Advanced Studies in Pure Mathematics 63, 2012
Galois–Teichmüller Theory and Arithmetic Geometry
pp. 59–89

Combinatorics of the double shuffle Lie algebra

Sarah Carr and Leila Schneps

Abstract.

In this article we give two combinatorial properties of elements satisfying the stuffle relations; one showing that double shuffle elements are determined by less than the full set of stuffle relations, and the other a cyclic property of their coefficients. Although simple, the properties have some useful applications, of which we give two. The first is a generalization of a theorem of Ihara on the abelianizations of elements of the Grothendieck-Teichmüller Lie algebra $\mathfrak{grt}$ to elements of the double shuffle Lie algebra in a much larger quotient of the polynomial algebra than the abelianization, namely the trace quotient introduced by Alekseev and Torossian. The second application is a proof that the Grothendieck–Teichmüller Lie algebra $\mathfrak{grt}$ injects into the double shuffle Lie algebra $\mathfrak{ds}$, based on the recent proof by H. Furusho of this theorem in the pro-unipotent situation, but in which the combinatorial properties provide a significant simplification.

§1. The cyclic property

Write Y for the alphabet $\{y_1, y_2, y_3, \ldots\}$ and U for the alphabet $\{u_1, u_2, u_3, \ldots\}$, where y_i and u_i are given the weight i, and these two alphabets are related by the expression

(1)

$$u_1+u_2+\cdots = \log(1+y_1+y_2+\cdots) = (y_1+y_2+\cdots)-\frac{1}{2}(y_1+y_2+\cdots)^2+\cdots,$$

Received January 12, 2011.
Revised February 10, 2012.
2000 *Mathematics Subject Classification.* Primary 17B40, 17B65, 17B70, 12Y05, 05E99.
Key words and phrases. Lie algebras, double shuffle, multiple zeta values.

with the parts of equal weight on each side identified, so that

$$(2)\qquad \begin{cases} u_1 = y_1 \\ u_2 = y_2 - \frac{1}{2}y_1^2 \\ u_3 = y_3 - \frac{1}{2}y_1y_2 - \frac{1}{2}y_2y_1 + \frac{1}{3}y_1^3 \dots . \end{cases}$$

Let $\mathbb{Q}[U]$ be the non-commutative polynomial ring freely generated by the u_i, and $\mathbb{Q}[Y]$ be that generated by the y_i.

Definition 1.1. *For any two sequences* $\mathbf{a} = (a_1, \dots, a_r)$ *and* $\mathbf{b} = (b_1, \dots, b_s)$ *of integers* $a_i, b_j \geq 1$*, define the shuffle of* $\mathbf{a}$ *and* $\mathbf{b}$*,* $sh(\mathbf{a}, \mathbf{b})$*, to be the list (i.e. a set which may contain repeated elements) of sequences obtained as follows. We write* $b_1 = a_{r+1}, \dots, b_s = a_{r+s}$*; for any permutation* $\sigma \in S_{r+s}$ *such that* $\sigma(1) < \cdots < \sigma(r)$ *and* $\sigma(r+1) < \dots < \sigma(r+s)$*, we define the sequence* $(a_{\sigma^{-1}(1)}, \dots, a_{\sigma^{-1}(r+s)})$*. Note that these sequences may not all be distinct: the shuffle* $sh(\mathbf{a}, \mathbf{b})$ *is the complete list, with repetitions, of these sequences.*

Similarly, for any two sequences $\mathbf{a} = (a_1, \dots, a_r)$ *and* $\mathbf{b} = (b_1, \dots, b_s)$ *of integers* $a_i, b_j \geq 1$*, define the stuffle of* $\mathbf{a}$ *and* $\mathbf{b}$*,* $st(\mathbf{a}, \mathbf{b})$*, to be the list of sequences obtained as follows. We again write* $b_1 = a_{r+1}, \dots, b_s = a_{r+s}$ *as above. Following Furusho's notation, let* $\mathrm{Sh}^{\leq}(r,s)$ *denote the set of surjective maps* $\sigma : \{1, \dots, r+s\} \twoheadrightarrow \{1, \dots, N\}$*,* $N \leq r+s$*, such that* $\sigma(1) < \dots < \sigma(r)$ *and* $\sigma(r+1) < \dots < \sigma(r+s)$*. For each* $\sigma \in \mathrm{Sh}^{\leq}(r,s)$*, we define the sequence* $c^\sigma(\mathbf{a}, \mathbf{b}) = (c_1, \dots, c_N)$ *by*

$$c_i = \begin{cases} a_k + b_{l-r} & \text{if } \sigma^{-1}(i) = \{k, l\} \text{ with } k \leq r < l \\ a_k & \text{if } \sigma^{-1}(i) = \{k\} \text{ with } k \leq r \\ b_{k-r} & \text{if } \sigma^{-1}(i) = \{k\} \text{ with } k > r. \end{cases}$$

Again, it may happen that $c^\sigma(\mathbf{a}, \mathbf{b}) = c^\tau(\mathbf{a}, \mathbf{b})$ *even if* $\sigma \neq \tau$*, so that the stuffle* $st(\mathbf{a}, \mathbf{b})$ *may contain repeated elements: it is given by the list*

$$st(\mathbf{a}, \mathbf{b}) = [c^\sigma(\mathbf{a}, \mathbf{b}) |\ \sigma \in \mathrm{Sh}^{\leq}(r,s)].$$

For any sequence $\mathbf{a} = (a_1, \dots, a_r)$, we write $u_\mathbf{a}$ for the associated word $u_\mathbf{a} = u_{a_1} \cdots u_{a_r} \in \mathbb{Q}[U]$, and $y_\mathbf{a} = y_{a_1} \cdots y_{a_r}$ for the associated word in $\mathbb{Q}[Y]$. We define the shuffle of two words $u_\mathbf{a}$ and $u_\mathbf{b}$ in $\mathbb{Q}[U]$ as the list

$$sh(u_\mathbf{a}, u_\mathbf{b}) = [u_\mathbf{c} | \mathbf{c} \in sh(\mathbf{a}, \mathbf{b})],$$

and the stuffle of two words $y_\mathbf{a}$ and $y_\mathbf{b}$ in $\mathbb{Q}[Y]$ as the list

$$st(y_\mathbf{a}, y_\mathbf{b}) = [y_\mathbf{c} | \mathbf{c} \in st(\mathbf{a}, \mathbf{b})].$$

Example. Let $\mathbf{a} = \mathbf{b} = (2)$. Then $sh\big((2),(2)\big) = [(2,2),(2,2)]$, so $sh(u_2, u_2) = u_{(2,2)} + u_{(2,2)} = 2u_2^2$, and $st\big((2),(2)\big) = [(2,2),(2,2),(4)]$, so $st(y_2, y_2) = 2y_{(2,2)} + y_4 = 2y_2^2 + y_4$. For any word w and any polynomial f, we write $(f|w)$ for the coefficient of w in f. We say that a polynomial f^U in the variables u_i satisfies the *shuffle relations* if

$$\sum_{w \in sh(u_{\mathbf{a}}, u_{\mathbf{b}})} (f^U|w) = 0$$

for all pairs of words $u_{\mathbf{a}}, u_{\mathbf{b}}$. A polynomial f^Y in the variables y_i is said to satisfy the *stuffle relations* if

$$\sum_{w \in st(y_{\mathbf{a}}, y_{\mathbf{b}})} (f^Y|w) = 0.$$

Let Δ denote the coproduct defined on $\mathbb{Q}[U]$ by $\Delta(u_i) = u_i \otimes 1 + 1 \otimes u_i$ for each $i \geq 1$, and setting $y_0 = 1$, let Δ_* denote the coproduct defined on $\mathbb{Q}[Y]$ by

$$\Delta_*(y_k) = \sum_{i=0}^{k} y_i \otimes y_{k-i}. \tag{3}$$

We introduce the following notation. Let $\iota : \mathbb{Q}[U] \to \mathbb{Q}[Y]$ be the map given by sending u_i to the right-hand side of the corresponding equality in (2) for all $i \geq 1$. If f^U is a polynomial in the variables u_i, i.e. $f^U \in \mathbb{Q}[U]$, we write $f^Y = \iota(f^U) \in \mathbb{Q}[Y]$; similarly if we are given a polynomial f^Y in the variables y_i, we write $f^U = \iota^{-1}(f^Y)$.

It follows easily from the definitions that if f^Y is a polynomial in the y_i, then

$$\big(\Delta_*(f^Y)\big)^U = \Delta(f^U). \tag{4}$$

Indeed, it can be checked directly for $f^Y = y_i$, $i \geq 1$, and then follows by the multiplicativity of the coproducts. We have

$$\begin{aligned}
\iota^{-1}\Big(\Delta_*(\sum_i y_i t^i)\Big) &= \iota^{-1}\Big((\sum_i y_i t^i)\Big) \otimes \iota^{-1}\Big((\sum_i y_i t^i)\Big) \\
&= \exp(\sum_i u_i t^i) \otimes \exp(\sum_i u_i t^i) \\
&= \exp(\sum_i (u_i \otimes 1 + 1 \otimes u_i) t^i) \\
&= \Delta\Big(\exp(\sum_i u_i t^i)\Big).
\end{aligned} \tag{5}$$

The results gathered in the following lemma are well-known (see [9] for example).

Lemma 1. *Let $f^U \in \mathbb{Q}[U]$, and write $f^Y = \iota^{-1}(f^U)$ as above. Then the following are equivalent:*
(i) f^U lies in the Lie algebra $\mathrm{Lie}[U] \subset \mathbb{Q}[U]$, the free Lie algebra on the u_i.
(ii) f^U satisfies the shuffle relations for all pairs of words $u_{\mathbf{a}}, u_{\mathbf{b}} \in \mathbb{Q}[U]$.
(iii) $\Delta(f^U) = f^U \otimes 1 + 1 \otimes f^U$.
(iv) $\Delta_*(f^Y) = f^Y \otimes 1 + 1 \otimes f^Y$.
(v) f^Y satisfies the stuffle relations.

We do not reproduce the proof in full detail. The equivalence of (i) and (iii) is shown for example in [10]. The equivalence of (ii) and (iii) follows from a direct computation which shows that the shuffle sums of the coefficients of f^U are equal to coefficients of the terms of $\Delta(f^U)$, i.e. for all non-trivial sequences $\mathbf{a}, \mathbf{b}$, we have

$$\big(\Delta(f^U)|u_{\mathbf{a}} \otimes u_{\mathbf{b}}\big) = \sum_{\mathbf{c} \in sh(\mathbf{a},\mathbf{b})} (f^U|\mathbf{c}). \tag{6}$$

Similarly, (iv) and (v) are equivalent because again, a direct computation shows that the stuffle sums of the coefficients of f^Y are equal to coefficients of $\Delta_*(f^Y)$, i.e. for all non-trivial sequences $\mathbf{a}, \mathbf{b}$, we have

$$\big(\Delta_*(f^Y)|y_{\mathbf{a}} \otimes y_{\mathbf{b}}\big) = \sum_{\mathbf{c} \in st(\mathbf{a},\mathbf{b})} (f^Y|\mathbf{c}). \tag{7}$$

Finally, the equivalence between (iii) and (iv) follows from (1), or from (5). The following lemma is also well-known, but we give its short proof here.

Lemma 2. *For every word w in the variables $u_1, u_2, \ldots$ having $d > 1$ letters, set $w^j = u_{i_j} \cdots u_{i_d} u_{i_1} \cdots u_{i_{j-1}}$ for $1 \leq j \leq d$; in particular $w^1 = w$, but note that the w^j are not necessarily all distinct if w has symmetries. Then for any linear combination f of expressions of the form $[g,h] = gh - hg$ where g and h are non-trivial words in the u_i, we have*

$$\sum_{j=1}^{d} (f|w^j) = 0 \quad \text{for all } w. \tag{8}$$

In particular, this holds for every $f \in \mathrm{Lie}[U]$ of degree > 1.

Proof. Then the only words to appear with non-zero coefficient in the polynomial $f = gh - hg$ are the words gh and hg, and we have $(f|gh) = 1$ and $(f|hg) = -1$. All other words appear with coefficient zero, and furthermore gh is a cyclic permutation of hg, so (8) holds for f. Then (8) obviously holds for all linear combinations of expressions of the form $gh - hg$ by additivity. In particular, every element of Lie$[U]$ apart from the u_i themselves is such a linear combination. Q.E.D.

