AGT

Algebraic & Geometric Topology

Volume 26 (2026)

Primitive Feynman diagrams and the rational Goussarov–Habiro Lie algebra of string links

BRUNO DULAR

msp

Algebraic & Geometric Topology **26**:3 (2026) 1037–1076
DOI: 10.2140/agt.2026.26.1037
Published: April 1, 2026

Primitive Feynman diagrams and the rational Goussarov–Habiro Lie algebra of string links

BRUNO DULAR

Goussarov–Habiro's theory of clasper surgeries defines a filtration of the monoid of string links $L(m)$ on m strands, in a way that geometrically realises the Feynman diagrams appearing in low-dimensional and quantum topology. Concretely, $L(m)$ is filtered by C_n-equivalence, for $n \geq 1$, which is defined via local moves that can be seen as higher-order crossing changes. The graded object associated to the Goussarov–Habiro filtration is the *Goussarov–Habiro Lie algebra of string links* $\mathcal{L}L(m)$. We give a concrete presentation, in terms of primitive Feynman (tree) diagrams and relations (1T, AS, IHX, STU^2), of the rational Goussarov–Habiro Lie algebra $\mathcal{L}L(m)_{\mathbb{Q}}$ and of the primitive Lie algebra of the Hopf algebra of Feynman diagrams. To that end, we investigate cycles in *graphs of forests*: *flip graphs* associated to forest diagrams and their STU relations. As an application, we give an alternative *diagrammatic* proof of Massuyeau's rational version of the Goussarov–Habiro conjecture for string links, which relates indistinguishability under finite type invariants of degree $< n$ and C_n-equivalence.

1 Introduction

1.1 Background

A *string link on m strands in the cylinder* is a smooth embedding of a disjoint union of m intervals into the cylinder

$$\gamma : \{1, \dots, m\} \times [0, 1] \hookrightarrow \mathbb{D}^2 \times [0, 1]$$

such that $\gamma(i, e) = (x_i, e)$ for all $i \in \{1, \dots, m\}$ and $e \in \{0, 1\}$, where $x_1, \dots, x_m \in \mathbb{D}^2$ are fixed points (usually ordered along the x-axis), and γ is perpendicular to the boundary near those points. String links on m strands are considered up to smooth isotopy, and $L(m)$ denotes the corresponding set of isotopy classes, also called *string links* when there is no ambiguity.

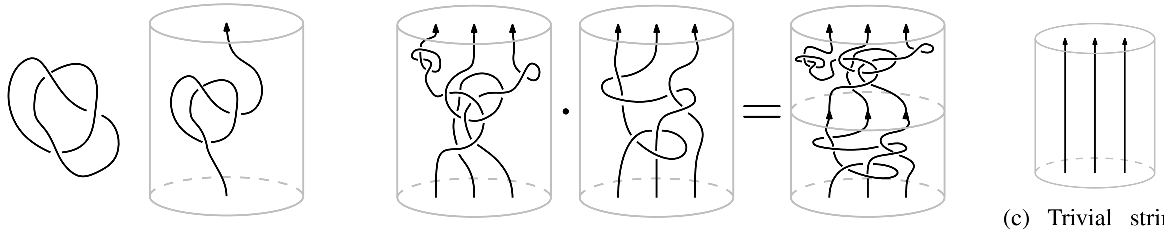
The set $L(m)$ is a *monoid* with multiplication given by vertical concatenation (Figure 1(b)), with the trivial string link as unit (Figure 1(c)).

When $m = 1$, elements of $L(1)$ are called *long knots* and *closing a long knot* yields an isomorphism between $L(1)$ and the monoid of knots under connected sum (Figure 1(a)). Thus, the study and classification of string links naturally belong to knot theory, and one of the main tools for probing them is the use of *invariants*.

A string link invariant $V : L(m) \to A$ valued in an abelian group A is a *Vassiliev invariant* or *finite type invariant* of *degree n* if it satisfies specific *skein relations* (see Vassiliev [43], Birman and Lin [3] and Bar-Natan [1]), which can be read as the condition that V vanishes on the $(n+1)$-st power of the

---

MSC2020: primary 57K16; secondary 16T30, 17B70.



(a) A knot and a long knot (b) Two string links and their concatenation in $L(3)$ (c) Trivial string link in $L(3)$

Figure 1: Long knots and string links.

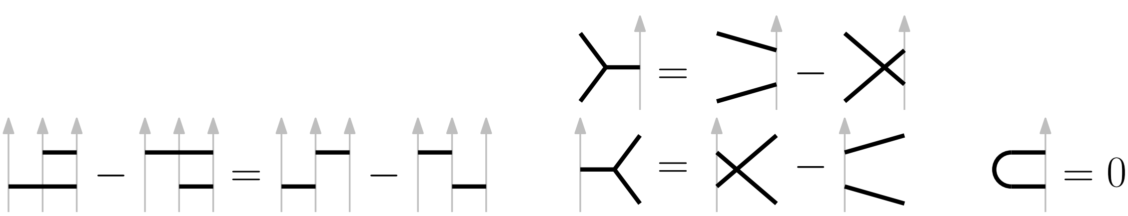
augmentation ideal in the monoid ring $\mathbb{Z}L(m)$. Vassiliev invariants dominate (dominate means that the collection of all Vassiliev invariants determines the mentioned known invariants) polynomial invariants (Conway, HOMFLY, etc.), quantum invariants (see Turaev [42]), Milnor invariants (see the survey by Meilhan [33]), etc. but it is an open question whether they separate string links (or even knots, in the case $m = 1$). One of the main results in that theory, due to Bar-Natan [1] and Kontsevich [25], is that any $\mathbb{Q}$-valued Vassiliev invariant factors through the Kontsevich integral

$$Z : L(m) \to \widehat{\mathcal{A}^{\mathrm{FI}}}(m)_{\mathbb{Q}}, \tag{1-1}$$

which is valued in the graded completion of the Hopf algebra $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ of *Feynman diagrams* on m strands, generated by uni-trivalent diagrams (with cyclic orientation at the trivalent vertices (the cyclic orientation is usually omitted, in which case it is assumed to be counterclockwise), and univalent vertices attached to m vertical strands; see Figure 4), modulo 1T and STU relations (which imply the relations 4T, AS, IHX; see Figure 2). The presence of the 1T relation translates the fact that we are working with *unframed string links*. This justifies the notation FI for *framing independence*, used throughout this paper.

For the above reason, the Kontsevich integral is called a *universal Vassiliev invariant over* $\mathbb{Q}$. The theory is much less understood over the integers.

A natural question about Vassiliev invariants is: what do they detect? In the case of knots ($m = 1$), Goussarov [17] and Habiro [22] independently answered this question by introducing C_n*-moves*, defined



(a) A 4T, *four-terms*, relation (b) STU relations (c) A 1T, *one-term*, relation

(d) An AS relation (e) An IHX relation in the *Kirchhoff form* [8, Section 5.2.7]

Figure 2: The 4T, STU, 1T, AS and IHX relations.

(a) A C_1-move (b) A C_2-move

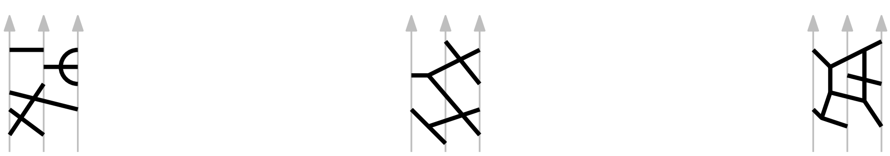
Figure 3: Diagrams, claspers realising them and C_n-moves.

using *clasper surgeries*. Those are local moves on string links that are modelled on *tree claspers*, ribbon trees with leaves attached to the strands, that geometrically realise the Feynman diagrams from above. See Figure 3 for examples of C_1- and C_2-moves. More precisely, C_n-moves generate the C_n-equivalence relation on $L(m)$, and they show that any two C_n-equivalent string links are V_n-equivalent, i.e., they cannot be distinguished by Vassiliev invariants of degree $< n$. Goussarov and Habiro conjecture that the converse holds as well: this is the *Goussarov–Habiro conjecture for string links in the cylinder* (which they prove for knots, i.e., the case $m = 1$). This is still wide open, with only progress made in small degrees by Meilhan and Yasuhara [34].

For each $n \geq 1$, the monoid $\mathcal{L}_n L(m) := L_n(m)/C_{n+1}$ of C_{n+1}-equivalence classes of C_n-trivial string links turns out to be a finitely generated abelian group. See the works of Goussarov, Polyak and Viro [18] and of Habiro [22]. Those combine into a graded Lie $\mathbb{Z}$-algebra $\mathcal{L}L(m) := \bigoplus_{n\geq 1} \mathcal{L}_n L(m)$, the *Goussarov–Habiro Lie algebra of string links on m strands*. From their work, it follows that there is a surjective $\mathbb{Z}$-linear graded *realisation map* $R : \mathbb{Z}\mathcal{D}^T(m) \twoheadrightarrow \mathcal{L}L(m)$ that sends a tree diagram to the result of surgery along a tree clasper realising it. Thus there is a presentation of the Goussarov–Habiro Lie algebra by tree diagrams on m strands modulo specific relations

$$\mathcal{L}L(m) \cong \mathbb{Z}\mathcal{D}^T(m)/\ker(R),$$

whose understanding would be a significant step towards the Goussarov–Habiro conjecture, while also being of its own interest. For example, it would provide a "*universal Goussarov–Habiro invariant*", valued in a Lie algebra of tree diagrams. This is related to the study of π_0 of the embedding calculus tower, as appearing in the works of Conant [9], Budney, Conant, Koytcheff and Sinha [5], and Kosanović [26]. The concordance analogue of this question, where C_n-equivalence is replaced by *C_n-concordance*, appears in works of Conant, Schneiderman and Teichner [11; 12], while the homotopy analogue has been settled by Habegger, Lin and Masbaum [20; 21]. Habiro and Massuyeau studied a similar question in the context of homology cylinders in [23]. Nozaki, Sato and Suzuki also studied $\ker(R)$ in [37].



(a) Chord diagram (b) Forest diagram (c) Feynman diagram

Figure 4: Examples of diagrams.

1.2 Results and strategy

The main contribution of the present paper is an explicit presentation of the *rational* Goussarov–Habiro Lie algebra $\mathcal{L}L(m)_{\mathbb{Q}} := \mathcal{L}L(m) \otimes \mathbb{Q}$, and of the primitive Lie algebra of the Hopf algebra of Feynman diagrams $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$. In other words, we identify $\ker(R)_{\mathbb{Q}}$ as the subgroup generated by specific relations in $\mathbb{Q}\mathcal{D}^T(m)$ (where $\mathcal{D}^T(m)$ is the set of tree diagrams on m strands): the usual $1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}$ relations, completed with Conant's STU^2 relation; see (1-2).

Theorem 1 *The three following graded Lie $\mathbb{Q}$-algebras are isomorphic to each other*:

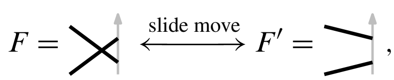
(i) *the Goussarov–Habiro Lie algebra of string links* $\mathcal{L}L(m)_{\mathbb{Q}}$;

(ii) *the primitive Lie algebra* $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m))_{\mathbb{Q}}$ *of the Hopf algebra of Feynman diagrams*;

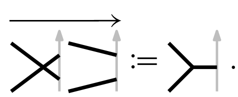
(iii) *the Lie algebra of tree diagrams* $\mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ *defined as* $\mathbb{Q}\mathcal{D}^T(m)/\langle 1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle$.

The isomorphism $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m))_{\mathbb{Q}} \cong \mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ is Theorem 2.4.5.1 in the text and is the focus of Section 2, which we outline now. It will be more convenient to see the algebra $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ as being generated by *forest diagrams*, i.e., Feynman diagrams without cycles, modulo 1T and STU. First, we prove that $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m))_{\mathbb{Q}}$ is generated by tree diagrams. To that end, we filter $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ by the *size* of forests, and we show that this size filtration coincides with the primitive filtration (Theorem 2.2.2.3). This requires averaging over permutations and is one of the main reasons why we cannot extend Theorem 1 over $\mathbb{Z}$ with the current methods.

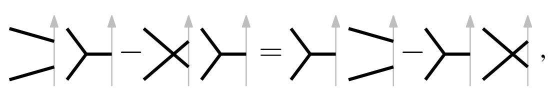
However, $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ is defined using the STU relation, which is not *size homogeneous*: it relates three forest diagrams, two of which with one more tree than the other. Here lies the main difficulty in finding a concrete presentation of $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$. This is resolved by introducing *graph of forests*, which are *flip graphs* with forest diagrams as vertices and with an edge for each *slide move*



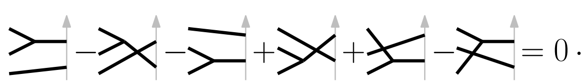
where it is understood that F and F' are identical outside of the shown part. Identifying relations in $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$ becomes a matter of finding cycles in such graphs of forests, and the STU relation becomes a way of assigning forests of smaller size to those cycles, via the assignment



We show that the first homology groups of graphs of forests with rational coefficients are generated by cycles of length 4 and cycles of length 6, respectively responsible for STU^2-relations

(1-2) 

where the four terms represent forest diagrams that are identical outside of the shown parts, and ⬡-relations

(1-3) 

In particular, the legs involved in each term of a ⬡-relation must belong to two distinct trees, for otherwise another term would contain a cycle. Thus, the ⬡-relations do not appear in the presentation of $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$, but they do appear in the presentation of the intermediate steps of the size filtration (more precisely, in size $n-1$ and degree n); see Section 2.4.

The isomorphism $\mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \cong \mathcal{L}L(m)_{\mathbb{Q}}$ follows from Theorems 3.1.3.2 and 3.2.2.1 in the text and is the focus of Section 3. In Theorem 3.1.3.2, we define a *realisation map* $\mathbb{Z}\mathcal{D}^T(m) \to \mathcal{L}L(m)$, which sends a tree diagram to the result of *clasper surgery* along a tree clasper realising it. This map is surjective and indeed satisfies the 1T, AS, IHX and STU^2 relations, by an application of clasper calculus. Similar realisation maps and clasper calculus computations appear in works of Ohtsuki, Conant, Teichner, Schneiderman and Kosanović [10; 13; 16; 26; 38], and hence we omit most details here.

By the above, there is a surjective realisation map $R : \mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \twoheadrightarrow \mathcal{L}L(m)_{\mathbb{Q}}$. To show that it is injective, we use the *Kontsevich integral* (1-1) [1; 25]. The Kontsevich integral indeed detects tree clasper surgeries, and hence it provides an inverse Z^{GH} to R, as follows from the following result.

Theorem 2 (tree preservation theorem (Theorem 3.2.2.1)) *Let $n \geq 1$ and $T \in \mathcal{D}_n^T(m)$ a degree-n tree diagram on m strands. Then*

$$Z(\sigma(C_T)) = 1 + T + O(n+1), \tag{1-4}$$

where $\sigma(C_T)$ denotes the string link obtained from the trivial one by surgery along a tree clasper C_T realising T, and $O(n+1)$ means that the equality holds modulo terms of degree $\geq n+1$.

A proof of this result is sketched by Ohtsuki in [38, Proposition E.24] but we provide a different detailed proof. The use of the Kontsevich integral is another main reason why we cannot extend Theorem 1 over $\mathbb{Z}$.

In Section 4, we apply Theorem 1 to give an alternative proof of Massuyeau's rational version of the Goussarov–Habiro conjecture for string links.

1.3 Further directions

We restricted our attention to unframed string links in order to make the presentation clearer. However, the identification of the primitive Lie algebra of $\mathcal{A}(m)_{\mathbb{Q}}$ holds with and without the 1T relation. By considering the Kontsevich integral for *framed string links*, as described by Le and Murakami [30], one should obtain a presentation of the framed version of the Goussarov–Habiro Lie algebra.

Conant's work [9] on the STU^2 relation was an important ingredient for identifying modules in the second page of the spectral sequence arising in the study of the *Goodwillie–Weiss embedding calculus tower*, as in the works of Budney, Conant, Koytcheff and Sinha [5], Kosanović [26] and Shi [41]. The present work motivates the investigation of the string link analogue of the embedding calculus tower.

Moreover, the Goodwillie–Weiss embedding calculus tower is known to be a universal Vassiliev invariant over $\mathbb{Z}_{(p)}$ in degree $\leq p+1$, by a theorem of Boavida de Brito and Horel [4, Theorem A]. Combining a *string link* version of this result with the present work could lead to a proof of the Goussarov–Habiro conjecture over $\mathbb{Z}_{(p)}$ in degree $n < p$.

1.4 Notation

Lower indices usually refer to *degrees* of diagrams and claspers, and upper indices usually refer to filtrations of diagrams by *size* (number of connected components). The notation FI, for *framing independence*, refers to the presence of the relation 1T. Sets of diagrams are denoted by $\mathcal{D}$ ornamented with various indices, Hopf algebras by $\mathcal{A}$ and Lie algebras by $\mathcal{L}$.

2 Primitive Lie algebra of Feynman diagrams

The goal of this section is to identify the primitive Lie algebra $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$ of the Hopf algebra $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ of Feynman diagrams on m strands.

In the case of knots, i.e., $m = 1$, tree diagrams generate the primitive Lie algebra $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(1)_{\mathbb{Q}})$ of $\mathcal{A}^{\mathrm{FI}}(1)_{\mathbb{Q}}$ [1], which, by [9], is isomorphic to the space of trees

$$\mathbb{Q}\mathcal{D}^{T}(1)/\langle 1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle.$$

See also [26, Theorem G3] for a concise presentation of these results. Commutativity of $\mathcal{A}^{\mathrm{FI}}(1)_{\mathbb{Q}}$ implies, by the Milnor–Moore theorem, that there is an isomorphism between the primitive part $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(1)_{\mathbb{Q}})$ and the inseparable quotient $(\mathcal{A}^{\mathrm{FI}}(1)_{\mathbb{Q}})^{I}$ (where any diagram which is a product of nontrivial diagrams is set to zero). This fact is essential in Conant's result [9], which actually holds over $\mathbb{Z}$.

In the case of string links, tree diagrams also generate the primitive part $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$. However, $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$ is not commutative when $m > 1$ and the argument of Conant does not translate a priori. In this section, we prove that the conclusion still holds: $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})$ is isomorphic to the Lie algebra $\mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} := \mathbb{Q}\mathcal{D}^{T}(m)/\langle 1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle$.

2.1 Algebras of diagrams

This section recalls the definition of chord and Feynman diagrams [1] on tangle skeletons, as well as their bialgebra structure.

2.1.1 Uni-trivalent graphs A *uni-trivalent graph*, also called *Feynman* or *Jacobi graph*, is an undirected graph $G = (V = U \sqcup T, E)$ consisting of univalent and trivalent vertices (forming the sets U and T, respectively). The edges are defined as pairs of distinct vertices.[1] The trivalent vertices are endowed with a cyclic orientation of their three adjacent edges. When a uni-trivalent graph is drawn, the cyclic orientations are the counterclockwise ones unless indicated otherwise. Also, four-valent vertices in drawings are not part of the graph, they are only artefacts of their planar representations.

The univalent and trivalent vertices are often called *leaves* and *nodes*, respectively. An edge adjacent to a leaf is called a *leg*.

The *degree* of a uni-trivalent graph G is $\deg(G) := \frac{|V|}{2}$.

