Zeroing the Output of Nonlinear Systems Without Relative Degree

W. Steven Gray
Department of Electrical and Computer Engineering
Old Dominion University
Norfolk, VA 23529 USA

Kurusch Ebrahimi-Fard
Department of Mathematical Sciences
Norwegian University of Science and Technology (NTNU)
7491 Trondheim, Norway

Alexander Schmeding
FLU, Nord University
Høgskoleveien 27
7601 Levanger, Norway

***Abstract*—The goal of this paper is to establish some facts concerning the problem of zeroing the output of an input-output system that does not have relative degree. The approach taken is to work with systems that have Chen-Fliess series representations. The main result is that a class of generating series called *primely nullable series* provides the building blocks for solving this problem using shuffle algebra. This is achieved by viewing the latter as the symmetric algebra over the commutative polynomials in Lyndon words in order to show that it is a unique factorization domain. Next, the focus turns to factoring generating series in the shuffle algebra into its irreducible elements. A specific algorithm based on the Chen-Fox-Lyndon factorization of words is given.**

***Keywords*—nonlinear control systems, zero dynamics, Chen-Fliess series**

I. Introduction

Consider a smooth control-affine state space realization

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

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 input-output map $F : u \mapsto y$ is left invertible [16], [21]. 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 [16], [21] or computed analytically using formal power series methods [10]. This construction leads to the notion of *zero dynamics* [12], [16], [17], [21]. When the system fails to have relative degree, there appears to be little known about the problem of zeroing the output. Take as a simple example the system

$$\dot{z}_1 = 1 - u, \quad \dot{z}_2 = z_3 - u, \quad \dot{z}_3 = 1, \quad z(0) = 0 \tag{2a}$$
$$y = z_1 z_2. \tag{2b}$$

It is easily verified that this realization does not have relative degree at the origin. Nevertheless, there are two inputs which give the zero output: $u^*(t) = 1$, $t \geq 0$ and $u^*(t) = t$, $t \geq 0$. The general goal of this paper is to establish some facts concerning how to zero the output of a system that does not have relative degree.

The approach taken will be to work purely in the input-output setting using Chen-Fliess series representations. 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 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 [14], 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 series are those having relative degree and one additional property. These will be called *linearly nullable*. While there is no known direct test for general nullability, linearly nullable series can be completely characterized, and in this case, the nulling input can be directly computed. It is shown that the shuffle product of two linearly nullable series is always strongly nullable but not linearly nullable. The shuffle product corresponds to the parallel product interconnection of two systems [6]. The focus then turns to an inverse problem, namely, factoring a polynomial into its irreducible elements in the shuffle algebra. It is first established that the shuffle algebra on the set of homogeneous noncommutative polynomials over $\mathbb{R}$ as a commutative ring is a unique factorization domain. This is achieved by assembling some existing results from algebra [4] and algebraic combinatorics [20]. Of particular importance is the fact that this shuffle algebra can be viewed as the symmetric algebra over the commutative polynomials in Lyndon words [23]. Once this factorization result is established, it is shown that irreducible nullable series are building blocks for constructing other nullable series. What is unknown at present is whether every nullable series can be written uniquely as the shuffle product of *primely* nullable series. Finally, an algorithm is given to factor a polynomial into its irreducible shuffle components. This is done by first mapping the given polynomial to the symmetric algebra using the Chen-Fox-Lyndon factorization of words [3], [15], [20]. The resulting polynomial is then factored using one of the many known algorithms for factoring multivariate commutative polynomials [26]. Then each factor is transformed 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 III, the concept of nullable generating series is presented. The subsequent section addresses the problem of factoring generating series in the shuffle algebra and irreducibility. The final section provides the main conclusions of the paper.

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. 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. A series c is called *proper* if $\emptyset \notin \mathrm{supp}(c)$. The *order* of c, $\mathrm{ord}(c)$, 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

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

where $x_i, x_j \in X$, $\eta, \xi \in X^*$ and with $\eta \sqcup\!\sqcup \emptyset = \emptyset \sqcup\!\sqcup \eta = \eta$ [6]. 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.

