The Hopf algebra of rooted trees, free Lie algebras, and Lie series

A. Murua*

October 14, 2013

Abstract

We present an approach that allows performing computations related to the Baker-Campbell-Haussdorff (BCH) formula and its generalizations in an arbitrary Hall basis, using labeled rooted trees. In particular, we provide explicit formulas (given in terms of the structure of certain labeled rooted trees) of the continuous BCH formula. We develop a rewriting algorithm (based on labeled rooted trees) in the dual Poincaré-Birkhoff-Witt (PBW) basis associated to an arbitrary Hall set, that allows handling Lie series, exponentials of Lie series, and related series written in the PBW basis. At the end of the paper we show that our approach is actually based on an explicit description of an epimorphism ν of Hopf algebras from the commutative Hopf algebra of labeled rooted trees to the shuffle Hopf algebra and its kernel $\ker \nu$.

1 Introduction, general setting, and examples

Consider a d-dimensional system of non-autonomous ODEs of the form

$$\frac{d}{d\tau} y = \lambda_1(\tau) f_1(y) + \lambda_2(\tau) f_2(y), \tag{1}$$

with smooth maps $f_1, f_2 : \mathbb{R}^d \to \mathbb{R}^d$, and integrable functions $\lambda_1, \lambda_2 : \mathbb{R}^+ \to \mathbb{R}$. Let E_1 and E_2 be the vector fields (or Lie operators) associated to f_1 and f_2 respectively, that is $E_a = \sum_{i=1}^d f_a^i \frac{\partial}{\partial y^i}$, for $a \in \{1, 2\}$. The solutions $y(\tau)$ of (1) can be expanded as follows. Given a smooth $g : \mathbb{R}^d \to \mathbb{R}$,

$$g(y(\tau)) = g(y(0)) + \sum_{m \geq 1} \sum_{i_1,\ldots,i_m \in A} \alpha_{i_1 \cdots i_m} E_{i_1} \cdots E_{i_m} g(y(0)), \tag{2}$$

*Konputazio Zientziak eta A. A. saila, Informatika Fakultatea, EHU/UPV, Donostia/San Sebastián, Spain (ander@si.ehu.es)

where, for each word $w = i_1 \cdots i_m \in A^*$ over the alphabet $A = \{1, 2\}$,

$$\alpha_{i_1 \cdots i_m}(\tau) = \int_0^{\tau} \int_0^{\tau_m} \cdots \int_0^{\tau_2} \lambda_{i_1}(\tau_1) \cdots \lambda_{i_m}(\tau_m) \, d\tau_1 \cdots d\tau_m. \tag{3}$$

Hereafter, A^* denotes the set of words on the alphabet A (i.e., the free monoid over the set A). Systems of ODEs of the form (1) arise in many applications, for instance, in non-linear control [15], where it can be used to model systems with two controls λ_1, λ_2 (or also, with $\lambda_1(\tau) \equiv 1$, systems with shift and one control λ_2). Series of the form (2)–(3) are referred to as Chen-Fliess series in this context. The series of linear differential operators on the right-hand side of (2) can be written in the form

$$\sum_{w \in A^*} \alpha_w \, E_w, \tag{4}$$

where A^* denotes the set of words on the alphabet $A = \{1, 2\}$, including the empty word e, the function α_w is given for each $w \in A^*$ in (3), and

$$E_e = I, \quad E_w = E_{a_1} \cdots E_{a_m}, \text{ if } w = a_1 \cdots a_m \in A^*, \tag{5}$$

I being the identity operator.

A well-know result due to Chen [6] implies that the series (4) can be formally rewritten as the exponential of a series of vector fields obtained as nested commutators of E_1 and E_2. One of the goals of the present work (Subsection 4.2) is to obtain explicit formulas for such series of vector fields (the continuous Baker-Campbell-Hausdorff formula), expressed in appropriate basis of the Lie algebra generated by the vector fields E_1 and E_2. A different approach that explicitly expresses the continuous BCH formula can be found in [22]. However, compared to our formula, in [22], the continuous BCH formula is written in terms of a spanning set of the free Lie algebra instead of a basis.

It is important to observe that expansions with very similar form with the same underlying combinatorics arise in many other applications. For instance, composition methods in numerical analysis of ODEs [18, 17] (Example 3 below), stochastic differential equations [2] (Example 2), or matrix differential equations on Lie groups [14] (Example 4).

We consider series of the form (4) where the coefficients α_w are elements in a certain commutative ring $\mathbb{K}$ with unit $1_{\mathbb{K}}$, A is a set of indices referred to as the *alphabet*, and the basic objects E_a, $a \in A$, are the generators of a (non-necessarily commutative) associative algebra $\mathcal{B}$ over $\mathbb{K}$. Obviously, all the elements of $\mathcal{B}$ can be written in the form (4), where $\alpha_w \in \mathbb{K}$ for each $w \in A^*$ (α_w being $\neq 0$ for a finite number of words w), and each E_w ($w \in A^*$) is given by (5), I being the unity in $\mathcal{B}$. If the algebra $\mathcal{B}$ is freely generated by $\{E_a : \ a \in A\}$, then such representation is unique. It is straightforward to check that

$$\left(\sum_{w \in A^*} \alpha_w \, E_w \right) \left(\sum_{w \in A^*} \beta_w \, E_w \right) = \left(\sum_{w \in A^*} (\alpha\beta)_w \, E_w \right),$$

where, for each word $w = a_1 \cdots a_m \in A^*$,

$$(\alpha\beta)_w = \alpha_w \beta_e + \alpha_e \beta_w + \sum_{j=1}^{m-1} \alpha_{a_1 \cdots a_j} \, \beta_{a_{j+1} \cdots a_m}. \tag{6}$$

In some applications, the alphabet A may be infinite, but then it typically makes sense assigning a weight (a positive integer) to each letter, and, when needed, truncating the series (4) according to the *weight* $||w||$ of each word $w \in A$. Note that the set A^* of words on the alphabet A with a prescribed weight is finite provided that the subsets $A_k \subset A$ of letters with weight k ($k \geq 1$) are finite.

Example 1 Let $\mathcal{M}$ be a smooth manifold. Consider $\mathbb{K}$ as the ring of piecewise continuous functions $\mathbb{R}^+ \to \mathbb{R}$, and let E_a, $a \in A$ (where A is a certain set of indices) be smooth vector fields on $\mathcal{M}$. Let $\mathcal{B}$ be the algebra over $\mathbb{K}$ of linear operators on $C^\infty(\mathcal{M})$ (the vector space of smooth functions on $\mathcal{M}$) generated by the vector fields E_a, $a \in A$. Given $\lambda_a \in \mathbb{K}$, $a \in A$, the solution operator of the time-varying vector field $E(\tau) = \sum_{a \in A} \lambda_a(\tau) E_a$ can be expanded as a series of linear operators in $\mathcal{B}$ of the form (4), with

$$\alpha_{a_1 \cdots a_m} = \textstyle\int_{a_m} \cdots \int_{a_1} 1_{\mathbb{K}}, \quad a_1, \ldots, a_m \in A, \tag{7}$$

where $1_{\mathbb{K}}$ is the unity function (the unity element in $\mathbb{K}$), and each $\int_a : \mathbb{K} \to \mathbb{K}$ is defined as

$$\{\int_a \mu\}(\tau) = \int_0^\tau \mu(s) \lambda_a(s) ds, \quad a \in A. \tag{8}$$

Notice that such formulas can make sense even in the case of an infinite alphabet A. For instance, consider a non-autonomous vector field $E(\tau) = E_1 + \tau E_2 + \tau^2 E_3 + \cdots$, so that $A = \{1, 2, 3, \ldots\}$, and $\lambda_a(\tau) = \tau^{a-1}$ for each $a \in A$. Consider also $\alpha_w \in \mathbb{K}$ and E_w given for each word w on the alphabet A by (7) and (5), respectively, and let $||w||$ denote the weight of the word $w \in A$ obtained by weighting the letters of the alphabet A as $||a|| = a$. Then, if $y(\tau)$ is an integral curve of the vector field $E(\tau)$ and g is an arbitrary smooth function on $\mathcal{M}$, it holds that $g(y(\tau)) - \sum_{||w||<n} \alpha_w(\tau) \, E_w g\,(y(0)) = \mathcal{O}(\tau^n)$ as $\tau \to 0$. Observe that such a statement makes sense, as the number of words with smaller weight than n is finite. □

Example 2 Consider a d-dimensional stochastic differential equation [2] rewritten in the form

$$y(\tau) = y(0) + \int_0^\tau f_0(y(\tau)) d\tau + \sum_{i=1}^m \int_0^\tau f_i(y(\tau)) dW_i(\tau), \tag{9}$$

where $W_i(\tau)$ ($i = 1, \ldots, m$) represent independent Wiener processes, and the stochastic integrals in the summation on the right-hand side of (9) are interpreted in the Stratonovich sense. In such a case, the formal expansions (2)-(7) hold for the alphabet $A = \{0, 1, \ldots, m\}$, with $\{\int_0 \mu\}(\tau) = \int_0^\tau \mu(s) ds$, and $\{\int_i \mu\}(\tau) = \int_0^\tau \mu(s) dW_i(s)$, for $i \geq 1$. The same combinatorics as in Example 1 arise here because the integration by parts formula also holds for Stratonovich integrals [2]. □

Example 3 An important family of numerical integrators for ODEs, which is particularly interesting when trying to conserve certain properties of the flow that are preserved under composition, can be applied if the original vector field E is split as the sum of two (or more) vector fields $E = E_1 + E_2$ such that the flow of E_1 and E_2 can be computed exactly (or at least very accurately). Each method of that family of integrators is determined by certain coefficients $\mu_1, \ldots, \mu_{2s} \in \mathbb{R}$ that can be chosen [18, 17] in such a way that

$$\exp(\mu_1 E_1)\exp(\mu_2 E_2)\cdots\exp(\mu_{2s-1}E_1)\exp(\mu_{2s}E_2) \tag{10}$$

approximates $\exp(E_1 + E_2)$ in some sense (with the exponentials defined as a power series expansion). It is obvious that both $\exp(E_1 + E_2)$ and (10) can be expanded as series of the form (4)–(5) for the alphabet $A = \{1, 2\}$. A typical requirement of such splitting integrators is that the expansions of $\exp(E_1 + E_2)$ and (10) coincide up to terms of a prescribed degree. (Actually, the same formal approximations (10) of $\exp(E_1 + E_2)$ are also used in many other applications, such as certain partial differential equations, where E_1, E_2 are, instead of vector fields, operator in some space.) More generally, one is required to expand formally expressions of the form

$$\exp\Big(\sum_{a\in A}\mu_{1,a}E_a\Big)\cdots\exp\Big(\sum_{a\in A}\mu_{s,a}E_a\Big), \tag{11}$$

where A is a certain set of indices, and $\mu_{j,a} \in \mathbb{R}$. Clearly, (11) can also be expanded in the form (4)–(5) with $\mathbb{K} = \mathbb{R}$. It is interesting to note that the expansion of (11) can be obtained as a particular case of Example 1, with $\mathbb{K}$ as the ring of piecewise continuous functions $\mathbb{R}^+ \to \mathbb{R}$, and each $\lambda_a \in \mathbb{K}$ defined as $\lambda_a(\tau) = \mu_{j,a}$ if $j - 1 \le \tau < j$ for $j \ge 1$. In such a case, the expansion of (11) corresponds to the series (4)–(5) where the functions $\alpha_w : \mathbb{R}^+ \to \mathbb{R}$, given by (3)–(8), are evaluated at $\tau = s$. $\square$

Example 4 Consider a matrix differential equation of the form $Y' = E(\tau)Y$, $Y(0) = Y_0$. It is often of interest (for instance, when the solution evolves in a Lie group [14]) to approximate the solution $Y(\tau)$ as the exponential $\exp\Omega(\tau)$ of an element $\Omega(\tau)$ of the Lie algebra of matrices generated by $E(c_i\tau)$, $i = 1, \ldots, s$, for appropriately chosen $c_i \in [0, 1]$. If $E(\tau)$ depends smoothly on τ, so that $E(\tau) = E_1 + \tau E_2 + \tau^2 E_3 + \cdots$, then $Y(\tau)$ and $\exp\Omega(\tau)$ can be formally interpreted as series of the form (4)–(5) for the weighted alphabet $A = \{1, 2, 3, \ldots\}$ with weights $||a|| = a$ for each $a \in A$ (the weights of each word $||w||$ here account for the power of τ in each term in (4)). Of course, in this case, we have that $\mathbb{K} = \mathbb{R}$, and $\mathcal{B}$ is the algebra of matrices generated by $\{E_a : \ a = 1, 2, 3, \ldots\}$. $\square$

In the general case where the coefficients α_w in the series (4) are not defined as iterated integrals of the form (3), particularly interesting are expressions (4) that are actually series of elements in the Lie algebra $\mathcal{L}(E_1, E_2)$ generated by $\{E_1, E_2\}$ (which we will loosely refer to as Lie series). If the basic elements E_a are vector fields, (4) being a Lie series implies that it represents a formal vector field (in the sense of a series of vector fields that does not necessarily converge). In general, it is often of interest in applications rewriting the series (4) in a Poincaré-Birkhoff-Witt (PBW) basis associated to a basis of $\mathcal{L}(E_1, E_2)$, and performing, in such basis, operations such as sum, product, commutator, formal exponential, and formal logarithm. (In particular, in such a PBW basis, checking if (4) is a Lie series or the exponential of a Lie series is a trivial task). The main goal of the present work is to give several results that allow performing such operations working in the dual of a PBW basis.

In order to do that, we will make use of rooted trees labeled by the alphabet A, and (although this will not be made clear until Section 7) of a commutative Hopf algebra structure on labeled rooted trees [8, 7, 9]. But how do labeled rooted trees come into play here? The origin of our approach is the observation [12] that the series expansion (4)–(3) of the solution operator of (1) can be alternatively written as

$$I + \sum_{u \in \mathcal{F}} \frac{\alpha(u)}{\sigma(u)} X(u), \tag{12}$$

where the summation is over the set $\mathcal{F}$ of forests of labeled rooted trees, $\sigma(u) \in \mathbb{Z}^+$ is a certain normalization factor (the *symmetry number* of the forest u to be defined in Subsection 2.1 below), each $X(u)$ is a linear differential operator (of order m if the forest u has m labeled rooted trees) acting on smooth functions on $\mathbb{R}^d$ (to be defined in Section 6), and each $\alpha(u)$ is a function on τ that is obtained from λ_a ($a \in A$) in terms of products and iterated integrals, and can be obtained from the functions α_w (3) in (4) by means of Definition 6 in Section 3.1 below. The main results of the present work arose from the study of the relation of $\alpha(u)$ in (12) with the coefficients α_w in (4)–(3). Although series of the form (12) do not make sense in the general case where E_a ($a \in A$) are not vector fields, all of our results, except those in Section 6, hold for the general case where E_a, $a \in A$, are the generators of an arbitrary non-commutative associative algebra $\mathcal{B}$.

The plan of the paper is as follows. In Section 2, labeled rooted trees and forests, and Hall sets of labeled rooted trees are introduced, as well as some definitions and a few results on them to be used in the rest of the paper. Some fundamental results (Theorem 3, Proposition 4, and Corollary 5) that relate a PBW basis of an arbitrary Hall set and labeled rooted trees are presented in Section 3. Such results are used to obtain explicit formulas to compute the coefficients in a PBW basis of the exponential of a Lie series and the logarithm (Subsection 4.1). The result obtained for the logarithm (Theorem 9) in particular provides explicit expressions for the coefficients of the continuous BCH formula (Subsection 4.2), useful for instance, in the contexts of Examples 1–4. Section 5 is devoted to the construction of rewriting algorithms (based on Algorithm 1), that together with Corollary 5 allows computing the coefficients of the product of two series written in a PBW basis. The results are applied in Subsection 5.3 to the computation of the Lie bracket of Lie series. The theory developed to construct our rewriting algorithm allows proving Theorem 3 (essentially Theorem 5.3 in [21]) in an alternative way (Subsection 5.4). The values $\alpha(u)$ for labeled forests u determined from the coefficients of the series (4) in Definition 6 (a key element in our approach) are interpreted in the context of series of vector fields (Section 6), which shows a connection with some results in [12]. In Section 7 we interpret the results in previous sections in the context of Hopf algebras, which we believe provides an interesting insight. Finally, some concluding remarks are given in Section 8.

2 Graph-theoretical tools

2.1 Rooted trees and forests labeled by A

Given an alphabet A, rooted trees and forests labeled by A can be defined as follows.

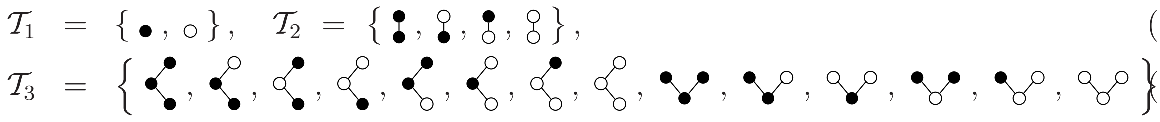
A partially ordered set labeled by A is a partially ordered set U together with a map from U to A. The elements of U are called vertices, and $|U|$ denotes the number of vertices of U. An edge $(x < y)$ of U is an ordered pair $(x, y) \in U \times U$ such that $x < y$ and there exists no $z \in U$ with $x < z < y$. In that case, it is said that x is a parent of y, and that y is a child of x. New labeled partially ordered sets can be obtained from U by adding and/or removing some vertices and/or edges. In particular, given a labeled partially ordered set U with vertices $\{x_1, \ldots, x_n\}$, a different labeled partially ordered subset $V \subset U$ is determined by each subset $\{x_{i_1}, \ldots, x_{i_m}\} \subset \{x_1, \ldots, x_n\}$, with the partial ordering and labelling inherited from U.

An isomorphism of two partially ordered sets labeled by A is a bijection of the underlying sets of vertices that preserves the partial order and the labelling. A forest labeled by A is an isomorphism class of finite partially ordered sets U labeled by A satisfying

$$x, y, z \in U, \ y < x, \ z < x \quad \Longrightarrow \quad \text{either} \ \ y < z \ \ \text{or} \ \ z < y \ \ \text{or} \ \ z = y. \tag{13}$$

The roots of a labeled partially ordered set U representing a labeled forest u are its minimal vertices. A rooted tree labeled by A is a forest represented by labeled partially ordered sets with only one root.

