Time-Varying Linear Systems: Relative Degree and Normal Form

Achim Ilchmann and Markus Mueller

***Abstract*—We define the relative degree of time-varying linear systems, show that it coincides with Isidori's and with Liberzon/Morse/Sontag's definition if the system is understood as a time-invariant nonlinear system, characterize it in terms of the system data and their derivatives, derive a normal form with respect to a time-varying linear coordinate transformation, and finally characterize the zero dynamics.**

***Index Terms*—Normal form, relative degree, time-varying systems.**

I. Introduction

THE concept of relative degree goes back to single-input single-output linear systems described in the frequency domain by a transfer function $p(s)/q(s)$ where the relative degree is defined by $r = \deg q - \deg p$; p, q denote polynomials with real coefficients. To derive a characterization in the time domain, take any realization of $p(s)/q(s)$, say

$$\left.\begin{aligned} \dot{x}(t) &= Ax(t) + bu(t) \\ y(t) &= cx(t) \end{aligned}\right\} \tag{1.1}$$

with $A \in \mathbb{R}^{n\times n}$, $b, c^T \in \mathbb{R}^n$. Then $p(s)/q(s) = c(sI - A)^{-1}b = \sum_{k=0}^{\infty} cA^k b s^{-(k+1)}$, and it is easy to see that $r = \deg q - \deg p$ if, and only if

$$\forall\, k = 0, \dots, r-2:\ cA^k b = 0 \quad cA^{r-1}b \neq 0. \tag{1.2}$$

Isidori [5, p. 137] generalizes the concept of relative degree to single-input–single-output time-invariant nonlinear systems, affine in the control, of the form

$$\left.\begin{aligned} \dot{x} &= f(x) + g(x)u(t) \\ y(t) &= h(x(t)), \end{aligned}\right\} \tag{1.3}$$

with $f, g \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^n)$, $h \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R})$, and $\ell \in \mathbb{N}$: System (1.3) has relative degree $r \in \{1, \dots, \ell\}$ at $x^0 \in \mathbb{R}^n$ if, and only if, there exists an open neighbourhood $\mathcal{X}$ of x^0, such that,

$$\forall\, x \in \mathcal{X} \qquad \forall\, k \in \{0, \dots, r-2\}: \quad L_g L_f^k h(x) = 0 \quad L_g L_f^{r-1} h(x^0) \neq 0 \tag{1.4}$$

where $L_f \lambda = (\partial\lambda/\partial x) f$ denotes the derivative of a function λ along a vector field f; see, for example, [5, Sec. 1.2].

The importance of the relative degree is that it leads to a normal form [5, Sec. 4.1]: If (1.3) has relative degree r at x^0, then there exists a diffeomorphism Φ, defined in a neighbourhood of x^0, which transforms (1.3) under $(\xi, \eta) = \Phi(x)$, $\xi = (y, \dots, y^{(r-1)})^T$, to

$$\left.\begin{aligned} y^{(r)} &= L_f^r h\left(\Phi^{-1}(\xi,\eta)\right) + L_g L_f^{(r-1)} h\left(\Phi^{-1}(\xi,\eta)\right) u(t) \\ \dot{\eta} &= q(\xi, \eta) \\ y(t) &= \xi_1(t), \end{aligned}\right\} \tag{1.5}$$

for some $q \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^{n-r})$. This form gives immediately that, for $x(t_0) = x^0$, "the relative degree r is exactly equal to the number of times one has to differentiate the output $y(t)$ at time $t = t_0$ in order to have the value $u(t_0)$ of the input explicitly appearing" [5, p. 139]. Moreover, u enters only in a single differential equation in (1.5) directly and it is possible to read off the zero dynamics, see [5, Sec. 4.3] and Section III of this paper.

The purpose of the present note is to introduce and characterize the concept of relative degree for time-varying linear systems of the form

$$\left.\begin{aligned} \dot{x} &= A(t)x + B(t)u(t) \\ y(t) &= C(t)x(t), \end{aligned}\right\} \tag{1.6}$$

with $A \in \mathcal{C}^\ell(\mathbb{R}, \mathbb{R}^{n\times n}), B, C^T \in \mathcal{C}^\ell(\mathbb{R}, \mathbb{R}^{n\times m})$, $\ell \in \mathbb{N}$, and to derive a time-varying linear coordinate transformation which takes (1.6) to a normal form.

The paper is organized as follows. In Section II, we present a definition of relative degree for time-varying nonlinear systems. It is shown that this definition coincides, if the system is time-invariant, with Isidori's definition [5, p. 220] respectively with the definition by Liberzon *et al.* [6, Def. 2]; furthermore, if the system is linear time-varying and viewed as a time-invariant nonlinear system, the definition coincides again with Isidori's definition. Our main result is a normal form for time-varying linear systems given in Section III. In Section IV, we parameterize the zero dynamics of time-varying linear systems and characterize their stability properties and show how a high-gain derivative output feedback controller stabilizes the linear system. We have relegated a refined version of Doležal's Theorem to the Appendix, which is used in the antecedent proofs.

We close this introduction with remarks on notation. Throughout, $\mathbb{R}_{\geq 0} := [0, \infty)$ and $\|\cdot\|$ is the Euclidean inner product or the induced norm on $\mathbb{R}^n$; if $M, N \in \mathbb{R}^{n\times n}$ are symmetric, then the notion $M \geq N$ means $x^T M x \geq x^T N x$ for all $x \in \mathbb{R}^n$; $\mathrm{Gl}_n(\mathbb{R})$ denotes the general linear group of invertible matrices $A \in \mathbb{R}^{n\times n}$; $C^\ell(U, W)$ is the vector space of ℓ-times differentiable functions $f : U \to W$, U and W are open sets; and $\mathcal{L}^\infty(U, W)$ the set of essentially bounded functions $f : U \to W$.

Manuscript received October 12, 2005; revised March 13, 2006 and September 25, 2006. Recommended by Associate Editor A. Hansson.

The authors are with the Institut für Mathematik, Technische Universität Ilmenau, 98693 Ilmenau, Germany (e-mail: achim.ilchmann@tu-ilmenau.de; markus.mueller@tu-ilmenau.de).

Digital Object Identifier 10.1109/TAC.2007.895843

II. Relative Degree: Definition and Characterizations

For time-invariant nonlinear multiple-input–multiple-output systems, affine in the control, of the form

$$\left.\begin{aligned}\dot{x} &= f(x) + g(x)u(t)\\ y(t) &= h(x(t)),\end{aligned}\right\} \tag{2.1}$$

with $f \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^n)$, $g \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^{n\times m})$, $h \in \mathcal{C}^\ell(\mathbb{R}^n, \mathbb{R}^m)$, $\ell \in \mathbb{N}$, the strict relative degree is defined as follows.

Definition 2.1: [5, p. 220]

Let $\mathcal{X} \subset \mathbb{R}^n$ be open and $r \in \{1, \ldots, \ell\}$. The time-invariant nonlinear system (2.1) has (*strict*) *relative degree* r on $\mathcal{X}$ if, and only if

i) $\forall\, \xi \in \mathcal{X}\ \forall\, k \in \{0, \ldots, r-2\} : L_g L_f^k h(\xi) = 0_{m\times m}$;

ii) $\forall\, \xi \in \mathcal{X} : L_g L_f^{r-1} h(\xi) \in \mathrm{Gl}_m(\mathbb{R})$.

This definition is due to Isidori [5, p. 220] who defines it more general for a vector relative degree; we consider only the "strict" relative degree, that is ii). In [5, p. 220], the relative degree is defined at a point $\xi^0 \in \mathbb{R}^n$; however, this is equivalent to Definition 2.1, the latter is technically easier to deal with in the following. In the single-input–single-output case, the notion of "strict" is redundant and Isidori shows that "the relative degree r is exactly equal to the number of times one has to differentiate the output $y(t)$ at time $t = t_0$ in order to have the value $u(t_0)$ of the input explicitly appearing" [5, p. 139]. This latter characterization is formalized by Liberzon *et al.* [6] and related to output-input stability. We extend their notion of functions H_k to time-varying nonlinear systems of the form

$$\left.\begin{aligned}\dot{x} &= F(t, x, u(t))\\ y(t) &= H(t, x(t))\end{aligned}\right\} \tag{2.2}$$

where

$$\begin{aligned}F &\in \mathcal{C}^\ell(\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m, \mathbb{R}^n)\\ H &\in \mathcal{C}^\ell(\mathbb{R}\times\mathbb{R}^n, \mathbb{R}^m), \qquad \ell \in \mathbb{N}\end{aligned}$$

and define recursively, for $k = 0, 1, 2, \ldots, \ell - 1$, the functions $H_0(t, x) := H(t, x)$

$$\left.\begin{aligned}&H_{k+1} : \mathbb{R}\times\mathbb{R}^n\times(\mathbb{R}^m)^{k+1} \to \mathbb{R}^m\\ &(t, x, u_0, \ldots, u_k) \mapsto \tfrac{\partial H_k}{\partial t} + \tfrac{\partial H_k}{\partial x} F(t, x, u_0)\\ &\qquad + \sum_{j=0}^{k-1} \tfrac{\partial H_k}{\partial u_j} u_{j+1}\end{aligned}\right\}. \tag{2.3}$$

This allows to express the kth derivative of $y(t)$ in terms of t, $x(t)$ and $u(t), \ldots, u^{(k-1)}(t)$

$$\forall\, t \in \mathbb{R}\quad \forall\, k \in \{1, \ldots, \ell-1\} : \quad y^{(k)}(t) = H_k\left(t, x(t), u(t), \ldots, u^{(k-1)}(t)\right).$$

Definition 2.2: (Generalization of [2, Def. 2])

Let $(\mathcal{T}, \mathcal{X}) \subset \mathbb{R}\times\mathbb{R}^n$ be open and $r \in \mathbb{N}$. Then, system (2.2) is said to have (*strict and uniform*) *relative degree* $r \in \mathbb{N}$ on $\mathcal{T}\times\mathcal{X}$ if, and only if

i) $\forall\ k = 1, \ldots, r-1\ \forall\ i = 0, \ldots, k-1\ \forall\ (t, x, u_0, \ldots, u_{k-1}) \in \mathcal{T}\times\mathcal{X}\times(\mathbb{R}^m)^k$: $(\partial H_k/\partial u_i)(t, x, u_0, \ldots, u_{k-1}) = 0_{m\times m}$;

ii) $\forall\ (t, x, u_0, \ldots, u_{r-1}) \in \mathcal{T}\times\mathcal{X}\times(\mathbb{R}^m)^r$: $(\partial H_r/\partial u_0)(t, x, u_0, \ldots, u_{r-1}) \in \mathrm{Gl}_m(\mathbb{R})$.

