2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Functional Derivatives of Chen-Fliess Series with Applications to Optimal Control

Luis A. Duffaut Espinosa† W. Steven Gray‡ Ivan Perez Avellaneda†

Abstract— **Functional optimization problems, such as those appearing in optimal control, are often stated in terms of finding the critical points of a variational derivative. The first goal of this paper is to describe the Fréchet derivative of a Chen-Fliess series and to provide an algebraic framework for computing it. The second goal is to show how to characterize and compute critical points of this Fréchet derivative both analytically and numerically. The former requires a certain *shuffle separability* property of the generating series for the Fréchet derivative and employs the concept of a *nullable series*. Finally, some simple examples are provided to show how these ideas can be applied to solve quadratic optimal control problems entirely in the context of Chen-Fliess series.**

Index Terms— **nonlinear control systems, optimal control, Chen-Fliess series, Fréchet and Gâteaux derivatives**

I. INTRODUCTION

Functional optimization problems, such as those appearing in optimal control, are often stated in terms of finding the critical points of a variational derivative. This yields the familiar Euler's equation [9]. Any system dynamics described by differential equations can be adjoined to the performance index a priori [1]. A less standard problem, however, is characterizing an optimal solution when the system dynamics are *not* described explicitly by differential equations but rather in terms of an input-output model, such as a Chen-Fliess series. Chen-Fliess series are weighted sums of iterated integrals defined locally within a Banach space that can be used to represent a broad class of nonlinear input-output systems [5], [13], [16]. In this situation, it is asserted that some notion of a variational derivative of a Chen-Fliess series will be required in order to characterize an optimal solution. This idea also appeared recently in a different application where the optimization problem was to find the minimum bounding box for the output reachable sets of a smooth, control-affine state space realization [18], [19]. In this context, it was necessary to determine the maxima and minima of the output of a Chen-Fliess series over a fixed time horizon and set of admissible inputs. The problem at its core required finding critical points of the Gâteaux derivatives of a Chen-Fliess series [4].

The first goal of this paper is to describe the Fréchet derivative of a Chen-Fliess series and to provide an algebraic framework for computing it. The earliest work in this direction is due to Fliess in [6], [8], who introduced the notion of a *causal derivative* of a Chen-Fliess series based on the work of Ree [20]. It employs the left-shift operator on a formal power series and ultimately produces the time derivative of the output function for a given input. This is distinct from a functional derivative on a Banach space, though there may be some relation between the two as suggested by this author. The next contribution in this direction, as mentioned above, has its origins in the work appearing in [18], [19] and involved defining the Gâteaux derivative of a Chen-Fliess series. This is the functional analogue of the directional derivative in finite dimensional calculus. It is a weaker notion of differentiation than the Fréchet derivative. The two always coincide when the latter exists. Next, the notion of a formal derivative of a generating series is defined. This is a purely algebraic device defined in terms of a derivation on the concatenation (Cauchy) product of formal power series. It defines a certain *differential language* which in some ways parallels the notion of a differential field in [2] and a differential algebra in [7]. (Any possible connections to these objects is beyond the scope of the present paper.) It is demonstrated that both functional derivatives described above, when they exist, can be computed using this formal derivative. Moreover, it is shown that the Gâteaux derivative always exist when the associated generating series is local convergent.

The second goal of the paper is to show how to characterize and compute critical points of the Fréchet derivative of a Chen-Fliess series. The first objective is straightforward, while the second is more interesting. In particular, it is shown that when the Fréchet derivative has a certain *shuffle separability* property, critical points can be characterized using the concept of a nullable series as described in [12]. In some cases, it is possible to explicitly compute the Taylor series of critical points. It is also shown how to numerically estimate these critical points via a practical algorithm based on the Newton-Householder algorithm [14]. Finally, the paper concludes with the original motivating application of optimal control. It is demonstrated via two simple examples how the ideas above can be applied to solve quadratic optimal control problems entirely in the context of Chen-Fliess series.

The paper is organized as follows. In the next section, some preliminaries on Chen-Fliess series are provided together with a brief overview of nullable formal power series. Section III presents the various notions of a derivative of a Chen-Fliess series considered in the paper. In Section IV, the concept of a critical point for a Chen-Fliess series is introduced along with a collection of examples. Section V describes one numerical method for estimating such critical points. Finally, Section VI gives an application of these results to optimal control. The conclusions of the paper are given in the last section.

†Department of Electrical and Biomedical Engineering, University of Vermont, Burlington, Vermont 05405 USA

‡Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia 23529 USA

II. Preliminaries

An *alphabet* $X = \{x_0, x_1, \ldots, x_m\}$ is any nonempty and finite set of 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 $X^* = \bigcup_{k\geq 0} X^k$, which is a monoid under the concatenation product. For any word $\eta \in X^*$, ηX^* is the set of all words having the prefix η. 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* $(c,\eta) \in \mathbb{R}^\ell$ 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 $\mathbb{R}^\ell\langle\langle X \rangle\rangle$. If $X_0 = \{x_0\}$, then $\mathbb{R}[[X_0]]$ is the set of all commutative series in x_0. $\mathbb{R}\langle\langle X \rangle\rangle$ 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

$$(x_i\eta) \sqcup\!\sqcup (x_j\xi) = x_i(\eta \sqcup\!\sqcup (x_j\xi)) + x_j((x_i\eta) \sqcup\!\sqcup \xi),$$

where $x_i, x_j \in X$, $\eta, \xi \in X^*$ and with $\eta \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup \eta = \eta$ [5], [20]. The shuffle of two languages $L_1, L_2 \subseteq X^*$ is the language $\mathbb{S}_{L_1,L_2} := \cup_{\eta_i \in L_i} \mathrm{supp}(\eta_1 \sqcup\!\sqcup \eta_2)$. Finally, the characteristic series of a language $L \subseteq X^*$ is the element in $\mathbb{R}\langle\langle X \rangle\rangle$ defined by $\mathrm{char}(L) = \sum_{\nu \in L} \nu$.

A. Chen-Fliess series

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 $B^m_\mathfrak{p}(R)[t_0, t_1] := \{u \in L^m_\mathfrak{p}[t_0, t_1] : \|u\|_\mathfrak{p} \leq R\}$. Assume $C[t_0, t_1]$ is the subset of continuous functions in $L^m_1[t_0, t_1]$. Define inductively for each word $\eta = x_i\bar{\eta} \in X^*$ 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 generating series c is

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

[5]. It can be shown that if there exist real numbers $K, M \geq 0$ such that

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

($|z| := \max_i |z_i|$ 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., $1/\mathfrak{p} + 1/\mathfrak{q} = 1$ [13]. Any such mapping is called a *locally convergent Chen-Fliess series*. If $X = \{x_0, x_1\}$ then F_c with $c \in \mathbb{R}\langle\langle X \rangle\rangle$ constitutes a single-input, single-output system.

Finally, a Chen-Fliess series 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

$$\dot{z}(t) = g_0(z(t)) + \sum_{i=1}^{m} g_i(z(t))\, u_i(t), \quad z(t_0) = z_0 \qquad (3a)$$

$$y_j(t) = h_j(z(t)), \quad j = 1, 2, \ldots, \ell, \qquad (3b)$$

where each g_i is an analytic vector field expressed in local coordinates on some neighborhood $\mathcal{W}$ of $z_0 \in \mathbb{R}^n$, and each output function h_j is an analytic function on $\mathcal{W}$ such that (3a) 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 $y_j(t) = F_{c_j}[u](t) = h_j(z(t))$, $j = 1, 2, \ldots, \ell$. It can be shown that for any word $\eta = x_{i_k} \cdots x_{i_1} \in X^*$

$$(c_j, \eta) = L_{g_\eta} h_j(z_0) := L_{g_{i_1}} \cdots L_{g_{i_k}} h_j(z_0), \qquad (4)$$

where $L_{g_i} h_j$ is the *Lie derivative* of h_j with respect to g_i [5], [16], [17].