We now come to the definition and proof of the *cyclic property* of polynomials in the y_i satisfying the stuffle relations. This proof is due to J. Ecalle[1] and was initially done for stuffle polynomials and given in terms of moulds, a context in which it appears very naturally (as does the equivalence of (ii) and (v) in Lemma 1). Here, the stuffle result appears in the corollary, and the theorem itself is stated in a more general case, suggested by the referee. We only consider polynomials in the y_i or the u_i which are homogeneous of a given weight n, which means that all the monomials $y_{i_1} \cdots y_{i_r}$ (or $u_{i_1} \ldots u_{i_r}$) appearing with non-zero coefficients satisfy $i_1 + \cdots + i_r = n$.

Theorem 1. *Let f^U be a polynomial in the u_i $(i > 0)$, of homogeneous weight n, which is a linear combination of expressions of the form $[g, h] = gh - hg$ where g and h are non-trivial words in the u_i. Let $f^Y = \iota(f^U)$ be the same polynomial written in the variables y_i. For every word $w = y_{i_1} \cdots y_{i_d}$ in the y_i $(i > 0)$, set $n = i_1 + \cdots + i_d$, and for $1 \le j \le d$, set $w^j = y_{i_j} \cdots y_{i_d} y_{i_1} \cdots y_{i_{j-1}}$, so the w^j are the cyclic permutations of $w = w^1$. Then f^Y satisfies the cyclic property*

$$\sum_{j=1}^{d} (f^Y|w^j) = (-1)^{d-1}(f^Y|y_n) \quad \text{for all } w.$$

Proof. From the expression (1), we know that

$$(f^U|u_n) = (f^Y|y_n). \tag{9}$$

For any word $w = y_{k_1} \cdots y_{k_r}$ in Y, with $k_1, \ldots, k_r > 0$, we write $||w|| = y_{k_1+\cdots+k_r}$, $v_w = u_{k_1} \cdots u_{k_r}$, and $||v_w|| = u_{k_1+\cdots+k_r}$. Let $l(w)$ denote the length of a word in the y_i, i.e. the number of letters y_i in the word. For

[1]The statement appears in [4], equation (3.46), in Ecalle's language. The proof can be deduced by twisting (2.73) by the flexion unit $\mathcal{E}$ defined by $1/u$. However, in personal communication Ecalle gave a somewhat different, more direct suggestion, which gave rise to the proof of Theorem 1 given here.

any word w in the y_i, the relation between the alphabets (1) yields the relation

$$(10)\quad (f^Y|w) = \sum_{s=1}^{l(w)} \sum_{v_w=v_1\cdots v_s} \frac{(-1)^{l(v_1)-1}}{l(v_1)}\frac{(-1)^{l(v_2)-1}}{l(v_2)}\cdots\frac{(-1)^{l(v_s)-1}}{l(v_s)}(f^U|||v_1||\cdots||v_s||),$$

where the v_i are non-trivial words, corresponding to composition of power series. Set $v^j = v_{w^j}$ for $1 \le j \le d$, i.e. $v^j = u_{i_j}\cdots u_{i_d}u_{i_1}\cdots u_{i_{j-1}}$. For the third equality below, we use the fact that for $1 \le j \le d$, we have $||v^j|| = u_{i_1+\cdots+i_d} = u_n$. Using (10), we find that

$$\begin{aligned}
&\sum_{j=1}^{d}(f^Y|w^j)\\
&=\sum_{j=1}^{d}\sum_{s=1}^{d}\sum_{v^j=v_1^j\cdots v_s^j}\frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)}\frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)}\cdots\frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)}(f^U|||v_1^j||\cdots||v_s^j||)\\
&=\sum_{s=2}^{d}\sum_{j=1}^{d}\sum_{v^j=v_1^j\cdots v_s^j}\frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)}\frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)}\cdots\frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)}(f^U|||v_1^j||\cdots||v_s^j||)\\
&\qquad+\sum_{j=1}^{d}\frac{(-1)^{d-1}}{d}(f^U|||v^j||)\\
&=\sum_{s=2}^{d}\sum_{j=1}^{d}\sum_{v^j=v_1^j\cdots v_s^j}\frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)}\frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)}\cdots\frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)}(f^U|||v_1^j||\cdots||v_s^j||)\\
&\qquad+(-1)^{d-1}(f^U|u_n)\\
&=\sum_{s=2}^{d}\sum_{j=1}^{d}\sum_{v^j=v_1^j\cdots v_s^j}\frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)}\frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)}\cdots\frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)}(f^U|||v_1^j||\cdots||v_s^j||)\\
&\qquad+(-1)^{d-1}(f^Y|y_n),
\end{aligned}$$

where the last equality follows from (9). In order to conclude the proof, we need to show that the complicated term

$$\sum_{s=2}^{d}\sum_{j=1}^{d}\sum_{v^j=v_1^j\cdots v_s^j}\frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)}\frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)}\cdots\frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)}(f^U|||v_1^j||\cdots||v_s^j||)$$

from the above expression is equal to zero, leaving the desired formula

$$\sum_{j=1}^{d}(f^Y|w^j) = (-1)^{d-1}(f^Y|y_n).$$

In fact we will show that

(11)

$$\sum_{j=1}^{d} \sum_{v^j=v_1^j\cdots v_s^j} \frac{(-1)^{l(v_1^j)-1}}{l(v_1^j)} \frac{(-1)^{l(v_2^j)-1}}{l(v_2^j)} \cdots \frac{(-1)^{l(v_s^j)-1}}{l(v_s^j)} (f^U|||v_1^j||\cdots||v_s^j||) = 0$$

for each s, $2 \leq s \leq d$. The point is that this sum breaks up into smaller sums over cyclic permutations of words, so that it is zero by Lemma 2.

Let

$$(12) \qquad (f^U|\,||v_1^j||\,||v_2^j||\cdots||v_s^j||)$$

be a term appearing in this sum. Then we only need to show that for every $1 \leq i \leq s$, the term

$$(13) \qquad (f^U|\,||v_i^j||\,||v_{i+1}^j||\cdots||v_s^j||\,||v_1^j||\cdots||v_{i-1}^j||)$$

appears with the same coefficient, since these words are exactly the cyclic permutations of the word $||v_1^j||\cdots||v_s^j||$, each $||v_i^j||$ being a single u_k.

The term (13) appears in the double sum (11) as

$$(14) \qquad (f^U|\,||v_1^{j'}||\,||v_2^{j'}||\cdots||v_s^{j'}||),$$

with $j' = j + l(v_1^j) + \cdots + l(v_{i-1}^j)$ (where j and j' are considered mod d and between 1 and d), and $v^{j'} = v_1^{j'}\cdots v_s^{j'}$ where the grouping into pieces $v_k^{j'}$ is determined by $l(v_k^{j'}) = l(v_{i+k-1}^j)$ (and i and k are considered mod s and between 1 and s). Thus, (13) and (14) are equal. Furthermore, the coefficients of (12) and (14) in (11) are obviously equal, because the set of lengths of the pieces v_k^j and the set of lengths of the pieces $v_k^{j'}$ are equal for $1 \leq k \leq s$. Thus the double sum (11) breaks up into cyclic subsums each of which is zero by Lemma 2. This concludes the proof of Theorem 1. Q.E.D.

The following corollary is an immediate consequence of Theorem 1, since by Lemma 1, f^Y is a polynomial in the y_i satisfying the stuffle relations if and only if $f^U = \iota^{-1}(f^Y)$ is a Lie polynomial in the u_i, which thus satisfies the hypothesis of Theorem 1.

Corollary 1. *Let f^Y be a polynomial in Y of homogeneous weight $n > 0$ satisfying the stuffle relations. Then with notation as in Theorem 1, we have*

$$\sum_{j=1}^{d} (f^Y|w^j) = (-1)^{d-1}(f^Y|y_n).$$

Our next theorem shows that if a polynomial in the y_i satisfies all the "non-trivial" stuffle relations, then it can be easily modified to obtain an actual stuffle polynomial.

Theorem 2. *Let f^Y be a polynomial in the y_i of homogeneous weight n satisfying the stuffle relations*

$$\sum_{\mathbf{c}\in st(\mathbf{a},\mathbf{b})} (f^Y|\mathbf{c}) = 0 \tag{15}$$

for all pairs of sequences $(\mathbf{a},\mathbf{b})$ which are not both sequences of 1's. Set $g^Y = f^Y - ay_1^n$. Then

$$a = \frac{(-1)^{n-2}}{n}(f^Y|y_n) + (f^Y|y_1^n)$$

is the unique constant for which g^Y satisfies the stuffle relations for all pairs, and for this value of a, we have

$$(g^Y|y_1^n) = \frac{(-1)^{n-1}}{n}(g^Y|y_n).$$

Proof. Assume that f^Y satisfies (15) for all pairs $(\mathbf{a},\mathbf{b})$ not both sequences of 1's. Then by equation (7), we know that the coefficients of pairs of words $y_\mathbf{a}\otimes y_\mathbf{b}$ in $\Delta_*(f^Y)$ are zero whenever $\mathbf{a}$ and $\mathbf{b}$ are not both sequences of 1's. In other words, we can write

$$\Delta_*(f^Y) = f^Y\otimes 1 + 1\otimes f^Y + \sum_{i=1}^{n} c_i y_1^i \otimes y_1^{n-i}.$$

Making the variable change to the u_i and using (4), we then have

$$\big(\Delta_*(f^Y)\big)^U = \Delta(f^U) = f^U\otimes 1 + 1\otimes f^U + \sum_{i=1}^{n-1} c_i u_1^i \otimes u_1^{n-i}.$$

Now, for any monomial $u_{i_1}\cdots u_{i_r}$, the definition of Δ implies that $\Delta(u_{i_1}\cdots u_{i_r})$ is a linear combination of terms of the form $u_{j_1}\cdots u_{j_s}\otimes u_{j_{s+1}}\cdots u_{j_r}$, where the list $u_{j_1},\ldots,u_{j_r}$ is the same as the list $u_{i_1},\ldots,u_{i_r}$ in a different order. In particular, only terms of the form $u_1^i\otimes u_1^j$ with $i+j=n$ will appear in $\Delta(u_1^n)$, but inversely, if $(i_1,\ldots,i_r)\neq(1,\ldots,1)$, then no term of that form can appear in $\Delta(u_{i_1}\ldots u_{i_r})$. Thus, if g^U is a homogeneous polynomial of weight n in the u_i with $(g^U|u_1^n)=0$, then $\big(\Delta(g^U)|u_1^i\otimes u_1^j\big)=0$ for all $i+j=n$. Let $k=(f^U|u_1^n)$, and set

$g^U = f^U - ku_1^n$, so $(g^U|u_1^n) = 0$. Then

$$\begin{aligned}
\Delta(g^U) &= \Delta(f^U) - k\Delta(u_1)^n \\
&= f^U \otimes 1 + 1 \otimes f^U + \sum_{i=1}^{n-1} c_i u_1^i \otimes u_1^{n-i} - k\Delta(u_1)^n \\
&= g^U \otimes 1 + 1 \otimes g^U + ku_1^n \otimes 1 + 1 \otimes ku_1^n \\
&\quad -k\Delta(u_1^n) + \sum_{i=1}^{n-1} c_i u_1^i \otimes u_1^{n-i} \\
&= g^U \otimes 1 + 1 \otimes g^U. \qquad (16)
\end{aligned}$$

Here, the last equality follows from the fact that since as we just saw, $\big(\Delta(g^U)|u_1^i \otimes u_1^j\big) = 0$, there can be no terms of the form $u_1^i \otimes u_1^j$ in $\Delta(g^U)$. Therefore the terms of that form in the third line above must either be canceled out by terms of that form in $g^U \otimes 1 + 1 \otimes g^U$, but there are no such terms since g^U does not contain a power of u_1, or sum to zero. Thus they sum to zero, yielding the last equality. Thus, $g^U = f^U - ku_1^n$ is an element of Lie$[U]$, and therefore by Lemma 1, $g^Y = f^Y - ky_1^n$ satisfies the stuffle relations.

