



ALGEBRAIC COMBINATORICS


Loïc Foissy, Frédéric Menous, Jean-Christophe Novelli & Jean-Yves Thibon

Quadri-algebras, preLie algebras, and the Catalan family of Lie idempotents

Volume 5, issue 4 (2022), p. 629-666.

<https://doi.org/10.5802/alco.224>

© The author(s), 2022.

 This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



*Algebraic Combinatorics is published by The Combinatorics Consortium
and is a member of the Centre Mersenne for Open Scientific Publishing*
www.tccpublishing.org www.centre-mersenne.org
e-ISSN: 2589-5486





Quadri-algebras, preLie algebras, and the Catalan family of Lie idempotents

Loïc Foissy, Frédéric Menous, Jean-Christophe Novelli
& Jean-Yves Thibon

ABSTRACT We compute the expansion of the Catalan family of Lie idempotents introduced in [Menous et al., *Adv. Applied Math.* 51 (2013), 177–22] on the PBW basis of the Lie module. It is found that the coefficient of a tree depends only on its number of left and right internal edges. In particular, the Catalan idempotents belong to a preLie algebra based on naked binary trees, of which we identify several Lie and preLie subalgebras.

1. INTRODUCTION

A *Lie idempotent* is an idempotent of the group algebra of the symmetric group which acts on the free associative algebra as a projector onto the free Lie algebra [30].

Historically, the first example of a Lie idempotent is provided by Dynkin’s theorem (1947, [9]): The linear map $\Theta_n : a_1 a_2 \cdots a_n \mapsto [\cdots [a_1, a_2], \cdots], a_n$, sending a word to its iterated bracketing, satisfies $\Theta_n \circ \Theta_n = n\Theta_n$. This result, originally intended as a mean to expanding the Hausdorff series $H(a_1, \dots, a_N) = \log(e^{a_1} \cdots e^{a_N})$ in terms of commutators, can also be used to give a new proof of the fact that H is a Lie series, and to prove the Poincaré–Birkhoff–Witt theorem. In particular one can deduce from it most basic facts about free Lie algebras, such as Friedrich’s criterion [4].

In 1969, Solomon [32] introduced another Lie idempotent, with the aim of providing a constructive proof of the Poincaré–Birkhoff–Witt theorem, that is, an explicit isomorphism between the universal enveloping algebra of a Lie algebra and its symmetric algebra. It turns out that this idempotent is also related to the Hausdorff series: it computes its expansion in terms of words, and has been rediscovered several times in this context (see, e.g. [3, 24]). Solomon’s idempotent is also known as the first Eulerian idempotent [16].

Finally, Witt’s formulas [34] for the dimensions of the multihomogeneous components of the free Lie algebra $L(V)$ generated by a vector space V show that, as a representation of $GL(V)$, $L_n(V)$ is isomorphic to the eigenspace with eigenvalue $e^{2i\pi/n}$ of the cyclic shift operator acting on $V^{\otimes n}$. An explicit intertwining operator for these representations is provided by Klyachko’s idempotent [14], introduced in 1974.

Prior to the introduction of noncommutative symmetric functions [12], these three examples, which are given by very different expressions, were the only known Lie

Manuscript received 29th April 2021, revised 9th January 2022, accepted 10th January 2022.

KEYWORDS. Noncommutative symmetric functions, Lie idempotents, Free Lie algebra, Dendriform algebras, PreLie algebras.

ACKNOWLEDGEMENTS. This research has been partially supported by the ANR program CARPLO (ANR-20-CE40-0007).

idempotents. They have however one important common point: they all belong to the *descent algebra*, a remarkable subalgebra of the group algebra of the symmetric group, introduced by Solomon [33] in 1976. The descent algebra is a noncommutative version of the character ring of the symmetric group, and noncommutative symmetric functions are to the descent algebra what ordinary symmetric functions are to the character ring. In particular, one can lift to the descent algebra various operations such as the $(1 - q)$ -transform and its inverse [15]. This allowed to give a complete characterization of Lie idempotents in the descent algebra, and to provide an explicit interpolation between all known examples.

It came therefore as a surprise that a sequence of operators defined in the context of resurgence theory [22], when interpreted as noncommutative symmetric functions, thanks to an isomorphism with Écalle’s Hopf algebra of alien operators [23], provided a new family of Lie idempotents of the descent algebras, not explained by the previous constructions.

In the sequel, we shall give a much simpler expression of these idempotents, in terms of the recently introduced PBW basis of the Lie module [31]. It is however not clear from this expression that the result belongs to the descent algebra. In the process of investigating the symmetries of Lie idempotents in this basis, we have been led to the discovery of various Lie and preLie subalgebras of the convolution Lie algebra of permutations. The PBW expansion of the Solomon idempotent, determined by Bandiera and Schätz [2], also satisfies the same type of symmetry. It involves a Lie algebra of binary trees, of which we provide an alternative interpretation. This Lie algebra have remarkable subalgebras, whose bases are parametrized by families of bicolored trees. We prove that Lie elements of the descent algebras all belong to one of these subalgebras and conjecture a similar statement for Lie elements of the Loday–Ronco algebra.

2. BACKGROUND

2.1. LIE IDEMPOTENTS. Let V be a vector space over some field \mathbb{K} of characteristic 0. Let $T(V)$ be its tensor algebra, and $L(V)$ the free Lie algebra generated by V . We denote by $L_n(V) = L(V) \cap V^{\otimes n}$ its homogeneous component of degree n . The group algebra $\mathbb{K}\mathfrak{S}_n$ of the symmetric group acts on the right on $V^{\otimes n}$ by

$$(1) \quad (v_1 \otimes v_2 \otimes \cdots \otimes v_n) \cdot \sigma = v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n)}.$$

This action commutes with the left action of $GL(V)$, and when $\dim V \geq n$, which we shall usually assume, these actions are the commutant of each other (Schur–Weyl duality).

Any $GL(V)$ -equivariant projector $\Pi_n : V^{\otimes n} \rightarrow L_n(V)$ can therefore be regarded as an idempotent π_n of $\mathbb{K}\mathfrak{S}_n$: $\Pi_n(\mathbf{v}) = \mathbf{v} \cdot \pi_n$. By definition (cf. [30]), such an element is called a *Lie idempotent* whenever its image is $L_n(V)$. Then, a homogeneous element $P_n \in V^{\otimes n}$ is in $L_n(V)$ if and only if $P_n \pi_n = P_n$.

From now on, we fix a basis $A = \{a_1, a_2, \dots\}$ of V . We identify $T(V)$ with the free associative algebra $\mathbb{K}\langle A \rangle$, and $L(V)$ with the free Lie algebra $L(A)$.

2.2. CONVOLUTION ALGEBRAS. For $\sigma \in \mathfrak{S}_n$, let g_σ be the endomorphism of $\mathbb{K}_n\langle A \rangle$ defined by

$$(2) \quad g_\sigma(a_1 a_2 \cdots a_n) = a_1 a_2 \cdots a_n \cdot \sigma = a_{\sigma(1)} a_{\sigma(2)} \cdots a_{\sigma(n)}.$$

The free algebra $\mathbb{K}\langle A \rangle$ is a graded bialgebra for the coproduct

$$(3) \quad \Delta(a) = a \otimes 1 + 1 \otimes a, \quad (a \in A),$$

so that a convolution product on the space of graded endomorphisms $\text{End}_{\text{gr}}(\mathbb{K}\langle A \rangle)$ can be defined by

$$(4) \quad f \star g(w) = \mu \circ (f \otimes g) \circ \Delta(w),$$

where μ denotes the multiplication of $\mathbb{K}\langle A \rangle$ (i.e. the concatenation product).

For permutations, this operation reads

$$(5) \quad g_\alpha \star g_\beta = \sum_{\substack{\text{std}(u)=\alpha \\ \text{std}(v)=\beta \\ \gamma=uv}} g_\gamma,$$

where $\text{std}(u)$ denotes the standardization of the word u , that is, the unique permutation having the same inversions as u .

We also define the \star product on permutations by the same formula. This convolution has been extensively studied by Reutenauer [29, 30] while investigating free Lie algebras and Lie idempotents.

When A is finite, the g_σ are not linearly independent anymore when the size of σ exceeds the cardinality of A . By taking an appropriate inverse limit over an increasing sequence of alphabets, one obtains a convolution algebra $\mathbb{K}\mathfrak{S}$ based on all permutations, and it has been shown by Malvenuto and Reutenauer [21] that this algebra acquires a Hopf algebra structure when endowed with the coproduct

$$(6) \quad \Delta\sigma = \sum_{u,v} \langle \sigma, u \sqcup v \rangle \text{std}(u) \otimes \text{std}(v),$$

where $\langle \sigma, u \sqcup v \rangle$ denotes the coefficient of σ in the shuffle product $u \sqcup v$ of the words u and v .

2.3. FREE QUASI-SYMMETRIC FUNCTIONS. The Malvenuto–Reutenauer Hopf algebra admits a convenient “polynomial realization”.

The algebra of *free quasi-symmetric functions* over a totally ordered alphabet A , denoted by $\mathbf{FQSym}(A)$, is the algebra spanned by the noncommutative “stable polynomials” (formal series of bounded degree, defined for any totally ordered alphabet)

$$(7) \quad \mathbf{G}_\sigma(A) := \sum_{\substack{w \in A^n \\ \text{std}(w)=\sigma}} w,$$

where σ is a permutation in the symmetric group \mathfrak{S}_n [8].

Its multiplication rule turns out to be given by

$$(8) \quad \mathbf{G}_\alpha \mathbf{G}_\beta = \sum_{\substack{\text{std}(u)=\alpha \\ \text{std}(v)=\beta \\ \gamma=uv}} \mathbf{G}_\gamma,$$

as in the Malvenuto–Reutenauer algebra $\mathbb{K}\mathfrak{S}$. Moreover, its natural coproduct, defined by the ordinal sum $A \oplus B$ of two mutually commuting alphabets

$$(9) \quad \Delta\mathbf{G}_\sigma := \mathbf{G}_\sigma(A \oplus B) = \sum_{u,v} \langle \sigma, u \sqcup v \rangle \mathbf{G}_{\text{std}(u)} \otimes \mathbf{G}_{\text{std}(v)},$$

also coincides with the Malvenuto–Reutenauer coproduct (identifying $F(A)G(B)$ with $F \otimes G$).

Thus, \mathbf{FQSym} is isomorphic to $\mathbb{K}\mathfrak{S}$ as a Hopf algebra.

In the sequel, we shall often identify permutations σ with the corresponding \mathbf{G}_σ without further notice.

2.4. DENDRIFORM STRUCTURE. A dendriform algebra [17, 18] is an associative algebra (\mathcal{A}, \cdot) endowed with two bilinear operations \prec, \succ , such that

$$(10) \quad a \cdot b = a \prec b + a \succ b,$$

satisfying the relations

$$(11) \quad (x \prec y) \prec z = x \prec (y \cdot z), (x \succ y) \prec z = x \succ (y \prec z), (x \cdot y) \succ z = x \succ (y \succ z).$$

The dendriform structure of **FQSym** is inherited from that of the free associative algebra over A , which is [25, 26]

$$(12) \quad u \prec v = \begin{cases} uv & \text{if } \max(v) < \max(u), \\ 0 & \text{otherwise,} \end{cases}$$

$$(13) \quad u \succ v = \begin{cases} uv & \text{if } \max(v) \geq \max(u), \\ 0 & \text{otherwise.} \end{cases}$$

This yields

$$(14) \quad \mathbf{G}_\alpha \mathbf{G}_\beta = \mathbf{G}_\alpha \prec \mathbf{G}_\beta + \mathbf{G}_\alpha \succ \mathbf{G}_\beta,$$

where

$$(15) \quad \mathbf{G}_\alpha \prec \mathbf{G}_\beta = \sum_{\substack{\gamma=uv \in \alpha \star \beta \\ |u|=|\alpha|; \max(v) < \max(u)}} \mathbf{G}_\gamma,$$

$$(16) \quad \mathbf{G}_\alpha \succ \mathbf{G}_\beta = \sum_{\substack{\gamma=u.v \in \alpha \star \beta \\ |u|=|\alpha|; \max(v) \geq \max(u)}} \mathbf{G}_\gamma,$$

where $\alpha \star \beta$ is interpreted as the set of permutations occurring in the convolution. Then $x = \mathbf{G}_1$ generates a free dendriform algebra in **FQSym**, isomorphic to **PBT**, the Loday–Ronco algebra of planar binary trees [19].

2.5. QUADRI-ALGEBRAS. The half-products of **FQSym**, which are defined by splitting the concatenation of two words according to the position of the greatest letter can again be refined according to the position of the smallest one: for $\alpha \in \mathfrak{S}_k, \beta \in \mathfrak{S}_l$ and $n = k + l$,

$$(17) \quad \mathbf{G}_\alpha \swarrow \mathbf{G}_\beta = \sum_{\substack{\gamma=uv \in \alpha \star \beta \\ 1, n \in u}} \mathbf{G}_\gamma,$$

$$(18) \quad \mathbf{G}_\alpha \swarrow \mathbf{G}_\beta = \sum_{\substack{\gamma=uv \in \alpha \star \beta \\ 1 \in v, n \in u}} \mathbf{G}_\gamma,$$

$$(19) \quad \mathbf{G}_\alpha \searrow \mathbf{G}_\beta = \sum_{\substack{\gamma=uv \in \alpha \star \beta \\ 1, n \in v}} \mathbf{G}_\gamma,$$

$$(20) \quad \mathbf{G}_\alpha \searrow \mathbf{G}_\beta = \sum_{\substack{\gamma=uv \in \alpha \star \beta \\ 1 \in u, n \in v}} \mathbf{G}_\gamma.$$

The relations satisfied by these partial products have led Aguiar and Loday to the notion of *quadri-algebra* [1].

A quadri-algebra is a family $(A, \leftarrow, \swarrow, \searrow, \nearrow)$, where A is a vector space and $\leftarrow, \swarrow, \searrow, \nearrow$ are products on A , such that for all $x, y, z \in A$:

$$(21) \quad \begin{aligned} (x \leftarrow y) \leftarrow z &= x \leftarrow (y \star z), \\ (x \nearrow y) \leftarrow z &= x \nearrow (y \leftarrow z), \\ (x \uparrow y) \nearrow z &= x \nearrow (y \rightarrow z), \end{aligned}$$

$$(22) \quad \begin{aligned} (x \swarrow y) \leftarrow z &= x \swarrow (y \uparrow z), \\ (x \searrow y) \leftarrow z &= x \searrow (y \leftarrow z), \\ (x \downarrow y) \nearrow z &= x \searrow (y \nearrow z), \end{aligned}$$

$$(23) \quad \begin{aligned} (x \leftarrow y) \swarrow z &= x \swarrow (y \downarrow z), \\ (x \rightarrow y) \swarrow z &= x \searrow (y \swarrow z), \\ (x \star y) \searrow z &= x \searrow (y \searrow z), \end{aligned}$$

where:

$$(24) \quad \leftarrow = \leftarrow + \swarrow, \quad \rightarrow = \nearrow + \searrow, \quad \uparrow = \leftarrow + \nearrow, \quad \downarrow = \swarrow + \searrow,$$

$$(25) \quad \star = \leftarrow + \swarrow + \searrow + \nearrow = \leftarrow + \rightarrow = \uparrow + \downarrow.$$

The augmentation ideal \mathbf{FQSym}_+ of the Hopf algebra \mathbf{FQSym} is a quadri-algebra for these operations [1], up to the involution $\sigma \mapsto \sigma^{-1}$.