Remark 2.3:

i) The notion "strict" refers to the multivariable case where we do not allow for a relative degree vector $(r_1, \ldots, r_m) \in \mathbb{N}^m$ with different entries, see [5, Sec. 5.1], but assume that the matrices $(\partial/\partial u_i)H_k$ are either "strictly" zero or invertible, globally on their domain. If the system (2.2) had p outputs, then one could recursively define $H_{k+1}^j : \mathbb{R}\times\mathbb{R}^n\times(\mathbb{R}^m)^{k+1} \to \mathbb{R}$ for $j = 1, \ldots, p$ and define a vector relative degree $(r_1, \ldots, r_p)$. This is omitted for brevity. Note that the vector relative degree, if it exists, is not necessarily unique; see [6, Sec. III.B].

ii) If i) of Definition 2.2 holds, then it is easy to see that the functions H_k do not depend on $(u_0, \ldots, u_{k-1})$ for $k = 1, \ldots, r-1$; thus only $\partial H_k/\partial u_0$ has to be checked in i).

Proposition 2.4: Let $\mathcal{X} \subset \mathbb{R}^n$ be an open set, $r, \ell \in \mathbb{N}$ with $r \le \ell$, and consider the time-invariant nonlinear system (2.1), affine in the control. Then (2.1) has relative degree r on $\mathcal{X}$ in the sense of Definition 2.1 if, and only if, (2.1) has relative degree r on $\mathbb{R}\times\mathcal{X}$ in the sense of Definition 2.2.

Proof: The following analysis is considered for $(t, x) \in \mathbb{R}\times\mathcal{X}$ only; for notational convenience, we omit to repeat this and also suppress the arguments of the functions for simplicity.

"$\Rightarrow$": Suppose that (2.1) has relative degree r in the sense of Definition 2.1. We first show by induction on k that the following holds:

$$\forall\, k \in \{1, \ldots, r-1\}\ \forall\, i \in \{1, \ldots, k-1\} : \quad H_k = L_f^k h \quad \frac{\partial H_k}{\partial u_i} = 0_{m\times m}. \tag{2.4}$$

We have, in view of i) in Definition 2.1

$$\begin{aligned}H_1 &= \frac{\partial H_0}{\partial x}(f + g u_0) + \frac{\partial H_0}{\partial u_0} u_1\\ &= L_f h + L_g h u_0\\ &= L_f h\end{aligned}$$

and so $\partial H_1/\partial u_0 = (\partial/\partial u_0) L_f h = 0_{m\times m}$. If (2.4) holds for all $k \le r-2$, then, in view of (i) in Definition 2.1,

$$\begin{aligned}H_{k+1} &= \frac{\partial H_k}{\partial x}(f + g u_0) + \sum_{j=0}^{k-1} \frac{\partial H_k}{\partial u_j} u_{j+1}\\ &= L_f^{k+1} h + L_g L_f^k h u_0\\ &= L_f^{k+1} h\end{aligned}$$

and so, for all $i \in \{1, \ldots, k\}$, $\partial H_{k+1}/\partial u_i = (\partial/\partial u_i) L_f^{k+1} h = 0_{m\times m}$. This proves (2.4) and, therefore, i) of Definition 2.2 holds.

To prove ii) in Definition 2.2, note that, in view of (2.4),

$$\begin{aligned}H_r &= \frac{\partial H_{r-1}}{\partial x}(f + g u_0) + \sum_{j=0}^{r-2} \frac{\partial H_{r-1}}{\partial u_j} u_{j+1}\\ &= L_f^r h + L_g L_f^{r-1} h u_0\end{aligned}$$

and thus, invoking ii) in Definition 2.1, $\partial H_r/\partial u_0 = L_g L_f^{r-1} h \in \mathrm{Gl}_m(\mathbb{R})$. This proves ii) in Definition 2.2.

"$\Leftarrow$": Suppose that (2.1) has relative degree r in the sense of Definition 2.2. We show first by induction on k that the following holds:

$$\forall\, k \in \{0,1,\dots,r-2\} : L_g L_f^k h = 0_{m\times m} \quad H_{k+1} = L_f^{k+1} h. \tag{2.5}$$

We have

$$\begin{aligned} H_1 &= \frac{\partial H_0}{\partial x}(f+gu_0) + \frac{\partial H_0}{\partial u_0} u_1 \\ &= \frac{\partial h}{\partial x} f + \frac{\partial h}{\partial x} g u_0 \\ &= L_f h + L_g h u_0 \end{aligned}$$

and so, in view of i) in Definition 2.2, $0_{m\times m} = \partial H_1/\partial u_0 = L_g h$, which gives $H_1 = L_f h$. If (2.5) holds for all $k \le r-3$, then, again in view of i) in Definition 2.2

$$\begin{aligned} H_{k+2} &= \frac{\partial H_{k+1}}{\partial x}(f+gu_0) + \sum_{j=0}^{k} \frac{\partial H_{k+1}}{\partial u_j} u_{j+1} \\ &= \frac{\partial}{\partial x} L_f^{k+1} h f + \frac{\partial}{\partial x} L_f^{k+1} h g u_0 \\ &= L_f^{k+2} h + L_g L_f^{k+1} h u_0 \end{aligned}$$

and furthermore $0_{m\times m} = \partial H_{k+2}/\partial u_0 = L_g L_f^{k+1} h$, which yields $H_{k+2} = L_f^{k+2} h$. This proves (2.5).

Finally, by (2.5), i) in Definition 2.1 follows. Applying (2.5) again gives

$$\begin{aligned} H_r &= \frac{\partial H_{r-1}}{\partial x}(f+gu_0) + \sum_{j=0}^{r-2} \frac{\partial H_{r-1}}{\partial u_j} u_{j+1} \\ &= L_f^r h + L_g L_f^{r-1} h u_0 \end{aligned}$$

and thus, invoking ii) in Definition 2.2, $\partial H_r/\partial u_0 = L_g L_f^{r-1} h \in \mathrm{Gl}_m(\mathbb{R})$. This proves ii) in Definition 2.1 and completes the proof of the proposition. ☐

Remark 2.5:

(i) It follows from the proof of Proposition 2.4 that the relative degree of the time-invariant system (2.1) does not depend on t: if its relative degree is defined on some open set $\mathcal{T}\times\mathcal{X} \subset \mathbb{R}\times\mathbb{R}^n$, then it is defined on $\mathbb{R}\times\mathcal{X}$.

(ii) It also follows from Proposition 2.4 and [6, Prop. 2] that for single-input single-output time-invariant systems (2.1), Definition 2.2 coincides with the definition of the relative degree given by [6, Def. 2].

We introduce the following right operator:

Notation 2.6: For $\ell \in \mathbb{N}$, $A \in \mathcal{C}^\ell(\mathbb{R}, \mathbb{R}^{n\times n})$, and $C \in \mathcal{C}^\ell(\mathbb{R}, \mathbb{R}^{m\times n})$ set

$$\begin{aligned} &\forall\, t\in\mathbb{R} : \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^0 (C(t)) := C(t) \\ &\forall\, t\in\mathbb{R} : \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right) (C(t)) := \dot{C}(t) + C(t)A(t) \\ &\forall\, t\in\mathbb{R} \quad \forall\, k\in\{1,\dots,\ell\} : \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^k (C(t)) \\ &\quad := \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right) \left(\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{k-1} (C(t)) \right). \end{aligned}$$

The subscript r in $A_r(C)$ indicates that A acts on C by multiplication from the right.

The operator $(\mathrm{d}/\mathrm{d}t + A(t)_r)$ had already been used by [3], [4], [7], and [9].

Theorem 2.7: Let $r,\ell \in \mathbb{N}$ with $r \le \ell$, and $\mathcal{T} \subset \mathbb{R}$ be an open set. Then, the time-varying linear system (1.6) has relative degree r on $\mathcal{T}\times\mathbb{R}^n$, if and only if, (A,B,C) satisfy

$$\left.\begin{aligned} &\forall\, t\in\mathcal{T} \quad \forall\, k=0,\dots,r-2 : \\ &\quad \left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^k (C(t))\, B(t) = 0_{m\times m} \\ &\forall\, t\in\mathcal{T} : \\ &\quad \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{r-1} (C(t))\, B(t)\right) \in \mathrm{Gl}_m(\mathbb{R}) \end{aligned}\right\}. \tag{2.6}$$

The proof of Theorem 2.7 depends on the following technicality.