Let us now show that $k = a$ and that

$$(g^Y|y_1^n) = \frac{(-1)^{n-1}}{n}(g^Y|y_n).$$

Since g^Y satisfies all the stuffle relations, in particular it satisfies the relation for $\mathbf{a} = (1)$, $b = (1, \ldots, 1)$ with $n-1$ 1's. Here, $st(\mathbf{a}, \mathbf{b}) = st\big((1), (1, \ldots, 1)\big)$ is given by the list

$$\big[\underbrace{(1, \ldots, 1), \ldots, (1, \ldots, 1)}_{n}, (2, 1, \ldots, 1), (1, 2, \ldots, 1) \cdots, (1, \ldots, 1, 2)\big],$$

so

$$\sum_{\mathbf{c} \in st\big((1),(1,\ldots,1)\big)} (g^Y|y_{\mathbf{c}}) =$$

$$n(g^Y|y_1^n) + (g^Y|y_2 y_1^{n-2}) + (g^Y|y_1 y_2 y_1 \cdots y_1) + \cdots + (g^Y|y_1^{n-2} y_2) = 0.$$

But the sum over the words $y_2 y_1^{n-2}$, $y_1 y_2 y_1 \cdots y_1$ etc. is a sum over a cyclic orbit of words of length $d = n-1$, so by the corollary to Theorem 1, it is equal to $(-1)^{n-2}(g^Y|y_n)$. Thus we obtain the desired identity

$$n(g^Y|y_1^n) + (-1)^{n-2}(f^Y|y_n) = 0.$$

Plugging into this equation the identity $g^Y = f^Y - ky_1^n$ yields the value $k = a$ given in the statement. Q.E.D.

§2. First application: a generalization of Ihara's abelianization theorem

The applications of Theorems 1 (or its corollary) and 2 to the double shuffle Lie algebra are straightforward, though they appear to be new and quite useful. In this section, we use them to give a simple proof of a quite surprising generalization of a theorem of Ihara concerning the Grothendieck–Teichmüller Lie algebra $\mathfrak{grt}$. Let $\mathrm{Lie}[x,y]$ denote the free Lie algebra on two generators, graded by degree, and let $\mathrm{Lie}_n[x,y]$ denote the subvector space of $\mathrm{Lie}[x,y]$ consisting of Lie polynomials of homogeneous degree n.

Definition 2.1. *Let* $\mathrm{Lie}\, P_5$ *denote the Lie algebra of the pure sphere 5-strand braid group. It is generated by* x_{ij}, $1 \le i,j \le 5$, *subject to the relations* $x_{ii} = 0$, $x_{ij} = x_{ji}$, $[x_{ij}, x_{kl}] = 0$ *if* $\{i,j\} \cap \{k,l\} = \emptyset$, $\sum_{i=1}^{5} x_{ij} = 0$ *for each fixed* $j \in \{1,\ldots,5\}$, *and* $[x_{ij}, x_{ij} + x_{ik} + x_{jk}] = 0$ *for any triple of indices* i,j,k.

The weight n *graded part* $\mathfrak{grt}_n$ *of the Grothendieck–Teichmüller Lie algebra* $\mathfrak{grt}$ *is defined to be the vector space of elements* $f \in \mathrm{Lie}_n[x,y]$ *such that*

$$f(x_{12},x_{23}) + f(x_{34},x_{45}) + f(x_{51},x_{12}) + f(x_{23},x_{34}) + f(x_{45},x_{51}) = 0, \tag{17}$$

where the defining "pentagon" relation takes place in $\mathrm{Lie}\, P_5$.[2] *We set* $\mathfrak{grt} = \oplus_{n\ge 3}\mathfrak{grt}_n$.

Ihara proved in [7] that $\mathfrak{grt}$ is a Lie algebra under the Poisson bracket

$$\{f,g\} = [f,g] + D_f(g) - D_g(f), \tag{18}$$

where for every $f \in \mathrm{Lie}[x,y]$, D_f denotes the derivation of $\mathrm{Lie}[x,y]$ defined by $D_f(x) = 0$ and $D_f(y) = [y,f]$. Now let us proceed to define the double shuffle Lie algebra. Let $\mathcal{A}$ denote the polynomial algebra $\mathcal{A} = \mathbb{Q}[x,y]$ on two non-commutative variables x, y, and let $\mathcal{B} \subset \mathcal{A}$ denote the subalgebra generated by $y_1, y_2, y_3, \ldots$, where $y_k = x^{k-1}y$. Set $y_0 = 1$.

Let Δ be the coproduct on $\mathcal{A}$ defined by $\Delta(x) = x\otimes 1 + 1\otimes x$ and $\Delta(y) = y \otimes 1 + 1 \otimes y$. Let Δ_* be the coproduct on $\mathcal{B}$ defined on the generators y_k by equation (1.2). Let $\mathrm{Lie}_n[x,y]$ denote the homogeneous

[2] Note that the definition classically contained three separate conditions on f, but H. Furusho in [5] gave a remarkable proof that the single pentagon condition implies both the others, making them unnecessary.

parts of weight n of $\mathrm{Lie}[x,y]$. For any $f \in \mathrm{Lie}_n[x,y]$, considered as a polynomial in x and y, we write $f = f_x x + f_y y$, and set

$$f_* = f_y y + \frac{(-1)^{n-1}}{n}(f|x^{n-1}y)y^n.$$

Since f_* is a polynomial ending in y, we can rewrite it in the variables y_k, so it lies in $\mathcal{B}$.

Definition 2.2. *The weight n graded part $\mathfrak{ds}_n$ of the double shuffle Lie algebra $\mathfrak{ds}$ is defined by*

$$\mathfrak{ds}_n = \{f \in \mathrm{Lie}_n[x,y] | \Delta_*(f_*) = f_* \otimes 1 + 1 \otimes f_*\}.$$

We set $\mathfrak{ds} = \oplus_{n\geq 3}\mathfrak{ds}_n$. It was shown in [9] that $\mathfrak{ds}$ is a Lie algebra under the Poisson bracket defined in (18).

By Lemma 1, if f_* satisfies $\Delta(f_*) = f_* \otimes 1 + 1 \otimes f_*$, then f_* satisfies the stuffle relations. Thus Theorem 1 applies to the elements f_* associated to elements $f \in \mathfrak{ds}$. The following statement is an immediate corollary of Theorem 2 and will be used in §3. For a sequence of strictly positive integers $\mathbf{c} = (c_1, \ldots, c_r)$, we use the notation

$$(f|y_{\mathbf{c}}) = (f|y_{c_1} \cdots y_{c_r}) = (f|x^{c_1-1}y \cdots x^{c_r-1}y).$$

Theorem 3. *Let $f \in \mathrm{Lie}_n[x,y]$, with $n \geq 3$. Then $f \in \mathfrak{ds}$ if and only if*

$$\sum_{\mathbf{c}\in st(\mathbf{a},\mathbf{b})} (f|y_{\mathbf{c}}) = 0 \qquad (19)$$

for all pairs $(\mathbf{a},\mathbf{b}) \neq \big((1,\ldots,1),(1,\ldots,1)\big)$.

Proof. Since all the words $y_{\mathbf{c}}$ end in y when considered in the variables x and y, we have $(f|y_{\mathbf{c}}) = (f_y y|y_{\mathbf{c}})$ whenever $\mathbf{c} \neq (1,\ldots,1)$. The sequence of 1's can only occur in the stuffle of $\mathbf{a}$ and $\mathbf{b}$ if both $\mathbf{a}$ and $\mathbf{b}$ are themselves sequences of 1's, so it never occurs in (19), and thus (19) is equivalent to the hypothesis of Theorem 2 on the polynomial $f_y y$, where we write $f = f_x x + f_y y$. Because f is a Lie element, we have $(f_y y|y^n) = 0$, and therefore the a of Theorem 2 is equal to $\frac{(-1)^n}{n}(f_y y|y_n)$, and Theorem 2 shows that $f_y y + \frac{(-1)^{n-1}}{n}(f_y y|y_n)y^n = f_*$ satisfies stuffle, so $f \in \mathfrak{ds}$. The useful point here is that this statement makes it possible to define elements of double shuffle via conditions on the Lie element f, making no reference to the much-studied "regularization" f_*. Q.E.D.

Now let us restate the corollary to Theorem 1 directly in the framework of the double shuffle Lie algebra.

Theorem 4. *Let* $f \in \mathfrak{ds} \subset \mathrm{Lie}[x,y]$, *let* $\mathbf{a} = (a_1,\ldots,a_d)$ *be a sequence of strictly positive integers, and for* $1 \le j \le d$, *set*

$$w^j = x^{a_j-1}y\cdots x^{a_d-1}yx^{a_1-1}y\cdots x^{a_{j-1}-1}y.$$

Then

$$\sum_{j=1}^{d}(f|w^j) = (-1)^{d-1}(f|x^{n-1}y). \tag{20}$$

We will use this theorem to generalize the following theorem, proved by Ihara in [8].

Theorem 5. (Ihara) *Let* $f \in \mathfrak{grt}_n$, *and write* $f = f_x x + f_y y$. *For any* $f \in \mathbb{Q}[x,y]$, *let* f^{ab} *denote the image of* f *in the abelianization of this ring; in particular, let* $X = x^{\mathrm{ab}}$ *and* $Y = y^{\mathrm{ab}}$. *Then*

$$(f_y y)^{\mathrm{ab}} = \frac{1}{n}(f|y_n)\big((X+Y)^n - X^n - Y^n\big), \tag{21}$$

Furthermore, if n *is even, then* $(f|y_n) = 0$, *so* $(f_y y)^{\mathrm{ab}} = 0$.

Our purpose in this section is to generalize this result of Ihara in two ways. To begin with, by Furusho's theorem from [6] (of which a simplified proof in the Lie case is given in §3), we know that $\mathfrak{grt}$ injects into $\mathfrak{ds}$ via the map $f(x,y) \mapsto f(x,-y)$. One conjectures that these two Lie algebras are isomorphic, but this is not known. We will prove our theorem for the a priori larger Lie algebra $\mathfrak{ds}$ rather than for $\mathfrak{grt}$.

But also, instead of working in the abelianization of $\mathbb{Q}[x,y]$, we prove the result in a much bigger quotient of $\mathbb{Q}[x,y]$, namely the trace quotient $Tr\langle x,y\rangle$ introduced by Alekseev and Torossian in [1]. The trace space is the quotient of $\mathbb{Q}[x,y]$ modulo the equivalence relation $uv \sim vu$ for every pair of monomials $u,v \in \mathbb{Q}[x,y]$. Although $xy = yx$ in $Tr\langle x,y\rangle$, it is not the abelianization; for example, we have $x^2y^2 \sim yx^2y \sim y^2x^2 \sim xy^2x$, but these words are not equivalent to $xyxy \sim yxyx$ which form a separate equivalence class. In fact, the equivalence classes of words under this relation are exactly the sets of cyclic permutations of words in x and y. The remarkable fact is that the statement of Ihara's theorem remains identical, not just when generalized from $\mathfrak{grt}$ to $\mathfrak{ds}$, which is natural considering that one believes the two Lie algebras to be isomorphic, but also when generalized from the abelianization to the trace quotient; even in this large quotient, all double shuffle elements of weight n become equal. In the following statement, we give the analog of (21), and afterwards we show that as in Ihara's theorem, this expression is equal to zero if n is even.

Theorem 6. *Let $g(x,y) \in \mathfrak{ds}$ be a homogeneous element of weight n, and set $f(x,y) = g(x,-y)$. Write $f = f_x x + f_y y$ and let $\overline{f_y y}$ denote the image of $f_y y$ in the quotient space $Tr\langle x,y\rangle$ of $\mathbb{Q}[x,y]$. Let $\bar{x}$ and $\bar{y}$ denote the images of x and y in $Tr\langle x,y\rangle$. Then*

$$\overline{f_y y} = \frac{1}{n}(f|y_n)\big((\bar{x}+\bar{y})^n - \bar{x}^n - \bar{y}^n\big). \tag{22}$$

Proof. The polynomial $(x+y)^n - x^n - y^n$ in $\mathbb{Q}[x,y]$ is equal to the sum of all words of weight n in x and y except for x^n and y^n. The image of this polynomial in $Tr\langle x,y\rangle$ is thus equal to a linear combination of the cyclic equivalence classes of these words, where the coefficient of each equivalence class is equal to the order of the cyclic class:

$$(\bar{x}+\bar{y})^n - \bar{x}^n - \bar{y}^n = \sum_C |C|C \tag{23}$$

in $Tr\langle x,y\rangle$, where the sum runs over the cyclic equivalences classes C of words of weight n different from x^n and y^n.