A uni-trivalent graph G is a *forest* if it has no cycles and a *tree* if in addition it has a single component. If G is a tree, then $\deg(G) = |U| - 1 = |T| + 1$.

[1] We could allow diagrams with edges attached to a single vertex, i.e., *tadpoles*, but those will vanish by AS.

2.1.2 Uni-trivalent forest and chord diagrams Let S be an oriented 1-manifold with boundary. In Section 3.2.1, S will be the underlying manifold of what will be referred to as a *tangle skeleton*. Since this work focuses on string links, S will usually be a disjoint union of intervals $\{1,\dots,m\}\times[0,1]$.

A *uni-trivalent* or *Feynman* or *Jacobi diagram* on S is a pair $D=(G,[u])$ where $G=(U\sqcup T,E)$ is a uni-trivalent graph and $[u]$ is the isotopy class of an embedding $u:U\hookrightarrow S-\partial S$ of the univalent vertices into the interior of S. The set of Feynman diagrams on S is denoted $\mathcal{D}^F(S)$.

The *degree* of a Feynman diagram is the degree of its underlying graph. Denote by $\mathcal{D}^F_n(S)\subset\mathcal{D}^F(S)$ the subset of degree-n Feynman diagrams. This endows $\mathcal{D}^F(S)$ with a graded set structure.

A Feynman diagram is a *forest*, resp. *tree*, diagram if its underlying graph is a forest, resp. a tree. A forest diagram is a *chord diagram* if it has no nodes. Denote by $\mathcal{D}(S)$, $\mathcal{D}^T(S)$, $\mathcal{D}^c(S)$ the corresponding graded subsets of $\mathcal{D}^F(S)$. The *size* of a forest diagram is its number of trees, i.e., the number of connected components of the underlying uni-trivalent graph. For example, the forest diagram on Figure 4(b) has size 3. For $s\geq 0$, the set of forest diagrams of size s is denoted $\mathcal{D}^s(S)$. Note that $\mathcal{D}^n_n(S)=\mathcal{D}^c_n(S)$ and $\mathcal{D}^T(S)=\mathcal{D}^1(S)$.

By [1], the inclusions $\mathcal{D}^c(S)\subset\mathcal{D}(S)\subset\mathcal{D}^F(S)$ induce isomorphisms[2] of graded $\mathbb{Z}$-modules

$$\mathcal{A}(S):=\frac{\mathbb{Z}\mathcal{D}^c(S)}{\langle 4\mathrm{T}\rangle}\cong\frac{\mathbb{Z}\mathcal{D}(S)}{\langle \mathrm{STU}\rangle}\cong\frac{\mathbb{Z}\mathbb{D}^F(S)}{\langle \mathrm{STU}\rangle} \tag{2-1}$$

and the relations AS, IHX are consequences of STU (see Figure 2). Unless stated otherwise, we use the presentation of $\mathcal{A}(S)$ using forest diagrams. The degree-n part of $\mathcal{A}(S)$ is denoted $\mathcal{A}_n(S)$. The only degree-0 diagram is the empty one, and hence $\mathcal{A}_0(S)\cong\mathbb{Z}$. In the present paper, we mostly consider those graded modules with rational coefficients, namely $\mathcal{A}(S)_{\mathbb{Q}}:=\mathcal{A}(S)\otimes\mathbb{Q}$, whose degree-$n$ part is $\mathcal{A}_n(S)_{\mathbb{Q}}=\mathcal{A}_n(S)\otimes\mathbb{Q}$.

2.1.3 Framed versus unframed There is an *unframed* or *framing independent* (FI) version of the module of Feynman diagrams

$$\mathcal{A}^{\mathrm{FI}}(S):=\mathcal{A}(S)/\langle 1\mathrm{T}\rangle, \tag{2-2}$$

where 1T is the graded submodule generated by all diagrams containing an *isolated chord*, i.e., a chord whose endpoints lie on the same strand with no other endpoint in between them.

Writing $\mathcal{A}^{(\mathrm{FI})}(S)$ indicates that the statement holds for both $\mathcal{A}(S)$ and $\mathcal{A}^{\mathrm{FI}}(S)$.

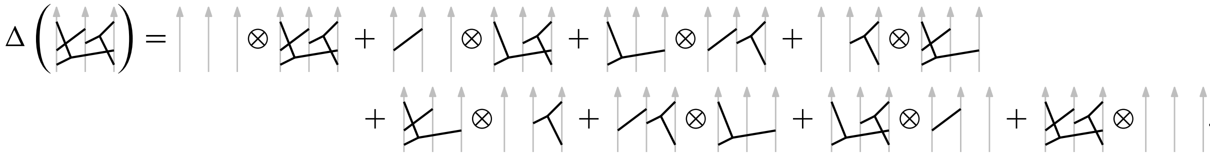
2.1.4 Coalgebra structure The graded $\mathbb{Z}$-module $\mathcal{A}^{(\mathrm{FI})}(S)$ has a structure of coalgebra, with comultiplication [1, Definition 3.7]

$$\Delta(D):=\sum_{J\subseteq\pi_0(D)} D_J\otimes D_{\pi_0(D)-J},$$

[2] In [1], it is shown that $\mathcal{D}^c(S)\hookrightarrow\mathcal{D}^F(S)$ induces an isomorphism $\mathbb{Z}\mathcal{D}^c(S)/\langle 4\mathrm{T}\rangle\cong\mathbb{Z}\mathcal{D}^F(S)/\langle\mathrm{STU}\rangle$. Thus, we get $\mathbb{Z}\mathcal{D}^c(S)/\langle 4\mathrm{T}\rangle\hookrightarrow\mathbb{Z}\mathcal{D}(S)/\langle\mathrm{STU}\rangle\twoheadrightarrow\mathbb{Z}\mathcal{D}^F(S)/\langle\mathrm{STU}\rangle$ and the first map is surjective since any forest can be written as a sum of chord diagrams by applying STU until it has no node left. Hence the first map is an isomorphism, so is the second since their composition is an isomorphism. The same holds with FI added everywhere.

where the sum runs over all the subsets J of the set $\pi_0(D)$ of connected components of D and D_J denotes the subdiagram of D consisting of those components contained in J. The counit ϵ is given by $\epsilon(D) = 1$ if D is the empty diagram, and it is zero otherwise.

For example, with $S = \{1, 2, 3\} \times [0, 1]$, one has



The comultiplication is coassociative and cocommutative.

2.1.5 Hopf algebra of diagrams on string links When $S = \{1, \dots, m\} \times [0, 1]$ is a disjoint union of intervals, we write $\mathcal{D}(m)$ for $\mathcal{D}(S)$, $\mathcal{A}(m)$ for $\mathcal{A}(S)$, $\mathcal{A}(m)_{\mathbb{Q}}$ for $\mathcal{A}(S)_{\mathbb{Q}}$ and so on. In that case, vertical stacking of diagrams endows $\mathcal{A}(m)$ and $\mathcal{A}^{\mathrm{FI}}(m)$ with a multiplication. More precisely, $D \cdot D'$ denotes the vertical concatenation of D and D', with D above D'. The unit 1 is the empty diagram.

Proposition 2.1.5.1 (Bar-Natan [1, Proposition 3.9]) *The comultiplication and multiplication defined above endow $\mathcal{A}^{(\mathrm{FI})}(m)$ with the structure of a connected cocommutative Hopf algebra. It is commutative when $m = 1$.*

2.2 Primitive filtration of forest diagrams

2.2.1 Primitive filtration of a Hopf algebra Let us recall some basic notions about primitive elements in Hopf algebras. See [6; 35; 36, Chapter 5; 40] for more details.

Fix a Hopf algebra A over $\mathbb{Q}$ with multiplication μ or $\cdot$, unit 1, comultiplication Δ and counit ϵ. An element $x \in A$ is *primitive* if $\Delta(x) = x \otimes 1 + 1 \otimes x$. A direct calculation shows that the commutator of two primitive elements is primitive as well. In particular, the set $\mathrm{Prim}(A)$ of primitive elements in A is a Lie algebra, called the *primitive Lie algebra of* A. Conversely, to any Lie algebra L one can associate its *universal enveloping algebra* $U(L)$, which is a Hopf algebra. The Milnor–Moore theorem states conditions under which the functors Prim and U are inverses of each other.

Theorem 2.2.1.1 (Milnor–Moore theorem [36, Theorem 5.6.5]) *Let A be a connected*[3] *and cocommutative Hopf algebra over a characteristic zero field. Then the inclusion* $\mathrm{Prim}(A) \hookrightarrow A$ *induces an isomorphism of Hopf algebras* $U(\mathrm{Prim}(A)) \cong A$.

Let us assume that A is a connected and cocommutative Hopf algebra over $\mathbb{Q}$, so that $U(\mathrm{Prim}(A)) \cong A$. The *primitive filtration* of A is the filtration

$$\mathbb{Q} \cdot 1 = P^0 A \hookrightarrow P^1 A \hookrightarrow P^2 A \hookrightarrow \cdots,$$

where $P^k A$ is the $\mathbb{Q}$-submodule of A additively generated by $\mathbb{Q} \cdot 1$ and all products of $\le k$-many primitive elements. In particular $P^1 A = \mathbb{Q} \oplus \mathrm{Prim}(A)$ and any element $x \in A$ is in some $P^k A$, for some k, since

[3] A Hopf algebra A is *connected* if $A_0 := \mathbb{Q} \cdot 1$ is its unique simple subcoalgebra.

A is primitively generated. Under the isomorphism $U(\mathrm{Prim}(A)) \cong A$, the primitive filtration coincides with the *Lie filtration* of $U(\mathrm{Prim}(A))$ [35, Proposition 5.17].

For all $k \geq 0$, we have $P^k A = \mathbb{Q} \oplus P^k A_{\geq 1}$ where $A_{\geq 1} := A_+$ is the kernel of the counit. The filtration

$$0 = P^0 A_{\geq 1} \hookrightarrow P^1 A_{\geq 1} \hookrightarrow P^2 A_{\geq 1} \hookrightarrow \cdots$$

is called the *reduced primitive filtration* of $A_{\geq 1}$.

2.2.2 Primitive filtration of forests The focus returns to the Hopf algebra of forest diagrams on m strands, $\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}}$, which is indeed connected and cocommutative, and hence the Milnor–Moore theorem applies, i.e., $U(\mathrm{Prim}(\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}})) \cong \mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}}$. It is also filtered by the primitive filtration.

In the case $m = 1$, Bar-Natan showed that $\mathrm{Prim}(\mathcal{A}(1)_{\mathbb{Q}})$ is (additively) generated by all connected diagrams [1] and his argument translates well for $m \geq 1$. Actually, this statement is also true over $\mathbb{Z}$, as shown by Lando in [28] via a beautiful counting argument. Unfortunately, the argument uses commutativity in $\mathcal{A}(1)$, and hence it does not translate directly to the general case $m \geq 1$. Instead, we give an argument that compares the primitive filtration to the size filtration defined below, in a way parallel to the ideas of the subsequent subsections in this section. First, let us introduce some notation.

Definition 2.2.2.1 The *size filtration* of $\mathcal{A}(m)$ is the filtration $\{F^k\mathcal{A}(m)\}_{k\geq 0}$ where $F^k\mathcal{A}(m)$ is the $\mathbb{Z}$-submodule generated by forest diagrams of size $\leq k$. In degree $n \geq 0$, it induces a filtration[4]

$$F^0\mathcal{A}_n(m) \hookrightarrow F^1\mathcal{A}_n(m) \hookrightarrow \cdots \hookrightarrow F^{n-1}\mathcal{A}_n(m) \hookrightarrow F^n\mathcal{A}_n(m) = \mathcal{A}_n(m)$$

of $\mathcal{A}_n(m)$. The *size filtration* of $\mathcal{A}^{\mathrm{FI}}(m)$ is defined in the same way, with FI added everywhere. Tensoring everything with $\mathbb{Q}$ yields the size filtration of $\mathcal{A}(m)_{\mathbb{Q}}$.

For $F \in \mathcal{D}(m)$ a forest diagram, define the *reduced comultiplication* $\overline{\Delta}(F)$ as

$$\overline{\Delta}(F) := \sum_{\varnothing \neq J \subsetneq \pi_0(F)} (F_J) \otimes (F_{\pi_0(F)-J}),$$

where F_J denotes the subdiagram of F that consists of the components contained in J. In other words, $\overline{\Delta}(F) = \Delta(F) - F \otimes 1 - 1 \otimes F$. Thus, $\overline{\Delta}$ satisfies AS, IHX, STU (and 1T) and descends to a well-defined map

$$\overline{\Delta} : \mathcal{A}_{\geq 1}(m) \to \mathcal{A}_{\geq 1}(m) \otimes \mathcal{A}_{\geq 1}(m).$$

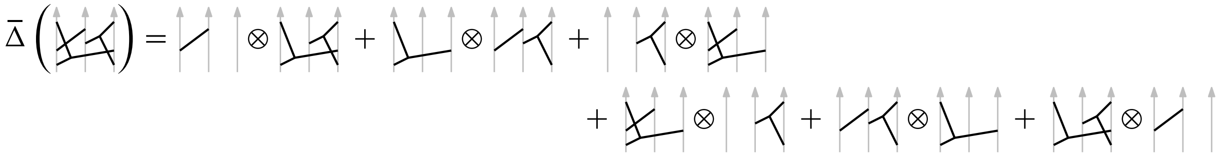
Define recursively $\overline{\Delta}^k := (\overline{\Delta}^{k-1} \otimes \mathrm{id}) \circ \overline{\Delta} : \mathcal{A}_{\geq 1}(m) \to (\mathcal{A}_{\geq 1}(m))^{\otimes k+1}$, for $k > 1$. Concretely, we have

$$\overline{\Delta}^k(F) = \sum_{\substack{J_0 \sqcup \cdots \sqcup J_k = \pi_0(F) \\ J_i \neq \varnothing}} (F_{J_0}) \otimes \cdots \otimes (F_{J_k}), \tag{2-3}$$

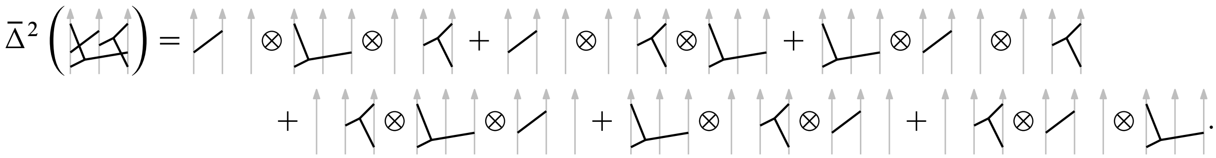
where the sum is over all possible partitions of the components of F into $k + 1$ nonempty subforests.

[4] The only forest diagram of size 0 is the empty diagram 1, and hence $F^0\mathcal{A}(m) = \mathbb{Z} \cdot 1$.

For example, we have



and



Note also that the reduced primitive filtration of $\mathcal{A}_{\geq 1}(m)_{\mathbb{Q}}$ can already be defined over $\mathbb{Z}$: for $k \geq 0$, $P^k \mathcal{A}_{\geq 1}(m)$ is generated by all products of $\leq k$-many primitive elements.

Lemma 2.2.2.2 *We have the following properties*:

(i) *for all* $x, y \in \mathcal{A}_{\geq 1}(m)$,

$$\overline{\Delta}(x \cdot y) = \overline{\Delta}(x) \cdot \overline{\Delta}(y) + \overline{\Delta}(x) \cdot (1 \otimes y + y \otimes 1) + (1 \otimes x + x \otimes 1) \cdot \overline{\Delta}(y) + x \otimes y + y \otimes x; \tag{2-4}$$

(ii) $\mathrm{Prim}(\mathcal{A}(m)) = \ker(\overline{\Delta})$;

(iii) $\overline{\Delta}(P^k \mathcal{A}_{\geq 1}(m)) \subset \sum_{i=1}^{k-1} P^i \mathcal{A}_{\geq 1}(m) \otimes P^{k-i} \mathcal{A}_{\geq 1}(m)$;

(iv) $P^k \mathcal{A}_{\geq_1}(m) \subset \ker(\overline{\Delta}^k)$;

(v) $F^k \mathcal{A}_{\geq_1}(m) \subset \ker(\overline{\Delta}^k)$, *i.e.*, $\overline{\Delta}^k$ *vanishes on forests of size* $\leq k$.

All of the above properties remain true over $\mathbb{Q}$, *and with* FI.

Proof Part (i) follows from the definition of $\overline{\Delta}$ and compatibility $\Delta(x \cdot y) = \Delta(x) \cdot \Delta(y)$. Part (ii) is immediate. Part (v) follows from the description (2-3), since for a forest F of size k the sum is empty. It remains to show (iii) and (iv).

We prove (iii) by induction on k. To lighten notation, we write P^i for $P^i \mathcal{A}_{\geq 1}(m)$. The base case, $k = 1$, is settled by (ii). Let $1 \leq k < n$ and assume that $\overline{\Delta}(P^k) \subset \sum_{i=1}^{k-1} P^i \otimes P^{k-i}$. The filtration stage P^{k+1} is generated by P^k together with all products of exactly $k+1$ primitive elements. Consider an element $z \in P^{k+1}$ of the latter sort, i.e., $z = xy$ for $x \in P^k$ and $y \in P^1 = \mathrm{Prim}(\mathcal{A}(m))$. Using $\overline{\Delta}(y) = 0$ (by (ii)), (2-4) becomes

$$\overline{\Delta}(z) = \overline{\Delta}(x) \cdot (1 \otimes y + y \otimes 1) + x \otimes y + y \otimes x.$$

Now, by induction hypothesis, we have $\overline{\Delta}(x) \in \sum_{i=1}^{k-1} P^i \otimes P^{k-i}$. Using $P^i \cdot y \subset P^{i+1}$, we finally obtain

$$\overline{\Delta}(z) \in \sum_{i=1}^{k-1} P^i \otimes P^{k+1-i} + \sum_{i=1}^{k-1} P^{i+1} \otimes P^{k-i} + P^k \otimes P^1 + P^1 \otimes P^k = \sum_{i=1}^{k} P^i \otimes P^{k+1-i},$$

which proves (iii).

The proof of (iv) also proceeds by induction on k, with the base case being (ii) again. Assume that $P^{k-1} \subset \ker(\overline{\Delta}^{k-1})$. Let $z \in P^k$. By (iii) we can write $\overline{\Delta}(z) = \sum_{i=1}^{k-1} \sum_j x_{ij} \otimes y_{ij}$ where $x_{ij} \in P^i$ and $y_{ij} \in P^{k-i}$. Then

$$\overline{\Delta}^k(z) := (\overline{\Delta}^{k-1} \otimes \mathrm{id}) \circ \overline{\Delta}(z) = \sum_{i=1}^{k-1} \sum_j \overline{\Delta}^{k-1}(x_{ij}) \otimes y_{ij} = 0$$

since all $\overline{\Delta}^{k-1}(x_{ij}) = 0$ by induction hypothesis. This concludes the proof of (iv). □

Theorem 2.2.2.3 *Over* $\mathbb{Q}$, *the reduced primitive filtration and the size filtration of* $\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}}$ *coincide. In particular*, $\mathrm{Prim}(\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}})$ *is the submodule of* $\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}}$ *generated by tree diagrams.*

When $m = 1$, *this already holds over* $\mathbb{Z}$.

Proof We omit FI in the notation since the proof for that case is the same. We write P^i for $P^i\mathcal{A}_{\geq 1}(m)$ and F^i for $F^i\mathcal{A}_{\geq 1}(m)$ to lighten notation.