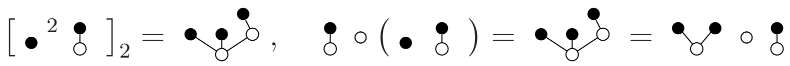
The degree $|u|$ of a forest u labeled by A represented by a labeled partially ordered set U is the number $|U|$ of vertices of U. Given a forest u labeled by A, the partial degree $|u|_a$ of u with respect to $a \in A$ is the number of vertices in U that are labeled by a. If A is a weighted alphabet, where a positive integer weight $||a||$ is associated to each letter $a \in A$, so that a vertex labeled by a has weight $||a||$, the weight of the forest u is defined as the sum of the weights of its vertices. We denote as $\mathcal{F}$ (resp., $\mathcal{T}$) the set of forests (resp., rooted trees) labeled by A. The empty forest is also included in $\mathcal{F}$, and we denote it as e. We denote $\mathcal{F}_k = \{u \in \mathcal{F} : \ |u| = k\}$ for each $k \geq 0$ and $\mathcal{T}_k = \{u \in \mathcal{T} : \ |u| = k\}$ for $k \geq 1$. When representing labeled rooted trees graphically, we will position the root at the bottom of the diagram. A vertex labeled by the letter $a \in A$ in a labeled rooted tree can be represented as a small circle with the letter a inside. Alternatively, we can assign a color to each letter in A, and then depict the vertex as a small circle in that color. In particular, in our examples with the two-letter alphabet $A = \{1, 2\}$, we assign 'black' to 1, and 'white' to 2. The first sets $\mathcal{T}_k$ for that alphabet are



(14)

(15)

We will later work with the ring $\mathbb{Z}[\mathcal{T}]$ of polynomials with the labeled rooted trees as commuting indeterminates, with $\mathbb{Z}$-basis $\{u_1^{r_1}\cdots u_m^{r_m},\ u_1,\ldots,u_m \in \mathcal{T}\}$. Clearly, each such expression $u_1^{r_1}\cdots u_m^{r_m}$ can be identified with a unique forest $u \in \mathcal{F}$, the forest obtained as the direct union, from $i=1$ to m of r_i copies of the labeled rooted tree u_i. We thus simply write $u = u_1^{r_1}\cdots u_m^{r_m}$, and denote as uv the direct union of two forests $u, v \in \mathcal{F}$ (if U and V are two labeled partially ordered sets representing u and v respectively, with disjoint sets of vertices, then uv is the forest represented by the union of U and V). We will consider several additional operations on the set $\mathcal{F}$ of labeled forests. Given $a \in A$, $t_1,\ldots,t_m \in \mathcal{T}$, $u = t_1\cdots t_m \in \mathcal{F}$, we denote by $[u]_a$ the labeled rooted tree of degree $|t_1|+\cdots+|t_m|+1$ obtained by grafting the roots of $t_1,\ldots,t_m$ to a new root labeled by a. That is, it corresponds to adding a new vertex r labeled by a to a partially ordered forest U representing u, and adding, for each t_i, a new edge $(r < r_i)$ connecting the root r_i of t_i with r. In particular, $[e]_a$ is the labeled rooted tree with only one vertex, labeled by a. We will identify $[e]_a$ simply with a when its meaning is clear from the context. Given a labeled rooted tree $t \in \mathcal{T}$ and a labeled forest $u \in \mathcal{F}$, we denote by $t \circ u$ the labeled rooted tree of degree $|u|+|t|$ obtained by grafting the labeled rooted trees in u to the root of t (this operation is often referred to in the context of numerical analysis of ordinary differential equations as the Butcher product). For instance, ($A=\{1,2\}$, $1\to\bullet$, $2\to\circ$)



In particular, $t \circ e = t$ and $[e]_a \circ u = [u]_a$ for $a \in A$, $t \in \mathcal{T}$, $u \in \mathcal{F}$. We also write $e \circ e = e$, $e \circ u = 0$, for each $u \in \mathcal{F}\backslash\{e\}$. The grafting operation $\circ$ is not associative, however,

$$(t \circ u) \circ v = t \circ (uv) = (t \circ v) \circ u \text{ for each } t \in \mathcal{T},\ u, v \in \mathcal{F}. \tag{16}$$

Given $t, t_1, \ldots, t_m \in \mathcal{T}$, $u \in \mathcal{F}$, we will use the notation

$$t_1 \circ t_2 \circ \cdots \circ t_m \circ u := t_1 \circ (t_2 \circ \cdots \circ (t_m \circ u)), \quad t^{\circ k} = \overbrace{t \circ \cdots \circ t}^{k},$$

that is, in the absence of parentheses, we interpret multiple grafting of rooted trees from right to left (notice that, by virtue of (16), for multiple grafting from left to right we have that $(\cdots((t_1 \circ t_2) \circ t_3) \cdots \circ t_m) = t_1 \circ (t_2 \cdots t_m))$.

The symmetry number $\sigma(u)$ of a labeled forest u is the number of different permutations of the set of vertices of a labeled partially ordered set representing u that are isomorphisms of labeled partially ordered sets. The symmetry number of forests can be recursively obtained as follows.

Lemma 1 *For each $a \in A$, $t_1, \dots, t_m \in \mathcal{T}$, $t_i \neq t_j$ if $i \neq j$,*

$$\sigma(e) = 1, \quad \sigma([u]_a) = \sigma(u), \quad \sigma(u) = \prod_{j=1}^{m} i_j!\sigma(t_j)^{i_j}, \quad \textit{if} \quad u = \prod_{j=1}^{m} t_j^{i_j}.$$

2.2 Hall sets of rooted trees

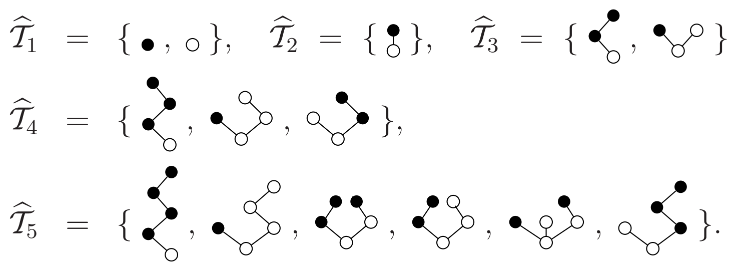
Definition 1 *A set $\widehat{\mathcal{T}}$ of rooted trees labeled by A together with a total order relation $>$ is a* Hall set *(of labeled rooted trees) over A, if the following conditions hold:*

1. *If $a \in A$, then $[e]_a \in \widehat{\mathcal{T}}$.*
2. *Given $a \in A$ and $u \in \mathcal{F}\backslash\{e\}$, where $u = t_1^{\circ r_1} \cdots t_m^{\circ r_m}$, $t_1, \dots, t_m \in \widehat{\mathcal{T}}$, $r_1, \dots, r_m \geq 1$, and $t_1 > \cdots > t_m$,*
$$[u]_a \in \widehat{\mathcal{T}} \quad \textit{if and only if} \quad t_m > [t_1^{\circ r_1} \cdots t_{m-1}^{\circ r_{m-1}}]_a \in \widehat{\mathcal{T}}. \tag{17}$$
3. *If $t = [t_1^{\circ r_1} \cdots t_m^{\circ r_m}]_a \in \widehat{\mathcal{T}}$, $t_1, \dots, t_m \in \widehat{\mathcal{T}}$, $r_1, \dots, r_m \geq 1$, $a \in A$, then $t_j > t$ for each $j = 1, \dots, m$.*

In (17), it is understood that $t_1^{\circ r_1} \cdots t_{m-1}^{\circ r_{m-1}} = e$ if $m = 1$. Thus, given $t \in \widehat{\mathcal{T}}$, $r \geq 1$, and $a \in A$, the labeled rooted tree $[t^{\circ r}]_a$ belongs to $\widehat{\mathcal{T}}$ if and only if $t > [e]_a$.

For instance, the sets of labeled rooted trees

$\widehat{\mathcal{T}}_1$ =, $\widehat{\mathcal{T}}_2$ =, $\widehat{\mathcal{T}}_3$ =, $\widehat{\mathcal{T}}_4$ =, $\widehat{\mathcal{T}}_5$ =



(18)

are the first homogeneous subsets $\widehat{\mathcal{T}}_k = \{t \in \widehat{\mathcal{T}} : \ |t| = k\}$ of a Hall set of labeled rooted trees over the alphabet $A = \{1, 2\}$ ($1 \to \bullet$, $2 \to \circ$). A total order of the displayed rooted trees that is compatible with the definition of Hall sets can be obtained by considering that the elements of each $\widehat{\mathcal{T}}_k$ ($k = 1, 2, 3, 4, 5$) are given in decreasing order in (18), and that $t > z$ if $|t| < |z|$.

Given a fixed Hall set $\widehat{\mathcal{T}}$ of labeled rooted trees, we consider the set of forests $\widehat{\mathcal{F}} = \{e\} \cup \{t_1^{\circ r_1} \cdots t_m^{\circ r_m} : \ r_1, \dots, r_m \geq 1, \ t_1, \dots, t_m \in \widehat{\mathcal{T}}, \ t_i \neq t_j \text{ if } i \neq j\}$. When $t \in \widehat{\mathcal{T}}$, we will say that t is a Hall rooted tree, and when $u \in \widehat{\mathcal{F}}$, that u is a Hall forest. Note that the labeled rooted trees of the form $t^{\circ r}$ with $t \in \widehat{\mathcal{T}}$ and $r > 1$ are Hall forests but not Hall rooted trees. We will also denote $\widehat{\mathcal{T}}_k = \{t \in \widehat{\mathcal{T}} : \ |t| = k\}$ and $\widehat{\mathcal{F}}_{k-1} = \{u \in \widehat{\mathcal{T}} : \ |u| = k-1\}$ for each $k \geq 1$.

Lemma 1 and Definition 1 show that $\sigma(u) = 1$ for each $u \in \widehat{\mathcal{F}}$, that is, Hall rooted trees and Hall forests have no symmetries (apart from the trivial identity permutation of its vertices).

Remark 1 There is a one-to-one correspondence between Hall sets of labeled rooted trees over $\widehat{\mathcal{T}}$ and a Hall set $\mathcal{W}$ of words [21] over A. Given a Hall set $\widehat{\mathcal{T}}$ of labeled rooted trees, consider the map Ω from $\widehat{\mathcal{T}}$ to the set of words A^* on the alphabet A defined as follows. For each $a \in A$, Ω assigns the one-letter word a to the rooted tree $[e]_a$, and given $t = [t_1^{\circ r_1} \cdots t_m^{\circ r_m}]_a \in \widehat{\mathcal{T}}$ such that $a \in A$, $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$, $r_1, \ldots, r_m \geq 1$, and $t_1 > \cdots > t_m$, Ω assigns the Hall word $\Omega(t) = \Omega(t_m)^{r_m} \cdots \Omega(t_1)^{r_1}\, a$ to the Hall rooted tree t. Here we use the standard notation $w_1 w_2$ for the concatenation of two words w_1 and w_2. Recall that this defines an associative binary operation that (contrary to the product $u_1 u_2$ of two forests $u_1, u_2 \in \mathcal{F}$) is not commutative.

It can be seen that such Ω is injective, and its image (together with the total order induced from that of $\widehat{\mathcal{T}}$) is a Hall set $\mathcal{W}$ of words over A (in the general sense of [21]).

The image by Ω of the sets H_k in (18) are the homogeneous subsets

$$\begin{aligned} \mathcal{W}_1 &= \{1,2\}, \quad \mathcal{W}_2 = \{12\}, \quad \mathcal{W}_3 = \{112,\ 212\}, \\ \mathcal{W}_4 &= \{1112,\ 2212,\ 1122\}, \quad \mathcal{W}_5 = \{11112,\ 22212,\ 12112,\ 22112,\ 12212,\ 11122\}, \end{aligned} \tag{19}$$

of a particular Hall set $\mathcal{W}$ of words over the alphabet $\{1,2\}$.

Hall sets of rooted trees labeled by A can be alternatively constructed from an arbitrary Hall set $\mathcal{W}$ of words on the alphabet A as the range of an injective map $\tau : \mathcal{W} \to \mathcal{T}$ recursively defined as follows: For words of degree 1, $\tau(a) = [e]_a$. Given a word $w \in \mathcal{W}$ of degree $|w| > 1$, consider $a \in A$ and the (non-necessarily Hall) word v on the alphabet A such that $w = va$. As any word on A, v can be written as a non-decreasing product of Hall words [21], $v = w_1^{r_1} \cdots w_m^{r_m}$, $w_1, \ldots, w_m \in \mathcal{W}$, $w_1 < \cdots < w_m$, $r_1 \ldots, r_m \geq 1$. Then, $\tau(w)$ is defined as $\tau(w) = [\tau(w_1) \cdots \tau(w_m)]_a$.

2.3 Technical results on Hall rooted trees

Definition 2 *The* standard decomposition $(t', t'') \in \widehat{\mathcal{T}} \times \widehat{\mathcal{T}}$ *of each* $t \in \widehat{\mathcal{T}}$ *is defined as follows. If* $|t| = 1$*, then* $t' = t$ *and* $t'' = e$*. If* $t = [t_1^{\circ r_1} \cdots t_m^{\circ r_m}]_a$*,* $a \in A$*,* $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$*,* $t_1 > \cdots > t_m$*,* $r_1, \ldots, r_m \geq 1$*, then* $t'' = t_m$*,* $t' = [t_1^{\circ r_1} \cdots t_{m-1}^{\circ r_{m-1}} t_m^{\circ (r_m - 1)}]_a$*. Similarly, the standard decomposition* $(u', u'') \in \widehat{\mathcal{F}} \times \widehat{\mathcal{T}}$ *of a Hall forest* $u \in \widehat{\mathcal{F}} \backslash \widehat{\mathcal{T}}$ *is given as follows. If* $u = t_1^{\circ r_1} \cdots t_m^{\circ r_m}$*,* $t_1 > \cdots > t_m$*, then* $u' = t_1^{\circ r_1} \cdots t_{m-1}^{\circ r_{m-1}} t_m^{\circ (r_m - 1)}$*,* $u'' = t_m$*.*

Definition 3 *We define a map* $\Gamma : \widehat{\mathcal{F}} \backslash \{e\} \to \mathcal{T}$ *as follows. Given* $u \in \widehat{\mathcal{F}}$*, where* $u = t_1^{\circ r_1} \cdots t_m^{\circ r_m}$*,* $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$*,* $m, r_1, \ldots, r_m \geq 1$*,* $t_1 > \cdots > t_m$*, then* $\Gamma(u) = t_1^{\circ r_1} \circ (t_2^{\circ r_2} \cdots t_m^{\circ r_m})$*.*

Lemma 2 *The map* Γ *in Definition 3 is injective, and its image is the set* $\{[u]_a \in \mathcal{T} : u \in \widehat{\mathcal{F}},\ a \in A\}$*. That is, given* $u \in \widehat{\mathcal{F}}$ *and* $a \in A$*, there exist unique* $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$ *and* $r_1, \ldots, r_m \geq 1$ *such that* $t_1 > \cdots > t_m$ *and* $t_1^{\circ r_1} \circ (t_2^{\circ r_2} \cdots t_m^{\circ r_m}) = [u]_a$*.*

Proof: Assume without loss of generality that $u = z_1^{\circ l_1} \cdots z_k^{\circ l_k}$, where $z_1, \ldots, z_k \in \widehat{\mathcal{T}}$, $z_1 > \cdots > z_k$, $l_1, \ldots, l_k \geq 1$.

We will first see that under the assumption that such $t_1^{\circ r_1}, \ldots, t_m^{\circ r_m}$ exist, they are unique. As $t_1 \in \widehat{\mathcal{T}}$, we have that $t_1 = [s_1^{\circ n_1} \cdots s_i^{\circ n_i}]_a$, where $s_1, \ldots, s_i \in \widehat{\mathcal{T}}$, $n_1, \ldots, n_i \geq 1$, and $s_1 > \cdots > s_i > t_1 > t_2 > \cdots > t_m$, and thus $(s_1^{\circ n_1}, \ldots, s_i^{\circ n_i}, t_1^{\circ (r_1 - 1)}, t_2^{\circ r_2}, \ldots, t_m^{\circ r_m}) = (z_1^{\circ l_1}, \ldots, z_k^{\circ l_k})$. Moreover, i is the largest integer such that $[z_1^{\circ l_1} \cdots z_i^{\circ l_i}]_a \in \widehat{\mathcal{T}}$ ($t_1 = [e]_a$ if $i = 0$), as $[s_1^{\circ n_1} \cdots s_i^{\circ n_i} t_1^{\circ (r_1 - 1)}]_a \notin \widehat{\mathcal{T}}$ and $[s_1^{\circ n_1} \cdots s_i^{\circ n_i} t_2^{\circ r_2}]_a \notin \widehat{\mathcal{T}}$ (due to the fact that $[s_1^{\circ n_1} \cdots s_i^{\circ n_i}]_a = t_1 > t_2$).

In order to prove the existence, let i be the largest integer such that $[z_1^{\circ l_1} \cdots z_i^{\circ l_i}]_a \in \widehat{\mathcal{T}}$ ($z = [e]_a$ if $i = 0$), and let $t_1 = [z_1^{\circ l_1} \cdots z_i^{\circ l_i}]_a$. It is enough proving that $t_1 \geq z_{i+1}$ (in such a case, if $t_1 = z_{i+1}$, then $(t_1^{\circ r_1}, \ldots, t_m^{\circ r_m}) = (z_{i+1}^{\circ (l_{i+1}+1)}, z_{i+2}^{\circ l_{i+2}}, \ldots, z_k^{\circ l_k})$, and otherwise, $(t_1^{\circ r_1}, \ldots, t_m^{\circ r_m}) = (t_1, z_{i+1}^{\circ l_{i+1}}, \ldots, z_k^{\circ l_k})$, and it obviously holds that $t_1 > \cdots > t_m$ if $z_1 > \ldots > z_k$.). The inequality $t_1 \geq z_{i+1}$ follows from the observation that, if $t_1 < z_{i+1}$, then, we have (as $z_1 > \cdots > z_{i+1} > t_1$), that $= [z_1^{\circ l_1} \cdots z_{i+1}^{\circ l_{i+1}}]_a \in \widehat{\mathcal{T}}$, which is not possible by definition of i. $\square$

Remark 2 Lemma 2 shows in particular that, for each $n \geq 1$, there exists a bijection between $\widehat{\mathcal{F}}_n \times A$ and $\widehat{\mathcal{F}}_{n+1}$. Furthermore, from its proof it is clear that such a bijection preserves all the partial degrees. Comparing that with the trivial bijection between $A^n \times A$ and the set A^{n+1} that gives the word wa of degree $|wa| = n + 1$ for each $(w, a) \in (A^n, A)$, it follows that the number of Hall forests $u \in \widehat{\mathcal{F}}$ with prescribed partial degrees $|u|_a = d_a$, $a \in A$, $d_a \geq 0$, coincides with the number of words $w \in A^*$ with $|w|_a = d_a$.

