Math. Control Signals Systems (1992) 5: 391–400

Mathematics of Control, Signals, and Systems

Continuity of Dynamical Systems: The Continuous-Time Case*

J. W. Nieuwenhuis† and J. C. Willems†

Abstract. The purpose of this paper is to study continuity of the parametrization of continuous-time linear time-invariant differential systems having a finite-dimensional state space. We show that convergence of the behavior of such systems corresponds to convergence of the coefficients of a set of associated differential equations. For this to hold, both the behavior and the convergence need to be appropriately defined.

Key words. Linear systems, Continuity of systems, Differential systems, Parametrization, Polynomial matrix description.

1. Introduction

This paper is a sequel to [NW] where we have studied the continuous parametrization of discrete-time linear systems. For easy reference, we repeat the main result of [NW]. The class of dynamical systems considered there consists of those defined by $\Sigma = (\mathbb{Z}, \mathbb{R}^q, \mathscr{B})$ with $\mathscr{B}$ a closed linear shift-invariant subspace of $(\mathbb{R}^q)^{\mathbb{Z}}$, equipped with the topology of pointwise convergence. We denote this family of dynamical systems (or their behavior) by $\mathscr{L}_d^q$ (q for the number of variables and d for discrete-time). A basic result [W] is that $\mathscr{B} \in \mathscr{L}_d^q$ if and only if it is the kernel of a polynomial operator in the shift: $\mathscr{B} = \ker R(\sigma, \sigma^{-1})$ for some $R(s, s^{-1}) \in \mathbb{R}^{\cdot \times q}[s, s^{-1}]$ (the family of polynomial matrices with q columns). Stated otherwise, $\mathbb{R}^{\cdot \times q}[s, s^{-1}]$ is a parametrization of $\mathscr{L}_d^q$. Let $\mathbb{R}_f^{\cdot \times q}[s, s^{-1}]$ denote the elements of $\mathbb{R}^{\cdot \times q}[s, s^{-1}]$ with full row rank. Now $\mathscr{B} \in \mathscr{L}_d^q$ if and only if $\mathscr{B} = \ker R(\sigma, \sigma^{-1})$ for some $R(s, s^{-1}) \in \mathbb{R}_f^{\cdot \times q}[s, s^{-1}]$. Let $R_\varepsilon(s, s^{-1}) \in \mathbb{R}_f^{\cdot \times q}[s, s^{-1}]$ for $\varepsilon \geq 0$.

Define the convergence $R_\varepsilon \xrightarrow[\varepsilon \to 0]{} R_0$ if there exists an L (independent of ε) such that the coefficient matrices of the same power of s converge, as $\varepsilon \to 0$, to the corresponding one of R_0, and if the coefficient matrix of s^ℓ is zero for all $\varepsilon > 0$ when $|\ell| \geq L$. Take $\mathscr{B}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathscr{B}_0$ to mean the following:

(i) $\mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon$, $\varepsilon > 0$, and $\mathbf{w} \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$ (convergence in the topology of pointwise convergence) implies $\mathbf{w}_0 \in \mathscr{B}_0$ and
(ii) for each $\mathbf{w}_0 \in \mathscr{B}_0$ there exists $\mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon$ such that $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$.

* Date received: October 29, 1990.

† Systems and Control Group, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands.

Finally, let us define the *lag* associated with an element $\Sigma \in \mathscr{L}_d^q$. For each $\mathscr{B} \in \mathscr{L}_d^q$ there exists a minimal number $L(\mathscr{B})$ such that $\mathscr{B} = \ker R(\sigma, \sigma^{-1})$ with $L(\mathscr{B})$ the degree of $R(s, s^{-1})$ (the degree is defined as the difference between the highest and lowest power with a nonzero coefficient). For simplicity we denote $L_\varepsilon := L(\mathscr{B}_\varepsilon)$.

In [NW] we proved that the parametrization $\mathbb{R}_f^{\cdot\times q}[s, s^{-1}]$ of $\mathscr{L}_d^q$ is a continuous one in the sense of the following theorem:

Theorem 1.

1. *Assume that* $R_\varepsilon(s, s^{-1}) \in \mathbb{R}_f^{g\times q}[s, s^{-1}]$, $\varepsilon \geq 0$, *satisfy* $R_\varepsilon \xrightarrow[\varepsilon\to 0]{} R_0$. *Let* $\mathscr{B}_\varepsilon := \ker R_\varepsilon(\sigma, \sigma^{-1})$. *Then* $\mathscr{B}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathscr{B}_0$.
2. *Assume that* $\mathscr{B}_\varepsilon \in \mathscr{L}_d^q$, $\varepsilon \geq 0$, *satisfy* $\mathscr{B}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathscr{B}_0$ *and that there exists an* $L \in \mathbb{Z}_+$ *such that* $L_\varepsilon \leq L$ *for all* ε *sufficiently small. Then there exists* $R_\varepsilon(s, s^{-1}) \in \mathbb{R}_f^{\cdot\times q}[s, s^{-1}]$, $\varepsilon \geq 0$, *such that* $R_\varepsilon \xrightarrow[\varepsilon\to 0]{} R_0$ *and* $\mathscr{B}_\varepsilon = \ker R_\varepsilon(\sigma, \sigma^{-1})$.

