

# Decompositions of Nonlinear Input-Output Systems to Zero the Output

W. Steven Gray<sup>a,\*</sup>, Kurusch Ebrahimi-Fard<sup>b</sup>, Alexander Schmeding<sup>b</sup>

<sup>a</sup>*Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, Virginia 23529, USA*

<sup>b</sup>*Department of Mathematical Sciences, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway*

---

## Abstract

Consider an input-output system where the output is the tracking error given some desired reference signal. It is natural to consider under what conditions the problem has an exact solution, that is, the tracking error is exactly the zero function. If the system has a well defined relative degree and the zero function is in the range of the input-output map, then it is well known that the system is locally left invertible, and thus, the problem has a unique exact solution. **A system will fail to have relative degree when more than one exact solution exists. The general goal of this paper is to describe a decomposition of an input-output system having a Chen-Fliess series representation into a parallel product of subsystems in order to identify possible solutions to the problem of zeroing the output.** For computational purposes, the focus is on systems whose generating series are polynomials. It is shown that the shuffle algebra on the set of generating polynomials is a unique factorization domain so that any polynomial can be uniquely factored modulo a permutation into its irreducible elements for the purpose of identifying the subsystems in a parallel product decomposition. This is achieved using the fact that this shuffle algebra is isomorphic to the symmetric algebra over the vector space spanned by Lyndon words. A specific algorithm for factoring generating polynomials into its irreducible factors is presented based on the Chen-Fox-Lyndon factorization of words.

*Key words:* nonlinear control systems, Chen-Fliess series, shuffle algebra

---

## 1. Introduction

Consider a smooth control-affine state space realization

$$\dot{z} = g_0(z) + g_1(z)u, \quad z(0) = z_0 \quad (1a)$$

$$y = h(z), \quad (1b)$$

where  $g_0$ ,  $g_1$ , and  $h$  are defined on  $W \subseteq \mathbb{R}^n$ . If the realization has a well defined relative degree at  $z_0 \in W$ , then it is a classical result that the corresponding single-input, single-output map  $F : u \mapsto y$  is left invertible on a neighborhood of  $z_0$  [20, 26]. If the zero output is known to be in the range of  $F$  for some class of inputs  $\mathcal{U}$ , then there exists a unique input  $u^* \in \mathcal{U}$  satisfying  $F[u^*] = 0$  which can be generated in real-time using feedback [20, 26] or computed analytically using formal power series methods [12]. This construction leads to the notion of *zero dynamics* [14, 20, 21, 26] and has well known applications in output tracking and path planning problems when the output is taken to be an error signal [12, 21, 26]. The problem of zeroing the output can also be applied in optimal control problems in order to determine critical points of functional derivatives [5].

System (1) can fail to have relative degree under a number of different circumstances. For example, it can turn out that the Lie derivative  $L_{g_1}h$  is exactly zero at  $z_0$ . As explained in [32, Remark 3], it is still possible for the system to be left invertible in a certain sense if the state trajectory immediately leaves the singularity after  $t = 0$ , and two inputs are said to be equivalent when they differ only on a set of measure zero. Another possibility is the case where the input-output map is simply not left-invertible. Take as a simple example the system

$$\dot{z}_1 = 1 - u, \quad \dot{z}_2 = z_3 - u, \quad \dot{z}_3 = 1, \quad z(0) = 0 \quad (2a)$$

$$y = z_1 z_2. \quad (2b)$$

It is easily verified that this realization does not have relative degree at the origin. The input-output map is not injective since there are two inputs which give the zero output:  $u_1^*(t) = 1, t \geq 0$  and  $u_2^*(t) = t, t \geq 0$ . This is a result of the fact that the system can be decomposed into a parallel product of two subsystems  $F_1 : u \mapsto y_1 = z_1$  and  $F_2 : u \mapsto y_2 = z_2$  as shown in Figure 1, where each subsystem has the zero function in its range. **The general goal of this paper is to describe how to perform this decomposition in order to identify possible solutions to the problem of zeroing the output.** This problem is only non-trivial in the nonlinear setting since linear time-invariant systems always have relative degree.

The approach taken will be to work purely in the input-output setting using Chen-Fliess series representations.

---

\*Corresponding author

Email addresses: `sgray@odu.edu` (W. Steven Gray),  
`kurusch.ebrahimi-fard@ntnu.no` (Kurusch Ebrahimi-Fard),  
`alexander.schmeding@ntnu.no` (Alexander Schmeding)

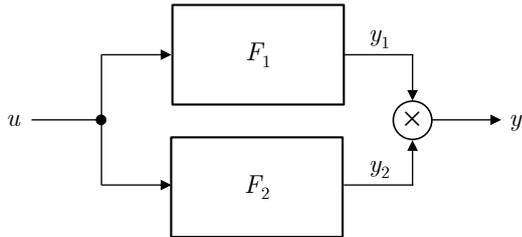


Figure 1: Parallel product connection of systems  $F_1$  and  $F_2$

One advantage to this point of view is that the nonuniqueness of coordinate systems can be avoided. That is, the generating series for the input-output map of a state space realization is invariant under coordinate transformation. In addition, this framework is more general as every analytic state space realization (1) has an input-output map with a Chen-Fliess series representation but not conversely. In order to avoid convergence issues associated with such series, the analysis will be done using *formal* Fliess operators [17], that is, maps that take an infinite jet representing a formal input function to an infinite jet representing a formal output function. In this context, the problem of zeroing the output boils down to a purely algebraic problem.

The concept of a *nullable* generating series is presented first. This is a formal power series representing a formal Fliess operator having the property that the zero output (jet) is in the range of the operator. A generating series is called *strongly nullable* if there is a nonzero input that maps to the zero output and *primely nullable* if this input is the only input with this property. A special class of primely nullable series are those having relative degree and one additional property. These will be called *linearly nullable*. While there is no known general test for nullability, linearly nullable series can be completely characterized, and their nulling inputs can be computed directly using the methods in [12]. A key fact is that the parallel product of two Chen-Fliess series has a Chen-Fliess series representation whose generating series is the shuffle product of the generating series of the subsystems [8]. It will be shown that the shuffle product of two linearly nullable series is always strongly nullable but never linearly nullable. On the other hand, a given series may not be linearly nullable but if any of its shuffle factors is linearly nullable, then the problem of zeroing the output can be directly solved since nulling one factor will zero the system output. The main issue then becomes how to factor a generating series into its irreducible elements in the shuffle algebra.

The algorithmic focus in this paper will be on the polynomial case. It is first established that the shuffle algebra as a commutative polynomial ring over a finite set of noncommuting indeterminates is a unique factorization domain. This is achieved by assembling existing results from algebra [4] and algebraic combinatorics [24]. Of particular importance is the fact that this shuffle algebra can be

viewed as the symmetric algebra over the  $\mathbb{R}$ -vector space spanned by Lyndon words [24, 28]. Next, an algorithm is given to factor a polynomial into its irreducible shuffle components. This is done by first mapping the polynomial to the symmetric algebra using the Chen-Fox-Lyndon factorization of words [3, 19, 24, 28, 30]. The resulting polynomial is then factored using one of the many known algorithms for factoring multivariate commutative polynomials [33]. Then each factor is mapped back to the shuffle algebra.

The paper is organized as follows. In the next section, a brief summary is given of the mathematical tools used to establish the main results of the paper. In Section 3, the concept of nullable generating series is presented. The subsequent section addresses the problem of factoring generating series in the shuffle algebra. The final section provides the main conclusions of the paper. It should be stated that a shorter preliminary version of this paper was presented as a conference paper [15]. The present version includes three additional examples (Examples 3.6, 3.10, and 4.1), a proof of identity (6) (see appendix), a new Corollary 3.1, a revised and expanded version of the proof of Theorem 4.1, an expanded description of the factorization algorithm in Section 4, two new figures, and seven additional references. In particular, Example 3.10 provides a new application of the main results in the context of optimal control, while the revised proof of Theorem 4.1 is simpler and more direct.