Remark 3 The bijection in Remark 2 is related to known results on Hall words and the maps Ω and τ in Remark 1 as follows: Given $a \in A$ and the Hall forest $u = z_1^{\circ l_1} \cdots z_k^{\circ l_k}$, where $z_1, \ldots, z_k \in \widehat{\mathcal{T}}$, $z_1 > \cdots > z_k$, $l_1, \ldots, l_k \geq 1$, consider the word $w = \Omega(z_k)^{l_k} \cdots \Omega(z_1)^{l_1}\, a$. Now, there exists [21] a unique decomposition of w as an increasing product of Hall words $w = w_m^{r_m} \cdots w_1^{r_1}$, with Hall words $w_m < \cdots < w_1$, $r_1, \ldots, r_m \geq 1$. Then it can be seen that $[z_1^{\circ l_1} \cdots z_k^{\circ l_k}]_a = t_1^{\circ r_1} \circ (t_2^{\circ r_2} \cdots t_m^{\circ r_m})$, where $t_i = \tau(w_i)$, $i = 1, \ldots, m$.

We define a partial order $\succ$ in the set $\widehat{\mathcal{F}}$ of Hall forests, to be used in Section 5 below as follows.

Definition 4 *Given $u, v \in \widehat{\mathcal{F}}$ with the same partial degrees, where $u = t_1^{\circ r_1} \cdots t_m^{\circ r_m}$, $v = z_1^{\circ l_1} \cdots z_k^{\circ l_k}$, $t_1, \ldots, t_m, z_1, \ldots, z_k \in \widehat{\mathcal{T}}$, $m, l, r_1, \ldots, r_m, l_1, \ldots, l_k \geq 1$, $t_1 > \cdots > t_m$ and $z_1 > \cdots > z_k$, If $t_1 \geq z_1'' > z_k \geq t_m$ (where (z_1', z_1'') is the standard decomposition of z_1), then $u \succ v$.*

Remark 4 Notice that, if $u \succ v$, $u, v \in \widehat{\mathcal{F}}$, then $\max(u) > \max(v) \geq \min(v) \geq \min(u)$, where we use the notation $\max(t_1^{\circ r_1} \cdots t_m^{\circ r_m}) = t_1$ if $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$, $t_1 > \cdots > t_m$, $r_1, \ldots, t_m \geq 1$.

3 Poincaré-Birkhoff-Witt basis and labeled rooted trees

3.1 Expressing series in a PBW basis associated to a set of Hall forests

Let $\mathcal{B}$ be a (non-necessarily commutative) associative algebra $\mathcal{B}$ generated by the elements $\{E_a :\ a \in A\}$. We next define one element $E(u) \in \mathcal{B}$ for each Hall forest of labeled rooted trees $u \in \mathcal{F}$. If the algebra $\mathcal{B}$ is freely generated by $\{E_a :\ a \in A\}$, that is, if $\{E_w :\ w \in A^*\}$ is a basis of $\mathcal{B}$, then $\{E(t) :\ t \in \widehat{\mathcal{T}}\}$ is a (Hall) basis (ordered according to the total order relation in $\widehat{\mathcal{T}}$) of the Lie algebra $\mathfrak{g}$ generated by $\{E_a :\ a \in A\}$, and $\{E(u) :\ u \in \widehat{\mathcal{F}}\}$ is the PBW basis associated to that ordered basis of $\mathfrak{g}$.

Definition 5 *Given a Hall set $\widehat{\mathcal{T}}$ of rooted trees over A and the corresponding set $\widehat{\mathcal{F}}$ of Hall forests, we assign an element $E(u) \in \mathcal{B}$ for each $u \in \widehat{\mathcal{F}}$ recursively as follows. For the empty forest, $E(e) = I$, for each $t \in \widehat{\mathcal{T}}$ with standard decomposition $(t', t'') \in \widehat{\mathcal{T}} \times \widehat{\mathcal{T}}$,*

$$E(t) = [E(t''), E(t')] = E(t'')E(t') - E(t')E(t''), \tag{20}$$

and given $u \in \widehat{\mathcal{F}} \backslash \widehat{\mathcal{T}}$ with standard decomposition $(u', u'') \in \widehat{\mathcal{F}} \times \widehat{\mathcal{T}}$,

$$E(u) = E(u'')E(u'). \tag{21}$$

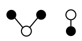
In the general case (where $\mathcal{B}$ may not be freely generated by $\{E_a :\ a \in \mathcal{B}\}$), $\mathfrak{g}$ is spanned by $\{E(t) :\ t \in \widehat{\mathcal{T}}\}$, and $\mathcal{B}$ is spanned by $\{E(u) :\ u \in \widehat{\mathcal{F}}\}$.

Theorem 3 below is an essential ingredient of the present work. We first need the following definition.

Definition 6 *Given a map that assigns $\alpha_w \in \mathbb{K}$ to each $w \in A^*$, we consider the map $\alpha : \mathcal{F} \to \mathbb{K}$ defined as follows. For the empty forest $\alpha(e) = \alpha_e$. Given $u \in \mathcal{F}\backslash\{e\}$, consider a labeled partially ordered set U representing the forest u, and let $x_1, \ldots, x_n$ be the vertices of U labeled as $l(x_i) = a_i \in A$ for each $i = 1, \ldots, n$. For each total order relation $>_U$ on the set of vertices of U, $x_{i_1} > \cdots > x_{i_n}$, that extends the partial order relation in U, we denote the word $a_{i_1} \cdots a_{i_n}$ as $w(>_U)$. Then*

$$\alpha(u) = \sum_{>_U} \alpha_{w(>_U)}, \tag{22}$$

where the summation goes over each total order relation $>_U$ on the set of vertices of U that extends its partial order relation.

Example 5 Consider the labeled forest $u =$  over the alphabet $A = \{1, 2\}$ ($1 \to \bullet$, $2 \to \circ$). It is represented, for instance, by the labeled partially ordered set $U = \{x_1, x_2, x_3, x_4, x_5\}$ with partial order $x_1 < x_2$, $x_1 < x_3$, and $x_4 < x_5$ and labeled as $l(x_2) = l(x_3) = l(x_4) = 1$, $l(x_1) = l(x_5) = 2$. In that case, each different total ordering $x_{i_1} > \cdots >$

x_{i_5} of U that extends its partial order relation is characterized by the 5-tuple $(i_1, \ldots, i_5)$. Such total orderings of U are obtained in this case for

$$\begin{aligned}(i_1,\ldots,i_5) \;=\;& (2,3,1,5,4),(2,3,5,4,1),(2,3,5,1,4),(2,5,4,3,1),(2,5,3,4,1),(2,5,3,1,4),\\ & (3,2,1,5,4),(3,2,5,1,4),(3,2,5,4,1),(3,5,4,2,1),(3,5,2,4,1),(3,5,2,1,4),\\ & (5,4,3,2,1),(5,4,2,3,1),(5,3,4,2,1),(5,3,2,4,1),\\ & (5,3,2,1,4),(5,2,4,3,1),(5,2,3,4,1),(5,2,3,1,4).\end{aligned}$$

□

Theorem 3 *Given a map that assigns $\alpha_w \in \mathbb{K}$ to each word $w \in A^*$, it holds that*

$$\sum_{w\in A^*} \alpha_w\, E_w = \sum_{u\in\widehat{\mathcal{F}}} \alpha(u)\, E(u), \tag{23}$$

where $\alpha : \mathcal{F} \to \mathbb{K}$ is given in Definition 6 above.

The discussion on the proof of Theorem 3 is postponed to Subsection 5.4, where it is shown that it is essentially equivalent to Theorem 5.3 in [21].

Remark 5 Theorem 3 implies that, given $\alpha : \mathcal{F} \to \mathbb{K}$ obtained by Definition 6 for some coefficients α_w, $w \in A^*$, then such α_w can be uniquelly written as a linear combination of the values $\alpha(u)$, $u \in \widehat{\mathcal{F}}$. To see this, it is enough to consider Theorem 3 with $\mathcal{B} = \mathbb{K}\langle A\rangle$ (the set of $\mathbb{K}$-linear combinations of words on the alphabet A), so that $\{E_w :\ w \in A^*\}$ is linearly independent, and the fact that by definition, each $E(u)$ is a linear combination of $\{E_w :\ w \in A^*\}$. We will give a direct proof (without using Theorem 3) of this statement in Section 5 (Corollary 19) when constructing an algorithm that allows, in particular, effectively rewriting α_w as a linear combination of the values $\alpha(u)$, $u \in \widehat{\mathcal{F}}$.

Remark 6 We have considered PBW basis of $\mathcal{B}$ (when $\mathcal{B}$ is freely generated by $\{E_a :\ a \in A\}$) consisting of increasing products of the ordered basis $\{E(t) :\ t \in \widehat{\mathcal{T}}\}$ of the Lie algebra generated by $\{E_a :\ a \in A\}$. PBW basis of $\mathcal{B}$ made of decreasing products can similarly be considered. One can translate from one case to the other by considering the opposite algebra of $\mathcal{B}$ (with product $*$ given by $X * Y = YX$ for $X, Y \in \mathcal{B}$). This simple trick gives

$$\sum_{w\in A^*} \alpha_w E_w = \sum_{u\in\widehat{\mathcal{F}}} \bar{\alpha}(u)\bar{E}(u),$$

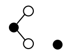
where $\bar{E}(t) = (-1)^{|t|}E(t)$, $\bar{E}(t_1^{\circ r_1} \cdots t_m^{\circ r_m}) = \bar{E}(t_1)^{r_1} \cdots \bar{E}(t_m)^{r_m}$ if $t_1 > \cdots > t_m$, and $\bar{\alpha} : \mathcal{F} \to \mathbb{K}$ is given by Definition 6 for $\bar{\alpha}_{a_1\cdots a_m} = \alpha_{a_m\cdots a_1}$, $a_1, \ldots, a_m \in A$.

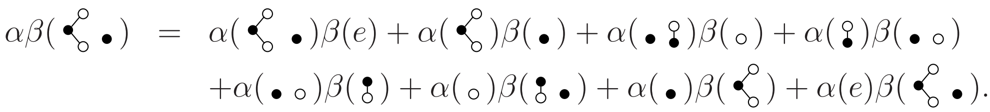
3.2 The product of two series in a PBW basis associated to a set of Hall forests

Definition 7 *Given two arbitrary maps $\alpha, \beta : \mathcal{F} \to \mathbb{K}$, we define $\alpha\beta : \mathcal{F} \to \mathbb{K}$ as follows. Given $u \in \mathcal{F}\backslash\{e\}$, let U be a labeled partially ordered set representing u, then*

$$\alpha\beta(u) = \sum_{(V,W)\in R(U)} \alpha(V)\,\beta(W), \tag{24}$$

where $(V, W) \in R(U)$ if V and W are labeled partially ordered subsets of U satisfying the following two conditions: (i) The set of vertices of U is the disjoint union of the sets of vertices of V and W; and (ii) Given $x, y \in U$ such that $x > y$, if $x \in W$, then $y \in W$.

Example 6 For instance, consider the labeled forest $u =$  over the alphabet $A = \{1, 2\}$. The labeled partially ordered set U with vertices x_1, x_2, x_3, x_4 partially ordered as $x_1 < x_3$, $x_3 < x_4$, and labeled as $l(x_2) = l(x_3) = 1$, $l(x_1) = l(x_4) = 2$, represents u. Each $(V, W) \in R(U)$ is determined by the set of vertices of V, and there are eight different possibilities for V, that are $\{x_1, x_2, x_3, x_4\}$, $\{x_1, x_3, x_4\}$, $\{x_2, x_3, x_4\}$, $\{x_3, x_4\}$, $\{x_2, x_4\}$, $\{x_4\}$, $\{x_2\}$, $\emptyset$, and thus,



□

Proposition 4 *Assume that $\alpha_w, \beta_w \in \mathbb{K}$ for each $w \in A^*$, and let $\alpha, \beta : \mathcal{F} \to \mathbb{K}$ be given by Definition 6. The map $\alpha\beta : \mathcal{F} \to \mathbb{K}$ given by Definition 7 and that given by Definition 6 for the coefficients $(\alpha\beta)_w$, $w \in A^*$ in (6) coincide.*

Proof: Here we give a direct proof of combinatorial nature. An algebraic proof is given in the discussion of Section 7 below (Corollary 28).

Consider a labeled partially ordered set U representing the forest u, and let $x_1, \ldots, x_n$ be the vertices of U labeled as $l(x_i) = a_i \in A$ for each $i = 1, \ldots, n$. Then, $\alpha\beta(u)$ given by Definition 6 for the coefficients $(\alpha\beta)_w$, $w \in A^*$, given in (6) is the sum, over each $j \in \{0, \ldots, n\}$ and each total ordering $x_{i_1} > \cdots > x_{i_n}$ of U that extends its partial order relation, of $\alpha(a_{i_1} \cdots a_{i_j})\beta(a_{i_{j+1}} \cdots a_{i_n})$. Each pair of such total ordering of U and j, determines different $(V, U, >_V, >_W)$ such that $(V, W) \in R(U)$ and $>_V$ (resp., $>_W$) is a total ordering of the set of vertices of V (resp., W) that extends its partial ordering, V (resp., W) being the labeled partially ordered subset of U determined by the vertices $\{x_{i_1,\ldots,x_{i_j}}\}$ (resp., $\{x_{i_{j+1},\ldots,x_{i_n}}\}$) totally ordered as $x_{i_1} > \cdots > x_{i_j}$ (resp., $x_{i_{j+1}} > \cdots > x_{i_n}$). Now, (24) follows from the observation that each $(V, U, >_V, >_W)$ such that $(V, W) \in R(U)$ and $>_V$ (resp., $>_W$) is a total ordering of the set of vertices of V (resp., W) that extends its partial ordering can be obtained in this way from a pair $(j, >_U)$ where $j \in \{0, \ldots, |U|\}$ and $>_U$ is a total ordering of the set of vertices of U that is compatible with its partial order relation. □

Theorem 3 and Proposition 4 imply the following:

Corollary 5 *Assume that $\alpha_w, \beta_w \in \mathbb{K}$ for each $w \in A^*$. Then it holds that*

$$\Big(\sum_{w\in A^*} \alpha_w\, E_w\Big)\Big(\sum_{w\in A^*} \beta_w\, E_w\Big) = \sum_{u\in\widehat{\mathcal{F}}} \alpha\beta(u)\, E(u). \tag{25}$$

Remark 7 A similar formula is obtained for 'PBW basis' of decreasing products using the notation and the trick in Remark 6, namely,

$$\Big(\sum_{w\in A^*} \alpha_w\, E_w\Big)\Big(\sum_{w\in A^*} \beta_w\, E_w\Big) = \sum_{u\in\widehat{\mathcal{F}}} \bar{\beta}\bar{\alpha}(u)\, \bar{E}(u).$$

If one explicitly has the values $\alpha(u), \beta(u)$ ($u \in \mathcal{F}$) given by Definition 6 (this is the case, for instance, of series with iterated integrals considered in Subsection 4.2 below), then Corollary 5 allows expressing the product of series of the form (4) in the PBW basis $\{E(u) : \ u \in \widehat{\mathcal{F}}\}$ directly, instead of first obtaining the coefficients $(\alpha\beta)_w$ for each word $w \in A^*$, and then computing the values $\alpha\beta(u)$ ($u \in \widehat{\mathcal{F}}$) given by Definition 7 in terms of the coefficients $(\alpha\beta)_w$. This will be exploited in Section 4, where, in particular, explicit formulas for the BCH formula and the continuous BCH formula are obtained.

Otherwise, assume that we want to compute the product of series that are written as in the right-hand side of (23). In such a case, one could always rewrite each α_w ($w \in A^*$) as a linear combination of the known values $\alpha(u)$, $u \in \widehat{\mathcal{F}}$ (see Remark 5 above), then obtain the coefficients $(\alpha\beta)_w$ by means of (6), and finally obtain the coefficients $\alpha\beta(u)$ ($u \in \widehat{\mathcal{F}}$) in (25) by applying Definition 6. Instead of that, one can directly obtain the coefficients $\alpha\beta(u)$ ($u \in \widehat{\mathcal{F}}$) from Definition 7. Unfortunately, given $u \in \widehat{\mathcal{F}}$ and $(V, W) \in R(U)$ in (24), the labeled partially ordered sets V and W may not represent Hall forests (V always represents a product of Hall rooted trees, but this is not the case for W). In order to overcome this difficulty, we need some rewriting algorithm that allows expressing $\alpha(v)$ for arbitrary $v \in \mathcal{F}$ as a linear combination of $\{\alpha(u) : \ u \in \widehat{\mathcal{F}}\}$. This will be accomplished in Section 5 below.

4 Exponentials of Lie series and explicit continuous BCH formulas

4.1 The exponential, the logarithm, and Lie series

In the present subsection we assume that $\mathbb{K}$ is a $\mathbb{Q}$-algebra. Recall that the formal exponential (defined as a power series) of a series (4) with $\alpha_e = 0$ is a new series of the form (4) with $\alpha_e = 1$, and that its inverse is the formal logarithm, defined for series (4) with $\alpha_e = 1$.

Definition 8 *Given a map* $\alpha : \mathcal{F} \to \mathbb{K}$*. If* $\alpha(e) = 0$*, then the exponential is defined as a new map* $\exp\alpha$ *determined as follows: For the empty forest,* $\exp\alpha(e) = 1$*, and for each* $u \in \mathcal{F}\backslash\{e\}$*,*

$$\exp\alpha(u) = \sum_{k=1}^{|u|} \frac{1}{k!}\alpha^k(u). \tag{26}$$

If $\alpha_e = 1$*, the logarithm* $\log\alpha$ *of* α *is a new map defined as follows:*

$$\log\alpha(e) \;=\; 0, \quad \log\alpha(u) = \sum_{k=1}^{|u|} \frac{(-1)^{k+1}}{k}(\alpha-\epsilon)^k(u), \quad \textit{for each } u \in \mathcal{F}\backslash\{e\}, \tag{27}$$

where $\epsilon(e) = 1$ *and* $\epsilon(u) = 0$ *for* $u \in \mathcal{F}\backslash\{e\}$*.*

Here we use the notation α^k referred to as the product given in Definition 7.

Remark 8 The fact that $\alpha^k(u) = 0$ if $\alpha(e) = 0$ and $|u| < k$ implies that, given $\alpha : \mathcal{F} \to \mathbb{K}$ such that $\alpha(e) = 0$, then

$$\exp\alpha = \epsilon + \sum_{k\geq 1} \frac{1}{k!}\,\alpha^k, \quad \log(\epsilon+\alpha) = \sum_{k\geq 1} \frac{(-1)^{k+1}}{k}\,\alpha^k. \tag{28}$$

The following is a consequence of (28) and Corollary 5.

