J. Mostovoy, J. M. Pérez-Izquierdo, I. P. Shestakov

A non-associative Baker-Campbell-Hausdorff formula

Proceedings of the American Mathematical Society

DOI: 10.1090/proc/13684

Accepted Manuscript

A NON-ASSOCIATIVE BAKER-CAMPBELL-HAUSDORFF FORMULA

J. MOSTOVOY, J. M. PÉREZ-IZQUIERDO, AND I. P. SHESTAKOV

Abstract. We address the problem of constructing the non-associative version of the Dynkin form of the Baker-Campbell-Hausdorff formula; that is, expressing $\log(\exp(x)\exp(y))$, where x and y are non-associative variables, in terms of the Shestakov-Umirbaev primitive operations. In particular, we obtain a recursive expression for the Magnus expansion of the Baker-Campbell-Hausdorff series and an explicit formula in degrees smaller than 5. Our main tool is a non-associative version of the Dynkin-Specht-Wever Lemma. A construction of Bernouilli numbers in terms of binary trees is also recovered.

1. Introduction

The Baker-Campbell-Hausdorff formula is the expansion of $\log(\exp(x)\exp(y))$ in terms of nested commutators for the non-commuting variables x and y, where the commutator of a and b is defined as $[a, b] := ab - ba$. The explicit combinatorial form of it was given by Dynkin in his 1947 paper [5]. By considering the linear extension of the map γ defined by $\gamma(1) := 0$, $\gamma(x) := x$, $\gamma(y) := y$, $\gamma(ux) := [\gamma(u), x]$ and $\gamma(uy) := [\gamma(u), y]$ he proved that

$$(1) \quad \log(\exp(x)\exp(y)) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \sum_{r_i+s_i \geq 1} \frac{(\sum_{j=1}^{n}(r_j + s_j))^{-1}}{r_1!s_1! \cdots r_n!s_n!} \gamma(x^{r_1} y^{s_1} \cdots x^{r_n} y^{s_n}).$$

This series is related to Lie's Third Theorem and the history around it is too rich to be retold here, so we refer the reader to the recent monograph [2] and references therein for a historical account.

The Baker-Campbell-Hausdorff formula, as well as many other results in Lie theory, firmly belongs to associative algebra. However, after the work of Mikheev and Sabinin on local analytic loops [12] and the description of the primitive operations in non-associative algebras by Shestakov and Umirbaev [13], associativity does not seem to be as essential for the Lie theory as previously thought [10].

In this paper we address the problem of determining $\mathrm{BCH}_l(x, y)$ in

$$\exp_l(x)\exp_l(y) = \exp_l(\mathrm{BCH}_l(x, y)),$$

2010 *Mathematics Subject Classification.* 17A50,20N05.

Key words and phrases. Baker-Campbell-Hausdorff formula, primitive elements, Sabinin algebras, Magnus expansion.

The authors acknowledge the support by the Spanish Ministerio de Ciencia e Innovación (MTM2013-45588-C3-3-P); J. Mostovoy was also supported by the CONACYT grant 168093-F; I. P. Shestakov also acknowledges support by FAPESP, processo 2014/09310-5 and CNPq, processos 303916/2014-1 and 456698/2014-0.

where

$$\exp_l(x) = \sum_{n\geq 0} \frac{1}{n!} \underbrace{(((xx)\cdots)x)x}_{n},$$

in terms of Shestakov-Umirbaev operations for the primitive elements of the non-associative algebra freely generated by x and y. Our approach uses a generalization of the Magnus expansion (see [1] for a readable survey), that is, we will study the differential equation

$$X'(t) = X(t)A(t),$$

where $X(t)$ stands for $\exp_l(\Omega(t))$ and both $A(t)$ and $\Omega(t)$ belong to a non-associative algebra. The differential equation

$$\Omega'(t) = A(t) + \sum_J n_J P_J(\Omega(t); A(t))$$

satisfied by $\Omega(t)$ (Corollary 3.7) is obtained with the help of a non-associative version of the Dynkin-Specht-Wever Lemma (Lemma 3.1). This equation leads to a recursive formula for computing the expansion of $\mathrm{BCH}_l(x, y)$, which gives, in degrees smaller than 5, the following expression:

$$\begin{aligned}
\mathrm{BCH}_l(x,y) = x + y &+ \frac{1}{2}[x,y] \\
&+ \frac{1}{12}[x,[x,y]] - \frac{1}{3}\langle x;x,y\rangle - \frac{1}{12}[y,[x,y]] - \frac{1}{6}\langle y;x,y\rangle - \frac{1}{2}\Phi(x;y,y) \\
&- \frac{1}{24}\langle x;x,[x,y]\rangle - \frac{1}{12}[x,\langle x;x,y\rangle] - \frac{1}{8}\langle x,x;x,y\rangle \\
&+ \frac{1}{24}[[x,[x,y]],y] - \frac{1}{24}[x,\langle y;x,y\rangle] - \frac{1}{4}\Phi(x,x;y,y) - \frac{1}{4}[x,\Phi(x;y,y)] \\
&- \frac{1}{24}[\langle x;x,y\rangle,y] - \frac{1}{24}\langle x;[x,y],y\rangle - \frac{1}{6}\langle x,y;x,y\rangle + \frac{1}{24}\langle y,x;x,y\rangle \\
&+ \frac{1}{12}[\Phi(x;y,y),y] + \frac{1}{24}\langle y;y,[x,y]\rangle - \frac{1}{24}\langle y,y;x,y\rangle - \frac{1}{6}\Phi(x;y,y,y) \\
&+ \dots
\end{aligned}$$

When all the operations apart from [,] vanish we recover the usual Baker-Campbell-Hausdorff formula. A different approach to the non-associative Baker-Campbell-Hausdorff formula has appeared in [8]; it does not explicitly use the Dynkin-Specht-Wever Lemma or the Magnus expansion. For the treatment of the subject from the point of view of differential geometry see [16]; actually, geometric considerations also motivate a different type of a Baker-Campbell-Hausdorff formula, see [11]; although it is of importance for the non-associative Lie theory, we will not consider it here.

Our results are presented for the unital $\boldsymbol{k}$-algebra of formal power series $\boldsymbol{k}\{\{x, y\}\}$ in two non-associative variables x and y. Readers with background in non-associative structures will realize that a more natural context for the Baker-Campbell-Hausdorff formula is the completion of the universal enveloping algebra of a relatively free Sabinin algebra on two generators. The extension of our results to that context is rather straightforward.