Lemma 2.8: Let $r,\ell \in \mathbb{N}$ with $r \le \ell$. Then the functions H_k defined in (2.3) and applied to the linear time-varying system (1.6) satisfy, for all $k \in \{0,\dots,\ell-1\}$ and all $(t,x,u_0,\dots,u_k) \in \mathbb{R}\times\mathbb{R}^n\times(\mathbb{R}^m)^{k+1}$

$$\begin{aligned} &H_{k+1}(t,x,u_0,\dots,u_k) \\ &\quad = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{k+1} (C(t))\, x + \sum_{j=0}^{k}\sum_{i=j}^{k} \binom{i}{j} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j} \\ &\qquad \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{k-i} (C(t))\, B(t) \right] u_j. \end{aligned} \tag{2.7}$$

Proof: Applying the definition of H_k to

$$\begin{aligned} &F(t,x,u) := A(t)x + B(t)u \\ \text{and} \quad &H(t,x) := C(t)x \\ &\qquad\qquad \text{for } (t,x,u)\in\mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \end{aligned}$$

we have

$$H_1(t,x,u_0) = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^1 (C(t))\, x$$

which shows (2.7) for $k=0$. To prove (2.7) by induction over $k \in \{0,\dots,\ell-1\}$, assume that (2.7) holds for all $k \in \{0,\dots,\ell-2\}$. Then, suppressing the arguments of H_k and A,B,C for simplicity, we calculate

$$\begin{aligned} H_{k+1} &= \frac{\partial}{\partial t}\Bigg[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^k (C)x + \sum_{j=0}^{k-1}\sum_{i=j}^{k-1} \binom{i}{j}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j} \\ &\qquad \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{k-1-i} (C)B\right] u_j \Bigg] \\ &\quad + \frac{\partial}{\partial x}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^k (C)x\right][Ax + Bu_0] \\ &\quad + \sum_{l=0}^{k-1} \frac{\partial}{\partial u_l}\Bigg[\sum_{j=0}^{k-1}\sum_{i=j}^{k-1} \binom{i}{j}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j} \\ &\qquad \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{k-1-i} (C)B\right] u_j \Bigg] u_{l+1} \\ &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^k (C)\right) + \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^k (C)A\right] x \\ &\quad + \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^k (C)Bu_0 + \sum_{j=0}^{k-1}\sum_{i=j}^{k-1} \binom{i}{j}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j+1} \end{aligned}$$

$$
\begin{aligned}
&\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_j\\
&+\sum_{j=1}^{k}\sum_{i=j-1}^{k-1}\binom{i}{j-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j+1}\\
&\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_j\\
=&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k+1}(C)x+\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k}(C)Bu_0\\
&+\sum_{i=0}^{k-1}\binom{i}{0}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i+1}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_0\\
&+\sum_{j=1}^{k-1}\left(\sum_{i=j}^{k-1}\left[\binom{i}{j}+\binom{i}{j-1}\right]\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j+1}\right.\\
&\qquad\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_j\\
&\qquad+\binom{j-1}{j-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-1-j+1}\\
&\qquad\left.\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-j+1}(C)B\right]u_j\right)\\
&+\sum_{i=k-1}^{k-1}\binom{i}{k-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-k+1}\\
&\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_k\\
=&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k+1}(C)x\\
&+\underbrace{\sum_{i=-1}^{k-1}\binom{i}{0}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i+1}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-1-i}(C)B\right]u_0}_{=\sum_{i=0}^{k}\binom{i}{0}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-0}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-i}(C)B\right]u_0}\\
&+\sum_{j=1}^{k-1}\left(\sum_{i=j}^{k-1}\binom{i+1}{j}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i+1-j}\right.\\
&\qquad\times\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-(i+1)}(C)B\right]u_j\\
&\qquad\left.+\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-j}(C)Bu_j\right)\\
&+\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{0}(C)Bu_k\\
=&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k+1}(C)x\\
&+\sum_{j=0}^{k}\sum_{i=j}^{k}\binom{i}{j}\left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{i-j}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^{k-i}(C)B\right]u_j.
\end{aligned}
$$

This shows (2.7) for $k+1$ and, therefore, the proof of the lemma is complete. □

Proof of Theorem 2.7: "$\Rightarrow$": Suppose that Definition 2.2 holds. We show the first condition in (2.6) by induction over $k \in \{1, \ldots, r-2\}$ (omitting the arguments of the operators). For $N = k = 0$ we have, by (2.7),

$$H_1 = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(C)x + CBu_0$$

and so, by Definition 2.2, $0_{m\times m} = \partial H_1/\partial u_0 = CB$. Suppose that the first condition in (2.6) holds for all $k = 0, \ldots, N$, where $N \le r-3$. Then (2.7) yields

$$H_{N+2} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{N+2}(C)x + \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{N+1}(C)Bu_0$$

and so, invoking i) of Definition 2.2

$$0_{m\times m} = \frac{\partial H_{N+2}}{\partial u_0} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{N+1}(C)B.$$

This proves the first condition in (2.6).

To see the second condition in (2.6), note that (2.7) yields

$$H_r = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r}(C)x + \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C)Bu_0$$

and so, by ii) of Definition 2.2

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C)B = \frac{\partial H_r}{\partial u_0} \in \mathrm{Gl}_m(\mathbb{R}).$$

This completes the proof of (2.6).

"$\Leftarrow$": Suppose that (2.6) holds. Then (2.7) yields

$$\forall\, k = 0, \ldots, r-2 : H_{k+1} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{k+1}(C)x$$

and thus i) in Definition 2.2 follows. Finally, the second statement in (2.6) together with (2.7) gives

$$
\begin{aligned}
\frac{\partial H_r}{\partial u_0} &= \frac{\partial}{\partial u_0}\left(\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r}(C)x + \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C)Bu_0\right)\\
&= \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C)B \in \mathrm{Gl}_m(\mathbb{R}).
\end{aligned}
$$

This proves ii) of Definition 2.2 and completes the proof of the theorem. □

Remark 2.9:

i) It follows from the proof of Theorem 2.7 that the relative degree of the time-varying linear system (1.6) does not depend on x: if its relative degree is defined on some open $\mathcal{T} \times \mathcal{X} \subset \mathbb{R} \times \mathbb{R}^n$, then it is defined on $\mathcal{T} \times \mathbb{R}^n$. We therefore omit, at most places in the following, the second component in $\mathcal{T} \times \mathbb{R}^n$.

ii) If A, B, C are real analytic matrices and the linear system (1.6) has relative degree r on $\mathcal{T} \times \mathbb{R}^n$ for some open $\mathcal{T} \subset \mathbb{R}$, then the Identity Theorem for analytic functions implies that (1.6) has relative degree r on $(\mathbb{R} \setminus D) \times \mathbb{R}^n$, where D denotes a discrete set.

iii) If the linear system (1.6) is time-invariant, then Theorem 2.7 yields that (1.6) has relative degree $r \in \mathbb{N}$ on $\mathbb{R} \times \mathbb{R}^n$ if, and only if

$$CA^kB = 0_{m\times m} \text{ for all } k = 0, \ldots, r-2 \quad \text{and} \quad CA^{r-1}B \in \mathrm{Gl}_m(\mathbb{R}).$$

This is the well known characterization of strict relative degree, see [5, Rem. 4.1.2] for single-input–single-output systems.

Instead of defining the relative degree of the linear time-varying system (1.6) as in Definition 2.2, we may consider the equivalent description of (1.6) as a time-invariant nonlinear system and determine the relative degree according to Definition 2.1. In the following we will show that both definitions coincide. Introducing an additional variable z with initial condition $z(0) = 0$, (1.6) is equivalent to

$$\left.\begin{aligned} \tfrac{\mathrm{d}}{\mathrm{d}t}\tbinom{x}{z} &= \tbinom{A(z)x}{1} + \tbinom{B(z)}{0}u(t) \\ y(t) &= C\left(z(t)\right)x(t) \end{aligned}\right\}. \tag{2.8}$$

Proposition 2.10: Let $\mathcal{T} \times \mathcal{X} \subset \mathbb{R} \times \mathbb{R}^n$ be open, $r, \ell \in \mathbb{N}$ with $r \leq \ell$. The time-varying linear system (1.6) has relative degree r on $\mathcal{T} \times \mathcal{X}$ in the sense of Definition 2.2 if, and only if, the equivalent nonlinear time-invariant system (2.8) has relative degree r on $\mathcal{X} \times \mathcal{T}$ in the sense of Definition 2.1.

Proof: Writing

$$f(x,z) = \begin{pmatrix} A(z)x \\ 1 \end{pmatrix} \quad g(x,z) = \begin{pmatrix} B(z) \\ 0 \end{pmatrix}$$
$$h(x,z) = C(z)x$$

we show by induction over $N \in \{0, \ldots, \ell\}$ that

$$\forall\, (x,z) \in \mathcal{X} \times \mathcal{T} \quad \forall\, k \in \{0, \ldots, \ell\}:$$
$$L_f^k h(x,z) = \left(\frac{\mathrm{d}}{\mathrm{d}z} + A(z)_r\right)^k (C(z))\, x. \tag{2.9}$$

For $k = N = 0$, we obviously have $L_f^0 h(x,z) = C(z)x$. Suppose that (2.9) holds for all $k \in \{0, \ldots, N\}$ for some $N \in \{0, \ldots, \ell - 1\}$. Then

$$\begin{aligned} L_f^{N+1}h(x,z) &= L_f\left(\left(\frac{\mathrm{d}}{\mathrm{d}z} + A_r\right)^N (C)x\right) \\ &= \left[\frac{\partial}{\partial x}\left(\left(\frac{\mathrm{d}}{\mathrm{d}z} + A_r\right)^N (C)x\right),\right. \\ &\qquad \left.\frac{\partial}{\partial z}\left(\left(\frac{\mathrm{d}}{\mathrm{d}z} + A_r\right)^N (C)x\right)\right] \begin{pmatrix} Ax \\ 1 \end{pmatrix} \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}z} + A_r\right)^{N+1} (C)x. \end{aligned}$$

This completes the proof of (2.9) and gives, for all $(x,z) \in \mathcal{X} \times \mathcal{T}$ and all $k \in \{0, \ldots, \ell\}$

$$\begin{aligned} L_g L_f^k h(x,z) &= \left[\frac{\partial}{\partial x}\left(\left(\frac{\mathrm{d}}{\mathrm{d}z} + A(z)_r\right)^k (C(z))\, x\right),\right. \\ &\qquad \left.\frac{\partial}{\partial z}\left(\left(\frac{\mathrm{d}}{\mathrm{d}z} + A(z)_r\right)^k (C(z))\, x\right)\right] \begin{pmatrix} B(z) \\ 0 \end{pmatrix} \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}z} + A(z)_r\right)^k (C(z))\, B(z). \end{aligned} \tag{2.10}$$

Finally, the proposition is a consequence of (2.9) and (2.10) applied to Theorem 2.7. □

III. Normal Form

In this section, we derive a normal form for time-varying linear systems (1.6). Theorem 2.7 may already indicate that the matrix function $(\mathrm{d}/\mathrm{d}t + A(\cdot)_r)^k (C(\cdot))$, $k = 0, \ldots, r-1$, are candidates for a new basis; however, this potential basis needs to be completed. We introduce the following matrix functions which will serve to derive a time-varying linear transformation.