Proposition 6 *Given* $\alpha_w \in \mathbb{K}$ *for each* $w \in A^*$*, let* $\alpha : \mathcal{F} \to \mathbb{K}$ *be given by Definition 6. If* $\alpha_e = 0$*, then*

$$\exp\Big(\sum_{w\in A^*} \alpha_w\, E_w\Big) = \sum_{u\in\widehat{\mathcal{F}}} \exp\alpha(u)\, E(u). \tag{29}$$

If $\alpha_e = 1$*, then*

$$\log\Big(\sum_{w\in A^*} \alpha_w\, E_w\Big) = \sum_{u\in\widehat{\mathcal{F}}} \log\alpha(u)\, E(u). \tag{30}$$

In many applications, series of the form (4) that can be written as a series of elements of the Lie algebra generated by $\{E_a : a \in A\}$ have a special relevance. Such series can be defined as follows.

Definition 9 *We say that a Lie series is a series of the form*

$$\sum_{t\in\widehat{\mathcal{T}}} \alpha(t)\, E(t), \tag{31}$$

where $\alpha(t) \in \mathbb{K}$ *for each* $t \in \widehat{\mathcal{T}}$*.*

Notice that this definition of Lie series coincides with that given in [21] when $E_a = a$ ($a \in A$) and $\mathcal{B}$ is the $\mathbb{K}$-algebra $\mathbb{K}\langle A\rangle$ of $\mathbb{K}$-linear combinations of words on the alphabet A with the concatenation product of words.

The next proposition is a standard result in the Hopf-algebraic context of Section 7.

Proposition 7 *Given an arbitrary map $\alpha : \mathcal{F} \to \mathbb{K}$, there exists $\beta : \mathcal{F} \to \mathbb{K}$ such that $\alpha = \exp\beta$ and $\beta(u) = 0$ whenever $u \in \mathcal{F}\backslash\mathcal{T}$, if and only if*

$$\alpha(e) = 1, \quad \textit{and} \quad \alpha(t_1 \cdots t_m) = \alpha(t_1)\cdots\alpha(t_m) \quad \textit{for} \quad t_1, \ldots, t_m \in \mathcal{F}. \tag{32}$$

In particular, a series of the form (4) is the exponential of a Lie series if (32) holds for the map $\alpha : \mathcal{F} \to \mathbb{K}$ given by Definition 6.

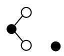
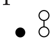
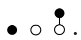
Remark 9 Lemma 15 in Subsection 5.1 below implies in particular that, if $\alpha : \mathcal{F} \to \mathbb{K}$ is given by Definition 6, then $\alpha(t^n) = n!\alpha(t^{\circ n})$ for each $t \in \mathcal{T}$. This implies that, if (32) holds, then for each $t \in \widehat{\mathcal{T}}$, $\exp(\alpha(t)E(t)) = I + \sum_{n\geq 1}\alpha(t^{\circ n})\, E(t)^n$, which shows that, in that case, the series (23) can be rewritten as an infinite directed product of exponentials of the form $\exp(\alpha(t)E(t))$, $t \in \widehat{\mathcal{T}}$.

We will give (under the assumptions of Proposition 7) explicit formulas for $\exp\beta(u)$ and $\log\alpha(u)$. First, we will fix some notation and give an auxiliary result.

Definition 10 *The factorial $u!$ of each forest $u \in \mathcal{F}$ is defined recursively as follows. For the empty forest, $e! = 1$. Given $a \in A$, $t_1, \ldots, t_m \in \mathcal{F}$, $v \in \mathcal{F}$, $t = [v]_a$, $u = t_1 \cdots t_m$,*

$$t! = v!|t|, \quad u! = t_1! \cdots t_m!.$$

Definition 11 *For each labeled partially ordered subset $V \subset U$ with vertices $\{x_1, \ldots, x_m\}$ ($|V| = m \leq |U|$), we consider two new labeled forests: The forest $v \in \mathcal{F}$ represented by V, and the forest $C^V(U)$ of labeled rooted trees obtained by removing from U all edges of the form ($y < x_i$), $1 \leq i \leq m$, $y \in U$. Note that if $r_1, \ldots, r_l$ are the roots of U, then $C^V(U) = C^{V\backslash\{r_1,\ldots,r_l\}}(U)$. If $V = \emptyset$, then $C^V(U) = u$.*

Example 7 Consider the labeled forest $u =$  over the alphabet $A = \{1, 2\}$. Let U be the labeled partially ordered set considered in Example 6. For the labeled partially ordered subset $V \subset U$ determined by the vertices $\{x_1, x_2, x_4\}$, we have that $v =$  and $C^V(U) =$ . □

Definition 12 *Given a labeled partially ordered set U representing a forest $u \in \mathcal{F}$, $K(U)$ will denote the family of labeled partially ordered subsets of U that include all the roots of U. For each $n \geq 1$, $K_n(U) = \{V \in K(U) : \ |V| = n\}$. The number of different total orderings of the set U that preserve the partial ordering in U will be denoted by $p(U)$. Given $n \geq 1$, $\omega_n(U)$ denotes the number of different ordered partitions $(U_1, \ldots, U_n)$ of the set of vertices of U satisfying that*

$$x \in U_i, \ y \in U_j, \ x < y \implies i < j. \tag{33}$$

Observe that, in the definition of $\omega_n(U)$, (33) excludes partitions $(U_1,\ldots,U_n)$ satisfying that $x > y$ for some $x, y \in U_i$ and some i. That is, for each i, the partially ordered subset of U determined by the vertices in U_i represents a forest of the form $\bullet^k$ (i.e., a forest with k vertices and no edges).

The next lemma follows from repeated application of Definition 7.

Lemma 8 *Consider a map $\alpha : \mathcal{F} \to \mathbb{K}$. Given $n \geq 1$ and $u \in \mathcal{F}$, let U be a labeled partially ordered set representing it.*

1. *If α is such that $\alpha(v) = 0$ whenever $v \in \mathcal{F}\backslash\mathcal{T}$, then*

$$\alpha^n(u) = \sum_{V \in K_n(U)} p(V)\alpha'(C^V(U)), \tag{34}$$

where we use the notation $\alpha'(v) = \alpha(t_1)\cdots\alpha(t_m)$ for arbitrary forests $v = t_1 \cdots t_m$, $(t_1,\ldots,t_m \in \mathcal{T})$.

2. *If α is such that $\alpha(e) = 0$ and $\alpha(t_1 \cdots t_m) = \alpha(t_1)\cdots\alpha(t_m)$ for arbitrary $t_1,\ldots,t_m \in \mathcal{T}$, then*

$$\alpha^n(u) = \sum_{V \in K_n(U)} \omega_n(V)\alpha(C^V(U)). \tag{35}$$

Theorem 9 *Let $\alpha : \mathcal{F} \to \mathbb{K}$ be a map such that $\alpha(u) = 0$ if $u \in \mathcal{F}\backslash\mathcal{T}$. Given $u \in \mathcal{F}$, let U be a labeled partially ordered set representing it. Then*

$$\exp\alpha(u) = \sum_{V \in K(U)} \frac{1}{v!}\alpha'(C^V(U)), \tag{36}$$

where v denotes the labeled rooted tree represented by V, $v!$ is given in Definition 10, and $\alpha'(w)$ $(w \in \mathcal{F})$ is interpreted as in Lemma 8.

Proof: By definition of $\exp\alpha$ (26) and Lemma 8 we have that

$$\exp\alpha(u) = \sum_{V \in K(U)} \frac{p(V)}{|V|!}\alpha'(C^V(U)). \tag{37}$$

Consider now $\phi : \mathcal{F} \to \mathbb{K}$ such that $\phi(u) = 1$ if $|u| = 1$ and $\phi(u) = 0$ otherwise. Then, (37) applied for $\alpha = \phi$ gives (as $\phi'(C^V(U)) = 1$ if $V = U$ and $\phi'(C^V(U)) = 0$ otherwise) that $\exp(\phi) = \theta$, where $\theta(u) = p(u)/|u|!$. We will prove that $|u|! = u!p(u)$ for each $u \in \mathcal{F}$, that is, $\theta(u) = 1/u!$. By Proposition 7, $\theta(e) = 1$ and $\theta(t_1 \cdots t_m) = \theta(t_1)\cdots\theta(t_m)$ for $t_1,\ldots,t_m \in \mathcal{T}$, and thus, according to Definition 10, the required result will follow if we prove that

$$|t|\,\theta(t) = \theta(u) \text{ whenever } t = [u]_a,\ u \in \mathcal{F},\ a \in A. \tag{38}$$

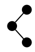
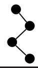
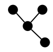
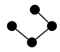
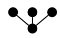
t								
$\|t\|$	1	2	3	3	4	4	4	4
$t!$	1	2	6	3	24	12	8	4
$\omega(t)$	1	$-\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$	$-\frac{1}{4}$	$-\frac{1}{6}$	$-\frac{1}{12}$	0

Table 1: Values of $t!$ and $\omega(t)$ for rooted trees of degree $|t| < 5$.

Consider the map $D\theta : \mathcal{F} \to \mathbb{K}$ given by $D\theta(u) = |u|\theta(u)$. It is straightforward to check that D can be considered as a derivation of maps $\beta : \mathcal{F} \to \mathbb{K}$ with respect to the product given in Definition 7, and thus $D\theta = D\exp\phi = (\exp\phi)\phi = \theta\phi$. This, together with $\theta\phi([u]_a) = \theta(u)$ for $u \in \mathcal{F}$ and $a \in A$ (by Definition 7 and the particular form of the map ϕ) finally leads to (38). □

The next result follows from the definition of $\log\alpha$ (27) and the second statement of Lemma 8.

Theorem 10 *Let $\alpha : \mathcal{F} \to \mathbb{K}$ be a map satisfying (32). Given $u \in \mathcal{F}$, let U be a labeled partially ordered set representing it. Then*

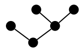
$$\log\alpha(u) = \sum_{V \in K(U)} \omega(v)\alpha(C^V(U)), \tag{39}$$

where v denotes the forest of labeled rooted trees represented by V, and

$$\omega(v) = \sum_{n=1}^{|v|} \frac{(-1)^{n+1}}{n}\,\omega_n(v). \tag{40}$$

Remark 10 Observe that, in the definition of $u!$, $\omega_n(u)$, and $\omega(u)$, the labelling of its vertices plays no role. Thus, $u!$, $\omega_n(u)$, and $\omega(u)$ are color-blind. We thus can consider that the values of $u!$ and $\omega(u)$ are defined for rooted trees, and whenever we write $u!$ (resp., $\omega(u)$) for a labeled rooted tree $u \in \mathcal{T}$, we refer to the value of ! (resp., ω) of the rooted tree obtained from u by forgetting its labeling. Notice that, by Proposition 7, one only needs to apply (36) and (39) for labeled rooted trees (as $\log\alpha(u) = 0$ if $u \in \mathcal{F}\backslash\mathcal{T}$, and $\exp\alpha(t_1\cdots t_m) = \exp\alpha(t_1)\cdots\exp\alpha(t_m)$). The values of $u!$ and $\omega(u)$ for rooted trees of degree $|u| < 5$ are displayed in Table 1.

Example 8 For each rooted tree t without ramifications ("tall rooted trees"), we have that $\omega(t) = (-1)^{|t|+1}/|t|$. This follows from (40) and the fact that, for such rooted trees, $t! = |t|!$, $\omega_{|t|}(t) = 1$, and $\omega_n(t) = 0$ if $n \neq |t|$.

As an additional example, consider the rooted tree $t =$  . Counting the number of ordered partitions $(U_1, \ldots, U_n)$ of the set of vertices of a labeled partially ordered set T representing t that satisfy condition (33), one gets that

$$\omega_1(t) = \omega_2(t) = 0, \quad \omega_3(t) = 2, \quad \omega_4(t) = 9, \quad \omega_5(t) = 8, \quad \omega(t) = \frac{2}{3} - \frac{9}{4} + \frac{8}{5} = \frac{1}{60}.$$

□

Remark 11 Let $\delta : \mathcal{F} \to \mathbb{K}$ be such that $\delta(e) = 1$, $\delta(t_1 \cdots t_m) = \delta(t_1) \cdots \delta(t_m)$ for $t_1, \ldots, t_m \in \mathcal{T}$, and given $t \in \mathcal{T}$, $\delta(t) = 1$ if $|t| = 1$ and $\delta(t) = 0$ otherwise. Then, (39) applied for $\alpha = \delta$ gives (as $\delta(C^V(U)) = 1$ if $V = U$ and $\delta(C^V(U)) = 0$ otherwise) that $\omega = \log \delta$ (in particular, according to Proposition 7, $\omega(u) = 0$ if $u \in \mathcal{F} \backslash \mathcal{T}$).

Thus, the values of $\omega(t)$ ($t \in \mathcal{T}$) can be obtained, as an alternative to (40), from solving for $\omega(t)$ in $\exp \omega = \delta$. Another alternative can be obtained from the expansion $z/(e^z - 1) = 1 + \sum_{k \geq 1} B_k/k!\, z^k$ (B_k being the Bernoulli numbers, in particular, $B_1 = -1/2$, $B_2 = 1/6$, and $B_{2k+1} = 0$ for $k \geq 1$), which gives $\omega = \delta + \sum_{k \geq 1} B_k/k!\, \omega^k \delta$, leading to

$$\omega([u]_a) = \sum_{k \geq 1} \frac{B_k}{k!}\, \omega^k(u), \;\; u \in \mathcal{F} \backslash \{e\}, \;\; a \in A.$$

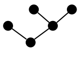
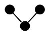
The first statement of Lemma 8 and the fact that $|v|! = v!\, p(v)$ for $v \in \mathcal{F}$ (proof of Theorem 9) then leads to

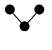
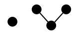
$$\omega([u]_a) = \sum_{V \in K(U)} \frac{B_{|v|}}{v!}\, \omega'(C^V(U)). \tag{41}$$

where U is a labeled partially ordered set representing $u \in \mathcal{F}$, and for each partially ordered set $V \in K(U)$, v is the forest that represents it, and $\omega'(t_1 \cdots t_m) = \omega(t_1) \cdots \omega(t_m)$. Observe that, since $B_k = 0$ for odd numbers $k > 1$, in (41) one only needs to consider $V \in K(U)$ with only one element (which is only possible if u is a rooted tree) and $V \in K(U)$ with an even number $|V|$ of elements.

Example 9 For rooted trees of the form $t = [\bullet^k]$, $k \geq 1$, ("bushy trees"), we have that $u = \bullet^k$ and $K(U) = \{U\}$, and hence

$$\omega([\bullet^k]) = B_k.$$

As an additional example of the application of (41), consider, as in Example 8, the rooted tree $t =$  , so that $u = \bullet$  . Then,

$\omega(t) = \frac{B_2}{\bullet^2!}\omega'(\bullet$  $) + \frac{B_4}{(\bullet \text{ } \ldots)!}$  $\omega'(\bullet^4) = B_2\, \omega($  $) + \frac{B_4}{3} = \frac{1}{36} + \frac{-1}{90} = \frac{1}{60}.$

□

Remark 12 In the literature of numerical analysis for ODEs [3, 16, 17], the factorial $t!$ of a rooted tree t is referred to as the density of t, and is denoted as $\gamma(t)$, while $\omega(t)$ ($t \in \mathcal{T}$) can be interpreted as the coefficient of the B-series corresponding to the modified equations of the explicit Euler method. The explicit formula (39) in Theorem 10 has been independently obtained in the context of numerical analysis of ODEs in [4], and additional formulas for $\omega(t)$, $t \in \mathcal{T}$ are given.

Example 10 Consider for each $a \in A$, the map $\phi_a : \mathcal{F} \to \mathbb{K}$ determined as follows. Given $u \in \mathcal{F}$, $\phi_a(u) = 1$ if $u = [e]_a$ and $\phi_a(u) = 0$ otherwise. It then trivially holds that $E_a = \sum_{u \in \widehat{\mathcal{F}}} \phi_a(u) E(u)$, and thus the BCH formula reads,

$$\exp(E_2)\exp(E_1) = \exp\left(\sum_{t \in \widehat{\mathcal{T}}} \log\alpha(t)\, E(t)\right),$$

where $A = \{1, 2\}$, and $\alpha = (\exp\phi_2)(\exp\phi_1)$. Recall that we represent as black vertices those labeled by 1, and those labeled by 2 as white vertices, The coefficients $\log\alpha(t)$ can be read from the structure of the labeled rooted tree t by application of Theorem 10 for the map α explicitly given as follows. Given $u \in \mathcal{T}$, let U be a labeled partially ordered set representing u. Let V (resp., W) be the labeled partially ordered subset of U determined by the white vertices of U (resp., black vertices). Then, V and W will represent two forests v and w, respectively. If $(V, W) \in R(U)$ (as given in Definition 7), then $\alpha(u) = 1/(v!w!)$, and $\alpha(u) = 0$ otherwise. This is a direct consequence of the fact that $\theta_a = \exp\phi_a$ is such that, given $u \in \mathcal{F}$, $\theta_a(u) = 1/u!$ if all the vertices of u are labeled by a, and $\theta_a(t) = 0$ otherwise. The latter statement can be proven in a very similar way to $\theta(t) := \exp\phi(t) = 1/t!$ in the proof of Theorem 9. Observe that $\alpha(t) = 0$ if t has some black vertex whose parent is white, so that is this case, one only needs to consider in (39) labeled partially ordered subsets of V including (in addition to all the roots of U) all the black vertices with white parents. For instance, for the Hall set of rooted trees in (18), we have that

$$\begin{aligned}\exp(E_2)\exp(E_1) \quad = \quad & \exp\Big(\beta(\bullet)E_1 + \beta(\circ)E_2 + \beta(\substack{\bullet\\ \circ})[E_1, E_2] \\ & +\beta(\substack{\bullet\\ \bullet\\ \circ})[E_1, [E_1, E_2]] + \beta(\bullet\!\!\substack{\\ \circ}\!\!\circ)[E_2, [E_1, E_2]] + \cdots\Big)\end{aligned}$$

(here, "$\cdots$" stands for terms of degree higher than three), where $\beta = \log\alpha$, and, in particular,

$$\begin{aligned}\beta(\bullet) &= \alpha(\bullet) = 1,\\ \beta(\circ) &= \alpha(\circ) = 1,\\ \beta(\substack{\bullet\\ \circ}) &= \omega(\substack{\bullet\\ \bullet})\alpha(\bullet)\alpha(\circ) = -\frac{1}{2},\\ \beta(\substack{\bullet\\ \bullet\\ \circ}) &= \omega(\substack{\bullet\\ \bullet})\alpha(\circ)\alpha(\substack{\bullet\\ \bullet}) + \omega(\substack{\bullet\\ \bullet\\ \bullet})\alpha(\bullet)^2\alpha(\circ) = -\frac{1}{2}\frac{1}{2} + \frac{1}{3} = \frac{1}{12},\\ \beta(\bullet\!\!\substack{\\ \circ}\!\!\circ) &= \omega(\substack{\bullet\\ \bullet})\alpha(\bullet)\alpha(\substack{\circ\\ \circ}) + \omega(\substack{\bullet\\ \bullet\\ \bullet})\alpha(\bullet)\alpha(\circ)^2 = -\frac{1}{2}\frac{1}{2} + \frac{1}{6} = -\frac{1}{12}.\end{aligned}$$

□

4.2 Iterated integrals and the continuous BCH formula

The proof of the next result will be postponed to Section 5 below.