As \mathbf{FQSym} is self-dual, its coproduct can also be split into four parts, making it a quadri-coalgebra. Dualizing, with the pairing defined by $\langle \mathbf{G}_\sigma, \mathbf{G}_\tau \rangle = \delta_{\sigma, \tau^{-1}}$, we obtain a quadri-coalgebra structure on \mathbf{FQSym}_+ , defined by:

$$(26) \quad \Delta_{\leftarrow}(\mathbf{G}_\sigma) = \sum_{\sigma(1), \sigma(n) \leq i < n} \mathbf{G}_{\sigma_{\{1, \dots, i\}}} \otimes \mathbf{G}_{\text{std}(\sigma_{\{1+i, \dots, n\}})}$$

$$(27) \quad \Delta_{\swarrow}(\mathbf{G}_\sigma) = \sum_{\sigma(n) \leq i < \sigma(1)} \mathbf{G}_{\sigma_{\{1, \dots, i\}}} \otimes \mathbf{G}_{\text{std}(\sigma_{\{1+i, \dots, n\}})}$$

$$(28) \quad \Delta_{\searrow}(\mathbf{G}_\sigma) = \sum_{1 \leq i < \sigma(1), \sigma(n)} \mathbf{G}_{\sigma_{\{1, \dots, i\}}} \otimes \mathbf{G}_{\text{std}(\sigma_{\{1+i, \dots, n\}})}$$

$$(29) \quad \Delta_{\nearrow}(\mathbf{G}_\sigma) = \sum_{\sigma(1) \leq i < \sigma(n)} \mathbf{G}_{\sigma_{\{1, \dots, i\}}} \otimes \mathbf{G}_{\text{std}(\sigma_{\{1+i, \dots, n\}})},$$

where for all $I \subseteq \{1, \dots, n\}$, σ_I is the word obtained by deleting in σ the letters which do not belong to I . The sum of the four coproducts is the usual coproduct $\tilde{\Delta}$:

$$(30) \quad \tilde{\Delta}(\mathbf{G}_\sigma) = \sum_{1 \leq i < n} \mathbf{G}_{\sigma_{\{1, \dots, i\}}} \otimes \mathbf{G}_{\text{std}(\sigma_{\{1+i, \dots, n\}})}.$$

Moreover, for any $a, b \in \mathbf{FQSym}$:

$$(31) \quad \tilde{\Delta}(a \leftarrow b) = a' \uparrow b \otimes a'' + a' \otimes a'' \leftarrow b + a' \uparrow b' \otimes a'' \leftarrow b'',$$

$$(32) \quad \tilde{\Delta}(a \swarrow b) = b \otimes a + a' \downarrow b \otimes a'' + b' \otimes a \leftarrow b'' + a' \downarrow b' \otimes a'' \leftarrow b'',$$

$$(33) \quad \tilde{\Delta}(a \searrow b) = a \downarrow b' \otimes b'' + b' \otimes a \rightarrow b'' + a' \downarrow b' \otimes a'' \rightarrow b'',$$

$$(34) \quad \tilde{\Delta}(a \nearrow b) = a \otimes b + a \uparrow b' \otimes b'' + a' \otimes a'' \rightarrow b + a' \uparrow b' \otimes a'' \rightarrow b''.$$

As a consequence, if a and b are two primitive elements of \mathbf{FQSym} :

$$(35) \quad \tilde{\Delta}(a \nearrow b - b \swarrow a) = a \otimes b - a \otimes b = 0,$$

so $a \nearrow b - b \swarrow a$ is also primitive.

PROPOSITION 2.1. *The operation*

$$(36) \quad x \nearrow \swarrow y = x \nearrow y - y \swarrow x$$

preserves primitive elements.

2.6. SOLOMON’S DESCENT ALGEBRA. The descent algebras have been introduced by Solomon [33] for general finite Coxeter groups in the following way. Let (W, S) be a Coxeter system. One says that $w \in W$ has a descent at $s \in S$ if w has a reduced word ending by s . For $W = \mathfrak{S}_n$ and $s_i = (i, i + 1)$, this means that $w(i) > w(i + 1)$, whence the terminology. In this case, we rather say that i is a descent of w . Let $\text{Des}(w)$ denote the descent set of w , and for a subset $E \subseteq S$, set

$$(37) \quad D_E = \sum_{\text{Des}(w)=E} w \in \mathbb{Z}W.$$

Solomon has shown that the D_E span a \mathbb{Z} -subalgebra Σ_n of $\mathbb{Z}W$. Moreover,

$$(38) \quad D_{E'} D_{E''} = \sum_E c_{E'E''}^E D_E,$$

where the coefficients $c_{E'E''}^E$ are nonnegative integers.

In the case of \mathfrak{S}_n , it is convenient to encode descent sets by compositions of n . If $E = \{d_1, \dots, d_{r-1}\}$, we set $d_0 = 0$, $d_r = n$ and $I = C(E) = (i_1, \dots, i_r)$, where $i_k = d_k - d_{k-1}$. We also say that E is the descent set of I . From now on, we shall write D_I instead of D_E . We shall also write $C(\sigma) = I$ if the descent set of σ is E .

Most known examples of Lie idempotents turn out to belong to the descent algebra Σ_n . The simplest (and oldest) example is the *Dynkin idempotent* [9]

$$(39) \quad \theta_n = \frac{1}{n} [\dots [[1, 2], 3], \dots, n] = \frac{1}{n} \sum_{k=0}^{n-1} (-1)^k D_{1^k, n-k}.$$

The Solomon idempotent [32], or first Eulerian idempotent is

$$(40) \quad \phi_n = \frac{1}{n} \sum_{\sigma \in \mathfrak{S}_n} \frac{(-1)^{d(\sigma)}}{\binom{n-1}{d(\sigma)}} \sigma$$

and the Klyachko idempotent [14] is

$$(41) \quad \kappa_n = \frac{1}{n} \sum_{\sigma \in \mathfrak{S}_n} \omega^{\text{maj}(\sigma)} \sigma,$$

where $d(\sigma)$ is the number of descents of σ , $\text{maj}(\sigma)$ is their sum, and $\omega = e^{2i\pi/n}$. Note that on these expressions, one easily sees that both the Solomon idempotent and the Klyachko idempotent live inside the descent algebra since the coefficient of a permutation σ only depends on its descents.

There is a one-parameter family (the q -Eulerians) interpolating between these three examples [5, 15], and more recently, another one-parameter family (the Catalan family) related to mould calculus and random walks on the line has been introduced [23]. This family is given by an explicit expansion in terms of descent classes. One of the aims of the present paper is to give its expansion on the so-called Poincaré–Birkhoff–Witt basis, to be defined below.

2.7. NONCOMMUTATIVE SYMMETRIC FUNCTIONS. The algebra of ordinary symmetric functions Sym can be regarded as the free associative and commutative algebra over an infinite sequence $(h_n)_{n \geq 1}$ of homogeneous generators (h_n is of degree n), so that its linear bases in degree n are naturally labelled by partitions of n (e.g. products $h_\mu = h_{\mu_1} \cdots h_{\mu_r}$ of complete homogeneous functions) [20].

Similarly, the algebra **Sym** of noncommutative symmetric functions is the free associative (but noncommutative) algebra over an infinite sequence $(S_n)_{n \geq 1}$ of homogeneous generators, endowed with a natural homomorphism $S_n \mapsto h_n$ (commutative image) [12].

Thus, linear bases of the homogeneous component \mathbf{Sym}_n of **Sym** are labelled by compositions of n , exactly as those of the descent algebra Σ_n of \mathfrak{S}_n .

Noncommutative symmetric functions can be realized in terms of an auxiliary (totally ordered) alphabet $A = \{a_1, a_2, \dots\}$ by setting

$$(42) \quad S_n(A) = \sum_{i_1 \leq i_2 \leq \dots \leq i_n} a_{i_1} a_{i_2} \cdots a_{i_n} = \mathbf{G}_{12 \dots n}(A),$$

that is, S_n is the sum of nondecreasing words, or, otherwise said, words with no descent. Then, obviously,

$$(43) \quad S^I = S_{i_1} S_{i_2} \cdots S_{i_r}$$

is the sum of words whose descent set is contained in $\text{Des}(I)$ (the descents of a word are defined as for permutations as those i such that $w_i > w_{i+1}$, and are similarly encoded as compositions $C(w)$ of n). Since $S_n(A) = \mathbf{G}_{12 \dots n}(A)$, **Sym**(A) is actually a subalgebra of **FQSym**(A).

Introducing the *noncommutative ribbon Schur functions*

$$(44) \quad R_I(A) = \sum_{C(w)=I} w = \sum_{C(\sigma)=I} \mathbf{G}_\sigma(A)$$

whose commutative image are indeed the skew Schur functions indexed by ribbon diagrams, we have

$$(45) \quad S^I = \sum_{J \leq I} R_J$$

where $J \leq I$ is the *reverse refinement order*, which means that $\text{Des}(J) \subseteq \text{Des}(I)$.

The linear map defined by $\alpha : D_I \rightarrow R_I$ appears therefore as a natural choice for a correspondence $\Sigma_n \rightarrow \mathbf{Sym}_n$. This choice is actually *canonical*. Indeed, there is a natural way to introduce an *internal product* $*$ on **Sym**, by dualizing a natural coproduct of its graded dual (which is the so-called Hopf algebra of quasi-symmetric functions).

For this product, α is an *anti-isomorphism*. This is convenient, because we want to interpret permutations as endomorphisms of tensor algebras: if $g_\sigma(w) = w\sigma$, then $g_\sigma \circ g_\tau = g_{\tau\sigma}$.

The *internal product* $*$ is consistently defined on **FQSym** by

$$(46) \quad \mathbf{G}_\sigma * \mathbf{G}_\tau = \begin{cases} \mathbf{G}_{\tau\sigma} & \text{if } |\tau| = |\sigma|, \\ 0 & \text{otherwise,} \end{cases}$$

where $\tau\sigma$ is the usual product of the symmetric group.

The fundamental property for computing with the internal product in **Sym** is the following *splitting formula*.

PROPOSITION 2.2 ([12]). *Let $F_1, F_2, \dots, F_r, G \in \mathbf{Sym}$. Then,*

$$(47) \quad (F_1 F_2 \cdots F_r) * G = \mu_r [(F_1 \otimes \cdots \otimes F_r) * \Delta^r G]$$

where in the right-hand side, μ_r denotes the r -fold ordinary multiplication and $*$ stands for the operation induced on $\mathbf{Sym}^{\otimes n}$ by $*$.

2.8. LIE IDEMPOTENTS AS NONCOMMUTATIVE SYMMETRIC FUNCTIONS. The first really interesting question about noncommutative symmetric functions is perhaps “what are the noncommutative power sums?”. Indeed, the answer to this question is far from being unique.

If one starts from the classical expression

$$(48) \quad \sigma_t(X) = \sum_{n \geq 0} h_n(X)t^n = \exp \left\{ \sum_{k \geq 1} p_k \frac{t^k}{k} \right\},$$

one can choose to define noncommutative power sums Φ_k by the same formula

$$(49) \quad \sigma_t(A) = \sum_{n \geq 0} S_n(A)t^n = \exp \left\{ \sum_{k \geq 1} \Phi_k \frac{t^k}{k} \right\},$$

but a noncommutative version of the Newton formulas

$$(50) \quad nh_n = h_{n-1}p_1 + h_{n-2}p_2 + \cdots + p_n$$

which are derived by taking the logarithmic derivative of (48) leads to different noncommutative power-sums Ψ_k inductively defined by

$$(51) \quad nS_n = S_{n-1}\Psi_1 + S_{n-2}\Psi_2 + \cdots + \Psi_n.$$

A bit of computation reveals then that

$$(52) \quad \Psi_n = R_n - R_{1,n-1} + R_{1,1,n-2} - \cdots = \sum_{k=0}^{n-1} (-1)^k R_{1^k, n-k},$$

which is analogous to the classical expression of p_n as the alternating sum of hook Schur functions. Therefore, in the descent algebra, Ψ_n corresponds to Dynkin’s element, $n\theta_n$ (see (39)).

The Φ_n can also be expressed on the ribbon basis without much difficulty, and one finds

$$(53) \quad \Phi_n = \sum_{|I|=n} \frac{(-1)^{l(I)-1}}{\binom{n-1}{l(I)-1}} R_I,$$

so that Φ_n corresponds to $n\phi_n$ (see (40)).

The case of Klyachko’s idempotent is even more interesting: it is related to the so-called $(1 - q)$ -transform (see [15]).

The following result is proved in [12].

THEOREM 2.3. *Let $F = \alpha(\pi)$ be an element of \mathbf{Sym}_n , where $\pi \in \Sigma_n$. The following assertions are equivalent:*

- (1) π is a Lie quasi-idempotent;
- (2) F is a primitive element for Δ ;
- (3) F belongs to the Lie algebra $L(\Psi)$ generated by the Ψ_n .

Moreover, π is a Lie idempotent iff $F - \frac{1}{n}\Psi_n$ is in the Lie ideal $[L(\Psi), L(\Psi)]$.

Thus, Lie idempotents are essentially the same thing as “noncommutative power sums” (up to a factor n), and we shall from now on identify both notions: a Lie idempotent in \mathbf{Sym}_n is a primitive element whose commutative image is p_n/n .

Noncommutative symmetric functions can also be regarded as elements of the free dendriform algebra on one generator, and Lie idempotents can be expressed in terms of its preLie structure. Indeed, there is a Hopf embedding $\iota : \mathbf{Sym} \rightarrow \mathbf{PBT}$ of noncommutative symmetric functions into \mathbf{PBT} [8, 12, 13], which is given by

$$(54) \quad \iota(S_n) = (\dots((x \succ x) \succ x) \dots) \succ x \quad (n \text{ times}).$$

For example, the Dynkin elements are

$$(55) \quad \iota(\Psi_n) = (\dots((x \triangleright x) \triangleright x) \dots) \triangleright x \quad (n \text{ times}),$$

where

$$(56) \quad a \triangleright b = a \succ b - b \prec a$$

is the preLie product defined in any dendriform algebra.

Using the embedding in \mathbf{FQSym} , the proof of this identity is remarkably simple. Indeed,

$$(57) \quad \mathbf{G}_\sigma \succ x = \mathbf{G}_{\sigma \cdot (n+1)} \text{ and } x \prec \mathbf{G}_\sigma = \mathbf{G}_{(n+1) \cdot \sigma},$$

so that $\iota(S_n) = \mathbf{G}_{12\dots n}$. In terms of permutations, this is therefore the standard embedding of \mathbf{Sym} into \mathbf{FQSym} as the descent algebra, for which, identifying \mathbf{G}_σ with σ ,

$$(58) \quad \Psi_n = [[\dots, [1, 2], \dots, n - 1], n].$$

It is then clear that

$$(59) \quad x \triangleright x = \mathbf{G}_{12} - \mathbf{G}_{21} = R_2 - R_{11} = \Psi_2$$

$$(60) \quad \Psi_2 \triangleright x = \mathbf{G}_{123} - \mathbf{G}_{213} - \mathbf{G}_{312} + \mathbf{G}_{321} = R_3 - R_{12} + R_{111} = \Psi_3$$

$$(61) \quad \Psi_{n-1} \triangleright x = \mathbf{G}_{12\dots n} - \dots \pm \mathbf{G}_{n\dots 21} = \sum_{k=0}^{n-1} (-1)^k R_{1^k, n-k} = \Psi_n.$$

2.9. THE CATALAN FAMILY. Recently, Theorem 2.3 has been used to identify a new family of Lie idempotents of the descent algebras [23]. Families of coefficients obtained from iterated integrals coming from resurgence theory turned out to be interpretable in terms of noncommutative symmetric functions, thanks to a highly nontrivial isomorphism with a certain Hopf algebra of Écalle’s alien operators.

