A BEHAVIOURAL APPROACH TO TIME-VARYING LINEAR SYSTEMS. PART 1: GENERAL THEORY

ACHIM ILCHMANN* AND VOLKER MEHRMANN†

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

Keywords: Time-varying linear systems, behavioural approach, controllability, observability, autonomous system, adjoint system, latent variables

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

Nomenclature

\( I_d := \text{diag}[1, \ldots, 1] \in \mathbb{R}^{d \times d} \)

\( 0_d := (0, \ldots, 0)^T \in \mathbb{R}^d \)

\( \mathcal{A} \) the ring of real analytic functions \( f : \mathbb{R} \to \mathbb{R} \)

\( \mathcal{M} \) the field of real meromorphic functions

\( \mathcal{A}[D], \mathcal{M}[D] \) the skew polynomial ring of differential polynomials with coefficients in \( \mathcal{A}, \mathcal{M} \) resp., 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 \( f : M \to \mathbb{R}^q \), \( M \subset \mathbb{R} \) an open set, \( N \in \mathbb{N} \cup \{ \infty \} \)

\( \mathcal{C}^\infty(\mathbb{I}, \mathbb{R}^q) \) the real vector space of real analytic functions \( f : \mathbb{I} \to \mathbb{R}^q \), \( \mathbb{I} \subset \mathbb{R} \) an open interval

\( \mathcal{C}^\infty_{\mathbb{I}}(\mathbb{R}^q) := \{ w \in \mathcal{C}^\infty(\mathbb{R} \setminus \mathcal{T}, \mathbb{R}^q) | \mathcal{T} \subset \mathbb{R} \text{ discrete} \} \)

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

\( \text{im}, M := \{ w \in \mathcal{C}^\infty_t(\mathbb{R}^q) | \exists l \in \mathcal{C}^\infty_t(\mathbb{R}^m) \forall \tau \in \text{dom } w \cap \text{dom } l : w(\tau) = M(\frac{d}{dt})l(\tau) \} \)

\( t \in \mathbb{R}, M(D) \in \mathcal{M}[D]^{q \times m} \)

\( \text{im } M := \{ w \in \mathcal{C}^\infty_{\text{pw}}(\mathbb{R}^q) | \exists l \in \mathcal{C}^\infty_{\text{pw}}(\mathbb{R}^m) \text{ for a.a. } \tau \in \text{dom } w \cap \text{dom } l : w(\tau) = M(\frac{d}{dt})l(\tau) \}, \quad M(D) \in \mathcal{M}[D]^{q \times m} \)

\( \ker R := \{ w \in \mathcal{C}^\infty_t(\mathbb{R}^q) | R(\frac{d}{dt})w(\tau) = 0 \text{ for all } \tau \in \text{dom } w \} , \quad t \in \mathbb{R}, \quad R(D) \in \mathcal{M}[D]^{q \times q}, \)

\( \ker R := \{ w \in \mathcal{C}^\infty_{\text{pw}}(\mathbb{R}^q) | R(\frac{d}{dt})w(\tau) = 0 \text{ for a.a. } \tau \in \mathbb{R} \} , \quad R(D) \in \mathcal{M}[D]^{q \times q}, \)

\( \text{dom } w \) the domain of a function \( w \)

1. Introduction.

*Institut für Mathematik, Technische Universität Ilmenau, Weimarer Straße 25, 98693 Ilmenau, FRG. achim.ilchmann@tu-ilmenau.de

†Institut für Mathematik, MA 4-5, TU Berlin, Straße des 17. Juni 136, D-10623 Berlin, FRG. This research was supported by DFG Research Center MATHEON Mathematics for key technologies in Berlin. mehrmann@math.tu-berlin.de
1.1. An algebraic approach and solution spaces. The aim of the present paper is to develop a behavioural approach to linear time-varying systems described by differential-algebraic equations of the form

\[ R(D)w = 0, \]  

(1.1)

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 \( I \subset \mathbb{R} \), this is omitted.