## 2. Preliminaries

An *alphabet*  $X = \{x_0, x_1, \dots, x_m\}$  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  $\eta$ , 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. Any mapping  $c : X^* \rightarrow \mathbb{R}^\ell$  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  $c$ ,  $\text{supp}(c)$ , is the set of all words having nonzero coefficients. A series  $c$  is called *proper* if  $\emptyset \notin \text{supp}(c)$ . The *order* of  $c$ ,  $\text{ord}(c)$ , is the length of the shortest word in its support. By definition the order of the zero series is  $+\infty$ . 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

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

where  $x_i, x_j \in X$ ,  $\eta, \xi \in X^*$  and with  $\eta \sqcup \emptyset = \emptyset \sqcup \eta = \eta$  [8]. For any letter  $x_i \in X$ , let  $x_i^{-1}$  denote the  $\mathbb{R}$ -linear *left-shift operator* defined by  $x_i^{-1}(\eta) = \eta'$  when  $\eta = x_i \eta'$  and zero otherwise. Higher order shifts are defined inductively via  $(x_i \xi)^{-1}(\cdot) = \xi^{-1} x_i^{-1}(\cdot)$ , where  $\xi \in X^*$ . It acts as a derivation on the shuffle product.

### 2.1. Chen-Fliess series

Given any  $c \in \mathbb{R}^\ell \langle\langle X \rangle\rangle$  one can associate a causal  $m$ -input,  $\ell$ -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  $u : [t_0, t_1] \rightarrow \mathbb{R}^m$ , 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  $L_{\mathfrak{p}}^m[t_0, t_1]$  denote the set of all measurable functions defined on  $[t_0, t_1]$  having a finite  $\|\cdot\|_{\mathfrak{p}}$  norm and  $B_{\mathfrak{p}}^m(R)[t_0, t_1] := \{u \in L_{\mathfrak{p}}^m[t_0, t_1] : \|u\|_{\mathfrak{p}} \leq R\}$ . Assume  $C[t_0, t_1]$  is the subset of continuous functions in  $L_1^m[t_0, t_1]$ . Define inductively for each word  $\eta = x_i \bar{\eta} \in X^*$  the map  $E_\eta : L_1^m[t_0, t_1] \rightarrow C[t_0, t_1]$  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  $c \in \mathbb{R}^\ell \langle\langle X \rangle\rangle$  is defined in [8] as

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

If there exist real numbers  $K_c, M_c > 0$  such that

$$|(c, \eta)| \leq K_c M_c^{|\eta|} |\eta|!, \quad \forall \eta \in X^*, \quad (3)$$

then  $F_c$  constitutes a well defined mapping from  $B_{\mathfrak{p}}^m(R)[t_0, t_0 + T]$  into  $B_{\mathfrak{q}}^\ell(S)[t_0, t_0 + T]$  for sufficiently small  $R, T > 0$  and some  $S > 0$ , where the numbers  $\mathfrak{p}, \mathfrak{q} \in [1, \infty]$  are conjugate exponents, i.e.,  $1/\mathfrak{p} + 1/\mathfrak{q} = 1$  [16]. (Here,  $|z| := \max_i |z_i|$  when  $z \in \mathbb{R}^\ell$ .) Any series  $c$  satisfying (3) is called *locally convergent*, and  $F_c$  is called a *Fliess operator*. The subset of all locally convergent series is denoted by  $\mathbb{R}_{LC}^\ell \langle\langle X \rangle\rangle$ .

A Fliess operator  $F_c$  defined on  $B_{\mathfrak{p}}^m(R)[t_0, t_0 + T]$  with  $\ell = 1$  is said to be *realizable* when there exists a state space realization (1) with each  $g_i$  being an analytic vector field expressed in local coordinates on some neighborhood  $W$  of  $z_0 \in \mathbb{R}^n$ , and the real-valued output function  $h$  is an analytic function on  $W$  such that (1a) has a well defined solution  $z(t)$ ,  $t \in [t_0, t_0 + T]$  for any given input  $u \in B_{\mathfrak{p}}^m(R)[t_0, t_0 + T]$ , and  $y(t) = F_c[u](t) = h(z(t))$ ,  $t \in [t_0, t_0 + T]$ . Denoting the *Lie derivative* of  $h$  with respect to  $g_i$  by  $L_{g_i} h$ , it can be shown that for any word  $\eta = x_{i_k} \cdots x_{i_1} \in X^*$

$$(c, \eta) = L_{g_\eta} h(z_0) := L_{g_{i_1}} \cdots L_{g_{i_k}} h(z_0) \quad (4)$$

[8, 20, 26].

### 2.2. System interconnections

Given Fliess operators  $F_c$  and  $F_d$ , where  $c, d \in \mathbb{R}_{LC}^\ell \langle\langle X \rangle\rangle$ , the parallel and product connections satisfy  $F_c + F_d = F_{c+d}$  and  $F_c F_d = F_{c \sqcup d}$ , respectively [8]. It is also known that the composition of two Fliess operators  $F_c$  and  $F_d$  with  $c \in \mathbb{R}_{LC}^\ell \langle\langle X \rangle\rangle$  and  $d \in \mathbb{R}_{LC}^m \langle\langle X \rangle\rangle$  always yields another Fliess operator with generating series  $c \circ d$ , where the *composition product* is given by

$$c \circ d = \sum_{\eta \in X^*} (c, \eta) \psi_d(\eta) \mathbf{1} \quad (5)$$

[7]. Here  $\psi_d$  is the continuous (in the ultrametric sense) algebra homomorphism from  $\mathbb{R} \langle\langle X \rangle\rangle$  to the vector space endomorphisms on  $\mathbb{R} \langle\langle X \rangle\rangle$ ,  $\text{End}(\mathbb{R} \langle\langle X \rangle\rangle)$ , uniquely specified by  $\psi_d(x_i \eta) = \psi_d(x_i) \circ \psi_d(\eta)$  with  $\psi_d(x_i)(e) = x_0(d_i \sqcup e)$ ,  $i = 0, 1, \dots, m$  for any  $e \in \mathbb{R} \langle\langle X \rangle\rangle$ , and where  $d_i$  is the  $i$ -th component series of  $d$  ( $d_0 := \mathbf{1} := 1\emptyset$ ). By definition,  $\psi_d(\emptyset)$  is the identity map on  $\mathbb{R} \langle\langle X \rangle\rangle$ . The left-shift operators distribute over the composition product as follows:

$$x_j^{-1}(c \circ d) = \begin{cases} x_0^{-1}(c) \circ d + \sum_{i=1}^m d_i \sqcup (x_i^{-1}(c) \circ d) & : j = 0 \\ 0 & : j \neq 0. \end{cases} \quad (6)$$

(See the appendix for a proof of this property.) If  $c, d \in \mathbb{R} \langle\langle X \rangle\rangle$  with  $m = \ell = 1$  and  $d$  non-proper, then one can define the quotient  $c/d = c \sqcup d^{\sqcup -1}$  so that  $F_c/F_d = F_{c/d}$  with the shuffle inverse of  $d$  defined as

$$d^{\sqcup -1} = ((d, \emptyset)(1 - d'))^{\sqcup -1} = (d, \emptyset)^{-1} (d')^{\sqcup *},$$

where  $d' = \mathbf{1} - (d/(d, \emptyset))$  is proper and  $(d')^{\sqcup *} := \sum_{k \geq 0} (d')^{\sqcup k}$  [12]. The following lemma will be useful.