1.1. **Notation.** Throughout this paper the characteristic of the base field $\boldsymbol{k}$ is zero. The unital associative $\boldsymbol{k}$-algebra freely generated by a set of generators X will be denoted by $\boldsymbol{k}\langle X\rangle$ while $\boldsymbol{k}\langle\langle X\rangle\rangle$ will stand for the unital associative algebra of formal power series on X with coefficients in $\boldsymbol{k}$. Their non-associative counterparts, namely, the unital non-associative $\boldsymbol{k}$-algebra freely generated by X and the unital non-associative $\boldsymbol{k}$-algebra of formal power series on X with coefficients in $\boldsymbol{k}$, will be denoted by $\boldsymbol{k}\{X\}$ and $\boldsymbol{k}\{\{X\}\}$ respectively. For any algebra H, $H[[t]]$ will denote the algebra of formal power series in t with coefficients in H. The parameter t commutes and associates with all the elements in $H[[t]]$. Finally, we will stick to the following order of parentheses for powers: $x^n := (((xx)\cdots)x)x$ (n times).

2. Fundamentals

2.1. **Non-associative Hopf algebras.** A coalgebra (C,Δ,ϵ) is a vector space equipped with two linear maps $\Delta\colon C\to C\otimes C$ (*comultiplicaton*) and $\epsilon\colon C\to\boldsymbol{k}$ (*counit*) such that

$$\sum \epsilon(x_{(1)})x_{(2)} = x = \sum \epsilon(x_{(2)})x_{(1)},$$

where $\sum x_{(1)}\otimes x_{(2)}$ stands for $\Delta(x)$ (*Sweedler notation*). Coassociative and cocommutative coalgebras are those coalgebras (C,Δ,ϵ) that, in addition, satisfy

$$(\Delta\otimes \mathrm{Id})\Delta = (\mathrm{Id}\otimes\Delta)\Delta$$

(*coassociativity*) and $\tau\Delta=\Delta$ (*cocommutativity*) where $\tau(x\otimes y)=y\otimes x$. Coassociativity ensures that

$$\sum {x_{(1)}}_{(1)}\otimes {x_{(1)}}_{(2)}\otimes x_{(2)} = \sum x_{(1)}\otimes {x_{(2)}}_{(1)}\otimes {x_{(2)}}_{(2)}$$

so we can safely write $\sum x_{(1)}\otimes x_{(2)}\otimes x_{(3)}$ for any of the sides of this equality. For coassociative coalgebras the result of the iterated application n times of Δ to x does not depend on the selected factors and it is denoted by $\sum x_{(1)}\otimes x_{(2)}\otimes\cdots\otimes x_{(n+1)}$. Cocommutativity ensures that we can freely permute the factors of $\sum x_{(1)}\otimes x_{(2)}\otimes\cdots\otimes x_{(n+1)}$ without altering the value of this expression [14].

In this paper, by a *(non-associative) Hopf algebra* $(H,m,u,\backslash,/,\Delta,\epsilon)$ we will mean a cocommutative and coassociative coalgebra (H,Δ,ϵ) endowed with the following linear maps: a *product* $m\colon H\otimes H\to H$, a *unit* $u\colon\boldsymbol{k}\to H$, a *left division* $\backslash\colon H\otimes H\to H$ and a *right division* $/\colon H\otimes H\to H$ so that $\Delta(xy)=\Delta(x)\Delta(y)$, $\Delta(1)=1\otimes 1$, $\epsilon(xy)=\epsilon(x)\epsilon(y)$, $\epsilon(1)=1$ and

$$\sum x_{(1)}\backslash(x_{(2)}y) = \epsilon(x)y = \sum x_{(1)}(x_{(2)}\backslash y) \tag{2}$$

$$\sum (yx_{(1)})/x_{(2)} = \epsilon(x)y = \sum (y/x_{(1)})x_{(2)} \tag{3}$$

where $xy := m(x\otimes y)$ and $1:=u(1)$ is the unit element (see [10] for a survey on non-associative Hopf algebras). In case that H is associative then the left and right divisions can be written as $x\backslash y = S(x)y$ and $x/y = xS(y)$ where S is the antipode. However, non-associative Hopf algebras lack antipodes in general.

2.2. **The free unital non-associative algebra $\boldsymbol{k}\{X\}$.** The most important example of a non-associative Hopf algebra in this paper is the unital non-associative algebra $\boldsymbol{k}\{X\}$ freely generated by $X:=\{x_1,x_2,\dots\}$. The maps $\Delta\colon x_i\mapsto x_i\otimes 1+1\otimes x_i$ and $\epsilon\colon x_i\mapsto 0$ $(i=1,2,\dots)$ induce homomorphisms of unital algebras $\Delta\colon\boldsymbol{k}\{X\}\to\boldsymbol{k}\{X\}\otimes\boldsymbol{k}\{X\}$ and $\epsilon\colon\boldsymbol{k}\{X\}\to\boldsymbol{k}$ so that $(\boldsymbol{k}\{X\},\Delta,\epsilon)$ is a coassociative and cocommutative coalgebra. By induction on the degree of x, the formulas (2) and

(3) uniquely determine the left and the right division in $\boldsymbol{k}\{X\}$. For instance, $1\backslash(1y) = \epsilon(1)y$ implies $1\backslash y = y$ and

$$x_i\backslash(1y) + 1\backslash(x_i y) = \epsilon(x_i)y = 0 \quad \text{implies} \quad x_i\backslash y = -x_i y$$

etc. The operations Δ, ϵ, $\backslash$ and $/$, together with the product and the unit, provide $\boldsymbol{k}\{X\}$ with the structure of a non-associative Hopf algebra. Far from being a fancy feature, the divisions are a valuable tool for computations.

2.3. **Primitive elements of $\boldsymbol{k}\{X\}$ and the Shestakov-Umirbaev operations.** An element a in a Hopf algebra H such that

$$\Delta(a) = a \otimes 1 + 1 \otimes a$$

is called *primitive*; the subspace of all such elements is denoted by $\mathrm{Prim}(H)$. While for associative Hopf algebras this subspace is a Lie algebra with the commutator product $[x, y] := xy - yx$, Shestakov and Umirbaev [13] realized that if H is non-associative, many more operations are required to describe its algebraic structure completely.