Consider the generating series

$$(62) \quad \mathbf{ca}(a, b, t) = \frac{1 - (a + b)t - \sqrt{1 - 2(a + b)t + (b - a)^2 t^2}}{2abt} = \sum_{n \geq 1} \mathbf{ca}_n(a, b) t^n.$$

The coefficients $\mathbf{ca}_n(a, b)$ are homogeneous and symmetric polynomials in a, b of degree $n - 1$, which reduce to the Catalan numbers for $a = b = 1$:

$$(63) \quad \begin{aligned} \mathbf{ca}_1(a, b) &= 1 \\ \mathbf{ca}_2(a, b) &= a + b \\ \mathbf{ca}_3(a, b) &= a^2 + 3ab + b^2 \\ \mathbf{ca}_4(a, b) &= a^3 + 6a^2b + 6ab^2 + b^3 \\ \mathbf{ca}_5(a, b) &= a^4 + 10a^3b + 20a^2b^2 + 10ab^3 + b^4. \end{aligned}$$

The coefficients of these polynomials are the Narayana numbers

$$(64) \quad T(n, k) = \frac{1}{k} \binom{n-1}{k-1} \binom{n}{k-1},$$

so that

$$(65) \quad \mathbf{ca}_n(a, b) = \sum_{i=0}^{n-1} T(n, i + 1) a^i b^{n-1-i}.$$

For any sequence of signs $\underline{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)$ ($n \geq 1$), consider its minimal decomposition into stacks of identical signs

$$(66) \quad \underline{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n) = (\eta_1)^{n_1} \dots (\eta_s)^{n_s},$$

with $\eta_i \neq \eta_{i+1}$ and $n_1 + \dots + n_s = n$.

Define the *signed ribbons* as

$$(67) \quad R_{\underline{\varepsilon}\bullet} = (-1)^{l(I)-1} R_I \quad (R_{\emptyset} = 1, R_{\bullet} = R_1),$$

where I is the composition such that

$$(68) \quad \text{Des}(I) = \{1 \leq i \leq n - 1 ; \varepsilon_i = -\}.$$

The following result is proved in [23].

THEOREM 2.4. *For $n \geq 1$, the element of \mathbf{Sym}_{n+1} defined by*

$$(69) \quad D_{a,b}^{n+1} = \sum_{\underline{\varepsilon}=(\eta_1)^{n_1}\dots(\eta_s)^{n_s}} \left(\prod_{\substack{\eta_i=+ \\ i < s}} a \right) \left(\prod_{\substack{\eta_i=- \\ i < s}} b \right) \mathbf{ca}_{n_1}(a, b) \dots \mathbf{ca}_{n_s}(a, b) R_{\underline{\varepsilon}\bullet}$$

is primitive, and corresponds (up to a scalar factor) to a Lie idempotent of the descent algebra.

For example, using the correspondence with the usual noncommutative ribbon Schur functions,

$$(70) \quad \begin{aligned} D_{a,b}^2 &= R_2 - R_{11} \\ D_{a,b}^3 &= (a + b)R_3 - aR_{21} - bR_{12} + (a + b)R_{111} \\ D_{a,b}^4 &= (a^2 + 3ab + b^2)R_4 - a(a + b)R_{31} - abR_{22} - (a + b)bR_{13} \\ &\quad + a(a + b)R_{211} + abR_{121} + (a + b)bR_{112} - (a^2 + 3ab + b^2)R_{1111}. \end{aligned}$$

3. THE LIE MODULE AND ITS BASES

3.1. THE LIE MODULE. Lie idempotents belong to a subspace of $\mathbb{K}\mathfrak{S}_n$ called the *Lie module*, denoted by $\text{Lie}(n)$. It is the linear span of the complete bracketing of permutations, regarded as words over the alphabet $\{1, 2, \dots, n\}$.

Of course, these bracketings are not linearly independent, and the dimension of $\text{Lie}(n)$ is actually $(n - 1)!$. The simplest basis of $\text{Lie}(n)$ is the Dynkin basis

$$(71) \quad D_{\sigma} = [\dots [[1, \sigma(2)], \sigma(3)], \dots, \sigma(n)]$$

parametrized by permutations fixing 1, or equivalently

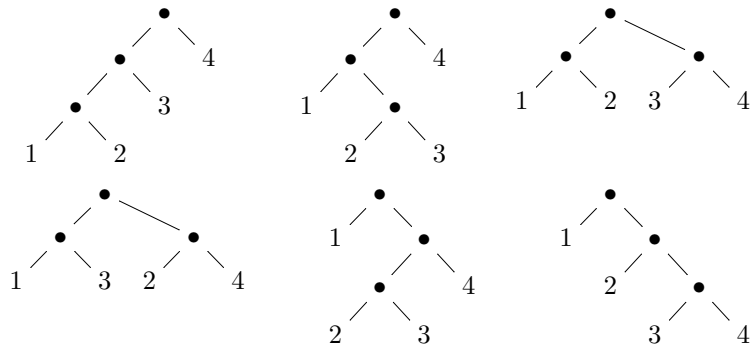
$$(72) \quad D'_{\sigma} = [\sigma(1), [\sigma(2), \dots, [\sigma(n - 1), n] \dots]]$$

parametrized by permutations fixing n . The expansions of various Lie idempotents in this basis can be found in [15].

3.2. THE POINCARÉ–BIRKHOFF–WITT BASIS. A less obvious basis is the so-called Poincaré–Birkhoff–Witt basis [2, 31]. Its elements are complete bracketings of permutations, represented by complete binary trees with labelled leaves, such that for each internal node, the smallest label is in the left subtree, and the greatest label is in the right subtree.

Such a labelling of the leaves will be said to be *admissible*. In particular, the leftmost leaf is always 1, and the rightmost one always n .

For example, the admissible labellings of all binary trees with three internal nodes are



The corresponding Lie elements are

$$(73) \quad \begin{aligned} &[[[1, 2], 3], 4], \quad [[1, [2, 3]], 4] \quad [[1, 2], [3, 4]], \\ &[[1, 3], [2, 4]], \quad [1, [[2, 3], 4]] \quad [1, [2, [3, 4]]], \end{aligned}$$

the smallest example of a tree with two admissible labellings being

$$(74) \quad [[1, 2], [3, 4]], \quad [[1, 3], [2, 4]].$$

Note in particular that there is only one admissible labelling of the left-comb tree and that this element is equal to the Dynkin element (39).

It is known [31] that the number of such labelled trees is indeed $(n - 1)!$ and that they are linearly independent. More precisely, one can count the number of admissible labellings of a given tree.

Recall that the *decreasing tree* $DT(w)$ of a word without repeated letters $w = unv$ is the binary tree with root $n = \max(w)$, left subtree $DT(u)$ and right subtree $DT(v)$.

The number of permutations of \mathfrak{S}_n with a given decreasing tree t is given by the *hook-length formula*

$$(75) \quad N(t) = n! \prod_{v \in t} \frac{1}{h(v)}$$

where the hook-length $h(v)$ of a node v is the number of nodes of its subtree.

One can show (see Appendix) that the admissible labellings of a complete binary tree T are in bijection with the permutations whose decreasing tree has shape t , the binary tree consisting of the internal nodes of T .

3.3. THE CONVOLUTION LIE ALGEBRA AND ITS CATALAN SUBALGEBRA. The following easy proposition appears to have remained unnoticed.

PROPOSITION 3.1.

(1) *The direct sum*

$$(76) \quad \text{Lie} = \bigoplus_{n \geq 1} \text{Lie}(n)$$

is a Lie algebra for the convolution bracket

$$(77) \quad [\alpha, \beta]_{\star} = \alpha \star \beta - \beta \star \alpha$$

of $\mathbb{K}\mathfrak{S}$, where \star is defined as in (5).

(2) *Representing its elements as complete binary trees with labelled leaves $T(\sigma)$, one has*

$$(78) \quad [T_1(\alpha), T_2(\beta)]_{\star} = \sum_{\gamma = \gamma_1 \cdot \gamma_2 \in \alpha \star \beta} [T_1(\gamma_1), T_2(\gamma_2)]$$

where γ_1 is the prefix of γ of size equal to the size of α . Note that in the r.h.s., the bracket is taken with respect to concatenation.

(3) *Lie is (strictly) contained in the primitive Lie algebra of $\mathbb{K}\mathfrak{S}$.*

For example

$$(79) \quad \begin{aligned} [[1, 2], 1]_{\star} &= [1, 2] \star 1 - 1 \star [1, 2] \\ &= [1, 2].3 + [1, 3].2 + [2, 3].1 - 1.[2, 3] - 2.[1, 3] - 3.[1, 2] \\ &= [[1, 2], 3] + [[1, 3], 2] + [[2, 3], 1]. \end{aligned}$$

Proof. Points (1) and (2) are consequences of the following identity. Let α and β be two permutations. Then,

$$(80) \quad [\alpha, \beta]_{\star} = \alpha \star \beta - \beta \star \alpha = \sum_{\gamma = \gamma_1 \cdot \gamma_2 \in \alpha \star \beta} (\gamma_1 \cdot \gamma_2 - \gamma_2 \cdot \gamma_1),$$

which is immediate from the expression of the convolution product (5).

Point (3) is a consequence of Ree’s criterion (cf. [30]): in the free associative algebra, the free Lie algebra is the orthogonal of the space of proper shuffles. But in $\mathbb{K}\mathfrak{S}$, an element is primitive iff it is orthogonal to special shuffles $u \sqcup v$, where u and v are permutations of consecutive intervals $[1, k]$ and $[k + 1, n]$. \square

As proved in [31], the PBW basis is a \mathbb{Z} -basis of $\text{Lie}(n)$.

It is known that the primitive Lie algebra of $\mathbb{K}\mathfrak{S}$ is free [10]. Therefore, Lie , the convolution Lie algebra, which is a Lie subalgebra of the previous one, is free as well (see [30, Section 2.2]).

Moreover, the PBW basis of Lie allows one to define an interesting Lie subalgebra implicitly defined in [2].

THEOREM 3.2 ([2]). *For a complete binary tree T , let*

$$(81) \quad c_T = \sum_{\sigma \text{ admissible}} T(\sigma).$$

Then, the c_T span a Lie subalgebra \mathfrak{C} of Lie.

Our proof of this fact relies upon the quadri-algebra structure, introduced in Section 4. Actually, \mathfrak{C} is also a preLie algebra as already observed in [2].

Let T and T' be two binary trees. We shall denote by $T \wedge T'$ the tree whose left (resp. right) subtree is T (resp. T'). This grafting operation is understood as bilinear. Let \triangleright be the preLie product as in Eq. (56).

EXAMPLE 3.3. Let us compute $1 \triangleright [1, 2]$. By definition,

$$(82) \quad 1 \triangleright [1, 2] = 1 \succ (12 - 21) - (12 - 21) \prec 1$$

and

$$(83) \quad 1 \succ 12 = 123 + 213,$$

$$(84) \quad 1 \prec 21 = 132 + 231,$$

$$(85) \quad 12 \prec 1 = 231 + 132,$$

$$(86) \quad 21 \prec 1 = 321 + 312,$$

so that

$$(87) \quad 1 \triangleright [1, 2] = 123 + 213 - 2 \times (132 + 231) + 312 + 321 = 2 \times [1, [2, 3]] - [[1, 2], 3]$$

is a \mathbb{Z} -linear combination of the two elements of degree 3 of the PBW basis.

PROPOSITION 3.4. \mathfrak{C} is stable for the preLie product \triangleright .

Writing T for c_T , the preLie product satisfies the recursion

$$(88) \quad T_1 \triangleright (T_2 \wedge T_3) = (T_1 \wedge (T_2 \wedge T_3)) + ((T_1 \triangleright T_2) \wedge T_3) + (T_2 \wedge (T_1 \triangleright T_3)) - ((T_1 \wedge T_2) \wedge T_3) - ((T_2 \wedge T_1) \wedge T_3),$$

or, pictorially,

$$(89) \quad \begin{array}{c} T_1 \triangleright \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_3 \end{array} = \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_3 \end{array} \end{array} + \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \triangleright T_2 \quad T_3 \end{array} + \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_1 \triangleright T_3 \end{array} \\ - \begin{array}{c} \bullet \\ / \quad \backslash \\ \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad T_2 \end{array} \quad T_3 \end{array} - \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad T_3 \end{array} \end{array} \end{array}$$

Proof. Let us define a new law \triangleright' as $T \triangleright' T' = T \triangleright T' - T \wedge T'$. Then the statement rewrites as

$$(90) \quad \begin{array}{c} T_1 \triangleright' \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_3 \end{array} = \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \triangleright' T_2 \quad T_3 \end{array} + \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_1 \triangleright' T_3 \end{array} \\ + \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad T_3 \end{array} \end{array} - \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad T_3 \end{array} \end{array} \end{array}$$

We have seen that the product of **FQSym** can be split into the sum of the four quadri-algebra products, so that the Lie bracket consists of eight different terms. Let us group the terms as follows:

$$(91) \quad \begin{array}{l} S_{E-W}(T_1, T_2) = T_1 \searrow T_2 - T_2 \swarrow T_1 \\ S_{W-E}(T_1, T_2) = T_1 \swarrow T_2 - T_2 \searrow T_1 \\ N_{W-E}(T_1, T_2) = T_1 \nwarrow T_2 - T_2 \nearrow T_1 \\ N_{E-W}(T_1, T_2) = T_1 \nearrow T_2 - T_2 \nwarrow T_1 \end{array}$$

Then,

$$(92) \quad S_{E-W}(T_1, T_2) = \begin{array}{c} \bullet \\ / \quad \backslash \\ T_1 \quad T_2 \end{array}, \quad S_{W-E}(T_1, T_2) = - \begin{array}{c} \bullet \\ / \quad \backslash \\ T_2 \quad T_1 \end{array},$$

(93)

$$\begin{aligned}
 N_{W-E} \left(T_1 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ T_2 \end{array}, T_3 \right) &= S_{E-W}(T_1, T_3) \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ T_2 \end{array} + N_{W-E}(T_1, T_3) \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ T_2 \end{array} \\
 &+ T_1 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ S_{W-E}(T_2, T_3) \end{array} T_1 \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ N_{W-E}(T_2, T_3) \end{array},
 \end{aligned}$$

and

(94)
$$N_{E-W}(T_1, T_2) = -N_{W-E}(T_2, T_1).$$

All these formulas are easily obtained by following carefully what happens to the smallest and largest values in the different products.

It follows that

(95)
$$T_1 \triangleright T_2 = T_1 \succ T_2 - T_2 \prec T_1 = S_{E-W}(T_1, T_2) + N_{E-W}(T_1, T_2),$$

whence the formula for the pre-Lie product of naked trees (89). □

In [2], a pre-Lie algebra structure is defined on the linear span of (abstract) complete binary trees, by means of a combinatorial formula defined in terms of graftings. We are now in a position to see that our pre-Lie structure on \mathfrak{C} coincides with that of [2].

PROPOSITION 3.5. *The Catalan preLie algebra \mathfrak{C} is isomorphic to the preLie algebra \mathcal{T}_{pb} defined in [2, Sec. 3.2].*

Proof. In the setting of [2], computing $T \triangleright T'$ amounts to taking the sum of all (naked) trees obtained by creating a new internal node n in the middle of any right branch and gluing T as left subtree of n , the right subtree of n being what was below n in the beginning, minus the similar sum where one creates a new internal node in the middle of any left branch and gluing T as right subtree. Indeed, Formula (90) consists in the creation of two nodes on the branches connected to the root (last two terms) and in the induction on the left and right subtrees (first two terms). □

Recall that the primitive Lie algebra of **Sym** is freely generated by the Ψ_n . Since $\Psi_n = (\cdots (\mathbf{G}_1 \triangleright \mathbf{G}_1) \triangleright \cdots) \triangleright \mathbf{G}_1 \in \mathfrak{C}_n$, we have:

PROPOSITION 3.6. *The primitive Lie algebra of **Sym** is contained in the Catalan Lie algebra \mathfrak{C} .*

It is therefore of interest to investigate the expansion of the various Lie idempotents of the descent algebra on the basis c_T . The Dynkin elements are obviously the left and right comb trees. The expansion of the Solomon idempotent is obtained in [2]. In the sequel, we shall obtain the following expansion for the Catalan idempotents of [23].

