2017

**SISO Output Affine Feedback Transformation Group and Its Faà di Bruno Hopf Algebra**

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

Kurusch Ebrahimi-Fard

Follow this and additional works at: [http://digitalcommons.odu.edu/ece_fac_pubs](http://digitalcommons.odu.edu/ece_fac_pubs)

Part of the [Algebra Commons](http://digitalcommons.odu.edu/ece_fac_pubs) and the [Electrical and Computer Engineering Commons](http://digitalcommons.odu.edu/ece_fac_pubs)

**Repository Citation**


[http://digitalcommons.odu.edu/ece_fac_pubs/124](http://digitalcommons.odu.edu/ece_fac_pubs/124)

**Original Publication Citation**


This Article is brought to you for free and open access by the Electrical & Computer Engineering at ODU Digital Commons. It has been accepted for inclusion in Electrical & Computer Engineering Faculty Publications by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.
SISO OUTPUT AFFINE FEEDBACK TRANSFORMATION GROUP AND ITS FA`A DI BRUNO HOPF ALGEBRA∗

W. STEVEN GRAY† AND KURUSCH EBRAMI-FARD‡

Abstract. The general goal of this paper is to identify a transformation group that can be used to describe a class of feedback interconnections involving subsystems which are modeled solely in terms of Chen–Fliess functional expansions or Fliess operators and are independent of the existence of any state space models. This interconnection, called an output affine feedback connection, is distinguished from conventional output feedback by the presence of a multiplier in an outer loop. Once this transformation group is established, three basic questions are addressed. How can this transformation group be used to provide an explicit Fliess operator representation of such a closed-loop system? Is it possible to use this feedback scheme to do system inversion purely in an input-output setting? In particular, can feedback input-output linearization be posed and solved entirely in this framework, i.e., without the need for any state space realization? Last, what can be said about feedback invariants under this transformation group? A final objective of the paper is to describe the Lie algebra of infinitesimal characters associated with the group in terms of a pre-Lie product.

Key words. nonlinear control systems, Chen–Fliess series, Hopf algebras, transformation groups

AMS subject classifications. 93C10, 93B18, 16T30

DOI. 10.1137/140997348

1. Introduction. Feedback transformation groups have been used extensively in control theory since its inception to explain the way that feedback can alter the nature of a system. The early work of Brockett and Krishnaprasad in the case of linear systems [3, 4] and Brockett, Jakubczyk, Respondek, and many others in the context of nonlinear state space systems [2, 25, 26, 37] has been important, for example, in identifying feedback invariants. The general goal of this paper is to identify a transformation group that can be used to describe a class of feedback interconnections involving subsystems which are modeled solely in terms of Chen–Fliess functional expansions or Fliess operators and are independent of the existence of any state space models [11, 12]. This interconnection, called an output affine feedback connection, is shown in Figure 1. The presence of a multiplier in an outer loop distinguishes this structure from conventional output feedback. Once this transformation group is established, three basic questions are addressed. How can this transformation group be used to provide an explicit Fliess operator representation of such a closed-loop system? Is it possible to use this feedback scheme to do system inversion purely in an input-output setting? In particular, can feedback input-output linearization be posed and solved entirely in this framework, i.e., without the need for any state space realization? Last, what can be said about feedback invariants under this transformation group? A final objective of the paper is to describe the Lie algebra of infinitesimal characters associated with the group in terms of a pre-Lie product.

∗Received by the editors November 25, 2014; accepted for publication (in revised form) December 2, 2016; published electronically March 28, 2017. A portion of this paper appeared in preliminary form in Proceedings of the International Symposium on the Mathematical Theory of Networks and Systems, 2014.

http://www.siam.org/journals/sicon/55-2/99734.html

Funding: The first author was supported by grant SEV-2011-0087 from the Severo Ochoa Excellence Program at the Instituto de Ciencias Matemáticas in Madrid, Spain. This research was also supported by a grant from the BBVA Foundation.

†Old Dominion University, Norfolk, VA 23529 (sgray@odu.edu).

‡Norwegian University of Science and Technology—NTNU, 7491 Trondheim, Norway (kurusch.ebrahimi-fard@math.ntnu.no). On leave from UHA, Mulhouse, France.
characters associated with this new transformation group and show that it is induced by a pre-Lie product analogous to what was found for the output feedback transformation group in \cite{13, 14}. This fact may give deeper insight into the combinatorial aspects of the theory and group invariants (see \cite{20}), but this aspect of the problem will be deferred to a later publication.

The basic approach to solving these problems is to use an operator algebra encoded by various products of formal power series. This is somewhat analogous to the familiar way in which interconnections of linear time-invariant systems are analyzed using only transfer functions and the algebra of rational functions. The feedback connection in general requires the use of a certain inverse operator that is naturally described in terms of a group. To do any calculations explicitly and efficiently here, it is necessary to present an associated Hopf algebra induced by the group product and its inverse. It has been known for a long time in some physics communities that Hopf algebras provide a natural computational framework for working with groups \cite{5, 10, 15}. For example, the group inverse can be computed explicitly in terms of the antipode, a particular antihomomorphism of the Hopf algebra. If the latter is both connected and graded, then this antipode has a recursive description which is ideal for software implementation. As will be seen, the Hopf algebra of interest here is directly related to a composition of Fliess operators. The inversion formula that it provides is a kind of functional analogue of Lagrange's formula for the inversion of smooth functions in terms of their Taylor series coefficients. This classical formula has a combinatorial interpretation in terms of a connected graded Hopf algebra \cite{27}. In particular, the coproduct of this Hopf algebra is defined in terms of the classical Faà di Bruno formula which gives the Taylor series coefficients for the composition of two sufficiently well behaved functions. Therefore, this Hopf algebra is often referred to as a Faà di Bruno type Hopf algebra \cite{10, 15}. The Hopf algebra presented in this paper is also of Faà di Bruno type and is best viewed as a generalization of existing work on output feedback (corresponding to setting $F_{d_L}[y] = 1$ for all $y$ in Figure 1) by the authors and others appearing in \cite{7, 13, 14, 16, 17, 18, 20, 22, 38}. The focus here will be on the single-input single-output (SISO) case as the multivariable case introduces additional technical issues that are nontrivial and require more advanced concepts from Hopf algebra theory.

The paper is organized as follows. In the next section, some preliminary concepts concerning Fliess operators, ultrametric spaces, and Hopf algebras are briefly outlined. In section 3, the output affine feedback transformation group is described. The problem of determining a Fliess operator representation of the output affine closed-loop system is then solved. The Hopf algebra is next developed in section 4 and subsequently applied in the following section to address the remaining two problems.
concerning system inversion/feedback linearization and feedback invariants. In section 6, the Lie algebra of infinitesimal characters is described.

2. Preliminaries. A finite nonempty set of noncommuting symbols \( X = \{ x_0, x_1, \ldots, x_m \} \) is called an alphabet. Each element of \( X \) is called a letter, and any finite sequence of letters from \( X \), \( \eta = x_{i_1} \cdots x_{i_k} \), is called a word over \( X \). The length of the word \( \eta \), \( |\eta| \), is given by the number of letters it contains. The set of all words with length \( k \) is denoted by \( X^k \). The set of all words including the empty word, \( \emptyset \), is designated by \( X^* \), while \( X^+ := X^* - \{ \emptyset \} \). The set \( X^* \) forms a monoid under concatenation. The set \( \eta X^* \) comprises all words with the prefix \( \eta \in X^* \). For any fixed integer \( \ell \geq 1 \), a mapping \( c : X^* \to \mathbb{R}^{\ell} \) is called a formal power series. The value of \( c \) at \( \eta \in X^* \) is written as \( \langle c, \eta \rangle \in \mathbb{R}^{\ell} \) and called the coefficient of \( \eta \) in \( c \). Typically, \( c \) is represented as the formal sum \( c = \sum_{\eta \in X^*} \langle c, \eta \rangle \eta \). If the constant term \( \langle c, \emptyset \rangle = 0 \), then \( c \) is said to be proper. The support of \( c \), \( \text{supp}(c) \), is the set of all words having nonzero coefficients. The order of \( c \), \( \text{ord}(c) \), is the length of the shortest word in its support \( (\text{ord}(0) := \infty) \). The collection of all formal power series over \( X^* \) is denoted by \( \mathbb{R}^{\ell}(\langle X \rangle) \). It forms an associative \( \mathbb{R} \)-algebra under the concatenation product and a commutative and associative \( \mathbb{R} \)-algebra under the shuffle product, denoted here by the shuffle symbol \( \shuffle \). The latter is the \( \mathbb{R} \)-bilinear extension of the shuffle product of two words, which is defined inductively by

\[
(\eta \shuffle \emptyset)(x_j, \xi) = x_i(\eta \shuffle (x_j, \xi)) + x_j((\eta \shuffle \emptyset)(x_j, \xi))
\]

with \( \emptyset \shuffle \emptyset = \emptyset \shuffle = \emptyset \) for all \( \eta, \xi \in X^* \) and \( x_i, x_j \in X \) [11].

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

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

where \( x_i \in X \), \( \bar{\eta} \in X^* \), and \( u_0 = 1 \). The input-output operator corresponding to \( c = \sum_{\eta \in X^*} \langle c, \eta \rangle \eta \in \mathbb{R}^{\ell}(\langle X \rangle) \) is the Fliess operator

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

[11, 12]. If there exist real numbers \( K_c, M_c > 0 \) such that

\[
|\langle c, \eta \rangle| \leq K_c M_c^{\|\eta\|_p}! \quad \forall \eta \in X^*,
\]

then \( F_c \) constitutes a well defined mapping from \( B^m_p(R)[t_0, t_0+T] \) into \( B^q_s(S)[t_0, t_0+T] \) for sufficiently small \( R, T > 0 \) and some \( S > 0 \), where the numbers \( p, q \in [1, \infty] \) are

\[1\] For notational convenience, \( p = (p, 0) \emptyset \in \mathbb{R}(X) \) is often abbreviated as \( p = (p, 0) \).
conjugate exponents, i.e., $1/p + 1/q = 1$ [21]. (Here, $|z| := \max |z_i|$ when $z \in \mathbb{R}^\ell$.)

The set of all such \textit{locally convergent} series is denoted by $\mathcal{B}_{LC}^\ell(\langle X \rangle)$. A Fliess operator $F_c$ defined on $B^n_m(R)[t_0, t_0 + T]$ is said to be \textit{realizable} when there exists a state space model consisting of $n$ differential equations and $\ell$ output functions

\begin{align}
(2.3a) & \quad \dot{z} = g_0(z) + \sum_{i=1}^{m} g_i(z)u_i, \quad z(t_0) = z_0, \\
(2.3b) & \quad y_j = h_j(z), \quad j = 1, 2, \ldots, \ell,
\end{align}

where each $g_i$ is an analytic vector field expressed in local coordinates on some neighborhood $\mathcal{W}$ of $z_0$, and each output function $h_j$ is an analytic function on $\mathcal{W}$ such that (2.3a) has a well defined solution $z(t), \quad t \in [t_0, t_0 + T]$ on $\mathcal{W}$ for any given input $u \in B^n_m(R)[t_0, t_0 + T]$, and $y_j(t) = F_c[u](t) = h_j(z(t)), \quad t \in [t_0, t_0 + T], \quad j = 1, 2, \ldots, \ell$ [11, 12, 21, 24, 32]. It can be shown that for any word $\eta = x_{i_\ell} \cdots x_{i_1} \in X^*$

\begin{equation}
(c, \eta) = L_{g_0}h_j(z_0) := L_{g_{i_\ell}} \cdots L_{g_{i_1}} h_j(z_0),
\end{equation}

where $L_{g_0}h_j$ is the Lie derivative of $h_j$ with respect to $g_i$. For any $c \in \mathbb{R}^\ell(\langle X \rangle)$, the $\mathbb{R}$-linear mapping $\mathcal{H}_c : \mathbb{R}(X) \to \mathbb{R}^\ell(\langle X \rangle)$ uniquely specified by $(\mathcal{H}_c(\eta), \xi) = (c, \xi\eta), \quad \xi, \eta \in X^*$ is called the \textit{ Hankel mapping} of $c$. The series $c$ is said to have \textit{Lie rank} $n$ when the range space of $\mathcal{H}_c$ restricted to the $\mathbb{R}$-vector space of Lie polynomials over $X$, i.e., the free Lie algebra $\mathcal{L}(X)$, has dimension $n$. It is well known that $F_c$ is realizable if and only if $c \in \mathcal{B}_{LC}^\ell(\langle X \rangle)$ has finite Lie rank [12, 24, 32].

When convergent Fliess operators are connected in a parallel-sum configuration, it is elementary to show that the composite system satisfies $F_c + F_d = F_{c+d}$. It was shown in [11] for the parallel-product connection that $F_cF_d = F_{c \cdot d}$. If $F_c$ and $F_d$ with $c \in \mathcal{B}_{LC}^\ell(\langle X \rangle)$ and $d \in \mathcal{B}_{LC}^\ell(\langle X \rangle)$ are interconnected in a cascade manner, the composite system $F_c \circ F_d$ has the Fliess operator representation $F_{cod}$, where $c \circ d$ denotes the \textit{composition product} of $c$ and $d$ as described in [9]. This product is associative and $\mathbb{R}$-linear in its left argument $c$. In the event that two convergent Fliess operators are interconnected to form a feedback system, the closed-loop system has a convergent Fliess operator representation whose generating series is the \textit{output feedback product} of $c$ and $d$, denoted by $c \circ d$ [17, 38]. Consider, for example, the SISO case where the alphabet $X = \{x_0, x_1\}$ and $\ell = 1$. Define the set of \textit{unital Fliess operators} $\mathcal{F}_\delta = \{ I + F_c : c \in \mathcal{B}_{LC}^\ell(\langle X \rangle) \}$, where $I$ denotes the identity map, i.e., $I[u] = u$. It is convenient to introduce the symbol $\delta$ as the (fictitious) generating series for $I$. That is, $F_\delta := I$ such that $I + F_c := F_{\delta+c} = F_{c_\delta}$ with $c_\delta := \delta + c$. The set of all such generating series for $\mathcal{F}_\delta$ is denoted by $\mathcal{B}_{LC}^\ell(\langle X_\delta \rangle)$ with the augmented alphabet $X_\delta := \{ \delta \} \cup X$. It has been proved that $\mathcal{F}_\delta$ forms a group under the composition $F_{c_\delta} \circ F_{d_\delta} = (I + F_{c_\delta}) \circ (I + F_{d_\delta}) = F_{c_\delta d_\delta}$, where $c_\delta \circ d_\delta := \delta + d + c \circ d$, and $\delta$ denotes the \textit{modified composition product} [19].\footnote{The same symbol will be used for composition on $\mathbb{R}(\langle X \rangle)$ and $\mathbb{R}(\langle X_\delta \rangle)$. As elements in these two sets have a distinct notation, i.e., $c$ versus $c_\delta$, respectively, it will always be clear which product is at play.} The unit element of this group is the identity map $I$. Of central importance is the so-called coordinate algebra of functions, $\mathcal{H}$, on the corresponding group $(\mathcal{B}_{LC}^\ell(\langle X_\delta \rangle), \circ, \delta)$. In this setting the group multiplication is dualized to a comultiplication on $\mathcal{H}$, which is captured through the notion of a Hopf algebra. In fact, the compositional nature of the product on $\mathcal{F}_\delta$ turns $\mathcal{H}$ into a Faà di Bruno type Hopf algebra. Thus, the group (composition) inverse in $(\mathcal{B}_{LC}^\ell(\langle X_\delta \rangle), \circ, \delta)$,
\( c^{-1} = \delta + c^{-1} \), can be computed efficiently by a recursive algorithm [7]. This inverse is also key in describing the output feedback product as shown in the following theorem.