B. Zeroing the output of a Chen-Fliess series

The problem of computing an input u^* such that $F_c[u^*] = 0$ on some interval $[0, T]$ is directly connected to the notion of *zero dynamics* when $y = F_c[u]$ has a state space realization (3) [16]. In this context, a sufficient condition for the existence of a u^* is that the realization has a well defined (vector) relative degree at a given initial condition. In [12], however, it is shown that the problem of zeroing the output of a single-input, single-output system F_c can have a solution even when F_c is not realizable and c does not have relative degree in the sense that there exists some $e \in \mathbb{R}\langle\langle X \rangle\rangle$ with $\mathrm{supp}(e) \subseteq X^*/\{X_0^*, x_1\}$ so that c can be written in the form

$$c = c_N + c_F = c_N + K x_0^{r-1} x_1 + x_0^{r-1} e,$$

where $c_N = \sum_{k\geq 0} (c, x_0^k) x_0^k$, $c_F = c - c_N$, and $K \in \mathbb{R}/\{0\}$ [10]. In fact, by employing a certain series composition product so that $F_c[u] = F_{c \circ c_u}[0]$, where c_u can be identified with the Taylor series of u at $t = 0$, it can be shown that the problem of determining c_{u^*} such that $c \circ c_{u^*} = 0$ is meaningful even when c and c_u do not converge in any sense. That is, the composition product is still well defined (locally finite), and setting $c \circ c_{u^*} = 0$ where c_{u^*} is unknown yields a purely algebraic problem which can be solved computationally using a variety of methods, see, for example, [11]. In this setting, the following concepts are useful.

Definition 1: A series $c \in \mathbb{R}\langle\langle X \rangle\rangle$ is said to be *nullable* if the zero series is in the range of the mapping $c\circ : \mathbb{R}[[X_0]] \to \mathbb{R}[[X_0]], c_u \mapsto c \circ c_u$. That is, there exists a *nulling series* $c_{u^*} \in \mathbb{R}[[X_0]]$ such that $c \circ c_{u^*} = 0$. The series is *strongly nullable* if it has a nonzero nulling series. A strongly nullable series is *primely nullable* if its nulling series is unique.

A sufficient condition for a series to be primely nullable is given in the following theorem.

Theorem 1: [12] If $c \in \mathbb{R}\langle\langle X \rangle\rangle$ has relative degree r, and $\mathrm{supp}(c_N) \subseteq x_0^r X_0^*$ is nonempty, then c is primely nullable.

Series satisfying Theorem 1 conditions will be referred to as *linearly nullable* since the linear word $x_0^{r-1} x_1$ in its support plays a key role in computing the nulling series [10].

III. Derivatives of Chen-Fliess Series

In this section, various notions of differentiability of Chen-Fliess series are presented. The following lemma providing a closed formula for a Chen-Fliess series driven by the sum of two inputs will be useful in this regard.

Lemma 1: [18] Let X and Y be alphabets associated with the inputs $u,\ v \in L_{\mathfrak{p}}^m[0,T]$, respectively. Define $Z = X \cup Y$. If $c \in \mathbb{R}_{LC}^{\ell}\langle\langle X\rangle\rangle$, then the Chen-Fliess series with input $u+v$ can be written as

$$F_c[u+v](t) = \sum_{k=0}^{\infty} \sum_{\xi \in \mathbb{S}_{X^*,Y^k}} (c, \sigma_X(\xi))\mathcal{E}_{\xi}[u,v](t), \quad (5)$$

where $\sigma_X : Z^* \to X^*$ is a substitution word homomorphism defined by its action on letters as $\sigma_X(x_i) = x_i$, $\sigma_X(y_i) = x_i$, and $\sigma_X(c) \in \mathbb{R}_{LC}\langle\langle Z\rangle\rangle$ such that $(\sigma_X(c), \xi) = (c, \sigma_X(\xi))$. The mapping $\mathcal{E}_\xi : L_{\mathfrak{p}}^m[0,T] \times L_{\mathfrak{p}}^m[0,T] \to \mathcal{C}[0,T]$ for $\xi \in Z^*$ is defined inductively by

$$\mathcal{E}_{z_i\bar{\xi}}[u,v](t) := \begin{cases} \displaystyle\int_0^t u_i(\tau)\mathcal{E}_{\bar{\xi}}[u,v](\tau)d\tau, & z_i \in X, \\ \displaystyle\int_0^t v_i(\tau)\mathcal{E}_{\bar{\xi}}[u,v](\tau)d\tau, & z_i \in Y, \end{cases}$$

and $\mathcal{E}_{\emptyset}[u,v](t) = 1$.

Observe that since $|(c,\sigma_X(\xi))| \leq KM^{\xi}|\xi|!$, the operator (5) preserves local convergence so that $u + v \in B_{\mathfrak{p}}^m(R)[0,T]$ is mapped to $B_{\mathfrak{q}}^{\ell}(S)[0,T]$ for some $R, S > 0$.

A. Fréchet derivative

A functional derivative on a Banach space is a classical concept in analysis [9], [15]. The focus here is on the existence and computation of functional derivatives of Chen-Fliess series. The following collection of results is adapted from the standard treatment.

Definition 2: Let $c \in \mathbb{R}_{LC}^{\ell}\langle\langle X\rangle\rangle$ and fix input $u \in B_{\mathfrak{p}}^m(R)[0,T]$. The Chen-Fliess series F_c is *Fréchet differentiable* at u if there exists $DF_c[u][\cdot] : B_{\mathfrak{p}}^m(R)[0,T] \to \mathbb{R}^{\ell}$ such that the following limit is satisfied:

$$\lim_{\|h\|_{\mathfrak{p}} \to 0} \frac{1}{\|h\|_{\mathfrak{p}}}\Big(F_c[u+h](t) - F_c[u](t) - DF_c[u][h](t)\Big) = 0 \quad (6)$$

for all $u+h, h \in B_{\mathfrak{p}}^m(R)[0,T]$.

Theorem 2: Let X and Y be alphabets associated with $u, h \in B_{\mathfrak{p}}^m(R)[0,T]$, respectively. Fix $c \in \mathbb{R}_{LC}^{\ell}\langle\langle X\rangle\rangle$. Then the Chen-Fliess series F_c is Fréchet differentiable at u if and only if

$$\lim_{\|h\|_{\mathfrak{p}} \to 0} \frac{1}{\|h\|_{\mathfrak{p}}}\left(\sum_{k=2}^{\infty}\sum_{\eta\in X^*}\sum_{\xi \in \mathbb{S}_{\eta,Y^k}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t)\right) = 0 \quad (7)$$

for all $h \in B_{\mathfrak{p}}^m(R)[0,T]$, and its Fréchet derivative is

$$DF_c[u][h](t) = \sum_{\eta\in X^*}\sum_{\xi\in\mathbb{S}_{\eta,Y}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t). \quad (8)$$

Proof: The proof follows from a direct application of Lemma 1 and Definition 2. Consider $\delta > 0$ and h such that $\|h\|_{\mathfrak{p}} < \delta$. From (5) it follows that

$$F_c[u+h](t) = \sum_{k=0}^{\infty}\sum_{\eta\in X^*}\sum_{\xi\in\mathbb{S}_{\eta,Y^k}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t).$$

