system in standard condensed

form

$$\begin{aligned}
 \sigma I_d \dot{x}_1 &= \tilde{A}_{11} x_1 + [\tilde{A}_{13}, \tilde{A}_{14}, \tilde{B}_{12}] \begin{bmatrix} x_3 \\ x_4 \\ u_2 \end{bmatrix}, \\
 \begin{bmatrix} x_2 \\ x_5 \\ u_1 \end{bmatrix} &= \begin{bmatrix} 0 & 0 & 0 & -\Sigma_a^{-1} \tilde{B}_{22} \\ 0 & 0 & 0 & 0 \\ -\Sigma_\gamma^{-1} \tilde{A}_{31} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_3 \\ x_4 \\ u_2 \end{bmatrix}, \\
 \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} 0 \\ \sigma^{-1} \tilde{C}_{21} \end{bmatrix} x_1 + \begin{bmatrix} \Sigma_\omega & 0 & \sigma^{-1} \tilde{F}_{12} \\ 0 & 0 & \sigma^{-1} \tilde{F}_{22} \end{bmatrix} \begin{bmatrix} x_3 \\ x_4 \\ u_2 \end{bmatrix},
 \end{aligned}
 \tag{2.6}$$

where

$$\sigma(t) := \det \Sigma_d(t) \det \Sigma_a(t) \det \Sigma_\gamma(t) \det \Sigma_f(t) \quad \text{for all } t \in \mathbb{R},
 \tag{2.7}$$

and all matrices are real analytic and of conforming formats. The integers d, a, γ, ω, f are invariants of (1.2).

Proof. The proof is constructive using a sequence of real analytic singular value decompositions. When multiplying with D , we will always use the product rule without saying so.

(i) Consider the first equation of (1.1) and choose orthogonal matrices $\tilde{U} \in \mathcal{A}^{l \times l}$, $\tilde{V} \in \mathcal{A}^{n \times n}$ so that

$$[\tilde{R}(D), -\tilde{B}] = \tilde{U}[ED - A, -B] \begin{bmatrix} \tilde{V} & 0 \\ 0 & I_m \end{bmatrix} = \left[\begin{bmatrix} \Sigma_d & 0 \\ 0 & 0 \end{bmatrix} D - \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}, \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix} \right],$$

where $\Sigma_d \in \mathcal{A}^{d \times d}$ with $d = \text{rk} E$ is diagonal.

Next, choose orthogonal matrices $\tilde{U} \in \mathcal{A}^{(l-d) \times (l-d)}$, $\tilde{V} \in \mathcal{A}^{(n-d) \times (n-d)}$ so that

$$\begin{aligned}
 [\bar{R}(D), -\bar{B}] &= \begin{bmatrix} I_d & 0 \\ 0 & \tilde{U} \end{bmatrix} [\tilde{R}(D), -\tilde{B}] \begin{bmatrix} I_d & 0 & 0 \\ 0 & \tilde{V} & 0 \\ 0 & 0 & I_m \end{bmatrix} \\
 &= \left[\begin{bmatrix} \Sigma_d & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} D - \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} & \bar{A}_{13} \\ \bar{A}_{21} & \Sigma_a & 0 \\ \bar{A}_{31} & 0 & 0 \end{bmatrix}, \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \\ \bar{B}_3 \end{bmatrix} \right],
 \end{aligned}$$

where $\Sigma_a \in \mathcal{A}^{a \times a}$ is diagonal and invertible over \mathcal{M} . Finally, choose an orthogonal $\hat{U} \in \mathcal{A}^{(l-d-a) \times (l-d-a)}$, so that $\begin{bmatrix} I_{d+a} & 0 \\ 0 & \hat{U} \end{bmatrix} [\bar{R}(D), -\bar{B}]$ has the form (2.2) with $B_3 \in \mathcal{A}^{\gamma \times m}$, $\gamma = \text{rk} B_3 = \text{rk} \bar{B}_3$, and $A_{41} \in \mathcal{A}^{f \times d}$, $f = \text{rk} A_{41}$. Performing all the transformations also on C and partitioning analogously shows (2.2).

(ii) We apply the so-called *index reduction process* as introduced in [18] to (2.4): Fix f variables of x_1 , corresponding to some f linearly independent columns of A_{41} , i.e., choose a unitary matrix $Q \in \mathcal{A}^{d \times d}$ such that $A_{41}Q = [A_{41}^\alpha, A_{41}^\beta]$ with $A_{41}^\alpha \in \mathcal{A}^{f \times f}$ is invertible over \mathcal{M} . Then

$$0 = A_{41}x_1 = A_{41}^\alpha x_1^\alpha + A_{41}^\beta x_1^\beta, \quad \begin{bmatrix} x_1^\alpha \\ x_1^\beta \end{bmatrix} := Qx_1,$$

and so

$$\dot{x}_1^\alpha = -(A_{41}^\alpha)^{-1}A_{41}^\beta \dot{x}_1^\beta - \frac{d}{dt} \left((A_{41}^\alpha)^{-1}A_{41}^\beta \right) x_1^\beta.$$