Let $r, \ell \in \mathbb{N}$ with $r \leq \ell$. Consider the system (1.6) and define, for $r \in \mathbb{N}$ and all $t \in \mathbb{R}$

$$\mathcal{C}(t) := \begin{bmatrix} C(t) \\ \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)(C(t)) \\ \vdots \\ \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{r-1}(C(t)) \end{bmatrix} \in \mathbb{R}^{rm \times n}$$

$$\begin{aligned} \mathcal{B}(t) := \Bigg[B(t), &\left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)(B(t)), \ldots, \\ &\left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^{r-1}(B(t))\Bigg] \in \mathbb{R}^{n \times rm} \end{aligned}$$

$$\Gamma(t) := \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{r-1}(C(t))\, B(t) \in \mathbb{R}^{m \times m}.$$

The following proposition presents two more characterizations for (1.6) having relative degree r. They are rather technical but essential to design the coordinate transformation for the normal form.

Proposition 3.1: Let $r, \ell \in \mathbb{N}$ with $r \leq \ell$ and $\mathcal{T} \subset \mathbb{R}$ be an open set. Then the following conditions are equivalent.

i) The system (1.6) has relative degree r on $\mathcal{T}$.

ii)

a) $\forall\, t \in \mathcal{T}$

$$\forall\, (i,j) \in \left\{(i,j) \in \mathbb{N}_0^2 \middle| 0 \leq i + j \leq r - 2\right\}:$$
$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^i (C(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^j (B(t)) = 0_{m \times m}.$$

b) $\forall\, t \in \mathcal{T}$

$$\forall\, (i,j) \in \left\{(i,j) \in \mathbb{N}_0^2 \middle| i + j = r - 1\right\}:$$
$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^i (C(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^j (B(t)) \in \mathrm{Gl}_m(\mathbb{R}).$$

iii) $\forall\, t \in \mathcal{T}$:

$$\mathcal{C}(t)\mathcal{B}(t) = \begin{bmatrix} 0 & & (-1)^{r-1}\Gamma(t) \\ & \cdot^{\cdot^{\cdot}} & \\ \Gamma(t) & & * \end{bmatrix} \in \mathrm{Gl}_{rm}(\mathbb{R}).$$

The proof of Proposition 3.1 iii) is related to some methods used in [3], [9, p. 37] but is different. The proof of Proposition 3.1 depends crucially on the following technical lemma.

Lemma 3.2: The linear time-varying system (1.6) satisfies, for all $i, j \in \mathbb{N}_0$ with $i + j \leq \ell$ and all $t \in \mathbb{R}$

$$\begin{aligned} &\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^i (C(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^j (B(t)) \\ &\quad = \frac{\mathrm{d}}{\mathrm{d}t}\left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^i (C(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^{j-1} (B(t))\right] \\ &\qquad - \left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{i+1} (C(t)) \\ &\qquad \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^{j-1} (B(t)) \end{aligned} \tag{3.1}$$

and

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^i (C(t)) \left(\frac{\mathrm{d}}{\mathrm{d}t} - A(t)\right)^j (B(t)) = \sum_{\mu=0}^{j} (-1)^\mu \binom{j}{\mu} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-\mu} \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A(t)_r\right)^{i+\mu} (C(t))\, B(t)\right]. \quad (3.2)$$

Proof: The equality (3.1) follows from the calculation:

$$\begin{aligned}
&\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^j (B) \\
&= \left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right) \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right) \\
&= \left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \Big\{ \cancel{\tfrac{\mathrm{d}}{\mathrm{d}t} \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right)} \\
&\qquad - A \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right) \Big\} \\
&\quad + \tfrac{\mathrm{d}}{\mathrm{d}t} \left[\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right)\right] \\
&\quad - \tfrac{\mathrm{d}}{\mathrm{d}t} \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C)\right) \cdot \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right) \\
&\quad - \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C)\right) \cdot \cancel{\tfrac{\mathrm{d}}{\mathrm{d}t} \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right)} \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t} \left[\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B)\right)\right] \\
&\quad - \underbrace{\left[\tfrac{\mathrm{d}}{\mathrm{d}t} \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C)\right) + \left(\tfrac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) A\right]}_{= \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)\left(\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C)\right)} \\
&\quad \cdot \left(\tfrac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j-1} (B).
\end{aligned}$$

We prove (3.2) by fixing $i \in \mathbb{N}_0$ and induction over $j = 0, \ldots, \ell - i$. For $j = 0$, (3.2) is obvious. Suppose that (3.2) holds for $j \le \ell - i - 1$. Then, invoking (3.1), it follows that

$$\begin{aligned}
&\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^i (C) \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)^{j+1} (B) \\
&= \frac{\mathrm{d}}{\mathrm{d}t} \left[\sum_{\mu=0}^{j} (-1)^\mu \binom{j}{\mu} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-\mu} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+\mu} (C) B\right]\right] \\
&\quad - \sum_{\mu=0}^{j} (-1)^\mu \binom{j}{\mu} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-\mu} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+1+\mu} (C) B\right] \\
&= \binom{j}{0} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+0} (C) B\right] \\
&\quad + \sum_{\mu=1}^{j} \left[(-1)^\mu \binom{j}{\mu} - (-1)^{\mu-1} \binom{j}{\mu-1}\right] \\
&\quad \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1-\mu} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+\mu} (C) B\right] \\
&\quad - (-1)^{j+1-1} \binom{j}{j+1-1} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j-j-1+1} \\
&\quad \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+j+1} (C) B\right] \\
&= (-1)^0 \binom{j}{0} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1-0} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+0} (C) B\right] \\
&\quad + \sum_{\mu=1}^{j} (-1)^\mu \binom{j+1}{\mu} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1-\mu} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+\mu} (C) B\right] \\
&\quad + (-1)^{j+1} \binom{j+1}{j+1} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1-(j+1)} \\
&\quad \times \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+j+1} (C) B\right] \\
&= \sum_{\mu=0}^{j+1} (-1)^\mu \binom{j+1}{\mu} \left(\frac{\mathrm{d}}{\mathrm{d}t}\right)^{j+1-\mu} \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{i+\mu} (C) B\right].
\end{aligned}$$

This completes the proof of the lemma. □

Proof of Proposition 3.1: The equivalence "i) ⇔ ii)" follows from (3.2) and (2.6), and the equivalence "ii) ⇔ iii)" follows from (3.2). □

The following corollary is a direct consequence of Proposition 3.1.

Corollary 3.3: If the linear time-varying system (1.6) has relative degree $r \in \mathbb{N}$ on some open set $\mathcal{T} \subset \mathbb{R}$, then the following hold.

i) $\forall\, t \in \mathcal{T}$: $\operatorname{rank} \mathcal{C}(t) = rm$ and $\operatorname{rank} \mathcal{B}(t) = rm$.

ii) The two sets of matrix functions $C(\cdot), (\mathrm{d}/\mathrm{d}t + A(\cdot)_r)(C(\cdot)), \ldots, (\mathrm{d}/\mathrm{d}t + A(\cdot)_r)^{r-1}(C(\cdot))$ and $B(\cdot), (\mathrm{d}/\mathrm{d}t - A(\cdot))(B(\cdot)), \ldots, (\mathrm{d}/\mathrm{d}t - A(\cdot))^{r-1}(B(\cdot))$ are both linearly independent over $\mathcal{T}$.

iii) $rm \le n$.

We are now in a position to design a time-varying linear basis transformation which will lead to a normal form.

Remark 3.4: Let $r, \ell \in \mathbb{N}$ with $r \le \ell$ and $\mathcal{T} \subset \mathbb{R}$ be an open set. Suppose that the time-varying linear system (1.6) has relative degree $r \in \mathbb{N}$ on $\mathcal{T}$. By Corollary 3.3, the rows in $\mathcal{C}$ qualify as new basis but the basis needs to be completed. By Theorem A.1, we may choose $T = [t_1, \ldots, t_n] \in \mathcal{C}^1(\mathcal{T}, \mathrm{Gl}_n(\mathbb{R}))$ such that

$$\begin{bmatrix} \mathcal{C} \\ 0 \end{bmatrix} T = \begin{bmatrix} F & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{C}^1(\mathcal{T}, \mathbb{R}^{n \times n}) \text{ with } F \in \mathcal{C}^1\left(\mathcal{T}, \mathrm{Gl}_{rm}(\mathbb{R})\right).$$

Defining

$$\mathcal{V} := [t_{rm+1}, \ldots, t_n] \in \mathcal{C}^1\left(\mathcal{T}, \mathbb{R}^{n \times (n-rm)}\right)$$

it also follows from Theorem A.1 that

$$\mathcal{V} \in \mathcal{L}^\infty\left(\mathcal{T}, \mathbb{R}^{n \times (n-rm)}\right) \quad \text{and} \quad (\mathcal{V}^T \mathcal{V})^{-1} \mathcal{V}^T \in \mathcal{L}^\infty\left(\mathcal{T}, \mathbb{R}^{(n-rm) \times n}\right)$$

and, moreover, we have

$$\forall\, t \in \mathcal{T} : \operatorname{im} \mathcal{V}(t) = \ker \mathcal{C}(t) \text{ and } \operatorname{rank} \mathcal{V}(t)^T \mathcal{V}(t) = n - rm \quad (3.3)$$

and writing

$$\mathcal{U} := \begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix} \in \mathcal{C}^1(\mathcal{T}, \mathbb{R}^{n\times n}) \quad \text{and}$$
$$\mathcal{N} := (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T \left[I - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right] \in \mathcal{C}^1\left(\mathcal{T}, \mathbb{R}^{(n-rm)\times n}\right)$$

it follows from

$$\begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix} \left[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \mathcal{V}\right] = I_n$$

that $\mathcal{U} \in \mathcal{C}^1(\mathcal{T}, \mathrm{Gl}_n(\mathbb{R}))$ with inverse

$$\mathcal{U}^{-1} = \left[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \mathcal{V}\right] \in \mathcal{C}^1\left(\mathcal{T}, \mathrm{Gl}_n(\mathbb{R})\right).$$

□

We are now in a position to derive the main result of this note, that is a normal form of the time-varying linear system (1.6).