**Lemma 2.1.** *For any  $c, d, e \in \mathbb{R} \langle\langle X \rangle\rangle$  with  $d$  non-proper, the following identity holds*

$$(c/d) \circ e = (c \circ e)/(d \circ e).$$

*Proof:* It can be shown directly from the definition of the composition product that if  $d$  is non-proper then so is  $d \circ e$ . In fact,  $(d \circ e, \emptyset) = (d, \emptyset) \neq 0$ . Thus, both sides of the equality in question are at least well defined formal power series. In light of the known identity

$$(c \sqcup d) \circ e = (c \circ e) \sqcup (d \circ e) \quad (7)$$

for any  $c, d, e \in \mathbb{R} \langle\langle X \rangle\rangle$  [9], it is sufficient to show that

$$d^{\sqcup -1} \circ e = (d \circ e)^{\sqcup -1}. \quad (8)$$

It is clear via induction that for any  $k \in \mathbb{N}$ ,

$$d^{\sqcup k} \circ e = (d \circ e)^{\sqcup k}.$$

Therefore, since  $d$  is non-proper, it follows that

$$d^{\sqcup -1} \circ e = (d, \emptyset)^{-1} \lim_{n \rightarrow \infty} \sum_{k=0}^n (d')^{\sqcup k} \circ e$$

$$\begin{aligned}
&= (d \circ e, \emptyset)^{-1} \lim_{n \rightarrow \infty} \sum_{k=0}^n (d' \circ e)^{\sqcup k} \\
&= (d \circ e)^{\sqcup -1}.
\end{aligned}$$

As  $d'$  and  $d' \circ e$  are both proper, all the limits above (in the ultrametric sense) exist, and thus, the claim is verified. ■

### 2.3. Formal Fliess operators

Suppose  $X = \{x_0, x_1\}$ ,  $X_0 = \{x_0\}$ , and define  $\mathbb{R}[[X_0]]$  to be the set of all (commutative) formal power series in the letter  $x_0$ . Then every series  $c_u \in \mathbb{R}[[X_0]]$  can be identified with an infinite jet  $j_{t_0}^\infty(u)$  for any fixed  $t_0 \in \mathbb{R}$ . By Borel's Lemma, there is a real-valued function  $u \in C^\infty(t_0)$  whose Taylor series corresponds to  $j_{t_0}^\infty(u)$ . In the event that the coefficients of  $c_u$  satisfy the growth bound (3), then  $u$  is real analytic. In which case, for any  $c \in \mathbb{R}_{LC}\langle\langle X \rangle\rangle$ ,  $F_{c_y}[v] = y = F_c[u] = F_c[F_{c_u}[v]] = F_{c \circ c_u}[v]$ , where  $v$  is just a placeholder in this chain of equalities. If the Taylor series for  $u$  does not converge, it is viewed as a formal function. Nevertheless, the mapping  $c \circ : \mathbb{R}[[X_0]] \rightarrow \mathbb{R}[[X_0]] : c_u \mapsto c_y = c \circ c_u$  is still well defined and takes the input infinite jet to the output infinite jet. This is called a *formal Fliess operator* [17]. The advantage of working with these formal objects is that their algebraic properties can be characterized independently of their analytic nature. This will be the approach taken below.

### 2.4. Relative degree of a generating series

Observe that  $c \in \mathbb{R}\langle\langle X \rangle\rangle$  can always be decomposed into its natural and forced components, that is,  $c = c_N + c_F$ , where  $c_N := \sum_{k \geq 0} (c, x_0^k) x_0^k$  and  $c_F := c - c_N$ .

**Definition 2.1.** [12] Given  $c \in \mathbb{R}\langle\langle X \rangle\rangle$  with  $X = \{x_0, x_1\}$ , let  $r \geq 1$  be the largest integer such that  $\text{supp}(c_F) \subseteq x_0^{r-1} X^*$ . Then  $c$  has relative degree  $r$  if the linear word  $x_0^{r-1} x_1 \in \text{supp}(c)$ , otherwise it is not well defined.

It is immediate that  $c$  has relative degree  $r$  if and only if there exists some  $e \in \mathbb{R}\langle\langle X \rangle\rangle$  with  $\text{supp}(e) \subseteq X^*/\{X_0^*, x_1\}$  such that

$$c = c_N + c_F = c_N + K x_0^{r-1} x_1 + x_0^{r-1} e \quad (9)$$

and  $K \neq 0$ . This notion of relative degree coincides with the usual definition given in a state space setting [13].

## 3. Nullable Generating Series

It is assumed for the remainder of the paper that all systems are single-input, single-output, i.e.,  $m = \ell = 1$  so that  $X = \{x_0, x_1\}$  and all series coefficients are real-valued. Consider the following classes of generating series.

**Definition 3.1.** 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  $c \circ : \mathbb{R}[[X_0]] \rightarrow \mathbb{R}[[X_0]]$ ,  $c_u \mapsto c \circ c_u$ . 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.

Observe that from (5) it follows that  $(c \circ c_u, \emptyset) = (c, \emptyset)$  for all  $c_u \in \mathbb{R}[[X_0]]$ . Thus, if  $c$  is nullable, then necessarily  $c$  must be proper. Also, every series  $c = c_F$  satisfies  $c \circ 0 = 0$ . Thus, it is nullable. If  $c = c_N + c_F$  with  $c_N \neq 0$ , then  $c \circ 0 = c_N$ . Therefore, if  $c$  is nullable, it must be strongly nullable.

**Example 3.1.** Observe that  $c = x_0^2 - x_1 x_0$  is primely nullable since  $c \circ \mathbf{1} = x_0^2 - x_0^2 = 0$ , and  $c_{u^*} = \mathbf{1}$  is the only series with this property. □

**Example 3.2.** The polynomial  $c = x_0 + x_0 x_1$  is not nullable since  $c \circ c_u = x_0 + x_0^2 c_u \neq 0$  for all  $c_u \in \mathbb{R}[[X_0]]$ . □

A sufficient (but not necessary) condition for a series to be primely nullable is given in the following theorem. It has its roots in [12]. A concise proof is given here to make its computational features explicit and the subsequent corollary more apparent.

**Theorem 3.1.** If  $c \in \mathbb{R}\langle\langle X \rangle\rangle$  has relative degree  $r$ , and  $\text{supp}(c_N) \subseteq x_0^r X_0^*$  is nonempty, then  $c$  is primely nullable.

*Proof:* Since  $c_N \neq 0$  by assumption, any nulling series must be nonzero. The claim is that  $c$  has a unique nonzero nulling series. Applying (6) to  $c_y = c \circ c_u$  with  $m = 1$  (let  $d_1 = d$ ) under the assumption that  $c$  has relative degree  $r$  gives

$$\begin{aligned}
c_y &= c \circ c_u \\
x_0^{-1}(c_y) &= x_0^{-1}(c) \circ c_u \\
&\vdots \\
x_0^{-r+1}(c_y) &= x_0^{-r+1}(c) \circ c_u \\
x_0^{-r}(c_y) &= x_0^{-r}(c) \circ c_u + c_u \sqcup ((x_0^{r-1} x_1)^{-1}(c) \circ c_u).
\end{aligned}$$