THEOREM 3.7. *For an element $T(\sigma)$ of the PBW basis of $\text{Lie}(n)$, denote by t the binary tree consisting of its internal nodes, and let $r(t)$ and $l(t)$ be respectively the number of right edges and left edges of t . Then,*

(96)
$$D_{a,b}^n = \sum_{T(\sigma)} a^{r(t)} b^{l(t)} T(\sigma),$$

where the sum is over all admissible labelled trees.

In particular, the sum of the PBW basis is a Lie idempotent of the descent algebra.

The proof of this result involves the fine structure of **FQSym**, as a dendriform algebra, and also as a quadri-algebra. As we shall see, it is not obvious to determine whether an element of \mathfrak{C} belongs to the descent algebra. This question motivates the developments of Section 6.

4. THE CATALAN IDEMPOTENTS IN THE PBW BASIS

This section is devoted to the proof of Theorem 3.7.

4.1. A FUNCTIONAL EQUATION. For a complete binary tree T , let $T \nearrow \swarrow$ be the evaluation of T with \mathbf{G}_1 in all the leaves, and the operation $\nearrow \swarrow$ in the internal nodes.

PROPOSITION 4.1. *The operation $\nearrow \swarrow$ is magmatic, i.e. satisfies no relation other than bilinearity, and*

$$(97) \quad T \nearrow \swarrow = \sum_{\sigma \text{ admissible}} T(\sigma) = c_T$$

is the sum of all trees of shape T in the PBW basis.

Proof. This follows from Prop. 7.2 with $B = \nearrow \swarrow$. □

It follows from the previous considerations that the formal sum X of the PBW basis satisfies the functional equation

$$(98) \quad X = \mathbf{G}_1 + B(X, X), \quad \text{where } B(X, Y) = X \nearrow \swarrow Y,$$

and setting $X = \mathbf{G}_1 + X_+$, we can introduce parameters to count left and right internal branches: the sum X of the expressions for the Catalan idempotents proposed in Theorem 3.7 is the unique solution of

$$(99) \quad X_+ = B(\mathbf{G}_1 + aX_+, \mathbf{G}_1 + bX_+).$$

From now on, we shall write σ for \mathbf{G}_σ , and set, for $n \geq 2$,

$$(100) \quad X_n = \sum_{\sigma \in \mathfrak{S}_n} c_\sigma \sigma.$$

Let $\sigma \in \mathfrak{S}_n$, with $n \geq 2$. Let us first consider the case where n precedes 1 in σ . Since $X = \mathbf{G}_1 + X \nearrow X - X \swarrow X$, the products contributing to c_σ correspond to products $\tau \swarrow \mu$ containing σ which amounts to considering all ways of writing σ as $u \cdot v$ with $1 \in v$ and $n \in u$. The other case for σ is similar so that c_σ is inductively

$$(101) \quad c_\sigma = \sum_{\substack{\sigma = u \cdot v \\ 1 \in u, n \in v}} a_{|u|} b_{|v|} c_{\text{std}(u)} c_{\text{std}(v)} - \sum_{\substack{\sigma = u \cdot v \\ n \in u, 1 \in v}} b_{|u|} a_{|v|} c_{\text{std}(u)} c_{\text{std}(v)}.$$

Note that only one sum contributes to a nonzero coefficient for c_σ .

LEMMA 4.2. *For $\sigma \in \mathfrak{S}_n$, define $\bar{\sigma} = \sigma \circ (n, \dots, 1) = (\sigma_n, \dots, \sigma_1)$. Then:*

$$c_\sigma = (-1)^{n-1} c_{\bar{\sigma}}.$$

Proof. This follows by induction from Eq. (101). Indeed,

$$(102) \quad \begin{aligned} c_{\bar{\sigma}} &= \sum_{\substack{\bar{\sigma} = u \cdot v \\ 1 \in u, n \in v}} a_{|u|} b_{|v|} c_{\text{std}(u)} c_{\text{std}(v)} - \sum_{\substack{\bar{\sigma} = u \cdot v \\ n \in u, 1 \in v}} b_{|u|} a_{|v|} c_{\text{std}(u)} c_{\text{std}(v)} \\ &= - \sum_{\substack{\sigma = u \cdot v \\ 1 \in u, n \in v}} a_{|u|} b_{|v|} c_{\text{std}(\bar{u})} c_{\text{std}(\bar{v})} + \sum_{\substack{\bar{\sigma} = u \cdot v \\ n \in u, 1 \in v}} b_{|u|} a_{|v|} c_{\text{std}(\bar{u})} c_{\text{std}(\bar{v})}. \end{aligned}$$

Now, $|\sigma|$ is equal to $|u| + |v|$ hence of the same parity so that we recover c_σ with the correct sign. \square

The Narayana polynomials $N_n(a, b) := \mathbf{ca}_{n-1}(a, b)$ satisfy the recurrence

$$(103) \quad N_n = \begin{cases} 1 & \text{if } n \leq 2, \\ (a + b)N_{n-1} + ab \sum_{k=2}^{n-2} N_k N_{n-k} & \text{if } n \geq 3. \end{cases}$$

Let $\sigma \in \mathfrak{S}_n$. We define the maximal runs of σ as the integers $i_0 = 1 < i_1 < \dots < i_p < n = i_{p+1}$ satisfying the conditions

- For any $0 \leq k \leq p$, $\sigma_{\{i_k, \dots, i_{k+1}\}}$ is monotonous.
- For any $0 \leq k \leq p - 1$, $\sigma_{\{i_k, \dots, i_{k+2}\}}$ is not monotonous.

4.2. PROOF OF THEOREM 3.7. We are now in a position to compute c_σ by induction. Thanks to Eq. (101), we know that c_σ is obtained by deconcatenating σ at all positions between 1 and n . The main ingredient of the proof consists in observing that summing over positions inside a given run gives a simple inductive formula.

PROPOSITION 4.3. Let $\sigma \in \mathfrak{S}_n$. Let (i_0, \dots, i_{p+1}) be its sequence of runs and, for $0 \leq k \leq p$, put:

$$m_k = \begin{cases} a & \text{if } \sigma_{\{i_k, \dots, i_{k+1}\}} \text{ is increasing,} \\ b & \text{if } \sigma_{\{i_k, \dots, i_{k+1}\}} \text{ is decreasing.} \end{cases}$$

Let $\text{des}(\sigma)$ denote the number of descents of σ . Then:

$$(104) \quad c_\sigma = (-1)^{\text{des}(\sigma)} \prod_{k=0}^{p-1} m_k \prod_{k=0}^p N_{i_{k+1} - i_k + 1}.$$

Proof. We proceed by induction on n . This is obvious if $n = 1$.

Let us first assume that $\sigma_1^{-1} < \sigma_n^{-1}$. There exists $0 \leq \alpha < \beta \leq k + 1$ such that $\sigma_1^{-1} = i_\alpha$ and $\sigma_n^{-1} = i_\beta$. Hence:

$$\begin{aligned} c_\sigma &= \sum_{k=\alpha}^{\beta-1} \sum_{\ell=0}^{i_{k+1} - i_k - 1} a_{i_1 + \dots + i_k + \ell} b_{n - i_1 - \dots - i_k - \ell} c_{\text{std}(\sigma_1, \dots, \sigma_{i_1 + \dots + i_k + \ell})} c_{\text{std}(\sigma_{i_1 + \dots + i_k + \ell + 1}, \dots, \sigma_n)} \\ &= (-1)^{\text{des}(\sigma)} \sum_{k=\alpha}^{\beta-1} \sum_{\ell=0}^{i_{k+1} - i_k - 1} \frac{M}{m_k} N_{i_1 - i_0 + 1} \cdots N_{i_k - i_{k-1} + 1} N_{i_{k+2} - i_{k+1} + 1} \cdots N_{i_{p+1} - i_p + 1} \varepsilon_k \mu_{k, \ell}, \end{aligned}$$

with

$$(105) \quad M = \prod_{i=0}^{p-1} m_i,$$

$$(106) \quad \varepsilon_k = \begin{cases} +1 & \text{if } \sigma_{\{i_k, \dots, i_{k+1}\}} \text{ is increasing,} \\ -1 & \text{if } \sigma_{\{i_k, \dots, i_{k+1}\}} \text{ is decreasing;} \end{cases}$$

$$(107) \quad \mu_{k, \ell} = \begin{cases} m_k N_{i_{k+1} - i_k + 1} & \text{if } \ell = 0, \\ m_{k-1} N_{i_{k+1} - i_k + 1} & \text{if } \ell = i_{k+1} - i_k - 1, \\ m_{k-1} m_k N_{\ell+1} N_{i_{k+1} - i_k - \ell} & \text{if } 1 \leq \ell \leq i_{k+1} - i_k - 2. \end{cases}$$

As $\{m_{k-1}, m_k\} = \{a, b\}$, summing over ℓ yields

$$(108) \quad \sum_{\ell=0}^{i_{k+1} - i_k - 1} \mu_{k, \ell} = N_{i_{k+1} - i_k + 1},$$

and we obtain

$$(109) \quad c_\sigma = (-1)^{\text{des}(\sigma)} \prod_{k=0}^{p-1} m_k \prod_{k=0}^p N_{i_{k+1}-i_k+1} \sum_{k=\alpha}^{\beta-1} \varepsilon_k.$$

As $i_\alpha = \sigma_1^{-1}$, $\sigma_{|\{i_\alpha, \dots, i_{\alpha+1}\}}$ is increasing, so $\varepsilon_k = (-1)^{k-\alpha}$. As $i_\beta = \sigma_n^{-1}$, $\sigma_{|\{i_{\beta-1}, \dots, i_\beta\}}$ is increasing, so $\beta - \alpha - 1$ is even. Hence,

$$(110) \quad \sum_{k=\alpha}^{\beta-1} \varepsilon_k = \sum_{k=0}^{\beta-\alpha-1} (-1)^k = \frac{1 - (-1)^{\beta-\alpha}}{2} = 1,$$

which finally gives the announced result for c_σ .

Let us now assume that $\sigma_n^{-1} < \sigma_1^{-1}$. Applying the previous result to $\bar{\sigma}$, we obtain:

$$(111) \quad c_\sigma = (-1)^{n-1-\text{des}(\bar{\sigma})} \prod_{k=0}^{p-1} m'_k \prod_{k=0}^p N_{i_{k+1}-i_k+1}.$$

If p is even, then

$$(112) \quad \prod_{k=0}^{p-1} m'_k = \prod_{k=0}^{p-1} m_k = (ab)^{\frac{p}{2}}.$$

If p is odd, then

$$(113) \quad \prod_{k=0}^{p-1} m_k = \begin{cases} (ab)^{\frac{p-1}{2}} a & \text{if } \sigma_{|\{i_0, \dots, i_1\}} \text{ is increasing,} \\ (ab)^{\frac{p-1}{2}} b & \text{if } \sigma_{|\{i_0, \dots, i_1\}} \text{ is decreasing,} \end{cases}$$

$$(114) \quad \prod_{k=0}^{p-1} m'_k = \begin{cases} (ab)^{\frac{p-1}{2}} a & \text{if } \sigma_{|\{i_p, \dots, i_{p+1}\}} \text{ is decreasing,} \\ (ab)^{\frac{p-1}{2}} b & \text{if } \sigma_{|\{i_p, \dots, i_{p+1}\}} \text{ is increasing.} \end{cases}$$

As p is odd, $\sigma_{|\{i_p, \dots, i_{p+1}\}}$ is decreasing if, and only if, $\sigma_{|\{i_0, \dots, i_1\}}$ is increasing. Therefore, in all cases,

$$(115) \quad \prod_{k=0}^{p-1} m'_k = \prod_{k=0}^{p-1} m_k,$$

and finally the result is proved for any σ . □

COROLLARY 4.4. *If the descent sets of σ and τ are equal, then $c_\sigma = c_\tau$.*

Comparing the explicit expressions for the coefficients, this proves Theorem 3.7.

5. A NEW BASIS OF \mathfrak{C}

5.1. FROM BINARY TREES TO PLANE TREES. In this section, it will be convenient to label the basis elements of \mathfrak{C} by plane rooted trees instead of binary trees.

The Butcher product $\mathbb{T}_1 \uparrow \mathbb{T}_2$ of two plane rooted trees $\mathbb{T}_1, \mathbb{T}_2$ is obtained by grafting \mathbb{T}_1 on the root of \mathbb{T}_2 , on the left. Then one defines the usual bijection from plane trees to binary trees as the Knuth rotation K recursively defined by $K(\bullet) = \epsilon$ (the empty binary tree) and for any plane rooted trees \mathbb{T} and \mathbb{T}' ,

$$(116) \quad K(\mathbb{T} \uparrow \mathbb{T}') = K(\mathbb{T}) \wedge K(\mathbb{T}').$$

This correspondence is illustrated on Figures 6 and 7 at the end of the paper, representing the Tamari order for $n = 3, 4$.

We shall make use of some typographical distinctions to help understand which object is of what type: if T is a complete binary tree, we denote by t the binary tree obtained by removing its leaves, and by \mathbb{T} the corresponding plane tree.

The cover relation of the Tamari order on plane trees is described as follows: starting from a tree T and a vertex x that is neither its root or a leaf, the trees $T' > T$ covering T are obtained by cutting off the leftmost subtree of x and grafting it back on the left of the parent of x .

Now, in the \mathbf{P} basis of \mathbf{PBT} (cf. [13] for background and notations), the product $\mathbf{P}_{t_1}\mathbf{P}_{t_2}$ is an interval of the Tamari order, which is described on plane trees as follows. Let us define $T_1 = B(U_1 \cdots U_r)$ where $U_1 \cdots U_r$ are the subtrees of T_1 and let b_1, \dots, b_n be the vertices of the leftmost branch of T_2 , b_1 being its root and b_n its leftmost leaf.

Then, any nondecreasing sequence $0 = i_0 \leq i_1 \leq i_2 \leq \dots \leq i_n = r$ of n indices corresponds to one term of the product $\mathbf{P}_{T_1}\mathbf{P}_{T_2}$: define the plane forest $F_1^{(k)} = (U_{i_{k-1}+1}, \dots, U_{i_k})$ for all $k \in [1, n]$. The corresponding term of the product (in the basis \mathbf{P}) is the tree obtained by grafting all the elements of the forest $F_1^{(i)}$, respecting their order, on the vertex b_i and to the left of b_{i+1} , see Figure 1.

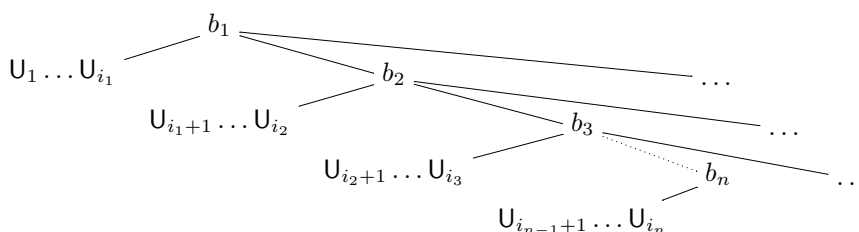


FIGURE 1. A generic element of the product $\mathbf{P}_{T_1}\mathbf{P}_{T_2}$, where $U_1, \dots, U_{i_n} = U_r$ are, in this order, the children of the root of T_1 and the b_i are the vertices of the left branch of T_2 .

5.2. A NEW BASIS OF \mathfrak{C} . Define a new basis x_t of \mathfrak{C}_n (indexed by incomplete binary trees of size $n - 1$) by the condition

$$(117) \quad c_t = \sum_{u \leq t} x_u \quad \Leftrightarrow \quad x_t = \sum_{u \leq t} \mu(u, t) c_u$$

where \leq is the Tamari order, and μ its Moebius function.