Theorem 3.5: Let $r, \ell \in \mathbb{N}$ with $r \le \ell$ and $\mathcal{T} \subset \mathbb{R}$ be an open set. Suppose the time-varying linear system (1.6) has relative degree r on $\mathcal{T}$ and choose $\mathcal{U} \in C(\mathcal{T}, \mathrm{Gl}_n(\mathbb{R}))$, $\mathcal{V} \in \mathcal{C}^1(\mathcal{T}, \mathbb{R}^{n\times(n-r)})$, and $\mathcal{N} \in \mathcal{C}^1(\mathcal{T}, \mathbb{R}^{(n-rm)\times n})$ as in Remark 3.4. Then the coordinate transformation

$$\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} := \mathcal{U}(t)x(t)$$
$$\xi(t) = \left(y(t)^T, \ldots, y^{(r-1)}(t)^T\right)^T \in \mathbb{R}^{rm}$$
$$\eta(t) \in \mathbb{R}^{n-rm}$$

converts (1.6) on $\mathcal{T}$ into

$$\left.\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} \xi \\ \eta \end{pmatrix} &= \hat{A}(t)\begin{pmatrix} \xi \\ \eta \end{pmatrix} + \hat{B}(t)u(t) \\ y(t) &= \hat{C}(t)\begin{pmatrix} \xi(t) \\ \eta(t) \end{pmatrix} \end{aligned}\right\} \tag{3.4}$$

where

$$\left.\begin{aligned} \hat{A}(t) &= \begin{bmatrix} 0 & I & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \ldots & 0 & I & 0 \\ R_1(t) & R_2(t) & \ldots & R_r(t) & S(t) \\ P(t) & 0 & \ldots & 0 & Q(t) \end{bmatrix} \\ \hat{B}(t) &= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \Gamma \\ 0 \end{bmatrix} \quad \hat{C}(t) = [I \quad 0 \quad \ldots \quad 0] \end{aligned}\right\} \tag{3.5}$$

and

$$\Gamma = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C)B \tag{3.6}$$

$$\begin{aligned} [R_1, \ldots, R_r, S] &= \left[\left[0_{m\times(rm-m)}, I_m\right]\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(\mathcal{C})\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \right. \\ &\qquad \left.\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(\mathcal{C})\mathcal{V}\right] && (3.7) \\ &= \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(C)\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(C)\mathcal{V}\right] && (3.8) \end{aligned}$$

$$\begin{aligned} Q &= \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(\mathcal{N})\mathcal{V} && (3.9) \\ &= -(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T \\ &\quad \cdot \left[\left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)\mathcal{V} - B\Gamma^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(C)\mathcal{V}\right] && (3.10) \\ P &= \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(\mathcal{N})\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\begin{bmatrix} I_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} && (3.11) \\ &= (-1)^{r-1}(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C} - I\right] \\ &\quad \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)^r(B)\Gamma^{-1}. && (3.12) \end{aligned}$$

Proof: The special form of $\mathcal{U}$ and $\mathcal{U}^{-1}$ gives immediately $\hat{B} = \mathcal{U}B$, $\hat{C}\mathcal{U} = C$ with the special structure as shown in (3.5) and Γ as in (3.6); furthermore

$$\hat{A} = [\mathcal{U}A + \dot{\mathcal{U}}]\mathcal{U}^{-1} = \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(\mathcal{U})\mathcal{U}^{-1}. \tag{3.13}$$

In view of Remark 3.4, it remains to show (3.8)–(3.12) and that

$$\begin{bmatrix} \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(C) \\ \vdots \\ \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^{r-1}(C) \\ \hline \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)^r(C) \\ \left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(\mathcal{N}) \end{bmatrix} = \begin{bmatrix} 0 & I & & & 0 \\ & \ddots & \ddots & & \vdots \\ & & 0 & I & 0 \\ \hline R_1 & R_2 & \ldots & R_r & S \\ P_1 & P_2 & \ldots & P_r & Q \end{bmatrix}\begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix}$$
$$\text{with } P_2 = \ldots = P_r = 0. \tag{3.14}$$

We first show (3.14). Since equality of the upper blocks in (3.14) is immediate, it remains to show that

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} + A_r\right)(\mathcal{N})\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} = [P_1, 0, \ldots, 0]. \tag{3.15}$$

Writing

$$\mathcal{C}\mathcal{B} = [\eta_1, \ldots, \eta_r] \quad (\mathcal{C}\mathcal{B})^{-1} = \begin{bmatrix} \psi^1 \\ \vdots \\ \psi^r \end{bmatrix}$$
$$\mathcal{B} = [\beta_1, \ldots, \beta_r]$$

we have

$$\begin{aligned} &\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)(\mathcal{B}) \\ &= \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\left[\beta_2, \ldots, \beta_r, \left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)^r(B)\right] \\ &= [\beta_1, \ldots, \beta_r] \cdot \begin{bmatrix} \psi^1 \\ \vdots \\ \psi^r \end{bmatrix} \\ &\qquad \cdot \left[\eta_2, \ldots, \eta_r, \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)^r(B)\right] \\ &= [\beta_1, \ldots, \beta_r]\begin{bmatrix} 0 & 0 & \ldots & 0 & * \\ I & 0 & \ldots & 0 & * \\ 0 & I & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & * \\ 0 & \ldots & 0 & I & * \end{bmatrix} \\ &= \left[\beta_2, \ldots, \beta_r, \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\left(\frac{\mathrm{d}}{\mathrm{d}t} - A\right)^r(B)\right] \end{aligned}$$

and, thus

$$\begin{aligned}&[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}-I_n]\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{B})\\&=\left[\beta_2,\dots,\beta_r,(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^r(B)\right]\\&\quad-\left[\beta_2,\dots,\beta_r,\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^r(B)\right]\\&=\left[0,\dots,0,[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}-I_n]\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^r(B)\right]\end{aligned}\tag{3.16}$$

which by invoking

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})&=\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right)\left(I-\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)\\&\quad+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(A-\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)\right)\end{aligned}\tag{3.17}$$

implies

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\mathcal{B}\\&=\Bigg[\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right)\left(I-\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)\\&\qquad+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(A-\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)\right)\Bigg]\mathcal{B}\\&=(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[A-\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)\right]\mathcal{B}\\&=(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}-I_n]\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{B})\\&=(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[0,\dots,0,[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}-I_n]\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^r(B)\right].\end{aligned}\tag{3.18}$$

Finally

$$(\mathcal{C}\mathcal{B})^{-1}=\begin{bmatrix} * & & \Gamma^{-1}\\ & \iddots & \\ (-1)^{r-1}\Gamma^{-1} & & 0\end{bmatrix}\tag{3.19}$$

applied to (3.18) yields (3.15), whence (3.14).

Next, we prove (3.8). By (3.13),

$$\begin{aligned}[R_1,\dots,R_r,S]&=[0_{m\times(rm-m)},I_m]\\&\quad\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{C})\,[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1},\mathcal{V}]\end{aligned}$$

and (3.8) follows from the definition of $\mathcal{C}$.

We show (3.9) and (3.10). Equality (3.9) follows immediately from the normal form (3.4) and (3.5). To see equality (3.9), note that (3.17) yields

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\mathcal{V}\\&=\Bigg[\frac{\mathrm{d}}{\mathrm{d}t}\left[(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right]\left[I_n-\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right]\\&\qquad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\frac{\mathrm{d}}{\mathrm{d}t}\left[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right]+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\\&\qquad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C})A\Bigg]\mathcal{V}\\&=\frac{\mathrm{d}}{\mathrm{d}t}\left[(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right]\Big[\mathcal{V}-\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\underbrace{\mathcal{C}\mathcal{V}}_{=0}\Big]\\&\quad+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\mathcal{V}\\&\quad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)+\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)A\right]\mathcal{V}\\&=\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\left[(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right]\mathcal{V}+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\mathcal{V}}_{=\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)((\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V})\mathcal{V}}\\&\quad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\Bigg[\dot{\mathcal{B}}(\mathcal{C}\mathcal{B})^{-1}\underbrace{\mathcal{C}\mathcal{V}}_{=0}+\mathcal{B}\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathcal{C}\mathcal{B})^{-1}\right)\underbrace{\mathcal{C}\mathcal{V}}_{=0}\\&\qquad+\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\dot{\mathcal{C}}\mathcal{V}+\left(\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\right)A\mathcal{V}\Bigg]\end{aligned}$$

which gives

$$\begin{aligned}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\mathcal{V}&=\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)\left((\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}\right)\mathcal{V}\\&\quad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{C})\mathcal{V}.\end{aligned}$$