Inserting \dot{x}_1^α into the differential equation of (2.2) leaves $d - f$ differential equations. Note that we may have introduced meromorphic functions by the inverse of A_{41}^α and its derivative. A multiplication from the left with a real analytic function yields a description in the form (1.1); however, the d differential equations have been reduced to $d - f$ differential equations and we may apply part (i) again. This index reduction process stops after finitely many iterations, and we arrive at the following condensed form:

$$(2.8) \quad \begin{aligned} \Sigma_d \dot{x}_1 &= \hat{A}_{11}x_1 + \hat{A}_{12}x_2 + \hat{A}_{13}x_3 + \hat{A}_{14}x_4 + \hat{B}_1u, \\ 0 &= \hat{A}_{21}x_1 + \Sigma_a x_2 + \hat{B}_2u, \\ 0 &= \hat{A}_{31}x_1 + \hat{B}_3u, \\ 0 &= \phantom{\hat{A}_{31}x_1} + \Sigma_f x_4, \\ 0 &= 0_{l-\nu}, \\ y &= \hat{C}_1x_1 + \hat{C}_2x_2 + \hat{C}_3x_3 + \hat{C}_4x_4 + Fu, \end{aligned}$$

where $\Sigma_d, \Sigma_a, \Sigma_f$ are diagonal matrices, invertible over \mathcal{M} , and of sizes d, a, f , respectively, and $\hat{B}_3 \in \mathcal{A}^{\gamma \times m}$ has full row rank over \mathcal{A} .

As a final step we perform an analytic singular value decomposition of \hat{C}_3, \hat{B}_3 , respectively, and derive (2.4).

(iii) Using the fact that the fourth equation in (2.4) implies that $x_5 \equiv 0$, which can be extended even at points where Σ_f is singular, we can eliminate all terms invoking x_5 from all the other equations. This corresponds to multiplying (2.4) from the left first by

$$\begin{bmatrix} I_d & -A_{12}\Sigma_a^{-1} & -[B_{11} - A_{12}\Sigma_a^{-1}B_{21}]\Sigma_\gamma^{-1} & 0 & 0 & 0 & 0 \\ 0 & I_a & -B_{21}\Sigma_\gamma^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & I_\gamma & -A_{15}\Sigma_f^{-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & I_f & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{l-\nu} & 0 & 0 \\ 0 & -C_{12}\Sigma_a^{-1} & -F_{11}\Sigma_\gamma^{-1} & 0 & 0 & I_\omega & 0 \\ 0 & -C_{22}\Sigma_a^{-1} & -F_{21}\Sigma_\gamma^{-1} & 0 & 0 & 0 & I_{p-\omega} \end{bmatrix}$$

and then by

$$\begin{bmatrix} \sigma \Sigma_d^{-1} & 0 \\ 0 & I_{l-d+p} \end{bmatrix}$$

and from the right by

$$\begin{bmatrix} I_d & 0 & 0 & 0 \\ -[A_{21} - B_{21}\Sigma_\gamma^{-1}A_{31}]\Sigma_a^{-1} & I_a & 0 & 0 \\ -C_{11}\Sigma_\omega^{-1} & 0 & I_\gamma & 0 \\ 0 & 0 & 0 & I_{(n-\nu+f+m+p) \times (n-\nu+f+m+p)} \end{bmatrix},$$

yielding the transformed system

$$\begin{aligned}
 (2.9) \quad & \sigma \dot{x}_1 = \tilde{A}_{11}x_1 + \tilde{A}_{13}x_3 + \tilde{A}_{14}x_4 + \tilde{B}_{12}u_2, \\
 & 0 = \Sigma_a x_2 + \sigma^{-1} \tilde{B}_{22}u_2, \\
 & 0 = \tilde{A}_{31}x_1 + \Sigma_\gamma u_1, \\
 & 0 = \Sigma_f x_5, \\
 & 0 = 0_{l-v}, \\
 & y_1 = \Sigma_\omega x_3 + \sigma^{-1} \tilde{F}_{12}u_2, \\
 & y_2 = \sigma^{-1} \tilde{C}_{21}x_1 + \sigma^{-1} \tilde{F}_{22}u_2,
 \end{aligned}$$

where all matrices are real analytic. This proves (2.6). \square

Remark 1.

- (i) If the descriptor system (1.2) is time-invariant, then all transformations in Theorem 2.1 may be chosen as constant matrices and $\sigma = 1$. In this case, the condensed forms in Theorem 2.1 are well known; see, for example [3].
- (ii) To derive (2.2), only an orthogonal transformation on the variables x in (2.1) has been applied. To derive (2.4), the transformations on the variables x and u have not been mixed.

To derive (2.6), we have used nonsingular transformations on x and orthogonal transformations on u . If we allow further linear combinations (which for classical systems where y , x , and u are fixed a priori as outputs, states, and controls, respectively, correspond to state feedback or output feedback), then we can simplify (2.6) further by removing blocks such as \tilde{A}_{31} or by introducing almost everywhere invertible diagonal blocks in diagonal positions of the transformed matrices E or A . Note that the transformation of derivative feedback is not an equivalence transformation, because under derivative feedback the characteristic quantities d, a, γ, f, w are not invariants and hence the properties of the system may be altered by this transformation completely; see [18].

- (iii) The description (2.6) is not quite of the form (1.1), since the coefficients of x_1 and u_2 in y_1 and y_2 may have poles at the zeros of σ .
- (iv) An immediate consequence of (2.6) is that the variables in x_1 represent couplings between algebraic equations and differential equations that are not influenced by u_1 . Systems where such couplings between differential equations and algebraic equations occur are typically called *high index systems*. For a detailed discussion of different index concepts see [1, 8, 17].
- (v) The transformation leading to (2.6) does not invoke any differentiation of u . Hence, if the variables denoted by u are classified as inputs a priori, then no extra differentiability conditions for these variables arise; see [4, 18].
- (vi) The condensed forms (2.1), (2.4), and (2.6) allow us to detect candidates for critical points, given by

$$(2.10) \quad \mathbb{T}_R^{\text{crit}} \subset \mathbb{T}_R := \{t' \in \mathbb{R} \mid \sigma(t') = 0\}.$$

As can be seen from the first system considered in Example 1(i), the set $\mathbb{T}_R^{\text{crit}} = \emptyset$ can be a strict subset of $\mathbb{T}_R = \{0\}$.

