2021

Continuity of Chen-Fliess Series for Applications in System Identification and Machine Learning

Rafael Dahmen

W. Steven Gray
Old Dominion University, sgray@odu.edu

Alexander Schmeding

Original Publication Citation

Dahmen, R., Gray, W. S., & Schmeding, A. (2021). Continuity of Chen-Fliess series for applications in system identification and machine learning. *IFAC-PapersOnLine*, *54*(9), 231-238. https://doi.org/10.1016/j.ifacol.2021.06.080

Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 54-9 (2021) 231–238

Continuity of Chen-Fliess Series for Applications in System Identification and Machine Learning ⋆

Rafael Dahmen * **W. Steven Gray** ** **Alexander Schmeding** ***

* *Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany, (e-mail: rafael.dahmen@kit.edu)*
** *Old Dominion University, Norfolk, Virginia 23529 USA, (e-mail: sgray@odu.edu)*
*** *FLU, Nord university, Høgskoleveien 27, 7601 Levanger, Norway, (e-mail: alexander.schmeding@nord.no)*

Abstract: Model continuity plays an important role in applications like system identification, adaptive control, and machine learning. This paper provides sufficient conditions under which input-output systems represented by locally convergent Chen-Fliess series are *jointly* continuous with respect to their generating series and as operators mapping a ball in an $L_{\mathfrak{p}}$-space to a ball in an $L_{\mathfrak{q}}$-space, where $\mathfrak{p}$ and $\mathfrak{q}$ are conjugate exponents. The starting point is to introduce a class of topological vector spaces known as *Silva spaces* to frame the problem and then to employ the concept of a *direct limit* to describe convergence. The proof of the main continuity result combines elements of proofs for other forms of continuity appearing in the literature to produce the desired conclusion.

Keywords: nonlinear systems, Chen-Fliess series, topological vector spaces, system identification, machine learning
AMS subject classification: 93C10, 46A04, 93B30, 68T07

1. INTRODUCTION

In applications involving system identification, adaptive control, and machine learning, a stream of input-output data is continually processed over time to produce a sequence of parameter/weight estimates so that an assumed model's behavior matches that of the data source. In the context of control, for example, this usually means that the dynamics of the model should asymptotically approach those of the plant. This can fail to happen when the model is incompatible with the plant or the data stream contains insufficient information. A more subtle mode of failure is one where the model's dynamics do not depend *continuously* on the parameters. In which case, it is possible for the sequence of parameter estimates to converge to some limit, while the corresponding sequence of approximations of the model's dynamics fail to converge in any sense.

The earliest work on the continuity of input-output systems was that of Hazewinkel (Hazewinkel, 1980). The focus there was on one parameter families of linear time-invariant systems and certain *degeneration* phenomena. Continuity of the same class of systems was later address from the behaviorial point of view in Nieuwenhuis and Willems (1988, 1992). Continuity of one parameter families of input-output systems with Chen-Fliess series representations (Fliess, 1981) was first characterized by Wang (1990). In this same work it was also shown that under certain growth conditions on the generating series such system are continuous as maps from $L_1[0,T]$ into $C[0,T]$ with the L_∞-norm for $T>0$ sufficiently small. More stringent growth conditions can even render an output function which is well defined and continuous on $[0,\infty)$ (Gray and Wang, 2002). Various improvements and generalizations of these result have appeared in Duffaut Espinosa (2009); Winter Arboleda (2019). In parallel with this development, continuity properties regarding control affine nonlinear state space models have appeared in Azhmyakov et al. (2009). The primary aim there was to characterize the continuity of flows with respect to the input and initial condition. Continuity with respect to the vector fields of the realization was not considered. As the coefficients of the corresponding Chen-Fliess depend explicitly on these vector fields and the initial condition, that analysis will not directly apply to the problems considered in this paper.

The main objective of this paper is provide sufficient conditions under which input-output systems represented by locally convergent Chen-Fliess series are *jointly* continuous with respect to their generating series and as operators mapping a ball in an $L_{\mathfrak{p}}$-space to a ball in an $L_{\mathfrak{q}}$-space, where $\mathfrak{p}$ and $\mathfrak{q}$ are conjugate exponents. Of course, continuity and convergence are ultimately topological concepts, so this phenomenon can only be understood precisely in a topological framework. The starting point is to introduce a class of topological vector spaces known as *Silva spaces* to frame the problem and then to employ the concept of a *direct limit* to describe convergence. The proof of the main continuity result combines elements of proofs for weaker forms of continuity appearing in Wang (1990), Gray and

⋆ The second author was supported by the National Science Foundation under grant CMMI-1839378.

10.1016/j.ifacol.2021.06.080

Wang (2002), and Duffaut Espinosa (2009) to produce the desired conclusion.

The paper is organized as follows. The next section gives a brief summary of the Chen-Fliess series mainly to establish the notation. The subsequent section describes the topological concepts used throughout the paper. The main continuity results appear in Section 4 along with some examples to illustrate their application. The final section summarizes the paper's main conclusions.

2. CHEN-FLIESS SERIES

An *alphabet* $X = \{x_0, x_1, \ldots, x_m\}$ is any nonempty and finite set of noncommuting symbols referred to as *letters*. A *word* $\eta = x_{i_1} \cdots x_{i_k}$ is a finite sequence of letters from X. The number of letters in a word η, written as $|\eta|$, is called its *length*. The empty word, $\emptyset$, is taken to have length zero. The collection of all words having length k is denoted by X^k. Define $X^* = \bigcup_{k\geq 0} X^k$ and $X^{\leq J} = \bigcup_{k=0}^{J} X^k$. The former is a monoid under the concatenation product. Any mapping $c : X^* \to \mathbb{R}^\ell$ is called a *formal power series*. Often c is written as the formal sum $c = \sum_{\eta\in X^*}(c,\eta)\eta$, where the *coefficient* (c,η) is the image of $\eta \in X^*$ under c. The *support* of c, $\mathrm{supp}(c)$, is the set of all words having nonzero coefficients. The set of all noncommutative formal power series over the alphabet X is denoted by $\mathbb{R}^\ell\langle\langle X\rangle\rangle$. The subset of series with finite support, i.e., polynomials, is represented by $\mathbb{R}^\ell\langle X\rangle$. Each set is an associative $\mathbb{R}$-algebra under the concatenation product and an associative and commutative $\mathbb{R}$-algebra under the *shuffle product*, that is, the bilinear product uniquely specified by the shuffle product of two words

$$(x_i\eta) ⧢ (x_j\xi) = x_i(\eta ⧢ (x_j\xi)) + x_j((x_i\eta) ⧢ \xi),$$

where $x_i, x_j \in X$, $\eta, \xi \in X^*$ and with $\eta ⧢ \emptyset = \emptyset ⧢ \eta = \eta$ (Fliess, 1981).