**Theorem 2.1** (see [17, 38]). For any \( c, d \in \mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle \) it follows that \( c \circ d = c \circ (-d \circ c)^{-1} \) with \( c \circ d \in \mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle \). In addition, \( c \circ d = c \circ (d \circ (c \circ d)) \).

It is known that the set of locally convergent generating series is closed under addition, the shuffle product [39], the composition product [19], the modified composition product [28], and the feedback product [17, 38]. Therefore, all the corresponding elementary interconnections of Fliess operators preserve local convergence.

Finally, the algebraic theory described above remains largely intact for formal power series that do not necessarily satisfy convergence condition (2.2). In place of a convergent Fliess operator one can introduce a formal Fliess operator which maps formal functions to formal functions [22]. The motivation for this is the observation that if \( c \in \mathbb{R}_{LC}\langle \langle X \rangle \rangle \) and \( u \) is analytic with a convergent Taylor series representation \( u = \sum_{n \geq 0} (c_u x_0^n)(t - t_0)^n / n! = \sum_{n \geq 0} (c_u x_0^n) E_{x_0^n}[u](t, t_0) \), then \( y = F_{c}[u] \) is also analytic with Taylor series coefficients given by \( c_y = c \circ c_u \) so that \( c_u \in \mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle \) and \( c_y \in \mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle \) with \( X_{\delta} := \{ x_0 \} \). But when \( c \) is not locally convergent, the mapping \( \mathbb{R}_{m}^{\mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle} : c_u \mapsto c_y = c \circ c_u \) is still well defined. It can be viewed as the formal version of \( F_c \). The set \( \mathbb{R}_{m}^{\mathbb{R}_{LC}\langle \langle X_{\delta} \rangle \rangle} \), which also forms a group under the product \( c \circ d \), is the corresponding set of generating series for these formal operators.

### 2.2. Shuffle product operations on ultrametric spaces

Given a set \( S \), a function \( \mathcal{U} : S \times S \to \mathbb{R} \) is called an ultrametric if it satisfies the following properties for all \( s, s', s'' \in S \):

1. \( \mathcal{U}(s, s') \geq 0 \).
2. \( \mathcal{U}(s, s') = 0 \) if and only if \( s = s' \).
3. \( \mathcal{U}(s, s') = \mathcal{U}(s', s) \).
4. \( \mathcal{U}(s, s') \leq \max\{ \mathcal{U}(s, s''), \mathcal{U}(s'', s') \} \).

The pair \( (S, \mathcal{U}) \) is referred to as an ultrametric space. In the event that property iv is replaced with the triangle inequality, \( (S, \mathcal{U}) \) reduces to the usual definition of a metric space. It is not difficult to show for any fixed \( 0 < \sigma < 1 \) that the \( \mathbb{R} \)-vector space \( \mathbb{R}\langle \langle X \rangle \rangle \) with the distance between two series defined as \( \text{dist}(c, d) = \sigma^{\text{ord}(c - d)} \) is a complete ultrametric space [1]. In this section two key lemmas are presented which describe how the distance between two series is altered by operations involving the shuffle product defined in (2.1). These results are employed in section 3 to prove the existence of a group inverse using ultrametric contractions on \( \mathbb{R}\langle \langle X \rangle \rangle, \text{dist} \).

**Lemma 2.2.** For any series \( c_i, d_i \in \mathbb{R}\langle \langle X \rangle \rangle, \) \( i = 1, 2, \)

\[
\text{dist}(c_1 \circ d_1, c_2 \circ d_2) \leq \max(\sigma^{\text{ord}(c_1)} \text{dist}(d_1, d_2), \sigma^{\text{ord}(d_2)} \text{dist}(c_1, c_2)).
\]

**Proof.** First observe that

\[
\text{dist}(c_i \circ d_1, c_i \circ d_2) = \sigma^{\text{ord}(c_i \circ (d_1 - d_2))} = \sigma^{\text{ord}(c_i) + \text{ord}(d_1 - d_2)} = \sigma^{\text{ord}(c_i)} \text{dist}(d_1, d_2).
\]

Therefore, it follows that

\[
\text{dist}(c_1 \circ d_1, c_2 \circ d_2) \leq \max(\text{dist}(c_1 \circ d_1, c_1 \circ d_2), \text{dist}(c_1 \circ d_2, c_2 \circ d_2)))
\]

\[
= \max(\sigma^{\text{ord}(c_1)} \text{dist}(d_1, d_2), \sigma^{\text{ord}(d_2)} \text{dist}(c_1, c_2)) \quad \Box
\]

**Corollary 2.3.** For a fixed \( c \in \mathbb{R}\langle \langle X \rangle \rangle \), the mapping \( d \mapsto c \circ d \) is an ultrametric contraction if \( c \) is proper and an isometry on \( \mathbb{R}\langle \langle X \rangle \rangle, \text{dist} \) otherwise.
The set of nonproper series $R_n((X)) \subset R((X))$ is a group under the shuffle product. In particular, the shuffle inverse of any such series $c \in R_n((X))$ is $c^{\shuffle -1} = ((c, \emptyset)(1-c'))^{\shuffle -1} := (c, \emptyset)^{-1}(c')^{\shuffle \ast}$, where $c' := 1-c/(c, \emptyset)$ is proper and $(c')^{\shuffle \ast} := \sum_{k \geq 0} (c')^{\shuffle k}$.

**Lemma 2.5.** The shuffle inverse is an isometry for any $c, d \in R_n((X))$ having identical constant terms.

**Proof.** For any $c, d \in R_n((X))$ with $(c, \emptyset) = (d, \emptyset)$ observe that

$$\text{ord}(c^{\shuffle -1} - d^{\shuffle -1}) = \text{ord} \left( \sum_{k=1}^{\infty} (c')^{\shuffle k} - (d')^{\shuffle k} \right) = \text{ord}(c' - d') = \text{ord}(c - d),$$

and hence the lemma is proved. □

### 2.3. Hopf algebras

Some definitions and basic facts concerning Hopf algebras are summarized in this section for later use. The treatment is based largely on [10, 23, 29, 34, 36]. In the following, a unital $\mathbb{R}$-algebra refers to a vector space $A$ over the base field $\mathbb{R}$ with associative product $m : A \otimes A \to A$ and unit map (element) $e_A : \mathbb{R} \to A$ ($e_A(1) = 1_A$). A counital coalgebra over $\mathbb{R}$ consists of a triple $(C, \Delta_C, \varepsilon_C)$. The coproduct $\Delta_C : C \to C \otimes C$ is coassociative, that is, $(\text{id}_C \otimes \Delta_C) \circ \Delta_C = (\Delta_C \otimes \text{id}_C) \circ \Delta_C$, where $\text{id}_C$ is the identity map on $C$ and $\varepsilon_C : C \to \mathbb{R}$ denotes the counit (augmentation) map. A coalgebra is cocommutative if $\sigma_C \circ \Delta_C = \Delta_C$, where $\tau_C : C \otimes C \to C \otimes C$ is the flip map, $\tau_C(x \otimes y) = y \otimes x$. A bialgebra $B$ is both a unital algebra and a counital coalgebra together with compatibility relations, such as both the algebra product, $xy := m_B(x \otimes y)$, and the unit map, $e_B : \mathbb{R} \to B$, are coalgebra morphisms. This provides, for example, the identity $\Delta_B(xy) = \Delta_B(x)\Delta_B(y)$. The unit of $B$ is denoted by $1_B = e_B(1)$. A bialgebra is called graded if there are $\mathbb{R}$-vector subspaces $B_n$, $n \in \mathbb{N}$ such that $B = \bigoplus_{n \geq 0} B_n$ with $m_B(B_k \otimes B_l) \subseteq B_{k+l}$ and $\Delta_B B_n \subseteq \bigoplus_{k+l=n} B_k \otimes B_l$. Elements $x \in B_n$ are given a degree $\deg(x) = n$. Moreover, $B$ is called connected if the $\mathbb{R}$-vector subspace $B_0 = R1_B$. Define $B_n = \bigoplus_{n \geq 0} B_n$. In a connected graded bialgebra, the coproduct for any $x \in B_n$ is of the form

$$\Delta_B(x) = x \otimes 1_B + 1_B \otimes x + \Delta'_B(x) \in \bigoplus_{k+l=n} B_k \otimes B_l,$$

where $\Delta'_B(x) := \Delta_B(x) - x \otimes 1_B - 1_B \otimes x \in \bigoplus_{k+l=n} B_k \otimes B_l$ is the reduced coproduct. Note that in the following the use of subscripts on the structure maps is limited to the cases where there is potential for notational confusion.

Suppose $B$ is a bialgebra and $A$ is an $\mathbb{R}$-algebra with product $m_A : A \otimes A \to A$ and unit map $e_A$, e.g., $A = \mathbb{R}$ or $A = B$. The vector space $L(B, A)$ of linear maps from the bialgebra $B$ to $A$ together with the convolution product

$$(2.5) \quad \Phi \ast \Psi := m_A \circ (\Phi \otimes \Psi) \circ \Delta : B \to A,$$

where $\Phi, \Psi \in L(B, A)$, is an associative algebra with unit $\iota := e_A \circ \varepsilon$. A Hopf algebra $H$ is a bialgebra together with a particular $\mathbb{R}$-linear map called an antipode $S : H \to H$ which satisfies the Hopf algebra axioms [34, 36] and has the property of being an antihomomorphism for both the algebra and the coalgebra structures, i.e., $S(xy) = S(y)S(x)$ and $\Delta \circ S = (S \otimes S) \circ \sigma \circ \Delta$. When $A = H$, the necessarily unique antipode $S \in L(H, H)$ is the inverse of the identity map id : $H \to H$ with respect to the convolution product, that is,

$$(2.6) \quad S \ast \text{id} \ast S := m \circ (S \otimes \text{id}) \circ \Delta = \varepsilon \circ \varepsilon.$$
Group-like and primitive elements in a Hopf algebra satisfy $\Delta(x) = x \otimes x$ and $\Delta(y) = y \otimes 1 + 1 \otimes y$, respectively. Note that if $x_1, x_2$ are primitive in $H$, then $[x_1, x_2] := m_H(x_1 \otimes x_2) - m_H(x_2 \otimes x_1)$ is primitive as well. In particular, the set of primitive elements $P(H) \subset H$ is a Lie algebra. Another important observation is that a connected graded bialgebra $H = \bigoplus_{n \geq 0} H_n$ is always a connected graded Hopf algebra.

Let $H = \bigoplus_{n \geq 0} H_n$ be a connected graded commutative bialgebra, and suppose $A$ is a commutative unital algebra. Any $\Phi \in L(H, A)$ is called a character if $\Phi(1) = 1_A$ and

$$\Phi(xy) = \Phi(x)\Phi(y)$$

for all $x, y \in H$. The set of characters is denoted by $G_A \subset L(H, A)$ and forms a group with respect to the convolution product (2.5). The neutral element $\iota := e_A \circ \varepsilon$ in $G_A$ is given by $\iota(1) = 1_A$ and $\iota(x) = 0$ for elements $x$ in the augmentation ideal $\Ker(\varepsilon) = H_+ = \bigoplus_{n > 0} H_n$. The inverse of $\Phi \in G_A$ is given by composition with the antipode

$$\Phi^{-1} = \Phi \circ S,$$

which follows from (2.6) and the fact that $\Phi$ is an algebra morphism. Specifically,

$$\Phi \ast (\Phi \circ S) = \Phi \circ (\id \ast S) = \Phi \circ (S \ast \id) = (\Phi \circ S) \ast \Phi = \Phi \circ e \circ \varepsilon = e \circ \varepsilon.$$

An infinitesimal character with values in $A$ is a linear map $\xi \in L(H, A)$ such that for $x, y \in H^+$, $\xi(m_H(x \otimes y)) = 0$, which implies $\xi(1) = 0$. The linear space of infinitesimal characters $g_A \subset L(H, A)$ forms a Lie algebra with respect to the Lie bracket defined in terms of the convolution product (2.5). For $\alpha \in g_A$ and any $x \in H_n$ the exponential $\exp^*(\alpha)(x) := \sum_{j \geq 0} \frac{1}{j!} \alpha^j(x)$ is a finite sum terminating at $j = n$.

**Proposition 2.6.** $\exp^*$ restricts to a bijection from $g_A$ onto $G_A$.

For details see references [8, 10]. The inverse of $\exp^*$ is given by $\log^*(\iota + \gamma)(x) = \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} \gamma^k(x)$, where again the sum terminates at $k = n$ for any $x \in H_n$.