Let $X := \{x, x_1, x_2, \dots\}$, $Y := \{y, y_1, y_2, \dots\}$ and $Z := \{z\}$ be disjoint sets of symbols that we take to be the free generators of $\boldsymbol{k}\{X \cup Y \cup Z\}$. Write $\underline{x} := ((x_1x_2)\cdots)x_m$, $\underline{y} := ((y_1y_1)\cdots)y_n$ and define

$$p(x_1, \dots, x_m; y_1, \dots, y_n; z) := p(\underline{x}, \underline{y}, z) := \sum (\underline{x}_{(1)}\underline{y}_{(1)})\backslash(\underline{x}_{(2)}, \underline{y}_{(2)}, z) \tag{4}$$

in $\boldsymbol{k}\{X \cup Y \cup Z\}$, where (x, y, z) denotes the associator $(xy)z - x(yz)$ of x, y and z. Each of the elements $p(\underline{x}, \underline{y}, z)$ is primitive. Considered as non-associative polynomials, $p(x_1, \dots, x_m; y_1, \dots, y_n; z)$ can be evaluated in any algebra A so we can think of them as of new multilinear operations derived from the binary product of A. Define

$$\begin{aligned}
[x, y] &:= xy - yx \\
\langle x_1, \dots, x_m; y, z\rangle &:= -p(x_1, \dots, x_m; y; z) + p(x_1, \dots, x_m; z; y) \\
\Phi(x_1, \dots, x_m; y_1, \dots, y_n, y_{n+1}) &:= \\
&\frac{1}{m!(n+1)!} \sum_{\sigma \in S_n, \tau \in S_{m+1}} p(x_{\sigma(1)}, \dots, x_{\sigma(m)}; y_{\tau(1)}, \dots, y_{\tau(n)}; y_{\tau(n+1)})
\end{aligned}$$

where $m, n \geq 1$ and S_k stands for the symmetric group on $\{1, \dots, k\}$. In order to simplify the notation, for $m = 0$ we write

$$\langle y, z\rangle := \langle x_1, \dots, x_m; y, z\rangle := \langle 1; y, z\rangle := -[y, z].$$

With this convention, (4) gives

$$(\underline{x}y)z - (\underline{x}z)y = -\sum \underline{x}_{(1)}\langle \underline{x}_{(2)}; y, z\rangle. \tag{5}$$

Shestakov and Umirbaev proved that

$$(\mathrm{Prim}(\boldsymbol{k}\{X\}), \langle\ ;\ ,\ \rangle, \Phi(\ ;\ )) \text{ is generated by } X.$$

Thus, while (1) can be written in terms of commutators, the natural language to write its non-associative counterpart uses $\langle\ ;\ ,\ \rangle$ and $\Phi(\ ;\ )$.

2.4. **Exponentials, logarithms and the Baker-Campbell-Hausdorff formula.** The algebra $\boldsymbol{k}\{\{x\}\}$ (respectively, $\boldsymbol{k}\langle\langle x\rangle\rangle$) of formal power series in x with coefficients in $\boldsymbol{k}$ is a topological Hopf algebra with the continuous extension of the operations of the Hopf algebra $\boldsymbol{k}\{x\}$ (respectively, $\boldsymbol{k}\langle x\rangle$). Since $\mathrm{Prim}(\boldsymbol{k}\langle x\rangle) = \boldsymbol{k}x$, the *group-like* elements of $\boldsymbol{k}\langle\langle x\rangle\rangle$, that is, the elements g such that $\Delta(g) = g \otimes g$ and $\epsilon(g) = 1$, are of the form $\exp(\alpha x)$ with $\alpha \in \boldsymbol{k}$. Therefore, $\exp(x)$ is, in a sense, canonical among all of them. However, $\mathrm{Prim}(\boldsymbol{k}\{\{x\}\})$ is infinite-dimensional and $\boldsymbol{k}\{\{x\}\}$ has an infinite number of group-like elements that could rightfully be considered as the non-associative analogs of the exponential series. Apart from the most obvious non-associative versions of the exponential

$$\exp_l(x) := \sum_{n\geq 0} \frac{1}{n!} \underbrace{(((xx)\cdots)x)x}_{n} \quad \text{and} \quad \exp_r(x) := \sum_{n\geq 0} \frac{1}{n!} \underbrace{x(x(\cdots(xx)))}_{n}$$

other series have been proposed as non-associative analogs of $\exp(x) := \sum_{n=0}^{\infty} x^n/n!$, each leading to a different logarithm [7].

Definition 2.1. A group-like element $e(x) \in \boldsymbol{k}\{\{x\}\}$ is a *base for logarithms* if its homogeneous component $e_1(x)$ of degree one in x is not zero. We say that the base for logarithms $e(x)$ is *normalized* if $e_1(x) = x$.

Associated with any base for logarithms $e(x)$ there exists a primitive element $\log_e(x) \in \boldsymbol{k}\{\{x\}\}$ determined by

$$e(\log_e(x)) = x = \log_e(e(x)).$$

The *exponentiation* on $\boldsymbol{k}\{\{X\}\}$ *with base* $e(x)$ and the *logarithm on* $\boldsymbol{k}\{\{X\}\}$ *to the base* $e(x)$ are the maps

$$\begin{aligned} e\colon \boldsymbol{k}\{\{X\}\}_+ &\to 1 + \boldsymbol{k}\{\{X\}\}_+ & \qquad \log_e\colon 1 + \boldsymbol{k}\{\{X\}\}_+ &\to \boldsymbol{k}\{\{X\}\}_+ \\ u &\mapsto e(u) & 1 + u &\mapsto \log_e(1+u) \end{aligned}$$

where $\boldsymbol{k}\{\{X\}\}_+$ denotes the space of formal power series with zero constant term. Both maps are inverse to each other and give a bijection between the primitive and the group-like elements in $\boldsymbol{k}\{\{X\}\}$. The logarithms to the bases $\exp_l(x)$ and $\exp_r(x)$ will be denoted by $\log_l$ and $\log_r$, respectively.