Since  $(x_0^{r-1} x_1)^{-1}(c)$  is non-proper (specifically,  $((x_0^{r-1} x_1)^{-1}(c), \emptyset) = K \neq 0$  in (9)) it can be shown that  $(x_0^{r-1} x_1)^{-1}(c) \circ c_u$  is also non-proper and thus has a shuffle inverse. Setting  $x_0^{-r}(c_y) = 0$  and dividing by  $(x_0^{r-1} x_1)^{-1}(c) \circ c_u$  gives

$$0 = (x_0^{-r}(c) \circ c_u) / ((x_0^{r-1} x_1)^{-1}(c) \circ c_u) + c_u.$$

Next, applying Lemma 2.1 yields

$$0 = (x_0^{-r}(c) / (x_0^{r-1} x_1)^{-1}(c)) \circ c_u + c_u.$$

Define a generalized series  $\delta$  with the defining property that  $F_\delta[u] = u$  for all admissible inputs  $u$ . Then it must

have the unital property  $\delta \circ c = c \circ \delta = c$  on the semigroup  $(\mathbb{R}\langle\langle X \rangle\rangle, \circ)$ . The previous equation can be written as

$$0 = \underbrace{(\delta + (x_0^{-r}(c)/(x_0^{r-1}x_1)^{-1}(c)))}_{:=d_\delta} \circ c_u.$$

It is known that the set of series  $\delta + \mathbb{R}\langle\langle X \rangle\rangle$  forms a group under the induced composition product [11]. Therefore, one can solve for  $c_u$  directly via left inversion to give  $c_u = d_\delta^{-1} \circ 0$ . In which case, there exists a unique  $c_u$  that will zero out all the coefficients of  $c_y$  with the exception of the first  $r$  coefficients. These initial coefficients are completely determined by  $c$  since

$$\begin{aligned} (c_y, x_0^k) &= (x_0^{-k}(c_y), \emptyset) = (x_0^{-k}(c) \circ c_u, \emptyset) \\ &= (x_0^{-k}(c), \emptyset) = (c, x_0^k), \quad k = 0, 1, \dots, r-1. \end{aligned}$$

By assumption  $\text{supp}(c_N) \subseteq x_0^r X_0^*$ . Hence, all the coefficients above must be zero so that  $c_y = 0$  as desired. ■

It is worth noting that  $(\mathbb{R}\langle\langle X \rangle\rangle, \circ, \delta)$  described above as well as  $(\mathbb{R}\langle\langle X \rangle\rangle, \sqcup, \mathbf{1})$  both include the monoids of characters over their respective graded connected bialgebras of coordinate functions. Identity (7), which is central in this work, can then be viewed in terms of the concept of two bialgebras in cointeraction [10]. In this respect, equation (8) is equivalent to stating that the right action of the character monoid  $(\mathbb{R}\langle\langle X \rangle\rangle, \circ, \delta)$  on the group of unital non-proper series  $(\mathbf{1} + \mathbb{R}_{np}\langle\langle X \rangle\rangle, \sqcup, \mathbf{1}) \subset (\mathbb{R}\langle\langle X \rangle\rangle, \sqcup, \mathbf{1})$  is compatible with the antipode of its Hopf algebra of coordinate functions.

Series satisfying the condition in Theorem 3.1 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. In light of (9), every such series has the form

$$c = x_0^r e_0 + K x_0^{r-1} x_1 + x_0^{r-1} e_1,$$

where  $r \in \mathbb{N}$ ,  $K \neq 0$ ,  $e_0 \in \mathbb{R}[[X_0]]/\{0\}$ , and  $\text{supp}(e_1) \subseteq X^*/\{X_0^*, x_1\}$ . From the general identity  $(x_0^k c) \circ d = x_0^k (c \circ d)$  and the fact that the composition product is left linear, it follows that

$$\begin{aligned} c \circ c_{u^*} &= (x_0^r e_0 + K x_0^{r-1} x_1 + x_0^{r-1} e_1) \circ c_{u^*} \\ &= x_0^{r-1} ((x_0 e_0 + K x_1 + e_1) \circ c_{u^*}) \\ &= 0. \end{aligned}$$

That is,  $c_{u^*}$  is a solution to

$$(x_0 e_0 + K x_1 + e_1) \circ c_{u^*} = x_0 (e_0 + K c_{u^*}) + e_1 \circ c_{u^*} = 0.$$

Central to the proof of Theorem 3.1 is the observation that mapping  $c \circ$  under the condition that  $c$  has relative degree is injective since it is left invertible. The following corollary, which also follows directly from this proof, states that  $c \circ$  is never surjective on  $\mathbb{R}[[X_0]]$  under this condition.

**Corollary 3.1.** *Suppose  $c$  has relative degree  $r$ . Define  $c_N^{r-1} = (c, \emptyset) + (c, x_0)x_0 + \dots + (c, x_0^{r-1})x_0^{r-1}$ . Then the range of the mapping  $c_u \mapsto c \circ c_u$  is the affine subspace of the  $\mathbb{R}$ -vector space  $\mathbb{R}[[X_0]]$*

$$R_c := \{c_y = c_N^{r-1} + x_0^r e : e \in \mathbb{R}[[X_0]]\}.$$

Therefore,  $c$  is nullable in this case only if  $c_N^{r-1} = 0$ .

**Example 3.3.** The polynomial  $c = x_0 + x_1$  has relative degree 1 and  $c_N = x_0 \in x_0 X_0^*$ . Therefore, it is linearly nullable. Specifically,  $c_{u^*} = -\mathbf{1}$  is the only series that yields  $c \circ c_{u^*} = 0$ . □

**Example 3.4.** The polynomial  $c = x_0^2 - x_1 x_0$  in Example 3.1 does not have relative degree since it does not have a linear word in its support (i.e.,  $K = 0$  in (9)). So it is primely nullable but not linearly nullable. □

**Example 3.5.** The polynomial  $c = x_0 + x_0 x_1$  in Example 3.2 has relative degree 2 and was shown not to be nullable. Observe  $c_N = x_0 \notin x_0^2 X_0^*$ , which is consistent with Theorem 3.1. □

**Example 3.6.** Consider the series  $c = \sum_{\eta \in X^+} |\eta|! \eta$ , where  $X^+ := X^*/\{\emptyset\}$ . The series has relative degree 1 and is linearly nullable. In this instance, the corresponding Chen-Fliess series has the closed-form expression

$$F_c[u] = \frac{F_{x_0+x_1}[u]}{1 - F_{x_0+x_1}[u]}.$$

Therefore, the unique nulling series for  $c$  is  $c_{u^*} = -\mathbf{1}$ . □

Let  $c \in \mathbb{R}\langle\langle X \rangle\rangle$  be nullable. Define the (two-sided) principal ideal

$$I_c = (c) := \{c \sqcup d : d \in \mathbb{R}\langle\langle X \rangle\rangle\}$$

in the shuffle algebra on  $\mathbb{R}\langle\langle X \rangle\rangle$ .

**Lemma 3.1.** *Every series in  $I_c$  is nullable. If  $c$  is strongly nullable, then every series in  $I_c$  is strongly nullable.*

*Proof:* Applying (7) it follows that  $(c \sqcup d) \circ c_{u^*} = (c \circ c_{u^*}) \sqcup (d \circ c_{u^*}) = 0$  if  $c_{u^*}$  is selected so that  $c \circ c_{u^*} = 0$ , which is always possible since  $c$  is nullable by assumption. The second claim is now obvious. ■

The first theorem below is obvious given the definition of primely nullable. The second theorem is less trivial but not unexpected. It confirms that the set of linearly nullable series is not closed under the shuffle product.