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

\[ Df = fD + \dot{f}. \]  

(1.2)

This is a consequence of assuming the associative rule \((Df)g = D(fg)\) for all differentiable functions \( f, g \) which yields 
\[ (Df)g = \frac{df}{dt} \cdot g + f \cdot \frac{dg}{dt} = (\frac{df}{dt} + fD)(g). \]

The non-commutativity 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 \( \mathcal{M}[D] \) has been introduced in [14] to study linear time-varying systems. We are interested in the behaviour 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 Sub-section 1.3). To be more precise, we allow for the solution space

\[ C_{pw}^{\infty} (\mathbb{R}^q) = \{ w \in C^{\infty}(\mathbb{R} \setminus T, \mathbb{R}^q) \mid T \subset \mathbb{R} \text{ discrete} \} \]

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

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

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

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

\[ \ker R = \{ w \in C_{pw}^{\infty}(\mathbb{R}^q) \mid R(\frac{d}{dt})w(\tau) = 0 \text{ for almost all } \tau \in \mathbb{R} \}, \]

and the \textit{local behaviour}

\[ \ker, R = \{ w \in C_{t}^{\infty}(\mathbb{R}^q) \mid R(\frac{d}{dt})w(\tau) = 0 \text{ for all } \tau \in \text{dom } w \}, \quad t \in \mathbb{R}. \]
1.2. Examples of system classes. Our approach generalizes results on the following sub-classes of systems.

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

\[
\frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t),
\]

\[
y(t) = C(t)x(t) + F(t)u(t),
\]

with real analytic matrices \( A \in \mathbb{A}^{n \times n}, B \in \mathbb{A}^{n \times m}, C \in \mathbb{A}^{p \times n} \) and \( F \in \mathbb{A}^{p \times m} \), are well studied, see for example the standard monograph [29].

(b) Time-varying descriptor systems of the form

\[
E(t) \frac{d}{dt}x(t) = A(t)x(t) + B(t)u(t),
\]

\[
y(t) = C(t)x(t) + F(t)u(t),
\]

with \( A \in \mathbb{A}^{\ell \times \ell}, B \in \mathbb{A}^{\ell \times m}, C \in \mathbb{A}^{p \times \ell}, F \in \mathbb{A}^{p \times m} \), where \( E \in \mathbb{A}^{\ell \times \ell} \) 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 have been studied in terms of derivative arrays. In [2] a first behaviour like approach to systems (1.4) with analytic coefficients has been discussed. A more general approach that allows for larger classes of coefficients and that can be implemented also numerically has been introduced in [19] and generalized partially to the nonlinear case in [18].

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

\[
P(\frac{d}{dt})z(t) = Q(\frac{d}{dt})u(t),
\]

\[
y(t) = V(\frac{d}{dt})z(t) + W(\frac{d}{dt})u(t),
\]

where \( P(D), Q(D), V(D) \) and \( W(D) \) are matrices of size \( r \times r, r \times m, p \times r, p \times m \), 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 \( P(\frac{d}{dt})z = 0 \) on some interval \( I \subseteq \mathbb{R} \), then this solution can be analytically extended to the whole of \( \mathbb{R} \).

- For every \( u \in C^\infty(\mathbb{R}, \mathbb{R}^m) \) with bounded support to the left, there exist some \( z \in C^\infty(\mathbb{R}, \mathbb{R}^r) \) and \( y \in C^\infty(\mathbb{R}, \mathbb{R}^p) \) 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 \( \text{det} P(\cdot) \neq 0 \), were introduced in [26], and are well studied, see for example [11, 37].

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

\[
R(\frac{d}{dt})w(t) = 0, \quad R(D) \in \mathbb{R}[D]^{p \times q}
\]

have been introduced by Willems in [33]; see also [34, 35, 36] and the monograph [23].