Finally, a brief description of the relationship between the group $G_A \subset L(H, A)$ and the Hopf algebra $H$ is given. The reader is referred to the paper of Frabetti and Manchon [15] for details and additional references. By definition, elements in $G_A$ map all of $H$ into the commutative unital algebra $A$. However, any element $x \in H$ can also be seen as an $A$-valued function on $G_A$. Indeed, let $\Phi \in G_A$; then $x(\Phi) := \Phi(x) \in A$ and the usual pointwise product of functions $(xy)(\Phi) = x(\Phi)y(\Phi)$ follows from (2.7) since $\Phi \in G_A$. The definition of the convolution product (2.5) in terms of the coproduct of $H$ implies a natural coproduct on functions $x \in H$, that is, $\Delta(x)(\Phi, \Psi) := (\Phi \ast \Psi)(x) \in A$. Similarly, the inverse of $G_A$ as well as its unit correspond naturally to the antipode and counit map on $H$, respectively. This reversed perspective on the relationship between $H$ and its group of characters $G_A$ allows one to interpret $H$ as the (Hopf) algebra of coordinate functions of the group $G_A$. More precisely, $H$ contains the representative functions over $G_A$. The reader is directed to Cartier’s work [5] for a comprehensive review of this topic. In this paper, the starting point will be a particular group whose product, unit, and inverse are used to identify its Hopf algebra of coordinate functions.

**3. Output affine feedback transformation group.** In general, if $G$ is a group and $S$ a given set, then $G$ is said to act as a transformation group on the right of $S$ if there exists a mapping $\phi : S \times G \rightarrow S : (h, g) \mapsto hg$ such that
i. \( h1 = h \), where 1 is the identity element of \( G \);

ii. \( h(g_1g_2) = (h(g_1))g_2 \) for all \( g_1, g_2 \in G \).

The action is said to be free if \( hg = h \) implies that \( g = 1 \). It will be evident in what follows that it is most convenient to redefine \( \mathbb{R} \langle L \rangle := \mathbb{R} \langle X \rangle \times \mathbb{R} \langle X \rangle \) and to identify the output affine feedback transformation group with those elements of \( \mathbb{R} \langle (X) \rangle \) which are invertible in some sense. Each tuple \( c_{d} := (c_L, c_R) \in \mathbb{R} \langle L \rangle \) gives rise to an operator of the form \( F_{c_{d}}[u] := F_{c_L}[u] + F_{c_R}[u] \), and by design the group product \( c_{d} \circ d_{s} \) for elements in \( \mathbb{R} \langle L \rangle \) will satisfy the identity \( F_{c_{d}} \circ F_{d_{s}} = F_{c_{d} \circ d_{s}} \) when \( c_{d} \circ d_{s} \in \mathbb{R} \langle L \rangle \). In addition, a second product \( c \circ d \) is defined which describes the action of the group \( \mathbb{R} \langle (X) \rangle \) on \( \mathbb{R} \langle (X) \rangle \) so that \( F_{c} \circ F_{d_{s}} = F_{c \circ d_{s}} \) when the series involved are all locally convergent. This product, referred to as the mixed composition product, plays a central role in the paper, so it will be the starting point for the section. Once these concepts are in place and the transformation group is established, the affine extension of the output feedback product in Theorem 2.1 will be given. Henceforth, the focus is on the SISO case.

For a fixed \( d_{s} = (d_L, d_R) \in \mathbb{R} \langle (X) \rangle \), the mixed composition product is defined in terms of an algebra morphism \( \phi_{d} \) which takes words in \( X^* \) to a composition of simple vector space endomorphisms involving only the shuffle product and catenation with a single letter \( x_0 \) or \( x_1 \). This sets up an inductive definition which provides some insight into the core combinatorial structures at play. The reader is referred in particular to [13, 14] for a broader theoretical view behind this approach.

**Definition 3.1.** Let \( d_{s} = (d_L, d_R) \in \mathbb{R} \langle (X) \rangle \). Define the mixed composition product mapping \( \mathbb{R} \langle (X) \rangle \times \mathbb{R} \langle (X) \rangle \) into \( \mathbb{R} \langle (X) \rangle \) as

\[
(3.1) \quad c \circ d_{s} = \phi_{d}(c)(1) = \sum_{\eta \in X^*} (c, \eta) \phi_{d}(\eta)(1),
\]

where \( \phi_{d} \) is the continuous (in the ultrametric sense) algebra homomorphism from \( \mathbb{R} \langle (X) \rangle \) to \( \text{End}(\mathbb{R} \langle (X) \rangle) \) uniquely specified by \( \phi_{d}(x_i \eta) = \phi_{d}(x_i) \circ \phi_{d}(\eta) \) with

\[
(3.2) \quad \phi_{d}(x_0 \eta)(e) = x_0 e, \quad \phi_{d}(x_1 \eta)(e) = x_1 (d_L \circ e) + x_0 (d_R \circ e)
\]

for any \( e \in \mathbb{R} \langle (X) \rangle \), and where \( \phi_{d}(\emptyset) \) denotes the identity map on \( \mathbb{R} \langle (X) \rangle \).

**Example 3.2.** Suppose \( c = 2x_0 + x_1x_0 \) and \( d_{s} = (d_L, d_R) = (3x_1, -x_1^2) \). Then

\[
(3.3) \quad c \circ d_{s} = 2\phi_{d}(x_0)(1) + \phi_{d}(x_1)(1),
\]

where \( \phi_{d}(x_0)(e) = x_0 e \) and \( \phi_{d}(x_1)(e) = x_1 (3x_1 \circ e) + x_0 (-x_1^2 \circ e) \). Therefore, \( c \circ d_{s} = 2x_0 + 3x_1^2x_0 + 3x_1x_0x_1 - x_0x_1^2x_0 - x_0x_1x_0x_1 - x_0^2x_0^2 \).

The modified composition product referred to in section 2 and used in earlier work on output feedback [17, 18, 19, 28, 38] corresponds here to the special case where \( d_L = 1 \). Some fundamental properties of the mixed composition product are given next.

**Lemma 3.3.** The mixed composition product (3.1) (1) is left \( R \)-linear;

(2) satisfies \( c \circ (1, 0) = c \);

(3) satisfies \( c \circ d_{s} = k \in R \) for any fixed \( d_{s} \) if and only if \( c = k \);

(4) satisfies \((x_0 c) \circ d_{s} = x_0 (c \circ d_{s}) \) and \((x_1 c) \circ d_{s} = x_1 (d_L \circ (c \circ d_{s})) + x_0 (d_R \circ (c \circ d_{s})) \);

(5) distributes to the left over the shuffle product.
Proof.
(1) This fact follows directly from the definition of the mixed composition product.
(2) The claim is immediate since \( \phi_{(1,0)}(\eta)(1) = \eta \).
(3) The only nontrivial assertion is that \( c \circ d_\delta = k \) implies \( c = k \). This claim is best handled in a Hopf algebra setting. So the proof is deferred until section 4.
(4) Observe
\[
(x_0 c) \circ d_\delta = \phi_d(x_0 c)(1) = \phi_d(x_0) \circ \phi_d(c)(1) = x_0 (c \circ d_\delta),
\]
\[
(x_1 c) \circ d_\delta = \phi_d(x_1 c)(1) = \phi_d(x_1) \circ \phi_d(c)(1) = x_1 (d_L \omega (c \circ d_\delta))
\]
\[+ x_0 (d_R \omega (c \circ d_\delta)).\]
(5) For any \( e_\delta \in \mathbb{R} \langle \langle X_\delta \rangle \rangle \), one can define a shuffle product on \( \text{End}(\mathbb{R} \langle \langle X \rangle \rangle) \) via
\[
\phi_e(x_\eta) \omega \phi_e(x_\xi) = \phi_e(x_i) \circ [\phi_e(\eta) \omega \phi_e(x_\xi)] + \phi_e(x_j) \circ [\phi_e(x_\eta) \omega \phi_e(\xi)].
\]
Therefore, \( \phi_e \) acts as an algebra map between the shuffle algebra on \( \mathbb{R} \langle \langle X \rangle \rangle \) and the shuffle algebra on \( \text{End}(\mathbb{R} \langle \langle X \rangle \rangle) \). That is, \( \phi_e(c \omega d) = \phi_e(c) \omega \phi_e(d) \).
Hence, \( (c \omega d) \circ e_\delta = \phi_e(c \omega d)(1) = \phi_e(c)(1) \omega \phi_e(d)(1) = (c \circ e_\delta) \omega (d \circ e_\delta). \]

It is easily checked that \( \text{dist}(c_\delta, d_\delta) := \max(\text{dist}(c_L, d_L), \text{dist}(c_R, d_R)) \) is an ultrametric on \( \mathbb{R} \langle \langle X_\delta \rangle \rangle \), and \( \mathbb{R} \langle \langle X_\delta \rangle \rangle \) is complete.\(^3\) The following lemma states that the mixed composition product reduces ultrametric distance in a certain sense.

**Lemma 3.4.** For any \( c \in \mathbb{R} \langle \langle X \rangle \rangle \) and \( d_{\delta,1}, d_{\delta,2} \in \mathbb{R} \langle \langle X_\delta \rangle \rangle \) it follows that
\[
\text{dist}(c \circ d_{\delta,1}, c \circ d_{\delta,2}) \leq \sigma^{\text{ord}(c')} \text{dist}(d_{\delta,1}, d_{\delta,2}),
\]
where \( c = (c, \emptyset) \emptyset + c' \) with \( c' \) proper.

**Proof.** For a fixed \( d_R \), consider the map \( d_L \mapsto c \circ (d_L, d_R) \). Likewise, for a fixed \( d_L \) there is a companion map \( d_R \mapsto c \circ (d_L, d_R) \). It is first shown that on the ultrametric space \( \langle \langle X \rangle \rangle \langle \langle X \rangle \rangle \), dist),
\[
(3.3a) \quad \text{dist}(c \circ (d_{L,1}, d_R), c \circ (d_{L,2}, d_R)) \leq \sigma^{\text{ord}(c')} \text{dist}(d_{L,1}, d_{L,2}),
\]
\[
(3.3b) \quad \text{dist}(c \circ (d_L, d_{R,1}), c \circ (d_L, d_{R,2})) \leq \sigma^{\text{ord}(c')} \text{dist}(d_{R,1}, d_{R,2}).
\]
Inequality (3.3a) is clearly true when \( d_{L,1} = d_{L,2} \). When \( d_{L,1} \neq d_{L,2} \), it is first necessary to show that for any \( \eta \in X^* \)
\[
(3.4) \quad \text{ord}(\phi_{d_\delta}(\eta)(1) - \phi_{d_\delta}(\eta)(1)) \geq |\eta| + \text{ord}(d_{L,1} - d_{L,2}).
\]
The proof is by induction on the length of \( \eta \). The claim is trivial when \( \eta \) is empty or a single letter. Assume the inequality holds for words up to length \( k \geq 0 \). For any \( x_0 \eta \) with \( \eta \in X^k \), inequality (3.4) follows directly from the induction hypothesis. The case for \( x_1 \eta \) is handled as follows:
\[
\text{ord}([\phi_{d_\delta}(x_1 \eta)(1) - \phi_{d_\delta}(x_1 \eta)(1)]
\]
\[
= \text{ord}([x_1 (d_{L,1} \omega \phi_{d_\delta}(\eta)(1) - d_{L,2} \omega \phi_{d_\delta}(\eta)(1))] + x_0 [d_R \omega (\phi_{d_\delta}(\eta)(1) - \phi_{d_\delta}(\eta)(1))]
\]
\[
= \text{ord}([x_1 (d_{L,1} - d_{L,2}) \omega \phi_{d_\delta}(\eta)(1) + d_{L,2} \omega (\phi_{d_\delta}(\eta)(1) - \phi_{d_\delta}(\eta)(1))] + x_0 [d_R \omega (\phi_{d_\delta}(\eta)(1) - \phi_{d_\delta}(\eta)(1))]
\]

\(^3\) Using dist for both the ultrametric on \( \mathbb{R} \langle \langle X \rangle \rangle \) and \( \mathbb{R} \langle \langle X_\delta \rangle \rangle \) should cause minimal confusion since their arguments are distinct.
Therefore, (3.4) holds for any \( \eta \in X^* \). The inequality (3.3a) is now derived. Observe

\[
\text{dist}(c \circ (d_{L,1}, d_R), c \circ (d_{L,2}, d_R)) \\
= \text{dist}(c' \circ (d_{L,1}, d_R), c' \circ (d_{L,2}, d_R)) = \sigma^{\text{ord}(\sum_{i} c'_i \circ (\phi_{d_i}(\eta)(1) - \phi_{d_j}(\eta)(1)))} \\
\leq \sigma^{\min_{\eta \in \text{supp}(c')}} \text{ord}(\phi_{d_i}(\eta)(1) - \phi_{d_j}(\eta)(1)) \\
= \sigma^{\text{ord}(c')} \text{dist}(d_{L,1} - d_{L,2}).
\]

The proof for (3.3b) is completely analogous. The final step of the proof is to employ the ultrametric property in conjunction with (3.3). Observe

\[
\text{dist}(c \circ d_\delta, c \circ d_\delta) \\
= \text{dist}(c \circ (d_{L,1}, d_{R,1}), c \circ (d_{L,2}, d_{R,2})) \\
\leq \max(\text{dist}(c \circ (d_{L,1}, d_{R,1}), c \circ (d_{L,2}, d_{R,1})), \text{dist}(c \circ (d_{L,2}, d_{R,1}), c \circ (d_{L,2}, d_{R,2}))) \\
\leq \sigma^{\text{ord}(c')} \max(\text{dist}(d_{L,1}, d_{L,2}), \text{dist}(d_{R,1}, d_{R,2})) = \sigma^{\text{ord}(c')} \text{dist}(d_{\delta,1}, d_{\delta,2}).
\]

The group product on \( \mathbb{R}(\langle X_\delta \rangle) \) is now defined in terms of the mixed composition product. Its basic properties are given in the subsequent lemma.

**Definition 3.5.** The composition product on \( \mathbb{R}(\langle X_\delta \rangle) \) is defined as

\[
c_\delta \circ d_\delta = ((c_L \circ d_\delta) \leftarrow d_L, (c_L \circ d_\delta) \leftarrow d_R + c_R \circ d_\delta).
\]

**Lemma 3.6.** The composition product on \( \mathbb{R}(\langle X_\delta \rangle) \)

1. is left \( \mathbb{R} \)-linear;
2. satisfies \((c \circ d_\delta) \circ e_\delta = c \circ (d_\delta \circ e_\delta) \) (mixed associativity);
3. is associative.