- (vii) The reader may wonder why we display equations of the form $0 = 0$ in the condensed form. These arise typically when automatic modelling systems are used and describe redundant equations in the system. \square

To characterize controllability we will need the following staircase form which generalizes the staircase form of Van Dooren [24] to systems with analytic coefficient matrices.

LEMMA 2.2. *For real analytic matrices $A \in \mathcal{A}^{n \times n}, B \in \mathcal{A}^{n \times m}$ there exist orthogonal matrices $P \in \mathcal{A}^{n \times n}$ and $Q \in \mathcal{A}^{m \times m}$ so that*

$$(2.11) \quad P [DI_n - A, -B] \begin{bmatrix} P^T & 0 \\ 0 & Q \end{bmatrix} \\ = \left[\begin{array}{cccc|ccc} DI_{n_1} - A_{11} & -A_{12} & \cdots & -A_{1,s-1} & -A_{1,s} & -B_1 & 0 \\ -[\hat{A}_{21}, 0] & \ddots & & \vdots & \vdots & 0 & 0 \\ & \ddots & \ddots & -A_{s-2,s-1} & \vdots & \vdots & \vdots \\ \hline & & -[\hat{A}_{s-1,s-2}, 0] & DI_{n_{s-1}} - A_{s-1,s-1} & -A_{s-1,s} & 0 & 0 \\ 0 & \cdots & 0 & 0 & DI_{n_s} - A_{s,s} & 0 & 0 \end{array} \right],$$

where $n_1 \geq n_2 \geq \dots \geq n_{s-1} \geq n_s \geq 0$, $n_{s-1} > 0$, and $B_1 \in \mathcal{A}^{n_1 \times n_1}$ and $\hat{A}_{i,i-1} \in \mathcal{A}^{n_i \times n_i}$ are invertible over \mathcal{M} for $i = 1, \dots, s-1$.

Proof. A constructive proof is given by the following generalization of the so-called *Staircase Algorithm* to systems with real analytic coefficients. Whenever we use Σ in the following, it denotes a diagonal matrix.

Step 0. Choose orthogonal $U_B \in \mathcal{A}^{n \times n}, V_B \in \mathcal{A}^{m \times m}$ so that

$$B = U_B^T \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix} V_B \in \mathcal{A}^{n \times m} \quad \text{with invertible } \Sigma_B \in \mathcal{A}^{n_1 \times n_1},$$

and set

$$A_0 := U_B A U_B^T + \dot{U}_B U_B^T = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{with } A_{21} \in \mathcal{A}^{(n-n_1) \times n_1}, \\ B_0 := U_B B V_B = \begin{bmatrix} \Sigma_B & 0 \\ 0 & 0 \end{bmatrix}.$$

Then, using the product rule, we have

$$U_B [DI_n - A, -B] \begin{bmatrix} U_B^T & 0 \\ 0 & V \end{bmatrix} = [DI_n - A_0, -B_0].$$

Step 1. If $n_1 < n$ and $A_{21} \neq 0$, then choose orthogonal $U_{21} \in \mathcal{A}^{(n-n_1) \times (n-n_1)}, V_{21} \in \mathcal{A}^{n_1 \times n_1}$ so that

$$A_{21} = U_{21} \begin{bmatrix} \Sigma_{21} & 0 \\ 0 & 0 \end{bmatrix} V_{21}^T \in \mathcal{A}^{(n-n_1) \times n_1} \quad \text{with invertible } \Sigma_{21} \in \mathcal{A}^{n_2 \times n_2},$$

and set

$$P_1 := \begin{bmatrix} V_{21}^T & 0 \\ 0 & U_{21}^T \end{bmatrix}, \\ A_1 := P_1 A_0 P_1^T + \dot{P}_1 P_1^T = \left[\begin{array}{cc|c} * & * & * \\ \Sigma_{21} & 0 & * \\ 0 & 0 & * \end{array} \right] + \left[\begin{array}{c|c} \dot{V}_{21}^T V_{21} & 0 \\ 0 & \dot{U}_{21}^T U_{21} \end{array} \right], \\ B_1 := V_{21}^T \Sigma_B, \\ \tilde{B}_1 := \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{A}^{n \times n}.$$

Using the product rule, for some $\tilde{A}_{32} \in \mathcal{A}^{(n-n_1-n_2) \times n_2}$ this gives

$$P_1 [DI_n - A_0, -B_0] \begin{bmatrix} P_1^T & 0 \\ 0 & I_m \end{bmatrix} = [DI_n - A_1, -\tilde{B}_1] \\ = \left[DI_n - \left[\begin{array}{c|cc} * & * & * \\ \hline [\Sigma_{21}, 0] & * & * \\ 0 & \tilde{A}_{32} & * \end{array} \right], - \left[\begin{array}{c|c} B_1 & 0 \\ \hline 0 & 0 \\ 0 & 0 \end{array} \right] \right].$$

