Journal of Algebra 367 (2012) 54–74

Contents lists available at SciVerse ScienceDirect

Journal of Algebra

www.elsevier.com/locate/jalgebra

# Double shuffle and Kashiwara–Vergne Lie algebras

Leila Schneps

*Institut de Mathématiques de Jussieu, 4 place Jussieu, Case 247, 75252 Paris Cedex, France*

ARTICLE INFO

*Article history:*
Received 7 June 2011
Available online 20 June 2012
Communicated by Masaki Kashiwara

*Keywords:*
Lie algebras
Kashiwara–Vergne problem
Double shuffle
Multiple zeta values

ABSTRACT

We prove that the double shuffle Lie algebra $\mathfrak{ds}$, dual to the space of new formal multiple zeta values, injects into the Kashiwara–Vergne Lie algebra $\mathfrak{krv}_2$ defined and studied by Alekseev and Torossian. The proof is based on a reformulation of the definition of $\mathfrak{krv}_2$, and uses a theorem of Ecalle on a property of elements of $\mathfrak{ds}$. In the final section, we lift this result to an injection $DS \hookrightarrow KRV_2$ of the corresponding pro-unipotent groups.



## 1. Definitions and main results

Let $\mathbb{Q}\langle x, y\rangle$ denote the ring of polynomials in non-commutative variables $x$ and $y$, and $\mathrm{Lie}[x, y]$ the Lie algebra of Lie polynomials inside it. For each $n \geqslant 1$, let $\mathbb{Q}_n\langle x, y\rangle$ (resp. $\mathrm{Lie}_n[x, y]$) denote the subspace of homogeneous polynomials (resp. Lie polynomials) of degree $n$. For $k \geqslant 1$, let $\mathbb{Q}_{n\geqslant k}\langle x, y\rangle$ (resp. $\mathrm{Lie}_{n\geqslant k}[x, y]$) denote the space of polynomials (resp. Lie polynomials) all of whose monomials are of degree $\geqslant k$, i.e. the direct sum of the $\mathbb{Q}_n\langle x, y\rangle$ (resp. $\mathrm{Lie}_n[x, y]$) for $n \geqslant k$.

The main theorem of this paper gives an injective map between two Lie algebras studied in the literature concerning formal multiple zeta values: the *double shuffle Lie algebra* $\mathfrak{ds}$, investigated in papers by Racinet and Ecalle amongst others (the associated graded of $\mathfrak{ds}$ is also studied in papers by Zagier, Kaneko and others), and the Kashiwara–Vergne Lie algebra introduced in work of Alekseev and Torossian (cf. [AT]). We begin by recalling the definitions of these two Lie algebras. As vector spaces, both are subspaces of the free Lie algebra $\mathrm{Lie}[x, y]$.

For any non-trivial monomial $w$ and polynomial $f \in \mathbb{Q}\langle x, y\rangle$, we use the notation $(f|w)$ for the coefficient of the monomial $w$ in the polynomial $f$, and extend it by linearity to polynomials $w$ without constant term. Set $y_i = x^{i-1}y$ for all $i \geqslant 1$; then all words ending in $y$ can be written as

*E-mail address:* leila@math.jussieu.fr.

0021-8693/$ – see front matter 
http://dx.doi.org/10.1016/j.jalgebra.2012.04.034

words in the variables $y_i$. The stuffle product $st(u,v) \in \mathbb{Q}\langle x,y\rangle$ of two such words $u$ and $v$ is defined recursively by

$$st(1,u) = st(u,1) = u \quad \text{and} \quad st(y_iu, y_jv) = y_i\, st(u, y_jv) + y_j\, st(y_iu, v) + y_{i+j}\, st(u,v).$$

**Definition 1.1.** The *double shuffle Lie algebra* $\mathfrak{ds}$[1] is the vector space of elements $f \in \mathrm{Lie}_{n\geqslant 3}[x,y]$ such that

$$\bigl(f \bigm| st(u,v)\bigr) = 0$$

for all words $u, v \in \mathbb{Q}\langle x,y\rangle$ ending in $y$ but not both simultaneously powers of $y$.

It has been shown by Racinet [R] (see also a simplified version of Racinet's proof in the appendix of [F]) and Ecalle [E] that $\mathfrak{ds}$ is actually closed, i.e. a Lie algebra, under the Poisson bracket defined on $\mathrm{Lie}[x,y]$ by

$$\{f,g\} = [f,g] + D_f(g) - D_g(f), \tag{1.1}$$

where for any $f \in \mathrm{Lie}[x,y]$, the associated derivation $D_f$ of $\mathrm{Lie}[x,y]$ is defined by $D_f(x) = 0$, $D_f(y) = [y,f]$. This Lie bracket corresponds to identifying $f$ with $D_f$ and taking the natural Lie bracket on derivations:

$$[D_f, D_g] = D_{\{f,g\}}. \tag{1.2}$$

Let us now recall the definition of $\mathfrak{krv}_2$. Following [AT], let $TR$ denote the vector space quotient of $\mathbb{Q}\langle x,y\rangle$ by relations $ab = ba$. The image in $TR$ of a monomial $w$ is the equivalence class of monomials obtained by cyclically permutating the letters of $w$. The trace map $\mathbb{Q}\langle x,y\rangle \to TR$ is denoted by $tr$.

**Definition 1.2.** For any pair of elements $F, G \in \mathrm{Lie}_n[x,y]$ with $n \geqslant 1$, let $D_{F,G}$ denote the derivation of $\mathrm{Lie}[x,y]$ defined by $x \mapsto [x,G]$ and $y \mapsto [y,F]$. Such a derivation is said to be *special* if $x+y \mapsto 0$, i.e. if $[x,G]+[y,F] = 0$. The underlying vector space of the Kashiwara–Vergne Lie algebra is spanned by those of these special derivations $D_{F,G}$ that also satisfy the property that writing $F = F_x x + F_y y$ and $G = G_x x + G_y y$, there exists a constant $A$ such that

$$tr(F_y y + G_x x) \equiv A\, tr\bigl((x+y)^n - x^n - y^n\bigr) \in TR. \tag{1.3}$$

It is shown in [AT] that $\mathfrak{krv}_2$ is a Lie algebra under the natural bracket on derivations. The degree provides a grading on $\mathfrak{krv}_2$, for which $(\mathfrak{krv}_2)_n$ is spanned by the $D_{F,G}$ with $F, G \in \mathrm{Lie}_n[x,y]$.

The first graded piece, $(\mathfrak{krv}_2)_1$, is 1-dimensional, generated by $D_{y,x}$. The second graded piece $(\mathfrak{krv}_2)_2 = 0$. Now let $n \geqslant 3$. Note that for any $F \in \mathrm{Lie}_n[x,y]$, if there exists $G \in \mathrm{Lie}_n[x,y]$ such that $[y,F]+[x,G] = 0$, then $G$ is unique. Indeed, $G$ is defined up to a centralizer of $x$, but that can only be $x$, which is of degree 1. One of the most useful results of this paper is the precise determination of the elements $F$ admitting such a $G$, together with an explicit formula for $G$ (Theorem 2.1, see also (1.5)).

[1] The equivalence of the present definition with the usual definition introduced in [R] is proven in [CS], Theorem 2, which proves that if a polynomial $f \in \mathrm{Lie}_n[x,y]$ has the property of the present definition, then $f + \frac{(-1)^{n-1}}{n}(f|x^{n-1}y)y^n$ satisfies the stuffle relations for all pairs of words $u, v$ ending in $y$. Since the words ending in $x$ are not involved in this condition, this is equivalent to the assertion that $\pi_y(f) + \frac{(-1)^{n-1}}{n}(f|x^{n-1}y)y^n$ satisfies stuffle, where $\pi_y(f)$ denotes the projection of $f$ onto just its words ending in $y$. This is the standard form of the defining property of elements of $\mathfrak{ds}$.

Let $\partial_x$ denote the derivation of $\mathbb{Q}\langle x,y\rangle$ defined by $\partial_x(x)=1$, $\partial_x(y)=0$. Following Racinet [R], for any polynomial $h$ in $x$ and $y$, set

$$s(h)=\sum_{i\geqslant 0}\frac{(-1)^i}{i!}\partial_x^i(h)yx^i. \tag{1.4}$$

Racinet shows that if $f=f_x x+f_y y$ is an element of $\mathrm{Lie}_{n\geqslant 2}[x,y]$, or indeed any polynomial such that $\partial_x(f)=0$, then

$$f=s(f_y). \tag{1.5}$$

The main result of this paper is the following.

**Theorem 1.1.** *Let $\tilde f(x,y)\in\mathfrak{ds}$, and set $f(x,y)=\tilde f(x,-y)$ and $F(x,y)=f(z,y)$ with $z=-x-y$. Write $F=F_x x+F_y y=xF^x+yF^y$ in $\mathbb{Q}\langle x,y\rangle$. Set $G=s(F^x)$. Then the map $\tilde f\mapsto D_{F,G}$ yields an injective map of Lie algebras*

$$\mathfrak{ds}\hookrightarrow\mathfrak{krv}_2.$$

**Remark.** The map defined in Theorem 1.1 from $\mathfrak{ds}$ to the space of derivations $D_{F,G}$ mapping $x\mapsto[x,G]$ and $y\mapsto[y,F]$ is injective. Indeed, because $D_{F,G}(y)=[y,F]$ is a Lie element in which no word starts and ends with $x$, we can recover $F$ from $D_{F,G}(y)$ by applying Proposition 2.2 (with $x$ and $y$ exchanged in the statement), and then we recover $\tilde f$ by $F(x,y)=\tilde f(z,-y)$.

Furthermore, this injection of vector spaces is in fact an injection of Lie algebras, since $\mathfrak{ds}$ is equipped with the Poisson bracket, which is compatible with the natural bracket on derivations (cf. (1.2)).

Thus, to prove Theorem 1.1, it remains only to prove that the derivations $D_{F,G}$ arising from elements $\tilde f\in\mathfrak{ds}$ actually lie in $\mathfrak{krv}_2$, i.e. are special and satisfy the trace formula (1.3).

One of the main ingredients in our proof of Theorem 1.1 is a combinatorial reformulation of the defining properties of $\mathfrak{krv}_2$, given in Theorem 1.2 below. First we need some definitions.