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 a special case 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

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

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

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

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

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 $t \mapsto w(t) = e^{t^2t}$ 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 $\lim_{t \to 0^-} w(t) = 0$ and $\lim_{t \to 0^+} w(t) = \infty$. Hence,

$$\ker_{\mathcal{A}_{\mid I}} r\left(\frac{d}{dt}\right) = \{0\}.$$

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

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

(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 $r\left(\frac{d}{dt}\right)w = 0$ on $(0, \infty)$, $(-\infty, 0)$, respectively. For every interval $I \subset \mathbb{R}$ with $0 \notin I$, we have

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

However,

$$\ker_{\mathcal{A}} r\left(\frac{d}{dt}\right) = \{0\}.$$
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 neighbourhood and going into the complex sphere, as suggested by Ilchmann et al. [13], does not work.\(^*\)

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

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

satisfies $w \in \ker_{C^\infty} r(D)$, 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 modelling multi-body systems, electric circuits, or coupled systems of partial differential equations, see [1, 10]. We present an application which first shows that modelling does not necessarily lead to a state space system; secondly, it illustrates a simple system where the notion of input, output, and state is not a priori clear; and thirdly, 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.

---

\*We are indebted to the anonymous referee of an earlier version of the present paper for pointing out this example to us.
The Lagrangian equations of motion take the form

\[
M(\theta) \ddot{\theta} + D(\theta, \dot{\theta}) \dot{\theta} + K(\theta) = u + F^T(\theta)\mu, \quad \psi(\theta) = 0,
\]

(1.8)

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, the maps \( M : \mathbb{R}^3 \to \mathbb{R}^{3 \times 3}, \ D : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{3 \times 3}, \ K : \mathbb{R}^3 \to \mathbb{R}^3 \) 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 \( \psi : \mathbb{R}^3 \to \mathbb{R}^2 \), \( F = \frac{\partial \psi}{\partial \theta} \), 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 behaviour, 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 [23, Sec. 6.2]. Under suitable smoothness assumptions of the involved functions, it can be shown (see for example [24, p. 62]) that there exists a local (possibly global) solution \( \theta(\cdot) \) of (1.8) on some open interval \( I \). Linearizing along this trajectory [4] and rewriting the system in Cartesian coordinates yields a model of the form

\[
M_0(t) \ddot{z}(t) + D_0(t) \dot{z}(t) + K_0(t) \dot{z}(t) = S_0 u(t) + F^T_0(t) \mu
\]

(1.9)

where \( M_0, D_0, K_0 \in C^\infty(I, \mathbb{R}^{3 \times 3}) \) and \( S_0 \in \mathbb{R}^{3 \times 3}, F^T_0 \in \mathbb{R}^{3 \times 2} \) with \( S_0 \) having full row rank. Introducing the 8 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

\[
E(t) \frac{d}{dt} x(t) = A(t)x(t) + B(t)u(t), \quad y(t) = C(t)x(t),
\]

(1.9)

where

\[
E(t) := \begin{bmatrix} I_3 & 0 & 0 \\ 0 & M_0(t) & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A(t) := \begin{bmatrix} 0 & I_3 & 0 \\ -K_0(t) & -D_0(t) & F^T_0(t) \\ F_0(t) & 0 & 0 \end{bmatrix}, \quad B(t) := \begin{bmatrix} 0 \\ S_0 \\ 0 \end{bmatrix},
\]

and \( C(\cdot) \) denotes a matrix with appropriate format, see [12] for explicit data. Actually in this example \( F_0 \) is not depending 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, 38, 39]; 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 [28] 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, \((t \mapsto \exp \left(\frac{1}{2t^2}\right)) \in \ker \left( t^3 \frac{d}{dt} + 1 \right)\) 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 [30], then [8] considers systems of the form (1.1) respectively behaviour in the kernel representation \(\ker R\), where the coefficient matrices of the polynomial \(R(D)\) are defined over rational analytic functions

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

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 [21], 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 \(\mathcal{M}[D]\) is 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-Euclidian 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 \(\text{im } Q(\frac{d}{dt}) \subseteq \text{im } P(\frac{d}{dt})\) and, most importantly, that \(P(D)\) is a “full” operator, i.e. every local analytic solution of \(P(D)z = 0\) 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 is presented for scalar systems. A behavioural approach to a certain class of time-varying systems is presented in [3].