Any base for logarithms $e(x)$ determines a Baker-Campbell-Hausdorff series in $\boldsymbol{k}\{\{x, y\}\}$:

$$\mathrm{BCH}_e(x, y) := \log_e(e(x)e(y)).$$

The element $\mathrm{BCH}_e(x, y)$ is primitive so it can be written in terms of the Shestakov-Umirbaev operations $\langle\ ;\ ,\ \rangle$ and $\Phi(\ ;\ )$. Since these operations are defined via the left-normed products $\underline{x}$ and $\underline{y}$, the base $\exp_l(x)$ is better adapted to recursive computations. In [9] $\log_l(1+x)$ has been described as follows. For $\tau = x$ set $B_\tau := \tau! := 1$. If $\tau \neq x$ is a non-associative monomial in x, there is only one way of writing τ as a product $(\ldots((x\tau_1)\tau_2)\ldots)\tau_k$. Set $B_\tau := B_k B_{\tau_1} \ldots B_{\tau_k}$ and $\tau! := k!\tau_1! \ldots \tau_k!$ where B_k is the kth Bernoulli number. With this notation we have

$$\log_l(1 + x) = \sum_\tau \frac{B_\tau}{\tau!}\tau \in \boldsymbol{k}\{\{x\}\}.$$

Given the explicit expressions for $\exp_l$ and $\log_l$, there is no difficulty writing down an explicit formula for $\mathrm{BCH}_l(x, y) := \mathrm{BCH}_{\exp_l}(x, y)$ in terms of the non-associative monomials in x and y. Namely, for a monomial $w(x, y)$ its coefficient in

$\mathrm{BCH}_l(x,y)$ is equal to

$$\sum_{(\tau, w_1,\dots,w_{|\tau|})} \frac{B_\tau}{\tau!} \frac{1}{i_1!\cdots i_{|\tau|}!} \frac{1}{j_1!\cdots j_{|\tau|}!}, \tag{6}$$

where the sum is taken over all the non-associative monomials τ in one variable and all the sets of monomials $w_q := x^{i_q} y^{j_q}$ such that the product $w_1 \dots w_{|\tau|}$ taken with the parentheses as in τ equals $w(x,y)$. The set of all such $(\tau, w_1, \dots, w_{|\tau|})$ can be easily described; however, we will not dwell on this subject since this formula is not nearly as useful as its associative version, for the reason explained in Remark 3.3 in the next section.

The Baker-Campbell-Hausdorff series for different bases are related in a straightforward manner. If e and f are two bases for logarithms, the series $h(x) := \log_f(e(x))$ is a primitive element of $\boldsymbol{k}\{\{x\}\}$ whose term of degree 1 is non-zero. In particular, it has a composition inverse $h^{-1}(x) = \log_e(f(x))$ such that $h^{-1}(h(x)) = x$. It is then clear that

$$\mathrm{BCH}_e(x,y) = h^{-1}(\mathrm{BCH}_f(h(x), h(y))).$$

Moreover, for any Baker-Campbell-Hausdorff series $\mathrm{BCH}(x,y)$ and any primitive $h \in \boldsymbol{k}\{\{x\}\}$ with $h_1 \neq 0$, the series $h^{-1}(\mathrm{BCH}(h(x), h(y)))$ is also a Baker-Campbell-Hausdorff series for some base.

3. A Nonassociative Baker-Campbell-Hausdorff formula

3.1. **A non-associative Dynkin-Specht-Wever Lemma.** Let d be a derivation of $\boldsymbol{k}\{X\}$ that preserves $\mathrm{Prim}(\boldsymbol{k}\{X\})$, that is

$$d(\mathrm{Prim}(\boldsymbol{k}\{X\})) \subseteq \mathrm{Prim}(\boldsymbol{k}\{X\}).$$

Define $\gamma_d(u) := \sum u_{(1)} \backslash d(u_{(2)})$; thus,

$$d(u) = \sum u_{(1)} \gamma_d(u_{(2)})$$

for all $u \in \boldsymbol{k}\{X\}$. The proof of the following result was inspired by [15].

Lemma 3.1 (The Dynkin-Specht-Wever Lemma)**.** *Let d be a derivation of $\boldsymbol{k}\{X\}$ that preserves $\mathrm{Prim}(\boldsymbol{k}\{X\})$, $u \in \boldsymbol{k}\{X\}$ and $a \in \mathrm{Prim}(\boldsymbol{k}\{X\})$. We have*

$$\gamma_d(ua) = \epsilon(u) d(a) + \sum \langle u_{(1)}; a, \gamma_d(u_{(2)}) \rangle.$$

Proof. Let us compute $d(ua)$ in two ways:

$$d(ua) = \begin{cases} \sum u_{(1)} \gamma_d(u_{(2)} a) + \sum (u_{(1)} a) \gamma_d(u_{(2)}) \\ \\ d(u)a + ud(a) = \sum (u_{(1)} \gamma_d(u_{(2)}))a + ud(a) \end{cases}$$

so that by (5)

$$\sum u_{(1)} \gamma_d(u_{(2)} a) = \sum u_{(1)} \langle u_{(2)}; a, \gamma_d(u_{(3)}) \rangle + \sum u_{(1)} \epsilon(u_{(2)}) d(a).$$

Using (2), divide by $u_{(1)}$ to get the result. □

Example 3.2. Let us compute the expansion $\log_l(\exp_l(x)\exp_l(y))$ up to degree 3 in terms of the Shestakov-Umirbaev operations with the help of the Dynkin-Specht-Wever Lemma. Since, up to the summands of degree ≥ 5, we have

$$\log_l(1+x) = \\ x - \frac{1}{2}x^2 + \frac{1}{12}x^2x + \frac{1}{4}xx^2 - \frac{1}{24}x(x^2x) - \frac{1}{8}x(xx^2) - \frac{1}{24}x^2x^2 - \frac{1}{24}(xx^2)x + \cdots,$$

the expansion of $\log_l(\exp_l(x)\exp_l(y))$ up to degree 3 is

$$x + y + \frac{1}{2}[x,y] + \frac{1}{3}x^2y - \frac{1}{4}x(xy) + \frac{1}{4}x(yx) - \frac{5}{12}(xy)x + \frac{1}{12}(yx)x \\ + \frac{1}{2}xy^2 - \frac{5}{12}(xy)y + \frac{1}{12}(yx)y - \frac{1}{4}y(xy) - \frac{1}{6}y^2x + \frac{1}{4}y(yx) + \cdots \tag{7}$$

Now, apply Lemma 3.1 with $d(u) := |u|u$, where $|u|$ denotes the degree of u, for homogeneous $u \in \boldsymbol{k}\{x,y\}$. First, observe that $\gamma_d(ab) = \langle b,a\rangle$,

$$\gamma_d((ab)c) = \langle c, \langle b,a\rangle\rangle + \langle a; c, b\rangle + \langle b; c, a\rangle$$

and

$$\gamma_d(a(bc)) = \gamma_d((ab)c - (a,b,c)) = \langle c, \langle b,a\rangle\rangle + \langle a; c, b\rangle + \langle b; c, a\rangle - 3(a,b,c).$$

Applying γ_d to the homogeneous summands in (7) and dividing by their degree, we can write (7) as

$$x+y+\frac{1}{2}[x,y]+\frac{1}{12}[x,[x,y]]-\frac{1}{3}\langle x;x,y\rangle-\frac{1}{12}[y,[x,y]]+\frac{1}{6}\langle y;y,x\rangle-\frac{1}{2}\Phi(x;y,y)+\cdots$$

□

Remark 3.3. In the associative case, the BCH formula (1) is obtained by applying the Dynkin-Specht-Wever Lemma to the expansion of $\log(\exp(x)\exp(y))$ in terms of the monomials in x and y. In our situation, we cannot apply Lemma 3.1 to the expansion given by (6) and get a closed formula since the monomials in (6) are not left-normed.