In degree $n = 0$, there is nothing to prove. Fix a degree $n > 0$. First, any forest diagram of size 1 is primitive by definition of the comultiplication, i.e., $F^1\mathcal{A}_n(m)_{\mathbb{Q}} \subset P^1\mathcal{A}_n(m)_{\mathbb{Q}}$.

In fact, $F^1\mathcal{A}_n(m)_{\mathbb{Q}} \subset P^1\mathcal{A}_n(m)_{\mathbb{Q}}$ implies that $F^k\mathcal{A}_n(m)_{\mathbb{Q}} \subset P^k\mathcal{A}_n(m)_{\mathbb{Q}}$ for each $k \geq 1$. We show this by induction.[5] The base case is the above paragraph. Assume that $F^k \subset P^k$ and consider a forest diagram F of size $k+1$. Pick any tree T in F and apply a sequence $F = F_0 \rightsquigarrow \cdots \rightsquigarrow F_r$ of STU-relations to slide the legs of T above the rest of the trees. Thus,

$$F_r = T \cdot (F \setminus T) \subset P^1 \cdot F^k \subset P^1 \cdot P^k \subset P^{k+1}.$$

For each i, $F_i - F_{i+1}$ is a forest of size k by STU, i.e., $F_i - F_{i+1} \in F^k \subset P^k$. Therefore,

$$F = \sum_{i=0}^{r-1} (F_i - F_{i+1}) + F_r \subset P^{k+1}.$$

Consequently, the size filtration injects into the primitive filtration:

$$\begin{tikzcd}
F^1\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] \arrow[d, hook] & F^2\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] \arrow[d, hook] & \cdots \arrow[r, hook] & F^{n-1}\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] \arrow[d, hook] & F^n\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[d, equal] \\
P^1\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] & P^2\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] & \cdots \arrow[r, hook] & P^{n-1}\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] & P^n\mathcal{A}_n(m)_{\mathbb{Q}}
\end{tikzcd}$$

We show that the vertical arrows are all equalities, inductively from right to left. Let $1 \leq k < n$ and assume that $F^{k+1} = P^{k+1}$. We seek for a section to the inclusion i_n^k as in the diagram

$$\begin{tikzcd}
F^k\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook, shift right, "i_n^k"'] \arrow[d, hook] & F^{k+1}\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[l, two heads, bend right, shift right, "s_n^{k+1}"'] \arrow[d, equal] \\
P^k\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook] & P^{k+1}\mathcal{A}_n(m)_{\mathbb{Q}}
\end{tikzcd} \tag{2-5}$$

[5]This argument is similar to the proof of Proposition 6.10 in [22], which deals with forest claspers instead of forest diagrams.

Consider the linear map

$$s_n^{k+1} : F^{k+1}\mathcal{A}_n(m)_{\mathbb{Q}} \to \mathcal{A}_n(m)_{\mathbb{Q}}, \quad F \mapsto F - \frac{1}{(k+1)!}\mu^k \circ \bar{\Delta}^k(F), \tag{2-6}$$

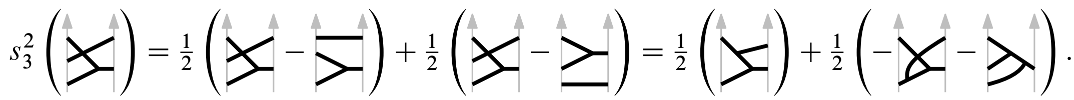
where $\mu^k : \mathcal{A}(m)^{\otimes k+1} \to \mathcal{A}(m)$ is the m-fold multiplication. We have:

- **The image of s_n^{k+1} lies in $F^k\mathcal{A}_n(m)_{\mathbb{Q}}$** Indeed, for a forest diagram F of size $k+1$, say $F = \bigcup_{i=1}^{k+1} T_i$, we have[6]

$$s_n^{k+1}(F) = \frac{1}{(k+1)!}\sum_{\sigma\in\mathfrak{S}_{k+1}} (F - (F)_\sigma), \tag{2-7}$$

where $(F)_\sigma \coloneqq T_{\sigma(1)} \dots T_{\sigma(k+1)}$ is obtained by stacking the trees of F in the order prescribed by σ. Each difference $(F - (F)_\sigma)$ can be obtained by a sequence of leg exchanges (seen later as a path in $\mathcal{F}(F)$; see Definition 2.3.3.1) from $(F)_\sigma$ to F, and hence we obtain $F - (F)_\sigma$ as a sum of forests of size k (each leg exchange yields a term where the two legs are attached together, by STU). Thus $s_n^{k+1}(F) \in F^k\mathcal{A}_n(m)_{\mathbb{Q}}$.

For example,



- **The map s_n^{k+1} is a section to the inclusion i_n^k** Indeed, for $F \in F^k\mathcal{A}_n(m)_{\mathbb{Q}}$, we have $\bar{\Delta}^k(F) = 0$ by Lemma 2.2.2.2(vi). Thus s_n^{k+1} restricts to the identity on $F^k\mathcal{A}_n(m)_{\mathbb{Q}}$.

By Lemma 2.2.2.2(v), s_n^{k+1} restricts to the identity on $P^k\mathcal{A}_n(m)_{\mathbb{Q}}$. Thus, starting with $x \in P^k\mathcal{A}_n(m)_{\mathbb{Q}}$ and going around the diagram (2-5) in the counterclockwise direction, we end up with $s_n^{k+1}(x) = x$, which shows that the vertical inclusion $F^k \hookrightarrow P^k$ is also surjective. This concludes the proof that the two filtrations coincide.

It follows that

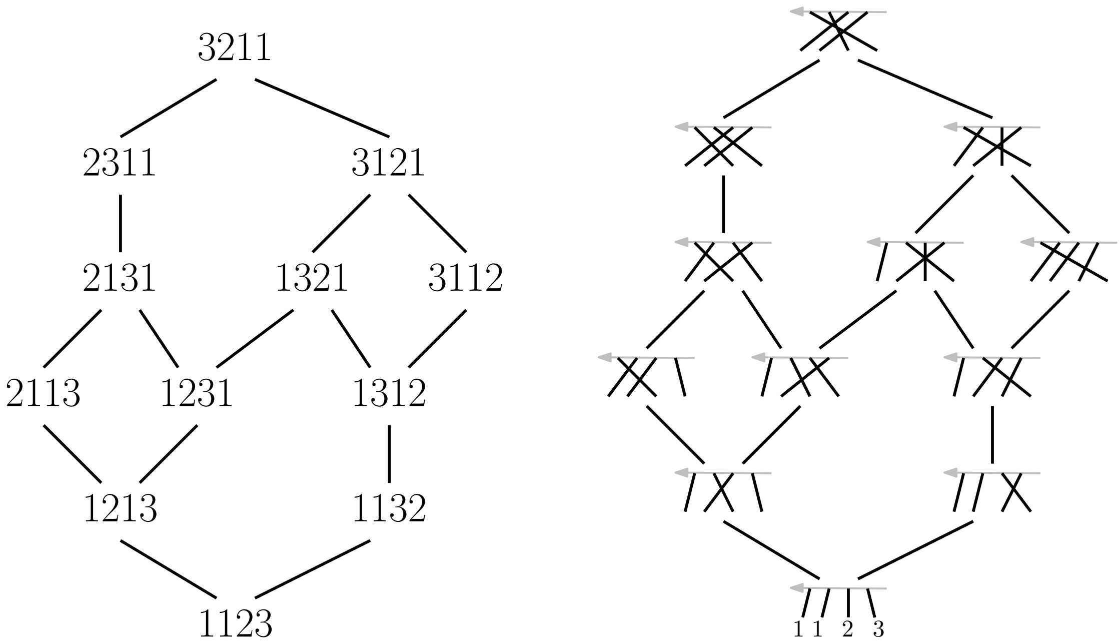
$$\mathrm{Prim}(\mathcal{A}(m)_{\mathbb{Q}}) = P^1\mathcal{A}_{\geq 1}(m)_{\mathbb{Q}} = F^1\mathcal{A}_{\geq 1}(m)_{\mathbb{Q}},$$

i.e., the primitive part of $\mathcal{A}(m)_{\mathbb{Q}}$ is additively generated by forests of size 1, that is, trees.

When $m = 1$, the bialgebra $\mathcal{A}(m)$ is commutative. Thus, if F is a forest of size $k+1$, then all $(F)_\sigma$, $\sigma \in \mathfrak{S}_{k+1}$, are equal and averaging is not needed. In particular, the section s_n^{k+1} defined in (2-6) is already defined over $\mathbb{Z}$. This yields an alternative proof of the fact that tree diagrams generate the primitive Lie algebra of $\mathcal{A}(1)$ over $\mathbb{Z}$. See [28] for a different argument. □

In order to better understand that filtration, and, in particular, the primitive Lie algebra $\mathrm{Prim}(\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}})$, a concrete description of the filtration steps, i.e., by generators and relations, is required. This is the topic of the next two sections.

[6] The choice of numbering of the trees in F is irrelevant, since the sum (2-7) is over all permutations.



(a) The permutograph $P_{(2,1,1)}$ (b) Graph of labelled legs on one strand

Figure 5: One slice of a graph of labelled forests.

2.3 Permutographs and graphs of forests

2.3.1 Permutographs For a tuple of nonnegative integers $\boldsymbol{n} := (n_1, \dots, n_s) \in \mathbb{N}^s$, consider the set $V_{\boldsymbol{n}}$ of all possible words made of n_i-many i, for each $i = 1, \dots, s$. In particular, we write

$$\underline{w}_{\boldsymbol{n}} := \underbrace{1 \dots 1}_{n_1} \dots \underbrace{s \dots s}_{n_s} \quad \text{and} \quad \overline{w}_{\boldsymbol{n}} := \underbrace{s \dots s}_{n_s} \dots \underbrace{1 \dots 1}_{n_1}.$$

The symmetric group $\mathfrak{S}_n$, where $n := \sum_i n_i$, acts on $V_{\boldsymbol{n}}$ by permutation, denoted by $\sigma \cdot w$ for $\sigma \in \mathfrak{S}_n$, $w \in V_{\boldsymbol{n}}$. The action is transitive but has many fixed points, unless all $n_i = 1$.

Definition 2.3.1.1 The $\boldsymbol{n}$*-permutograph* is the undirected graph $P_{\boldsymbol{n}} := (V_{\boldsymbol{n}}, E_{\boldsymbol{n}})$, whose set of vertices is $V_{\boldsymbol{n}}$ and two distinct vertices $v, w \in V_{\boldsymbol{n}}$ are connected by an undirected edge if there is a basic transposition $\tau_i = (i, i+1) \in \mathfrak{S}_{\boldsymbol{n}}$ such that $w = \tau \cdot v$ (hence also $v = \tau \cdot w$), i.e., $v = \dots jk \dots$ and $w = \dots kj \dots$ where $1 \le j \ne k \le s$ are the i-th and $(i+1)$-st entries of v, w.

There is at most one edge joining any pair of vertices and at most $n-1$ edges adjacent to a given vertex, since there are only $n-1$ basic transpositions $\tau_1, \dots, \tau_{n-1}$. Permutographs are connected, since the $\mathfrak{S}_{\boldsymbol{n}}$-action on $V_{\boldsymbol{n}}$ is transitive.

Remark 2.3.1.2 Permutographs are actually the 1-skeletons of permutohedra, which are rich combinatorial objects [27; 39; 44]. △

Proposition 2.3.1.3 *The first homology group of the permutograph P_n is generated by length* 4 *and* 6 *cycles. In other words, any* 1*-cycle in P_n is the boundary of a union of squares and hexagons.*

Proof To prove this, we scan P_n using the *height function*

$$h : V_n \to \mathbb{N} : v \mapsto \sum_{i=1}^{n} \#\{j > i \mid v_j < v_i\},$$

which assigns to a word v the number $h(v)$ of pairs of letters $\dots v_i \dots v_j \dots$ where $v_j < v_i$ appearing in it, i.e., the number of inversions in v. For example, $h(\underline{w}_n) = 0$ is minimal and $h(\overline{w}_n)$ is maximal in P_n. The function h is a Morse function in the sense that, for any edge $\{v, w\}$, we have $h(w) = h(v) \pm 1$. Pictorially, when drawing P_n with all vertices of height i at height i, no edge is horizontal (see Figure 5(a)).

We show that any 1-cycle C in P_n is a sum of length 4 and 6 cycles, by induction on $\max_{v \in C} h(v)$. Up to decomposing C into smaller subcycles and removing paths going back and forth, we can assume that $C = \overrightarrow{v^0 v^1} + \dots + \overrightarrow{v^{l-1} v^l}$ where $(v^0, \dots, v^l = v^0)$ is a sequence of adjacent vertices without repetitions (except $v^l = v^0$). If $\max_{v \in C} h(v) = 0$ then there is nothing to prove since $C = 0$. This proves the base case.

Consider a 1-cycle C with $M := \max_{v \in C} h(v) > 0$ and assume the statement to be true for all cycles of maximal height $< M$. Pick any vertex $v^M \in C$ that reaches the maximum M, then we must have $h(v^{M \pm 1}) = h(v^M) - 1$. By definition of the permutograph, $v^{M-1} = \tau_i \cdot v^M$ and $v^{M+1} = \tau_j \cdot v^M$ for some i, j such that $i \neq j$. We assume $i < j$ to simplify notation. There are two cases to consider:

(1) If $|j - i| > 1$, then τ_i and τ_j have disjoint support and the vertex $\widetilde{v}^M := \tau_i \tau_j \cdot v^M = \tau_i \cdot v^{M+1} = \tau_j \cdot v^{M-1}$ forms a square together with v^M, v^{M+1}, v^{M-1} (we use the relation $\tau_i \tau_j = \tau_j \tau_i$ in $\mathfrak{S}_n$). In particular,

$$\begin{aligned} C &= \overrightarrow{v^0 v^1} + \dots + \overrightarrow{v^{M-1} v^M} + \overrightarrow{v^M v^{M+1}} + \dots + \overrightarrow{v^{l-1} v^l} \\ &= \underbrace{(\overrightarrow{v^0 v^1} + \dots + \overrightarrow{v^{M-1} \widetilde{v}^M} + \overrightarrow{\widetilde{v}^M v^{M+1}} + \dots + \overrightarrow{v^{l-1} v^l})}_{=:C'} + c = C' + c, \end{aligned}$$

where $c := \overrightarrow{v^{M-1} v^M} + \overrightarrow{v^M v^{M+1}} - \overrightarrow{v^{M-1} \widetilde{v}^M} - \overrightarrow{\widetilde{v}^M v^{M+1}}$ is a length 4 cycle. Moreover, $\widetilde{v}^M$ has height $M - 2$ since it has exactly two inversions less than v^M.

(2) If $|j - i| = 1$, and hence $j = i + 1$, then the above does not work since the vertices $\tau_i \tau_{i+1} \cdot v^M$ and $\tau_{i+1} \tau_i \cdot v^M$ are distinct. In this case we need to dive one step lower in the permutograph and use instead the relation $\tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1}$ in $\mathfrak{S}_n$.

Denote the i-th, $(i+1)$-st and $(i+2)$-nd entries of v^M by c, b, a, respectively. These are the entries that are permuted by τ_i and τ_{i+1}. Given a permutation (x_1, x_2, x_3) of (c, b, a), denote by $v^M_{x_1 x_2 x_3}$ the word obtained from v^M by replacing the i-th, $(i+1)$-st and $(i+2)$-nd entries by x_1, x_2, x_3, respectively. In particular, $v^{M-1} = v^M_{bca}$, $v^M = v^M_{cba}$ and $v^{M+1} = v^M_{cab}$.

Define C' from C by replacing the vertex v^M with the vertices v^M_{bac}, v^M_{abc}, v^M_{acb}, in this order. Those new vertices have height $< M$.

Then $C = C' + c$ where c is a length 6 cycle going through the vertices v^M_{bca}, v^M_{cba}, v^M_{cab}, v^M_{acb}, v^M_{abc}, v^M_{bac}, v^M_{bca}, in this order.

See the hexagon at the bottom of Figure 5(a) for an example with $(c, b, a) = (3, 2, 1)$.

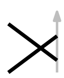
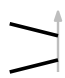
In both cases, $C = C' + c$ for a length 4 or 6 cycle c, and C' is a 1-cycle with one less vertex of height M. Repeating this process for each vertex of height M in C, we obtain $C = C'' + c_1 + \cdots + c_k$ where $c_1, \dots, c_k$ are length 4 or 6 cycles, and where C'' has no more vertex of height M. That is, C'' has maximal height $< M$, hence C'' is a sum of length 4 and 6 cycles by the induction hypothesis, so is C. □

Definition 2.3.1.4 For two graphs G_1, G_2, their *cartesian* or *box product*, denoted by $G_1 \square G_2$, is the 1-skeleton of $G_1 \times G_2$. The box product of graphs is associative.

Corollary 2.3.1.5 *The first homology of a box product of permutographs $\square_{j=1}^m P_{\boldsymbol{n}^j}$ is generated by length* 4 *and* 6 *cycles.*

Proof The exact same proof as Proposition 2.3.1.3 works, where now the two transpositions in case 1 can happen in different factors. Consequently, there are two types of length 4 cycles: product of edges from different factors and squares contained in one factor. □

2.3.2 Graph of labelled forests A *labelled forest diagram* is simply a forest diagram F in $\mathcal{D}(m)$ together with a labelling of its trees by $1, \dots, s$, where s is the size of the forest, i.e., a choice of identification $\pi_0(F) \cong \{1, \dots, s\}$. We say that two labelled forests F, F' are related by a *slide move* if F' is obtained from F by permuting adjacent legs belonging to distinct trees, that is,

$F =$  $\xleftrightarrow{\text{slide move}}$ $F' =$  ,

where it is understood, as usual, that F and F' are identical outside of the part shown.

Definition 2.3.2.1 Given s-many tree diagrams $T_1, \dots, T_s \in \mathcal{D}^T(m)$ on m strands ($s > 0$), the *graph of labelled forests on* $T_1, \dots, T_s$ is the undirected graph whose set of vertices is the set

$$\widetilde{\mathcal{F}}(T_1, \dots, T_s)$$

of all labelled forests on m strands, whose i-th tree is T_i, and where two labelled forests are connected by an edge if they are related by a slide move.

A part of a graph of labelled forests, corresponding to what happens on one strand, is shown on Figure 5(b). On this figure, the two leftmost legs (e.g., in the bottom diagram) belong to T_1, while the next ones belong to T_2 and T_3, respectively.[7] The legs from T_1 are not allowed to slide across one another, but they can slide across the two other legs. Note that any two vertices in $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$ are connected by *at most* one edge. There are no edges when $s = 1$.

[7] For simplicity, we have labelled only the forest at the bottom, but all forests on the figure should be labelled accordingly.

Proposition 2.3.2.2 *Suppose that the tree T_i has n_i^j legs on the j-th strand, for $1 \le i \le s$ and $1 \le j \le m$. Then there is a graph isomorphism defined on vertices as*[8]

$$W : \widetilde{\mathcal{F}}(T_1, \dots, T_s) \xrightarrow{\cong} \square_{j=1}^m P_{\boldsymbol{n}^j}, \tag{2-8}$$

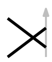
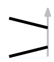
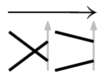
where $\boldsymbol{n}^j = (n_1^j, \dots, n_s^j)$, sending a labelled forest to $(w^j)_j$ where w^j encodes the ordering of the legs on the j-strand, from top to bottom.