**Definition 1.3.** Let $w=x^{a_0}y\cdots yx^{a_r}$ be a monomial in $\mathbb{Q}\langle x,y\rangle$ of depth $r$ (i.e. containing $r$ $y$'s), with $a_i\geqslant 0$ for $0\leqslant i\leqslant r$. Let *anti* denote the palindrome or backwards-writing operator on monomials, and let *push* denote the cyclic permutation of $x$-powers operator on monomials, defined respectively by

$$anti\big(x^{a_0}y\cdots yx^{a_{r-1}}yx^{a_r}\big)=x^{a_r}yx^{a_{r-1}}y\cdots yx^{a_0}, \tag{1.6}$$

$$push\big(x^{a_0}y\cdots yx^{a_{r-1}}yx^{a_r}\big)=x^{a_r}yx^{a_0}\cdots yx^{a_{r-1}}. \tag{1.7}$$

For any word $w$, we define the list $Push(w)$ to be the list of $(r+1)$ words obtained from $w$ by iterating the *push* operator. Note that $Push(w)$ is a list, not a set; it may contain repeated words. For example, if $w=x^2yxy$, then $Push(w)=[x^2yxy, yx^2yx, xy^2x^2]$, and if $w=xyxyx$, then $Push(w)=[xyxyx, xyxyx, xyxyx]$.

**Definition 1.4.** We extend the *anti* and *push* operators to operators on polynomials by linearity; it makes sense to apply these operators to a polynomial even if the monomials in the polynomial have different degrees and depths. If $f$ is a polynomial in $x$ and $y$ of homogeneous degree $n\geqslant 3$, we say that $f$ is

- *palindromic* if $f=(-1)^{n-1}anti(f)$,
- *antipalindromic* if $f=(-1)^n anti(f)$,

- *push-invariant* if $push(f) = f$,
- *push-constant* if there exists a constant $A$ such that $\sum_{v \in Push(w)} (f|v) = A$ for all $w \neq y^n$, and $(f|y^n) = 0$.

The following statement contains our reformulation of the definition of $\mathfrak{krv}_2$ that appears in [AT].

**Theorem 1.2.** *Let $V_{kv}$ be the vector space spanned by all polynomials $F \in \mathrm{Lie}_n[x, y]$ for $n \geqslant 3$ such that, writing $F = F_x x + F_y y$, we have*

i) *$F_y$ is antipalindromic, or equivalently, $F$ is push-invariant*;
ii) *$F_y - F_x$ is push-constant.*

*For each such $F$, set $G = s(F^x)$. Then the map $F \mapsto D_{F,G}$ extends to a vector space isomorphism*

$$V_{kv} \xrightarrow{\sim} \mathfrak{krv}_2. \tag{1.8}$$

The main result of Section 2, Theorem 2.1, is an enumeration of several conditions equivalent to the specialness property. Using this result, Theorems 1.1 and 1.2 are proved in Section 3. The proof of Theorem 1.1 is based on two previously known results for $\mathfrak{ds}$, each implying one of the two properties of Theorem 1.2. The first of these theorems, Theorem 3.1, is a translation into the standard terms of $x$, $y$ variables of a theorem due to Ecalle [E]. Because this result is couched in Ecalle's own original language, we give not only the reference to the precise statement, but also Appendix A giving the complete calculation-translation which brings it to the form of Theorem 3.1. The second, Theorem 3.2, appeared as Theorem 1 of [CS], with a complete elementary proof which was also based on an idea of Ecalle.

## 2. Characterizing special derivations

The main theorem of this section characterizes special derivations $D_{F,G}$ of $\mathrm{Lie}[x, y]$. From now on, if $f$ is an element of $\mathrm{Lie}[x, y]$, we say that $f$ is *special* if setting $F = f(z, y)$ with $z = -x - y$, there exists a $G \in \mathrm{Lie}[x, y]$ such that $D_{F,G}$ is special. By additivity, we may restrict ourselves to homogeneous Lie elements.

**Notation.** For any $f \in \mathbb{Q}\langle x, y\rangle$, we will use the notation

$$f = f_x x + f_y y = x f^x + y f^y.$$

Observe that since every Lie element is palindromic, if $f \in \mathrm{Lie}_n[x, y]$, we have

$$f = (-1)^{n-1} anti(f) = f_x x + f_y y = (-1)^{n-1} x\, anti(f_x) + (-1)^{n-1} y\, anti(f_y) = x f^x + y f^y,$$

so in fact

$$f^x = (-1)^{n-1} anti(f_x), \qquad f^y = (-1)^{n-1} anti(f_y). \tag{2.1}$$

Recall also the definition of the map $s : \mathbb{Q}\langle x, y\rangle \to \mathbb{Q}\langle x, y\rangle$ from (1.4). We will also use the similar map

$$s'(h) = \sum_{i \geqslant 0} \frac{(-1)^i}{i!} x^i y \partial_x^i(h). \tag{2.2}$$

When $f \in \mathrm{Lie}_n[x,y]$ ($n \geqslant 2$), it follows by symmetry from Racinet's result $f = s(f_y)$ that if we write $f = xf^x + yf^y$, then $f = s'(f^y)$.

**Theorem 2.1.** *Let $n \geqslant 3$, and let $f \in \mathrm{Lie}_n[x,y]$. Set $F = f(-x-y, y)$, and write $f = f_x x + f_y y$ and $F = F_x x + F_y y$. Then the following are equivalent:*

i) *$f$ is special, i.e. there exists a unique $G \in \mathrm{Lie}_n[x,y]$ such that $[y,F] + [x,G] = 0$.*
ii) *Setting $G = s'(F_x)$, the derivation $D_{F,G}$ is special.*
iii) *$F_y$ is antipalindromic.*
iv) *$F$ is push-invariant.*
v) *$f_y - f_x$ is antipalindromic.*

The equivalence of i), ii) and iii) is given in Proposition 2.3. The equivalence of iii) and iv) is proven in Proposition 2.4, and the equivalence of iii) and v) is given in the following Proposition 2.6. Some of these results, in particular Propositions 2.2 and 2.6, will also be used in the proofs of the main theorems in Section 3.

**Proposition 2.2.** *Let $n \geqslant 3$, and let $f \in \mathrm{Lie}_n[x,y]$ have the property that expanded as a polynomial, $f$ has no terms that start and end in $y$, so that writing $f = f_x x + f_y y$, we have $f_y y = xPy$. Then $s(P) \in \mathrm{Lie}_{n-1}[x,y]$ and $f = [x, s(P)]$.*

**Proof.** By hypothesis, $f$ has no terms starting and ending in $y$, so we can write $f_y y = xPy$. By Racinet's result, we have $g = s(g_y)$ for all $g \in \mathrm{Lie}_n[x,y]$ with $n \geqslant 2$, so in particular we have $f = s(xP)$. Now, since the partial derivative satisfies $\partial^i(xP) = i\partial^{i-1}(P) + x\partial^i(P)$, and $\partial^n(P) = 0$ since $P$ is of degree $n-1$, we compute

$$
\begin{aligned}
f = s(xP) &= \sum_{i=0}^{n} \frac{(-1)^i}{i!} \partial^i(xP) y x^i \\
&= \sum_{i=0}^{n} \frac{(-1)^i}{i!} \big(i\partial^{i-1}(P) + x\partial^i(P)\big) y x^i \\
&= \sum_{i=0}^{n} \frac{(-1)^i}{(i-1)!} \partial^{i-1}(P) y x^i + \sum_{i=0}^{n} \frac{(-1)^i}{i!} \big(x\partial^i(P)\big) y x^i \\
&= \sum_{i=0}^{n-1} \frac{(-1)^{i-1}}{i!} \partial^{i}(P) y x^i + \sum_{i=0}^{n-1} \frac{(-1)^i}{i!} \big(x\partial^i(P)\big) y x^i \\
&= -s(P)x + xs(P).
\end{aligned}
$$

Thus, $f = [x, s(P)]$.

It remains only to show that $s(P)$ is a Lie element. Let $\Phi : \mathbb{Q}_{n\geqslant 1}\langle x,y\rangle \to \mathrm{Lie}[x,y]$ be the linear map sending a non-trivial word $w = x_1x_2x_3\cdots x_m$ to $[x_1,[x_2,[x_3,\ldots]]]$, where $x_i \in \{x,y\}$, and let $\theta : \mathbb{Q}\langle x,y\rangle \to \mathrm{End}_{\mathbb{Q}}\mathrm{Lie}[x,y]$ be the algebra homomorphism mapping $x$ to $ad(x)$ and $y$ to $ad(y)$. By [B, Chap. 2, §3, no. 2] the following properties hold:

- a polynomial $h \in \mathbb{Q}_n\langle x,y\rangle$ is Lie if and only if $\Phi(h) = nh$;
- $\Phi(uv) = \theta(u)\Phi(v)$ for $u \in \mathbb{Q}\langle x,y\rangle$ and $v \in \mathbb{Q}_{n\geqslant 1}\langle x,y\rangle$;
- $\theta(u)(v) = [u,v]$ if $u$ is Lie.

Since $f \in \mathrm{Lie}[x, y]$, we have

$$[f, x] = \theta(f)(x) = \theta\big(\big[x, s(P)\big]\big)(x) = \big[ad(x), \theta\big(s(P)\big)\big](x) = \big[x, \theta\big(s(P)\big)(x)\big] = -\big[\theta\big(s(P)\big)(x), x\big].$$

Thus, $[f + \theta(s(P))(x), x] = 0$, so since both $f$ and $\theta(s(P))(x)$ are Lie elements of degree $> 1$, we have $f = -\theta(s(P))(x)$. Thus,

$$\begin{aligned} nf &= \Phi(f) = \Phi\big(\big[x, s(P)\big]\big) = \theta(x)\Phi\big(s(P)\big) - \theta\big(s(P)\big)\Phi(x) \\ &= \big[x, \Phi\big(s(P)\big)\big] - \theta\big(s(P)\big)(x) = \big[x, \Phi\big(s(P)\big)\big] + f. \end{aligned}$$

Thus $[x, \Phi(s(P))] = (n-1)f = (n-1)[x, s(P)]$, so $[x, \Phi(s(P)) - (n-1)s(P)] = 0$. Since $s(P)$ is of degree $n-1 > 1$, we must have $\Phi(s(P)) = (n-1)s(P)$, but this means that $s(P) \in \mathrm{Lie}_{n-1}[x, y]$. $\square$

**Proposition 2.3.** *Let $f \in \mathrm{Lie}_{n \geqslant 3}[x, y]$, and set $F = f(z, y) = F_x x + F_y y$ and $G = s'(F_x)$. Then $D_{F,G}$ is special if and only if $f$ is special, and this is the case if and only if $F_y$ is antipalindromic.*

**Proof.** If setting $G = s'(F_x)$, the derivation $D_{F,G}$ is special, then $f$ is special by definition. Conversely, if $f$ is special, there exists a unique $G \in \mathrm{Lie}_n[x, y]$ such that $[y, F] + [x, G] = 0$. Setting $H = yF - Fy = Gx - xG$ and writing $F = F_x x + F_y y = xF^x + yF^y$ and $G = G_x x + G_y y = xG^x + yG^y$, this means that

$$H = yF_y y + yF_x x - yF^y y - xF^x y = xG^x x + yG^y x - xG_x x - xG_y y, \tag{2.3}$$

so comparing the terms starting with $x$ and ending with $y$, we find that $-xF^x y = -xG_y y$, so $F^x = G_y$. By a result of Racinet [R], since $G$ is a Lie element, we must have $G = s(G_y) = s(F^x) = s'(F_x)$. This proves the first equivalence.