For example,

$$(118) \quad x_{\text{tree}} = c_{\text{tree}}$$

$$(119) \quad x_{\text{tree}} = c_{\text{tree}} - c_{\text{tree}}$$

$$(120) \quad x_{\text{tree}} = c_{\text{tree}} - c_{\text{tree}}$$

$$(121) \quad x_{\text{tree}} = c_{\text{tree}} - c_{\text{tree}}$$

$$(122) \quad x_{\text{tree}} = c_{\text{tree}} - c_{\text{tree}} - c_{\text{tree}} + c_{\text{tree}}$$

For a binary tree t of size $n - 1$, denote by $T = K^{-1}(t)$ the plane tree of size n corresponding to t , and set $X_T = x_t$ and $C_T = c_t$. Then on plane trees, the previous equations read as (see Figure 6 for the bijection between incomplete binary trees of

size 3 and plane trees of size 4)

$$\begin{aligned}
 (123) \quad X &= C \\
 (124) \quad X &= C - C \\
 (125) \quad X &= C - C \\
 (126) \quad X &= C - C \\
 (127) \quad X &= C - C - C + C
 \end{aligned}$$

THEOREM 5.1. The preLie product in the X basis is given by

$$(128) \quad X_{T_1} \triangleright X_{T_2} = \sum_{T \in G(T_1, T_2)} X_T$$

where $G(T_1, T_2)$ is the multiset of trees obtained by grafting T_1 on all nodes of T_2 in all possible ways.

For example,

$$\begin{aligned}
 (129) \quad X \triangleright X &= X + X + X \\
 (130) \quad X \triangleright X &= X + X + X + X + X
 \end{aligned}$$

To prove this theorem, we shall show that Equations (128) and (89) are equivalent. To this aim, we define a new product \blacktriangleright by the condition that it satisfies (128) and then show that it also satisfies (89). This is done in the following section.

5.3. A PRELIE STRUCTURE ON **PBT**. Relation (117) between the bases x and c is the same as the one between the natural basis \mathbf{P}_t of **PBT** and the multiplicative basis \mathbf{H}_t defined in [13, Eq. (46)]. We can therefore define a linear map

$$\begin{aligned}
 (131) \quad \psi : \mathbf{PBT}_{n-1} &\longrightarrow \text{Lie}(n) \\
 \mathbf{P}_t &\longmapsto x_t \\
 \mathbf{H}_t &\longmapsto c_t,
 \end{aligned}$$

and define a preLie product on **PBT** by requiring that

$$(132) \quad \psi(\mathbf{H}_{t_1} \triangleright \mathbf{H}_{t_2}) = c_{t_1} \triangleright c_{t_2}.$$

Theorem 5.1 can then be derived from a compatibility property of the usual product of **PBT** with this preLie product. Indexing as above the bases of **PBT** by plane trees instead of binary trees, we define a new product \blacktriangleright by

$$(133) \quad \mathbf{P}_T \blacktriangleright \mathbf{P}_{T'} = \sum_{T'' \in G(T, T')} \mathbf{P}_{T''}.$$

NOTE 5.2. We shall use the facts that $\mathbf{P}_T \blacktriangleright \mathbf{P}_\bullet = \mathbf{P}_{B(T)}$ (by definition since B amounts to add a root on a sequence of trees) and that $\mathbf{H}_T \blacktriangleright \mathbf{H}_\bullet = \mathbf{H}_{B(T)}$ since the trees smaller than $B(T)$ in the Tamari order are the $B(T')$ with $T' \leq T$.

PROPOSITION 5.3. For $U, V, W \in \mathbf{PBT}$,

$$(134) \quad U \blacktriangleright (VW) = (U \blacktriangleright V)W + V(U \blacktriangleright W) - V(U \blacktriangleright \bullet)W,$$

where $\bullet = \mathbf{P}_\bullet = \mathbf{H}_\bullet$.

The idea behind this formula is fairly simple: the left hand-side consists in gluing U either on W or on V . The first two terms amount to doing essentially that, except that an extra term appears that should not be there, where V is glued on U , hence the corrective term with a minus sign.

Proof. Since (134) is linear in U, V, W , we can assume that

$$(135) \quad U = \mathbf{P}_{T_0}, \quad V = \mathbf{P}_{T_1}, \quad W = \mathbf{P}_{T_2}.$$

In that case, let b_1, \dots, b_n be as above the vertices the leftmost branch of T_2 . Then the first product $\mathbf{P}_{T_0} \blacktriangleright (\mathbf{P}_{T_1} \mathbf{P}_{T_2})$ is obtained by summing over three disjoint sets of grafting patterns that we will denote by (a), (b) and (c).

Set (a) consists of the grafting patterns where T_0 is grafted on a b_i and to the left of b_{i+1} . Set (b) consists of those where T_0 is grafted on a vertex belonging to T_1 . Set (c) consists of all other elements. We also denote by (d) the set of trees obtained in the product $\mathbf{P}_{T_1}(\mathbf{P}_{T_0} \blacktriangleright \bullet)\mathbf{P}_{T_2}$.

Let us now consider the other two terms. The term $(\mathbf{P}_{T_0} \blacktriangleright \mathbf{P}_{T_1})\mathbf{P}_{T_2}$ also gives rise to two cases: either T_0 is grafted on the root of T_1 (set (a')), or not (set (b')). Finally, the term $\mathbf{P}_{T_1}(\mathbf{P}_{T_0} \blacktriangleright \mathbf{P}_{T_2})$ splits into two cases: either T_0 is grafted on a b_i to the left of b_{i+1} (set (d')) or T_0 is grafted somewhere else (set (c')).

Let us show that the sets labelled with the same letters coincide.

First, consider an element T of (a'). From the product formula on plane trees, T is a tree obtained by grafting T_0 on the root of T_1 , and then grafting all the children of the root of this new tree, respecting their order, on the b_i s of T_2 , to the left of b_{i+1} . Whether we put the children of T_1 on the left of the b_i s and then T_0 on the same spots or put them all together is irrelevant, so that sets (a) and (a') coincide.

Let us now prove that (b) = (b'). Let T be in (b'). It has been obtained by grafting T_0 somewhere on a subtree U_i of T_1 and then by grafting all subtrees of T_1 on the b_i s as usual. This is the same as first grafting all subtrees of T_1 on the b_i s and then grafting T_0 on a subtree that was initially a subtree of T_1 . So (b) = (b').

The equality (c) = (c') is easy: the operations of grafting the U_i s on the b_i s to the left of b_{i+1} and then grafting T_0 not on that spots clearly commute, which means precisely that (c) = (c').

Finally, let us prove that (d) = (d'). Observe that the product $\mathbf{P}_{T_0} \blacktriangleright \mathbf{P}_{T_2}$ decomposes into two cases, depending on whether T_0 is grafted on a b_i left of b_{i+1} or not. If it is the case (from which case (d') is derived by multiplying by \mathbf{P}_{T_1} on the left), it corresponds to a term of the product $(\mathbf{P}_{T_0} \blacktriangleright \bullet)\mathbf{P}_{T_2}$ (from which case (d) is derived) also by multiplying by \mathbf{P}_{T_1} on the left. Hence (d) = (d').

This proves that

$$(136) \quad \mathbf{P}_{T_0} \blacktriangleright (\mathbf{P}_{T_1} \mathbf{P}_{T_2}) + \mathbf{P}_{T_1} (\mathbf{P}_{T_0} \blacktriangleright \bullet) \mathbf{P}_{T_2} = (\mathbf{P}_{T_0} \blacktriangleright \mathbf{P}_{T_1}) \mathbf{P}_{T_2} + \mathbf{P}_{T_2} (\mathbf{P}_{T_0} \blacktriangleright \mathbf{P}_{T_2})$$

which is equivalent to (134). □

LEMMA 5.4. *Let T be a tree of the form $T = B(T')$ and T_0 be an arbitrary tree. Then,*

$$(137) \quad \mathbf{H}_{T_0} \blacktriangleright \mathbf{H}_T = \mathbf{H}_{B(T_0 T')} + \mathbf{H}_{B(T' T_0)} - \mathbf{H}_{B(T_0 \uparrow T')} - \mathbf{H}_{B(T' \uparrow T_0)} + \mathbf{H}_{B(T_0 \blacktriangleright T')}.$$

Proof. In the \mathbf{P} basis,

$$(138) \quad \mathbf{H}_{T_0} \blacktriangleright \mathbf{H}_T = \left(\sum_{T'_0 \leq T_0} \mathbf{P}_{T'_0} \right) \blacktriangleright \left(\sum_{T'' \leq T'} \mathbf{P}_{B(T'')} \right).$$

Expanding the product, we find three kinds of terms: either T'_0 is grafted on the root of $B(T'')$ to the left or to the right of T'' (two cases) or it is grafted below it. This last case sums up to $\mathbf{H}_{B(T_0 \blacktriangleright T')}$. The other two cases are identical up to exchanging the roles of T_0 and T' . Let us deal with the first case. All trees appearing by grafting a $T'_0 \leq T_0$ to the left of the root of a $B(T'')$ with $T'' \leq T'$, are all smaller in the Tamari order than $B(T_0 T')$. Conversely, all trees smaller than $B(T_0 T')$ belong to the first case, except those where T'_0 was grafted below the root. But these trees are precisely the elements smaller in the Tamari order than $B(T_0 \uparrow T')$. So the terms belonging to the first case sum up to $\mathbf{H}_{B(T_0 T')} - \mathbf{H}_{B(T_0 \uparrow T')}$.

Analogously, the terms of the second case sum up to $\mathbf{H}_{B(T' T_0)} - \mathbf{H}_{B(T' \uparrow T_0)}$. Adding the contributions of the three cases, we obtain Formula (137). □

Proof of Theorem 5.1. We are now in a position to prove Theorem 5.1 by showing that \blacktriangleright and \triangleright coincide.

We want to compute a generic product $\mathbf{H}_{T_1} \blacktriangleright \mathbf{H}_T$. Since the \mathbf{H} basis is multiplicative, we can replace \mathbf{H}_T by $\mathbf{H}_{T_2} \mathbf{H}_{T_3}$ where T_2 has a single tree attached: $T_2 = B(T'_2)$.

Now let us apply Formula (134) on the \mathbf{H} basis:

$$(139) \quad \mathbf{H}_{T_1} \blacktriangleright (\mathbf{H}_{T_2} \mathbf{H}_{T_3}) = (\mathbf{H}_{T_1} \blacktriangleright \mathbf{H}_{T_2}) \mathbf{H}_{T_3} + \mathbf{H}_{T_2} (\mathbf{H}_{T_1} \blacktriangleright \mathbf{H}_{T_3}) - \mathbf{H}_{T_2} (\mathbf{H}_{T_1} \blacktriangleright \bullet) \mathbf{H}_{T_3}.$$

By (137), the first term can be rewritten as

$$(140) \quad (\mathbf{H}_{B(T_1 T'_2)} + \mathbf{H}_{B(T'_2 T_1)} - \mathbf{H}_{B(T_1 \uparrow T'_2)} - \mathbf{H}_{B(T'_2 \uparrow T_1)} + \mathbf{H}_{B(T_1 \blacktriangleright T'_2)}) \mathbf{H}_{T_3},$$

and now its second term $\mathbf{H}_{B(T'_2 T_1)} \mathbf{H}_{T_3}$ cancels with $\mathbf{H}_{T_2} (\mathbf{H}_{T_1} \blacktriangleright \bullet) \mathbf{H}_{T_3}$, so that

$$(141) \quad \begin{aligned} \mathbf{H}_{T_1} \blacktriangleright (\mathbf{H}_{T_2} \mathbf{H}_{T_3}) &= \mathbf{H}_{B(T_1 T'_2)} \mathbf{H}_{T_3} + \mathbf{H}_{B(T_1 \blacktriangleright T'_2)} \mathbf{H}_{T_3} + \mathbf{H}_{T_2} (\mathbf{H}_{T_1} \blacktriangleright \mathbf{H}_{T_3}) \\ &\quad - \mathbf{H}_{B(T_1 \uparrow T'_2)} \mathbf{H}_{T_3} - \mathbf{H}_{B(T'_2 \uparrow T_1)} \mathbf{H}_{T_3}, \end{aligned}$$

which is exactly (89) expressed on plane trees. This proves that $\blacktriangleright = \triangleright$ by induction on the sizes of the trees. □

NOTE 5.5. The preLie product in the X -basis coincides with that of the free brace algebra on one generator, which is the linear span of all plane trees, endowed with the brace product

$$(142) \quad \langle T_1 T_2 \cdots T_k, T_{k+1} \rangle = \sum_{T \in G(T_1, T_2, \dots, T_k; T_{k+1})} T,$$

where $G(T_1, T_2, \dots, T_k; T_{k+1})$ is the multiset of trees obtained by grafting T_1, \dots, T_k in this order on all nodes of T_{k+1} in all possible ways [11].

This brace product does not coincide with the one induced by the dendriform structure of \mathbf{FQSym} . Indeed, one can check that \mathfrak{C} is not stable for this one: $\langle \langle \bullet, \bullet \rangle, \bullet \rangle$

is not in Lie. However, both products induce the same preLie structure, as well as the extended preLie products

$$(143) \quad \{\mathbb{T}_1 \mathbb{T}_2 \cdots \mathbb{T}_r; \mathbb{T}\} = \sum_{\sigma \in \mathfrak{S}_r} \langle \mathbb{T}_{\sigma(1)} \mathbb{T}_{\sigma(2)} \cdots \mathbb{T}_{\sigma(r)}; \mathbb{T} \rangle$$

defined by their symmetrized versions (the Oudom–Guin construction [27]).

5.4. THE CATALAN IDEMPOTENT IN THE X -BASIS. Let PL be the free preLie algebra on one generator, and let p_τ be its Chapoton–Livernet basis [7], indexed by rooted trees. Recall that the preLie product in this basis is given by

$$(144) \quad p_{\tau_1} \triangleright p_{\tau_2} = \sum_{\tau \in g(\tau_1, \tau_2)} p_\tau,$$

where $g(\tau_1, \tau_2)$ is the multiset of trees obtained by grafting the root of τ_1 on all nodes of τ_2 .

By definition, \mathfrak{C} contains the free preLie algebra generated by $X_\bullet = \mathbf{G}_1$. As an easy consequence of Theorem 5.1, we can express its Chapoton–Livernet basis in terms of the X -basis of \mathfrak{C} .

LEMMA 5.6. *For a rooted plane tree \mathbb{T} , denote by $\bar{\mathbb{T}}$ the underlying non-plane rooted tree. The map*

$$(145) \quad \iota : p_\tau \longmapsto |\text{Aut}(\tau)| \sum_{\bar{\mathbb{T}}=\tau} X_{\bar{\mathbb{T}}}$$

is an embedding of preLie algebras.

In terms of the symmetrized brace product (143),

$$(146) \quad p_{B(\tau_1 \cdots \tau_k)} = \{p_{\tau_1} \cdots p_{\tau_k}, X_\bullet\}.$$

We shall therefore identify PL with its image under ι .

As already mentioned, the primitive Lie algebra of **Sym**, being generated by the Ψ_n , is a Lie subalgebra of PL. In [6], Chapoton gives an expression of the Catalan idempotent on the basis p_τ : setting the unnecessary parameter a to 1, the coefficient of p_τ in $D_{1,b}^n$ is equal to the generating polynomial of *small closed flows* on τ by *size*.

A small closed flow of size k on τ is equivalent to the following data:

- a set O of k distinct vertices (outputs), which cannot be the root;
- a multiset I of k vertices (inputs);
- a perfect matching between I and O such that each pair (i, o) determines a path from i to o directed towards the root.

Denoting by $\mathbf{F}'(\tau, k)$ the set of small flows of size k on τ and by $d_\tau(b)$ the polynomial

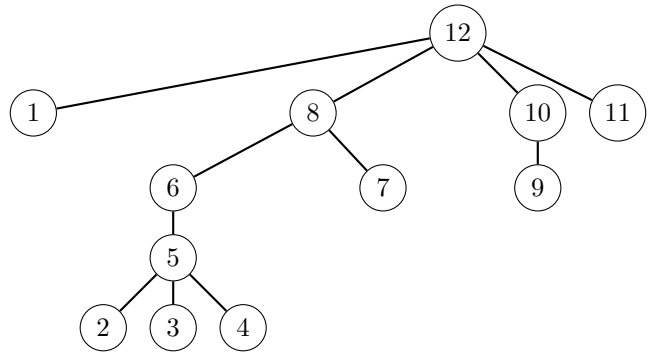
$$(147) \quad d_\tau(b) = \sum_k \sum_{\phi \in \mathbf{F}'(\tau, k)} b^k,$$

we have [6]

$$(148) \quad D_{1,b}^n = \sum_{|\tau|=n} d_\tau(b) p_\tau.$$

The *rate* of a vertex is the sum of the multiplicities of the inputs minus the number of outputs in its subtree.