The purpose of this paper is to study the analogous problem for continuous-time systems. The results obtained are very similar to the discrete-time case, but due to the nature of continuous-time systems the statement of the results as well as the proofs are much more technical in nature. Because of the intrinsic importance of continuous-time systems in engineering and physics, we feel that it is important to have these results available in the literature. Our exposition is rather brief since much of the paper is an adaptation of [NW]. However, both the methods and the techniques used in the proofs depart in essential ways from the discrete-time case.

2. Linear Time-Invariant Differential Systems

We study continuous-time linear time-invariant dynamical systems $\Sigma = (\mathbb{R}, \mathbb{R}^q, \mathscr{B})$ with, as indicated, time-axis $\mathbb{R}$, signal space $\mathbb{R}^q$, and with behavior $\mathscr{B} \subseteq (\mathbb{R}^q)^{\mathbb{R}}$ described by a finite number of linear differential equations. That is, we assume that there exists a polynomial matrix $R(s) \in \mathbb{R}^{\cdot\times q}[s]$ such that $\mathbf{w} \in \mathscr{B}$ if and only if $R(d/dt)\mathbf{w} = \mathbf{0}$, i.e., if and only if $\mathbf{w} \in \ker R(d/dt)$. We have to clarify what this means. In fact, we may consider various choices for this kernel but we need to take an appropriate choice, partly dictated by what we are able to prove later. We consider the following class of solutions of the set of differential equations $R(d/dt)\mathbf{w} = \mathbf{0}$ as the behavior $\mathscr{B}$:

$$\left\{\mathbf{w} \in \ker R\left(\frac{d}{dt}\right)\right\} \quad \text{if and only if } \{\mathbf{w} \in \mathscr{C}^\infty(\mathbb{R}; \mathbb{R}^q),$$

$$\mathbf{w} \text{ is of class } \mathscr{E}, \text{ and } R\left(\frac{d}{dt}\right)\mathbf{w} = \mathbf{0}\right\}.$$

We say that $\mathbf{w} \in \mathscr{C}^\infty(\mathbb{R}; \mathbb{C}^q)$ is of class $\mathscr{E}$ if for all $n \in \mathbb{Z}_+$ there exists $\alpha \in \mathbb{R}$ such that $d^n\mathbf{w}/dt^n \in L_1^\alpha(\mathbb{R}; \mathbb{C}^q)$ with $L_1^\alpha(\mathbb{R}; \mathbb{C}^q) := \{\mathbf{f}: \mathbb{R} \to \mathbb{C}^q \mid \int_{-\infty}^{\infty} \|\mathbf{f}(t)\| e^{-\alpha|t|}\, dt < \infty\}$. We denote the family of dynamical systems (or their behavior thus obtained) by $\mathscr{L}_c^q$ (q for the number of variables, c for continuous-time). The requirement that $\mathbf{w}$ be of class $\mathcal{E}$ and in particular be of exponential growth is a bit annoying, and we would like (but were unable) to prove the main result without this assumption. It is, however, very similar to what Hazewinkel [H] had to impose in an analogous context.

It is easy to see (for example from the Smith form) that $\mathcal{B} \in \mathcal{L}_c^q$ if and only if there exists a $R(s) \in \mathbb{R}_f^{\cdot \times q}[s]$ (the full row rank polynomial matrices with q columns) such that $\mathcal{B}$ will be described by $R(s)$. This implies that $\mathbb{R}_f^{\cdot \times q}[s]$ is a parametrization of $\mathcal{L}_c^q$. Moreover, if $R_1(s)$ and $R_2(s)$ are in $\mathbb{R}_f^{\cdot \times q}[s]$, then $\ker R_1(d/dt) = \ker R_2(d/dt)$ if and only if there exists a unimodular polynomial matrix $U(s)$ such that $R_2(s) = U(s)R_1(s)$.

Let us next consider families of such systems and define their convergence. We consider dependence on a real parameter $\varepsilon \geq 0$ (see, however, remark 4.4). Let $\mathcal{B}_\varepsilon \in \mathcal{L}_c^q$, for all $\varepsilon \geq 0$. Then we define convergence $\mathcal{B}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathcal{B}_0$ to mean

(i) $\mathbf{w}_\varepsilon \in \mathcal{B}_\varepsilon$ and $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$ imply $\mathbf{w}_0 \in \mathcal{B}_0$, and
(ii) $\mathbf{w}_0 \in \mathcal{B}_0$ implies that there exist $\mathbf{w}_\varepsilon \in \mathcal{B}_\varepsilon$ such that $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$.

It remains to define what we mean by $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$. We say that $\mathbf{w}_\varepsilon$ converges to $\mathbf{w}_0$ as $\varepsilon \to 0$, written as $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathbf{w}_0$, if, for all $n \in \mathbb{Z}_+$,

$$\frac{d^n \mathbf{w}_\varepsilon}{dt^n}(t) \xrightarrow[\varepsilon \to 0]{} \frac{d^n \mathbf{w}_0}{dt^n}(t),$$

uniformly on bounded subsets of $\mathbb{R}$.

We now define convergence of polynomial matrices. Denote $R_\varepsilon(s) = R_\varepsilon^{L_\varepsilon} s^{L_\varepsilon} + R_\varepsilon^{L_\varepsilon - 1} s^{L_\varepsilon - 1} + \cdots + R_\varepsilon^1 s + R_\varepsilon^0$. Then we say that $R_\varepsilon(s) \xrightarrow[\varepsilon \to 0]{} R_0(s)$ if

(i) there exists $L \in \mathbb{Z}_+$ such that $L_\varepsilon \leq L$ for $\varepsilon > 0$ sufficiently small.
(ii) $R_\varepsilon^k \xrightarrow[\varepsilon \to 0]{} R_0^k$ for all k.