Let us now assume that $F_y$ is antipalindromic, i.e. by (2.1), $F^y = F_y$. Set

$$H = yF - Fy = y(F_y y + F_x x) - \big(yF^y + xF^x\big)y = yF_y y - yF^y y + yF_x x - xF^x y. \tag{2.4}$$

This shows that $H$ has no words starting and ending in $y$, so by Proposition 2.2, there exists $G \in \mathrm{Lie}_{n-1}[x, y]$ such that $H = Gx - xG$. But then the derivation $D_{F,G}$ is special, so $f$ is special.

Finally, assume that $f$ is special, and set $H = yF - Fy$, so that there exists $G$ with $H = yF - Fy = Gx - xG$. Then (2.3) holds. The expression $H = Gx - xG$ shows that $H$ can have no terms starting and ending in $y$, and the left-hand expression for $H$ in (2.3) then shows that we must have $F_y = F^y$, i.e. by (2.1), $F_y$ is antipalindromic. $\square$

**Proposition 2.4.** *Let $F \in \mathrm{Lie}_n\langle x, y\rangle$. Then $F_y$ is antipalindromic if and only if $F$ is push-invariant.*

**Proof.** As usual, we write $F = F_x x + F_y y = xF^x + yF^y$. Assume first that $F_y$ is antipalindromic, i.e. that $F_y = F^y$. Since $F$ is a Lie polynomial, we have $F = s(F_y) = s'(F^y) = s'(F_y)$, i.e.

$$F = \sum_{i \geqslant 0} \frac{(-1)^i}{i!} \partial_x^i(F_y) y x^i = \sum_{i \geqslant 0} \frac{(-1)^i}{i!} x^i y \partial_x^i\big(F^y\big) = \sum_{i \geqslant 0} \frac{(-1)^i}{i!} x^i y \partial_x^i(F_y). \tag{2.5}$$

Using the second and fourth terms of (2.5), we compute the coefficient of a word in $F$ as

$$\begin{aligned}\big(F\big|x^{a_0}y\cdots x^{a_{r-1}}yx^{a_r}\big) &= \frac{(-1)^{a_r}}{(a_r)!}\big(\partial_x^{a_r}(F_y)yx^{a_r}\big|x^{a_0}y\cdots yx^{a_{r-1}}yx^{a_r}\big)\\ &= \frac{(-1)^{a_r}}{(a_r)!}\big(\partial_x^{a_r}(F_y)\big|x^{a_0}y\cdots yx^{a_{r-1}}\big)\\ &= \frac{(-1)^{a_r}}{(a_r)!}\big(x^{a_r}y\partial_x^{a_r}(F_y)\big|x^{a_r}yx^{a_0}y\cdots yx^{a_{r-1}}\big)\\ &= \big(F\big|x^{a_r}yx^{a_0}y\cdots yx^{a_{r-1}}\big),\end{aligned}$$

so $F$ is push-invariant.

In the other direction, suppose that $F$ is push-invariant, and let's show that $F_y = F^y$. By assumption, we have

$$\big(F\big|x^{a_0}y\cdots yx^{a_r}\big) = \big(F\big|x^{a_r}yx^{a_0}y\cdots x^{a_{r-1}}\big).$$

In particular, for all words with $a_r = 0$, we have $(F|x^{a_0}y\cdots yx^{a_{r-1}}y) = (F|yx^{a_0}y\cdots yx^{a_{r-1}})$, i.e.

$$\big(F_y y\big|x^{a_0}y\cdots yx^{a_{r-1}}y\big) = \big(yF^y\big|yx^{a_0}y\cdots yx^{a_{r-1}}\big),$$

so

$$\big(F_y\big|x^{a_0}y\cdots yx^{a_{r-1}}\big) = \big(F^y\big|x^{a_0}y\cdots yx^{a_{r-1}}\big).$$

Thus $F_y = F^y$. $\square$

**Lemma 2.5.** *Let $g \in \mathbb{Q}_n\langle x, y\rangle$, let $\phi(x, y)$ and $\psi(x, y)$ be linear expressions of the form $ax + by$, $a, b \in \mathbb{Q}$, and let $h(x, y) = g(\phi(x, y), \psi(x, y))$. If $g$ is antipalindromic, then $h$ is antipalindromic.*

**Proof.** The operator *anti* is an anti-automorphism of the ring $\mathbb{Q}\langle x, y\rangle$, so

$$anti(h) = anti\big(g\big(anti(\phi), anti(\psi)\big)\big).$$

But *anti* fixes linear expressions $ax + by$, so since $g$ is antipalindromic, we have

$$anti(h) = anti\big(g(\phi, \psi)\big) = anti(g)(\phi, \psi) = (-1)^{n-1}g(\phi, \psi) = (-1)^{n-1}h.$$

Thus $h$ is antipalindromic. $\square$

**Proposition 2.6.** *For any $g \in \mathrm{Lie}_n[x, y]$, set $z = -x - y$ and $G = g(z, y)$. Write $g = g_x x + g_y y$ and $G = G_x x + G_y y$. Then*

$$G_y - G_x = g_y(z, y).$$

*In particular, $g_y$ is antipalindromic if and only if $G_y - G_x$ is antipalindromic.*

**Proof.** We have $g(x, y) = g_x(x, y)x + g_y(x, y)y$, so

$$G = g(z, y) = g_x(z, y)z + g_y(z, y)y = -g_x(z, y)x - g_x(z, y)y + g_y(z, y)y.$$

Thus $G_y = -g_x(z, y) + g_y(z, y)$ and $G_x = -g_x(z, y)$, so $G_y - G_x = g_y(z, y)$. Then by Lemma 2.5, since $g_y$ is antipalindromic, so is $G_y - G_x$, and the converse holds as well since $(G_y - G_x)(z, y) = g_y$. □

We can now conclude the proof of Theorem 2.1 by showing the equivalence of iii) and v). To do this, we simply apply Proposition 2.6 with $f = G$ and $g = F$, to see that $F_y$ is antipalindromic if and only if $f_y - f_x$ is antipalindromic. This completes the proof.

## 3. Proofs of Theorems 1.2 and 1.1

**Proof of Theorem 1.2.** Let $F \in V_{kv}$. We may assume that $F$ is homogeneous of degree $n \geqslant 3$, i.e. $F \in \mathrm{Lie}_n[x, y]$ with $n \geqslant 3$. Set $G = s'(F_x) = s(F^x)$. By Theorem 2.1, $F_y$ is antipalindromic if and only if $F$ is push-invariant, and these conditions are equivalent to the fact that $G \in \mathrm{Lie}_n[x, y]$ and $D_{F,G}$ is special.

Now consider the map $F \mapsto D_{F,G}$ from $V_{kv}$ to the vector space of special derivations, and let us show that it is injective. Suppose that $F, F' \in V_{kv}$ and $D_{F,G} = D_{F',G'}$. Then $D_{F,G}(y) = D_{F',G'}(y)$, i.e. $[y, F] = [y, F']$, so $F - F'$ commutes with $y$. Since $F - F'$ is of degree $> 1$, this means that $F - F' = 0$.

Let us now show that $D_{F,G}$ satisfies the trace formula (1.3). Note that by (2.1), $F_y - F_x = (-1)^{n-1} anti(F^y - F^x)$, so by symmetry, $F_y - F_x$ is push-constant if and only if $F^y - F^x$ is push-constant. It is convenient to use the latter condition.

Since any Lie polynomial of degree $> 1$ is a sum of terms of the form $fg - gf$, Lie polynomials map to zero in *TR*. Thus, we have $tr(G_x x) = -tr(G_y y)$, so

$$\begin{aligned} tr(F_y y + G_x x) &= tr(F_y y - G_y y) \\ &= tr\big(F_y y - F^x y\big) \quad \text{since } G_y = F^x \\ &= tr\big(F^y y - F^x y\big) \quad \text{since } F^y = F_y \\ &= tr\big(\big(F^y - F^x\big)y\big). \end{aligned} \tag{3.1}$$

Rephrasing the trace formula (1.3) via (3.1) as

$$tr\big(\big(F^y - F^x\big)y\big) = A\, tr\big((x + y)^n - x^n - y^n\big), \tag{3.2}$$

we can now show that a special derivation $D_{F,G}$ satisfies the trace formula in *TR* if and only if $F^y - F^x$ is push-constant. In fact, these are just two ways of making the identical statement. To see this, let $\bar{C}$ denote the *list* of words in the cyclic permutation class of $w$, so that $\bar{C}$ contains exactly $n$ words; then $\bar{C}$ consists of $n/|C|$ copies of $C$. For any word $v = uy$ ending in $y$, let $\bar{C}$ denote the associated cyclic permutation list, and $\bar{C}_y$ the list obtained from $\bar{C}$ by removing all the words ending in $x$. Write $\bar{C}_y = [u_1 y, \dots, u_r y]$. Then by definition, we have the equality of lists

$$[u_1, \dots, u_r] = Push(u). \tag{3.3}$$

Now, the trace condition $tr((F^y - F^x)y) = A\, tr((x + y)^n - x^n - y^n)$ means firstly that $((F^y - F^x)y|y^n) = 0$, which is equivalent to $(F^y - F^x|y^{n-1}) = 0$, and secondly that for each equivalence class $C$ of cyclic permutations of a given word $w \neq y^n$, we have

$$\big(tr\big(\big(F^y - F^x\big)y\big)\big|C\big) := \sum_{v \in C} \big(\big(F^y - F^x\big)y\big|v\big) = |C|A,$$

where the first equality is just the definition of the coefficient of an equivalence class in a trace polynomial. Using the notation $\bar{C}$ and $\bar{C}_y$ as above and (3.3), this means that for every word $v \neq y^n$ ending in $y$, writing $v = uy$, we have

$$|C|A = \frac{|C|}{n}\sum_{v\in\bar{C}}\big(\big(F^y - F^x\big)y\big|v\big) = \frac{|C|}{n}\sum_{v\in\bar{C}_y}\big(\big(F^y - F^x\big)y\big|v\big) = \frac{|C|}{n}\sum_{u'\in Push(u)}\big(\big(F^y - F^x\big)\big|u'\big). \quad (3.4)$$

But this is equivalent to

$$\sum_{u'\in Push(u)}\big(\big(F^y - F^x\big)\big|u'\big) = nA \quad (3.5)$$

for all $u' \neq y^{n-1}$, which, together with the fact that $(F^y - F^x|y^{n-1}) = 0$, is precisely equivalent to the statement that $F^y - F^x$ is push-constant (for the constant $nA$).