Given any $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ one can associate a causal m-input, ℓ-output operator, F_c, in the following manner. Let $\mathfrak{p} \geq 1$ and $t_0 < t_1$ be given. For a Lebesgue measurable function $u : [t_0, t_1] \to \mathbb{R}^m$, define $\|u\|_\mathfrak{p} = \max\{\|u_i\|_\mathfrak{p} : 1 \leq i \leq m\}$, where $\|u_i\|_\mathfrak{p}$ is the usual $L_\mathfrak{p}$-norm for a measurable real-valued function, u_i, defined on $[t_0, t_1]$. Let $L^m_\mathfrak{p}[t_0, t_1]$ denote the set of all measurable functions defined on $[t_0, t_1]$ having a finite $\|\cdot\|_\mathfrak{p}$ norm and $B^m_\mathfrak{p}(R_u)[t_0, t_1] := \{u \in L^m_\mathfrak{p}[t_0, t_1] : \|u\|_\mathfrak{p} \leq R_u\}$. Assume $C[t_0, t_1]$ is the subset of continuous functions in $L^m_1[t_0, t_1]$. Define inductively for each $\eta \in X^*$ the map $E_\eta : L^m_1[t_0, t_1] \to C[t_0, t_1]$ by setting $E_\emptyset[u] = 1$ and letting

$$E_{x_i\bar{\eta}}[u](t, t_0) = \int_{t_0}^{t} u_i(\tau)E_{\bar{\eta}}[u](\tau, t_0)\, d\tau,$$

where $x_i \in X$, $\bar{\eta} \in X^*$, and $u_0 = 1$. The *Chen-Fliess series* corresponding to c is

$$y(t) = F_c[u](t) = \sum_{\eta\in X^*} (c, \eta)\, E_\eta[u](t, t_0) \qquad (1)$$

(Fliess, 1981, 1983). It can be shown that if there exists real numbers $K, M \geq 0$ such that

$$|(c, \eta)| \leq KM^{|\eta|}\, |\eta|!, \quad \forall \eta \in X^* \qquad (2)$$

($|z| := \max_i |z_i|$ when $z \in \mathbb{R}^\ell$) then the series (1) converges absolutely and uniformly for sufficiently small $R, T > 0$ and constitutes a well defined mapping from $B^m_\mathfrak{p}(R)[t_0, t_0 + T]$ into $B^\ell_\mathfrak{q}(S)[t_0, t_0 + T]$, where the numbers $\mathfrak{p}, \mathfrak{q} \in [1, +\infty]$ are conjugate exponents, i.e., $1/\mathfrak{p} + 1/\mathfrak{q} = 1$ (Gray and Wang, 2002). Any such mapping is called a *locally convergent Fliess operator*.

A more refined convergence analysis of Chen-Fliess series appears in Winter Arboleda (2019) utilizing the notion of *Gevrey order*. A series $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ is said to have Gevrey order $s \in [0, \infty)$ if there exists constants $K, M > 0$ such that

$$|(c, \eta)| \leq KM^{|\eta|}(|\eta|!)^s, \quad \forall \eta \in X^*. \qquad (3)$$

Clearly, if c has Gevrey order s then it is also has Gevrey order s', where $s' > s$. Define for a given c the real number $\gamma_c = \min\{s \in [0, \infty) : s \text{ satisfies } (3)\}$ and the set of all generating series with minimum Gevrey order γ as $\mathbb{R}^\ell_\gamma\langle\langle X\rangle\rangle$. In this context, the set of all generating series for locally convergent Fliess operators as described above is

$$\mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle := \bigcup_{0\leq\gamma\leq 1} \mathbb{R}^\ell_\gamma\langle\langle X\rangle\rangle,$$

while a subset of series (note the upper bound on γ)

$$\mathbb{R}^\ell_{GC}\langle\langle X\rangle\rangle := \bigcup_{0\leq\gamma< 1} \mathbb{R}^\ell_\gamma\langle\langle X\rangle\rangle$$

can be shown to yield a type of *global convergence* on the extended space $L^m_{\mathfrak{p},e}(t_0)$ into $C[t_0, \infty)$, where

$$L^m_{\mathfrak{p},e}(t_0) := \bigcup_{T>0} L^m_\mathfrak{p}[t_0, t_0 + T]$$

(Winter Arboleda et al., 2015). Interestingly, this latter set of generating series does *not* constitute all of those that provide a globally defined Fliess operator as shown by example in Winter Arboleda (2019).

Finally, a Fliess operator F_c defined on $B^m_\mathfrak{p}(R)[t_0, t_0+T]$ is said to be *realizable* when there exists a state space model

$$\dot{z}(t) = g_0(z(t)) + \sum_{i=1}^{m} g_i(z(t))\, u_i(t), \quad z(t_0) = z_0 \qquad (4a)$$

$$y_j(t) = h_j(z(t)), \quad j = 1, 2, \ldots, \ell, \qquad (4b)$$

where each g_i is an analytic vector field expressed in local coordinates on some neighborhood $\mathcal{W}$ of z_0, and each output function h_j is an analytic function on $\mathcal{W}$ such that (4a) has a well defined solution $z(t)$, $t \in [t_0, t_0+T]$ for any given input $u \in B^m_\mathfrak{p}(R)[t_0, t_0+T]$, and $y_j(t) = F_{c_j}[u](t) = h_j(z(t))$, $t \in [t_0, t_0 + T]$, $j = 1, 2, \ldots, \ell$. It can be shown that for any word $\eta = x_{i_k} \cdots x_{i_1} \in X^*$

$$(c_j, \eta) = L_{g_\eta} h_j(z_0) := L_{g_{i_1}} \cdots L_{g_{i_k}} h_j(z_0), \qquad (5)$$

where $L_{g_i} h_j$ is the *Lie derivative* of h_j with respect to g_i.

3. TOPOLOGICAL SUBSPACES OF $\mathbb{R}^\ell\langle\langle X\rangle\rangle$

Suppose a sequence of generating series $\{c_j\}_{j\geq 1}$ is produced in real-time by processing a stream of input-output data in some manner. The corresponding sequence of Chen-Fliess series is taken to be $\{F_{c_j}\}_{j\geq 1}$. If the estimation or learning algorithm producing these generating series ensures that $c_j \to c$ in some sense, then it is desirable that $F_{c_j} \to F_c$ in some fashion as well. Perhaps the most obvious way in which one series can approach another is in the ultrametric sense. Specifically, for any fixed real number σ such that $0 < \sigma < 1$, consider the mapping

$$\mathrm{dist} : \mathbb{R}\langle\langle X\rangle\rangle \times \mathbb{R}\langle\langle X\rangle\rangle \to \mathbb{R},$$

$$(c, d) \mapsto \sigma^{\mathrm{ord}(c-d)},$$

where $\mathrm{ord}(c)$ is the length of the shortest word in the support of c ($\mathrm{ord}(0) := \infty$). The $\mathbb{R}$-vector space $\mathbb{R}^\ell\langle\langle X\rangle\rangle$ with mapping dist is known to be a complete ultrametric space (Berstel and Reutenauer, 1988). If each series $c_j \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$, the following simple example illustrates that in the limit there is not always a well defined operator to which a given sequence of locally convergent Fliess operators is converging.

Example 1. Let $X = \{x_1\}$ and consider the sequence of polynomials

$$c_j = x_1 + (2!)^2\, x_1^2 + (3!)^2\, x_1^3 + \cdots + (j!)^2\, x_1^j,\ \ j \geq 1.$$

Clearly, each polynomial c_j is locally convergent. Thus, each Fliess operator F_{c_j} is well defined on some ball of input functions in $L_{\mathfrak{p}}^m[t_0, t_1]$. Furthermore, the sequence $(c_j)_j$ converges to $c = \sum_{k\geq 1}(k!)^2 x_1^k$ in the ultrametric topology. But the limiting Chen-Fliess series F_c is not well defined in any obvious sense.