For $k=0$, one has that

$$F_c[u](t) = \sum_{\eta\in X^*}\sum_{\xi\in\mathbb{S}_{\eta,Y^0}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t). \quad (9)$$

Note here that $\mathcal{E}_\xi[u,h] = E_\xi[u]$ since $\xi \in X^*$, and thus, the left-hand side of (9) does not depend on h. It then follows that

$$F_c[u+h](t) - F_c[u](t) - \sum_{\eta\in X^*}\sum_{\xi\in\mathbb{S}_{\eta,Y}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t) = \sum_{k=2}^{\infty}\sum_{\eta\in X^*}\sum_{\xi\in\mathbb{S}_{\eta,Y^k}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t).$$

Multiplying both sides above by $1/\|h\|_{\mathfrak{p}}$ and letting $\|h\|_{\mathfrak{p}} \to 0$ gives the desired result. Observe that if c satisfies (2), then $|(c,\sigma_X(\xi))| \leq KM^{|\xi|}|\xi|!$ since $\sigma_X(\xi) \in \mathrm{supp}(c)$. Thus, the generating series for $DF_c[u][h](t)$ inherits a local convergent bound from the original series F_c. Therefore, the Fréchet derivative $DF_c[u][h](t)$ maps $u,h \in B_{\mathfrak{p}}^m(R)[0,T]$ to an output in $B_{\mathfrak{q}}^{\ell}(S)[0,T]$, which completes the proof. ■

Example 1: Let $c = x_1^2$ and $h \in B_{\mathfrak{p}}(R)[0,T]$ such that (7) holds. From Theorem 2, the Fréchet derivative of c is computed using (8). The only words $\xi \in Z^*$ such that $\sigma_X(\xi) = x_1^2$ are those in $\mathbb{S}_{\eta,Y} = \{x_1y_1, y_1x_1\}$ with $\eta = x_1$. Therefore,

$$\begin{aligned} DF_c[u][h](t) &= \sum_{\xi\in\mathbb{S}_{x_1,Y}} (c,\sigma_X(\xi))\mathcal{E}_\xi[u,h](t) \\ &= \mathcal{E}_{x_1y_1}[u,h](t) + \mathcal{E}_{y_1x_1}[u,h](t). \end{aligned} \quad (10)$$

As a check, one can compute the Fréchet derivative of $y(t) = \int_0^t u(\tau)\int_0^\tau u(\sigma)\, d\sigma d\tau$ directly from first principles. Necessarily from (6),

$$\begin{aligned} 0 &= \lim_{\|h\|_{\mathfrak{p}}\to 0}\frac{1}{\|h\|_{\mathfrak{p}}}\Big(\int_0^t (u+h)(\tau)\int_0^\tau (u+h)(\sigma)\, d\sigma d\tau \\ &\quad - \int_0^t u(\tau)\int_0^\tau u(\sigma)\, d\sigma d\tau - DF_c[u][h](t)\Big) \\ &= \lim_{\|h\|_{\mathfrak{p}}\to 0}\frac{1}{\|h\|_{\mathfrak{p}}}\Big(\int_0^t u(\tau)\int_0^\tau h(\sigma)\, d\sigma d\tau \\ &\quad + \int_0^t h(\tau)\int_0^\tau u(\sigma)\, d\sigma d\tau + \int_0^t h(\tau)\int_0^\tau h(\sigma)\, d\sigma d\tau \\ &\quad - DF_c[u][h](t)\Big). \end{aligned}$$

The assertion is that in the limit, the third term above is zero. Clearly,

$$\left\|\mathcal{E}_{y_1^2}[u,h]\right\|_{\mathfrak{q}} \leq \left(\int_0^T\left(\int_0^t |h(\tau)|\int_0^\tau |h(\sigma)|\, d\sigma d\tau\right)^q dt\right)^{1/q},$$

and from Hölder's inequality $\|h\|_1 \leq \|h\|_{\mathfrak{p}} T^{1/q}$. A second application of Hölder's inequality yields

$$\left\|\mathcal{E}_{y_1^2}[u,h]\right\|_{\mathfrak{q}} \leq \|h\|_{\mathfrak{p}} T^{1/q}\left(\int_0^T\left(\int_0^T |h(\sigma)|\, d\sigma\right)^q d\tau\right)^{1/q}$$

$$= \|h\|_{\mathfrak{p}}^2 T^{3/q}.$$

Hence,

$$\lim_{\|h\|_{\mathfrak{p}}\to 0} \frac{1}{\|h\|_{\mathfrak{p}}} |\mathcal{E}_{y_1^2}[u,h]| \leq \lim_{\|h\|_{\mathfrak{p}}\to 0} \frac{1}{\|h\|_{\mathfrak{p}}} \left\| \mathcal{E}_{y_1^2}[u,h] \right\|_{\mathfrak{q}} \leq \lim_{\|h\|_{\mathfrak{p}}\to 0} \|h\|_{\mathfrak{p}} T^{3/q} = 0.$$

By the continuity result in [3, Theorem 4] and

$$\lim_{\|h\|_{\mathfrak{p}}\to 0} \frac{1}{\|h\|_{\mathfrak{p}}} \int_0^t h(\tau) \int_0^\tau h(\sigma)\ d\sigma d\tau = 0, \tag{11}$$

one has that

$$DF_c[u][h](t) = \int_0^t u(\tau) \int_0^\tau h(\sigma)\ d\sigma d\tau + \int_0^t h(\tau) \int_0^\tau u(\sigma)\ d\sigma d\tau$$

as expected from (10), and (11) is exactly condition (7) in Theorem 2. □

Example 2: Consider the locally convergent generating series $c = \sum_{\eta\in X^*} KM^{|\eta|}\, |\eta|!\, \eta$ for some fixed real numbers $K, M > 0$. In this instance, the Chen-Fliess series has a closed-form expression

$$F_c[u] = \frac{K}{1 - M\sum_{i=0}^m E_{x_i}[u]}.$$

Therefore, the Fréchet derivative is computed directly as

$$DF_c[u][h] = \frac{KM}{(1 - M\sum_{i=0}^m E_{x_i}[u])^2} \sum_{i=1}^m E_{x_i}[h].$$

□

B. Gâteaux Derivative

The derivative of a Chen-Fliess series in a specific direction in $L_{\mathfrak{p}}^m[0,T]$ is given next. This is the Gâteaux derivative. It is a weaker notion of differentiability than that of Fréchet.

Definition 3: Given $c \in \mathbb{R}_{LC}^\ell\langle\langle X\rangle\rangle$ and the input functions $u,\ v \in B_{\mathfrak{p}}^m(R)[0,T]$, the Chen-Fliess series is *Gâteaux differentiable* at u in the direction of v if and only if there exists for each $t \in [0,T]$ an $\mathbb{R}^\ell$-vector denoted by $\frac{\partial}{\partial v} F_c[u](t)$ such that the following limit holds

$$\lim_{\varepsilon\to 0} \frac{1}{\varepsilon}\Big(F_c[u+\varepsilon v](t) - F_c[u](t) - \frac{\partial}{\partial v} F_c[u](t)\varepsilon \Big) = 0.$$

The next result is a stronger version of that given in [19, Theorem 1]. It is a consequence of Theorem 2 when a fixed direction $h = v$ is chosen.