Define the canonical labelling $\text{can}(\mathbb{T})$ of the vertices of a plane tree \mathbb{T} as the one which yields the identity permutation when the tree is traversed in postfix order: the label of a node is the cardinality of its subtree plus the number of nodes strictly to its left. For example,

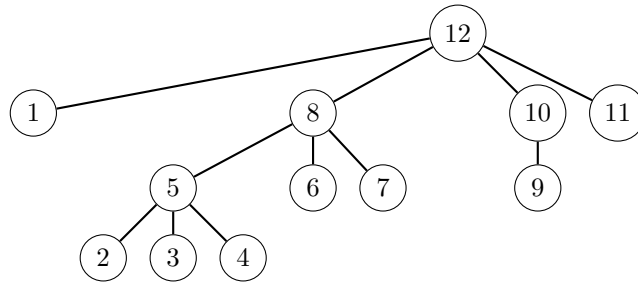


Extending the notion of flow to plane trees, we define the vector of a flow ϕ on T as the tuple $V_\phi = (\phi(v_1), \dots, \phi(v_n))$, where v_i are the vertices of T numbered by their canonical labelling. The complete example of all small closed flows on plane trees on 4 vertices can be found on Figure 8.

PROPOSITION 5.7. *If $T' \geq T$ in the Tamari order, the vector of any small closed flow on T' is also the vector of a small closed flow on T .*

For example, on Figure 8, one can check that there are five different flow vectors on plane trees of size 4 and that they all are flow vectors of the minimal plane tree, the chain.

Proof. Let v be the vector of a small closed flow on T' . If $T' \geq T$ are canonically labelled, and if T' covers T , the description of the Tamari order on plane trees shows that the subtree of a node labelled x in T' , contains all the vertices of the subtree of the node labelled x in T . Indeed, T' is obtained from T by cutting the subtree U of a vertex u , and grafting it again on the left of its parent. The labels of the vertices of U , being equal to the cardinality of their subtree plus the number of vertices to their left, remain unchanged by this operation, and the other labels remain the same as well. For example, the following tree covers the above one. The subtree of 5 has been cut and grafted back on 8 without changing its labelling.



Thus, the rate of x in T is at least the rate of x in T' , so that it is in particular non-negative. Hence, v is the vector of a flow on T , which is obviously small and closed. \square

Conversely, if a perfect matching between I and O never go through a left edge of a tree T , then the same values on the same elements give rise to another perfect matching on the tree T' (a covering tree of T) obtained by cutting this edge and gluing it on the node immediately above.

Define the maximal small closed flow $\phi_0(T)$ by choosing as outputs all the internal vertices except the root, and the leftmost leaf of each subtree as inputs with the

maximal allowed value. This is a flow of maximal size, whose vector is lexicographically maximal among all flows of T .

For example, the maximal flow on the first tree above is

$$(149) \quad [0, 3, 0, 0, -1, -1, 0, -1, 1, -1, 0, 0],$$

and the maximal flow of the second one is

$$(150) \quad [0, 2, 0, 0, -1, 0, 0, -1, 1, -1, 0, 0].$$

One easily checks that both flows are flows of the first tree but not of the second one (the leaf 6 has a -1 as value).

Let $k_0(T)$ be the size of $\phi_0(T)$.

PROPOSITION 5.8. *Given a plane tree T , there is a bijection between all small closed flow on T and the set of maximal flows on all trees $T' \geq T$. The flow vector is preserved in the bijection.*

Proof. We have already seen that each flow vector of $T' > T$ is a flow vector of T . So in particular, each maximal flow vector of a $T' > T$ is a flow vector of T .

We need to prove that all maximal flow vectors are distinct and that each flow vector of T is indeed a maximal flow vector of a $T' > T$.

First, the maximal flow vectors are distinct since one can rebuild a tree given its maximal flow vector v : the nodes i such that $v_i = -1$ are the internal nodes, those with $v_i > 0$ are the leftmost leaves of a non-root node, and the nodes with $v_i = 0$ are the other leaves and the root. So, reading the flow vector from right to left, there cannot be any ambiguity on where node i should go, and since nodes are added from right to left, no two flows can give rise to the same tree.

Finally, let ϕ be a flow on a tree T . If it is its maximal flow then the statement is a tautology, so we assume that it is not maximal. We will prove that ϕ is a flow on a tree $T' > T$ and conclude by induction on the maximal distance from T to the maximal element of the Tamari lattice, the corolla. Consider the maximal flow ϕ' of T . There is a node that is nonzero in ϕ' and that does not have the same value in ϕ . If it is an internal node, it has 0 value in ϕ and since it is the leftmost child of its parent, we can cut its connection to its father and graft it on its grand-parent, then obtaining a flow on a tree $T' > T$. If there is no such internal node differing between ϕ and ϕ' , there is a leaf that does not have its maximal possible value in ϕ . Then, since all the internal nodes above it except the root have value -1 , let n be the bottom-most such node that has another flow coming to it from below. Then cut the left child of n and graft it on the parent of n . In the resulting tree T' , the flow is still valid. Indeed, by definition of a flow, if all matchings between inputs and outputs correspond to a path towards to the root, then it is valid. In our case, n can be matched to this *other source*, so that all elements matching to the leaf are either above or below n and hence remain valid matchings in T' .

So we have proven that any flow of a tree T is either its maximal flow or a flow of a tree $T' > T$. Iterating this process with T' , it will have to stop since there are only a finite number of trees greater than a given tree. \square

Note that $k_0(T)$ is the number of non-root internal vertices so that $k_0(T) = i(T) - 1$, where $i(T)$ is the number of internal vertices. It is also the number of left edges of the corresponding binary tree. Thus, the polynomials $d_T(b)$ enumerate Tamari intervals according to these statistics.

PROPOSITION 5.9.

$$(151) \quad d_T(b) := \sum_k \sum_{\phi \in \mathbf{F}'(T,k)} b^k = \sum_{T' \geq T} b^{k_0(T')}.$$

As a consequence, we recover Theorem 3.7.

COROLLARY 5.10.

$$(152) \quad D_{1,b}^n = \sum_{|t|=n-1} b^{l(t)} b_t.$$

Thus, we have now two quite different explanations of the curious fact that the Catalan idempotent is the weighted sum of the PBW basis. The first one relies upon the functional equation (99), which involves the quadri-algebra structure of **FQSym**, and the second one, which is derived from the preLie expansion given by Chapoton in terms of small closed flows. It would be interesting to investigate in the same manner the PBW expansions of the new Lie idempotents introduced in [6, Prop. 6.8 and Conj. 6.10].

6. LIE AND PRE-LIE SUBALGEBRAS OF \mathfrak{C}

6.1. LEFT EQUIVALENCE. As we have seen on the case of the Catalan idempotent, it is not easy to decide whether an element of $\text{Lie}(n)$ expressed on the PBW basis belongs to the descent algebra **Sym**_n. A similar question arises with **PBT**. For example, an old conjecture by Écalle, as reformulated by Chapoton, states that $\text{PBT} \cap \text{Lie}$ is the free preLie algebra generated by $\mathbf{G}_1 = \mathbf{P}_\bullet$.

The equivalence classes to be defined below arose from the following question: in a generic linear combination $g = \sum_t \alpha_t c_t$ of the basis c_t of \mathfrak{C} , which trees must have the same coefficient if one requires that $g \in \text{PBT}$ or $g \in \text{Sym}$?

It turns out that the sums over equivalence classes span Lie subalgebras of Lie, which are conjectured to contain respectively $\text{PBT} \cap \text{Lie}$ (for the L-classes) and proved to contain $\text{Sym} \cap \text{Lie}$ (for the LR-classes). In both cases, the bracket of these subalgebras admits a remarkable combinatorial description.

6.1.1. *An exchange rule.* Consider the following transformation on a binary tree T : choose a vertex v that can either be the root of T or the right child of its parent. Let T_1 and T_2 be the subtrees depicted below

$$(153) \quad \begin{array}{c} v \\ / \quad \backslash \\ T_1 \quad \bullet \\ \backslash \\ T_2 \end{array}$$

and define $T' = L_v(T)$ as the result of exchanging T_1 and T_2 :

$$(154) \quad \begin{array}{c} v \\ / \quad \backslash \\ T_1 \quad \bullet \\ \backslash \\ T_2 \end{array} \mapsto \begin{array}{c} v \\ / \quad \backslash \\ T_2 \quad \bullet \\ \backslash \\ T_1 \end{array}$$

Define an equivalence relation by $T \equiv_L T'$ iff $T' = L_v(T)$ for some vertex v . The equivalence classes will be called L-classes.

Note that the numbers of left and right branches are constant in each equivalence class.

For example, here are the L-classes of (incomplete) trees of sizes 3 and 4. In size 3, there are four L-classes for five binary trees so that only one class contains two trees:

$$(155) \quad \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array} \text{ and } \begin{array}{c} \bullet \\ \bullet \\ \bullet \end{array}$$

In size 4, there are ten classes. Here are the three non-trivial ones.

$$(156) \quad \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array}, \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array}, \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array} \right\}, \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array}, \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array} \right\}, \left\{ \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array}, \begin{array}{c} \bullet \\ \bullet \\ \bullet \\ \bullet \end{array} \right\}.$$

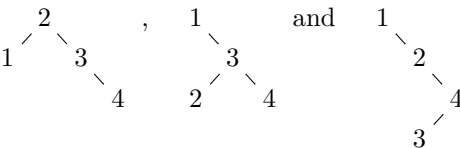
6.1.2. *Combinatorial encoding of the L-classes.* The L-classes can be parametrized by certain bicolored plane trees. To see this, let C be an L-class and let $T \in C$.

Number its vertices in infix order (recursively label by consecutive integers the left subtree, then the root, then the right subtree) so that flattening the tree results in the identity permutation, and build two set partitions $L(T)$ and $R(T)$ where $L(T)$ (resp. $R(T)$) consists of the sets of labels of nodes along left (resp. right) branches. One may observe that both partitions are noncrossing, and are related by the operation of Kreweras complement.

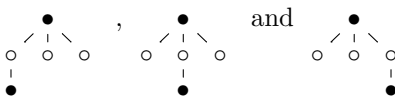
The *left* subsets will correspond to white nodes and the right ones to black nodes.

Now build a bipartite plane tree, whose root is representing the sequence of $R(T)$ containing the root of T and such that all nodes below a given node are the subsets of the other color (except the one encoding its own parent) that have an intersection with the given node in increasing order of the intersection.

For example, let us consider the trees of the L-class (156). Their labelings are

(157) 

For the first tree, the white nodes will be the sets $\{1, 2\}$, $\{3\}$ and $\{4\}$, and the black nodes the sets $\{1\}$ and $\{2, 3, 4\}$. These five sets are the vertices of an undirected bipartite graph, where there is an edge between a white node and a black node whenever the intersection of the underlying sets is nonempty. This graph is necessarily a tree, because a cycle would imply a nonempty intersection between blocks of a partition. The corresponding bipartite trees are thus

(158) 

As a more substantial example, consider the labeled tree

(159) 

for which $L(T) = \{125, 34, 6, 7\}$ and $R(T) = \{567, 24, 1, 3\}$ so that the corresponding bipartite tree is

(160) 

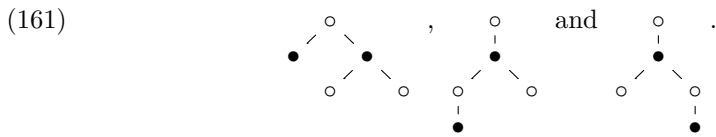
Note that in that case, the resulting bipartite tree is not the image of the initial binary tree by the classical bijection between plane trees and binary trees (Knuth rotation). However, this operation is a bijection since the number of children of the (black) root gives the length of the right branch starting from the root and each subtree corresponds itself recursively to a binary tree.

Actually, this correspondence coincides with the twisted Knuth rotation defined in [2].

THEOREM 6.1. *The L-classes are indexed by bipartite trees with black roots up to reordering of the children of the black nodes.*

Proof. The exchange rule between two binary trees along a right edge amounts to the exchange of the corresponding subtrees of the black node corresponding to the right branch containing this right edge. \square

We shall need at some point a similar encoding, now with a white root. The process is the same except that the root is the part of $L(T)$ that contains the label of the root of T . These trees can be obtained from the black-rooted trees by moving above the root its first (white-rooted) subtree. Again with the same three bipartite trees, we get



Note that in the case of the white-rooted bipartite trees we find in general several classes of bipartite trees up to reordering of the children of the black nodes.

6.1.3. The Écalle-Chapoton conjecture. The L-classes arose while investigating the following question: given a generic linear combination C of the $T(\sigma)$, regroup the $T(\sigma)$ that always have same coefficient if one requires that C belong to **PBT**. They conjecturally are the L-classes.

CONJECTURE 6.2. *Any Lie element of **PBT** is a linear combination of L-classes.*

The number of L-classes in size n is greater than the size of $\mathbf{PBT}_n \cap \text{Lie}(n)$ so that the conjecture provides only a necessary condition.

Here are some tables to clarify this point. Define $a_{n,k}$ as the dimension of $\text{Lie}(n) \cap \mathbf{PBT}_{n,k}$, where $\mathbf{PBT}_{n,k}$ is the subspace of \mathbf{PBT}_n spanned by the binary trees with k right branches.

The $a_{n,k}$ are conjectured to be Sequence A055277, which is the number of rooted trees with n nodes (hence the dimension of the free preLie algebra on one generator as conjectured by Écalle and Chapoton) refined by their number of leaves (parameter k).

(162)

| $n \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------|---|----|----|----|----|---|---|
| 2 | 1 | | | | | | |
| 3 | 1 | 1 | | | | | |
| 4 | 1 | 2 | 1 | | | | |
| 5 | 1 | 4 | 3 | 1 | | | |
| 6 | 1 | 6 | 8 | 4 | 1 | | |
| 7 | 1 | 9 | 18 | 14 | 5 | 1 | |
| 8 | 1 | 12 | 35 | 39 | 21 | 6 | 1 |

A055277

FIGURE 2. The $a_{n,k}$.

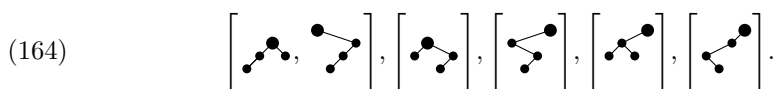
Let now $b_{n,k}$ be the number of L-classes with $n - 1$ nodes and k right branches.

(163)

| $n \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------|---|----|----|----|----|---|---|
| 2 | 1 | | | | | | |
| 3 | 1 | 1 | | | | | |
| 4 | 1 | 2 | 1 | | | | |
| 5 | 1 | 5 | 3 | 1 | | | |
| 6 | 1 | 8 | 11 | 4 | 1 | | |
| 7 | 1 | 13 | 30 | 20 | 5 | 1 | |
| 8 | 1 | 18 | 67 | 73 | 31 | 6 | 1 |

FIGURE 3. The $b_{n,k}$.

The first discrepancy between $a_{n,k}$ and $b_{n,k}$ occurs at $n = 5$. Indeed, all five L-classes do not belong to **PBT**. The five L-classes with one right branch are (represented as incomplete binary trees)



Let us denote these L-classes by C_1, \dots, C_5 . Expanding these as combinations of permutations in **FQSym**₅, one finds that only C_5 belongs to **PBT**₅. The linear span of the other ones has a 3-dimensional intersection with **PBT**₅. A linear basis of this intersection is given, e.g. by

(165) $C_1 + C_4, \quad C_2 - C_4, \quad C_3 + C_4.$

Even if many properties of L-classes are conjectural, the following result is already of interest.