Invoking

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)\left((\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\right)\mathcal{V}\\&=\frac{\mathrm{d}}{\mathrm{d}t}\left((\mathcal{V}^T\mathcal{V})^{-1}\right)\mathcal{V}^T\mathcal{V}+(\mathcal{V}^T\mathcal{V})^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V}^T)\mathcal{V}\\&\quad+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\mathcal{V}\\&=-(\mathcal{V}^T\mathcal{V})^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V}^T\mathcal{V})(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{V}\\&\quad+(\mathcal{V}^T\mathcal{V})^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V}^T)\mathcal{V}+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\mathcal{V}\\&=-(\mathcal{V}^T\mathcal{V})^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V}^T)\mathcal{V}-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V})\\&\quad+(\mathcal{V}^T\mathcal{V})^{-1}\frac{\mathrm{d}}{\mathrm{d}t}(\mathcal{V}^T)\mathcal{V}+(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^TA\mathcal{V}\\&=-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V})\end{aligned}$$

we arrive at

$$\begin{aligned}&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\mathcal{V}\\&=-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V})\\&\quad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{C})\mathcal{V}\\&=-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V})\\&\quad-(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\begin{bmatrix}0\\ \vdots\\ 0\\ \left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^r(C)\mathcal{V}\end{bmatrix}\end{aligned}$$

$$
\begin{aligned}
&= -(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V}) - (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{B} \\
&\quad\cdot\begin{bmatrix} * & & \Gamma^{-1} \\ & \cdot^{\cdot^{\cdot}} & \\ (-1)^{r-1}\Gamma^{-1} & & 0\end{bmatrix}\cdot\begin{bmatrix} 0 \\ \vdots \\ 0 \\ \left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^r(C)\mathcal{V}\end{bmatrix} \\
&= -(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V}) \\
&\quad - (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[B,\ldots,\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^{r-1}(B)\right] \\
&\quad\cdot\begin{bmatrix}\Gamma^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^r(C)\mathcal{V}0 \\ \vdots \\ 0\end{bmatrix} \\
&= -(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T \\
&\quad\cdot\left[\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)(\mathcal{V}) - B\Gamma^{-1}\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)^r(\mathcal{C})\mathcal{V}\right].
\end{aligned}
$$

This proves (3.10).

Finally, we show (3.11) and (3.12) for $P := P_1$. First, (3.14) together with Remark 3.4 yields

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\left[\mathcal{B}(\mathcal{CB})^{-1},\mathcal{V}\right] = [P,0,\ldots,0] \qquad (3.20)$$

and hence (3.11) follows.

To see (3.12), note that (3.18) gives

$$
\begin{aligned}
&\left(\frac{\mathrm{d}}{\mathrm{d}t}+A_r\right)(\mathcal{N})\mathcal{B}(\mathcal{CB})^{-1}\begin{bmatrix} I_m \\ 0 \\ \vdots \\ 0\end{bmatrix} \\
&\quad= (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T \\
&\quad\quad\cdot\left[0,\ldots,0,\left[\mathcal{B}(\mathcal{CB})^{-1}\mathcal{C}-I_n\right]\left(\frac{\mathrm{d}}{\mathrm{d}t}-A\right)^r(B)\right] \\
&\quad\quad\cdot\begin{bmatrix} * \\ \vdots \\ * \\ (-1)^{r-1}\Gamma^{-1}\end{bmatrix}
\end{aligned}
$$

which proves (3.12). This completes the proof of the theorem.□

As a direct consequence of Proposition 3.6 we obtain the following corollary.

Corollary 3.6: Under the assumptions of Theorem 3.5 and using the same notation, the normal form (3.4) may be written as

$$
\begin{aligned}
y^{(r)} &= \sum_{i=1}^{r} R_i(t)y^{(i-1)} + S(t)\eta + \Gamma(t)u(t) \\
\dot{\eta} &= Q(t)\eta + P(t)y
\end{aligned}
$$

and, if in addition (1.6) is time-invariant

$$
\begin{aligned}
y^{(r)} &= \sum_{i=1}^{r} R_i y^{(i-1)} + S\eta + CA^{r-1}B\,u(t) \\
\dot{\eta} &= \mathcal{N}A\mathcal{V}\eta + (-1)^{r-1}(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\mathcal{C}A^rB(CA^{r-1}B)^{-1}y.
\end{aligned}
$$

IV. Zero Dynamics

For time-invariant linear (1.1) or nonlinear (1.3) systems, the zero dynamics can be read off from the normal form of the system; see [5, Sec. 4.2]. The normal form is of particular help for control objectives such as output feedback stabilization or tracking; see [5, Sec. 4]. However, for nonlinear systems the normal form is local, even if the relative degree holds globally, and only under additional assumptions the diffeomorphism which describes the coordinate change is globally defined. Although for time-varying linear systems with fixed global relative degree the linear coordinate transformation $\mathcal{U}$ yielding the normal form (3.4) is globally defined, it is only useful for control objectives such stabilization if $\mathcal{U}$ is bounded and has bounded inverse. The latter will be guaranteed by additional assumptions on the systems' data. However, without further assumptions, Theorem A.1—an extended version of Doležal's Theorem—ensures that $\mathcal{V}$ and its left inverse $(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T$ are bounded matrix functions and therefore the stability properties of the zero dynamics are equivalent to the stability properties of $\dot{\eta} = Q\eta$ [see (3.10)]. To be precise, we first define the zero dynamics of a linear time-varying systems.

Definition 4.1:

i) The system $\dot{x} = A(t)x$, for $A \in \mathcal{C}^1(\mathbb{R},\mathbb{R}^{n\times n})$, is called *uniformly asymptotically stable* on $[0,\infty)$ if, and only if, there exist $M,\lambda > 0$ such that

$$\forall\, t \geq t_0 \geq 0 : \|x(t)\| \leq Me^{-\lambda(t-t_0)}\,\|x(t_0)\|$$

for any solution $x \in \mathcal{C}^1(\mathbb{R}_{\geq 0},\mathbb{R}^n)$ of $\dot{x} = A(t)x$.

ii) The *zero dynamics of the system* (1.6) on the open set $\mathcal{T} \subset \mathbb{R}$ are defined as the real vector space of trajectories

$$
\begin{aligned}
\mathcal{ZD}_{\mathcal{T}}(A,B,C) := \big\{&(x,u) \in \mathcal{C}^1(\mathcal{T},\mathbb{R}^n)\times\mathcal{C}^1(\mathcal{T},\mathbb{R}^m) \\
&\big|(x,u) \text{ solves (1.6) with } y \equiv 0 \text{ on } \mathcal{T}\big\}.
\end{aligned}
$$

Also for time-varying systems, as known for time-invariant systems, the zero dynamics can be read off the normal form (3.4). This is shown in the following proposition. In fact, the zero dynamics can be parameterized.

Proposition 4.2: Let $r,\ell \in \mathbb{N}$ with $r \leq \ell$ and $\mathcal{T} \subset \mathbb{R}$ be an open set. Then for any system (1.6) with relative degree $r \leq \ell$ on $\mathcal{T}$ and normal form (3.4) the following holds:

$$
\begin{aligned}
&\mathcal{ZD}_{\mathcal{T}}(A,B,C) = \\
&\big\{(\mathcal{V}\eta,-\Gamma^{-1}S\eta) \in \mathcal{C}^1(\mathcal{T},\mathbb{R}^n)\times\mathcal{C}^1(\mathcal{T},\mathbb{R}^m)\big|\,\dot{\eta} = Q\eta\big\}. \qquad (4.1)
\end{aligned}
$$

Proof: Set

$$\mathcal{Z} = \big\{(\mathcal{V}\eta,-\Gamma^{-1}S\eta) \in \mathcal{C}^1(\mathcal{T},\mathbb{R}^n)\times\mathcal{C}^1(\mathcal{T},\mathbb{R}^m)\big|\,\dot{\eta} = Q\eta\big\}.$$

"$\subseteq$": If $(x,u) \in \mathcal{ZD}_{\mathcal{T}}(A,B,C)$, then on $\mathcal{T}$ we have that $y = 0$ and so

$$\xi = \left(y^T,\ldots,\left(y^{(r-1)}\right)^T\right)^T = 0$$

which yields, in view of (3.4)

$$x = \mathcal{V}\eta = \mathcal{U}^{-1}\begin{pmatrix}0\\ \eta\end{pmatrix} \quad\text{and}\quad 0 = S\eta + \Gamma u$$

and therefore, $(x,u) \in \mathcal{Z}$.

"$\supseteq$": If $(\tilde{x}, \tilde{u}) = (\mathcal{V}\eta, -\Gamma^{-1}S\eta) \in \mathcal{Z}$, then on $\mathcal{T}$ we have that

$$\tilde{y} := C\tilde{x} = C\mathcal{V}\eta = 0$$

and so

$$\tilde{\xi} = \left(\tilde{y}^T, \ldots, \left(\tilde{y}^{(r-1)}\right)^T\right)^T = 0$$

and therefore $\left(\binom{0}{\eta}, \tilde{u}\right)$ solves the first equation in (3.4) with $\tilde{y} = 0$, and it follows that

$$(\tilde{x}, \tilde{u}) = \left(\mathcal{U}^{-1}\begin{pmatrix}0\\ \eta\end{pmatrix}, \tilde{u}\right) = (\mathcal{V}\eta, \tilde{u}) \in \mathcal{ZD}_{\mathcal{T}}(A, B, C).$$

This completes the proof of the proposition. □

In the remainder of this section we study stability of the zero dynamics.