Step 2. If $n_1 + n_2 < n$ and $\tilde{A}_{32} \neq 0$, then choose orthogonal matrices $U_{32} \in \mathcal{A}^{(n-n_1-n_2) \times (n-n_1-n_2)}$, $V_{32}^T \in \mathcal{A}^{n_2 \times n_2}$ so that

$$\tilde{A}_{32} = U_{32} \begin{bmatrix} \Sigma_{32} & 0 \\ 0 & 0 \end{bmatrix} V_{32}^T \in \mathcal{A}^{(n-n_1-n_2) \times n_2} \quad \text{with invertible } \Sigma_{32} \in \mathcal{A}^{n_3 \times n_3},$$

and set

$$P_2 := \text{diag} \{ I_{n_1}, V_{32}^T, U_{32}^T \}, \\ \hat{A}_{21} := V_{32}^T \Sigma_{21}, \\ A_2 := P_2 A_1 P_2^T + \dot{P}_2 P_2^T \\ = \left[\begin{array}{c|cc} * & * & * \\ \hline V_{32}^T [\Sigma_{21}, 0] & * & * \\ 0 & U_{32}^T \tilde{A}_{32} V_{32} & * \end{array} \right] + \left[\begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & \dot{V}_{32}^T V_{32} & 0 \\ 0 & 0 & \dot{U}_{32}^T U_{32} \end{array} \right] \\ = \left[\begin{array}{cc|cc|c} * & * & * & * & * \\ \hline \hat{A}_{21} & 0 & * & * & * \\ \hline 0 & 0 & \Sigma_{32} & 0 & * \\ 0 & 0 & 0 & 0 & * \end{array} \right].$$

Then, for some $\tilde{A}_{43} \in \mathcal{A}^{(n-n_1-n_2-n_3) \times n_3}$,

$$P_2 P_1 [DI_n - A_0, -B_0] \begin{bmatrix} P_1^T P_2^T & 0 \\ 0 & I_m \end{bmatrix} \\ = [DI_n - A_2, -\tilde{B}_1] \\ = \left[DI_n - \left[\begin{array}{cc|cc|c} * & * & * & * & * \\ \hline \hat{A}_{21} & 0 & * & * & * \\ \hline 0 & 0 & \Sigma_{32} & 0 & * \\ 0 & 0 & 0 & 0 & \tilde{A}_{43} & * \end{array} \right], - \left[\begin{array}{c|c} B_1 & 0 \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array} \right] \right].$$

Step 3. In the remainder of the proof we proceed analogously as in Step 2 and terminate after finitely many steps with the form (2.11). This completes the proof. \square

Example 2. As an example consider the model of a two-dimensional, three-link constrained mobile manipulator studied in [10]; see also [12]. This model leads, after linearization along a trajectory, to a system of the form

$$(2.12) \quad \begin{aligned} M_0(t) \ddot{z}(t) + D_0(t) \dot{z}(t) + K_0(t) z(t) &= S_0 u(t) + F_0^T \mu(t), \\ F_0 z(t) &= 0, \end{aligned}$$

where $M_0, D_0, K_0 \in \mathcal{C}^\omega(\mathbb{I}, \mathbb{R}^{3 \times 3})$ and $S_0, F_0^T \in \mathbb{R}^{3 \times 2}$ with S_0 having full 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 (1.1) with $F \equiv 0$,

$$(2.13) \quad E(t) = \begin{bmatrix} I_3 & 0 & 0 \\ 0 & M_0(t) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A(t) = \begin{bmatrix} 0 & I_3 & 0 \\ -K_0(t) & -D_0(t) & F_0^T \\ F_0 & 0 & 0 \end{bmatrix}, \quad B \equiv \begin{bmatrix} 0 \\ S_0 \\ 0 \end{bmatrix},$$

and the specification of C is left open for the time being.

The critical points of (2.13) include those values of t where the mass matrix $M_0(t)$ changes rank. This happens, for example, when two arms of the manipulator are in one straight line.

Without loss of generality (by using an appropriate permutation of the basis), 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 if we partition

$$\begin{aligned} -K_0 &= \begin{bmatrix} K_{11}(t) & K_{12}(t) \\ K_{21}(t) & K_{22}(t) \end{bmatrix}, \quad M_0 = \begin{bmatrix} M_{11}(t) & M_{12}(t) \\ M_{21}(t) & M_{22}(t) \end{bmatrix}, \\ -D_0 &= \begin{bmatrix} D_{11}(t) & D_{12}(t) \\ D_{21}(t) & D_{22}(t) \end{bmatrix}, \quad S_0 = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix}, \end{aligned}$$

with $K_{11}(t), M_{11}(t), D_{11}(t), S_1 \in \mathbb{R}^{2 \times 2}$ and all other formats accordingly, then system (2.13) may be written as

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

Since F_1 is constant and nonsingular, we obtain $x_1 = 0$ and $\dot{x}_1 = 0$. Inserting this

and changing the order of equations and blocks leads to

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

We can repeat the reduction process once more by using that $x_3 = 0$, and hence $\dot{x}_3 = 0$, which gives a system

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

Since the mass matrix M_0 is positive definite almost everywhere, we can eliminate the block M_{12} and obtain the system

$$\begin{aligned} & \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 0 & 0 \\ 0 & M_{22}(t) & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{bmatrix} \dot{x}_2 \\ \dot{x}_4 \\ \dot{x}_3 \\ \dot{x}_5 \\ \dot{x}_1 \end{bmatrix} \\ (2.14) \quad &= \left[\begin{array}{cc|ccc} 0 & 1 & 0 & 0 & 0 \\ K_{22}(t) & D_{22}(t) & 0 & 0 & 0 \\ \hline \tilde{K}_{12}(t) & \tilde{D}_{12}(t) & 0 & F_1^T & 0 \\ 0 & 0 & 0 & 0 & F_1 \\ 0 & 0 & I_2 & 0 & 0 \end{array} \right] \begin{bmatrix} x_2 \\ x_4 \\ x_3 \\ x_5 \\ x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ S_2 \\ \tilde{S}_1 \\ 0 \\ 0 \end{bmatrix} u. \end{aligned}$$

This system is essentially (apart from diagonal matrices Σ) in the condensed form (2.2), with

$$\Sigma_d = \begin{bmatrix} 1 & 0 \\ 0 & M_{22}(t) \end{bmatrix}, \quad \Sigma_a = \begin{bmatrix} 0 & F_1^T & 0 \\ 0 & 0 & F_1 \\ I_2 & 0 & 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} \tilde{S}_1 \\ 0 \\ 0 \end{bmatrix}.$$

It is then obvious how the more refined forms (2.4) and (2.6) can be determined. □

3. Controllability. In this section we discuss the concept of controllability for descriptor systems of the form (1.2). Recall that (local) controllability for general systems of the form $R(\frac{d}{dt})w = 0$, where $R(D) \in \mathcal{M}[D]^{g \times q}$, is introduced in [12, Def. 3.1] and discussed in [12, Rem. 3.2].