**Theorem 3.2.** *If  $c, d \in \mathbb{R}\langle\langle X \rangle\rangle$  are primely nullable with  $c_{u^*} \neq d_{u^*}$ , then  $c \sqcup d$  is strongly nullable but not primely nullable.*

**Theorem 3.3.** *If  $c, d \in \mathbb{R}\langle\langle X \rangle\rangle$  are linearly nullable, then  $c \sqcup d$  is strongly nullable but not linearly nullable.*

*Proof:* The strong nullability property follows directly from the lemma above. Regarding the second assertion, if  $c \sqcup d$  is linearly nullable, then necessarily  $c \sqcup d$  must have relative degree, say  $s$ , and  $(c \sqcup d)_N \in x_0^s X_0^*$ . Observe that

$$c \sqcup d = (x_0^{r_c} e_0 + K_c x_0^{r_c-1} x_1 + x_0^{r_c-1} e_1) \sqcup (x_0^{r_d} f_0 + K_d x_0^{r_d-1} x_1 + x_0^{r_d-1} f_1)$$

has the property that  $(c \sqcup d)_N \in x_0^{r_c+r_d} X_0^*$ . But the assertion is that  $c \sqcup d$  cannot have relative degree  $r_c + r_d$ . This would require that the shortest linear word in  $\text{supp}(c \sqcup d)_F$  be  $x_0^{r_c+r_d-1} x_1$  and all other words in  $\text{supp}(c \sqcup d)_F$  must have the prefix  $x_0^{r_c+r_d-1}$ . This linear word will only be present if

$$K_c(f_0, \emptyset) + K_d(e_0, \emptyset) \neq 0. \quad (10)$$

This means that at least one of the constant terms  $(e_0, \emptyset)$  or  $(f_0, \emptyset)$  must be nonzero. In addition, note that every word in the support of

$$(e_0, \emptyset) x_0^{r_c} \sqcup K_d x_0^{r_d-1} x_1 + (f_0, \emptyset) x_0^{r_d} \sqcup K_c x_0^{r_c-1} x_1 \\ = K_d(e_0, \emptyset) (x_0^{r_c} \sqcup x_0^{r_d-1} x_1) + K_c(f_0, \emptyset) (x_0^{r_d} \sqcup x_0^{r_c-1} x_1)$$

has length  $r_c + r_d$ , and these words must have the required prefix  $x_0^{r_c+r_d-1}$  since no other words in the larger shuffle product are short enough to cancel these words. But the only way to remove an illegal word would violate (10). For example, if  $r_c = r_d = 1$ , then

$$(e_0, \emptyset) x_0 \sqcup K_d x_1 + (f_0, \emptyset) x_0 \sqcup K_c x_1 \\ = K_d(e_0, \emptyset) (x_0 x_1 + x_1 x_0) + K_c(f_0, \emptyset) (x_0 x_1 + x_1 x_0).$$

The illegal word  $x_1 x_0$  cannot be canceled without removing the required linear word  $x_0 x_1$ . Thus,  $c \sqcup d$  cannot be linearly nullable.  $\blacksquare$

**Example 3.7.** Suppose  $c = x_0 - x_1$  and  $d = x_0^2 - x_1$ . Both series are linearly nullable with relative degree 1. The nulling series for  $c$  is  $c_{u^*} = \mathbf{1}$ , and the nulling series for  $d$  is  $d_{u^*} = x_0$ . Observe

$$c \sqcup d = -x_0 x_1 - x_1 x_0 + 2x_1^2 + 3x_0^3 - x_0^2 x_1 - x_0 x_1 x_0 - x_1 x_0^2$$

does not have relative degree. Therefore  $c \sqcup d$  is strongly nullable, but not linearly nullable and not primely nullable. In fact, if the coefficients for the realization (2) are computed from (4), one will find directly that the generating series is the polynomial given above. This is the origin of the example given in the introduction.  $\square$

**Example 3.8.** Suppose  $c = x_0 + x_1$  and  $d = \mathbf{1} + x_1$ . In this case,  $c$  is linearly nullable with relative degree 1, and  $d$  also has relative degree 1 but is not nullable as it is not proper. Observe

$$c \sqcup d = x_0 + x_1 + x_0 x_1 + x_1 x_0 + 2x_1^2$$

is also linearly nullable with relative degree 1. That is, Theorem 3.3 does not preclude the possibility that primely nullable series can have shuffle factors that are not nullable.  $\square$

**Example 3.9.** Suppose  $c = d = x_0 - x_1$  so that both series are linearly nullable with relative degree 1. As expected,

$$c \sqcup d = 2x_0^2 - 2x_0 x_1 - 2x_1 x_0 - 2x_1^2$$

is not linearly nullable as it does not have relative degree, but it is primely nullable since  $c_{u^*} = d_{u^*} = \mathbf{1}$  is the only nulling series for  $c \sqcup d$  as the shuffle product is an integral domain. That is, in general  $(c \sqcup d) \circ e_u = (c \circ e_u) \sqcup (d \circ e_u) = 0$  if and only if at least one argument in the second shuffle product is the zero series.  $\square$

In summary, if  $\mathbb{R}_p \langle \langle X \rangle \rangle$  is the set of all proper series in  $\mathbb{R} \langle \langle X \rangle \rangle$ , then the following inclusions hold:

$$\mathbb{R}_p \langle \langle X \rangle \rangle \supset \text{nullable series} \supset \text{strongly nullable series} \\ \supset \text{primely nullable series} \supset \text{linearly nullable series}.$$

In light of Theorems 3.2 and 3.3, only the set of nullable series and strongly nullable series are closed under the shuffle product.

The final example provides an application of nullable series.

**Example 3.10.** In optimal control problems, it is often necessary to determine critical points of quadratic objective functions. From the calculus of variations, this is accomplished by computing the critical points of a variational derivative. For a system described only in terms of a Chen-Fliess series, this would involve determining the critical points of a variational derivative of a Chen-Fliess series. The Fréchet derivative of  $F_c$ , for example, can be computed by introducing a variational alphabet associated with  $X = \{x_0, x_1\}$ , say  $\delta X = \{\delta x_0, \delta x_1\}$ . Define the mapping  $\delta : X \rightarrow \delta X$  by  $\delta(x_0) = \delta x_0 = 0$  and  $\delta(x_1) = \delta x_1$ . Extend the definition of  $\delta$  to  $X^*$  by letting it act as a derivation with respect to concatenation. Further extend the definition to  $\mathbb{R} \langle \langle X \rangle \rangle$  by linearity. In which case, the Fréchet derivative of  $F_c$  at  $u$  is the linear functional  $DF_c[u][h] = F_{\delta(c)}[u, h]$  [5]. Consider the simple example where  $c = x_0 x_1 + x_1 x_0 + x_1^2$  so that  $\delta(c) = (x_0 + x_1) \sqcup \delta x_1$ . Identifying  $u$  with  $x_1$  and  $h$  with  $\delta x_1$  from some admissible set of functions  $\mathcal{U}$ , it follows that

$$DF_c[u][h] = F_{x_0+x_1}[u] E_{\delta x_1}[h],$$

where  $DF_c[u][\cdot]$  is clearly linear. Critical points in this context are the inputs  $u^* \in \mathcal{U}$  such that  $DF_c[u^*][h] = 0$  for all  $h \in \mathcal{U}$ . Here it is evident since  $x_0 + x_1$  is linearly nullable that  $u^*(t) = -1, t \geq 0$ .  $\square$

#### 4. Factorizations in the Shuffle Algebra

The shuffle product on  $\mathbb{R} \langle X \rangle$  forms a commutative ring. Such structures appear in the following chain of class inclusions:

commutative rings  $\supset$  integral domains  $\supset$  integrally closed domains  $\supset$  GCD domains  $\supset$  unique factorization domains  $\supset$  principal ideal domains  $\supset$  Euclidean domains  $\supset$  fields

[1, 23]. The integral domain property of the shuffle algebra was proved in [29, Theorem 3.2]. The following theorem identifies the strongest structure available on this ring.