Now, for each $\mathcal{B} \in \mathcal{L}_c^q$ there exists a minimal L such that $\mathcal{B}$ is described by $R(s)$ with $R(s) \in \mathbb{R}_f^{\cdot \times q}[s]$ of degree L. We denote this minimal L as $L(\mathcal{B})$ and call it the *differential order* of $(\mathbb{R}, \mathbb{R}^q, \mathcal{B})$. L_ε is shorthand for $L(\mathcal{B}_\varepsilon)$.

3. The Main Result

The main result of this paper is stated as follows:

Theorem 2.

1. *Assume that* $R_\varepsilon(s) \in \mathbb{R}_f^{\cdot \times q}[s]$, $\varepsilon \geq 0$, *and that* $R_\varepsilon \xrightarrow[\varepsilon \to 0]{} R_0$. *Let* $\mathcal{B}_\varepsilon = \ker R_\varepsilon(d/dt)$. *Then* $\mathcal{B}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathcal{B}_0$.
2. *Assume that* $\mathcal{B}_\varepsilon \in \mathcal{L}_c^q$, *for all* $\varepsilon \geq 0$, *and that for some* $L \in \mathbb{Z}_+$ *we have* $L_\varepsilon \leq L$, *for all* $\varepsilon \geq 0$. *Then* $\mathcal{B}_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mathcal{B}_0$ *implies the existence of polynomial matrices* $R_\varepsilon(s) \in \mathbb{R}_f^{\cdot \times q}[s]$ *such that* $\mathcal{B}_\varepsilon = \ker R_\varepsilon(d/dt)$ *and such that* $R_\varepsilon \xrightarrow[\varepsilon \to 0]{} R_0$.

The proof of our main result uses the following lemmas, some of which we believe to be important in their own right.

Let $R(s) \in \mathbb{R}_f^{\cdot \times q}[s]$. Define the *McMillan degree* of R, $\mathbf{n}(R)$, to be the smallest degree of all the $q \times q$ minors of $R(s)$. If $R(s) \in \mathbb{R}_f^{q \times q}[s]$, then $n(R) = degree$ (det R). For M a matrix with entries in $\mathbb{C}$, take $M^* := \overline{M}^{\mathrm{T}}$, where $\overline{M}$ is the complex conjugate of M and the superscript T denotes transpose.

Lemma 1. *Let $P(s) \in \mathbb{R}_f^{q \times q}[s]$ be such that* $\det P(-\lambda) = 0$ *for some $\lambda \in \mathbb{C}$. Let $a \in \mathbb{C}^q$ have norm 1, and be such that $a^*P(-\lambda) = 0$. Then there is a unique polynomial matrix $\tilde{P}(s)$, with $\mathbf{n}(\tilde{P}) = \mathbf{n}(P) - 1$ such that*

$$P(s) = [I + aa^*(s - 1 + \lambda)]\tilde{P}(s). \tag{1}$$

Next, assume that $0 \neq \mu \in \mathbb{C}$ and $\det P(-1/\mu) = 0$. *Let $b \in \mathbb{C}^q$ have norm 1 and be such that $b^*P(-1/\mu) = 0$. Then there is a unique polynomial matrix $\hat{P}(s)$ with $\mathbf{n}(\hat{P}) = \mathbf{n}(P) - 1$ such that*

$$P(s) = [I + \mu bb^*s]\hat{P}(s). \tag{2}$$

Proof. Define the rational matrices $\tilde{P}(s)$ and $\hat{P}(s)$ by formulae (1) and (2). We now prove that these matrices are polynomials.

1. Let $V \in \mathbb{C}^{q \times q}$ be such that $VV^* = I$ and such that $Va = e := (1, 0, \ldots, 0)^{\mathrm{T}} \in \mathbb{C}^q$. Then $VP(s)V^* = [I + ee^{\mathrm{T}}(s - 1 + \lambda)]V\tilde{P}(s)V^*$. Notice that $e^{\mathrm{T}}VP(s)V^* = a^*P(s)V^*$ is the first row of $VP(s)V^*$ and is equal to zero for $s = -\lambda$. Therefore this row is divisible by $s + \lambda$ and hence $V\tilde{P}(s)V^*$ is a polynomial matrix and consequently so is $\tilde{P}(s)$.
2. Let $W \in \mathbb{C}^{q \times q}$ be such that $WW^* = I$ and such that $Wb = e$. Then $WP(s)W^* = [I + \mu ee^{\mathrm{T}}s]W\tilde{P}(s)W^*$. Therefore

$$W\hat{P}(s)W^* = \begin{bmatrix} (1 + \mu s)^{-1} & 0 & \cdots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} WP(s)W^*.$$

The first row of $WP(s)W^*$ is zero for $s = -1/\mu$ and hence is divisible by $1 + \mu s$. However, this implies that $W\tilde{P}(s)W^*$, and hence $\hat{P}(s)$ as well, is polynomial.

The following generalization to matrix polynomials of the factorization of scalar polynomials can be deduced from Lemma 1. Let $P(s) \in \mathbb{R}_f^{q \times q}[s]$ have determinant $p_n s^n + \cdots + p_0$ with $p_n \neq 0$, and roots $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$. Then $P(s)$ may be factored as

$$P(s) = P_1(s)P_2(s) \cdots P_n(s) \tag{3}$$

with all the P_i's of McMillan degree 1 and with $\det P_i(\lambda_i) = 0$.