Remark 2. For descriptor systems with constant coefficients, several different controllability concepts have been introduced; see [3, 7, 19].

- (i) System (1.1) with constant coefficients is called
 - R-controllable* iff $\text{rk}[\lambda E - A, B]$ is full for all $\lambda \in \mathbb{C}$,
 - I-controllable* iff $\text{rk}[E, AS_\infty, B]$ is full,
 - where S_∞ spans the kernel of E ,

strongly controllable iff the system is *R-controllable* and *I-controllable*. We stress that these algebraic characterizations are sometimes misleading in the literature, since it is sometimes assumed that the rank of $[E, B]$ is full and sometimes not.

It follows that if system (1.1) is square and time-invariant (thus, in particular, $l = n$), then system (1.1) is *I-controllable* if and only if $n - (d + a + \gamma + f) = 0$. The constants a, d, f, γ are defined in Theorem 2.1(ii). *I-controllability* is related to regularization and index reduction; i.e., in particular it is needed to avoid impulsive solutions in the case of nondifferentiable input functions. In our framework this concept is not relevant.

- (ii) If the descriptor system (1.2) is time-varying, then [12, Def. 3.1] is new; see [5, 22, 18] for a discussion of different controllability concepts for time-varying descriptor systems.
- (iii) For time-invariant state-space systems, i.e. (1.1) with $E = I_n$, the algebraic conditions can be checked numerically via the Staircase Algorithm of [24]. In a similar fashion Lemma 2.2 may be used to check controllability for time-varying systems.
- (iv) For time-invariant systems (1.2), [12, Def. 3.1] corresponds to the concept of *R-controllability*. This follows from Theorem 3.1 below. □

Theorem 2.1 and Lemma 2.2 put us in a position to characterize controllability of time-varying descriptor systems (1.2).

THEOREM 3.1. *Consider a time-varying descriptor system (1.2) and assume that $R(D)$ has full row rank over $\mathcal{M}[D]$. Consider the condensed form (2.6) and σ as defined in (2.7). Set, for notational convenience,*

$$G(t) := \tilde{A}_{11}(t), \quad S(t) := [\tilde{A}_{13}(t), \tilde{A}_{14}(t), \tilde{B}_{12}(t)], \quad v(t) := [x_3(t)^T, x_4(t)^T, u_2(t)^T]^T.$$

Then the following conditions are equivalent:

- (i) (1.2) is locally controllable almost everywhere.
- (ii) $R(D)$ is right invertible over $\mathcal{M}[D]$.
- (iii) (2.3), respectively (2.4), is locally controllable almost everywhere.
- (iv) $\hat{R}(D) := [\sigma DI_d - G, S]$ is right invertible over $\mathcal{M}[D]$.
- (v) In the staircase form (2.11) of the pair $[DI_d - G, S]$, the lower block is not present; i.e., $n_s = 0$.
- (vi) There exists a discrete set $\mathbb{T} \subset \mathbb{R}$ such that for every

$$\begin{bmatrix} x_1^0 \\ v^0 \end{bmatrix}, \begin{bmatrix} x_1^1 \\ v^1 \end{bmatrix} \in \ker_t \hat{R}$$

and for every open interval $\mathbb{I} \subset \mathbb{R} \setminus \mathbb{T}$ and all $t_0 \in \mathbb{I}$, there exists $t_1 > t_0$, $t_1 \in \mathbb{I}$, and $[x_1^T, v^T]^T \in \ker_t \hat{R}$, such that

$$\begin{bmatrix} x_1(t) \\ v(t) \end{bmatrix} = \begin{cases} \begin{bmatrix} x_1^0(t) \\ v^0(t) \end{bmatrix} & \text{if } t \in (-\infty, t_0] \cap \mathbb{R} \setminus \mathbb{T}, \\ \begin{bmatrix} x_1^1(t) \\ v^1(t) \end{bmatrix} & \text{if } t \in [t_1, \infty) \cap \mathbb{R} \setminus \mathbb{T}. \end{cases}$$

Proof.

(i) \Leftrightarrow (ii): This is proved in [12, Prop. 3.6].

(ii) \Leftrightarrow (iii): The equivalence of local controllability almost everywhere of (1.2) and (2.4), respectively (1.1) and (2.6), follows from (2.3) by invoking orthogonality of U_2, V_2, W, Z .