3.2. **A non-associative Magnus expansion.** The differential equation

$$X'(t) = A(t)X(t)$$

when $X(t)$ and $A(t)$ do not necessarily commute (for instance, $X(t)$ may belong to a matrix Lie group and $A(t)$ to the corresponding Lie algebra) has been studied since long ago [1]. A fruitful approach is to look for solutions of the form $X(t) = \exp(\Omega(t))$ for some $\Omega(t)$, where $\exp(x)$ denotes the usual exponential. The solution $\Omega(t)$ is determined by the initial condition and by the differential equation

$$\Omega'(t) = \frac{\mathrm{ad}_{\Omega(t)}}{\exp(\mathrm{ad}_{\Omega(t)}) - \mathrm{Id}}(A(t)) = \sum_{n=0}^{\infty} \frac{B_n}{n!}\,\mathrm{ad}^n_{\Omega(t)}(A(t)), \tag{8}$$

where B_n denotes the n-th Bernoulli number. Take $X(t) := \exp(tx)\exp(y)$; then

$$X'(t) = (x\exp(tx))\exp(y) \overset{<1>}{=} x(\exp(tx)\exp(y)) = xX(t)$$

so we can use (8) in order to study $\Omega(t) = \log(\exp(tx)\exp(y))$. However, in a non-associative setting there are some details to be taken care of, since, for instance, equality <1> above requires the associativity.

Proposition 3.4. *Let H be a unital algebra, $e(x) \in \boldsymbol{k}\{\{x\}\}$ a base for logarithms and $X(t) := e(\Omega(t))$ with $\Omega(t) \in H[[t]]$ such that $\Omega(0) = 0$. For any $A(t) \in H[[t]]$ the solution $\Omega(t)$ to the equation*

$$X'(t) = X(t)A(t)$$

satisfies

$$\Omega'(t) = (\tau^{e(\Omega(t))})^{-1}(A(t))$$

where $\tau^{e(x)}$ is defined by

$$\tau^{e(x)}(y) := e(x) \Bigg\backslash \frac{d}{ds}\Big|_{s=0} e(x+sy) \in \boldsymbol{k}\{\{x,y\}\}.$$

Proof. Evaluating at $x = \Omega(t)$ and $y = \Omega'(t)$ we get

$$\begin{aligned}\tau^{e(\Omega(t))}(\Omega'(t)) &= e(\Omega(t)) \Bigg\backslash \frac{d}{ds}\Big|_{s=0} e(\Omega(t)+s\Omega'(t)) = e(\Omega(t)) \Bigg\backslash \frac{d}{dt} e(\Omega(t)) \\ &= X(t)\backslash X'(t) = A(t).\end{aligned}$$

If $x = 0$ then $\tau^{e(x)}(y) = \frac{d}{ds}\big|_{s=0} e(sy) = \alpha y$ for some $0 \neq \alpha \in \boldsymbol{k}$ and there exists $(\tau^{e(x)})^{-1}(y) \in \boldsymbol{k}\{\{x,y\}\}$ such that $(\tau^{e(x)})^{-1}(\tau^{e(x)}(y)) = y$. Therefore $\Omega'(t) = (\tau^{e(\Omega(t))})^{-1}(A(t))$. $\square$

In order to compute $(\tau^{e(x)})^{-1}(y)$ in terms of the Shestakov-Umirbaev operations, we will use the Dynkin-Specht-Wever Lemma. Consider the derivation $y\partial_x$ of $\boldsymbol{k}\{\{x,y\}\}$ determined by

$$(y\partial_x)(x) := y \quad \text{and} \quad (y\partial_x)(y) := 0. \tag{9}$$

By induction on the degree $|u|$ of u we can check that

$$\Delta((y\partial_x)(u)) = \sum (y\partial_x)(u_{(1)}) \otimes u_{(2)} + u_{(1)} \otimes (y\partial_x)(u_{(2)})$$

so that $(y\partial_x)$ preserves $\mathrm{Prim}(\boldsymbol{k}\{\{x,y\}\})$ and it is related to $\tau^{e(x)}(y)$ via

$$\tau^{e(x)}(y) = e(x) \Bigg\backslash \frac{d}{ds}\Big|_{s=0} e(x+sy) = e(x)\backslash(y\partial_x)(e(x)) = \gamma_{y\partial_x}(e(x)).$$

Now, in order to apply the Dynkin-Specht-Wever Lemma recursively $e(x)$ should be a linear combination of left-normed products of primitive elements. This is the main reason for restricting ourselves to $\exp_l(x)$.

Lemma 3.5. *The component τ_n of degree n in x of $\tau^{\exp_l(x)}(y)$ is*

$$\sum_{i=1}^{n} \frac{1}{n+1}\frac{1}{(n-i)!}\langle x^{n-i}; x, \tau_{i-1}\rangle$$

where $\tau_0 := y$.

The expansion of $(\tau^{\exp_l(x)})^{-1}(y)$ can be easily obtained from the expansion of $\tau^{\exp_l(x)}(y)$. Given a tuple $J = (j_1, \dots, j_s) \in \mathbb{Z}^s$ with $j_1, \dots, j_s \geq 1$ define

$$P_J(x;y) := \langle \underbrace{x,\dots,x}_{j_1-1}; x, \langle \underbrace{x,\dots,x}_{j_2-1}; x, \langle \dots \langle \underbrace{x,\dots,x}_{j_s-1}; x, y\rangle\rangle\rangle \quad \text{and}$$

$$m_J := \frac{1}{j_1+\dots+j_s+1}\frac{1}{(j_1-1)!}\frac{1}{j_2+\dots+j_s+1}\frac{1}{(j_2-1)!}\cdots\frac{1}{j_s+1}\frac{1}{(j_s-1)!}.$$

The concatenation $(i_1,\dots,i_r,j_1,\dots,j_s)$ of $(i_1,\dots,i_r)$ and $(j_1,\dots,j_s)$ will be denoted by $(i_1,\dots,i_r)||(j_1,\dots,j_s)$.