Lemma 2 considers polynomial matrices depending on a parameter and the behavior of the factorization in McMillan degree 1 factors which was obtained in Lemma 1.

Lemma 2. *Let $P_\varepsilon(s) \in \mathbb{R}_f^{q\times q}[s]$ and degree* (det P_ε) $= n$, *for all* $\varepsilon > 0$. *Assume further that $P_\varepsilon \xrightarrow[\varepsilon\to 0]{} P_0$. Then $P_\varepsilon(s)$, $\varepsilon \geq 0$, admits the following factorization*:

$$P_\varepsilon(s) = \prod_{i=1}^{k} [a_\varepsilon^i a_\varepsilon^{i*}(s - \lambda_\varepsilon^i) + (I - a_\varepsilon^i a_\varepsilon^{i*})] \cdot \prod_{i=k+1}^{n} [I - \mu_\varepsilon^i b_\varepsilon^i b_\varepsilon^{i*}] \cdot C_\varepsilon(s), \tag{4}$$

where $C_\varepsilon(s)$ is unimodular and the vectors a_ε^i and b_ε^i all have norm 1. Further, the following limits hold: $\lambda_\varepsilon^i \xrightarrow[\varepsilon\to 0]{} \lambda_0^i$, $i = 1, 2, \ldots, k$, and $\mu_\varepsilon^i \xrightarrow[\varepsilon\to 0]{} \mu_0^i$, $i = k+1, k+2, \ldots, n$, $\mu_0^i = 0$, $i = k+1, k+2, \ldots, n$. Moreover, every sequence $\{\varepsilon' \to 0\}$ will contain a subsequence $\{\varepsilon'' \to 0\}$ such that $a_{\varepsilon''}^i$, $b_{\varepsilon''}^i$, and *$C_{\varepsilon''}(s)$ converge. If, in addition, $q = 1$, then all the entities indexed by ε converge.*

Proof. The factorization is a consequence of Lemma 1. Since det $P_\varepsilon(s) \xrightarrow[\varepsilon\to 0]{}$ det $P_0(s)$, the convergence of the λ_ε^i's and μ_ε^i's follows. It remains to prove the convergence of $C_{\varepsilon''}(s)$. In order to do that, write $P_\varepsilon(s) = V_\varepsilon(s)C_\varepsilon(s)$, $\varepsilon \geq 0$. We may assume that $V_{\varepsilon''}(s) \xrightarrow[\varepsilon''\to 0]{} V_0(s)$. Since by assumption $P_\varepsilon(s) \xrightarrow[\varepsilon\to 0]{} P_0(s)$, it now easily follows that $C_{\varepsilon''}(\lambda) \xrightarrow[\varepsilon''\to 0]{} C_0(\lambda)$ for all but a finite number of points $\lambda \in \mathbb{C}$. From the proof of Lemma 1 it also follows that the degrees of the polynomials in $V_\varepsilon(s)$ and $C_\varepsilon(s)$ are uniformly bounded. This yields that $C_{\varepsilon''}(s) \xrightarrow[\varepsilon''\to 0]{} C_0(s)$. In order to see this, write a typical element of $C_\varepsilon(s)$ as $f_\varepsilon(s) = f_\varepsilon^n s^n + \cdots + f_\varepsilon^0$, for all $\varepsilon > 0$, with $f_\varepsilon^n \neq 0$. Now assume that $f_{\varepsilon''}(s) \not\xrightarrow[\varepsilon''\to 0]{} f_0(s)$. Then we may assume without loss of generality that there is a subsequence $\tilde\varepsilon \to 0$ and an integer $k \in \{0, 1, 2, \ldots, n\}$ such that $f_{\tilde\varepsilon}^k \xrightarrow[\tilde\varepsilon\to 0]{} +\infty$ and such that $(f_{\tilde\varepsilon}^k)^{-1} f_{\tilde\varepsilon}^i$ converges for $\tilde\varepsilon \to 0$ and for all $i \in \{0, 1, \ldots, n\}$. Let $\lambda \in \mathbb{C}$ be such that $f_{\varepsilon''}(\lambda) \xrightarrow[\varepsilon''\to 0]{} f_0(\lambda)$. This yields a contradiction since at the same time we will have that $(f_{\varepsilon''}^k)^{-1} f_{\varepsilon''}(\lambda) \xrightarrow[\varepsilon''\to 0]{} 0$ and that $(f_{\varepsilon''}^k)^{-1} f_{\tilde\varepsilon}(s)$ converges to a nonzero polynomial.

In the next lemma we consider some convolution integrals appearing as basic building blocks in the proof of our main result.