Proof This is immediate, since a labelled forest is entirely determined by its trees $T_1, \dots, T_s$ and the relative positions of their legs on the strands. $\square$

Corollary 2.3.2.3 *The first homology of the graph of labelled forests $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$ is generated by length 4 and 6 cycles. In other words, any 1-cycle in it is the boundary of a union of squares and hexagons.*

Proof This follows immediately from Proposition 2.3.2.2 and Corollary 2.3.1.5. $\square$

Definition 2.3.2.4 To a directed edge $\overrightarrow{FF'}$ in $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$, $s \ge 2$, one can assign the third term appearing in the STU relation, which we denote by $\overrightarrow{FF'} \in \mathcal{D}^{s-1}(m)$:

If $F =$  and $F' =$ , then $\overrightarrow{FF'} =$  $:=$ .

In particular, $\overrightarrow{FF'} = F' - F$ in $\mathcal{A}(m)$.

The fact that the legs involved in a slide move belong to distinct trees is important here, otherwise $\overrightarrow{FF'}$ would have a cycle and would not be a forest diagram.

There are many paths joining forests $F, F' \in \widetilde{\mathcal{F}}(T_1, \dots, T_s)$. Hence, $\overrightarrow{FF'}$ is not well defined for nonadjacent forests. However, any two paths differ by a cycle in $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$, which by Corollary 2.3.2.3 can be written as a sum of back-and-forth paths, length 4 and 6 cycles. Thus, modding out the target $\mathbb{Z}\mathcal{D}^{s-1}(m)$ by those cycles will give us a well-defined map $\overrightarrow{\cdot}$.

Definition 2.3.2.5 A $\square$ *relation* in $\mathbb{Z}\mathcal{D}^{s-1}(m)$ is a relation of the form

$$\overrightarrow{F_1F_2} + \overrightarrow{F_2F_3} + \overrightarrow{F_3F_4} + \overrightarrow{F_4F_1}, \tag{2-9}$$

where (F_1, F_2, F_3, F_4) is a length 4 cycle in $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$ for some trees T_i. See Figure 6(a).

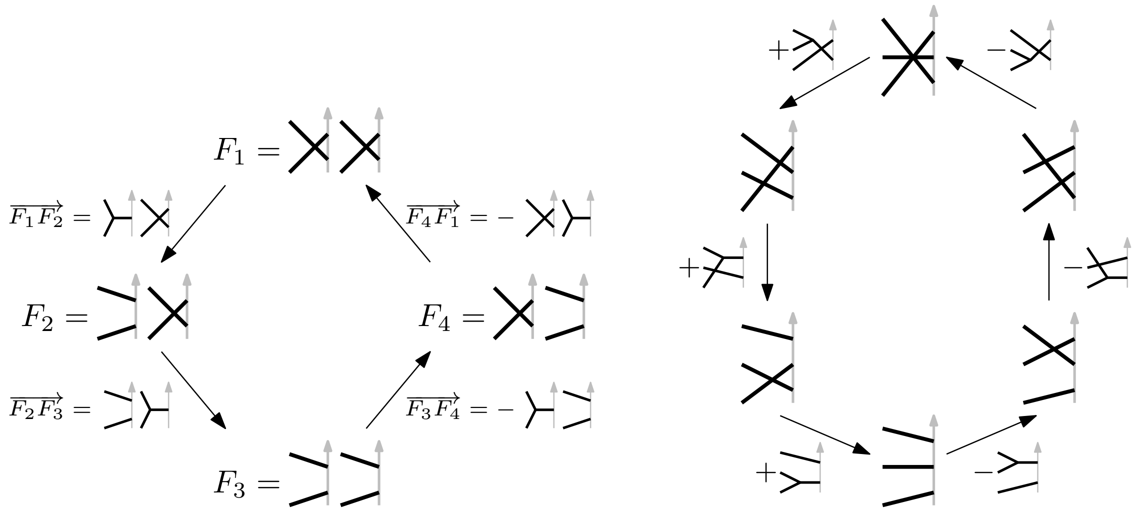
Similarly, a $\hexagon$ *relation* in $\mathbb{Z}\mathcal{D}^{s-1}(m)$ is a relation of the form

$$\overrightarrow{F_1F_2} + \overrightarrow{F_2F_3} + \overrightarrow{F_3F_4} + \overrightarrow{F_4F_5} + \overrightarrow{F_5F_6} + \overrightarrow{F_6F_1}, \tag{2-10}$$

where $(F_1, \dots, F_6)$ is a length 6 cycle in $\widetilde{\mathcal{F}}(T_1, \dots, T_s)$ for some trees T_i. See Figure 6(b).

Let $\square R$, resp. $\hexagon R$, denote the submodule of $\mathbb{Z}\mathcal{D}(m)$ generated by all $\square$, resp. $\hexagon$, relations.

Remark 2.3.2.6 A $\hexagon$ relation can be obtained in the following way: start with an IHX relation that involves a leg, and break apart that leg in the three terms. The resulting signed sum of six terms is a $\hexagon$ relation, and all $\hexagon$ relations can be obtained in this way. $\triangle$



(a) A □ relation (b) A ⬡ relation

Figure 6: Relations obtained from cycles of length 4 and 6.

By the above definition and Corollary 2.3.2.3, the following definition is well defined (AS takes care of the back-and-forth paths).

Definition 2.3.2.7 For two labelled forests $F, F' \in \widetilde{\mathcal{F}}(T_1,\dots,T_s)$, $s \geq 2$, define

$$\overrightarrow{FF'} := \overrightarrow{FF_1} + \cdots + \overrightarrow{F_{l-1}F'} \in \mathbb{Z}\mathcal{D}^{s-1}(m)/\langle \mathrm{AS}, \mathrm{STU}^2, \hexagon R\rangle, \tag{2-11}$$

where $F \rightsquigarrow F_1 \rightsquigarrow \cdots \rightsquigarrow F_{l-1} \rightsquigarrow F'$ is a path from F to F' in $\widetilde{\mathcal{F}}(T_1,\dots,T_s)$.

Proposition 2.3.2.8 *Square relations and* STU^2 *relations coincide, i.e.,* $\square R = \mathrm{STU}^2$ *as submodules of* $\mathbb{Z}\mathcal{D}^s(m)$.

Proof Recall [9, Definition 3.1] that STU^2 relations among size s forests are defined starting with a *template*, i.e., a size $s-1$ forest diagram or a size s Feynman diagram with a single cycle, then choosing two legs (which must be adjacent to the cycle in the latter case) and breaking one or the other apart using STU. Breaking apart both legs yields four possible size $s+1$ forest diagrams, which form the corners of a square in a graph of labelled forests and the corresponding □ relation coincides with the STU^2 relation.

Conversely, consider a square as on Figure 6(a) and the associated □ relation. Then the template yields a STU^2 relation which coincides with that □ relation. □

2.3.3 Graph of forests In the previous section we investigated the graph of labelled forests and the relations that are required to define differences between labelled forests that are joined by a path. However, forests diagrams in the bialgebra $\mathcal{A}(m)$ are not labelled. The *true* graph of forests is a quotient of the graph of labelled forests, where the quotient map forgets the labelling.

[8] Since there is at most one edge connecting two vertices, this determines the morphism on edges as well.

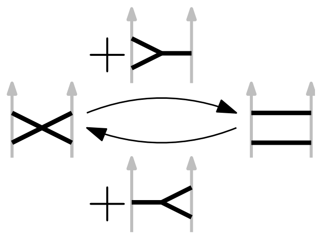


Figure 7: Two different leg moves joining two forests.

Definition 2.3.3.1 The *graph of forests* on trees $T_1, \dots, T_s$ is the graph $\mathcal{F}(T_1, \dots, T_s)$ whose vertices are all forest diagrams in $\mathcal{D}^s(m)$ consisting of those trees, and to each slide move relating two forests there is an associated edge joining the corresponding vertices.

Note that $\mathcal{F}(T_1, \dots, T_s)$ does not depend on the ordering of the trees T_i. If $F = \bigcup_{i=1}^s T_i$ is a forest diagram consisting of trees $T_1, \dots, T_s$, we write $\mathcal{F}(F) := \mathcal{F}(T_1, \dots, T_s)$.

There is a natural map

$$\widetilde{\mathcal{F}}(T_1, \dots, T_s) \twoheadrightarrow \mathcal{F}(T_1, \dots, T_s) \tag{2-12}$$

that forgets the labelling of forests. It is surjective but not injective if there are identical trees among $T_1, \dots, T_s$.

Remark 2.3.3.2 In $\mathcal{F}(T_1, \dots, T_s)$, it can happen that two vertices are joined by multiple edges. See Figure 7 for an example: the top, resp. bottom, edge corresponds to exchanging the legs on the right, resp. left, strand.

Thus, the notation $\overrightarrow{FF'}$ is ambiguous. Given a directed edge $F \overset{e}{\rightsquigarrow} F'$ we write $\overrightarrow{e}$ to denote the third term of the corresponding STU relation, i.e., $\overrightarrow{e} = F' - F$ in $\mathcal{A}(m)$. Similarly, if P is a path consisting of directed edges $e_1, \dots, e_l$, then define $\overrightarrow{P} := \sum_{i=1}^{l} \overrightarrow{e_i}$. △

Lemma 2.3.3.3 *For any labelled forest diagram F, the forgetful map $\widetilde{\mathcal{F}}(F) \twoheadrightarrow \mathcal{F}(F)$ satisfies the following* **path-lifting property**: *For any path $P = F_0 \overset{e_0}{\rightsquigarrow} F_1 \overset{e_1}{\rightsquigarrow} \cdots \overset{e_{l-1}}{\rightsquigarrow} F_l$ in $\mathcal{F}(F)$ and any choice of labelled forest $\widetilde{F}_0$ lifting F_0, there exists a unique path $\widetilde{P} = \widetilde{F}_0 \overset{\widetilde{e}_0}{\rightsquigarrow} \widetilde{F}_1 \overset{\widetilde{e}_1}{\rightsquigarrow} \cdots \overset{\widetilde{e}_{l-1}}{\rightsquigarrow} \widetilde{F}_l$ in $\widetilde{\mathcal{F}}(F)$ that lifts P. In other words, the surjection $\widetilde{\mathcal{F}}(F) \twoheadrightarrow \mathcal{F}(F)$ is a finite covering.*

Moreover, $\overrightarrow{\widetilde{e}_i} = \overrightarrow{e_i}$ for each $i = 0, \dots, l-1$.

Proof Starting with $\widetilde{F}_0$ and applying the sequence of slide moves prescribed by P yields such a lift $\widetilde{P}$, which is uniquely determined by this process. □

Proposition 2.3.3.4 *Any cycle in $\mathcal{F}(T_1, \dots, T_s)$ has a multiple that lifts along the covering*

$$\widetilde{\mathcal{F}}(T_1, \dots, T_s) \twoheadrightarrow \mathcal{F}(T_1, \dots, T_s).$$

In particular, over $\mathbb{Q}$, for any C, the element $\overrightarrow{C}$ belongs to the submodule $\langle \mathrm{AS}, \square R, \bigcirc R \rangle \subset \mathbb{Q}\mathcal{D}^{s-1}(m)$.

Proof Let G be the group of deck transformations of the finite covering $p:\widetilde{\mathcal{F}}(T_1,\dots,T_s)\twoheadrightarrow\mathcal{F}(T_1,\dots,T_s)$, and $|G|$ its size. Consider the transfer homomorphism $\tau: H_1(\mathcal{F}(T_1,\dots,T_s))\to H_1(\widetilde{\mathcal{F}}(T_1,\dots,T_s))$. Then $p_*\circ\tau$ equals multiplication by $|G|$. In particular, the cokernel of p_* is $|G|$-torsion and p_* becomes surjective when passing to coefficients in $\mathbb{Q}$. This concludes.

The second statement then follows from the first and Corollary 2.3.2.3. □

With the above result in hands, we can adapt Definition 2.3.2.7 to the graph of forests, but with rational coefficients. See Remark 2.3.3.2 for the notation $\overrightarrow{e}$ and $\overrightarrow{P}$.

Definition 2.3.3.5 For two forests $F, F'\in\mathcal{F}(T_1,\dots,T_s)$, define

$$\overrightarrow{FF'} := \overrightarrow{P}\in\mathbb{Q}\mathcal{D}^{s-1}(m)/\langle \mathrm{AS},\mathrm{STU}^2,\bigcirc R\rangle, \tag{2-13}$$

where P is a path from F to F' in $\mathcal{F}(T_1,\dots,T_s)$.

2.4 Presentation of the primitive filtration of $\mathcal{A}(m)_{\mathbb{Q}}$

2.4.1 Splitting trees

Definition 2.4.1.1 For $1\le s\le n-1$, the $\mathbb{Q}$*-module of size s forest diagrams* $\mathcal{F}^s(m)_{\mathbb{Q}}$ is defined by

$$\mathcal{F}^s(m)_{\mathbb{Q}} := \mathbb{Q}\mathcal{D}^s(m)/\langle \mathrm{AS},\mathrm{IHX},\mathrm{STU}^2,\bigcirc R\rangle.$$

It is graded by the degree of diagrams, with degree-n graded part denoted by $\mathcal{F}^s_n(m)_{\mathbb{Q}}$. The *framing independent version* is defined as $\mathcal{F}^{\mathrm{FI},s}(m)_{\mathbb{Q}} := \mathcal{F}^s(m)_{\mathbb{Q}}/\langle 1\mathrm{T}\rangle$. We use the notation (FI) as explained in Section 2.1.3.

Remark 2.4.1.2 When $s=1$, the submodule $\bigcirc R$ is actually trivial, since a $\bigcirc$ relation requires at least two trees. Thus, $\mathcal{F}^1(m)_{\mathbb{Q}} = \mathcal{L}(m)_{\mathbb{Q}}$ as $\mathbb{Q}$-modules (see Theorem 2.4.5.1 for the definition of $\mathcal{L}(m)_{\mathbb{Q}}$).

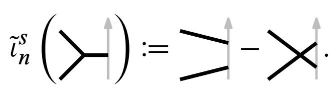
In degree n and with $s=n$, STU^2 becomes 4T while $\bigcirc R$ is trivial; hence we have $\mathcal{F}^n_n(m)_{\mathbb{Q}}\cong\mathcal{A}_n(m)_{\mathbb{Q}}$.

If $s<n-1$, the $\bigcirc$ relations are actually implied by STU^2 and IHX. Indeed, by Remark 2.3.2.6 any $\bigcirc$ relation can be obtained by breaking apart a leg in an IHX relation. Since $s<n-1$, the terms of the $\bigcirc$ relation must contain a node not involved in the relation. Apply STU^2 to break that additional node and rebuild a sum of two IHX relations. Thus $\bigcirc R\subset\langle \mathrm{STU}^2,\mathrm{IHX}\rangle$ as submodules of $\mathbb{Q}\mathcal{D}^s_n(m)$ when $s<n-1$. △

For each $1\le s<n$, there is a linear map

$$\tilde{\iota}^s_n:\mathbb{Q}\mathcal{D}^s_n(m)\to\mathcal{F}^{s+1}_n(m)_{\mathbb{Q}} \tag{2-14}$$

defined as follows: given a size s forest diagram $F\in\mathcal{D}^s_n(m)$, pick a trivalent vertex (which must exist since $s<n$) adjacent to a strand and break it apart according to STU, i.e.,

(2-15)


By STU^2, this does not depend on the choice of trivalent vertex to break apart.

Proposition 2.4.1.3 *The linear map* $\tilde{\iota}_n^s := \mathbb{Q}\mathcal{D}_n^s(m) \to \mathcal{F}_n^{(\mathrm{FI}),s+1}(m)_{\mathbb{Q}}$ *factors through the linear map* $\iota_n^s : \mathcal{F}_n^{(\mathrm{FI}),s}(m)_{\mathbb{Q}} \to \mathcal{F}_n^{(\mathrm{FI}),s+1}(m)_{\mathbb{Q}}$.

Proof This proof is very similar to the proof of Claim 3.4 in [9].

We need to show that $\tilde{\iota}_n^s$ vanishes on AS, IHX, STU^2, $\hexagon R$. If $s < n-2$, then forest diagrams in $\mathcal{D}_n^s(m)$ have at least three nodes, while the relations AS, IHX, STU^2, $\hexagon R$ involve at most 2 nodes. Hence it suffices to pick another node to break apart and the image under $\tilde{\iota}_n^s$ is a difference of sums that belong to the respective submodules.

For $s < n-1$, the above argument also applies to STU^2, $\hexagon R$, AS. For $s = n-2$ and IHX, observe that an IHX relation is sent to a $\hexagon$ relation.

For $s = n-1$, there are no IHX relations, while a direct calculation shows that STU^2 and AS relations are sent to zero.

This concludes the proof that $\tilde{\iota}_n^s$ factors through $\mathcal{F}_n^s(m)_{\mathbb{Q}}$. The FI case holds as well: if D has an isolated chord, then so do the two terms in $\tilde{\iota}_n^s(D)$. □

Remark 2.4.1.4 In [9, Claim 3.4], the proof restricts to the case $s \neq n-1$ because the $\hexagon$ relation is not taken into account there. △

The linear maps ι_n^s, $s = 1, \dots, n-1$, provide a factorisation of the map $\iota_n : \mathcal{F}_n^1(m)_{\mathbb{Q}} \to \mathcal{A}_n(m)_{\mathbb{Q}}$,

$$\begin{tikzcd} \mathcal{F}_n^1(m)_{\mathbb{Q}} \arrow[r, "\iota_n^1"] \arrow[rrrr, bend right, "\iota_n"'] & \mathcal{F}_n^2(m)_{\mathbb{Q}} \arrow[r, "\iota_n^2"] & \cdots \arrow[r, "\iota_s^{n-1}"] & \mathcal{F}_n^n(m)_{\mathbb{Q}} = \mathcal{A}_n(m)_{\mathbb{Q}}, \end{tikzcd}$$

where nodes are broken apart one after the other, instead of all at once

2.4.2 Barycenters of forests

Let $T_1, \dots, T_s \in \mathcal{D}^T(m)$ be tree diagrams. Definition 2.3.3.5 defines a map

(2-16) $$\vec{\cdot} : \mathcal{F}(T_1, \dots, T_s) \times \mathcal{F}(T_1, \dots, T_s) \to \mathcal{F}^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}} : (F, F') \mapsto \overrightarrow{FF'}.$$

The vector notation is justified by the fact that $\mathcal{F}(T_1, \dots, T_s)$ inherits some properties of affine spaces. In particular, we can still talk about barycenters of vertices of $\mathcal{F}(T_1, \dots, T_s)$.

Definition 2.4.2.1 The *barycenter* of forests $F_1, \dots, F_k \in \mathcal{F}(T_1, \dots, T_s)$ with coefficients $\lambda_1, \dots, \lambda_k \in \mathbb{Q}$ summing to $\sum_i \lambda_i = 1$ is the formal sum $\sum_i \lambda_i F_i$.

Denote by $\mathrm{Bar}(\mathcal{F}(T_1, \dots, T_s))$ the set of all barycenters of forests in $\mathcal{F}(T_1, \dots, T_s)$. There is a natural inclusion $\mathcal{F}(T_1, \dots, T_s) \hookrightarrow \mathrm{Bar}(\mathcal{F}(T_1, \dots, T_s))$ sending a forest F to $1 \cdot F$.