Now consider the image $\overline{f_y y}$ of $f_y y$ in $Tr\langle x,y\rangle$. The coefficient of the cyclic equivalence class C in $\overline{f_y y}$ is exactly given by $\sum_{w\in C}(f_y y|w)$, i.e. we have

$$\overline{f_y y} = \sum_C \Big(\sum_{w\in C}(f_y y|w)\Big)C \tag{24}$$

in $Tr\langle x,y\rangle$. We can apply Theorem 4 to compute this coefficient, paying attention to the fact that the sum in (20) is over the n cyclic permutations of w, even if some of them are repeated. If $|C| = n$, then the n cyclic permutations form one copy of the class C, but if $|C| < n$, as for instance the class $\{xyxy, yxyx\}$ where $|C| = 2$ and $n = 4$, then the complete list of n cyclic permutations forms $n/|C|$ copies of C. Fix an element $w^1 \in C$ ending in y, and let d be the number of y's in w^1. Write $w^1 = x^{a_1-1}y\cdots x^{a_d-1}y$ and $w^j = x^{a_j-1}y\cdots x^{a_d-1}yx^{a_1-1}y\cdots x^{a_{j-1}-1}y$ for $2 \le j \le d$. Then by Theorem 4 applied to $g \in \mathfrak{ds}$, we obtain

$$\begin{aligned}\frac{n}{|C|}\sum_{w\in C}(f_y y|w) = \sum_{j=1}^{d}(f_y y|w^j) &= (-1)^d\sum_{j=1}^{d}(g_y y|w^j)\\ &= -(g_y y|x^{n-1}y) = (f_y y|x^{n-1}y).\end{aligned}$$

From this we obtain

$$\sum_{w\in C}(f_y y|w) = \frac{|C|}{n}(f|x^{n-1}y). \tag{25}$$

Putting this back into (24) yields

$$\overline{f_y y} = \frac{1}{n}(f|x^{n-1}y)\sum_C |C|C = \frac{1}{n}(f|x^{n-1}y)\big(\bar{x}^n + \bar{y}^n - (\bar{x}+\bar{y})^n\big)$$

by (23), proving the theorem. Q.E.D.

The generalization of Ihara's statement for even n is contained in the following proposition.

Proposition 1. *Let n be even, and let $f \in \mathfrak{ds}_n$. Then $(f|y_n) = 0$, so in particular, $\overline{f_y y} = 0$ in $Tr\langle x, y\rangle$.*

Proof. The statement that $(f|y_n) = 0$ when n is even has been proved in various forms in various places (for example [11]). The proof we give here comes from the unpublished thesis [3], and seems worth reproducing here for its combinatorial interest. We set $y_i = x^{i-1}y$ and compute in terms of x, y. An easy argument by induction on n shows that for every $n \geq 1$, every monomial of weight n and every Lie element $f \in \mathrm{Lie}_n[x, y]$, we have $(f|w) = (-1)^{n-1}(f|\overleftarrow{w})$, where $\overleftarrow{w}$ is the word w written backwards. In particular, for even n, we have $(f|yx^{n-2}y) = 0$. The stuffle relation associated to the sequences (1) and $(n-1)$ is given by $(f|yx^{n-2}y) + (f|x^{n-2}y^2) + (f|x^{n-1}y) = 0$, so if n is even, we see that

$$(f|x^{n-2}y^2) = -(f|x^{n-1}y). \tag{26}$$

Let us write $[x^i y]$ for the depth 1 Lie element $[x, [x, \cdots, [x, y]\cdots]] = \mathrm{ad}(x)^i(y)$. The element $[x^{n-1}y]$ forms a basis for the 1-dimensional space $\mathrm{Lie}^1_n[x, y]$ of depth 1 elements of $\mathrm{Lie}_n[x, y]$, and the elements $[[x^{n-j-2}y], [x^j y]]$, $0 \leq j \leq [\frac{n-1}{2}]-1$, form a basis (known as the Lyndon–Lie basis) for the $\left[\frac{n-1}{2}\right]$-dimensional space $\mathrm{Lie}^2_n[x, y]$ of Lie elements of weight n and depth 2. Thus we can write

$$f = A[x^{n-1}y] + \sum_{j=0}^{\frac{n-4}{2}} a_j[[x^{n-j-2}y], [x^j y]] + \cdots \tag{27}$$

where $A = (f|x^{n-1}y)$. We can expand the Lie brackets $[[x^{n-j-2}y], [x^j y]]$ explicitly as polynomials using the binomial identity

$$[x^j y] = \sum_{i=0}^{j}\binom{j}{i}(-1)^{i-j}x^i y x^{j-i},$$

and we find that the coefficient of the word $x^{n-i-2}yx^iy$ in the Lie bracket $[[x^{n-j-2}y],[x^jy]]$ is equal to

$$(-1)^{i-j}\left[\binom{n-j-2}{n-i-2}-\binom{j}{n-i-2}\right], \tag{28}$$

where binomial coefficients are considered to be zero whenever the top entry is zero and the bottom entry non-zero or when the bottom entry is negative or greater than the top entry.

From this and (27), we obtain

$$(f|x^{n-i-2}yx^iy)=\sum_{j=0}^{n-2}(-1)^{i-j}a_j\left(\binom{n-j-2}{n-i-2}-\binom{j}{n-i-2}\right), \tag{29}$$

where $a_j=0$ if $j>\frac{n-4}{2}$.

Now we add up the coefficients of all the words of depth 2 in f ending in y, obtaining

$$\begin{aligned}\sum_{i=0}^{n-2}(f|x^{n-i-2}yx^iy)&=\sum_{i=0}^{n-2}\sum_{j=0}^{n-2}(-1)^{i-j}a_j\left(\binom{n-j-2}{n-i-2}-\binom{j}{n-i-2}\right)\\&=\sum_{j=0}^{n-2}a_j\sum_{i=0}^{n-2}(-1)^{i-j}\left(\binom{n-j-2}{n-i-2}-\binom{j}{n-i-2}\right)\\&=-a_0.\end{aligned} \tag{30}$$

Indeed, the term in the sum in (30) for $j=0$ is given by

$$\begin{aligned}&a_0\sum_{i=0}^{n-2}(-1)^i\left(\binom{n-2}{n-i-2}-\binom{0}{n-i-2}\right)\\&\quad=a_0\sum_{i=0}^{n-3}(-1)^i\left(\binom{n-2}{n-i-2}\right)+a_0\left(\binom{n-2}{0}-\binom{0}{0}\right)\\&\quad=-a_0,\end{aligned}$$

whereas each term in (30) for $j>0$ is zero, since if $j>0$, then

$$\begin{aligned}
&\sum_{i=0}^{n-2}(-1)^{i-j}\left(\binom{n-j-2}{n-i-2}-\binom{j}{n-i-2}\right)\\
&=\sum_{i=0}^{n-2}(-1)^{i-j}\binom{n-j-2}{n-i-2}-\sum_{i=0}^{n-2}(-1)^{i-j'}\binom{n-j'-2}{n-i-2}\\
&\qquad\qquad\text{with } j'=n-j-2\\
&=\sum_{i=j}^{n-2}(-1)^{i-j}\binom{n-j-2}{n-i-2}-\sum_{i=j'}^{n-2}(-1)^{i-j'}\binom{n-j'-2}{n-i-2}\\
&=\sum_{m=0}^{n-j-2}(-1)^{m}\binom{n-j-2}{n-j-2-m}-\sum_{m'=0}^{n-j'-2}(-1)^{m'}\binom{n-j'-2}{n-2-j'-m'}\\
&=\sum_{m=0}^{n-j-2}(-1)^{m}\binom{n-j-2}{m}-\sum_{m'=0}^{n-j'-2}(-1)^{m'}\binom{n-j'-2}{m'}=0,
\end{aligned}$$

where the second equality comes from removing the indices in the sums which give terms equal to zero, the third equality is obtained by reindexing the first sum over $m=i-j$ and the second one over $m'=i-j'$, and the final one is zero since in fact each of the two sums separately is already zero.

This proves (30), which allows us to easily finish the proof. Indeed, the $n/2$ stuffle relations in depth 2 are given by

$$(f|x^{n-i-2}yx^iy)+(f|x^iyx^{n-i-2}y)=-A$$

for $0\le i\le(n-4)/2$ (note that when $i=(n-2)/2$ the relation is $2(f|x^{(n-2)/2}yx^{(n-2)/2}y)=-A$), and taking their sum thus yields

$$\sum_{i=0}^{n-2}(f|x^{n-i-2}yx^iy)=-\frac{(n-2)}{2}A-\frac{A}{2}=-\frac{n-1}{2}A. \tag{31}$$

Comparing this with (30), we see that $a_0=\frac{n-1}{2}A$. But (29) shows that $(f|x^{n-2}y^2)=a_0$, so since $A=(f|x^{n-1}y)$, we finally obtain

$$(f|x^{n-2}y^2)=\frac{n-1}{2}(f|x^{n-1}y). \tag{32}$$

Comparing this with (26), since $n\neq-1$, shows that $(f|x^{n-1}y)=(f|y_n)=0$. Q.E.D.

§3. Second application: Furusho's theorem

In this section we use Theorem 2 to give a very simple proof of the Lie version of an important theorem recently proved by H. Furusho [6] in the more general pro-unipotent setting. Of course the Lie statement is implied by Furusho's proof, but Theorem 2 provides a significant simplication of the proof in the Lie case which seems worth explaining here.

Theorem 7. (Furusho) *Let* $f(x,y) \in \mathfrak{grt}$. *Then* $f(x,-y) \in \mathfrak{ds}$.

3.1. Basic setup of the proof.

Furusho's article [6] gives the complete geometric framework for his proof, whose essential idea is to adapt the known stuffle-like relations for double polylogarithms to imply the desired stuffle relations on an element of $\mathrm{Lie}\, P_5$ satisfying the pentagon relation. We do not explain this geometric background here. The purpose of the present exposition is to show that Theorem 2 yields a simplification of the proof of Furusho's theorem in the Lie situation of Theorem 3.1, with respect to the proof that he gives in the pro-unipotent situation in [6]. Therefore our exposition is as minimal as possible, and self-contained with the exception of the main background theorem (theorem 3.2 below) following from Chen's theory of iterated integrals.

Let $\mathrm{Lie}\, P_5$ be the pure sphere 5-strand braid Lie algebra whose definition was recalled at the beginning of §2. Recall that it can be generated by five of the elements x_{ij}. Following Furusho, we fix here the choice of x_{12}, x_{23}, x_{34}, x_{45} and x_{24} as generators. Let $\mathrm{Lie}\, P_5^\vee$ be the dual Lie coalgebra, and write $\omega_{12}, \omega_{23}, \omega_{34}, \omega_{45}, \omega_{24}$ for the duals of the corresponding x_{ij}.