**Theorem 4.1.** *The shuffle algebra on  $\mathbb{R}\langle X \rangle$  is a unique factorization domain but not a principal ideal domain.*

*Proof:* The claim that the shuffle algebra on  $\mathbb{R}\langle X \rangle$  is a unique factorization domain follows from existing results. It is known from [28] (see also [18, Section 6] and [24, Chapter 5]) that the shuffle algebra on  $\mathbb{R}\langle X \rangle$  is isomorphic to the symmetric algebra on the  $\mathbb{R}$ -vector space  $V$  having basis  $L = \{l_i\}_{i \geq 0}$ , the set of Lyndon words. The symmetric algebra  $S(V)$  is in turn canonically isomorphic to the free polynomial algebra  $\mathbb{R}[L]$ . Thus, there exists an  $\mathbb{R}$ -linear map  $\mathcal{L} : \mathbb{R}\langle X \rangle \rightarrow \mathbb{R}[L]$  such that

$$\mathcal{L}(\eta_1 \sqcup \eta_2) = \mathcal{L}(\eta_1)\mathcal{L}(\eta_2), \quad \forall \eta_i \in X^* \quad (11)$$

with  $\mathcal{L}(\mathbf{1}) = \mathbf{1}$ , and, in particular,  $\mathcal{L}(l_i) = l_i$ . Put another way, the shuffle algebra on  $\mathbb{R}\langle X \rangle$  is freely generated by the set of all Lyndon words [30, Theorem 6.1]. It is shown in [4, Section 4, Corollary 1] that any such polynomial ring is a unique factorization domain (see also [2]).

To be a principal ideal domain, it is necessary that every ideal in  $\mathbb{R}\langle X \rangle$  be generated by a single element. The classical argument that this is not the case in the present context goes as follows (e.g., see [25, p. 153]). The assertion is that the set of all proper polynomials in  $\mathbb{R}\langle X \rangle$ ,  $\mathbb{R}_p\langle X \rangle$ , is an ideal which is not principal. It is clear that  $\mathbb{R}_p\langle X \rangle$  is an ideal. Now suppose  $\mathbb{R}_p\langle X \rangle$  has a single generator  $p$  in the shuffle algebra, i.e.,  $\mathbb{R}_p\langle X \rangle = (p) := \{p \sqcup q : q \in \mathbb{R}\langle X \rangle\}$ . Since  $x_0, x_1 \in \mathbb{R}_p\langle X \rangle$ , there must exist  $q_0, q_1 \in \mathbb{R}\langle X \rangle$  such that  $x_0 = p \sqcup q_0$  and  $x_1 = p \sqcup q_1$ . In light of the degrees of  $x_0$  and  $x_1$ , this would require a generator of the form  $p = \alpha_0 x_0 + \alpha_1 x_1$ ,  $\alpha_i \in \mathbb{R}$ . If  $\alpha_0 = 0$ , then  $p$  will generate  $x_1$  but not  $x_0$ . Likewise, if  $\alpha_1 = 0$  then  $p$  will generate  $x_0$  but not  $x_1$ . Thus, the ideal  $\mathbb{R}_p\langle X \rangle$  has two basis elements, that is,  $\mathbb{R}_p\langle X \rangle = (x_0, x_1) := \{x_0 \sqcup q_0 + x_1 \sqcup q_1 : q_0, q_1 \in \mathbb{R}\langle X \rangle\}$ , and thus is not principal. ■

The main theorem of this section is presented next. It describes how to zero the output of a formal Fliess operator  $F_c$  using its parallel product decomposition as shown in Figure 2.

**Theorem 4.2.** *Let  $c \in \mathbb{R}\langle X \rangle$  with  $c_N \neq 0$  and unique factorization  $c = c_1 \sqcup c_2 \sqcup \dots \sqcup c_n$  (modulo a permutation), where each  $c_i$  is irreducible as a polynomial in the shuffle algebra. Then  $c_{u^*} \neq 0$  is a nulling series for  $c$  if and only if it is a nulling series for at least one of the factors  $c_i$ .*

*Proof:* If  $c_{u^*} \neq 0$  is a nulling series for  $c_i$ , then directly from Lemma 3.1 it is a nulling series for  $c$ . Conversely, if

$$c \circ c_{u^*} = (c_1 \circ c_{u^*}) \sqcup (c_2 \circ c_{u^*}) \sqcup \dots \sqcup (c_n \circ c_{u^*}) = 0$$

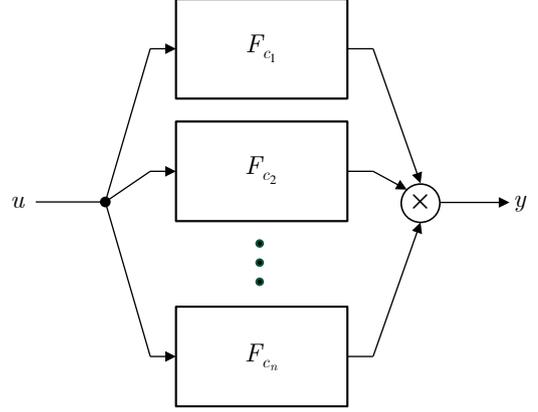


Figure 2: Parallel product decomposition of  $F_c$

for some  $c_{u^*} \neq 0$ , then since the shuffle algebra is an integral domain, at least one series  $c_i \circ c_{u^*}$  must be the zero series, and the theorem is proved. ■

It is important to point out what the theorem above is not saying, namely, that every nullable series can be factored into a shuffle product of *primely* nullable series. While it is easy to demonstrate that a primely nullable series need not be irreducible (Examples 3.8 and 4.3), it is unknown at present whether a nullable and irreducible series is always primely nullable. This is a much deeper problem.

Next, an algorithm is given to shuffle factorize any polynomial  $c \in \mathbb{R}\langle X \rangle$  into its irreducible components. This result follows from combining the proof of the previous theorem and various known properties of Lyndon words. Suppose  $X = \{x_0, x_1\}$  is ordered with  $x_0 < x_1$  so as to induce a corresponding lexicographical ordering on  $X^+$ . Recall that a word  $\eta \in X^+$  is called a *Lyndon word* if all factorizations  $\eta = \xi\nu$  with  $\xi, \nu \in X^+$  have the property that  $\eta < \nu\xi$ . In this case, the first few Lyndon words are  $L = \{l_i\}_{i \geq 0} = \{x_0, x_1, x_0x_1, x_0^2x_1, x_0x_1^2, x_0^3x_1, x_0^2x_1^2, x_0x_1^3, \dots\}$ , where here the ordering is by increasing word length and then lexicographically among words of the same length.<sup>1</sup> The Chen-Fox-Lyndon factorization of a word  $\eta \in X^+$  is a unique non-increasing product of Lyndon words so that

$$\eta = l_{i_1} l_{i_2} \dots l_{i_n}, \quad l_{i_1} \geq l_{i_2} \geq \dots \geq l_{i_n}$$

[3, 19, 24, 30]. A consequence of this factorization is that

$$l_{i_1} \sqcup l_{i_2} \sqcup \dots \sqcup l_{i_n} = a\eta + R, \quad (12)$$

where  $a$  is a nonzero element in the set of rational numbers,  $Q$ , and all the words in the remainder  $R$  are lexicographically smaller than  $\eta$  [19, Theorem 5.6]. Applying (11) to both sides of (12) provides the recursive formula

$$\mathcal{L}(l_{i_1} l_{i_2} \dots l_{i_n}) = \frac{1}{a}(l_{i_1} l_{i_2} \dots l_{i_n} - \mathcal{L}(R)). \quad (13)$$

<sup>1</sup>This ordering is only for convenience in displaying results and does not play any mathematical role in this presentation.

The recursion will terminate when all the remainders are zero.