This example motivates the following fundamental problem: On what topological subspaces of $\mathbb{R}^\ell\langle\langle X\rangle\rangle$ does $c_j \to c$ imply that $F_{c_j} \to F_c$ in some sense with the limit point F_c being a well defined operator? The following subsections lay the foundation for addressing this problem by presenting what subspaces are available for consideration.

3.1 Fixed $M > 0$ (Banach Spaces)

As a first step, consider the following interpretation of condition (2). Fix $M > 0$ and define

$$\|c\|_{\ell_{\infty,M}} := \sup\left\{\frac{|(c,\eta)|}{M^{|\eta|}\,|\eta|!} : \eta \in X^*\right\} \in [0, \infty]$$

for each $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$. The set of all c with $\|c\|_{\ell_{\infty,M}} < \infty$ is denoted by $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$. It is straightforward to check that $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ is a vector subspace of $\mathbb{R}^\ell\langle\langle X\rangle\rangle$. The function $\|\cdot\|_{\ell_{\infty,M}}$ is a norm on $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$. The following assignment is an isometry of normed spaces:

$$\ell_{\infty,M}(X^*, \mathbb{R}^\ell) \longrightarrow \ell_\infty(X^*, \mathbb{R}^\ell) : c \mapsto \frac{c}{M^{|\eta|}\,|\eta|!},$$

where $\ell_\infty(X^*, \mathbb{R}^\ell) := \left\{c\colon X^* \to \mathbb{R}^\ell : \sup_\eta |(c,\eta)| < \infty\right\}$ is the Banach space of all *bounded* functions from X^* to $\mathbb{R}^\ell$. This shows that for each fixed $M > 0$ the space $(\ell_{\infty,M}(X^*, \mathbb{R}^\ell), \|\cdot\|_{\ell_{\infty,M}})$ is a Banach space. A series $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ belongs to $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ if and only if the bound (2) holds for some $K \geq 0$ and the fixed number M. In fact, the norm $\|c\|_{\ell_{\infty,M}}$ is the smallest number $K \geq 0$ such that (2) is satisfied.

As $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ is a Banach space, and, in particular, a metric space, the topology of $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ can be recovered from convergent sequences, where a sequence $(c_j)_j$ in $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ converges to $c \in \ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ if and only if

$$\lim_{j\to\infty} \|c_j - c\|_{\ell_{\infty,M}} = 0.$$

Given that $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ is an infinite dimensional Banach space, the Bolzano-Weierstrass theorem fails to hold, i.e., not every $\|\cdot\|_{\ell_{\infty,M}}$-bounded sequence has a $\|\cdot\|_{\ell_{\infty,M}}$-convergent subsequence (Werner, 2000, Satz I.2.7). Furthermore, the space is not separable, i.e., there is no countable dense subset. Given M_1 and M_2 such that $M_1 \leq M_2$, it is clear that $\|\cdot\|_{\ell_{\infty,M_1}} \geq \|\cdot\|_{\ell_{\infty,M_2}}$, and thus the inclusion (as vector spaces)

$$\ell_{\infty,M_1}(X^*, \mathbb{R}^\ell) \subseteq \ell_{\infty,M_2}(X^*, \mathbb{R}^\ell)$$

holds. This inclusion is not a topological embedding as the topology induced by $\ell_{\infty,M_2}(X^*, \mathbb{R}^\ell)$ is coarser than the one induced by $\ell_{\infty,M_1}(X^*, \mathbb{R}^\ell)$. It turns out for $M_1 < M_2$ that the inclusion map

$$\ell_{\infty,M_1}(X^*, \mathbb{R}^\ell) \to \ell_{\infty,M_2}(X^*, \mathbb{R}^\ell)$$

is a *compact operator* (Dahmen and Schmeding, 2018, Lemma B.6), i.e., it maps bounded sets to relatively compact sets. In particular, this shows for $M_1 < M_2$ that every sequence which is bounded in the $\|\cdot\|_{\ell_{\infty,M_1}}$-norm has a subsequence which converges in the coarser $\|\cdot\|_{\ell_{\infty,M_2}}$-topology.

3.2 The projective limit $M \to 0$ (Fréchet–Schwartz Spaces)

Consider next those $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ for which $\|c\|_{\ell_{\infty,M}}$ is finite for all $M > 0$. This means that for each $M > 0$ there is a $K = \|c\|_{\ell_{\infty,M}} \geq 0$ satisfying (2). Algebraically, this corresponds to the *intersection* of all vector spaces $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$, namely,

$$\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell) := \bigcap_{M>0} \ell_{\infty,M}(X^*, \mathbb{R}^\ell).$$

On spaces like these, there is a natural topology which turns this space into a locally convex topological vector space. In the functional analysis literature, this object is called the *projective limit* (or *inverse limit* or *categorical limit*) of the system $\left(\ell_{\infty,M}(X^*, \mathbb{R}^\ell)\right)_{M>0}$ and denoted also by

$$\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell) := \varprojlim_{M\to 0}\, \ell_{\infty,M}(X^*, \mathbb{R}^\ell) = \bigcap_{M>0} \ell_{\infty,M}(X^*, \mathbb{R}^\ell).$$

For a given $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, one can check whether it belongs to this space in the following way:

$$c \in \ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell) \iff \|c\|_{\ell_{\infty,M}} < \infty,\ \forall M > 0. \tag{6}$$

The sequence $M_k = 1/k$, $k \in \mathbb{N}$ is cofinal, hence it suffices to check (6) only for M of the form M_k.

Now $\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell)$ is the projective limit of countably many Banach spaces. Thus it becomes a *Fréchet space*, i.e., a complete metrisable space. Fréchet spaces share many nice properties with Banach spaces. For example, their topology is determined by sequences, where a sequence $(c_j)_j$ in $\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell)$ converges to $c \in \ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell)$ if and only if

$$\lim_{j\to\infty} \|c_j - c\|_{\ell_{\infty,M}} = 0,\ \forall M > 0.$$

(Again it suffices to check this only for all $M = 1/k$, $k \in \mathbb{N}$.) Since the inclusion maps are all compact operators, $\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell)$ is even a Fréchet–Schwartz space, (Pérez Carreras and Bonet, 1987, Definition 8.5.2). Hence, it behaves much nicer than the Banach spaces from which it was built. In particular, the space $\ell_{\infty,\leftarrow}(X^*, \mathbb{R}^\ell)$ satisfies a version of the Bolzano-Weierstrass theorem, namely, every $\ell_{\infty,\leftarrow}$-bounded sequence has a $\ell_{\infty,\leftarrow}$-convergent subsequence. Here, a sequence $(c_j)_j$ is called $\ell_{\infty,\leftarrow}$-bounded if $\sup_j \|c_j\|_{\ell_{\infty,M}} < \infty$ for *all* $M > 0$. This follows from

(Pérez Carreras and Bonet, 1987, Proposition 8.5.9), which furthermore implies that $\ell_{\infty,\leftarrow}(X^*,\mathbb{R}^\ell)$ is separable, i.e., there is countable dense subset. In (Winter Arboleda, 2019, Theorem 3.4.5) it is shown that

$$\ell_{\infty,\leftarrow}(X^*,\mathbb{R}^\ell) = \overline{\mathbb{R}^\ell_{GC}\langle\langle X\rangle\rangle},$$

where the closure on the right is taken with respect to the $\ell_{\infty,\leftarrow}$-topology (called the *semi-norm topology* in loc. cit.). In other words, there are some generating series with minimum Gevrey order $\gamma = 1$ that yield globally defined Fliess operators.