Theorem 3.6. *In $\boldsymbol{k}\{\{x,y\}\}$ we have*

$$(\tau^{\exp_l(x)})^{-1}(y) = y + \sum_J n_J P_J(x;y)$$

where J runs over all possible tuples with entries ≥ 1 and

$$n_J := \sum_{J=J_1||\cdots||J_l} (-1)^l m_{J_1}\cdots m_{J_l}.$$

Proof. Let $y\partial_x$ be the derivation of $\boldsymbol{k}\{\{x,y\}\}$ determined by (9). The Dynkin-Specht-Wever Lemma implies

$$\frac{1}{(n+1)!}\gamma_{y\partial_x}(x^{n+1}) = \frac{1}{(n+1)!}\sum_{\substack{J=(j_1,\dots,j_s)\\ j_1+\cdots+j_s=n}} \binom{n}{j_1-1}\binom{n-j_1}{j_2-1}\cdots$$
$$\cdots\binom{n-j_1-\cdots-j_s}{j_s-1}P_J(x;y)$$

so $\tau^{\exp_l(x)}(y) = y + \sum_J m_J P_J(x;y)$ and

$$(\tau^{\exp_l(x)})^{-1}(y) = y + \sum_{\substack{l\geq 1\\ J_1,\dots,J_l}} (-1)^l m_{J_1}\cdots m_{J_l} P_{J_1}(x;P_{J_2}(x;\cdots(P_{J_l}(x;y)))).$$

Since $P_{J_1}(x;P_{J_2}(x;\cdots(P_{J_l}(x;y)))) = P_{J_1||\cdots||J_l}(x;y)$, the result follows. $\square$

Corollary 3.7. *Let H be a unital algebra and $A(t)\in H[[t]]$. The solution $\Omega(t)\in H[[t]]$ of the equation*

$$X'(t) = X(t)A(t)$$

with $X(t) := \exp_l(\Omega(t))$ and $\Omega(0)=0$ satisfies

$$\Omega'(t) = A(t) + \sum_J n_J P_J(\Omega(t);A(t)) \tag{10}$$

where J runs over all possible tuples with the components ≥ 1.

3.3. **A non-associative Baker-Campbell-Hausdorff formula.** We will use the formula for $(\tau^{\exp_l(x)})^{-1}$ in Theorem 3.6 to describe, in terms of the Shestakov-Umirbaev operations, the differential equation satisfied by $\log_l(\exp_l(x)\exp_l(ty))$.

Proposition 3.8. *Let $e(x)$ be a normalized base for logarithms. In $\boldsymbol{k}\{\{x,y\}\}$ we have*

$$\log_e(e(x)e(y)) = x + (\tau^{e(x)})^{-1}(y) + O(y^2).$$

Proof. Consider $\Omega(t) := \log_e(e(x)e(ty)) = x + \Omega'(0)t + O(t^2)$. Since

$$\tau^{e(\Omega(t))}(\Omega'(t)) = e(\Omega(t))\Big\backslash \frac{d}{dt}e(\Omega(t)) = (e(x)e(ty))\Big\backslash\frac{d}{dt}(e(x)e(ty))$$
$$= (e(x)e(ty))\Big\backslash\left(e(x)\frac{d}{dt}e(ty)\right),$$

evaluating at $t=0$, we get $\tau^{e(x)}(\Omega'(0)) = y$ so $\Omega'(0) = (\tau^{e(x)})^{-1}(y)$. $\square$

In the case when $e(x)$ is $\exp_l(x)$ or $\exp_r(x)$, Proposition 3.8 was proved in [16].

Example 3.9. The components of degree $0,1,2$ and 3 of $\tau^{\exp_l(x)}(y)$ are $\tau_0 = y$, $\tau_1 = \frac{1}{2}\langle x,y\rangle$, $\tau_2 = \frac{1}{3}\langle x;x,y\rangle + \frac{1}{6}\langle x,\langle x,y\rangle\rangle$ and $\tau_3 = \frac{1}{8}\langle x,x;x,y\rangle + \frac{1}{8}\langle x;x,\langle x,y\rangle\rangle + \frac{1}{12}\langle x,\langle x;x,y\rangle\rangle + \frac{1}{24}\langle x,\langle x,\langle x,y\rangle\rangle\rangle$. Thus, the component of degree one in y in $\log_l(\exp_l(x)\exp_l(y))$ is

$$y - \frac{1}{2}\langle x,y\rangle + \left(\frac{1}{12}\langle x,\langle x,y\rangle\rangle - \frac{1}{3}\langle x;x,y\rangle\right) + \left(\frac{1}{12}\langle x,\langle x;x,y\rangle\rangle + \frac{1}{24}\langle x;x,\langle x,y\rangle\rangle - \frac{1}{8}\langle x,x;x,y\rangle\right) + \cdots$$

We can compute directly the coefficient of $\langle x;x,\langle x,y\rangle\rangle$ in $\log_l(\exp_l(x)\exp_l(y))$, for instance. Since $\langle x;x,\langle x,y\rangle\rangle = P_{(2,1)}(x;y)$ then Theorem 3.6 ensures that this coefficient equals $n_{(2,1)} = m_{(2)}m_{(1)} - m_{(2,1)} = \frac{1}{3}\frac{1}{1!}\frac{1}{2}\frac{1}{0!} - \frac{1}{4}\frac{1}{1!}\frac{1}{2}\frac{1}{0!} = \frac{1}{24}$. □

Proposition 3.10 (Magnus expansion for the Baker-Campbell-Hausdorff formula)**.** *Let $\Omega(t) := \log_l(\exp_l(x)\exp_l(ty))$. In $\boldsymbol{k}\{\{x,y\}\}[[t]]$ we have*

$$\Omega'(t) = \left(y + \sum_J n_J P_J(\Omega(t);y)\right) - \left(\Phi(\exp_l(x);\exp_l(ty),y) + \sum_J P_J(\Omega(t);\Phi(\exp_l(x);\exp_l(ty),y)\right).$$

Proof. We have

$$\begin{aligned}
\tau^{\exp_l(\Omega(t))}(\Omega'(t)) &= \exp_l(\Omega(t)) \Big\backslash \frac{d}{dt}\exp_l(\Omega(t)) \\
&= (\exp_l(x)\exp_l(ty)) \Big\backslash \left(\exp_l(x)\frac{d}{dt}\exp_l(ty)\right) \\
&= (\exp_l(x)\exp_l(ty)) \backslash (\exp_l(x)(\exp_l(ty)y)) \\
&= y - p(\exp_l(x);\exp_l(ty);y) \\
&= y - \Phi(\exp_l(x);\exp_l(ty),y)
\end{aligned}$$

so $\Omega'(t) = (\tau^{\exp_l(\Omega(t))})^{-1}(y) - (\tau^{\exp_l(\Omega(t))})^{-1}(\Phi(\exp_l(x);\exp_l(ty),y))$. The result follows from Theorem 3.6. □