The dual V_5 of the enveloping algebra $U\mathrm{Lie}\, P_5$ is isomorphic to a subspace of the freely generated polynomial ring $\Omega = \mathbb{Q}[\omega_{ij}]$. A word in the ω_{ij} in this ring is written using the bar-notation $[\omega_{i_1j_1}|\cdots|\omega_{i_rj_r}]$ and called a bar-word. Linear combinations of bar-words are called bar-symbols. Multiplication in the ring Ω is commutative, given by the shuffle operation on words. For example,

$$[\omega|\omega'] \cdot [\omega''] = [\omega|\omega'|\omega''] + [\omega|\omega''|\omega'] + [\omega''|\omega|\omega'].$$

The grading on $U\mathrm{Lie}\, P_5$ given by letting all x_{ij} be of weight 1 translates to a grading on V_5 given by the lengths of the words in ω_{ij}. The ω_{ij} can be identified with differential 1-forms on the moduli space

$$M_{0,5} \simeq \left(\mathbb{P}^1\mathbb{C} - \{0,1,\infty\}\right)^2 - \{XY = 1\}$$

as follows:

$$\omega_{12} = \frac{dX}{X}, \quad \omega_{23} = \frac{dX}{X-1}, \quad \omega_{34} = \frac{dY}{Y-1},$$

$$\omega_{45} = \frac{dY}{Y}, \quad \omega_{24} = \frac{XdY + YdX}{XY-1}. \tag{33}$$

Let Ω_r denote the subvector space of Ω consisting of polynomials in the ω_{ij} of homogeneous degree r. Then the homogeneous subspace $(V_5)_r \subset \Omega_r$ is characterized by the following property: (P) Let $\sigma = \big((i_1,j_1),\ldots,(i_r,j_r)\big)$ denote the r-tuples of pairs $(i_m,j_m) \in \{(1,2),(2,3),(3,4),(4,5),(2,4)\}$, and let $W = \sum_\sigma a_\sigma[\omega_{i_1,j_1}|\cdots|\omega_{i_r,j_r}] \in \Omega_r$. Then W lies in V_5 if and only if the $(r-1)$ sums

$$\sum_\sigma a_\sigma[\omega_{i_1,j_1}|\cdots|\omega_{i_k,j_k} \wedge \omega_{i_{k+1},j_{k+1}}|\cdots|\omega_{i_r,j_r}]$$

are equal to zero as elements of $(V_5)_1^{\otimes k-1} \otimes H^2_{DR}(M_{0,5}) \otimes (V_5)_1^{\otimes r-k-1}$ for $1 \le k \le r-1$, where the ω_{ij} are wedged as the differential 1-forms in (33). For example, $\omega_{12}\wedge\omega_{23} = \omega_{34}\wedge\omega_{45} = 0$ and $\omega_{12}\wedge\omega_{45} = -\omega_{45}\wedge\omega_{24}$. More specifically, (P) can be understood by separately considering triples

$$T = \Big(k, \big((a_1,b_1),\ldots,(a_{k-1},b_{k-1})\big), \big((a_{k+2},b_{k+2}),\ldots,(a_r,b_r)\big)\Big),$$

where $k \in \{1,\ldots,r-1\}$ and the pairs (a_m,b_m) all lie in $\{(1,2),(2,3),(3,4),(4,5),(2,4)\}$. For every such triple T, let S_T denote the set of r-tuples of pairs $\big((i_1,j_1),\ldots,(i_r,j_r)\big)$ such that $(i_m,j_m) = (a_m,b_m)$ for $1 \le m \le k-1$ and $k+2 \le m \le r$. The condition (P) for W to lie in V_5 is then that for each triple T,

$$\sum_{\sigma \in S_T} a_\sigma \omega_{i_k,j_k} \wedge \omega_{i_{k+1},j_{k+1}} = 0.$$

The Lie coalgebra $\mathrm{Lie}\, P_5^\vee$ is isomorphic to the quotient of V_5 modulo (shuffle) products. In other words, every shuffle sum of bar-words in $\mathrm{Lie}\, P_5^\vee$ is equal to zero. In [6], Furusho introduced particular elements in V_5, called

$$l^X_{\mathbf{a}}, l^Y_{\mathbf{a}}, l^{XY}_{\mathbf{a}}, l^{X,Y}_{\mathbf{a},\mathbf{b}}, l^{Y,X}_{\mathbf{a},\mathbf{b}},$$

where X and Y are free commutative variables, and $\mathbf{a} = (a_1,\ldots,a_r)$ and $\mathbf{b} = (b_1,\ldots,b_s)$ are tuples of strictly positive integers. We will give a direct recursive definition of these elements here. In order for our notation to correspond more closely to Furusho's, we need the following change of notation with respect to the two previous sections.

Change of notation. Let $\mathbf{a} = (a_1, \dots, a_r)$ be a sequence of strictly positive integers. We now define $w_{\mathbf{a}}$ to be the word in non-commutative variables x and y given by

$$w_{\mathbf{a}} = x^{a_r-1}y \cdots x^{a_2-1}yx^{a_1-1}y.$$

With this notation, we have $w_{\mathbf{a}}w_{\mathbf{b}} = w_{\mathbf{ba}}$. Let us now define Furusho's symbols, using this new notation.

Definition 3.1. *For any element* $\varphi \in \mathrm{Lie}[x,y]$, *we write* $(\varphi|w)$ *for the coefficient of the word* w *in the polynomial* φ. *Define the element* $l_{\mathbf{a}} \in \mathrm{Lie}[x,y]^\vee$ *by*

$$l_{\mathbf{a}}(\varphi) = (-1)^r(\varphi|w_{\mathbf{a}}). \tag{34}$$

• The element $l_{\mathbf{a}}^X \in \mathrm{Lie}\,P_5^\vee$ is the bar-word defined by replacing every x in the word $w_{\mathbf{a}} = x^{a_r-1}y \cdots x^{a_1-1}y$ by ω_{12} and every y by ω_{23}. • The element $l_{\mathbf{a}}^Y \in \mathrm{Lie}\,P_5^\vee$ is the bar-symbol defined by replacing every x in the word $w_{\mathbf{a}}$ by ω_{45} and every y by ω_{34}. • The element $l_{\mathbf{a}}^{XY} \in \mathrm{Lie}\,P_5^\vee$ is the bar-word defined by replacing every x in the word $w_{\mathbf{a}}$ by $\omega_{12}+\omega_{45}$ and every y by ω_{24}. • The element $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ is defined recursively according to the form of the tuples $\mathbf{a}$ and $\mathbf{b}$. Let $\mathbf{a} = (a_1, \dots, a_r)$, $\mathbf{b} = (b_1, \dots, b_s)$. If $a_r > 1$, set $\mathbf{a}' = (a_1, \dots, a_r - 1)$; if $a_r = 1$ but $r > 1$, set $\mathbf{a}' = (a_1, \dots, a_{r-1})$ (with this notation, $xw_{\mathbf{a}'} = w_{\mathbf{a}}$). Use the same notation for $\mathbf{b}$. If $r > 1$, $a_r = 1$ and $s > 1$, set $\mathbf{a}'' = (a_1, \dots, a_{r-1}, b_1)$ and $\mathbf{b}'' = (b_2, \dots, b_s)$. The element $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ is defined by:

(35)

$$\begin{cases}
[\omega_{12}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] + [\omega_{45}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r > 1, b_s > 1\\
[\omega_{12}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] + [\omega_{34}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r > 1, b_s = 1, s > 1\\
[\omega_{12}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] + [\omega_{34}|l_{\mathbf{a}}^{XY}] & \text{if } a_r > 1, b_s = s = 1\\
[\omega_{23}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{a}'',\mathbf{b}''}^{X,Y}] + [\omega_{45}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r > 1, b_s > 1, s > 1\\
[\omega_{23}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{a}''}^{XY}] + [\omega_{45}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r > 1, b_s > 1, s = 1\\
[\omega_{23}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{a}'',\mathbf{b}''}^{X,Y}] + [\omega_{34}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r > 1, b_s = 1, s > 1\\
[\omega_{23}|l_{\mathbf{a}',\mathbf{b}}^{X,Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{a}''}^{XY}] + [\omega_{34}|l_{\mathbf{a}}^{XY}] & \text{if } a_r = 1, r > 1, b_s = 1, s = 1\\
[\omega_{23}|l_{\mathbf{b}}^{Y}] - [\omega_{12}+\omega_{23}|l_{(b_1),\mathbf{b}''}^{X,Y}] + [\omega_{45}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r = 1, b_s > 1, s > 1\\
[\omega_{23}|l_{\mathbf{b}}^{Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{b}}^{XY}] + [\omega_{45}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r = 1, b_s > 1, s = 1\\
[\omega_{23}|l_{\mathbf{b}}^{Y}] - [\omega_{12}+\omega_{23}|l_{(b_1),\mathbf{b}''}^{X,Y}] + [\omega_{34}|l_{\mathbf{a},\mathbf{b}'}^{X,Y}] & \text{if } a_r = 1, r = 1, b_s = 1, s > 1\\
[\omega_{23}|l_{\mathbf{b}}^{Y}] - [\omega_{12}+\omega_{23}|l_{\mathbf{b}}^{XY}] + [\omega_{34}|l_{\mathbf{a}}^{XY}] & \text{if } a_r = 1, r = 1, b_s = 1, s = 1.
\end{cases}$$

• Finally, the element $l_{\mathbf{a},\mathbf{b}}^{Y,X}$ is defined by computing $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ and then applying the order 2 automorphism ρ which exchanges the pair ω_{45} and ω_{12}, and the pair ω_{23} and ω_{34}, while fixing ω_{24}.

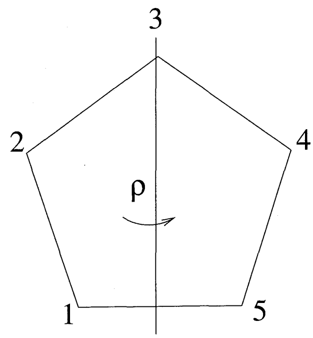

Fig. 1

Examples. Let $\mathbf{a} = (2,1)$. Then $w_{\mathbf{a}} = yxy$ and

$$l^{XY}_{(2,1)} = [\omega_{24}|\omega_{12} + \omega_{45}|\omega_{24}] = [\omega_{24}|\omega_{12}|\omega_{24}] + [\omega_{24}|\omega_{45}|\omega_{24}].$$

This polynomial lies in V_5 since it satisfies the property (P): indeed, as we saw above, $\frac{dX}{X} \wedge \frac{YdX+XdY}{XY-1} + \frac{dY}{Y} \wedge \frac{XdY+YdX}{XY-1} = 0$, which ensures that both the first sum $\omega_{24} \wedge \omega_{12} + \omega_{24} \wedge \omega_{45}$ and the second sum $\omega_{12} \wedge \omega_{24} + \omega_{45} \wedge \omega_{24}$ are zero.

Now let $\mathbf{a} = (1)$, $\mathbf{b} = (1)$. We have

$$\begin{aligned} l^{X,Y}_{(1),(1)} &= [\omega_{23}|l^{Y}_{(1)}] - [\omega_{12} + \omega_{23}|l^{XY}_{(1)}] + [\omega_{34}|l^{XY}_{(1)}] \\ &= [\omega_{23}|\omega_{34}] - [\omega_{12} + \omega_{23}|\omega_{24}] + [\omega_{34}|\omega_{24}]. \end{aligned}$$

If $\mathbf{a} = (2)$ and $\mathbf{b} = (1)$, we have

$$\begin{aligned} l^{X,Y}_{(2),(1)} &= [\omega_{12}|l^{X,Y}_{(1),(1)}] + [\omega_{34}|l^{XY}_{(2)}] \\ &= [\omega_{12}|\omega_{23}|\omega_{34}] - [\omega_{12}|\omega_{12} + \omega_{23}|\omega_{24}] + [\omega_{12}|\omega_{34}|\omega_{24}] \\ &\qquad + [\omega_{34}|\omega_{12} + \omega_{45}|\omega_{24}]. \end{aligned}$$

We now introduce the fundamental "stuffle-type" relations satisfied by these elements of Lie $P_5^{\vee}$.

Definition 3.2. *Furusho gives a generalization of* $st(\mathbf{a},\mathbf{b})$ *to the* Lie P_5 *situation as follows. Recall from §1 the definition of the set of maps* $\mathrm{Sh}^{\leq}(r,s)$ *and the stuffle set* $st(\mathbf{a},\mathbf{b})$*. Let* $ST(\mathbf{a},\mathbf{b})$ *be the set of pairs of sequences* $\sigma(\mathbf{a},\mathbf{b}) = \big((c_1,\ldots,c_j),(c_{j+1},\ldots,c_N)\big)$*, where*

$$(c_1,\ldots,c_N) = c^{\sigma}(\mathbf{a},\mathbf{b}) \in st(\mathbf{a},\mathbf{b})$$

and $j = \min\big(\sigma(r), \sigma(r+s)\big)$*. Also, for each* $\sigma \in \mathrm{Sh}^{\leq}(r,s)$*, set* $\sigma(X,Y) = XY$ *if* $\sigma^{-1}(N) = \{r, r+s\}$*,* $\sigma(X,Y) = (X,Y)$ *if* $\sigma^{-1}(N) = r+s$ *and* $\sigma(X,Y) = (Y,X)$ *if* $\sigma^{-1}(N) = r$*.*

Furusho bases the proof of his theorem on the following fundamental set of stuffle identities.