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 $c \in \mathbb{R}^\ell\langle\langle X\rangle\rangle$ is

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

[6]. If there exist real numbers $K_c, M_c > 0$ such that

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

then F_c constitutes a well defined mapping from $B^m_\mathfrak{p}(R)[t_0, t_0 + T]$ into $B^\ell_\mathfrak{q}(S)[t_0, t_0 + T]$ 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., $1/\mathfrak{p} + 1/\mathfrak{q} = 1$ [13]. (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 $\mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$.

A Fliess operator F_c defined on $B^m_\mathfrak{p}(R)[t_0, t_0 + T]$ 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 $u \in B^m_\mathfrak{p}(R)[t_0, t_0 + T]$, 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 $\eta = x_{i_k} \cdots x_{i_1} \in X^*$

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

[6], [16], [21].

B. System interconnections

Given Fliess operators F_c and F_d, where $c, d \in \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$, the parallel and product connections satisfy $F_c + F_d = F_{c+d}$ and $F_cF_d = F_{c \sqcup\!\sqcup d}$, respectively [6]. It is also known that the composition of two Fliess operators F_c and F_d with $c \in \mathbb{R}^\ell_{LC}\langle\langle X\rangle\rangle$ and $d \in \mathbb{R}^m_{LC}\langle\langle X\rangle\rangle$ always yields another Fliess operator with generating series $c \circ d$, where the *composition product* is given by

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

[5]. Here ψ_d is the continuous (in the ultrametric sense) algebra homomorphism from $\mathbb{R}\langle\langle X\rangle\rangle$ to the vector space endomorphisms on $\mathbb{R}\langle\langle X\rangle\rangle$, $\mathrm{End}(\mathbb{R}\langle\langle X\rangle\rangle)$, uniquely specified by $\psi_d(x_i\eta) = \psi_d(x_i) \circ \psi_d(\eta)$ with $\psi_d(x_i)(e) = x_0(d_i \sqcup\!\sqcup e)$, $i = 0, 1, \ldots, m$ 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 $\mathbb{R}\langle\langle X\rangle\rangle$. It can be verified directly that

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

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 $c/d = c \sqcup\!\sqcup d^{\sqcup\!\sqcup -1}$ so that $F_c/F_d = F_{c/d}$ with the shuffle inverse of d defined as

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

where $d' = \mathbf{1} - (d/(d,\emptyset))$ is proper and $(d')^{\sqcup\!\sqcup *} := \sum_{k\geq 0}(d')^{\sqcup\!\sqcup k}$ [10]. 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

$$(c{\sqcup\!\sqcup}d)\circ e=(c\circ e){\sqcup\!\sqcup}(d\circ e) \tag{7}$$

for any $c,d,e\in\mathbb{R}\langle\langle X\rangle\rangle$ [7], it is sufficient to show that

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

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

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

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

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

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

C. Formal Fliess operators

Suppose $X=\{x_0,x_1\}$ and define $X_0=\{x_0\}$. Then every series $c_u\in\mathbb{R}[[X_0]]$ can be identified with an infinite jet $j^{\infty}_{t_0}(u)$ 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 $j^{\infty}_{t_0}(u)$. 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 $c\circ:\mathbb{R}[[X_0]]\to\mathbb{R}[[X_0]]:c_u\mapsto c_y=c\circ c_u$ is still well defined and takes the input infinite jet to the output infinite jet. This is called a *formal Fliess operator* [14]. The advantage to 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.

D. 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: [10] 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 $\mathrm{supp}(c_F)\subseteq x_0^{r-1}X^*$. Then c has *relative degree* r if the linear word $x_0^{r-1}x_1\in\mathrm{supp}(c)$, 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 $\mathrm{supp}(e)\subseteq X^*/\{X_0^*,x_1\}$ such that

$$c=c_N+c_F=c_N+Kx_0^{r-1}x_1+x_0^{r-1}e \tag{9}$$

and $K\neq 0$. This notion of relative degree coincides with the usual definition given in a state space setting [11].