Definition 4.3: For any $t_0 \in \mathbb{R}$ and $\ell \in \mathbb{N}$ consider (1.6) on $\mathcal{T} = (t_0, \infty)$. Then the zero dynamics of (1.6) are called *uniformly asymptotically stable* if, and only if, there exist $M, \lambda > 0$ such that

$$\forall\, (x, u) \in \mathcal{ZD}_{\mathcal{T}}(A, B, C) \quad \forall\, t \geq t_0 \geq 0 \quad : \|x(t)\| \leq M e^{\lambda(t-t_0)} \|x(t_0)\|.$$

Although the time-varying coordinate transformation $\mathcal{U}$ which converts (1.6) to (3.4) may be unbounded or its inverse may be unbounded, we will show in the following proposition that—surprisingly—the zero dynamics of (1.6) are uniformly asymptotically stable if, and only if, $\dot{\eta} = Q\eta$ is a uniformly asymptotically stable system.

Proposition 4.4: Let $r, \ell \in \mathbb{N}$ with $r \leq \ell$. Suppose that the system (1.6) has relative degree $r \leq \ell$ on $\mathbb{R}_{\geq 0}$ and consider its normal form (3.4). Then the zero dynamics of (1.6) are uniformly asymptotically stable if, and only if, $\dot{\eta} = Q(t)\eta$ is uniformly asymptotically stable system.

Proof: The matrix function $\mathcal{V}$, as noted in Remark 3.4, satisfies $\mathcal{V} \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times(n-rm)})$ and $(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{(n-rm)\times n})$. Now the claim of the theorem is a consequence of the fact that if $(x, u) \in \mathcal{ZD}_{\mathbb{R}_{\geq 0}}(A, B, C)$, then Proposition 4.2 yields

$$x = \mathcal{V}\eta \quad \text{and} \quad \eta = (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T x.$$

□

In the following, we show how the above findings can be exploited to design a derivative feedback controller for time-varying systems, provided some boundedness assumptions are satisfied. Due to brevity, we restrict our attention to the single-input–single-output case.

Theorem 4.5: Let $r, \ell \in \mathbb{N}$ with $r \leq \ell$. Suppose the system (1.6) is single-input–single-output (i.e. $m = 1$), has relative degree $r \leq \ell$ on $\mathbb{R}_{\geq 0}$, has uniformly asymptotically stable zero dynamics, and the matrix functions (A, B, C), $\mathcal{B}$, $\mathcal{C}$, $(\mathcal{CB})^{-1}$ are bounded on $\mathbb{R}_{\geq 0}$. Then for any Hurwitz polynomial $(s \mapsto s^r + \sum_{i=0}^{r-1} k_i s^i) \in \mathbb{R}[s]$, there exists $\kappa^* \geq 1$ such that, for all $\kappa \geq \kappa^*$, the feedback

$$u(t) = -\Gamma(t)^{-1} \sum_{i=0}^{r-1} \kappa^{r-i} k_i y^{(i)}(t) \tag{4.2}$$

applied to (1.6) yields a uniformly asymptotically stable closed-loop system.

Proof: Choose $\mathcal{U}$ as in Remark 3.4. Then the assumptions on boundedness guarantee, in view of Remark 3.4, that $\mathcal{U}, \mathcal{U}^{-1} \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times n})$.

Since boundedness of $(\mathcal{CB})^{-1}$ together with (3.19) gives, in particular, boundedness of Γ^{-1}, it follows that (4.2) is well-defined.

Therefore, it remains to show that the closed-loop system (3.4), (4.2), i.e.,

$$\left.\begin{aligned}
\dot{\xi}_i &= \xi_{i+1} \qquad \text{for } i = 1, \ldots, r-1\\
\dot{\xi}_r &= \sum_{i=0}^{r-1} (R_{i+1} - \kappa^{r-i} k_i)\xi_{i+1} + S\eta\\
\dot{\eta} &= Q\eta + P\xi_1
\end{aligned}\right\} \tag{4.3}$$

is uniformly asymptotically stable.

Setting

$$\zeta_i(t) := \kappa^{-i+1}\xi_i(t) \quad \text{for} \quad i = 1, \ldots, r \quad \text{and} \quad t \geq 0$$

yields

$$\begin{aligned}
\dot{\zeta}_i &= \kappa\zeta_{i+1} \text{ for } i = 1, \ldots, r-1\\
\dot{\zeta}_r &= \sum_{i=0}^{r-1} (\kappa^{-r+i+1} R_{i+1} - \kappa k_i)\zeta_{i+1} + \kappa^{-r+1} S\eta
\end{aligned}$$

and therefore, (4.3) is equivalent, for

$$F_\kappa(t) := \begin{bmatrix} 0 & & \\ & \ddots & \\ & & 0 \\ \kappa^{-r+1}R_1(t) & \ldots & \kappa^0 R_r(t) \end{bmatrix}$$

$$\hat{A} := \begin{bmatrix} 0 & I & & \\ \vdots & \ddots & \ddots & \\ 0 & \ldots & 0 & I \\ -k_0 & & \ldots & -k_{r-1} \end{bmatrix}$$

to

$$\left.\begin{aligned}
\dot{\zeta} &= F_\kappa\zeta + \kappa\hat{A}\zeta + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \kappa^{-r+1}S\eta \end{bmatrix}\\
\dot{\eta} &= Q\eta + P\zeta_1
\end{aligned}\right\}. \tag{4.4}$$

In passing, note that the boundedness assumptions of the theorem together with boundedness of $\mathcal{U}$ and $\mathcal{U}^{-1}$ yield boundedness of F_κ, independent of $\kappa \geq \kappa^* \geq 1$, $\hat{A}$, S, Q, and P. Since $s \mapsto \det(sI - \hat{A}) = s^r + \sum_{i=0}^{r-1} k_i s^i$ is Hurwitz and the zero dynamics are uniformly asymptotically stable, we may choose positive definite solutions $N_\zeta = N_\zeta^T \in \mathbb{R}^{r\times r}$ and $N_\eta = N_\eta^T \in \mathbb{R}^{(n-r)\times(n-r)}$ of

$$\left.\begin{aligned}
\hat{A}N_\zeta + N_\zeta\hat{A}^T &= I_r,\\
QN_\eta + N_\eta Q^T &= I_{n-r}, \text{ resp.}
\end{aligned}\right\} \tag{4.5}$$

Let $(\xi^T, \eta^T)^T$ denote an arbitrary solution of (4.4). Then, the derivative of

$$V(t) := \zeta(t)^T N_\zeta \zeta(t) + \eta(t)^T N_\eta \eta(t)$$

along (4.4) yields, in view of (4.5), for all $t \geq 0$ and omitting the arguments for brevity

$$\begin{aligned}\dot{V}(t) =& 2\zeta^T N_\zeta \left[F_\kappa \zeta + \kappa \hat{A}\zeta + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \kappa^{-r+1} S\eta \end{bmatrix} \right] \\ &+ 2\eta^T N_\eta [Q\eta + P\zeta_1] \\ \leq& -\kappa\|\zeta\|^2 - \|\eta\|^2 + 2\|N_\zeta F_\kappa\|\|\zeta\|^2 \\ &+ 2\kappa^{-r+1}\|N_\zeta S\|\|\zeta\|\|\eta\| + 2\|N_\eta P\|\|\eta\|\|\zeta_1\|\end{aligned}$$

and since $\kappa \geq \kappa^* \geq 1$, there exists $L > 0$ independent of κ such that

$$\begin{aligned}\dot{V}(t) \leq& -(\kappa - L)\|\zeta\|^2 - \|\eta\|^2 + L\sqrt{2}\|\zeta\|\|\eta\|\frac{1}{\sqrt{2}} \\ \leq& -(\kappa - L - 2L^2)\|\zeta\|^2 - \frac{1}{2}\|\eta\|^2.\end{aligned}$$

We conclude, for $\kappa - L - 2L^2 \geq 1/2$ and $\hat{L} := \min\{1/2\|N_\zeta\|, 1/2\|N_\eta\|\}$

$$\begin{aligned}\dot{V}(t) \leq& -\frac{1}{2}\|\zeta\|^2 - \frac{1}{2}\|\eta\|^2 \\ \leq& -\frac{1}{2\|N_\zeta\|}\zeta^T N_\zeta \zeta - \frac{1}{2\|N_\eta\|}\eta^T N_\eta \eta \\ \leq& -\hat{L}V(t).\end{aligned}$$

Therefore, for all $t \geq t_0 \geq 0$

$$V(t) \leq e^{-\hat{L}(t-t_0)} V(t_0)$$

where

$$\left\| \begin{bmatrix} \zeta(t) \\ \eta(t) \end{bmatrix} \right\| \leq \sqrt{\frac{\sigma_{\max}\left(\begin{pmatrix} N_\zeta & 0 \\ 0 & N_\eta \end{pmatrix} \right)}{\sigma_{\min}\left(\begin{pmatrix} N_\zeta & 0 \\ 0 & N_\eta \end{pmatrix} \right)}} \left\| \begin{bmatrix} \zeta(t_0) \\ \eta(t_0) \end{bmatrix} \right\|.$$

This proves uniform asymptotic stability of (4.4) and the proof of the theorem is complete. □

APPENDIX
DOLEŽAL'S THEOREM RE-REVISITED

Doležal's Theorem [1], which states convenient representations of range and kernels of time-varying matrices, has found numerous applications in systems theory and has been generalized and improved in various directions [2], [8], [10], [11]. In the following we give a generalization of [8, Th. 2] which is tailored for the needs of this paper.

Theorem A.1: Let $\ell \in \mathbb{N}$, $M \in \mathcal{C}^\ell(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times n})$ and suppose that there exists $r \in \{1,\dots,n\}$ such that, for all $t \geq 0$, $\operatorname{rank} M(t) = r$. Then, there exists $T \in \mathcal{C}^\ell(\mathbb{R}_{\geq 0}, \mathrm{Gl}_n(\mathbb{R}))$ such that

$$\forall\, t \geq 0 : M(t)T(t) = \left[F(t), 0_{n\times(n-r)}\right] \tag{A.1}$$

$$\exists \beta > 0\, \forall\, t \geq 0 : \|T(t)\| \leq \beta \tag{A.2}$$

$$\exists \varepsilon \in (0,1)\, \forall\, t \geq 0 : \varepsilon \leq T^T(t)T(t) \leq \frac{1}{\varepsilon} \tag{A.3}$$

where, obviously, $\operatorname{rank} F(t) = r$ for all $t \geq 0$.