Theorem 8. *For all tuples of strictly positive integers* $\mathbf{a} = (a_1, \ldots, a_r)$ *and* $\mathbf{b} = (b_1, \ldots, b_s)$*, the elements* $l_{\mathbf{a},\mathbf{b}}^{X,Y}$, $l_{\mathbf{a},\mathbf{b}}^{Y,X}$, $l_{\mathbf{a}}^{XY} \in \mathrm{Lie}\, P_5^\vee$ *defined above satisfy the relation*

$$\sum_{\sigma(\mathbf{a},\mathbf{b}) \in \mathrm{ST}(\mathbf{a},\mathbf{b})} l_{\sigma(\mathbf{a},\mathbf{b})}^{\sigma(X,Y)} = 0. \tag{36}$$

Sketch of proof. The proof of this result follows from Chen's theory of iterated integrals. This is explained in more detail in [6] (see also [2] for some of the proofs), so we only sketch the situation here. In this theory, the dual elements ω_{ij} of the x_{ij} are identified with the 1-forms on the moduli space $M_{0,5} \simeq (\mathbb{P}^1 - \{0, 1, \infty\})^2 - \{XY = 1\}$ to the x_{ij}, given by the expressions in (33), and (linear combinations of) bar-words in the ω_{ij} are identified with iterated integrals of the entries along a path on $M_{0,5}$ from $(0, 0)$ to (X, Y). The condition above defining an element of V_5 is precisely the "integrability" condition of a linear combination of bar-words, ensuring that the value of the integral depends only on the homotopy class of the chosen path, and Chen's theory (see also [2]) shows that the map from V_5 to iterated integrals is injective. An easy computation shows that the iterated integrals associated to the elements $l_{\mathbf{a}}^{XY}$, $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ and $l_{\mathbf{a},\mathbf{b}}^{Y,X}$ are single and double polylogarithm functions $Li_{\mathbf{a}}(XY)$, $Li_{\mathbf{a},\mathbf{b}}(X, Y)$ and $Li_{\mathbf{a},\mathbf{b}}(Y, X)$ (see [6] for their explicit expressions), and these are classically known to satisfy the equalities

$$\sum_{\sigma(\mathbf{a},\mathbf{b}) \in \mathrm{ST}(\mathbf{a},\mathbf{b})} Li_{\sigma(\mathbf{a},\mathbf{b})}(\sigma(X, Y)) = Li_{\mathbf{a}}(X) Li_{\mathbf{a}}(Y).$$

Thus by injectivity of the iterated integral map from V_5 to functions of X and Y, we see that

$$\sum_{\sigma \in \mathrm{ST}(\mathbf{a},\mathbf{b})} l_{\sigma(\mathbf{a},\mathbf{b})}^{\sigma(X,Y)} = l_{\mathbf{a}}^X l_{\mathbf{b}}^Y$$

in V_5. Thus when these elements are considered in the quotient $\mathrm{Lie}\, P_5^\vee$, we recover (36). Q.E.D.

3.2. Furusho's lemma

In §4 of [6], Furusho states and proves two lemmas in the pro-unipotent situation, making use of a regularization defined in the body of the paper. The statement of Lemma 3 summarizes the essence of the the Lie part of Furusho's statements, but the lemma is slightly stronger than the one in [6], in that the hypothesis on $\mathbf{a}, \mathbf{b}$ in the sixth statement is weaker than the one there. This is one of the the main points of simplication of the Lie proof.

Apart from the sixth statement and its proof, the rest of the statements and proofs are exact Lie analogs of those given in [6]. However, the terminology is different and the proofs there are partly left to the reader, so in the interest of completeness, we give the full proof in detail here.

Lemma 3. *Let $p_3 : B_5 \to F_2$ be the map defined by $p_3 : x_{12} \mapsto x$, $x_{24} \mapsto -y$, $x_{i3} \mapsto 0$. and let the maps $i_{jkl} : F_2 \to B_5$ be defined by $i_{jkl}(x) = x_{jk}$, $i_{jkl}(y) = -x_{kl}$. We have the following six identities:*

$$\begin{cases} l_{\mathbf{a}}^{XY} = l_{\mathbf{a}} \circ p_3 & \textit{for all } (\mathbf{a}, \mathbf{b}) \\ l_{\mathbf{a},\mathbf{b}}^{X,Y} \circ i_{123} = 0 & \textit{for all } (\mathbf{a}, \mathbf{b}) \\ l_{\mathbf{a},\mathbf{b}}^{Y,X} \circ i_{543} = 0 & \textit{for all } (\mathbf{a}, \mathbf{b}) \\ l_{\mathbf{a},\mathbf{b}}^{X,Y} \circ i_{451} = l_{\mathbf{ab}} & \textit{for all } (\mathbf{a}, \mathbf{b}) \\ l_{\mathbf{a},\mathbf{b}}^{Y,X} \circ i_{215} = l_{\mathbf{ab}} & \textit{for all } (\mathbf{a}, \mathbf{b}) \\ l_{\mathbf{a},\mathbf{b}}^{Y,X} \circ i_{432} = 0 & \textit{for all } (\mathbf{a}, \mathbf{b}) \neq \big((1, \dots, 1), (1, \dots, 1)\big). \end{cases}$$

Proof. Recall that if $\mathbf{a} = (a_1, \dots, a_r)$, we set $w_{\mathbf{a}} = x^{a_r - 1}y \cdots x^{a_1 - 1}y$; Then $l_{\mathbf{a}}(w_{\mathbf{a}}) = (-1)^r$ and $l_{\mathbf{a}}(v) = 0$ for all words $v \neq w_{\mathbf{a}}$. By definition, $l_{\mathbf{a}}^{XY}$ is the bar-word obtained from $w_{\mathbf{a}}$ by replacing x by $\omega_{12} + \omega_{45}$ and y by ω_{24}. Expand this word out as a polynomial as follows. Let $n = a_1 + \cdots + a_r$ and let $I \subset \{1, \dots, n\}$ be given by $I = \{a_1, a_1 + a_2, \dots, a_1 + \cdots + a_{r-1}, a_1 + \cdots + a_r\}$. Let $\mathcal{E}$ denote the set of distinct tuples $(\epsilon_1, \dots, \epsilon_n)$ such that

$$\begin{cases} \epsilon_i = (24) & i \in I \\ \epsilon_j \in \{(12), (45)\} & j \in \{1, \dots, n\} \setminus I. \end{cases}$$

Then the expansion is given by

$$l_{\mathbf{a}}^{XY} = \sum_{(\epsilon_1, \dots, \epsilon_n) \in \mathcal{E}} [\omega_{\epsilon_1} | \cdots | \omega_{\epsilon_n}].$$

Thus, $l_{\mathbf{a}}^{XY}$ takes the value zero on all words of length n in the generators $x_{12}, x_{23}, x_{34}, x_{45}, x_{24}$ of $\mathrm{Lie}\, P_5$ except for the ones of the form $W = x_{\epsilon_1} \cdots x_{\epsilon_n}$ for $(\epsilon_1, \dots, \epsilon_n) \in \mathcal{E}$.

Now, $x_{45} = x_{12} + x_{13} + x_{23}$ in $\mathrm{Lie}\, P_5$, so $p_3(x_{45}) = p_3(x_{12}) = x$, and $p_3(x_{24}) = -y$, so $p_3(W) = (-1)^r w_{\mathbf{a}}$. Thus for every word $W = x_{\epsilon_1} \cdots x_{\epsilon_n}$, we have $l_{\mathbf{a}}^{XY}(W) = 1 = l_{\mathbf{a}}((-1)^r w_{\mathbf{a}}) = l_{\mathbf{a}}(p_3(W))$. The words W are exactly the complete set of words such that $p_3(W) = (-1)^r w_{\mathbf{a}}$, so if V is not of the form W, we have $l_{\mathbf{a}}^{XY}(V) = 0 = l_{\mathbf{a}}(p_3(V))$. This proves the first statement.

The other statements are proved using induction on the length $r+s$ of the sequences $\mathbf{a}$ and $\mathbf{b}$ together, where $r, s \geq 1$. Let $\mathbf{b} = (b_1, \ldots, b_s)$, and consider the second statement. It comes down to saying that for any pair $(\mathbf{a}, \mathbf{b})$, the symbol $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ cannot contain any bar-words in the two variables ω_{12} and ω_{23} only. The base case, $r+s=2$, was computed in the examples in §3.1:

$$(37) \qquad l_{(1),(1)}^{X,Y} = [\omega_{23}|\omega_{34}] - [\omega_{12}|\omega_{24}] - [\omega_{23}|\omega_{24}] + [\omega_{34}|\omega_{24}].$$

Now assume $r+s > 2$, make the induction hypothesis that $l_{\mathbf{c},\mathbf{d}}^{X,Y}$ contains no such bar-word when the sum of the lengths of $\mathbf{c}$ and $\mathbf{d}$ is less than $r+s$, and fix a pair $(\mathbf{a}, \mathbf{b})$ of lengths r and s. Consider the definition of $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ in (35). A word in ω_{12} and ω_{23} only would have to come from the term(s) in each line of (35) of the form $[\omega_{12}|\cdots]$ or $[\omega_{23}|\cdots]$. But for each such term, the right-hand part of the term is either one of $l_{\mathbf{a}',\mathbf{b}}^{X,Y}$, $l_{\mathbf{a}'',\mathbf{b}''}^{X,Y}$, $l_{(b_1),\mathbf{b}''}^{X,Y}$, none of which contain a word in only ω_{12}, ω_{23} by the induction hypothesis, or $l_{\mathbf{b}}^{XY}$, $l_{\mathbf{a}''}^{XY}$ or $l_{\mathbf{b}}^{Y}$, which do not contain any such word by definition. This proves the second statement.

The third statement is equivalent to the second, given the definition of $l_{\mathbf{a},\mathbf{b}}^{Y,X}$, which is obtained from $l_{\mathbf{a},\mathbf{b}}^{X,Y}$ by applying the automorphism ρ of the bar-construction defined in figure 1, since $i_{123} = \rho \circ i_{543}$.

The fifth statement follows similarly from the fourth by applying ρ, so let us prove the fourth statement. We first note that

$$(38) \qquad l_{\mathbf{a}}^{Y}(i_{451}(w)) = l_{\mathbf{a}}^{XY}(i_{451}(w)) = l_{\mathbf{a}}(w)$$

for all sequences $\mathbf{b}$. Indeed, $l_{\mathbf{a}}^{Y}$ is a bar-word in ω_{34} and ω_{45} only, so in computing the left hand term one can ignore all the terms containing x_{23} or x_{24} that appear in $i_{451}(w)$. Similarly, $l_{\mathbf{a}}^{XY}$ is a bar-symbol in ω_{12}, ω_{45} and ω_{24}, so in computing the middle member of (38), one can ignore the ω_{12} that appear there, and all the terms in i_{451} that contain x_{23} or x_{34}.

We first take care of the base case $\mathbf{a} = (1), \mathbf{b} = (1)$, and show that $l_{(1),(1)}^{X,Y}(i_{451}(w)) = l_{(1,1)}(w)$ for any word w of length 2 in x and y. We have $l_{(1,1)}(x^2) = l_{(1,1)}(xy) = l_{(1,1)}(yx) = 0$, $l_{(1,1)}(y^2) = 1$. By observing (37) and the equality $\omega_{51} = \omega_{23} + \omega_{24} + \omega_{34}$ in $\mathrm{Lie}\, P_5^{\vee}$, we see that $l_{\mathbf{a},\mathbf{b}}^{X,Y} = l_{(1),(1)}^{X,Y}$ has value zero on the bar-words $i_{451}(x^2) = [\omega_{45}|\omega_{45}]$, $i_{451}(xy) = -[\omega_{45}|\omega_{51}]$, and $i_{451}(yx) = -[\omega_{51}|\omega_{45}]$, since they all contain an x_{45} which doesn't appear in (37). But in $i_{451}(y^2)=[\omega_{51}|\omega_{51}]$, there appear three words on which $l^{X,Y}_{(1),(1)}$ has non-zero value, namely $[\omega_{23}|\omega_{34}]$, $[\omega_{23}|\omega_{24}]$ and $[\omega_{34}|\omega_{24}]$. The values are 1, -1 and 1 respectively, so $l^{X,Y}_{(1),(1)}([\omega_{51}|\omega_{51}])=1=l_{(1,1)}(y^2)$. This settles the base case.