(ii) \Leftrightarrow (iv): By (2.5), there exist invertible matrices $\tilde{U} \in \mathcal{M}^{(l+p) \times (l+p)}$, $\tilde{V} \in \mathcal{M}^{(n+m+p) \times (n+m+p)}$ so that (1.1) is related to (2.6) in the form (1.2) by the transformation

$$(3.1) \quad \tilde{U} \begin{bmatrix} E D - A & -B & 0 \\ -C & -F & I_p \end{bmatrix} \tilde{V} = \begin{bmatrix} \sigma D I_d - \tilde{A}_{11} & 0 & -\tilde{A}_{13} & -\tilde{A}_{14} & 0 & 0 & -\tilde{B}_{12} & 0 & 0 \\ 0 & -\Sigma_a & 0 & 0 & 0 & 0 & -\tilde{B}_{22} & 0 & 0 \\ -\tilde{A}_{31} & 0 & 0 & 0 & 0 & -\Sigma_\gamma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\Sigma_f & 0 & 0 & 0 & 0 \\ 0_{(l-\nu) \times d} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\sigma \Sigma_\omega & 0 & 0 & 0 & -\tilde{F}_{12} & \sigma I_\omega & 0 \\ -\tilde{C}_{21} & 0 & 0 & 0 & 0 & 0 & -\tilde{F}_{22} & 0 & \sigma I_{p-\omega} \end{bmatrix}.$$

The right-hand side is right invertible if and only if $l - \nu = l - d - a - \gamma - f = 0$ (which is a consequence of the full row rank assumption) and $[\sigma D I_d - G, S]$ is right invertible over $\mathcal{M}[D]$.

(iv) \Leftrightarrow (v): By Lemma 2.2 there exist orthogonal matrices P and Q so that $P[\sigma D I_d - G, S][\begin{smallmatrix} P^T & 0 \\ 0 & Q \end{smallmatrix}]$ is of the staircase form (2.11). Note that σ does not affect the staircase form. Now the equivalence (iv) \Leftrightarrow (v) follows immediately since B_1 and $\hat{A}_{i,i-1}$ are invertible over \mathcal{M} for $i = 1, \dots, s - 1$.

(ii) \Leftrightarrow (vi): This equivalence follows readily from [12, Def. 3.1] and from (3.1), since the set of zeros and poles of the coefficients of \tilde{U} and \tilde{V} is a discrete set. \square

Note that the assumption that (1.2) has full row rank over $\mathcal{M}[D]$ is equivalent to $l - d - a - \gamma - f = 0$ in (2.6).

Note further that the characterization in Theorem 3.1(ii) does not require a reinterpretation of variables as is done in [4]. Moreover, in contrast to the case of controllability of state space systems, here $u_1(\cdot)$ in (1.1) is not a “free input” variable.

For standard time-invariant state space systems (i.e., $E = I_n$), the right invertibility of $R(D)$ in Theorem 3.1 is derived differently in [21, Thm. 5.2.10].

Remark 3. For time-varying systems (1.1) with $E = I_n$, i.e., state space systems, it is well known that controllability of the system yields that it can be controlled in an arbitrary short time. The interval \mathbb{I} in [12, Def. 3.1] can be replaced by any arbitrary short open interval $\hat{\mathbb{I}} \subset \mathbb{I}$. This also holds true for descriptor systems (1.2), since $\hat{R}(D)$ in Theorem 3.1(iv) can be viewed locally as a state space system, namely, at those $t \in \mathbb{R}$ where $\sigma(t) \neq 0$; note that the zeros of σ are a discrete set. An alternative and constructive proof is given in [12, Thm. 3.3] for general systems of the form $R(\frac{d}{dt})w = 0$. \square

Example 3.

- (i) $R(D) = [t^2D + 1, 1]$ has right inverse $[0, 1]^T$, and hence, by Theorem 3.1, $R(\frac{d}{dt})w = 0$ is controllable.
- (ii) Revisit the linearized model (2.13) of the three-link constrained mobile manipulator. In Example 2 it is shown that (2.13) is equivalent to (2.11). Rewriting (2.11) in the form (1.2) and invoking that M_{22} is invertible over \mathcal{M} and F_1 is nonsingular, it is easy to see that the corresponding $R(D)$ is right invertible. Therefore, by Theorem 3.1, the linearized model (2.13) is controllable. \square

4. Observability and autonomous behavior. In [12, sect. 4], the concept of autonomous behavior $\ker_t^{\text{aut}} R$ also has been introduced, and it has been shown that the behavior of a system (1.2) (and hence also of (1.1)) can be decomposed into the direct sum of a controllable and an autonomous behavior. It also immediately follows from the results in [12] that an autonomous behavior $\ker_t^{\text{aut}} R$ of the system (1.2) is invariant under all transformations (2.1), (2.2), (2.4), (2.10). Loosely speaking, an autonomous behavior consists of those solutions which are uniquely determined if they are known on an arbitrarily small open interval. For systems (1.2) we have to cope with the problem of finite escape time.

Example 4. Consider a time-varying state space system (1.2) with $E = I_n$. By [14] there exists $T \in \mathcal{A}^{n \times n}$ invertible over \mathcal{A} so that the coordinate transformation $z := T^{-1}x$ converts (1.1) into

$$(4.1) \quad \begin{aligned} \frac{d}{dt}z_1(t) &= A_{11}(t)z_1(t) + A_{12}(t)z_2(t) + B_1(t)u(t), \\ \frac{d}{dt}z_2(t) &= A_{22}(t)z_2(t), \\ y(t) &= C_1(t)z_1(t) + C_2(t)z_2(t) + F(t)u(t), \end{aligned}$$

with all matrices real analytic of conforming formats, and controllable subsystem $\frac{d}{dt}z_1(t) = A_{11}(t)z_1(t) + B_1(t)u(t)$. Since (4.1) is a state space system, finite escape time does not occur and the controllable and autonomous subspaces can be described globally. Set

$$\hat{R}(D) := \begin{bmatrix} DI - A_{11} & -A_{12} & -B_1 & 0 \\ 0 & DI - A_{22} & 0 & 0 \\ -C_1 & -C_2 & -F & -I_p \end{bmatrix}.$$