**Proof.**

1. This property follows from the left linearity of the mixed composition product.
2. In light of the first item, it is sufficient to prove the claim only for \( c = \eta \in X^k \), \( k \geq 0 \). The cases \( k = 0 \) and \( k = 1 \) are trivial. Assume the claim holds up to some fixed \( k \geq 0 \). Then via Lemma 3.3(4) and the induction hypothesis it follows that

\[
((x_0 \eta) \circ d_\delta) \circ e_\delta = (x_0(\eta \circ d_\delta)) \circ e_\delta = x_0(\eta \circ d_\delta) \circ e_\delta \\
= x_0(\eta \circ (d_\delta \circ e_\delta)) = (x_0 \eta) \circ (d_\delta \circ e_\delta).
\]

In a similar fashion, apply the properties in Lemma 3.3(1), (4), and (5) to get

\[
\begin{align*}
((x_1 \eta) \circ d_\delta) \circ e_\delta \\
= [x_1(d_L \leftarrow (\eta \circ d_\delta)) + x_0(d_R \leftarrow (\eta \circ d_\delta))] \circ e_\delta \\
= [x_1(d_L \leftarrow (\eta \circ d_\delta))] \circ e_\delta + [x_0(d_R \leftarrow (\eta \circ d_\delta))] \circ e_\delta \\
= x_1[e_L \leftarrow ((d_L \leftarrow (\eta \circ d_\delta)) \circ e_\delta)] + x_0[e_R \leftarrow ((d_L \leftarrow (\eta \circ d_\delta)) \circ e_\delta)] \\
+ x_0[(d_R \leftarrow (\eta \circ d_\delta)) \circ e_\delta]
\end{align*}
\]
The Fliess operator interpretations of the two formal power series products described above are given in the next theorem.

\textbf{Theorem 3.7.} For any \( c \in \mathbb{R}_LC\langle \langle X \rangle \rangle \) and \( d_3, d_5 \in \mathbb{R}_LC\langle \langle X \rangle \rangle \) such that \( c \circ d_3 \in \mathbb{R}_LC\langle \langle X \rangle \rangle \) and \( c \circ d_5 \in \mathbb{R}_LC\langle \langle X \rangle \rangle \), the following identities hold:

1. \( F_c \circ F_{d_3} = F_{c \circ d_3} \),
2. \( F_{c_3} \circ F_{d_5} = F_{c_3 \circ d_5} \).

\textit{Proof.}

1. It is sufficient to prove the claim for \( c = \eta \in X^* \). This is done by induction on the length of \( \eta \). The case for the empty word is trivial. Assume the identity holds for words \( \eta \in X^k \) up to some fixed length \( k \geq 0 \). Then

\[
E_{x_0 \eta} \circ F_{d_3}[u](t) = \int_{t_0}^{t} E_{\eta}(F_{d_3}[u])(\tau, t_0) \, d\tau = \int_{t_0}^{t} F_{\eta \circ d_3}[u](\tau) \, d\tau = F_{x_0(\eta \circ d_3)}[u](t) = F_{(x_0 \eta) \circ d_3}[u](t).
\]
Similarly,

\[
E_{x, \eta} \circ F_{d_L}[u](t) = \int_{t_0}^{t} F_{\eta \circ d_L[u]}[u](\tau)[F_{d_L}[u](\tau)u(\tau) + F_{d_R}[u](\tau)] d\tau
\]

\[
= \int_{t_0}^{t} F_{d_L \omega (\eta \circ d_L)[u]}[u](\tau)u(\tau) + F_{(\eta \circ d_L) \omega d_L}[u](\tau) d\tau
\]

\[
= F_{x(\eta \circ d_L)) \omega (\eta \circ d_L) \omega d_R}[u](t) = F_{(x, \eta)\circ d_L}[u](t).
\]

Hence, the claim holds for all \( \eta \in X^* \).

(2) Observe

\[
F_{c_L} \circ F_{d_L}[u] = F_{c_L}[u] + F_{d_R}[u]
\]

\[
= F_{c_L}(F_{d_L}[u])F_{d_L}[u] + F_{c_L}[F_{d_L}[u]]F_{d_R}[u] + F_{c_R}[F_{d_L}[u]]
\]

\[
= F_{(c_L \circ c_R) \circ d_L}[u] + F_{(c_L \circ c_R) \circ d_R}[u] + F_{c_R \circ d_R}[u]
\]

\[
= F_{(c_L \circ c_R) \circ d_L}[u] + F_{(c_L \circ c_R) \circ d_R}[u] + F_{c_R \circ d_R}[u] = F_{c_L \circ c_R \circ d_L}[u].
\]

Let \( \mathbb{R}^{\text{np}}(\langle X_d \rangle) \) denote the subset of \( \mathbb{R}(\langle X_d \rangle) \) whose left series are nonproper.

**Theorem 3.8.** The set \( \mathbb{R}^{\text{np}}(\langle X_d \rangle) \) is a group.

**Proof.** From Lemma 3.3(2), it follows directly that \( c\circ (1,0) = c\circ. \) Using the identities \( 1 \circ c = 1 \) and \( 0 \circ c = 0 \) from Lemma 3.3(3), it is straightforward to see that \( (1,0) \circ c = c \circ. \) Associativity of the group product was established in Lemma 3.6(3).

So the only open issue is the existence of inverses. Suppose \( c \circ \) is fixed, and one seeks a right inverse \( c^{-1} = (c_L^{-1}, c_R^{-1}) \), that is, \( c \circ c^{-1} = (1,0). \)

Then it follows directly from Theorem 2.4 and Definition 3.5 that

\[
c_L^{-1} = (c_L \circ (c_L, c_R^{-1})) \circ (c_L, c_R) = (c_L, c_R^{-1}) \circ (c_L, c_R) = c_L, \quad c_R^{-1} = (c_R \circ (c_L^{-1}, c_R)) \circ (c_L, c_R) = (c_R, c_L^{-1}) \circ (c_L, c_R) = c_R.
\]

It is first shown that the mapping

\[
S_R : (c_L, c_R) \mapsto ((c_L \circ (c_L, c_R)) \circ (c_L, c_R)) = (c_L, c_R)^{-1}
\]

is an ultra-metric contraction on \( \mathbb{R}(\langle X_d \rangle) \) and therefore has a unique fixed point, which by design is a right inverse, \( c\circ^{-1} \). Note that for any \( c \circ \) it follows that \( (S_R(c_L, c_R), \emptyset) = (c_L, \emptyset)^{-1} \neq 0. \) Thus, the fixed point will always be in the group. Then it is shown that this same series is also a left inverse, that is, \( c^{-1} \circ c = (1,0) \), or equivalently,

\[
c_L = (c_L^{-1} \circ (c_L, c_R)) \circ (c_L, c_R) = c_L, \quad c_R = (c_R^{-1} \circ (c_L, c_R)) \circ (c_L, c_R) = c_R.
\]

To establish the first claim, observe via Corollary 2.3 and Lemma 2.5 that for any \( c \circ, \tilde{c} \circ \in \mathbb{R}(\langle X_d \rangle) \)

\[
\text{dist}(S_R(c\circ), S_R(\tilde{c} \circ)) = \max(\text{dist}((c_L \circ (c_L, c_R)) \circ (c_L, c_R), (c_L \circ (\tilde{c}_L, \tilde{c}_R))) \circ (c_L, c_R))
\]

\[
\leq \max(\text{dist}((c_L \circ (c_L, c_R), c_L \circ (\tilde{c}_L, \tilde{c}_R)), (c_R \circ (c_L, c_R), c_R \circ (\tilde{c}_L, \tilde{c}_R)))
\]

\[
= \max(\text{dist}(c_L \circ (c_L, c_R), c_L \circ (\tilde{c}_L, \tilde{c}_R)), (c_R \circ (c_L, c_R), c_R \circ (\tilde{c}_L, \tilde{c}_R))).
\]

4The notation \( c^{-1} \), \( c \in \mathbb{R}(\langle X \rangle) \), is used in this situation instead of \( c^{-1} \) to distinguish it from \( c^{-1} \).
Thus, from Lemma 3.4 it follows that
\[ \text{dist}(S_R(e_\delta), S_R(\bar{e}_\delta)) \leq \max(\sigma^{\text{ord}}(c'), \text{dist}((e_L, e_R), (\bar{e}_L, \bar{e}_R))), \]
setting
\[ \sigma^{\text{ord}}(c') \text{dist}((e_L, e_R), (\bar{e}_L, \bar{e}_R))) \]
\[ \leq \sigma \text{dist}(e_\delta, \bar{e}_\delta). \]

To address the second claim, suppose \( c^{-1}_\delta \) satisfies (3.5a). Then it follows that
\[ (c_L \circ c^{-1}_\delta) \circ \bar{c}_L^{-1} = 1, \]
\[ (c_L \circ \bar{c}_\delta^{-1}) \circ (c_L \circ \bar{c}_\delta^{-1}) = 1. \]

Using Lemma 3.6(2) and Lemma 3.3(5) gives
\[ (c_L \circ \bar{c}_\delta^{-1}) \circ ((c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_\delta^{-1}) = 1, \]
\[ (c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_\delta^{-1} = 1. \]

Applying Lemma 3.3(3) then yields \( c_L \circ \bar{c}_\delta^{-1} = 1, \) and thus, \( c_L = (c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_\delta^{-1}, \) which is (3.6a). If, in addition, \( c^{-1}_\delta \) also satisfies (3.5b), then substituting (3.5a) into (3.5b) gives \( c_R^{-1} = -(c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_\delta^{-1} \). Therefore, in a similar fashion,
\[ -(c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_L^{-1} = c_R \circ \bar{c}_\delta^{-1}, \]
\[ (c_L \circ \bar{c}_\delta^{-1}) \circ \bar{c}_\delta^{-1} = c_R \circ \bar{c}_\delta^{-1}, \]
\[ (c_R + c_L \circ (c_R^{-1} \circ \bar{c}_\delta^{-1})) \circ \bar{c}_\delta^{-1} = 0. \]

Once again applying Lemma 3.3(3) gives \( c_R + c_L \circ (c_R^{-1} \circ \bar{c}_\delta) = 0, \) which is equivalent to (3.6b).

The following theorem is the first main result of this section. The issue of freeness will be addressed in section 5.

**Theorem 3.9.** The group \( (\mathbb{R}_{np} X^{\delta}, \circ, (1, 0)) \) acts as a right transformation group on \( \mathbb{R} \langle X \rangle \) via the action \( c \circ d_\delta. \)

**Proof.** See Lemma 3.3(2) and Lemma 3.6(2).

The second main result of the section is a formula for the group element \( e_\delta \) which describes the output affine feedback connection shown in Figure 1. This defines the output affine feedback product.

**Theorem 3.10.** Let \( c \in \mathbb{R}_{LC} \langle X \rangle \) and \( d_\delta = (d_L, d_R) \in \mathbb{R}_{LC} \langle X \rangle \). Then the generating series for the output affine feedback connected system is \( c @ d_\delta := c \circ e_\delta^{-1}, \) where
\[ e_\delta = ((d_L \circ c) \circ \bar{c})^{-1}, - (d_L \circ c) \circ \bar{c}^{-1} \circ (d_R \circ c), \]
provided that \( e_\delta^{-1} \in \mathbb{R}_{LC} \langle X \rangle \) and \( c @ d_\delta \in \mathbb{R}_{LC} \langle X \rangle \). In addition, \( c @ d_\delta \) satisfies the fixed-point equation
\[ c @ d_\delta = c \circ (d_L \circ (c @ d_\delta), d_R \circ (c @ d_\delta)). \]

**Proof.** From Figure 1 it is clear that \( u = F_{d_L} [y] + F_{d_R} [y] = F_{d_L \circ c} [u] + F_{d_R \circ c} [u]. \) Therefore, using the nonproperness of \( d_L \circ c, \) and the fact that the shuffle product, shuffle inverse, and composition product on \( \mathbb{R}_{LC} \langle X \rangle \) all preserve local convergence, it follows that
\[ u - F_{d_R \circ c} [u] = F_{d_L \circ c} [u], \]
respectively. the antipode of this Hopf algebra, and some illustrative examples are given.

\( R \) can assume this form by simply rescaling \( y \) is no loss of generality since the generating series of any Fliess operator \( c \) reduces to \( R \) the Hopf algebra of coordinate functions for the group \( \langle \langle X \rangle \rangle \). It is first necessary to restrict the setup to the subset of series having the

\( \eta \) observe that the output affine feedback equation is

\( \text{(3.7)} \)–\( \text{(3.8)} \) reduce to their expected forms

\( \text{(4.1)} \) \( b_\eta : \mathbb{R}_{\text{np}} \langle \langle X \rangle \rangle \to \mathbb{R} : c_\delta \mapsto (c_L, \eta) \), \( a_\eta : \mathbb{R}_{\text{np}} \langle \langle X \rangle \rangle \to \mathbb{R} : c_\delta \mapsto (c_R, \eta) \), respectively.\(^5\) Let \( V \) denote the \( \mathbb{R} \)-vector space spanned by these maps. Define the corresponding free commutative algebra, \( H \), with product \( \mu : h_\eta \otimes \tilde{h}_\xi \mapsto h_\eta \tilde{h}_\xi \), \( h, \tilde{h} \in \{a, b\} \) and unit \( 1 \) which maps every \( c_\delta \in \mathbb{R}_{\text{np}} \langle \langle X \rangle \rangle \) to 1. Consequently, definition \( \text{(4.1)} \) is extended multiplicatively, that is, for \( h, \tilde{h} \in \{a, b\} \) the product \( h_\eta \tilde{h}_\xi \) corresponds to the pointwise product of coordinate functions with respect to \( c_\delta = (c_L, c_R) \in \mathbb{R}_{\text{np}} \langle \langle X \rangle \rangle \)

\( \text{(4.2)} \) \( h_\eta \tilde{h}_\xi (c_\delta) := h_\eta (c_\delta) \tilde{h}_\xi (c_\delta) \).

\(^5\)This terminology will be justified later.
The degree of a coordinate function is taken as $\deg(b_\eta) = 2|\eta|_{x_0} + |\eta|_{x_1}$, $\deg(a_\eta) = 2|\eta|_{x_0} + |\eta|_{x_1} + 1$, and $\deg(1) = 0$. Therefore, $V$ is a connected graded vector space, that is, $V = \bigoplus_{n \geq 0} V_n$ with $V_n$ denoting the span of all coordinate functions of degree $n$ and $V_0 = \mathbb{R}1$. Let $V_+ = \bigoplus_{n \geq 1} V_n$. Similarly, $H$ has the connected graduation $H = \bigoplus_{n \geq 0} H_n$ with $H_0 = \mathbb{R}1$.