Make the induction hypothesis that for any pair $\mathbf{c},\mathbf{d}$ of length $r+s-1$ and any word v of length $r+s-1$, we have $l^{X,Y}_{\mathbf{c},\mathbf{d}}(i_{451}(v))=l_{\mathbf{cd}}(v)$. Note that since $i_{451}(x)=x_{45}$ and $i_{451}(y)=-x_{51}=-x_{23}-x_{24}-x_{34}$, for any one of Furusho's symbols L, we have

$$(39)\qquad \begin{cases} [\omega_{45}|L](i_{451}(w))=\begin{cases} L(i_{451}(v)) & \text{if } w=xv\\ 0 & \text{otherwise}\end{cases}\\ [\omega_{23}|L](i_{451}(w))=\begin{cases} -L(i_{451}(v)) & \text{if } w=yv\\ 0 & \text{otherwise}\end{cases}\\ [\omega_{24}|L](i_{451}(w))=\begin{cases} -L(i_{451}(v)) & \text{if } w=yv\\ 0 & \text{otherwise}\end{cases}\\ [\omega_{34}|L](i_{451}(w))=\begin{cases} -L(i_{451}(v)) & \text{if } w=yv\\ 0 & \text{otherwise.}\end{cases}\end{cases}$$

Fix $\mathbf{a},\mathbf{b}$ of lengths r, s, and let us consider the eleven cases of (35); let $\Omega=\{\omega_{45},\omega_{23},\omega_{24},\omega_{34}\}$. Only the terms of (35) starting with an $\omega\in\Omega$ can have a non-zero value on $i_{451}(w)$, since $i_{451}(w)$ is a polynomial in the variables of Ω only. Let us now prove the desired identity $l^{X,Y}_{\mathbf{a},\mathbf{b}}\circ i_{451}=l_{\mathbf{ab}}$ on a case by case basis according to the nature of the sequences $\mathbf{a},\mathbf{b}$ as in (35).

Case 1. The element ω_{12} is not in Ω, so we only need to consider the term $[\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}]$. By the first entry of (39), applying this term to $i_{451}(w)$ yields 0 if $w=yv$; but also $l_{\mathbf{ab}}(yv)=0$ since $b_s>1$, settling the desired equality for all words starting in y. If on the other hand $w=xv$, then we have

$$\begin{aligned} l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) &= [\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}](i_{451}(w))\\ &= l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) \quad \text{by (39)}\\ &= l_{\mathbf{ab}'}(v) \quad \text{by the induction hypothesis}\\ &= l_{\mathbf{ab}}(w).\end{aligned}$$

Case 2. The only relevant term is $[\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}]$, so by (39), if $w=xv$, we have $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w))=0=l_{\mathbf{ab}}(w)$ in this case (since $b_s=1$). If $w=yv$, we have

$$\begin{aligned} l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) &= [\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}](i_{451}(w)) \\ &= -l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) \text{ by (39)} \\ &= -l_{\mathbf{a}\mathbf{b}'}(v) \text{ by the induction hypothesis} \\ &= l_{\mathbf{a}\mathbf{b}}(w) \end{aligned}$$

since w contains one more y than v.

Case 3. There are no relevant terms since $[\omega_{34}|l^{XY}_{\mathbf{a}}]$ contains ω_{24}.

Case 4. There are three relevant terms:

$$[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}], \quad -[\omega_{23}|l^{X,Y}_{\mathbf{a}'',\mathbf{b}''}], \quad [\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}].$$

If $w = xv$, the first two have value 0, and by (39) the third has value $l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) = l_{\mathbf{a}\mathbf{b}'}(v) = l_{\mathbf{a}\mathbf{b}}(w)$, by the induction hypothesis and the fact that $b_s > 1$. If $w = yv$, the third has value 0, and by (39) and induction, the sum of the first two has value $-l_{\mathbf{a}'\mathbf{b}}(v) + l_{\mathbf{a}''\mathbf{b}''}(v) = 0$, since the concatenated sequences $\mathbf{a}'\mathbf{b}$ and $\mathbf{a}''\mathbf{b}''$ are both equal to $(a_1, \ldots, a_{r-1}, b_1, \ldots, b_s)$ when $a_r = 1$, so $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = 0$ if $w = yv$. But we also have $l_{\mathbf{a}\mathbf{b}}(w) = 0$ in this case, since $b_s > 1$.

Case 5. Here the relevant terms are

$$[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}], \quad -[\omega_{23}|l^{XY}_{\mathbf{a}''}], \quad [\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}].$$

If $w = xv$, the first two terms take the value zero, so

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(xv)) = l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) = l_{\mathbf{a}\mathbf{b}'}(v) = l_{\mathbf{a}\mathbf{b}}(w)$$

by the induction hypothesis and the fact that $b_s > 1$. If $w = yv$, the third term takes the value zero, so

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(yv)) = -l^{X,Y}_{\mathbf{a}',\mathbf{b}}(i_{451}(v)) + l^{XY}_{\mathbf{a}''}(i_{451}(v)).$$

But $l^{XY}_{\mathbf{a}''}$ is a bar-word in ω_{12}, ω_{45} and ω_{24}, so since $i_{451}(x) = x_{45}$ and $i_{451}(y) = -x_{23} - x_{24} - x_{34}$, we can ignore the ω_{12} in $l_{\mathbf{a}''}$ and the x_{23}, x_{34} in $i_{451}(v)$, to obtain $l^{XY}_{\mathbf{a}''}(i_{451}(v)) = l_{\mathbf{a}''}(v)$. Using induction, we have $l^{X,Y}_{\mathbf{a}',\mathbf{b}}(i_{451}(v)) = l_{\mathbf{a}'\mathbf{b}}(v) = l_{\mathbf{a}''}(v)$, where the last equality follows since the concatenation $\mathbf{a}'\mathbf{b} = \mathbf{a}''$ in this case. Thus $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(v)) = 0$. But this is equal to $l_{\mathbf{a}\mathbf{b}}(w) = l_{\mathbf{a}\mathbf{b}}(yv)$, since $\mathbf{b} = (b_1)$ with $b_1 > 1$, so $l_{\mathbf{a}\mathbf{b}}$ can only take a non-zero value on a word starting with x.

Case 6. Here the relevant terms are

$$[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}],\quad -[\omega_{23}|l^{X,Y}_{\mathbf{a}'',\mathbf{b}''}],\quad [\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}],$$

so if $w = xv$, we have $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = 0$, but also $l_{\mathbf{ab}}(xv) = 0$ since $b_s = 1$, so $w_{\mathbf{ab}}$ begins with a y. If $w = yv$, we have

$$\begin{aligned} l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) &= -l^{X,Y}_{\mathbf{a}',\mathbf{b}}(i_{451}(v)) + l^{X,Y}_{\mathbf{a}'',\mathbf{b}''}(i_{451}(v)) - l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) \\ &= -l_{\mathbf{a}'\mathbf{b}}(v) + l_{\mathbf{a}''\mathbf{b}''}(v) - l_{\mathbf{ab}'}(v) \end{aligned}$$

by the induction hypothesis, but $\mathbf{a}'\mathbf{b} = \mathbf{a}''\mathbf{b}'' = (a_1, \ldots, a_{r-1}, b_1, \ldots, b_s)$ in this case, so this is equal to

$$-l_{\mathbf{ab}'}(v) = l_{\mathbf{ab}}(yv) = l_{\mathbf{ab}}(w).$$

Case 7. Here the relevant terms are

$$[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}],\quad -[\omega_{23}|l^{XY}_{\mathbf{a}''}],\quad [\omega_{34}|l^{XY}_{\mathbf{a}}].$$

If $w = xv$, we have $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = 0$, but also $l_{\mathbf{ab}}(w) = 0$ since $\mathbf{b} = (1)$. If $w = yv$, we have

$$\begin{aligned} l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) &= -l^{X,Y}_{\mathbf{a}',\mathbf{b}}(i_{451}(v)) + l^{XY}_{\mathbf{a}''}(i_{451}(v)) - l^{Y}_{\mathbf{a}}(i_{451}(v)) \\ &= -l_{\mathbf{a}'\mathbf{b}}(v) + l_{\mathbf{a}}(v) - l_{\mathbf{a}}(v) = -l_{\mathbf{a}'\mathbf{b}}(v) \end{aligned}$$

by induction and because $\mathbf{a}'' = \mathbf{a}$ in this case, as $a_r = 1 = b_1$. But $-l_{\mathbf{a}'\mathbf{b}}(v) = l_{\mathbf{ab}}(yv) = l_{\mathbf{ab}}(w)$ since $\mathbf{a} = (a_1, \ldots, a_{r-1}, 1)$ and $\mathbf{b} = (1)$, so $\mathbf{a}'\mathbf{b} = (a_1, \ldots, a_{r-1}, 1)$ and $\mathbf{ab} = (a_1, \ldots, a_{r-1}, 1, 1)$.

Case 8. Here the relevant terms are

$$[\omega_{23}|l^{Y}_{\mathbf{b}}],\quad -[\omega_{23}|l^{X,Y}_{(b_1),\mathbf{b}''}],\quad [\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}],$$

If $w = xv$, the first and second terms take value zero on w, so

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) = l_{\mathbf{ab}'}(v) = l_{\mathbf{ab}}(w)$$

by induction and because $b_s > 1$. If $w = yv$, the third term takes value zero; then $l_{\mathbf{ab}}(w) = 0$ since $b_s > 1$, and also

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = -l^{Y}_{\mathbf{b}}(i_{451}(v)) + l^{X,Y}_{(b_1),\mathbf{b}''}(i_{451}(v)) = -l_{\mathbf{b}}(v) + l_{\mathbf{b}}(v) = 0$$

by (38) for the first term and induction for the second.

Case 9. Here the relevant terms are

$$[\omega_{23}|l^{Y}_{\mathbf{b}}],\quad -[\omega_{23}|l^{XY}_{\mathbf{b}}],\quad [\omega_{45}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}].$$

If $w = xv$, the first and second terms take the value zero, so

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = l_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) = l_{\mathbf{ab}'}(v) = l_{\mathbf{ab}}(w)$$

by induction and because $b_s > 1$. If $w = yv$, the third term takes value zero, so we have

$$l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = -l^{Y}_{\mathbf{b}}(i_{451}(v)) + l^{XY}_{\mathbf{b}}(i_{451}(v)) = 0$$

by (38). But also $l_{\mathbf{ab}}(yv) = 0$ in this case since $\mathbf{b} = (b_1)$ with $b_1 > 1$.

Case 10. Here the relevant terms are

$$[\omega_{23}|l^{Y}_{\mathbf{b}}],\quad -[\omega_{23}|l^{X,Y}_{(b_1),\mathbf{b}''}],\quad [\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}].$$

Thus, if $w = xv$, we have $l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) = 0$, and also $l_{\mathbf{ab}}(w) = 0$ since $b_s = 1$. If $w = yv$, we have

$$\begin{aligned} l^{X,Y}_{\mathbf{a},\mathbf{b}}(i_{451}(w)) &= -l^{Y}_{\mathbf{b}}(i_{451}(v)) + l^{X,Y}_{(b_1),\mathbf{b}''}(i_{451}(v)) - l^{X,Y}_{\mathbf{a},\mathbf{b}'}(i_{451}(v)) \\ &= -l_{\mathbf{b}}(v) + l_{\mathbf{b}}(v) - l_{\mathbf{ab}'}(v) = -l_{\mathbf{ab}'}(v) \\ &= l_{\mathbf{ab}}(w) \end{aligned}$$

by (38) for the left-hand term, induction for the middle term (since (b_1) concatenated with $\mathbf{b}''$ is just $\mathbf{b}$), and induction also for the last term. The final equality works because $b_s = 1$ and w begins with y.

Case 11. This is the case $\mathbf{a} = (1), \mathbf{b} = (1)$ and was already treated as the base case for the induction.