3.3 The direct limit $M \to \infty$ (Silva Spaces)

Consider next a series $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ where there exists at least one number $M > 0$ such that $K = \|c\|_{\ell_{\infty,M}} \geq 0$ satisfies (2). Algebraically, this case corresponds to the *union* of all vector spaces $\ell_{\infty,M}(X^*,\mathbb{R}^\ell)$, that is,

$$\ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell) := \bigcup_{M>0} \ell_{\infty,M}(X^*,\mathbb{R}^\ell).$$

As with the intersection, there is also a natural topology turning this space into a locally convex topological vector space. This object is called the *direct limit* (or *inductive limit* or *categorical colimit*) of the system $\big(\ell_{\infty,M}(X^*,\mathbb{R}^\ell)\big)_{M>0}$ and denoted by

$$\begin{aligned}\ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell) &= \varinjlim_{M\to\infty} \ell_{\infty,M}(X^*,\mathbb{R}^\ell)\\ &= \bigcup_{M>0} \ell_{\infty,M}(X^*,\mathbb{R}^\ell).\end{aligned}$$

This construction can also be found in Bogfjellmo and Schmeding (2016); Dahmen and Schmeding (2018). For a given $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$, one can check whether it belongs to this space in the following way:

$$c \in \ell_{\infty,\leftarrow}(X^*,\mathbb{R}^\ell) \iff \exists M > 0 \text{ such that } \|c\|_{\ell_{\infty,M}} < \infty.$$

Since the sequence $M_k = k$, $k \in \mathbb{N}$ is cofinal, there exists an $M \in \mathbb{N}$ for which $\|c\|_{\ell_{\infty,M}} < \infty$. Thus, one could equivalently work only with $M \in \mathbb{N}$. In general, direct limits are more difficult to work with than projective limits. Fortunately, this particular direct limit is a countable direct limit of Banach spaces with compact operators as inclusion maps. Direct limit spaces like these are called *Silva spaces.*

Although Silva spaces are not metrizable, they are always *sequential* (Yoshinaga, 1957, Proposition 6). This means that as in the Banach space case, the topology is determined by sequences, i.e., sets are closed if and only if they are sequentially closed. A sequence $(c_j)_j$ in $\ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell)$ converges to $c \in \ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell)$ if and only if

$$\lim_{j\to\infty} \|c_j - c\|_{\ell_{\infty,M}} = 0 \text{ for } \textit{one fixed } M > 0.$$

In other words, a sequence in a Silva space converges if there exists one fixed $M > 0$ for the whole sequence such that $(c_j)_j$ converges in the Banach space $\ell_{\infty,M}(X^*,\mathbb{R}^\ell)$ (Yoshinaga, 1957, Theorem 1). In particular, note that for a sequence to converge, all terms must lie in one of the spaces $\ell_{\infty,M}(X^*,\mathbb{R}^\ell)$, i.e., one $M > 0$ has to work for the whole sequence. The sequence in Example 1 fails to converge in the Silva topology since there is no M for which the sequence is Cauchy.

Using again the compactness of the inclusion maps, it follows that a sequence which is bounded in one $\|\cdot\|_{\ell_{\infty,M_1}}$-norm (for a given $M_1 > 0$) has a subsequence which converges in the coarser $\|\cdot\|_{\ell_{\infty,M_2}}$-topology for all $M_2 > M_1$. As earlier, there is a version of the Bolzano-Weierstrass theorem, namely, every $\ell_{\infty,\rightarrow}$-bounded sequence has a $\ell_{\infty,\rightarrow}$-convergent subsequence. In this case, a sequence $(c_j)_j$ is called $\ell_{\infty,\rightarrow}$-bounded if there is at least one $M > 0$ with $\sup_j \|c_j\|_{\ell_{\infty,M}} < \infty$. Therefore, a Silva space has better topological properties than the Banach spaces from which it is constructed. Furthermore, every Silva space is separable. Finally, it is shown in (Winter Arboleda, 2019, Theorem 3.2.7) that

$$\ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell) = \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle.$$

Example 2. Let $X = \{x_1\}$. The sequence $c_j := j!\, x_1^j$, $j \in \mathbb{N}$ has norm $\|c_j\|_{\ell_{\infty,1}} = 1$. Therefore, $(c_j)_j$ does not converge to zero in $\ell_{\infty,1}(X^*,\mathbb{R})$. However, since $\|c_j\|_{\ell_{\infty,2}} = 1/2^j$, it follows that $c_j \xrightarrow{j\to\infty} 0$ in $\ell_{\infty,2}(X^*,\mathbb{R})$ and also in the Silva topology.

Example 3. Let $X = \{x_1\}$. Define for $n, j \in \mathbb{N}$ the sequence $d_{n,j} := n^{(5j-2)/2j} C_n$, where $C_n := (2n)!/((n+1)!n!)$ is the nth Catalan number.[1] Recall that the asymptotic growth of the Catalan numbers is $C_n \sim 4^n/(n^{3/2}\sqrt{\pi})$. Thus, for $d_j := \sum_{n=1}^\infty d_{n,j} x_1^n$, it is clear that $d_1 \in \ell_{\infty,4}(X^*,\mathbb{R})$, but $d_j \notin \ell_{\infty,4}(X^*,\mathbb{R})$ for $j > 1$. However, since $\|d_j\|_{\ell_{\infty,5}} < \infty$, it does hold that $(d_j)_j \subseteq \ell_{\infty,\rightarrow}(X^*,\mathbb{R})$. Furthermore, it is easily checked that $\lim_{j\to\infty} \|d_j - d\|_{\ell_{\infty,5}} = 0$, where $d = \sum_{n=1}^\infty n^{5/2} C_n x_1^n$. Thus, d_j, $j \in \mathbb{N}$ converges to zero in the Silva topology.