Proposition 11 *Consider the particular case of (4)–(5) where each α_w is an iterated integral of the form (7), where $\mathbb{K}$ is an arbitrary commutative ring, $1_{\mathbb{K}}$ is the unity in $\mathbb{K}$, and $\int_a : \mathbb{K} \to \mathbb{K}$, $a \in A$, are $\mathbb{Z}$-linear endomorphisms satisfying*

$$(\textstyle\int_a \mu)(\int_b \lambda) = \int_a(\mu \int_b \lambda) + \int_b(\lambda \int_a \mu) \tag{42}$$

for each $a, b \in A$ and arbitrary $\mu, \lambda \in \mathbb{K}$ (the "integration by parts" property). The map $\alpha : \mathcal{F} \to \mathbb{K}$ given by Definition 6 for α_w in (7) satisfies

$$\alpha(e) = 1, \quad \alpha(t_1 \cdots t_m) = \alpha(t_1) \cdots \alpha(t_m), \quad \alpha([t_1 \cdots t_m]_a) = \int_a \alpha(t_1) \cdots \alpha(t_m) \tag{43}$$

whenever $t_1, \ldots, t_m \in \mathcal{T}$, $a \in A$. Such $\alpha : \mathcal{F} \to \mathbb{K}$ is uniquely determined by (43).

The assumptions of Proposition 11 hold, for instance in Examples 1–4.

Remark 13 By virtue of Proposition 7, the second equality in (43) implies that (4)–(7) can be written as the exponential of a Lie series. According to Remark 9, (44) can alternatively be expressed as an infinite directed product of exponentials of the form $\exp(\alpha(t)E(t))$, $t \in \widehat{\mathcal{T}}$, each coefficient $\alpha(t)$ being a generalized iterated integral that reflects the structure of the labeled rooted tree t (see [15] for this result in the context of non-linear control theory).

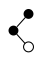
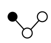
We finally have that, under the assumptions of Proposition 11,

$$\sum_{w \in A^*} \alpha_w \, E_w = \exp\Big(\sum_{t \in \widehat{\mathcal{T}}} \beta(t) \, E(t) \Big), \tag{44}$$

where $\beta = \log \alpha$ is explicitly given for $\alpha : \mathcal{F} \to \mathbb{K}$ in (43) by means of Theorem 10. That is, we have obtained an explicit expression of the so-called continuous BCH formula. Recall that, compared to the results in [22], our formula is written in a basis (an arbitrary Hall basis) of the free Lie algebra, while in [22], a spanning set of the free Lie algebra is used to express the formula (and thus requires the use of some rewriting algorithm in order to express it in a basis).

Example 11 Under the assumptions of Proposition 11, the application of Theorems 3 and 10 gives, for instance, for the Hall set of rooted trees in (18), that

$$\sum_{w \in A^*} \alpha_w \, E_w = \exp\Big(\beta(\bullet) E_1 + \beta(\circ) E_2 + \beta($$  $$)[E_1, E_2]$$

$$+\beta($$  $$)[E_1, [E_1, E_2]] + \beta($$  $$)[E_2, [E_1, E_2]] + \cdots \Big)$$

where $\beta = \log\alpha$. Using the notation $\alpha_a = \int_a 1_{\mathbb{K}}$ for $a \in A = \{1,2\}$, we have, in particular,

$$\begin{aligned}
\beta(\bullet) &= \alpha(\bullet) = \alpha_1, \\
\beta(\circ) &= \alpha(\circ) = \alpha_2, \\
\beta(\genfrac{}{}{0pt}{}{\bullet}{\circ}) &= \omega(\bullet)\alpha(\genfrac{}{}{0pt}{}{\bullet}{\circ}) + \omega(\genfrac{}{}{0pt}{}{\bullet}{\bullet})\alpha(\bullet)\alpha(\circ) \ = \ \int_2 \alpha_1 - \tfrac{1}{2}\,\alpha_1\alpha_2, \\
\beta(\bullet\!<\!\genfrac{}{}{0pt}{}{\bullet}{\circ}) &= \omega(\bullet)\alpha(\bullet\!<\!\genfrac{}{}{0pt}{}{\bullet}{\circ}) + \omega(\genfrac{}{}{0pt}{}{\bullet}{\bullet})(\alpha(\circ)\alpha(\genfrac{}{}{0pt}{}{\bullet}{\bullet}) + \alpha(\bullet)\alpha(\genfrac{}{}{0pt}{}{\bullet}{\circ})) + \omega(\bullet\!<\!\genfrac{}{}{0pt}{}{\bullet}{\bullet})\alpha(\bullet)^2\alpha(\circ) \\
&= \int_2\int_1 \alpha_1 - \tfrac{1}{2}(\alpha_2\int_1\alpha_1 + \alpha_1\int_2\alpha_1) + \tfrac{1}{3}\alpha_1^2\alpha_2, \\
\beta(\genfrac{}{}{0pt}{}{\bullet\;\circ}{\circ}) &= \omega(\bullet)\alpha(\genfrac{}{}{0pt}{}{\bullet\;\circ}{\circ}) + \omega(\genfrac{}{}{0pt}{}{\bullet}{\bullet})(\alpha(\bullet)\alpha(\genfrac{}{}{0pt}{}{\circ}{\circ}) + \alpha(\circ)\alpha(\genfrac{}{}{0pt}{}{\bullet}{\circ})) + \omega(\bullet\!<\!\genfrac{}{}{0pt}{}{\bullet}{\bullet})\alpha(\bullet)\alpha(\circ)^2 \\
&= \int_2 \alpha_1\alpha_2 - \tfrac{1}{2}(\alpha_1\int_2\alpha_2 + \alpha_2\int_2\alpha_1) + \tfrac{1}{6}\alpha_1\alpha_2^2.
\end{aligned}$$

□

5 Rewriting labeled rooted trees as linear combinations of Hall forests

5.1 A congruence on linear combinations of forests

Here we assume that the characteristic of the ring $\mathbb{K}$ is 0 (if the characteristic of $\mathbb{K}$ is $k \neq 0$, one should replace $\mathbb{Z}$ by $\mathbb{Z}/(k)$ in what follows).

Consider the ring $\mathbb{Z}[\mathcal{T}]$ of polynomials with integer coefficients in the commuting indeterminates $u \in \mathcal{T}$, where the products of labeled rooted trees (i.e. the monomials in $\mathbb{Z}[\mathcal{T}]$) are identified with forests $u \in \mathcal{F}$. The unity element is the empty forest, and the elements in $\mathbb{Z}[\mathcal{T}]$ are $\mathbb{Z}$-linear combinations of forests $u \in \mathcal{F}$. We extend by bilinearity the grafting operation $t \circ u$ for $\mathbb{Z}$-linear combinations of labeled rooted trees t and $\mathbb{Z}$-linear combinations of labeled forests u. We also extend arbitrary maps $\alpha : \mathcal{F} \to \mathbb{K}$ by linearity to $\mathbb{Z}[\mathcal{T}]$.

Definition 13 *Given $u, v \in \mathbb{Z}[\mathcal{T}]$, we say that u and v are* congruent*, and write $u \equiv v$, if for every map $\widehat{\alpha} : A^* \to \mathbb{K}$ that assigns $\widehat{\alpha}(w) = \alpha_w$ to each word w, it holds that $\alpha(u) = \alpha(v)$ for the map $\alpha : \mathcal{F} \to \mathbb{K}$ given by Definition 6.*

Proposition 12 *Given $m \geq 2$, $t_1, \ldots, t_m \in \mathcal{T}$,*

$$t_1 \cdots t_m \equiv \sum_{i=1}^{m} t_i \circ \prod_{j \neq i} t_j. \tag{45}$$

Proof: Given a labeled partially ordered set U representing a forest $u = t_1 \cdots t_m$, $t_1, \ldots, t_m$ with roots $r_1, \ldots, r_m$. Consider for each $i = 1, \ldots, m$ the labeled partially ordered set Z_i obtained from U by adding the edges $r_i < r_j$ for each $j \in \{1, \ldots, m\}\backslash\{i\}$ (thus, Z_i representing the labeled rooted tree $z_i = t_i \circ \prod_{j \neq i} t_j$). Then, it is straightforward to check that each total order $<_U$ of the set of vertices of U that extends the partial order in U extends the partial order of one and only one of the Z_i, which shows that $\alpha(t_1 \cdots t_m) = \alpha(z_1) + \cdots + \alpha(z_m)$ for arbitrary $\alpha : \mathcal{F} \to \mathbb{K}$ given by Definition 6 for some $\alpha_w \in \mathbb{K}$, $w \in A^*$. □

Proof of Proposition 11 Given a labeled partially ordered set T representing a labeled rooted tree $t = [u]_a$, $u \in \mathcal{F}$, with vertices $\{x_1, \ldots, x_n\}$ and root x_n labeled as $l(x_n) = a$, then the labeled partially ordered set $U \subset T$ determined by the set of vertices $\{x_1, \ldots, x_{n-1}\}$ (i.e., obtaining from T by removing the root x_n) represents the forest u. Now, for each total ordering $>_U$ of $\{x_1, \ldots, x_{n-1}\}$ that extends the partial order in U, there is clearly a unique total ordering $>_T$ of $\{x_1, \ldots, x_n\}$ that extends both the partial order in T and the total ordering $>_U$, and in that case, $w(>_T) = w(>_U)a$. This defines a one-to-one correspondence between the total order $>_T$ and $>_U$, and thus

$$\alpha([u]_a) = \sum_{>_T} \alpha_{w(>_T)} = \sum_{>_U} \alpha_{w(>_U)a} = \int_a \sum_{>_U} \alpha_{w(>_U)} = \int_a \alpha(u). \tag{46}$$

The "integration by parts" property (42) can be generalized (by applying induction on m) as

$$\prod_{i=1}^{m} \int_{a_i} \mu_i = \sum_{i=1}^{m} \int_{a_i} \left(\mu_i \prod_{j \neq i} \int_{a_j} \mu_j \right), \quad a_i \in A,\ \mu_i \in \mathbb{K},\ i = 1, \ldots, m. \tag{47}$$

The second equality in (43) can by proven by induction on the degree of $u = t_1 \cdots t_m$ (and Proposition 12) as follows. Let $t_i = [v_i]_{a_i}$, where $a_i \in A$ and $v_i \in \mathcal{F}$ for each $i = 1, \ldots, m$, then, by virtue of Proposition 12 and (46), we have that

$$\alpha(u) = \sum_{j=1}^{m} \alpha([v_i \prod_{j \neq i} t_j]_{a_i}) = \sum_{j=1}^{m} \int_{a_i} \alpha(v_i \prod_{j \neq i} t_j),$$

and by induction hypothesis and applying (46) again, we obtain that

$$\alpha(u) = \sum_{j=1}^{m} \int_{a_i} \alpha(v_i) \prod_{j \neq i} \alpha(t_j) = \sum_{j=1}^{m} \int_{a_i} \alpha(v_i) \prod_{j \neq i} \int_{a_j} \alpha(v_j),$$

and (47) with $\mu_i = \alpha(v_i)$ finally leads to $\alpha(u) = \alpha(t_1) \ldots \alpha(t_m)$.

Induction on the degree of forests shows that (43) uniquely determines the map $\alpha : \mathcal{F} \to \mathbb{K}$. □

Consider the $\mathbb{Z}$-module $\mathbb{Z}\langle A \rangle$ of $\mathbb{Z}$-linear combinations of words, and let us denote the shuffle product of $w_1, w_2 \in \mathbb{Z}\langle A \rangle$ [21] as $w_1 ⧢ w_2$.

Proposition 13 *Consider the $\mathbb{Z}$-linear map $\nu : \mathbb{Z}[\mathcal{T}] \to \mathbb{Z}\langle A \rangle$ given as follows. For each $u \in \mathcal{F}$,*

$$\nu(u) = \sum_{>_U} w(>_U) \tag{48}$$

where we adopt the same notation as in Definition 6. Then it holds that

$$\nu([u]_a) = \nu(u)a, \quad \nu(uv) = \nu(u) ⧢ \nu(v), \quad \textit{for each } u, v \in \mathcal{F},\ a \in A. \tag{49}$$

Proof: Consider $\mathbb{K} = \mathbb{Z}\langle A\rangle$, and $\int_a w = wa$ for each $w \in A^*$, $a \in A$. The recursive definition of the shuffle product [21] then just reads (42). Proposition 11 applied in this case gives the required result. $\square$

Remark 14 It obviously holds that a map $\alpha : \mathcal{F} \to \mathbb{K}$ is given by Definition 6 for some coefficients $\{\alpha_w : \ w \in A^*\}$ if and only if

$$\alpha(u) = \widehat{\alpha}(\nu(u)), \quad \text{for each } u \in \mathcal{F}, \tag{50}$$

where $\widehat{\alpha} : \mathbb{Z}\langle A\rangle \to \mathbb{K}$ is the $\mathbb{Z}$-linear map given by $\widehat{\alpha}(w) = \alpha_w$ for each $w \in A^*$. We thus have that $u \equiv v$ if and only if $u - v \in \ker \nu$ (i.e., u and v are congruent modulo $\ker \nu$).

According to Remark 14, Proposition 13 implies the following:

Proposition 14 *Given $u, v \in \mathbb{Z}[\mathcal{T}]$, $a \in A$, if $u \equiv v$, then $[u]_a \equiv [v]_a$. Furthermore, if $\bar{u} \equiv \bar{v}$ for $\bar{u}, \bar{v} \in \mathcal{F}$, then $u\bar{u} \equiv v\bar{v}$. Moreover, if under such conditions $t_1, \ldots, t_m \in \mathcal{T}$, then $t_1 \circ \cdots \circ t_m \circ u \equiv t_1 \circ \cdots \circ t_m \circ v$.*

The following result follows from Proposition 12 and from third statement of Proposition 14.

Lemma 15 *For arbitrary $t \in \mathcal{T}$, $n \geq 1$, it holds that $t^n \equiv n! t^{\circ n}$.*

Corollary 16 *For arbitrary $t \in \mathcal{T}$, $i, j \geq 1$, it holds that $t^{\circ i} t^{\circ j} \equiv \frac{(i+j)!}{i!j!} t^{\circ(i+j)}$.*

Remark 15 Repeated application of Corollary 16 shows that the product of arbitrary Hall forests is congruent to a Hall forest multiplied by a positive integer. In particular, this implies that, given $t \in \widehat{\mathcal{T}}$, $u \in \widehat{\mathcal{F}}$, and $n \geq 1$, there exist $\lambda \in \mathbb{Z}$ and $w \in \widehat{\mathcal{F}}$ such that $t^{\circ n} \circ u \equiv \lambda\, [w]_a$. More precisely, w is such that $v t^{\circ(n-1)} u \equiv \lambda w$, where $t = [v]_a$, $v \in \widehat{\mathcal{F}}$, $a \in A$.

5.2 A rewriting algorithm

We will give an algorithm that allows rewriting each $u \in \mathcal{F}$ as $u \equiv v$, where $v \in \mathbb{Z}\widehat{\mathcal{F}}$ (i.e., v is a $\mathbb{Z}$-linear combination of Hall forests), in a finite number of recursion steps. The main tool is an algorithm (Algorithm 1 below) that rewrites any labeled rooted tree of the form $[u]_a$ where $a \in A$ and $u \in \widehat{\mathcal{F}}$ as $[u]_a \equiv v$, where $v \in \mathbb{Z}\widehat{\mathcal{F}}$.

Algorithm 1 Given $t = [u]_a \in \mathcal{T}$, where $u \in \widehat{\mathcal{F}}$ and $a \in A$,

1. Find (following the constructive proof of Lemma 2) the Hall forest $v \in \widehat{\mathcal{F}}$ such that $\Gamma(v) = [u]_a$, that is, find $v = t_1^{\circ r_1} \cdots t_m^{\circ r_m}$ such that $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$, $m, r_1, \ldots, r_m \geq 1$, $t_1 > \cdots > t_m$, and $\Gamma(v) := t_1^{\circ r_1} \circ (t_2^{\circ r_2} \cdots t_m^{\circ r_m}) = [u]_a$.

2. If $m = 1$, then $t = t_1^{\text{o}r_1} \in \widehat{\mathcal{F}}$, and the algorithm stops. Otherwise, set

$$t \equiv t_1^{\text{o}r_1} \cdots t_m^{\text{o}r_m} - \sum_{i=2}^{m} t_i^{\text{o}r_i} \circ (t_1^{\text{o}r_1} \cdots t_{i-1}^{\text{o}r_{i-1}} t_{i+1}^{\text{o}r_{i+1}} \cdots t_m^{\text{o}r_m}), \tag{51}$$

rewrite, following Lemma 17 below, each term in the summation of the right-hand side of (51) in the form $\lambda_i \Gamma(v_i)$, $\lambda_i \in \mathbb{Z}$, $v_i \in \widehat{\mathcal{F}}$, and recursively go to the second step in Algorithm 1 (with v replaced by v_i) for each $i = 2, \ldots, m$.

Lemma 17 *Consider* $v = t_1^{\text{o}r_1} \cdots t_m^{\text{o}r_m} \in \widehat{\mathcal{F}}$, *where* $m > 1$, $t_1, \ldots, t_m \in \widehat{\mathcal{T}}$, $r_1, \ldots, r_m \geq 1$, $t_1 > \cdots > t_m$. *Given* $i \in \{2, \ldots, m\}$, *there exist* $\lambda_i \geq 1$, $v_i \in \widehat{\mathcal{F}}$, *such that* $t_i^{\text{o}r_i} \circ (t_1^{\text{o}r_1} \cdots t_{i-1}^{\text{o}r_{i-1}} t_{i+1}^{\text{o}r_{i+1}} \cdots t_m^{\text{o}r_m}) \equiv \lambda_i \Gamma(v_i)$ *and* $v \succ v_i$.