Theorem 3: If $c \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$, then F_c is always Gâteaux differentiable at u in the direction of v for any admissible u, v. Specifically,

$$\frac{\partial}{\partial v} F_c[u](t) = \sum_{\eta\in X^*} \sum_{\xi\in \mathbb{S}_{\eta,Y}} (c, \sigma_X(\xi)) \mathcal{E}_\xi[u,v](t). \tag{12}$$

Proof: Due to space constraints, only the outline of the proof is given. From Lemma 1 and Definition 3, the Gâteaux derivative exist if and only if

$$\lim_{\varepsilon\to 0} \sum_{k=2}^\infty \sum_{\eta\in X^*} \sum_{\xi\in \mathbb{S}_{\eta,Y^k}} (c, \sigma_X(\xi)) \mathcal{E}_\xi[u,v](t)\, \varepsilon^k\ =\ 0 \tag{13}$$

holds. Since $c \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$, $(c, \sigma_X(\xi))$ satisfies (2), and there exist a nonzero radius of convergence for F_c [13]. By counting the terms in the summations and using the fact that $\|\mathcal{E}_\xi[u,v]\| \propto 1/|\xi|!$, it is not difficult to show that (13) converges to 0 within the radius of convergence of F_c and that $\frac{\partial}{\partial v} F_c[u](t)$ is written as in (12). ■

Example 3: Reconsider the system in Example 1. Fixing $v \in B_{\mathfrak{p}}(R)[0,T]$, one can compute the corresponding Gâteaux derivative as

$$\begin{aligned}
&\frac{\partial}{\partial v} F_c[u](t) \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \bigg(\int_0^t (u+\varepsilon v)(\tau) \int_0^\tau (u+\varepsilon v)(\sigma)\ d\sigma d\tau \\
&\quad - \int_0^t u(\tau) \int_0^\tau u(\sigma)\ d\sigma d\tau \bigg) \\
&= \lim_{\varepsilon\to 0} \frac{1}{\varepsilon} \bigg(\varepsilon \int_0^t u(\tau) \int_0^\tau v(\sigma)\ d\sigma d\tau + \varepsilon \int_0^t v(\tau) \int_0^\tau u(\sigma)\ d\sigma d\tau \\
&\quad + \varepsilon^2 \int_0^t v(\tau) \int_0^\tau v(\sigma)\ d\sigma d\tau \bigg) \\
&= \int_0^t u(\tau) \int_0^\tau v(\sigma)\ d\sigma d\tau + \int_0^t v(\tau) \int_0^\tau u(\sigma)\ d\sigma d\tau \\
&= \mathcal{E}_{x_1 y_1}[u,v](t) + \mathcal{E}_{y_1 x_1}[u,v](t).
\end{aligned}$$

Moreover, from the definition of the set $\mathbb{S}_{x_1,Y}$ with $Y = \{y_1\}$, one can write

$$\frac{\partial}{\partial v} F_c[u](t) = F_{x_1 \sqcup\!\sqcup\, y_1}[u,v](t).$$

□

Example 4: Let $c = \sum_{n\geq 0} (c, x_0^n x_1) x_0^n x_1$. This is called a *linear series* because the words $x_0^n x_1$ result from applying (4) to a linear time-invariant system. In this case,

$$\frac{\partial}{\partial v} F_c[u](t) = \sum_{n=0}^\infty (c, x_0^n x_1) \mathcal{E}_{x_0^n y_1}[u,v](t).$$

Suppose $\dot z = z + u$, $z(0) = 0$ with output $y = z$. The nonzero coefficients of the Chen-Fliess series are $(c, x_0^n x_1) = 1$ for $n \geq 0$ and $\mathcal{E}_{x_0^n y_1}[u,v](t) = \int_0^t \frac{(t-\tau)^n}{n!} v(\tau)\ d\tau$. Thus,

$$\frac{\partial}{\partial v} F_c[u](t) = \sum_{n=0}^\infty \mathcal{E}_{x_0^n y_1}[u,v](t) = \int_0^t e^{t-\tau}\ v(\tau)\ d\tau.$$

□

The previous examples suggest that one can compute the Fréchet and Gâteaux derivatives of a Chen-Fliess series using only its generating series as considered next.

C. Formal derivative

Let $X = \{x_0, x_1, \cdots, x_m\}$, and define alphabets $\delta X = \{\delta x_1, \cdots, \delta x_m\}$ and $Z = X \cup \delta X$. Define the map $\delta : Z \cup \{0\} \to Z \cup \{0\}$ such that $\delta(x_i) = \delta x_i$ for $x_i \neq x_0$ and $\delta(x_0) = \delta(\emptyset) = \delta(0) = 0$. Extend the definition of δ to Z^* by letting it act as a derivation with respect to concatenation so that

$$\delta(z_i \eta) = \delta(z_i)\eta + z_i \delta(\eta), \tag{14}$$

where $z_i \in Z$ and $\eta \in Z^*$. Assume δ acts linearly on polynomials and series in Z and extend it componentwise to $\mathbb{R}^\ell\langle\langle Z\rangle\rangle$. Finally, define $\delta^2(z_i) = 0$. The operator δ will be referred to as the *formal derivative*.

Example 5: If $c = x_0 + 2x_0x_1 + x_1^2$, then

$$\delta(c) = 2x_0\delta x_1 + \delta x_1 x_1 + x_1\delta x_1, \quad \delta^2(c) = 2(\delta x_1)^2,$$

and $\delta^k(c) = 0$ for all $k > 2$. □

The next lemma shows that δ also acts as a derivation on the shuffle product.

Lemma 2: If $c, d \in \mathbb{R}\langle\langle Z\rangle\rangle$, then the operator δ acts as a derivation on the shuffle product $c \sqcup\!\sqcup\, d$. That is,

$$\delta(c \sqcup\!\sqcup\, d) = \delta(c) \sqcup\!\sqcup\, d + c \sqcup\!\sqcup\, \delta(d). \tag{15}$$

Proof: First observe that if (15) is true for words in Z^*, that is,

$$\delta(\eta \sqcup\!\sqcup\, \xi) = \delta(\eta) \sqcup\!\sqcup\, \xi + \eta \sqcup\!\sqcup\, \delta(\xi), \tag{16}$$

then

$$\begin{aligned}\delta(c \sqcup\!\sqcup\, d) &= \sum_{\eta,\xi\in Z^*} (c,\eta)(d,\xi)\delta(\eta \sqcup\!\sqcup\, \xi) \\ &= \sum_{\eta,\xi\in Z^*} (c,\eta)(d,\xi)(\delta(\eta) \sqcup\!\sqcup\, \xi + \eta \sqcup\!\sqcup\, \delta(\xi)) \\ &= \sum_{\eta,\xi\in Z^*} (c,\eta)(d,\xi)\delta(\eta) \sqcup\!\sqcup\, \xi \\ &\quad + \sum_{\eta,\xi\in Z^*} (c,\eta)(d,\xi)\eta \sqcup\!\sqcup\, \delta(\xi) \\ &= \delta(c) \sqcup\!\sqcup\, d + c \sqcup\!\sqcup\, \delta(d).\end{aligned}$$

Thus, it is sufficient to show that (16) holds. The proof is done by induction on the word length $n = |\eta| + |\xi|$. The cases $n = 0, 1, 2$ are trivial. Assuming (16) holds for $n \geq 2$, consider $\eta = z_i\eta'$ and $\xi = z_j\xi'$ with $z_i, z_j \in Z$ and $\eta', \xi' \in Z^*$. From the shuffle product definition and (14), one has

$$\begin{aligned}\delta(\eta \sqcup\!\sqcup\, \xi) &= \delta(z_i(\eta' \sqcup\!\sqcup\, \xi) + z_j(\eta \sqcup\!\sqcup\, \xi')) \\ &= \delta(z_i)(\eta' \sqcup\!\sqcup\, \xi) + z_i\delta(\eta' \sqcup\!\sqcup\, \xi) \\ &\quad + \delta(z_j)(\eta \sqcup\!\sqcup\, \xi') + z_j\delta(\eta \sqcup\!\sqcup\, \xi').\end{aligned}$$