So far we have proven that $F \mapsto D_{F,s(F^x)}$ for homogeneous $F$ extends to an injective map $V_{kv} \hookrightarrow \mathfrak{krv}_2$. Let us show that it is an isomorphism, i.e. also surjective. It is enough to consider derivations $D_{F,G} \in \mathfrak{krv}_2$ with $F$, $G$ homogeneous of degree $n$. Then $D_{F,G}$ is special, so $F_y$ is antipalindromic by Theorem 2.1, and $D_{F,G}$ satisfies the trace formula (3.2), which as we just saw is equivalent to the property that $F^y - F^x$ is push-constant. Finally, since $F_y - F_x = (-1)^{n-1} anti(F^y - F^x)$, we see that $F_y - F_x$ is also push-constant, so $F \in V_{kv}$, completing the proof. □

Let us now prove Theorem 1.1. The proof is based on the fact that two previously known combinatorial results about double shuffle elements $\tilde{f} \in \mathfrak{ds}$ make it possible to deduce that $F = \tilde{f}(x, -y)$ satisfies the two defining properties of $V_{kv}$ given in Theorem 1.2. Thus $\tilde{f} \mapsto F$ yields an injection $\mathfrak{ds} \hookrightarrow V_{kv}$, and the injection $V_{kv} \hookrightarrow \mathfrak{krv}_2$ of Theorem 1.2 completes the argument.

The two known results are given in Theorems 3.1 and 3.2. As the original statement of Theorem 3.1 is extremely different in appearance (Theorem A.1 below), the translation from the original terminology to the statement given here is provided in Appendix A, which also serves as an initiation to Ecalle's language. We write $\mathfrak{ds}_n$ for the homogeneous weight $n$ part of $\mathfrak{ds}$, consisting of polynomials in $\mathfrak{ds}$ which are of homogeneous degree $n$.

**Theorem 3.1.** *(See [E, cf. Appendix].) Let $\tilde{f} \in \mathfrak{ds}_n$, and write $\tilde{f} = \tilde{f}_x x + \tilde{f}_y y$. Then $\tilde{f}_x + \tilde{f}_y$ is antipalindromic.*

**Theorem 3.2.** *(See [CS].) Let $\tilde{f} = \tilde{f}_x x + \tilde{f}_y y \in \mathfrak{ds}_n$, and set $A = (\tilde{f}|x^{n-1}y)$. Then $\tilde{f}_y$ satisfies the property that $(\tilde{f}_y|y^{n-1}) = 0$ and for each degree $n$ monomial $w \neq y^{n-1}$ containing $r$ $y$'s, we have*

$$\sum_{v\in Push(w)} (\tilde{f}_y|v) = (-1)^r A.$$

**Proof of Theorem 1.1.** Let $n \geqslant 3$ and assume that $\tilde{f} \in \mathfrak{ds}_n$, i.e. $\tilde{f}$ is a homogeneous Lie polynomial of degree $n$. Set $f(x, y) = \tilde{f}(x, -y)$. It follows directly from Theorem 3.2 that $f_y$ is push-constant. Let us deduce from Theorem 3.1 that $f_y - f_x$ is antipalindromic. Indeed, $f(x, y) = f_x(x, y)x + f_y(x, y)y$ and $\tilde{f}(x, y) = f(x, -y)$, so $\tilde{f}(x, y) = f_x(x, -y)x - f_y(x, -y)y$, i.e. $\tilde{f}_x = f_x(x, -y)$, $\tilde{f}_y = -f_y(x, -y)$. Thus

$$\tilde{f}_x + \tilde{f}_y = f_x(x, -y) - f_y(x, -y) = (f_x - f_y)(x, -y). \quad (3.6)$$

The left-hand side is antipalindromic by Theorem 3.1, so the right-hand side is antipalindromic, and then by Lemma 2.5 $f_x - f_y$ and thus also $f_y - f_x$ are antipalindromic.

Set $F = f(z, y)$. We will use the two properties on $f$ to show that $f \mapsto F$ is an injection from $\mathfrak{ds}$ into $V_{kv}$. By Proposition 2.6 with $g = F$ and $G = f$, we see that $f_y - f_x$ antipalindromic implies that $F_y$ is antipalindromic. It remains only to show that $f_y$ push-constant implies that $F_y - F_x$ is push-constant, which is a little more delicate. We prove it in the following lemma.

**Lemma 3.3.** *For any $f \in \mathrm{Lie}_n[x, y]$, set $F = f(z, y)$ and write $f = f_x x + f_y y$ and $F = F_x x + F_y y$. Suppose that $f_y$ is push-constant for a constant $A$, and that $A = 0$ if $n$ is even. Then $F_y - F_x$ is also push-constant for $A$.*

**Proof.** To show that $F_y - F_x$ is push-constant for $A$, let us first show that $(F_y - F_x|y^{n-1}) = 0$. As we saw in the proof of Theorem 1.2, the condition that $f_y$ is push-constant is equivalent to the condition that $tr(f_y y) = A\,tr((x+y)^n - x^n - y^n)$. By Proposition 2.6 with $g = f$ and $G = F$, we have $F_y - F_x = f_y(z, y)$, so $(F_y - F_x)(z, y) = f_y$. Multiplying by $y$ on the right of both sides and taking the trace yields

$$tr\big((F_y - F_x)(z, y)y\big) = tr(f_y y) = A\,tr\big((x+y)^n - x^n - y^n\big).$$

Making the variable change $x \mapsto z = -x - y$ on both sides, this gives

$$tr\big((F_y - F_x)y\big) = A\,tr\big((-1)^n x^n - (-1)^n (x+y)^n - y^n\big).$$

When $n$ is odd, the right-hand side does not contain the equivalence class of $y^n$, so the left-hand side cannot contain it either, which means that $(F_y - F_x|y^{n-1}) = 0$. When $n$ is even, $A = 0$ by assumption, so the equivalence class of $y^{n-1}$ cannot appear in the left-hand side, which again means that $(F_y - F_x|y^{n-1}) = 0$.

Now let us prove that $F_y - F_x$ is push-constant. Write

$$f_y = \sum_{\mathbf{a}} c_{\mathbf{a}} x^{a_0} y \cdots y x^{a_r} = \sum_v c_v v,$$

where $\mathbf{a}$ runs over the tuples $\mathbf{a} = (a_0, \ldots, a_r)$ with $r \geqslant 1$ and $a_0 + \cdots + a_r = n - r - 1$, and $v$ runs over degree $n-1$ words. If $v = x^{a_0} y \cdots y x^{a_r}$, we write $c_v = c_{\mathbf{a}}$. For a given tuple $\mathbf{a} = (a_0, \ldots, a_r)$, let

$$Push(\mathbf{a}) = \big[(a_0, \ldots, a_r), (a_r, a_0, \ldots, a_{r-1}), \ldots, (a_1, \ldots, a_r, a_0)\big]$$

be the list of its $r+1$ cyclic permutations. The fact that $f_y$ is push-constant means that for all $w \neq y^{n-1}$, we have

$$\sum_{v \in Push(w)} (f_y|v) = \sum_{v \in Push(w)} c_v = \sum_{\mathbf{a}' \in Push(\mathbf{a})} c_{\mathbf{a}'} = A. \tag{3.7}$$

Let us now compute the coefficient in $F_y - F_x$ of a given word $w = x^{b_0} y \cdots y x^{b_d}$, $w \neq y^{n-1}$. By Proposition 2.6, we have

$$F_y - F_x = f_y(z, y) = \sum_{\mathbf{a}} c_{\mathbf{a}} z^{a_0} y \cdots y z^{a_r} = \sum_{\mathbf{a}} (-1)^{n-r-1} c_{\mathbf{a}} (x+y)^{a_0} y \cdots y (x+y)^{a_r}, \tag{3.8}$$

so

$$(F_y - F_x|w) = (-1)^{n-1}\left(\sum_{\mathbf{a}} (-1)^r c_{\mathbf{a}}\, (x+y)^{a_0} y \cdots y (x+y)^{a_r} \Big| w\right). \tag{3.9}$$

Clearly if $r > d$ then the expansion of $(x+y)^{a_0} y \cdots y (x+y)^{a_r}$ cannot contain the word $w$, so (3.9) is equal to

$$(F_y - F_x|w) = (-1)^{n-1} \sum_{\mathbf{a} \text{ s.t. } 0 \leqslant r \leqslant d} (-1)^r c_{\mathbf{a}} \big((x+y)^{a_0} y \cdots y (x+y)^{a_r} \big| x^{b_0} y \cdots y x^{b_d}\big). \tag{3.10}$$

The only terms $(x+y)^{a_0} y \cdots (x+y)^{a_r}$ in which $w$ will appear with a positive coefficient (necessarily equal to 1) are the $2^d$ terms $(x+y)^{a_0} y \cdots (x+y)^{a_r}$ constructed as follows: choose any of the $2^d$

subsets of the $y$'s in $w$, and change the $y$'s in that subset to $x$'s; then substitute $x \mapsto (x+y)$ in the resulting word.

Let $w = x^{b_0} y \cdots y x^{b_d}$ be a monomial, and set $\mathbf{b} = (b_0, \ldots, b_d)$. Write $X_{\mathbf{b}}$ for the set of $2^d$ sequences $(a_0, \ldots, a_r)$, $0 \leqslant r \leqslant d$, such that the corresponding word $x^{a_0} y \cdots y x^{a_r}$ is obtained from $w$ by changing any subset of $y$'s into $x$'s. Then the coefficient (3.10) is equal to

$$(F_y - F_x | w) = (-1)^{n-1} \sum_{\mathbf{a} \in X_{\mathbf{b}}} (-1)^r c_{\mathbf{a}}. \tag{3.11}$$

By (3.11), we have

$$\sum_{v \in Push(w)} (F_y - F_x | v) = (-1)^{n-1} \sum_{\mathbf{c} \in Push(\mathbf{b})} \sum_{\mathbf{a} \in X_{\mathbf{c}}} (-1)^r c_{\mathbf{a}}. \tag{3.12}$$

Let us write

$$\mathcal{X}_{\mathbf{b}} = \coprod_{\mathbf{c} \in Push(\mathbf{b})} X_{\mathbf{c}}$$

for the disjoint union, i.e. the list-union of the words in the lists $X_{\mathbf{c}}$, where $\mathbf{c}$ runs through the cyclic permutations of $\mathbf{b}$. There are $(d+1)2^d$ words in $\mathcal{X}_{\mathbf{b}}$. Let us count the words in $\mathcal{X}_{\mathbf{b}}$ of each given depth $0 \leqslant r \leqslant d$.