Proof: Given $i \in \{2, \ldots, m\}$, according to Remark 15, there exist $\lambda_i \geq 1$, $a_i \in A$, and $w_i \in \widehat{\mathcal{F}}$ such that $t_i^{\text{o}r_i} \circ (t_1^{\text{o}r_1} \cdots t_{i-1}^{\text{o}r_{i-1}} t_{i+1}^{\text{o}r_{i+1}} \cdots t_m^{\text{o}r_m}) \equiv \lambda_i [w_i]_{a_i}$. Consider $v_i = \Gamma^{-1}([w_i]_{a_i}) \in \widehat{\mathcal{F}}$, that is, the unique $v_i = z_1^{\text{o}l_1} \circ (z_2^{\text{o}l_2} \cdots z_k^{\text{o}l_k})$ with $l_1, \ldots, l_k \geq 1$, $z_1, \ldots, z_k \in \widehat{\mathcal{T}}$, and $z_1 > \cdots > z_k$ such that $[w_i]_{a_i} = z_1^{\text{o}l_1} \circ (z_2^{\text{o}l_2} \cdots z_k^{\text{o}l_k})$.

1. If $k = l_1 = 1$, that is, $v = z_1 \in \widehat{\mathcal{T}}$, then $z_1'' = \min(t_{m-1}, t_m'')$ provided that $i = m$ and $r_m = 1$, and $z_1'' = t_m$, otherwise, and hence $t_1 \geq z_1'' \geq t_m$, that is, $v \succ z_1 = v_i$.

2. Otherwise, from the proof of Lemma 2 it is clear that $z_k = t_m$ (and $l_k = r_m$) if $j \neq m$ of $r_m > 1$. If $j = m$ and $r_m = 1$, then $z_k = \min(t_m'', t_{m-1}) \geq t_m$. It then remains to prove that $z_1'' \leq t_1$. We will show that $z_1 = [t_1^{\text{o}r} v]_{a_i}$, where $r \geq 1$ and $v \in \widehat{\mathcal{F}}$, which implies that $z_1'' \leq t_1$.

   By definition of Hall rooted trees applied for $t_i \in \widehat{\mathcal{T}}$, there exist $a \in A$, $n_1, \ldots, n_p \geq 1$, and $s_1, \ldots, s_p \in \widehat{\mathcal{T}}$ such that $s_1 > \cdots > s_p > t_i = [s_1^{\text{o}n_1} \cdots s_p^{\text{o}n_p}]_{a_i}$.

   If $s_p > t_1$ (resp., $s_p = t_1$), then it implies, together with $t_i = [s_1^{\text{o}n_1} \cdots s_p^{\text{o}n_p}]_{a_i} < t_1$, that $[s_1^{\text{o}n_1} \cdots s_p^{\text{o}n_p} t_1^{\text{o}r_1}]_{a_i} \in \widehat{\mathcal{T}}$ (resp., $[s_1^{\text{o}n_1} \cdots s_{p-1}^{\text{o}n_{p-1}} t_1^{\text{o}(s_p+r_1)}]_{a_i} \in \widehat{\mathcal{T}}$), and the proof of Lemma 2 shows that z_1 is of the form $z_1 = [s_1^{\text{o}n_1} \cdots s_p^{\text{o}n_p} t_1^{\text{o}r_1} u]_{a_i}$ (resp., $z_1 = [s_1^{\text{o}n_1} \cdots s_{p-1}^{\text{o}n_{p-1}} t_1^{\text{o}(s_p+r_1)} u]_{a_i}$), where $u \in \widehat{\mathcal{F}}$.

   Otherwise, if $s_p < t_1$, let j be the smallest integer such that $s_j < t_1$. If z_1 is not of the form $z_1 = [t_1^{\text{o}r} v]_{a_i}$, $v \in \widehat{\mathcal{F}}$, then the proof of Lemma 2 shows that $z_1 = [s_1^{\text{o}n_1} \cdots s_{j-1}^{\text{o}n_{j-1}}]_{a_i} > z_2 = t_1$. But, the fact that $[s_1^{\text{o}n_1} \cdots s_j^{\text{o}n_j}]_{a_i} \in \widehat{\mathcal{T}}$ together with $s_j < t_1$ implies that $z_1 = [s_1^{\text{o}n_1} \cdots s_{j-1}^{\text{o}n_{j-1}}]_{a_i} < s_j < t_1 = z_2$, which leads to a contradiction.

□

Proposition 18 *Algorithm 1 when applied to an arbitrary labeled rooted tree of the form* $t = [u]_a$, $u \in \widehat{\mathcal{F}}$, $a \in A$, *rewrites* t *as* $t \equiv v$, *where* $v \in \mathbb{Z}\widehat{\mathcal{F}}$, *in a finite number of recursion steps. Moreover, such* v *is a* $\mathbb{Z}$*-linear combination of Hall forests* $v_i \in \widehat{\mathcal{F}}$ *with the same partial degrees as* t *and satisfying that* $\Gamma^{-1}(t) \succeq v_i$.

Proof: According to Proposition 12, if Algorithm 1 stops when applied for a labeled rooted tree of the form $t = [u]_a$, $u \in \widehat{\mathcal{F}}$, $a \in A$, then it succeeds in rewriting it as $t \equiv v \in \mathbb{Z}\widehat{\mathcal{F}}$. To see that the algorithm stops after a finite number of recursion steps, first notice that the set $\{z \in \mathcal{T} : \ |z|_a = |t|_a \text{ for each } a \in A\}$ of labeled rooted trees with the same partial degrees as an arbitrary $t \in \mathcal{T}$ is always finite. And, finally, observe that each of the labeled rooted trees $\Gamma(v_i)$ in Algorithm 1 has the same partial degrees as t, and is, according to Lemma 17, such that $\Gamma^{-1}(t) \succ v_i$, where the partial order on Hall forests $\succ$ is given in Definition 4. □

Corollary 19 *Under the assumptions of Theorem 3, for each word $w \in A^n$, there exists $v \in \mathbb{Z}\widehat{\mathcal{F}}$ of homogeneous degree n such that $\alpha_w = \alpha(v)$.*

Proof: According to Remark 14, one only needs to show that $w = \nu(v)$ for some $v \in \mathbb{Z}\widehat{\mathcal{F}}$. This follows applying induction on the degree $|w|$ from Propositions 18 and 14. It is trivial for $w = a \in A$, and given $w \in A$, by induction hypothesis and Proposition 14 we have that $wa = \nu(v)a = \nu([v]_a)$ and the required result follows by applying Algorithm 1 to each $[v_i]_a$, where $v = \sum \lambda_i v_i$, $v_i \in \widehat{\mathcal{F}}$. □

Remark 16 The proof of Corollary 19 gives a recursive algorithm to rewrite each word w as $w = \nu(v)$ where $v \in \mathbb{Z}\widehat{\mathcal{F}}$. This together with (48) allows rewriting an arbitrary forest $u \in \mathcal{F}$ as $u \equiv v$, $v \in \mathbb{Z}\widehat{\mathcal{F}}$. However, this can be done more efficiently, directly using Algorithm 1 and the results in Proposition 14. Application of the second statement in Proposition 14 allows rewriting any forest $u = t_1 \cdots t_m$ provided that each $t_1, \ldots, t_m \in \mathcal{T}$ has been previously rewritten. For labeled rooted trees $t \in \mathcal{T}$, one can either use recursion and the first two statements in Proposition 14, or directly use the third statement in Proposition 14 combined with Algorithm 1 to reduce in step-by-step manner each subtree of t of the form $[u]_a$, $u \in \widehat{\mathcal{F}}$ (i.e., such that $t = t_1 \circ \cdots \circ t_m \circ [u]_a$).

Remark 17 In [20], a different definition of Hall sets of labeled rooted trees is given. Such a set of labeled rooted trees can be obtained from a Hall set $\widehat{\mathcal{T}}$ (as given in Definition 1) as $\widehat{\mathcal{T}}^* = \{t^* \in \mathcal{T} : \ t \in \widehat{\mathcal{T}}\}$, where for each Hall forest $u = t_1^{\circ r_1} \cdots t_m^{\circ r_m} \in \widehat{\mathcal{F}}$ and for the Hall rooted tree $t = [u]_a \in \widehat{\mathcal{T}}$, u^* and t^* are given recursively as $u^* = (t_1^*)^{r_1} \cdots (t_m^*)^{r_m}$ and $t^* = [u^*]_a$ (with $e^* = e$). It follows from Lemma 15 that $u^* = \sigma(u^*)u$. Hence, Corollary 19 also holds with $\mathbb{Z}\widehat{\mathcal{F}}$ replaced by $\mathbb{Q}[\widehat{\mathcal{T}}^*]$. An equivalent version of Algorithm 1 can be easily obtained with $\widehat{\mathcal{F}}^*$ instead of $\widehat{\mathcal{F}}$, which is slightly simpler in some sense (instead of Corollary 16 we just need the trivial $t^i t^j = t^{i+j}$), but it has the drawback of requiring the use of rational numbers.

5.3 Application to computations in the Lie algebra of Lie series

Proposition 20 *Under the assumptions of Proposition 4, if $\alpha(u) = 0$ for $u \in \mathcal{F}\backslash\mathcal{T}$, then*

$$\Big(\sum_{t \in \widehat{\mathcal{T}}} \alpha(t)\, E(t)\Big)\Big(\sum_{u \in \widehat{\mathcal{F}}} \beta(u)\, E(u)\Big) = \sum_{u \in \widehat{\mathcal{F}}} \alpha\beta(u)\, E(u), \tag{52}$$

where given $u \in \mathcal{F}$ and a labeled partially ordered set U representing u,

$$\alpha\beta(u) = \sum_{(Z,V)\in R_\tau(U)} \alpha(Z)\,\beta(V), \tag{53}$$

where $R_\tau(U)$ is the set of pairs $(Z, V) \in R(U)$ such that the labeled partially ordered set Z represents a Hall rooted tree $z \in \widehat{\mathcal{T}}$.

Proof: Corollary 5 implies that (52)–(53) hold with $R_\tau(U)$ the set of pairs $(Z, V) \in R(U)$ such that the labeled partially ordered set Z represents a labeled rooted tree $z \in \widehat{\mathcal{T}}$. With such a definition of $R_\tau(U)$, there is one pair $(Z, V) \in R_\tau(U)$ per edge $(x < y)$ of U, where Z and V are the two connected components of the labeled partially ordered set obtained from U by removing the edge $(x < y)$. The definition of Hall rooted trees shows that such Z can only represent labeled rooted trees of the form $z^{\circ r}$ where $z \in \widehat{\mathcal{T}}$ and $r \geq 1$. But, by assumption, $\alpha(z^{\circ r}) = 0$ if $r > 1$. □

The following result gives formulas to compute the Lie bracket of two Lie series (as given in Definition 9).

Corollary 21 *Under the assumptions of Proposition 4, if $\alpha(u) = \beta(u) = 0$ for $u \in \mathcal{F}\backslash\mathcal{T}$ (i.e., $\alpha(u) = \beta(u) = 0$ for $u \in \widehat{\mathcal{F}}\backslash\widehat{\mathcal{T}}$), then*

$$\Big[\sum_{t\in\widehat{\mathcal{T}}} \alpha(t)\,E(t), \sum_{t\in\widehat{\mathcal{T}}} \beta(t)\,E(t)\Big] = \sum_{t\in\widehat{\mathcal{T}}} [\alpha,\beta](t)\,E(t), \tag{54}$$

where given $t \in \mathcal{T}$ and a labeled partially ordered set T representing t,

$$[\alpha,\beta](t) = \sum_{(Z,S)\in R_\tau(T)} \Big(\alpha(Z)\,\beta(S) - \beta(Z)\,\alpha(S)\Big). \tag{55}$$

When computing the Lie bracket (54) of two Lie series, one only needs to compute $[\alpha, \beta](t)$ for $t \in \widehat{\mathcal{T}}$, but S in the summation of (55) does not necessarily represent a labeled rooted tree in $\widehat{\mathcal{T}}$, and thus one will need to apply some rewriting algorithm (based, for instance, on Algorithm 1) unless the values of $\alpha(s), \beta(s)$ for arbitrary $s \in \mathcal{T}$ are previously known.

Example 12 For instance, for the Hall set of rooted trees in (18), we have that $[\alpha, \beta](\bullet) = [\alpha, \beta](\circ) = 0$, $[\alpha, \beta](\overset{\bullet}{\circ}) = \alpha(\bullet)\beta(\circ) - \beta(\bullet)\alpha(\circ)$, $[\alpha, \beta](\overset{\bullet}{\overset{\bullet}{\circ}}) = \alpha(\bullet)\beta(\overset{\bullet}{\circ}) - \beta(\bullet)\alpha(\overset{\bullet}{\circ})$, and $[\alpha, \beta](\bullet\!\!\searrow\!\circ\!\swarrow\!\!\circ) = \alpha(\circ)\beta(\overset{\bullet}{\circ}) - \beta(\circ)\alpha(\overset{\bullet}{\circ})$ (recall that $\overset{\bullet}{\bullet}, \overset{\circ}{\circ} \in \widehat{\mathcal{F}}\backslash\widehat{\mathcal{T}}$, and thus $\alpha(u) = \beta(u) = 0$ for $u = \overset{\bullet}{\bullet}, \overset{\circ}{\circ}$). □

Remark 18 Formulas (54)–(55) can be used to do computations in an arbitrary Lie algebra $\mathfrak{g}$ as follows. Let the Lie algebra $\mathfrak{g}$ be generated by the set $\{E_a : a \in A\} \subset \mathfrak{g}$ (A is a set of indices). Consider a Hall set $\widehat{\mathcal{T}}$ of rooted trees labeled by A. Recursively define $E(t) \in \mathfrak{g}$, $t \in \widehat{\mathcal{T}}$, by (20). Then any element of $\mathfrak{g}$ can be represented by a Lie series (31) (with a finite number of non-vanishing coefficients $\alpha(t)$), and the Lie bracket of two arbitrary elements of $\mathfrak{g}$ can be computed using (54)–(55). This is a direct consequence of previous results applied with $\mathcal{B}$ being the enveloping universal algebra of the Lie algebra $\mathfrak{g}$.

5.4 The proof of Theorem 3

In order to prove Theorem 3 it is clearly sufficient to show that the result holds in the case $\mathbb{K} = \mathbb{Z}\langle A \rangle$ (with the shuffle product) and $\alpha_w = w$ for each $w \in A^*$. According to Proposition 13, Remark 6, and Lemma 15, Theorem 5.3 in [21] is equivalent to the required result.

For completeness, we next outline an alternative proof that uses the results in the preceding two subsections.

Proposition 22 *Under the assumptions of Proposition 20, for each $u \in \widehat{\mathcal{F}}$ it holds that $\alpha\beta(u) - \alpha(u'')\beta(u')$ is a $\mathbb{Z}$-linear combination of terms of the form $\alpha(z)\beta(v)$, such that $z \in \widehat{\mathcal{T}}$, $v \in \widehat{\mathcal{F}}$, and $z > v''$.*

Proof: Given a labeled partially ordered set T representing a labeled rooted tree $t \in \mathcal{T}$, each pair $(Z, S) \in R(T)$ such that Z and S represent two labeled rooted trees z and s, is associated to a unique decomposition of T of the form $T = T_m \circ \cdots \circ T_1 \circ T_0$ ($m \geq 1$, and each T_i representing a labeled rooted tree $t_i \in \mathcal{T}$) with $Z = T_0$ and $S = T_m \circ \cdots \circ T_1$, which is a one-to-one correspondence. Equivalently, there is a one-to-one correspondence between such decompositions of T and the edges $(x < y)$ of T (where x and y are the roots of T_1 and T_0, respectively). Thus, each edge of T is associated to a decomposition of t of the form

$$t = t_m \circ \cdots \circ t_0, \tag{56}$$

In general, different edges of T can give the same decomposition (56). However, if $\sigma(t) = 1$, and in particular, if $t \in \widehat{\mathcal{T}}$, then there is a one-to-one correspondence between the pairs $(Z, S) \in R(T)$ such that Z and S represent two labeled rooted trees and the different decompositions (56) of t. That correspondence associates each decomposition (56) to a pair pair $(Z, S) \in R(T)$ such that Z and S represent the labeled rooted trees t_0 and $t_m \circ \cdots \circ t_1$ respectively. By definition of Hall rooted trees, if $t \in \widehat{\mathcal{T}}$, then t_0 is of the form $t_0 = z^{\circ r}$, $r \geq 1$, $z \in \widehat{\mathcal{T}}$. Exactly the same argument holds with $t \in \widehat{\mathcal{T}}$ replaced by $t^{\circ n}$, where $t \in \widehat{\mathcal{T}}$ and $n \geq 1$. Thus, according to Proposition 20, given $t \in \widehat{\mathcal{T}}$, $n \geq 1$, $\alpha\beta(t^{\circ n})$ is the sum, over all decompositions

$$t^{\circ n} = t_m \circ \cdots \circ t_1 \circ z^{\circ r},$$

with $r \geq 1$, $z \in \widehat{\mathcal{T}}$, $t_1, \ldots, t_m \in \mathcal{T}$, of $\alpha(z)\beta(t_m \circ \cdots \circ t_1 \circ z^{\circ(r-1)})$. There is a unique such decomposition of $t^{\circ n}$ such that $m = 1$ and $(t_1 \circ z^{\circ(r-1)}, z)$ is the standard decomposition of $t^{\circ n}$. It is not difficult to check (using the properties of Hall rooted trees) that the remaining decompositions are such that $t^{\circ n} = t_m \circ \cdots \circ t_1 \circ z^{\circ(r-1)}$ is of the form

$$(z_m^{\circ r_m} \circ u_m) \circ \cdots \circ (z_1^{\circ r_1} \circ u_1), \quad \text{where} \quad z > \max(z_1 \cdots z_m u_1 \cdots u_m) \tag{57}$$

$r_1, \ldots, r_m \geq 1$, $z_1, \ldots, z_m \in \widehat{\mathcal{T}}$, and $u_1, \ldots, u_m \in \widehat{\mathcal{F}}$. Here, we have used the notation adopted in Remark 4 for the maximum Hall rooted tree $\max(u)$ of a product u of Hall forests. If $m > 1$, the third statement in Proposition 14, Proposition 18, Lemma 17, and Remark 4 show that (57) is congruent to a linear combination of labeled rooted trees of the form (57) with m replaced by $m-1$. If $m = 1$, Proposition 18 and Lemma 17 show that (57) is congruent to a linear combination of Hall forests $v \in \widehat{\mathcal{F}}$ such that $z_1 u_1 \succ v$, and, in particular, $z > v''$ (as $z > \max(z_1 u_1)$). We have thus proven the statement of Proposition 22 when $u = t^{\circ n}$, $t \in \widehat{\mathcal{T}}$, and $n \geq 1$. The required result for general Hall forests $u \in \widehat{\mathcal{F}}$ follows from the next statement. Given a labeled partially ordered set U representing a forest $u = t_1 \cdots t_m \in \mathcal{F}$, $t_1, \ldots, t_k \in \mathcal{T}$, if $T_1, \ldots, T_m$ (resp., $U_1, \ldots, U_m$) are labeled partially ordered subsets of U such that T_i represents the labeled rooted tree t_i (resp., U_i represents the labeled forest $t_1 \cdots t_{i-1} t_{i+1} \cdots t_m$), then $(Z, V) \in R(U)$ is and only if either $Z = T_i$ and $V = U_i$, or V is the direct union of U_i and S, where $(Z, S) \in R(T_i)$. □