This concludes the proof of the fourth statement, which as noted above immediately implies the fifth by symmetry. To complete the proof of the lemma, it thus remains only to prove the sixth statement. It is enough to prove that if $(\mathbf{a}, \mathbf{b})$ are not all 1's, there is no bar-word in just ω_{23} and ω_{34} appearing in $l^{Y,X}_{\mathbf{a},\mathbf{b}}$. Since the definition of $l^{Y,X}_{\mathbf{a},\mathbf{b}}$ as ρ applied to (35), this is equivalent to proving that $l^{X,Y}_{\mathbf{a},\mathbf{b}} \circ i_{234} = 0$, i.e. that $l^{X,Y}_{\mathbf{a},\mathbf{b}}$ has no bar-words in just ω_{23}, ω_{34}, for pairs $(\mathbf{a}, \mathbf{b})$ not all 1's. This is more convenient as we can stare at (35). The base case for the induction here is given by $l^{X,Y}_{(2),(1)}$ which was computed above and contains no such terms, and $l^{X,Y}_{(1),(2)}$, which is given by

$$\begin{aligned} l^{X,Y}_{(1),(2)} &= [\omega_{23}|l^{Y}_{(2)}] - [\omega_{12}+\omega_{23}|l^{XY}_{(2)}] + [\omega_{45}|l^{X,Y}_{(1),(1)}] \\ &= [\omega_{23}|\omega_{45}|\omega_{34}] - [\omega_{12}+\omega_{23}|\omega_{12}+\omega_{45}|\omega_{24}] + [\omega_{45}|\omega_{23}|\omega_{34}] \\ &\quad - [\omega_{45}|\omega_{12}|\omega_{24}] - [\omega_{45}|\omega_{23}|\omega_{24}] + [\omega_{45}|\omega_{34}|\omega_{24}] \end{aligned}$$

and also contains no such terms. Make the induction hypothesis that for all $\mathbf{a}, \mathbf{b}$ not both sequences of 1's with total length $< r+s$, then $l^{X,Y}_{\mathbf{a},\mathbf{b}}$ has no bar-words in only ω_{23}, ω_{34}. Consider a pair $(\mathbf{a}, \mathbf{b})$, not all 1's, of length $r+s$. The only terms which could occur with only ω_{23} and ω_{34} would come from the terms starting with one of these two elements in (35). Furthermore, those terms in (35) starting with ω_{23} or ω_{34} but followed by a term

$$l^{XY}_{\mathbf{a}}, \quad l^{XY}_{\mathbf{a}''} \quad \text{or} \quad l^{XY}_{\mathbf{b}} \tag{40}$$

cannot yield bar-words with just ω_{23} and ω_{34}, since $l^{XY}_{\mathbf{a}}$ always contains at least one ω_{24}. Thus in each of the eleven cases it is necessary to check that the remaining "risky" terms (the ones starting with ω_{23} or ω_{34} followed by a term not in (40)), can never yield a bar-word in ω_{23}, ω_{34}. There are no risky terms in the first case. In the second case, $[\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}]$ is risky, but in fact it cannot yield a word in ω_{23}, ω_{34} only by the induction hypothesis, since $a_r > 1$ appears in the pair $(\mathbf{a}, \mathbf{b}')$. The third case has no risky terms, and in the fourth, which contains both $[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}]$ and $[\omega_{23}|l^{X,Y}_{\mathbf{a}'',\mathbf{b}''}]$, we see that bad words cannot appear by induction, since $b_s > 1$ appears in both the pairs $(\mathbf{a}', \mathbf{b})$ and $(\mathbf{a}'', \mathbf{b}'')$. In the fifth case, the only risky term is $[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}]$, but by induction, this contains no bad terms since $b_s > 1$ appears. In the sixth case, there are three risky terms, $[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}]$, $[\omega_{23}|l^{X,Y}_{\mathbf{a}'',\mathbf{b}''}]$ and $[\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}]$, but the induction hypothesis works for all three again because since $(\mathbf{a}, \mathbf{b})$ are not all 1's and $a_r = b_s = 1$, none of the sequences $(\mathbf{a}', \mathbf{b})$, $(\mathbf{a}'', \mathbf{b}'')$ and $(\mathbf{a}, \mathbf{b}')$ can be all 1's.

In the seventh case, the risky term is $[\omega_{23}|l^{X,Y}_{\mathbf{a}',\mathbf{b}}]$ but again, $(\mathbf{a}', \mathbf{b})$ cannot be all 1's since $a_r = 1$. In the eighth case, there are two risky terms, $[\omega_{23}|l^{Y}_{\mathbf{b}}]$ and $-[\omega_{23}|l^{X,Y}_{(b_1),\mathbf{b}''}]$. The first term can contain only ω_{23} and ω_{34} only if $l^{Y}_{\mathbf{b}} = [\omega_{34}|\cdots|\omega_{34}]$, i.e. if $\mathbf{b} = (1, \ldots, 1)$, which is impossible since $b_s > 1$. The second works by induction since the pair $\mathbf{b}''$ is not all 1's, as $b_s > 1$. In the ninth case, the only risky term is $[\omega_{23}|l^{Y}_{\mathbf{b}}]$, which again can only be a word in ω_{23} and ω_{34} if $l^{Y}_{\mathbf{b}} = [\omega_{34}|\cdots|\omega_{34}]$, i.e. $\mathbf{b} = (1, \ldots, 1)$, which is impossible since $b_s = 1$ in this case. In the tenth case, the risky terms are $[\omega_{23}|l^{Y}_{\mathbf{b}}]$, $[\omega_{23}|l^{X,Y}_{(b_1),\mathbf{b}''}]$, and $[\omega_{34}|l^{X,Y}_{\mathbf{a},\mathbf{b}'}]$.

For either of the first two terms to have a word in ω_{23} and ω_{34} only, the sequence $\mathbf{b}$ would have to be all 1's, which is impossible since $\mathbf{a} = (1)$ in this case. As for the third term, $l^{X,Y}_{\mathbf{a},\mathbf{b}'}$ would have to be a word in ω_{23} and ω_{34} only, which is impossible since by definition ω_{12} and/or ω_{45} appear in this expression. Finally, the last case is excluded because $\mathbf{a} = (1)$, $\mathbf{b} = (1)$. This concludes the proof of the sixth and final statement of the lemma. Q.E.D.

3.3. Proof of Theorem 7

The version of the statement of Furusho's lemma given in the previous section, and the proof, are essentially just complete versions of the Lie part of the proof sketched in [6].

It is in the application of the lemma to the proof of the Lie version of his main theorem that the simplification is strong.

In [6], the proof of the main theorem is done in the unipotent situation, by generalizing these lemmas to that situation, defining a notion of regularization, and using a computation on regularizations due to Goncharov. In the Lie situation, however, thanks to Theorem 2 and the slight generalization of Furusho's lemma given in the sixth statement of Lemma 3, none of this is necessary. The desired result stated in Theorem 7 comes out immediately, as follows.

Proof of Theorem 7. Let $f \in \mathfrak{grt}_n$ and let $\mathbf{a} = (a_1, \ldots, a_r)$, $\mathbf{b} = (b_1, \ldots, b_s)$. Then by Lemma 3, we find that as long as $\mathbf{a}$ and $\mathbf{b}$ are not both sequences of 1's, we have:

$$\begin{cases} l^{XY}_{\mathbf{a}}\big(f(x_{45}, x_{51}) + f(x_{12}, x_{23})\big) = l_{\mathbf{a}}(f) \\ l^{X,Y}_{\mathbf{a},\mathbf{b}}\big(f(x_{45}, x_{51}) + f(x_{12}, x_{23})\big) = l^{X,Y}_{\mathbf{a},\mathbf{b}}\big(f(x_{45}, x_{51})\big) = l_{\mathbf{ab}}(f) \\ l^{Y,X}_{\mathbf{a},\mathbf{b}}\big(f(x_{45}, x_{51}) + f(x_{12}, x_{23})\big) \\ \qquad = l^{Y,X}_{\mathbf{a},\mathbf{b}}\big(f(x_{43}, x_{32}) + f(x_{21}, x_{15}) + f(x_{54}, x_{43})\big) \\ \qquad = l^{Y,X}_{\mathbf{a},\mathbf{b}}\big(f(x_{21}, x_{15})\big) = l_{\mathbf{ab}}(f). \end{cases}$$

So applying (36) to $f(x_{45}, x_{51}) + f(x_{12}, x_{23}) \in \mathrm{Lie}\, P_5$, we obtain the following identities for all pairs $(\mathbf{a}, \mathbf{b}) \neq \big((1, \ldots, 1), (1, \ldots, 1)\big)$, where $st(\mathbf{a}, \mathbf{b})$ and $ST(\mathbf{a}, \mathbf{b})$ are as in Definitions 1.1 and 3.2:

$$(41) \quad 0 = \sum_{\sigma(\mathbf{a},\mathbf{b}) \in \mathrm{ST}(\mathbf{a},\mathbf{b})} l^{\sigma(X,Y)}_{\sigma(\mathbf{a},\mathbf{b})}\big(f(x_{45}, x_{51}) + f(x_{12}, x_{23})\big)$$
$$= \sum_{\sigma \in \mathrm{Sh}^{\leq}(r,s)} l_{c^{\sigma}(\mathbf{a},\mathbf{b})}(f) = \sum_{\mathbf{c} \in st(\mathbf{a},\mathbf{b})} l_{\mathbf{c}}(f).$$

Set $F(x,y) = f(x,-y)$, and let F^Y denote the part of F consisting of words ending in y, rewritten in the variables y_i. Then by (34) and the fact that all words $w_{\mathbf{c}}$ end in y, we have $l_{\mathbf{c}}(f) = (-1)^r(f|w_{\mathbf{c}}) = (F|w_{\mathbf{c}}) = (F^Y|w_{\mathbf{c}})$, where $\mathbf{c} = (c_1, \dots, c_r)$, i.e. r is the "depth" of $\mathbf{c}$. So (41) yields

$$\sum_{\mathbf{c} \in st(\mathbf{a},\mathbf{b})} (F^Y|w_{\mathbf{c}}) = 0 \quad \text{for} \quad (\mathbf{a},\mathbf{b}) \neq \big((1,\dots,1),(1,\dots,1)\big).$$

Thus, F^Y satisfies the hypothesis of Theorem 2. Note also that $(F^Y|y_1^n) = 0$ since F is a Lie polynomial. So, setting $F_* = F^Y + \frac{(-1)^n}{n}(F|y_n)y_1^n$, Theorem 2 shows that F_* satisfies the stuffle relations for all pairs $(\mathbf{a},\mathbf{b})$. This means precisely that $F = f(x,-y) \in \mathfrak{ds}$, concluding the proof of Theorem 7. Q.E.D.

Acknowledgments. The authors warmly thank Jean Ecalle for taking the time to answer their questions and indicate proofs of symmetry properties of moulds. We also thank the referee for the large number of useful suggestions and corrections.

References

[1] A. Alekseev and C. Torossian, The Kashiwara–Vergne conjecture and Drinfel'd's associators, preprint, 2007, arXiv:0802.4300.

[2] F. Brown, Multiple zeta values and periods of moduli spaces $\overline{M}_{0,n}$, Ann. Sci. Éc. Norm. Supér. (4), **42** (2009), 371–489.

[3] S. Carr, Ph. D. thesis, 2008, arXiv:0911.2643.

[4] J. Ecalle, The flexion structure and dimorphy: flexion units, singulators, generators, and the enumeration of multizeta irreducibles, In: Asymptotics in Dynamics, Geometry and PDE's; Generalized Borel Summation, vol. II, (eds. O. Costin, F. Fauvet, F. Menous and D. Sauzin), CRM Series, Scuola Normale di Pisa, Edizioni della Normale, 2011.

[5] H. Furusho, Pentagon and hexagon equations, preprint, 2007, arXiv:0702128; Ann. of Math. (2), **171** (2010), 545–556.

[6] H. Furusho, Double shuffle relation for associators, preprint, 2009, arXiv:0808.0319; Ann. of Math. (2), **174** (2011), 341–360.

[7] Y. Ihara, On the stable derivation algebra associated with some braid groups, Israel J. Math., **80** (1992), 135–153.

[8] Y. Ihara, Some arithmetic aspects of Galois actions in the pro-p fundamental group of $\mathbb{P}^1 - \{0,1,\infty\}$, In: Arithmetic Fundamental Groups and Noncommutative Algebra, Proc. Sympos. Pure. Math., **70**, Amer. Math. Soc., Providence, RI, 2002, pp. 247–273.

[9] G. Racinet, Doubles mélanges des polylogarithmes multiples aux racines de l'unité, Publ. Math. Inst. Hautes Études Sci., **95** (2002), 185–231.

[10] J.-P. Serre, Lie Algebras and Lie Groups, Lecture Notes in Math., **1500**, Springer-Verlag, 2006 (orig. 1965).
[11] D. Zagier, Periods of modular forms, traces of Hecke operators, and multiple zeta values, Sûrikaisekikenkyûsho Kôkyûroku, **843** (1993), 162–170.

Sarah Carr
Ludwig-Maximilians-Universität München
Mathematisches Institut
Theresienstr. 39, D-80333 München
Germany

Leila Schneps
Equipe Analyse Algébrique, Institut de Mathématiques de Jussieu
4 place Jussieu, Case 247, 75252 Paris Cedex
France

E-mail address: sarah.c.carr@gmail.com,
leila@math.jussieu.fr