For each tuple $\mathbf{c} \in Push(\mathbf{b})$, let $w_{\mathbf{c}}$ be the word associated to $\mathbf{c}$. The list $\mathcal{X}_{\mathbf{b}}$ is exactly the list of all words obtained by changing $k$ of the $d$ $y$'s in $w_{\mathbf{c}}$ to $x$'s, for all $0 \leqslant k \leqslant d$ and all $\mathbf{c} \in Push(\mathbf{b})$. Thus, $\mathcal{X}_{\mathbf{b}}$ contains $(d+1)$ words of depth $d$, which are the words $w_{\mathbf{c}}$ for $\mathbf{c} \in Push(\mathbf{b})$, and for each smaller depth $r = d-k$ for $1 \leqslant k \leqslant d$, $\mathcal{X}_{\mathbf{b}}$ contains the words obtained by changing $k$ $y$'s to $x$'s in each of the $d+1$ words (all of depth $d$) of $Push(\mathbf{b})$. Thus, there are exactly $(d+1)\binom{d}{k}$ words of depth $r = d-k$ in $\mathcal{X}_{\mathbf{b}}$, and these words fall into exactly

$$\frac{d+1}{d-k+1}\binom{d}{k} = \binom{d+1}{k}$$

cycles of length $r+1 = d-k+1$, of words of depth $r = d-k$.

Since $f_y$ is push-constant, the coefficients $c_{\mathbf{a}}$ of each of the $\binom{d+1}{k}$ cycles of depth $r = d-k$ in $f_y$ add up to $A$. Thus, for all $\mathbf{b} \neq (1, \ldots, 1)$, (3.12) is given by

$$\begin{aligned}(-1)^{n-1} \sum_{\mathbf{a} \in \mathcal{X}_b} (-1)^r c_{\mathbf{a}} &= (-1)^{n-1} \sum_{k=0}^{d} \binom{d+1}{k} (-1)^{d-k} A \\ &= (-1)^n \sum_{k=0}^{d} \binom{d+1}{k} (-1)^{d+1-k} A = (-1)^{n-1} A.\end{aligned}$$

This proves that $F_y - F_x$ is push-constant for the value $(-1)^{n-1}A$. $\square$

We can now conclude the proof of Theorem 1.1. Using the well-known result on $\mathfrak{ds}$ (cf. [E,R,IKZ,...]) that the coefficient of $x^{n-1}y$ is zero for all even-degree elements of $\mathfrak{ds}$, we see that when $n$ is even, $A = 0$ in Theorem 3.2, so if $\tilde{f} \in \mathfrak{ds}$, then $f = \tilde{f}(x, -y)$ satisfies the hypotheses of Lemma 3.3.

Thus, we have shown so far that if $\tilde{f} \in \mathfrak{ds}$, setting $f(x,y) = \tilde{f}(x,-y)$ and $F = f(z,y)$, $F_y$ is antipalindromic by the argument of the first paragraph of the proof of Theorem 1.1, and $F_y - F_x$ is push-constant by Lemma 3.3. Thus, the map $\tilde{f} \mapsto F$ is an injective map from $\mathfrak{ds} \to V_{kv}$. By (1.8), we then have an injective composition of maps

$$\mathfrak{ds} \hookrightarrow V_{kv} \xrightarrow{\sim} \mathfrak{krv}_2.$$

This concludes the proof of Theorem 1.1. □

## 4. The prounipotent version

Let $V = \bigoplus_{n=0}^{\infty} V_n$ be a graded $\mathbb{Q}$-vector space, and let $\mathfrak{un}(V)$ denote the Lie algebra of (pro)unipotent endomorphisms of $V$, i.e. linear endomorphisms $D$ such that $D(V_{\geqslant n}) \subset V_{\geqslant n+1}$ for all $n \geqslant 0$. The usual exponentiation

$$\exp(D) = \sum_{n\geqslant 0} \frac{1}{n!} D^n \tag{4.1}$$

maps $\mathfrak{un}(V)$ bijectively to the group $UN(V)$ of (pro)unipotent linear automorphisms of $V$.

Suppose we now have a Lie algebra $\mathfrak{g}$ equipped with an injective Lie algebra map $\mathfrak{g} \overset{\rho}{\hookrightarrow} \mathfrak{un}(V)$. The universal enveloping algebra $\mathcal{U}\mathfrak{g}$ is a ring whose multiplication we denote by $\odot$. The exponential associated to $\mathfrak{g}$ is given by the formula

$$\exp^{\odot}(f) = \sum_{n\geqslant 0} \frac{1}{n!} f^{\odot n}; \tag{4.2}$$

it maps $\mathfrak{g}$ bijectively to the associated group $G \subset \widehat{\mathcal{U}\mathfrak{g}}$, and the following diagram commutes:

$$\begin{array}{ccc} G & \longrightarrow & UN(V) \\ {\scriptstyle \exp^{\odot}}\uparrow & & \uparrow{\scriptstyle \exp} \\ \mathfrak{g} & \xrightarrow{\rho} & \mathfrak{un}(V). \end{array} \tag{4.3}$$

Now let $V$ denote the underlying vector space of $\mathrm{Lie}[x, y]$. Following the notation of [AT], let $\mathfrak{tder}_2$ denote the Lie algebra of *tangential derivations* of $\mathrm{Lie}[x, y]$, i.e. derivations $D$ having the property that $D(x) = [x, a]$ and $D(y) = [y, b]$ for elements $a, b \in \mathrm{Lie}[x, y]$. There is an injective map of Lie algebras $\mathfrak{tder}_2 \hookrightarrow \mathfrak{un}(V)$. Indeed, if $V = \mathrm{Lie}[x, y]$ is equipped with the grading given by the degree, then any derivation $D \in \mathfrak{tder}_2$ increases the degree, i.e. $D(V_{\geqslant n}) \subset V_{\geqslant n+1}$. Let $TAut_2$ denote the group of automorphisms of $V$ obtained by exponentiating $\mathfrak{tder}_2$:

$$\begin{aligned} \exp : \mathfrak{tder}_2 &\to TAut_2 \subset UN(V) \\ D &\mapsto \exp(D) = \sum_{n\geqslant 0} \frac{1}{n!} D^n. \end{aligned} \tag{4.4}$$

Let $\mathfrak{sder}_2$ denote the subalgebra of $\mathfrak{tder}_2$ consisting of derivations $D$ such that $D(x+y) = 0$, and $SAut_2$ the corresponding subgroup of $TAut_2$ consisting of automorphisms $A$ such that $A(x+y) = x+y$, so that $\exp(\mathfrak{sder}_2) = SAut_2$. According to [AT], the exponential map (4.1) not only restricts to (4.4), but also to maps from the following subspaces to subgroups:

$$\begin{array}{ccccccc} KRV_2 & \longrightarrow & SAut_2 & \longrightarrow & TAut_2 & \longrightarrow & UN(V) \\ \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} & & \uparrow{\scriptstyle \exp} \\ \mathfrak{krv}_2 & \longrightarrow & \mathfrak{sder}_2 & \longrightarrow & \mathfrak{tder}_2 & \longrightarrow & \mathfrak{un}(V) \end{array} \tag{4.5}$$

where the upper left-hand group, $KRV_2$, is the prounipotent group actually defined as the image in $SAut_2$ of $\mathfrak{krv}_2 \subset \mathfrak{sder}_2$ under the exponential map, although the authors then go on to also provide a direct description of $KRV_2$ [AT, §5.1].

Let us now recall the definition of the prounipotent group version $DS$ of the double shuffle Lie algebra $\mathfrak{ds}$ originally given by Racinet in [R, Chap. 4, §1]; this is the group that Racinet denotes $DM_0(\mathbf{k})$, but we take the base field $\mathbf{k} = \mathbb{Q}$; note that he also writes $\mathfrak{dm}_0(\mathbf{k})$ for $\mathfrak{ds}$.

For any monomials $u, v \in \mathbb{Q}\langle x, y\rangle$, let the *shuffle product* $sh(u, v) \in \mathbb{Q}\langle x, y\rangle$ be defined recursively by

$$sh(1, u) = sh(u, 1) = u, \qquad sh(xu, yv) = x\,sh(u, yv) + y\,sh(xu, v). \tag{4.6}$$

It is well-known that the condition for a polynomial $f \in \mathbb{Q}\langle x, y\rangle$ to be a Lie polynomial is equivalent to the condition

$$\big(f \big| sh(u, v)\big) = 0 \tag{4.7}$$

for all pairs of words $(u, v)$. The elements of the double shuffle Lie algebra $\mathfrak{ds}$ are thus defined by (4.7) and the stuffle condition

$$\big(f_* \big| st(u, v)\big) = 0 \tag{4.8}$$

for all words $u$, $v$ ending in $y$, where $f_* = \pi_y(f) + \sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}(f|x^{n-1}y)y^n$ (cf. footnote to Section 1).

Let $DS$ be the group consisting of power series in $\Phi \in \mathbb{Q}\langle\langle x, y\rangle\rangle$ having constant term 1, no degree 1 or 2 terms, and satisfying two properties, which are essentially group-like analogs of (4.7) and (4.8), namely

$$\big(\Phi \big| sh(u, v)\big) = \Phi(u)\Phi(v) \tag{4.9}$$

for all pairs of words $(u, v)$ and

$$\big(\Phi_* \big| st(u, v)\big) = \Phi_*(u)\Phi_*(v) \tag{4.10}$$

for all pairs of words $(u, v)$ both ending in $y$, where

$$\Phi_* = \exp\left(\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}\big(\Phi \big| x^{n-1}y\big)y^n\right)\pi_y(\Phi).$$

The elements of $\mathfrak{ds}$ are Lie polynomials; as we saw in Section 1, the main result of [R] states that $\mathfrak{ds}$ is a Lie algebra under the Poisson bracket (1.1). If $f \in \mathfrak{ds}$, then for any $g$ in the universal enveloping algebra $\mathcal{U}\mathfrak{ds}$, the multiplication in $\mathcal{U}\mathfrak{ds}$ is given by the explicit formula $f \odot g = fg + D_f(g)$. Thus for $f \in \mathfrak{ds}$, one can define $f^{\odot n} = f \odot f^{\odot n-1}$, which gives an explicit polynomial formula for $f^{\odot n}$. The exponential map of the Lie algebra is then given by $\exp^{\odot}(f) = \sum_{n\geqslant 0} \frac{1}{n!} f^{\odot n}$ as in (4.2). In [R, Chap. 4, §3.3, Corollaire 3.11] Racinet showed, using a method based on induction on the degree, that

$$\exp^{\odot}(\mathfrak{ds}) \simeq DS. \tag{4.11}$$

The next theorem shows that there exists an injective map $DS \to KRV_2$, the group analog of the Lie algebra map of Theorem 1.1. Given the results above on $KRV_2$ and $DS$, this is in fact nothing more than an immediate corollary of Theorem 1.1.