Proof of Theorem 3: According to Corollary 19, for each $u \in \widehat{\mathcal{F}}$, there exist $\alpha_w^{(u)} \in \mathbb{K}$, $w \in A^*$, such that the map $\alpha^{(u)} : \mathcal{F} \to \mathbb{K}$ given by Definition 6 is such that, for each $v \in \widehat{\mathcal{F}}$, $\alpha^{(u)}(v) = 1$ if $v = u$ and $\alpha^{(u)}(v) = 0$ otherwise. Corollary 19 implies that (23) and (25) hold with some $E(u) \in \mathcal{B}$ (certain $\mathbb{K}$-linear combination of E_w for words w of degree $|u|$). It then remains to show that such $E(u)$ ($u \in \widehat{\mathcal{F}}$) satisfy (20)–(21), which is clearly equivalent to

$$\alpha^{(t)} = \alpha^{(t'')}\alpha^{(t')} - \alpha^{(t')}\alpha^{(t'')}, \quad \alpha^{(u)} = \alpha^{(u'')}\alpha^{(u')} \quad \text{where} \quad t \in \widehat{\mathcal{T}},\ u \in \widehat{\mathcal{F}}, \tag{58}$$

and $(t', t'') \in \widehat{\mathcal{T}} \times \widehat{\mathcal{T}}$ and $(u', u'') \in \widehat{\mathcal{F}} \times \widehat{\mathcal{T}}$ are the standard decompositions of t and u, respectively. Finally, (58) is implied by the following. According to Proposition 22, given $u, v \in \widehat{\mathcal{F}}$ and $t \in \widehat{\mathcal{T}}$, if $\alpha^{(u'')}\alpha^{(u')}(v) - \alpha^{(u'')}(v'')\alpha^{(u')}(v') \neq 0$, then $u'' > (u')''$ (which is false by definition of standard decomposition), and, similarly, $\alpha^{(t')}\alpha^{(t'')}(v) \neq 0$ implies that $t' > (t'')''$ (which is again false). □

6 Interpretation in terms of vector fields

The map $\alpha : \mathcal{F} \to \mathbb{K}$ in Definition 6, which has a central role in the present work, lacks, in the context of Sections 3–5, a direct interpretation of the coefficients $\alpha(u)$ for general $u \in \mathcal{F}$.

Assume that $\{E_a : \ a \in A\}$ are smooth vector fields (viewed as linear differential operators) over a d-dimensional smooth manifold $\mathcal{M}$. Let $\mathcal{B}$ be the vector space over $\mathbb{R}$ of linear differential operators acting on smooth functions on $\mathcal{M}$ spanned by the identity operator I and the set $\{E_{a_1} \cdots E_{a_m} : \ m \geq 1,\ a_1, \ldots, a_m \in A\}$. Here the expression $E_{a_1} \cdots E_{a_m}$ represents a composition of operators (so that the action of $E_{a_1} \cdots E_{a_m}$ on a smooth function $g \in C^\infty(\mathcal{M})$ is recursively given by $E_{a_1} \cdots E_{a_m} g := E_{a_1} \cdots E_{a_{m-1}}(E_{a_m} g)$). Thus $\mathcal{B}$ is generated as an algebra by the vector fields $\{E_a : \ a \in A\}$). Obviously, all the results shown in the preceding sections are valid in this case with $\mathbb{K} = \mathbb{R}$. In the present section, we will consider $\mathbb{K} = \mathbb{R}$.

Theorems 24 and 25 below give to the map $\alpha : \mathcal{F} \to \mathbb{K}$ associated in Definition 6 to each series of the form (4), a more concrete meaning. The product of arbitrary maps from $\mathcal{F}$ to $\mathbb{K}$, that we have so far applied only to maps associated to a series of the form $\sum_{w\in A^*} \alpha_w E_w$, can now be interpreted with Theorem 24 for general maps $\alpha : \mathcal{F} \to \mathbb{K}$.

The results in the present section are closely related to the results in [12] and references therein.

In local coordinates $(x_1, \ldots, x_d)$, defined in an open set $\mathcal{U} \subset \mathcal{M}$, the smooth vector fields E_a can be written as

$$E_a = \sum_{i=1}^{d} f_a^i \, D_i \quad \text{for each} \quad a \in A, \tag{59}$$

with suitable smooth functions $f_a^i : \mathcal{U} \subset \mathcal{M} \to \mathbb{K}$ and $D_i = \partial/\partial x_i$, $i = 1, \ldots, d$. In what follows, $C^\infty(\mathcal{U})$ denotes the set of smooth functions on $\mathcal{U} \subset \mathcal{M}$.

Definition 14 *A linear differential operator $X(u)$ acting on $C^\infty(\mathcal{U})$ is assigned to each forest $u \in \mathcal{F}$ of rooted trees labeled by A as follows. For the empty forest, $X(e) = I$, and*

$$X(t_1 \cdots t_m) \;=\; \sum_{i_1,\ldots,i_m=1}^{d} F(t_1)^{i_1} \cdots F(t_m)^{i_m} \, D_{i_1} \cdots D_{i_m}, \quad \text{if} \quad t_1, \ldots, t_m \in \mathcal{T}, \tag{60}$$

where, for each $t \in \mathcal{T}$ and each $i \in \{1, \ldots, m\}$, $F(t)^i$ is a smooth function on $\mathcal{U} \subset \mathcal{M}$ given by

$$F(t)^i \;=\; X(u) f_a^i \quad \text{if} \quad t = [u]_a, \; u \in \mathcal{F}, \; a \in A. \tag{61}$$

Theorem 24 below, which gives a rule to perform the product of series of linear differential operators of the form

$$\sum_{u\in\mathcal{F}} \frac{\alpha(u)}{\sigma(u)} X(u), \tag{62}$$

can be proven in an indirect way using some results available in the literature (see Remark 23 in Section 7 below). We give for completeness a direct proof. We first need the following result, due to Grossman and Larson [12, Proposition 2].

Lemma 23 *Given two labeled partially ordered sets V and W representing, respectively, the labeled forests $v, w \in \mathcal{F}$ and having disjoint sets of vertices, it holds that*

$$X(v)X(u) = \sum_{W} X(W)$$

where the summation goes over all labeled partially ordered sets W representing labeled forests such that $(V, U) \in R(W)$, and $X(W) = X(w)$ if W represents $w \in \mathcal{F}$.

Proof: We first observe that, given a forest $u \in \mathcal{F}_n$ $(n \geq 1)$ represented by a labeled partially ordered set U with vertices $1, \ldots, n$ labeled by $a_1, \ldots, a_m \in A$, respectively, and roots $\{1, \ldots, k\}$ $(k \leq n)$, for each $g \in C^\infty(\mathcal{U})$,

$$X(u)g = \sum_{i_1,\ldots,i_n=1}^{d} \left(\prod_{j=1}^{n} D_{i_{\chi(j,1)}} \cdots D_{i_{\chi(j,n_j)}} f_{a_j}^{i_j} \right) D_{i_1} \cdots D_{i_k} g, \tag{63}$$

where for each $j \in \{1, \ldots, n\}$, $\{\chi(j,1), \ldots, \chi(j,n_j)\}$ is the set of children of the vertex j in U. Notice that $X(u)g$ is a summation of products of terms having one factor per vertex of U plus one more for a partial derivative of g. This shows that, given $v = t_1 \ldots, t_m$, $t_1, \ldots, t_m \in \mathcal{T}$, and $u \in \mathcal{F}$, $X(v)X(u)g$ can be expanded as a sum of $(|u|+1)^m$ terms obtained from (63) by letting each $\sum_{l=1}^{d} F(t_r)^l D_l$ $(r = 1, \ldots, m)$ act on either the factor in (63) corresponding to one vertex of U or on $D_{i_1} \cdots D_{i_k} g$. But each of such terms is of the form $X(W)g$, with one term per labeled partially ordered set W satisfying that the set of vertices of W is (following the notation in the statement of Lemma 23) the union of the sets of vertices of U and V, and the set of edges of W is obtained as the union of the edges of U and V and possibly some edges of the form $(x < r)$, $x \in U$, $r \in \text{roots}(V)$, with at most one such edge per root of V. Or, equivalently, $X(v)X(u)g$ is the sum of $X(W)g$ over all labeled partially ordered sets W satisfying that $(V,U) \in R(W)$. □

Theorem 24 *Given two arbitrary maps $\alpha, \beta : \mathcal{F} \to \mathbb{K}$, it formally holds that*

$$\left(\sum_{u \in \mathcal{F}} \frac{\alpha(u)}{\sigma(u)} X(u) \right) \left(\sum_{u \in \mathcal{F}} \frac{\beta(u)}{\sigma(u)} X(u) \right) = \sum_{u \in \mathcal{F}} \frac{\alpha\beta(u)}{\sigma(u)} X(u), \tag{64}$$

where $\alpha\beta : \mathcal{F} \to \mathbb{K}$ is given in Definition 7.

Proof: It is clearly sufficient to prove the statement for α and β defined, for arbitrary $u, v \in \mathcal{F}$, as follows. Given $w \in \mathcal{F}$, $\alpha(w) = \sigma(v)$ if $w = v$ and $\alpha(w) = 0$ otherwise, and $\beta(w) = \sigma(u)$ if $w = u$ and $\beta(w) = 0$ otherwise. In that case, the left-hand side of (64) is $X(v)X(u)$. By Lemma 23, we have that

$$X(v)X(u) = \sum_{w \in \mathcal{F}} \frac{\gamma(u)}{\sigma(w)} X(w),$$

where $\gamma(w) = k(w,v,u)\,\sigma(w)$ for each $w \in \mathcal{F}$, where $k(w,v,u)$ is a non-negative integer determined as follows. Given V, U (with disjoint sets of vertices) representing v, u, respectively, $k(w,v,u)$ is the number of different labeled partially ordered sets W representing w such that $(V,U) \in R(W)$. We thus need to prove that $\alpha\beta = \gamma$ for these particular α and β associated to v and u, respectively. By Definition 7 we have that $\gamma(w) = l(w,v,u)\,\sigma(v)\sigma(u)$, where given a labeled partially ordered set W representing w, $l(w,v,u)$ is the number of different pairs $(V,W) \in R(W)$ that represents, respectively, v and u. It thus remains to show that $k(w,v,u)\sigma(w) = l(w,v,u)\sigma(v)\sigma(u)$, which follows from the following observation.

If $(V,U) \in R(W)$ for some labeled partially ordered sets V, U, W representing, respectively, v, u, w, then $k(w,v,u) = \sigma(v)\sigma(u)/\sigma(W,V,U)$ and $l(w,v,u) = \sigma(w)/\sigma(W,V,U)$, and $k(w,v,u) = l(w,v,u) = 0$ otherwise, where $\sigma(W,V,U)$ denotes the number of different pairs of permutations of the vertices of V and U, respectively, which determine an isomorphism of the labeled partially ordered set W. □

Theorem 25 *Given a map that assigns $\alpha_w \in \mathbb{K}$ to each word $w \in A^*$. Let $\alpha : \mathcal{F} \to \mathbb{K}$ be given by Definition 6, then*

$$\sum_{w \in A^*} \alpha_w \, E_w = \sum_{u \in \mathcal{F}} \frac{\alpha(u)}{\sigma(u)} X(u). \tag{65}$$

Proof: It is clearly sufficient to prove the statement with $\alpha_w = \alpha_w^{(a_1 \cdots a_m)}$ given for arbitrarily fixed $a_1, \ldots, a_m \in A$ as follows: $\alpha_w^{(a_1 \cdots a_m)} = 1$ if $w = a_1 \cdots a_m$ and $\alpha_w^{(a_1 \cdots a_m)} = 0$ otherwise. We will prove by induction on m that

$$\sum_{w \in A^*} \alpha_w^{(a_1 \cdots a_m)} \, E_w = \sum_{u \in \mathcal{F}} \frac{\alpha^{(a_1 \cdots a_m)}(u)}{\sigma(u)} X(u) \tag{66}$$

holds for $\alpha^{(a_1 \cdots a_m)} : \mathcal{F} \to \mathbb{K}$ given by Definition 6.

For $m = 1$, the map $\alpha^{(a_1)} : \mathcal{F} \to \mathbb{K}$ given by Definition 6 is trivially such that, for $u \in \mathcal{F}$, $\alpha^{(a_1)}(u) = 1$ if $u = [e]_{a_1}$ and $\alpha^{(a_1)}(u) = 0$ otherwise. But, by definition, $E_a = X([e]_a)$, which leads to (66) for $m = 1$.

For $m > 1$, by induction hypothesis and Theorem 24, we have that (66) holds for $\alpha^{(a_1 \cdots a_m)} : \mathcal{F} \to \mathbb{K}$ given by Definition 7 as the product of $\alpha = \alpha^{(a_1 \cdots a_{m-1})}$ and $\beta = \alpha^{(a_m)}$. On the other hand, it is trivially checked that $\alpha_w^{(a_1 \cdots a_m)} = (\alpha\beta)_w$ in (6) for $\alpha = \alpha_w^{(a_1 \cdots a_{m-1})}$ and $\beta_w = \alpha^{(a_m)}$. The required result then follows from Corollary 5. □

Remark 19 It is straightforward to check that Lemma 23 and Theorems 24 and 25 still hold (with exactly the same proofs) in the general situation where $\mathbb{K}$ is an arbitrary commutative ring with unit, $C^\infty(\mathcal{U})$ is replaced by an arbitrary commutative $\mathbb{K}$-algebra $\mathcal{R}$, and for each $a \in A$ and each $i = 1, \ldots, d$, $f_a^i \in \mathcal{R}$, and E_a and D_i are $\mathbb{K}$-linear derivations of $\mathcal{R}$ such that $D_1, \ldots, D_d$ commute with each other, and (59) holds.

Remark 20 Theorem 25 gives, in particular, an explicit description of the map ϕ in [12, Proposition 3].

Remark 21 Theorem 30 in Section 7 below implies that, if $\alpha : \mathcal{F} \to \mathbb{K}$ is such that $\alpha(e) = 1$ and

$$\alpha(t_1 \cdots t_m) = \alpha(t_1) \cdots \alpha(t_m) = \sum_{i=1}^{m} \alpha(t_i \circ \prod_{j \neq i} t_j) \tag{67}$$

for each $m \geq 1$, $t_1, \ldots, t_m \in \mathcal{T}$, then the series of differential operators (62) is the exponential of a series of vector fields in the Lie algebra generated by the basic vector fields $\{E_a : \ a \in A\}$, more precisely,

$$\sum_{u \in \mathcal{F}} \frac{\alpha(u)}{\sigma(u)} X(u) = \exp\left(\sum_{t \in \widehat{\mathcal{T}}} \log \alpha(t)\, E(t)\right).$$

Remark 22 Series of differential operators of the form (62) (typically with the presence of a factor of the form $h^{|u|}$ in each term, associated to the time-discretization parameter h, which can be incorporated in the definition of the functions f_a^i) are particularly useful in the theoretical study of large classes of numerical integrators for ODEs [19, 20, 5]. In the case of a map $\alpha : \mathcal{F} \to \mathbb{K}$ satisfying (32), the series (62) is equivalent to the more standard so-called B-series associated to a numerical integrator [16, 3, 17], or its generalizations to labeled (coloured) rooted trees [16, 17, 1, 20]. More precisely, $\mathcal{S}(\alpha)g$ ($g \in C^\infty(\mathcal{U})$) is formally equivalent to the composition of g with a B-series with coefficients $\alpha(t)$, $t \in \mathcal{T}$ [19, 5]). Theorem 24 in the particular case of α and β satisfying (32) is equivalent to the composition of two B-series (generalized to labeled rooted trees) with coefficients $\beta(t)$ and $\alpha(t)$ ($t \in \mathcal{F}$), respectively. The formal vector field whose 1-flow interpolates the numerical solution given by a one-step method expanded with such a B-series (the modified equation in formal backward error analysis) is just the series $\sum_{t \in \mathcal{T}} \log \alpha(t)/\sigma(t)\, X(t)$. See [5] for a recent work where series of the form (62) are exploited, and, in particular, maps $\alpha : \mathcal{F} \to \mathbb{K}$ satisfying (67) are interpreted in the context of preservation of first integrals of numerical integrators for ODEs.

7 Hopf-algebraic interpretation

We refer to [23] for the basic theory of Hopf algebras. Consider the commutative $\mathbb{K}$-algebra structure given by the shuffle product to $\mathbb{K}\langle A\rangle$, with the empty word as the unity of the shuffle product. In order to distinguish it from the empty labeled forest e, we will hereafter denote the empty word as $\widehat{e}$. It is well known that the shuffle algebra $\mathbb{K}\langle A\rangle$ has a commutative Hopf algebra structure with co-product $\widehat{\Delta} : \mathbb{K}\langle A\rangle \to \mathbb{K}\langle A\rangle \otimes \mathbb{K}\langle A\rangle$ given for each word $w = a_1 \cdots a_m$ ($a_1, \ldots, a_m \in A$) as

$$\widehat{\Delta} w = w \otimes \widehat{e} + \widehat{e} \otimes w + \sum_{j=1}^{m-1} a_1 \cdots a_j \otimes a_{j+1} \cdots a_m. \tag{68}$$

The co-unit $\widehat{\epsilon} : \mathbb{K}\langle A\rangle \to \mathbb{K}$ is given by $\widehat{\epsilon}(\widehat{e}) = 1$ and $\widehat{\epsilon}(w) = 0$ for $w \in A^*$. The interpretation of that co-product in our context is as follows. Each element $\widehat{\alpha}$ in the (algebraic) linear dual $\mathbb{K}\langle A\rangle^*$ of $\mathbb{K}\langle A\rangle$ gives rise to a series of the form (4) with $\alpha_w = \langle \widehat{\alpha}, w\rangle$ for each $w \in \mathbb{K}\langle A\rangle$, and the product of two such series associated to $\widehat{\alpha}, \widehat{\beta} \in \mathbb{K}\langle A\rangle^*$ is the series associated to $\widehat{\alpha}\widehat{\beta} \in \mathbb{K}\langle A\rangle^*$ given as

$$\langle \widehat{\alpha}\widehat{\beta}, w\rangle = \langle \widehat{\alpha} \otimes \widehat{\beta}, \widehat{\Delta} w\rangle, \quad w \in \mathbb{K}\langle A\rangle. \tag{69}$$

(If, for each $w \in A^*$, $\langle \widehat{\alpha}, w\rangle = \alpha_w$ and $\langle \widehat{\beta}, w\rangle = \beta_w$, then $\langle \widehat{\alpha}\widehat{\beta}, w\rangle = (\alpha\beta)_w$ given in (6)).