Three coproducts are now introduced. The first coproduct is used to define the Hopf algebra on $H$. The remaining two products provide an inductive method in which to compute it. Recalling that $(c_\circ \delta)_{L} = (c_L \circ \delta)_L \omega d_L$ and $(c_\circ \delta)_R = (c_L \circ \delta)_R \omega d_R + c_R \circ \delta)_R$, define coproduct $\Delta$ for any $b_\eta, a_\eta \in V_+$ such that

\begin{align*}
\Delta b_\eta(c_\circ \delta) = b_\eta(c_\circ \delta) = ((c_L \circ \delta)_L \omega d_L, \eta), \\
\Delta a_\eta(c_\circ \delta) = a_\eta(c_\circ \delta) = ((c_L \circ \delta)_R \omega d_R, \eta) + (c_R \circ \delta)_R, \eta).
\end{align*}

The coassociativity of $\Delta$ follows directly from the associativity of the group product on $\mathbb{R}_n((X))$. The second coproduct is $\Delta_{\omega \iota}(V_+) \subset V_+ \otimes V_+$, which is defined for any $\hat{h}, \hat{h} \in V_+$ by

\begin{align*}
\Delta_{\omega \iota}(\hat{h}_\eta) &= \hat{h}_\eta \otimes \hat{h}_\eta, \\
\Delta_{\omega \iota}(\hat{h}_\eta) \circ \theta_k &= (\theta_k \otimes 1 + 1 \otimes \theta_k) \circ \Delta_{\omega \iota}(\hat{h}_\eta),
\end{align*}

where $\theta_k$ denotes the endomorphism on $V_+$ specified by $\theta_k h_\eta = h_{\omega x^k \eta} \eta = h_\eta$ for $k = 0, 1$. Clearly this coproduct can be computed recursively. Finally, consider for any $\eta \in X^*$ the coproduct $\Delta_{\omega \iota}(V_+) \subset V_+ \otimes H$, where

$$\Delta b_\eta(c_\circ \delta, \eta) = (c_L \circ \delta, \eta), \quad \Delta a_\eta(c_\circ \delta) = (c_R \circ \delta, \eta).$$

In either case, using the notation of Sweedler [36],

$$(\Delta b_\eta(c_\circ \delta), d_\delta) = \sum h_{\eta(1)}(c_\circ \delta) h_{\eta(2)}(d_\delta) = \sum h_{\eta(1)} \otimes h_{\eta(2)}(c_\circ \delta), d_\delta),$$

where $h_{\eta(1)} \in V_+$ and $h_{\eta(2)} \in H$. The sums are taken over all the terms that appear in the respective composition in (4.5), and the specific nature of the factors $h_{\eta(1)}$ and $h_{\eta(2)}$ is not important here. Like the coproduct $\Delta_{\omega \iota}$, this coproduct can also be computed inductively as described next.

**Lemma 4.1.** 
The following identities hold:

1. $\Delta h_\eta = h_\eta \otimes 1$,
2. $\Delta \circ \theta_1 = (\theta_1 \otimes \mu) \circ (\Delta \otimes \text{id}) \circ \Delta_{\omega \iota}$,
3. $\Delta \circ \theta_0 = (\theta_0 \otimes \mu) \circ \Delta + (\theta_1 \otimes \mu) \circ (\Delta \otimes \text{id}) \circ \Delta_{\omega \iota}$,

where id denotes the identity map on $H$.

**Proof.**

1. Assume $h = a$ and write $c_R = x_0 c_R^0 + x_1 c_R^1 R + (c_R, \emptyset)$ with $c_R^0, c_R^1 \in \mathbb{R}((X))$.

Then from Lemma 3.3(1) and (4) it follows that

\begin{align*}
(c_R \circ \delta)_R \delta d_\delta &= (x_0 c_R^0 \circ \delta)_R \delta d_\delta + (x_1 c_R^1 \circ \delta)_R \delta d_\delta + (c_R, \emptyset) \circ \delta d_\delta \\
&= x_0 (c_R^0 \circ \delta)_R \delta d_\delta + x_1 (c_R^1 \circ \delta)_R \delta d_\delta + x_0 (c_R \circ \delta)_R \omega (c_R^1 \circ \delta)_R \delta d_\delta) + (c_R, \emptyset).
\end{align*}

Therefore, $\Delta a_\eta(c_\circ \delta, \emptyset) = (c_R \circ \delta, \emptyset) = (a_\emptyset \otimes 1)(c_\circ \delta)$. A similar argument holds when $h = b$. 


(2) If \( h = a \), then
\[
(\Delta \circ \Theta_i) a_\eta(c_\delta, d_\delta) = \Delta a_{\eta,\eta}(c_\delta, d_\delta) = (c_R \circ d_\delta, x, \eta)
\]
\[
= (x_i^{-1}[x_0(c_R^0 \circ d_\delta) + x_1(d_L \circ (c_R^1 \circ d_\delta)) + x_0(d_R \circ (c_R^1 \circ d_\delta)), \eta)
\]
\[
= (1_{i0}[c_R^1 \circ d_\delta + d_R \circ (c_R^1 \circ d_\delta)] + 1_{i1}[d_L \circ (c_R^1 \circ d_\delta)], \eta),
\]
where \( x_i^{-1}(\cdot) \) is the \( \mathbb{R} \)-linear left-shift operator specified by \( x_i^{-1}(\eta) = \eta' \) when \( \eta = x_i\eta' \) with \( \eta' \in X^* \) and zero otherwise, and \( 1_{xy} \) is the indicator function. So \( 1_{xy} = 1 \) when \( x = y \) and zero otherwise. Letting \( c_\delta^1 = (c_L, c_R^1) \), it follows that
\[
(\Delta \circ \Theta_i) a_\eta(c_\delta, d_\delta) = (d_L \circ (c_R^1 \circ d_\delta), \eta) = \sum_{\xi, \nu \in X^*} (c_R^1 \circ d_\delta, \xi)(d_L, \nu)(\xi \circ \nu, \eta)
\]
\[
= \sum_{\xi, \nu \in X^*} \Delta a_\xi(c_\delta^1, d_\delta) b_\nu(d_\delta)(\xi \circ \nu, \eta)
\]
\[
= \sum_{\xi, \nu \in X^*} \sum_{\xi(1)} a_{\xi(1)} \otimes a_{\xi(2)}(c_\delta^1, d_\delta) b_\nu(d_\delta)(\xi \circ \nu, \eta)
\]
\[
= \sum_{\xi, \nu \in X^*} \sum_{\xi(1)} \theta_1(a_{\xi(1)} \otimes a_{\xi(2)}(c_\delta^1, d_\delta) b_\nu(d_\delta)(\xi \circ \nu, \eta)
\]
\[
= (\theta_1 \circ \text{id}) \circ \sum_{\xi, \nu \in X^*} \Delta a_\xi \otimes \nu(c_\delta^1, d_\delta)(\xi \circ \nu, \eta)
\]
\[
= (\theta_1 \otimes \mu) \circ (\Delta \otimes \text{id}) \circ \Delta^b a_\eta(c_\delta, d_\delta).
\]
The proof when \( h = b \) is perfectly analogous.

(3) If \( h = a \), then \( (\Delta \circ \Theta_0) a_\eta(c_\delta, d_\delta) = (c_R^0 \circ d_\delta, \eta) + (d_R \circ (c_R^1 \circ d_\delta), \eta) \). From this point, the method of proof is exactly the same as that in part (2) modulo the fact that \( \Delta^a_{\alpha,\alpha} \) is used here due to the presence of \( d_R \) instead of \( d_L \) in the shuffle product.

Example 4.2. Applying the identities in Lemma 4.1 gives the first few \( \Delta \) coproduct terms ordered by degree \( n_b \) (\( h = b \)), \( n_a \) (\( h = a \)):

\[
n_b, n_a = 0, 1 : \Delta h_0 = h_0 \otimes 1,
\]
\[
n_b, n_a = 1, 2 : \Delta h_{x_1} = h_{x_1} \otimes 1,
\]
\[
n_b, n_a = 2, 3 : \Delta h_{x_0} = h_{x_0} \otimes 1 + h_{x_1} \otimes a_\theta,
\]
\[
n_b, n_a = 2, 3 : \Delta h_{x_1^2} = h_{x_1^2} \otimes 1 + h_{x_1} \otimes b_{x_1},
\]
\[
n_b, n_a = 3, 4 : \Delta h_{x_0x_1} = h_{x_0x_1} \otimes 1 + h_{x_1} \otimes a_{x_1} + h_{x_1^2} \otimes a_\theta,
\]
\[
n_b, n_a = 3, 4 : \Delta h_{x_1x_0} = h_{x_1x_0} \otimes 1 + h_{x_1} \otimes b_{x_0} + h_{x_1^2} \otimes a_\theta,
\]
\[
n_b, n_a = 3, 4 : \Delta h_{x_1^3} = h_{x_1^3} \otimes 1 + 3h_{x_1^2} \otimes b_{x_1} + h_{x_1} \otimes b_{x_1^2},
\]
\[
n_b, n_a = 4, 5 : \Delta h_{x_0^2} = h_{x_0^2} \otimes 1 + h_{x_1} \otimes a_{x_0} + h_{x_0x_1} \otimes a_\theta + h_{x_1x_0} \otimes a_\theta + h_{x_1^2} \otimes (a_\theta)^2.
\]

The next lemma provides a grading for the coproduct \( \Delta \), which is clearly evident in the example above. Its particular form, namely, having \( V \) on the left, implies that \( H \) is a right-sided Hopf algebra [29, 31], a property that is essential in section 6.
Lemma 4.3. For any $h_\eta \in V_n$

\begin{equation}
\tilde{\Delta} h_\eta \in \bigoplus_{j+k=n} V_j \otimes H_k =: (V \otimes H)_n.
\end{equation}

Proof. The following facts are essential:
(1) $\deg(\theta_1 h) = \deg(h) + 1$,
(2) $\deg(\theta_0 h) = \deg(h) + 2$,
(3) $\Delta^b h_\eta \in (V \otimes V)_{\deg(h)+1}$.

The proof is via induction on the length of $\eta$. When $|\eta| = 0$ then clearly $\tilde{\Delta} h_\emptyset = h_\emptyset \otimes 1 \in V_n \otimes H_0$, where $n = \deg(h_\emptyset) \in \{0, 1\}$ (noting that $b_\emptyset \sim 1$). Assume now that (4.6) holds for words up to length $|\eta| \geq 0$. Let $n = \deg(h_\eta)$. There are two ways to increase the length of $\eta$. First consider $h x_1 \eta$. From item (1) above $\deg(h x_1 \eta) = n + 1$. Now apply item (3), the induction hypothesis, and Lemma 4.1 in that order:

\[ \Delta^b h_\eta \in (V \otimes V)_n, \]
\[ (\tilde{\Delta} \otimes \text{id}) \circ \Delta^b h_\eta \in (V \otimes H \otimes V)_n, \]
\[ (\theta_1 \otimes \mu) \circ (\tilde{\Delta} \otimes \text{id}) \circ \Delta^b h_\eta \in \bigoplus_{j+k=1}^{n} V_{j+1} \otimes H_k, \]
\[ \tilde{\Delta} h_{x_1 \eta} \in (V \otimes H)_{n+1}, \]

which proves the assertion. Consider next $h x_0 \eta$. From item (2) above $\deg(h x_0 \eta) = n + 2$. In this case, repeat the first two steps of the previous case and apply item (1) to get

\[ (\theta_1 \otimes \mu) \circ (\tilde{\Delta} \otimes \text{id}) \circ \Delta^b h_\eta \in \bigoplus_{j+k=1}^{n+1} V_{j+1} \otimes H_k \subset (V \otimes H)_{n+2}. \]

In addition, from the induction hypothesis and item (2) it follows that

\[ (\theta_0 \otimes \text{id}) \circ \tilde{\Delta} h_\eta \in \bigoplus_{j+k=1}^{n} V_{j+2} \otimes H_k \subset (V \otimes H)_{n+2}. \]

Thus, applying Lemma 4.1, $\tilde{\Delta} h_{x_0 \eta} \in (V \otimes H)_{n+2}$, which again proves the assertion and completes the proof. \qed

The following lemma is of central importance as it shows how to compute the Hopf algebra coproduct $\Delta$ in terms of $\Delta$ and $\Delta^\mu$.

Lemma 4.4. The following identities hold:
(1) $\Delta h_\eta = (\text{id} \otimes \mu) \circ (\tilde{\Delta} \otimes \text{id}) \circ \Delta^b h_\eta$,
(2) $\Delta a_\eta = (\text{id} \otimes \mu) \circ (\tilde{\Delta} \otimes \text{id}) \circ \Delta^a h_\eta + \tilde{\Delta} a_\eta$.

Proof.
(1) Observe

\[ \Delta b_\eta(c_\delta, d_\delta) = \sum_{\xi, \nu \in X^*} (c_L \tilde{\delta} d_\delta, \xi)(d_L, \nu)(\xi \bowtie \nu, \eta) = \sum_{\xi, \nu \in X^*} \tilde{\Delta} b_\xi(c_\delta, d_\delta) b_\nu(d_\delta)(\xi \bowtie \nu, \eta). \]
coproduct terms, namely, \( \Delta^9_0 \), given concerning the nature of the antipode. Observe that for any \( S \)-algebra \((A, \Delta^c)\) together with the coproduct \( \Delta \) is a connected graded bialgebra and thus one can naturally associate a Hopf algebra character \( \Phi \).