**Theorem 4.1.** *There is an injective homomorphism of prounipotent groups $DS \hookrightarrow KRV_2$ making the following diagram commute*:

$$\begin{tikzcd} DS \arrow[r, hook] & KRV_2 \\ \mathfrak{ds} \arrow[u, "\exp^{\odot}"] \arrow[r, hook] & \mathfrak{krv}_2. \arrow[u, "\exp"'] \end{tikzcd}$$

**Proof.** Let $\rho : \mathfrak{ds} \hookrightarrow \mathfrak{krv}_2 \subset \mathfrak{sder}_2$ be the map of Theorem 1.1; then the following diagram commutes by definition:

$$\begin{tikzcd} \exp(\rho(\mathfrak{ds})) \arrow[r, hook] & KRV_2 \arrow[r, hook] & SAut_2 \\ \rho(\mathfrak{ds}) \arrow[u, "\exp"] \arrow[r, hook] & \mathfrak{krv}_2 \arrow[u, "\exp"'] \arrow[r, hook] & \mathfrak{sder}_2, \arrow[u, "\exp"'] \end{tikzcd} \tag{4.12}$$

where all the horizontal injections are just inclusions.

We also have the commutative diagram

$$\begin{tikzcd} DS \arrow[r] & \exp(\rho(\mathfrak{ds})) \\ \mathfrak{ds} \arrow[u, "\exp^{\odot}"] \arrow[r] & \rho(\mathfrak{ds}), \arrow[u, "\exp"'] \end{tikzcd}$$

where the left vertical arrow is Racinet's isomorphism (4.11), the right vertical arrow is the exponential isomorphism from (4.12), the bottom arrow is the isomorphism $\rho$ from Theorem 1.1, and the top arrow is simply the isomorphism defined by these other three arrows. Then the composition

$$DS \xrightarrow{\sim} \exp\bigl(\rho(\mathfrak{ds})\bigr) \subset KRV_2$$

is the desired injection. $\square$

## Acknowledgments

Pierre Lochak and Samuel Baumard both provided arguments for the second half of Proposition 2.2, the latter being eventually used as it was shorter. Much of the spirit of the approach introduced here emerges from the reading of the works of Jean Ecalle, who always insists that the situation must be studied entirely via the symmetries that occur. The terminology *anti*, *push* etc. is introduced purposely here with a view to eventually providing a more general introduction to his papers. Finally, we warmly thank the referee for a very detailed and patient job with a great many useful suggestions and corrections, in particular the addition of the final section of this paper.

## Appendix A. Ecalle's theorem

Let $n \geqslant 1$, and let $f \in \mathbb{Q}_n\langle x, y\rangle$ be a weight $n$ homogeneous polynomial. For $0 \leqslant r \leqslant n$, let $\mathcal{E}_n^r$ be the set of $r$-tuples

$$\mathcal{E}_n^r = \left\{ \mathbf{e} = (e_0, \ldots, e_r) \;\middle|\; e_i \geqslant 0,\ \sum_{i=0}^{r} e_i = n \right\},$$

and set $\mathcal{E}_n = \bigcup_{r=1}^{n} \mathcal{E}_n^r$. Then we can write

$$f^r = \sum_{\mathcal{E}_n^r} a_{\mathbf{e}} x^{e_0} y \cdots y x^{e_r} \quad \text{for } 0 \leqslant r \leqslant n, \quad \text{and} \quad f = \sum_{r=0}^{n} f^r = \sum_{\mathcal{E}_n} a_{\mathbf{e}} x^{e_0} y \cdots y x^{e_r}. \tag{A.1}$$

To each such $f$, we associate two families of polynomials. The first family, $vimo_f$, is a set of polynomials in commutative variables $z_i$, and the second family, $ma_f$, is in commutative variables $u_i$. We set

$$vimo_f^r(z_0, \ldots, z_r) = \sum_{\mathbf{e}=(e_0,\ldots,e_r)} a_{\mathbf{e}} z_0^{e_0} \cdots z_r^{e_r}, \quad 0 \leqslant r \leqslant n, \tag{A.2}$$

$$ma_f^r(u_1, \ldots, u_r) = vimo_f^r(0, u_1, u_1 + u_2, \ldots, u_1 + \cdots + u_r), \quad 1 \leqslant r \leqslant n. \tag{A.3}$$

Note that if $f$ is a Lie polynomial, $vimo_f^0 = vimo_f^n = ma_f^n = 0$.

Ecalle calls a *mould* any family $fa$ of functions $fa^r(u_1, \ldots, u_r)$, $r \geqslant 0$, with $fa^0$ being a constant in a specified field. He considers arbitrary functions, but in this appendix it is enough to consider only polynomial-valued moulds $ma$ which are such that for each depth $r$, we have $ma^r(u_1, \ldots, u_r) \in \mathbb{Q}[u_1, \ldots, u_r]$ (and $ma^0 = 0$). For any fixed integer $n \geqslant 1$, such a polynomial mould is said to be *homogeneous of degree n* if $ma^r(u_1, \ldots, u_r)$ is a homogeneous polynomial of degree $n - r$.

Ecalle defines the following transformations of a mould $ma$ with $ma^0 = 0$: $swap(ma)$, $mantar(ma)$, $push(ma)$ and $teru(ma)$, by specifying their depth $r$ parts as follows:

$$swap(ma)^r(v_1, \ldots, v_r) = ma^r(v_r, v_{r-1} - v_r, \ldots, v_1 - v_2), \tag{A.4}$$

$$mantar(ma)^r(u_1, \ldots, u_r) = (-1)^{r-1} ma^r(u_r, \ldots, u_1), \tag{A.5}$$

$$push(ma)^r(u_1, \ldots, u_r) = ma^r(-u_1 - \cdots - u_r, u_1, \ldots, u_{r-1}), \tag{A.6}$$

$$\begin{aligned} teru(ma)^r(u_1, \ldots, u_r) = ma^r(u_1, \ldots, u_r) + \frac{1}{u_r}\big(&ma^{r-1}(u_1, \ldots, u_{r-2}, u_{r-1} + u_r) \\ &- ma^{r-1}(u_1, \ldots, u_{r-2}, u_{r-1})\big). \end{aligned} \tag{A.7}$$

The result of Ecalle that we use here is the following.

**Theorem A.1.** *(See Ecalle [E, §3.5, (3.64)].) Let $n \geqslant 3$ and let $\tilde{f} \in \mathfrak{ds}_n$, so that $\tilde{f}$ is a homogeneous polynomial of degree $n$; in particular $\tilde{f}^r = 0$ if $r = 0$ or $r \geqslant n$. Let $ma$ be the mould $ma_{\tilde{f}}$ associated to $\tilde{f}$ as in* (A.3)*. Then $ma$ is a homogeneous mould of degree $n$, and for $1 \leqslant r \leqslant n$, we have*

$$teru(ma)^r = push \circ mantar \circ teru \circ mantar(ma)^r. \tag{A.8}$$

The purpose of this appendix is to show that this theorem is equivalent to Theorem 3.1, by translating Ecalle's language back into terms of the non-commutative variables $x$, $y$. The first observation is that $mantar(ma) = ma$, because $ma$ comes from a Lie polynomial.

**Lemma A.2.** *Let $f \in \mathrm{Lie}_n[x, y]$ be a polynomial of homogeneous depth $r \geqslant 1$, and let ma be the mould associated to $f$ as in* (A.3)*. Then $mantar(ma) = ma$.*

**Proof.** Let $f$ be a polynomial of homogeneous degree $n \geqslant 3$ all of whose terms are of fixed depth $r$; we write it as in (A.1) (with only the fixed value of $r$ giving non-zero terms). By the Lazard elimination theorem, any Lie polynomial belongs to the polynomial ring generated by the polynomials $ad(x)^{i-1}(y)$ for $i \geqslant 1$. Thus, we can write

$$f = \sum_{\mathbf{c}} b_{\mathbf{c}}\, ad(x)^{c_1}(y) \cdots ad(x)^{c_r}(y).$$

We can show that we then have

$$ma_f^r(u_1, \dots, u_r) = \sum_{\mathbf{c}} b_{\mathbf{c}} u_1^{c_1} \cdots u_r^{c_r};$$

in other words, the meaning of the coefficients of the mould $ma_f$ is that they reflect the expression of $f$ as a polynomial in the $C_i$. This idea was expressed by Racinet in [R, Appendix A], but the proof is not given there. It can be done by induction; the complete proof is given in Chapter 3 of the unpublished manuscript [S].