Consider the graded commutative $\mathbb{K}$-algebra $\mathbb{K}[\mathcal{T}]$ (graded by the degree of forests, or, more generally, by the weight of forests if A is a weighted alphabet), where the unity element is represented by the empty forest, and each monomial $t_1 \cdots t_m$ with $t_1, \ldots, t_m \in \mathcal{T}$ is associated to a forest $u \in \mathcal{F}$. It is well known [7, 9] that the graded commutative $\mathbb{K}$-algebra $\mathbb{K}[\mathcal{T}]$ can be given a commutative Hopf algebra structure over $\mathbb{K}$ compatible to the grading given to $\mathbb{K}[\mathcal{T}]$. Such a graded commutative Hopf algebra structure is uniquely determined by the co-product $\Delta : \mathbb{K}[\mathcal{T}] \to \mathbb{K}[\mathcal{T}] \otimes \mathbb{K}[\mathcal{T}]$ which is defined for each $u \in \mathcal{F}$ as follows. Given a labeled partially ordered set U representing u,

$$\Delta(u) = \sum_{(V,W)\in R(U)} v \otimes w, \tag{70}$$

where $R(U)$ is given in Definition 7, and for each pair $(V, W) \in R(U)$ of labeled partially ordered sets, the labeled forests v and w are represented by V and W, respectively. The co-unit $\epsilon : \mathbb{K}[\mathcal{T}] \to \mathbb{K}$ is given by $\epsilon(e) = 1$ and $\epsilon(u) = 0$ for $u \in \mathcal{F}\backslash\{e\}$. From now on, we will refer to such a commutative Hopf algebra simply as $\mathbb{K}[\mathcal{T}]$.

Clearly, each $\alpha \in \mathbb{K}[\mathcal{T}]^*$ is determined by its values $\langle\alpha, u\rangle = \alpha(u)$ for $u \in \mathcal{F}$, and the product in Definition 7 exactly corresponds to the product in the $\mathbb{K}$-algebra structure of $\mathbb{K}[\mathcal{T}]^*$ dual to the coalgebra $(\mathbb{K}[\mathcal{T}], \Delta, \epsilon)$, that is,

$$\langle\alpha\beta, u\rangle = \langle\alpha \otimes \beta, \Delta u\rangle, \quad u \in \mathbb{K}[\mathcal{T}]. \tag{71}$$

The maps $\alpha : \mathcal{F} \to \mathbb{K}$ such that $\alpha(u) = 0$ if $u \in \mathcal{F}\backslash\mathcal{T}$ (resp., such that (32) holds) correspond to $\alpha \in \mathbb{K}[\mathcal{T}]^*$ in the Lie algebra $P(K[\mathcal{T}]^\circ)$ of primitive elements (resp., in the group $G(K[\mathcal{T}]^\circ)$ of group-like elements) of the dual Hopf algebra $\mathbb{K}[\mathcal{T}]^\circ$ (in the sense of Sweedler [23]) of the commutative Hopf algebra on $\mathbb{K}[\mathcal{T}]$. Proposition 7 corresponds to the standard result that (if the base ring $\mathbb{K}$ is a $\mathbb{Q}$-algebra) the exponential defines a bijection $\exp : P(K[\mathcal{T}]^\circ) \to G(K[\mathcal{T}]^\circ)$ whose inverse is the logarithm.

Remark 23 The cocommutative Hopf algebra $\mathcal{H}_{GL}$ on labeled rooted trees of Grossman and Larson [11] is a Hopf subalgebra of the Hopf algebra $\mathbb{K}[\mathcal{T}]^\circ$. Actually, $\mathcal{H}_{GL}$ and $\mathbb{K}[\mathcal{T}]$ are graded dual (graded by the degree of forests if A is finite, and with other more general grading in the general case) to each other [10, 13]. That duality together with the results in [12] gives an alternative proof of Theorem 24 (actually, the proof we give in Section 6 is essentially a proof of the duality between the coalgebra structure of $\mathbb{K}[\mathcal{T}]$ and the algebra structure of $\mathcal{H}_{GL}$).

It is straightforward to check that the following holds for the co-product (70) in $\mathbb{K}[\mathcal{T}]$. Given $u \in \mathcal{F}$, $a \in A$, then

$$\Delta[u]_a = [u]_a \otimes e + (\mathrm{id} \otimes [\mathrm{id}]_a)\Delta u. \tag{72}$$

This, together with the fact that Δ is an algebra map, can be used to recursively compute the coproduct for all forests. Equivalently, (72) can be written as follows. Given $t \in \mathcal{T}$, $u, v, w \in \mathcal{F}$, we define $(u \otimes t) \circ (v \otimes w) := (uv) \otimes (t \circ w)$, and then

$$\Delta(t) = t \otimes e + \bar{\Delta}(t), \quad \bar{\Delta}([e]_a) = e \otimes [e]_a, \quad \bar{\Delta}(t \circ u) = \bar{\Delta}(t) \circ \Delta(u)$$

for each $t \in \mathcal{T}$, $u \in \mathcal{F}$, $a \in A$,

The graded commutative Hopf algebra $\mathbb{K}[\mathcal{T}]$ of rooted trees labeled by A can be characterized by the following universal property [7, 9].

Theorem 26 *Given a commutative algebra $\mathcal{C}$ over $\mathbb{K}$ and a family of $\mathbb{K}$-module maps $L_a : \mathcal{C} \to \mathcal{C}$, $a \in A$, there exists a unique $\mathbb{K}$-algebra homomorphism $\psi : \mathbb{K}[\mathcal{T}] \to \mathcal{C}$ such that $\psi([u]_a) = L_a(u)$ for each $u \in \mathbb{K}[\mathcal{T}]$, $a \in A$.*

If $\mathcal{C}$ has a Hopf algebra structure (with unity element $1_{\mathcal{C}}$ and coalgebra structure $(\mathcal{C}, \Delta_{\mathcal{C}}, \epsilon_{\mathcal{C}})$) satisfying $\cup_{a\in A} \mathrm{Im}\, L_a \subset \ker \epsilon_{\mathcal{C}}$ and

$$\Delta_{\mathcal{C}} L_d(c) = L_d(c) \otimes 1_{\mathcal{C}} + (\mathrm{id}_{\mathcal{C}} \otimes L_d)(\Delta_{\mathcal{C}}(c)) \tag{73}$$

for each $c \in \mathcal{C}$, $a \in A$, then ψ is a Hopf algebra homomorphism.

Corollary 27 *The $\mathbb{K}$-linear map $\nu : \mathbb{K}[\mathcal{T}] \to \mathbb{K}\langle A\rangle$ defined for each forest $u \in \mathcal{F}$ as in (48), is a Hopf algebra homomorphism over $\mathbb{K}$.*

Proof: Consider $\mathcal{C}$ as the shuffle algebra $\mathbb{K}\langle A\rangle$ and $L_a(w) = wa$ for each $w \in A^*$. We know from Proposition 13 that ν is a $\mathbb{K}$-algebra map, and that $\nu([u]_a) = L_a(u)$ for each $u \in \mathbb{K}[\mathcal{T}]$, $a \in A$. Whence, $\psi = \nu$. Furthermore, consider the commutative Hopf algebra structure on $\mathcal{C} = \mathbb{K}\langle A\rangle$ given by the co-product (68), it is straightforward to check that the assumptions on the second statement of Theorem 26 hold, and thus ν is an homomorphism of Hopf algebras over $\mathbb{K}$. □

The next result (equivalent to Proposition 4) is a direct consequence of Corollary 27.

Corollary 28 *Given $\widehat{\alpha}, \widehat{\beta} \in \mathbb{K}\langle A\rangle^*$, it holds that*

$$\langle \widehat{\alpha}\widehat{\beta}, \nu(u)\rangle = \langle \alpha\beta, u\rangle, \quad u \in \mathbb{K}[\mathcal{T}], \tag{74}$$

where $\alpha, \beta \in \mathbb{K}[\mathcal{T}]^*$ are given by $\alpha(u) = \widehat{\alpha}(\nu(u))$ and $\beta(u) = \widehat{\beta}(\nu(u))$ for each $u \in \mathbb{K}[\mathcal{T}]$.

Since any word $w \in A$ is the image by $\nu : \mathbb{K}[\mathcal{T}] \to \mathbb{K}$ of a labeled rooted tree without ramifications, we have that ν is an epimorphism of Hopf algebras. Thus, the shuffle Hopf algebra $\mathbb{K}\langle A\rangle$ is isomorphic to the quotient Hopf algebra $\mathbb{K}[\mathcal{T}]/(\ker \nu)$.

Lemma 29 *The Hopf ideal $\ker \nu$ coincides with the ideal $\mathcal{I}$ of the commutative algebra $\mathbb{K}[\mathcal{T}]$ generated by the set*

$$\left\{ \prod_{i=1}^{m} t_i - \sum_{i=1}^{m} t_i \circ \prod_{j\neq i} t_j : \quad m > 1,\ t_1, \ldots, t_m \in \mathcal{T} \right\}. \tag{75}$$

Proof: The set (75) can also be written as $\{\xi(v)-v : u \in \mathcal{F}\}$, where $\xi(v)$ for forests $v \in \mathcal{F}$ is recursively defined as follows: If $t \in \mathcal{T}$ and $v \in \mathcal{F}$, then $\xi(t) = t$ and $\xi(vt) = t \circ v + \xi(v) \circ t$. We first show that, if $u \in \mathcal{I}$ and $t \in \mathcal{T}$, then $t \circ u \in \mathcal{I}$. It is clearly sufficient to show that this is true for u of the form $u = \xi(v) - v$, $v \in \mathcal{F}$, which is obtained from the following. Given $t \in \mathcal{T}$ and $v \in \mathcal{F}$,

$$\begin{aligned} t \circ (v - \xi(v)) &= t \circ v + \xi(v) \circ t - (t \circ \xi(v) + \xi(v) \circ t - t\xi(v)) - t\xi(v) \\ &= (\xi(tv) - tv) + (\xi(\xi(v)t) - \xi(v)t) - (\xi(v) - v)t \in \mathcal{I}. \end{aligned}$$

Then it is easy to show by induction that any $u \in \mathcal{F}$ is congruent modulo $\mathcal{I}$ to a $\mathbb{Z}$-linear combination of rooted trees of the form $a_1 \circ \cdots \circ a_m$, $a_1, \ldots, a_m \in A$ (labeled rooted trees without ramifications). Clearly, $\mathcal{I} \subset \ker \nu$ and thus the $\mathbb{K}$-module $\mathbb{K}\langle A \rangle$ is isomorphic to some $\mathbb{K}$-submodule of $\mathbb{K}[\mathcal{T}]/\mathcal{I}$. Since the set of labeled rooted trees without ramifications can be identified with the set of words on the alphabet A, we have by a dimensional argument that $\mathbb{K}[\mathcal{T}]/\mathcal{I}$ is isomorphic to the $\mathbb{K}$-module $\mathbb{K}\langle A \rangle$, and thus, $\mathcal{I} = \ker \nu$. □

Now the results obtained so far lead to the following:

Theorem 30 *The shuffle Hopf algebra $\mathbb{K}\langle A \rangle$ is isomorphic to the quotient Hopf algebra $\mathbb{K}[\mathcal{T}]/\mathcal{I}$, where the Hopf ideal $\mathcal{I}$ is given in Lemma 29. The cosets $\{u + \mathcal{I} : u \in \widehat{\mathcal{F}}\}$ form a basis of the free $\mathbb{K}$-module $\mathbb{K}[\mathcal{T}]/\mathcal{I}$, and as a $\mathbb{K}$-algebra, $\mathbb{K}[\mathcal{T}]/\mathcal{I}$ is freely generated by $\{t + \mathcal{I} : t \in \widehat{\mathcal{T}}\}$ provided that $\mathbb{K}$ is a $\mathbb{Q}$-algebra. The dual basis of $\{u + \mathcal{I} : u \in \widehat{\mathcal{F}}\}$ is a PBW basis associated to a Hall basis of the free Lie algebra over the alphabet A.*

Notice that (72) together with Algorithm 1 gives a recursive way to describe the co-product in $\mathbb{K}[\mathcal{T}]/\mathcal{I}$ in terms of the basis $\{u+\mathcal{I} : u \in \widehat{\mathcal{F}}\}$, as an alternative to using (70) with a full rewriting algorithm as in Remark 16. According to the last statement in Theorem 30, such a description of the co-product in $\mathbb{K}[\mathcal{T}]/\mathcal{I}$ provides a direct way of computing the product of series written in the PBW basis $\{E(u) : u \in \widehat{\mathcal{F}}\}$.

Remark 24 It can be seen that the kernel $\mathcal{I} = \ker \nu$ is the smallest ideal $\mathcal{I}$ of $\mathbb{K}[\mathcal{T}]$ satisfying that $\{t \circ z + z \circ t - tz : t, z \in \mathcal{T}\} \subset \mathcal{I}$ and $t \circ u \in \mathcal{I}$ whenever $u \in \mathcal{I}$ and $t \in \mathcal{T}$. Moreover, it is not difficult to show that $\mathcal{I}$ is the ideal of the algebra $\mathbb{K}[\mathcal{T}]$ generated by the set

$$\{t \circ z + z \circ t - tz : t, z \in \mathcal{T}\} \cup \{s \circ t \circ z + s \circ z \circ t - s \circ (tz) : t, z, s \in \mathcal{T}\},$$

or, alternatively, by the set

$$\{t \circ z + z \circ t - tz : t, z \in \mathcal{T}\} \cup \{s \circ (tz) + z \circ (ts) + t \circ (sz) - tzs : t, z, s \in \mathcal{T}\}.$$

8 Concluding remarks

We have presented a new approach to dealing with Lie series, exponentials of Lie series, and related series in a PBW basis associated to an arbitrary Hall set that uses labeled rooted trees.

Some of the results we present are equivalent to (or can be derived from) known results, for instance, Theorem 3. The main original results of our work are Proposition 4 (equivalently, Corollary 27), Theorems 9 and 10, and as a by-product, the continuous BCH formula written in terms of Hall rooted trees, the rewriting Algorithm 1 and related results, Corollary 21, Theorems 24 and 25, and the explicit description of the epimorphism ν of Hopf algebras from the commutative Hopf algebra of labeled rooted trees and the shuffle Hopf algebra and its kernel.

References

[1] A. L. Araujo, A. Murua, and J. M. Sanz-Serna, *Symplectic methods based on decompositions*, SIAM Journal of Numerical Analysis, Vol 34, No. 5 (1997), 1926-1947.

[2] K. Burrage and P. M. Burrage, *High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula* in *Predictability: Quantifying Uncertainty in Models of Complex Phenomena* (S. Chen, L. Margolin, D. Sharp, eds.) Physica D 133 1-4 (1999), 34–48.

[3] J. C. Butcher, *Numerical Methods for Ordinary Differential Equations*, Wiley, Chichester, UK, (2003).

[4] P. Chartier, E. Hairer, and G. Vilmart, *A substitution law for B-series vector fields*, INRIA report No. 5498, 2005.

[5] P. Chartier, E. Faou, and A. Murua, *An algebraic approach to invariant preserving integrators: The case of quadratic and Hamiltonian invariants*, to appear in Numerische Mathematik (2006).

[6] K. T. Chen, *Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula*, Annals of Mathematics, Vol. 65, No. 1, (1957), 163–178.

[7] A. Connes and D. Kreimer, *Hopf algebras, renormalization, and non-commutative geometry*, Commun. Math. Phys. 199 (1998), 203–242.

[8] A. Dür, *Mobius Functions, Incidence Algebras and Power-Series Representations*, Lecture Notes in Mathematics, Vol. 1202, Springer-Verlag, Berlin, 1986.

[9] L. Foissy, *Les algèbres de Hopf des arbres enracinés dècorés, I*, Bulletin des Sciences Mathematiques, 126, 3 (2002), 193–239.

[10] L. Foissy, *Les algèbres de Hopf des arbres enracinés dècorés, II*, Bulletin des Sciences Mathematiques, 126, 4 (2002), 249–288.

[11] R. Grossman and R. G. Larson, *Hopf-algebraic structure of families of trees*, J. Algebra 126 (1989), 184–210.

[12] R. Grossman and R. G. Larson, *Solving nonlinear equations from higher order derivations in linear stages*, Advances in Mathematics, Vol. 82 (1990), 180–202.

[13] M. E. Hoffman, *Combinatorics of rooted trees and Hopf algebras*, Trans. Amer. Math. Soc. 355 (2003), 3795–3811.

[14] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, *Lie-group methods*, Acta Numerica 9 (2000), 215–365.

[15] M. Kawski and H. Sussmann, *Noncommutative power series and formal Lie-algebraic thecniques in nonlinear control theory*, in Operators, Systems and Linear Algebra: Three Decades of Algebraic Systems Theory (U. Helmke, D. Praetzel-Wolters, and E. Zerz eds.), B. G. Teubner, Stuttgart, (1997), pp. 111-129.

[16] E. Hairer, S.P. Nørset, and G. Wanner: *Solving ordinary differential equations I. Nonstiff problems*, 2nd. edition, Springer-Verlag, New York, 1993.

[17] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Springer-Verlag, Berlin, 2002.

[18] R.I. McLachlan and G.R.W. Quispel, *Splitting Methods*, Acta Numerica, 11 (2002), 341–434.

[19] A. Murua, *Formal series and numerical integrators. Part I: Systems of ODEs and Symplectic integrators*, Appl. Num. Math. 29 (1999), 221–251.

[20] A. Murua and J.M. Sanz-Serna, *Order conditions for numerical integrators obtained by composing simpler integrators*, Philosophical Trans. Royal Soc. A 357 (1999), 1079–1100.

[21] C. Reutenauer, *Free Lie Algebras*, London Math. Soc. monographs (new series), Vol. 7, Oxford University Press, USA, 1993.

[22] E. M. Rocha, On computation of the logarithm of the Chen-Fliess series for nonlinear systems, in Nonlinear and adaptive control(A. Zinober et al., eds.), Lecture Notes in Control and Inform. Sci. 281 (2003), 317-326.

[23] M. Sweedler, *Hopf Algebras*, Benjamin, New York, 1969.