A completely different approach results from the study of differential-algebraic equations introduced in [1, 9]. A general solvability theory for non-square linear time-varying systems was first given in [16] and analysed for control problems in a behavioural context in [2, 19, 25], see also [18] for the general nonlinear case.

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

2. Behaviour. In this section we present the Teichmüller-Nakayama normal form for matrices over \(\mathcal{M}[D]\). This will be the main tool for analysing \(\ker R\). To this end we recall some results on matrices over the skew polynomial ring \(\mathcal{M}[D]\):

7
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, \( q_1 a = bq_2 \) for some \( a, b \in \mathcal{M}[D] \) for which \( q_1 \) and \( b \) \((q_2 \text{ and } a) \) 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 factorisation of the ring elements cannot be expected. However, Ore [22] 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, Ch. 8]. The proof is constructive using elementary matrices and 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 factorised into

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

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 non-zero, unique up to similarity, and of unique degree.

**Remark 2.2.** Let \( R(D) \in \mathcal{M}[D]^{g \times q} \) and consider the factorization (2.1).

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

\[
\forall w \in C^\infty_t(\mathbb{R}^g) : w \in \ker_t R \iff \begin{bmatrix} I_{\ell - 1} & 0_{\ell \times q} \\ r(D) & V^{-1} \end{bmatrix} \in \ker_t(\mathbb{R}^{g+q})\]

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

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

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

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

Furthermore,

\[
\dim \ker_t R = \begin{cases} 
\deg r(D) & \text{for a.a. } t \in \mathbb{R}, \text{ if } \text{rk } R(D) = q \\
\infty & \text{for all } t \in \mathbb{R}, \text{ if } \text{rk } R(D) < q.
\end{cases}
\]
The latter is a simple consequence of (2.1) and the fact that the set of \( t \) where \( r_i(\frac{t}{T}) \psi(t) = 0 \) 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 [32, Ch. 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 non-zero 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 non-empty 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, 18, 19, 17] 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 behaviour is non-trivial, i.e. \( \{ t \in \mathbb{R} \mid \ker_t R \neq \{ 0 \} \} \), is discrete. □

Remark 2.3. 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

\[
R(D) = U(D)^{-1} \begin{bmatrix}
\text{diag}\{r_1(D), \ldots, r_l(D)\} & 0_{\ell \times (g-l)} \\
0_{(g-l) \times \ell} & 0_{(g-l) \times (g-l)}
\end{bmatrix} V(D)^{-1}, \quad (2.2)
\]

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 non-zero monic polynomials with \( r_i|_{r_{i+1}} \), \( i = 1, \ldots, \ell - 1 \), where \( \ell = \text{rk}_{\mathbb{R}[D]} R(D) \) and \( r_1(D) = \psi(1(D)) \psi_{\ell-1}(D) \psi_0(\cdot) \equiv 1 \) and \( \psi_i(D) \) is the greatest common divisor of minors of order \( i \) of \( R(D) \); see for example [27, 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 \( w \in \mathcal{C}_N^N(\mathbb{R}^n) \) of \( R(D)w = 0 \), where \( N \) is sufficiently large depending on \( \deg R(D) \) and the degrees of the transformation matrices, can be continued to a global solution on \( \mathbb{R} \) and it 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 \( \ker_t R = \ker_t R \) for any \( t \in \mathbb{R} \) and it follows that \( \text{dim} \ker_t R = \sum_{i=1}^{\ell} \deg r_i(D) \) for all \( t \in \mathbb{R} \). □