THEOREM 6.3. *Let $\mathcal{L}(n)$ be the linear subspace of $\text{Lie}(n)$ (and therefore of the Catalan subalgebra \mathfrak{C}) spanned by the L-classes. Then $\mathcal{L}(n)$ is a sub pre-Lie algebra of \mathfrak{C} .*

The pre-Lie product $s_1 \triangleright s_2$ of two L-classes can be computed as follows: let B_1 and B_2 be their respective representatives as black-rooted bipartite trees and let B'_1 be the non-equivalent representatives as white-rooted bipartite trees of s_1 . Then the product is obtained as the following sum:

- with a minus sign, all trees obtained by gluing B_1 as a child of a white node of B_2 ;
- with a plus sign, all trees obtained by gluing any element of B'_1 as a child of a black node B_2 .

The multiplicity of a given tree is equal to the number of ways to cut it into two parts such that s_2 is the part containing the root of the tree and s_1 is the other part.

Proof. Recall that the preLie product of two naked trees $T' \triangleright T''$ is obtained by gluing in all possible ways T' on the middle of all branches of T'' (adding an invisible root to T'' so that it is its right subtree), the sign of the result depending on the branch being left (-1) or right ($+1$), see (89).


Consider two L-classes C' and C'' and let T_1 be a tree occurring in $C' \triangleright C''$. Let now T_2 be a tree obtained by a right branch exchange from T_1 (hence in the same L-class as T_1). We want to show that the coefficient of T_2 in the preLie product $C' \triangleright C''$ is the same as the coefficient of T_1 in the same product.

To do this, let us encode all the ways of obtaining T_1 in $T'_1 \triangleright T''_1$ with $T'_1 \in C'$ and $T''_1 \in C''$ as a set of triples (T'_1, T''_1, a) where $a \in T''_1$ is the branch used to insert T'_1 inside T''_1 .

Now, follow the branch a through the right branch exchange sending T_1 to T_2 and contract it (that is, reverse the gluing process on it). We then get two trees, T'_2 and T''_2 . Now, several cases arise. If a was not itself the right branch on which the exchange occurred, it is obvious that: either $T'_2 = T'_1$ and T''_2 is obtained from T''_1 by an exchange along a right branch, or $T''_2 = T''_1$ and T'_2 is obtained from T'_1 by an exchange along a right branch. Moreover, if a was the exchange branch, then the parent branch of a was itself a right branch (condition to be allowed to exchange subtrees) and T_2 is then obtained by gluing T'_1 onto the father branch of a in T''_1 .

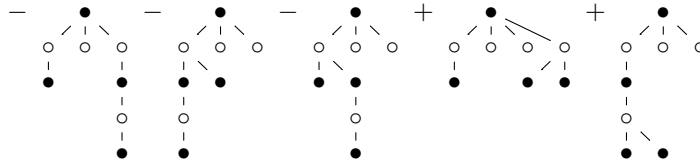
So each triple of T_1 is sent to a different triple of T_2 . Since their roles can be reversed, this injection is in fact a bijection, which proves the required result. \square

Let us now describe the pre-Lie product of two L-classes on their natural combinatorial encodings.

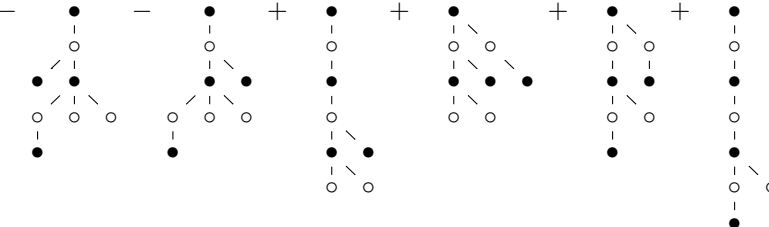
For example, let s_1 be the L-class of  and s_2 be the L-class (156). Then B_1 and B_2 are respectively

(166)  and 

and the product $s_1 \triangleright s_2$ is

(167) 

whereas the product $s_2 \triangleright s_1$ is

(168) 

6.2. LR-CLASSES OF BINARY TREES.

6.2.1. *R-equivalence.* Exchanging left and right, we have a symmetrical notion of *R-equivalence*. For a binary tree T , choose a vertex v that can either be the root of T

or the left child of its parent. Then let T_1 and T_2 be as depicted below

$$(169) \quad \begin{array}{c} v \\ \swarrow \quad \searrow \\ \bullet \quad T_1 \\ \quad \quad \quad \searrow \\ \quad \quad \quad T_2 \end{array}$$

and define $T' = R_v(T)$ as the result of exchanging T_1 and T_2 :

$$(170) \quad \begin{array}{c} v \\ \swarrow \quad \searrow \\ \bullet \quad T_1 \\ \quad \quad \quad \searrow \\ \quad \quad \quad T_2 \end{array} \mapsto \begin{array}{c} v \\ \swarrow \quad \searrow \\ \bullet \quad T_2 \\ \quad \quad \quad \searrow \\ \quad \quad \quad T_1 \end{array}.$$

The union of this relation and its twin relation (154) defines the LR-equivalence: $T' \equiv_{LR} T$ iff $T' = L_v(T)$ or $T' = R_v(T)$ for some vertex v .

Again, the numbers of left and right branches are constant in equivalence classes.

For example, here are all LR-classes of (incomplete) trees of sizes 3 and 4. In size 3, there are three LR-classes for five binary trees and one class contains three trees:

$$(171) \quad \left\{ \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array} \right\}$$

On size 4, there are six LR-classes. Here are the four non-trivial ones.

$$(172) \quad \left\{ \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array} \right\}, \left\{ \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array} \right\}, \left\{ \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array} \right\}, \left\{ \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array}, \begin{array}{c} \bullet \\ \swarrow \quad \searrow \\ \bullet \quad \bullet \\ \quad \quad \quad \bullet \end{array} \right\}.$$

6.2.2. *Combinatorial encoding of the LR-classes.* Consider an LR-class C and a tree T belonging to it. Number its nodes as above in infix order and build two set partitions $L(T)$ and $R(T)$ where $L(T)$ (resp. $R(T)$) consists of the subsets of numbers of nodes along left (resp. right) branches. The left subsets will correspond to white nodes and the right ones to black nodes.

Now build a bipartite free tree (i.e. neither rooted nor ordered) whose nodes represent the blocks of $L(T)$ or $R(T)$ (with the appropriate colors) such that two nodes are connected iff their intersection is nonempty.

For example, consider the trees of the LR-class (172). The first three give rise to the same tree (see Equations (156) and (158)) and the last one is labelled as

$$(173) \quad \begin{array}{c} 4 \\ \swarrow \\ 1 \\ \searrow \\ 2 \\ \quad \quad \quad \searrow \\ \quad \quad \quad 3 \end{array}$$

so that its (plane bipartite) tree is

$$(174) \quad \begin{array}{c} \bullet \\ \vdots \\ \circ \\ \vdots \\ \bullet \\ \vdots \\ \circ \end{array} \begin{array}{c} \\ \\ \\ \\ \searrow \\ \circ \end{array}$$

which is again topologically equivalent to the three previous trees.

THEOREM 6.4. *The LR-classes are parametrized by bipartite free trees.*

Proof. LR-classes are built by successively applying either an exchange of subtrees along a left or a right branch. None of these change the tree associated with their plane trees (defined as the encoding of their associated L-classes). Indeed, a right exchange corresponds to an exchange of children of a black node whereas a left exchange

corresponds to an exchange of children of a white node. The statement then follows since this is an equivalence. \square

6.3. LR-CLASSES AND THE DESCENT ALGEBRA. The LR-classes arose while investigating the following question: given a generic linear combination g of the c_t , regroup the c_t that must have the same coefficient if one requires that C belongs to **Sym**.

In this case, one can prove the following result.

THEOREM 6.5. *The intersection $\text{Lie} \cap \mathbf{Sym}$ is contained in the linear span of the LR-classes.*

Proof. It is known that Lie is (strictly) contained in the primitive Lie algebra of **FQSym** (cf. [28]; this is also an easy consequence of Ree’s criterion). The coproduct of **Sym** is the restriction of that of **FQSym**, so all elements of $\text{Lie} \cap \mathbf{Sym}$ are primitive. Conversely, the primitive Lie algebra of **Sym** is generated by the Dynkin elements Ψ_n , which is an element of the PBW basis, and is alone in its LR-class (its bipartite free tree is the star with a black vertex at the center). Hence, the result follows from Proposition 6.6 below. \square

The number of LR-classes in size n is greater than the dimension of $\mathbf{Sym}_n \cap \text{Lie}(n)$, so that an LR-class is not necessarily in **Sym**.

Here are some tables to clarify this point.

Let us define $a'_{n,k}$ as the dimension of $\text{Lie}(n)$ intersected with $\mathbf{Sym}_{n,k}$ where $\mathbf{Sym}_{n,k}$ is the subspace of **Sym** spanned by all ribbons with k parts.

Note that the $a'_{n,k}$ are known (from external considerations) to be Sequence A055277, which is the number of Lyndon words in two letters a and b with k times letter b .

(175)

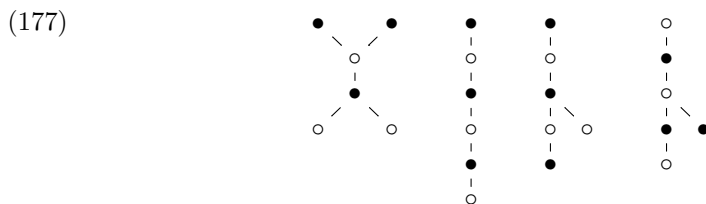
| $n \backslash k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|------------------|---|---|---|---|---|---|---|
| 2 | 1 | | | | | | |
| 3 | 1 | 1 | | | | | |
| 4 | 1 | 1 | 1 | | | | |
| 5 | 1 | 2 | 2 | 1 | | | |
| 6 | 1 | 2 | 3 | 2 | 1 | | |
| 7 | 1 | 3 | 5 | 5 | 3 | 1 | |
| 8 | 1 | 3 | 7 | 8 | 7 | 3 | 1 |

A245558

FIGURE 4. The $a'_{n,k}$.

Let now $b'_{n,k}$ be the number of LR-classes with $n - 1$ nodes and k right branches.

The first discrepancy between $a'_{n,k}$ and $b'_{n,k}$ occurs at $n = 6$. Indeed, all ten LR-classes do not belong to **Sym**. The discrepancy occurs with $k = 2$, hence on free trees with 3 black and 3 white nodes. They are



(176)

| $n \setminus k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|-----------------|---|---|----|----|----|---|---|
| 2 | 1 | | | | | | |
| 3 | 1 | 1 | | | | | |
| 4 | 1 | 1 | 1 | | | | |
| 5 | 1 | 2 | 2 | 1 | | | |
| 6 | 1 | 2 | 4 | 2 | 1 | | |
| 7 | 1 | 3 | 7 | 7 | 3 | 1 | |
| 8 | 1 | 3 | 10 | 14 | 10 | 3 | 1 |

A122085

FIGURE 5. The $b'_{n,k}$.

Let us denote these LR-classes by L_1, \dots, L_4 . Expanding these as combinations of permutations in \mathbf{FQSym}_6 , one finds that none belongs to \mathbf{Sym}_6 . However, the intersection of their linear span with \mathbf{Sym}_6 has dimension 3, and a linear basis is given, e.g. by

(178)
$$L_1 + L_4, \quad L_2 + L_4, \quad L_3 - L_4.$$

PROPOSITION 6.6. *Let $\mathcal{B}(n)$ be the (vector) subspace of $\text{Lie}(n)$ (and of naked trees) generated by LR-classes. Then $\mathcal{B}(n)$ is a Lie algebra.*

Proof. This follows directly from the fact that the subspace generated by L-classes is a preLie algebra with a simple product rule. □

Again, directly from the product rule of the L-classes expressed as bipartite black-rooted trees, the Lie product of the LR-classes is quite simple.

THEOREM 6.7. *The Lie product $[s_1, s_2]$ of two LR-classes can be computed as follows: let B_1 and B_2 be their respective representatives as free bipartite trees. Then the product is obtained as the following sum: all trees obtained by gluing B_1 as a child of a white (resp. black) node of B_2 with a minus (resp. plus) sign. The multiplicity of a given tree is equal to the number of ways to cut it into two parts such that one is s_1 and the other one s_2 .*

7. APPENDIX

7.1. MISCELLANEOUS REMARKS. The LR-classes admit several alternative definitions. The original one involves elementary moves on binary trees. Another one relies on two non-crossing partitions which can be read on a complete binary tree with n leaves: label the sectors by $1, \dots, n - 1$ from left to right. The blocks of the first partition π consist of the sectors which are separated by a left branch. The blocks of the second one π' consist of the sectors which are separated by a right branch. Then, π is the Kreweras complement of π' . Define a directed graph whose vertices are blocks of b' of π' and blocks b of π . The oriented branches are the pairs (b', b) such that $b' \cap b \neq \emptyset$. This graph can be encoded by a bicolored tree, by coloring sinks and source in black and white, and forgetting the arrows and the labels.

CONJECTURE 7.1. *The bracket of \mathcal{B} is*

(179)
$$[x, y] = x \vdash y - y \vdash x$$

where \vdash is a Lie admissible operation, computed on bicolored trees from a coproduct à la Connes–Kreimer: the coefficient of z in $x \vdash y$ is equal to the number of subgraphs of z isomorphic to x such that removing its vertices yields a graph isomorphic to y .

Let us also observe that the main result of [2], the expression of the Solomon idempotent on the basis c_t of \mathfrak{C} , is actually (as expected) an expression of the LR -classes, and that moreover, the coefficient of a free tree is the same for its two possible bicolourings.

7.2. ADMISSIBLE LABELLINGS AND DECREASING TREES.

PROPOSITION 7.2. *The admissible labellings of a complete binary tree T are in bijection with the permutations whose decreasing tree has shape t , the binary tree consisting of the internal nodes of T . Their number is therefore given by the hook-length formula.*

Proof. Let $T_1(\alpha), T_2(\beta)$ be two admissible labellings. Define

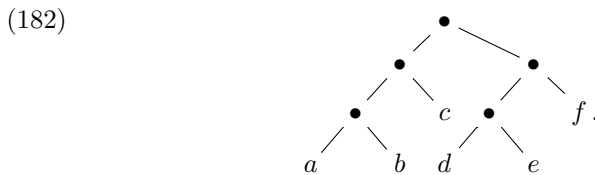
$$(180) \quad B(T_1(\alpha), T_2(\beta)) = \sum_{\substack{\text{std}(u)=\alpha, \text{std}(v)=\beta \\ 1 \in u, n \in v, \gamma=uv}} T(\gamma),$$

where T is the complete binary tree with left and right subtrees T_1 and T_2 . The sum runs over a subset of the admissible labellings of T , and each admissible labelling is obtained in one, and only one sum $B(T_1(\alpha), T_2(\beta))$ for some α and β . Thus, the sum of all admissible labellings of T is $B_T(\bullet_1)$, where for a complete binary tree T and a bilinear map B , $B_T(a)$ means the evaluation of the same tree with B at internal nodes and a in all the leaves. This yields a bijection between these labellings and the permutations of \mathfrak{S}_{n-1} occurring in $B'_T(1)$, where

$$(181) \quad B'(\mathbf{G}_\alpha, \mathbf{G}_\beta) = \sum_{\substack{\text{std}(u)=\alpha, \text{std}(v)=\beta \\ 1 \in u, n \in v, \gamma=uv}} \mathbf{G}\gamma = \mathbf{G}_\alpha \succ 1 \prec \mathbf{G}_\beta$$

in \mathbf{FQSym} (see (15) and (16) below for the last expression), which are precisely the permutations whose decreasing tree is t . □

For example, consider the complete binary tree T



The constraints on the labellings impose that $a = 1, f = 6, b < c$, and $d < e$. There are therefore 6 different labellings:

$$(183) \quad 123456, 124356, 125346, 134256, 135246, 145236.$$

These labellings are indeed in bijection with the decreasing trees of shape



where the constraints fix that $F = 5$, $B < C$, and $D < E$. Note that the inequalities are exactly the same between capital and non-capital letters.