In the continuity theorems presented in the next section, every sequences $(c_j)_j$ will be assumed a priori to be entirely contained in some Banach space $\ell_{\infty,M}(X^*,\mathbb{R})$, $M > 0$, thus avoiding the phenomenon shown in the previous example. Therefore, only a Banach topology is really needed. However, for applications such as the interconnection of Chen-Fliess series, the Silva topology is more applicable in the corresponding continuity analysis. For example, one can define for generating series $c, d \in \ell_{\infty,M}(X^*,\mathbb{R}^\ell)$ a product $c \circ d$ such that the composition satisfies $F_c \circ F_d = F_{c\circ d}$. It is known that in general $c \circ d$ will not be contained in $\ell_{\infty,M}(X^*,\mathbb{R}^\ell)$ (Thitsa and Gray, 2012). However, there does exist a $K(M,N,R) < \infty$ for all $M, N, R \in [0,\infty[$ such that for $B_{\infty,N}(R) := \{c \in \ell_{\infty,N}(X^*,\mathbb{R}^\ell) : \|c\|_{\ell_{\infty,N}} < R\}$ the map

$$\circ\colon \ell_{\infty,M}(X^*,\mathbb{R}^\ell) \times B_{\infty,N}(R) \to \ell_{\infty,K(M,N,R)}(X^*,\mathbb{R}^\ell)$$

is well defined (Gray et al., 2021, Lemma 3.4). Using these estimates, the interconnection of Chen-Fliess series induces a continuous product on the Silva space $\ell_{\infty,\rightarrow}(X^*,\mathbb{R}^\ell)$. Hence, the Silva topology is the natural topology for describing the continuity of such interconnections. This is investigated further in Gray et al. (2021).

4. MAIN CONTINUITY THEOREMS

The continuity problem for a Chen-Fliess series $F_c[u]$ is approached incrementally. It is first assumed that the input u is fixed and the generating series c is variable (series to output continuity). Then the case where c is

[1] Sequence A000108 in OEIS (2021).

fixed and u is variable is presented (input-output operator continuity). Finally, the two cases are combined. For notational convenience, define the space of $L_\mathfrak{p}$-germs $L_\mathfrak{p}^m(t_0) := \{[u] \mid u \in L_\mathfrak{p}^m[t_0, t_1] \text{ for some } t_1 > t_0\}$, where the class $[u]$ contains all functions equal to u in some neighborhood of t_0. Note that this space can *not* be endowed with any useful topology making the inclusion $L_\mathfrak{p}^m[t_0, t_1]$ continuous (as this would automatically be non-Hausdorff).

Theorem 4. (series to output continuity). The map

$$L_\mathfrak{p}^m(0) \times \mathbb{R}_{LC}^\ell\langle\langle X\rangle\rangle \to L_\mathfrak{q}^\ell(0), \quad (u, c) \mapsto y = F_c[u]$$

is well defined. Moreover, for every $M > 0$ and fixed $u \in L_\mathfrak{p}^m(0)$, there exists a $T > 0$ such that

$$\ell_{\infty,M}(X^*, \mathbb{R}^\ell) \to L_\mathfrak{q}^\ell[0, T], \quad c \mapsto y = F_c[u]$$

is continuous.

The following two lemmas are needed for the proof (Duffaut Espinosa, 2009).

Lemma 5. Let $X = \{x_0, x_1, \ldots, x_m\}$. For any $k \in \mathbb{N}_0$, the characteristic polynomial $\mathrm{char}(X^k)$ of X^k, i.e. the sum of all words of length k, satisfies

$$\mathrm{char}(X^k) = \sum_{\substack{r_0, r_1, \ldots, r_m \geq 0 \\ r_0 + r_1 + \cdots + r_m = k}} x_0^{r_0} ⧢ x_1^{r_1} ⧢ \cdots ⧢ x_m^{r_m},$$

where ⧢ denotes the shuffle product.

Lemma 6. Let $X = \{x_0, x_1, \ldots, x_m\}$. For any $u \in L_1^m[0, T]$ and $\eta \in X^*$

$$|E_\eta[u](t)| \leq E_\eta[\bar{u}](t), \quad 0 \leq t \leq T,$$

where $\bar{u} \in L_1^m[0, T]$ has components $\bar{u}_j := |u_j|$, $j = 1, 2, \ldots, m$. Furthermore, for any integers $r_j \geq 0$ it follows that

$$\left| E_{x_0^{r_0} ⧢ x_1^{r_1} ⧢ \cdots ⧢ x_m^{r_m}}[u](t) \right| \leq \prod_{j=0}^{m} \frac{U_j^{r_j}(t)}{r_j!}, \quad 0 \leq t \leq T,$$

where $U_j(t) := \int_0^t |u_j(\tau)|\, d\tau$.[2] In particular, if on $[0, T]$ it is assumed that $\max\{\|u\|_1, T\} \leq R$ then

$$\left| E_{x_0^{r_0} ⧢ x_1^{r_1} ⧢ \cdots ⧢ x_m^{r_m}}[u](t) \right| \leq \frac{R^k}{r_0!\, r_1! \cdots r_m!}, \quad 0 \leq t \leq T,$$

where $k = \sum_j r_j$.

Now the proof of Theorem 4.

Proof: If $c \in \mathbb{R}_{LC}^\ell\langle\langle X\rangle\rangle$, then there exists $K, M \geq 0$ satisfying (2). Fix $u \in L_1^m(0)$ (without loss of generality $\mathfrak{p} = 1$ and $t_0 = 0$) so that for some $T > 0$, $u \in L_1^m[0, T]$. Define $R = \max\{\|u\|_1, T\}$. Applying Lemmas 5 and 6 it then follows that:

$$\begin{aligned}
|y(t)| &\leq \sum_{\eta \in X^*} |(c, \eta) E_\eta[u](t)| \\
&\leq \sum_{k=0}^{\infty} \sum_{\eta \in X^k} |(c, \eta)|\, E_\eta[\bar{u}](t) \\
&\leq \sum_{k=0}^{\infty} K M^k k! \sum_{\substack{r_0, r_1, \ldots, r_m \geq 0 \\ r_0 + r_1 + \cdots + r_m = k}} E_{x_0^{r_0} ⧢ x_1^{r_1} ⧢ \cdots ⧢ x_m^{r_m}}[\bar{u}](t) \\
&\leq \sum_{k=0}^{\infty} K M^k k! \sum_{\substack{r_0, r_1, \ldots, r_m \geq 0 \\ r_0 + r_1 + \cdots + r_m = k}} \frac{R^k}{r_0! r_1! \cdots r_m!}
\end{aligned}$$

[2] For notational convenience, occasionally F_p will denoted by E_p when $p \in \mathbb{R}\langle X\rangle$.

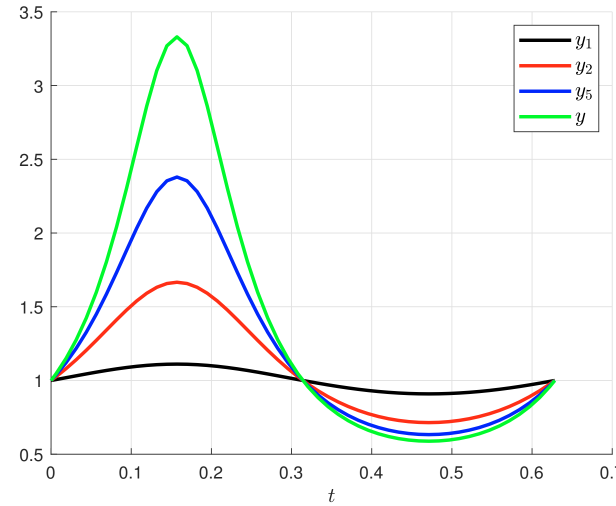

Fig. 1. Convergence of y_j to y in Example 7 when $M_a = 1$, $M_b = 7$, and $u(t) = \cos(10t)$ with $t \in [0, 0.2\pi]$.