By the induction hypothesis and using (14) a second time, it follows that

$$\begin{aligned}\delta(\eta \sqcup\!\sqcup\, \xi) &= \delta(z_i)(\eta' \sqcup\!\sqcup\, \xi) + z_i(\delta(\eta') \sqcup\!\sqcup\, \xi) + z_i(\eta' \sqcup\!\sqcup\, \delta(\xi)) \\ &\quad + \delta(z_j)(\eta \sqcup\!\sqcup\, \xi') + z_j(\delta(\eta) \sqcup\!\sqcup\, \xi') + z_j(\eta \sqcup\!\sqcup\, \delta(\xi')) \\ &= \delta(z_i)(\eta' \sqcup\!\sqcup\, \xi) + z_i(\delta(\eta') \sqcup\!\sqcup\, \xi) \\ &\quad + z_i(\eta' \sqcup\!\sqcup\, (\delta(z_j)\xi' + z_j\delta(\xi'))) \\ &\quad + \delta(z_j)(\eta \sqcup\!\sqcup\, \xi') + z_j(\eta \sqcup\!\sqcup\, \delta(\xi')) \\ &\quad + z_j((\delta(z_i)\eta' + z_i\delta(\eta')) \sqcup\!\sqcup\, \xi').\end{aligned}$$

From the linearity of the shuffle product, one can re-arrange the terms of the expression above as

$$\begin{aligned}\delta(\eta \sqcup\!\sqcup\, \xi) &= \delta(z_i)\eta' \sqcup\!\sqcup\, \xi + \eta \sqcup\!\sqcup\, \delta(z_j)\xi' \\ &\quad z_i\delta(\eta') \sqcup\!\sqcup\, \xi + \eta \sqcup\!\sqcup\, z_j\delta(\xi') \\ &= \delta(\eta) \sqcup\!\sqcup\, \xi + \eta \sqcup\!\sqcup\, \delta(\xi),\end{aligned}$$

which completes the proof. ■

Theorem 4: Let $c \in \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$ and assume that (7) holds so that $DF_c[u][h]$ is well-defined. Then

$$DF_c[u][h] = F_{\delta(c)}[u, h]$$

for all admissible u, h.

Proof: It was established in the proof of Theorem 2 that $\delta(c) \in \mathbb{R}^\ell_{LC}\langle\langle Z\rangle\rangle$, so the operator on the right side is at least well defined. The equivalence of the two operators on some admissible set follows from the same argument given in [19, Theorem 2] due to the fact that (12) coincides with (8) when (7) holds. ■

Example 6: Reconsider the series in Example 2. Note that the Fréchet derivative can be re-written as

$$DF_c[u][h] = \frac{M}{K}F_c^2[u]F_{\mathrm{char}(\delta X)}[h].$$

Since generating series are unique, this implies that $\delta(c) = (M/K)c \sqcup\!\sqcup\, c \sqcup\!\sqcup\, \mathrm{char}(\delta X)$. This is an example of a case where $\delta(c)$ is *shuffle separable*, that is, $\delta(c) = d \sqcup\!\sqcup\, e$ for some $d \in \mathbb{R}\langle\langle X\rangle\rangle$ and $e \in \mathbb{R}\langle\langle \delta X\rangle\rangle$. □

A natural consequence of Lemma 2 and Theorem 4 is the following result.

Theorem 5: Let $c, d \in \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$. If (7) holds for $F_c[u]$ and $F_d[u]$ so that $DF_c[u][h]$ and $DF_d[u][h]$ are both well-defined, then

$$D(F_cF_d)[u][h] = F_{\delta(c)}[u, h]F_d[u] + F_c[u]F_{\delta(d)}[u, h]$$

for all admissible u, h.

Proof: Observe

$$\begin{aligned}D(F_cF_d)[u][h] &= DF_{c \sqcup\!\sqcup\, d}[u][h] \\ &= F_{\delta(c \sqcup\!\sqcup\, d)}[u, h] \\ &= F_{\delta(c) \sqcup\!\sqcup\, d + c \sqcup\!\sqcup\, \delta(d)}[u, h] \\ &= F_{\delta(c)}[u, h]F_d[u] + F_c[u]F_{\delta(d)}[u, h],\end{aligned}$$

which completes the proof. ■

This section concludes with a brief summary of useful results recently appearing in [19] regarding the computation of the gradient of Chen-Fliess series reformulated in terms of the δ operator. If $c \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$ and $v = u_i$, then the corresponding Gâteaux derivative is

$$\frac{\partial}{\partial u_i}F_c[u](t) = F_{\delta_{x_i}(c)}[u](t).$$

Define the gradient of a Chen-Fliess series as $\nabla F_c : B^m_{\mathfrak{p}}(R)[t_0, t_1] \to B^m_{\mathfrak{q}}(S)[t_0, t_1]$ such that

$$\nabla F_c[u](t) = \left(\frac{\partial}{\partial u_1}F_c[u](t) \cdots \frac{\partial}{\partial u_m}F_c[u](t)\right)^T.$$

Given a *constant* direction $v = \sum_{i=1}^m v_ie_i \in \mathbb{R}^m$ with e_i the i-th elementary vector in $\mathbb{R}^m$, it follows that

$$\frac{\partial}{\partial v}F_c[u](t) = v^T\nabla F_c[u](t). \tag{17}$$

IV. Critical Points of Chen-Fliess Series

The objective of this section is to define the notion of a critical point for a Chen-Fliess series and provide some specific examples. The main definition is given first.

Definition 4: Let F_c with $c \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$ be given. An input $u^* \in B^m_{\mathfrak{p}}(R)[0, T]$ constitutes a *critical point* of F_c if $DF_c[u^*][h] = 0$ for all $t \in [0, T]$ and for all admissible h.

Of course, the main interest is in critical points that identify maxima and minima of F_c, that is, u^* such that

$$DF_c[u^*][h] = 0 \text{ and } D^2F_c[u^*][h] \lesseqgtr 0, \ \forall h, \text{ respectively.}$$

Here the generating series of $D^2F_c[u^*][h]$ is $\delta^2(c)$, which is also a consequence of [19, Theorem 2] as in the proof of Theorem 4. In the event that $\delta(c)$ is shuffle separable so that

$\delta(c) = d \sqcup\!\sqcup\, e$ for some $d \in \mathbb{R}_{LC}\langle\langle X\rangle\rangle$ and $e \in \mathbb{R}_{LC}\langle\langle \delta X\rangle\rangle$, then critical points must satisfy $F_d[u^*] = 0$. That is, d must be a nullable series as described in Definition 1. One example of this separability property is given in Example 6. Additional examples are provided below.