**Example 4.1.** Let  $\eta_1 = x_0x_1x_0$ ,  $\eta_2 = x_0^2x_1x_0$ ,  $\eta_3 = x_0x_1x_0^2$ . Their Chen-Fox-Lyndon factorizations are, respectively,  $\eta_1 = (x_0x_1)x_0 = l_2l_0$ ,  $\eta_2 = (x_0^2x_1)(x_0) = l_3l_0$ , and  $\eta_3 = (x_0x_1)(x_0)(x_0) = l_2l_0^2$ . Observe, for example, that

$$l_2 \sqcup l_0 = x_0x_1x_0 + 2x_0^2x_1 = \eta_1 + R,$$

where  $R = 2l_3$ . Therefore, from (13)

$$\mathcal{L}(\eta_1) = l_0l_2 - 2l_3.$$

Likewise,

$$\begin{aligned} \mathcal{L}(\eta_2) &= l_0l_3 - 3l_5 \\ \mathcal{L}(\eta_3) &= \frac{1}{2}l_0^2l_2 - 2l_0l_3 + 3l_5. \end{aligned}$$

To illustrate the isomorphism between the shuffle algebra on  $\mathbb{R}\langle X \rangle$  and the free algebra  $R[L]$ , observe that

$$\eta_1 \sqcup x_0 = 2\eta_2 + 2\eta_3,$$

so that as expected from (11)

$$\begin{aligned} \mathcal{L}(\eta_1 \sqcup x_0) &= 2\mathcal{L}(\eta_2) + 2\mathcal{L}(\eta_3) \\ &= (l_0l_2 - 2l_3)l_0 \\ &= \mathcal{L}(\eta_1)\mathcal{L}(x_0). \end{aligned}$$

□

The proposed algorithm for shuffle factoring a proper polynomial  $c$  is summarized below:

1. Compute  $c_L = \mathcal{L}(c)$  using (12)-(13).
2. Factor  $c_L$  using Mathematica's `Factor` command [34].
3. Apply the map  $\mathcal{L}^{-1}$  to each factor in  $R[L]$  from the previous step. In particular,  $\mathcal{L}^{-1}(l_{i_1}l_{i_2}\cdots l_{i_k}) = l_{i_1} \sqcup l_{i_2} \sqcup \cdots \sqcup l_{i_k}$ .

An efficient algorithm for the Chen-Fox-Lyndon factorizations in step 1 is given in [6] (see also [31]). Mathematica's implementation notes for `Factor` provide a description of the specific algorithms used to factor multivariate polynomials. For a more general treatment of the subject see [33].

**Example 4.2.** Consider the polynomial

$$\begin{aligned} c &= 2x_0^2 - 2x_1^2 + 2x_0^2x_1x_0 + 2x_0x_1x_0^2 - 2x_0x_1^2x_0 \\ &\quad + 2x_1x_0^2x_1 + 2x_1x_0x_1^2 + 2x_1^2x_0x_1 + 2x_0x_1x_0x_1x_0x_1 \\ &\quad + 2x_0x_1x_0x_1^2x_0 + 4x_0x_1^2x_0^2x_1 + 2x_0x_1^2x_0x_1x_0 \\ &\quad + 2x_1x_0^2x_1x_0x_1 + 4x_1x_0^2x_1^2x_0 + 2x_1x_0x_1x_0^2x_1 \\ &\quad + 2x_1x_0x_1x_0x_1x_0, \end{aligned}$$

which does not have relative degree since it has no linear words of the form  $x_0^{r-1}x_1$  in its support. The algorithm above is applied to  $c$  with the help of the Mathematica NCFPS package [27]. The underlined terms correspond to Example 4.1 which is embedded in this example.

Step 1: Observe

$$\begin{aligned} c_L = \mathcal{L}(c) &= l_0^2 - l_1^2 + \underline{l_0^2l_2} + l_1^2l_2 + l_0l_1l_2^2 - \underline{2l_0l_3} \\ &\quad + 2l_1l_3 - 2l_1l_2l_3 - 2l_0l_4 - 2l_1l_4 - 2l_0l_2l_4 + 4l_3l_4. \end{aligned}$$

Step 2: Using the `Factor` command in Mathematica gives

$$c_L = (l_0 + l_1 + \underline{l_0l_2 - 2l_3})(\underline{l_0 - l_1 + l_1l_2 - 2l_4}).$$

Step 3: Mapping each factor of  $c_L$  back to  $\mathbb{R}_p\langle X \rangle$  yields

$$c = c_1 \sqcup c_2 = (x_0 + x_1 + \underline{x_0x_1x_0}) \sqcup (\underline{x_0 - x_1 + x_1x_0x_1}).$$

Observe that the two factors of  $c$  are distinct and linearly nullable with relative degree  $r = 1$ . Hence, there exist two distinct nulling inputs  $c_{u_1^*}$  and  $c_{u_2^*}$  for this polynomial. Each input can be computed via the algorithm in [12] or by solving an initial value problem which follows from setting  $F_{c_i}[u] = 0$  and then repeatedly differentiating with respect to time. For  $c_1$  the latter approach yields

$$u\ddot{u} - 2\dot{u}^2 - u^4 = 0, \quad u(0) = -1, \quad \dot{u}(0) = 0$$

so that

$$c_{u_1^*} = \mathbf{1} + x_0^2 + 7x_0^4 + 127x_0^6 + 4369x_0^8 + \cdots$$

Similarly, for  $c_2$  the corresponding initial value problem is

$$\dot{u} + tu = 0, \quad u(0) = -1,$$

which gives

$$c_{u_2^*} = -\mathbf{1} + x_0^2 - 3x_0^4 + 15x_0^6 - 105x_0^8 + \cdots$$

To empirically verify that  $c \circ c_{u_i^*} = 0$ , it is necessary to truncate  $c_{u_i^*}$ . This means that  $c \circ c_{u_i^*}$  will not be exactly zero, but instead zero up to some word length depending on the number of terms retained in  $c_{u_i^*}$ . For example, truncating both  $c_{u_i^*}$  to words of maximum length six gives

$$\begin{aligned} c \circ c_{u_1^*} &= 87380x_0^{10} + 2946560x_0^{12} + 153856528x_0^{14} + O(x_0^{16}) \\ c \circ c_{u_2^*} &= 2100x_0^{10} - 840840x_0^{14} + 57657600x_0^{16} - O(x_0^{18}). \end{aligned}$$

□

**Example 4.3.** Reconsider Example 3.8 where  $c_L = l_0 + l_1$  and  $d_L = \mathbf{1} + l_1$ . As observed earlier,  $c \sqcup d$  is primely nullable and linearly nullable. Clearly  $(c \sqcup d)_L = c_L d_L$  is reducible with one linearly nullable factor  $c_L$ . □

**Example 4.4.** Recall that for polynomials in one variable, the class of irreducible polynomials depends on the base field. For example, over the real field, the irreducible polynomials are either of degree 1 or degree 2

(e.g.,  $x_0^2 + 1$ ). Over the complex field, there are only degree 1 irreducibles [22, Chapter IV.1]. However, in every multivariate polynomial ring there are irreducible elements of higher degree. Consider the polynomial  $c = 6x_1^3 - 2x_1x_0^2 - 2x_0x_1x_0 - 2x_0^2x_1 - 24x_0^4 \in \mathbb{R}_p\langle X \rangle$ . It does not have relative degree, and thus, it is not linearly nullable. There is at present no direct test for any other form of nullability. In the Lyndon basis, it follows that  $c_L = l_1^3 - l_0^2l_1 - l_0^4 \in \mathbb{R}[L]$ . Now if  $c_L$  is reducible, one could write

$$c_L = (l_1 - p_1(l_0))(l_1^2 + p_2(l_0)l_1 + p_3(l_0)) \quad (14)$$

for some polynomials  $p_i(l_0)$ . Since  $l_0^4 = p_1(l_0)p_3(l_0)$ , necessarily  $p_1(l_0) = al_0^n$  and  $p_3(l_0) = bl_0^{4-n}$  for some  $n \in \{0, 1, 2, 3, 4\}$  and  $a, b \in \mathbb{R}$  with  $ab = 1$ . Substituting these forms into (14) shows directly that there are no values of  $n$  that can yield  $c_L$ . Thus,  $c_L$  is an irreducible multivariate polynomial of degree 4 as an element in  $\mathbb{R}[L]$ .  $\square$

## 5. Conclusions

Working entirely in a Chen-Fliess series setting, it was shown that the problem of zeroing the output can be solved for the class of systems whose generating series can be factored via the shuffle algebra into terms where at least one factor is nullable. The shuffle algebra on  $\mathbb{R}\langle X \rangle$  was shown to be a unique factorization domain so that any polynomial can be uniquely factored into its irreducible elements for this purpose. The factorization is done by viewing this shuffle algebra as the symmetric algebra over the vector space spanned by Lyndon words. A specific polynomial factorization algorithm was given based on the Chen-Fox-Lyndon factorization of words.