Lemma 2.4.2.2 *We have the following properties*:

(i) *If* $F_0, F_1, F_2 \in \mathcal{F}(T_1, \dots, T_s)$, *then* $\overrightarrow{F_0F_2} = \overrightarrow{F_0F_1} + \overrightarrow{F_1F_2}$ *(Chasles relation).*

(ii) *Let* $F_0, F_0', F_1, \dots, F_k \in \mathcal{F}(T_1, \dots, T_s)$ *and* $\lambda_1, \dots, \lambda_k \in \mathbb{Q}$ *such that* $\sum_i \lambda_i = 0$. *Then* $\sum_i \lambda_i \overrightarrow{F_0F_i} = \sum_i \lambda_i \overrightarrow{F_0'F_i}$ *in* $\mathcal{F}^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}}$.

(iii) *The map*

$$\text{(2-17)}\qquad \begin{gathered}\overrightarrow{\cdot\,\cdot} : \mathrm{Bar}(\mathcal{F}(T_1,\dots,T_s)) \times \mathrm{Bar}(\mathcal{F}(T_1,\dots,T_s)) \to \mathcal{F}^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}},\\ \Bigl(B=\sum_i \lambda_i F_i,\ B'=\sum_j \lambda'_j F'_j\Bigr) \mapsto \sum_j \lambda'_j \overrightarrow{F_0F'_j} - \sum_i \lambda_i \overrightarrow{F_0F_i},\end{gathered}$$

where $F_0 \in \mathcal{F}(T_1,\dots,T_s)$ is any forest, does not depend on the choice of F_0, extends the map (2-16) *and satisfies the Chasles relation.*

Proof Part (i) follows by choosing a path from F_0 to F_2 passing through F_1. Part (ii) follows directly from (i) and implies that (2-17) does not depend on the choice of F_0. The rest is direct. □

By Lemma 2.4.2.2, any choice of barycenter $B_0 \in \mathrm{Bar}(\mathcal{F}(T_1,\dots,T_s))$ defines a map

$$\mathcal{F}(T_1,\dots,T_s) \to \mathcal{F}^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}} : F \mapsto \overrightarrow{B_0F},$$

which sends a forest of size s to a linear combination of forests of size $s-1$. In order to obtain sections to the maps ι_n^s, we need to consistently pick barycenters in each graph of forest.

Definition 2.4.2.3 For F a forest of size s consisting of trees $T_1,\dots,T_s$, define the *average barycenter*, or *average basepoint*, associated to F as

$$\text{(2-18)}\qquad (F)_{\mathrm{avg}} := \frac{1}{s!}\sum_{\sigma\in\mathfrak{S}_s} T_{\sigma(1)}\dots T_{\sigma(s)},$$

i.e., the average of all possible ways of stacking the trees of F one above the other.

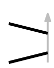
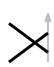
Note that $(F)_{\mathrm{avg}} = (F')_{\mathrm{avg}}$ whenever F and F' consist of the same trees, i.e., they belong to a common graph of forests.

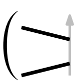
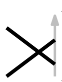
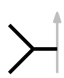
2.4.3 Merging trees

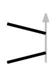
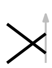
Consider the linear map

$$\text{(2-19)}\qquad \widetilde{\pi}_n^s : \mathbb{Q}\mathcal{D}_n^s(m) \to \mathcal{F}_n^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}} : F \mapsto \overrightarrow{(F)_{\mathrm{avg}}F},$$

which is well defined by Lemma 2.4.2.2.

Lemma 2.4.3.1 (diagrammatic STU) *For any adjacent forests $F^{=} =$*  *and $F^{\times} =$*  *in a graph of forests, we have*

$$\widetilde{\pi}_n^s\Bigl($$  $$-$$  $$\Bigr) =$$  ,

where  $-$  $+$  *is an* STU *relation.*

Proof Since $F^{=}$ and $F^{\times}$ belong to the same graph of forests, we have $(F^{=})_{\mathrm{avg}} = (F^{\times})_{\mathrm{avg}}$ and

$$\widetilde{\pi}_n^s(F^{=}-F^{\times}) = \widetilde{\pi}_n^s(F^{=}) - \widetilde{\pi}_n^s(F^{\times}) = \overrightarrow{(F^{=})_{\mathrm{avg}}F^{=}} - \overrightarrow{(F^{\times})_{\mathrm{avg}}F^{\times}} = \overrightarrow{F^{\times}F^{=}} = F^{Y}$$

by the Chasles relation. □

Proposition 2.4.3.2 *The map $\widetilde{\pi}_n^s : \mathbb{Q}\mathcal{D}_n^s(m) \to \mathcal{F}_n^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}}$ factors through $\pi_n^s : \mathcal{F}_n^{(\mathrm{FI}),s}(m)_{\mathbb{Q}} \to \mathcal{F}_n^{(\mathrm{FI}),s-1}(m)_{\mathbb{Q}}$.*

Proof We need to show that $\widetilde{\pi}_n^s$ vanishes on (1T), AS, IHX, STU^2 and $\bigcirc R$. The argument is similar to [9, Claim 3.6].

Start with an AS relation $F + F'$ in $\mathcal{D}_n^s(m)$. We can write $(F^{(')})_{\mathrm{avg}} = \sum_\sigma \lambda_\sigma F_\sigma^{(')}$ where $F_\sigma + F'_\sigma$ are AS relations (the notation $(')$ indicates that we can read it with or without $'$). Pick a path $F_\sigma \overset{e_0}{\rightsquigarrow} \cdots \overset{e_{l-1}}{\rightsquigarrow} F$ of leg slides in $\mathcal{F}(F)$. Then we can follow the parallel path

$$F'_\sigma \overset{e'_0}{\rightsquigarrow} \cdots \overset{e'_{l-1}}{\rightsquigarrow} F'$$

in $\mathcal{F}(F')$, obtained by performing the exact same leg slides. Notice that for each j, the origin and extremities of e_j and e'_j only differ by the cyclic ordering of a node, in the same trees as for F and F'. Consequently, $\overrightarrow{e_j} + \overrightarrow{e'_j}$ is an AS relation for each j and $\overrightarrow{F_\sigma F} + \overrightarrow{F'_\sigma F'}$ is a sum of those. In particular,

$$\overrightarrow{(F)_{\mathrm{avg}} F} + \overrightarrow{(F')_{\mathrm{avg}} F'} = \sum_\sigma \lambda_\sigma (\overrightarrow{F_\sigma F} + \overrightarrow{F'_\sigma F'}) = 0.$$

For an IHX relation $F^I + F^H + F^X$, the argument is the same. As above, picking parallel paths from $F_\sigma^\bullet$ to $F^\bullet$, for $\bullet \in \{I, H, X\}$, we obtain

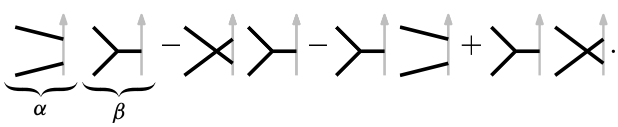
$$\overrightarrow{(F^I)_{\mathrm{avg}} F^I} + \overrightarrow{(F^H)_{\mathrm{avg}} F^H} + \overrightarrow{(F^X)_{\mathrm{avg}} F^X} = 0$$

since it is a sum of IHX relations.

Unlike AS and IHX which happen far from the strands, STU^2 happens close to the strands and we need to distinguish two cases. If the STU^2 relation $F^{Y=} - F^{Y\times} - F^{=Y} + F^{\times Y}$ involves two or three trees, then the legs forming $=$ or $\times$ must belong to distinct trees.[9] Then, two applications of the diagrammatic STU (Lemma 2.4.3.1) yield

$$\widetilde{\pi}_n^s(F^{Y=} - F^{Y\times}) = F^{YY} = \widetilde{\pi}_n^s(F^{=Y} - F^{\times Y}),$$

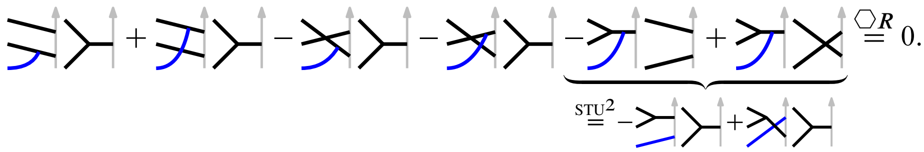
and hence the STU^2 relation is satisfied by $\widetilde{\pi}_n^s$ in this case. Let us turn to the second case, where the STU^2 relation involves a single tree, i.e., all the legs involved belong to the same tree. We argue as for the AS and IHX relations, by picking four parallel paths of slide moves from $F_\sigma^\bullet$ to $F^\bullet$, for $\bullet \in \{Y=, Y\times, =Y, \times Y\}$. More precisely, call the places involved in the STU^2 relation α and β,



A leg sliding across site α induces one or two edges in the four parallel paths, depending on whether site β contains a "Y", "$=$" or "$\times$" configuration. The corresponding terms in $\sum_\bullet \overrightarrow{F_\sigma^\bullet F^\bullet}$ add up to zero

[9] Otherwise, joining them in the term with Y would form a cycle.

by STU^2 and $\hexagon R$:



When a leg slides across site β, the argument is the same. Thus $\widetilde{\pi}_n^s$ satisfies STU^2.

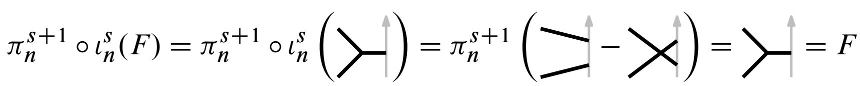
It remains to show that $\hexagon$-relations are satisfied. This immediately follows from three applications of the diagrammatic STU lemma (Lemma 2.4.3.1) and the IHX relation.

This concludes the proof that $\widetilde{\pi}_n^s$ factors through $\mathcal{F}_n^s(m)_{\mathbb{Q}}$.

In the FI case, we need to check that $\widetilde{\pi}_n^s$ vanishes on 1T. Let $F \in \mathcal{D}_n^s(m)$ be a forest diagram with an isolated chord. For $\sigma \in \mathfrak{S}_s$, the element $\overrightarrow{(F)_\sigma F}$ is a sum of diagrams in which this isolated chord remains, together with diagrams coming from sliding that isolated chord across another leg. Those terms of the latter type vanish in pairs by AS. This also follows from the commutativity property, shown in the proof of [1, Lemma 3.1]. □

Proposition 2.4.3.3 *For each $1 \le s < n$, the map π_n^{s+1} is a section of ι_n^s, i.e., $\pi_n^{s+1} \circ \iota_n^s = \mathrm{id}_{\mathcal{F}_n^{(\mathrm{FI}),s}(m)_{\mathbb{Q}}}$. In particular, ι_n^s is injective.*

Proof Given a size s and degree-n forest diagram $F =$  (represented by one of its nodes adjacent to a strand, which exists since $s < n$), we have



by the diagrammatic STU (Lemma 2.4.3.1). □

2.4.4 From $\mathcal{F}^k(m)_{\mathbb{Q}}$ to $F^k\mathcal{A}(m)_{\mathbb{Q}}$

Combining the results of Sections 2.2 and 2.4.3, we obtain the following theorem.

Theorem 2.4.4.1 *For each $1 \le k \le n$, there is a natural linear isomorphism*

$$\mathcal{F}_n^{(\mathrm{FI}),k}(m)_{\mathbb{Q}} \xrightarrow{\cong} F^k\mathcal{A}_n^{(\mathrm{FI})}(m)_{\mathbb{Q}}.$$

More precisely, we have the isomorphism of filtrations (with FI *as well)*

$$\begin{tikzcd} \mathcal{F}_n^1(m)_{\mathbb{Q}} \arrow[r, hook, "\iota_n^1"] \arrow[d, dashed, "\cong"] & \mathcal{F}_n^2(m)_{\mathbb{Q}} \arrow[r, hook, "\iota_n^2"] \arrow[d, dashed, "\cong"] & \cdots \arrow[r, hook, "\iota_n^{n-2}"] & \mathcal{F}_n^{n-1}(m)_{\mathbb{Q}} \arrow[r, hook, "\iota_n^{n-2}"] \arrow[d, dashed, "\cong"] & \mathcal{F}_n^n(m)_{\mathbb{Q}} = \mathcal{A}_n(m)_{\mathbb{Q}} \arrow[d, equal] \\ F^1\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook, "i_n^1"'] & F^2\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook, "i_n^2"'] & \cdots \arrow[r, hook, "i_n^{n-2}"'] & F^{n-1}\mathcal{A}_n(m)_{\mathbb{Q}} \arrow[r, hook, "i_n^{n-1}"'] & F^n\mathcal{A}_n(m)_{\mathbb{Q}} = \mathcal{A}_n(m)_{\mathbb{Q}} \end{tikzcd}$$

under which the sections s_n^k and π_n^k coincide.

Proof We omit FI. By definition, $F^k\mathcal{A}_n(m)_{\mathbb{Q}}$ is generated by size k forests in $\mathcal{A}_n(m)_{\mathbb{Q}}$, i.e., there is a linear surjection $\mathbb{Q}\mathcal{D}^k(m) \twoheadrightarrow F^k\mathcal{A}_n(m)_{\mathbb{Q}}$. The relations AS, IHX, STU^2, $\bigcirc R$ in $\mathbb{Q}\mathcal{D}^k(m)$ all hold in $\mathcal{A}_n(m)_{\mathbb{Q}}$, and hence this surjection descends to a surjection

$$\mathcal{F}_n^k(m)_{\mathbb{Q}} \twoheadrightarrow F^k\mathcal{A}_n(m)_{\mathbb{Q}}.$$

Those surjections, for $1 \le k \le n$, combine into the vertical maps in the commutative diagram of the statement and commute with the maps i_n^k and ι_n^k. By Proposition 2.4.3.3, the maps ι are injective, and hence so are the vertical maps.

To see that the sections s_n^k and π_n^k coincide, compare (2-7) with (2-19). □

2.4.5 The Lie algebra of trees We directly obtain, from the presentation of the primitive filtration of $\mathcal{A}_n(m)_{\mathbb{Q}}$ in Theorem 2.4.4.1, a concrete presentation of the primitive Lie algebra $\mathrm{Prim}(\mathcal{A}(m)_{\mathbb{Q}})$.

Theorem 2.4.5.1 *The primitive Lie algebra* $\mathrm{Prim}(\mathcal{A}^{(\mathrm{FI})}(m)_{\mathbb{Q}})$ *is naturally isomorphic to the* ***Lie algebra of trees***

$$\mathcal{L}^{(\mathrm{FI})}(m)_{\mathbb{Q}} := \mathcal{F}^{(\mathrm{FI}),1}(m)_{\mathbb{Q}} = \mathbb{Q}\mathcal{D}^T(m)/\langle (1\mathrm{T}), \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle$$

endowed with the bracket

$$[\cdot,\cdot] : \mathcal{L}^{(\mathrm{FI})}(m)_{\mathbb{Q}} \times \mathcal{L}^{(\mathrm{FI})}(m)_{\mathbb{Q}} \to \mathcal{L}^{(\mathrm{FI})}(m)_{\mathbb{Q}}, \quad (T,T') \mapsto [T,T'] := \overrightarrow{(T'\cdot T)(T\cdot T')},$$

for any $T, T' \in \mathcal{D}^T(m)$ *and extended linearly.*

Proof Under the isomorphism of filtrations from Theorem 2.4.4.1 and by Theorem 2.2.2.3, we have an isomorphism of $\mathbb{Q}$-modules (we omit (FI))

$$\mathcal{L}(m)_{\mathbb{Q}} \cong F^1\mathcal{A}_{\ge 1}(m)_{\mathbb{Q}} \cong \mathrm{Prim}(\mathcal{A}(m)_{\mathbb{Q}})$$

and $\overrightarrow{(T'\cdot T)(T\cdot T')}$ in $\mathcal{L}(m)_{\mathbb{Q}}$ corresponds to $T\cdot T' - T'\cdot T$ in $\mathrm{Prim}(\mathcal{A}(m)_{\mathbb{Q}})$ by definition of $\overrightarrow{(\cdot)}$. Hence $\mathcal{L}(m)_{\mathbb{Q}} \cong \mathrm{Prim}(\mathcal{A}(m)_{\mathbb{Q}})$ as Lie algebras as well. □

3 The rational Goussarov–Habiro Lie algebra of string links

3.1 Clasper calculus and realisation of diagrams

In this section, we describe the *realisation map* $R : \mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \to \mathcal{L}L(m)_{\mathbb{Q}}$ which sends a tree diagram T to the string link obtained by clasper surgery along a tree clasper realising T. First, we need to make precise what is meant by a tree clasper realising a tree diagram. Then, we need to show that the relations 1T, AS, IHX and STU^2 are satisfied. Most are known and recalled in Proposition 3.1.3.1.

For complete definitions of claspers and clasper surgeries, we refer the reader to the original papers of Habiro [22] and Goussarov [17; 18].

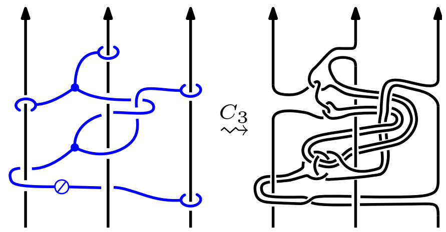


Figure 8: A degree-3 simple tree clasper and the resulting clasper surgery.

3.1.1 Clasper calculus A *simple tree clasper* C on a string link γ is a ribbon uni-trivalent tree diagram embedded in the complement of γ, with *leaves* attached to the strands. See Figure 8, on the left, where γ is the trivial string link on three strands. We use Habiro's drawing convention [22, Figure 7]: nodes (trivalent vertices) are dots, leaves (univalent vertices) are boundaries of disks, each meeting γ transversely exactly once. Ribbon edges have the blackboard framing with half-twists represented by small barred disks, where the bar indicates whether the half-twist is positive or negative (e.g., the clasper in Figure 8 has a *negative* half-twist).

A tree clasper C has a degree $\deg(C) = \#\text{leaves} - 1$, defined as for uni-trivalent tree diagrams. For example, the tree claspers in Figure 3 have degree 1 and 2, respectively, while the one in Figure 8 has degree 3.

Habiro actually defines more general claspers than just simple tree claspers, with boxes and more general leaves. Those appear when manipulating tree claspers. A clasper C determines a *framed* link L_C. Dehn surgery along L_C is called *clasper surgery along C*. A clasper surgery (along a simple clasper) does not modify the ambient manifold (here, the cylinder), but does modify the string link. The result of clasper surgery along a clasper C on γ is denoted γ^C. When γ is the trivial string link $\gamma_0 = 1$, we write $\sigma(C) := \gamma_0^C = 1^C$. See Figures 3 and 8 for examples of clasper surgeries.

Clasper calculus consists in a collection of *moves*, i.e., modifications on claspers that do not alter the resulting surgery γ^C. Those moves are pictured in [22, Proposition 2.7].

Clasper surgeries are entirely supported in a regular neighbourhood of the clasper; hence they define *local* operations on the set of string links and naturally lead to the study of equivalence classes of string links under those moves.