Example 7: Consider a system described by the generating series

$$c = 3x_0^4 - x_0^2x_1^2 - x_0x_1x_0x_1 - x_0x_1^2x_0 - x_1x_0^2x_1 - x_1x_0x_1x_0 - x_1^2x_0^2 + 3x_1^4.$$

Observe that

$$c = \frac{1}{8}\left((x_0+x_1) \sqcup\!\sqcup\, (x_0-x_1)\right)^{\sqcup\!\sqcup\, 2}.$$

From (15), it follows that

$$\begin{aligned}\delta(c) &= \frac{1}{4}\left((x_0+x_1) \sqcup\!\sqcup\, (x_0-x_1)\right) \\ &\quad \sqcup\!\sqcup\, \delta\left((x_0+x_1) \sqcup\!\sqcup\, (x_0-x_1)\right) \\ &= -\frac{1}{2}((x_0+x_1) \sqcup\!\sqcup\, (x_0-x_1) \sqcup\!\sqcup\, x_1 \sqcup\!\sqcup\, \delta x_1).\end{aligned}$$

Thus, $\delta(c)$ is shuffle separable. It follows from the shuffle factors that c has three critical points: $u^*(t) = \pm 1$, $t \geq 0$ and $u^*(t) = 0$, $t \geq 0$. Similarly, the generating series of $D^2F_c[u^*][h]$ is

$$\begin{aligned}\delta^2(c) = -\frac{1}{2}&((x_0-x_1) \sqcup\!\sqcup\, x_1 - (x_0+x_1) \sqcup\!\sqcup\, x_1 \\ &+ (x_0+x_1) \sqcup\!\sqcup\, (x_0-x_1)) \sqcup\!\sqcup\, \delta x_1 \sqcup\!\sqcup\, \delta x_1,\end{aligned}$$

which is also shuffle separable. If $u^*(t) = 0$, $t \geq 0$, then

$$\begin{aligned}F_{\delta^2(c)}[u^*,h](t) &= -\frac{1}{2}F_{x_0 \sqcup\!\sqcup\, x_0}[0](t)\left(F_{\delta x_1}[h](t)\right)^2 \\ &= -\frac{1}{2}t^2\left(F_{\delta x_1}[h](t)\right)^2 < 0.\end{aligned}$$

So this critical point is a maximum. If $u^*(t) = 1$, $t \geq 0$, then

$$\begin{aligned}F_{\delta^2(c)}[u^*,h](t) &= \frac{1}{2}F_{(x_0+x_1) \sqcup\!\sqcup\, x_1}[1](t)\left(F_{\delta x_1}[h](t)\right)^2 \\ &= t^2\left(F_{\delta x_1}[h](t)\right)^2 > 0.\end{aligned}$$

So this critical point is a minimum. Finally, if $u^*(t) = -1$, $t \geq 0$, then

$$\begin{aligned}F_{\delta^2(c)}[u^*,h](t) &= -\frac{1}{2}F_{(x_0-x_1) \sqcup\!\sqcup\, x_1}[1](t)\left(F_{\delta x_1}[h](t)\right)^2 \\ &= t^2\left(F_{\delta x_1}[h](t)\right)^2 > 0.\end{aligned}$$

So this critical point is also a minimum. □

Example 8: Consider the generating series

$$c = (x_0^2 - x_1) \sqcup\!\sqcup\, (2x_0 + x_1)$$

and its derivative

$$\delta(c) = (x_0^2 - 2x_0 - 2x_1) \sqcup\!\sqcup\, \delta x_1,$$

which is shuffle separable. In this case, the series $x_0^2 - 2x_0 - 2x_1$ is nullable as per Theorem 1. A straightforward calculation gives the unique critical point $u^*(t) = t/2 - 1$, $t \geq 1$. Furthermore, the generating series for the $D^2F_c[u^*][h]$ is $\delta^2(c) = -4\delta x_1^2$, which indicates that this critical point of F_c is a maximum. □

Example 9: Consider the state space system

$$\dot{z} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} z + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u, \quad z(0) = z_0, \quad y = z_2^2.$$

Let (A, b) denotes the linear state equation. The generating series for the input-to-state map is

$$c_z = \begin{bmatrix} c_{z,1} \\ c_{z,2} \end{bmatrix} = \sum_{n=0}^{\infty} (c_z, x_0^n)x_0^n + (c_z, x_0^n x_1)x_0^n x_1,$$

where $(c_z, x_0^n) = A^n z_0$ and $(c_z, x_0^n x_1) = A^n b$, $n \geq 0$. Therefore, the generating series for the input-output map is $c_y = c_{z,2} \sqcup\!\sqcup\, c_{z,2}$, and

$$\delta(c_y) = 2(\delta(c_{z,2}) \sqcup\!\sqcup\, c_{z,2}).$$

For the case where $z_0 = 0$, $F_{c_{z,0}}[u^*]$ is the zero function if and only if u^* is the zero input. In addition,

$$\delta^2(c_y) = 2(\delta(c_{z,2}) \sqcup\!\sqcup\, \delta(c_{z,2})),$$

which implies that

$$F_{\delta^2(c_y)}[u^*, h] = 2F^2_{\delta(c_{z,2})}[h] > 0$$

for all h. Hence, $u^*(t) = 0$, $t \geq 0$ is a minimum of F_{c_y}. □

V. Estimating Critical Points Numerically

Critical points of a Chen-Fliess series can be estimated numerically under certain conditions. The objective of this section is to describe one such approach based on the well-known Newton's root finding method [14]. Since searching for critical points directly in $L_{\mathfrak{p}}^m[0,T]$ is not feasible, a heuristic sequential search for critical points within a dense subset of $L_{\mathfrak{p}}^m[0,T]$ is proposed. Specifically, consider the set of piece-wise constant functions taking values in $\mathbb{R}^m$ over a partition $\{t_k\}_{k=0,1,\ldots,N_t}$ of $[0,T]$ as shown in Fig. 1 when $m = 1$. A critical point is assumed to have the form

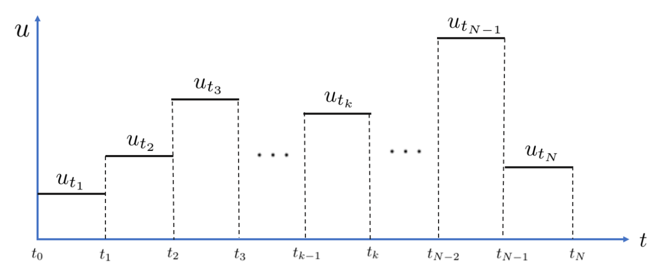

Fig. 1. Piecewise constant function u for partition $\{t_k\}_{k=0,\ldots,N_t}$.

$u^*(t) = u^*_{t_k} \in \mathbb{R}^m$ for $t \in [t_{k-1}, t_k]$ and $k \in \{1, 2, \ldots, N_t\}$. A recursion to find such critical points can be formulated using Newton's root finding method to estimate each of these constant pieces. Starting with the interval $[t_0, t_1]$, choose an initial constant $u^0_{t_1}$ and perform the iteration

$$u^{i+1}(t_1) = u^i(t_1) - (\nabla F_c[u^i](t_1))^{-1} F_c[u^i](t_1), \qquad (18)$$

where $F_c[u^i](t_1)$ and $\nabla F_c[u^i](t_1)$ are evaluated at the constant input $u^i_{t_1}$ over $[0, t_1]$. ∇F_c is computed using (17) in the constant directions of the elementary basis vectors, e_i, for the vector space $\mathbb{R}^m$. After N_r iterations, the value $u^*_{t_1} := u^{N_r}_{t_1}$ gives $F_c[u^*](t_1) \approx 0$ when N_r is chosen sufficiently large. Similarly, the iteration for the second interval in the partition of $[0,T]$ uses (18) to find $u^*_{t_2} = u^{N_r}_{t_2}$, but $F_c[u^i](t_2)$ and

$\nabla F_c[u^i](t_2)$ are now driven by the piecewise function u^i that takes the value of $u^*_{t_1}$ over $[0, t_1]$ and $u^i_{t_2}$ over $(t_1, t_2]$. Only $u^i_{t_2}$ is updated by (18). The procedure is repeated sequentially for all t_k in the partition of $[0, T]$, where $u^*_{t_j}$ is fixed for $j = 1, \ldots, k-1$ and only $u^i_{t_k}$ is updated. The method is summarized in Algorithm 1 below. Note in step 4 of this algorithm that $F_c[u^i](t_k)$ and $\nabla F_c[u^i](t_k)$ are driven by the piecewise constant input defined by $u^i(t) = u^*_{t_j}$ for $t \in [t_{j-1}, t_j]$, $j = 1, \ldots, k-1$, and $u^i(t) = u^i_{t_k}$ for $t \in [t_{k-1}, t_k]$.