III. 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 $c\circ:\mathbb{R}[[X_0]]\to\mathbb{R}[[X_0]],c_u\mapsto c\circ c_u$. 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.

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 $c=x_0^2-x_1x_0$ 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 $c=x_0+x_0x_1$ is not nullable since $c\circ c_u=x_0+x_0^2c_u\neq 0$ for all $c_u\in\mathbb{R}[[X_0]]$. □

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

Theorem 3.1: If $c\in\mathbb{R}\langle\langle X\rangle\rangle$ has relative degree r, and $\mathrm{supp}(c_N)\subseteq x_0^rX_0^*$ 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

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

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 δ 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 [9]. Therefore, one can solve for c_u directly via left inversion to give $c_u = d_\delta^{\circ -1} \circ 0$. In which case, there exists a unique c_u which will zero out all coefficients of c_y with the exception of the first r coefficients. These initial coefficients are completely determined by c since

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

By assumption $\mathrm{supp}(c_N) \subseteq x_0^r X_0^*$. 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 $(\mathbb{R}\langle\langle X\rangle\rangle, ⧢, \mathbf{1})$ 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 [8]. 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 $(\mathbf{1} + \mathbb{R}_{np}\langle\langle X\rangle\rangle, ⧢, \mathbf{1}) \subset (\mathbb{R}\langle\langle X\rangle\rangle, ⧢, \mathbf{1})$ 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 $\mathrm{supp}(e_1) \subseteq X^*/\{X_0^*, x_1\}$.

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. 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. □

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

$$I_c = \{c ⧢ d : d \in \mathbb{R}\langle\langle X\rangle\rangle\}$$

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 $(c ⧢ d) \circ c_{u^*} = (c \circ c_{u^*}) ⧢ (d \circ c_{u^*}) = 0$ 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.

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

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

$$\begin{aligned}c ⧢ d = (&x_0^{r_c} e_0 + K_c x_0^{r_c - 1} x_1 + x_0^{r_c - 1} e_1) ⧢ \\ &(x_0^{r_d} f_0 + K_d x_0^{r_d - 1} x_1 + x_0^{r_d - 1} f_1)\end{aligned}$$

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

$$K_c(f_0, \emptyset) + K_d(e_0, \emptyset) \neq 0. \tag{10}$$

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

$$\begin{aligned}&(e_0, \emptyset) x_0^{r_c} ⧢ K_d x_0^{r_d - 1} x_1 + (f_0, \emptyset) x_0^{r_d} ⧢ K_c x_0^{r_c - 1} x_1 \\ &= K_d(e_0, \emptyset)(x_0^{r_c} ⧢ x_0^{r_d - 1} x_1) + K_c(f_0, \emptyset)(x_0^{r_d} ⧢ x_0^{r_c - 1} x_1)\end{aligned}$$

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

$$\begin{aligned}&(e_0, \emptyset) x_0 ⧢ K_d x_1 + (f_0, \emptyset) x_0 ⧢ K_c x_1 \\ &= K_d(e_0, \emptyset)(x_0 x_1 + x_1 x_0) + K_c(f_0, \emptyset)(x_0 x_1 + x_1 x_0).\end{aligned}$$

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

Example 3.6: 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

$$c ⧢ d = -x_0 x_1 - x_1 x_0 + 2x_1^2 + 3x_0^3 - x_0^2 x_1 - x_0 x_1 x_0 - x_1 x_0^2$$

does not have relative degree. Therefore $c ⧢ d$ 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. □

Example 3.7: 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

$$c ⧢ d = x_0 + x_1 + x_0 x_1 + x_1 x_0 + 2x_1^2$$

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. □

Example 3.8: Suppose $c = d = x_0 - x_1$ so that both series are linearly nullable with relative degree 1. As expected,

$$c ⧢ d = 2x_0^2 - 2x_0 x_1 - 2x_1 x_0 - 2x_1^2$$

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 $c \sqcup\!\sqcup d$ as the shuffle product is an integral domain. That is, in general $(c \sqcup\!\sqcup d) \circ e_u = (c \sqcup\!\sqcup e_u) \sqcup\!\sqcup (d \circ e_u) = 0$ 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:

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

IV. Factorizations in the Shuffle Algebra

Let $\mathbb{R}_p\langle X\rangle$ denote the set of all proper polynomials in $\mathbb{R}\langle X\rangle$. The shuffle product on $\mathbb{R}_p\langle X\rangle$ 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 $\supset$ <u>unique factorization domains</u> $\supset$ principal ideal domains $\supset$ Euclidean domains

[1], [19]. The integral domain property of the shuffle algebra was proved in [24, Theorem 3.2]. The following theorem identifies the strongest structure available on this ring.

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

Proof: The claim that the shuffle algebra on $\mathbb{R}_p\langle X\rangle$ is a unique factorization domain follows from existing results. Specifically, it is known from [20, Chapter 5] (see also [23]) that this shuffle algebra is isomorphic to the symmetric algebra over the vector space $\mathbb{R}[L]$, where L is the set of Lyndon words. This algebra is in turn isomorphic to $\mathbb{R}[L]$ as a commutative polynomial algebra. It is shown in [4, Corollary 1] that any such polynomial algebra is a unique factorization domain (see also [2]).

To show that the shuffle algebra is not a principal ideal domain, the following inclusions are useful:

> GCD domains $\supset$ Bézout domains $\supset$ principal ideal domains.

The assertion is that the shuffle algebra is not a Bézout domain and thus not a principal ideal domain. Observe that x_0 and x_1 are coprime and yet the Bézout identity $x_0 \sqcup\!\sqcup c + x_1 \sqcup\!\sqcup d = \mathbf{1}$ has no solution (c,d) in the shuffle algebra since in general $\mathrm{ord}(c \sqcup\!\sqcup d) \geq \mathrm{ord}(c) + \mathrm{ord}(d)$. In fact, the unit $\mathbf{1} \notin \mathbb{R}_p\langle X\rangle$. $\blacksquare$

The main theorem of this section is presented next.

Theorem 4.2: Let $c \in \mathbb{R}_p\langle X\rangle$ with $c_N \neq 0$ and unique factorization $c = c_1 \sqcup\!\sqcup c_2 \sqcup\!\sqcup \cdots \sqcup\!\sqcup c_n$ (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

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

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. $\blacksquare$

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 (Example 3.7), it is unknown at present whether a nullable and irreducible series is always primely nullable. This is a much deeper problem.

The following algorithm to factor a given $c \in \mathbb{R}_p\langle X\rangle$ into its irreducible components follows directly from the proof of the previous theorem and existing results:

1. Map c to $c_L \in \mathbb{R}[L]$ using the Chen-Fox-Lyndon factorization to map each word in $\mathrm{supp}(c)$ to a unique non-decreasing product of Lyndon words [3], [15], [20].
2. Factor c_L using Mathematica's `Factor` command [27].
3. Map each factor in $\mathbb{R}[L]$ from the previous step back to $\mathbb{R}_p\langle X\rangle$ using the map $L^* \to \mathbb{R}_p\langle X\rangle : L_{i_1}L_{i_2}\cdots L_{i_k} \mapsto L_{i_1} \sqcup\!\sqcup L_{i_2} \sqcup\!\sqcup L_{i_k}$.

An efficient algorithm for implementing step 1 is given in [25]. Mathematica's implementation notes for `Factor` provide a description of the specific algorithms used to factor multivariate polynomials. Also see [26] for a more general treatment of the subject.

Example 4.1: Consider the polynomial

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

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 [22].