Example 3.11. The component of $\log_l(\exp_l(x)\exp_l(y))$ of degree 1 in x and degree 2 in y is

$$\begin{aligned}
\Omega_{1,2} &:= \frac{1}{2}\left(n_{(1)}\langle\Omega_{1,1},y\rangle + n_{(1,1)}\langle y,\langle x,y\rangle\rangle + n_{(2)}\langle y;x,y\rangle - \Phi(x;y,y)\right) \\
&= -\frac{1}{12}[y,[x,y]] + \frac{1}{6}\langle y;y,x\rangle - \frac{1}{2}\Phi(x;y,y).
\end{aligned}$$

□

Based on Proposition 3.10 we can compute the initial terms of the expansion of $\log_l(\exp_l(x)\exp_l(y))$.

Theorem 3.12 (Non-associative Baker-Campbell-Hausdorff Formula)**.** *The expansion of* $\log_l(\exp_l(x)\exp_l(y))$ *in* $\boldsymbol{k}\{\{x,y\}\}$ *is*

$$
\begin{aligned}
& x+y+\frac{1}{2}[x,y] \\
&\quad +\frac{1}{12}[x,[x,y]]-\frac{1}{3}\langle x;x,y\rangle-\frac{1}{12}[y,[x,y]]-\frac{1}{6}\langle y;x,y\rangle-\frac{1}{2}\Phi(x;y,y) \\
&\quad -\frac{1}{24}\langle x;x,[x,y]\rangle-\frac{1}{12}[x,\langle x;x,y\rangle]-\frac{1}{8}\langle x,x;x,y\rangle \\
&\quad +\frac{1}{24}[[x,[x,y]],y]-\frac{1}{24}[x,\langle y;x,y\rangle]-\frac{1}{4}\Phi(x,x;y,y)-\frac{1}{4}[x,\Phi(x;y,y)] \\
&\quad -\frac{1}{24}[\langle x;x,y\rangle,y]-\frac{1}{24}\langle x;[x,y],y\rangle-\frac{1}{6}\langle x,y;x,y\rangle+\frac{1}{24}\langle y,x;x,y\rangle \\
&\quad +\frac{1}{12}[\Phi(x;y,y),y]+\frac{1}{24}\langle y;y,[x,y]\rangle-\frac{1}{24}\langle y,y;x,y\rangle-\frac{1}{6}\Phi(x;y,y,y)
\end{aligned}
$$

plus terms of degree ≥ 5.

4. A connection with Bernoulli numbers and binary trees

Formulas (8) and (10) give an alternative point of view on the relation between the numbers $\{n_J\}_J$ and $\{B_k\}_k$ which has been established in [6, 17].

Theorem 4.1. *We have*

$$\frac{B_k}{k!}=n_{\underbrace{(1,\dots,1)}_{k}}.$$

Proof. By definition, $\langle x_1,\dots,x_m;y,z\rangle$ and $\Phi(x_1,\dots,x_m;y_1,\dots,y_n,y_{n+1})$ vanish in any associative algebra, with the only exception of $\langle y,z\rangle$. Thus, after projecting from $\boldsymbol{k}\{\{x,y\}\}$, in $\boldsymbol{k}\langle\langle x,y\rangle\rangle$ we get

$$(\tau^{\exp(x)})^{-1}(y)=y+\sum_{J=(1,\dots,1)} n_J P_J(x;y)=\sum_{k=0}^{\infty} n_{\underbrace{(1,\dots,1)}_{k}}(-1)^k \operatorname{ad}_x^k(y).$$

Since it is well-known that $(\tau^{\exp(x)})^{-1}(y)=\sum_{k=0}^{\infty}(-1)^k\frac{B_k}{k!}\operatorname{ad}_x^k(y)$ holds in $\boldsymbol{k}\langle\langle x,y\rangle\rangle$ we get the result; the sign $(-1)^k$ in the latter formula comes from our choice $\tau^{\exp(x)}:=\exp(x)\backslash\left.\frac{d}{ds}\right|_{s=0}\exp(x+sy)$ instead of $\left.\frac{d}{ds}\right|_{s=0}\exp(x+sy)/\exp(x)$. $\square$

In [17] Woon gave an algorithm to compute $B_n/n!$ with the help of the binary tree

$$
\begin{array}{ccccccccc}
 & & & [1,2] & & & & & \text{level 1} \\
 & & \swarrow & & \searrow & & & & \\
 & [-1,3] & & & & [1,2,2] & & & \text{level 2} \\
 \swarrow & & \searrow & & \swarrow & & \searrow & & \\
[1,4] & & [-1,2,3] & \vdots & [-1,3,2] & & [1,2,2,2] & & \vdots
\end{array}
$$

Here, the nodes are labeled by $[a_1,\dots,a_r]$; the root is $[1,2]$ and at any node we have

$$\begin{array}{ccc} & [a_1,\dots,a_r] & \\ \swarrow & & \searrow \\ [-a_1,a_2+1,\dots,a_r] & & [a_1,2,a_2,\dots,a_r] \end{array}$$

The factorial of the node $N=[a_1,\dots,a_r]$ is $N! := a_1(a_2!\cdots a_r!)$. Woon proved the equality

$$\frac{B_k}{k!} = \sum_N \frac{1}{N!}$$

for $k \geq 2$, where N runs over the nodes in the level k. In [6] Fuchs extended this construction as follows. Consider the *general PI binary tree* (see [6]):

$$\begin{array}{ccccccccc} & & & (1) & & & & & \text{level 1} \\ & \swarrow & & & & \searrow & & & \\ & (1,1) & & & & (2) & & & \text{level 2} \\ \swarrow & & \searrow & & \swarrow & & \searrow & & \\ (1,1,1) & & (2,1) & \vdots & (1,2) & & (3) & & \vdots \end{array}$$

with root (1) and at each node

$$\begin{array}{ccc} & (a_1,\dots,a_r) & \\ \swarrow & & \searrow \\ (1,a_1,a_2,\dots,a_r) & & (a_1+1,a_2,\dots,a_r) \end{array}$$

For any sequence $(c_n)_{n\geq 1}$ of complex numbers change the node $(a_1,\dots,a_r)$ by $c_{a_1}\cdots c_{a_r}$. Then define x_k to be the sum of the nodes in the k-th level. This value depends on the sequence $(c_n)_{n\geq 1}$. In the case when $c_n = \frac{-1}{n+1!}$ we get the tree

$$\begin{array}{ccccccc} & & & -\frac{1}{2!} & & & \\ & \swarrow & & & & \searrow & \\ & \frac{1}{2!2!} & & & & -\frac{1}{3!} & \\ \swarrow & & \searrow & & \swarrow & & \searrow \\ -\frac{1}{2!2!2!} & & \frac{1}{3!2!} & \vdots & \frac{1}{2!3!} & & -\frac{1}{4!} \end{array}$$

and for each k we have $x_k = B_k/k!$.