3.1.2 The Goussarov–Habiro Lie algebra of string links Let γ be a string link. A *C_n-move* on γ is a surgery along a degree-n tree clasper on γ. Two string links γ, γ' are said to be *C_n-equivalent*, denoted by $\gamma \overset{C_n}{\sim} \gamma'$, if one can be obtained from the other by applying a sequence of C_n-moves. The C_n-equivalence relation is symmetric by [22, Proposition 3.23]. This defines the Goussarov–Habiro filtration, of the monoid of string links $L(m)$,

$$L(m) = L_1(m) \supset L_2(m) \supset L_3(m) \supset \cdots, \tag{3-1}$$

where $L_n(m) := \{\gamma \in L(m) \mid \gamma \overset{C_n}{\sim} 1\}$ is the submonoid of C_n-trivial string links, i.e., string links that are C_n-equivalent to the trivial string link 1. The quotient monoids $L_k(m)/C_{n+1}$ are finitely generated nilpotent groups [22, Theorem 5.4]. They are abelian groups when $k = n$, in which case they are denoted by

$$\mathcal{L}_n L(m) := \overline{L_n}(m) := L_n(m)/C_{n+1}, \tag{3-2}$$

and more generally, for any $1 \leq k, k' \leq n$, one has

$$[L_k(m)/C_{n+1}, L_{k'}(m)/C_{n+1}] \subset L_{k+k'}(m)/C_{n+1}. \tag{3-3}$$

Thus the abelian groups (3-2) combine into a graded Lie algebra, called the *Goussarov–Habiro Lie algebra of string links on m strands*

$$\mathcal{L}L(m) := \bigoplus_{n\geq 1} \mathcal{L}_n L(m) := \bigoplus_{n\geq 1} \overline{L_n}(m),$$

whose Lie bracket is induced from (3-3). Alternatively, the truncation $\mathcal{L}_{\leq n} L(m)$ is the graded Lie algebra associated to the N-series $\{L_k(m)/C_{n+1}\}_{1\leq k\leq n}$ of the group $L(m)/C_{n+1}$. See [29, Theorem 2.1; 31].

3.1.3 Realising tree diagrams in the Goussarov–Habiro Lie algebra The following proposition gathers known relations between string links obtained from clasper surgeries along tree claspers "realising" the relations AS, IHX, STU^2 and STU between diagrams.

Proposition 3.1.3.1 *Let $n \geq 1$. The following relations, involving claspers on the trivial string link 1, hold in $\overline{L_n}(m)$, where the additive notation is used and the C_{n+1}-equivalence class of a string link 1^C is denoted by $[1^C]$:*

(i) *If C and C' are simple tree claspers of degree n that are **homotopic** with respect to the trivial string link (see [22, §4.1]), then $[1^C] = [1^{C'}]$.*

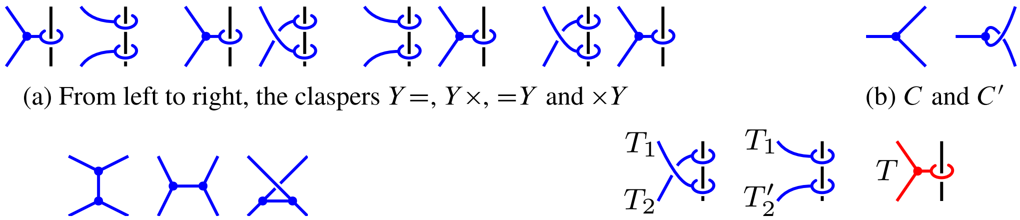
(ii) *If C and C' are simple tree claspers of degree n, that only differ at one place as in Figure 9(b), then $[1^C] + [1^{C'}] = 0$ (AS relation). The same holds if C' is obtained from C by adding a half-twist to an edge.*

(iii) *If I, H and X are three simple tree claspers of degree n, that only differ at one place as in Figure 9(c), then $[1^I] - [1^H] + [1^X] = 0$ (IHX relation).*

(iv) *If $Y=$, $Y\times$, $=Y$ and $=\times$ are four simple tree claspers of degree n, that only differ at the two places shown in Figure 9(a), then $[1^{Y=}] - [1^{Y\times}] = [1^{=Y}] - [1^{Y\times}]$ (STU^2 relation).*

We also have:

(v) *Let $F^{\times} = T_1 \cup T_2$, $F^{=} = T_1 \cup T_2'$, where T_1, T_2, T_2' are three simple tree claspers of degree n_1 (T_1) and n_2 (T_2, T_2'), be such that $F^{\times}$ and $F^{=}$ are identical outside of the part shown in Figure 9(d), which involves one leg of T_1 and another of T_2, respectively T_2' (i.e., T_2 and T_2' only differ by the position of one leg with respect to the shown leg of T_1). Then $1^{F_\times} \overset{C_{n+1}}{\sim} 1^{F^=} \cdot 1^T$ where T is the simple tree clasper (in red) of degree $n := n_1 + n_2$ obtained by joining T_1 and T_2. We say that $F^=$ and $F^\times$ are related by a **slide move**.*



(a) From left to right, the claspers $Y{=}$, $Y\times$, ${=}Y$ and $\times Y$

(b) C and C'

(c) I, H and X

(d) $F^{\times}$, $F^{=}$ and T

Figure 9: Illustration of claspers mentioned in Proposition 3.1.3.1.

Proof See [22, Theorem 4.3] for (i), and [22, Theorem 4.7; 38, E.9] for (ii). Proofs and sketches of (iii) can be found in [10; 16, Proposition 3.4.4.1; 19, Theorem 6.6; 32, Lemma 2.9; 38, E.11]. In [26, Section 10.3], (iv) is proved for knots, but the argument applies for string links. For (v), follow the proof of [22, Proposition 4.4] then "separate" T from $F^{=}$ using a homotopy [22, Theorem 4.3] and then remove the two half-twists, which does not change the C_{n+1}-equivalence class of the result. □

We are now able to describe the realisation map from tree diagrams to string links. Similar realisation maps are mentioned by Habiro [22, §8.2] and described by Ohtsuki [38], Conant–Teichner [13; 14] and Kosanovic [26].

Theorem 3.1.3.2 *For each $n \geq 1$, there is a surjective $\mathbb{Q}$-linear morphism*

$$R_n : \mathcal{L}_n^{\mathrm{FI}}(m)_{\mathbb{Q}} = \mathbb{Q}\mathcal{D}_n^T(m)/\langle 1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle \to \mathcal{L}_n L(m)_{\mathbb{Q}}$$

that sends a tree diagram T to the class of the string link obtained by performing a clasper surgery along a clasper C whose underlying diagram is T (defined in the proof). Those morphisms combine into a surjective morphism of Lie algebras, over $\mathbb{Q}$,

$$R : \mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \twoheadrightarrow \mathcal{L}L(m)_{\mathbb{Q}}$$

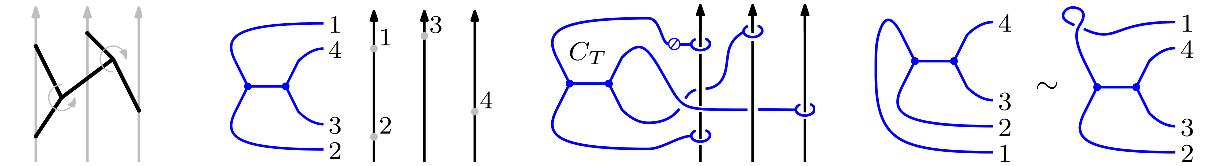
from the primitive Lie algebra of trees (Theorem 2.4.5.1) onto the rational Goussarov–Habiro Lie algebra.

Proof First, we define the realisation map

$$\widetilde{R} : \mathcal{D}_n^T(m) \to \overline{L_n}(m)$$

as follows: Start with a tree diagram $T \in \mathcal{D}_n^T(m)$ as on Figure 10(a). Since T is a tree, its underlying graph is planar and we can pick a planar projection where the cyclic orderings of the nodes are counterclockise. Place that planar projection of the tree to the left of the m vertical strands, with the univalent vertices lying on a vertical line parallel to the strands, as on Figure 10(b). Attach the leaves of the tree to their respective positions on the strand, as given by the diagram T, with an additional (negative or positive) half-twist added to a single chosen leaf[10] (see Figure 10(c)). Outside of that additional half-twist, the attachment

[10]The addition of that additional half-twist is related to the footnote on page 68 in [22]. Although it is not crucial in the case of realising trees, it will be when realising forests later on, in order for the STU relation to hold with the usual signs.



(a) Diagram T (b) Planar tree (c) Clasper C_T (d) Changing the order

Figure 10: Realising a tree diagram.

must be made without introducing half-twists, always going from left to right, and the disk-leaves must clasp to the tangle strands as on Figure 10(c). This yields a simple tree clasper C_T on m strands.

The choices made in the above construction do not alter the result, modulo C_{n+1}-equivalence:

(1) Changing the choice of ordering of the univalent vertices on the vertical line in Figure 10(b) introduces full twists, which do not matter modulo C_{n+1}-moves by (6) in the proof of [22, Theorem 4.3]. See Figure 10(d).

(2) Any two choices of embeddings of the edges joining the planar graph to the disk-leaves are related by a homotopy, which does not change the C_{n+1}-equivalence class of the result by Proposition 3.1.3.1(i).

(3) Adding the half-twist to a different leaf can be obtained by introducing two new half-twists: one to cancel out with the former one, and a new one. Since adding a half-twist corresponds to inversing modulo C_{n+1}-equivalence (Proposition 3.1.3.1(ii)) this choice does not matter modulo C_{n+1}-equivalence. For the same reason, it does not matter whether we chose a positive or negative half-twist.

For a tree clasper C_T in such a position, we say that its *underlying diagram* is $D(C_T) := T$.

Define $\widetilde{R}(T) := \gamma_0^{C_T} = \sigma(C_T)$. Since C_T is a degree-n tree clasper, the C_{n+1}-equivalence class of $\widetilde{R}(T)$ is indeed an element of $\overline{L_n}(m)$.

Since $\overline{L_n}(m)$ is an abelian group, the map $\widetilde{R}$ induces a morphism of abelian groups

$$\widetilde{R} : \mathbb{Z}\mathcal{D}_n^T(m) \to \overline{L_n}(m). \tag{3-4}$$

This morphism sends AS, IHX and STU^2 to zero by Proposition 3.1.3.1(ii), (iii) and (iv), respectively. It also sends 1T to zero; see Figure 11. Thus $\widetilde{R}$ factors through the quotient $\mathbb{Z}\mathcal{D}_n^T(m)/\langle 1\mathrm{T}, \mathrm{AS}, \mathrm{IHX}, \mathrm{STU}^2\rangle$. Moreover, this map is surjective, since $\overline{L_n}(m)$ is generated by string links of the form 1^T where T is a simple tree clasper of degree n [22, Lemma 5.5], and any such clasper can be homotoped [22, Theorem 4.3] to a clasper in the position described in the construction of $\widetilde{R}$ (up to adding and removing half-twists, that is, up to a sign).

After passing to rational coefficients, we obtain the map R_n as in the statement. Those combine into a surjective $\mathbb{Q}$-linear morphism $R : \mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \to \mathcal{L}L(m)_{\mathbb{Q}}$. It remains to show that R is compatible with the Lie brackets.

Pick two tree diagrams T, T' of respective degrees k, l, and let $n := k + l$. Denote by the same letters tree claspers realising them.

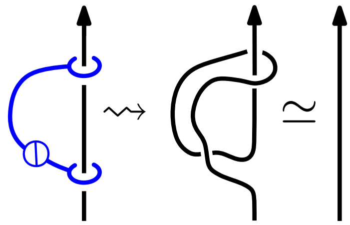


Figure 11: The morphism induced by $\widetilde{R}$ sends 1T to zero.

The commutator $[R(T), R(T')]$ is equal to the C_{n+1}-equivalence class of $\sigma(\overline{T}\,\overline{T'}\,T\,T')$ where $\overline{T}$ represents the inverse of T modulo C_{n+1}-equivalence, idem for T', the product is vertical concatenation and σ denotes surgery modulo C_{n+1}-equivalence. The strategy is to go from the clasper $\overline{T}\,\overline{T'}\,T\,T'$ to the clasper $\overline{T}\,\overline{T'}\,T'\,T$ by a sequence $\epsilon_0, \dots, \epsilon_{r-1}$ of slide moves; see Figure 12 (top). On one hand, $\overline{T}\,\overline{T'}\,T'\,T \overset{C_{n+1}}{\sim} 1$. On the other hand, each slide move creates a red tree clasper $\overrightarrow{\epsilon_i}$ (Proposition 3.1.3.1(v)). All in all, we obtain

$$[R(T), R(T')] = \sigma(\overline{T}\,\overline{T'}\,T\,T') = \sum_{i=1}^{r-1} \sigma(\overrightarrow{\epsilon_i}). \tag{3-5}$$

Let us compare (3-5) with the commutator of the original tree diagrams T, T'. We can follow the exact same path of slide moves on the diagrammatic level to get

$$[T', T] = \overrightarrow{(T\,T')(T'\,T)} = \sum_{i=1}^{r-1} \overrightarrow{e_i} \tag{3-6}$$

by Definition 2.3.2.7.

In (3-5) and (3-6) the roles of T and T' are reversed. This justifies the additional half-twist in the definition of the realisation map. More precisely, the tree claspers realising T, T' both have an additional

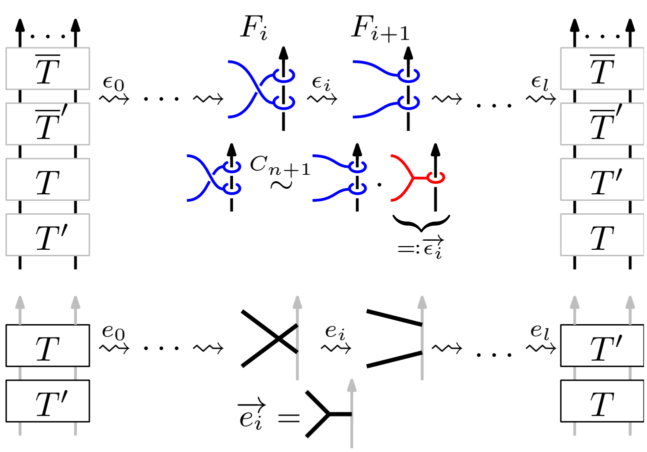


Figure 12: Commutator of string links (top) versus tree diagrams (bottom).

half-twist introduced somewhere. Thus, the red tree clasper $\overrightarrow{\epsilon_i}$ has two half-twists, or zero, modulo C_{n+1}. This means that $R(\overrightarrow{e_i}) = -\sigma(\overrightarrow{\epsilon_i})$ and therefore

$$R([T,T']) = -R([T',T]) = -\sum_i R(\overrightarrow{e_i}) = \sum_i \sigma(\overrightarrow{\epsilon_i}) = [R(T), R(T')],$$

which concludes the proof that R is a surjective morphism of Lie algebras over $\mathbb{Q}$. □

3.2 The Kontsevich integral and the tree preservation theorem

The *Kontsevich integral* is the last ingredient of our identification of the rational Goussarov–Habiro Lie algebra. Discovered by Kontsevich [25] as a universal Vassiliev invariant over $\mathbb{Q}$, it is a morphism

$$Z : L(m) \to \widehat{\mathcal{A}^{\mathrm{FI}}}(m)_{\mathbb{Q}},$$

which to a string link $\gamma \in L(m)$ associates an element $Z(\gamma)$ in the graded completion $\widehat{\mathcal{A}^{\mathrm{FI}}}(m)_{\mathbb{Q}}$ of $\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}$. The Kontsevich integral of a string link $Z(\gamma)$ can be thought of as a *series expansion* of γ, such that degree $\leq n$ Vassiliev invariants only depend on the degree $\leq n$ part $Z_{\leq n}(\gamma)$. However, it is an open question whether Z is injective on $L(m)$, i.e., whether the Kontsevich integral classifies string links.

Since the Kontsevich integral turns string links into diagrams, it is a good candidate for inducing an inverse to the realisation map R. This is indeed the case, as we will see in this section.

3.2.1 The combinatorial Kontsevich integral The combinatorial version of the Kontsevich integral [2] is easier to work with. The combinatorial Kontsevich integral is understood as a functor whose domain is the category **PTangles** of *parenthesized tangles* up to isotopy (denoted by **PT** in [2]), i.e., whose objects are parenthesized $\updownarrow$-words (i.e., finite words in the alphabet $\{\uparrow, \downarrow\}$, together with a parenthesizing of their letters [2, §2.1]) and whose morphisms are isotopy classes of unframed tangles with the adequate $\updownarrow$-words as boundaries, also referred to as *q-tangles* [30]. Its target is the category $\widehat{\mathbf{Diag}}_{\mathbb{Q}}$ of uni-trivalent diagrams [2, Definition 3.1] on tangle skeletons, whose objects are (nonparenthesized) $\updownarrow$-words and morphisms are formal series in $\widehat{\mathcal{A}^{\mathrm{FI}}}(S)_{\mathbb{Q}}$ for tangle skeleta S with the adequate $\updownarrow$-words as boundaries. Composition in **PTangles** is simply vertical concatenation. Composition in $\widehat{\mathbf{Diag}}_{\mathbb{Q}}$ is given by concatenation of the tangle skeleta and concatenation of the diagrams (extended linearly).

Theorem 3.2.1.1 [2; 7; 15; 38] *The Kontsevich integral of unframed parenthesized tangles*

$$Z : \mathbf{PTangles} \to \widehat{\mathbf{Diag}}_{\mathbb{Q}}$$

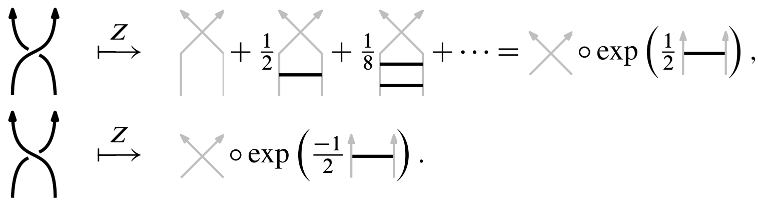
has the following properties:

(i) *It is a* ***functor***: *If $\gamma = \mathrm{id}_w$ for some parenthesized $\updownarrow$-word w, i.e., it consists of parallel vertical strands, then $Z(\gamma) = 1 \in \widehat{\mathcal{A}^{\mathrm{FI}}}(m)$ is the empty diagram. If γ_1, γ_2 are two concatenable tangles (i.e., composable in* **PTangles**), *then $Z(\gamma_1 \circ \gamma_2) = Z(\gamma_1) \circ Z(\gamma_2)$.*

(ii) *It is a* ***monoidal*** *functor, where the monoidal structure on both sides is given by juxtaposition, i.e., horizontal concatenation:*[11] $Z(\gamma_1 \otimes \gamma_2) = Z(\gamma_1) \otimes Z(\gamma_2)$.

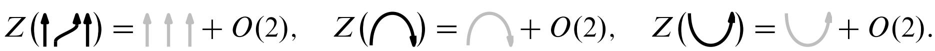
[11] In **PTangles**, the parenthesizing of the tensor product is given by $u \otimes v := ((u)(v))$.