Lemma 3. *Let $\lambda \in \mathbb{C}$ and $a \in \mathbb{C}^q$, with $a^*a = 1$, be given. Let $\mathbf{u}(\cdot) \in L_1^{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^q)$ and $\mathbf{y}(\cdot) \in L_1^{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^q)$ be related by*:

1. $\mathbf{y}(t) = (I - aa^*)\mathbf{u}(t) + \int_0^t aa^* e^{-\lambda(t-\tau)}\mathbf{u}(\tau)\, d\tau$ *Then $\mathbf{u}(\cdot)$ and $\mathbf{y}(\cdot)$ defined by this expression obey the following differential equation*:

$$\left(I + aa^*\left(\frac{d}{dt} - 1 + \lambda\right)\right)\mathbf{y} = \mathbf{u}. \tag{5}$$

2. $\mathbf{y}(t) := (I - aa^*)\mathbf{u}(t) + \lambda^{-1}\int_{-\infty}^t aa^* e^{-\lambda^{-1}(t-\tau)}\mathbf{u}(\tau)\, d\tau$, *with $\mathbf{u}(\cdot) \in L_1^{\mathrm{loc}}(\mathbb{R}; \mathbb{R}^q)$ such that this infinite integral exists for all $t \in \mathbb{R}$. Then $\mathbf{u}(\cdot)$ and $\mathbf{y}(\cdot)$ related by this expression obey the following differential equation*:

$$\left(I + \lambda aa^*\frac{d}{dt}\right)\mathbf{y} = \mathbf{u}. \tag{6}$$

Proof. These conclusions follow from straightforward computations.

In the next two lemmas we study convolution integrals of the types introduced in Lemma 3, parametrized by $\varepsilon \geq 0$. We only consider the case $q = 1$ (scalar-valued functions) since the extension to the case $q > 1$ is easy.

We call a collection of functions $\mathbf{f}_\varepsilon: \mathbb{R} \to \mathbb{C}$, $\varepsilon \geq 0$, *boundedly uniformly converging* (buc) if:

(i) $\mathbf{f}_\varepsilon \in C^\infty(\mathbb{R}; \mathbb{C})$ for all $\varepsilon \geq 0$.
(ii) For all $n \in \mathbb{Z}_+$, there exists $\alpha \in \mathbb{R}$ and $C \in \mathbb{R}_+$ such that

$$\int_{-\infty}^{\infty} \left\| \frac{d\mathbf{f}_\varepsilon}{dt^n}(t) \right\| e^{-\alpha|t|}\, dt < C,$$

for all $\varepsilon \geq 0$.
(iii) For all $n \in \mathbb{Z}_+$, $(d^n\mathbf{f}_\varepsilon/dt^n)(t) \xrightarrow[\varepsilon \to 0]{} (d^n\mathbf{f}_0/dt)(t)$, uniformly on bounded intervals of $\mathbb{R}$.

Lemma 4. *Consider for $\varepsilon \geq 0$ the following scalar differential equation,*

$$\frac{d\mathbf{y}_\varepsilon}{dt} + \lambda_\varepsilon \mathbf{y}_\varepsilon = \mathbf{u}_\varepsilon, \tag{7}$$

and assume that $\lambda_\varepsilon \in \mathbb{C}$, $\lambda_\varepsilon \xrightarrow[\varepsilon \to 0]{} \lambda_0$. Let $\{\mathbf{u}_\varepsilon, \varepsilon \geq 0\}$ be buc. *Then there exists a solution $\hat{\mathbf{y}}_\varepsilon$ of this differential equation such that $\hat{\mathbf{y}}_\varepsilon$, $\varepsilon \geq 0$, is* buc.

Proof. Consider the following solution to (7):

$$\hat{\mathbf{y}}_\varepsilon(t) = \int_0^t e^{-\lambda_\varepsilon(t-\tau)}\mathbf{u}_\varepsilon(\tau)\, d\tau.$$

The result then follows from straightforward estimates.

Lemma 5. *Consider for $\varepsilon \geq 0$ the following scalar differential equation,*

$$\mu_\varepsilon \frac{d\mathbf{y}_\varepsilon}{dt} + \mathbf{y}_\varepsilon = \mathbf{u}_\varepsilon, \tag{8}$$

and assume that $\mu_\varepsilon \in \mathbb{C}_\varepsilon$, $\mu_\varepsilon \xrightarrow[\varepsilon \to 0]{} \mu_0 = 0$. Let $\mathbf{u}_\varepsilon$, $\varepsilon \geq 0$, be buc. *Then in each of the four following situations there exists a solution $\hat{\mathbf{y}}_\varepsilon$ to* (8) *such that $\hat{\mathbf{y}}_\varepsilon$, $\varepsilon \geq 0$, is* buc:

1. $\operatorname{Re} \mu_\varepsilon^{-1} \to \infty$.
2. $\operatorname{Re} \mu_\varepsilon^{-1} \to -\infty$.
3. $\operatorname{Re} \mu_\varepsilon^{-1} \to c$ *and* $\operatorname{Im} \mu_\varepsilon^{-1} \to \infty$.
4. $\operatorname{Re} \mu_\varepsilon^{-1} \to c$ *and* $\operatorname{Im} \mu_\varepsilon^{-1} \to -\infty$.