To relate these constructions to the numbers n_J in Theorem 3.6 we use a binary tree to collect the summands involved in $n_J = \sum_{J=J_1||\cdots||J_l}(-1)^l m_{J_1}\cdots m_{J_l}$. Consider associative but non-commutative indeterminates $x_1,x_2,\dots$ and the tree

$$\begin{array}{ccccccccc} & & & (x_1) & & & & & \text{level 1} \\ & \swarrow & & & & \searrow & & & \\ & (x_1,x_2) & & & & (x_1x_2) & & & \text{level 2} \\ \swarrow & & \searrow & & \swarrow & & \searrow & & \\ (x_1,x_2,x_3) & & (x_1,x_2x_3) & \vdots & (x_1x_2,x_3) & & (x_1x_2x_3) & & \vdots \end{array}$$

where at any node on the level $n-1$ we have

$$\begin{array}{ccc} & (w_1,\dots,w_r) & \\ (w_1,\dots,w_r,x_n) & & (w_1,\dots,w_r x_n) \end{array}$$

Consider a sequence of numbers $a_1, a_2, \dots$. Define for $w = x_{i_1}\cdots x_{i_s}$ the number $m_w = m_{x_{i_1}\cdots x_{i_s}} = -m_{(a_{i_1},\dots,a_{i_s})}$ and replace any node $(w_1,\dots,w_r)$ with $m_{(w_1,\dots,w_r)} = m_{w_1}\cdots m_{w_r}$. The sum of the nodes in the level n of the resulting tree is $n_{(a_1,\dots,a_n)}$. In case that $a_1 = a_2 = \cdots = 1$, in the previous construction we can replace the label $x_{i_1}\cdots x_{i_s}$ by s without losing information. With these new labels, at any node on the level $n-1$ of the tree we have

$$\begin{array}{ccc} & (a_1,\dots,a_r) & \\ (a_1,\dots,a_r,1) & & (a_1,\dots,a_r+1) \end{array}$$

which essentially gives the general PI binary tree. The number that we attach to the node $(a_1,\dots,a_r)$ is $(-m_{\underbrace{(1,\dots,1)}_{a_1}})\cdots(-m_{\underbrace{(1,\dots,1)}_{a_r}}) = \frac{-1}{(a_1+1)!}\cdots\frac{-1}{(a_r+1)!}$, so we recover the construction of Fuchs.

References

[1] S. Blanes, F. Casas, J. A. Oteo, and J. Ros, *The Magnus expansion and some of its applications*, Phys. Rep. **470** (2009), no. 5-6, 151–238.

[2] A. Bonfiglioli and R. Fulci, *Topics in noncommutative algebra*, Lecture Notes in Mathematics, vol. 2034, Springer, Heidelberg, 2012. The theorem of Campbell, Baker, Hausdorff and Dynkin.

[3] F. Chapoton and F. Patras, *Enveloping algebras of preLie algebras, Solomon idempotents and the Magnus formula*, Internat. J. Algebra Comput. **23** (2013), no. 4, 853–861.

[4] V. Drensky and L. Gerritzen, *Nonassociative exponential and logarithm*, J. Algebra **272** (2004), no. 1, 311–320.

[5] E. B. Dynkin, *Calculation of the coefficients in the Campbell-Hausdorff formula*, Doklady Akad. Nauk SSSR (N.S.) **57** (1947), 323–326 (Russian).

[6] P. Fuchs, *Bernoulli numbers and binary trees*, Tatra Mt. Math. Publ. **20** (2000), 111–117. Number theory (Liptovský Ján, 1999).

[7] L. Gerritzen, *Planar rooted trees and non-associative exponential series*, Adv. in Appl. Math. **33** (2004), no. 2, 342–365.

[8] L. Gerritzen and R. Holtkamp, *Hopf co-addition for free magma algebras and the non-associative Hausdorff series*, J. Algebra **265** (2003), no. 1, 264–284.

[9] J. Mostovoy and J. M. Pérez-Izquierdo, *Formal multiplications, bialgebras of distributions and nonassociative Lie theory*, Transformation Groups **15** (2010), no. 3, 625-653.

[10] J. Mostovoy, J. M. Pérez-Izquierdo, and I. P. Shestakov, *Hopf algebras in non-associative Lie theory*, Bulletin of Mathematical Sciences **4** (2014), no. 1, 129-173.

[11] ———, *Nilpotent Sabinin algebras*, Journal of Algebra **419** (2014), 95-123.

[12] L. V. Sabinin and P. O. Mikheev, *Infinitesimal theory of local analytic loops*, Dokl. Akad. Nauk SSSR **297** (1987), no. 4, 801–804 (Russian); English transl., Soviet Math. Dokl. **36** (1988), no. 3, 545–548.

[13] I. P. Shestakov and U. U. Umirbaev, *Free Akivis algebras, primitive elements, and hyperalgebras*, J. Algebra **250** (2002), no. 2, 533–548.

[14] M. E. Sweedler, *Hopf algebras*, Mathematics Lecture Note Series, W. A. Benjamin, Inc., New York, 1969.

[15] D. Wigner, *An identity in the free Lie algebra*, Proc. Amer. Math. Soc. **106** (1989), no. 3, 639–640.

[16] G. Weingart, *On the axioms for Sabinin algebras*, Adv. Geom. **16** (2016), no. 2, 205–229.

[17] S. C. Woon, *A tree for generating Bernoulli numbers*, Math. Mag. **70** (1997), no. 1, 51–56.

Departamento de Matemáticas, CINVESTAV-IPN, Apartado Postal 14–740, 07000 México D.F., Mexico
E-mail address: jacob@math.cinvestav.mx

Departamento de Matemáticas y Computación, Universidad de La Rioja, 26004, Logroño, Spain
E-mail address: jm.perez@unirioja.es

Instituto de Matemática e Estatística, Universidade de São Paulo, Caixa Postal 66281, São Paulo, SP 05311-970, Brazil
E-mail address: shestak@ime.usp.br