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

$$\ker_t^{\text{contr}} \hat{R} = \left\{ w = [z_1^T, z_2^T, u^T, y^T]^T \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{(n+m+p)}) \mid \hat{R}(\frac{d}{dt})w = 0 \wedge z_2 = 0 \right\}$$

and

$$\ker_t^{\text{aut}} \hat{R} = \left\{ w = [z_1^T, z_2^T, u^T, y^T]^T \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^{(n+m+p)}) \mid \begin{aligned} &\hat{R}(\frac{d}{dt})w = 0, z_1 = 0, \\ &u = 0, \dot{z}_2 = A_{22} z_2 \end{aligned} \right\}$$

is an autonomous behavior, and, hence, in the original coordinates, we have

$$\ker_t^{\text{contr}} R = \begin{bmatrix} T(t) & 0 \\ 0 & I_{m+p} \end{bmatrix} \ker_t^{\text{aut}} R \oplus \begin{bmatrix} T(t) & 0 \\ 0 & I_{m+p} \end{bmatrix} \ker_t^{\text{contr}} \hat{R} \quad \forall t \in \mathbb{R}.$$

Remark 4. Consider a time-varying descriptor system (1.1) in the condensed form (2.6). If (2.6) were controllable, then $\ker^{\text{aut}} R = \{0\}$ would be the only autonomous

behavior of (2.6). To see this, note that x_3, x_4, u_2 are free to choose and hence cannot be a nonzero component of an autonomous behavior. Furthermore, since $[\sigma DI_d - G, S]$ is controllable by Theorem 3.1(iii), it follows that x_1 is uniquely (modulo initial condition) determined by x_3, x_4, u_2 , and hence also not a nontrivial component of an autonomous behavior. Finally, (2.6) yields that the remaining components x_2, x_5, u_1, y_1, y_2 are uniquely determined by x_3, x_4, u_2, x_1 . This shows $\ker^{\text{aut}} R = \{0\}$.

If (2.6) is not controllable but has a nontrivial uncontrollable subspace, then there exists $\ker^{\text{aut}} R \neq \{0\}$ which is determined by the uncontrollable subspace as for state space systems; see Example 4. \square

In [12] it has also been discussed how one behavior can be observed from another. We refer to this paper for the definition of adjoints and observable behavior which generalize well-known concepts of observability, such as for time-varying state space systems (see, for example, [23]) and time-varying Rosenbrock systems (see [13]). It has also been shown that local observability and local controllability are dual concepts.

An application of [12, Thms. 5.5 and 5.6] to descriptor systems (1.2) yields the following result.

THEOREM 4.1. *Consider a descriptor system (1.2) with $R(D) = [R_1(D), R_2(D)]$ partitioned as*

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

Then the following are equivalent:

- (i) The trajectory x is locally observable from (u, y) almost everywhere.
- (ii) $R_1(D)$ is left invertible over $\mathcal{M}[D]$.
- (iii) The matrix

$$(4.2) \quad \begin{bmatrix} \sigma DI_d - \tilde{A}_{11} & -\tilde{A}_{13} & -\tilde{A}_{14} \\ -\tilde{A}_{31} & 0 & 0 \\ 0 & -\Sigma_\omega & 0 \\ -\sigma^{-1}\tilde{C}_{21} & 0 & 0 \end{bmatrix}$$

is left invertible over $\mathcal{M}[D]$, where the matrices in (4.2) are from the condensed form (2.4).

Proof. The equivalence (i) \leftrightarrow (ii) follows from [12, Thms. 5.5 and 5.6]. To see (ii) \leftrightarrow (iii), note that left invertibility of $R_2(D)$ is equivalent to

$$\begin{bmatrix} U & 0 \\ X & W \end{bmatrix} R_2(D)V = \begin{bmatrix} \sigma DI_d - \tilde{A}_{11} & 0 & -\tilde{A}_{13} & -\tilde{A}_{14} & 0 \\ 0 & -\Sigma_a & 0 & 0 & 0 \\ -\tilde{A}_{31} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\Sigma_f \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\Sigma_\omega & 0 & 0 \\ -\sigma^{-1}\tilde{C}_{21} & 0 & 0 & 0 & 0 \end{bmatrix}$$

being left invertible, where U, X, V, W are specified in Theorem 2.1(iii). Since $\begin{bmatrix} U & 0 \\ X & W \end{bmatrix}$ and V are invertible over \mathcal{M} , the latter holds true if and only if (4.2) is left invertible. This completes the proof. \square

Example 5. Consider again the linearized model (2.13) of the three-link constrained mobile manipulator. Suppose that the positions can be measured, corre-

sponding to the additional equation

$$(4.3) \quad y = \begin{bmatrix} 0 & 0 & I_2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}.$$

In Example 2 we have shown that $x_1 = 0$ and $x_3 = 0$ and thus $\dot{x}_1 = 0$ and $\dot{x}_3 = 0$, and permuting the variables accordingly to (2.14), we obtain

$$(4.4) \quad y = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_2 \\ x_4 \\ x_3 \\ x_5 \\ x_1 \end{bmatrix}.$$

Hence by Theorem 4.1, x is observable from (u, y) with respect to the system (2.13), (4.3) or, equivalently, system (2.11), (4.4) if and only if

$$(4.5) \quad \left[\begin{array}{cc|ccc} D-1 & 0 & 0 & 0 & 0 \\ -K_{22} & M_{22}D - D_{22} & 0 & 0 & 0 \\ \hline -\tilde{K}_{12} & -\tilde{D}_{12} & 0 & -F_1^T & 0 \\ 0 & 0 & 0 & 0 & F_1 \\ 0 & 0 & -I_2 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

is left invertible over $\mathcal{M}[D]$. Since F_1 is invertible over \mathcal{M} , (4.5) is left invertible if and only if $\begin{bmatrix} D-1 & 0 \\ -K_{22} & M_{22}D - D_{22} \end{bmatrix}$ is invertible over $\mathcal{M}[D]$. Summarizing, x is observable from (u, v) almost everywhere if and only if K_{22} is invertible over \mathcal{M} . \square

5. Conclusion. We have introduced a general behavioral approach to linear descriptor systems with real analytic coefficients. We have characterized autonomous, controllable, and observable behavior and have generalized results on time-varying ordinary differential equations and on time-invariant linear algebraic-differential equations. The results have been illustrated by several examples, which demonstrates that the approach also helps in understanding practical problems such as constrained multibody systems.