(iii) *It takes, on the standard* ***braiding*** *tangles* (*which are elements of* $\mathbf{PTangles}(\uparrow\uparrow,\uparrow\uparrow)$), *the values*



(iv) *On the standard* ***associator***, ***cap*** *and* ***cup*** *parenthesized tangles*,

it is trivial modulo terms of degree ≥ 2:



(v) *It is compatible with the operation of* ***reversing a strand***. *If* $\boldsymbol{R}_C\gamma$ *is obtained from* γ *by reversing the orientation of the component* C, *then*

$$Z(\boldsymbol{R}_C\gamma)=(\boldsymbol{R}_C)_*Z(\gamma),$$

where $(R_C)_*:\widehat{\mathcal{A}^{\mathrm{FI}}}(S)_{\mathbb{Q}}\to\widehat{\mathcal{A}^{\mathrm{FI}}}(\boldsymbol{R}_C S)_{\mathbb{Q}}$ *is defined as follows. For a diagram* D *on the tangle skeleton* $S=S(\gamma)$,

$$(\boldsymbol{R}_C)_*(D)=(-1)^{\#\{\text{endpoints on } C\}}D'\in\widehat{\mathcal{A}^{\mathrm{FI}}}(\boldsymbol{R}_C S)_{\mathbb{Q}},$$

where D' *is the same diagram as* D *on the tangle skeleton* $\boldsymbol{R}_C S=S(\boldsymbol{R}_C\gamma)$.

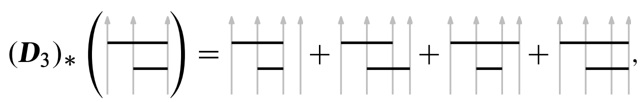
(vi) *It is compatible with the operation of* ***doubling a strand***. *If* $\boldsymbol{D}_C\gamma$ *is obtained from* γ *by doubling*[12] *the strand* C, *then*

$$Z(\boldsymbol{D}_C\gamma)=(\boldsymbol{D}_C)_*Z(\gamma),$$

where $(D_C)_*:\widehat{\mathcal{A}^{\mathrm{FI}}}(S)_{\mathbb{Q}}\to\widehat{\mathcal{A}^{\mathrm{FI}}}(\boldsymbol{D}_C S)_{\mathbb{Q}}$ *is defined as follows. For a diagram* D *on the tangle skeleton* $S=S(\gamma)$,

$$(\boldsymbol{D}_C)_*(D)=\sum_{D' \text{ lifts } D}D'\in\widehat{\mathcal{A}^{\mathrm{FI}}}(\boldsymbol{R}_C S)$$

is the sum of all the $2^{\#\{\text{endpoints on } C\}}$ *ways of "lifting"* D *to a diagram on* $\boldsymbol{D}_C\gamma$. *For example*,



where $\boldsymbol{D}_3$ *is the operation of doubling the third strand.*

As explained in [8, Section 10.3.2], any parenthesized tangle can be decomposed into a product of tensor products of basic tangles: braidings, cups, caps, associators, or images of those under reversing

[12] If $\uparrow$ is an endpoint of C one of the end $\updownarrow$-words, then it is replaced by $(\uparrow\uparrow)$ in $\mathbf{D}_C\gamma$, and similarly for $\downarrow$.

and doubling operations. Knowing the value of Z on each of those standard tangles suffices to deduce the value of Z on any tangle.

It should be mentioned that different choices of value of Z on the associator tangle lead to different combinatorial Kontsevich integrals. The possible choices are called *Drinfeld's associators*, of which there is a deep theory (see [8] for more details). For our purpose, it only suffices to know that they are all trivial modulo terms of degree ≥ 2. This also allows us to ignore the parenthesizing of the extremities of the string links considered in the theorem below.

Denote by $Z_{\leq n}$, Z_n and $Z_{\geq n}$ the compositions $\mathrm{proj} \circ Z$, where proj is the projection onto the corresponding direct summands.

3.2.2 Tree preservation theorem

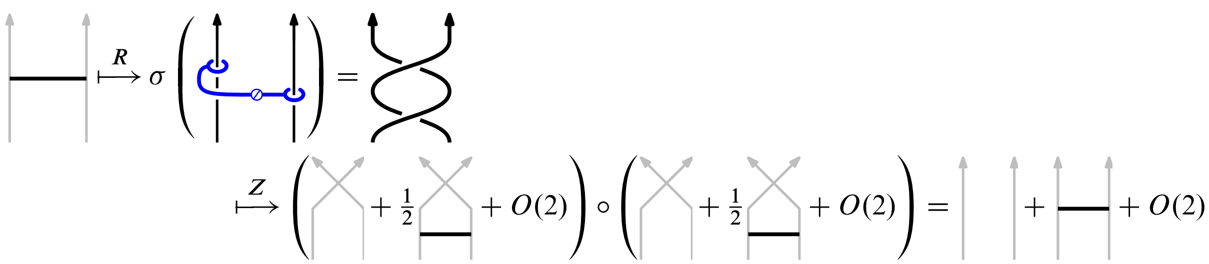
Theorem 3.2.2.1 *Let $n \geq 1$ and let $T \in \mathcal{D}_n^T(m)$ be a degree-n tree diagram on m strands. Then*

$$Z(R(T)) = 1 + T + O(n+1), \tag{3-7}$$

where $R(T) = \sigma(C_T)$ is the result of surgery along a tree clasper C_T realising T (see Theorem 3.1.3.2), and the notation $O(n+1)$ means that the equality holds modulo terms of degree $\geq n+1$. In particular, the Kontsevich integral induces an inverse to R, which then provides an isomorphism of graded Lie $\mathbb{Q}$-algebras $\mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \cong \mathcal{L}L(m)_{\mathbb{Q}}$.

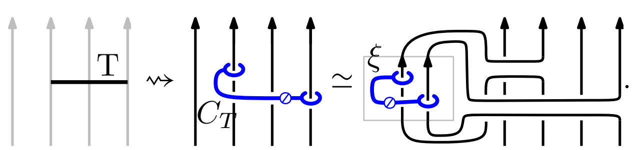
Proof We proceed by induction on n. Moreover, at each step we first make the simplifying assumption (**A**): $m = n+1$, T has one leg on each strand and (if $n > 1$) the legs attached to the rightmost two strands are adjacent to a common node. Then we show the result for the general case.

For $n = 1$ and assuming (**A**), there is only one possible diagram $T \in \mathcal{D}_1^T(2)$ with a leg on each strand. Its realisation is a full positive braiding between the two strands, and hence we obtain the desired result by Theorem 3.2.1.1(iii):

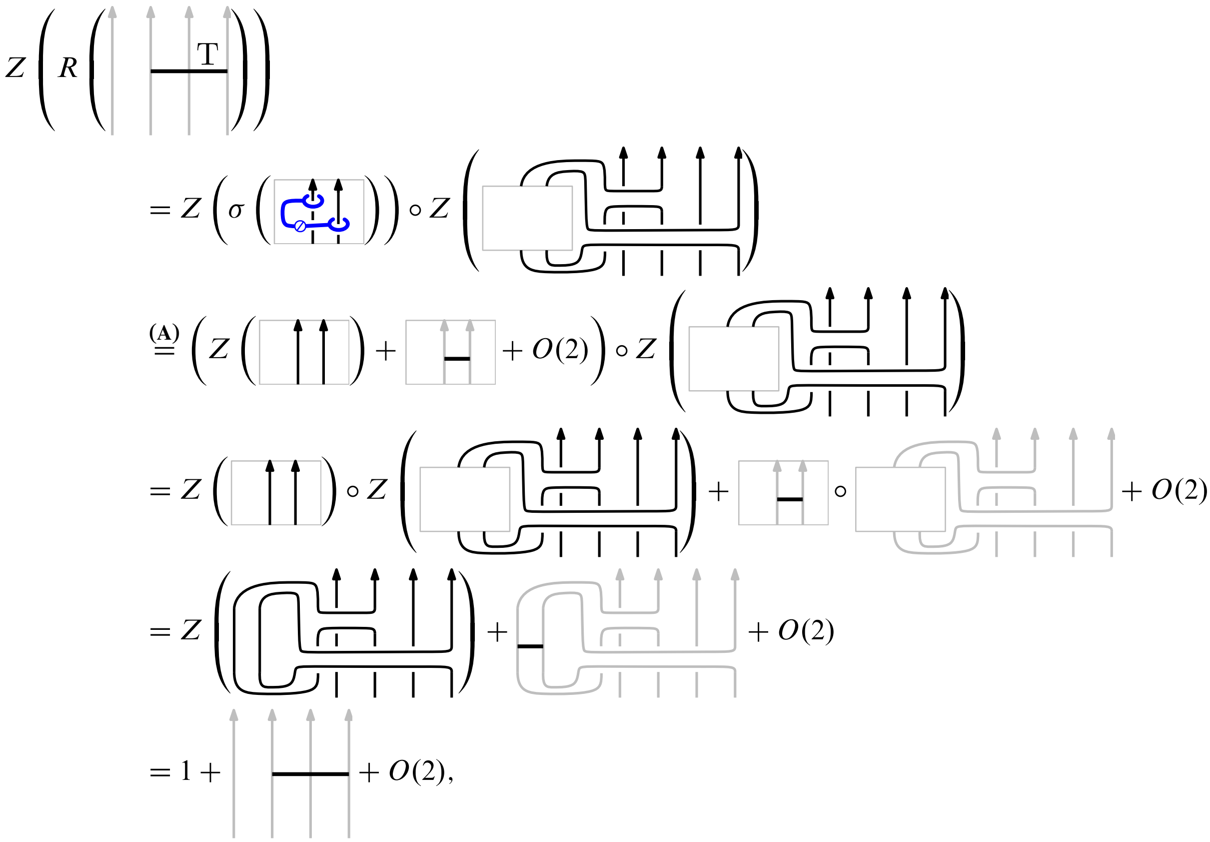


Now we remove the assumption (**A**), and hence the single chord of T can have its legs attached anywhere on the m strands. Consider a clasper C_T realising T, as in the construction of the realising map R; see Figure 10. By an isotopy, we can pull the clasper out and bring it into a small box, such that the box contains 2 parallel vertical strands and the clasper C_T is attached to them satisfying (**A**):

(3-8)



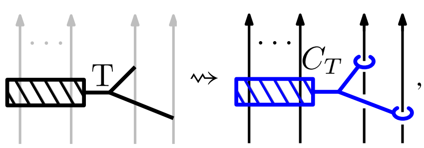
The grey box consists of a string link ξ with a degree-1 tree clasper C_T on it, satisfying (**A**). Thus we compute



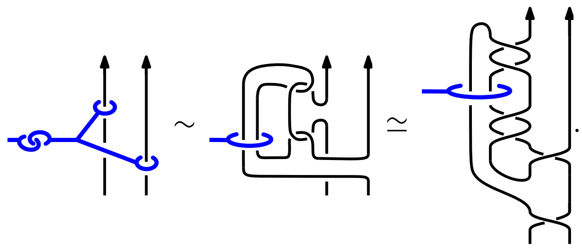
where the first equality uses (3-8) and multiplicativity. Then the left factor can be computed by the above since (**A**) is satisfied in the small grey box. Then we distribute the product over the sum on the left to obtain the second line, where everything that has degree ≥ 2 is contained in $O(2)$. Use multiplicativity again to reconstruct the first term as the Kontsevich integral of a string link isotopic to the unlink, which becomes 1, while the second term gives the diagram T when attached to the outside. This concludes the proof of the case $n = 1$.

Now we show the induction step. Let $n \geq 2$ and assume that the statement holds in degree $n - 1$.

First, consider a degree-n tree diagram $T \in \mathcal{D}_n^T(n+1)$ on $n+1$ strands and satisfying (**A**). It looks like the diagram on the left of



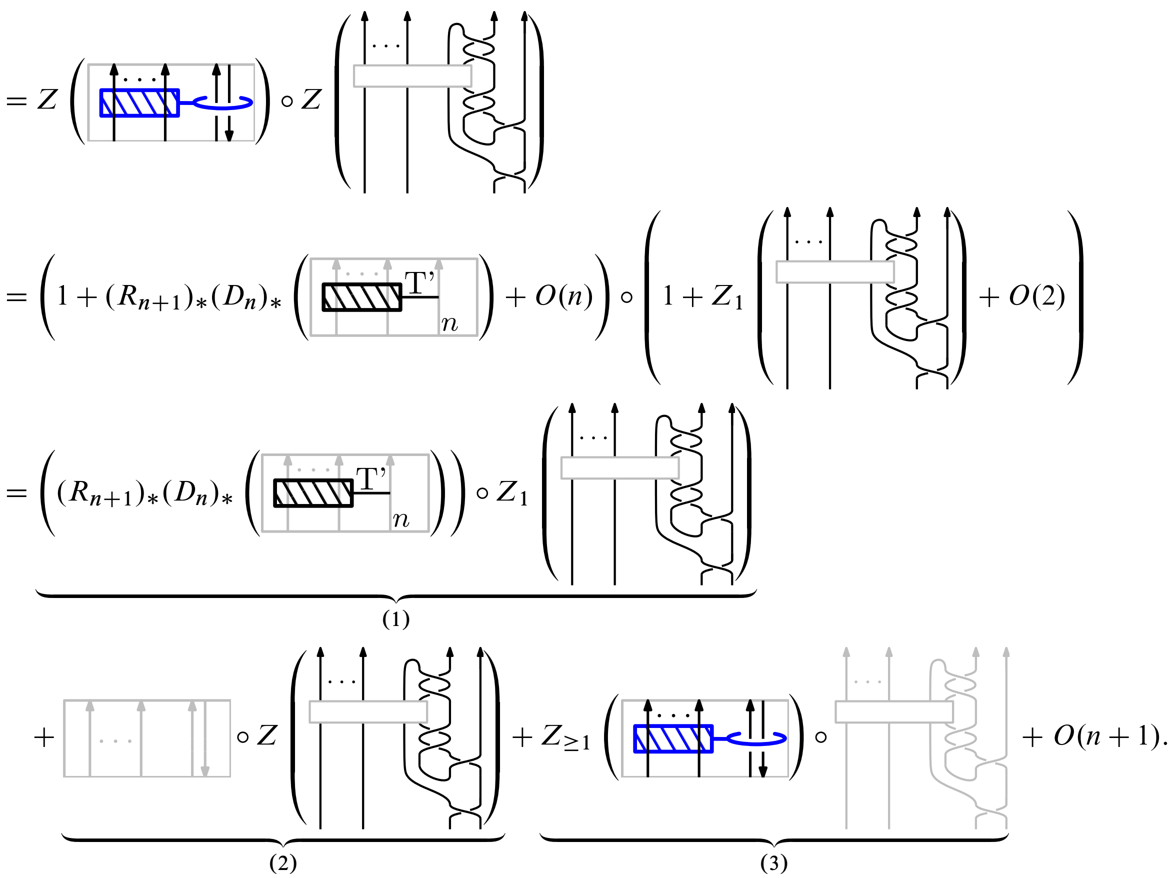
where the black dashed rectangle contains the $n - 1$ other legs. Realise it as a tree clasper C_T, assuming that the additional half-twist is in the blue blob. The core of the induction step is to apply Habiro's move 2 and 10 [22, pages 14, 15] to the visible tripod:



The right-hand side is obtained by isotopy. The new tree clasper C' (C_T minus the tripod) realises a degree-($n-1$) tree diagram T'.

Thus, to compute $R(T)$ we can first apply moves 2 and 10 as above, then perform surgery on C'. This surgery is entirely contained in the grey rectangle shown in the first line below:

$Z(R(T))$

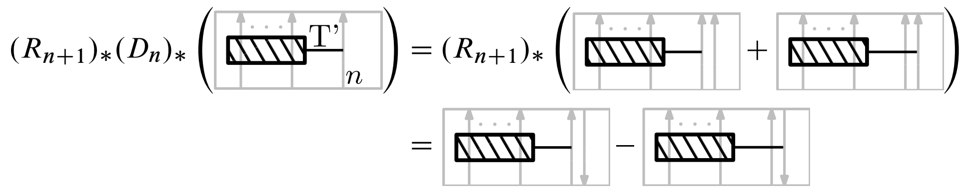


We obtain the first equality by multiplicativity of Z. From the second to the third line, the induction hypothesis

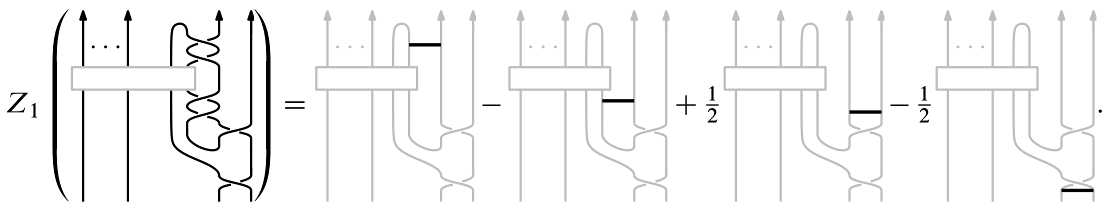
$$Z(R(T')) = 1 + T' + O(n)$$

is used, with additional strand doubling and reversing. Then the fourth line is obtained by distributing this product (Recall that 1 stands for the empty diagram on the adequate tangle skeleton), and putting all the terms of degree $\geq n+1$ inside $O(n+1)$. Let us compute the resulting three terms, denoted (1), (2) and (3), respectively:

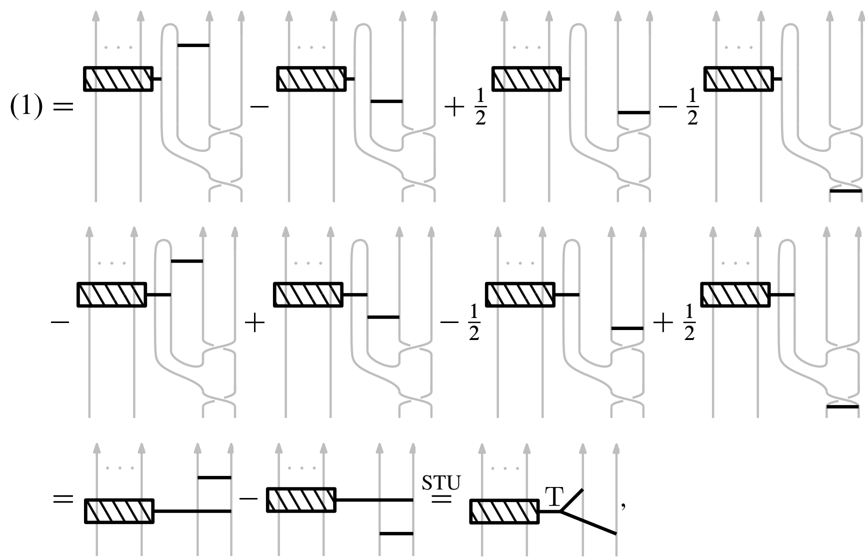
(1) The first term is a product of degree-$(n-1)$ terms with degree-1 terms. First, we compute those two factors individually. Since we are only computing the degree-1 part of Z, only the braidings matter. The first factor is obtained by doubling the n-th strand and reversing the new $(n+1)$-st strand



and the second factor is a sum of Kontsevich integrals of braidings

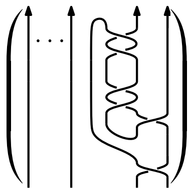
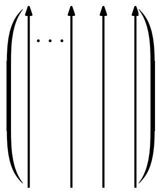


Combining them yields



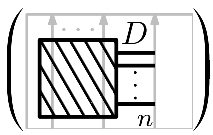
where the last three columns vanish, since the leg attached to the blob is free to slide across the maximum. In the first column, the additional chord prevents that leg from sliding. This column gives the tree diagram T by STU.