Before stating the actual theorem, some additional motivation is given concerning the nature of the antipode. Observe that for any \( c, d \in R_{\text{up}}(\langle X \rangle) \) one can naturally associate a Hopf algebra character \( \Phi_c \in L(H, \mathbb{R}) \) by letting \( \Phi_c : b \mapsto (c, b) \), and \( \Phi_c(1) = 1 \). Note that (4.2) implies \( \Phi_c(h_{\eta}h_{\xi}) = \Phi_c(h_{\eta}) \Phi_c(h_{\xi}) \) for any \( h_{\eta}, h_{\xi} \in \{a, b\} \). Since the coproduct \( \Delta \) of \( H \) has been constructed to satisfy (4.3), the following simple calculation shows that for all \( c, d \in R_{\text{up}}(\langle X \rangle) \), the convolution product of \( \Phi_c, \Phi_d \in L(H, \mathbb{R}) \) corresponds to composition in \( R_{\text{up}}(\langle X \rangle) \):

\[
(\Phi_c \cdot \Phi_d)(h_{\eta}) := m_R(\Phi_c \otimes \Phi_d) \Delta(h_{\eta}) = h_{\eta}(c \circ d) = (c \circ d, h_{\eta}).
\]

Example 4.5. Applying the identities in Lemma 4.4 gives the first few reduced coproduct terms, namely, \( \Delta^b h_{\eta} := \Delta h_{\eta} - h_{\eta} \otimes 1 - 1 \otimes h_{\eta} \):

\[
\begin{align*}
n &= 1 : \Delta^b a_0 = 0, \\
n &= 2 : \Delta^b a_1 = b_1 \otimes a_0, \\
n &= 3 : \Delta^b a_2 = b_0 \otimes a_0 + a_x a_0 + 2b_1 \otimes a_0 + 3b_2 \otimes a_0 + b_3 \otimes a_0, \\
n &= 4 : \Delta^b a_3 = b_0 \otimes a_0 + b_1 \otimes a_0 + 2b_2 \otimes a_0 + 3b_3 \otimes a_0 + 4b_4 \otimes a_0 + b_5 \otimes a_0 + b_6 \otimes a_0 + b_7 \otimes a_0 + b_8 \otimes a_0.
\end{align*}
\]

The main result of this section is given next, namely, that the commutative unital algebra \((H, \mu)\) together with the coproduct \( \Delta \) is a connected graded bialgebra and thus has an antipode \( S \). Before stating the actual theorem, some additional motivation is given concerning the nature of the antipode. Observe that for any \( c, d \in R_{\text{up}}(\langle X \rangle) \) one can naturally associate a Hopf algebra character \( \Phi_c \in L(H, \mathbb{R}) \) by letting \( \Phi_c : b \mapsto (c, b) \), and \( \Phi_c(1) = 1 \). Note that (4.2) implies \( \Phi_c(h_{\eta}h_{\xi}) = \Phi_c(h_{\eta}) \Phi_c(h_{\xi}) \) for any \( h_{\eta}, h_{\xi} \in \{a, b\} \). Since the coproduct \( \Delta \) of \( H \) has been constructed to satisfy (4.3), the following simple calculation shows that for all \( c, d \in R_{\text{up}}(\langle X \rangle) \), the convolution product of \( \Phi_c, \Phi_d \in L(H, \mathbb{R}) \) corresponds to composition in \( R_{\text{up}}(\langle X \rangle) \):

\[
(\Phi_c \cdot \Phi_d)(h_{\eta}) := m_R(\Phi_c \otimes \Phi_d) \Delta(h_{\eta}) = h_{\eta}(c \circ d) = (c \circ d, h_{\eta}).
\]
(See [17, Lemma 2] for a similar calculation.) In light of (2.8), it then follows that
\[ h \eta (c_s^{-1}) = (Sh \eta)(c_s) \ \forall \eta \in X^*, \ h \in \{ b, a \}. \]

Thus, the antipode provides an explicit way to compute the group inverse.

**Theorem 4.6.** \((H, \mu, \Delta)\) is a connected graded commutative unital Hopf algebra.

**Proof.** From the development above, it is clear that \((H, \mu, \Delta)\) is a connected bialgebra with unit \(1\) and a counit \(\varepsilon\) defined by \(\varepsilon(a_\eta) = 0\) for all \(\eta \in X^*\), \(\varepsilon(b_\eta) = 0\) for all nonempty \(\eta \in X^*\), and \(\varepsilon(1) = 1\). Here it is shown that this bialgebra is also graded and thus is automatically a Hopf algebra, i.e., has a well defined antipode, \(S\) [10]. It only remains to be shown for any \(n \geq 0\) that \(\Delta H_n \subseteq (H \otimes H)_n\). It is well known if \(h \in V_n\), then \(\Delta h, h \in (V \otimes V)_n\). Therefore, it follows directly from Lemmas 4.3 and 4.4 that \(\Delta h \in (V \otimes H)_n\). Thus, via the identity \(\Delta(a_\eta a_\xi^j) = \Delta a_\eta^i \Delta a_\xi^j\), it must hold that \(\Delta H_n \subseteq (H \otimes H)_n\), \(n \geq 0\). \(\square\)

The next theorem says that the antipode of any connected graded Hopf algebra can be computed in a recursive manner once the coproduct is computed.

**Theorem 4.7** (see [10]). The antipode, \(S\), of any connected graded Hopf algebra \((H, \mu, \Delta)\) can be computed for any \(a \in H_k\), \(k \geq 1\), by
\[ S a = -a - \sum (Sa'_{(1)} a'_{(2)}) = -a - \sum a'_{(1)} Sa'_{(2)}, \]

where the reduced coproduct is \(\Delta' a = \Delta a - a \otimes 1 - 1 \otimes a = \sum a'_{(1)} a'_{(2)}\).

As noted earlier, the coproducts \(\Delta_{\mu}^1\) and \(\tilde{\Delta}\) can be computed recursively, and \(\Delta\) is computed directly in terms of \(\Delta_{\mu}^1\) and \(\tilde{\Delta}\). So in fact the antipode of \(H\) can be computed in a fully recursive manner as described next.

**Theorem 4.8.** The antipode, \(S\), of any \(h \eta \in V_+\) can be computed by the following algorithm:

i. Recursively compute \(\tilde{\Delta}_{h,\mu}^1\) via (4.4).
ii. Recursively compute \(\tilde{\Delta}\) via Lemma 4.1.
iii. Compute \(\Delta\) via Lemma 4.4.
iv. Recursively compute \(S\) via Theorem 4.7.

**Example 4.9.** The first few antipode terms computed via Theorem 4.8 are

\[
\begin{align*}
n = 1: & \ Sb_{x_1} = -b_{x_1}, \\
n = 2: & \ Sb_{x_0} = -b_{x_0} + b_{x_1} a_\emptyset, \\
n = 2: & \ Sb_{x_1}^3 = -b_{x_2}^3 + 3(b_{x_1})^2, \\
n = 3: & \ Sb_{x_0 x_1} = -b_{x_0 x_1} + b_{x_1} a_\emptyset - 3(b_{x_1})^2 a_\emptyset + 2b_{x_0} b_{x_1} + b_{x_1} a_{x_1}, \\
n = 3: & \ Sb_{x_1 x_0} = -b_{x_1 x_0} + b_{x_2} a_\emptyset + 3b_{x_0} b_{x_1} - 3(b_{x_1})^2 a_\emptyset, \\
n = 3: & \ Sb_{x_1} = -b_{x_1}^3 + 10b_{x_1} b_{x_1} - 15(b_{x_1})^3, \\
n = 4: & \ Sb_{x_2} = -b_{x_0} + b_{x_0 x_2} a_\emptyset + b_{x_1 x_2} a_\emptyset - b_{x_2} (a_\emptyset)^2 + 3(b_{x_1})^2 (a_\emptyset)^2 + 2(b_{x_0})^2 \\
& - 5b_{x_0} b_{x_1} a_\emptyset + b_{x_2} a_\emptyset a_{x_1} + b_{x_1} a_{x_0}, \\
n = 1: & \ Sa_0 = -a_0, \\
n = 2: & \ Sa_{x_1} = -a_{x_1} + b_{x_1} a_\emptyset,
\end{align*}
\]
This information could be useful in determining the complexity of the antipode recursion. In this case, \( (3.8) \) reduces to

\[
- b_{x_2}^2 (a_{\emptyset})^2 + a_{\emptyset} a_{x_2}^2 + b_{x_2} a_{x_1} + b_{x_1} a_{x_0},
\]

\[
+ 3(b_{x_1})^2 (a_{\emptyset})^2 + 2b_{x_2} a_{x_1} + b_{x_1} a_{x_0},
\]

\[
n = 4 : S a_{x_1} = a_{x_0} - b_{x_0} a_{x_0} a_{x_0} - 2b_{x_0} b_{x_1} a_{\emptyset} - 4b_{x_1} a_{\emptyset} a_{x_1} + 3(b_{x_1})^2 (a_{\emptyset})^2
\]

\[
+ 6b_{x_1} a_{x_2}^2,
\]

\[
n = 5 : S a_{x_0} = a_{x_0} + b_{x_2} a_{\emptyset} - 2(b_{x_0})^2 a_{\emptyset} a_{x_0} a_{x_0} + a_{\emptyset} a_{x_0} a_{x_0} - 3b_{x_1} a_{x_0} a_{\emptyset}
\]

\[
- 3b_{x_0} a_{x_1} a_{\emptyset} - b_{x_0} a_{x_1} (a_{\emptyset})^2 + b_{x_0} a_{x_1} (a_{\emptyset})^2 + 5b_{x_2} b_{x_1} (a_{\emptyset})^2 - (a_{\emptyset})^2 a_{x_2}^2
\]

\[
+ 4b_{x_1} (a_{\emptyset})^2 a_{x_1} + b_{x_2} (a_{\emptyset})^2 - 3(b_{x_1})^2 (a_{\emptyset})^2 + 2b_{x_0} a_{x_0} + a_{x_0} a_{x_1}.
\]

**Example 4.10.** Reconsider the multiplicative feedback system described in Example 3.12. Suppose \( c = \sum_{k \geq 0} k! x_1^k = \sum_{k \geq 0} x_1^k \) and \( F_{d_5} = I \) (unity feedback). In this case, \( (3.8) \) reduces to \( c @ d_5 = c \circ (d_L \circ c @ d_5, 0) \) so that

\[
c @ d_5 = c \circ (c_{\emptyset}^{-1}, 0) = (c_{\emptyset}^{-1}, 0) = (1 - x_1, 0)^{-1}.
\]

The composition inverse can be computed directly from (4.7) and the antipode formulas above. Namely, taking into account only the antipode terms containing powers of the coordinate function \( b_{x_1} \) (all others will yield zero), it follows that

\[
c @ d_5 = (1 - x_1, 0)^{-1} = \left( \frac{1 + \sum_{\eta \in X^+} \eta S b_{x_1} \eta (1 - x_1), \sum_{\eta \in X^*} \eta S a_{x_1} (0)}{1 + x_1 S b_{x_1} + x_1^2 S b_{x_2} + x_1^3 S b_{x_3} + \ldots} \right) (1 - x_1, 0)
\]

\[
= (1 + x_1 (-b_{x_1} (1 - x_1)) + x_1^2 \left( -b_{x_2} + 3(b_{x_1})^2 \right) (1 - x_1)
\]

\[
+ x_1^3 \left( b_{x_3} + 3b_{x_2} b_{x_1} - 15(b_{x_1})^3 \right) (1 - x_1), 0)
\]

\[
= (1 + x_1 + 3x_2 + 15x_3 + \ldots, 0).
\]

The deferred proof from section 3 is presented next.

**Proof of Lemma 3.3.3(3).** The assertion is that \( c \circ d_5 = k \) implies \( c = k \). The proof is by induction on the grading of \( H \). If \( c \circ d_5 = k \), then clearly \( k = a_{\emptyset} (c_3 \circ d_5) = \Delta a_{\emptyset} (c_3, d_5) = a_{c_3} (c_\emptyset) \) assuming without loss of generality that \( c_\emptyset = (1, c) \). Therefore, \( c = k \).

Similarly, it follows that \( 0 = a_{x_1} (c_3 \circ d_5) = \Delta a_{x_1} (c_3, d_5) = a_{x_1} (c_\emptyset) \). Thus, \( (c, x_1) = 0 \). Now suppose \( a_{x_\eta} (c_\emptyset) = 0 \) for all \( 0 \neq \eta \in H \). Then for any \( x \neq X \)

\[
0 = \Delta a_{x_\eta} (c_\emptyset, d_5) = a_{x_\eta} (c_\emptyset) + \sum_{a_{x_\eta (2)} \neq 1} a_{x_\eta (1)} (c_\emptyset) a_{x_\eta (2)} (d_5),
\]

where in general \( a_{x, \eta (1)} (c_\emptyset) \neq a_{0} \). Therefore, \( a_{x_\eta} (c_\emptyset) = 0 \), or equivalently, \( (c, x_\eta) = 0 \). Thus, \( c = k \).
the scope of this paper. Let $V_{h,k}$ denote the subspace of $V_k$ spanned by the coordinate functions $h_\eta$ of degree $k$, where $h \in \{a,b\}$. Define $p_{h,k} = \dim(V_{h,k})$, $p_k = \dim(V_k)$, and the corresponding generating functions $F_{V_k} = \sum_{k\geq 1} p_{h,k} X^k$, $F_V = \sum_{k\geq 1} p_k X^k$. Analogous definitions apply when $V$ is replaced by $H$.

**Theorem 4.11.** The following identities hold:

\[
F_{V_a} = \frac{X}{1 - X - X^2} = X + X^2 + 2X^3 + 3X^4 + 5X^5 + 8X^6 + 13X^7 + 21X^8 + \ldots,
\]

\[
F_{V_b} = \frac{X + X^2}{1 - X - X^2} = X + 2X^2 + 3X^3 + 5X^4 + 8X^5 + 13X^6 + 21X^7 + 34X^8 + \ldots,
\]

\[
F_V = F_{V_a} + F_{V_b} = \frac{2X + X^2}{1 - X - X^2} = 2X + 3X^2 + 5X^3 + 8X^4 + 13X^5 + 21X^6 + \ldots,
\]

\[
F_{H_a} = \prod_{k=1}^{\infty} \frac{1}{(1 - X)^{p_{a,k}}} = 1 + X + 2X^2 + 4X^3 + 8X^4 + 15X^5 + 30X^6 + 56X^7 + \ldots,
\]

\[
F_{H_b} = \prod_{k=1}^{\infty} \frac{1}{(1 - X)^{p_{b,k}}} = 1 + X + 3X^2 + 6X^3 + 14X^4 + 28X^5 + 61X^6 + 122X^7 + \ldots,
\]

\[
F_H = \prod_{k=1}^{\infty} \frac{1}{(1 - X)^{p_k}} = F_{H_a} F_{H_b} = 1 + 2X + 6X^2 + 15X^3 + 38X^4 + 89X^5 + \ldots.
\]

**Proof.** The identity for $F_{V_a}$ is proved in [13, Proposition 8]; the proof for $F_{V_b}$ is perfectly analogous. The identity for $F_V$ follows directly from the fact that $V = V_a \oplus V_b$. It is worth noting that the coefficients of all three series come from the Fibonacci sequence. The identity for $F_{H_a}$ was also proved in [13], and again the proof for $F_{H_b}$ is very similar. The factorization of $F_H$ is a consequence of the fact that $p_k = p_{a,k} + p_{b,k}$. In this case, the coefficients of $F_{H_a}$ and $F_{H_b}$ are integer sequences A168861 and A200544, respectively, in [35], while the sequence for $F_H$ appears to be new.