Moreover, for any $m \in \{1,\dots,n\}$ and partition

$$T(t) = [X(t), V(t)] \quad X(t) \in \mathbb{R}^{n\times(n-m)} \quad V(t) \in \mathbb{R}^{n\times m} \tag{A.4}$$

we have

$$\exists \delta \in (0,1)\, \forall\, t \geq 0 : \delta \leq V^T(t)V(t) \leq \frac{1}{\delta} \tag{A.5}$$

and $V \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times m})$ with left inverse $(V^TV)^{-1}V^T \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{m\times n})$.

We preface the proof with the following lemma.

Lemma A.2: Let $\ell \in \mathbb{N}$ and $P \in \mathcal{C}^\ell(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times n})$ such that, for all $t \geq 0$, $P(t) = P(t)^T > 0$. Then the following statements are equivalent.

i) $\exists \varepsilon \in (0,1)\, \forall\, t \geq 0 : \varepsilon \leq \det P(t)$ and $\|P(t)\| \leq 1/\varepsilon$.
ii) $\exists \delta \in (0,1)\, \forall\, t \geq 0 : \delta \leq P(t) \leq 1/\delta$.
iii) $\exists \delta \in (0,1)\, \forall\, t \geq 0 : \delta \leq P(t)^{-1} \leq 1/\delta$.

Proof: Note that positivity of $P(t)$ yields

$$\begin{aligned}&\forall\, t \geq 0\, \exists U(t) \in \mathbb{R}^{n\times n} \text{ and } p_1(t),\dots,p_n(t) > 0 : \\ &\quad U^T(t)U(t) = I_n \text{ and} \\ &\quad P(t) = U(t)\operatorname{diag}(p_1(t),\dots,p_n(t))\, U^T(t)\end{aligned} \tag{A.6}$$

and, therefore

$$\begin{aligned}&\forall\, t \geq 0 : \det P(t) = \prod_{i=1}^n p_i(t) \\ &\quad\text{and} \quad \|\operatorname{diag}(p_1(t),\dots,p_n(t))\| = \|P(t)\|.\end{aligned} \tag{A.7}$$

Hence, we may conclude that

$$\begin{aligned}\text{i)} &\overset{(A.7)}{\iff} \exists \delta \in (0,1)\, \forall\, t \geq 0\, \forall\, i \in \{1,\dots,n\} : \\ &\qquad \delta \leq p_i(t) \leq 1/\delta \\ &\iff \exists \delta \in (0,1)\, \forall\, t \geq 0 : \\ &\qquad \delta \leq \operatorname{diag}(p_1(t),\dots,p_n(t)) \leq 1/\delta \\ &\overset{(A.6)}{\iff} \text{ii)} \\ &\iff \exists \delta \in (0,1)\, \forall\, t \geq 0 : \\ &\qquad \delta \leq \operatorname{diag}(p_1(t),\dots,p_n(t)) \leq 1/\delta \\ &\iff \exists \delta \in (0,1)\, \forall\, t \geq 0\, \forall\, i \in \{1,\dots,n\} : \\ &\qquad \delta \leq p_i(t) \leq 1/\delta \\ &\iff \exists \delta \in (0,1)\, \forall\, t \geq 0\, \forall\, i \in \{1,\dots,n\} : \\ &\qquad \delta \leq p_i(t)^{-1} \leq 1/\delta \\ &\overset{(A.6)}{\iff} \text{iii)}.\end{aligned}$$

This completes the proof of the lemma. □

Proof of Theorem A.1: In [8, Th. 2] it is shown that there exists $T \in \mathcal{C}^\ell(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times n})$ such that

$$\begin{aligned}&\exists \alpha \in (0,1)\, \forall\, t \geq 0 : M(t)T(t) = \left[F(t), 0_{n\times(n-r)}\right] \quad \text{and} \\ &\qquad \alpha \leq |\det T(t)| \quad \text{and} \quad \|T(t)\| \leq 1/\alpha.\end{aligned}$$

This yields (A.1) and (A.2), and furthermore

$$\begin{aligned}&\exists \alpha \in (0,1)\, \forall\, t \geq 0 : \alpha^2 \leq \det T^T(t)T(t) \\ &\qquad \text{and} \quad \left\|T^T(t)T(t)\right\| \leq 1/\alpha^2\end{aligned}$$

which implies, in view of Lemma A.2, (A.3).

Let, for $m \in \{1,\dots,n\}$, T be partitioned as in (A.4). The second inequality in (A.5) is a direct consequence of (A.2) and it remains to show the first inequality in (A.5). Seeking a contradiction, suppose that

$$\exists (\eta_i)_{i\in\mathbb{N}} \in (\mathcal{S}^{m-1})^{\mathbb{N}} \exists (t_i)_{i\in\mathbb{N}} \in (\mathbb{R}_{\geq 0})^{\mathbb{N}} \\ : \lim_{i\to\infty} \eta_i^T V^T(t_i) V(t_i) \eta_i = 0$$

where $\mathcal{S}^{m-1} := \{\eta \in \mathbb{R}^m \mid \|\eta\| = 1\}$. Then

$$\begin{aligned} 0 &= \lim_{i\to\infty} \eta_i^T V^T(t_i) V(t_i) \eta_i \\ &= \lim_{i\to\infty} \left(0_{1\times(n-m)}, \eta_i^T\right) \\ &\quad \cdot \begin{bmatrix} X^T(t_i)X(t_i) & X^T(t_i)V(t_i) \\ V^T(t_i)X(t_i) & V^T(t_i)V(t_i) \end{bmatrix} \begin{pmatrix} 0 \\ \eta_i \end{pmatrix} \\ &= \lim_{i\to\infty} \left(0_{1\times(n-m)}, \eta_i^T\right) T^T(t_i) T(t_i) \begin{pmatrix} 0_{n-m} \\ \eta_i \end{pmatrix} \end{aligned}$$

and this contradicts the first inequality in (A.3).

Finally, $V \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{n\times m})$ is immediate from (A.2), and (A.5) together with Lemma A.2 gives $V^T V \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{m\times m})$, and therefore, again by Lemma A.2, $(V^T V)^{-1} V^T \in \mathcal{L}^\infty(\mathbb{R}_{\geq 0}, \mathbb{R}^{m\times n})$. This completes the proof of the theorem. $\square$

ACKNOWLEDGMENT

We are indebted to P.C. Müller (Wuppertal) and J. Rudolph (Dresden) for pointing out to us early references from 1968–1971 where similar methods were used for the analysis of time-varying systems.

REFERENCES

[1] V. Doležal, "The existence of a continuous basis of a certain linear subspace of E_r, which depends on a parameter," *Časopis pro Pestovani Matematiky*, vol. 89, pp. 466–468, 1964.

[2] V. Doležal, *Dynamics of Linear Systems*. Prague, Czech Republic: Academic, 1967.

[3] E. Freund, "Über die synthese linearer zeitvarianter systeme," Ph.D. dissertation, Fak. für Elektrotechnik der Technischen Universität, Berlin, Germany, 1968.

[4] E. Freund, "Design of time-varying multivariable systems by decoupling and by the inverse," *IEEE Trans. Autom. Control*, vol. AC-16, no. 2, pp. 183–185, Apr. 1971.

[5] A. Isidori, *Nonlinear Control Systems*, 3rd ed. London, U.K.: Springer-Verlag, 1995.

[6] D. Liberzon, A. S. Morse, and E. D. Sontag, "Output-input stability and minimum-phase nonlinear systems," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 422–436, Mar. 2002.

[7] W. A. Porter, "Decoupling of and inverses for time-varying linear systems," *IEEE Trans. Autom. Control*, vol. AC-14, no. 4, pp. 378–380, Aug. 1969.

[8] P. Sen and M. R. Chidambara, "Doležal's theorem revisited," *Math. Syst. Theory*, vol. 13, pp. 67–79, 1979.

[9] L. M. Silverman, "Properties and application of inverse systems," *IEEE Trans. Autom. Control*, vol. AC-13, no. 4, pp. 436–437, Aug. 1968.

[10] L. M. Silverman and R. S. Bucy, "Generalizations of a theorem of Doležal," *Math. Syst. Theory*, vol. 4, pp. 334–339, 1970.

[11] L. Weiss and P. L. Falb, "Doležal's theorem, linear algebra with continously parametrized elements, and time-varying systems," *Math. Syst. Theory*, vol. 3, no. 1, pp. 67–75, 1969.

Achim Ilchmann received the Diploma from the University of Oldenburg, Oldenburg, Germany, in 1983, the Ph.D. degree from the University of Bremen, Bremen, Germany, in 1987, and the Habilitation from the University of Hamburg, Hamburg, Germany, in 1993, all in mathematics.

He has had teaching positions at Hamburg, Lübeck (both FRG), and Exeter (U.K.). Since 2000, he has been a Professor in Mathematics at the Technical University Ilmenau (FRG), Germany. His research interests include adaptive control of nonlinear systems, modelling and applications to bio-chemical reactions, and stability and robustness of time-varying linear systems.

Markus Mueller received the Diploma in mathematics from Ilmenau Technical University (TU Ilmenau), Germany, in 2005. Since 2006, he has been a TU Ilmenau Scholarship Holder and Graduate Assistant at TU Ilmenau where he is currently working toward the Ph.D. degree in mathematics.

His research interests include linear and nonlinear control theory, especially system analysis, controllability, and stability of time-varying systems.