<u>Step 1</u>: Assuming $x_0 < x_1$, the first few Lyndon words are: $L_0 = x_0$, $L_1 = x_1$, $L_2 = x_0x_1$, $L_3 = x_0^2x_1$, $L_4 = x_0x_1^2$. In this case, c maps to

$$\begin{aligned} c_L = {} & L_0^2 - L_1^2 + L_0^2L_2 + L_1^2L_2 + L_0L_1L_2^2 - 2L_0L_3 + 2L_1L_3 \\ & - 2L_1L_2L_3 - 2L_0L_4 - 2L_1L_4 - 2L_0L_2L_4 + 4L_3L_4. \end{aligned}$$

<u>Step 2</u>: Using the `Factor` command in Mathematica gives

$$c_L = (L_0 + L_1 + L_0L_2 - 2L_3)(L_0 - L_1 + L_1L_2 - 2L_4).$$

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

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

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 [10] 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

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

so that

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

Similarly, for c_2 the corresponding initial value problem is

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

which gives

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

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

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

□

Example 4.2: Reconsider Example 3.7 where $c_L=L_0+L_1$ and $d_L=\mathbf{1}+L_1$ (slightly abusing the notation since $d\notin\mathbb{R}_p\langle X\rangle$). As observed earlier, $c⧢d$ is primely nullable but not linearly nullable. Clearly $(c⧢d)_L=c_Ld_L$ is reducible with one linearly nullable factor c_L. □

Example 4.3: 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 [18, 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

$$c_L=(L_1-p_1(L_0))(L_1^2+p_2(L_0)L_1+p_3(L_0)) \qquad (11)$$

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 (11) 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 $\mathbb{R}[L]$. □

V. Conclusions

Working entirely in a Chen-Fliess series setting, it was shown that a class of generating series called primely nullable series provides building blocks in the shuffle algebra for the problem of zeroing the output. This was accomplished by showing that the shuffle algebra over $\mathbb{R}$ is a unique factorization domain so that any nullable series can be uniquely factored into its irreducible elements for the purpose of identifying any nullable factors. This factorization is done by viewing the shuffle algebra as the symmetric algebra over the commutative polynomials in Lyndon words. A specific algorithm based on the Chen-Fox-Lyndon factorization of words was given.

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] A. Ferfera, Combinatoire du monoïde libre et composition de certains systèmes non linéaires, *Astérisque*, 75-76 (1980) 87–93.

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

[7] L. Foissy, The Hopf algebra of Fliess operators and its dual pre-Lie algebra, *Comm. Algebra*, 43 (2015) 4528–4552.

[8] L. Foissy, Bialgebras in cointeraction, the antipode and the eulerian idempotent, arXiv:2201.11974, 2022.

[9] W. S. Gray, L. A. Duffaut Espinosa, and K. Ebrahimi-Fard, Faá di Bruno Hopf algebra of the output feedback group for multivariable Fliess operators, *Systems Control Lett.*, 74 (2014) 64–73.

[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 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.

[12] 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.

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

[14] 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.

[15] M. Hazewinkel, The algebra of quasi-symmetric functions is free over the integers, *Adv. Math.*, 164 (2001) 283–300.

[16] A. Isidori, *Nonlinear Control Systems*, 3rd Ed., Springer, London, 1995.

[17] 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.

[18] S. Lang, *Algebra*, 3rd Ed., Springer, New York, 2002.

[19] H. Lombardi and C. Quitté, *Commutative Algebra: Constructive Methods*, Springer, Dordrecht, 2015.

[20] M. Lothaire, *Combinatorics on Words*, Cambridge Univesity Press, Cambridge, United Kingdom, 1983.

[21] H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer, New York, 1990.

[22] NonCommutative Formal Power Series (NCFPS), Mathematica package available at: https://github.com/LBerlin/NCFPS-MMA, 2021.

[23] D. E. Radford, A natural ring basis for the shuffle algebra and an application to group schemes, *J. Algebra*, 58 (1979) 432–454.

[24] R. Ree, Lie elements and an algebra associated with shuffles, *Ann. of Math. (2)*, 68 (1958) 210–220.

[25] A. C. J. Scholtens, On the graded dual of the noncommutative universal Leibniz Hopf algebra Z, preprint, Math. Seminar of the Free University of Amsterdam, 1994.

[26] J. von zur Gathen, Irreducibility of multivariate polynomials, *J. Comput. System Sci.*, 31 (1985) 225–264.

[27] Wolfram Research, Factor, Wolfram Language function, https:// reference.wolfram.com/language/ref/Factor.html, 2007.