This observation allows a more direct description of the bijection. Start with a complete binary tree T and an admissible labelling. Record into each internal node the set of values of the leaves below it. Then remove 1 from all subsets containing it, send each value i to $i - 1$ in all subsets and then select in each subset (starting from the leaves and moving towards the root) the smallest value that has not yet been used by the nodes below it.

One easily checks that the resulting tree is indeed decreasing, that the process can be reversed, and that starting with a decreasing tree, one indeed obtains an admissible labelling of the corresponding complete binary tree.

For example, this bijection sends each labelling $(1, b, c, d, e, 6)$ of the above complete binary tree to the labelling $(B, C, D, E, F) = (b - 1, c - 1, d - 1, e - 1, 5)$ of the corresponding incomplete binary tree.

7.3. THE TAMARI ORDER ON PLANE TREES. Recall the cover relation of the Tamari order on plane trees: starting from a tree T and a vertex x that is neither its root or a leaf, the trees $T' > T$ covering T are obtained by cutting off the leftmost subtree of x and grafting it back on the left of the parent of x .

The Hasse diagrams here are drawn upside down with respect to the usual convention: the smallest element is at the top.

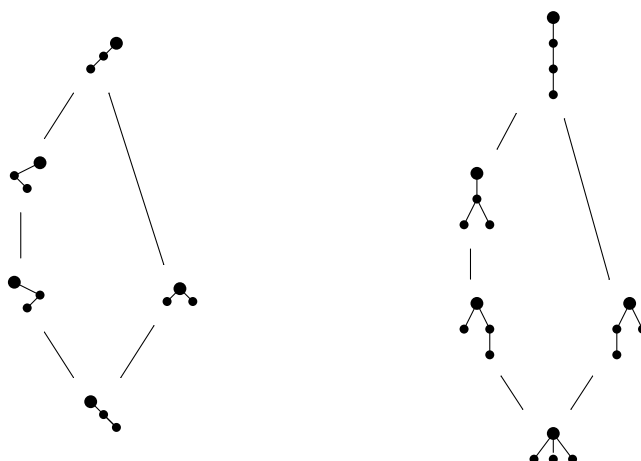


FIGURE 6. The Tamari order on incomplete binary trees and on the corresponding plane trees (size 3).

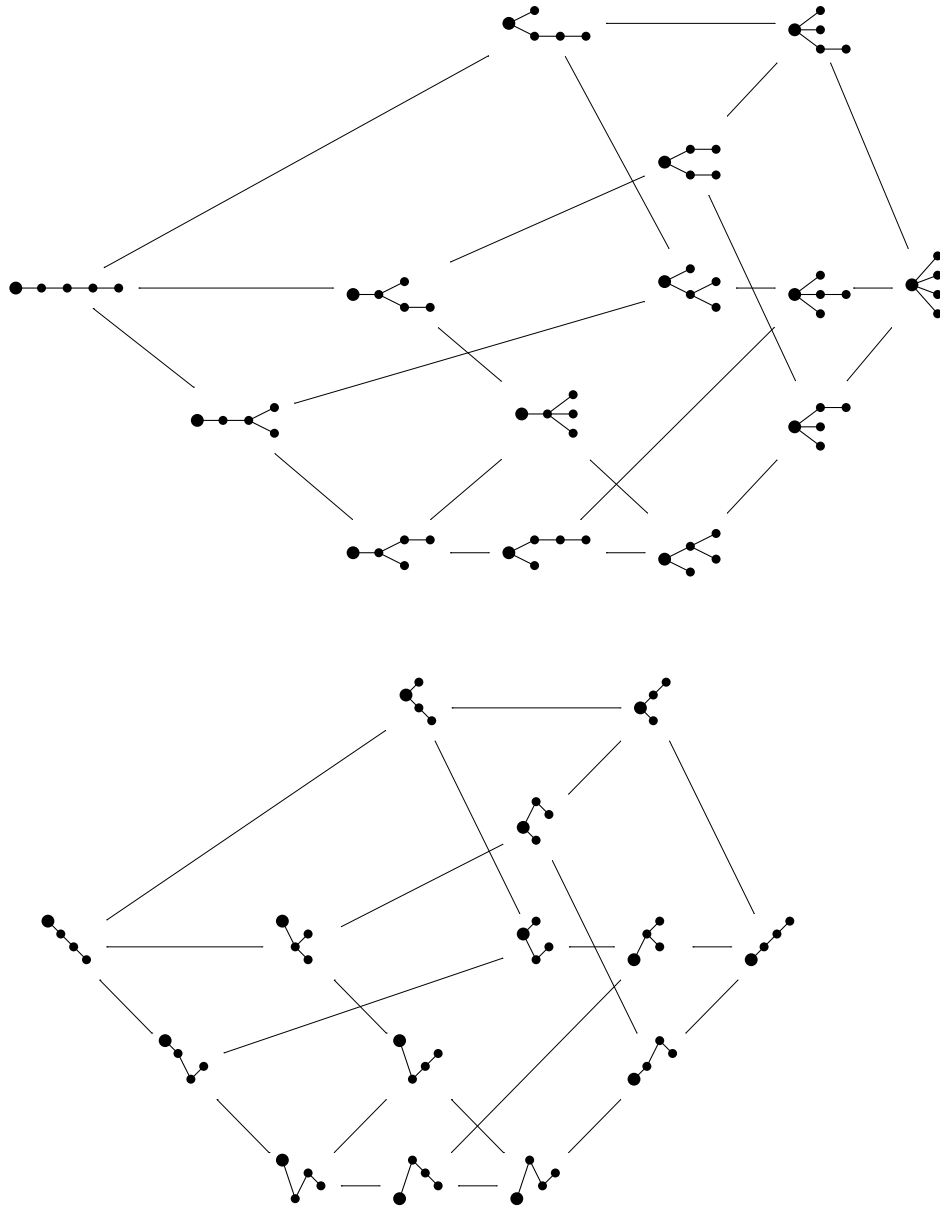


FIGURE 7. The Tamari order on incomplete binary trees and on the corresponding plane trees (size 4).

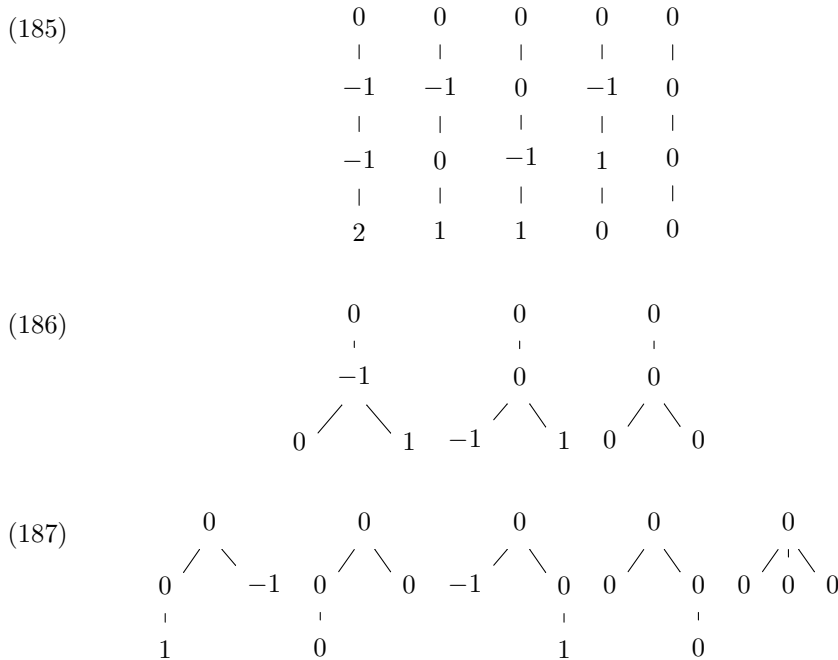


FIGURE 8. Small closed flows on plane trees with 4 nodes

REFERENCES

[1] Marcelo Aguiar and Jean-Louis Loday, *Quadri-algebras*, J. Pure Appl. Algebra **191** (2004), no. 3, 205–221.

[2] Ruggero Bandiera and Florian Schaetz, *Eulerian idempotent, pre-Lie logarithm and combinatorics of trees*, preprint, <https://arxiv.org/abs/1702.08907>, February 2017.

[3] I. Białynicki-Birula, Bogdan Mielnik, and Jerzy Plebański, *Explicit solution of the continuous Baker–Campbell–Hausdorff problem and a new expression for the phase operator*, Annals of Physics **51** (1969), no. 1, 187–200.

[4] Pierre Cartier, *Démonstration algébrique de la formule de Hausdorff*, Bull. Soc. Math. France **84** (1956), 241–249.

[5] Frédéric Chapoton, *A rooted-trees q-series lifting a one-parameter family of Lie idempotents*, Algebra Number Theory **3** (2009), no. 6, 611–636.

[6] ———, *Flows on rooted trees and the Menous–Novelli–Thibon idempotents*, Math. Scand. **115** (2014), no. 1, 20–61.

[7] Frédéric Chapoton and Muriel Livernet, *Pre-Lie algebras and the rooted trees operad*, Internat. Math. Res. Notices (2001), no. 8, 395–408.

[8] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon, *Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras*, Internat. J. Algebra Comput. **12** (2002), no. 5, 671–717.

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

[10] Loïc Foissy, *Bidendriform bialgebras, trees, and free quasi-symmetric functions*, J. Pure Appl. Algebra **209** (2007), no. 2, 439–459.

[11] ———, *Free brace algebras are free pre-Lie algebras*, Comm. Algebra **38** (2010), no. 9, 3358–3369.

[12] Israel M. Gel’fand, Daniel Krob, Alain Lascoux, Bernard Leclerc, Vladimir S. Retakh, and Jean-Yves Thibon, *Noncommutative symmetric functions*, Adv. Math. **112** (1995), no. 2, 218–348.

[13] Florent Hivert, Jean-Christophe Novelli, and Jean-Yves Thibon, *The algebra of binary search trees*, Theoret. Comput. Sci. **339** (2005), no. 1, 129–165.

- [14] Alexander A. Klyachko, *Lie elements in a tensor algebra*, Sibirsk. Mat. Ž. **15** (1974), 1296–1304, 1430.
- [15] Daniel Krob, Bernard Leclerc, and Jean-Yves Thibon, *Noncommutative symmetric functions. II. Transformations of alphabets*, Internat. J. Algebra Comput. **7** (1997), no. 2, 181–264.
- [16] Jean-Louis Loday, *Opérations sur l’homologie cyclique des algèbres commutatives*, Invent. Math. **96** (1989), no. 1, 205–230.
- [17] ———, *Scindement d’associativité et algèbres de Hopf*, in Actes des Journées Mathématiques à la Mémoire de Jean Leray, Sémin. Congr., vol. 9, Soc. Math. France, Paris, 2004, pp. 155–172.
- [18] Jean-Louis Loday, Alessandra Frabetti, Frédéric Chapoton, and François Goichot, *Dialgebras and related operads*, Lecture Notes in Mathematics, vol. 1763, Springer-Verlag, Berlin, 2001.
- [19] Jean-Louis Loday and María O. Ronco, *Hopf algebra of the planar binary trees*, Adv. Math. **139** (1998), no. 2, 293–309.
- [20] Ian G. Macdonald, *Symmetric functions and Hall polynomials. With contribution by A. V. Zelevinsky and a foreword by Richard Stanley. Reprint of the 2008 paperback edition*, second ed., Oxford Classic Texts in the Physical Sciences, The Clarendon Press, Oxford University Press, New York, 2015.
- [21] Claudi Malvenuto and Christophe Reutenauer, *Duality between quasi-symmetric functions and the Solomon descent algebra*, J. Algebra **177** (1995), no. 3, 967–982.
- [22] Frédéric Menous, *The well-behaved Catalan and Brownian averages and their applications to real resummation*, Proceedings of the Symposium on Planar Vector Fields (Lleida, 1996), vol. 41, 1997, pp. 209–222.
- [23] Frédéric Menous, Jean-Christophe Novelli, and Jean-Yves Thibon, *Mould calculus, polyhedral cones, and characters of combinatorial Hopf algebras*, Adv. in Appl. Math. **51** (2013), no. 2, 177–227.
- [24] Bogdan Mielnik and Jerzy Plebański, *Combinatorial approach to Baker–Campbell–Hausdorff exponents*, Ann. Inst. H. Poincaré Sect. A (N.S.) **12** (1970), 215–254.
- [25] Jean-Christophe Novelli and Jean-Yves Thibon, *Construction de trigèbres dendrifformes*, C. R. Math. Acad. Sci. Paris **342** (2006), no. 6, 365–369.
- [26] ———, *Polynomial realizations of some trialgebras*, 18th Formal Power Series and Algebraic Combinatorics (FPSAC’06) (San Diego, USA, France), no. 1, 2006, pp. 243–254.
- [27] Jean-Michel Oudom and Daniel Guin, *Sur l’algèbre enveloppante d’une algèbre pré-Lie*, C. R. Math. Acad. Sci. Paris **340** (2005), no. 5, 331–336.
- [28] Frédéric Patras and Christophe Reutenauer, *Lie representations and an algebra containing Solomon’s*, J. Algebraic Combin. **16** (2002), no. 3, 301–314 (2003).
- [29] Christophe Reutenauer, *Theorem of Poincaré–Birkhoff–Witt, logarithm and symmetric group representations of degrees equal to Stirling numbers*, in Combinatoire énumérative (Montreal, Que., 1985/Quebec, Que., 1985), Lecture Notes in Math., vol. 1234, Springer, Berlin, 1986, pp. 267–284.
- [30] ———, *Free Lie algebras*, London Mathematical Society Monographs. New Series, vol. 7, The Clarendon Press, Oxford University Press, New York, 1993, Oxford Science Publications.
- [31] Paolo Salvatore and Roberto Tauraso, *The operad Lie is free*, J. Pure Appl. Algebra **213** (2009), no. 2, 224–230.
- [32] Louis Solomon, *On the Poincaré–Birkhoff–Witt theorem*, J. Combinatorial Theory **4** (1968), 363–375.
- [33] ———, *A Mackey formula in the group ring of a Coxeter group*, J. Algebra **41** (1976), no. 2, 255–264.
- [34] Ernst Witt, *Treue Darstellung Liescher Ringe*, J. Reine Angew. Math. **177** (1937), 152–160.

LOÏC FOISSY, Fédération de Recherche Mathématique du Nord Pas de Calais FR 2956, Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville, Université du Littoral Côte d'Opale-Centre Universitaire de la Mi-Voix, 50, rue Ferdinand Buisson, CS 80699, 62228 Calais Cedex, France

E-mail : foissy@univ-littoral.fr

FRÉDÉRIC MENOUS, Laboratoire de Mathématiques, Bât. 425, Université Paris-Sud, 91405 Orsay Cedex, France

E-mail : Frederic.Menous@u-psud.fr

JEAN-CHRISTOPHE NOVELLI, Laboratoire d'Informatique Gaspard-Monge, Université Gustave Eiffel, CNRS, ENPC, ESIEE-Paris,, 5 Boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, FRANCE

E-mail : novelli@univ-mlv.fr

JEAN-YVES THIBON, Laboratoire d'Informatique Gaspard-Monge, Université Gustave Eiffel, CNRS, ENPC, ESIEE-Paris,, 5 Boulevard Descartes, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, FRANCE

E-mail : jyt@univ-mlv.fr