Algorithm 1 Newton's root finding method

Input: N_r, N_t,
1: **for** $k = 1$ to N_t **do**
2: *Initialize*: $u^0_{t_j}$
3: **for** $i = 1$ to N_r **do**
4: $u^{i+1}_{t_k} = u^i_{t_k} - (\nabla F_c[u^i](t_k))^{-1} F_c[u^i](t_k)$
5: **end for**
6: $u^*_{t_k} \leftarrow u^{N_r}_{t_k}$
7: **end for**
8: **return** $(u^*_{t_1}, \ldots, u^*_{t_N})$

Describing specific conditions under which Algorithm 1 converges is outside the scope of this paper. But the conjecture is that such a proof would follow along the lines of that for the standard Newton-Kantorovich Theorem since it is based on (18) [21].

Example 10: Algorithm 1 is applied to Examples 8 and 9 to estimate their critical points. An estimate of the critical point of F_c in Example 8 is shown in Fig. 2 when $N_r = 100$ and $u^0_0 = -1$ in step 2 of the algorithm. It is compared against the theoretical critical point $u^* = t/2 - 1$ giving an RMS error $= 9.34 \times 10^{-4}$.

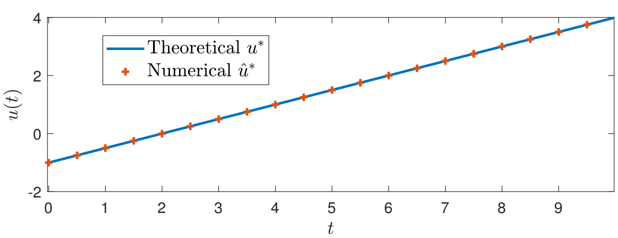

Fig. 2. Comparison of the theoretical and numerical u^* in Example 8.

The Wiener system in Example 9 was shown to have a critical point $u^*(t) = 0$, $t \geq 0$. The output of Algorithm 1 is shown in Fig. 3 and verifies this result when truncating the generating series to $N = 5$ terms. Here the RMS error is $= 3.93 \times 10^{-13}$. □

VI. Application to Optimal Control

Consider an input-output system $y = F_c[u]$, where $c \in \mathbb{R}_{LC}\langle\langle X \rangle\rangle$, and F_c is convergent on $B_{\mathfrak{p}}(R)[0, T]$. Define for some fixed weight $W > 0$ the quadratic performance index

$$J[u] = \int_0^T \frac{1}{2} F_c^2[u](\tau) + \frac{1}{2} W u^2(\tau)\, d\tau.$$

The general goal is to determine the extremals, u^*, of J. A necessary condition follows directly from the calculus of

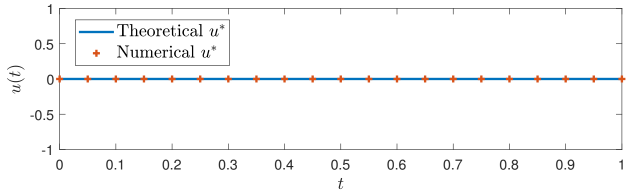

Fig. 3. Comparison of the theoretical and numerical u^* in Example 9.

variations, namely, that the first variation at $u = u^*$ should satisfy

$$\delta J[u^*][h] = \int_0^T F_c[u^*](\tau) F_{\delta(c)}[u^*][h](\tau) + W u^*(\tau) h(\tau)\, d\tau = 0$$

for all admissible h. The proposition is that u^* can often be found by solving an associated equation of the form $F_d[u^*] = 0$, where d is a linearly nullable series. In which case, the unique nulling input u^* for d can be directly computed using the formal power series methods in [10]. The method is illustrated by two simple examples involving a linear and a nonlinear system. A more general and complete treatment will be deferred to a future publication.

Example 11: Let F_c be a double integrator system with $c = \alpha_0 + \alpha_1 x_0 x_1$ and $\alpha_i \in \mathbb{R}$. Since $\delta c = \alpha_1 x_0 \delta x_1$, it follows that

$$\delta J[u][h] = \int_0^T \alpha_1 F_c[u](\tau) F_{x_0 \delta x_1}[h](\tau) + W u(\tau) h(\tau)\, d\tau.$$

Defining $\tilde{h} = F_{x_0 \delta x_1}[h]$ so that $\tilde{h}'' = h$, the corresponding Euler's equation is

$$\alpha_1 F_c[u^*] + W \frac{d^2 u^*}{dt^2} = 0,$$

where the only fixed boundary condition is $u^*(T) = 0$ since $\tilde{h}(T) \neq 0$ (see [9, p. 26, p. 42]). Integrating both sides of this integro-differential equation three times yields $F_d[u^*] = 0$, where

$$d = a_0 + a_1 x_0 + a_2 x_0^2 + \frac{\alpha_0 \alpha_1}{W} x_0^3 + x_1 + \frac{\alpha_1^2}{W} x_0^4 x_1,$$

and the $a_i \in \mathbb{R}$ are free integration constants. Observe that d has relative degree 1. So if $a_0 = 0$ and $\alpha_0 \alpha_1 \neq 0$, then by Theorem 1 the series d is linearly nullable. With $b := \alpha_1^2 / W$ and any $T > 0$ such that $u^*(T) = 0$, the optimal solution is

$$u^*(t) = \sum_{k=0}^{\infty} (-b)^k \left(-a_1 \frac{t^{4k}}{(4k)!} - a_2 \frac{t^{4k+1}}{(4k+1)!} - \frac{\alpha_0 \alpha_1}{W} \frac{t^{4k+2}}{(4k+2)!} \right) \quad (19)$$

with $u^*(0) = -a_1$ and $du^*(0)/dt = -a_2$. This result can be checked independently using standard LQR theory for the case where $T = \infty$. Suppose $\alpha_0 = 2$, $\alpha_1 = 2$, and $W = 4$ as in the MATLAB code below.

```
alpha0=2; alpha1=2; W=4;
A=[0 1;0 0]; B=[0 alpha1]'; C=[1 0]; z0=[alpha0 0]';
[K,S,P] = lqry(ss(A,B,C,0),1/2,W/2);
[NU,DU]=ss2tf(A-B*K,z0,-1*K,0)
```

The computed $U^*(s) = -s/(s^2 + \sqrt{2}s + 1)$ yields $u^*(0) = -1$, $du^*(0)/dt = \sqrt{2}$, and $u^*(\infty) = 0$. Using these initial conditions, the Taylor series for u^* about $t = 0$ is exactly as

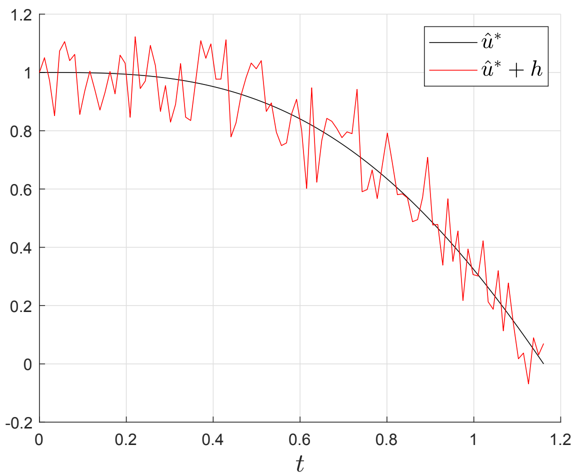

Fig. 4. Plot of $\hat{u}^*$ and a perturbation $\hat{u}^* + h$ in Example 12.