5. **System inversion and feedback invariants.** In this section two system theoretic problems connected with the output affine feedback transformation group are addressed. First it is shown how to pose the dynamic inversion in terms of this transformation group and its inverse. This naturally leads to the topic of feedback input-output linearization via nonlinear feedback. This then takes the presentation to the second problem of feedback invariants under the output affine feedback transformation group.

5.1. **System inversion.** A problem often encountered in control applications is the following system inverse problem. An input $u$ is mapped to an output $y$ by an operator $F$ defined in terms of a dynamical system

\[
\begin{aligned}
(5.1a) \quad \dot{z} &= g_0(z) + g_1(z)u, \quad z(t_0) = z_0, \\
(5.1b) \quad y &= h_L(z)u + h_R(z),
\end{aligned}
\]

where all the functions of the state $z$ are assumed to be analytic on some neighborhood $W$ of $z_0$. If $h_L(z_0) \neq 0$, then $F$ is locally invertible. The left inverse, $F^{-1} : y \mapsto u$, can be determined dynamically by first solving the output equation for $u$, $u = h^{-1}_L(z)[y - h_R(z)]$, and then substituting the result into the state equation so that the system

\[
\begin{aligned}
\dot{z} &= [g_0(z) - g_1(z)h^{-1}_L(z)h_R(z)] + g_1(z)h^{-1}_L(z)y, \quad z(t_0) = z_0, \\
u &= h^{-1}_L(z)y - h^{-1}_L(z)h_R(z)
\end{aligned}
\]
has the property that if $y$ is a given function in the range of $F$, then $y = F[u]$ maps to $u$ over an interval $[t_0, t_0 + T]$ with $T > 0$. In short, input-output inversion is done by mapping the given system $(g_0, g_1, h_L, h_R, z_0)$ to its inverse system

\[(g_0, g_1, h_L, h_R, z_0) := (g_0 - g_1 h_L^{-1} h_R, g_1 h_L^{-1}, h_R, -h_L^{-1} h_R, z_0). \tag{5.2}\]

It is not difficult to see that the mapping $F : u \mapsto y$ can be described independent of the state $z$ and its dynamics by an element $c_\delta = (c_L, c_R)$ of the output affine feedback transformation group so that $y = F[u] = F_{c_L}[u] + F_{c_R}[u] =: F_{c_\delta}[u]$, where $c_L, c_R \in \mathbb{R}[X]$ are each computed from (2.4) using the respective outputs $h_L$ and $h_R$. The condition that $h_L(z_0) \neq 0$ ensures that $c_L$ is nonproper. Therefore, the inverse input-output map is given by $F_{c_\delta}^{-1}$, where the generating series $c_\delta^{-1}$ can be computed using the antipode of the underlying Hopf algebra described in Section 4.

In fact, the following example demonstrates that this relation goes in both directions when $c_\delta = (c_L, c_R)$ has a realization of the form (5.1).

**Example 5.1.** Reconsider Example 4.10, where it was necessary to compute $(1 - x_1, 0)^{-1}$. It is straightforward to show that $y_L = F_{c_L}[u] = F_{1 - x_1}[u], y_R = 0$ is realized by $(g_0, g_1, h_L, h_R, z_0) = (0, -1, z, 0, 1)$ so that by (5.2) it follows that

\[c_\delta^{-1} = 1 + x_1 L_g \bar{h}_L(z_0) + x_1^2 L_{g_1}^2 \bar{h}_L(z_0) + x_1^3 L_{g_1}^3 \bar{h}_L(z_0) + \cdots \]

\[= 1 + x_1 + 3x_1^2 + 15x_1^3 + \ldots, \]

as computed earlier.

Given the connection between dynamic inversion and feedback linearization [24, 32], it is natural to consider the latter in terms of the output affine feedback transformation group. Consider a Fliess operator $y = F_c[u]$ which has an n-dimensional analytic state space realization of the form

\[\dot{z} = g_0(z) + g_1(z) u, \quad z(0) = z_0, \quad y = h(z), \tag{5.3}\]

with relative degree $r$. Then, it follows that $y^{(r)} = L_{g_0}^r h(z) + L_{g_1} L_{g_0}^{r-1} h(z) u$ with $L_{g_1} L_{g_0}^{r-1} h(z)$ being nonzero on some neighborhood of $z_0$ [24, 32]. Since the state can be viewed in terms of a Fliess operator for some generating series $c_\delta \in \mathbb{R}[X]$, it follows from Fliess’s fundamental formula that there are generating series $c_\delta \in \mathbb{R}[X]$ such that $F_{c_\delta}[u] = L_{g_0}^r h(F_{c_\delta}[u])$ and $F_{c_\delta}[u] = L_{g_1} L_{g_0}^{r-1} h(F_{c_\delta}[u])$, respectively, so that $y^{(r)} = F_{c_\delta}[u] + F_{c_\delta}[u]$. Since $(c_\delta, 0) = L_{g_0} L_{g_0}^{r-1} h(F_{c_\delta}[u])(0) \neq 0$, it follows that $c_\delta = (c_L, c_R) \in \mathbb{R}[X]$ and therefore, if $v := y^{(r)}$, then the feedback linearization law can be written in the form $u = F_{c_\delta}^{-1}[v]$. When this feedback is applied to the plant, $F_c$, the feedback linearized system is described by

\[y = F_c[u] = F_c[F_{c_\delta}^{-1}[v]] = F_{c_\delta}^{-1} c_\delta^{-1}[v] = u^{(-r)}. \]

So it is evident that the output affine feedback transformation group is at play in this problem. The proposition, however, is that this analysis holds even when $F_c$ does not have a state space realization.

The starting point for a realization free framework for feedback input-output linearization is the following definition which describes relative degree from a generating series point of view. It uses the notion of a linear word, that is, any word in the language $\mathcal{L} = \{ n \in X^+ : n = x_0^n x_1 x_0^n, n_1, n_0 \geq 0 \}$. Furthermore, note that every $c \in \mathbb{R}[X]$ can be decomposed into its natural and forced components, that is,
$c = c_N + c_F$, where $c_F := \sum_{k \geq 0} (c, x^k_0)x^k_0$ and $c_F := c - c_N$. Finally, for any letter $x_i \in X$, the left-shift operator $x_i^{-1}(-)$ introduced earlier is defined inductively for higher order shifts via $(x_i\xi)^{-1}(-) = \xi^{-1}x_i^{-1}(-)$, where $\xi \in X^*$.

**Definition 5.2** (see [18]). Given $c \in \mathbb{R}(\mathcal{L}(X))$, let $r \geq 1$ be the largest integer such that $\text{supp}(c_F) \subseteq x_0^{-r}X^*$. Then $c$ has relative degree $r$ if the linear word $x_0^{-r}x_1 \in \text{supp}(c)$, otherwise it is not well defined.

Observe that $c$ having relative degree $r$ is equivalent to saying that

$$c = c_N + c_F = c_N + Kx_0^{-r}x_1 + x_0^{-r}e$$

for some $K \neq 0$ and some proper $e \in \mathbb{R}(\mathcal{L}(X))$ with $x_1 \not\in \text{supp}(e)$. Furthermore, this notion coincides with the classical definition when $y = F_c[u]$ has an analytic state space realization of the form given in (5.3). Specifically, in light of the identity

$$\dot{y} = F_{x_0^{-1}x_1}[u] + uF_{x_0^{-1}x_1}[u]$$

and (5.4), it follows that if $c$ has relative degree $r$, then on some interval $[0, T)$ with $T > 0$ and for any sufficiently small $u \in L_1[0, T]$, $u$ has relative degree $r$.

$$\dot{y}(k) = F_{(x_0^k)}^{-1}(-)[u] \quad k = 0, 1, \ldots, r-1,$$

(5.5a)

$$\dot{y}(r) = F_{(x_0^r)}^{-1}(-)[u] + uF_{(x_0^r}^{-1}x_1)[-1(c)].u$$

(5.5b)

This would imply that $L_{g_1}L_{g_0}h(z(t)) = F_{(x_0^r)\bar{x}_{z_1}^{-1}(-)}(u) = 0$ on $[0, T)$ for $k = 0, 1, \ldots, r-2$. In addition, $L_{g_1}L_{g_0}^{-1}(z_0) = F_{(x_0^r)\bar{x}_{z_1}^{-1}(-)}(u)(0) = (c, x_0^{-r}x_1) = K \neq 0$. Hence, the realization has relative degree $r$ at $z_0$ in the classical sense. Furthermore, if the word $x_0^{-r}x_1 \not\in \text{supp}(c)$, then $F_{(x_0^r)\bar{x}_{z_1}^{-1}(-)}(u)(0) = 0$, but $F_{(x_0^r)\bar{x}_{z_1}^{-1}(-)}(u)(0)$ is not identically zero on some interval $[0, T_1)$ with $T_1 > 0$ unless $(x_0^{-r}x_1)^{-1}(c) = x_1^{-1}(c) = 0$ (since generating series are unique). But this can happen only if $\text{supp}(c) \subseteq x_0X^*$, which contradicts the assumption that $r$ is the largest integer such that $\text{supp}(c_F) \subseteq x_0^{-r}X^*$. Therefore, the realization cannot have a well defined relative degree in this case. Conversely, if the realization (5.3) has relative degree $r$ at $z_0$, then for $k = 0, 1, \ldots, r-2$ it follows that $L_{g_1}L_{g_0}h(z(t)) = F_{(x_0^k)}^{-1}(-)[-1(c)](u) = 0$ on a neighborhood of $z_0$. So necessarily $(x_0^x)\bar{x}_{z_1}^{-1}(-)(c) = 0$ for $k = 0, 1, \ldots, r-2$, or equivalently, no word in the support of $c_F$ can have a prefix of the form $x_0^r x_1$ with $k < r-1$. Therefore, $\text{supp}(c_F) \subseteq x_0^{-r}X^*$. In addition, since $L_{g_1}L_{g_0}^{-1}(z_0) = F_{(x_0^r)\bar{x}_{z_1}^{-1}(-)}(u)(0) = (c, x_0^{-r}x_1) \neq 0$, then $x_0^{-r}x_1 \in \text{supp}(c_F)$, and $x_0^{-r}$ is the longest such prefix that all the words in $\text{supp}(c_F)$ can share. So $c$ must have the form of (5.4), implying that $c$ has relative degree $r$.

**Example 5.3.** Consider the state space realization (5.3), where

$$g_0(z) = 
\begin{bmatrix}
  z_2^3 + 2z_3z + 3 \\
  z_1^5 + 3z_3 \\
  z_1^2
\end{bmatrix},
\quad g_1(z) = 
\begin{bmatrix}
  0 \\
  0 \\
  1
\end{bmatrix},
\quad h(z) = z_1.$$

Observe that $L_{g_1}h(z) = 0$ and $L_{g_1}L_{g_0}h(z) = z_2 + 1$, so that the realization has relative degree 2 at $z_0 = [z_{01} z_{02} z_{03}]^T$ if $z_{02} \neq -1$ and is undefined otherwise. So, for example, the generating series for $y = F_c[u]$ when $z = [0 1 1]^T$ is

$$c = 3x_0 + 3x_0^3 + 2x_0x_1 + 2x_0^3 + 4x_0^2x_1 + x_0x_1x_0 + 36x_0^2 + 4x_0^3x_1 + 2x_0^2x_1x_0 + \ldots,$$

which has relative degree 2 since $\text{supp}(c_F) \subseteq x_0X^*$, $\text{supp}(c_F) \subseteq x_0^{-r}X^*$ for $r > 2$, and $x_0x_1 \in \text{supp}(c)$. But when $z = [0 -1 1]^T$, the generating series becomes

$$c = x_0 - x_0^2 + 2x_0^3 + x_0x_1x_0 + 4x_0^3x_1 + 2x_0^2x_1x_0 + 2x_0^3x_1^2 + x_0x_1x_0x_1 + 6x_0^5 + 4x_0^3x_1^2 + \ldots,$$
which fails to have a well defined relative degree since \( \text{supp}(c_F) \subset x_0X^* \) and \( \text{supp}(c_F) \not\subset x_0^{-1}X^* \) for \( r > 2 \), but \( x_0x_1 \not\subset \text{supp}(c) \).

The following theorem describes feedback input-output linearization without any requirement that the plant has a state space realization.

**Theorem 5.4.** Suppose \( c \in \mathbb{R}_{LC}\langle \langle X \rangle \rangle \) has relative degree \( r \). Then \( y = F_{e\delta}\langle [u] \rangle \) is feedback input-output linearized by \( u = F_{e\delta}^{-1}\langle v \rangle \), where \( e\delta = \langle (x_0^{-1}x_1)^{-1}(c), (x_0^r)^{-1}(c) \rangle \) and provided that \( e\delta^{-1} \in \mathbb{R}_{LC}\langle \langle X \rangle \rangle \), giving the closed-loop system \( y = v^{(r)} \).

**Proof.** It follows from (5.5b) that \( y^{(r)} = F_{e\delta}\langle [u] \rangle \). Therefore, setting \( u = F_{e\delta}^{-1}\langle [v] \rangle \) gives \( y^{(r)} = F_{e\delta} \circ F_{e\delta}^{-1}\langle [v] \rangle = F_{e\delta}\circ e\delta^{-1}\langle [v] \rangle = F_{(1,0)}[v] = v \), which yields the desired result.

**Example 5.5.** Consider the polynomial \( c = x_1 + x_1^2 \), which has relative degree 1. Setting \( e\delta = (x_1^{-1}(c), x_0^{-1}(c)) = (1 + x_1, 0) \), the inverse of \( e\delta \) can be computed by the same computational approach as in Example 4.10:

\[
(1 + x_1, 0)^{-1} = \left( [1 + x_1 Sb_{x_1} + x_1^2 Sb_{x_1^2} + x_1^3 Sb_{x_1^3} + \cdots ](1 + x_1), 0 \right) \\
= (1 - x_1 + 3x_1^2 - 15x_1^3 + \ldots, 0).
\]

Therefore, a direct calculation gives \( c \circ (1 + x_1, 0)^{-1} = x_1 \) as desired. This can also be verified by observing that \( c = x_1 \circ (1 + x_1, 0) = \phi_\delta(1)(1) + \phi_\delta(x_1^2)(1) = x_1 + x_1^2 \). In addition, \( c \circ e\delta^{-1} \) clearly has the same relative degree as \( c \), which in the next section is shown to hold in general.