$$\begin{aligned}
&= \sum_{k=0}^{\infty} K(MR)^k \sum_{\substack{r_0, r_1, \ldots, r_m \geq 0 \\ r_0 + r_1 + \cdots + r_m = k}} \frac{k!}{r_0! r_1! \cdots r_m!} \\
&= \sum_{k=0}^{\infty} K(MR(m+1))^k. \qquad (7)
\end{aligned}$$

Therefore, if $R < 1/(M(m+1))$, i.e., if

$$\max\{\|u\|_1, T\} < \frac{1}{M(m+1)}, \qquad (8)$$

then the series (1) converges absolutely and uniformly on $[0, T]$ so that y is well defined as an L_∞-germ, specifically, $y \in L_\infty^m[0, T]$. Now since the mapping $c \mapsto y = F_c[u]$ is linear, it is sufficient for proving continuity to show that it is bounded as a mapping from the Banach space $\ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ into the Banach space $L_\infty^\ell[0, T]$, i.e., $\|y\|_\infty / \|c\|_{\ell_{\infty,M}} < \infty$. Observe that $K = \|c\|_{\ell_{\infty,M}} > 0$ is a valid choice, and therefore, $\|y\|_\infty / K \leq 1/(1 - MR(m+1)) < \infty$ as claimed. ■

Example 7. Let $X = \{x_1\}$ and consider the sequence $c_j = \sum_{k \geq 0} M_j^k k!\, x_1^k$, $j \geq 1$ in $\mathbb{R}_{LC}\langle\langle X\rangle\rangle$, where

$$M_j = M_b \theta_j + M_a(1 - \theta_j)$$

and $\theta_j = (j-1)/j$. In which case, $c_1 = \sum_{k \geq 0} M_a^k k!\, x_1^k$ and define $c = \sum_{k \geq 0} M_b^k k!\, x_1^k$. To see that $c_j \xrightarrow{j \to \infty} c$ as a direct limit, set $M = \max\{M_a, M_b\}$ and observe that

$$\|c_j - c\|_{\ell_M^\infty} = \sup_{k \geq 0} \left| \frac{M_j^k - M_b^k}{M^k} \right| = \left| \frac{M_j - M_b}{M} \right|.$$

Thus, the conclusion follows directly. Now for the given M pick T and u to satisfy (8). Therefore,

$$\begin{aligned}
y_j(t) &= \sum_{k=0}^{\infty} M_j^k k! E_{x_1^k}[u](t) = \sum_{k=0}^{\infty} M_j^k k! \frac{E_{x_1}^k[u](t)}{k!} \\
&= \frac{1}{1 - M_j E_{x_1}[u](t)},
\end{aligned}$$

and likewise

$$y(t) = \frac{1}{1 - M_b E_{x_1}[u](t)},$$

are both well defined on $[0, T]$. In addition,

$$\|y_j - y\|_\infty = \sup_{t\in[0,T]} \left\| \frac{(M_j - M_b)E_{x_1}[u](t)}{(1-M_jE_{x_1}[u](t))(1-M_bE_{x_1}[u](t))} \right\|$$

so that $y_j \xrightarrow{j\to\infty} y$ in the $L_\infty[0,T]$ norm sense. The specific example where $M_a = 1$, $M_b = 7$, and $u(t) = \cos(10t)$ with $t \in [0, 0.2\pi]$ is shown in Figure 1. Here $T = 0.6283$, $\|u\|_1 = 0.4$ and $1/\max(M_a, M_b) = 0.1429$ (note $m+1 = 1$ as X has only one letter), which shows that condition (8) is conservative in this instance.

Input-output operator continuity is addressed in the next theorem. The proof is inspired by results appearing in Wang (1990) except that certain details have to be handled differently in order to use the result in the proof of the final continuity theorem.

Theorem 8. (input-output operator continuity) Suppose $c \in \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$ and select any pair of conjugate exponents $\mathfrak{p}, \mathfrak{q} \in [1,\infty]$. If $0 < T \le R < 1/(M(m+1))$ such that $c \in \ell_{\infty,M}(X^*, \mathbb{R}^\ell)$, then the operator

$$F_c\colon B^m_\mathfrak{p}(R)[t_0,\, t_0+T] \to B^\ell_\mathfrak{q}(S)[t_0, t_0+T]$$

for some $S > 0$ is continuous with respect to the $L_\mathfrak{p}$ and $L_\mathfrak{q}$ norms.

Proof: It needs to be shown for any $\epsilon > 0$ that there exists a $\delta > 0$ such that if $v, u \in B^m_\mathfrak{p}(R)[t_0,\, t_0+T]$ satisfy $\|v-u\|_\mathfrak{p} < \delta$, then $\|F_c[v] - F_c[u]\|_\mathfrak{q} < \epsilon$. It is first proved by induction on the length of the word $\eta \in X^*$ that the mapping

$$E_\eta : B^m_\mathfrak{p}(R)[t_0,\, t_0+T] \to B^\ell_\mathfrak{q}(S)[t_0, t_0+T]$$

has the desired continuity property. The focus is on the case where $\mathfrak{p}, \mathfrak{q} \in (1,\infty)$ (the remaining case is handled similarly). Without loss of generality, assume $t_0 = 0$. The claim is trivial when η is the empty word. If $\eta = x_i$, then

$$\begin{aligned}
\|E_{x_i}[v] - E_{x_i}[u]\|_\mathfrak{q} &= \left(\int_0^T |E_{x_i}[v](t) - E_{x_i}[u](t)|^\mathfrak{q}\, dt\right)^{\frac{1}{\mathfrak{q}}} \\
&\le \left(\int_0^T \left(\int_0^T |v_i(\tau) - u_i(\tau)|\, d\tau\right)^\mathfrak{q} dt\right)^{\frac{1}{\mathfrak{q}}} \\
&= \int_0^T |v_i(\tau) - u_i(\tau)|\, d\tau\, T^{1/\mathfrak{q}} \\
&\le \|v_i - u_i\|_p\, T^{2/\mathfrak{q}} \\
&\le \|v - u\|_p\, T^{2/\mathfrak{q}},
\end{aligned}$$

where Hölder's inequality has been used in the second to the last step above. Thus, if $\|v-u\|_p < \delta_{x_i} := \epsilon/T^{2/\mathfrak{q}}$, then clearly

$$\|E_{x_i}[v] - E_{x_i}[u]\|_\mathfrak{q} < \epsilon.$$

Now suppose the claim holds for all words up to some fixed length $k \ge 0$. Then for any $x_i \in X$ and $\eta \in X^k$ observe