Now, if $P = ad(x)^{c-1}(y)$ and $anti(P)$ is as usual the polynomial obtained from $P$ by writing all its words backwards, then $anti(P) = (-1)^{c-1}P$. It follows that if $P$ is a product $P = ad(x)^{c_1-1}(y) \cdots ad(x)^{c_r-1}(y)$ and $P' = ad(x)^{c_r-1}(y) \cdots ad(x)^{c_1-1}(y)$, we have $P' = (-1)^{c_1+\cdots+c_r-r} anti(P)$. Now assume that $f \in \mathrm{Lie}_n[x, y]$, so $(-1)^{n-1} anti(f) = f$. This means that

$$\begin{aligned} f = (-1)^{n-1} anti(f) &= (-1)^{n-1} \sum_{\mathbf{c}} b_{\mathbf{c}} (-1)^{c_1+\cdots+c_r-r} ad(x)^{c_r-1}(y) \cdots ad(x)^{c_1-1}(y) \\ &= (-1)^{r-1} \sum_{\mathbf{c}} b_{\mathbf{c}'}\, ad(x)^{c_1-1}(y) \cdots ad(x)^{c_r-1}(y), \end{aligned}$$

where if $\mathbf{c} = (c_1, \dots, c_r)$, we write $\mathbf{c}' = (c_r, \dots, c_1)$, so that in particular $b_{\mathbf{c}'} = (-1)^{r-1} b_{\mathbf{c}}$. Then

$$\begin{aligned} mantar(ma_f)^r(u_1, \dots, u_r) = (-1)^{r-1} ma_f^r(u_r, \dots, u_1) &= (-1)^{r-1} \sum_{\mathbf{c}} b_{\mathbf{c}} u_1^{c_r} \cdots u_r^{c_1} \\ &= \sum_{\mathbf{c}} b_{\mathbf{c}'} u_1^{c_r} \cdots u_r^{c_1} = \sum_{\mathbf{c}} b_{\mathbf{c}} u_1^{c_1} \cdots u_r^{c_r} = ma_f^r(u_1, \dots, u_r). \end{aligned}$$

This concludes the proof. $\square$

The statement of Ecalle's theorem (A.8) for $r = 1$ is easy to prove, since by (A.7), $teru(ma)^1(u_1) = ma^1(u_1)$, and $push(ma^1(u_1)) = ma^1(-u_1)$. Now, if $n$ is even, it is well-known that if $\tilde{f} \in \mathfrak{ds}_n$, then $\tilde{f}^1 = 0$, so $ma^1(u_1) = 0$ and (A.8) holds. If $n$ is odd, then either $\tilde{f} \in \mathfrak{ds}_n$ also satisfies $\tilde{f}^1 = 0$, so that again (A.8) holds, or $\tilde{f}^1 = a\, ad(x)^{n-1} y$, in which case $ma^1(u_1) = a u_1^{n-1}$, so $push(ma^1(u_1)) = ma^1(-u_1) = ma^1(u_1)$.

Let us now give a reformulation of (A.8) for $2 \leqslant r \leqslant n$. By Lemma A.2, we can rewrite (A.8) as

$$swap \circ teru(ma)^r = swap \circ push \circ mantar \circ teru(ma)^r. \tag{A.9}$$

The swap is obviously not necessary in the equality, but useful for the computation below as it is easier to compute both sides as polynomials in the commutative variables $v_i$.

By applying (A.4) to (A.7), we see that for $2 \leqslant r \leqslant n$, the left-hand side is given by

$$\begin{aligned}
&swap\big(teru(ma)^r\big)(v_1,\ldots,v_r)\\
&\quad= ma^r(v_r, v_{r-1}-v_r,\ldots,v_1-v_2) + \frac{1}{v_1-v_2}\big(ma^{r-1}(v_r, v_{r-1}-v_r,\ldots,v_3-v_4, v_1-v_3)\\
&\qquad - ma^{r-1}(v_r, v_{r-1}-v_r,\ldots,v_3-v_4,v_2-v_3)\big)\\
&\quad= vimo^r(0,v_r,\ldots,v_1) + \frac{1}{v_1-v_2}\big(vimo^{r-1}(0,v_r,\ldots,v_3,v_1) - vimo^{r-1}(0,v_r,\ldots,v_3,v_2)\big),
\end{aligned} \tag{A.10}$$

where *vimo* is the mould associated to $\tilde{f}$ as in (A.2).

Let us calculate the right-hand side of (A.9) one step at a time using (A.4)–(A.7).

$$\begin{aligned}
&swap \circ push \circ mantar \circ teru(ma)^r\\
&\quad= swap\circ push\circ mantar\big(ma^r(u_1,\ldots,u_r)\big)\\
&\qquad + swap\circ push\circ mantar\left(\frac{1}{u_r}\big(ma^{r-1}(u_1,\ldots,u_{r-2},u_{r-1}+u_r) - ma^{r-1}(u_1,\ldots,u_{r-2},u_{r-1})\big)\right)\\
&\quad= (-1)^{r-1} swap\circ push\big(ma^r(u_r,\ldots,u_1)\big)\\
&\qquad + (-1)^{r-1} swap\circ push\left(\frac{1}{u_1}\big(ma^{r-1}(u_r,\ldots,u_3,u_1+u_2) - ma^{r-1}(u_r,\ldots,u_3,u_2)\big)\right)\\
&\quad= (-1)^{r-1} swap\big(ma^r(u_{r-1},\ldots,u_2,u_1,-u_1-\cdots-u_r)\big)\\
&\qquad + (-1)^{r-1} swap\left(\frac{1}{(-u_1-\cdots-u_r)}\big(ma^{r-1}(u_{r-1},\ldots,u_2,-u_2-\cdots-u_r)\right.\\
&\qquad \left. - ma^{r-1}(u_{r-1},\ldots,u_2,u_1)\big)\right)\\
&\quad= (-1)^{r-1} ma^r(v_2-v_3,\ldots,v_{r-1}-v_r,v_r,-v_1)\\
&\qquad + (-1)^{r-1}\frac{1}{-v_1}\big(ma^{r-1}(v_2-v_3,\ldots,v_{r-1}-v_r,v_r-v_1)\\
&\qquad - ma^{r-1}(v_2-v_3,\ldots,v_{r-1}-v_r,v_r)\big)\\
&\quad= (-1)^{r-1} vimo^r(0,v_2-v_3,\ldots,v_2-v_r,v_2,v_2-v_1)\\
&\qquad + \frac{(-1)^r}{v_1}\big(vimo^{r-1}(0,v_2-v_3,\ldots,v_2-v_r,v_2-v_1) - vimo^{r-1}(0,v_2-v_3,\ldots,v_2-v_r,v_2)\big).
\end{aligned} \tag{A.11}$$

The following useful elementary identities will simplify the form of (A.11): for any *vimo* associated to a polynomial as in (A.2), we have

$$vimo^r(z_0,\ldots,z_r) = (-1)^{n-r} vimo^r(-z_0,\ldots,-z_r), \tag{A.12}$$

and if *vimo* is associated to a Lie polynomial, then

$$vimo^r(z_0, z_1, \dots, z_r) = vimo^r(0, z_1 - z_0, \dots, z_r - z_0). \tag{A.13}$$

Note that the meaning of (A.13) is that any value (called $z_0$) can be added to each argument of $vimo^r$ without changing the value of the function. Let us quickly indicate the easy proof of (A.13) by induction. For $r = 1$, up to scalar multiple, we must have

$$f^1 = ad(x)^m(y) = \sum_{i=0}^{m}(-1)^i\binom{m}{i}x^{m-i}yx^i, \quad \text{so } vimo^1(z_0, z_1) = \sum_{i=0}^{m}(-1)^i\binom{m}{i}z_0^{m-i}z_1^i,$$

which is equal to $(-1)^m(z_1 - z_0)^m = vimo^1(0, z_1 - z_0)$. Now assume that (A.13) holds up to depth $r - 1$ and consider a Lie polynomial $f$ of homogeneous depth $r$. By linearity, we may assume that $f = [g, h]$, where $g$ and $h$ are homogeneous depths $s < r$ and $t < r$ respectively, with $r = s + t$. Then we have

$$vimo_f^r(z_0, \dots, z_r) = vimo_g^s(z_0, \dots, z_s)vimo_h^t(z_s, \dots, z_{s+t}) - vimo_h^t(z_0, \dots, z_t)vimo_g^s(z_t, \dots, z_{s+t}),$$

so using repeated applications of (A.13) to the $vimo_g$ and $vimo_h$ factors by the induction hypothesis, we have

$$\begin{aligned}
vimo_f^r(0, z_1 - z_0, \dots, z_r - z_0) &= vimo_g^s(0, z_1 - z_0, \dots, z_s - z_0)vimo_h^t(z_s - z_0, \dots, z_{s+t} - z_0) \\
&\quad - vimo_h^t(0, z_1 - z_0, \dots, z_t - z_0)vimo_g^s(z_t - z_0, \dots, z_{s+t} - z_0) \\
&= vimo_g^s(z_0, z_1, \dots, z_s)vimo_h^t(0, z_{s+1} - z_s, \dots, z_{s+t} - z_s) \\
&\quad - vimo_h^t(z_0, z_1, \dots, z_t)vimo_g^s(0, z_{t+1} - z_t, \dots, z_{s+t} - z_t) \\
&= vimo_g^s(z_0, z_1, \dots, z_s)vimo_h^t(z_s, z_{s+1}, \dots, z_{s+t}) \\
&\quad - vimo_h^t(z_0, z_1, \dots, z_t)vimo_g^s(z_t, z_{t+1}, \dots, z_{s+t}) \\
&= vimo_f^r(z_0, \dots, z_r).
\end{aligned}$$

This yields (A.13).

Now, applying (A.12) to (A.11) yields

$$\begin{aligned}
&(-1)^{n-1}vimo^r(0, v_3 - v_2, \dots, v_r - v_2, -v_2, v_1 - v_2) \\
&\quad + \frac{(-1)^{n-1}}{v_1}\big(vimo^{r-1}(0, v_3 - v_2, \dots, v_r - v_2, v_1 - v_2) \\
&\quad - vimo^{r-1}(0, v_3 - v_2, \dots, v_r - v_2, -v_2)\big).
\end{aligned}$$

Applying (A.13) to this with $z_0 = v_2$, i.e. adding $v_2$ to each argument of $vimo^r$, then yields

$$\begin{aligned}
(-1)^{n-1}\bigg[&vimo^r(v_2, v_3, \dots, v_r, 0, v_1) + \frac{1}{v_1}\big(vimo^{r-1}(v_2, v_3, \dots, v_r, v_1) \\
&- vimo^{r-1}(v_2, v_3, \dots, v_r, 0)\big)\bigg].
\end{aligned} \tag{A.14}$$

Since if *vimo* is the mould associated to a polynomial $\tilde{f} \in \mathfrak{ds}_n$, then $vimo^0 = vimo^n = 0$, Ecalle's theorem can be expressed by the equalities (A.10) = (A.14) for $1 \leqslant r \leqslant n$.

Let us now show that the statement of Theorem 3.1 can be deduced from the equalities (A.10) = (A.14) for $2 \leqslant r \leqslant n$.

**Proposition A.3.** *Let $\tilde{f} \in \mathfrak{ds}_n$ for $n \geqslant 3$, and write $\tilde{f} = \tilde{f}_x x + \tilde{f}_y y$. Then $\tilde{f}_x + \tilde{f}_y$ is antipalindromic.*

**Proof.** We will show the identity

$$(\tilde{f}_x + \tilde{f}_y)^r = (-1)^{n-1} anti(\tilde{f}_x + \tilde{f}_y)^r,$$

separately for each depth $1 \leqslant r \leqslant n-1$ occurring in $\tilde{f}$.