### 5.2. Feedback invariants

It is shown next that relative degree as defined in Definition 5.2 is invariant under the action of the output affine feedback transformation group. In addition, this action is free when restricted to all series having well defined relative degree.

**Theorem 5.6.** A series \( c \) has relative degree \( r \) if and only if it is on the orbit of \( c_N + x_0^{-1}x_1 \) under \( \mathbb{R}_{np}\langle \langle X \rangle \rangle \).

**Proof.** If \( c \) has well defined relative degree \( r \), then it can be decomposed as in (5.4), where without loss of generality \( c = x_0e_0 + x_1e_1 \) with \( e_1 \) proper. Then, setting \( e\delta := (K + e_1, e_0) \in \mathbb{R}_{np}\langle \langle X \rangle \rangle \) (since \( K + e_1 \) is nonproper), it follows from (3.2) that

\[
c = cN + x_0^{-1}x_1(K + e_1) + x_0^{-1}e_0 = cN + \phi_\delta(x_0^{-1}x_1)(1) = (cN + x_0^{-1}x_1) \circ e\delta.
\]

Therefore, \( c \circ e\delta^{-1} = cN + x_0^{-1}x_1 \), or equivalently, \( c \) is on the orbit of \( cN + x_0^{-1}x_1 \) under \( \mathbb{R}_{np}\langle \langle X \rangle \rangle \). The converse holds since all the steps above are reversible.

When the series \( e\delta \) and \( e\delta^{-1} \) above are both locally convergent then it follows directly that if \( u = F_{e\delta}^{-1}\langle [v] \rangle \), then \( y = F_{e\delta}\langle [u] \rangle = F_{e\delta} \circ e\delta^{-1}\langle [v] \rangle = F_{cN + x_0^{-1}x_1} \circ [v] \), as expected.

**Theorem 5.7.** The transformation group \( \mathbb{R}_{np}\langle \langle X \rangle \rangle \) acts freely on the subset of \( \mathbb{R}\langle \langle X \rangle \rangle \) having well defined relative degree.

**Proof.** Assume \( c \) has relative degree \( r \). Without loss of generality let \( c_N = 0 \). Then there exists an \( e\delta \in \mathbb{R}_{np}\langle \langle X \rangle \rangle \) such that \( c \circ e\delta^{-1} = x_0^{-1}x_1 \). So if \( c \circ d\delta = c \) for some \( d\delta \in \mathbb{R}_{np}\langle \langle X \rangle \rangle \), then it follows immediately that \( (c \circ d\delta) \circ e\delta^{-1} = c \circ e\delta^{-1} \) and \( (c \circ e\delta^{-1}) \circ d\delta^{-1} = c \circ e\delta^{-1} \), where \( d\delta \) corresponds to the conjugate action \( e\delta \circ d\delta \circ e\delta^{-1} \). Hence, \( x_0^{-1}x_1 \circ d\delta^{-1} = x_0^{-1}x_1 \), or equivalently, \( x_0^{-1}x_1 d\delta^L = x_0^{-1}x_1 \), and therefore, \( d\delta^L := (d\delta^L, d\delta^R) = (1, 0) \), the identity element of \( \mathbb{R}_{np}\langle \langle X \rangle \rangle \). Thus, \( e\delta \circ d\delta \circ e\delta^{-1} = (1, 0) \), which gives the desired conclusion that \( d\delta = (1, 0) \).
6. The Lie algebra $\text{Lie}(\mathbb{R}_n \langle X \rangle)$. The main goal in this section is to describe the Lie algebra $\text{Lie}(\mathbb{R}_n \langle X \rangle)$ of infinitesimal characters associated with the group $\mathbb{R}_n \langle X \rangle$ (see Proposition 2.6) in terms of a pre-Lie product. The latter is considered the most elementary combinatorial building block since the group can be fully reconstructed solely in terms of its pre-Lie product [14, 33]. The reader is referred to [6, 30] for details on pre-Lie algebras. There are at least two ways to approach this problem. One way is to view $\mathbb{R}_n \langle X \rangle$ as an affine group scheme, which in turn always corresponds to a Hopf algebra [40, p. 9], in this case the Hopf algebra $H$ as described in Theorem 4.6. In this setting, an $\mathbb{R}$-derivation $D$ is any linear mapping taking $H$ into an $H$-module $M$ satisfying $D(h_1 h_2) = D(h_1) h_2 + h_1 D(h_2)$ for all $h_1, h_2 \in H$ with $D(k) = 0$ when $k \in \mathbb{R}$. It is described in terms of the universal derivation on $H$ $d = (\text{id} \otimes \pi) \circ \Delta$, where $\pi : H \to H_+$ with $H_+ = \bigoplus_{n>0} H_n$ being the augmentation ideal $\varepsilon : H \to \mathbb{R}$ of a Hopf algebra. A linear operator $T : H \to H$ is said to be left-invariant when $\Delta \circ T = (\text{id} \otimes T) \circ \Delta$. The Lie algebra of the group represented by $H$ is the $\mathbb{R}$-vector space of all left-invariant derivations on $H$. This is the approach taken in [13] for the case of output feedback. The Lie algebra $\text{Lie}(\mathbb{R}_n \langle X \rangle)$ is then determined by dualizing this approach using the fact that $\alpha_n(\nu) = \mathbb{I}_{\nu^*}$ for all $\alpha_n \in H$ and $\nu \in X^*$. A second approach is to treat $\mathbb{R}_n \langle X \rangle$ as though it were a Lie group and to show that the Lie algebra associated with the set of all left-invariant vector fields at the identity element can be described in terms of an $X^*$-derivation of the pre-Lie product. Of course, this in no way implies that $\mathbb{R}_n \langle X \rangle$ is a Lie group, but it provides a complementary point of view to the purely algebraic first approach. It is convenient in this context to write $c_\delta = (c_L, c_R) \in \mathbb{R}_n \langle X \rangle$ as $c_\delta = \delta c_L + c_R$, so that the symbol $\delta$ is treated more like a letter in $X$. The first task is to describe a left-invariant vector field on $\mathbb{R}_n \langle X \rangle$. The left translation of $d_\delta$ by $c_\delta$ is $c_\delta \cdot d_\delta = \delta([c_L \circ d_\delta] \omega d_L) + ([c_L \circ d_\delta] \omega d_R + c_R \circ d_\delta]$, which is denoted by $(c_\delta \circ) : \mathbb{R}_n \langle X \rangle \to \mathbb{R}_n \langle X \rangle : d_\delta \to c_\delta \cdot d_\delta$. Since composition is left linear, there is no loss of generality in setting $c_\delta = \xi_\delta := \delta \xi_L + \xi_R$, where $\xi_L, \xi_R \in X^*$. The differential of $(\xi_\delta \circ) : \mathbb{R}_n \langle X \rangle \to \mathbb{R}_n \langle X \rangle$ at the identity element $\delta$ is the linear map $(\xi_\delta \circ)_* : T_{\xi_\delta} \mathbb{R}_n \langle X \rangle \to T_{\xi_\delta} \mathbb{R}_n \langle X \rangle$. Consider for some $\epsilon > 0$ a differentiable path $\gamma : (-\epsilon, \epsilon) \to \mathbb{R}_n \langle X \rangle : t \mapsto v_\delta(t)$ such that $v_\delta(0) = 0$. Define the velocity vector at $t = 0$ as the series in $\mathbb{R}(\langle X \rangle)$ of the form $v_\delta = v_\delta(0) = \delta d_L(0) + d_R(0) = \delta v_L + v_R$, where $v_L \in \mathbb{R}(\langle X \rangle)$ is proper. Then specifically the differential of $\xi_\delta \circ$ at $\delta$ in the direction of $v_\delta$ is

$$
(\xi_\delta \circ)_* (v_\delta) = \left. \frac{d}{dt} \xi_\delta \circ d_\delta(t) \right|_{t=0} = \left. \frac{d}{dt} \delta([\xi_L \circ d_\delta(t)] \omega d_L(t) + (\xi_L \circ d_\delta(t)) \omega d_R(t) + \xi_R \circ d_\delta(t)) \right|_{t=0} = \delta \left[ \left. \frac{d}{dt} \xi_L \circ d_\delta(t) \right|_{t=0} + \xi_L \omega v_L + \xi_L \omega v_R + \left. \frac{d}{dt} \xi_R \circ d_\delta(t) \right|_{t=0} \right].
$$

The time derivative of the product $\xi \circ d_\delta$ is computed inductively. It is clearly zero when $\xi = \emptyset$. Otherwise, using Lemma 3.3(4),

$$
\left. \frac{d}{dt} (x_0 \xi \circ d_\delta(t)) \right|_{t=0} = x_0 \left. \frac{d}{dt} \xi \circ d_\delta(t) \right|_{t=0}
$$

and

$$
\left. \frac{d}{dt} (x_1 \xi \circ d_\delta(t)) \right|_{t=0} = x_1 \left. \frac{d}{dt} (d_L(t) \omega (\xi \circ d_\delta(t))) + x_0 (d_R(t) \omega (\xi \circ d_\delta(t))) \right|_{t=0}
$$

are computed inductively.
Therefore, \( \frac{d}{dt} \xi \partial \delta(t) \bigg|_{t=0} = \xi \bullet v_\delta \), where \( \emptyset \bullet v_\delta = 0 \) and

\[
(x_0 \xi) \bullet v_\delta = x_0 (\xi \bullet v_\delta), \quad (x_1 \xi) \bullet v_\delta = x_1 (v_L \omega \xi + \xi \bullet v_\delta) + x_0 (v_R \omega \xi).
\]

So the differential in question is

\[
(\xi_\delta \circ_\bullet v_\delta) = \delta[\xi_L \bullet v_\delta + \xi_L \omega v_L] + \xi_L \omega v_R + \xi_R \bullet v_\delta = (\delta \xi_L) \bullet v_\delta + \xi_R \bullet v_\delta = \xi_\delta \bullet v_\delta,
\]

where the definition in (6.1) is extended to treat the letter \( \delta \) as \( (\delta \xi) \bullet v_\delta = \delta(v_L \omega \xi + \xi \bullet v_\delta) + (v_R \omega \xi) \).

Thus, the left-invariant vector field satisfying \( \chi^v \circ_\bullet v_\delta = v_\delta \) is

\[
\chi^v : \mathbb{R}_{np}\langle\langle X_\delta \rangle\rangle \to T\mathbb{R}_{np}\langle\langle X_\delta \rangle\rangle : c_\delta \mapsto c_\delta \bullet v_\delta.
\]

Now \( H \) is free as a commutative algebra, that is, \( H = \bigwedge(V) \), where \( \bigwedge(V) \) is the symmetric algebra over the vector space \( V \). Also the mixed composition product is left linear, and in light of (4.6) the corresponding coproduct satisfies \( \Delta V \subseteq V \otimes H \).

Hence, \( H \) is a free commutative, right-sided Hopf algebra. Therefore, the graded dual of \( V \), namely, the vector space \( V^\ast \) formed from the \( \mathbb{R} \) span over words in \( X^\ast \) (the duals of the \( \eta, \xi \)'s) and nonempty words with prefix \( \delta \), denoted by \( \delta X^+ \) (the duals of \( b_\eta, b_\xi \)'s; recall \( b_\eta \sim 1 \)), must have a right pre-Lie product \([6, 29, 30, 31]\).

**Lemma 6.1.** The vector space \( V^\ast \) with product \( \bullet \) is a right pre-Lie algebra, i.e.,

\[
(v_0^1 \bullet v_0^2) \bullet v_0^3 - v_0^1 \bullet (v_0^2 \bullet v_0^3) = (v_0^1 \bullet v_0^3) \bullet v_0^2 - v_0^1 \bullet (v_0^2 \bullet v_0^3).
\]

**Proof.** The identity can be verified directly using the distributive property \((\eta \omega \xi) \bullet v_\delta = (\eta \bullet v_\delta) \omega \xi + \eta \omega (\xi \bullet v_\delta)\), which can be proved by induction on the sum of the lengths of \( \eta, \xi \in X^\ast \).

**Example 6.2.** Consider (6.3), where \( v_0^1 = \delta x_1, v_0^2 = x_1, \) and \( v_0^3 = \delta x_0 \). Then \( \delta x_1 \bullet x_1 = \delta x_0 x_1 + 2x_1^2, x_1 \bullet \delta x_0 = x_1 x_0, \delta x_1 \bullet \delta x_0 = \delta(2x_1 x_0 + x_0 x_1), \delta x_0 \bullet x_1 = x_0 x_1 + x_1 x_0, \) and both sides of (6.3) equal \( \delta(2x_0^2 x_1 + x_0 x_1 x_0) + 2x_1^2 x_0 + x_1 x_0 x_1) \).

Finally, a pre-Lie product is Lie admissible, that is, it always defines a proper Lie bracket \([,]_\bullet\) under antisymmetrization. In the present case, this is equivalent to the Lie algebra induced by (6.2) since

\[
[v_0^1, v_0^2] = \left[ \chi^v \right] \bigg|_{\delta} = \partial \chi^v(c_\delta \bullet v_0^3) - \partial \chi^v(c_\delta \bullet v_0^1) \bigg|_{c_\delta = \delta} = v_0^2 \bullet v_0^1 - v_0^1 \bullet v_0^2.
\]

The next theorem follows directly from what is known for free commutative right- (or left-) sided polynomial Hopf algebras (see [29, Theorem 5.8], [31, Proposition 5], and [33, section 3]).

**Theorem 6.3.** The Lie algebra \( \text{Lie}(\mathbb{R}_{np}\langle\langle X_\delta \rangle\rangle) \) of infinitesimal characters associated with the group \( \mathbb{R}_{np}\langle\langle X_\delta \rangle\rangle \) is isomorphic as a Lie algebra to \((V^\ast, [,]_\bullet)\).

**Example 6.4.** In the special case where \( c_\delta = \delta + c_R \) and \( d_\delta = \delta + d_R \), the corresponding subspace of \( T_0 \mathbb{R}_{np}\langle\langle X_\delta \rangle\rangle \) is spanned by vectors of the form \( v_\delta = \delta 0 + v_R \).

Thus, the pre-Lie product and Lie bracket above reduce to those described in [13] and [20], respectively.
REFERENCES