$$\begin{aligned}
&\|E_{x_i\eta}[v] - E_{x_i\eta}[u]\|_\mathfrak{q} \\
&= \left\| \left(E_{x_i\eta}[v] - \int_0^\cdot u_i(\tau)E_\eta[v](\tau)\, d\tau\right) + \left(\int_0^\cdot u_i(\tau)E_\eta[v](\tau)\, d\tau - E_{x_i\eta}[u]\right)\right\|_\mathfrak{q} \\
&\le \left\|E_{x_i\eta}[v] - \int_0^\cdot u_i(\tau)E_\eta[v](\tau)\, d\tau\right\|_\mathfrak{q} + \left\|\int_0^\cdot u_i(\tau)E_\eta[v](\tau)\, d\tau - E_{x_i\eta}[u]\right\|_\mathfrak{q} \\
&\le \left(\int_0^T \left(\int_0^T |v_i(\tau) - u_i(\tau)|\,|E_\eta[v](\tau)|\, d\tau\right)^\mathfrak{q} dt\right)^{\frac{1}{\mathfrak{q}}} + \left(\int_0^T \left(\int_0^T |u_i(\tau)|\,|E_\eta[v](\tau) - E_\eta[u](\tau)|\, d\tau\right)^\mathfrak{q} dt\right)^{\frac{1}{\mathfrak{q}}} \\
&\le \int_0^T |v_i(\tau) - u_i(\tau)|\,|E_\eta[v](\tau)|\, d\tau\, T^{1/\mathfrak{q}} + \int_0^T |u_i(\tau)|\,|E_\eta[v](\tau) - E_\eta[u](\tau)|\, d\tau\, T^{1/\mathfrak{q}} \\
&\le \|v-u\|_\mathfrak{p} \|E_\eta[v]\|_\mathfrak{q} T^{1/\mathfrak{q}} + \|u\|_\mathfrak{p} \|E_\eta[v] - E_\eta[u]\|_\mathfrak{q}\, T^{1/\mathfrak{q}}.
\end{aligned}$$

From the induction hypothesis, E_η is continuous in the desired sense. Thus, it follows that for any $\epsilon > 0$, there exists a $\delta'_\eta > 0$ such that

$$\|E_\eta[v]\|_\mathfrak{q} \le \|E_\eta[u]\|_\mathfrak{q} + 1$$

and

$$\|u\|_\mathfrak{p} \|E_\eta[v] - E_\eta[u]\|_\mathfrak{q} T^{1/\mathfrak{q}} < \epsilon/2$$

for all v in a ball centered at u of radius $\delta'_\eta > 0$.[3] In which case, choose

$$\delta_{x_i\eta} = \min\left\{\delta'_\eta, \frac{\epsilon/2}{(\|E_\eta[u]\|_\mathfrak{q} + 1)T^{1/\mathfrak{q}}}\right\}$$

so that if $\|u - v\|_\mathfrak{p} < \delta_{x_i\eta}$, then

$$\|E_{x_i\eta}[v] - E_{x_i\eta}[u]\|_\mathfrak{q} < \epsilon.$$

Hence, by induction, E_η is continuous with respect to the $L_\mathfrak{p}$ and $L_\mathfrak{q}$ norms for every $\eta \in X^*$.

To show that F_c is also continuous in the desired sense, observe that for any integer $N > 0$

$$\begin{aligned}
\|F_c[v] - F_c[u]\|_\mathfrak{q} &= \left\|\sum_{k=0}^\infty \sum_{\eta\in X^k} (c,\eta)(E_\eta[v] - E_\eta[u])\right\|_\mathfrak{q} \\
&\le \left\|\sum_{k=0}^{N-1} \sum_{\eta\in X^k} (c,\eta)(E_\eta[v] - E_\eta[u])\right\|_\mathfrak{q} + \left\|\sum_{k=N}^{\infty} \sum_{\eta\in X^k} (c,\eta)(E_\eta[v] - E_\eta[u])\right\|_\mathfrak{q} \\
&\le \left\|\sum_{k=0}^{N-1} \sum_{\eta\in X^k} (c,\eta)(E_\eta[v] - E_\eta[u])\right\|_\mathfrak{q} + 2\sum_{k=N}^\infty K(MR(m+1))^k,
\end{aligned}$$

where $c \in \ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ and $K = \|c\|_{\ell_{\infty,M}} > 0$. Clearly the second term above can be bounded by $\epsilon/2$ by selecting N to be sufficiently large. Having done this, take $\delta := \min_{|\eta|\le N} \delta_\eta$, where the δ_η have been chosen as above to bound the first term by $\epsilon/2$. This establishes the continuity of the map to $L^\ell_\mathfrak{q}[t_0, t_0+T]$. Moreover, in light of (7), it follows that $\|F_c[u]\|_\mathfrak{q} \le \|c\|_{\ell_{\infty,M}} \sum_{n\ge 0}(MR(m+1))^n$, where the series is a convergent geometric series. Hence,

[3] Of course, δ'_η must be selected so that this ball is contained inside $B^m_\mathfrak{p}(R)[0,\,T]$. It is also being tacitly assumed that u is not on the boundary of $B^m_\mathfrak{p}(R)[0,\,T]$. Otherwise, this argument needs a few minor adjustments.

there exists a constant $S > 0$ depending only on $\|c\|_{\ell_{\infty,M}}$ bounding the $L_{\mathfrak{q}}$-norm of $F_c[u]$. ■

Now the stronger property of joint continuity is derived using some of concepts developed for the previous two theorems. First recall that for Banach spaces V, W and $U \subseteq V$ open, the following spaces are Banach spaces:

- $\mathrm{BC}(U, W)$ the space of bounded continuous functions with the supremum norm $\|\cdot\|_\infty$.
- $L(V, W)$ the space of bounded linear functions with the operator norm $\|\cdot\|_{\mathrm{op}}$.

Theorem 9. (joint continuity) Let $M \in \mathbb{R}^+$, $\mathfrak{p}, \mathfrak{q}$ be conjugate exponents, and $0 < T \leq R < 1/(M(m+1))$. The maps

$$\Phi\colon \ell_{\infty,M}(X^*, \mathbb{R}^\ell) \to \mathrm{BC}(B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T], L_{\mathfrak{q}}^\ell[t_0, t_0+T]),$$
$$c \mapsto F_c$$
$$\Psi\colon B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T] \to L(\ell_{\infty,M}(X^*, \mathbb{R}^\ell), L_{\mathfrak{q}}^\ell[t_0, t_0+T]),$$
$$u \mapsto (c \mapsto F_c[u])$$

are well defined and continuous. Therefore, the joint map

$$\ell_{\infty,M}(X^*, \mathbb{R}^\ell) \times B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T] \to L_{\mathfrak{q}}^\ell[t_0, t_0+T],$$
$$(c, u) \mapsto F_c[u] \tag{9}$$

is also continuous.

Proof: First observe from Theorem 8 that for every $c \in \ell_{\infty,M}(X^*, \mathbb{R}^\ell)$, F_c is bounded and continuous on $B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T]$. Hence, Φ is well defined and clearly linear. Furthermore, from (7),

$$\|\Phi(c)\|_\infty = \sup_{u \in B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T]} \|F_c[u]\|_{\mathfrak{q}}$$
$$\leq \sum_{k=0}^{\infty} \|c\|_{\ell_{\infty,M}} (MR(m+1))^k.$$

Suitably choosing R, the right-hand side will be bounded by a finite constant times the factor $\|c\|_{\ell_{\infty,M}}$. Hence, $\|\Phi\|_{\mathrm{op}} < \infty$, and Φ is continuous.