Proof. We only consider cases 1, 2, and 3, since 4 is similar to 3. In case 1 consider the following solution to (8):

$$\hat{\mathbf{y}}_\varepsilon(t) := \int_{-\infty}^t \mu_\varepsilon^{-1} \exp(\mu_\varepsilon^{-1}(t-\tau))\mathbf{u}_\varepsilon(\tau)\, d\tau, \qquad t \in \mathbb{R}. \tag{9}$$

In case 2 consider the solution

$$\hat{\mathbf{y}}_\varepsilon(t) := \int_t^\infty \mu_\varepsilon^{-1} \exp(-\mu_\varepsilon^{-1}(t-\tau))\mathbf{u}_\varepsilon(\tau)\,d\tau, \qquad t \in \mathbb{R}. \tag{10}$$

In case 3 define first $\mathbf{\Delta}_\varepsilon := \mathbf{y}_\varepsilon - \mathbf{u}_0$. The equation for $\mathbf{\Delta}_\varepsilon$ becomes

$$\mu_\varepsilon \frac{d\mathbf{\Delta}_\varepsilon}{dt} + \mathbf{\Delta}_\varepsilon = \mathbf{v}_\varepsilon \tag{11}$$

with $\mathbf{v}_\varepsilon := \mu_\varepsilon(d\mathbf{u}_0/dt) + \mathbf{u}_\varepsilon - \mathbf{u}_0$. Hence $\mathbf{v}_\varepsilon$, $\varepsilon \geq 0$, is buc and $\mathbf{v}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{0}$. Now consider the following solution to (11):

$$\hat{\mathbf{\Delta}}_\varepsilon(t) = \int_0^t \mu_0^{-1} \exp(-\mu_\varepsilon^{-1}(t-\tau))\mathbf{v}_\varepsilon(\tau)\,d\tau, \qquad t \in \mathbb{R}. \tag{12}$$

Let us now analyze case 1 in more detail. Integration by parts yields

$$\hat{\mathbf{y}}_\varepsilon(t) = \mathbf{u}_\varepsilon(t) - \int_{-\infty}^t \exp(-\mu_\varepsilon^{-1}(t-\tau))\frac{d\mathbf{u}_\varepsilon}{dt}(\tau)\,d\tau. \tag{13}$$

Differentiating and once more integrating by parts yields:

$$\frac{d^n\hat{\mathbf{y}}_\varepsilon(t)}{dt^n} = \int_{-\infty}^t \mu_\varepsilon^{-1} \exp(-\mu_\varepsilon^{-1}(t-\tau))\frac{d^n\mathbf{u}_\varepsilon}{dt^n}(\tau)\,d\tau, \tag{14}$$

$$\frac{d^n\hat{\mathbf{y}}_\varepsilon(t)}{dt^n} = \frac{d^n\mathbf{u}_\varepsilon}{dt^n} - \int_{-\infty}^t \exp(-\mu_\varepsilon^{-1}(t-\tau))\frac{d^{n+1}\mathbf{u}_\varepsilon}{dt^{n+1}}(\tau)\,d\tau. \tag{15}$$

Using (14) it follows that $\hat{\mathbf{y}}_\varepsilon$, $\varepsilon \geq 0$, is buc. This can be seen as follows. Since the sequence $\mathbf{u}_\varepsilon$, $\varepsilon \geq 0$, is buc, there is an $\alpha > 0$ and a $C \in \mathbb{R}$ such that

$$\int_{-\infty}^\infty |\mathbf{u}_\varepsilon(t)|e^{-\alpha|t|}\,dt < C \qquad \text{for all} \quad \varepsilon \geq 0.$$

Write $\mathrm{Re}(1/\mu_\varepsilon) = 1/\bar{\mu}_\varepsilon$, hence $1/\bar{\mu}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \infty$ in case 1. Take $\hat{\alpha} > \alpha$ and consider

$$\begin{aligned}
&\int_{-\infty}^\infty |\mathbf{y}(t)|e^{-\hat{\alpha}|t|}\,dt \\
&\leq \int_{-\infty}^\infty \exp\left(-\hat{\alpha}|t| - \frac{1}{\bar{\mu}_\varepsilon}t\right)\int_{-\infty}^t \exp\left(\frac{1}{\bar{\mu}_\varepsilon}t' + \alpha|t'| - \alpha|t'|\right)|\mathbf{u}_\varepsilon(t')|\,dt'\,dt \\
&\leq C\int_{-\infty}^\infty e^{(\alpha-\hat{\alpha})|t|}\,dt \qquad \text{for } \varepsilon \text{ sufficiently small.}
\end{aligned}$$

Applying the same argument to the derivatives of $\hat{\mathbf{y}}_\varepsilon$ and using (15) in a similar way to prove that $(d^n\hat{\mathbf{y}}_\varepsilon/dt^n)(t) \xrightarrow[\varepsilon\to 0]{} (d^n\mathbf{u}_0/dt^n)(t)$ it follows that $\hat{\mathbf{y}}_\varepsilon$, $\varepsilon \geq 0$, is buc.
Similar manipulations yield cases 2 and 3.

Proof of Theorem 2. 1. We need to show that $\mathscr{B}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathscr{B}_0$. This requires proving

(i) $\{\mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon, \varepsilon > 0, \mathbf{w}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{w}_0\} \Rightarrow \{\mathbf{w}_0 \in \mathscr{B}_0\}$, and
(ii) $\{\mathbf{w}_0 \in \mathscr{B}_0\} \Rightarrow \{\exists \mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon \text{ such that } \mathbf{w}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{w}_0\}$.