given in (19). □

Example 12: Suppose F_c has a quadratic nonlinearity with $c = \alpha_0 x_0 + \alpha_1 x_1^2$. Then $\delta c = \alpha_1 x_1 \sqcup\!\sqcup\, \delta x_1$ and

$$\begin{aligned}\delta J[u][h] &= \int_0^T \alpha_1 F_{c \sqcup\!\sqcup\, x_1}[u](\tau) F_{\delta x_1}[h](\tau) + W u(\tau) h(\tau)\, d\tau \\ &= \int_0^T (\alpha_0 \alpha_1 F_{x_0 x_1 + x_1 x_0}[u](\tau) + 3\alpha_1^2 F_{x_1^3}) \tilde{h}(\tau) \\ &\qquad + W u(\tau) \tilde{h}'(\tau)\, d\tau,\end{aligned}$$

where $\tilde{h} := F_{\delta x_1}[h]$. The corresponding Euler's equation is

$$\alpha_0 \alpha_1 F_{x_0 x_1 + x_1 x_0}[u^*] + 3\alpha_1^2 F_{x_1^3}[u^*] - W \frac{du^*}{dt} = 0$$

with $u^*(T) = 0$. Integrating both side of this equation twice gives $F_d[u^*] = 0$, where

$$d = a_0 + a_1 x_0 - x_1 + \frac{\alpha_0 \alpha_1}{W}(x_0^3 x_1 + x_0^2 x_1 x_0) + \frac{3\alpha_1^2}{W} x_0^2 x_1^3.$$

Series d has relative degree 1 so that if $a_0 = 0$ and $a_1 \neq 0$, then d is linearly nullable. For example, when $\alpha_0 = -7$, $\alpha_1 = 1$, $W = 3$, $a_0 = 0$, and $a_1 = 1$, it follows that

$$u^*(t) = 1 - \frac{14}{3}\frac{t^3}{3!} + \frac{t^4}{4!} + \frac{490}{9}\frac{t^6}{6!} - 84\frac{t^7}{7!} + 21\frac{t^8}{8!} - \cdots$$

Fig. 4 shows a polynomial approximation of u^*, namely, $\hat{u}^*$, derived by truncating the series starting at degree 13. In this case, $\hat{u}^*(T) = 5.40865 \times 10^{-7}$ when $T = 1.160697$. A numerical estimate of the extremal value is $J[\hat{u}^*] = 12.82142$. To verify that u^* is an extremal, $\delta J[\hat{u}^*][h] = J[\hat{u}^* + h] - J[\hat{u}^*]$ was computed numerically for 100 Monte Carlo runs, where h is a zero mean Gaussian random process on $(0, T]$ with standard deviation $\sigma = 0.1$ (see Fig. 4). The mean value of $\delta J[\hat{u}^*][h]$ was on the order of 10^{-3}, and the standard deviation was on the order of 10^{-2}. Thus, $\delta J[\hat{u}^*][h] \approx 0$ as expected. □

VII. Conclusions

The Fréchet derivative of a Chen-Fliess series was described along with an algebraic tool for computing it, namely the formal derivative of its generating series. It was next shown how to characterize and compute critical points of this Fréchet derivative both analytically and numerically. The former required that the generating series for the Fréchet derivative be shuffle separable and employed the concept of a nullable series. If the shuffle factor over $\mathbb{R}\langle\langle X \rangle\rangle$ was linearly nullable, for example, then the Taylor series for the unique critical point could be computed explicitly. Finally, two simple examples were provided to show how these ideas can be applied to solve quadratic optimal control problems entirely in the context of Chen-Fliess series.

Acknowledgment

The first and third authors were supported by the Broad Agency Announcement Program and the Cold Regions Research and Engineering Laboratory (ERDC-CRREL) under Contract No. W913E521C0003.

References

[1] A. E. Bryson, Jr. and Y.-C. Ho, *Applied Optimal Control*, Hemisphere Publishing Company, Washington, DC, 1975.
[2] T. Crespo and Z. Hajto, *Algebraic Groups and Differential Galois Theory*, American Mathematical Society, Providence, RI, 2011.
[3] R. Dahmen, W. S. Gray and A. Schmeding, Continuity of Chen-Fliess series in identification and machine learning, *Proc. 24th Inter. Symp. on Mathematical Theory of Networks and Systems*, Cambridge, UK, 2021, pp. 231–238.
[4] J. Dieudonné, *Foundations of Modern Analysis*, Academic Press, New York, 1969.
[5] M. Fliess, Fonctionnelles causales non linéaires et indéterminées non commutatives, *Bull. Soc. Math. France*, 109 (1981) 3–40.
[6] M. Fliess, On the concept of derivatives and Taylor expansions for nonlinear input-output systems, *Proc. 22nd IEEE Conf. on Decision and Control*, San Antonio, TX, 1983, pp. 643–646.
[7] M. Fliess, Nonlinear control theory and differential algebra, in *Modelling and Adaptive Control*, C. I. Byrnes and A. B. Kurzhanski, Eds., Springer, Berlin Heidelberg, 1988, pp. 134–145.
[8] M. Fliess, Vers une notion de dérivation fonctionnelle causale, *Ann. Inst. Henri Poincaré*, 3 (1986) 67–76.
[9] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice-Hall, Englewood Cliffs, NJ, 1963.
[10] W. S. Gray, L. A. Duffaut Espinosa, and M. Thitsa, Left inversion of analytic nonlinear SISO systems via formal power series methods, *Automatica*, 50 (2014) 2381–2388.
[11] W. S. Gray, K. Ebrahimi-Fard, and A. Schmeding, Universal zero dynamics: SISO case, *Proc. 55th Conf. on Information Sciences and Systems*, Baltimore, MD, 2021.
[12] W. S. Gray, K. Ebrahimi-Fard, and A. Schmeding, Zeroing the output of nonlinear systems without relative degree, *Proc. 57th Conf. on Information Science and Systems*, Baltimore, MD, 2023.
[13] W. S. Gray and Y. Wang, Fliess operators on L_p spaces: Convergence and continuity, *Systems Control Lett.*, 46 (2002) 67–74.
[14] A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill, New York, 1970.
[15] J. K. Hunter and B. Nachtergaele, *Applied Analysis*, World Scientific, Singapore, 2001.
[16] A. Isidori, *Nonlinear Control Systems*, 3rd Ed., Springer, London, 1995.
[17] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer, New York, 1990.
[18] I. Perez Avellaneda and L. A. Duffaut Espinosa, On mixed-monotonicity of Chen-Fliess series, *Proc. 26th Inter. Conf. on System Theory, Control and Computing*, Sinaia, Romania, 2022, pp. 98–103.
[19] I. Perez Avellaneda and L. A. Duffaut Espinosa, Output reachability of Chen-Fliess series: A Newton-Raphson approach, *Proc. 57th Conf. on Information Science and Systems*, Baltimore, MD, 2023.
[20] R. Ree, Lie elements and an algebra associated with shuffles, *Ann. of Math. (2)*, 68 (1958) 210–220.
[21] E. Zeidler, *Nonlinear Functional Analysis and its Applications: Part 1: Fixed-Point Theorems*, Springer, New York, 1985.