In light of Theorem 4, Ψ is also well defined. To see that Ψ is continuous, let $\epsilon > 0$ and fix $u, v \in B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T]$. Then the estimate in the proof of Theorem 8 yields

$$\|\Psi(u) - \Psi(v)\|_{\mathrm{op}} = \sup_{\|c\|_{\ell_{\infty,M}}=1} \|F_c[u] - F_c[v]\|_{\mathfrak{q}}$$
$$\leq \sum_{k=0}^{N-1} \sum_{\eta \in X^k} \underbrace{|(c,\eta)|}_{\leq M^k k!} \|E_\eta[v] - E_\eta[u]\|_{\mathfrak{q}} +$$
$$2 \sum_{k=N}^{\infty} (MR(m+1))^k.$$

Choosing N sufficiently large, the second term is smaller than $\epsilon/2$. It is known from the proof of Theorem 8 that every E_η is continuous, hence one can choose $\delta > 0$ such that the first term is less than $\epsilon/2$ if v is in the δ-ball around u. Therefore, Ψ is continuous.

The continuity of the joint map (9) follows directly from the continuity of Ψ by (Lang, 1999, I. §3 Proposition 3.10). However, it can also be easily derived from the continuity of Φ as shown next. Select $c_1, c_2 \in \ell_{\infty,M}(X^*, \mathbb{R}^\ell)$ and $u, v \in B_{\mathfrak{p}}^m(R)[t_0,\, t_0+T]$. Applying the triangle inequality, observe that $\|F_{c_1}[u] - F_{c_2}[v]\|_{\mathfrak{q}}$ is dominated by

$$\|\underbrace{F_{c_1}[u_1] - F_{c_2}[v]}_{=\Phi(c_1-c_2)(u)}\|_{\mathfrak{q}} + \|F_{c_1}[u] - F_{c_2}[v]\|_{\mathfrak{q}}$$
$$\leq \|\Phi\|_{\mathrm{op}} \|c_1 - c_2\|_{\ell_{\infty,M}} + \|F_{c_2}[u] - F_{c_2}[v]\|_{\mathfrak{q}}. \tag{10}$$

Finally, Theorem 8 shows that (10) converges to 0 as c_1 tends to c_2 and u to v, hence (9) is continuous. ■

Example 10. A nonlinear system identification problem is solved in Gray et al. (2020) by truncating (1) up to words of length J and then applying a recursive least-squares algorithm to identify the coefficients of the generating polynomial $p := \sum_{\eta \in X^{\leq J}} (c, \eta)\eta$. In this case, Theorem 9 applies directly to the sequence of estimates $(\hat{p}_j)_j$ with, for example, $M = 1$.

Example 11. Cuchiero et al. (2019) describe a model for deep neural networks using (4) with parameter dependent vector fields $g_i(z, \theta)$. Here θ is assumed to be the set of fixed parameters of the network, while the inputs u_i correspond to the *trainable* parameters. In light of (5), if these vector fields are analytic in the state, such networks constitute a family of Chen-Fliess series $C = \{c_\theta \in \mathbb{R}_{LC}^\ell\langle\langle X \rangle\rangle : \theta \in \Theta\}$. In which case, each generating series would have a θ dependent growth parameter $M(\theta)$. For a fixed $\theta \in \Theta$, Theorem 8 ensures that the output of the network is a continuous function of the trainable parameters. In addition, if $\sup_{\theta \in \Theta} M(\theta)$ is finite, then Theorem 9 guarantees joint continuity in both the trainable parameters and over the set of design parameters Θ.

5. CONCLUSIONS

Sufficient conditions were given under which input-output systems represented by locally convergent Chen-Fliess series are *jointly* continuous with respect to their generating series and as operators mapping a ball in an $L_{\mathfrak{p}}$-space to a ball in an $L_{\mathfrak{q}}$-space, where $\mathfrak{p}$ and $\mathfrak{q}$ are conjugate exponents. Continuity with respect to the generating series was characterized using Banach topologies on subsets of $\mathbb{R}^\ell\langle\langle X \rangle\rangle$. These results were then combined with elements of proofs for other forms of continuity appearing in the literature to produce the desired joint continuity result.

REFERENCES

Azhmyakov, V., Egerstedt, M., Fridman, L., and Poznyak, A. (2009). Continuity properties of nonlinear affine control systems: Applications to hybrid and sliding mode dynamics. *IFAC Proceedings Volumes*, 42, 204–209.

Berstel, J. and Reutenauer, C. (1988). *Rational Series and Their Languages*. Springer-Verlag, Berlin.

Bogfjellmo, G. and Schmeding, A. (2016). The tame Butcher group. *J. Lie Theory*, 26, 1107–1144.

Cuchiero, C., Larsson, M., and Teichmann, J. (2019). Deep neural networks, generic universal interpolation, and controlled ODEs. arxiv:1908.07838.

Dahmen, R. and Schmeding, A. (2018). Lie groups of controlled characters of combinatorial Hopf algebras. *Ann. Inst. Henri Poincaré D, Comb. Phys. Interact. (AIHPD)*.

Duffaut Espinosa, L.A. (2009). *Interconnections of Nonlinear Systems Driven by L_2-Itô Stochastic Processes*. Ph.D. dissertation, Old Dominion University.

Fliess, M. (1981). Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bull. Soc. Math. France*, 109, 3–40.

Fliess, M. (1983). Réalisation locale des systèmes non linéaires, algèbres de lie filtrées transitives et séries génératrices non commutatives. *Invent. Math.*, 71, 521–537.

Gray, W.S., Palmstrøm, M., and Schmeding, A. (2021). Continuity of formal power series products in nonlinear control theory. arxiv:2102.05732.

Gray, W.S., Venkatesh, G.S., and Duffaut Espinosa, L.A. (2020). Nonlinear system identification for multivariable control via discrete-time Chen-Fliess series. *Automatica*, 119. article 109085.

Gray, W.S. and Wang, Y. (2002). Fliess operators on L_p spaces: Convergence and continuity. *Systems Control Lett.*, 46, 67–74.

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

Lang, S. (1999). *Fundamentals of Differential Geometry*, volume 191 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

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

Nieuwenhuis, J.W. and Willems, J.C. (1992). Continuity of dynamical systems: The continuous-time case. *Math. Control Signals Systems*, 5, 391–400.

OEIS (2021). The On-line Encyclopedia of Integer Sequences. https://oeis.org.

Pérez Carreras, P. and Bonet, J. (1987). *Barrelled locally convex spaces*, volume 131 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam. Notas de Matemática, 113.

Thitsa, M. and Gray, W.S. (2012). On the radius of convergence of interconnected analytic nonlinear input-output systems. *SIAM J. Control Optim.*, 50, 2786–2813.

Wang, Y. (1990). *Differential Equations and Nonlinear Control Systems*. Ph.D. dissertation, Rutgers University, New Brunswick, NJ.

Werner, D. (2000). *Funktionalanalysis*. Springer-Verlag, Berlin, extended edition.

Winter Arboleda, I.M. (2019). *On Analytic Nonlinear Input-output Systems: Expanded Global Convergence and System Interconnections*. Ph.D. dissertation, Old Dominion University.

Winter Arboleda, I.M., Gray, W.S., and Duffaut Espinosa, L.A. (2015). Fractional Fliess operators: Two approaches. In *Proc. 49th Conference on Information Sciences and Systems*, pp. 1–6. Baltimore, Maryland.

Yoshinaga, K. (1957). On a locally convex space introduced by J. S. E. Silva. *J. Sci. Hiroshima Univ. Ser. A*, 21, 89–98.