Part (i) is trivial. Part (ii) is proved by induction. Let $R_0(s) = [P_0(s), -Q_0(s)]$ with $\det P_0(s) \neq 0$ (this can always be achieved by permuting, if need be, the columns of $R_0(s)$). Let $R_\varepsilon(s) = [P_\varepsilon(s), -Q_\varepsilon(s)]$ and assume that degree $(\det P_\varepsilon(s)) \leq N$ for all $\varepsilon \geq 0$. Our proof is inductive in N. Let $N = 0$. Then $P_\varepsilon(s)$ is unimodular for $\varepsilon \geq 0$ and $\mathscr{B}_\varepsilon$ is described by

$$\mathbf{y}_\varepsilon = P_\varepsilon^{-1}\left(\frac{d}{dt}\right)Q_\varepsilon\left(\frac{d}{dt}\right)\mathbf{u}_\varepsilon.$$

Let

$$\begin{pmatrix}\mathbf{u}_0\\ \mathbf{y}_0\end{pmatrix} \in \mathscr{B}_0$$

and observe that

$$\begin{bmatrix}\mathbf{u}_0\\ P_\varepsilon^{-1}\left(\frac{d}{dt}\right)Q_\varepsilon\left(\frac{d}{dt}\right)\mathbf{u}_0\end{bmatrix} \xrightarrow[\varepsilon\to 0]{} \begin{bmatrix}\mathbf{u}_0\\ P_0^{-1}\left(\frac{d}{dt}\right)Q_0\left(\frac{d}{dt}\right)\mathbf{u}_0\end{bmatrix}.$$

Next, assume that if degree $(\det P_\varepsilon(s)) \leq N$, the following induction hypothesis holds: $P_0(d/dt)\mathbf{y}_0 = \mathbf{v}_0$ and $\mathbf{v}_\varepsilon$, $\varepsilon \geq 0$, buc, imply the existence of a collection $\mathbf{y}_\varepsilon$, $\varepsilon \geq 0$, buc, such that $P_\varepsilon(d/dt)\mathbf{y}_\varepsilon = \mathbf{v}_\varepsilon$. Now assume that degree $(\det P_\varepsilon(s)) = N + 1$, and write, following Lemma 1 or 2,

$$P_\varepsilon(s) = F_\varepsilon(s)P_0'(s) \qquad \text{with} \quad P_\varepsilon'(s) \xrightarrow[\varepsilon\to 0]{} P_0'(s) \quad \text{and} \quad \text{degree det } P_\varepsilon'(s)) = N$$

and

$$F_\varepsilon(s) = a_\varepsilon a_\varepsilon^*(s - \lambda_\varepsilon) + (I - a_\varepsilon a_\varepsilon^*), \qquad \lambda_\varepsilon \to \lambda_0, \qquad a_\varepsilon \to a_0 \neq 0,$$

or

$$F_\varepsilon(s) = \mu_\varepsilon b_\varepsilon b_\varepsilon^* s + I, \qquad \mu_\varepsilon \to 0, \quad b_\varepsilon \to b_0 \neq 0.$$

Now write the behavioral differential equation for $\mathscr{B}_\varepsilon$ as

$$F_\varepsilon\left(\frac{d}{dt}\right)P_\varepsilon'\left(\frac{d}{dt}\right)\mathbf{y}_\varepsilon = Q_\varepsilon\left(\frac{d}{dt}\right)\mathbf{u}_\varepsilon$$

and express it as

$$F_\varepsilon\left(\frac{d}{dt}\right)\mathbf{y}_\varepsilon' = Q_\varepsilon\left(\frac{d}{dt}\right)\mathbf{u}_\varepsilon, \tag{16}$$

$$P_\varepsilon'\left(\frac{d}{dt}\right)\mathbf{y}_\varepsilon = \mathbf{y}_\varepsilon'. \tag{17}$$

Notice that $\{\mathbf{y} \mid F_\varepsilon(d/dt)\mathbf{y} = \mathbf{0}\}$ is a one-dimensional vector space and that for all $\mathbf{y}_0$ such that $F_0(d/dt)\mathbf{y}_0 = \mathbf{0}$ there is a collection $\mathbf{y}_\varepsilon$, $\varepsilon \geq 0$, buc, such that $\mathbf{y}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{y}_0$ and $F_\varepsilon(d/dt)\mathbf{y}_\varepsilon = \mathbf{0}$. In order to see this, note that

$$\left(I + a_\varepsilon a_\varepsilon^*\left(\frac{d}{dt} - 1 + \lambda_\varepsilon\right)\right)a_\varepsilon e^{-\lambda_\varepsilon t} = 0 \qquad \text{for all} \quad t \in \mathbb{R}.$$

Let $\mathbf{y}_0$ and $\mathbf{u}_0$ be C^∞ functions satisfying

$$P_0\left(\frac{d}{dt}\right)\mathbf{y}_0 = Q\left(\frac{d}{dt}\right)\mathbf{u}_0. \tag{18}$$

Rewrite (18) as

$$P_0'\left(\frac{d}{dt}\right)\mathbf{y}_0 = \mathbf{y}_0',$$

$$F_0\left(\frac{d}{dt}\right)\mathbf{y}_0' = Q_0\left(\frac{d}{dt}\right)\mathbf{u}_0.$$

Take in (16) $\mathbf{u}_\varepsilon = \mathbf{u}_0$ for all $\varepsilon > 0$. By Lemmas 1–4 and the observation above about the kernels of F_ε there is a buc sequence $\mathbf{y}_\varepsilon' \xrightarrow[\varepsilon\to 0]{} \mathbf{y}'$ such that $F_\varepsilon(d/dt)\mathbf{y}_\varepsilon' = Q_\varepsilon(d/dt)\mathbf{u}_0$. Applying the induction hypothesis yields a sequence $\mathbf{y}_\varepsilon$, $\varepsilon > 0$, buc, such that $\mathbf{y}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{y}_0$ with $P_\varepsilon(d/dt)\mathbf{y}_\varepsilon = \mathbf{y}_\varepsilon'$.