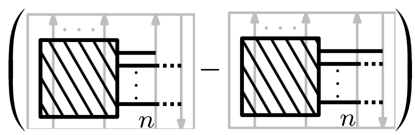
(2) The left factor is Z applied to a trivial string link, with reversed rightmost strand. By multiplicativity of Z, we can combine the two factors into

Z  $= Z$  $= 1.$

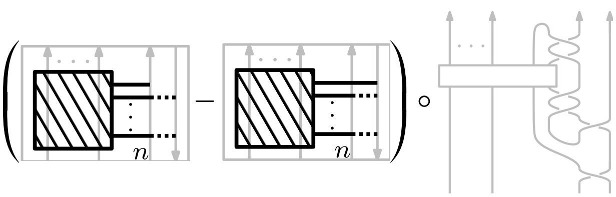
(3) Since terms (1) and (2) equal T and 1, respectively, we expect (3) to vanish. Although the first factor might be very complicated, all we know is that it is obtained from a sum of diagrams by doubling the last strand, then reversing the new last strand. This fact alone actually suffices to show that (3) vanishes.

Indeed, suppose that D is any nonempty diagram on n strands, with $k \geq 1$ legs attached to the last strand. Doubling the last strand turns that diagram into a sum of all its 2^k lifts. Such a lift is indexed by a word $e_1 \dots e_k \in \{0,1\}^k$ where an $e_i = 0$, resp. $e_i = 1$, means that the i-th leg is attached to the n-th, resp. $(n+1)$-st strand. Reversing the last strand has the effect of multiplying by ± 1, depending on the number of legs attached to the last strand. The resulting sum

$(R_{n+1})_*(D_n)_*$ 

$= \sum_{e_2,\dots,e_k \in \{0,1\}} (-1)^{\sum_{j \geq 2} e_j}$ 

is a sum of differences of the form "$e_1 = 0$" minus "$e_1 = 1$". Multiplying each such difference with the second factor of (3) gives 0 since the first leg is now free to go from one position to the other:



$= 0.$

The above shows that all terms in $Z(R(T'))$ that have at least one leg on the last strand vanish once we apply $(R_{n+1})_*(D_n)_*$ and multiply the result with the second factor.

In fact, all the terms in $Z(R(T'))$ must have at least 1 leg on the last strand. Indeed, by definition of the Kontsevich integral, the sum of all terms with no leg on the last strand is equal to the Kontsevich integral of $R(T')$ minus the last strand. If the last strand is removed, then T' gets a trivial leaf and the surgery is trivial. In other words, this comes from the fact that a string link obtained by surgery along a tree clasper with exactly one leg on each strand is a *Brunnian string link* [24].

This concludes the proof that (3) vanishes.

By the above computations of (1), (2) and (3), we obtain

$$Z(R(T)) = 1 + T + O(n+1)$$

as desired. This concludes the case (**A**) in degree n.

The general case is shown exactly as in degree 1, by pulling the clasper out to extract a box containing the clasper (as in (3-8)), in which assumption (**A**) is satisfied. This concludes the induction step, and the proof of (3-7).

It remains to show that Z induces an inverse to R.

By (3-7), the degree-n part of Z induces a map $Z_n : L_n(m) \to \mathcal{A}_n^{\mathrm{FI}}(m)_{\mathbb{Q}}$, which is constant on C_{n+1}-equivalence classes. Thus it factors through

$$Z_n^{GH} : \mathcal{L}_n L(m) \to \mathcal{A}_n^{\mathrm{FI}}(m)_{\mathbb{Q}},$$

whose image lies in $\mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}})_n$. Identifying the latter space with $\mathcal{L}_n^{\mathrm{FI}}(m)_{\mathbb{Q}}$ by Theorem 2.4.5.1, direct summing all the degrees at once and tensoring with $\mathbb{Q}$, we finally obtain

$$Z^{GH} : \mathcal{L}L(m)_{\mathbb{Q}} \to \mathcal{L}_n^{\mathrm{FI}}(m)_{\mathbb{Q}},$$

which is an inverse to R by (3-7). □

4 Application to the rational Goussarov–Habiro conjecture

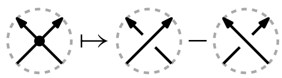
4.1 The rational Goussarov–Habiro conjecture

4.1.1 Vassiliev filtration and associated graded algebra The monoid ring $\mathbb{Z}L(m)$ of string links is filtered by the *Vassiliev filtration* (4-1), with associated graded algebra $\mathcal{A}L(m)$ (4-2):

$$\mathbb{Z}L(m) \supset \partial_1(\mathbb{Z}S_1L(m)) \supset \partial_2(\mathbb{Z}S_2L(m)) \supset \cdots, \tag{4-1}$$

$$\mathcal{A}L(m) := \bigoplus_{n\geq 0} \mathcal{A}_n L(m) := \bigoplus_{n\geq 0} \frac{\partial_n(\mathbb{Z}S_nL(m))}{\partial_{n+1}(\mathbb{Z}S_{n+1}L(m))}, \tag{4-2}$$

where $S_\bullet L(m)$ denotes the graded monoid of (isotopy classes of) singular string links (where the only allowed singularities are transversal double points), graded by the number of double points. The $\mathbb{Z}$-linear maps $\partial_n : \mathbb{Z}S_nL(m) \to \mathbb{Z}L(m)$ are defined by applying the *Skein relation* [3]

(4-3) 

to each double point, where it is understood that the terms are identical outside of the part shown. Thus, for $\gamma \in S_nL(m)$ a singular string link with n double points, its image $\partial_n(\gamma) \in \mathbb{Z}L(m)$ is an alternating sum of all possible resolutions of the n double points.

A string link invariant $V : L(m) \to A$ valued in an abelian group A naturally extends to a Z-linear map $V : \mathbb{Z}L(m) \to A$. We say that V is a *Vassiliev invariant* or *finite type invariant* of degree n if it vanishes on $\partial_{n+1}(\mathbb{Z}S_{n+1}L(m))$. In particular, a Vassiliev invariant of degree n induces a functional on $\mathcal{A}_nL(m)$. Two string links $\gamma, \gamma' \in L(m)$ are *V_n-equivalent* if they cannot be distinguished by Vassiliev invariants of degree $< n$ or, equivalently, if $\gamma - \gamma'$ belongs to $\partial_n(\mathbb{Z}S_nL(m))$.

There is a realisation map $R^v : \mathcal{A}^{\mathrm{FI}}(m) \twoheadrightarrow \mathcal{A}L(m)$ (v refers to *Vassiliev*), which is a morphism of $\mathbb{Z}$-algebras. Originally, R^v was defined on the presentation of $\mathcal{A}^{\mathrm{FI}}(m)$ by chord diagrams, by sending a chord diagram to a singular string link realising it [1]. However, one can define this realisation map on forest diagrams using claspers [13, Theorem 1.1; 22, Section 6], in order to construct a surjective morphism of graded $\mathbb{Z}$-algebras

$$R^v : \mathcal{A}^{\mathrm{FI}}(m) \twoheadrightarrow \mathcal{A}L(m), \tag{4-4}$$

which in degree n sends a forest diagram $F \in \mathcal{D}_n(m)$ to

$$[\gamma_0; T_1, \dots, T_s] := \sum_{\mathcal{S}' \subset \{T_1,\dots,T_s\}} (-1)^{s-|\mathcal{S}'|} (\gamma_0^{\mathcal{S}'}) \in \mathcal{A}_n L(m),$$

where $T_1 \cup \dots \cup T_s$ is a forest clasper realising F, γ_0 denotes the trivial string link on m strands and $\gamma_0^{\mathcal{S}'}$ denotes the result of clasper surgery on γ_0 along the trees contained in $\mathcal{S}'$.

Remark 4.1.1.1 A forest clasper that realises a forest diagram F is constructed as in the proof of Theorem 3.1.3.2, but with multiple trees at the same time and with an additional positive half-twist near a chosen leaf of each tree. Those half-twists are important in order for STU relations to be realised with correct signs. This explains the footnote on page 68 in [22]. △

Over $\mathbb{Q}$, the realisation map $R^v : \mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}} \xrightarrow{\cong} \mathcal{A}L(m)_{\mathbb{Q}}$ is an isomorphism, with inverse induced by the Kontsevich integral [1; 25].

4.1.2 The Goussarov–Habiro conjecture From the definition of the Vassiliev filtration using claspers, it follows that C_n-equivalence implies V_n-equivalence for string links in $L(m)$. Goussarov and Habiro independently conjectured the converse:

Conjecture (Goussarov–Habiro conjecture for string links in the cylinder) *For all $m \geq 1$ and $n \geq 1$, any V_n-equivalent string links on m strands are C_n-equivalent.*

Since C_n-equivalence implies V_n-equivalence, any C_n-trivial string link $\gamma \in L_n(m)$ is automatically V_n-equivalent to the trivial string link 1, and hence $\gamma - 1$ belongs to $\partial_n(\mathbb{Z}S_n L(m))$. This yields the *comparison map* [22, Section 8.2]

$$\chi : \mathcal{L}L(m) \to \mathcal{A}L(m) : [\gamma] \mapsto [\gamma - 1],$$

which is injective if and only if the Goussarov–Habiro conjecture is true. In particular, it is injective when $m = 1$ by the Goussarov–Habiro theorem for knots [22, Theorem 6.18].

Exploiting the *dimension subgroup property* in characteristic zero, Massuyeau showed the following rational version of the Goussarov–Habiro conjecture:

Theorem 4.1.2.1 (Massuyeau [31]) *Over $\mathbb{Q}$, the map $\chi : \mathcal{L}L(m)_{\mathbb{Q}} \to \mathcal{A}L(m)_{\mathbb{Q}}$ is injective.*

Let us now give an alternative proof of Massuyeau's theorem, following from the main result of this paper. Once we pass to rational coefficients, the realisation maps R, R^v and the comparison map χ

combine into the commutative square of graded (Lie or Hopf) algebras

$$\begin{tikzcd}
\mathcal{L}^{\mathrm{FI}}(m)_{\mathbb{Q}} \arrow[r, "\iota", "\cong"'] \arrow[d, "R"', "\cong"] & \mathrm{Prim}(\mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}}) \arrow[r, hook] & \mathcal{A}^{\mathrm{FI}}(m)_{\mathbb{Q}} \arrow[d, bend right, "R^v"', "\cong"] \\
\mathcal{L}L(m)_{\mathbb{Q}} \arrow[ru, "Z^{GH}"'] \arrow[rr, hook, "\chi"] & & \mathcal{A}L(m)_{\mathbb{Q}} \arrow[u, bend right, "Z^v"']
\end{tikzcd}$$

where Z^v is the inverse to R^v by Kontsevich's theorem, ι is an isomorphism by Theorem 2.4.5.1, the left triangle commutes and R becomes an isomorphism by Theorem 3.2.2.1. Therefore, χ is injective, which concludes the alternative proof of the rational Goussarov–Habiro conjecture.

Acknowledgements

This work was done under the supervision of Peter Teichner, whom I wish to thank for his guidance. I also wish to thank James Conant for helpful discussions and comments. I also thank the anonymous referee for helpful comments and corrections. Part of the writing was supported by FNR AFR grant 18890152.

References

[1] **D Bar-Natan**, *On the Vassiliev knot invariants*, Topology 34:2 (1995) 423–472 MR

[2] **D Bar-Natan**, *Non-associative tangles*, from "Geometric topology" (Athens, GA, 1993) (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2.1, Amer. Math. Soc., Providence, RI (1997) 139–183 MR

[3] **J S Birman**, **X-S Lin**, *Knot polynomials and Vassiliev's invariants*, Invent. Math. 111:2 (1993) 225–270 MR

[4] **P Boavida de Brito**, **G Horel**, *Galois symmetries of knot spaces*, Compos. Math. 157:5 (2021) 997–1021 MR

[5] **R Budney**, **J Conant**, **R Koytcheff**, **D Sinha**, *Embedding calculus knot invariants are of finite type*, Algebr. Geom. Topol. 17:3 (2017) 1701–1742 MR

[6] **P Cartier**, *A primer of Hopf algebras*, from "Frontiers in number theory, physics, and geometry, II" (P Cartier, B Julia, P Moussa, P Vanhove, editors), Springer (2007) 537–615 MR

[7] **S Chmutov**, **S Duzhin**, *The Kontsevich integral*, Acta Appl. Math. 66:2 (2001) 155–190 MR

[8] **S Chmutov**, **S Duzhin**, **J Mostovoy**, *Introduction to Vassiliev knot invariants*, Cambridge Univ. Press (2012) MR

[9] **J Conant**, *Homotopy approximations to the space of knots, Feynman diagrams, and a conjecture of Scannell and Sinha*, Amer. J. Math. 130:2 (2008) 341–357 MR

[10] **J Conant**, **R Schneiderman**, **P Teichner**, *Jacobi identities in low-dimensional topology*, Compos. Math. 143:3 (2007) 780–810 MR

[11] **J Conant**, **R Schneiderman**, **P Teichner**, *Whitney tower concordance of classical links*, Geom. Topol. 16:3 (2012) 1419–1479 MR

[12] **J Conant**, **R Schneiderman**, **P Teichner**, *Geometric filtrations of string links and homology cylinders*, Quantum Topol. 7:2 (2016) 281–328 MR

[13] **J Conant**, **P Teichner**, *Grope cobordism and Feynman diagrams*, Math. Ann. 328:1-2 (2004) 135–171 MR

[14] **J Conant**, **P Teichner**, *Grope cobordism of classical knots*, Topology 43:1 (2004) 119–156 MR

[15] **V G Drinfeld**, *On quasi-triangular quasi-Hopf algfebras and a group closely connected with* $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, Leningrad Math. J. 2:4 (1991) 829–860

[16] **B Dular**, *The Lie algebra of tree diagrams and the rational Goussarov–Habiro conjecture*, Master's thesis (2023) Available at https://sites.google.com/view/brunodular/research

[17] **M N Goussarov**, *Interdependent modifications of links and invariants of finite degree*, Topology 37:3 (1998) 595–602 MR

[18] **M Goussarov**, **M Polyak**, **O Viro**, *Finite-type invariants of classical and virtual knots*, Topology 39:5 (2000) 1045–1068 MR

[19] **M N Gusarov**, *Variations of knotted graphs: the geometric technique of n-equivalence*, Algebra i Analiz 12:4 (2000) 79–125 MR In Russian; translated in St. Petersburg Math. J. 12 (2001) 569–604

[20] **N Habegger**, **X-S Lin**, *The classification of links up to link-homotopy*, J. Amer. Math. Soc. 3:2 (1990) 389–419 MR

[21] **N Habegger**, **G Masbaum**, *The Kontsevich integral and Milnor's invariants*, Topology 39:6 (2000) 1253–1289 MR

[22] **K Habiro**, *Claspers and finite type invariants of links*, Geom. Topol. 4 (2000) 1–83 MR

[23] **K Habiro**, **G Massuyeau**, *Symplectic Jacobi diagrams and the Lie algebra of homology cylinders*, J. Topol. 2:3 (2009) 527–569 MR

[24] **K Habiro**, **J-B Meilhan**, *On the Kontsevich integral of Brunnian links*, Algebr. Geom. Topol. 6 (2006) 1399–1412 MR

[25] **M Kontsevich**, *Vassiliev's knot invariants*, from "I M Gelfand Seminar, II" (S Gelfand, S Gindikin, editors), Adv. Soviet Math. 16, Amer. Math. Soc., Providence, RI (1993) 137–150 MR

[26] **D Kosanović**, *A geometric approach to the embedding calculus knot invariants*, PhD thesis, Universität Bonn (2020) Available at https://bonndoc.ulb.uni-bonn.de/xmlui/bitstream/handle/20.500.11811/8651/5979.pdf

[27] **P Lambrechts**, **V Turchin**, **I Volić**, *Associahedron, cyclohedron and permutohedron as compactifications of configuration spaces*, Bull. Belg. Math. Soc. Simon Stevin 17:2 (2010) 303–332 MR

[28] **S K Lando**, *On primitive elements in the bialgebra of chord diagrams*, from "Topics in singularity theory" (A Khovanskiĭ, A Varchenko, V Vassiliev, A B Sossinsky, editors), Amer. Math. Soc. Transl. Ser. 2 180, Amer. Math. Soc., Providence, RI (1997) 167–174 MR

[29] **M Lazard**, *Sur les groupes nilpotents et les anneaux de Lie*, Ann. Sci. École Norm. Sup. (3) 71 (1954) 101–190 MR

[30] **T Q T Le**, **J Murakami**, *The universal Vassiliev–Kontsevich invariant for framed oriented links*, Compositio Math. 102:1 (1996) 41–64 MR

[31] **G Massuyeau**, *Finite-type invariants of* 3*-manifolds and the dimension subgroup problem*, J. Lond. Math. Soc. (2) 75:3 (2007) 791–811 MR

[32] **J-B Meilhan**, *On surgery along Brunnian links in* 3*-manifolds*, Algebr. Geom. Topol. 6 (2006) 2417–2453 MR

[33] **J-B Meilhan**, *Linking number and milnor invariants*, from "Encyclopedia of knot theory" (C Adams, E Flapan, A Henrich, L H Kauffman, L D Ludwig, S Nelson, editors), Chapman and Hall/CRC (2021) 817–830

[34] **J-B Meilhan**, **A Yasuhara**, *Characterization of finite type string link invariants of degree* < 5, Math. Proc. Cambridge Philos. Soc. 148:3 (2010) 439–472 MR

[35] **J W Milnor**, **J C Moore**, *On the structure of Hopf algebras*, Ann. of Math. (2) 81 (1965) 211–264 MR

[36] **S Montgomery**, *Hopf algebras and their actions on rings*, CBMS Regional Conference Series in Mathematics 82, Amer. Math. Soc., Providence, RI (1993) MR

[37] **Y Nozaki**, **M Sato**, **M Suzuki**, *On the kernel of the surgery map restricted to the* 1*-loop part*, J. Topol. 15:2 (2022) 587–619 MR

[38] **T Ohtsuki**, *Quantum invariants: a study of knots,* 3*-manifolds, and their sets*, Series on Knots and Everything 29, World Scientific, River Edge, NJ (2002) MR

[39] **A Postnikov**, *Permutohedra, associahedra, and beyond*, Int. Math. Res. Not. 2009:6 (2009) 1026–1106 MR

[40] **M Ronco**, *Shuffle bialgebras*, Ann. Inst. Fourier (Grenoble) 61:3 (2011) 799–850 MR

[41] **Y Shi**, *Goodwillie's cosimplicial model for the space of long knots and its applications*, J. Homotopy Relat. Struct. 18:2–3 (2023) 265–312 MR

[42] **V G Turaev**, *Quantum invariants of knots and* 3*-manifolds*, 3rd edition, De Gruyter Studies in Mathematics 18, De Gruyter, Berlin (2016) MR

[43] **V A Vassiliev**, *Cohomology of knot spaces*, from "Theory of singularities and its applications" (V I Arnold, editor), Adv. Soviet Math. 1, Amer. Math. Soc., Providence, RI (1990) 23–69 MR

[44] **G M Ziegler**, *Lectures on polytopes*, Graduate Texts in Mathematics 152, Springer (1995) MR

BRUNO DULAR bruno.dular@uni.lu

Department of Mathematics, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Received: August 6, 2024 Revised: March 17, 2025

Algebraic & Geometric Topology

msp.org/agt

See inside back cover or msp.org/agt for submission instructions.

The subscription price for 2026 is US $795/year for the electronic version, and $1170/year (+$80, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Algebraic & Geometric Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Algebraic & Geometric Topology (ISSN 1472-2747 printed, 1472-2739 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, 2000 Allston Way # 59, Berkeley, CA 94701-4004.

AGT peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

https://msp.org/

ALGEBRAIC & GEOMETRIC TOPOLOGY

Volume 26 Issue 3 (pages 825–1227) 2026