## Acknowledgments

KEF and AS were supported by the Research Council of Norway through project 302831 Computational Dynamics and Stochastics on Manifolds (CODYSMA).

The authors wish to thank Lance Berlin for maintaining the Mathematica software package Noncommutative Formal Power Series (NCFPS), which was used extensively in this paper to do the formal power series calculations.

## Appendix. Proof of Identity (6)

Observe for any  $c, d \in \mathbb{R}\langle\langle X \rangle\rangle$  that

$$\begin{aligned} x_0^{-1}(c \circ d) &= \sum_{\eta \in X^*} (c, \eta) x_0^{-1}(\eta \circ d) \\ &= \sum_{\eta \in X^*} (c, x_0\eta) x_0^{-1}(x_0\eta \circ d) + \\ &\quad \sum_{i=1}^m \sum_{\eta \in X^*} (c, x_i\eta) x_0^{-1}(x_i\eta \circ d) \end{aligned}$$

$$\begin{aligned} &= \sum_{\eta \in X^*} (x_0^{-1}(c), \eta) \eta \circ d + \\ &\quad \sum_{i=1}^m \sum_{\eta \in X^*} (x_i^{-1}(c), \eta) (d_i \sqcup (\eta \circ d)) \\ &= x_0^{-1}(c) \circ d + \sum_{i=1}^m d_i \sqcup [x_i^{-1}(c) \circ d]. \end{aligned}$$

For any  $i \neq 0$ , note that

$$x_i^{-1}(c \circ d) = \sum_{\eta \in X^*} (c, \eta) x_i^{-1}(\eta \circ d),$$

and  $\eta \circ d$  always has a leading  $x_0$  when  $\eta$  is nonempty. Thus,  $x_i^{-1}(\eta \circ d) = 0$  for all  $\eta \in X^*$ , and the identity is verified.

## References

- [1] M. Artin, *Algebra*, 2nd Ed., Prentice Hall, Boston, MA, 2011.
- [2] J. Berstel and L. Boasson, Shuffle factorization is unique, *Theoret. Comput. Sci.*, 273 (2002) 47–67.
- [3] K.-T. Chen, R. H. Fox, and R. C. Lyndon, Free differential calculus, IV, *Ann. Math.*, 68 (1958) 81–95.
- [4] P. M. Cohn, Unique factorization domains, *Amer. Math. Monthly*, 80 (1973) 1–18.
- [5] L. A. Duffaut Espinosa, W. S. Gray, and I. Perez Avellaneda, Derivatives of Chen-Fliess series with applications to optimal control, *Proc. 62nd IEEE Conf. on Decision and Control*, Singapore, pp. 5945–5952.
- [6] J.-P. Duval, Factorizing words over an ordered alphabet, *J. of Algorithms*, 4 (1983) 363–381.
- [7] A. Ferfera, Combinatoire du monoïde libre et composition de certains systèmes non linéaires, *Astérisque*, 75-76 (1980) 87–93.
- [8] M. Fliess, Fonctions causales non linéaires et indéterminées non commutatives, *Bull. Soc. Math. France*, 109 (1981) 3–40.
- [9] L. Foissy, The Hopf algebra of Fliess operators and its dual pre-Lie algebra, *Comm. Algebra*, 43 (2015) 4528–4552.
- [10] L. Foissy, Bialgebras in cointeraction, the antipode and the eulerian idempotent, arXiv:2201.11974, 2022.
- [11] W. S. Gray, L. A. Duffaut Espinosa, and K. Ebrahimi-Fard, Faà di Bruno Hopf algebra of the output feedback group for multivariable Fliess operators, *Systems Control Lett.*, 74 (2014) 64–73.
- [12] 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.
- [13] W. S. Gray and K. Ebrahimi-Fard, SISO output affine feedback transformation group and its Faà di Bruno Hopf algebra, *SIAM J. Control Optim.*, 55 (2017) 885–912.
- [14] 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, 6 pages.
- [15] W. S. Gray, K. Ebrahimi-Fard, and A. Schmeding, Zeroing the output of nonlinear systems without relative degree, *Proc. 57th Conf. on Information Sciences and Systems*, Baltimore, MD, 2021, 6 pages.
- [16] W. S. Gray and Y. Wang, Fliess operators on  $L_p$  spaces: Convergence and continuity, *Systems Control Lett.*, 46 (2002) 67–74.
- [17] W. S. Gray and Y. Wang, Formal Fliess operators with applications to feedback interconnections, *Proc. 18th Inter. Symp. on the Mathematical Theory of Networks and Systems*, Blacksburg, VA, 2008, 12 pages.
- [18] D. Grinberg and V. Reiner, Hopf algebras in combinatorics, arXiv:1409.8356, 2020.
- [19] M. Hazewinkel, The algebra of quasi-symmetric functions is free over the integers, *Adv. Math.*, 164 (2001) 283–300.

- [20] A. Isidori, *Nonlinear Control Systems*, 3rd Ed., Springer, London, 1995.
- [21] A. Isidori, The zero dynamics of a nonlinear system: From the origin to the latest progresses of a long successful story, *European J. Control*, 19 (2013) 369–378.
- [22] S. Lang, *Algebra*, 3rd Ed., Springer, New York, 2002.
- [23] H. Lombardi and C. Quitté, *Commutative Algebra: Constructive Methods*, Springer, Dordrecht, 2015.
- [24] M. Lothaire, *Combinatorics on Words*, Cambridge University Press, Cambridge, United Kingdom, 1983.
- [25] S. MacLane and G. Birkhoff, *Algebra*, Macmillan Co., New York, 1967.
- [26] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer, New York, 1990.
- [27] NonCommutative Formal Power Series (NCFPS), Mathematica package available at: <https://github.com/LBerlin/NCFPS-MMA>, 2023.
- [28] D. E. Radford, A natural ring basis for the shuffle algebra and an application to group schemes, *J. Algebra*, 58 (1979) 432–454.
- [29] R. Ree, Lie elements and an algebra associated with shuffles, *Ann. of Math. (2)*, 68 (1958) 210–220.
- [30] C. Reutenauer, *Free Lie Algebras*, Oxford University Press, New York, 1993.
- [31] A. C. J. Scholtens, On the graded dual of the non-commutative universal Leibniz Hopf algebra  $\mathcal{Z}$ , Faculty of Mathematics and Computer Science, series no. 419, Free University, Amsterdam, 1994.
- [32] A. Tanwani and L. Liberzon, Invertibility of switched nonlinear systems, *Automatica*, 46 (2010) 1962–1973.
- [33] J. von zur Gathen, Irreducibility of multivariate polynomials, *J. Comput. System Sci.*, 31 (1985) 225–264.
- [34] Wolfram Research, Factor, Wolfram Language function, <https://reference.wolfram.com/language/ref/Factor.html>, 2007.