REFERENCES

- [1] K. E. BRENNAN, S. L. CAMPBELL, AND L. R. PETZOLD, *Numerical Solution of Initial-Value Problems in Differential Algebraic Equations*, 2nd ed., Elsevier-North Holland, New York, 1996.
- [2] A. BUNSE-GERSTNER, R. BYERS, V. MEHRMANN, AND N. K. NICHOLS, *Numerical computation of an analytic singular value decomposition of a matrix valued function*, Numer. Math., 60 (1991), pp. 1–40.
- [3] A. BUNSE-GERSTNER, R. BYERS, V. MEHRMANN, AND N. K. NICHOLS, *Feedback design for regularizing descriptor systems*, Linear Algebra Appl., 299 (1999), pp. 119–151.
- [4] R. BYERS, P. KUNKEL, AND V. MEHRMANN, *Regularization of linear descriptor systems with variable coefficients*, SIAM J. Control Optim., 35 (1997), pp. 117–133.

- [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*, London Math. Soc. Monogr. (NS) 2, Academic Press, London, New York, 1971.
- [7] L. DAI, *Singular Control Systems*, Springer-Verlag, Berlin, 1989.
- [8] E. GRIEPENTROG AND R. MÄRZ, *Differential-Algebraic Equations and Their Numerical Treatment*, Teubner-Texte Math. 88, Teubner Verlag, Leipzig, Germany, 1986.
- [9] E. HAIRER AND G. WANNER, *Solving Ordinary Differential Equations. II: Stiff and Differential-Algebraic Problems*, Springer Ser. Comput. Math. 14, 2nd ed., Springer-Verlag, Berlin, 1996.
- [10] M. HOU AND P. C. MÜLLER, *LQ and Tracking Control of Descriptor Systems with Application to Constrained Manipulator*, Tech. report, Sicherheitstechnische Regelungs- und Meßtechnik, Universität Wuppertal, Wuppertal, Germany, 1994.
- [11] A. ILCHMANN, Y. KUANG, M. KUIJPER, AND C. ZHANG, *Continuous time-varying scalar systems—A behavioral approach*, in *Proceedings of the Third Asian Control Conference*, Shanghai, China, 2000, pp. 429–433.
- [12] A. ILCHMANN AND V. MEHRMANN, *A behavioral approach to time-varying linear systems. Part 1: General theory*, *SIAM J. Control Optim.*, 44 (2005), pp. 1725–1747.
- [13] A. ILCHMANN, I. NÜRNBERGER, AND W. SCHMALE, *Time-varying polynomial matrix systems*, *Internat. J. Control*, 40 (1984), pp. 329–362.
- [14] R. E. KALMAN, *Canonical structure of linear dynamical systems*, *Proc. Natl. Acad. Sci. USA*, 48 (1962), pp. 596–600.
- [15] P. KUNKEL AND V. MEHRMANN, *A new look at pencils of matrix valued functions*, *Linear Algebra Appl.*, 212/213 (1994), pp. 215–248.
- [16] 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.
- [17] P. KUNKEL AND V. MEHRMANN, *Differential-Algebraic Equations—Analysis and Numerical Solution*, EMS Publishing House, Zürich, Switzerland, to appear.
- [18] 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.
- [19] V. MEHRMANN, *The Autonomous Linear Quadratic Control Problem: Theory and Numerical Algorithms*, Lecture Notes in Control and Inform. Sci. 163, Springer-Verlag, Berlin, 1991.
- [20] V. MEHRMANN AND W. RATH, *Numerical methods for the computation of analytic singular value decompositions*, *Electron Trans. Numer. Anal.*, 1 (1993), pp. 72–88.
- [21] J. W. POLDERMAN AND J. C. WILLEMS, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Texts Appl. Math. 26, Springer-Verlag, New York, 1998.
- [22] W. RATH, *Feedback Design and Regularization for Linear Descriptor Systems with Variable Coefficients*, Ph.D. thesis, TU Chemnitz-Zwickau, Shaker Verlag, Aachen, Germany, 1997.
- [23] W. J. RUGH, *Linear System Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [24] P. VAN DOOREN, *The computation of Kronecker's canonical form of a singular pencil*, *Linear Algebra Appl.*, 27 (1979), pp. 103–121.
- [25] J. C. WILLEMS, *System theoretic models for the analysis of physical systems*, *Ricerche Automat.*, 10 (1979), pp. 71–106.
- [26] J. C. WILLEMS, *From time series to linear system. I. Finite-dimensional linear time invariant systems*, *Automatica J. IFAC*, 22 (1986), pp. 561–580.
- [27] J. C. WILLEMS, *From time series to linear system. II. Exact modelling*, *Automatica J. IFAC*, 22 (1986), pp. 675–694.
- [28] J. C. WILLEMS, *From time series to linear system. III. Approximate modelling*, *Automatica J. IFAC*, 23 (1987), pp. 87–115.