We conclude that for all sequences $\{\varepsilon' \to 0\}$ there is a subsequence $\{\varepsilon'' \to 0\}$ such that there exist solutions such that $\mathbf{w}_\varepsilon'' \xrightarrow[\varepsilon''\to 0]{} \mathbf{w}_0$. Scrutinizing the proof we can however deduce the following:

Take an arbitrary collection of finite intervals I_i in $\mathbb{R}$ such that $\bigcup I_i = \mathbb{R}$, and define

$$\delta_i(\mathscr{B}_\varepsilon) := \inf_{\mathbf{w}_\varepsilon}\left\{\sup_t \|\mathbf{w}_\varepsilon(t) - \mathbf{w}_0(t)\|,\ t \in I_i,\ \mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon\right\}.$$

Then the proof (all the sequences are buc!) shows that $\lim_{\varepsilon\to 0}\delta_i(\mathscr{B}_\varepsilon) = 0$, for all i.

It is now straightforward to construct a sequence $\hat{\mathbf{w}}_\varepsilon \in \mathscr{B}_\varepsilon$ such that $\hat{\mathbf{w}}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{w}_0$.

2. The proof of this part is completely analogous to the proof of the main result in [NW] and is therefore be omitted.

4. Extensions

4.1. Whether our main result remains valid with the behavior ker $R(d/dt)$ defined as the C^∞ solutions, *without* the exponential growth conditions imposed by $\mathscr{E}$, or defined simply as distributions satisfying $R(d/dt)\mathbf{w} = \mathbf{0}$, remains a matter of conjecture. In this case we were unfortunately unable to prove that $\mathbf{w}_0 \in \mathscr{B}_0$ implies the existence of $\mathbf{w}_\varepsilon \in \mathscr{B}_\varepsilon$ such that $\mathbf{w}_\varepsilon \xrightarrow[\varepsilon\to 0]{} \mathbf{w}_0$. It is clear that a proof without involving class $\mathscr{E}$ will require different methods than those used here.

4.2. Consider the systems

$$P_\varepsilon\left(\frac{d}{dt}\right)\mathbf{y} = Q_\varepsilon\left(\frac{d}{dt}\right)\mathbf{u}, \qquad \mathbf{w} = \begin{bmatrix}\mathbf{u}\\ \mathbf{y}\end{bmatrix},$$

with $P_\varepsilon(s) \in \mathbb{R}^{p\times p}[s]$, $Q_\varepsilon(s) \in \mathbb{R}^{p\times m}[s]$, and $\det P_\varepsilon(s) \neq 0$, $\varepsilon > 0$. Now assume that $P_\varepsilon(s) \xrightarrow[\varepsilon\to 0]{} P_0(s)$ and $Q_\varepsilon(s) \xrightarrow[\varepsilon\to 0]{} Q_0(s)$ with $[P_0(s), Q_0(s)] \in \mathbb{R}_f^{p\times(p+m)}[s]$. Then three things may happen:

1. det $P_0(s) \neq 0$ and $P_0^{-1}(s)Q_0(s)$ is strictly proper.
2. det $P_0(s) \neq 0$ and $P_0^{-1}(s)Q_0(s)$ is not strictly proper.
3. det $P_0(s) = 0$.

According to our main results the behaviors converge in all three cases. What can we conclude about the behavior of the impulse response? In case 1 the impulse response converges in the sense of L_1^{loc}, in case 2 in the sense of distributions. In case 3, however, the impulse response from $\mathbf{u}$ to $\mathbf{y}$ is not really defined for the limit system $\varepsilon = 0$. This shows that our notion of convergence is distinct from what would have been obtained in an input–output approach, as for instance adopted by Hazewinkel in [H], where systems are viewed as input–output mappings, with signals starting at time zero. Despite the difference in our approach, the function-spaces and notions of convergence used here are very similar to those considered by Hazewinkel.

4.3. The results of [NW] and our main result remain valid when instead of $\varepsilon \in \mathbb{R}$ we take a vector consisting of a finite number of parameters $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$, and let $\varepsilon \to \mathbf{0}$.

4.4. From the proofs it follows that we could enlarge the solution space of $R(d/dt)\mathbf{w} = \mathbf{0}$ without losing our continuity results. In particular, we only need to require convergence of a finite number of derivatives. However, this number will depend on the matrices $R_\varepsilon(s)$, and so we have chosen to state the results as we did.

4.5. For a treatment of continuity of dynamical systems with latent variables, see [WN].

References

[H] M. Hazewinkel, On families of linear systems: degeneration phenomena, *Algebraic and Geometric Methods in Linear Systems Theory* (C. I. Byrnes and C. F. Martin, eds.), pp. 157–189, Lectures in Applied Mathematics, Vol. 18, American Mathematical Society, Providence, RI, 1980.

[NW] J. W. Nieuwenhuis and J. C. Willems, Continuity of dynamical systems: a system theoretic approach, *Math. Control Signals Systems*, **1** (1988), 147–165.

[W] J. C. Willems, Paradigms and puzzles in the theory of dynamical systems, *IEEE Trans. Automat. Control*, **36** (1991), 259–294.

[WN] J. C. Willems and J. W. Nieuwenhuis, Continuity of latent variable models, *IEEE Trans. Automat. Control*, **36** (1991), 528–538.