These equalities are equivalent to the equalities of polynomials in commutative variables

$$vimo_{\tilde{f}_x^r + \tilde{f}_y^r}(z_0, \ldots, z_r) = (-1)^{n-1} vimo^r_{anti(\tilde{f}_x^r + \tilde{f}_y^r)}(z_0, \ldots, z_r) \tag{A.15}$$

for $1 \leqslant r \leqslant n-1$.

To prove the proposition, we will deduce (A.15) from Ecalle's theorem, i.e. from the set of equalities (A.10) = (A.14) for $2 \leqslant r \leqslant n$. To do this, we explicitly compute both sides of (A.15).

Write $(\tilde{f}_x)^r$ (resp. $(\tilde{f}_y)^r$) for the depth $r$ part of the polynomial $\tilde{f}_x$ (resp. $\tilde{f}_y$), $1 \leqslant r \leqslant n-1$. Each term of the polynomial $(\tilde{f}_x + \tilde{f}_y)^r$ comes either from a term in $\tilde{f}^r$ ending with $x$, i.e. from $(\tilde{f}_x)^r x$, or from a term in $\tilde{f}^{r+1}$ ending with $y$, i.e. from $(\tilde{f}_y)^r y$, by cutting off the final letter. Let us find the *vimo* polynomials associated to $(\tilde{f}_y)^r y$ and $(\tilde{f}_x)^r x$.

Write $\tilde{f}^{r+1} = \sum_{\mathbf{e}=(e_0,\ldots,e_{r+1})} a_{\mathbf{e}} x^{e_0} y \cdots y x^{e_{r+1}}$. Since $\tilde{f}^{r+1}$ is homogeneous in depth $r+1$, we have

$$vimo_{\tilde{f}^{r+1}}(z_0, \ldots, z_{r+1}) = \sum_{\mathbf{e}=(e_0,\ldots,e_{r+1})} a_{\mathbf{e}} z_0^{e_0} \cdots z_{r+1}^{e_{r+1}}.$$

The polynomial $(\tilde{f}_y)^r y$ consists of the terms of $\tilde{f}^{r+1}$ ending in $y$, so we have $(\tilde{f}_y)^r y = \sum_{\mathbf{e}=(e_0,\ldots,e_r,0)} a_{\mathbf{e}} x^{e_0} y \cdots x^{e_r} y$, and the depth $r+1$ polynomial $vimo_{(\tilde{f}_y)^r y}$ is given by

$$vimo_{(\tilde{f}_y)^r y}(z_0, \ldots, z_{r+1}) = \sum_{\mathbf{e}=(e_0,\ldots,e_r,0)} a_{\mathbf{e}} z_0^{e_0} \cdots z_r^{e_r} = vimo_{\tilde{f}^{r+1}}(z_0, \ldots, z_r, 0).$$

Since we have $(\tilde{f}_y)^r = \sum_{\mathbf{e}=(e_0,\ldots,e_r)} a_{\mathbf{e}} x^{e_0} y \cdots y x^{e_r}$, we see that $vimo_{(\tilde{f}_y)^r}(z_0, \ldots, z_r) = vimo_{(\tilde{f}_y)^r y}(z_0, \ldots, z_{r+1})$, i.e.

$$vimo_{(\tilde{f}_y)^r}(z_0, \ldots, z_r) = vimo_{(\tilde{f}_y)^r y}(z_0, \ldots, z_{r+1}) = vimo_{\tilde{f}^{r+1}}(z_0, \ldots, z_r, 0). \tag{A.16}$$

To find the *vimo* associated to $\tilde{f}_x^r x$, we consider this polynomial as the difference $(\tilde{f}_x)^r x = \tilde{f}^r - (\tilde{f}_y)^{r-1} y$. Thus, using (A.16) for $r-1$ instead of $r$, we have

$$\begin{aligned} vimo_{(\tilde{f}_x)^r x}(z_0, \ldots, z_r) &= vimo_{\tilde{f}^r - (\tilde{f}_y)^{r-1} y}(z_0, \ldots, z_r) \\ &= vimo_{\tilde{f}^r}(z_0, \ldots, z_{r-1}, z_r) - vimo_{\tilde{f}^r}(z_0, \ldots, z_{r-1}, 0). \end{aligned} \tag{A.17}$$

Because we know that there is an $x$ at the end of every word of the polynomial $(\tilde f_x)^r x$, the polynomial in (A.17) is divisible by $z_r$, and we have

$$vimo_{(\tilde f_x)^r}(z_0,\dots,z_r)=\frac{1}{z_r}\big(vimo_{\tilde f^r}(z_0,\dots,z_{r-1},z_r)-vimo_{\tilde f^r}(z_0,\dots,z_{r-1},0)\big). \tag{A.18}$$

Putting (A.16) and (A.18) together yields the following expression for the left-hand side of the desired equality (A.15):

$$\begin{aligned}vimo_{(\tilde f_x+\tilde f_y)^r}(z_0,\dots,z_r)&=vimo_{\tilde f^{r+1}}(z_0,\dots,z_r,0)\\&\quad+\frac{1}{z_r}\big(vimo_{\tilde f^r}(z_0,\dots,z_{r-1},z_r)-vimo_{\tilde f^r}(z_0,\dots,z_{r-1},0)\big).\end{aligned} \tag{A.19}$$

Since *anti* corresponds to reversing the order of $z_0,\dots,z_r$, the right-hand side of (A.15) is then given by

$$\begin{aligned}&(-1)^{n-1}vimo_{anti((\tilde f_x+\tilde f_y)^r)}(z_0,\dots,z_r)\\&\quad=(-1)^{n-1}\left[vimo_{\tilde f^{r+1}}(z_r,\dots,z_0,0)+\frac{1}{z_0}\big(vimo_{\tilde f^r}(z_r,\dots,z_1,z_0)-vimo_{\tilde f^r}(z_r,\dots,z_1,0)\big)\right],\end{aligned} \tag{A.20}$$

so the statement of the proposition is equivalent to the set of equalities (A.19) = (A.20) for $1\leqslant r\leqslant n-1$.

Thus it remains only to show that Ecalle's set of equalities (A.10) = (A.14) for $2\leqslant r\leqslant n$ implies the set of equalities (A.19) = (A.20) for $1\leqslant r\leqslant n-1$. By (A.13), we can add the same quantity to every argument of *vimo* and not change its value, so we first use this to rewrite (A.10), by adding the quantity $-v_1$ to every argument of the three *vimo* terms in (A.10):

$$\begin{aligned}(\text{A.10})=\;&vimo_{\tilde f^r}(-v_1,v_r-v_1,\dots,v_2-v_1,0)\\&+\frac{1}{v_1-v_2}\big(vimo_{\tilde f^{r-1}}(-v_1,v_r-v_1,\dots,v_3-v_1,0)\\&-vimo_{\tilde f^{r-1}}(-v_1,v_r-v_1,\dots,v_3-v_1,v_2-v_1)\big).\end{aligned}$$

Now we apply the variable change

$$z_0=-v_1,\qquad z_1=v_r-v_1,\quad\dots,\quad z_{r-1}=v_2-v_1 \tag{A.21}$$

to this, to obtain

$$\begin{aligned}&=vimo_{\tilde f^r}(z_0,z_1,\dots,z_{r-1},0)\\&\quad+\frac{1}{-z_{r-1}}\big(vimo_{\tilde f^{r-1}}(z_0,z_1,\dots,z_{r-2},0)-vimo_{\tilde f^{r-1}}(z_0,\dots,z_{r-2},z_{r-1})\big).\end{aligned} \tag{A.22}$$

This is equivalent to (A.19), for $r-1$ instead of $r$.

Next, we use (A.13) to rewrite (A.14), adding the quantity $-v_1$ to every argument in the three *vimo* terms that appear in (A.14):

$$(-1)^{n-1}\Big[vimo_{\tilde{f}^r}(v_2-v_1,v_3-v_1,\dots,v_r-v_1,-v_1,0)$$
$$+\frac{1}{v_1}\big(vimo_{\tilde{f}^{r-1}}(v_2-v_1,v_3-v_1,\dots,v_r-v_1,0)$$
$$-vimo_{\tilde{f}^{r-1}}(v_2-v_1,v_3-v_1,\dots,v_r-v_1,-v_1)\big)\Big],$$

and then the variable change (A.21), which yields

$$(-1)^{n-1}\Big[vimo_{\tilde{f}^r}(z_{r-1},z_{r-2},\dots,z_1,z_0,0)$$
$$-\frac{1}{z_0}\big(vimo_{\tilde{f}^{r-1}}(z_{r-1},z_{r-2},\dots,z_1,0)-vimo_{\tilde{f}^{r-1}}(z_{r-1},z_{r-2},\dots,z_1,z_0)\big)\Big].$$

This is exactly (A.20) for $r-1$ instead of $r$. Thus Ecalle's equalities (A.10) = (A.14) for $2\leqslant r\leqslant n$ imply the desired equalities (A.15) for $1\leqslant r\leqslant n-1$ as desired. $\square$

## References


[AT] A. Alekseev, C. Torossian, The Kashiwara–Vergne Conjecture and Drinfeld's associators, Ann. of Math. 175 (2) (2012) 415–463.

[B] N. Bourbaki, Groupes et algèbres de Lie, Chapitres 2 et 3, Springer-Verlag, Berlin, 1982.

[CS] S. Carr, L. Schneps, Combinatorics of the double shuffle Lie algebra, in: H. Nakamura, A. Tamagawa (Eds.), Development of Galois–Teichmüller Theory and Anabelian Geometry, in: Adv. Stud. Pure Math., Math. Soc. Japan, 2013, in press.

[E] J. Ecalle, The flexion structure and dimorphy: flexion units, singulators, generators, and the enumeration of multizeta irreducibles, in: O. Costin, F. Fauvet, F. Menous, D. Sauzin (Eds.), Aymptotics in Dynamics, Geometry and PDEs; Generalized Borel Summation, vol. II, in: CRM Series, vol. 12, Edizioni della Normale, Pisa, 2011, pp. 27–211.

[F] H. Furusho, Double shuffle relations for associators, Ann. of Math. 174 (1) (2011) 341–360.

[IKZ] K. Ihara, M. Kaneko, D. Zagier, Derivation and double shuffle relations for multiple zeta values, Compos. Math. 142 (2) (2006) 307–338.

[R] G. Racinet, Séries génératrices non-commutatives de polyzêtas et associateurs de Drinfel'd, Ph.D. thesis, 2000.

[S] L. Schneps, ARI, GARI Zig and Zag, Ecalle's theory of multiple zeta values, book manuscript, 2011.